23. Ulusal ve 6. Uluslararası

Biyoistatistik Kongresi

26-29 Ekim 2022

Ankara Üniversitesi Tıp Fakültesi,
Morfoloji Yerleşkesi

Özet Bildiri Kitabı

Değerli Meslektaşlarımız,

Biyoistatistik Derneği ve Ankara Üniversitesi Tıp Fakültesi Biyoistatistik Anabilim Dalı olarak **26-29 Ekim 2022** tarihlerinde Ankara'da düzenleyeceğimiz **23. Ulusal ve 6. Uluslararası Biyoistatistik kongresi'**ne sizleri davet etmekten büyük mutluluk duyuyoruz.

Tüm dünyayı etkileyen Covid-19 Pandemisi nedeniyle geçtiğimiz yıl kongremizi çevrim içi olarak gerçekleştirebilmişken, en son 26-29 Ekim 2019'da Antalya'da düzenlenen kongrede bir araya gelebilmiştik. Üç yılın ardından tekrar buluşacak olmanın heyecanıyla düzenleme kurulu olarak bilimsel ve sosyal açıdan zengin içerikli bir kongre hazırlamak için tüm gücümüzle çalışıyoruz.

Kongrelerin bilimsel niteliğini artıran en önemli etkinliklerden birisi, alanında uzman eğiticilerin güncel konularda kısa süre içinde bilgi paylaşımı yaptıkları kurslardır. Kongremiz kapsamında 26 Ekim 2022 tarihinde uygulamalı olarak düzenlenecek dört farklı kurs ile katılımcıların bilgi ve program kullanma becerilerini geliştirmeyi hedeflemekteyiz. 27-29 Ekim 2022 tarihlerinde ise sözlü ve poster sunumlarının yer aldığı bir bilimsel program ile sosyal gezi programı planlamaktayız.

Siz değerli meslektaşlarımızın katılımı ile bilimsel açıdan zengin ve güncel bilgilerin yer aldığı çalışmaların paylaşıldığı başarılı bir kongre geçireceğimizden hiç şüphemiz yok. Birlikteliğimizi pekiştirmek ve gücümüze güç katmak için tüm meslektaşlarımızı kongremizde buluşmaya davet ediyoruz,

Saygılarımızla,

Prof. Dr. Yasemin Yavuz
Kongre Başkanı

Prof. Dr. Erdem Karabulut
Biyoistatistik Derneği Başkanı

Doç. Dr. Derya Gökmen
Kongre Sekreteri

Profesör **Alan TENNANT**
Profesör **Ahmet DİRİCAN**
Profesör **Ahmet ÖZTÜRK**
Profesör **Atilla Halil ELHAN**
Profesör **Ayşe Canan YAZICI GÜVERCİN**
Profesör **Beyza ŞAHİN**
Profesör **Bülent ÇELİK**
Profesör **Canan BAYDEMİR**
Profesör **Celal Reha ALPAR**
Profesör **Erdem KARABULUT**
Profesör **Ersin ÖĞÜŞ**
Profesör **Ferhan ELMALI**
Profesör **Gülhan TEMEL**
Profesör **Hafize SEZER**
Profesör **İmran ÖMÜRLÜ**
Profesör **Kadir SÜMBÜLOĞLU**
Profesör **Mehtap AKÇİL OK**
Profesör **Melody GHAHRAMANİ**
Profesör **Peter W. ROSE**
Profesör **Recai YÜCEL**
Profesör **Sanjib BASU**
Profesör **Seval KUL**
Profesör **Sıddık KESKİN**
Profesör **Şengül CANGÖR**
Profesör **Ünal ERKORKMAZ**
Profesör **Vildan SÜMBÜLOĞLU**
Profesör **Yasemin YAVUZ**
Profesör **Yusuf ÇELİK**

Profesör **Z. Nazan ALPARSLAN**
Profesör **Zeki AKKUŞ**
Profesör **Mithat GÖNEN**
Doçent Doktor **Adem DOĞANER**
Doçent Doktor **Alex R. de LEON**
Doçent Doktor **Beyza DOĞANAY ERDOĞAN**
Doçent Doktor **Can ATEŞ**
Doçent Doktor **Derya GÖKMEN**
Doçent Doktor **Emre DEMİR**
Doçent Doktor **Erdal COŞGUN**
Doçent Doktor **Gökmen ZARARSIZ**
Doçent Doktor **Harika GÖZÜKARA BAĞ**
Doçent Doktor **İlkay DOĞAN**
Doçent Doktor **İlker ÜNAL**
Doçent Doktor **Jale KARAKAYA**
Doçent Doktor **Kenan KÖSE**
Doçent Doktor **Leman TOMAK**
Doçent Doktor **Lynn SLEEPER**
Doçent Doktor **Mehmet KARADAĞ**
Doçent Doktor **Mustafa Agah TEKİNDAL**
Doçent Doktor **Pınar GÜNEL**
Doçent Doktor **Selcen PEHLİVAN**
Doçent Doktor **Selçuk KORKMAZ**
Doçent Doktor **Şirin ÇETİN**
Doçent Doktor **Umut ÖZBEK**
Yrd. Doçent Doktor **Duchwan RYU**
Dr. Öğr. Üyesi **Tolga KASKATI**
Dr. **Bernd KLAUS**

Kongre Başkanı

Prof. Dr. Yasemin YAVUZ, Ankara Üniversitesi

Düzenleme Kurulu

Prof. Dr. Yasemin YAVUZ, Ankara Üniversitesi

Prof. Dr. Erdem KARABULUT, Hacettepe Üniversitesi

Prof. Dr. Mehtap AKÇİL OK, Başkent Üniversitesi

Prof. Dr. Seval KUL, Gaziantep Üniversitesi

Doç. Dr. Derya GÖKMEN, Ankara Üniversitesi

Doç. Dr. İlker ÜNAL, Çukurova Üniversitesi

Dr. İrem KAR, Ankara Üniversitesi

Dr. Batuhan BAKIRARAR, Ankara Üniversitesi

26.10.2022 Çarşamba		
09.00-17.00	**Kongre Kayıt**	
Salon	**Prof. Dr. Lütfi Tat Konferans Salonu (Mavi Salon)**	**Pembe Salon**
09.30-13.00	**RevMan ve R ile Uygulamalı Meta Analizi** Prof. Dr. Erdem KARABULUT, Doç. Dr. Can ATEŞ	**JAMOVİ ve JASP ile Veri Analitiği** Doç. Dr. Beyza DOĞANAY ERDOĞAN, Doç. Dr. Mustafa Agah TEKİNDAL
13.00-14.00	**Öğle Yemeği Arası**	
14.00-17.30	**R ve SPSS uygulamaları ile Propensity Skor Eşleştirme Analizi** Doç. Dr. Emre DEMİR	**Yüksek Boyutlu Verilerin R Programında Analizi** Prof. Dr. Seval KUL

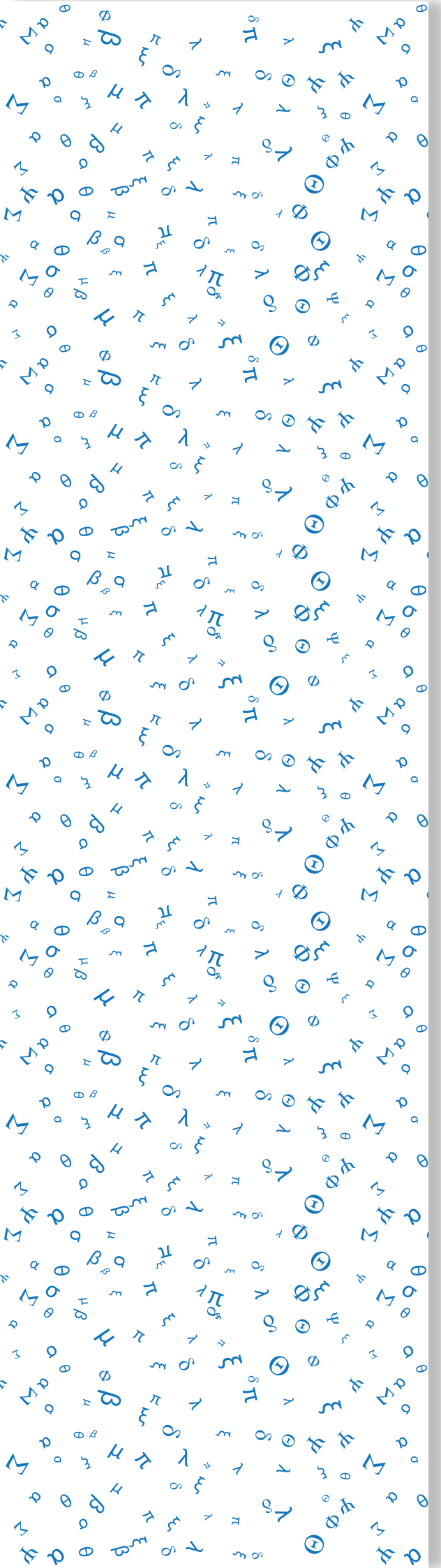
27.10.2022 Perşembe			
09.00-17.00	**Kongre Kayıt**		
09.30-10.45	**Açılış Programı (Ord. Prof. Dr. Abdülkadir Noyan Konferans Salonu)**		
10.45-11.00	**Çay/Kahve Arası**		
11.00-12.00	**Bayesçi Yöntemlerin Klinik Denemelerdeki Uygulamaları (Ord. Prof. Dr. Abdülkadir Noyan Konferans Salonu)** **Konuşmacı: Prof. Dr. Mithat Gönen, Memorial Sloan Kettering Cancer Center**		
12.00-13.30	**Öğle Yemeği Arası**		
Salon	**Ord. Prof. Dr. Abdülkadir Noyan Konferans Salonu**	**Sarı Salon**	**Pembe Salon**
Oturum Başkanları	**Prof. Dr. Seval KUL** **Prof. Dr. Ayşe Canan YAZICI GÜVERCİN** **Doç. Dr. Didem DERİCİ YILDIRIM**	**Prof. Dr. Kevser Setenay ÖNER** **Prof. Dr. Pınar ÖZDEMİR** **Doç. Dr. İlkay DOĞAN**	**Prof. Dr. G. Nural BEKİROĞLU** **Prof. Dr. Duru KARASOY** **Doç. Dr. S. Kenan KÖSE**
13.30-13.45	Genom-boyu İlişki Çalışmalarında Poligenik Risk Skorunun Makine Öğrenimi ve Derin Öğrenme Yöntemleri ile Tahmin Edilmesi Ragıp Onur ÖZTORNACI, Erdal COŞGUN, Bahar TAŞDELEN, Cemil ÇOLAK	Türkiye Covid-19 Verileri için Parçalı Fonksiyona Dayalı Büyüme Modelleri Leyla BAKACAK KARABENLİ, Serpil AKTAŞ ALTUNAY	Periton Diyaliz Hastalarında Sağkalım ve Uzunlamasına Verilerin Birlikte Modellemesi: Kişiye Özgü Dinamik Risk Kestirimleri Merve Başol GÖKSÜLÜK, Dinçer GÖKSÜLÜK, Ergun KARAAĞAOĞLU
13.45-14.00	Recurrent Neural Network for Compex Survival Problems Pius MARTHIN	Psikiyatrik Hastalıklarda, Psikopatoloji ve Çevresel Etken İlişkisinin Özyinelemesiz Model ile Değerlendirilmesi Selen Begüm UZUN, Derya GÖKMEN, Meram Can SAKA	Boylamsal ve Yaşam Verilerinin Parametrik Bileşik Modellemesinde Parametre Tahmin Yaklaşımlarının Karşılaştırılması Elif YILDIRIM, Duru KARASOY
14.00-14.15	RNA-Dizileme Verilerinde Sağkalım Algoritmalarının Kapsamlı Bir Karşılaştırması Ahu CEPHE, Ahmet SEZGİN, Necla KOÇHAN, Gözde ERTÜRK ZARARSIZ, Vahap ELDEM, Erdem KARABULUT, Gökmen ZARARSIZ	The Interaction Term and Its Graphical Interpretation In Regression Models Handan ANKARALI, Özge PASİN, Senem GÖNENÇ	Sağkalım Analizi için Veri Simülasyonu İrem KAR, Erdal COŞGUN, Atilla Halil ELHAN
14.15-14.30	Biocrates Mxp® Quant 500 Kiti Kullanılarak Elde Edilen Metabolomik Verilerinin Laboratuvarlar Arası Karşılaştırması Gözde ERTÜRK ZARARSIZ, Alexander CECIL, Jutta LINTELMANN, Gernot POSCHET, Jennifer KIRWAN, Sven SCHUCHARDT, Xue L. GUAN, Daisuke SAIGUSA, David WISHART, Jiamin ZHENG, Rupasri MANDAL, Lisa ST. JOHN-WILLIAMS, Kendra ADAMS, J. Will THOMPSON, Michael P. SNYDER, Kevin CONTREPOIS, S Songjie CHEN, Nadia ASHRAFI, Sumeyya AKYOL, Ali YILMAZ, Stewart GRAHAM, Thomas M. O'CONNELL, Karel KALECKÝ, T. Hai PHAM, Therese KOAL, Teodoro BOTTIGLIERI, Jerzy ADAMSKI, Gabi KASTENMÜLLER	Kantil Regresyon Yöntemi İle Oluşturulan Pediatrik Büyüme Eğrilerinin Model Uyumunun Worm Grafiği İle Değerlendirilmesi Eda ÇAKMAK, Ergun KARAAĞAOĞLU, Serhat KILIÇ, Pınar ÖZDEMİR	Sağkalım Analizinde Kopula Entropi Tabanlı Değişken Seçimi Aslıhan ALHAN

14.30-14.45	RNA-Dizi Verilerinde İntron Tutulumunun Belirlenmesinde Kulllanılan Biyoinformatik Algoritmaların Değerlendirilmesi Esma Gamze AKSEL, Vahap ELDEM, Gökmen ZARARSIZ	Küçük Küme ve Periyotlu Boylamsal Çalışmalarda Değişken ve Model Seçimi Merve TURKEGÜN, Bahar TAŞDELEN, Saim YOLOĞLU	Kadaverik Böbrek Naklinde Yaşam Süresini Belirleyen Nakil Sırasında ve Sonrasındaki Önemli Değişkenler Pembe KESKİNOĞLU, Kemal OLÇA, Duygu KARAKAYA, Ebru ÖZTÜRK, Mehmet ORMAN, Timur KÖSE, Erdem KARABULUT, Ahmet DİRİCAN
14.45-15.00	**Çay/Kahve Arası**		
Oturum Başkanları	**Prof. Dr. Canan BAYDEMİR Prof. Dr. Gökmen ZARARSIZ Doç. Dr. Pınar GÜNEL**	**Prof. Dr. Ahmet ÖZTÜRK Prof. Dr. Hafize SEZER Doç. Dr. Harika Gözde GÖZÜKARA BAĞ**	**Prof. Dr. Ertuğrul ÇOLAK Prof. Dr. Fezan MUTLU Doç. Dr. Emre DEMİR**
15.00-15.15	Machine Learning Sensors for Diagnosis of COVID-19 Disease Using Routine Blood Values for Internet of Things Application Mehmet Tahir HUYUT, Andrei VELICHKO, Belyaev MAKSIM	Efor Kullanımlı ve Üç Parametreli Lojistik Madde Yanıt Teorisi Modellerinin Bilgisayar Uyarlamalı Test Performanslarının Karşılaştırılması Yusuf Kemal ARSLAN, Afra ALKAN, Atilla Halil ELHAN	Referans Aralığı Belirlemede Kullanılan Tek Değişkenli ve Çok Değişkenli Yöntemlerin Karşılaştırılması Esra Kutsal MERGEN, Sevilay KARAHAN
15.15-15.30	Makine Öğrenmesinde Sınıf Dengesizliği Sorunu için Bazı Veri Ön İşleme Yöntemlerinin Karşılaştırılması: Bir Uygulama Hülya ÖZEN, Doğukan ÖZEN	Ölçme Araçlarının Boyutluluk Yapısının Belirlenmesinde Yapısal Eşitlik Modellemesine Dayalı Yöntemler Serhat HAYME, Derya GÖKMEN, Şehim KUTLAY, Ayşe Adile KÜÇÜKDEVECİ	Mamografi Görüntülerinde Farklı Ön İşleme Kombinasyonlarının Sınıflandırma Performansları Üzerine Etkisi Hanife AVCI, Jale KARAKAYA
15.30-15.45	Makine Öğrenme Yöntemlerinin Başarılarını Değerlendirirken Hangi Performans Ölçütleri Kullanılmalıdır. Yaşar SERTDEMİR, İlker ÜNAL, Hülya BİNOKAY	Doğrusal Olmayan Yapısal Eşitlik Modellemesi Yaklaşımlarının İncelenmesi ve Bir Uygulama Nur Efşan TIĞLI, Şengül CANGÜR	Obuchowski Parametrik olmayan ROC Eğrisi Analizi ve Küme Bootstrap Yaklaşımları ile İlişkili Verilerin Değerlendirilmesi: Multipl Skleroz Hastalarında Optik Nöropati Teşhisi Üzerine Bir Uygulama Büşra EMİR, Fatma Ezgi CAN, Mustafa Agah TEKİNDAL, Ferhan ELMALI, Ertuğrul ÇOLAK
15.45-16.00	Veri Madenciliğinde Hibrit Model Yaklaşımı Batuhan BAKIRARAR, Atilla Halil ELHAN	Biyoistatistik Dersine Yönelik İstatistik Öz Yeterlilik İnanç Ölçeği, İstatistik Kaygı Ölçeği ve İstatistik Tutum Ölçeği Arasındaki Aracılık İlişkisinin Araştırılması: Bir Yapısal Eşitlik Modellemesi Çalışması (Path Analizi) Kamber KAŞALİ, Senem GÖNENÇ	Sıralı Bir Yanıt Değişkenini Sınıflamada Sayısal Bir Ölçüm İçin Optimal Kesim Noktalarını Belirleme İlker ÜNAL, Yaşar SERTDEMIR, H. Esin ÜNAL
16.00-16.15	Yayın Yanlılığını Önlemede P-Eğrisi Yöntemi Arzu BAYGÜL EDEN, Alev BAKIR KAYI	Yapısal Eşitlik Modellemesi Uygulamalarının Raporlanması: Bir Kontrol Listesi Zeliha AYDIN KASAP, Burçin KURT, Kemal TURHAN	The Effectiveness of Blood Routine Parameters and Some Biomarkers as a Potential Diagnostic Tool in the Diagnosis and Prognosis of COVID-19 Disease Mehmet Tahir HUYUT, Sıddık KESKİN

28.10.2022 Cuma			
Salon	**Ord. Prof. Dr. Abdülkadir Noyan Konferans Salonu**	**Sarı Salon**	**Pembe Salon**
Oturum Başkanları		**Prof. Dr. Atilla Halil ELHAN Prof. Dr. Şengül CANGÜR Doç. Dr. Özlem KAYMAZ**	**Prof.Dr. Gülşah SEYDAOĞLU Prof. Dr. Sinan SARAÇLI Doç. Dr. Mehmet KARADAĞ**
09.00-09.05		Sobel ve Baron-Kenny Yöntemlerinin Performanslarının Karşılaştırılması Hülya BİNOKAY, Yaşar SERTDEMİR	Lökosit Antijen Molekülleri Toplum Doku Tipi Sıklığı Kemal OLÇA, Ahmet KESKİNOĞLU, Duygu KARAKAYA, Ebru ÖZTÜRK, Pembe KESKİNOĞLU, Mehmet ORMAN, Timur KÖSE, Ahmet DİRİCAN, Erdem KARABULUT
09.05-09.10		Mamografi Görüntülerinde K-ortalamalar Kümeleme ve Canny Kenar Algılama Bölütleme Yöntemlerinin Sınıflandırma Performanslarının Karşılaştırılması Jale KARAKAYA, Hanife AVCI	Çok Boyutlu Ölçekleme Yöntemi ile Trakya'da Tarımsal Üretim Bölgelerinde Ağır Metal Yoğunluğunun Gösterimi Büşra AYDIN, Özgecan KORKMAZ AĞAOĞLU, İsmayıl Safa GURCAN
09.10-09.15		Log Multinomial Regresyon Modeline Hosmer-Lemeshow Uyum İyiliği Testinin Uyarlanması Yasemin ÖZTÜRK, Erdem KARABULUT	Lojistik Regresyon Modeli ile Kardiyovasküler Kalp Hastalığı Sınıflandırılması ve Tahmini Mehmet KIVRAK
09.15-09.20		Diyabetik Gözlerde Multifokal Elektroretinografide Halka Amplitüdünün ve Halka Amplitüdünü Etkileyen Faktörlerin Değerlendirilmesi Ayşe Büşra GÜNAY ŞENER, Hidayet ŞENER	Önceden Eğitilmiş Evrişimsel Sinir Ağlarından Elde Edilen Öznitelikler ile Sperm Morfolojisinin Julia Tabanlı Web Uygulaması ile Sınıflandırılması Hüseyin KUTLU, Emek GÜLDOĞAN, Cemil ÇOLAK
09.20-10.00	**Çay/Kahve Arası**		
10.00-12.00	**Panel (Ord. Prof. Dr. Abdülkadir Noyan Konferans Salonu) Moderatörler: Prof. Dr. Z. Nazan ALPARSLAN, Prof. Dr. Yasemin YAVUZ**		
12.30-16.30	**Gezi Programı**		

29.10.2022 Cumartesi			
Salon	**Ord. Prof. Dr. Abdülkadir Noyan Konferans Salonu**	**Sarı Salon**	**Pembe Salon**
Oturum Başkanları	**Prof. Dr. Mehtap AKÇİL OK Doç. Dr. Jale KARAKAYA KARABULUT Doç. Dr. Beyza DOĞANAY ERDOĞAN**	**Prof. Dr. Yusuf ÇELİK Prof. Dr. Nurhan DOĞAN Doç. Dr. Cengiz BAL**	**Prof. Dr. Ismayıl Safa GÜRCAN Prof. Dr. S. Yavuz SANİSOĞLU Doç. Dr. Yaşar SERTDEMİR**
09.30-9.45	Makine Öğrenmesi Algoritmalarının Performanslarının Farklı Koşullar Altında Karşılaştırılması: Bir Monte Carlo Benzetim Çalışması Müge COŞKUN, Osman DAĞ	dtComb: İki Sürekli Tanı Testinin Birleştirilmesine Yönelik Geliştirilmiş Kapsamlı Bir R Paketi Serra İlayda YERLİTAŞ, Serra Bersan GENGEÇ, Gözde ERTÜRK ZARARSIZ, Selçuk KORKMAZ, Gökmen ZARARSIZ	Çok Değişkenli Varyans Analizinde Kullanılan Test İstatistiklerinin Tip I Hata Oranları Bakımından Karşılaştırılması Fatma Ezgi CAN, Büşra EMİR, Ferhan ELMALI, Can ATEŞ, Mustafa Agah TEKİNDAL
09.45-10.00	Boyut İndirgeme Uygulanan ve Uygulanmayan Verilerin Analizinde Makine Öğrenimi Yöntemlerinin Karşılaştırılması Feyza İNCEOĞLU, Cemil ÇOLAK	Propensity Skora Dayalı Eşleştirme Analizi İçin Yeni Bir R Shiny Web Uygulaması Emre DEMİR, Gülçin AYDOĞDU, Muhammed Ali YILMAZ, Yasemin YAVUZ	Kanıtsal Güç ve Literatür Bütünlüğünün p-Eğrisi ile İncelenmesi Alev BAKIR KAYI, Arzu BAYGÜL EDEN

10.00-10.15	Yüksek Boyutlu Verilerde Eksik Veri Değer Atama Yöntemlerinin Tahmin Performanslarının Aşırı Öğrenme Makineleriyle İncelenmesi Buğra VAROL, İmran ÖMÜRLÜ, Mevlüt TÜRE	R Shiny ile Zaman Serileri Analizi için İnteraktif Arayüz Tasarımı Esra GÜLTÜRK, Abdulvahap PINAR, Ahmet Kadir ARSLAN, Pınar YILMAZ	Taguchi ve Faktöriyel Deney Tasarım Metotlarının Karşılaştırması: Lactobacillus spp. ERIC-PCR Optimizasyonu Örneği Selen YILMAZ IŞIKHAN, Ceren ÖZKUL
10.15-10.30	Sınıflamada Kullanılan Veri Madenciliği Yöntemlerinin Performanslarının D Vitamini Veri Seti Üzerinde Karşılaştırılması Özlem ARIK, İnci ARIKAN, Erdem KARABULUT	Twowaytests: İki Yönlü Bağımsız Grup Tasarımı İçin Bir R Paketi Muhammed Ali YILMAZ, Osman DAĞ, Samaradasa WEERAHANDI, Malwane M.A. ANANDA	Eğilim Skoru Eşleştirmesinde Greedy ve Optimal Yöntemlerin Karşılaştırılması Gülden HAKVERDİ, Timur KÖSE, Cemil ÇOLAK
10.30-10.45	Bulanık Kümeleme Yöntemi ile Biyoistatistik Dersi Alan Öğrencilerin Değerlendirilmesi Senem GÖNENÇ, Kamber KAŞALİ, Özge PASİN	Kantil Regresyon Yaklaşımı için İnteraktif Bir Arayüz Tasarımı Abdulvahap PINAR, Ahmet Kadir ARSLAN	KOAH Hastalarında Fonksiyonel Kapasite Üzerindeki Etkilerin İstatistiksel Yöntemlerle Analizi Sinan SARAÇLI, Aydın BALCI, Berkalp TUNCA
10.45-11.00	**Çay/Kahve Arası**		
Oturum Başkanları	**Prof. Dr. Ferhan ELMALI** **Doç. Dr. İlker ÜNAL** **Doç. Dr. Selen YILMAZ IŞIKHAN**	**Prof. Dr. Ünal ERKORKMAZ** **Doç. Dr. Selçuk KORKMAZ** **Doç. Dr. M. Agah TEKİNDAL**	**Prof. Dr. Erdem KARABULUT** **Doç. Dr. Selcen YÜKSEL** **Doç. Dr. Can ATEŞ**
11.00-11.15	Covid-19 Tanısında Konvolüsyonel Sinir Ağları Mimarilerinin Karşılaştırılması Işıl ÜNALDI, Leman TOMAK	Performance Analysis of Simulated Data Sets by Machine Learning Algorithms Senem KOC, Leman TOMAK, Erdem KARABULUT	Holstein Irkı İneklerde Güç Doğum Üzerinde Etkili Olan Faktörlerin Belirlenmesinde FP Growth Algoritmasının Kullanımı Elif ÇELİK GÜRBULAK, Murat ABAY, Aytaç AKÇAY, Funda İPEKTEN, Gökmen ZARARSIZ, Esma Gamze AKSEL
11.15-11.30	Görüntü İşleme Yöntemlerinin COVID-19 Pozitif Tespitinde Klasik Veri Madenciliği Yöntemleri ile Sınıflandırmaya Etkisi Fatma Gül KURT, Jale KARAKAYA	Propensity Skora Dayalı Eşleştirme Analizinde Optimum Caliper Belirlenmesi İçin Bir Benzetim Çalışması Gülçin AYDOĞDU, Muhammed Ali YILMAZ, Emre DEMİR, Yasemin YAVUZ	Kardiyoverter Defibrilatör (ICD) İmplantasyonu Yapılan Hastalarda Yaşam Durumu Tahmin Modellemesi Hatice AKTAŞ GÖKÇE, Amine BAYRAKLI, Emre YAŞAR
11.30-11.45	Diagnosis and Prognosis of COVID-19 Disease Using Routine Blood Values and LogNNet Neural Network Mehmet Tahir HUYUT, Andrei VELICHKO	Assessing and comparison of various biomarker scenarios for guiding treatment selection: Simulation results and recommendations Hülya KOÇYİĞİT	Erken Evre Meme Kanseri Hastalarına ait Hazard Faktörlerinin Uygulanan Çeşitli Model Sonuçlarının Karşılaştırılması Nural BEKİROĞLU, Ayşe ÜLGEN, Emrah Gökay ÖZGÜR, Sinan UZUN, Wentian LI
11.45-12.00	Önerilen Derin Öğrenme Mimarisi ile MR Görüntülerine Dayalı Alzheimer Hastalığının Sınıflandırılması Mehmet KIVRAK, Cemil ÇOLAK	BKSY: Bilgi Keşfi Süreci Yazılımı Ahmet Kadir ARSLAN, Cemil ÇOLAK	MEVARYA Mutlu UMAROĞLU
12.00-12.15	**Çay/Kahve Arası**		
12.15-13.00	**Kapanış Konuşmaları (Ord. Prof. Dr. Abdülkadir Noyan Konferans Salonu)**		



Sözel Bildileriler

SB – 1

Yapısal Eşitlik Modellemesi Uygulamalarının Raporlanması: Bir Kontrol Listesi

[1]Zeliha AYDIN KASAP, [2]Burçin KURT, [3]Kemal TURHAN

[1]Karadeniz Teknik Üniversitesi, Sağlık Bilimleri Enstitüsü, Biyoistatistik ve Tıp Bilişimi Anabilim Dalı
[2]Karadeniz Teknik Üniversitesi, Sağlık Bilimleri Enstitüsü, Biyoistatistik ve Tıp Bilişimi Anabilim Dalı
[3]Karadeniz Teknik Üniversitesi, Sağlık Bilimleri Enstitüsü, Biyoistatistik ve Tıp Bilişimi Anabilim Dalı

Amaç: Bu çalışmada, Yapısal Eşitlik Modellemesi (YEM) uygulamalarında içerik, örneklem boyutu, eksik verilerin yönetimi, modellerin özellikleri ve tanımlanması, parametre tahmin yöntemi seçimleri ile model performanslarında uygun uyum indeksleri ve etkilendiği durumların göz önüne alınarak raporlamada yapılan ve sık karşılaşılan hataların azaltılmasına yönelik araştırmacı/okuyucu/değerlendiricilere rehberlik edecek bir kontrol listesinin geliştirilmesi amaçlanmıştır.

Gereç ve Yöntem: YEM uygulamalarında araştırma sorularının açık ve yalın bir dil ile belirlenmesi ve geliştirilecek ölçüm/yapısal modelin kavramsal çerçevesinin teorik temele dayandırılması oldukça önemlidir. Çalışmanın başında sürece hakim olunması, çalışmanın kalite standartlarını arttırmada ve süreç yönetiminin hızlandırılmasında etkin rol oynamaktadır. Literatürde, YEM için ihtiyaç duyulan örneklem sayısının, 5 ile 20 gözlem arasında değiştiği görülmektedir. Bunun yanında çalışmalarda eksik verilerin üstesinden nasıl gelindiği konusu da kapsamlı bir tartışmayı içermelidir. Bu çalışmada, literatürde gözlenen kontrol listelerindeki uygulama basamakları detaylıca incelenmiş ve geliştirilen kontrol listesinin ana hatları modelin tasarımı, modelin değerlendirilmesi, modelin modifikasyonu ve YEM tabanlı çalışmaların yorumlanması/raporlanması olmak üzere 4 temel başlık altında listelenerek güncellenmiştir.

Bulgular: YEM uygulamalarında, varsayımların kontrol edilmemesi, önsel modelin teorik temele dayandırılmaması, ölçüm sonuçlarının geçerli ve güvenilir ölçeklerle elde edilmemesi, modelin teorisinde girdi matrisine alınacak elemanlar ile çıktı matrisinde oluşacak duruma dikkat edilmemesive sonucunda ilgili metriklerin uygun şekilde hesaplanamaması (underidentified/overidentified) sık görülen hatalardan bazılarıdır. Ayrıca, YEM modeli için yapılan açıklayıcı ve doğrulayıcı faktör analizlerinde aynı veri setinin kullanılması sonucu uyum indekslerinin yanıltıcı sonuçlar vermesi ve modelde uzman görüşüne başvurulmadan modifikasyon indekslerinin iyileştirilmeye çalışılması YEM uygulamalarında göz ardı edilen durumlar arasında gelmektedir. İstatistiksel bir model oluşturmanın amacının karmaşık veri yapısını basitleştirmek olduğu düşünüldüğünde, değişken sayısını azaltmaya çalışmamış bir modelin parsimoni ilkesi açısından daha az değerli olduğu unutulmamalıdır.

Sonuç: YEM kavramı, ölçüm, regresyon, yol analizi ve faktör modelleri dahil olmak üzere çok geniş bir model yelpazesini kapsar. Bu nedenle, araştırmacıların yöntemi anlayabilmeleri ve uygun şekilde raporlayabilmeleri için önemli düzeyde istatistiksel bilgiye sahip olmaları gerekmektedir. Bu çalışma, karmaşıklığı nedeniyle YEM modellerinin her yönünü kapsayamasa da, bir YEM çalışmasının; araştırmacılar tarafından ele alınmasında, gerek çalışmanın zaman yönetimine gerekse raporlamaların standart ve etkin bir şekilde gerçekleştirilmesine katkı sağlayacağı düşünülmektedir. Böylece, bu çalışma, YEM tabanlı uygulamaların araştırma kalitesinin iyileştirilmesine ve raporlama sürecinde modelin bir bütün halinde, kontrol listesindeki adımlar ile değerlendirmesinde araştırmacılara yardımcı olacaktır.

Anahtar Kelimeler: Faktör Analizi, Yapısal Eşitlik Modellemesi, Kontrol Listesi

SB – 2

Kadaverik Böbrek Naklinde Yaşam Süresini Belirleyen Nakil Sırasında ve Sonrasındaki Önemli Değişkenler

[1]Pembe KESKİNOĞLU, [2]Kemal OLÇA, [3]Duygu KARAKAYA, [4]Ebru ÖZTÜRK, [5]Mehmet ORMAN, [6]Timur KÖSE, [7]Erdem KARABULUT, [8]Ahmet DİRİCAN

[1]Dokuz Eylül Üniversitesi Tıp Fakültesi Biyoistatistik ve Tıbbi Bilişim AD
[2]İstanbul Üniversitesi Cerrahpaşa Lisansüstü Eğitim Enstitüsü, Biyoistatistik ve Tıbbi Bilişim AD
[3]Ege Üniversitesi Tıp Fakültesi Biyoistatistik ve Tıbbi Bilişim AD
[4]Hacettepe Üniversitesi Tıp Fakültesi Biyoistatistik AD
[5]Ege Üniversitesi Tıp Fakültesi Biyoistatistik ve Tıbbi Bilişim AD
[6]Ege Üniversitesi Tıp Fakültesi Biyoistatistik ve Tıbbi Bilişim AD
[7]Hacettepe Üniversitesi Tıp Fakültesi Biyoistatistik AD
[8]İstanbul Üniversitesi Cerrahpaşa Lisansüstü Eğitim Enstitüsü, Biyoistatistik ve Tıbbi Bilişim AD

Amaç: Son dönem böbrek yetmezliği gelişmiş hastalarda temel hedef tedavi, böbrek naklidir. Ülkemizde ve dünyada organ bekleme listelerinin çoğunu böbrek nakli bekleyen hastalar oluşturmaktadır. Ülkemizde organ nakli bekleyen hastaların %86.4'ü böbrek beklemektedir ve kadavra donör sayısı diğer ülkelerden çok daha sınırlıdır. Optimal eşleştirme için öncelikle klinik gözlem ve kararlar ön plana çıkmıştır. Sonradan doku-organ reddi gibi başarısızlık sonuçları biyoistatistik yaklaşımlarla değerlendirilmeye başlanmıştır. Konu ile ilgili biyoistatistik değerlendirmeler, nakilli böbreğin yaşam süresini etkili değişkenlerle birlikte inceleyen cox regresyon analizi ile yapılır. Böbrek naklinde red veya olumsuz sonuçların tahminlenmesi, alıcı, verici ve doku tipi ve klinik özelliklerle gerçekleştirilir. Bu çalışmada kadaverik organ nakli yapılmış çok merkez verileri rerospektif değerlendirilerek organ veya hasta ölümüne neden olan eşleştirme ile ilgili değişkenlerin etkisi araştırılmıştır.

Gereç ve Yöntem: Üç büyük ilde (İzmir, İstanbul ve Ankara) kadaverik nakil sayısı yüksek olan 9 organ nakli merkezinin verileri (hardcopy hasta dosyaları, elektronik ortamda ulaşılan veriler) kaydedilmiştir. Merkezler arası farklı kayıt özelliği olabilecek verilerin girişleri için standardizasyonu eğitimler yapılmıştır. Amerika ve Avrupa organ nakil sistemleri (UNOS ve Eurotransplant) incelenmiş, klinik alandaki deneyimler ve ulaşılabilen böbrek nakline ait veriler ile cinsiyet, yaş, birden fazla transplantasyon geçirme durumu, nakil öncesi hipertansiyon ve diyabet varlığı, nakil sonrası gecikmiş greft fonksiyonu yaşanması durumu ve HLA doku uyum değişkenlerinin nakil edilen organın yaşam süresine etkileri araştırılmıştır. Açıklayıcı değişkenler stepwise (adım adım) seçme yöntemi ile modele dahil edilip çıkarılarak yaşam süresine en anlamlı etki eden en anlamlı açıklayıcı değişkenler araştırılmıştır. Bu çalışmada sunulan veriler proje verilerinin tamamını içermemektedir.

Bulgular: Üç ilde dokuz organ nakli merkez verilerinden elde edilen 645 hastanın verisi analiz edilmiştir. Kadaverik böbrek nakli yapılan bu hastaların yaş ortalaması 40,3 yıldır ve 74'ü 18 yaş altında (çocuk), 73'ü 35-50 yaş aralığında ve 413'ü 50 yaş üzerinde olduğu saptandı. Alıcı hastaların 201'inde nakil öncesi hipertansiyon ve 48'inde nakil öncesi diyabet tanısı bulunmaktadır. Nakil anı ve hemen sonrası gelişen gecikmiş greft fonksiyonu 137 hastada gerçekleşmiştir, 29 alıcının birden fazla nakil geçirmiş olduğu belirlenmiştir. Parametrik modeller ve Cox oransal tehlikeler modelleri uygulanarak, alıcının gecikmiş greft fonksiyonu yaşanması ve yaş artışının organ yaşam süresini açıklamada diğer açıklayıcı değişkenlere göre daha etkili olduğu bulunmuştur.

Sonuç: Yalnızca kadaverik böbrek nakilli hastaların verilerinden eşleştirme için önemli olan değişkenlerin dahil edildiği biyoistatistik modellerle analiz edildiğinde, yaş ve gecikmiş greft fonksiyon varlığı değişkenleri böbrek yaşam süresi için önemli bulunmuştur. Ancak bu değişkenler klinik bilgi ile de birlikte değerlendirilmelidir. Gecikmiş gretf fonksiyonu ara sonuç ise neden-sonuç ilişkileri tekrar değerlendirilerek analiz edilmelidir. Teşekkür: Bu çalışma TÜBİTAK-SBAG- 318S164 Nolu proje verilerinden gerçekleştirilmiştir.

Anahtar Kelimeler: Kadaverik Böbrek Nakli, Sağkalım analizi, Cox regresyon

SB – 3

Assessing and Comparison of Various Biomarker Scenarios for Guiding Treatment Selection: Simulation Results and Recommendations

Hülya KOÇYİĞİT

Karamanoglu Mehmetbey University

Aim: The use of treatment selection biomarkers is essential for assessing whether or not a patient improves after receiving a certain treatment. Oncology has performed a lot of research on tumor markers throughout the years, however, few of these markers have been shown to be useful medically. Therefore, it is crucial to identify the appropriate therapy biomarker to verify the patient's benefits. This study suggests a set of descriptive and inferential techniques for assessing individual markers and evaluating potential markers.

Materials and Methods: It is frequently required to take into account factors related to the target marker when assessing these markers. This study determines the best decision rule based on the marker to categorize patients following the benefits of their treatments, and we characterize the marker's performance using the associated classification accuracy as well as the classifier's overall distribution.

Results: Across this paper, Monte Carlo simulation, which imitates breast cancer datasets, has been performed to illustrate treatment selection based on the biomarkers.

Conclusion: This paper is a rare study in observational studies investigating and presenting summary measures for MC simulation based on the cancer markers. Few studies in the understanding of biomarker rely on observational studies instead of randomized control trial. But, in a non-randomized observational study, researchers have no control over the treatment. Thus, this paper performs various simulation scenario to illustrate confounding effects when used in observational data sets. Thus, generating different treatment, outcome or both have been performed significantly different impact on treatment selection.

Keywords: Treatment selection, biomarker, observational study

SB – 4

Boyut İndirgeme Uygulanan ve Uygulanmayan Verilerin Analizinde Makine Öğrenimi Yöntemlerinin Karşılaştırılması

[1]Feyza INCEOĞLU, [2]Cemil ÇOLAK

[1]İnönü Üniversitesi Tıp Fakültesi Biyoistatistik ve Tıp Bilişimi Anabilim Dalı, Malatya, Türkiye
[2]İnönü Üniversitesi Tıp Fakültesi Biyoistatistik ve Tıp Bilişimi Anabilim Dalı, Malatya, Türkiye

Amaç: Çok boyutlu verilerde araştırmacılar veri seti hakkında bilgi sahibi olmadığı durumlarda veriyi yorumlayabilmek için boyut indirgeme yöntemlerini kullanmaktadır. Boyut indirgeme yöntemleri ile çok boyutlu verilerde daha küçük yapılar elde edilerek veri hakkında daha kolay yorumlar yapılmaktadır. Temel amaç veri kaybını minimum düzeyde tutarak veriden yüksek düzeyde bilgi sağlamaktır. Boyut indirgeme yöntemi için en çok tercih edilen yöntemlerden biri de "Temel Bileşenler Analizi"dir. Gerçek veri ile izdüşümleri arasındaki ortalama uzaklık minimum yapılarak analiz gerçekleştirilmekte ve çok boyutlu veriler küçük boyutlu homojen yapılara dönüştürülmektedir. Bu çalışmanın amacı açık kaynak kodlu bir veri setinde boyut indirgeme yöntemlerinden olan temel bileşenler analiz yöntemi uygulanan ve uygulanmayan verilerde Random Forest ve Destek Vektör Makinesi yöntem sonuçlarının doğruluk oranlarını karşılaştırmaktadır.

Gereç ve Yöntem: Çalışmada Kaggle kardiyovasküler veri seti kullanılmıştır. 462 bireyden oluşan veri setine ilk olarak boyut indirgeme analizi uygulanmadan önce Random Forest (RF) ve Destek Vektör Makinesi (SVM) doğruluk oranları hesaplanmıştır. Daha sonra temel bileşenler analizi yöntemi (PCA) uygulanarak boyut indirgeme analizi uygulanmıştır. PCA sonrası Random Forest (RF) ve Destek Vektör Makinesi (SVM) doğruluk oranları hesaplanmıştır. Elde edilen sonuçlar karşılaştırılmıştır. Bütün hesaplamalarda Python programlama dili ve paketleri kullanılmıştır.

Bulgular: Çalışmada kullanılan veri setinde yer alan katılımcılara ait yaş ortalaması 42.82 ± 14.61 standart sapmadır. Katılımcıların 302'sinde (%64.4) kardiyovasküler hastalık bulunmaz iken 160'ında (%34.6) bulunmaktadır. Veri setinde kayıp gözlem bulunmamaktadır. Temel bileşenler analizi (PCA) sonucunda Bartlett Küresellik test sonucu 936.247 ($p=0.001<0.05$) olarak ve Kaiser-Meyer-Olkin (KMO) test değeri ise 0.69 olarak hesaplanmıştır. PCA %86.6 açıklanan varyans oranına sahip 5 faktörlü yapı elde edilmiştir. 5 faktörlü yapının elde edilen doğruluk oranı RF ile 0.60, SVM ile 0.71 olarak hesaplanmıştır. Gözetimli makine öğrenmesi ila gerçekleştirilen analiz sonucunda doğruluk oranı RF ile 0.70, SVM ile 0.73 olarak hesaplanmıştır.

Sonuç: Çok sınıflı yapılarda elde edilen doğruluk oranları iki sınıflı yapılara göre daha düşük bulunmuştur. Gözetimsiz makine öğrenme yöntemleri veri yapısı bilinmediği ve dağılımlarının orantılığı olmadığı durumlarda dezavantajlıdır. Veri yapısında orantısal dağılımlar sağlandığı zaman sonuçların doğruluk oranı da artacaktır. Yüksek boyutlu veriler ile çalışmak hem zaman hem de maddi açıdan kayıplara neden olmaktadır. Bu nedenle boyut indirgeme yöntemleri avantaj sağlamaktadır. Bununla birlikte boyut indirgeme analizleri veri görselleştirmede de araştırmacılara sunum açısından kolaylık sağlamaktadır.

Anahtar Kelimeler: Boyut İndirgeme, Temel bileşenler Analizi, Random Forest, Destek Vektör Makinesi.

SB – 5

Holstein Irkı İneklerde Güç Doğum Üzerinde Etkili Olan Faktörlerin Belirlenmesinde FP Growth Algoritmasının Kullanımı

[1]Elif ÇELİK GÜRBULAK, [2]Murat ABAY, [3]Aytaç AKÇAY, [4]Funda İPEKTEN, [5]Gökmen ZARARSIZ, [6]Esma Gamze AKSEL

[1]Erciyes Üniversitesi Veteriner Fakültesi Biyometri Anabilim Dalı Kayseri
[2]Erciyes Üniversitesi Veteriner Fakültesi, Doğum ve Jinekoloji Anabilim Dalı, Kayseri
[3]Ankara Üniversitesi Veteriner Fakültesi, Biyoistatistik Anabilim Dalı, Ankara
[4]Erciyes Üniversitesi Sağlık Bilimleri Enstitüsü, Kayseri
[5]Erciyes Üniversitesi Tıp Fakültesi, Biyoistatistik Anabilim Dalı, Kayseri
[6]Erciyes Üniversitesi Veteriner Fakültesi, Genetik Anabilim Dalı, Kayseri

Amaç: Güç doğum, anne ve yavru kaynaklı olup, gerek anne ve yavrunun sağlığı bakımından gerekse ekonomik açıdan önemli bir problemdir. Bu nedenle, güç doğum özellikle süt verimi yüksek olan işletmelerdeönemli bir sorun olarak ortaya çıkmaktadır. Literatürlere göre güç doğum üzerinde etkili faktörlerin başında hayvanın ırkı, cinsiyeti, beslenme, buzağılama mevsimi, yaş değişkenleri ve yönetimsel faktörler gelmektedir. Güç doğumla ilişkili nedenler çoğunlukla bilinmekle beraber özellikle büyük işletmelerde klasik ilişki analizleri ile tespit edilemeyen örüntüleri bulmak önem arz etmektedir. Bu çalışmada, güç doğum üzerinde etkili olan bazı faktörlerin birliktelik kuralları algoritmalarından biri olan FrequentPatternGrowth(FP Growth) algoritması ile belirlenmesi amaçlanmıştır.

Gereç ve Yöntem: Birliktelik kuralları alışılmış istatistiksel yöntemler yerine kullanılabilecek analitik bir yöntemdir. Birliktelik kuralları ile tüketicilerin satın aldığı ürünler arasındaki ilişkilerin ortaya çıkarılarak satın alma eğilimleri belirlenmektedir. FP Growth ve Apriori algoritmaları birliktelik kuralları algoritmaları arasında en sık kullanılan algoritmalardır. Algoritmanın performans açısından diğer algoritmalardan daha işlevsel olduğu daha önceki çalışmalarda gösterilmiştir. FP Growth Algoritmasını diğer algoritmalardan ayıran en temel karakteristikleri yaygın nesne kümelerini dolaysız olarak test edebilmesi ve büyük veri kümelerinde kısa sürede sonuç verebilmesidir. Bu çalışmada güç doğum, metritis, mastitis, ayak hastalığı, buzağılama mevsimi ve buzağılama yaşı değişkenlerinden oluşan 108 baş Holstein ineğe ait veriler kullanılmıştır. Minimum destek değeri ve minimum güven değeri 0.03 olarak seçilmiştir. Veriler R 4.2.1 (https://cran.r-project.org) yazılımında ve "rCBA" ve "arules" paketleri kullanılarak analiz edilmiştir.

Bulgular: Uygulanan FP Growth algoritması ile 176 kural 0.03 saniyede oluşturulmuştur. Elde edilen sonuçlara göre Holstein ineklerde 3-4 yaşlı olma ve normal doğumun birlikte görülme olasılığı %10 iken, 3-4 yaşlı Holstein ineklerinde normal doğum görülme olasılığı %65'tir. 9-10 yaşlı ve mastitisli olma durumu ile güç doğumun birlikte görülme olasılığı %11 iken; 9-10 yaşlı ve mastitisli ineklerde güç doğumun da görülme olasılığı %71 olarak belirlenmiştir.

Sonuç: FP Growth algoritması ile elde edilen sonuçların Apriori algoritması ile elde edilen sonuçlara benzerlik gösterdiği görülmüştür. Ayrıca, birliktelik kurallarının veteriner hekimlik alanında işlevsel bir yöntem olduğu sonucuna varılmıştır.

Anahtar Kelimeler: Holstein, güç doğum, FP Growth, apriori, birliktelik kuralları

SB – 6

Mamografi Görüntülerinde Farklı Ön İşleme Kombinasyonlarının Sınıflandırma Performansları Üzerine Etkisi

[1]Hanife AVCI, [2]Jale KARAKAYA

[1]Hacettepe Üniversitesi Tıp Fakültesi, Biyoistatistik Anabilim Dalı, Ankara
[2]Hacettepe Üniversitesi Tıp Fakültesi, Biyoistatistik Anabilim Dalı, Ankara

Amaç: Bir görüntü içerisinde çevresel faktörler, nem, elektromanyetik dalgalar ya da görüntü yakalama aygıtlarından kaynaklanan görüntüde istenmeyen parazitler oluşur. Görüntüler üzerinde ortaya çıkan istenmeyen işaretler (gürültü, noise), bölütleme ve özellik çıkarımı adımlarından çıkarılacak sonuçları etkiler. Bu nedenle görüntülerin bu gürültülerden temizlenmesi gerekmektedir. Alanyazında görüntülerdeki gürültülerin yok edilmesi ve görüntülerin iyileştirilmesi için farklı ön işleme yöntemleri önerilmiştir. Bu nedenle çalışmanın amacı, görüntü ön işleme yöntemleri ile, mamografi görüntüleri üzerinde netleştirme, belirli ayrıntıları ortaya çıkarma, görüntüyü yumuşatma ve kenar keskinleştirme gibi işlemler gerçekleştirmek ve ön işleme kombinasyonlarının sınıflandırma performansına katkısını incelemektedir.

Yöntem: Mamografi, meme kanseri taraması ve anormalliklerin tespiti için en etkili görüntüleme yöntemidir. Mamografi görüntülerinin okunmasının zor olması ve görüntü kalitesinin düşük olması nedeniyle bilgisayar destekli tanı (CAD) modelleri ile ilgili çalışmalar son yıllarda artmıştır. Bu çalışmada, mamografi görüntülerinde şüpheli alanların tespiti için bilgisayar destekli tanı sistemleri kullanılmıştır. Bu çalışmadaki CAD sistemi beş ana adımdan oluşmaktadır: ön işleme, bölütleme, ilgi bölgesi (ROI), özellik çıkarma ve sınıflandırma. Ön işleme algoritmaları, bölütleme ve sınıflandırma yöntemlerinin doğruluğunu önemli ölçüde etkiler. Bu çalışmada, mini-MIAS veri tabanından elde edilen 322 mamografi görüntüsü kullanılarak kontrast sınırlı uyarlamalı histogram eşitleme (CLAHE), Medyan filtresi ve Keskinliği Azaltma maskeleme gibi farklı ön işleme yöntemlerinin farklı kombinasyonlarının sınıflandırma performansı üzerindeki etkisi değerlendirilmiştir. ROI örnekleri, k-ortalamalar kümeleme bölütleme yöntemiyle elde edilmiştir. Gri seviye eş oluşum matrisi (GLCM) ve Gri seviye çalışma uzunluğu matrisi (GRLM) teknikleri kullanılarak ROI örneklerinden özellikler çıkarılmıştır. Korelasyon matrisi ile öznitelik seçimi yapıldıktan sonra bu özniteliklerin sınıflandırma başarısına etkileri 6 farklı klasik veri madenciliği yöntemi ile incelenmiştir. Görüntü işleme sonucu elde edilen sayısal veriler sınıflandırma yöntemlerinde giriş değişkeni olarak kullanılmıştır. Bu çalışmada hesaplanan özellikler ile ROI örnekleri normal-anormal doku olarak sınıflandırılmıştır. Sınıflandırma yöntemlerinin performansları, doğruluk, duyarlılık ve AUC gibi ölçüler kullanılarak karşılaştırılmıştır.

Bulgular: Çalışmada yapılan sınıflandırmada CLAHE algoritması ön işleme yöntemi olarak tek başına kullanıldığında diğer ön işleme kombinasyonlarına göre daha düşük sınıflandırma performans değerleri elde edilmiştir. CLAHE algoritmasına Medyan filtresi ve Unsharp maskeleme algoritmaları eklendiğinde, sınıflandırma yöntemlerinin performansı artmıştır. Sınıflandırma başarısı açısından en iyi algoritmaya baktığımızda Destek Vektör Makineleri (DVM), Rastgele Orman (RO), Yapay Sinir Ağları (YSA), Naive Bayes (NB) ve Karar Ağaçları (KA) yöntemleri k-En Yakın Komşuluk (k-NN) algoritmasına göre her iki sınıflandırmada da yüksek sınıflandırma sonuçları vermiştir.

Sonuç: Üçlü kombinasyonun ikili ve tekli kombinasyonlardan daha iyi performans göstermediği gözlemlenmiştir. Sınıflandırma yöntemleri açısından en yüksek performansı DVM, RO ve YSA göstermiştir.

Anahtar Kelimeler: Mamografi görüntüleri, ön işleme algoritmaları, sınıflandırma performansı.

Bu çalışma Hacettepe Üniversitesi Bilimsel Araştırma Projeleri Koordinasyon Birimi tarafından desteklenmiştir. Hızlı Destek Projesi (THD-2020-18735)

SB – 7

SHINY İle Zaman Serileri Analizi İçin İnteraktif Arayüz Tasarımı

[1]Esra GÜLTÜRK, [2]Abdulvahap PINAR, [3]Ahmet Kadir ARSLAN, [4]Pınar YILMAZ

[1]Cumhuriyet Üniversitesi Tıp Fakültesi Biyoistatistik Anabilim Dalı, Sivas
[2]İnönü Üniversitesi Tıp Fakültesi Biyoistatistik ve Tıbbi Bilişim Anabilim Dalı, Malatya
[3]Cumhuriyet Üniversitesi Tıp Fakültesi Biyoistatistik Anabilim Dalı, Sivas
[4]İnönü Üniversitesi Tıp Fakültesi Tıbbi Biyoloji ve Genetik Anabilim Dalı, Malatya

Amaç: Günümüzde ileriye dönük güvenilir karar alabilmek için geleceği doğru tahmin etmek son derece önemlidir. Zaman serileri analizi zamanla değişen gözlemlerin geçmişteki değerleri kullanılarak gelecekte alacağı değerleri tahmin edebilmek için kullanılan yöntemlerdir. Bu çalışmada R programlama dili kullanılarak, zaman serileri analizinde kullanılabilecek bir program geliştirmek ve ayrıca bu programa web üzerinden erişebilen bir arayüz oluşturmak amaçlanmıştır.

Gereç ve Yöntem: Bu çalışmada shiny paketi kullanılarak bir arayüz oluşturulmuştur. Geliştirilen arayüz xls/.xlsx, .arff, .sav ve .csv/. txt uzantılı veri dosyalarını desteklemektedir. Tek değişkenli zaman serileri yöntemleri ile ilgili "Karma otoregresif-hareketli ortalamalar (ARMA), Filtre (süzme) modelleri, Düzleştirme yöntemleri, Zaman serisi ayrıştırma yöntemleri ve yaygın olarak kullanılan zaman serisi teknikleri için geçmiş gözlemlerin ağırlıklandırılmasından oluşan kısa dönemli dalgalanmaları önleyen istatistiksel tekniklerdir. Bu tekniklerin temelinde bazı düzgün eğrilerden rasgele ayrılmaları gösteren zaman serisine ait geçmiş dönemlerde yaşanan dalgalanmalar gelecek için bir öngörü niteliğinde olacaktır. Buna göre en iyi öngörüye karar vermek için geliştirilen yazılım örnek veri setleri kullanılarak test edilmiştir.

Bulgular: En uygun modeli belirlemek için zaman yolu grafiği çizilmiştir. Durağanlığı bozan unsurlar var ise (trend ve mevsimsellik gibi) durağanlaştırma yapılmıştır. Veri setinin normal dağılıma uygunluğu test edilmiştir. En son uygun öngörüye göre model oluşturulmuştur.

Sonuç: Yazılım, sağlık araştırmalarında, ,ekonomide, tarım gibi birçok alanda kolaylıkla tek değişkenli zaman serisi analizini uygulayabilmek için geliştirilmiştir. İlerleyen çalışmalarda ilgili yazılımın çok değişkenli zaman serilerini analiz etme imkanı sağlayan sürümü de geliştirilecektir.

Anahtar Kelimeler: Arayüz, Shiny, Zaman serileri

SB – 8

Kardiyoverter Defibrilatör (Icd) İmplantasyonu Yapılan Hastalarda Yaşam Durumu Tahmin Modellemesi

[1]Hatice AKTAŞ GÖKÇE, [2]Amine BAYRAKLI, [3]Emre YAŞAR

[1]TC Sağlık Bakanlığı
[2]Karadeniz Teknik Üniversitesi, Sağlık Bilimleri Enstitüsü, Biyoistatistik ve Tıp Bilişimi Anabilim Dalı
[3]Ankara Yıldırım Beyazıt Üniversitesi Tıp Fakültesi, Biyoistatistik Ana Bilim Dalı, Ankara, Türkiye

Amaç: Geçici kalp pili implante edilen hastaların klinik özellikleri, komorbiditeleri, altta yatan hastalıkları saptanıp, has¬tane içi mortalitelerinin tahminlemesi amaçlanmıştır.

Gereç ve Yöntem: Çalışmaya farklı sebeplerle kardiyoverter defibrilatör (ICD) implantasyonu işlemi yapılan 2019-2021 tarihleri arasında ülkemizde bulunan 2. Ve 3. Basamak sağlık tesislerinde takip edilen 25233 hasta dahil edilmiştir. Çalışmada Lojistik Regresyon yöntemi ile ölüm durumunu tahminleme amaçlanmıştır.

Bulgular: Çalışmaya geriye dönük olarak taranan 18768 erkek birey, 6464 kadın birey katılım sağlamıştır. Bireylerin tamamı kardiyoverter defibrilatör (ICD) implantasyonu işlemi gerçekleşmiş hastalar olup, kadınların yaş ortalaması 63.22±14.83, erkeklerin yaş ortalaması 62.72±12.93 olarak saptanmıştır. En sık kalp pili işlemi Implantable Cardioverter Defibrillator takılması, tek elektrot olarak belirlenmiştir. Ölüm durumunu bağımlı değişken olarak belirlediğimizde modele eklediğimiz yaş, cinsiyet, işlem tipi, ölüm zamanı (ölüm zamanı-işlem zamanı) gibi bağımsız değişkenler ile bağımlı değişkenin %78.7'sinin lojistik regresyon yöntemi ile açıklandığı saptandı. Çalışma Phyton dilinde NumPy, Matplotlib kütüphaneleri ve Scikit-learn paketleri de kullanılarak ilave bağımsız değişkenler ile lojistik regresyon yöntemi ve Karar Ağaçları yöntemi uygulanarak tekrar değerlendirilip karşılaştırılacaktır.

Sonuç: Kardiyoverter defibrilatör (ICD) implantasyonu işlemi gerçekleşmiş bireylerin yaşa, cinsiyete, işlem tipine bağlı olarak yaşama süreleri farklılık göstermektedir. Farklılığın sonucu ise ölümcül seyirdir.

Anahtar Kelimeler: Tahmin modelleme, yaşam durumu, kalp pili

SB – 9

Sağkalım Analizi için Veri Simülasyonu

[1]İrem KAR, [2]Erdal COŞGUN, [3]Atilla Halil ELHAN

[1]Ankara Üniversitesi Tıp Fakültesi, Biyoistatistik Anabilim Dalı, Ankara
[2]GenomicsTeam, Microsoft Research&AI, Redmond, WA, USA
[3]Ankara Üniversitesi Tıp Fakültesi, Biyoistatistik Anabilim Dalı, Ankara

Amaç: Sağkalım analizi çalışmalarında gerçek veriye ulaşmak oldukça zordur. Bu nedenle, veri analistleri simülasyon çalışmalarına yönelmişlerdir. Bu çalışmada, sağkalım analizinde derin öğrenme için veri simüle ederken dikkat edilmesi gereken parametreleri vurgulamak ve bu parametrelerin DeepSurv yönteminin tahminleme başarısına etkisini ortaya koymak amaçlanmıştır.

Gereç ve Yöntem: Burada, sağkalım analizleri için veri simülasyonunda en çok tercih edilen R programlama dilindeki "coxed" paketinde üzerinde durulmuş olup, simüle edilmiş veri setleri kullanılarak bir benzetim çalışması gerçekleştirilmiştir. Bu kapsamda örneklem büyüklüğünün, sansür oranının, bağımsız değişkenlere ilişkin katsayıların ve çeşitli istatistiksel dağılımlardan bağımsız değişkenlerin simüle edildiği pek çok senaryo oluşturulmuştur. Oluşturulan her senaryo ise 100 kez tekrarlanmıştır. Ayrıca, simüle verilere ek olarak gerçek veri üzerinde de DeepSurv yöntemi uygulanmıştır. Yöntemlerin karşılaştırılmasında Harrell'in C-indeks değeri kullanılmıştır. Verilerin türetilmesi R programlama dili (ver. 3.6.3) ile derin öğrenmeye ilişkin analizler ise Python programlama dilinde gerçekleştirilmiştir.

Bulgular: Yapılan benzetim çalışması sonucunda, örneklem büyüklüğünün, sansür oranının, bağımsız değişkenlere ilişkin katsayıların ve bağımsız değişkenlerin simülasyonunda kullanılan istatistiksel dağılımların tahminleme başarısını etkilediği bulunmuştur. Örneklem büyüklüğü arttıkça yöntemlerin performansı artmıştır. Sansür oranının artması tahminleme başarısını azaltmıştır. Bağımsız değişkenlere ilişkin katsayıların artması yöntemin sınıflama performansını arttırmıştır. Bağımsız değişken katsayılarının "coxed" paketindeki mevcut haliyle kullanıldığı durumda ise 0,50-0,60 aralığında bir C-indeks değeri elde edilmiştir.

Sonuç: Bu çalışmada, sağkalım verisi simülasyonunda kullanılan parametrelerin önemi ve tahminleme başarısına etkisi açık bir şekilde ortaya konmuştur. Sonuç olarak, sağkalım analizinde derin öğrenme için kullanılan DeepSurv algoritmasında çok sayıda hiperparametre kullanılması model başarısını arttırmak için avantaj olmasına rağmen, veri simüle ederken göz ardı edilen ya da yanlış kullanılan parametrelerin tahminleme başarısı üzerindeki etkisinin önemi anlaşılmıştır.

Anahtar Kelimeler: DeepSurv, harrell'in C-indeksi, sağkalım analizi

SB – 10

R Shiny İle Zaman Serisi Analizzi İçin İnteraktif Arayüz Tasarımı

[1]Esra GÜLTÜRK, [2]Abdulvahap PINAR, [3]Ahmet Kadir ARSLAN, [4]Pınar YILMAZ

[1]Cumhuriyet Üniversitesi Tıp Fakültesi, Biyoistatistik Ana Bilim Dalı, Sivas
[2]İnönü Üniversitesi Tıp Fakültesi Biyoistatistik ve Tıp Bilişimi Malatya Türkiye
[3]İnönü Üniversitesi Tıp Fakültesi Biyoistatistik ve Tıp Bilişimi Malatya Türkiye
[4]İnönü Üniversitesi Tıp Fakültesi Tıbbi Biyoloji ve Genetik Anabilim Dalı, Malatya

Amaç: Günümüzde ileriye dönük güvenilir karar alabilmek için geleceği doğru tahmin etmek son derece önemlidir. Zaman serileri analizi zamanla değişen gözlemlerin geçmişteki değerleri kullanılarak gelecekte alacağı değerleri tahmin edebilmek için kullanılan yöntemlerdir. Bu çalışma R programlama dili kullanılarak, zaman serileri analizinde kullanılabilecek bir program geliştirmek ve ayrıca bu programa web üzerinden erişebilen bir arayüz oluşturmak amaçlanmıştır.

Gereç ve Yöntem: Bu Çalışmada R-Studio ile shiny paketi kullanılarak bir arayüz oluşturulmuştur. Geliştirilen arayüz xls/.xlsx, .arff, .sav ve .csv/.txt uzantılı veri dosyalarını desteklemektedir. En çok kullanılan öngörülerden tek değişkenli zaman serileri yöntemleri ile ilgili "Karma otoregresif-hareketli ortalamalar (ARMA), Filtre (süzme) modelleri, Düzleştirme yöntemleri, Zaman serisi ayrıştırma yöntemleri ve yaygın olarak kullanılan zaman serisi teknikleri için geçmiş gözlemlerin ağırlıklandırılmasından oluşan kısa dönemli dalgalanmaları önleyen istatistiksel metotlardır. Bu metotların temelinde bazı düzgün eğrilerden tesadüfi ayrılmaları gösteren zaman serisine ait geçmiş dönemlerde yaşanan dalgalanmalar gelecek için bir öngörü niteliğinde olacaktır. Buna göre en iyi öngörüye karar vermek için hazırlanan program ve arayüz örnek veri setleri kullanılarak test edilmiştir.

Bulgular: En uygun modeli belirlemek için zaman yolu grafiği çizilmiştir. Durağanlığı bozan unsurlar var ise (trend ve mevsimsellik gibi) durağanlaştırma yapılmıştır. Veri setinin normal olup olmadığı test edilmiştir. En son uygun öngörüye göre model oluşturulmuştur.

Sonuç: Yazılım, sağlık araştırmalarında, ,ekonomide, tarım gibi birçok alanda kolaylıkla tek değişkenli zaman serisi analizini uygulayabilmek için geliştirilmiştir. İlerleyen çalışmalarda çok değişkenli zaman serileri ile ilgili yazılım geliştirilecektir.

Anahtar Kelimeler: Arayüz, Shiny, Zaman serileri

SB – 11

Periton Diyaliz Hastalarında Sağkalım ve Uzunlamasına Verilerin Birlikte Modellemesi: Kişiye Özgü Dinamik Risk Kestirimleri

[1]Merve Başol GÖKSÜLÜK, [2]Dinçer GÖKSÜLÜK, [3]Ergun KARAAĞAOĞLU

[1]Erciyes Üniversitesi Tıp Fakültesi, Biyoistatistik ve Tıp Bilişimi ABD, Kayseri, Türkiye
[2]Erciyes Üniversitesi Tıp Fakültesi, Biyoistatistik ve Tıp Bilişimi ABD, Kayseri, Türkiye
[3]Erciyes Üniversitesi Tıp Fakültesi, Biyoistatistik ve Tıp Bilişimi ABD, Kayseri, Türkiye

Amaç: Periton diyaliz (PD) tedavisi, geç dönem böbrek yetmezliği olan hastalarda sıklıkla uygulanmaktadır ve nefrolji alanında PD hastalarının sağkalım durumları güncel olarak çalışılmaya devam etmektedir. Hastalığın seyrini değerlendirebilmek için böbrek fonksiyonu ile ilişkili olan biyobelirteçler ve çeşitli biyokimsayasal parametreler izlem süresinde tekrarlı olarak ölçülmektedir. Literatürdeki çalışmalar incelendiğinde sağkalım analizlerinde tekrarlı alınan bu ölçümlerin başlangıç ve/veya izlem süresi içindeki ortalama değerlerinin çoğunlukla dikkate alındığı görülmektedir. Bu yaklaşım yanlış olmamakla birlikte tekrarlı ölçümlerin zaman içerisindeki değişimini gözardı etmekte ve yanlı sonuçlar elde edilmesine yol açabilmektedir. Bu çalışma, PD hastalarında çeşitli biyobelirteçlerin mortalite ile olan ilişkisini tekrarlı ölçüm yapısını dikkate alarak ortaya koymayı amaçlamaktadır.

Gereç ve Yöntem: Bu çalışmada 1995 – 2007 yılları arasında Erciyes Üniversitesi Nefroloji bölümünde periton diyaliz (PD) tedavisine başlayan 511 hastaya ait veriler geriye dönük olarak değerlendirilmiştir [1]. Hastalar, PD başlangıcından itibaren ölüm, böbrek nakli, hemodiyalize geçiş, izlemden çıkma, PD tedavisinin sonlandırılması gibi sebeplerden herhangi birisi gözleninceye kadar takip edilmiştir. Mortalite yanıt değişkeni PD sürecinde gözlenen tüm sebeplerden kaynaklı ölümdür. Biyobelirteçlerin zaman içerisindeki değişiminin sağkalım üzerindeki etkisi birleşik modelleme (joint modelin) yaklaşımı kullanılarak değerlendirilmiştir [2, 3]. Sağkalım modeli için Cox regresyon yöntemi ve uzunlamasına veriler için karma etkili doğrusal modelden yararlanılmıştır. Çalışmada tekrarlı ölçüm olarak serum albumin değerleri dikkate alınmıştır.

Bulgular: Elde edilen sonuçlara göre ortalama serum albumin değerlerinin mortalite ile ilişki olmadığı görülmüştür. Ancak, tekrarlı olarak ölçülen serum albumin değerlerinin mortalite ile anlamlı ve ters yönlü bir ilişkisi olduğu gözlenmiştir (HR: 2.43, %95 G.A.: 1.48 – 4.16). Peritonit hızı, hemodiyaliz geçmişi ve toplam komorbit hastalık sayısı ile mortalite arasında pozitif yönlü bir ilişkinin olduğu ve ölüm oranlarının sırasıyla 1.74, 3.21 ve 1.41 kat artış gösterdiği belirlenmiştir. Sağkalım durumları hastalar arasında anlamlı bir değişkenlik göstermektedir. Kişiselleştirilmiş dinamik risk kestirimleri dikkate alındığında Cox regresyon modelinden elde edilen genel sağkalım kestirimlerinin tüm hastaları temsil edemediği, hastaların karama etkili modelden elde edilen bireysel sağkalım kestirimlerinin genel tahmin değerlerinden önemli ölçüde farklılaştığı görülmüştür.

Sonuç: Serum albumin değerlerinin sağkalım ile olan ilişkisi tekrarlı ölçümler üzerinden uygun istatistiksel modelleme teknikleri ile değerlendirildiğinde daha doğru ve yansız kestirimler elde edilmektedir. Kişi bazlı elde edilen dinamik risk kestirimleri hastalığın seyrini tahmin etmede ve hastaların takibinde önemli bir rol oynamaktadır. Ayrıca, izlem süresince yeni veriler elde edildikçe risk kestirimleri dinamik olarak güncellenebilmekte ve kişiye özel kararlar alınabilmektedir.

Anahtar Kelimeler: Mortalite, kişiselleştirilmiş tıp, sağkalım, birleşik model, periton diyaliz

SB – 12

Boylamsal ve Yaşam Verilerinin Parametrik Bileşik Modellemesinde Parametre Tahmin Yaklaşımlarının Karşılaştırılması

[1]Elif YILDIRIM, [2]Duru KARASOY

[1]Konya Teknik Üniversitesi, Rektörlük, Kalite Koordinatörlüğü, Konya
[2]Hacettepe Üniversitesi, Fen Fakültesi, İstatistik Bölümü, Ankara

Amaç: Boylamsal ve yaşam verilerinin farklı yöntemler ile modellenmesi üzerine yapılan çalışmaların sayısı son yirmi yılda artmıştır. Farklı modellemelerin yanı sıra parametre tahminlerini daha yansız olarak elde edebilmek için farklı parametre tahmin yöntemlerinin de geliştirildiği görülmektedir. Bileşik modelleme ve parametre tahmin yaklaşımları üzerine yapılan çalışmaların çoğu boylamsal veriler için doğrusal karma etkili model ve yaşam verilerinin Cox regresyon modelinin paylaşılmış rastgele etkiler ile birleştirilmesinden elde edilmektedir. Yaşam verilerinin orantılı tehlikeler varsayımını sağlamadığı durumlarda ise parametrik bileşik modellemeler kullanılabilmektedir. Bu çalışmanın amacı, yaşam verilerinin orantılı tehlikeler varsayımını sağlamadığı durumlarda da kullanılabilen parametrik bileşik modellemede daha yansız tahminler elde etmek için parametre tahmin yöntemlerini farklı simülasyon senaryoları ile incelemek ve en uygun parametre tahmin yöntemini belirlemektir.

Gereç ve Yöntem: Parametrik bileşik modellemede parametre tahmin yaklaşımlarını incelemek amacıyla, yaşam verileri için Weibull hızlandırılmış başarısızlık süresi (AFT) modeli ile boylamsal veriler için doğrusal karma etkili modelin parametrik bileşik modellemesi ele alınmıştır. Elde edilen modelin parametrelerini elde etmek için kullanılan standart Gauss Hermite ve Uyarlanmış (adaptive) Gauss Hermite yaklaşımları farklı örneklem büyüklüklerinde simülasyon çalışmasıyla incelenmiş ve iki yöntem karşılaştırılmıştır.

Bulgular: Çalışmada, Weibull AFT modeli ve doğrusal karma etkili modelin rastgele etkiler ile birleştirilmesinden elde edilen bileşik modelin parametre tahminlerini elde etmek için kullanılan ve integrasyon yaklaşımı olan standart Gauss Hermite ve uyarlanmış Gauss Hermite yaklaşımlarının performanslarını karşılaştırmak için Monte Carlo simülasyonu çalıştırılmıştır. Simülasyon kodları R programında yazılmış olup, model tahminlerini elde etmek için "JM" paketi kullanılmıştır. Çalışmada durdurma oranı %50 olarak belirlenmiş ve n=50, 100 ve 500 için iki parametre tahmin yaklaşımının performans karşılaştırması yapılmıştır. Boylamsal alt modeldeki intercept, zaman, grup*zaman ve grup değişkenlerinin gerçek değerleri sırasıyla 2, 1.5, 0.5 ve 0.8 olarak; yaşam alt modeldeki grup değişkeninin gerçek değeri -0.3 ve birleştirme katsayısı 1.6 olarak belirlenmiştir. Aynı zamanda boylamsal gözlemin zaman boyutu 0, 2, 4, 6 ve 8 olmak üzere 5 boyuttan oluşacak şekilde üretilmiştir. Yöntemlerin performanslarının karşılaştırmak için parametre tahminlerinin yan ve hata kareler ortalamalarının karekök (RMSE) değerleri kullanılmıştır.

Sonuç: Parametrik bileşik modellemede parametre yaklaşım yöntemlerini incelemek için yapılan simülasyon sonuçlarına göre; örneklem genişliği küçük olduğunda uyarlanmış Gauss Hermite yaklaşımının RMSE ve yan değerlerinin standart Gauss Hermite yaklaşımına göre daha küçük olduğu gözlemlenirken, örneklem genişliği arttıkça Gauss Hermite yaklaşımının, uyarlanmış Gauss Hermite yaklaşımına göre daha az yanlı tahminler verdiği görülmüştür.

Anahtar Kelimeler: Parametrik Bileşik Model, Parametre Tahmin Yaklaşımı, Gauss Hermite

SB – 13

Önerilen Derin Öğrenme Mimarisi ile MR Görüntülerine Dayalı Alzheimer Hastalığının Sınıflandırılması

[1]Mehmet KIVRAK, [2]Cemil ÇOLAK

[1]Recep Tayyip Erdogan Üniversitesi Biyoistatistik ve Tıp Bilişimi
[2]İnönü Üniversitesi Tıp Fakültesi Biyoistatistik ve Tıbbi Bilişim Anabilim Dalı, Malatya

Amaç: Bu çalışma, keras kütüphanesi ve evrişimsel sinir ağları (CNN'ler) kullanarak derin öğrenme yaklaşımı ile MRI görüntülerinden oluşan çok hafif/hafif/orta/demanssız hastaları yüksek doğrulukla sınıflandırmayı amaçlamaktadır.

Gereç ve Yöntem: İncelenen görüntü veri seti, 179 hafif demanslı, 12 orta demans, 640 demanssız ve 448 çok hafif demanslı olmak üzere toplam 1279 MRI hasta taramasından oluşmaktadır. Araştırmada geriye dönük vaka-kontrol tasarımı kullanılmıştır. Keras kitaplığı (TensorFlow), kullanıcının evrişimsel sinir ağı modellerini daha basit bir şekilde oluşturmasını sağlar ve kullanıcıyı bu düşük seviyeli kitaplıkların karmaşıklığından kurtarır. GPU donanımı olarak CUDA destekli yapay zeka ve analitik taleplerini karşılamak üzere geliştirilen NVIDIA DGX™ Sistemleri kullanıldı. Modellerin sınıflandırma başarımı çeşitli metrikler ile değerlendirilmiştir.

Bulgular: CNNs modelini geliştirmek için kullanılan veri seti, eğitim (1024) ve test (255) veri setlerine ayrılmıştır. Modelin eğitim aşamasındaki en yüksek doğruluk oranı 0.96 olarak tahmin edilmiştir. Test verilerinin doğruluk, duyarlılık, özgüllük, pozitif ve negatif tahmin değerlerine göre en yüksek sınıflandırma performans oranı 0.93 ile pozitif tahmin değer olmuştur.

Sonuç: Deneysel sonuçlar, MRI görüntülerine dayalı Alzheimer sınıflandırılmasında derin öğrenme yöntemlerinin (CNN'ler) oldukça başarılı olduğunu ve benzer çalışmaların diğer hastalıklar için de yapılabileceğini göstermiştir.

Anahtar Kelimeler: Alzheimer, Görüntü İşleme, Derin Öğrenme, Evrişimsel Sinir Ağları.

SB – 14

Yüksek Boyutlu Verilerde Eksik Veri Değer Atama Yöntemlerinin Tahmin Performanslarının Aşırı Öğrenme Makineleriyle İncelenmesi

[1]Buğra VAROL, [2]İmran ÖMÜRLÜ, [3]Mevlüt TÜRE

[1]Adnan Menderes Üniversitesi, Sağlık Bilimleri Enstitüsü, Biyoistatistik Anabilim Dalı, Aydın
[2]Adnan Menderes Üniversitesi, Tıp Fakültesi, Biyoistatistik Anabilim Dalı, Aydın
[3]Adnan Menderes Üniversitesi, Tıp Fakültesi, Biyoistatistik Anabilim Dalı, Aydın

Amaç: Eksik değerli veri setleri yüksek boyutlu verilerin analizi ve sınıflandırılması açısından temel bir sorundur. Bu çalışmanın amacı, yüksek boyutlu verilerde farklı eksik veri değer atama yöntemlerinin tahmin performanslarının türetilmiş veriler üzerinde aşırı öğrenme makinalarıyla (ELM) karşılaştırılmasıdır.

Gereç ve Yöntem: Çalışmada, öncelikle n=150 gözlemden oluşan iki kategorili bağımlı değişken ve karma ilişki yapılı p=500 değişkenden oluşan rastgele veriler türetilmiştir. Daha sonra bağımsız değişkenlerden %10, %20, %30, %40, %50 oranında eksik değerler oluşturulmuştur. Eksik veri değer atama yöntemlerinden; ortalama, medyan, rastgele, k en yakın komşu (KNN), sınıflandırma ve regresyon ağaçları (CART), rasgele orman (RF) değer atama yöntemlerinin yanı sıra yüksek boyutlu veriler için geliştirilen düzenlileştirilmiş regresyonun doğrudan kullanımı (DURR) ve düzenlileştirilmiş regresyonun dolaylı kullanımı (IURR) yöntemleri ile eksik değerler tahmin edilmiştir. 1000 döngü ile yapılan simülasyon sonunda yöntemlerin ELM ile sınıflandırma skorları elde edilmiş ve bu skorların referansa olan yakınlığına göre eksik değer tahmin performansları değerlendirilmiştir.

Bulgular: Simülasyon bulguları incelendiğinde düşük eksik oranlarında tüm yöntemler arasından CART, KNN, RF, DURR ve IURR yöntemlerinin birbirine ve referansa yakın iyi performans gösteren yöntemler olduğu belirlenmiş; bunun yanı sıra CART, KNN ve RF yöntemlerinin eksik oranı arttıkça performanslarının genel olarak düşüş eğiliminde olduğu; DURR ve IURR yöntemlerinin ise yüksek eksik oranlarında bile istikrarlı bir şekilde iyi performans göstermeye devam ettiği gözlenmiştir.

Sonuç: Eksik değer oranı arttıkça ortalama, medyan, rastgele, CART, RF ve KNN yöntemlerinin tahmin performansında düşüşler gözlenirken, DURR ve IURR yöntemlerinin eksik değer oranından etkilenmediği belirlenmiştir. Eksik değer oranındaki artışa karşı daha kararlı sonuçlar vermeleri nedeniyle DURR ve IURR yöntemleri, yüksek boyutlu verilerde daha avantajlıdır.

Anahtar Kelimeler: Aşırı öğrenme makineleri, Eksik veri, Değer atama, Sınıflandırma, Simülasyon.

SB – 15

Biocrates Mxp® Quant 500 Kiti Kullanılarak Elde Edilen Metabolomik Verilerinin Laboratuvarlar Arası Karşılaştırması

[1]Gözde ERTÜRK ZARARSIZ, [2]Gözde ERTÜRK ZARARSIZ, [3]Alexander CECIL, [4]Jutta LINTELMANN, [5]Gernot POSCHET, [6]Jennifer KIRWAN, [7]Sven SCHUCHARDT, [8]Xue L. GUAN, [9]Daisuke SAIGUSA, [10]David WISHART, [11]Jiamin ZHENG, [12]Rupasri MANDAL, [13]Lisa ST. JOHN-WILLIAMS, [14]Kendra ADAMS, [15]J. Will THOMPSON, [16]Michael P. SNYDER, [17]Kevin CONTREPOIS, [18]S Songjie CHEN, [19]Nadia ASHRAFI, [20]Sumeyya AKYOL, [21]Ali YILMAZ, [22]Stewart GRAHAM, [23]Thomas M. O'CONNELL, [24-25]Karel KALECKÝ, [26]T. Hai PHAM, [27]Therese KOAL, [28]Teodoro BOTTIGLIERI, [29]Teodoro BOTTIGLIERI, [30]Jerzy ADAMSKI, [31]Gabi KASTENMÜLLER

[1]Erciyes Üniversitesi Tıp Fakültesi, Biyoistatistik ve Tıp Bilişimi ABD, Kayseri, Türkiye
[2]Institute of Computational Biology, Helmholtz Zentrum München, Almanya,
[3]Metabolomics and Proteomics Core, Helmholtz Zentrum München, Almanya
[4]Metabolomics and Proteomics Core, Helmholtz Zentrum München, Almanya
[5]Metabolomics Core Technology Platform, Heidelberg University, Almanya
[6]BIH Metabolomics Platform, Berlin Institute of Health@Charité, Almanya
[7]Fraunhofer-Institute for Toxicology and Experimental Medicine, Hannover, Almanya
[8]Lee Kong Chian School of Medicine, Nanyang Technological University, Singapur
[9]Laboratory of Biomedical and Analytical Sciences, Faculty of Pharma-Science, Teikyo University, Japonya
[10]Departments of Biological Sciences and Computing Science
[11]University of Alberta, Alberta, Kanada
[12]University of Alberta, Alberta, Kanada
[13]Duke University, Durham, NC. Amerika
[14]Duke University, Durham, NC. Amerika
[15]Duke University, Durham, NC. Amerika
[16]Stanford University, Stanford, CA, Amerika
[17]Stanford University, Stanford, CA, Amerika
[18]Stanford University, Stanford, CA, Amerika
[19]Beaumont Health System, Royal Oak, MI, Amerika
[20]Beaumont Health System, Royal Oak, MI, Amerika
[21]Beaumont Health System, Royal Oak, MI, Amerika
[22]Beaumont Health System, Royal Oak, MI, Amerika
[23]Indiana University, Indianapolis, IN, Amerika
[24]Center of Metabolomics, Institute of Metabolic Disease, Baylor Scott & White Research Institute, Dallas, TX, Amerika
[25]Institute of Biomedical Studies, Baylor University, Waco, TX, Amerika
[26]Biocrates life sciences ag, Innsbruck, Avustralya
[27]Biocrates life sciences ag, Innsbruck, Avustralya
[28]Center of Metabolomics, Institute of Metabolic Disease, Baylor Scott & White Research Institute, Dallas, TX, Amerika
[29]Institute of Biomedical Studies, Baylor University, Waco, TX, Amerika
[30]Institute of Experimental Genetics, Helmholtz Zentrum München, German Research Center for Environmental Health, Almanya
[31]Institute of Computational Biology, Helmholtz Zentrum München, Almanya,

Amaç: Amaç: Metabolomik, yüzlerce düşük ağırlıklı organik molekülü (metabolit) tek bir biyo numunede paralel olarak ölçerek, belirli hücre içi biyokimyasal süreçlerin geride bıraktığı kimyasal parmak izlerini inceler. Bu projede, 14 uluslararası laboratuvarda 634'e yakın metabolitin kantifikasyonunda kullanılan MxP® Quant 500 (Biocrates Life Science AG, Innsbruck, Avusturya) metabolomik kitinin analitik performansının değerlendirilmesituvarlar arasındaki analitik değişimin değerlendirilmesi hedeflendi.

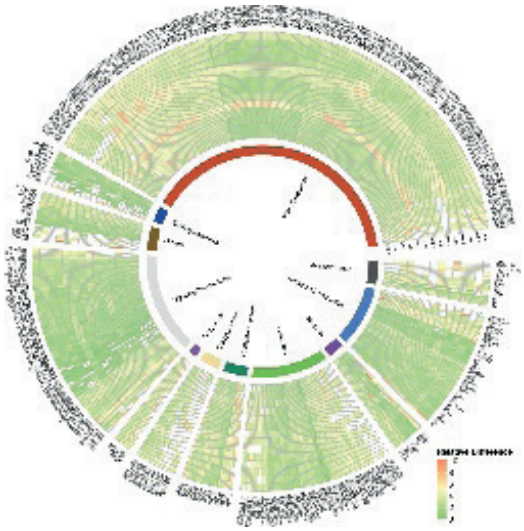
Gereç ve Yöntem: Yöntem: Laboratuvarlarda aynı insan serumu ve plazma setleri (kadın ve erkek), NIST 1950 insan plazması, lipemik insan plazması ve fare ve rat plazma örnekleri ile kalite kontrol örnekleri (QC2) analiz edildi. MxP® Quant 500 kiti bu laboratuvarlarda bulunan dört farklı kütle spektrometrisi ve dört farklı sıvı kramatografisi sisteminde kullanıldı. Çalışmaya katılan her laboratuvara A'dan N'ye kadar tek bir harf etiketi verildi. Deneylerin gerçekleştirildiği laboratuvarlardan veriler toplandı. Laboratuvarlar arası tekrar üretilebilirlik varyasyon katsayıları ile değerlendirildi. Her bir laboratuvar arasındaki farklılık yöntem karşılaştırma ile değerlendirildi. Analitik ölçüm hata oranlarının laboratuvarlar arasında sabit olduğu varsayıldı. Laboratuvar çiftleri arasındaki karşılaştırmalarda Deming regresyon yöntemi kullanıldı. Ölçüm hataları her metabolitin üç tekrarlı ölçümlerinden tahmin edildi. Güven aralıklarının tespitinde jackknife

yöntemi uygulandı. Doğrusal regresyon analizleri ile birlikte mutlak ve göreli farklılıklar hesaplandı. Sonuçlar her bir metabolitin en küçük, ortanca ve en büyük değerleri için ayrı ayrı değerlendirildi.

Bulgular: Bulgular: Çalışma kayıp verilerin temizlenmesi ile 634'ten 561 metabolite analize hazır hale getirildi. Metabolite düzeylerine göre laboratuvarlar arasındaki sabit hata miktarı 574 iken oransal hata miktarı 595'tir. Sistematik hata miktarı 600 olarak bulunmuştur. En yüksek hata miktarı PC.aa.C36.0 metabolitinde bulunmuştur. Yapılan detaylı karşılaştırmalar sonucunda metabolit konsantrasyonlarına göre laboratuvarlar arasında sistematik farklılıklar bulundu. Metabolitlerin en küçük değerleri için bağıl farklılığın daha fazla olduğu tespit edildi. Laboratuvar J için triaçilgliserit metabolit sınıfında yüksek bağıl farklılıklar elde edildi. Laboratuvar sonuçları Varyasyon katsayısı %5'in altında 6, %10'un altında 122, %20'nin altında ise 392 metabolit tespit edildi.

Sonuç: Sonuç: Laboratuvarlar arasında belirli metabolit ve metabolit grupları için sistematik farklılıkların olduğu görülmüştür. Bu doğrultuda kullanılan MxP® Quant 500 kitinin hangi laboratuvarlarda farklı sonuçlar verdiği ve bu farklılığın ne kadar olduğu belirlenmiştir. Çalışmada kullanılan bu yaklaşım ile birbirlerinden uzak ama aynı işlemi yapan laboratuvarlarda MxP® Quant 500 kiti kullanılarak metabolomik analizleri gerçekleştirilebilir. Gözlem sayısının fazla olduğu ve örneklerin tekrarlı çalışıldığı farklı laboratuvarlarda veriler birlikte değerlendirilerek analiz edilebilir.

Anahtar Kelimeler: Anahtar kelimeler: Bağıl farklılık, Deming regresyon, metabolomik, sistematik farklılık, yöntem karşılaştırma



Şekil-1. Dairesel ısı grafiği ile labaratuvarlar arasında metabolitlerin rölatif sistematik farklılıklarının gösterilmesi

SB – 16

Sağkalım Analizinde Kopula Entropi Tabanlı Değişken Seçimi

Aslıhan ALHAN

Ufuk Üniversitesi Tıp Fakültesi, Biyoistatistik Anabilim Dalı, Ankara

Amaç: Veri analizlerinde modele seçilen değişkenlerle, oluşturulan son modelin yorumlanabilir olması kritik öneme sahiptir. Bir modelin yorumlanabilir olması, modelin sonucunu tutarlı bir şekilde tahmin edebilme ve güvenirlik derecesidir. Yüksek boyutlu değişkenlerin modellenmesinde, modelin yorumlanabilir olması amacıyla daha az değişken kullanılarak modelleme yapılmaktadır. Bu durum ise modele alınabilecek aday değişkenlerin seçiminde önemli bir problem oluşturmaktadır. Son zamanlarda değişken seçiminde uygulanan kopula entropisi, değişkenler ile olaya kadar geçen süre arasındaki korelasyonu elde eden ve istatistiksel bağımsızlığı göstermek amacıyla kullanılan bir matematiksel kavramdır. Sağkalım analizinde kullanılan Cox regresyon, Random Survival Forest (RSF) ve Lasso-Cox değişken seçim yöntemlerine alternatif olacak bir yaklaşım olanağı sunan kopula entropi tabanlı değişken seçim yöntemini tanıtmak bu çalışmanın amacıdır.

Gereç ve Yöntem: Kopula entropi ile bağımlı ve bağımsız değişkenler arasındaki korelasyon ölçülerek değişken seçimi yapılmıştır. Sağkalım analizinde değişken seçimini uygulamak amacıyla Jian Ma'nın (2022) ABD Kuzey Merkez Kanser Tedavi Grubu'nda ileri evre akciğer kanserli hastalardan elde edilen 10 değişkenli 167 hastanın verilerinden yararlanılmıştır. İstatistiksel analizler için "copent", "survival", "randomForestSRC" ve "fastcox" R paket programı kullanılmıştır.

Bulgular: Model elde etmek amacıyla kullanılan 10 değişken (hastane türü, hayatta kalma süresi, sansürleme durumu, yaş, cinsiyet, ECOG performans puanı ve doktorlar tarafından değerlendirilen Karnofsky performans puanı, hastaların kendi Karnofsky performans puanı, öğünlerde tüketilen kaloriler ve son altı ayda kilo kaybı) için, seçilen değişkenler Cox regresyon, Random Survival Forest (RSF), Lasso-Cox ve Kopula Entropi modellerinde farklı bulunmuştur. Tüm modeller tarafından seçilen cinsiyet, akciğer kanserinin iyi bilinen bir prognostik faktörüdür. Performans skorları, akciğer kanseri için ortak faktörlerdir ve diğer faktörlerden daha önemli olarak kabul edilmektedir. Kopula entropi modeli, klinik kullanım için etkinliğini gösteren üç performans faktörünün tümünü seçerken RSF ve Lasso-Cox ise iki performans faktörünü seçmiştir (hastalara göre ECOG performans puanı ve Karnofsky performans puanı). Risk faktörü için bir gösterge olarak kabul edilen kilo kaybı ise sadece RSF modeli tarafından seçilmiştir.

Sonuç: Kopula entropi yönteminde, değişkenler ve olaya kadar geçen süre arasındaki korelasyon ölçülür ve ardından daha küçük kopula entropi değerleriyle hayatta kalma süresi ile ilişkilendirilen değişkenler seçilir. Kopula entropi, istatistiksel bağımsızlık için modelsiz bir ölçüdür ve parametrik olmayan bir tahmin yöntemine sahiptir. Bu nedenle, herhangi bir varsayım olmaksızın sağkalım analizine uygulanabilir. Kopula entropiyi tahmin etmek için kullanılan yöntemin hiper-parametre seçimlerine karşı duyarsız olması nedeniyle ayar gerektirmemesi diğer bir avantajdır. Yapılan analizler sonucunda, kopula entropi ile tahmin performansından ödün vermeden yorumlanabilir değişkenlerin seçilebileceği görülmektedir.

Anahtar Kelimeler: Sağkalım analizi, kopula entropisi, değişken seçimi, yorumlanabilirlik.

SB – 17

Propensity Skora Dayalı Eşleştirme Analizinde Optimum Caliper Belirlenmesi İçin Bir Benzetim Çalışması

[1]Gülçin AYDOĞDU, [2]Muhammed Ali YILMAZ, [3]Emre DEMİR, [4]Yasemin YAVUZ

[1]Hitit Üniversitesi Tıp Fakültesi, Biyoistatistik Anabilim Dalı, Çorum
[2]Sağlık Bakanlığı, Çağrı Hizmetleri Dairesi Başkanlığı, Ankara
[3]Hitit Üniversitesi Tıp Fakültesi, Biyoistatistik Anabilim Dalı, Çorum
[4]Ankara Üniversitesi Tıp Fakültesi, Biyoistatistik Anabilim Dalı, Ankara

Amaç: Randomizasyonun yapılamadığı gözlemsel araştırmalarda vaka ve kontrol grupları arasında karıştırıcı değişken/faktörlerin benzer dağılmasını sağlamak amacıyla vaka-kontrol eşleştirmesi için literatürde yaygın olarak lojistik regresyon ile hesaplanan Propensity Skor (PS) tahminine dayalı eşleştirme analizi kullanılmaktadır. Eşleştirme analizinde caliper seçimi eşleştirmenin başarısını doğrudan etkileyen önemli bir parametredir. Bu çalışmanın amacı, Propensity skora dayalı eşleştirme analizinde farklı senaryolar oluşturarak optimum caliper değerinin belirlenmesi amacıyla bir benzetim çalışması gerçekleştirmektir.

Gereç ve Yöntem: Eşleştirme analizini ve PS tahminini gerçekleştirmek için "matchIt" kütüphanesi kullanıldı. Ortak değişkenlere ilişkin eşleştirme dengesinin değerlendirilmesi için Standardize Ortalama Farklar (SOF), Varyans Oranları (VO), Kolmogorov Smirnov (KS) test istatistikleri ve çeşitli grafiksel yaklaşımlar kullanıldı. Farklı senaryolar ile üretilen veri setleri ile eşleştirme analizi yapıldı. Propensity skorlar lojistik regresyon kullanılarak hesaplandı. Eşleştirme analizlerinde en yakın komşuluk algoritması yöntemi 1:1 eşleştirme oranında gerçekleştirildi. Benzetim senaryolarında sırasıyla 0.1, 0.15, 0.2, 0.25, 0.3, 0.35, 0.40, 0.45, 0.5 ve propensity skorların standart sapmasından hesaplanan caliper değerleri kullanıldı.

Bulgular: Tüm benzetim senaryolarında SOF, KS ve VO denge değerlendirmesi bulguları Şekil 1 ve Şekil 2'de gösterildi.

Sonuç: Sadece kategorik değişkenlerden oluşan ve hem kategorik hem sayısal değişkenlerden oluşan her iki modelde de eşleştirme sonrasında denge değerlendirmelerine göre en başarılı caliper değeri 0.1 olarak belirlendi.

Anahtar Kelimeler: Propensity skor, eşleştirme, caliper, gözlemsel çalışmalar

Parametreler		Standartlaştırılmış Ortalamalar Arası Fark						Kolmogorov Smirnov İstatistikleri					
Senaryo	Örneklem Büyüklüğü	X1	X4	X5	X6	X8	X9	X1	X4	X5	X6	X8	X9
A	200	0,1	0,1	0,15	0,1	0,25	0,1	0,2	0,35	0,25	0,25	0,4	0,2
B	500	0,1	0,1	0,1	0,1	0,1	0,1	Sd of Ps	Sd of Ps	0,1	Sd of Ps	0,5	0,35
A	500	0,1	0,1	0,1	0,1	0,2	0,1	0,2	0,25	0,15	0,1	0,15	0,25
B	500	0,1	0,15	0,15	0,1	0,2	0,5	0,2	0,1	0,2	0,2	0,5	0,1
A	1000	0,1	0,3	0,1	0,1	0,45	0,35	0,15	0,1	0,1	0,1	0,5	0,1
B	1000	0,1	0,1	0,1	0,1	0,45	0,5	0,1	0,1	0,1	0,1	0,5	0,1

Şekil-1. Sadece kategorik değişkenden oluşan modele ilişkin farklı senaryolarda en başarılı eşleştirme dengesine sahip caliper değerleri (Sd of Ps: propensity skorların standart sapması)

Parametreler		Standartlaştırılmış Ortalamalar Arası Fark									
Senaryo	Örneklem Büyüklüğü	X1	X2	X3	X4	X5	X6	X7	X8	X9	X10
A	200	0,1	0,1	0,1	0,1	0,1	0,1	0,1	0,25	0,1	0,1
B	200	0,1	0,1	0,1	0,1	0,1	0,1	0,1	0,1	0,1	0,2
A	500	0,1	0,1	0,1	0,1	0,1	0,1	0,1	0,1	0,1	0,3
B	500	0,1	0,1	0,1	0,1	0,1	0,1	0,1	0,1	0,1	0,2
A	1000	0,1	0,1	0,1	0,1	0,1	0,1	0,1	0,1	0,1	0,5
B	1000	0,1	0,1	0,1	0,1	0,1	0,1	0,1	0,1	0,1	0,4
		Kolmogorov-Smirnov İstatistikleri									
Senaryo	Örneklem Büyüklüğü	X1	X2	X3	X4	X5	X6	X7	X8	X9	X10
A	200	0,15	0,25	0,15	Sd of Ps	0,15	0,2	0,15	0,4	0,2	0,4
B	200	0,2	Sd of Ps	0,2	0,2	Sd of Ps	0,15	0,15	0,4	0,25	0,5
A	500	0,1	0,1	0,1	0,15	0,1	0,1	0,1	0,4	0,1	0,5
B	500	0,1	0,15	0,1	0,1	0,1	0,1	0,1	0,3	0,2	0,35
A	1000	0,1	0,1	0,1	0,1	0,1	0,1	0,1	Sd of Ps	0,1	0,5
B	1000	0,1	0,15	0,1	0,1	0,1	0,1	0,1	Sd of Ps	0,1	0,5
		Varyans Oranları									
Senaryo	Örneklem Büyüklüğü	X1	X2	X3	X4	X5	X6	X7	X8	X9	X10
A	200		Sd of Ps		0,15			0,1			0,4
B	200		0,25		0,15			0,1			0,5
A	500		0,15		0,1			0,1			0,15
B	500		0,1		0,15			0,1			0,35
A	1000		0,1		0,1			0,1			0,3
B	1000		0,1		0,1			0,1			0,3

Şekil-2. Hem kategorik hem sayısal değişkenlerden oluşan modele ilişkin farklı senaryolarda en başarılı eşleştirme dengesine sahip caliper değerleri (Sd of Ps: propensity skorların standart sapması)

SB – 18

Görüntü İşleme Yöntemlerinin COVID-19 Pozitif Tespitinde Klasik Veri Madenciliği Yöntemleri ile Sınıflandırmaya Etkisi

[1]Fatma Gül KURT, [2]Jale KARAKAYA

[1,2] Hacettepe Üniversitesi Tıp Fakültesi, Biyoistatistik Anabilim Dalı, Ankara

Amaç: Dünya Sağlık Örgütü (DSÖ), Mart 2020'de SARS-CoV-2'nin neden olduğu bulaşıcı bir salgın olan yeni koronavirüsü (COVID-19) pandemi olarak ilan etmiştir. Bu virüsten etkilenen tüm ülkelerde, enfekte olmuş ve ölen hastaların sayısı her geçen gün artmaktadır [1]. Erken tanı virüsün yayılmasını azaltabileceğinden, akıllı tahmin ve tanı araçlarına sahip olmak önemlidir. Günümüzde tıbbi görüntüler yardımı ile birçok sağlık problemi tespitinde (beyin tümörü tespiti, meme kanseri tespiti, zatürre tespiti vb.) yapay zeka tabanlı çözümlerden yararlanılmaktadır [2–4]. İçerisinde bulunduğumuz bu pandemi sürecinde de COVID-19'un tanı süresini kısaltmak ve doğru tanı oranını arttırmak için yapay zeka tabanlı birçok çalışma yapılmıştır [5–9]. Ancak görüntüler üzerinde gürültü (noise) olarak adlandırılan bozukluklara, lekelere vb. durumlara çok sık rastlanılmaktadır. Bu gürültüler çalışmalarda kullanılan öznitelik çıkarma ve sınıflandırma aşamalarında başarı performanslarını etkilemektedir. Görüntülerdeki yüksek frekanslı nesne kenarlarının ya da ayrıntıların geliştirilmesi görüntü kalitesinde iyileşmeye yol açar. Böylece öznitelik çıkarımı aşamasında nesnelerin belli sınıflara atanması kolaylaşır bu da sınıflandırma başarısını arttırır [10]. Çalışmanın amacı görüntü işleme tekniklerinin COVID-19 ve Non-COVID akciğer grafi gruplarını sınıflandırma performansına etkisini incelemektir.

Yöntem: Çalışmada, açık erişimli Kaggle platformundan elde edilen, akciğer grafilerinin bulunduğu COVID-19 Radiography veri tabanından yararlanılmıştır [11]. COVID-19 ve Non-COVID akciğer grafi gruplarının her birinden rastgele 2784 akciğer grafisi seçilmiştir. Akciğer grafileri tıbbi görüntü işleme ile sayısallaştırılmıştır. Sayısal görüntülere Bulanık Maskeleme (USM) görüntü işleme yöntemi uygulanarak görüntü kalitesinin iyileştirilmesi hedeflenmiştir. Bir derin öğrenme yöntemi olan VGGNet-16 (VGG16) evrişimli sinir ağı mimarisi ile görüntülerden öznitelik çıkarımı gerçekleştirilmiştir. Destek Vektör Makineleri (DVM), Rastgele Orman (RO) ve Yapay Sinir Ağları (YSA) algoritmaları ile görüntüler sınıflandırılırken uygulanan görüntü işleme tekniklerinin sınıflandırma performansına etkisi incelenmiştir. Algoritmaların sınıflandırma performansları doğruluk, kesinlik, duyarlılık ve F1 ölçüleri ile değerlendirilmiştir.

Bulgular: Klasik veri madenciliği algoritmalarında sınıflandırma performansı başarısı sırasıyla RO>DVM>YSA iken görüntülerden öznitelik çıkarımı VGGNet-16 evrişimli sinir ağı mimarisi ile gerçekleştirildiğinde algoritmaların sınıflandırma başarısı sırasıyla YSA>DVM>RO şeklindedir. Orijinal görüntüler sınıflandırıldığında en yüksek sınıflandırma performansı VGG16+YSA algoritması uygulandığında %95,617 doğruluk ile elde edilmiştir. Görüntülere USM görüntü işleme yöntemi uygulandığında ise en yüksek sınıflandırma performansı %97,09 doğruluk ile VGG16+YSA algoritması sonucunda elde edilmiştir.

Sonuç: VGG16 evrişimli sinir ağı mimarisi ile öznitelikler çıkarıldığında elde edilen sınıflandırma performansları sadece klasik veri madenciliği algoritmaları ile elde edilen sonuçlara göre oldukça yüksek bulunmuştur. Uygulanan USM görüntü işleme tekniği, orijinal görüntülere göre bütün sınıflandırma algoritmalarında sınıflandırma performansını arttırmıştır.

Anahtar Kelimeler: Görüntü İşleme, Veri Madenciliği, Evrişimli Sinir Ağları, COVID-19

SB – 21

Eğilim Skoru Eşleştirmesinde Greedy ve Optimal Yöntemlerin Karşılaştırılması

[1]Gülden HAKVERDİ, [2]Timur KÖSE, [3]Cemil ÇOLAK

[1]Ege Üniversitesi Tıp Fakültesi Biyoistatistik ve Tıbbi Bilişim Anabilim Dalı,İzmir
[2]Ege Üniversitesi Tıp Fakültesi Biyoistatistik ve Tıbbi Bilişim Anabilim Dalı,İzmir
[3]İnönü Üniversitesi Tıp Fakültesi Biyoistatistik ve Tıp Bilişimi Anabilim Dalı, Malatya

Amaç: Eğilim skoru, randomizasyonun sağlanamadığı gözlemsel çalışmalarda vaka ve kontrol gruplarında ortak değişkenler bakımından var olan dengesizliği gidermek için kullanılan bir ölçüdür. Ortak değişkenler açısından vaka-kontrol gruplarında dengesizlik olduğunda sonuç hakkında yorum yapmak yanlılığa neden olmaktadır. Bu yanlılığı gidermek için sıklıkla eğilim skoru eşleştirmesi kullanılmaktadır. Bu çalışmanın amacı açık kaynak (hipotiroit) veri setinde eğilim skoru eşleştirme yöntemlerinden greedy ve optimal eşleştirme yöntemlerini kullanarak çalışma gruplarını dengelemek ve sonuçları karşılaştırmaktır.

Gereç ve Yöntem: Mevcut araştırmada UCI hipotiroid veri seti kullanılmış olup, veri seti 135 vaka ve 2138 kontrol olmak üzere toplam 2273 bireyden oluşmaktadır. Yaş, cinsiyet, hipotiroid, tiroksin, t4u ortak değişkenleri için çoklu lojistik regresyon analizi ile eğilim skoru tahmin edildikten sonra gruplar greedy ve optimal eşleştirmesi ile sırasıyla 1:1, 1:2 ve 1:3 oranında eşleştirilmiştir. Greedy ve optimal eşleştirme sonuçları denge bakımından karşılaştırılmıştır. Bütün sonuçlar için standartlaştırılmış ortalama farkları, varyans oranı, vb. hesaplanarak değerlendirilmiştir. Bütün hesaplamalarda R programlama dili ve paketleri kullanılmıştır.

Bulgular: Çalışmada kullanılan yöntemlerin bulguları incelendiğinde; optimal ve greedy eğilim skoru eşleştirme yöntemi kullanılarak 1:1 (135 vaka, 135 kontrol), 1:2 (135 vaka, 270 kontrol) ve 1:3 (135 vaka, 405 kontrol) oranında eşleştirme yapılmıştır. Eşleştirme sonuçlarına göre, yaş değişkeni dışındaki diğer değişkenlerdeki ortalama farklar ve varyans oranları, optimal eşleştirmede greedy eşleştirmeye göre daha küçük olarak elde edilmiştir. Tüm değişkenler için standartlaştırılmış ortalama farkları <0.1 ve varyans oranı <2 elde edilerek gruplar arasında denge sağlanmıştır.

Sonuç: Vaka ve kontrol grupları arasında ortak değişkenler açısından dengesizlik olması yanlı sonuçlar elde edilmesine sebep olur. Bu durum sağlık alanında yapılan tedavi veya müdahalelerin sonuçlarını olumsuz etkilemektedir. Çalışmada elde edilen bulgulara göre optimal eşleştirme greedy eşleştirme yöntemine göre ortalama farklar daha küçük elde edilmiştir. Dolayısıyla optimal eşleştirmede eşleşen çiftler uzaklık olarak birbirine daha yakın bulunmuştur. Bu çalışmadaki veriler açık kaynaklı olmakla birlikte elde edilen sonuçların literatürle benzer olduğu belirlenmiştir.

Anahtar Kelimeler: Eğilim skoru, eşleştirme, ortak değişken, vaka-kontrol

SB – 25

MEVARYA

Mutlu UMAROĞLU

Orta Doğu Teknik Üniversitesi

Amaç: Tıp, sağlık bilimleri, sosyal bilimler ve eğitim bilimleri gibi alanlarda ölçek geliştirme çalışmaları sıklıkla yapılmaktadır. Yapılan bu ölçek geliştirme çalışmaları farklı ülkelerde kullanılan ölçeklerin anadile uyarlaması olabileceği gibi belirli bir konuya açıklık getirmek amacı ile yeni bir ölçek geliştirme şeklinde de olabilmektedir. Ölçek geliştirme çalışmalarında izlenmesi gereken bazı adımlar ve değerlendirilmesi gereken bazı istatistikler bulunmaktadır. Bu adımlar ölçme aracının ilgili özelliği doğru bir şekilde ölçüp ölçmediğinin değerlendirildiği geçerlik ve ölçeğin aynı durumdaki kişilere uygulandığında alınan sonuçların benzerliğinin değerlendirildiği güvenilirlik olarak adlandırılmaktadır. Ancak ölçek geliştirme çalışmalarında bazı geçerlik-güvenirlik adımlarının uygulanmadığı ve bazı istatistikler sunulmadığı görülmektedir.

Gereç ve Yöntem: MEVARYA ölçek geliştirme çalışmalarında sunulması gereken istatistiklerin hesaplanmasını sağlayan kapsamlı bir web uygulamasıdır. MEVARYA geçerlik-güvenirlik çalışmalarında kodlarla ve karmaşık menülerle uğraşmadan sistematik bir şekilde uygulanması gereken tüm analizlerin gerçekleştirilmesine olanak sağlar.

Bulgular: MEVARYA tanımlayıcı istatistikler, açıklayıcı ve doğrulayıcı faktör analizi, Cronbach alfa katsayısı, ICC, Krippendorff alfa katsayısı, Bland-Altman grafiği gibi araçları bünyesinde barındırır. Bu web uygulamasının diğer uygulamalara göre avantajları ise web tabanlı olması nedeni ile bilgisayardan ve işletim sisteminden bağımsız olarak çalışabilmesidir.

Sonuç: Sonuç olarak piyasada bulunan ücretsiz ve ticari pek çok yazılım ile geçerlik-güvenirlik çalışmalarında kullanılan istatistikleri hesaplamak mümkündür. Kullanılacak istatistiklerin seçimi ve yorumlanması konusunda rehber içermeyen bu yazılımlara alternatif olarak MEVARYA geliştirilmiştir.

Anahtar Kelimeler: Ölçek geliştirme, geçerlik, güvenilirlik, web uygulaması

SB – 26

Makine Öğrenmesi Algoritmalarının Performanslarının Farklı Koşullar Altında Karşılaştırılması: Bir Monte Carlo Benzetim Çalışması

[1]Müge COŞKUN, [2]Osman DAĞ

[1]Biyoistatistik Uygulama ve Araştırma Birimi, Lokman Hekim Üniversitesi, Ankara, Türkiye
[2]Biyoistatistik Anabilim Dalı, Hacettepe Üniversitesi, Ankara, Türkiye.

Abstract

Data mining aims to analyze the data obtained in various ways and transform it into an understandable structure. The datasets of studies in fields such as medicine, health, and biology are quite complex. At this point, data mining is a useful set of methods that use statistical methods and machine learning algorithms to transform raw data into information. In this study, machine learning algorithms- xgboost (eXtreme Gradient Boosting), lightgbm (Light Gradient Boosting Machine), random forest, support vector machines, ridge regression, and GMDH (Group Method of Data Handling) type of neural network- are compared with the Monte Carlo simulation study. As a result of the simulation study, the performances of the related methods are compared with accuracy, MCC, F1, sensitivity, and specificity.

Keywords: Data Mining, Machine Learning Algorithms, Monte Carlo Simulation Study.

Özet

Amaç: Bu çalışmada, makine öğrenmesi algoritmalarını karşılaştırmak için bir Monte Carlo benzetim çalışması yapılarak ilgili yöntemlerin performanslarının karşılaştırılması amaçlanmıştır.

Gereç ve Yöntem: Araştırmada farklı koşulların makine öğrenmesi algoritmaları üzerine etkisini araştırmak için bir Monte Carlo benzetim çalışması yapılmıştır. Benzetim çalışması 162 farklı senaryo altında oluşturulmuştur. BinNor R paketi altında jointly. generate.binary. normal fonksiyonu kullanılarak veri türetilmiştir. Bağımsız değişkenler iki sınıflı (%40) ve sürekli (%60) değişken olarak türetilmiştir. Benzetim çalışması aşağıdakilerin tüm olası kombinasyonlarını içerir (Dağ ve ark., 2019). • Bağımsız değişkenlerin sayısı: 5, 10, 15; • Gözlem sayısı: 50, 100, 500; • Bağımlı ve bağımsız değişkenler arasındaki korelasyon: 0.2 (düşük), 0.5 (orta), 0.8 (yüksek); • Bağımsız değişkenler arasındaki korelasyon: 0- 0.1 (düşük), 0.4- 0.5 (orta), 0.8- 0.9 (yüksek); • Pozitiflerin oranı: 0.3, 0.5. Model performanslarını değerlendirebilmek için veri, eğitim (%80) ve test (%20) olarak ikiye ayrılmıştır. Test kümeleri için gerçek ve kestirilen sınıflar kullanılarak elde edilen karışıklık matrislerine dayalı olarak doğru sınıflama oranı, F1, duyarlılık ve seçicilik ile yöntemlerin performansı araştırılmaktadır. Karışıklık matrisi için GMDH2 paketin altında confMat fonksiyonu kullanılmıştır. Benzetim senaryoları 10,000 kez tekrarlanmıştır.

Bulgular: Gözlem sayısı ile bağımlı ve bağımsız değişkenler arasındaki korelasyon arttıkça doğru sınıflama oranı, duyarlılık ve seçicilik artış göstermektedir. Pozitiflerin oranı 0.5 iken çoğu senaryo altında doğru sınıflama oranı, duyarlılık ve seçicilik bakımından ridge regresyon daha iyi performans vermiştir. Bağımsız değişkenler arası korelasyon orta iken gözlem sayısı düşük olduğunda rastgele orman daha iyi performans vermiştir. Bağımsız değişkenler arası korelasyon 0- 0.1 ve pozitiflerin oranı 0.3 iken çoğu senaryo altında doğru sınıflama oranı ve duyarlılık bakımından ridge regresyon daha iyi performans verirken, seçicilik bakımından rastgele orman daha iyi performans vermiştir. Bağımsız değişkenler arası korelasyon 0.4- 0.5 ve pozitiflerin oranı 0.3 iken çoğu senaryo altında doğru sınıflama oranı ve duyarlılık bakımından ridge regresyon daha iyi performans verirken, seçicilik bakımından rastgele orman daha iyi performans vermiştir. Ancak aynı koşullarda gözlem sayısı küçükken rasgele orman daha iyi performans vermiştir. Bağımsız değişkenler arası korelasyon 0.8- 0.9 ve pozitiflerin oranı 0.3 iken çoğu senaryo altında doğru sınıflama oranı ve seçicilik bakımından rastgele orman, duyarlılık bakımından lightgbm daha iyi performans vermiştir.

Sonuç: Bu çalışmada; xgboost, lightgbm, rastgele orman, destek vektör makineleri, ridge regresyon ve GMDH yöntemleri Monte Carlo benzetim çalışmasıyla karşılaştırılmıştır. Bu benzetim çalışması ışığında ridge regresyon algoritmasının çoğu senaryo altında daha iyi performans gösterdiği saptanmıştır.

Anahtar Kelimeler: Makine öğrenmesi algoritmaları, veri madenciliği, monte carlo benzetim çalışması

SB – 28

Kantil Regresyon Yaklaşımı İçin İnteraktif Bir Arayüz Tasarımı

[1]Abdulvahap PINAR, [2]Ahmet Kadir ARSLAN

[1]İnönü Üniversitesi Tıp Fakültesi Biyoistatistik ve Tıbbi Bilişim Anabilim Dalı, Malatya
[2]İnönü Üniversitesi Tıp Fakültesi Biyoistatistik ve Tıbbi Bilişim Anabilim Dalı, Malatya

Amaç: Regresyon analizi, bağımlı değişken ile bağımsız değişken(ler) arasındaki ilişkinin matematiksel yorumu olarak adlandırılması yönünde En Küçük Kareler (EKK) yöntemine dayalı istatistiksel bir yöntemdir. Kantil regresyon, klasik regresyon varsayımı olan hata teriminin normal dağılmaması ve uç değerlere karşı etkin olması bakımından sağlam bir tekniktir. Klasik regresyon yöntemlerinin uygulanamadığı durumlarda kuantil regresyon sıklıkla kullanılan yöntemlerin başında gelir. Tıpta referans grafikleri uygun kantillerden oluşan normalden farklı bazı özel ölçüm değeri olan kişileri belirlemek için kullanılır. Yaşam analizleri için uygulamalarda ise bağımsız değişkenler orta, düşük ve yüksek riskli bireyler üzerinde farklı bir etkiye sahip olabilir. Bu etkileri yaşam süresi için farklı kantil değerleri ele alınarak anlaşılabilir. Sağlık çalışmalarında (uzunluk, kilo, hastalıkların artışı, uyuşturucu kullanım ile ilgili) tedavi yöntemlerin belirlenmesi gibi konular için web üzerinde çalışabil

Gereç ve Yöntem: Detaylı Kantil Regresyon Analiz Yazılımı, Shiny paketi yardımı ile web tabanlı arayüzü oluşturulmuştur. Geliştirilen yazılım, xls/.xlsx, .arff, .sav ve .csv/.txt uzantılı veri dosyalarını desteklemektedir. Geliştirilen yazılım, bağımsız değişkenler tek tek ya da istenirse etkileşimli olarak modele eklenebilmesine imkân tanımaktadır. Ayrıca, model performansını artırmaya yönelik olarak Çapraz Doğrulama (Cross-Validation) ve Düzenlileştirme (Lasso, Elastic-net) yöntemleri de geliştirilen yazılımda uygulanabilmektedir.

Bulgular: Geliştirilen yazılımda bulgular, modele ilişkin genel bulgular, varyans analizi tablosu, ortalama karesel hata (MSE), modele ilişkin grafikler, model katsayıları başlıkları altında raporlanmaktadır. Ayrıca bulgular, PDF ve HTML uzantılı raporlar olarak kişisel bilgisayarlara kaydedilebilmektedir.

Sonuç: Yazılım, sağlık araştırmalarında, özellikle uç değerlerin varlığı durumunda klasik doğrusal regresyonun kullanılamadığı durumlar için kolaylıkla kantil regresyon analizini uygulayabilmek amacıyla geliştirilmiştir. İlerleyen çalışmalarda, kantil regresyonunun farklı varyantlarının da (Derin Kuantil Regresyon vb.) yazılıma eklenmesi hedeflenmektedir.

Anahtar Kelimeler: Kantil regresyon, web-tabanlı yazılım, shiny

SB – 29

RNA-Dizileme Verilerinde Sağkalım Algoritmalarının Kapsamlı Bir Karşılaştırması

[1]Ahu CEPHE, [2]Ahmet SEZGIN, [3]Necla KOÇHAN, [4]Gözde ERTÜRK ZARARSIZ, [5]Vahap ELDEM, [6]Erdem KARABULUT, [7]Gökmen ZARARSIZ

[1]Erciyes Üniversitesi, Rektörlük, Kurumsal Veri Yönetimi ve Analitiği Koordinatörlüğü, Kayseri, Türkiye
[2]Abdullah Gül Üniversitesi, Bilgisayar Mühendisliği
[3]İzmir Ekonomi Üniversitesi, Fen-Edebiyat Fakültesi, Matematik Bölümü
[4]Erciyes Üniversitesi Tıp Fakültesi, Biyoistatistik ve Tıp Bilişimi ABD, Kayseri, Türkiye
[5]İstanbul Üniversitesi, Fen Fakültesi, Biyoloji Bölümü, Zooloji Anabilim Dalı, İstanbul
[6]Hacettepe Üniversitesi Tıp Fakültesi Biyoistatistik Anabilim Dalı, Ankara
[7]Erciyes Üniversitesi Tıp Fakültesi, Biyoistatistik ve Tıp Bilişimi ABD, Kayseri, Türkiye

Amaç: Son yıllarda Rna-dizileme gibi yüksek-boyutlu gen ifade verileri kullanılarak çeşitli hastalıklarda sağkalım tahminleme çalışmaları yapılmaktadır (1-4).Belirli bir olay olana kadar geçen süre ve gen ifade profili arasındaki ilişki bulunup daha kesin prognoz tahminleri yapılarak daha iyi tedavi stratejileri geliştirilebilmektedir. Düşük-boyutlu sağkalım verilerinde zaman ve sansür değişkeni dikkate alınarak birçok geleneksel sağkalım yöntemi ile analizler yapılabilirken, RNA-dizileme verilerinde kullanılacak olan sağkalım algoritmalarının bu değişkenlere ek olarak yüksek-boyutluluk, heterojenlik ve yüksek-korelasyonlu genlerle de başa çıkabilmesi gerekmektedir (1).Bu nedenle, bu verilerde sağkalım analizleri yapılırken düzenlileştirilmiş Cox ve sağkalım verilerine uyarlanmış makine öğrenme yaklaşımları kullanılmaktadır. Bu çalışmada, çeşitli kanser hastalarına ait 20 farklı RNA-dizileme verisinde mevcut sağkalım algoritmalarının performanslarının kapsamlı bir karşılaştırılması yapılmıştır.

Gereç ve Yöntem: Karşılaştırmalar yapılırken cezalı Cox, rasgele sağkalım ormanı, destek vektör makineleri, yapay sınır ağları yöntemlerini içeren 25 farklı algoritma kullanılmıştır. Sağkalım zamanı ve sansür değişkeni içeren RNA-dizileme veri setleri Kanser Genom Atlas Portalı (TCGA, https://portal.gdc.cancer.gov) veri tabanından TCGAbiolinks R paketi kullanılarak indirilmiştir. Sağkalım analizlerinde performans karşılaştırmaları için Harrell'in C-indeksi kullanılmıştır. Tüm analizler R programlama dilinin 4.2.1 versiyonu ile mlr3, mlr3proba, survival, survminer, edgeR, DESeq2 paketleri kullanılarak yapılmıştır.

Bulgular: Çalışma sonucları incelendiğinde rasgele orman (5-7) algoritmalarının genelikle yüksek performans gösterdiği görülmüştür. Bu algoritmalardan sonra lasso, ridge ve elastic-net regresyonlarının (6-8) RNA-dizileme verilerinin sağkalım analizlerinde yüksek C-indeksi sonuçları verdiği görülmüştür.

Sonuç: Yuksek-boyutlu, heterojen ve yüksek-korelasyonlu genler içeren RNA-dizileme verilerinin sağkalım analizlerinde rasgele ormanlar ile lasso, ridge ve elastic-net regresyonu algoritmaları yüksek performans göstermektedir.

Anahtar Kelimeler: Cox, gen ifade verileri, makine öğrenmesi, RNA-dizileme, sağkalım analizi

SB – 30

Sıralı Bir Yanıt Değişkenini Sınıflamada Sayısal Bir Ölçüm İçin Optimal Kesim Noktalarını Belirleme

[1]İlker ÜNAL, [2]Yaşar SERTDEMİR, [3]H. Esin ÜNAL

[1]Çukurova Üniversitesi Tıp Fakültesi Biyoistatistik Anabilim Dalı, Adana, Türkiye
[2]Çukurova Üniversitesi Tıp Fakültesi Biyoistatistik Anabilim Dalı, Adana, Türkiye
[3]Çukurova Üniversitesi Mühendislik Fakültesi Bilgisayar Mühendisliği Bölümü, Adana, Türkiye

Amaç: Sınıflama grupların uygun kriterler ile ayrıştırılması olarak tanımlanabilir. Her ne kadar sınıflama denince iki grup akla gelse de ikiden fazla grup için de aynı terminoloji kullanır. Bu çalışmada üç grubun sınıflamasında bir sayısal ölçüm için uygun kesim noktası belirlenme yöntemleri karşılaştırılmıştır.

Gereç ve Yöntem: Literatürde sıralı bir yanıt değişkenin sınıflamasında sayısal ölçümler için kesim noktası belirleme işlemi, iki sınıflı durum için kullanılan Youden endeksi, sol üst köşeye en yakın nokta, uyum olasılığı gibi metotların genelleştirilmesi ile yapılmaktadır. Bu çalışmada aynı amaçla kullanılan IU metodunun üçlü bir yanıt değişkeni için genelleştirilmesi yapılarak kesim noktaları belirlenmiştir. Çalışmada farklı senaryolarda üretilmiş veri ve literatürden alınan gerçek veri seti kullanılmıştır. Veri üretiminde örneklem büyüklüğünün etkisi, sayısal ölçümün dağılımının etkisi, gruplardaki gözlem sayısının dengesi gibi özellikler incelenmiştir.

Bulgular: Yapılan analizler sonucunda IU metodunun genelleştirilmiş halinin uygun kesim noktasını belirlemede diğer yöntemlerle benzer ve kimi durumlarda da daha iyi sonuç verdiği görülmüştür.

Sonuç: IU metodu kullanım kolaylığı ve hesap gerektirmeyen yapısı nedeniyle sınıflama için uygun kesim noktası belirlemede oldukça etkili ve kullanışlı bir yöntemdir.

Anahtar Kelimeler: Sınıflama, sıralı yanıt, IU metodu, Youden indeksi

SB – 33

Efor Kullanımlı ve Üç Parametreli Lojistik Madde Yanıt Teorisi Modellerinin Bilgisayar Uyarlamalı Test Performanslarının Karşılaştırılması

[1]Yusuf Kemal ARSLAN, [2]Afra ALKAN, [3]Atilla Halil ELHAN

[1]Çukurova Üniversitesi Tıp Fakültesi Biyoistatistik Anabilim Dalı, Adana, Türkiye
[2]Ankara Yıldırım Beyazıt Üniversitesi Tıp Fakültesi, Biyoistatistik Ana Bilim Dalı, Ankara, Türkiye
[3]Ankara Üniversitesi Tıp Fakültesi Biyoistatistik Anabilim Dalı, Ankara, Türkiye

Amaç: Bu çalışmanın asıl amacı, üç parametreli lojistik (3PL) madde yanıt teorisi modeli ve Efor kullanımlı (EM) madde yanıt teorisi modelinin bilgisayar uyarlamalı test (BUT) performanslarını karşılaştırmaktır. Primer amaca ek olarak klasik kağıt kalem testleri ve BUT performansları da karşılaştırılacaktır.

Gereç ve Yöntem: Çalışma bir benzetim çalışmasıdır. Kişi sayısı 1000; soru bankasındaki madde sayısı 50, 100 ve 250; BUT durdurma kriteri için standart hata <0.3 ve <0.5 olacak şeklinde toplam 6 senaryo için veri setleri türetilmiştir. Her senaryo için 1000 tekrar uygulanmıştır. Benzetim çalışmasında 3PL model ve EM model için düzenlenen kodlar kullanılmıştır. Bilgi fonksiyonları için uyarlanan kodların doğruluğu kontrol edilmiştir. Her bir senaryoda, BUT 3PL ve BUT EM›den elde edilen θ kestirimleri kaydedilmiştir. BUT performansını belirlemek üzere BUT kestirimlerinin yanlılıkları, standart hataları, uygulanan madde sayısı ortalaması ve standart sapması ile soru bankasındaki maddelerin uygulanma frekansları hesaplanmıştır. Ayrıca kestirimlerin gerçek θ değerleri ile uyumları değerlendirilmiştir. Oluşturulan soru bankasındaki madde parametreleri uygun dağılımlardan türetilmiştir. Sonsal beklenti kestirimleri için [-4, 4] aralığında 0.1 artışla 81 adet θ oluşturulmuştur. Yanıt süresi için her bir senaryoda kişilerin %10'u hızlı yanıt davranışı, %90'u çözümsel davranış gösterecek şekilde b'lerle ilişkili olarak türetilmiştir. BUT uygulaması, tüm kişilere orta düzeyli madde sorularak başlatılmıştır, modeller için sonraki madde seçimi en yüksek bilgiyi veren maddeyi seçerek ilerlenmesi şeklinde gerçekleştirilmiş ve süreç durdurma kriteri sağlanana kadar devam ettirilmiştir. Yetenek düzeyi kestirimlerinde en çok olabilirlik (ML) ve sonsal beklenti (EAP) kestirimleri kullanılmıştır.

Bulgular: Benzetim çalışması sonucunda yöntemden bağımsız olarak soru bankasındaki madde sayısı artarken durdurma kriteri olarak belirlenen her iki standart hata düzeyi için de uygulanan ortalama soru sayısı azalmıştır. Durdurma kriteri olarak alınan standart hatanın 0,3'ten 0,5'e çıkması uygulanan soru sayısının artmasına sebep olmuştur. BUT EM'de uygulanan ortalama soru sayılarının BUT 3PL'ye göre daha düşük olduğu görülmüştür. Ortalama standart hata için elde edilen bulgular incelendiğinde madde havuzundaki madde sayısının artmasının her iki yöntemde de ortalama standart hatayı azalttığı sonucuna varıldı. Senaryolardan 50 madde içeren ve standart hata 0,3 olan durumda her iki yöntemde de ortalama standart hatanın minimum değeri beklenen standart hata düzeyine ulaşamadı. Madde sayısındaki artışın gerçek yetenek düzeyi ve BUT yöntemleri ile kestirilen yetenek düzeyleri arasındaki SKK'nın artmasına sebep olduğu görüldü. Benzer şekilde standart hatadaki azalışın SKK'nın artmasına sebep olduğu sonucuna varıldı. Gerçek yetenek düzeyleri ve kestirilen yetenek düzeyleri arasındaki uyum incelendiğinde, standart hatanın 0,3 olduğu durumda, BUT 3PL yönteminde madde sayısı arttıkça uyum sınırları daralmış ve uyum sınırları dışında kalan kişilerin yüzdesi azalmıştır. BUT 3PL'ye benzer şekilde BUT EM yönteminde de durum benzerdi.

Sonuç: BUT EM yöntemi yetenek düzeyini BUT 3PL yöntemine göre önemsiz düzeyde düşük kestirme eğilimindedir. Her iki yöntemde de standart hata düzeyinin düşmesi sorulan soru sayısını arttırmıştır ve gerçek yetenek düzeyi ile daha uyumlu kestirimler yapılmasını sağlamıştır. Klasik kağıt-kalem yöntemleri ile BUT uygulamaları benzer sonuçlar vermiştir. Dijitalleşmenin hızlandığı günümüzde, kişi yetenek düzeyini daha belirgin şekilde belirlemek ve kişilerin sınav stratejilerini çözümlemek adına BUT EM'ler kullanılabilir.

Anahtar Kelimeler: Efor kullanımlı model, üç parametreli lojistik model, yanıt süresi, bilgisayar uyarlamalı test

SB – 34

dtComb: İki Sürekli Tanı Testinin Birleştirilmesine Yönelik Geliştirilmiş Kapsamlı Bir R Paketi

[1,2]Serra İlayda YERLITAŞ, , [3]Serra Bersan GENGEÇ, [4,5]Gözde ERTÜRK ZARARSIZ, [6]Selçuk KORKMAZ, [7-9]Gökmen ZARARSIZ,

[1]Erciyes Üniversitesi Tıp Fakültesi, Biyoistatistik ve Tıp Bilişimi ABD, Kayseri, Türkiye
[2]Erciyes Üniversitesi, İlaç Uygulama ve Araştırma Merkezi (ERFARMA), Kayseri, Türkiye
[3]Erciyes Üniversitesi, İlaç Uygulama ve Araştırma Merkezi (ERFARMA), Kayseri, Türkiye
[4]Erciyes Üniversitesi Tıp Fakültesi, Biyoistatistik ve Tıp Bilişimi ABD, Kayseri, Türkiye
[5]Erciyes Üniversitesi, İlaç Uygulama ve Araştırma Merkezi (ERFARMA), Kayseri, Türkiye
[6]Trakya Üniversitesi, Tıp Fakültesi, Biyoistatistik ve Tıbbi Bilişim Anabilim Dalı, Edirne, Türkiye
[7]Erciyes Üniversitesi, Rektörlük, Kurumsal Veri Yönetimi ve Analitiği Koordinatörlüğü, Kayseri, Türkiye
[8]Erciyes Üniversitesi Tıp Fakültesi, Biyoistatistik ve Tıp Bilişimi ABD, Kayseri, Türkiye
[9]Erciyes Üniversitesi, İlaç Uygulama ve Araştırma Merkezi (ERFARMA), Kayseri, Türkiye

Amaç: Hastalıkların teşhisi ve birbirlerinden ayırt edilmesi için tanı testlerinin geliştirilmesi sağlık alanında önemli bir araştırma alanıdır. Birden fazla tanı testi sonuçlarının birleştirilmesi ile daha yüksek tanı doğruluğu performansı ve güvenilirlik elde edilebilir. Bu çalışmada, dağınık biçimde yer alan çok sayıda kombinasyon yönteminin bir araya getirilmesi, kolay erişim ve kullanım ile araştırmacılara sunulması amacıyla dtComb R paketi ve web yazılımının geliştirilmesi amaçlanmıştır.

Gereç ve Yöntem: Tanı testlerinin performansını artırmaya yönelik birleştirilmesi için literatürde çok sayıda yöntem mevcuttur. dtComb paketi kapsamında bu kombinasyon yöntemleri doğrusal yöntemler, doğrusal olmayan yöntemler, matematiksel operatörler ve makine öğrenmesi algoritmaları olarak dört gruba ayrılmıştır. Bu dört grup içerisinde 143 adet tanı testi birleştirme yöntemine ait kodlamalar, 20 farklı R fonksiyonu içerisinde gerçekleştirilmiştir. Paketin oluşturulma aşamaları devtools R paketi ile gerçekleştirilmiştir. Geliştirilen fonksiyonların her biriminin beklendiği gibi çalıştığını doğrulamakta kullanılan birim testler için testthat R paketinden yararlanılmıştır. Oluşturulan fonksiyonların yazılım testleri 270 adet birim test ile tamamlanmıştır. Ayrıca, R kullanıcısı olmayanlar için shiny R paketi kullanılarak dtComb paketinin web yazılımı geliştirilmiştir.

Bulgular: Tanı testlerini birleştirmek için oluşturulan dört ana fonksiyon altın standart durumun ve iki tanı testinin sürekli değerlerini içeren bir veriyi girdi olarak kullanmaktadır. Bu fonksiyonlar yardımıyla kullanıcı seçtiği kombinasyon yöntemini çeşitli standartlaştırma, yeniden örnekleme ve kesim noktası yöntemleriyle birlikte kullanarak tanı testlerini birleştirebilir. Oluşturulan model ile birleştirilmiş tanı testi skorları, tekli ve birleştirilmiş tanı testi istatistikleri, ROC eğrisi ve eğri altında kalan alan istatistikleri, birleştirilmiş tanı testinin tekli tanı testi skorları ile karşılaştırmasına ilişkin test sonuçları, birleştirilmiş ve tekli tanı testi dağılım ve duyarlılık&seçicilik grafikleri raporlanabilmektedir. Ayrıca kurulan modelden alınan parametreler ile yeni veriler için tahminleme yapabilmektedir.

Sonuç: Çalışmanın kapsamında iki sürekli tanı testini birleştirmek amacıyla mevcut kombinasyon yöntemlerinin uygulanabilmesi için bir R paketi ve web yazılımı geliştirilmiştir. Geliştirilen dtComb R paketine izleyen adresten erişilebilir: http://biosoft.erciyes.edu.tr/app/dtComb.

Anahtar Kelimeler: Makine öğrenmesi, R programlama dili, ROC analizi, tanı testlerinin birleştirilmesi

SB – 35

RNA-Dizi Verilerinde İntron Tutulumunun Belirlenmesinde Kulllanılan Biyoinfotmatik Algoritmaların Değerlendirilmesi

[1]Esma Gamze AKSEL, [2]Vahap ELDEM, [3]Gökmen ZARARSIZ

[1]Erciyes Üniversitesi, Veteriner Fakültesi, Genetik Anabilim Dalı, Kayseri
[2]İstanbul Üniversitesi, Fen Fakültesi, Biyoloji Bölümü, Zooloji Anabilim Dalı, İstanbul
[3]Erciyes Üniversitesi, Tıp Fakültesi, Biyoistatistik Anabilim Dalı, Kayseri

Amaç: Alternatif ekleme tiplerinden biri olan intron tutulumu (intron retention, IR), bir intron pre-mRNA'ya kopyalandığında ve son mRNA'da kaldığında meydana gelir. İntron tutulumunun güvenilir bir şekilde ölçülmesi zor olduğu için genellikle ihmal edilen bir alternatif ekleme sınıfını oluşturur. Biyoinformatik algoritma araçları ile intron tutulumunun tespitinde karşılaşılabilecek birçok karıştırıcı faktör bulunmaktadır. Dolayısıyla IR'nin belirlenmesi için bu faktörlerin tespiti önemlidir. Bu çalışmada RNA-dizileme verilerinde intron tutulumunun belirlenmesinde kullanılan algoritmalar sunulacaktır. Bunlardan bazıları; Mixture-of-isoforms model (MISO), Keep Me Around (KMA), IRFinder, Vast-tools, MAJIQ-SPEL, Whippet, Junction usage model (JUM), IntEREst, IREAD, IRFinder-S algoritmalarıdır. Ayrıca post-kovid ve kontrol hastalarının Periferal tek çekirdekli kan hücrelerine (Peripheral Blood Mononuclear Cells, PBMC) ait 3'er örnek RNA-dizi verilerinin IRFinder-S algoritmasında yapılacak intron tutulumu analizi ile örneklendirilmesi planlanmıştır.

Yöntem: Post-kovid ve kontrol hastalarının PBMC'lerine ait RNA-dizileme verileri paired-end, fastq formatında elde edilmiştir. Analizler için Linux tabanlı sanal sunucuda Python 3.9 programlama dili ile çalışılmıştır. Referans genom olarak insan genomuna ait GRChg38, STAR programı ile indekslenmiştir. Genomik indexleme sonrası IRFinder-S programı kullanılarak RNA-dizileme verileri referans genoma hizalanarak tutulan intronların sayısal verileri elde edilmiştir. Kontrol ve post kovid PBMC RNA-dizilerinde intron tutulumuna ilişkin elde edilen sayısal verilerin istatistiksel anlamlılığı DESeq2 paketi ile belirlenmiştir.

Bulgular: Elde edilen bulgulara göre post-kovid ve kontrol PBMC'lerine ait belirli genlerde intron tutulumu oranının istatistiksel olarak anlamlı bulunduğu belirlenmiştir.

Sonuç: Özellikle hastalık, kanser türlerinin oluşum ve gelişim safhalarında intron tutulumu rolünün belirlenmesinde bahsi geçen algoritmaların tanınması, kullanımı ve yaygınlaştırılması hastalık ve kanser oluşum yolağının belirlenmesinde önem arz etmektedir.

Teşekkür: Bu çalışma Erciyes Üniversitesi Bilimsel Araştırmalar Koordinasyon Birimi tarafından TYL-2022- 12231 proje numarası ile desteklenmiştir.

Anahtar Kelimeler: Gen ekspresyonu, intron tutulumu, RNA-dizileme

SB – 36

Biyoistatistik Dersine Yönelik İstatistik Öz Yeterlilik İnanç Ölçeği, İstatistik Kaygı Ölçeği ve İstatistik Tutum Ölçeği Arasındaki Aracılık İlişkisinin Araştırılması: Bir Yapısal Eşitlik Modellemesi Çalışması (Path Analizi)

[1]Kamber KAŞALİ, [2]Senem GÖNENÇ

[1]Atatürk Üniversitesi, Tıp Fakültesi, Biyoistatistik ve Tıp Bilişimi Anabilim Dalı, Erzurum
[2]Atatürk Üniversitesi, Fen Fakültesi, İstatistik Bölümü, Erzurum

Amaç: Araştırmacıların istatistik konularına karşı korkuları akademik hayatlarını doğrudan etkilemektedir. Çünkü akademik çalışmalar en başından en sonuna kadar istatistiksel bir metodoloji ile hazırlanmalıdır. Bu sebeple, istatistik konularına karşı tutum, kaygı ve yetersizlikleri belirlemek için ölçekler geliştirilmiştir. Çalışmanın amacı; istatistik öz yeterlilik inanç ölçeği, istatistik kaygı ölçeği ve istatistik tutum ölçeği arasındaki aracılık ilişkisinin belirlenmesidir.

Gereç ve Yöntem: Çalışmamızda; istatistik öz yeterlilik inanç ölçeği, istatistik kaygı ölçeği (Alt boyutlar; Endişe, Kaçınma ve Duygusallık) ve istatistik tutum ölçeği (Alt boyutlar; İlgi duyma, Kaygı ve Motivasyon) kullanılarak öğrencilerin biyoistatistik dersine karşı durumları belirlenmiştir. Veriler dersin öncesinde (T0), ortasında (T1) ve sonunda (T2) olmak üzere 3 farklı zamanda toplanmıştır. Çalışmaya biyoistatistik dersini alan lisans, yüksek lisans ve doktora öğrencilerinden oluşan toplam 74 kişi katılmıştır. Nicel değişkenler arasındaki ilişki için korelasyon testleri uygulanmış olup gruplar arası kıyaslamalarda 2 grup için İndependent Samples t testi, Mann Whitney u testi; 3 grup için ANOVA ve Kruskal Wallis testi kullanılmıştır. Oluşturulan modelde yer alan dolaylı etkilerin anlamlılık düzeyi yapısal eşitlik modellemesi Bootstarpping yöntemi ile test edilmiş ve modele ilişkin yol katsayıları (β) hesaplanmıştır. Analizlerde, IBM SPSS 20 ve JAMOVİ 2.2.2 programları kullanılmıştır.

Bulgular: Araştırmamızda ders öncesinde, ders ortasında ve ders sonunda ölçek alt boyutları arasındaki aracılık etkisi incelenmiştir. Path analizi sonuçlarına bakıldığı zaman ders öncesinde istatistik kaygı ile istatistik öz yeterlilik inanç durumu arasında istatistiksel endişenin aracılık etkisi olduğu belirlenmiştir (β=0,20, p=0,019). Ders ortasında istatistik kaygı ile istatistik öz yeterlilik inanç durumu arasında istatistiksel duygusallığın kısmen aracılık etkisi olduğu belirlenmiştir (β=0,19, p=0,057). Ders sonunda istatistik kaygı ile istatistik öz yeterlilik inanç durumu arasında istatistiksel duygusallığın aracılık etkisi olduğu belirlenmiştir (β=0,23, p=0,049).

Sonuç: Çalışmamızda sonuç olarak Biyoistatistik dersini alan öğrencilerin istatistik öz yeterlilik ve inançları üzerinde kaygının dolaylı bir etkisinin olduğu ve bu etkide endişe ve duygusallığın aracı bir etkisi olduğu belirlenmiştir.

Anahtar Kelimeler: Yapısal Eşitlik Modellemesi, Aracılık Etkisi, Biyoistatistik

SB – 37

Bulanık Kümeleme Yöntemi ile Biyoistatistik Dersi Alan Öğrencilerin Değerlendirilmesi

[1]Senem GÖNENÇ, [2]Kamber KAŞALİ, [3]Özge PASİN

[1]Atatürk Üniversitesi, Fen Fakültesi, İstatistik Bölümü, Erzurum
[2]Atatürk Üniversitesi, Tıp Fakültesi, Biyoistatistik ve Tıp Bilişimi Anabilim Dalı, Erzurum
[3]Bezmiâlem Vakıf Üniversitesi, Tıp Fakültesi, Biyoistatistik ve Tıp Bilişimi Anabilim Dalı, İstanbul

Giriş ve Amaç: İstatistik dersi akademik hayatımızdaki başarıyı doğrudan etkilemektedir. Ancak sağlık alanında çalışma yürüten araştırmacılar genelde istatistik konularına karşı olumsuz bir tutum, kaygı ve yetersizlik içerisine girebilmektedir. Bu gibi hususların belirlenmesi ve ortaya çıkarılması akademik başarı açısından oldukça önemli bir konudur. Bu çalışmada biyoistatistik dersi alan lisans, yüksek lisans ve doktora öğrencilerinden oluşan araştırma grubu bulanık kümeleme analizi ile incelenmiştir.

Materyal ve Metot: Çalışmanın amacını ortaya koyabilmek adına araştırma grubuna ilişkin eğitim gördükleri bölümler (beslenme ve diyetetik, hemşirelik, diğer) ve daha önce istatistik dersi alma durumlarına ait bilgilere ulaşılarak öğrencilerin biyoistatistik dersine karşı durumları; istatistik öz yeterlilik inanç ölçeği, istatistik kaygı ölçeği ve istatistik tutum ölçeği kullanılarak değerlendirilmeye çalışılmıştır. Veriler dersin öncesinde ve sonrasında olmak üzere 2 farklı zamanda toplanmıştır. Söz konusu bu iki farklı zaman grubunda elde edilen veriler, bulanık kümeleme yöntemleri içerisinde en çok tercih edilen yöntem olan Bulanık k-Ortalamalar algoritması (FkM) kullanılarak analiz edilmiştir. İstatistiksel olarak bulanık kümeleme sonuçları R programında 'cluster', 'fclust' ve 'factoextra' gibi paketler vasıtası ile elde edilmiş olup her bir farklı zaman grubunda ortaya çıkan kümeleme yapısı, öğrencilerin okuduğu bölümler ve daha önce istatistik dersi alıp almama durumlarına göre SPSS programında karşılaştırılmıştır.

Bulgular: Araştırma sonucunda bulanık modelden elde edilen kümeleme sonuçları ile çalışma öncesindeki gruplar tahmin edilmeye çalışılmıştır. Öğrencilerin okuduğu bölümler açısından doğru sınıflandırma yüzdelerine bakıldığında ders öncesi grupta bulanık yöntemden elde edilen 3 küme; bölümlerdeki bireyleri sırasıyla %47, %61 ve %20 oranında doğru tespit ederken ders sonrası grupta bu oranlar; %30, %43 ve %0 olarak bulunmuştur. Daha önce istatistik dersi alma durumuna göre ders öncesi grup için bulanık kümeleme sonucu oluşan 2 küme; bireyleri sırasıyla %29 ve %43 şeklinde, ders sonrası ise %56 ve %46 gibi genel olarak artan doğru sınıflandırma oranları ile tespit etmiştir.

Sonuç: Bulgular kısmı incelendiğinde, bölüm bakımından biyoistatistik dersine başlamadan önce ölçeklerdeki sorulara verilen cevaplar arasında ders sonrasına göre daha belirgin bir fark olduğu belirlenmiştir. Biyoistatistik dersinin bitimine doğru bölümlerdeki öğrencilerin vermiş oldukları cevaplar birbirine benzemeye başladığından bulanık kümeleme sonuçlarına ait doğru sınıflandırma oranları da giderek azalmıştır. İstatistik dersini daha önce alıp almama durumunda ise ders öncesinde ölçeklerdeki sorulara verilen cevapların ders sonrasına göre daha fazla benzer oldukları (istatistik dersi almış ama konuları unutmuş) belirlenmiş ve dolayısıyla bulanık kümeleme sonuçları daha düşük doğruluk oranlarında sınıfları tahmin etmiştir. Biyoistatistik dersi devam ederken daha önce aldığı istatistik dersi konularını hatırlayıp üzerine yeni bilgiler eklediklerinden ölçek sorularına verilen cevaplar arasında ders sonuna doğru daha belirgin bir fark oluşmuştur. Bununla birlikte bulanık kümeleme sonuçları da daha yüksek doğruluk oranı gösteren tahminlerde bulunmuştur. Ayrıca bulanık algoritmanın kümeleme işlemi için kullanılmasıyla; ortaya çıkan üyelik katsayıları matrisinden hareketle aynı kümede yer alan bireyler, önem derecelerine göre kendi aralarında sıralanabilir. Böylece kümelerde yer alan her birey için tek tek ayrıntılı bir şekilde çıkarımda bulunma ve yorum yapabilme imkânı elde edilir. Çalışma sonucunda, bulanık kümelemenin istatistiksel olarak anlamlı ve yorumlanabilir sonuçlar ortaya koyduğu bununla birlikte bu tarz çalışmalarda kullanılabilir bir yöntem olarak tercih edilmesinin uygun olacağı düşünülmektedir. Bu çalışmadan başka özellikle hastaların klinik özelliklerinin gruplandırmasına ilişkin çalışmalarda da kümeleme yöntemlerinin kullanılması, klinisyenlerin karar vermesinde oldukça yardımcı olacaktır.

Anahtar Kelimeler: Kümeleme Yöntemleri, Bulanık Kümeleme, Biyoistatistik, İstatistiksel Analiz, Sınıflandırma.

SB – 38

Erken Evre Meme Kanseri Hastalarına Ait Hazard Faktörlerinin Uygulanan Çeşitli Model Sonuçlarının Karşılaştırılması

[1]Nural BEKIROGLU, [2]Ayse ULGEN, [3]Emrah Gokay ÖZGÜR, [4]Sinan UZUN, [5]Wentian LI

[1]Marmara Üniversitesi, Tıp Fakültesi, Biyoistatistik Anabilim Dalı, İstanbul
[2]Girne Amerikan Üniversitesi, Tıp Fakültesi, Biyoistatistik Anabilim Dalı, Karmi, KKTC
[3]Marmara Üniversitesi, Tıp Fakültesi, Biyoistatistik Anabilim Dalı, İstanbul
[4]Marmara Üniversitesi, Sağlık Bilimleri Enstitüsü, Biyoistatistik Anabilim Dalı, İstanbul
[5]Robert S Boas Center for Genomics and Human Genetics, The Feinstein Institute for Medical Research, Northwell Health, Manhassat, New York, ABD

Amaç: Cox Orantılısal Hazard (COH) sağkalım analizinde kullanılan çok değişkenli regresyon modelidir. Orantılı hazard (OH) varsayımını sağlamak her zaman mümkün olmamaktadır. Bu çalışmada, erken evre meme kanseri hastaların klinik bilgilerine göre, tedavi sonrası sağkalım olasılıklarına ilişkin R yazılımlı COH regresyon, makine ögrenmesi teknikleri, parametrik Hızlandırılmış Başarısızlık (ölüm) Zamanı (HBZ) modeli ile hazard faktörler açısından tahminlenmesi ve karşılaştırılması amaçlanmıştır.

Gereç ve Yöntem: Çalışmaya Marmara Üniversitesi Tıbbi Onkoloji polikliniğinde erken evre meme kanseri tanısı konulan 697 hasta dâhil edilmiştir. Erken evre meme kanseri hastaların 5 ve 10 yıllık sağkalım olasılıkları Kaplan-Meier yöntemi ile hesaplanmıştır. Sağkalıma etkili olan anlamlı hazard faktörlerinin belirlenmesinde; tek değişkenli analizlerde anlamlı hazard faktörlerin tespiti için Log-Rank testi, çok değişkenli analizlerde ise COH regresyon çözümlenmesi yapılmıştır. Aynı veri setine makine öğrenmesi algortimaları olan Rastgele Orman, Xgboost ve Gbm modelleri uygulanarak, 5 ve 10 yıllık sağkalım olasılıkları hesaplanmış olup önem derecesine göre sağkalıma etkili olan hazard faktörler sıralanmıştır. Son aşamada da; genel sağkalıma ait en uygun hazard fonksiyonu dağılımının Log-lojistik dağılıma uygun olduğu görülmüş dolayısıyla yine aynı veri setine Log-lojistik HBZ modeli uygulanmıştır. Modellerin anlamlı çıkan ilk 5 hazard faktörü tespit edilerek en uygun model saptanmaya çalışılmıştır.

Bulgular: Erken evre meme kanseri tanısı konulan 697 hastanın makine öğrenmesi algoritmalarıyla ile Kaplan-Meier sağkalım yöntemleriyle hesaplanan 5 ve 10 yıllık sağkalım olasılıkları hemen hemen aynı değerlere sahiptir. COH modelinde hazard faktörlerinin etkisi HO (hazard oranı) ile yorumlanırken, HBZ modellerinde ZO (zaman oranı) şeklinde yorumlanmaktadır. ZO, deneklerin sağkalım eğrisinde katettiği mesafenin oranlarının bir karşılaştırması olduğundan HO gibi yorumlanmasa da; hazard faktörleri açısından yapılan karşılaştırmada, meme kanseri hastaları için klinik açıdan önemli bir risk göstergesi olan Her-2 faktörü makine öğrenme algoritmalarıyla oluşturulan sağkalım modellerinin ikisinde (Gbm, Xgboost) önemli ilk 5 hazard faktörün içerisinde sıralanmış olasına rağmen COH Regresyon ve HBZ modelinde anlamlı çıkmamıştır.

Sonuç: Sağlık alanına ait, özellikle takibe dayalı çalışmalarda eksik gözlem veya dengesiz dağılım sergileyen veri setlerinde seçenek yöntemlerin tercih edilmesi analiz sonuçlarını klinik açıdan daha güvenilir yapacaktır.

Anahtar Kelimeler: Erken Evre Meme Kanseri, Cox Orantısal Hazard Regresyon, Parametrik Hızlandırılmış Başarısızlık Zamanı Modeli, Makine Öğrenmesi Teknikleri

SB – 39

Makine Öğrenme Yöntemlerinin Başarılarını Değerlendirirken Hangi Performans Ölçütleri Kullanılmalıdır

[1]Yaşar SERTDEMİR, [2]İlker ÜNAL, [3]Hülya BİNOKAY

[1]Çukurova Üniversitesi Tıp Fakültesi Biyoistatistik Anabilim Dalı, Adana
[2]Çukurova Üniversitesi Tıp Fakültesi Biyoistatistik Anabilim Dalı, Adana
[3]Çukurova Üniversitesi Tıp Fakültesi Biyoistatistik Anabilim Dalı, Adana

Amaç: Bu çalışmadaki amaç, makine öğrenme yöntemlerinin değerlendirmesinde kullanılan performans ölçütlerinden hangilerinin benzer ve hangilerinin farklı bilgi içerdiğini belirleyerek performans değerlendirmesinde kullanılacak ölçüt gruplarını belirlemektir.

Gereç ve Yöntem: Literatürde makine öğrenme yöntemlerinin karşılaştırılmasında kullanılan ve özellik olarak birbirinden 41 farklı gerçek veri setinin her biri %70'i eğitim ve %30'u test seti olarak ayrıldı. Bu ayırma işlemi rasgele olarak 300 kere tekrarlandı. Her bir tekrarda lojistik regresyon analizi uygulanarak oluşturulan modellerin test veri setlerinde PÖ'leri elde edildi. Her bir veri seti için, ölçütler üzerinde faktör analizi yapılarak hangi PÖ'lerinin birliktelik gösterdiği kayıt edildi.

Bulgular: Performans ölçütleri veri setlerindeki bağımlı değişkenin prevalansına göre <0.25, 0.25-0.39 ve 0.40-0.5 setlerine ayrılarak değerlendirildiğinde Performans ölçütleri grup sayısı ve bu gruplarda yer alan performans ölçütleri farklılık göstermektedir. Bu PÖ arasında prevelansa göre farklı PÖ gruplarında yer alan ölçütleri; RMSE, MCC, Cohen Kappa ve YI olduğu gözlenmiştir.

Sonuç: Makine öğrenme yöntemlerinin performanslarını karşılaştırırken ACC, Adj_F ve (AUCROC veya RMSE) den biri kullanıldığında üç PÖ den ikisinin aynı grupta yer alması nedeniyle ilk iki PÖ başarısına daha fazla ağırlık verilmesi neden ile yanlış yöntem başarılı bulunabilmektedir. Önerimiz bağımlı değişkenin prevalasına göre belirlenen gruplarda oluşan PÖ gruplarından birer PÖ seçilerek yöntemlerin karşılaştırılmasıdır.

Anahtar Kelimeler: Performans ölçütleri, makine öğrenme

SB – 40

Propensity Skora Dayalı Eşleştirme Analizi İçin Yeni Bir R Shiny Web Uygulaması (A New R Shiny Web Application for Propensity Score Matching Analysis)

[1]Emre DEMİR, [2]Gülçin AYDOĞDU, [3]Muhammed Ali YILMAZ, [4]Yasemin YAVUZ

[1]Hitit Üniversitesi Tıp Fakültesi, Biyoistatistik Anabilim Dalı, Çorum
[2]Hitit Üniversitesi Tıp Fakültesi, Biyoistatistik Anabilim Dalı, Çorum
[3]Sağlık Bakanlığı, Çağrı Hizmetleri Dairesi Başkanlığı, Ankara
[4]Ankara Üniversitesi Tıp Fakültesi, Biyoistatistik Anabilim Dalı, Ankara

Amaç: Birçok tıp alanında randomizasyonun gerçekleştirilemediği gözlemsel çalışmalarda bir maruziyetin veya tedavinin gerçek etkisinin ortaya konabilmesi amacıyla vaka ve kontrol gruplarının ortak değişkenler (covariate) açısından benzer dağılması istenilmektedir. Gözlemsel araştırmalarda, vaka ve kontrol grupları arasında birçok karıştırıcı (confounder) değişken açısından farklılıklar bulunabilir. Gözlemsel araştırmalarda karıştırıcı değişkenleri kontrol altında tutarak yansız tahminler elde etmek için literatürde yaygın olarak Lojistik Regresyon analizi ile hesaplanan Propensity skor (PS) tahminlerine dayalı eşleştirme analizi kullanılmaktadır. Eşleştirme analizinin amacı vaka grubundaki bireylerin sahip olduğu ortak değişken özelliklerine benzer olan bireyleri kontrol grubu içerisinden seçmektir. Bu çalışmanın amacı PS'ye dayalı eşleştirme analizi yapan ve eşleştirme sonrası birçok farklı denge değerlendirme yöntemi içeren web tabanlı yeni bir interaktif R Shiny uygulaması geliştirmektir.

Gereç ve Yöntem: Web tabanlı uygulama R Shiny (R Core Team, 2019) kullanılarak geliştirilmiştir. Eşleştirme analizi ve PS tahmini için "matchIt" kütüphanesi kullanıldı. Dengeyi değerlendirmek ve grafikler oluşturmak için "matchIt" kütüphanesine ek olarak "cobalt" kütüphanesi kullanıldı. Ortak değişkenlere/faktörlere göre eşleştirme dengesinin değerlendirilmesi için, Standardize Ortalama Farklar (SOF), Varyans Oranları (VO), Kolmogorov Smirnov (KS) test istatistikleri ve çeşitli grafiksel yaklaşımlar kullanıldı. İnteraktif uygulamanın kullanımı bir benzetim veri seti üzerinde açıklanmıştır. Araştırmacılar "https://psmatching.shinyapps.io/propensityscore" web adresinden uygulamaya erişebilirler.

Bulgular: Üretilen benzetim veri seti ile eşleştirme analizi gerçekleştirildi. Propensity skorlar lojistik regresyon analizi ile hesaplandı. Eşleştirme analizi için en yakın komşuluk algoritması, optimal eşleştirme ve tam eşleştirme yöntemleri, 1:1 ve 1:2 oranlarında sırasıyla 0.1, 0.15, 0.2, 0.25 ve 0.3 caliper değerleri kullanılarak hesaplandı. Farklı istatistiksel ve grafiksel yaklaşımlar ile gerçekleştirilen denge değerlendirilmesi sonucunda en iyi dengeye sahip eşleştirilmiş veri seti elde edildi.

Sonuç: Randomizasyonun yapılamadığı gözlemsel araştırmalarda PS'ye göre vaka ve kontrol gruplarının eşleştirilmesi karıştırıcı faktörlerin etkisini kontrol ederek maruziyetin saf etkisini belirlemek amacıyla araştırmalarda dikkate alınması gereken ve son dekatta önemi artan bir yöntemdir. Geliştirmiş olduğumuz interaktif uygulamanın küresel olarak araştırmacılar tarafından gözlemsel çalışmalarda vaka grubuna en uygun kontrol grubunu seçmek amacıyla rahatlıkla kullanılabilecek bir uygulama olacağını düşünmekteyiz.

Anahtar Kelimeler: Propensity skor, eşleştirme, R Shiny, gözlemsel çalışmalar, randomizasyon

Şekil-1. Uygulamada denge değerlendirmesini gösteren bulgular

SB – 41

Kantil Regresyon Yöntemi İle Oluşturulan Pediatrik Büyüme Eğrilerinin Model Uyumunun Worm Grafiği İle Değerlendirilmesi

[1]Eda ÇAKMAK, [2]Ergun KARAAĞAOĞLU, [3]Serhat KILIÇ, [4]Pınar ÖZDEMİR

[1]Başkent Üniversitesi Sağlık Bilimleri Fakültesi, Ankara
[2]Hacettepe Üniversitesi Tıp Fakültesi, Biyoistatistik Anabilim Dalı, Ankara
[3]Başkent Üniversitesi Tıp Fakültesi, Çocuk Sağlığı ve Hastalıkları Anabilim Dalı, Ankara
[4]Hacettepe Üniversitesi Tıp Fakültesi, Biyoistatistik Anabilim Dalı, Ankara

Amaç: Bir çocuğun büyüme ve gelişiminin incelenmesinde, hastalık durumunun erken tanılanmasında en iyi araçlardan biri pediatrik büyüme eğrileridir. Büyüme eğrilerinin oluşturulmasında kullanılan yöntemlerden biri de kantil regresyon yöntemidir. Pediatrik büyüme eğrileri oluşturulurken model yeterliğinin incelenmesinde worm grafiğinden yararlanılmaktadır. Bu çalışmanın amacı, kantil regresyon yöntemi aracılığıyla oluşturulan büyüme eğrilerinin model yeterliğinin worm grafiği ile incelenmesidir.

Gereç ve Yöntem: Kantil regresyon yöntemi veride herhangi bir dağılım varsayımı gerektirmediğinden dolayı worm grafiğini elde ederken veride bu şart aranmamaktadır. Kantil regresyon yöntemi ile oluşturulan büyüme eğrilerinin veriye ne derece uyduğunu göstermek için kullanılan worm grafiği yararlı bir görsel araçtır. Worm grafiği her bir yaş grubu için uydurulan (fitting) büyüme eğrilerinin model uyumunun değerlendirilmesinde kullanılmaktadır. Bu çalışmada Başkent Üniversitesi Ankara Hastanesi'nde 0-3 yaş aralığındaki 3320 çocuğun boy uzunluğu ve vücut ağırlığı ölçümleri kullanılarak kantil regresyon yöntemine göre büyüme eğrileri oluşturulmuş ve büyüme eğrilerinin model yeterliğinin incelenmesinde worm grafiği kullanılmıştır. Çalışmada R 4.1.2 versiyonu kullanılarak kantil regresyon yöntemi ile büyüme eğrilerinin elde edilmesinde "quantregGrowth" paketi ve kantil regresyon yöntemine bağlı olarak worm grafiklerinin oluşturulmasında "AGD" ve "quantreg" paketleri kullanılmıştır.

Bulgular: Gerçek veri seti kullanılarak kantil regresyon yöntemi ile cinsiyete göre boy uzunluğu ve vücut ağırlığı ölçümlerine ilişkin uydurulan büyüme eğrilerinin model yeterliği worm grafiği ile değerlendirilmiştir. Kız ve erkek çocuklarının 0-3 yaş aralığındaki boy uzunluğu ve vücut ağırlığı verileri için uydurulan büyüme eğrilerinin model uyumunun yeterli olduğu söylenebilmektedir. Yaş grupları için ayrı ayrı model uyumu incelendiğinde ise bazı yaş gruplarında hafif sapmalar mevcut ancak genel olarak wormlar %95 güven aralığında yer almaktadır.

Sonuç: Özellikle dağılım varsayımı gerektirmeyen kantil regresyon yöntemi ile uydurulan büyüme eğrilerinin model yeterliğinin incelenmesinde worm grafikleri kolaylık sağlayan görsel bir araçtır. Worm grafiği belirli bir yaş aralığında tüm gruba ilişkin model uyumunu değerlendirmekle birlikte, bu yaş aralığının her bir alt grupları için de model uyumunu değerlendirmek amacıyla kullanılmaktadır.

Anahtar Kelimeler: Kantil regresyon, worm grafiği, büyüme eğrileri

SB – 42

gw-Exp-Tool: A Snakemake Workflow for Genome-wide Characterization and Expression Profiling of Transcription Factors or Gene Families

Vahap ELDEM

İstanbul Üniversitesi Fen Fakültesi Biyoloji Bölümü

Amaç: Genome-wide characterization and gene expression profiling of any protein-coding genes of interest, particularly for transcription factor (TF) genes, attract great interest among biologist and medical parasitologist becuase these analyses serve as a quick solution for gene characterization at the genome-wide level, expression profiling and gene function estimation. Therefore, determining the TFs might particularly hold promise for understanding molecular and genetic determinants of pathogenesis.

Gereç ve Yöntem: We present a Snakemake workflow called "gw-Exp-Tool" built primarily for biologists to run a wide variety of useful genome-wide characterization, expression analyses and statistical test to determine whether or not TFs are expressed correctly in terms of biological meaning. Analyses consisted of several steps: (i) homology search of genes of interest against well-annotated proteomes of closely-related or model species, (ii) searching for presence of conserved domain of TFs, (iii) multiple sequence alignment and a neighbour joining (NJ) phylogenetic tree construction, (iv) motif annotation, chromosome distribution and collinearity analysis, (v) transcript abundance estimation and differential gene expression analysis, (vi) Markov-chain based rank aggregation method called Weighted MC test for most biologically relevant TFs.

Bulgular: We run the tools on four genomes including serious human pathogens (Anopheles gambiae, Ixodes scapularis, Leishmania major) and ecologically important insects (Bombus terrestris) to determine TFs, expression profiles and statistical evaluations. We successfully extracted nearly all TFs (>94%) in four genomes and further identified all TF motifs. The tools (DESeq2, edgeR, and limma) were used for the differential gene expression of TFs at genome-wide level and edgeR and limma identified the largest number of significantly differentially expressed TFs ($p<0.01$). As for biological relevance, the final list of TFs by weighted rank aggregation showed that the TFs stayed on top of the list assessed as mostly compatible with DeSeq2 results.

Sonuç: In this study, we presented a Snakemake workflow called "gw-Exp-Tool" built primarily for biologists to run a wide variety of useful genome-wide characterization, expression and statistical analysis of TFs in a manner easy enough to be employed by users without significant computational training. By providing only three inputs including genome and annotation (proteins), conserved domain IDs and transcriptomic data (or SRR accession numbers), one can easily implement and run.

Anahtar Kelimeler: Transcription factors, genome-wide analysis, snakemake workflow, rank

SB – 43

Performance Analysis of Simulated Data Sets by Machine Learning Algorithms

[1]Senem KOC, [2]Leman TOMAK, [3]Erdem KARABULUT

[1]Nişantaşı Üniversitesi Tıbbi İstatistik Anabilim Dalı, İstanbul
[2]Ondokuz Mayıs Üniversitesi Biyoistatistik ve Tıp Bilişimi Anabilim Dalı, Samsun
[3]Hacettepe Üniversitesi Biyoistatistik Anabilim Dalı, Ankara

Amaç: Machine Learning provides new tools for interpreting the high-dimensional and complex data sets that clinicians are confronted with. The aim of this study is comparing different sample size of simulated data sets by using machine learning algorithms for the performance analysis.

Gereç ve Yöntem: The simulated data were created using an original data set. This data set contain genetic information with a sample size of 174 which consist of 10 factors of periodontitis patients. Correlation matrices were created according to the peridontitis data set with the help of mean and variance values. Within these results, two data sets of 500 and 1000 sizes were created. Decision trees, K Nearest Neighborhood, Naive Bayes, Support Vector Machines, and Random Forest algorithm were compared as classification models. In addition, the 10-fold cross-validation method was used to save classifiers from side effects such as overlearning during the training phase. Accuracy, sensitivity, and specificity criteria were used to evaluate the classification performance of the models. The analyses were performed using R software.

Bulgular: As a result of the analysis, the accuracy was obtained as 70% with decision trees and nearest k neighborhood for the peridontitis data set, 71% accuracy for support vector machines, 68% accuracy for random forest and 72% accuracy with naive bayes algorithm. For the first simulated data set, the accuracy was obtained as 68% with decision trees, 70% accuracy with k nearest neighborhood and naive bayes, 72% accuracy with random forest and 73% accuracy with support vector machines. For the second simulated data set, the accuracy for naive bayes with decision trees was obtained as 75%, the accuracy for the k nearest neighborhood was 73%, the accuracy for random forest was 76%, and the accuracy for support vector machines was 78% as shown in Table 1.

Sonuç: This study contained different sample size of simulated data set. The results show that the second simulated data show the best performance by support vector machines algorithm. Which shows that by enlarging the sample sizes the performance can be improved by using machine learning algorithms.

Anahtar Kelimeler: Simulated data, machine learning, support vector machines, random forest, naive bayes.

SB – 44

Makine Öğrenmesinde Sınıf Dengesizliği Sorunu için Bazı Veri Ön İşleme Yöntemlerinin Karşılaştırılması: Bir Uygulama

[1]Hülya ÖZEN, [2]Doğukan ÖZEN

[1]Sağlık Bilimleri Üniversitesi Gülhane Tıp Fakültesi, Tıp Bilişimi Anabilim Dalı, Ankara
[2]Ankara Üniversitesi Veteriner Fakültesi, Biyoistatistik Anabilim Dalı, Ankara

Amaç: Dengesiz veriler tipik olarak sınıfların eşit olarak temsil edilmediği sınıflandırma problemleriyle ilgili bir sorunu ifade eder. Bu çalışmada; operasyon sonrası hasta dataları içeren dengesiz bir veri seti üzerinde, sınıflandırmada kullanılan veri ön işleme yöntemleri ve sonrasında uygulanan sınıflandırma algoritmalarının performanslarının karşılaştırılması amaçlanmaktadır.

Gereç ve Yöntem: Bu çalışmada postoperatif iyileşme alanındaki hastaların bir sonraki aşamada hastane içerisinde bir birime sevk edilmesi ya da taburcu olmaları durumunu tahmin edecek uygun bir sınıflandırma algoritması belirlenmeye çalışılmıştır. Bir birime sevk edilen hastalar çoğunluk, taburcu olan hastalar ise azınlık sınıfını oluşturmaktadır. Dengesiz veri setleri kullanılarak doğrudan oluşturulan sınıflandırma algoritmaları, çoğunluk sınıfı yönünde yanlı tahminlerde bulunur. Bu nedenle sınıf dengesizliğini düzenlemeye yönelik veri ön işleme yöntemleri geliştirilmiştir. Bu çalışmada aşağı örnekleme (rasgele,tek taraflı seçim, Tomek links), yukarı örnekleme (rasgele, ADASYN, SMOTE) ve hibrit yöntemler kullanılmıştır.Veri ön işleme yöntemleri, eğitim veri setlerine uygulandıktan sonra, oluşan yeni veri setlerinin sınıflandırılması için C5.0, Random Forests ve Adaboost algoritmaları kullanılmıştır. Kullanılan algoritmaların performansları kesinlik, duyarlılık, F1 skoru metrikleri ile kıyaslanmıştır.

Bulgular: Tek taraflı seçim veri ön işleme yönteminin kullanılmasının ardından Random Forests algoritmasının uygulanması ile en yüksek tahmin sonuçları elde edilmiştir. Azınlık sınıfına ait kesinlik 0.667, duyarlılık 0.571; çoğunluk sınıfına ait kesinlik 0.842, duyarlılık ise 0.889 bulunmuştur. Modele ait F1 skoru 0.615, dengeli doğruluk (balanced accuracy) değeri ise 0.730 olarak elde edilmiştir.

Sonuç: Bu çalışma; dengesiz veri setleri için belirlenecek uygun sınıflandırma algoritmasının tahmin geçerliliğinin veri ön işleme yöntemleri ile arttırılabileceğini ortaya koymaktadır. Veri setine uygun ön işleme yönteminin seçimi doğru sonuca ulaşmak için oldukça önemlidir.

Anahtar Kelimeler: Dengesiz veri seti, sınıflandırma, veri ön işleme, postoperatif veri

SB – 45

Twowaytests: İki Yönlü Bağımsız Grup Tasarımı İçin Bir R Paketi

[1]Muhammed Ali YILMAZ, [2]Osman DAĞ, [3]Samaradasa WEERAHANDI, [4]Malwane M.A. ANANDA

[1]Hacettepe Üniversitesi Tıp Fakültesi Biyoistatistik Anabilim Dalı, Ankara
[2]Hacettepe Üniversitesi Tıp Fakültesi Biyoistatistik Anabilim Dalı, Ankara
[3]Department of Mathematical Statistics, University of Nevada, Las Vegas, NV, USA
[4]Department of Mathematical Statistics, University of Nevada, Las Vegas, NV, USA

Amaç: Bu çalışmada ise iki yönlü bağımsız grup tasarımında kullanılabilecek tüm yöntemler ve analizler tek bir R paketi altında toplanarak, araştırmacılara kolaylık sağlamak amaçlanmıştır.

Gereç ve Yöntem: İki faktörün hedef değişken üzerindeki etkisi araştırılırken en çok kullanılan yöntemlerden birisi iki yönlü varyans analizi (ANOVA)'dir. ANOVA yönteminin uygulanabilmesi için normallik ve varyans homojenliği gibi iki temel varsayımın sağlanması gerekir. Bu varsayımlar sağlanmadığı durumda, iki yönlü varyans analizine alternatif yöntemler geliştirilmiştir. Geliştirilen bu yöntemler gPQ tabanlı değişen varyanslı iki yönlü ANOVA, PB tabanlı değişen varyanslı iki yönlü ANOVA, ortanca için iki yönlü ANOVA, kırpılmış ortalama için iki yönlü ANOVA ve M-kestiricileri için iki yönlü ANOVA yöntemleridir. Bahsi geçen yedi farklı test yönteminin fonksiyonları "twowaytests" R paketi içerisinde yayınlanmıştır.

Bulgular: "twowaytests" paketi içerisinde bulunan her analiz yöntemi için oluşturulan fonksiyonun çıktısında faktörlere ait test istatistiklerine ve bu test istatistiklerine ait p değerlerine yer verilmektedir. Araştırmacılar alfa değerini kendileri belirleyebilme ve analiz sonuçlarının yazılı çıktısını elde etme imkanına sahiptirler. Kütüphane içerisinde yer alan çeşitli fonksiyonlarla araştırmacılar tanımlayıcı istatistiklerle ve çeşitli grafik türleriyle veri üzerinde ilk incelemelerini yapabilirler.

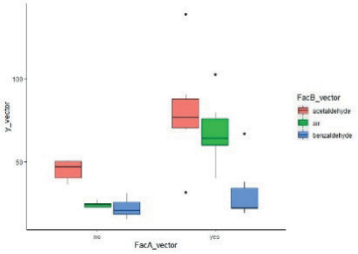
Sonuç: Bu çalışmada iki yönlü ANOVA yapmak isteyen araştırmacıların kullanması amacıyla geliştirdiğimiz "twowaytests" R paketi hazırlanmıştır. İlerleye çalışmalarda, bu fonksiyonlara ek olarak post-hoc testlerine de yer verilmesi öngörülmektedir.

Anahtar Kelimeler: İki yönlü ANOVA, ortanca, kırpılmış ortalama

```
R> library(twowaytests)
R> data("alveolar")
R> gpTwoWay(count ~ ovalbumin*treatment, data = alveolar, method = "gPB")


    Generalized p-value by PB method (alpha = 0.05)
  ===================================================
               Factor   P-value   Result
            ovalbumin    0.0044   Reject
            treatment    0.0002   Reject
  ovalbumin:treatment    0.0404   Reject
  ===================================================
```

Analiz-1. Kütüphane içerisinde yer alan örnek bir fonksiyon ile yapılan bir analizin çıktısına yer verilmiştir.



Grafik-1. Paket sayesinde örnek bir veri seti kullanılarak elde edilen bir grafiğe yer verilmiştir.

SB – 46

Referans Aralığı Belirlemede Kullanılan Tek Değişkenli ve Çok Değişkenli Yöntemlerin Karşılaştırılması

[1]Esra Kutsal MERGEN, [2]Sevilay KARAHAN

[1]Hacettepe Üniversitesi Tıp Fakültesi Biyoistatistik Anabilim Dalı, Ankara
[2]Hacettepe Üniversitesi Tıp Fakültesi Biyoistatistik Anabilim Dalı, Ankara

Amaç: Çok değişkenli referans alan yöntemi, klinik laboratuvar literatüründe tek değişkenli referans aralık yöntemine bir alternatiftir. Çok değişkenli referans alanı elde etme ve analiz etmedeki zorluklar nedeniyle bu çalışma, klinikte uygulanması kolay ve yanlış pozitif oranını azaltan çok değişkenli referans aralık yöntemleri önermektedir.

Gereç ve Yöntem: Bu çalışmada; sağlık alanında anemi teşhisi için kullanılan Serum Ferritin ve Transferrin Satürasyon yüzdesi laboratuvar değerlerinin tek değişkenli referans aralık yöntemi ile önerilen çok değişkenli referans aralık yöntemlerinin kapsayıcılığını karşılaştırmak için R Studio programı kullanılarak farklı korelasyon ve örneklem genişliği senaryoları için benzetim çalışması yapılmıştır. Anemisi olmayan, 10-20 yaş grubunda bulunan kadın ve erkekler iki ayrı alt grup olarak seçilmiştir. Bu iki laboratuvar değerinin kadın ve erkeklere ilişkin ortalama ve standart sapma değerleri kullanılarak iki değişkenli normal dağılıma sahip veriler türetilmiştir. Benzetim çalışmasında 9 farklı senaryo ve 5 farklı referans aralık yöntemi ile bireyi ilgilenilen hastalık bakımından hasta ve sağlıklı olarak sınıflandırarak karşılaştırmalar yapılmıştır.

Bulgular: Tüm senaryolarda hastalık sınıflandırma oranı en yüksek Uni1 ve en düşük Multi2 yöntemlerinde bulundu. Karşılaştırıldığında, Multi1 yöntemi, Mahalanobis yöntemine yaklaşık olarak benzer bir hastalık oranına sahipti. Tüm yöntemlerde örneklem sayısı arttıkça bulunan hasta sayısının değişkenliği azalmaktadır.

Sonuç: Referans aralıklar, sağlıklı kitlenin merkezdeki %95'ini karakterize etmek için kullanılan istatistiksel olarak türetilmiş araçlar olduğu için, Mahalanobis yöntemiyle her senaryoda ulaşılan yaklaşık %5'lik hasta oranı bu yöntemin çok değişkenli referans alanın belirsizliği ve uygulama zorluğu karşısında önerilebileceğini kanıtlamaktadır. Mahalanobis %5 değerine en yakın hasta sınıflama oranını vermesine rağmen hekimin kolaylıkla değerlendirme yapabileceği alt ve üst referans sınırını vermediği için bu yönteme en yakın sonuçları veren çok değişkenli 1 yönteminin kullanılması klinik değerlendirmede kolaylık sağlayacaktır.

Anahtar Kelimeler: Mahalanobis Uzaklığı, referans aralığı, çok değişkenli referans bölgesi, tek değişkenli referans aralığı, simülasyon çalışması

SB – 47

Yayın Yanlılığını Önlemede P-Eğrisi Yöntemi

[1]Arzu BAYGÜL EDEN, [2]Alev BAKIR KAYI

[1]Koç Üniversitesi Tıp Fakültesi
[2]İstanbul Üniversitesi, Çocuk Sağlığı Enstitüsü, Sosyal Pediatri Anabilim Dalı, İstanbul

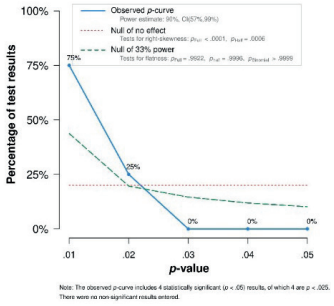
Amaç: Yayın yanlılığı ve P-korsanlığı özellikle meta analizlerde rastlanılan etik dışı ve gerçek dışı sonuçların elde edilmesini sağlayan ve direkt olarak insan sağlığını etkileyebilecek sonuçlara yolaçabilecek kavramlardır. Bu sorunları bertaraf edebilmek adına günümüzde farklı yöntemler kullanılmaktadır. En sık kullanılan yöntemler Funnel plot, Eggers rank korelasyon testi gibi yöntemlerdir. Çalışmamızda bu yöntemlere ek olarak p-eğrisi yöntemini inceleyeceğiz.

Gereç ve Yöntem: P-eğrisi yöntemi ilk olarak Simonsohn ve arkadaşları tarafından 2014 yılında, araştırıcıların özellikle kendi tarafında olan sonuçları yayınlama eğilimlerini fark edip, "Acaba bu etkiler doğru mudur, yoksa bir selektif yayın sonucu mu elde edilmiştir?" sorusuna bir cevap bulma amacıyla ortaya atılmıştır. Bu yöntemde herhangi bir konuda yayınlanmış çalışmalardan elde edilen istatistiksel olarak anlamlı "p" değerlerinden (p<0.05) oluşan bir dağılım incelenir. Bu dağılımın şekli bize çalışmaların kanıtsal bir bilgi sağlayıp sağlamayacağını söyleyecektir. Başka bir deyişle elde edilen şekillere göre tekrar edilebilir sonuçlar elde edilip edilmeyeceği hakkında bir fikir sahibi olunacaktır.

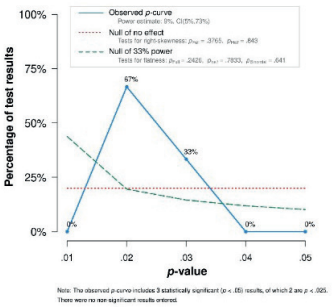
Bulgular: Uygulama1 Uygulama çevrimiçi P-curve.com/app uygulaması kullanılarak (http://www.p-curve.com/app4/pcurve4.php) gerçekleştirilmiştir. Farklı çalışmaların p değerlerini ekleyerek p- eğrisi çizdirilmiş ve elde edilen şekle göre sonuçların "kanıtsal" olup olmaması incelenmiştir. Girdiler: r(147)=0,246; F(1,100)=9,1; f(2,210)=4,45; t(200)=5.2 Sonuçların kanıtsal olması için yarı p-eğrisinin sağa çarpık olması için yapılan testte p<0,05 veya yarı ve tam p eğrisinin sağa çarpık olması için yapılan testte p<0,10 olmalıdır. Bu şartlar sağlandığı için çalışmaların kanıtsal sonuç sağladığı söylenebilir. Uygulama 2 Girdi: f(2,210)=4; r(77)=0,25; chi(2)=8 Sonuçların kanıtsal olduğunu gösteren şartlar sağlanamadığından bu çalışmalar açısından bir yayın biasından söz edilebilir.

Sonuç: p-eğrisi yayın hatasını ve selektif raporlamayı belirlemede önemli ve etkin rol oynamaktadır. Özellikle meta analiz çalışmalarında çalışmaya dahil edilecek yayınların bulunması ya da ortak bir etki arandığında "gerçek" etkiyi ifade eden çalışmaların bulunması için p-eğrisi yöntemine başvurulması iyi bir tercih olacaktır.

Anahtar Kelimeler: p-eğrisi, yayın biası, meta analiz



Şekil-1. Uygulama1



Şekil-2. Uygulama 2

Tablo-1. Uygulama 1

	Binomial Test (Share of results p<.025)	Continuous Test (Aggregate with Stouffer Method)	
		Full p-curve (p's<.05)	Half p-curve (p's<.025)
1) Studies contain evidential value. (Right skew)	p=.0625	Z=-4.04, p<.0001	Z=-3.25, p=.0006

	Statistical Power
Power of tests included in p-curve (correcting for selective reporting)	Estimate: 90% 90% Confidence interval: (57% , 99%)

Tablo-2. Uygulama 2

	Binomial Test (Share of results p<.025)	Continuous Test (Aggregate with Stouffer Method)	
		Full p-curve (p's<.05)	Half p-curve (p's<.025)
1) Studies contain evidential value. (Right skew)	p=.5	Z=-0.31, p=.3765	Z=1.01, p=.843

	Statistical Power
Power of tests included in p-curve (correcting for selective reporting)	Estimate: 9% 90% Confidence interval: (5% , 73%)

SB – 48

Kanıtsal Güç ve Literatür Bütünlüğünün p-Eğrisi ile İncelenmesi

[1]Alev BAKIR KAYI, [2]Arzu BAYGÜL EDEN

[1]İstanbul Üniversitesi, Çocuk Sağlığı Enstitüsü, Sosyal Pediatri Anabilim Dalı, İstanbul
[2]Koç Üniversitesi, Tıp Fakültesi, Biyoistatistik Bölümü, İstanbul

Amaç: Kanıt piramidinin en üstünde yer alan, tıbbi kararlar alırken en çok güven duyulan ve kanıta dayalı tıbbın bel kemiği olan meta-analiz çalışmalarında yanlı seçimler ve p-korsanlığı (p-hacking) yapılabilme ihtimalinden dolayı, meta-analiz sonuçlarının şüpheli bir geçerliliğe sahip olabileceğine dair endişeler artmaktadır. Meta-analiz sonucunda anlamlı ortak bir etki kararına varılsa da sonuçların kanıtsal olup olmadığını, tekrarlanabilirliğini p-eğrisi yöntemi ile incelenmesi amaçlanmaktadır.

Gereç ve Yöntem: Farkındalık temelli müdahalelerin dikkat eksikliği hiperaktivite bozukluğu semptomları üzerindeki etkisi meta analizi ve diyabetin transkateter aort kapak replasmanından sonra artan 1 yıllık mortalite ile ilişkisi meta-analizi çalışmalarındaki test istatistikleri kullanılarak elde edilen p dağılımı ile meta-analizlerin kanıt güçleri p-eğrisi yöntemi ile incelendi. P-eğrisi yönteminin R kodu ile çalışan hazır bir web uygulamasına başvurulmuştur.

Bulgular: Farkındalık temelli müdahalelerin dikkat eksikliği tedavisinde etkili olduğu ile ilgili derlenmiş 682 katılımcıyı içeren 11 makale ile yapılan meta-analizinin, p-eğrisi yöntemi ile kanıtsal sonuç sağladığı görülmüştür (yarı p-eğrisi için Z=-6,97 , p<0,0001, tam p-eğrisi için Z=-6,46, p<0,0001 ve istatistiksel güç %96 (%90 G.A.: %85-%99)). Diyabetin, transkateter aort kapak replasmanından sonra 1 yıllık mortalite ile ilişkisini incelemek için 17.503 börek hastasını kapsayan 22 çalışmayı içeren meta-analiz sonucunda anlamlı ortak bir karara varılmasına rağmen, p-eğrisi yarı p-eğrisi Z=0,29 , p=0,614, tam p-eğrisi Z=1,03 , p=0,847 ve istatistiksel güç %5 (%90 G.A.: %5-%22) şeklinde olup kanıtsal bir sonuç sağlamadığı bulunmuştur.

Sonuç: Meta-analiz sonuçlarının p-eğrisi yöntemi ile de değerlendirilmesi hem kanıtsal bir sonuç elde edilip edilmediğini hem de bulunan kanıtsal güce olan güveni arttıracağı için literatüre katkı sağlayacaktır.

Anahtar Kelimeler: p-eğrisi, meta-analiz, kanıtsal güç

SB – 49

The Effectiveness of Blood Routine Parameters and Some Biomarkers as a Potential Diagnostic Tool in the Diagnosis and Prognosis of COVID-19 Disease

[1]Mehmet Tahir HUYUT, [2]Sıddık KESKİN

[1]Erzincan Binali Yıldırım Üniversitesi, Biyoistatistik ve Tıbbi Bilişim A.D, Erzincan
[2]Yüzüncü Yıl Üniversitesi, Biyoistatistik A.D, Van

Aim: Due to the complex pathogenesis of COVID-19, early diagnosis and timely treatment are important. There is still a lack of information about the effects of routine blood values (RBV) on the disease process, and it is necessary to generalize and increase the reliability of existing studies. The aim of this study is to evaluate the effectiveness of new biomarkers derived from these blood values instead of using direct blood values for the diagnosis and prognosis of COVID-19 in a large patient group

Materials and Methods: Two groups were reported in this study, those treated in the intensive care unit and subjects in the all services unit. In addition, the same number of control groups was determined as the patient group. The following criteria were applied when designing derived biomarkers to determine the diagnosis and prognosis of the disease: 1) Having a significant probability ratio as a result of Multivariate Logistic Regression analysis, 2) Having significant AUC values as a result of ROC analysis, 3) Reporting of biomarkers that meet the Inflation Factor of Variance (VIF) <10 criterion among variables. In addition, a decision tree suitable for clinician use was created to check whether the evaluation of predictors together is more successful than evaluations with a single feature in determining the diagnosis and prognosis of the disease. The accuracy rate of the decision tree in predicting the diagnosis prognosis of the disease was calculated.

Results: Low lymphocyte (LYM) and white-blood-cell (WBC), high CRP and Ferritin were effective in the diagnosis of the disease. The (d-CWL) = CRP WBC*LYM and (d- CFL) = CRP*Ferritin*LYM biomarkers derived from them were the most important risk factors in diagnosing the disease and were more successful than direct RBV values. High d-CWL and d CFL values largely confirmed the Covid-19 diagnosis. The most effective RBV in the prognosis of the disease was CRP. (d-CIT) = CRP*INR*Troponin; (d-CT) = CRP*Troponin; (d-PPT) = PT*Troponin*Procalcitonin biomarkers were found to be more successful than direct RBV values and biomarkers used in previous studies in the prognosis of the disease. In this study, biomarkers derived from RBV were found to be more successful in both diagnosis and prognosis of COVID-19 than previously used direct RBV and biomarkers.

Conclusion: It was observed that the biomarkers we derived from routine laboratory tests determined the diagnosis and prognosis of COVID-19 more successfully than direct blood values and biomarkers previously entered in the literature. The independent biomarkers derived in this study are thought to provide an alternative to clinical approaches in the diagnosis and prediction of the prognosis of COVID-19.

Keywords: Covid-19, biochemical, hematological and inflammatory biomarkers, blood routine parameters, covid-19's ındependent biomarkers.

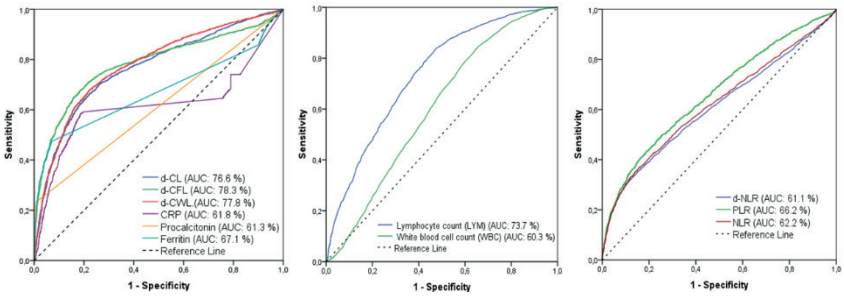


Fig-1. The left figure and the middle figure are the ROC curves for the diagnosis of Covid-19 of biomarkers derived with blood parameters. The figure on the right is the ROC curves of the biomarkers used in the literature for the diagnosis of this study.

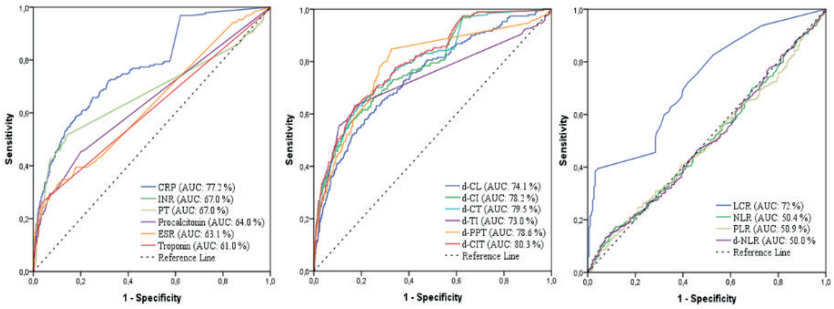


Fig-2. In the distinction between non-ICU and ICU (ie, prognosis of the disease): the left figure is the blood parameters; the middle figure is of newly derived biomarkers; The figure on the right is the ROC curves of biomarkers in the literature.

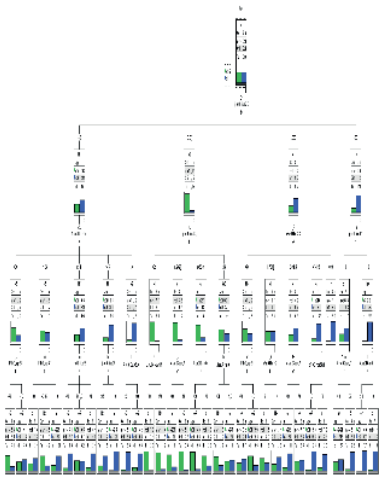


Fig-3. Decision tree evaluating the effects of predictive variables together in the diagnosis of Covid-19.

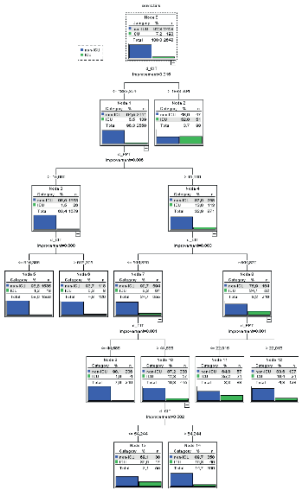


Fig-4. Decision tree evaluating the combined effect of predictive variables in determining patients to non-ICU and ICU in Covid-19.

SB – 50

Çok Değişkenli Varyans Analizinde Kullanılan Test İstatistiklerinin Tip I Hata Oranları Bakımından Karşılaştırılması

[1]Fatma Ezgi CAN, [2]Büşra EMIR, [3]Ferhan ELMALI, [4]Can ATEŞ, [5]Mustafa Agah TEKINDAL

[1]İzmir Kâtip Çelebi Üniversitesi Tıp Fakültesi, Biyoistatistik Anabilim Dalı, İzmir
[2]İzmir Kâtip Çelebi Üniversitesi Tıp Fakültesi, Biyoistatistik Anabilim Dalı, İzmir
[3]İzmir Kâtip Çelebi Üniversitesi Tıp Fakültesi, Biyoistatistik Anabilim Dalı, İzmir
[4]Aksaray Üniversitesi, Tıp Fakültesi, Biyoistatistik ve Tıp Bilişimi Anabilim Dalı, Aksaray
[5]İzmir Kâtip Çelebi Üniversitesi Tıp Fakültesi, Biyoistatistik Anabilim Dalı, İzmir

Amaç: Çok değişkenli varyans analizi, MANOVA (Multivariate Analysis of Variance), iki veya daha fazla bağımlı değişkenden oluşan verilerin analizinde kullanılmaktadır. Bu çalışmada dengeli ve dengesiz örneklemlerde MANOVA yönteminin normallik ve varyansların homojenliği varsayımlarının sağlanmadığı durumlarda Wilks Lamda, Pillai iz, Hotelling iz ve Roy'un en büyük özdeğer test istatistiklerinin Tip I hata oranları araştırılmıştır.

Gereç ve Yöntem: Test istatistiklerinin Tip I Hata oranlarını karşılaştırmak amacıyla 10000 tekrarlı farklı senaryolardan oluşan simülasyon çalışması yapılmıştır. Lognormal (1;0,5), Weibull (2,1;1) ve Weibull (0,68;1) dağılımlarından farklı grup, farklı değişken sayıları, dengeli ve dengesiz örneklem büyüklüklerinde, sabit ve artan grup varyansları dikkate alınarak rasgele sayılar türetilmiştir. $\alpha = 0,05$ alınarak her test istatistiği için Tip I hata oranları hesaplanmıştır. Simülasyon çalışmasında R v.4.2.1 program dili ve MASS v. 7.3-57 paketi kullanılmıştır.

Bulgular: Yapılan simülasyon çalışması sonucunda Lognormal dağılımda sabit varyans ve dengesiz örneklemlerde Pillai iz test istatistiğinin Tip I hata oranları $\alpha = 0,05$ değerine en yakın sonuçları vermiştir. Değişen varyans, 2 ve 3 değişkenli dengeli örneklemlerde ise Hotelling iz istatistiği en yakın sonuçları göstermiştir. Değişen varyans, 3 değişkenli dengesiz örneklemlerdeyse Roy'un en büyük özdeğer test istatistiği en yakın sonuçları vermiştir. Weibull dağılımında sabit varyans ve 2 değişkenli dengesiz örneklemlerde $\alpha = 0,05$ değerine en yakın sonuçlar Hotelling iz test istatistiğinden elde edilirken; 3 değişkenli olduğu durumda Wilks Lamda istatistiği en yakın sonucu göstermiştir. Weibull (0,68;1) dağılımından elde edilen sonuçlarda 2 değişkenli dengeli ve dengesiz örneklemlerde Pillai iz, 3 değişkenli dengeli örneklemlerde Hotelling iz, 3 değişkenli dengesiz örneklemlerdeyse Pillai iz test istatistiği en yakın sonuçları göstermiştir.

Sonuç: Simülasyon çalışmasında elde edilen sonuçlara göre Lognormal dağılımın Weibull dağılımına göre daha tutarlı sonuçlar verdiği görülmüştür.

Anahtar Kelimeler: Wilks lamda istatistiği, pillai iz istatistiği, hotelling iz istatistiği, roy'un en büyük özdeğer istatistiği, MANOVA

SB – 51

Obuchowski Parametrik olmayan ROC Eğrisi Analizi ve Küme Bootstrap Yaklaşımları ile İlişkili Verilerin Değerlendirilmesi: Multipl Skleroz Hastalarında Optik Nöropati Teşhisi Üzerine Bir Uygulama

[1]Büşra EMIR, [2]Fatma Ezgi CAN, [3]Mustafa Agah TEKINDAL, [4]Ferhan ELMALI, [5]Ertuğrul ÇOLAK

[1]İzmir Katip Çelebi Üniversitesi Tıp Fakültesi, Biyoistatistik Anabilim Dalı
[2]İzmir Katip Çelebi Üniversitesi Tıp Fakültesi, Biyoistatistik Anabilim Dalı
[3]İzmir Katip Çelebi Üniversitesi Tıp Fakültesi, Biyoistatistik Anabilim Dalı
[4]İzmir Katip Çelebi Üniversitesi Tıp Fakültesi, Biyoistatistik Anabilim Dalı
[5]Eskişehir Osmangazi Üniversitesi Tıp Fakültesi, Biyoistatistik Anabilim Dalı

Amaç: Bu çalışmada, ilişkili verilerin ROC (Receiver Operating Characteristic) eğrisi ile değerlendirilmesinde önerilen Obuchowski parametrik olmayan ROC eğrisi analizi ve Küme Bootstrap yaklaşımları incelenmiştir. Temel motivasyonumuz bağımlı kümelenmiş verilerin değerlendirilmesinde üç yöntemin karşılaştırılmasıdır.

Gereç ve Yöntem: Multipl Skleroz hastalarında Optik Nöropati (ON) teşhisi üzerine (ON +/-), iki farklı görüntüleme tekniğine (OCT, GDx) ait sağ ve sol göz retina sinir lifi tabakası kalınlığı ölçümleri için Obuchowski parametrik olmayan ROC eğrisi analizi ve Küme Bootstrap yöntemlerine ait eğri altında kalan alan (AUC) değerleri ve %95 güven aralıkları hesaplanmıştır. Orijinal veri setine ait sağ ve sol göz RNFL kalınlığı ölçüm ortalama ve standart sapma değerleri, örneklem büyüklükleri ve değişen gözler arası korelasyon yapısı (r=0.25, 0.50 ve 0.75) kullanılarak çok değişkenli normal dağılımdan üç farklı veri seti oluşturulmuştur. Değişen gözler arası korelasyonu dahil etmeden gerçekleştirilen parametrik ROC eğrisi analizi sonuçları ile hesaplanan yaklaşımlar karşılaştırılmıştır. Obuchowski parametrik olmayan ROC eğrisi analizi RStudio, parametrik ROC eğrisi analizi ve küme bootstrap yaklaşımı SAS University Edition programı kullanılarak gerçekleştirilmiştir.

Bulgular: Yetmiş iki hasta ve 144 retina sinir lifi tabakası kalınlığı ölçümlerinin yer aldığı optik nöropati (ON) çalışmasında, ON teşhisi için optik koherens tomografi görüntüleme tekniği (OCT) ile elde edilen retina sinir lifi tabakası (RNFL) kalınlığının AUC değeri 0.705 (%95 güven aralığı (%95 GA): 0.616, 0.793) olarak elde edilmiştir. Gözler arası korelasyonu hesaba katmadan parametrik yaklaşım ile elde edilen AUC değeri ve güven aralığı, Obuchowski parametrik olmayan ROC eğrisi analizi (%95 GA: 0.607, 0.803) ve küme bootstrap (%95 GA: 0.600, 0.800) yaklaşımlardan daha dardı. Tarayıcı lazer polimetri görüntüleme tekniği (GDx) ile elde edilen RNFL kalınlığının AUC değeri 0.666 (%95 GA : 0.576, 0.757) olarak elde edilmiştir. Gözler arası korelasyonu hesaba katmadan parametrik yaklaşım ile elde edilen AUC değeri, Obuchowski parametrik olmayan ROC eğrisi analizi (%95 GA: 0.564, 0.769) ve küme bootstrap (%95 GA: 0.555, 0.764) yaklaşımlardan daha dardı.

Sonuç: İlişkili verilerin ROC analizinde, veriler arası korelasyonun göz ardı edilerek parametrik ROC eğrisi analizinin kullanılması, korelasyonun büyüklüğüne bağlı olarak %95 güven aralığının daha dar olmasına sebep olmuştur. İlişkili verilerin analizinde Obuchowski parametrik olmayan ROC eğrisi analizi ya da küme bootstrap yaklaşımlarının kullanılması ile daha geniş güven aralığı elde edilebilir.

Anahtar Kelimeler: Küme bootstrap, obuchowski, parametrik olmayan roc eğrisi

SB – 53

Doğrusal Olmayan Yapısal Eşitlik Modellemesi Yaklaşımlarının İncelenmesi ve Bir Uygulama

[1]Nur Efşan TIĞLI, [2]Şengül CANGÜR

[1]Düzce üniversitesi Biyoistatistik Anabilim Dalı, Düzce
[2]Düzce üniversitesi Biyoistatistik Anabilim Dalı, Düzce

Amaç: Bu çalışmada değişkenler arasındaki ilişkilerin doğrusal olmadığı ya da normal dağılım varsayımı sağlanamadığı durumlarda kullanılan gizil moderatörlü yapısal eşitlikler (GMYE), yapısal eşitlik karışım modelleri (YEKM) ve doğrusal olmayan yapısal eşitlik karışım modellemesi (DOYEKM) yaklaşımlarının tanıtılması ve gerçek veri seti üzerinde uygulamasının gösterilmesi amaçlanmıştır.

Gereç ve Yöntem: Doğrusal olmayan YEM kapsamında üç yaklaşım üzerinde durulmuştur. Birincisi, etkileşim etkisi veya ikinci dereceden etkiler gibi gizil değişkenleri kullanarak doğrusal olmayan ilişki türlerini analiz etmek için kullanılan GMYE yaklaşımı, ikincisi ise doğrusal olmayan gizil değişkenlerin doğrusal ilişkilerin karışımları ile yaklaşık olarak tahmin edilmesi için kullanılan YEKM yaklaşımıdır. Bu iki yaklaşımın birleşimini gösteren üçüncü bir yaklaşım olarak da yarı parametrik yöntem olarak geliştirilen DOYEKM yaklaşımı ele alınmıştır. Bu üç yaklaşım kullanılarak Düzce Üniversitesi Tıp Fakültesi öğrencilerinin (n=272) yalnızlık ve yaşam doyumunun internet bağımlılığı üzerindeki etkilerini incelemek için yapısal modeller oluşturulmuştur. Model karşılaştırmaları için Akaike bilgi kriteri (AIC), Schwarz Bayes bilgi kriteri (SBIC) ve loglikelihood değerleri hesaplanmıştır. Verilerin analizi için R-CRAN 3.6.2 ve SPSS 24 programları kullanılmıştır.

Bulgular: AIC ve SBIC kriterlerine göre en küçük AIC ve SBIC değerlerine sahip model en uygun model olarak kabul edildiğinden, DOYEKM yaklaşımının diğer iki yaklaşıma göre daha kullanışlı olduğu görülmüştür. Yalnızlık gizil değişkeninin internet bağımlılığı üzerinde anlamlı pozitif bir etkisi olduğu, yaşam doyumu gizil değişkeninin ise anlamlı negatif bir etkisi olduğu saptanmıştır. Ayrıca hem yalnızlık ve yaşam doyumu gizil değişkenlerinin etkileşim etkisinin hem de yaşam doyumu gizil değişkeninin karesel etkisinin anlamlı düzeyde negatif etkilere sahip oldukları bulunmuştur.

Sonuç: Değişkenler arasındaki ilişkilerin doğrusal olmadığı ya da normal dağılım sağlanamadığı durumlarda klasik YEM yaklaşımlarının yetersiz kalmasına karşı geliştirilen doğrusal olmayan YEM yaklaşımlarının özellikle sağlık ve davranış bilimleri başta olmak üzere tüm alanlarda yaygın bir şekilde kullanılmasını önermekteyiz.

Anahtar Kelimeler: Gizil moderatörlü yapısal eşitlikler, Yapısal eşitlik karışım modelleri, Doğrusal olmayan yapısal eşitlik karışım modellemesi

SB – 54

Machine Learning Sensors for Diagnosis of COVID-19 Disease Using Routine Blood Values for Internet of Things Application

[1]Mehmet Tahir HUYUT, [2]Andrei VELICHKO, [3]Belyaev MAKSIM

[1]Erzincan Binali Yıldırım Üniversitesi, Biyoistatistik ve Tıbbi Bilişim A.D, Erzincan
[2]Petrozavodsk State University,Institute of Physics and Technology,33 Lenin Str., 185910 Petrozavodsk, Russia
[3]Petrozavodsk State University,Institute of Physics and Technology,33 Lenin Str., 185910 Petrozavodsk, Russia

Aim: Healthcare digitalization needs effective methods of human sensorics, when various parameters of the human body are instantly monitored in everyday life and connected to the Internet of Things (IoT). In particular, Machine Learning (ML) sensors for the prompt diagnosis of COVID-19 is an important case for IoT application in healthcare and Ambient Assistance Living. This study aims to provide a fast, reliable and economical tool for the diagnosis of COVID-19 based on routine blood values.

Materials and Methods: In this study, 51 features belonging to the same number of patients with positive and negative COVID-19 test results were used (a total of 5296 patients). In this study, the performance of 13 popular supervised ML models and LogNNet deep neural network models in the diagnosis of the disease was investigated. Before the models were trained, the data were normalized and then additional features were generated. Quantitative transformer (QT) for the normalization process; Robust scaler (RS); MinMax (MM) methods were used. Polynomial (PN) to create additional properties; Random trees embedding (RTE); Extra trees preprocessor (ETP); Linear SVM preprocessor (LSVMP); Independent component analysis (ICA); Nystroem sampler (NS) methods were used. The "auto-sklearn" software package adjusted to the target indicator-classification accuracy was used for optimization of each classifier model. The accuracy of the models was evaluated by the K-fold cross validation method.

Results: The most successful classifier model in terms of time and accuracy in the detection of the disease was the Histogram-based Gradient Boosting (HGB) (accuracy: 100%, time: 6.39 sec). The HGB classifier identified the 11 most important features to detect the disease with 100% accuracy. In addition, the accuracy rate of the single, double and triple combinations of these features operated with HGB in the diagnosis of the disease was determined. Paired combinations of these features run with HGB achieved higher accuracy in diagnosing the disease than the direct features. Also, the implementation of the LogNNet network on the arduino demonstrated the practical applicability of the model in low-power devices (ie the concept of the Internet of Things).

Conclusion: We propose to use these 11 traits and their binary combinations as important biomarkers for ML sensors in diagnosis of the disease, supporting edge computing on Arduino and cloud IoT service. The results make it possible to create the concept of Machine Learning Sensors for the Diagnosis of COVID-19 Disease. In addition, the results of this study can be effectively used in IoT(IoTM) medical peripheral devices with low RAM resources, ML sensors, portable point-of-care blood testing devices, clinical decision support systems, remote internet medicine and telemedicine.

Keywords: COVID-19; biochemical and hematological biomarkers; routine blood values; feature selection method; LogNNet neural network; Machine Learning sensors, Internet of Medical Things; IoT.

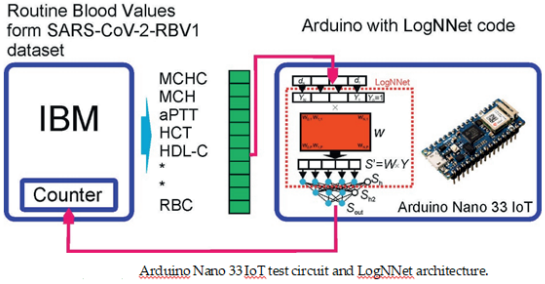
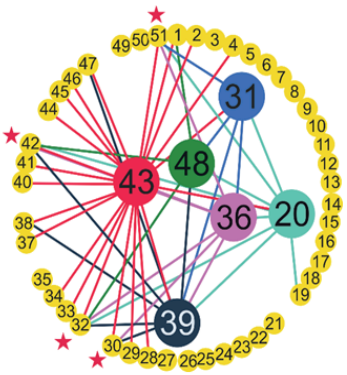


Fig-1. Arduino Nano 33 IoT test devresi, LogNNet mimarisi ve kartı Fig-1'de gösterilmiştir.

Table-1. Results of assessing the classification accuracy of machine learning models for the SARS-CoV-2-RBV1 dataset.

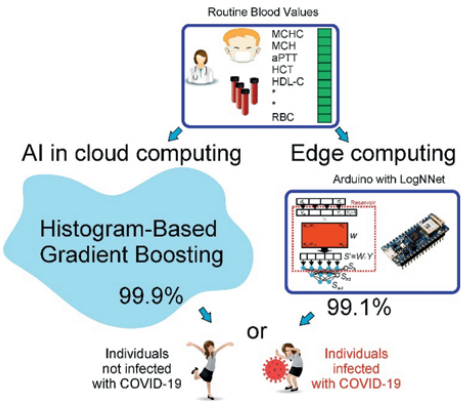
Results of assessing the classification accuracy of machine learning models for the SARS-CoV-2-RBV1 dataset.

Classification algorithm	Average model accuracy A_{51}, %	Average learning time, s	Normalization method	Methods for generating additional features
Histogram based Gradient Boosting	100	6.39	-	-
Random Forest	99.943	13.15	QT	-
K-nearest neighbors	99.924	3.17	QT	ETP
Extra Trees classifier	99.905	18.73	RS	-
Multilayer Perceptron	99.886	3.99	RS	LSVMP
Multinomial Naive Bayes	99.792	2.48	QT	RTE
Linear Discriminant Analysis	99.773	9.15	QT	PN
Support Vector Machine with non-linear kernel	99.754	222.41	QT	NS
Decision Tree	99.660	1.46	RS	LSVMP
Passive-Aggressive	99.641	2.91	QT	RTE
Bernoulli Naive Bayes	99.622	2.59	QT	RTE
Support Vector Machine with linear kernel	99.584	5.21	MM	PN
Gaussian Naive Bayes	98.565	1.68	QT	ICA
LogNNet [37]	99.509	100	-	-



Pairs of features with high classification efficiency SARS-CoV-2-RBV1 datasets for the Histogram based Gradient Boosting classifier.

Fig-2. Pairs of features with high classification efficiency SARS-CoV-2-RBV1 datasets for the Histogram based Gradient Boosting classifier.



Two architectures of the IoT system, which includes an IoT device with LogNNet implementation (edge computing) and a cloud service containing a trained HGB model (AI computing).

Fig-3. Two architectures of the IoT system, which includes an IoT device with LogNNet implementation (edge computing) and a cloud service containing a trained HGB model (AI computing).

SB – 55

Sınıflamada Kullanılan Veri Madenciliği Yöntemlerinin Performanslarının D Vitamini Veri Seti Üzerinde Karşılaştırılması

[1]Özlem ARIK, [2]İnci ARIKAN, [3]Erdem KARABULUT

[1]Kütahya Sağlık Bilimleri Üniversitesi, Tıp Fakültesi, Biyoistatistik Anabilim Dalı, Kütahya
[2]Kütahya Sağlık Bilimleri Üniversitesi, Tıp Fakültesi, Halk Sağlığı Anabilim Dalı, Kütahya
[3]Hacettepe Üniversitesi, Tıp Fakültesi, Biyoistatistik Anabilim Dalı, Kütahya

Amaç: Büyük miktarda biriken ham veri setlerinden anlamlı, değerli ve faydalı bilgilerin ortaya çıkarılmasına veri madenciliği denir. Tahmin edici yöntemler arasında sınıflama yöntemleri bulunmaktadır. Bu çalışmadaki temel amaç, Kütahya Sağlık Bilimleri Üniversitesi Evliya Çelebi Eğitim ve Araştırma Hastanesine 2017-2021 yılları arasında ayaktan başvuran 121727 hastaya ait bazı bilgiler kullanılarak hastanın D vitamini düzeyini sınıflamak ve sınıflama yöntemlerinin performansını ölçmektir.

Gereç ve Yöntem: Çalışmada, hastaneye ayaktan başvuran hastalardan 25-OH VitD düzeyi çalışılanlar hastane veri tabanı siste-minden geriye dönük olarak incelendi. Hastaların yaş, cinsiyet, mevsim ve 25-OH VitD düzeyi bilgileri kullanıldı. Veri ile ilgili ön işlemeden sonra veri madenciliği sınıflama yöntemlerinden k en-yakın komşu (kNN), karar ağaçları, destek vektör makineleri (DVM), random forest, yapay sinir ağları (YSA), naive bayes (NB) ve lojistik regresyon kullanıldı. Doğru sınıflama oranı, duyarlılık, seçicilik, F1 ölçüsü, kesinlik ve ROC eğrisi altında kalan alan (EAKA) ölçüleri aracılığıyla yöntemlerin performansları belirlendi.

Bulgular: D vitamini verisi için 25-OH VitD'nin 20 ng/ml'den (50 nmol/L) küçük düzeyleri Vitamin D eksikliği olarak ifade edilmiştir. Vitamin D eksikliği var-yok şeklinde ve hedef sınıf: Vitamin D eksikliği var olduğunda sınıflama modellerine ait elde edilen bulgular Tablo 1`de yer almaktadır.

Sonuç: Verilerin analizinin Orange 3.32 programı aracılığıyla yapıldığı çalışmada, Tablo 1`de verilen doğru sınıflama oranı, duyarlılık, F1 ölçüsü, kesinlik ve EAKA değerlerine göre karar ağaçları, random forest ve yapay sinir ağları yöntemlerinin performansları birbirine yakındır ve diğerlerine göre daha iyidir. ROC eğrisi altında kalan alan (EAKA) değerleri incelendiğinde ise hastaları D vitamini düzeylerine göre sınıflara ayırmak için kullanılan yöntemler içerisinde DVM modelinin performansının diğerlerine göre daha düşük olduğu sonucuna ulaşılmıştır.

Anahtar Kelimeler: Sınıflama, performans, D vitamini

SB – 56

Psikiyatrik Hastalıklarda, Psikopatoloji ve Çevresel Etken İlişkisinin Özyinelemesiz Model ile Değerlendirilmesi

[1]Selen Begüm UZUN, [2]Derya GÖKMEN, [3]Meram Can SAKA

[1]T.C. Sağlık Bakanlığı, Sağlık Bilgi Sistemleri Genel Müdürlüğü, Sağlık İstatistikleri Dairesi Başkanlığı, Ankara
[2]Ankara Üniversitesi, Tıp Fakültesi, Biyoistatistik Anabilim Dalı, Ankara
[3]Ankara Üniversitesi, Tıp Fakültesi, Ruh Sağlığı ve Hastalıkları Anabilim Dalı, Ankara

Amaç: Psikiyatri alanında; çevresel etmenler, hastalık durumu ve şiddeti, hastaların işlevselliği, algıladıkları sosyal destek ve sosyal ilişkileri psikopatoloji üzerinde etkili değişkenler olarak değerlendirilmektedir. Bu çalışmada, çevresel etmenler, hastalık durumu, hastalık şiddeti gibi değişkenlerin psikopatoloji ile ilişkisinin özyinelemesiz eş zamanlı denklem modelleri ile değerlendirilmesi amaçlanmıştır.

Gereç ve Yöntem: Çalışma, Mayıs-Haziran 2017 tarihleri arasında Ankara Üniversitesi Tıp Fakültesi Ruh Sağlığı ve Hastalıkları Anabilim Dalı Polikliniğine başvuran ve çalışmaya katılmayı kabul eden 378 hasta üzerinde gerçekleştirilmiştir. Çalışma kapsamında; hastaların demografik özelliklerinin yanısıra, Çocukluk Çağı Travma Ölçeği (ÇÇT), Rosenberg Benlik Saygısı, Çok Boyutlu Algılanan Sosyal Destek Ölçeği (ÇBASD), Bireysel ve Sosyal Performans Ölçeğine (PSP) verdikleri yanıtlar bulunmaktadır. Hastalık durumu; hekim tarafından yapılan konsültasyon sonrasında ICD-10 kodları kullanılarak; belirti/semptom şiddeti ise Klinik Global İzlenim Ölçeği kullanılarak belirlenmiştir. Çalışmada araç değişkenler yardımıyla iki aşamalı probit en küçük kareler (2SPLS), iki aşamalı koşullu en çok olabilirlik (2SCML) ve üç aşamalı en küçük kareler (3SLS) yöntemleri Stata programı aracılığı ile gerçekleştirilmiştir.

Bulgular: Çalışmaya katılan hastaların yaş ortalaması (standart sapma:SS) 37,3 (11,6) olup; % 61,6'sı kadın, % 58,8'i lise ve altı eğitime sahip, %48,7'si evli, % 72,1'i asgari ücretin üstünde gelire sahiptir. Hastaların ölçek puanlarına ilişkin tanımlayıcı istatistikler değerlendirildiğinde; bireylerin, aile/arkadaş veya özel bir kişiden gördüklerini düşündükleri/algıladıkları puanı ortalaması 53,9 (18,9), benlik saygısı puan ortalaması 22,8 (5,7), çocukluklarında istismar/ihmal nedeniyle yaşadıkları travma şiddetinin ortalaması 44,5 (12,1) olarak bulunmuştur. Araç değişkenler yardımıyla uygulanan 2SPLS ve 2SCML yöntemleri sonucunda; kişilerin hastalık grubuna atanmasında-tanı almasında ÇBASD ve PSP ve Rosenberg değerlerinin anlamlı etkisi olmadığı; 3SLS sonucunda ise benlik saygısı için çocukluk çağı travması ve yaş, algılanan destek için çocukluk çağı travması, bireysel ve sosyal işlevsellikte ise semptom şiddetinin anlamlı değişken olarak belirlenebileceğine işaret etmektedir.

Sonuç: Çocukluk çağında yaşanan istismar şiddeti, benlik saygısı ve destek algısının azalmasına sebep olmakta, bu konuda özellikle eğitimcilerin ve ebeveynlerin bilinçlendirilmesi gerekmektedir. Hastalık durumu ve şiddetinin sosyal işlevselliği etkilemesi nedeniyle bu hastalara gerekli psikolojik desteğin verilmesinin hastanın psikopatoloji açısından iyilik hali üzerinde olumlu etkilere sahip olacağı düşünülmektedir. Teorik altyapısı açıklanan özyinelemesiz modellere özgü yöntemlerden 3SLS, 2SPLS ve 2SCML için sonuçlar elde edilmiş ve oluşturulan diyagram ve ilişki yapısı göz önüne alındığında, sağlık alanında özyinelemesiz modellerin uygulanabileceği öngörülmüştür.

Anahtar Kelimeler: Psikopatoloji, üç aşamalı en küçük kareler, iki aşamalı probit en küçük kareler, iki aşamalı koşullu en çok olabilirlik, yapısal eşitlik modellemesi, özyinelemesiz modeller, araç değişkeni

SB – 57

Diagnosis and Prognosis of COVID-19 Disease Using Routine Blood Values and LogNNet Neural Network

[1]Mehmet Tahir HUYUT, [2]Andrei VELICHKO

[1]Erzincan Binali Yıldırım Üniversitesi, Biyoistatistik ve Tıbbi Bilişim A.D, Erzincan
[2]Petrozavodsk State University,Institute of Physics and Technology, 33 Lenin Str., 185910 Petrozavodsk, Russia

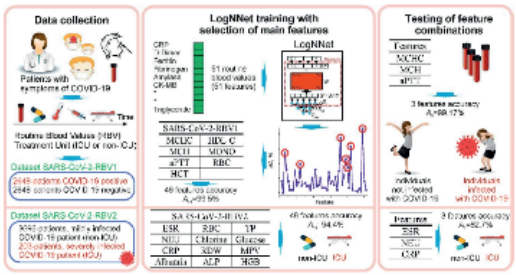
Aim: Since February 2020, the world has embarked on an intense struggle with the COVID-19 disease, and health systems have come under a tragic pressure as the disease has turned into a pandemic. The aim of this study is to determine the most effective routine blood values (RBV) in the diagnosis and prognosis of COVID-19 by using a retrospective feature elimination algorithm for the LogNNet reservoir neural network and to evaluate the classification performance of the method.

Materials and Methods: This study consists of two data sets. The first dataset includes data from 2648 patients with positive RT-PCR test results and 2648 patients with negative results. The LogNNet neural network is a feedforward network that improves classification accuracy by passing the feature vector through a special reservoir matrix and transforming it into a feature vector of different size. Previous studies have shown that the higher the entropy of a chaotic mapping that fills a reservoir matrix, the better the classification accuracy. Therefore, the procedure for optimizing chaotic map parameters plays an important role in the presented data analysis method using the LogNNet neural network. In addition, due to the characteristics of chaotic mapping, RAM usage by a neural network can be significantly reduced.

Results: The LogNNet-model achieved 99.5% accuracy in diagnosing the disease with 46 features. Moreover, the LogNNet model in diagnosing the disease achieved an accuracy of 99.17 with only three features (mean corpuscular hemoglobin concentration, mean corpuscular hemoglobin and activated partial prothrombin time). The second data set consists of the data of 3899 patients, 203 of whom were severe and 3696 of whom were mild, with a diagnosis of COVID-19 who were treated in hospital. The LogNNet model reached 94.4% accuracy in determining the prognosis of the disease with 48 features. Moreover, in predicting the prognosis of the disease, the LogNNet model achieved 82.7% accuracy with only three features (erythrocyte sedimentation ratio, neutrophil count, and C reactive protein).

Conclusion: This study provides a fast, reliable, and economic alternative mobile tool for the diagnosis and prognosis of COVID-19 based on the RBV values measured only at the time of admission to the hospital. The results of this study can be effectively used in medical peripheral devices of the IoT (IoTM) with low RAM resources, including clinical decision support systems, remote internet medicine, and telemedicine.

Keywords: COVID-19; biochemical and hematological biomarkers; routine blood values; feature selection method; LogNNet neural network; Internet of Medical Things; IoT



Work flow chart

The main stages of the study for the diagnosis and prognosis of COVID-19 using the routine blood values: data collection, LogNNet training with the selection of main features, testing combinations of the most important features that influence the diagnosis and prognosis of the disease.

SB – 58

Ölçme Araçlarının Boyutluluk Yapısının Belirlenmesinde Yapısal Eşitlik Modellemesine Dayalı Yöntemler

[1]Serhat HAYME, [2]Derya GÖKMEN, [3]Şehim KUTLAY, [4]Ayşe Adile KÜÇÜKDEVECİ

[1]Erzincan Binali Yıldırım Üniversitesi, Biyoistatistik ve Tıbbi Bilişim A.D, Erzincan
[2]Ankara Üniversitesi Tıp Fakültesi, Biyoistatistik A.D, Ankara
[3]Ankara Üniversitesi Tıp Fakültesi Fiziksel Tıp ve Rehabilitasyon A.D, Ankara
[4]Ankara Üniversitesi Tıp Fakültesi Fiziksel Tıp ve Rehabilitasyon A.D, Ankara

Amaç: Çok boyutluluk kavramsal olarak iki farklı şekilde tanımlanır: gerçek çok boyutluluk ve psikometrik çok boyutluluk. Gerçek çok boyutluluk, birden fazla boyutu farklı sorularla değerlendirmek için tasarlanmış ölçme araçlarında gözlenen çok boyutluluk türü iken; psikometrik çok boyutluluk, soruların birden fazla boyut ile ilişkili olabilmesi durumudur. Sağlık alanında kullanılan ölçme araçlarında psikometrik çok boyutluluğun her iki kaynağının varlığı, farklı modellerle değerlendirilmiştir.

Gereç ve Yöntem: Psikometrik çok boyutluluğun yapısal olarak iki kaynağı vardır. Bunlardan ilki "Soruların, tek bir yapının mükemmel göstergeleri gibi değerlendirilmesinden kaynaklı yanılabilir doğası", ikincisi ise "Değerlendirilen yapıların hiyerarşik doğası" şeklindedir. Bu çok boyutluluğun ilk kaynağı doğrulayıcı faktör analizi (DFA) ile açıklayıcı yapısal eşitlik modellemesi (AYEM) çözümlerinin karşılaştırılması ile değerlendirilirken, ikinci kaynağı ise birinci mertebe ve hiyerarşik ölçüm modellerinin karşılaştırılması ile belirlenir. Hiyerarşik ölçüm modelleri kapsamında yer alan ikinci mertebe modeller ve bifaktör modeller yardımıyla soruların altında yatan global bir yapının olup olmadığı test edilir. Her iki çok boyutluluk kaynağının bir arada olduğu durumlarda hiyerarşik modellerde AYEM modellerinin tercih edildiğini gösteren çalışmalara literatürde yaygın olarak rastlanmaktadır. Veri seti, Romatoid Artridli (RA) 270 hastanın, 109 soruluk ölçme aracına verdikleri cevaplardan oluşmaktadır

Bulgular: Faktör yapısı açıklayıcı yaklaşım altında hem birinci mertebe AYEM modeli ile hem de global yapı varlığında Bifaktör AYEM(-B-AYEM) modeli ile incelenmiştir. İlk model kapsamında veri yapısının iki faktörlü bir yapıya uygun olduğu gösterilmiş, faktörler "fiziksel durum" ve "ağrı-bilişsel-psikososyal durum" olarak alan uzmanları tarafından isimlendirilmiştir. Psikometrik çok boyutluluğun ikinci kaynağı bakımından değerlendirme yapıldığında ise, bu faktörler alan uzmanları tarafından global faktör "işlevsellik", birinci spesifik faktör "kişilerarası ve sosyal etkileşim/ilişki" ve ikinci spesifik faktör "mobiliteyle ilişkili zorluk/ağrı" olacak şekilde tanımlanmıştır.

Sonuç: Bu çalışma kapsamında elde edilen bulgular, psikometrik çok boyutluluğun her iki kaynağının da mevcut olduğunu B-AYEM modeli ile ortaya koymuştur. Daha açık olarak ifade edilirse, bu çalışma ile; AFA, DFA, AYEM ve B-AYEM, psikometrik çok boyutluluğun her iki kaynağının beraber incelenmesine olanak sağlayan genel bir çerçeve sağladığı gösterilmiştir. B-AYEM modelinin özellikle karmaşık çok boyutlu ölçme araçlarıyla çalışıldığında göz önünde bulundurulması gereken bir model olduğu düşünülmektedir.

Anahtar Kelimeler: Açıklayıcı yapısal eşitlik modellemesi, psikometrik çok boyutluluk, bifaktör model

SB – 59

The Interaction Term and Its Graphical Interpretation in Regression Models

[1]Handan ANKARALI, [2]Özge PASİN, [3]Senem GÖNENÇ

[1]İstanbul Medeniyet Üniversitesi, Tıp Fakültesi, Temel Tıp Bilimleri, Biyoistatistik Ve Tıp Bilişimi Anabilim Dalı, İstanbul
[2]Bezmiâlem Vakıf Üniversitesi, Tıp Fakültesi, Temel Tıp Bilimleri Bölümü, Biyoistatistik Ve Tıp Bilişimi Anabilim Dalı, İstanbul
[3]Atatürk Üniversitesi, Fen Fakültesi, İstatistik Bölümü, Erzurum

Aim: One of the most important steps in the models to be established in order to define the relationships between the measured characteristics in health field is to correctly define the effects. One of these effects is the interaction effect, which is rarely used in practice, but on the contrary, which is predicted to be used frequently. In this study, the concept of interaction between independent variables in a continuous structure, which is little known and used very rarely in regression-like models, is to be presented with an easily interpretable graphical result.

Materials and Methods: The data used to emphasize the interpretation and importance of the interaction in the regression model were produced by simulation based on the descriptive statistics and distribution patterns of the data in a real study. The data set includes the Systolic blood pressures of 167 people aged between 40 and 76 and a body mass index between 21 and 52. Age and body mass index were defined as independent variables and systolic blood pressure as dependent variables.

Results: In the non-interaction model, an increase in body mass index was increased systolic blood pressure when age was kept constant, and an increase in age increased systolic blood pressure when body mass index was kept constant. Although this result is sufficient, appropriate, and meaningful for the practitioner, it will not make sense without knowing the importance and meaning of the interaction between body mass index and age. When the interaction term was added to the model, the effect of 1-unit change observed in body mass index on systolic blood pressure differs at different ages and the effect of 1-year increase in age on systolic blood pressure differs according to body mass index values. In this case, it has emerged that a physician who will make a clinical decision should also consider age when deciding on systolic blood pressure according to body mass index. In addition, the contour graphic method, which will facilitate the work of the practitioners in the interpretation of the interaction, will make a significant contribution to the evaluation of this term in models.

Discussion: Faulty model design means producing faulty results. The modeling process in healthcare research involving complex relationships requires substantial knowledge, domain knowledge, modeling knowledge, and accurate interpretation of results. Examining the interaction terms is of great importance in the modeling process. If this effect is significant, the actual effects of the interacting effects are meaningless and their interpretation will yield erroneous results.

Keywords: Regression model, interaction, contour chart, independent variable

SB – 60

BKSY: Bilgi Keşfi Süreci Yazılımı

[1]Ahmet Kadir ARSLAN, [2]Cemil ÇOLAK

[1]İnönü Üniversitesi Tıp Fakültesi, Biyoistatistik ve Tıp Bilişimi Anabilim Dalı, Malatya
[2]İnönü Üniversitesi Tıp Fakültesi, Biyoistatistik ve Tıp Bilişimi Anabilim Dalı, Malatya

Amaç: Bu çalışmada, veri bilimi süreçlerine yönelik olarak bilgi keşfi süreci (BKS, Knowledge Discovery Process) tabanlı, uçtan uca veri analizi yapan kapsamlı bir web-tabanlı aracın geliştirilmesi amaçlanmıştır.

Gereç ve Yöntem: BKS, üzerinde çalışılacak veri setinde saklı olarak bulunan bilgi, örüntü ve özellikleri ortaya çıkarmak için uygulanan ve beş adımdan oluşan bir süreçtir. BKS, analize konu veri setinin ilgili veri tabanlarından temin edilmesi, veri önişlemesi, veri dönüşümü, değişken seçimi, veri madenciliği ve veri madenciliği çıktılarının değerlendirilmesi ve yorumlanması aşamalarından oluşur. BKS temelli geliştirilen yazılımda, analize konu veri setinin yazılıma yüklemesi yapılır. Veri önişleme aşamasında, veri setindeki kayıp değerlere çeşitli yöntemler kullanılarak değer ataması yapılır, çeşitli istatistiksel teknikler kullanılarak uç değerlerin tespiti ve veri setinden çıkarılması işlemi gerçekleştirilir. Veri dönüşümü aşamasında veri setine standardizasyon (z-dönüşüm) ve diğer normalizasyon dönüşümleri uygulanabilmektedir. Değişken seçimi aşamasında, Boruta veya LASSO regresyon tekniklerinden biri kullanılarak veri setindeki değişken sayısı en uygun sayıya indirgeme yapılır. Veri madenciliği aşamasında ise sınıflandırma ve regresyon görevlerinde kullanılan çeşitli makine öğrenmesi modelleri kullanılarak veri setindeki potansiyel desen ve örüntülerden bilgi çıkarma işlemi gerçekleştirilir. Yorumlama ve değerlendirme aşamasında ise veri madenciliği çıktıları yorumlanır ve modelin performansı çeşitli ölçütlerle ölçülür.

Bulgular: Geliştirilen yazılımda, kullanılan makine öğrenmesi yönteminin amacına göre (sınıflandırma veya regresyon) bulgular oluşturulmaktadır. Sınıflandırma problemlerinde genel doğruluk, duyarlılık, seçicilik, ROC eğrisi altına kalan alan vb. bulgular elde edilebilirken, regresyon problemlerinde ise ortalama karesel hata, hata kareler ortalamasının karekökü ve ortalama mutlak hata gibi bulgular sunulmaktadır. Ayrıca modellere ilişkin değişken önemlilik skorları da geliştirilen yazılımda raporlanmaktadır.

Sonuç: Geliştirilen yazılım ile, • Tıp, biyoloji ve diğer alanlara yönelik çeşitli dosya uzantılarına sahip (.xlsx, .csv, .txt, .arff, .sav, vb.) veri dosyalarının analizine imkân tanıması, • Yüklenen veri setlerinin kapsamlı önişlemesi, • Önişlemesi yapılan veri setinden belirli bir bağımlı değişkene (hastalık vb.) göre tahmin sonucu üretmesi, • Sonuç/çıktı değişkeni ile ilişkili olabilecek faktörleri modelleme öncesinde tespit etmesi, • Yapılan tahmin sonuçlarının açıklanabilmesi sayesinde ilgili veri setindeki hastalık risk faktörlerinin/biyobelirteçlerin tahmini ve değişken önemliliklerinin (variable importance) belirlenmesi, • Açık erişimli ve web-tabanlı olması nedeniyle tüm kişisel bilgisayarlarda kullanılabilir olması, • Sınıflandırma ve regresyon problemlerine uygulanabilmesi gibi özgün özelliklerinin olması nedeniyle veri bilimi ve ilişkili alanlarda çalışan araştırmacılar için faydalı bir analiz aracı olması beklenmektedir.

Anahtar Kelimeler: Bilgi keşfi süreci, shiny, web-tabanlı yazılım

Şekil-1. Keşfi Süreci Yazılımı'nın genel görünümü

SB – 61

KOAH Hastalarında Fonksiyonel Kapasite Üzerindeki Etkilerin İstatistiksel Yöntemlerle Analizi

[1]Sinan SARAÇLI, [2]Aydın BALCI, [3]Berkalp TUNCA

[1]Balıkesir Üniversitesi Tıp Fakültesi, Biyoistatistik ve Tıbbi Bilişim ABD
[2]Afyon Sağlık Bilimleri Üniversitesi Tıp Fakültesi Göğüs Hastalıkları ABD
[3]Eskişehir Osmangazi Üniversitesi Fen Bilimleri Enstitüsü İstatistik ABD

Amaç: Bu çalışmanın amacı, KOAH Hastalarında Fonksiyonel Kapasite Üzerindeki Etkili olan faktörleri İstatistiksel Yöntemlerle Analiz etmektir.

Gereç ve Yöntem: Çalışmada Afyonkarahisar Sağlık Bilimleri Üniversitesi Tıp Fakültesi Hastanesi göğüs hastalıkları polikliniğine Haziran 2021 - Haziran 2022 tarihleri arasında başvuran 120 Amfizem fenotipli KOAH hastasının fonksiyonel kapasitelerini ölçmek için 6 dakika yürüme testi, 1 dakika otur kalk testi ve 5 tekrarlı otur kalk testi uygulanarak ilgili veriler derlenmiştir. İlgili verilere ilişkin betimleyici istatistikler verildikten sonra KOAH hastalarının fonksiyonel kapasiteleri üzerinde etkili olan etmenler CRT analizi ve PLS analizi ile incelenmiştir.

Bulgular: Her bir fonksiyonel kapasite ölçüm testi için uygulanan CRT analizi sonucunda 6 dakika testine bağlı fonksiyonel kapasite üzerinde BORG indeksi, 1 dakika otur kalk testine bağlı fonksiyonel kapasite üzerinde nabız değerleri, 5 dakika otur kalk testine bağlı fonksiyonel kapasite üzerinde ise saturasyon değerleri en etkili değişkenler olarak belirlenmiştir. Tüm fonksiyonel kapasite testlerine bağlı olarak kurulan PLS modeli üzerinde ise BORG indeksi, diğer değişkenlere göre en etkili değişken olarak belirlenmiştir.

Sonuç: KOAH hastalarının 6 dakika yürüme testi sonucunda, BORG indeksi için kritik değer 4,5 olarak belirlenmiş, BORG indeksi 4,5'in altında olan hastaların fonksiyonel kapasitesi diğerlerine göre daha yüksek olarak belirlenirken, BORG indeksi 4,5'in üzerinde olan hastalar için saturasyon değeri anlamlı olarak bulunmuş, saturasyonu 89'un üzerinde olan hastaların fonksiyonel kapasitesinin diğerlerine göre yüksek olduğu belirlenmiştir. 1 dakika otur kalk testi sonucunda, nabız değeri 139,5'in üzerinde kalanların fonksiyonel kapasitesi diğerlerine göre daha yüksek olarak belirlenmiş, nabız değeri 139,5'in altında kalan nabız değerleri üzerinde BORG indeksi etkili bulunmuş ve BORG indeks değeri 3,5'in altında kalan hastaların fonksiyonel kapasitesinin daha yüksek olduğu belirlenmiştir. 5 dakika otur kalk testi sonucunda, saturasyon değişkeni en etkili değişken olarak belirlenmiş, saturasyon değeri 89,5'in altında olan hastaların fonksiyonel kapasitelerinin değerlerine göre daha yüksek olduğu belirlenmiştir. Tüm fonksiyonel kapasite testlerine bağlı olarak kurulan PLS modeli sonucunda ise, BORG indeksindeki 1 birimlik artışın fonksiyonel kapasite üzerinde 0,42'lik katsayı ile azalışa sahip olan en etkili değişken olduğu belirlenmiştir.

Anahtar Kelimeler: Koah, fonksiyonel kapasite, CRT, PLS.

SB – 62

Türkiye Covid-19 Verileri için Parçalı Fonksiyona Dayalı Büyüme Modelleri

[1]Leyla BAKACAK KARABENLI, [2]Serpil AKTAŞ ALTUNAY

[1]Hacettepe Üniversitesi, İstatistik Bölümü, Ankara
[2]Hacettepe Üniversitesi, İstatistik Bölümü, Ankara

Amaç: Türkiye'nin Covid-19 verilerindeki ölüm sayısı belli bir trend içinde olmayıp dalgalı, sürekli artan ve azalan bir trend göstermekte ve homojen olmayan bir dağılım sergilemektedir. Bu nedenle, tek bir büyüme modeli üzerinden ölüm sayısını tahmin etmek zorlaşmaktadır. Bu çalışmada, çeşitli büyüme modelleri dikkate alınarak Covid-19 verileri üzerinden ölüm sayılarını tahmin etmek amaçlanmıştır.

Gereç ve Yöntem: Büyüme modellerinde tek bir model kullanarak ölüm sayısını tahmin etmek yerine, dalgalanmaları da dikkate alan parçalı bir fonksiyon tanımlanmıştır. Bu fonksiyonlar log-üstel ve log-Gompertz modeller ile tanımlanmıştır. Parçalı fonksiyonların belirlenebilmesi için değişim noktaları tahmin edilmiştir. Böylece parçalı büyüme fonksiyonları sayesinde veri seti modellenmiştir. Değişim noktalarının tahmininde bilgi kriterlerinden yararlanılmıştır.

Bulgular: İncelenen zaman aralığında farklı eğilimlere sahip Covid-19 verileri, değişim noktalarına göre farklı büyüme fonksiyonları ile temsil edilmiştir. Bu yöntem tahminlerde daha fazla esneklik sağlamış ve daha iyi tahmin eğrilerinin elde edilmesine olanak vermiştir. 17/03/2020-21/03/2021 tarihleri arasındaki birbirinden farklı eğilimlere sahip kümülatif ölüm sayıları, beş parçalı log-Gompertz modeli ile temsil edilmiştir.

Sonuç: Bu çalışmada ölüm sayıları incelenmiş ve verilerdeki dalgalanmalardan dolayı ölüm sayılarını tahmin etmek için tek bir modelin yeterli olmayacağı düşünülmüştür. Bu nedenle verilerin değişim noktaları araştırılmış ve bu noktalara göre en uygun büyüme modelleri seçilmiştir.

Anahtar Kelimeler: Büyüme modelleri, parçalı fonksiyon, Türkiye Covid-19 verileri

SB – 63

Recurrent Neural Network for Complex Survival Problems

[1]Pius MARTHIN, [2]Nihal ATA TUTKUN

[1]Hacettepe university
[2]Hacettepe university

Aim: Survival analysis has become one of the paramount procedures in the modeling of time-to-event data. When we encounter complex survival problems, the traditional approach remains limited in accounting for the complex correlational structure between the covariates and the outcome due to the strong assumptions that limit the inference and prediction ability of the resulting models. There exist several studies on the deep learning approach to survival modeling nevertheless, the application for the case of complex survival problem remain limited. In addition, the existing models do not fully address the complexity of the data structure and are subject to noise and redundancy information. In this study, we design a deep learning technique (CmpXRnnSurv_AE) that obliterates the limitations imposed by traditional approaches and addresses the above issues to jointly predict the risk-specific probabilities and survival function for recurrent events with competing risks.

Materials and Methods: We introduce the component termed Risks Information Weights (RIW) as an attention mechanism to compute the weighted cumulative incidence function (WCIF) and an external auto-encoder (ExternalAE) as a feature selector to extract complex characteristics among the set of covariates responsible for the cause-specific events.

Results: We train our model using synthetic and real data sets. For the performance evaluation, we compute the time-dependent concordant index for complex survival models. We select both traditional and machine learning models as benchmarks. Our model demonstrates better performance across all datasets with the maximum concordance indices of 0.86, 0.89, and 0.92 on the synthetic, MIMIC-III clinical, and the UNOS-OPTN(KIDPAN) datasets accordingly.

Conclusion: We developed a novel deep learning model termed CmpXRnnSurv-AE for the complex survival problem. Using the external autoencoder we addressed the noise invariant problem in the complex survival data. By computing the weighted cumulative incidence function (WCIF) we managed the complex correlation structure in recurrent events with competing risks survival data. Since deep learning models are overparameterized and deterministic in nature, the point estimates provided are not reliable for uncertainty measurements in the predictions. In addition, the RNN is subjective to temporal inconsistency. In future work, we propose to apply the Historical Consistent Neural Networks (HCNN) in the Bayesian framework to quantify uncertainties in the model's predictions for the case of complex survival problems.

Keywords: Cumulative incidence function (CIF) Risk Information Weight (RIW), Autoencoders (AE), Survival analysis, Recurrent events with competing risks, Recurrent neural networks (RNN), Long short term memory (LSTM), Self attention, Multilayers perceptrons.

KS – 1

Önceden Eğitilmiş Evrişimsel Sinir Ağlarından Elde Edilen Öznitelikler ile Sperm Morfolojisinin Julia Tabanlı Web Uygulaması ile Sınıflandırılması

[1]Hüseyin KUTLU, [2]Emek GÜLDOĞAN, [3]Cemil ÇOLAK

[1]Adıyaman Üniversitesi
[2]Adıyaman Üniversitesi
[3]Adıyaman Üniversitesi

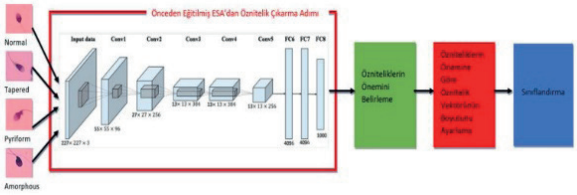
Amaç: Derin öğrenme yöntemleri, eğitim esnasında, binlerce, hatta milyonlarca parametre güncellemesi yaparlar. Dolayısıyla büyük miktarda eğitim verisine ihtiyaç duyarlar. Aksi bir durumda test veri kümesinde düşük doğruluklu, örneklem dışı uygulamada yanlış sonuçlar veren bir model ortaya çıktığı görülür. Tıbbi imgeler gerek yeteri kadar etiketlendirilmiş veri olmayışı, gerek negatif örneğin pozitif örneğe göre düşük oranda oluşu sebebi ile bu durumdan mustariptir. İmge işlemede bu tür problemler için başka bir veri seti ile önceden eğitilmiş bir evrişimsel sinir ağı (ESA) kullanmaktadır. Bu çalışmada erkek infertilitesinin teçhiz ve tedavisinde önemli bir yere sahip olan sperm baş morfolojisi imgelerinin önceden eğitilmiş ESA mimarilerinden elde edilen özniteliklerle sınıflandırması amaçlanmıştır.

Gereç ve Yöntem: Çalışmada halka açık bir veri kümesi olan HuSHem sperm baş morfolojisi veri setinden elde edilen imgeler kullanılmıştır. İmgelere herhangi bir ön işleme ve çoğaltma uygulanmadan, imgeler önceden eğitilmiş ESA mimarilerinden (squeezenet, googlenet, inceptionv3, densenet201, mobilnetv2, resnet18, resnet50, resnet101, xception, inceptionresnetv2, shufflenet, nesnetmobile, nasnetlarge, darknet19, efficientnet, alexnet, vgg16, vgg19) geçirilmiş ESA'ların son katmanlarından imgelerin öznitelikleri alınmıştır. Her imgeden elde edilen 1000 adet özniteliğin önem değerleri Lojistik Regresyon katsayıları temelli öznitelik önemi belirleme yöntemi, Ağaç tabanlı XGBoost algoritması gini indeksi temelli öznitelik önemi belirleme yöntemi ve Temel Bileşenler Analizi (Principal Component Analysis – PCA) öznitelik yük değeri (factor loading) temelli öznitelik önemi belirleme yöntemi ile hesaplanmıştır. Öznitelikler öznitelik önemi belirleme yöntemi ile indirgenerek sınıflandırılmıştır.

Bulgular: Önceden eğitilmiş ESA modellerinden sperm morfolojik imgelerinin en etkili öznitelikleri GoogleNet ESA mimarisinden elde edilmiştir. Öznitelik önem değerlerine göre öz nitelikler önem sırasına göre sıralanmış ve en önemli öznitelikten başlayarak bir artırımla öznitelik alt kümeleri sınıflandırılmıştır. Sınıflandırmada İnönü Üniversitesi Biyoistatistik Anabilim dalının Geliştirdiği Julia tabanlı web uygulaması kullanılmıştır. Lojistik regresyon katsayıları ile elde edilen öznitelikler 161 öznitelikte en yüksek (%81.01) sınıflandırma doğruluğuna ulaşmıştır. XGBoost algoritmasından elde edilen öznitelik değerleri 248 öznitelikte en yüksek (%79.13) sınıflandırma doğruluğuna ulaşmıştır. PCA faktör yük değerlerinden (elde edilen öznitelik önemleri 403 öznitelikte en yüksek sınıflandırma doğruluğuna (%79.58) ulaşmıştır.

Sonuç: İmge sayısının yetersiz olduğu problemlerde derin öğrenmenin avantajlarından faydalanmak için Önceden eğitilmiş ESA mimarilerinin kullanılabilir bir yöntem olduğu görülmüştür. Derin özniteliklerin önem değerini belirlemede Lojistik regresyon katsayıları diğer yöntemlere göre öne çıkmıştır.

Anahtar Kelimeler: Derin Öznitelik Çıkarımı, Önceden eğitilmiş ESA'lar, Öznitelik Önemi, Julia, Sınıflandırıcı



Sperm Önerilen Yöntemin Adımlarını Gösteren Diyagram

KS – 2

Lökosit Antijen Molekülleri Toplum Doku Tipi Sıklığı

[1]Kemal OLÇA, [2]Ahmet KESKİNOĞLU, [3]Duygu KARAKAYA, [4]Ebru ÖZTÜRK, [5]Pembe KESKİNOĞLU, [6]Mehmet ORMAN, [7]Timur KÖSE, [8]Ahmet DİRİCAN, [9]Erdem KARABULUT

[1]İstanbul Üniversitesi Lisansüstü Eğitim Enstitüsü, Cerrahpaşa Tıp Fakültesi Biyoistatistik ve Tıp Bilişim AD, İstanbul
[2]Ege Üniversitesi Tıp Fakültesi, Çocuk Sağlığı Nefroloji AD, İzmir
[3]Dokuz Eylül Üniversitesi Tıp Fakültesi, Biyoistatistik ve Tıbbi Bilişim AD, İzmir
[4]Hacettepe Üniversitesi Tıp Fakültesi Biyoistatistik AD, Ankara
[5]Dokuz Eylül Üniversitesi Tıp Fakültesi, Biyoistatistik ve Tıbbi Bilişim AD, İzmir
[6]Ege Üniversitesi Tıp Fakültesi Biyoistatistik ve Tıbbi Bilişim AD, İzmir
[7]Ege Üniversitesi Tıp Fakültesi Biyoistatistik ve Tıbbi Bilişim AD, İzmir
[8]İstanbul Üniversitesi Lisansüstü Eğitim Enstitüsü, Cerrahpaşa Tıp Fakültesi Biyoistatistik ve Tıp Bilişim AD, İstanbul
[9]Hacettepe Üniversitesi Tıp Fakültesi Biyoistatistik AD, Ankara

Amaç: Lökosit antijen molekülleri (human leukocyte Antigen-HLA) en çeşitli ve polimorfik molekülerdendir. HLA molekülleri, kromozom 6'nın kısa kolunda bulunur, temel işlevi, yabancı cisimlerin tanınması ve immünojenisitedir. Yapısal ve fonksiyonel özelliklerine göre 3 sınıfa ayrılır, sınıf 1: HLA-A, -B, ve sınıf 2; HLA-DR, -DQ antijenleri önemlidir. Toplumlara göre HLA sıklıkları değişmektedir. HLA doku tipi sıklıkları, antropolojik olarak değerlendirmeler dışında hastalıklara yatkınlık ve organ naklindeki rolü de önemli araştırma konusudur. Ülkemizde hastalıklara yatkınlık ile ilgili sınırlı sayıda hasta popülasyonlarında HLA doku tipleri çalışılmıştır, ayrıca donörler üzerinden sağlıklı kişileri içeren HLA sıklıklarının da araştırıldığı görülmektedir. Ancak bu çalışmalarda toplam "n" sayısı çok sınırlıdır. Bu çalışmada, ülkemizin her yerinden hasta ve yakınlarının başvurduğu bir üniversite hastanesi doku tipi laboratuvarında donörlerin HLA doku tipi sıklıkları değerlendirilmiştir.

Gereç ve Yöntem: Retrospektif olarak, ülkemizin batısında ve organ nakli çok fazla gerçekleştiren donör sayısı çok yüksek olan üniversite hastanesi doku tipi laboratuvarının 40 yıllık verisi taranmıştır. Bu verilerim çoğunluğuna elektronik ortamda düz metin şeklinde elde edilmiştir. Bu metin raporlar kod yazılarak HLA-A, B ve DR için ikişer lokus bilgisi ayrılmış ve veri tabanına ayrı 6 sütun olarak kaydedilmiştir. Her üç HLA doku tipi için allel frekansları belirlendi ve Hardy Weinderg eşitliğime (HWE) uyumları da test edildi.

Bulgular: Doku tipi laboratuvar verilerinde toplam 17130 donörün HLA bilgisine ulaşıldı. HLA-A doku tipleri en sık A*02(%22), A*24 (%16,4), en sık HLA-B için B-35 (%17,4) ve HLA-DR için DRB1*11 (%22,6) ve DRB1-04 (%14) olarak saptandı.

Sonuç: Yüksek sayıdaki donör HLA verilerinden A, B ve DR için allel sıklıkları toplum doku sıklığı olarak değerlendirilebilir. Göçler, epigenetik etkilenimler vb. sonuçlar nedeni ile çok polimormik olduğu bilinen HLA tipleri belirli aralarla rutin hizmet verilerinden sıklıklarının belirlenmesi bu doku tiplerinin önemli olduğu pek çok hastalık için gerekli olduğu düşünülmüştür. Teşekkür: Bu çalışma 318S166 nolu TÜBİTAK-SBAG projesi verilerinden gerçekleştirilmiştir.

Anahtar Kelimeler: Lökosit Antijen Molekülleri, HLA, Doku Tipi, Hardy Weinderg Eşitliği, HWE

KS – 3

Mamografi Görüntülerinde K-ortalamalar Kümeleme ve Canny Kenar Algılama Bölütleme Yöntemlerinin Sınıflandırma Performanslarının Karşılaştırılması

[1]Jale KARAKAYA, [2]Hanife AVCI

[1]Hacettepe Üniversitesi Tıp Fakültesi, Biyoistatistik Anabilim Dalı, Ankara
[2]Hacettepe Üniversitesi Tıp Fakültesi, Biyoistatistik Anabilim Dalı, Ankara

Amaç: Bu çalışmanın amacı, mamografi görüntülerini çeşitli ön işleme yöntemleriyle iyileştirdikten sonra, görüntülerden çıkarılan özniteliklerin hastalık durumunu belirlemede iki farklı bölütleme yönteminin başarısını araştırmaktır.

Yöntem: Çalışmada açık erişimli mini-MIAS veri tabanı kullanıldı. Bu veri seti, 161 hastanın sağ ve sol meme görüntülerini içeren 322 sayısallaştırılmış mamografi görüntüsünden oluşmaktadır. Çalışmada ön işleme yöntemleri olarak sırasıyla medyan filtre, kontrast sınırlı uyarlanabilir histogram eşitleme (CLAHE) ve keskin olmayan maskeleme görüntü iyileştirme teknikleri kullanılmıştır. Ön işleme yöntemleriyle etiketlerden ve yapay çizgilerden arındırılmış mamografi görüntüleri için bir sonraki önemli adım, uygun segmentasyon yöntemiyle görüntülerden pektoral kas vb. temizleyerek ilgi bölgesi (region of interest, ROI) belirlemektir. Bu çalışmada, ön işleme yöntemlerinin performansını incelemek için k-ortalamalar kümeleme algoritması seçilmiştir. Daha sonra, ikinci bir bölütleme yönteminin görüntülerden çıkarılan öznitelikler üzerindeki etkisini görmek için mamografi görüntülerine kenar tabanlı görüntü bölütleme yöntemlerinden biri olan Canny kenar algılama algoritması uygulanmıştır. Görüntü bölütleme algoritmalarının uygulanmasından sonra elde edilen görüntüler, orijinal görüntüler üzerine maskeler olarak uygulanmakta ve daha sonra gerçek zemin görüntüleri elde edilmektedir. Mamografi görüntülerine görüntü ön işleme ve bölütleme işlemleri uygulandıktan sonra özellik çıkarımı yapılmıştır. Gri seviye eş oluşum matrisi (GLCM) ve Gri seviye çalışma uzunluğu matrisi (GRLM) teknikleri kullanılarak ROI örneklerinden özellikler çıkarılmıştır. Korelasyon matrisi ile öznitelik seçimi yapıldıktan sonra bu özniteliklerin sınıflandırma başarısına etkileri 6 farklı klasik veri madenciliği yöntemi ile incelenmiştir. Görüntü işleme sonucu elde edilen sayısal veriler sınıflandırma yöntemlerinde giriş değişkeni olarak kullanılmıştır. Bu çalışmada hesaplanan özellikler ile ROI örnekleri normal-anormal doku olarak sınıflandırılmıştır. Sınıflandırma yöntemlerinin performansları, doğruluk, duyarlılık gibi ölçüler kullanılarak karşılaştırılmıştır.

Bulgular: Normal-anormal doku sınıflandırması için Canny kenar algılama algoritması, tüm sınıflandırma yöntemlerinde k-ortalamalar kümeleme algoritmasından daha düşük performans göstermiştir. Bu sonuçlara göre doğruluk, duyarlılık, eğri altında kalan alan (AUC) gibi ölçülere baktığımızda normal-anormal sınıflandırmasında Destek Vektör Makinesi, Rastgele Orman, Yapay Sinir Ağı ve Naive Bayes sınıflandırma yöntemlerinin performansları k-en yakın komşuluk ve Karar Ağacı göre daha başarılıdır.

Sonuç: Bu çalışmada, kenar tabanlı görüntü bölütleme tekniği olan Canny kenar algılama yöntemi ile bölge tabanlı görüntü bölütleme tekniği olan k-ortalamalar kümeleme yönteminin sınıflandırma performansları karşılaştırılmıştır. K-ortalamalar kümeleme algoritmasının Canny kenar algılama algoritmasına göre daha yüksek bir sınıflandırma özelliğine sahip olduğu tespit edilmiştir.

Anahtar Kelimeler: Mamografi görüntüleri, bölütleme yöntemleri, sınıflandırma performansı.

Bu çalışma Hacettepe Üniversitesi Bilimsel Araştırma Projeleri Koordinasyon Birimi tarafından desteklenmiştir. Hızlı Destek Projesi (THD-2020-18735)

KS – 4

Log Multinomial Regresyon Modeline Hosmer-Lemeshow Uyum İyiliği Testinin Uyarlanması

[1]Yasemin ÖZTÜRK, [2]Erdem KARABULUT

[1]T.C. Sosyal Güvenlik Kurumu, Genel Sağlık Sigortası Genel Müdürlüğü, Ankara
[2]Hacettepe Üniversitesi Tıp Fakültesi, Biyoistatistik Anabilim Dalı, Ankara

Amaç: İki durumlu veya multinomial lojistik regresyon modelinin uyumunun değerlendirilmesinde Hosmer ve Lemeshow tarafından önerilen uyum iyiliği testi kullanılmaktadır. İleriye yönelik çalışmalarda göreli risk kestirimi elde edilebildiğinden, iki kategorili bağımlı değişken olduğunda log-binomial ve 3 ya da daha fazla kategorili bağımlı değişken olduğunda log-multinomial regresyon modelinin kullanılması önerilmektedir. Ancak, log-multinomial regresyon modelinin veriye uyumunun değerlendirilmesine yönelik herhangi bir metot bulunmamaktadır. Bu çalışmada Hosmer Lemeshow uyum iyiliği testi log multinomial regresyon modeline uyarlanmış ve benzetim çalışması ile performansı değerlendirilmiştir.

Gereç ve Yöntem: Uyarlanan Hosmer Lemeshow uyum iyiliği testinin uygulanabilirliği çeşitli senaryolar altında benzetim çalışması ile incelenmiştir. Söz konusu senaryolarda bağımlı ve bağımsız değişkenlerin farklı yapıları, çeşitli örneklem genişliklerinde benzetim çalışması yardımıyla karşılaştırılmıştır. Benzetim çalışmasında, ikiden fazla kategorili bağımlı kategorik değişkenler, sürekli ya da kategorik bağımsız değişkenler ve farklı örneklem genişlikleri (200, 400, 600, 800 ve 1000) dikkate alınarak toplamda 70 farklı senaryo kullanılmıştır. Benzetim senaryolarındaki modellerde kullanılan regresyon katsayıları, kullanılan gerçek veri setlerinde yer alan değişkenler yardımı ile kestirilen log-multinomial modellerden elde edilmiştir. Her bir senaryo 1000 kez tekrar edilmiş ve her bir tekrarda uyum iyiliği testi sonuçları tip I ve tip II hata sıklıkları elde edilerek değerlendirilmiştir. Belirlenen senaryolar için benzetim çalışması, modelleme ve uyum iyiliği testinin sonuçlarının incelenmesi R yazılımında yer alan "RStata", "general hoslem" ve "MASS" paketleri ve STATA yazılımı kullanılarak gerçekleştirilmiştir.

Bulgular: Uyarlanan Hosmer-Lemeshow testi, sürekli sayısal bağımsız değişkene ilişkin karesel terim ve bağımsız değişkenler arasındaki etkileşim teriminin dahil edilmemesi gibi çeşitli uyumsuzlukları belirleyebildiği görülmektedir. Tüm senaryolarda örneklem büyüklüğü arttıkça tip I ve tip II hata oranının düştüğü ve çoğunlukla nominal düzeyin altında olduğu sonucu elde edilmiştir.

Sonuç: Hosmer-Lemeshow uyum iyiliği testinin log multinomial regresyon modelinin uyum iyiliğinin tespit edilmesinde kullanılabilir yapıda olduğu görülmektedir.

Anahtar Kelimeler: Log multinomial regresyon, Hosmer Lemeshow test istatistiği, uyum iyiliği.

KS – 5

Sobel ve Baron-Kenny Yöntemlerinin Performanslarının Karşılaştırılması

[1]Hülya BİNOKAY, [2]Yaşar SERTDEMİR

[1]Çukurova Üniversitesi Tıp Fakültesi Biyoistatistik Anabilim Dalı, Adana
[2]Çukurova Üniversitesi Tıp Fakültesi Biyoistatistik Anabilim Dalı, Adana

Amaç: Bu çalışmada, Sobel ve nedensel adım yaklaşımı yöntemlerinin farklı koşullar altında performanslarının karşılaştırılması amaçlanmıştır. Ayrıca literatürden alınan gerçek bir veri setinde sonuçlar değerlendirilecektir

Gereç ve Yöntem: Farklı örnek büyüklüğü (100,150,200,250,300,350,400) ve PM=(dolaylı etki)/(toplam etki)=(a*b)/c (aracı değişken oranı)(0.05,0.10,0.15,0.20,0.25) olduğu durumlar için normal dağılımdan veri üretilmiştir. Y=β0+cX+e1 M=β0′+aX+e2 Y=5+β0″+c′ X+bM+e3 Farklı örnek büyüklüğü ve aracı değişken oranı için 4 farklı senaryo ele alınmıştır. Senaryo 1: sadece a katsayısı anlamlı, Senaryo 2: sadece b katsayısı anlamlı, Senaryo 3: a ve b katsayılarının ikisi de anlamlı değil, Senaryo 4: a ve b katsayılarının ikisi de anlamlı

Bulgular: Aracı değişken oranları için, senaryo 3 te, M değişkeninin, nedensel adım yaklaşımı ve Sobel yöntemlerine göre aracı değişken olmadığı gözlenmiştir. Senaryo 1, 2 ve 4 te, Sobel yöntemine göre M aracı değişken olduğu kararı verilirken, nedensel adım yaklaşımı yöntemine göre aracı değişken olmadığı kararı verilmektedir. Örnek büyüklüğü arttıkça, M aracı değişkeninin sobel yöntemine göre aracı değişken olduğu, nedensel adım yaklaşımı yöntemine göre aracı değişken olmadığı gözlenmiştir.

Sonuç: Farklı senaryolar incelendiğinde, Sobel yöntemine göre aracı değişken olan M değişkeni, nedensel adım yaklaşımına göre aracı değişken değildir. Böyle bir durumda bir değişkenin aracı değişken olup olmadığına karar vermek net değildir. Bundan dolayı yeni bir yöntemin önerilmesi gerekmektedir.

Anahtar Kelimeler: Aracılık, Sobel, Baron-Kenny

KS – 6

Lojistik Regresyon Modeli ile Kardiyovasküler Kalp Hastalığı Sınıflandırılması ve Tahmini

Mehmet KIVRAK

Recep Tayyip Erdogan Üniversitesi Biyoistatistik ve Tıp Bilişimi

Amaç: Kardiyovasküler hastalıklar dünya çapında önde gelen ölüm nedenidir. Bu nedenle kalp krizlerini etkili bir şekilde öngörmek için nedenlerin belirlenmesi ve bir sistem geliştirilmesi gerekmektedir. Bu çalışma, lojistik regresyon yöntemi ile açık erişimli kardiyovasküler kalp rahatsızlığı verilerini sınıflandırmayı ve etkili faktörleri tahminlemeyi amaçlamaktadır.

Gereç ve Yöntem: Çalışmada açık erişimli kardiyo-kalp veri seti kullanıldı. Veri seti, yaş, cinsiyet, göğüs ağrı tipi, dinlenme anındaki kan basıncı, serum kolestrol düzeyi, açlık kan şekeri, elektrokardiyografi sonuçları, kalp atış hızı, egsersize bağlı göğüs ağrısı, ST segmentinin yükselmesi, floroskopi ile boyalı damar sayısı, talasemi gibi risk faktörlerinden oluşmaktadır. Hastalığı sınıflandırmak için lojistik regresyon modeli kullanıldı. Model performansı için doğruluk, duyarlılık, seçicilik, pozitif tahmin değeri ve negatif tahmin değeri performans metrikleri değerlendirildi. Model anlamlılığı için Hosmer-Lemeshow testi, katsayıların değerlendirilmesinde Odds oranı ve uyum iyiliği için Cox-Snell R2 değerleri incelendi.

Bulgular: Deneysel bulgulara göre model anlamlılığı sağlanırken (p=0.03) cp2, cp3'ün hastalığı artırıcı, cinsiyet1, ca1, ca2, ca3'ün ise hastalığı azaltıcı etkisi bulunmaktadır. Y aş, cp1, trestbps, kolestrol,fbs1, restecg1, restecg2, thalach, exang1, oldpeak, slope1, slope2, ca4, thal1, thal2, thal3 değişkenlerinin ise hastalığa anlamlı etkisi görülmedi. Cox-Snell R2 değeri (bağımsız değişkenlerin bağımlı değişkendeki değişimi açıklama yüzdesi) ise yeterli düzeydedir (R2 =0.544). Modelin sınıflandırma performansına bakıldığında, 5 katlı çapraz geçerlilik ile parametre optimizasyonu yapıldı (Grid search taramalı yüzeysel). Modelleme sonuçlarından elde edilen doğruluk, duyarlılık, özgüllük, pozitif prediktif değer, negatif prediktif değeri sırasıyla % 88.4, % 88.0, % 83.3, % 92.7, % 90.6 ve % 86.9 olarak bulundu.

Sonuç: Bu araştırmanın bir sonucu olarak, lojistik regresyon yöntemi ile açık erişimli kardiyovasküler kalp rahatsızlığı verilerini sınıflandırma ve etkili faktörleri tahminleme yapıldı. Gelecekte, bu tarz yöntemlerin benzer çalışmalar ile güvenilirliği doğrulanarak bu yöntemlere dayalı terapötik prosedürler oluşturulabilir ve klinik pratikteki yararları belgelenebilir.

Anahtar Kelimeler: Kardiyovasküler Kalp Hastalığı, Sınıflandırma, Tahmin, Lojistik Regresyon.

KS – 7

Çok Boyutlu Ölçekleme Yöntemi ile Trakya'da Tarımsal Üretim Bölgelerinde Ağır Metal Yoğunluğunun Gösterimi

[1]Büşra AYDIN, [2]Özgecan KORKMAZ AĞAOĞLU, [3] Ismayıl SAFA GURCAN

[1]Ankara Üniversitesi Veteriner Fakültesi, Biyoistatistik Anabilim Dalı
[2]Ankara Üniversitesi Veteriner Fakültesi, Biyoistatistik Anabilim Dalı
[3]Ankara Üniversitesi Veteriner Fakültesi, Biyoistatistik Anabilim Dalı

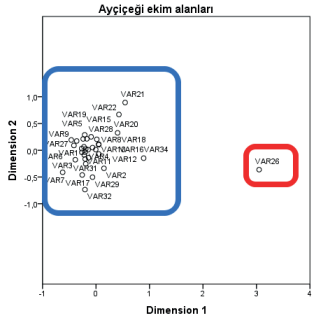
Amaç: Bu çalışma ile havzada tarım yapılan farklı bölgelerden alınan Ayçiçek, Buğday ve Çeltik örneklerinden ağır metal değerlerine göre, bölgeler arası ağır metal yükü bakımından meydana gelen kümelenme, X-Y düzlemi üzerinde incelenmiştir.

Gereç ve Yöntem: Tanımlayıcı istatistikler kullanılarak ağır metal yükleri hakkında bilgiler ışığında uzaklık matrisleri Eucledian uzaklığı kullanılarak çok boyutlu ölçekleme ile uyum iyiliği ve stress değerleri hesaplanarak düzlem üzerinde bölgelerin işaretlemesi yapılmıştır.

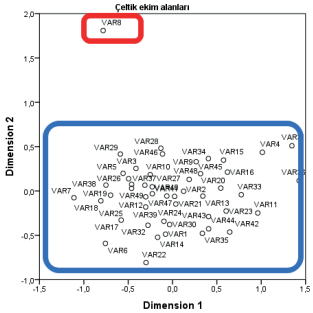
Bulgular: Varyasyon katsayısı yüksek bulunan ağır metaller Buğday için; sırası ile Alüminyum (%188), Krom (%162), Arsenik (%149), Kurşun (%138), Kobalt (%117) ve Demir (%115) ve Nikel (%92), Ayçiçeği için; Alüminyum (%105), Kurşun (%69), Arsenik (%66) ve Kalay (%56), Çeltik için Nikel (%282),ve Kadmiyum (%65) olmuştur. Çok Boyutlu Ölçekleme yöntemi ile Buğday için yapılan hesaplamalar sonucunda elde edilen koordinat noktalarına göre Yeniköy ve Merkez ve Bayramlı tarım bölgeleri diğer tarım bölgelerinden uzak kalmıştır. Ayçiçeği tarımı yapılan alanlarda Kırcasalih bölgesi düzlem üzerinde ana kümeden uzakta yer almıştır. Çeltik ekimi yapılan bölgelerde; Uzunköprü Merkez bölgesi küme oluşturan diğer bölgelerden uzak konumlanmıştır.

Sonuç: Genel olarak sanayi atıklarının yer aldığı havzada belirli dönemlerde nehir taşması yaşanan dönemlerde sulak alan ve yakınlarında tarımsal bölgelerde üretimi yapılan Buğday, Ayçiçeği ve Çeltikte saptanan ağır metal düzeylerinin, iç kesimlerde tarım üretimi yapılan bölgelere göre daha yüksek düzeyde olduğu söylenebilir.

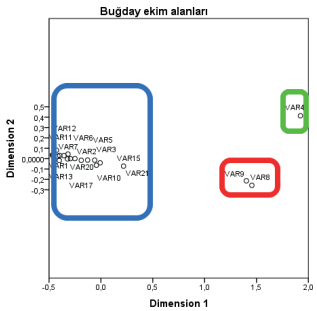
Anahtar Kelimeler: Ağır metaller, çok boyutlu ölçekleme, halk sağlığı, tarımsal bitki



Ayçiçeği için çok boyutlu ölçekleme haritası



Çeltik için çok boyutlu ölçekleme haritası



Buğday için çok boyutlu ölçekleme haritası

KS – 8

Diyabetik Gözlerde Multifokal Elektroretinografide Halka Amplitüdünün ve Halka Amplitüdünü Etkileyen Faktörlerin Değerlendirilmesi

[1]Ayşe Büşra GÜNAY ŞENER, [2]Hidayet SENER

[1]Erciyes Üniversitesi Tıp Fakültesi, Biyoistatistik ve Tıp Bilişimi ABD, Kayseri, Türkiye
[2]Erciyes Üniversitesi Tıp Fakültesi, Göz Hastalıkları ABD, Kayseri, Türkiye

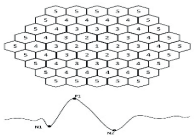
Amaç: Multifokal ERG (MFERG), retinanın elektrofizyolojik yanıtların topografik olarak değerlendirilmesine izin verir. Diyabetik hastalarda halka genliklerini değerlendirmeyi ve diyabetik retinopati için risk faktörlerinin halka genlikleri üzerindeki etkisini değerlendirmeyi amaçladık. Ayrıca halka genliklerinin tanısal değerini araştırmayı amaçladık.

Gereç ve Yöntem: Çalışmada retinopati olmayan (DM) 32 diyabetik hastanın 32 gözü, hafif proliferatif olmayan diyabetik retinopatisi (NPDR) olan 34 hastanın 34 gözü ve 62 kontrolün (CG) 62 gözü dahil edildi. MFERG kayıtları Uluslararası Klinik Elektrofizyoloji Derneği (ISCEV) önerileri›ne göre yapıldı. Halka genlikleri ve ring alanları Şekil 1 de tanımlanmıştır. Halka genlikleri tek yönlü varyans analizi ve posthoc Scheffe testi karşılaştırıldı. Diyabet ve hipertansiyonun halka genlikleri üzerindeki etkisi iki yönlü ANOVA ile değerlendirildi. Yaş, diyabet süresi, HBA1C ve halka genlikleri arasındaki ilişkiyi belirlemek için stepwise multivariate regresyon analizi yapıldı. Halka genliklerinin tanısal değerini değerlenditmek için üç yönlü bir ROC analizi yapıldı. Analiz sonucunda, ROC yüzeyi altındaki hacim (VUS) hesaplandı. VUS değeri 0.17'den büyükse, test başarılı olarak kabul edilmiştir.

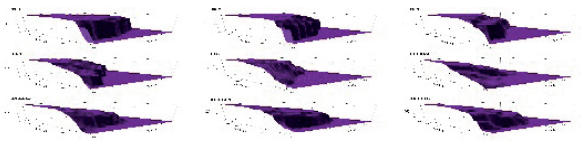
Bulgular: Halka genliklerinin (R1-R5) ve halka genlik oranlarının karşılaştırılmasında, DM ve NPDR gruplarının verileri CG'ye kıyasla anlamlı olarak azalmıştır. Diyabetin genlikler üzerinde etkili olduğu bulunmuştur, ancak hipertansiyon ve diyabet-hipertansiyonun etkileşimi etkili değildir. Ղ2 değerlendirildiğinde, diyabetin etkisinin periferik halkalara doğru azaldığı bulunmuştur. HbA1c, tüm halka genlikleri (R1-R5) ile korelasyon göstermiştir ve en etkili faktördür. Öngörücü değeri periferik halkalara doğru azalmıştır. Diyabet süresi ilk üç halka genliği (R1-R3) ile korelasyon gösterdi. Yaş sadece R1 ile ilişkiliydi. Üç yönlü ROC analizinde değerlendirilen tüm parametrelerin (halka genlikleri ve oranları) VUS değeri 0.17›den büyüktü ve anlamlıydı. VUS(R1) = 0.65, P <0.001 ve [0.50-0.72] % 95 güven aralığına sahipti ve üç sınıfı ayırt etmek için en iyi kestirim noktasının C1 = 217.3, C2 = 151.2 olduğu bulundu (senstivite 0.72, spesifite 0.59). VUS sonuçları Şekil 2 de sunulmuştur.

Sonuç: Halka oranı karşılaştırması, diyabetin şiddeti arttıkça retina nörodejenerasyonunun arttığını ve merkezi topografik lokalizasyonun daha fazla etkilendiğini ve daha objektif kanıtlar sağladığını göstermiştir. Tüm halka genlikleri diyabet nedeniyle etkilenir, ancak halka oranlarındaki değişiklik, diyabetle ilişkili retinopatide nispi retinal topografik etkinin daha kolay yorumlanmasını sağlayabilir. Grupların ayırımının yapılmasında R1 / R4 ve R1 / R5 oranları R1 / R2 ve R1 / R3 oranlarından daha iyi tanısal performans gösterdi. Bu, merkezden periferik halkaya değişen diyabetin etki büyüklüğü ile açıklanabilir. R1 genliğinin periferik genliklerden daha fazla etkilendiği görülmüştür, kronik hipergliseminin bir sonucu olabilir. Kronik hipergliseminin bir belirteci olan HbA1c seviyeleri varyansın% 40'ını açıklamıştır. R1 genliği, en iyi tanısal değere sahipti (VUS = 0.65). VUS 0.55 değeri AUC için 0.8 değerine karşılık gelir. Bu nedenle, VUS'ları kendi sınırları içinde değerlendirdiğimizde, R1 ve R2 en iyi tanısal değere sahiptir. Sonuç olarak, diyabetin etkileri retina üzerinde topografik olarak eşit olmayan bir şekilde dağılmıştır. HBA1C kontrol edilebilir bir faktördür ve retina fonksiyonu düşüşü ile ilişkisi nedeniyle sıkı kontrolü önemlidir. R1 ve R2 genlikleri, retina fonksiyonundaki erken değişiklikleri tanımlamanın etkili yollarıdır.

Anahtar Kelimeler: Diyabetik retinopati, multifokal elektroretinografi, HBA1C, diyabet süresi, diyabetes mellitus



Şekil-1. Halkaların gruplandırılması. Retina eksantrikliği gösterilmemiştir. Yanıt yoğunlukları, her halka için P1 - N1 olarak hesaplandı.



Şekil-2. Halka genliklerinin ve halka oranlarının ROC yüzeyi altındaki hacimleri

Tam Metinler

Comparison of Convolutional Neural Network Architectures in the Diagnosis of COVID-19

[1]Işıl ÜNALDI, [2]Leman TOMAK

[1,2] Ondokuz Mayıs University. Faculty of Medicine. Department of Biostatistics and Medical Informatics, Samsun

Aim: In order to prevent the transmission rate of COVID-19, early diagnosis with high accuracy is essential. In this study, it is aimed to determine the disease using pre-trained CNN (convolutional neural network) architectures, which allow accurate and rapid detection of COVID-19 on chest CT (computed tomography) images, and to evaluate the classification success by comparing with various performance measures.

Method: The open source SARS-COV-2 Ct-Scan Dataset accessed via Kaggle was used in the study. In the data, a total of 2482 patients have CT images, of which 1252 patients were diagnosed with COVID-19 and 1230 patients were diagnosed with non-COVID-19. Classification of the disease was done with CNN architectures InceptionResNetV2, DenseNet169, InceptionV3, MobileNetV2, ResNet50, VGG16 and VGG19. The performances of the architectures were evaluated by accuracy, sensitivity, specificity and F1 score. The Python programming language and the Keras library were used to analyze the data.

Results: In this study, the architectures with the highest performance are; they were InceptionResNetV2 with accuracy of 98.8%, sensitivity of 99.1%, specificity of 98.5%, F1 score of 98.7% and DenseNet169 with accuracy of 99.0%, sensitivity of 99.1%, specificity of 98.9%, F1 score of 98.9%. Among the architectures, VGG19 showed the lowest performance with accuracy of 93.0%, sensitivity of 96.0%, specificity of 90.5% and F1 score of 92.4%. The performance values of other architectures were determined as follows; accuracy of 98.0%, sensitivity of 96.4%, specificity of 99.3% and F1 score of 97.7% for MobileNetV2; accuracy of 98.0%, sensitivity of 98.7%, specificity of 97.4% and F1 score of 97.8% for InceptionV3; accuracy of 98.4%, sensitivity of 97.3%, specificity of 99.3% and F1 score of 98.2% for ResNet50; accuracy of 97.4%, sensitivity of 99.1%, specificity of 96.0% and F1 score of 97.1% for VGG16.

Conclusion: In the study, when the performances of the architectures used in the diagnosis of COVID-19 were compared, the most successful architectures were DenseNet169 and InceptionResNetV2 with accuracies of 99%, and the performance metrics for other CNN architectures were in the range of 90.5% - 98%. When the results obtained were evaluated, it was determined that the CNN architectures used in the diagnosis of COVID-19 worked with over 90% high performance for each performance metrics considered. In this study, it was revealed that the CNN method in detecting COVID-19 using CT images is an automatic decision mechanism that will help decision makers, and thus the disease can be diagnosed as soon as possible, thus increasing the effectiveness of the treatment and preventing the rate of transmission and deaths.

Keywords: Covid-19, Deep learning, Convolutional Neural Networks, Image classification, Computed tomography

Introduction

The COVID-19 pandemic has affected the lives of millions of people, caused difficulties in their daily lives and caused serious damage to the economies of many countries around the world. It also caused the death of millions of people and brought a great burden and difficulty especially to healthcare systems. Therefore, timely and cost- effective screening of affected patients is essential to struggle this disease [1, 2].

Due to the rapid spread of this highly contagious disease worldwide, the necessity of early diagnosis with high accuracy has emerged in order to prevent the rate of transmission. Currently, the RT-PCR test (reverse- transcriptase polymerase chain reaction), taken as a throat swab, is the main screening method for COVID-19. However, this method is costly, time consuming, and test kits are limited. Diagnostic tests, which give results within minutes, also sometimes give false negative results. Therefore, it is very important to develop rapid, low-cost and reliable methods for the automatic diagnosis of COVID-19 [3].

CT (computed tomography), which is a routine imaging tool for the diagnosis of pneumonia, is easy to apply and can provide rapid diagnosis. As it is known, the ground glass opacities and tissue changes in the lungs are seen in almost all COVID-19 patients and can be easily noticed with CT images. These features are also seen in patients with negative RT-PCR results despite having clinical symptoms. CT and X-ray imaging methods are considered among the most effective techniques, especially for the diagnosis of COVID-19. CT scan is preferred over X-ray because of its versatility and three-dimensional pulmonary appearance [4, 5, 6].

In this context, it is very important and necessary to use rapid diagnosis systems based on deep learning through CT images. Early diagnosis and rapid treatment solutions of COVID-19 disease will be further developed with studies to provide clinicians with a more

objective interpretation. With the use and development of deep learning methods and CNN (convolutional neural network) architectures to detect anomalies in radiological images and diagnose COVID-19, image classification will be possible faster than a few hours or almost as soon as the image is taken [7].

The aim of this study is to compare the classification successes of CNN architectures with various performance metrics, which allow the accurate and rapid detection of COVID-19 on chest CT images, which are pre-trained on the ImageNet dataset and which we can easily apply to other data sets with the transfer learning method.

Method

The open source SARS-COV-2 Ct-Scan Dataset accessed via Kaggle was used in the study. The classification of the disease was performed with the architectures InceptionResNetV2, DenseNet169, InceptionV3, MobileNetV2, ResNet50, VGG16 and VGG19. The performances of the architectures were evaluated by accuracy, sensitivity, specificity and F1 score. The Python programming and the Keras library were used to analyze the data.

Dataset

In the data, a total of 2482 patients have CT images, of which 1252 patients were diagnosed with COVID-19 and 1230 patients were diagnosed with non-COVID-19. Data were collected from real patients in hospitals in Sao Paulo, Brazil. The data set was collected for the purpose of researching and developing artificial intelligence methods that can determine whether the patient is infected with SARS-CoV-2 with the help of analysis of CT images [8].

Convolutional Neural Networks

Artificial intelligence-assisted method has emerged for problems in various fields such as medical image processing, processing of large data sets, processing of digital images and supporting the movement of autonomous cars. One of these methods is CNNs, which is a class of artificial neural networks used for image classification. CNNs are based on deep learning algorithms, especially analysis using image data. In CNNs, some special layers are used to vectorize the image data by making some transformations. CNNs consist of three basic layers: the convolution layer, the pooling layer and the fully-connected layer [9].

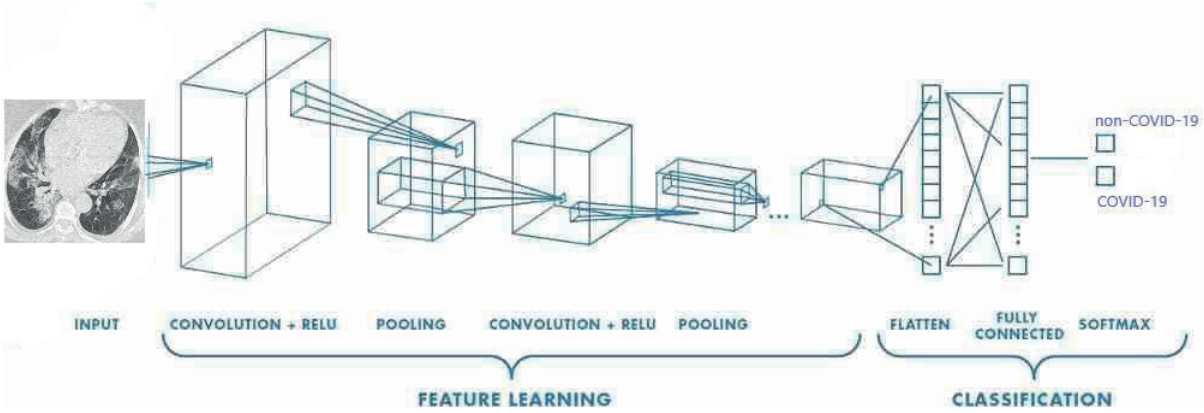


Figure-1. CNN layers and general operation

Classification Methods

Disease classification was performed using CNN architectures InceptionResNetV2, DenseNet169, InceptionV3, MobileNetV2, ResNet50, VGG16 and VGG19.

InceptionResNetV2 is an efficient CNN architecture that uses residual links with good efficiency. Because of the residual learning framework plays an important role in improving the training speed for the Inception architecture, the InceptionResNetV2 network outperforms similar expensive networks that without residual connections. Also, an Inception network that now uses connections has deeper layers of convolution to efficiently extract high-level features from images [10].

DenseNet is a highly advantageous architecture in alleviate the vanishing gradient problem, enhancing feature propagation, promoting feature reuse, and significantly reducing the number of parameters. Although DenseNet169 has a layer depth of 169, it has fewer parameters compared to other models, and the architecture handles the vanish gradient problem well [11].

InceptionV3 is a CNN architecture that provides several improvements, including label smoothing, factorized 7×7 convolutions, and the

use of a auxiliary classifier to propagate label information down the network. The InceptionV3 architecture has 44 layers with 21 million learnable parameters. The InceptionV3 model was trained using a 1000-class dataset from the ImageNet database [12].

MobileNet is based on a modern architecture that uses deeply separable convolution layers to create deep neural networks. Compared to other models in the ImageNet classification, it is an efficient architecture in terms of resource-accuracy balance and strong performance. The MobileNetV2 architecture includes an initial fully convolution layer with 32 filters followed by 19 residual bottleneck layers [13, 14].

ResNet architecture was first proposed to overcome common problems encountered when training multilayer deep neural networks. It adds jump links between layers, which reduces the problem of gradient lost in deep networks and improves overall performance. ResNet50 is a 50-layer network type grouped into residual blocks and is very good at training very deep architectures [15, 16].

Based on the deep concept of CNNs, the VGG architecture can create even deeper networks using small convolution filters. The VGG architecture is the first to use 3×3 filters with 1 step at each convolution layer where the depth of the feature-map gradually increases over the network [17]. The VGG16 network consists of 16 layers, including 13 convolution layers and 3 fully-connected layers. To reduce the computational overhead and control overfitting, there are groups of 5 convolution layers, each followed by a maximum pooling layer with a filter size of 2×2 [18]. VGG19 consists of 19 layers, 16 of which are convolutional and 3 are fully-connected layers. In this architecture, filters are used to reduce the number of parameters in the convolution layers. The VGG19 architecture contains almost 138 million computational parameters [19].

Results

In this study, the architectures with the highest performance are; they were InceptionResNetV2 with accuracy of 98.8%, sensitivity of 99.1%, specificity of 98.5%, F1 score of 98.7% and DenseNet169 with accuracy of 99.0%, sensitivity of 99.1%, specificity of 98.9%, F1 score of 98.9%. Among the architectures, VGG19 showed the lowest performance with accuracy of 93.0%, sensitivity of 96.0%, specificity of 90.5% and F1 score of 92.4%. The performance values of other architectures were determined as follows; accuracy of 98.0%, sensitivity of 96.4%, specificity of 99.3% and F1 score of 97.7% for MobileNetV2; accuracy of 98.0%, sensitivity of 98.7%, specificity of 97.4% and F1 score of 97.8% for InceptionV3; accuracy of 98.4%, sensitivity of 97.3%, specificity of 99.3% and F1 score of 98.2% for ResNet50; accuracy of 97.4%, sensitivity of 99.1%, specificity of 96.0% and F1 score of 97.1% for VGG16. The results for the architectures are given in Table-1.

Table-1. Comparison of CNN architectures

	Accuracy	Sensitivity	Specificity	F1 Score
InceptionResNetV2	0.991	0.985	0.988	0.987
DenseNet169	0.990	0.991	0.989	0.989
InceptionV3	0.980	0.987	0.974	0.978
MobileNetV2	0.980	0.964	0.993	0.977
ResNet50	0.984	0.973	0.993	0.982
VGG16	0.974	0.991	0.960	0.971
VGG19	0.930	0.960	0.905	0.924

The learning curves of the architectures are given in Figures-2-8.

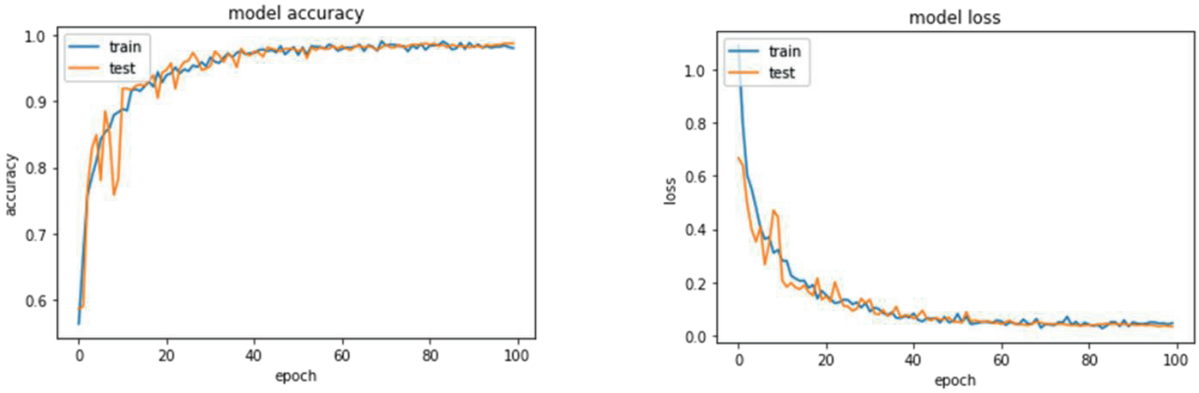


Figure-2. Learning curves of InceptionResNetV2 architecture

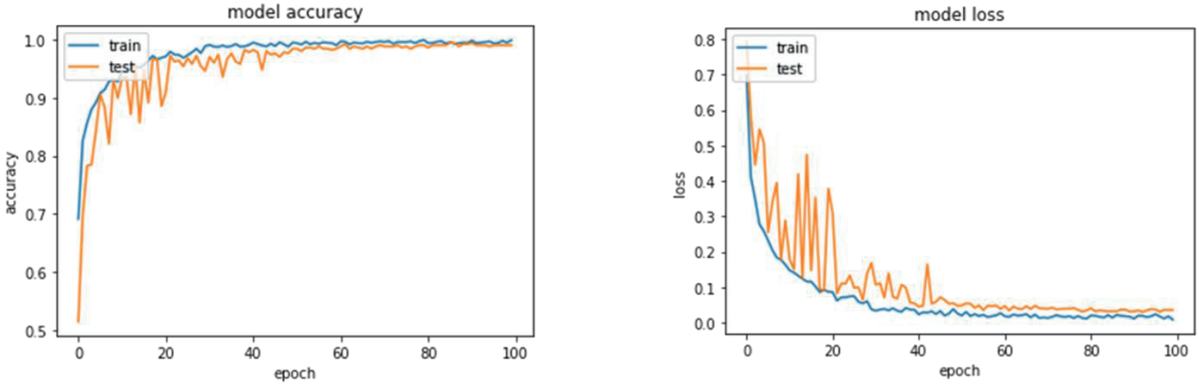


Figure-3. Learning curves of DenseNet169 architecture

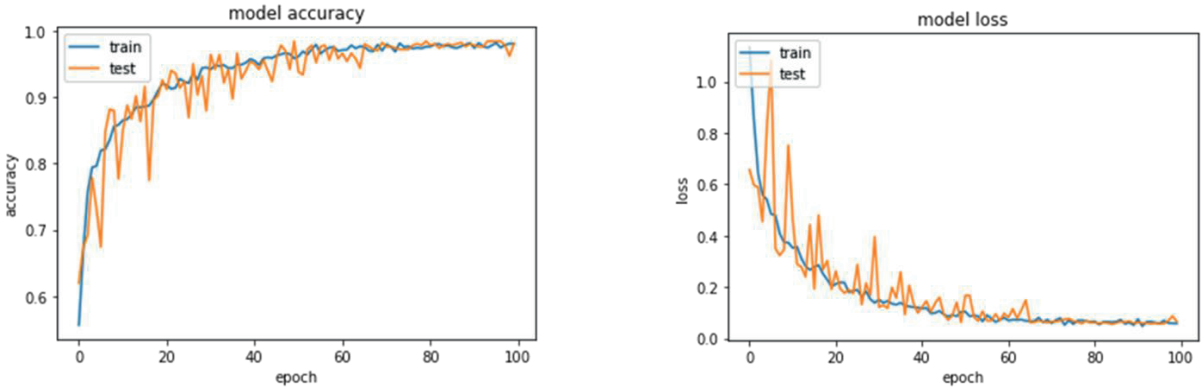


Figure-4. Learning curves of InceptionV3 architecture

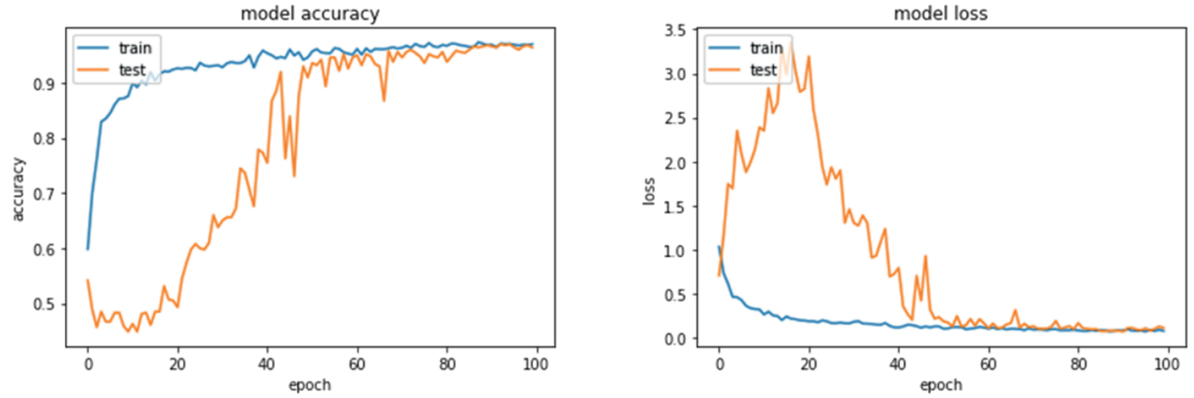


Figure-5. Learning curves of MobileNetV2 architecture

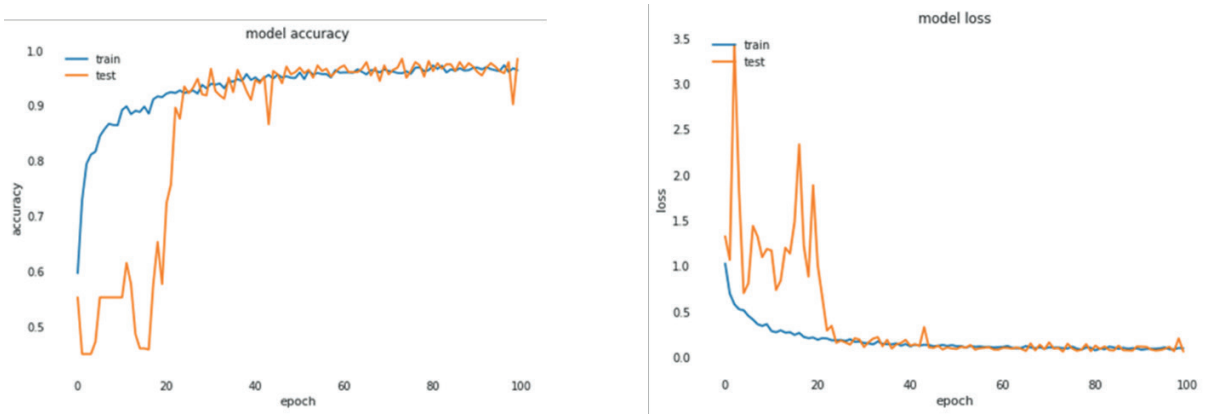


Figure-6. Learning curves of ResNet50 architecture

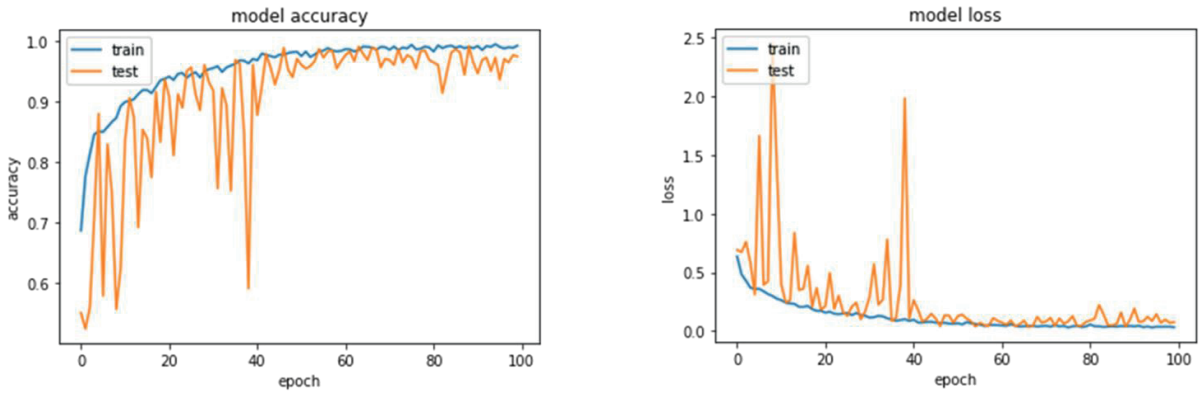


Figure-7. Learning curves of VGG16 architecture

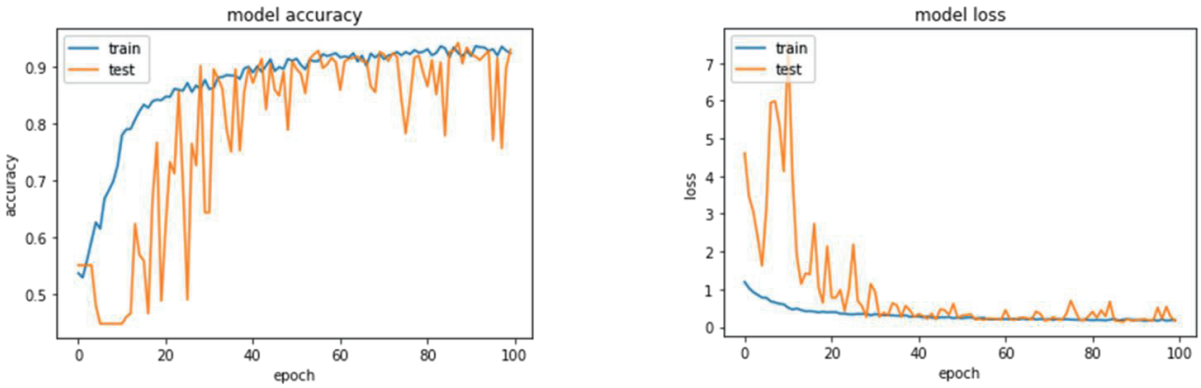


Figure-8. Learning curves of VGG19 architecture

The structure of the learning curves can be used to examine the behavior of the model and make suggestions for adjustments that can be made to improve training and the performance. When the curves obtained for each architecture above are examined, it can be said that the dataset is suitable and generalizable to these architectures, since there is not much deviation in the curves of InceptionResNetV2 and DenseNet169, which have achieved high performance metrics.

Discussion and Conclusion

In this study, an experimental evaluation of pre-trained CNN-based image classification architectures for COVID- 19 diagnosis from chest CT images is presented. When the results obtained were examined, it was seen that each performance metrics had a value of over 90% for the CNN architectures used in the diagnosis of COVID-19.

Both CT and X-ray images are used for the diagnosis of COVID-19 in deep learning-based decision support systems. Ardakani et al. [20] examined the architectures of AlexNet, VGG-16, VGG-19, SqueezeNet, GoogleNet, MobileNetV2, ResNet18, ResNet50, ResNet101 and Xception for the diagnosis of COVID-19 from CT images. The dataset they examined consists of a total of 1020 CT samples from COVID-19 and non-COVID-19 cases. They achieved accuracy of over 90% with MobileNetV2, ResNet50, ResNet18, ResNet101 and Xception. Apostolopoulos and Bessiana [21] used pre-trained VGG19, MobileNetV2, Inception, Xception and InceptionResNetV2 architectures for automatic diagnosis of COVID-19 cases. They ran the architectures they examined on two datasets. Using MobileNetV2, they achieved the highest performance with accuracy of 96.78% for the second dataset. Horry et al. [22] worked with X-ray images and used VGG19, Inception, Xception and ResNet architectures. The dataset they used consists of 100 COVID-19 cases, 100 pneumonia and 200 healthy cases. They considered precision, recall and F1 score as performance metrics. As a result, they obtained the highest performance from VGG19 architecture with precision of 83% in their study for three-class data. Punn and Agarwal [23] developed an automated COVID-19 diagnostic system using the architectures ResNet, InceptionV3, InceptionResNetV2, DenseNet169 and NASNetLarge with a small number of X-ray images. In terms of fine- tuning in transfer learning, 6 scenarios are discussed. A total of 1076 chest X-ray images were considered for the experiments. As a result of their studies, the best performance values obtained with InceptionResNetV2 and DenseNet169 were AUC of 99% and recall of 96%, respectively. Narin et al. [24] used a total of 100 X-ray image data, 50 from COVID-19 patients and the remaining 50 from healthy individuals, to detect patients infected with COVID-19. The pre-trained models used were ResNet50, ResNet101, ResNet152, InceptionV3 and InceptionResNetV2. With the InceptionV3, ResNet50 and InceptionResNetV2 architectures, they achieved average accuracies of 95.4%, 96.1% and 94.2%, respectively. Asif et al. [25] used a combined dataset containing a total of 2541 chest X-ray images from two different sources. For the diagnosis of COVID-19, performance evaluation was made using the InceptionV3, Xception, MobileNetV2, NasNet and DenseNet201 architectures. Each of the architectures they used showed over 90% performance.

Considering the literature and the results of this study, it was seen that the architectures used in this study were frequently used in the classification of COVID-19 and achieved successful results. It was concluded that performances could be improved by fine-tuning for future studies. In addition, with this study, it has been revealed that the CNN method is an automatic decision mechanism that will help decision makers in detecting COVID-19 using CT images, so that the disease can be diagnosed as soon as possible and thus deaths can be prevented by increasing the effectiveness of the treatment.

REFERENCES

[1] Khan, S. H., Sohail, A., Khan, A., Hassan, M., Lee, Y. S., Alam, J., ... & Zubair, S. (2021). COVID-19 detection in chest X-ray images using deep boosted hybrid learning. *Computers in Biology and Medicine,* 137, 104816.
[2] Meem, A. T., Khan, M. M., Masud, M., & Aljahdali, S. (2022). Prediction of Covid-19 Based on Chest X- Ray Images Using Deep Learning with CNN. *Computer Systems Science and Engineering,* 1223-1240.
[3] Hosseinzadeh, H. (2022). Deep multi-view feature learning for detecting COVID-19 based on chest X-ray images. Biomedical Signal Processing and Control, 75, 103595.
[4] Ai, T., Yang, Z., Hou, H., Zhan, C., Chen, C., Lv, W., ... & Xia, L. (2020). Correlation of chest CT and RT- PCR testing for coronavirus disease 2019 (COVID-19) in China: a report of 1014 cases. *Radiology, 296*(2), E32-E40.
[5] Dong, D., Tang, Z., Wang, S., Hui, H., Gong, L., Lu, Y., ... & Li, H. (2020). The role of imaging in the detection and management of COVID-19: a review. IEEE reviews in biomedical engineering, 14, 16-29.
[6] Islam, M. M., Karray, F., Alhajj, R., & Zeng, J. (2021). A review on deep learning techniques for the diagnosis of novel coronavirus (COVID-19). Ieee Access, 9, 30551-30572.
[7] Clark, S., Kamalinejad, E., Magpantay, C., Sahota, S., Zhong, J., & Hu, Y. (2021, November). A Review of CNN on Medical Imaging to Diagnose COVID-19 Infections. In Proceedings of ISCA 34th International Conference on (Vol. 79, pp. 91-98).
[8] Soares, E., Angelov, P., Biaso, S., Froes, M. H., & Abe, D. K. (2020). SARS-CoV-2 CT-scan dataset: A large dataset of real patients CT scans for SARS-CoV-2 identification. *MedRxiv*.
[9] Özkan, Y. (2021). Uygulamalı Derin Öğrenme: Yapay Zeka-Makine Öğrenmesi-Yapay Sinir Ağları. *Papatya Bilim Akademik Yayınevi.*
[10] Vo, D. M., Nguyen, N. Q., & Lee, S. W. (2019). Classification of breast cancer histology images using incremental boosting convolution networks. *Information Sciences,* 482, 123-138.
[11] Huang, G., Liu, Z., Van Der Maaten, L., & Weinberger, K. Q. (2017). Densely connected convolutional networks. In *Proceedings of the IEEE conference on computer vision and pattern recognition* (pp. 4700- 4708).
[12] Szegedy, C., Vanhoucke, V., Ioffe, S., Shlens, J., & Wojna, Z. (2016). Rethinking the inception architecture for computer vision. In *Proceedings of the IEEE conference on computer vision and pattern recognition* (pp. 2818-2826).
[13] Howard, A. G., Zhu, M., Chen, B., Kalenichenko, D., Wang, W., Weyand, T., ... & Adam, H. (2017). Mobilenets: Efficient convolutional neural networks for mobile vision applications. *arXiv preprint arXiv:1704.04861*.
[14] Sandler, M., Howard, A., Zhu, M., Zhmoginov, A., & Chen, L. C. (2018). Mobilenetv2: Inverted residuals and linear bottlenecks. In *Proceedings of the IEEE conference on computer vision and pattern recognition* (pp. 4510-4520).
[15] Basu, A., Sheikh, K. H., Cuevas, E., & Sarkar, R. (2022). COVID-19 detection from CT scans using a two- stage framework. *Expert Systems with Applications,* 193, 116377.
[16] Kamble, R. M., Chan, G. C., Perdomo, O., Kokare, M., Gonzalez, F. A., Müller, H., & Mériaudeau, F. (2018, December). Automated diabetic macular edema (DME) analysis using fine tuning with inception-resnet-v2 on OCT images. In *2018 IEEE-EMBS Conference on Biomedical Engineering and Sciences (IECBES)* (pp. 442-446). IEEE.
[17] Jakubec, M., Lieskovska, E., & Jarina, R. (2021, April). Speaker Recognition with ResNet and VGG Networks. In 2021 31st *International Conference Radioelektronika (RADIOELEKTRONIKA)* (pp. 1-5). IEEE.
[18] Zhen, X., Chen, J., Zhong, Z., Hrycushko, B., Zhou, L., Jiang, S., ... & Gu, X. (2017). Deep convolutional neural network with transfer learning for rectum toxicity prediction in cervical cancer radiotherapy: a feasibility study. *Physics in Medicine & Biology, 62*(21), 8246.
[19] Toğaçar, M., Ergen, B., & Cömert, Z. (2020). Application of breast cancer diagnosis based on a combination of convolutional neural networks, ridge regression and linear discriminant analysis using invasive breast cancer images processed with autoencoders. *Medical hypotheses*, 135, 109503.
[20] Ardakani, A. A., Kanafi, A. R., Acharya, U. R., Khadem, N., & Mohammadi, A. (2020). Application of deep learning technique to manage COVID-19 in routine clinical practice using CT images: Results of 10 convolutional neural networks. *Computers in biology and medicine, 121*, 103795.
[21] Apostolopoulos, I. D., & Mpesiana, T. A. (2020). Covid-19: automatic detection from x-ray images utilizing transfer learning with convolutional neural networks. *Physical and engineering sciences in medicine*, 43(2), 635-640.
[22] Horry, M. J., Chakraborty, S., Paul, M., Ulhaq, A., Pradhan, B., Saha, M., & Shukla, N. (2020). X-ray image based COVID-19 detection using pre-trained deep learning models.
[23] Punn, N. S., & Agarwal, S. (2021). Automated diagnosis of COVID-19 with limited posteroanterior chest X- ray images using fine-tuned deep neural networks. *Applied Intelligence*, 51(5), 2689-2702.
[24] Narin, A., Kaya, C. & Pamuk, Z. (2021). Automatic detection of coronavirus disease (COVID-19) using X- ray images and deep convolutional neural networks. *Pattern Anal Applic* 24, 1207–1220.
[25] Asif, S., Zhao, M., Tang, F. et al. (2022). A deep learning-based framework for detecting COVID-19 patients using chest X-rays. *Multimedia Systems* 28, 1495–1513.

Taguchi ve Faktöriyel Deney Tasarım Metotlarının Karşılaştırması: *Lactobacillus* spp. ERIC-PCR Optimizasyonu Örneği

[1]Selen YILMAZ IŞIKHAN, [2]Ceren ÖZKUL

[1]Hacettepe Üniversitesi Sosyal Bilimler Meslek Yüksekokulu
[2]Hacettepe Üniversitesi Eczacılık Fakültesi, Farmasötik Mikrobiyoloji Anabilim Dalı

Amaç: Faktöriyel tasarımlar çok güçlü istatistiksel araçlardır çünkü bir araştırmacının sonuç değişkeni üzerindeki çoklu faktör düzeyindeki kombinasyonların etkilerini aynı anda test etmesine izin verir. Ancak bazen faktöriyel deneylerde kullanılan faktör sayıları ve seviyeleri zor ve pahalıdır. Taguchi yöntemi, ortogonal diziler kullanılarak, faktöriyel tasarıma göre deney zamanını ve maliyetini en aza indiren bir tasarımdır. Ayrıca, deney birimlerinin sayısının azaltılması, kontrol edilemeyen faktörlerin etkilerinin azaltılmasına izin verir. Bu çalışmada deney tasarım metotlarından Taguchi ve faktöriyel tasarımların performansını Lactobacillus spp.'nin ERIC-PCR ile parmak izi analizleri üzerinde karşılaştırmak hedeflenmiştir. Probiyotik mikroorganizmaları içeren Lactobacillus genusu içerisindeki türlerin, DNA parmak izi analizleri ile profillendirilmesi ve farklı türlerin tespiti özellikle anne sütü gibi klinik örneklerde kullanılan yaklaşımlardır.

Yöntem: Uygulamada *L. plantarum* 299v suşu için ERIC-PCR optimizasyon verileri kullanılmıştır. Saf L. plantarum 299v kültüründen DNA izolasyonu sonrası, DNA miktar ve saflığı ölçülmüştür. PCR basamakları için 4 faktörün varyasyonları kullanılmıştır (MgCl2, dNTP, primer, DNA miktarı). ERIC primer dizileri kullanılarak gerçekleştirilen PCR sonrası, amplikonlar %1.8 agaroz jelde yürütülerek Etidyum bromür varlığında bantlar görüntülenmiş ve bant sayıları belirlenmiştir. Bu deneyde temel amaç en yüksek bant sayısını, dolayısıyla ayrımı sağlayan faktör düzeylerini tespit etmektir. Verilerin deneysel tasarımı 3x3x3x3 (L9, 3^4) faktöriyel tasarımdır. Tam faktöriyel deney tasarımı ile 81 deney yapılması gerekirken, Taguchi L9 deney tasarım tablosu kullanılarak 9 farklı değer içeren deney seti ile çalışma yapılmıştır. Taguchi yöntemi ve Faktöriyel deney sonuçları karşılaştırmalı olarak incelenmiştir.

Bulgular: Taguchi yöntemi, daha az sayıda deney ünitesi ile Lactobacillus spp. bant sayısına etkisi olan dNTP(µM), $MgCl_2$ (mM), DNA (ng) ve Primer (µM)'ün en iyi kombinasyonunu belirlemiştir. Bant sayısı üzerinde en etkili ilk faktör Primer (µM), ikincisi dNTP (µM) olarak belirlenmiştir. $MgCl_2$ (mM) ile DNA (ng) ise aynı Delta değerini (rank) verdiklerinden istatistiksel olarak anlamlı olmayan eşit ve daha az bir etkiye sahiptirler. Taguchi modeli sonucunda her bir primer konsantrasyonun 0.9 µM olduğu koşulda en yüksek bant sayısı 12 olarak elde edilmiştir (p=0.033).

Sonuç: Taguchi deney tasarım metodu faktöriyel tasarıma kıyasla, hangi faktörlerin sonuç değişkeni üzerinde daha etkili olduğunu, ayrıca hangi alt düzeylerin minimum maliyet ve optimum performansı sağladığını tasarımın başında hızlıca anlamamızı sağlamaktadır. Günümüzde yaygın olarak kullanılmayan Taguchi yöntemi, ERIC-PCR gibi DNA parmak izi analizlerinde PCR reaktiflerinin konsantrasyonunu belirlemek için maliyet ve süreyi azaltan bir yaklaşım olarak kullanılabilir.

Anahtar Kelimeler: Taguchi deney tasarımı, faktöriyel tasarım, ortogonal dizi optimizasyonu, Lactobacillus, ERIC-PCR.

GİRİŞ

Faktöriyel deney tasarımları tüm çalışma alanlarında araştırmacılar ve mühendisler tarafından yaygın olarak kullanılan, özellikle endüstriyel süreçlerin deneylerini incelemek için çok büyük ölçekte kullanılan güçlü istatistiksel araçlardır. Bir araştırmacının sonuç değişkeni üzerindeki çoklu faktör düzeyindeki kombinasyonların etkilerini aynı anda test etmesine izin verir. Ancak bazen faktöriyel deneylerde kullanılan faktör sayıları ve seviyeleri zor ve pahalı olabilmektedir.

Taguchi yöntemi, ortogonal diziler kullanılarak, faktöriyel tasarıma göre deney zamanını ve maliyetini en aza indiren bir tasarımdır. Ayrıca, deney birimlerinin sayısının azaltılması, kontrol edilemeyen faktörlerin etkilerinin azaltılmasına izin verir. Bu çalışmada deney tasarım metotlarından Taguchi ve faktöriyel tasarımların performansını Lactobacillus spp.'nin ERIC-PCR ile parmak izi analizleri üzerinde karşılaştırmak hedeflenmiştir. ERIC-PCR gibi DNA parmak izi analizlerinde PCR koşullarının optimizasyonu önem taşımaktadır.

YÖNTEM

Taguchi Deney Tasarımı

Ar-Ge verimliliğini ve ürün kalitesini artırmakla görevli Dr. Genichi Taguchi Mühendislik deneyleri ve testleri için deney sayısını ve dolayısıyla maliyeti azaltan tekniğini 1966'da öne sürdü. Taguchi, "ortogonal diziler" (OA) olarak bilinen özel oluşturulmuş tabloları kullanarak deneyler tasarlamıştır. Deney tasarımında klasik yöntemlerin yetersizliği deney tasarım yöntemleri ile giderilmiştir. İstatistiksel deney tasarımındaki yöntemler Tam Faktöriyel deney tasarımı, Kesirli Faktöriyel deney tasarımı ve Taguchi metotudur.

Deneylerin Faktöriyel Tasarımı ilk kez

R.A. Fisher (1920) tarafından önerilmiştir.

Tam faktöriyel tasarım: k: faktör seviyelerinin sayısı, n: faktörlerin sayısı iken k^n deneyden oluşur ve tüm olası etkileşimlerin incelenmesini sağlar.

Kesirli faktöriyel tasarım: Maliyetten ve zamandan kazanmak için deney sayısı azaltılarak en fazla bilgi üreten küme seçilir. Örneğin; 7 parametre ve 2'şer düzey olsun. Tam faktöriyel denemede $2^7 = 128$ deney yerine ½:64,

¼:32, 1/8:16 şeklinde kesirli olarak deney sayısı azaltılabilir. Taguchi, çok sayıda deneysel durumu açıklamak için özel bir ortogonal diziler (OA) seti oluşturmuştur. Ortogonal dizinin en önemli özelliği birçok faktörün en az sayıda test edilmesi ve faktör seviyelerini eş zamanlı olarak değişiminin sağlanmasıdır. Ortogonal diziler 2 veya 3 kademeli olarak belirlenmektedir. 2 Seviyelilerde en sık kullanılanlar L4, L8, L16 iken 3 seviyelilerde L9 ve L27'dir.

A	B	C	AB	BC	AC	ABC
1	1	1	1	1	1	1
2	1	1	2	1	2	2
1	2	1	2	2	1	2
2	2	1	1	2	2	1
1	1	2	1	2	2	2
2	1	2	2	2	1	1
1	2	2	2	1	2	1
2	2	2	1	1	1	2

Table 1.
Full fractional factor assignments to experimental array columns.

A	B	C	D	E	F	G
1	1	1	1	1	1	1
2	1	1	2	1	2	2
1	2	1	2	2	1	2
2	2	1	1	2	2	1
1	1	2	1	2	2	2
2	1	2	2	2	1	1
1	2	2	2	1	2	1
2	2	2	1	1	1	2

Table 2.
Orthogonal array factor assignments to experimental array columns.

Taguchi metodundaki temel adımlar:

1. Sürecin performans ölçümü için bir hedef değeri tanımlayın (örn; akış hızı, sıcaklık değişimi, bant sayısı vb.) İşlemin hedefi maksimum ya da minimum olabilir, akış hızını maksimize etmek gibi.
2. Süreci etkileyen tasarım parametrelerini belirleyin. Parametreler, kolayca kontrol edilebilen sıcaklık, basınç vb. gibi performans ölçümünü etkileyen süreç içindeki değişkenlerdir. Parametrelerin değişen düzey sayıları belirlenmelidir. Örneğin, bir sıcaklık 40°C ve 80°C şeklinde düşük ve yüksek olarak seçilebilir.
3. Parametre sayısı ve düzey sayısına bağlı olarak ortogonal dizi seçilmelidir.
4. Tamamlanmış dizide belirtilen deneylerin gerçekleştirilmesi.
5. Performans ölçüsü üzerinde veri analizinin tamamlanması.

Taguchi metodunda bir sürecin performans karakteristiğinin hedef değeri, τ, ve ölçülen değer y iken kayıp fonksiyon: $l(y) = k_c (y-\tau)^2$ şeklindedir. k_c sabiti belirleme limitleri ya da kabul edilebilir bir delta aralığı dikkate alınarak belirlenebilir.

Eğer amaç performans karakteristik değerini minimize etmek ise kayıp fonksiyon:

$l(y) = k_c y^2$, $\tau = 0$

Eğer amaç performans karakteristik değerini maksimize etmek ise kayıp fonksiyon:

$l(y) = k_c / y^2$ şeklindedir (Fraley, 2006).

Deney sonuçlarına uygun amaç fonksiyonuna göre Sinyal/Gürültü (S/N) oranlarının bulunması gerekir. En temel üç amaç fonksiyonu vardır. Bunlar:

1. En büyük en iyi ($S/N = -10 \log [\sum(1/Y^2)/n]$)
2. En küçük en iyi ($S/N = -10 \log [\sum Y^2 / n]$)
3. Nominal en iyi ($S/N = 10 \log(\bar{Y}^2/S^2)$)

Parametre tasarımı-ortogonal diziyi belirleme

Yoğun bir deney setinde birçok farklı parametrenin performans karakteristiği üzerindeki etkisi, Taguchi tarafından önerilen ortogonal dizi deney tasarımı kullanılarak incelenebilir. Bunun için önce parametrelerin düzeyleri belirlenir. Tipik olarak, deney tasarımındaki tüm parametreler için seviye sayısı, uygun ortogonal dizinin seçimine yardımcı olmak için aynı olacak şekilde seçilir. Aşağıda gösterilen dizi seçici tablosu kullanılarak parametre sayısı ve seviye sayısına karşılık gelen uygun ortogonal dizi seçilebilir.

Tablo-3. Ortogonal dizi seçici

Number of Levels	Number of Parameters (P)																													
	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31
2	L4	L4	L8	L8	L8	L8	L12	L12	L12	L12	L16	L16	L16	L16	L32	L32	L32	L32	L32	L32	L32	L32	L32	L32	L32	L32	L32	L32	L32	L32
3	L9	L9	L9	L18	L18	L18	L18	L27	L27	L27	L27	L27	L36	L36	L36	L36	L36	L36	L36	L36	L36	L36								
4	L'16	L'16	L'16	L'16	L'32	L'32	L'32	L'32	L'32																					
5	L25	L25	L25	L25	L25	L50	L50	L50	L50	L50	L50																			

Her parametrenin aynı sayıda düzeye sahip olmadığı durumda genellikle en yüksek değer alınır veya fark alınır. bir ortogonal diziyi seçme ve uygun şekilde kullanma konusunda fikir vermektedir.

Örnek 1: # parametre: A, B, C, D = 4 # seviye: 3, 3, 3, 2 = ~3 dizi: **L9**

Örnek 2: # parametre: A, B, C, D, E, F = 6 # seviye: 4, 5, 3, 2, 2, 2 = ~3 dizi: **L18**

L9 (3 Düzeyli) Ortogonal Dizi Tasarımı

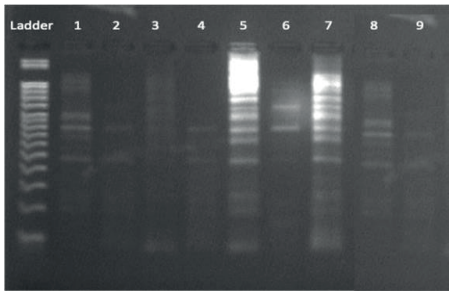
Ortogonal dizi ürün ortalaması ve varyansında etkili olan bir çok faktörü aynı anda ve daha kısa sürede çalışmayı sağlamaktadır. L9 (3^4) tasarım matrisi aşağıda verildiği gibidir:

Run	Columns			
	1	2	3	4
1	1	1	1	1
2	1	2	2	2
3	1	3	3	3
4	2	1	2	3
5	2	2	3	1
6	2	3	1	2
7	3	1	3	2
8	3	2	1	3
9	3	3	2	1

Materyal metot

Uygulama Verisi Tanımı:

Uygulamada *L. plantarum* 299v suşu için ERIC-PCR optimizasyon verileri kullanılmıştır. Saf L. plantarum 299v kültüründen DNA izolasyonu sonrası, DNA miktar ve saflığı ölçülmüştür. PCR basamakları için 4 faktörün varyasyonları kullanılmıştır (MgCl2, dNTP, primer, DNA miktarı). Bu deneyde temel amaç en yüksek bant sayısını, dolayısıyla ayrımı sağlayan faktör düzeylerini tespit etmektir. ERIC PCR için PCR reaksiyon karışımı değişen konsantrasyonlarda ERIC-1R ve ERIC2 primerleri, MgCl2, dNTP içerecek şekilde hazırlanmıştır. Termal döngü cihazında ERIC primerleri ile PCR sonrası amplikonlar %1.8 agaroz jelde yürütülerek bant sayıları belirlenmiştir.



Bantların agaroz jelde görüntüsü

Veri Düzeni-L9 Ortogonal Dizisi

Verilerin deneysel tasarımı 3x3x3x3 (L9, 3^4) faktöriyel tasarımdır. Tam faktöriyel deney tasarımı ile 81 deney yapılması gerekirken, Taguchi L9 deney tasarım tablosu kullanılarak 9 farklı değer içeren deney seti ile çalışma yapılmıştır. Çalışmada maksimum bant sayısını sağlayan (MgCl2 (mM), dNTPs(mM), Each Primer(pmol) ve DNA(ng) seviyeleri araştırılmıştır.

Tablo. Üç Düzeyli 4 Faktörün L9 ortogonal dizisi ve bant sayıları

Deney No	MgCl2 (mM)	dNTPs (mM)	Each Primer (pmol)	DNA (ng)	Output(Band number)
1	2,5	200	0,3	120	8
2	2,5	400	0,6	12	3
3	2,5	600	0,9	6	6
4	3	200	0,6	6	3
5	3	400	0,9	120	15
6	3	600	0,3	12	2
7	3,5	200	0,9	12	15
8	3,5	400	0,3	6	8
9	3,5	600	0,6	120	3

BULGULAR

İlk elde edilen modelde hem S/N oranları hem de ortalama değerler için max-min formülüne dayalı olarak Delta değerleri hesaplanmıştır. Ancak burada artık hatası ve serbestlik derecesi 0 çıktığı için hiçbir faktörün bant sayısı üzerinde önemli etkisi bulunmadığı gibi bir sonuç gözlenmiş hatta, en önemli ilk faktör Primer, ikincisi dNTPS iken MgCL2 ve DNA'nın Delta değerleri aynı olduğu için aynı sıra değerine sahip olmuşlardır.

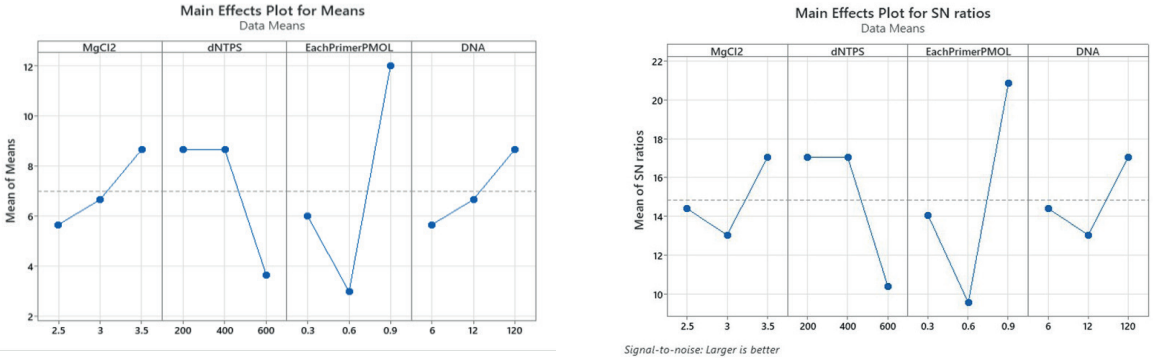
S	R-Sq	R-Sq(adj)
*	100,00%	*

Analysis of Variance for SN ratios

Source	DF	Seq SS	Adj SS	Adj MS	F	P
MgCl2	2	25,000	25,000	12,4999	*	*
dNTPS	2	88,889	88,889	44,4444	*	*
EachPrimerPMOL	2	195,113	195,113	97,5567	*	*
DNA	2	25,000	25,000	12,4999	*	*
Residual Error	0	*	*	*		
Total	8	334,002				

Analysis of Variance for Means

Source	DF	Seq SS	Adj SS	Adj MS	F	P
MgCl2	2	14,000	14,000	7,0000	*	*
dNTPS	2	50,000	50,000	25,0000	*	*
EachPrimerPMOL	2	126,000	126,000	63,0000	*	*
DNA	2	14,000	14,000	7,0000	*	*
Residual Error	0	*	*	*		
Total	8	204,000				



Şekil. S/N oranı ve ortalama için etki grafikleri

Aynı ranka sahip MgCl2 ve DNA faktörleri çıkarılıp Taguchi analizi tekrarlandığında, ayrıca etkileşimli modele de bakıldığında sonuçlar aşağıdaki gibi elde edilmiştir. 9 Deney sayısı toplam 8 serbestlik dereceli model çözümüne yeterli olamamıştır. İki faktörlü modelin R^2'si %86.27 elde edilmiştir.

Analysis of Variance for SN ratios

Source	DF	Seq SS	Adj SS	Adj MS	F	P
dNTPS	2	88,89	88,89	44,44	3,56	0,130
EachPrimerPMOL	2	195,11	195,11	97,56	7,80	0,042
Residual Error	4	50,00	50,00	12,50		
Total	8	334,00				

Analysis of Variance for Means

Source	DF	Seq SS	Adj SS	Adj MS	F	P
dNTPS	2	50,000	50,000	25,0000	*	*
EachPrimerPMOL	2	126,000	126,000	63,0000	*	*
dNTPS*EachPrimerPMOL	4	28,000	28,000	7,0000	*	*
Residual Error	0	*	*	*		
Total	8	204,000				

Buna rağmen yine 3 düzey 4 parametreli L18 düzenine göre çözülen mamografi görüntü verisi için bazı ikili etkileşim sonuçları elde edilebilmiştir.

Analysis of Variance for SN ratios

Source	DF	Seq SS	Adj SS	Adj MS	F	P
Anode	1	0,009	0,009	0,009	1,27	0,323
kVp	2	0,134	0,134	0,067	8,80	0,034
mAs	2	0,047	0,041	0,020	2,72	0,179
FOV	2	0,068	0,063	0,031	4,13	0,107
Anode*kVp	2	0,028	0,028	0,014	1,87	0,267
Anode*mAs	2	0,006	0,006	0,003	0,40	0,694
Anode*FOV	2	0,009	0,009	0,004	0,65	0,570
Residual Error	4	0,030	0,030	0,007		
Total	17	0,336				

SONUÇ

Taguchi deney tasarım metodu faktöriyel tasarıma kıyasla, hangi faktörlerin sonuç değişkeni üzerinde daha etkili olduğunu, ayrıca hangi alt düzeylerin minimum maliyet ve optimum performansı sağladığını tasarımın başında hızlıca anlamamızı sağlamaktadır. Yöntemin zayıf yönü etkileşimleri sağlamamasıdır. Chen ve ark'nın 2016'da yaptıkları çalışmalarında, Taguchi analizini kullanarak bir X-ışını mamografi görüntüleme sisteminin faktörlerini modüle ederek lezyonların görselleştirilmesinin iyileştirildiğini göstermiştir. (Chen, 2016). Bu çalışmadaki L18 veri düzenine Taguchi analizi uygulanmış ve sonucunda 4 ana etki ve en az 3 etkileşim için p değeri elde edilebilmiştir. Dolayısıyla, en az iki değişken arasında etkileşim etkisinden şüpheleniyor ve modele dahil etmek istiyorsak; 2 düzeyliler için L4 yerine L16 ya da L24, 3 düzeyli tasarımlarda ise L9 yerine L18 ya da L27 gibi daha üst ortogonal dizilerin seçilmesi avantaj sağlayacaktır.

Günümüzde yaygın olarak kullanılmayan Taguchi yöntemi, ERIC-PCR gibi DNA parmak izi analizlerinde PCR reaktiflerinin konsantrasyonunu belirlemek için maliyet ve süreyi azaltan bir yaklaşım olarak kullanılabilir.

KAYNAKLAR

Sabırlı, A., & Fığlalı, A. Optimum Çekirdek Çapını Elde Etmek İçin Elektrik Direnç Kaynak Parametrelerinin Taguchi Metoduyla Optimizasyonu. *Kocaeli Üniversitesi Fen Bilimleri Dergisi*, 3(2), 223-229.

Fraley, S., Oom, M., Terrien, B., & Date, J. Z. (2006). Design of experiments via Taguchi methods: orthogonal arrays. *The Michigan chemical process dynamic and controls open text book, USA,* 2(3), 4.

https://support.minitab.com/en-us/minitab/21/help-and-how-to/statistical-modeling/doe/how-to/taguchi/analyze-taguchi- design/before-you-start/example-of-analyzing-a-dynamic-design, Erişim: 20.10.2022.

Stefańska, I., Kwiecień, E., Górzyńska, M., Sałamaszyńska-Guz, A., & Rzewuska, M. (2022). RAPD-PCR-Based Fingerprinting Method as a Tool for Epidemiological Analysis of Trueperella pyogenes Infections. *Pathogens*, 11(5), 562.

Kechagias, J. D., Aslani, K. E., Fountas, N. A., Vaxevanidis, N. M., & Manolakos, D. E. (2020). A comparative investigation of Taguchi and full factorial design for machinability prediction in turning of a titanium alloy. *Measurement*, 151, 107213.

Chen, C. Y., Pan, L. F., Chiang, F. T., Yeh, D. M., & Pan, L. K. (2016). Optimizing quality of digital mammographic imaging using Taguchi analysis with an ACR accreditation phantom. *Bioengineered*, 7(4), 226-234.

Islam, M. N., & Pramanik, A. (2016). Comparison of design of experiments via traditional and Taguchi method. *Journal of Advanced Manufacturing Systems,* 15(03), 151-160.

Peng Y et al. Orthogonal array design in optimizing ERIC-PCR system for fingerprinting rat's intestinal microflora. *Journal of Applied Microbiology* 103 (2007) 2095–2101.

Medan, N., Lobontiu, M., Nagy, S. R., & Dezső, G. (2017). Taguchi versus full factorial design to determine the equation of impact forces produced by water jets used in sewer cleaning. In *MATEC Web of Conferences* (Vol. 112, p. 03007). EDP Sciences.

Veri Madenciliğinde Hibrit Model Yaklaşımı

[1]Batuhan BAKIRARAR, [2]Atilla Halil ELHAN

[1]Ankara Üniversitesi Tıp Fakültesi Biyoistatistik Anabilim Dalı, Ankara, Türkiye
[2]Ankara Üniversitesi Tıp Fakültesi Biyoistatistik Anabilim Dalı, Ankara, Türkiye

Özet

Amaç: Çalışmanın amacı, son teknoloji veri madenciliği algoritmalarını ve uygulamalarını sunmak ve tıbbi verilerin kümelenmesi ve sınıflandırılması için yeni bir hibrit veri madenciliği yaklaşımı önermektir. Ayrıca çalışmada, dengeli ve dengesiz dağılıma sahip veri setlerinde, farklı örneklem büyüklüklerinde ve farklı değişkenler arası ilişkiler olması durumunda performans ölçütlerinin hesaplanması ve bu ölçütlerin hibrit modelden elde edilen ölçütler ile karşılaştırılması amaçlanmıştır.

Gereç ve Yöntem: Çalışmada UCI veri tabanından alınan hepatit veri seti kullanılmıştır. Sık kullanılan ve veri setlerinde en iyi performansa sahip denetimli öğrenme algoritmalarından Karar Ağaçları, Destek Vektör Makinesi, Random Forest, Naive Bayes ve K-en yakın komşu'nun yanı sıra Lojistik Regresyon ve Yapay Sinir Ağları algoritmaları, denetimsiz öğrenme algoritmalarından ise K-ortalama kullanılmıştır. Ayrıca kullanılan denetimli ve denetimsiz öğrenme algoritmaları birleştirilerek hibrit modeller oluşturulmuştur. Performans ölçütü olarak dengesiz veri setlerinde daha yansız sonuçlar verdiği için MCC belirlenmiştir. MCC değeri benzer olan modellerde ise bu ölçüte ek olarak Accuracy ve F-measure değerlerine bakılması uygun görülmüştür. Veri setinde ilk aşamada gerçek değerler kullanılarak değerlendirme yapılmış, daha sonra bu verinin tüm özellikleri korunarak örneklem büyüklüğü 250 ve 500'e yükseltilmiştir.

Bulgular: Gerçek veri setinde ve 250, 500 örnekleme sahip veri setlerinde performans ölçütleri hesaplanıp karşılaştırmalar yapıldığında ise örneklem büyüklüğü arttıkça performans ölçütlerinin de arttığı görülmüştür. Gerçek veri seti için MCC değeri 0,730 olarak bulunurken, bu değer 250 örnekleme sahip veri setinde 0,798 ve 500 örnekleme sahip veri setinde 0,920 olarak bulunmuştur.

Sonuç: Tüm veri setlerinde hibrit model daha iyi sonuçlar vermiştir. Uygun veri madenciliği yöntemlerinin birlikte kullanımı durumunda, yapılacak tüm çalışmalarda hibrit model yaklaşımının daha iyi modeller elde etmemizi sağlayacağı sonucuna ulaşılmaktadır.

Anahtar Kelimeler: Hibrit model, veri madenciliği, performans ölçütleri

Giriş: Veri madenciliği algoritmalarındaki son gelişmeler, her türlü ham verinin üst düzey bilgiye dönüştürülmesini mümkün kılmıştır. Ancak asıl sorun, her yöntemin veri yapısı, şekli ve geçerliliği ile ilgili kendi yaklaşımına sahip olmasıdır. Bu sınırlama sınıflandırma yöntemlerinin performansını etkiler. Sonuç olarak, hibrit bir veri madenciliği yaklaşımına ihtiyaç duyulmaktadır ve bu konuda son yıllarda yapılan çalışma sayısı gittikçe artmaktadır.

Amaç: Çalışmanın amacı, son teknoloji veri madenciliği algoritmalarını ve uygulamalarını sunmak ve tıbbi verilerin kümelenmesi ve sınıflandırılması için yeni bir hibrit veri madenciliği yaklaşımı önermektir. Ayrıca çalışmada, dengeli ve dengesiz dağılıma sahip veri setlerinde, farklı örneklem büyüklüklerinde ve farklı değişkenler arası ilişkiler olması durumunda performans ölçütlerinin hesaplanması ve bu ölçütlerin hibrit modelden elde edilen ölçütler ile karşılaştırılması amaçlanmıştır.

Gereç ve Yöntem: Çalışmada UCI veri tabanından alınan hepatit veri seti kullanılmıştır. Sık kullanılan ve veri setlerinde en iyi performansa sahip denetimli öğrenme algoritmalarından Karar Ağaçları, Destek Vektör Makinesi, Random Forest, Naive Bayes ve K-en yakın komşu'nun yanı sıra Lojistik Regresyon ve Yapay Sinir Ağları algoritmaları, denetimsiz öğrenme algoritmalarından ise K-ortalama kullanılmıştır.Ayrıca kullanılan denetimli ve denetimsiz öğrenme algoritmaları birleştirilerek hibrit modeller oluşturulmuştur. Performans ölçütü olarak dengesiz veri setlerinde daha yansız sonuçlar verdiği için MCC belirlenmiştir. MCC değeri benzer olan modellerde ise bu ölçüte ek olarak Doğru sınıflama Oranı ve F-ölçütü değerlerine bakılması uygun görülmüştür. Veri setinde ilk aşamada gerçek değerler kullanılarak değerlendirme yapılmış, daha sonra bu verinin tüm özellikleri korunarak örneklem büyüklüğü 250 ve 500'e yükseltilmiştir.

Bulgular: Gerçek veri setinde ve 250, 500 örnekleme sahip veri setlerinde performans ölçütleri hesaplanıp karşılaştırmalar yapıldığında ise örneklem büyüklüğü arttıkça performans ölçütlerinin de arttığı görülmüştür. Gerçek veri seti için MCC değeri 0,730 olarak bulunurken, bu değer 250 örnekleme sahip veri setinde 0,798 ve 500 örnekleme sahip veri setinde 0,920 olarak bulunmuştur.

Sonuç: Tüm veri setlerinde hibrit model daha iyi sonuçlar vermiştir. Uygun veri madenciliği yöntemlerinin birlikte kullanımı durumunda, yapılacak tüm çalışmalarda hibrit model yaklaşımının daha iyi modeller elde etmemizi sağlayacağı sonucuna ulaşılmaktadır.

Anahtar Kelimeler: Hibrit Model, Veri Madenciliği, Performans Ölçütleri

GİRİŞ

Bilgisayarların kullanımının artmasıyla, bilgisayar tabanlı sistemler tarafından büyük miktarda veri üretilmektedir. Devlet kurumları, bilimsel kurumlar ve işletmeler, veri toplamak ve depolamak için muazzam kaynaklara sahiptir. Sağlık sektörü, büyük veriye sahip kaynaklar arasında yer almaktadır. Tıbbi bilgi, hasta bilgileri ve diğer hasta/hastane verileri günlük olarak büyümeye devam etmektedir. Örneğin, bir hastanenin yılda beş terabayt veri üretebileceği tahmin edilmektedir. Bu verileri kaliteli sağlık hizmetleri adına yararlı bilgiler elde etmek için kullanabilmek çok önemlidir (1).

Veri madenciliği, büyük hacimli verilerde yeni, değerli ve keşfedilmemiş bilgi arayışıdır. Bu, insanların ve bilgisayarların ortak bir çabasıdır. Makine öğrenimi, veri tabanı teknolojisi, istatistik, yüksek performanslı bilgi işlem, görselleştirme ve matematik gibi mevcut teknolojilerin sinerjisinden doğan hızlı büyüyen bir teknolojidir. Daha önce bilinmeyen, potansiyel olarak yararlı ve ilginç bilgileri büyük ve sıklıkla farklı, geçmiş veri kaynaklarından ayıklama süreci olarak da tanımlanır (2).

Veri madenciliği algoritmalarındaki son gelişmeler, her türlü ham verinin üst düzey bilgiye dönüştürülmesini mümkün kılmıştır. Ancak asıl sorun, her yöntemin veri yapısı, şekli ve geçerliliği ile ilgili kendi yaklaşımına sahip olmasıdır. Bu sınırlama sınıflandırma yöntemlerinin performansını etkiler. Sonuç olarak, hibrit bir veri madenciliği yaklaşımına ihtiyaç duyulmaktadır ve bu konuda son yıllarda yapılan çalışma sayısı gittikçe artmaktadır. Goonatilake ve arkadaşları hibrit sistemi, iki veya daha fazla tekniğin bireysel tekniklerin sınırlamalarının üstesinden gelecek şekilde birleştirildiği bir sistem olarak tanımlar (3).

Çalışmanın amacı, son teknoloji veri madenciliği algoritmalarını ve uygulamalarını sunmak ve tıbbi verilerin kümelenmesi ve sınıflandırılması için yeni bir hibrit veri madenciliği yaklaşımı önermektir. Ayrıca çalışmada, dengeli ve dengesiz dağılıma sahip veri setlerinde, farklı örneklem büyüklüklerinde ve farklı değişkenler arası ilişkiler olması durumunda performans ölçütlerinin hesaplanması ve bu ölçütlerin hibrit modelden elde edilen ölçütler ile karşılaştırılması amaçlanmıştır.

GEREÇ VE YÖNTEM

Veri Seti

Çalışmada UCI veri tabanından alınan hepatit veri seti kullanılmıştır. 155 hastaya ait yaş, cinsiyet, bilirubin, steroid kullanımı, antiviral kullanımı, tükenmişlik, kırgınlık, iştahsızlık, karaciğerde büyüme, karaciğerde sertlik gibi değişkenler kullanılmıştır.

Hibrit Model Yapısı

Hibrit model için ilk önce kullanılan veri seti için en iyi performansa sahip beş veri madenciliği yöntemi oluşturulan java kaynak kodlu yazılım sayesinde seçilmektedir. İkinci aşama olarak seçilen yöntemler en iyi performansa sahip yöntemden en kötü performansa sahip yönteme doğru sıralanmaktadır. Sonraki aşamada en iyi performansa sahip yöntem hibrit model için ilk seçilen yöntem olacak şekilde, bu yöntemlerden sırasıyla ikili, üçlü ve dörtlü yöntemler grubu oluşturulmaktadır. Oluşturulan bu gruplara ait performans ölçütleri teker teker hesaplanmakta ve en iyi sonuç veren grup baz alınarak hibrit model oluşturulmaktadır.

İstatistiksel Analiz

Verilerin analizinde R 4.0.2 programlama dilinden faydalanılmıştır. Tanımlayıcı olarak nicel değişkenler için ortalama±standart sapma ve ortanca (minimum-maksimum), nitel değişkenler için ise hasta sayısı (yüzde) verilmiştir. Nicel değişken bakımından iki kategoriye sahip nitel değişkenin kategorileri arasında fark olup olmadığına; normal dağılım varsayımları sağlanmadığı için Mann-Whitney U testi kullanılarak bakılmıştır. İki nitel değişken arasındaki ilişki incelenmek ise istendiğinde Ki-kare ve Fisher-exact testleri kullanılmıştır. İstatistiksel anlamlılık düzeyi 0,05 olarak alınmıştır. Veri madenciliği yöntemlerinin hepsi denenerek her analiz için en iyi sonuç veren 5 yöntem kullanılmıştır. Verilerin analizleri 1000 tekrar üzerinden yapılmıştır ve genelleme yöntemi olarak 10-kat çapraz geçerlik kullanılmıştır.

Performans ölçütü olarak doğru sınıflama oranı, F-ölçütü, kesinlik, Mathews Korelasyon Katsayısı(MKK), PRC alanı ve EAKA kullanılmıştır. R programlama dilinde ise e1071, pROC, RSNNS, JWileymisc, ppclust, factoextra, pheatmap, dplyr ve fclust paketleri kullanılmıştır.

BULGULAR

Tablo-1 ve Tablo-2'de hepatit veri seti gerçek verilerine ait tanımlayıcılar verilmiştir. Nicel değişkenler için sınıf değişkeni kategorilerilerine (Yaşıyor, Ölü) göre istatistiksel olarak anlamlı farklılıklar ve tanımlayıcı istatistikler verilmiştir. Nicel değişkenlerden yaş, bilirubin, Alk fosfat, Albumin ve Protime için farklar istatistiksel olarak anlamlı bulunmuştur (sırasıyla $p=0,002$, $p<0,001$, $p=0,034$, $p<0,001$ ve $p<0,001$).

Tablo-1. Nicel Değişkenler için Tanımlayıcılar ve İstatistiksel Karşılaştırmalar

Değişkenler	Sınıf				p değeri
	Yaşıyor		Ölü		
	Ort.±SS	Ortanca (Min.-Maks.)	Ort.±SS	Ortanca (Min.-Maks.)	
Yaş, (n=155)	39,80±12,83	38,00 (7,00-78,00)	46,59±9,95	46,50 (30,00-70,00)	0,002[a]
Bilirubin, (n=149)	1,15±0,72	1,00 (0,30-4,60)	2,54±1,94	1,95 (0,40-8,00)	<0,001[a]
Alk Fosfat, (n=126)	101,31±50,26	85,00 (26,00-295,00)	122,38±54,35	113,50 (62,00-280,00)	0,034[a]
Sgot, (n=151)	82,44±86,51	55,00 (14,00-648,00)	99,83±101,77	66,00 (16,00-528,00)	0,222[a]
Albumin, (n=139)	3,98±0,56	4,00 (2,70-6,40)	3,15±0,60	3,30 (2,10-4,20)	<0,001[a]
Protime, (n=88)	66,57±21,91	66,00 (0,00-100,00)	43,50±16,76	39,00 (29,00-90,00)	<0,001[a]

a:Mann-Whitney U testi, Ort.:Ortalama, SS:Standart Sapma, Min.:Minimum, Maks.: Maksimum

Tablo 2'de nitel değişkenler ile Class değişkeni kategorilerileri (Yaşıyor, Ölü) arasında istatistiksel olarak anlamlı ilişkilere bakılmış ve tanımlayıcı istatistikler verilmiştir. Nitel değişkenlerden Cinsiyet, Tükenmişlik, Kırgınlık, Spleen Palpable, Spiders, Karın İltihabı, Varis ve Histoloji ile Sınıf değişkeni arasında istatistiksel olarak anlamlı ilişki bulunmuştur (sırasıyla p=0,044, p<0,001, p<0,001, p=0,003, p<0,001, p<0,001, p<0,001 ve p<0,001).

Tablo-2. Nitel Değişkenler için Tanımlayıcılar ve İstatistiksel Karşılaştırmalar.

Değişkenler		Sınıf				p değeri
		Yaşıyor		Ölü		
		Sayı	%	Sayı	%	
Cinsiyet	Erkek	107	87,0	32	100,0	0,044[b]
	Kadın	16	13,0	0	0,0	
Steroid	Yok	56	45,9	20	62,5	0,095[a]
	Var	66	54,1	12	37,5	
Antiviraller	Yok	22	17,9	2	6,2	0,168[b]
	Var	101	82,1	30	93,8	
Tükenmişlik	Yok	70	57,4	30	93,8	<0,001[a]
	Var	52	42,6	2	6,2	
Kırgınlık	Yok	38	31,1	23	71,9	<0,001[a]
	Var	84	68,9	9	28,1	
İştahsızlık	Yok	22	18,0	10	31,2	0,101[a]
	Var	100	82,0	22	68,8	
Karaciğerde Büyüme	Yok	22	18,6	3	11,1	0,572[b]
	Var	96	81,4	24	88,9	
Karaciğerde Sertlik	Yok	47	40,2	13	48,1	0,449[a]
	Var	70	59,8	14	51,9	
Spleen Palpable	Yok	18	15,1	12	38,7	0,003[a]
	Var	101	84,9	19	61,3	
Spiders	Yok	29	24,4	22	71,0	<0,001[a]
	Var	90	75,6	9	29,0	
Karın İltihabı	Yok	6	5,0	14	45,2	<0,001[b]
	Var	113	95,0	17	54,8	

Varis	Yok	7	5,9	11	35,5	<0,001[b]
	Var	112	94,1	20	64,5	
Histoloji	Yok	78	63,4	7	21,9	<0,001[a]
	Var	45	36,6	25	78,1	

a:Ki-kare testi, **b:**Fisher-exact testi

Tablo 3'teki veri madenciliği yöntemlerinin performanslarını karşılaştırmak için literatürle uyumlu olması açısından en önemli karar verici performans ölçütü olarak MKK belirlenmiş ve MKK değeri benzer (yakın) olan yöntemler için bu ölçüte ek olarak doğru sınıflama oranı ve F-ölçütü değerlerine bakılması uygun görülmüştür. Bu ölçütlere bakıldığında en iyi performansa Hibrit Model ile ulaşıldığı görülmektedir. Hibrit Model'i sırasıyla Destek Vektör Makinesi, Random Forest, J48, Çok Katmanlı Algılayıcı ve Lojistik Regresyon yöntemleri izlemektedir.

Tablo-3. Hepatit Gerçek Veri Seti için Veri Madenciliği Yöntemlerine ait Performans Ölçütleri.

Yöntemler		**Performans Ölçütleri**					
		Doğru Sınıflama Oranı	**F-ölçütü**	**Kesinlik**	**MKK**	**PRC Alanı**	**EAKA**
Çok Katmanlı Algılayıcı	**Yaşıyor**	0,894	0,891	0,887	0,462	0,958	0,855
	Ölü	0,563	0,571	0,581	0,462	0,600	0,855
	Genel	0,826	0,825	0,824	0,462	0,884	0,855
Destek Vektör Makinesi	**Yaşıyor**	0,919	0,908	0,897	0,532	0,888	0,756
	Ölü	0,594	0,623	0,655	0,532	0,473	0,756
	Genel	0,852	0,849	0,847	0,532	0,803	0,756
Lojistik Regresyon	**Yaşıyor**	0,919	0,893	0,869	0,426	0,898	0,802
	Ölü	0,469	0,526	0,600	0,426	0,562	0,802
	Genel	0,826	0,818	0,814	0,426	0,829	0,802
J48	**Yaşıyor**	0,943	0,903	0,866	0,450	0,856	0,708
	Ölü	0,438	0,528	0,667	0,450	0,585	0,708
	Genel	0,839	0,825	0,825	0,450	0,800	0,708
Random Forest	**Yaşıyor**	0,951	0,914	0,880	0,523	0,961	0,869
	Ölü	0,500	0,593	0,727	0,523	0,654	0,869
	Genel	0,858	0,848	0,848	0,523	0,897	0,869
Hibrit Model	**Yaşıyor**	0,976	0,949	0,923	0,730	0,991	0,832
	Ölü	0,688	0,772	0,880	0,730	0,836	0,832
	Genel	0,916	0,912	0,914	0,730	0,959	0,832

Tablo 4'teki veri madenciliği yöntemlerinin performanslarını karşılaştırmak için literatürle uyumlu olması açısından en önemli karar verici performans ölçütü olarak MKK belirlenmiş ve MKK değeri benzer (yakın) olan yöntemler için bu ölçüte ek olarak Accuracy ve F- measure değerlerine bakılması uygun görülmüştür. Bu ölçütlere bakıldığında en iyi performansa Hibrit Model ile ulaşıldığı görülmektedir. Hibrit Model'i sırasıyla Destek Vektör Makinesi, Lojistik Regresyon, Çok Katmanlı Algılayıcı, Random Forest ve J48 yöntemleri izlemektedir.

Tablo-4. Hepatit 250 Hastadan Oluşan Simüle Veri Seti için Veri Madenciliği Yöntemlerine ait Performans Ölçütleri.

Yöntemler		**Performans Ölçütleri**					
		Doğru Sınıflama Oranı	**F-ölçütü**	**Kesinlik**	**MKK**	**PRC Alanı**	**EAKA**
Çok Katmanlı Algılayıcı	**Yaşıyor**	0,934	0,927	0,920	0,641	0,970	0,907
	Ölü	0,692	0,713	0,735	0,641	0,712	0,907
	Genel	0,884	0,883	0,882	0,641	0,916	0,907
Destek Vektör Makinesi	**Yaşıyor**	0,949	0,938	0,926	0,687	0,919	0,831
	Ölü	0,712	0,747	0,787	0,687	0,620	0,831
	Genel	0,900	0,898	0,897	0,687	0,857	0,831

Lojistik Regresyon	**Yaşıyor**	0,939	0,930	0,921	0,651	0,977	0,928
	Ölü	0,692	0,720	0,750	0,651	0,735	0,928
	Genel	0,888	0,886	0,885	0,651	0,928	0,928
J48	**Yaşıyor**	0,899	0,890	0,881	0,451	0,857	0,732
	Ölü	0,538	0,560	0,583	0,451	0,543	0,732
	Genel	0,824	0,821	0,819	0,451	0,792	0,732
Random Forest	**Yaşıyor**	0,965	0,920	0,880	0,557	0,977	0,926
	Ölü	0,500	0,612	0,788	0,557	0,774	0,926
	Genel	0,868	0,856	0,861	0,557	0,935	0,926
Hibrit Model	**Yaşıyor**	0,980	0,960	0,942	0,798	0,998	0,875
	Ölü	0,769	0,833	0,909	0,798	0,857	0,875
	Genel	0,936	0,934	0,935	0,798	0,969	0,875

Tablo 5'teki veri madenciliği yöntemlerinin performanslarını karşılaştırmak için literatürle uyumlu olması açısından en önemli karar verici performans ölçütü olarak MKK belirlenmiş ve MKK değeri benzer (yakın) olan yöntemler için bu ölçüte ek olarak Accuracy ve F-measure değerlerine bakılması uygun görülmüştür. Bu ölçütlere bakıldığında en iyi performansa Hibrit Model ile ulaşıldığı görülmektedir. Hibrit Model'i sırasıyla Destek Vektör Makinesi, Lojistik Regresyon, Çok Katmanlı Algılayıcı, Random Forest ve J48 yöntemleri izlemektedir.

Tablo-5. Hepatit 500 Hastadan Oluşan Simüle Veri Seti.

Yöntemler		**Performans Ölçütleri**					
		Doğru Sınıflama Oranı	**F-ölçütü**	**Kesinlik**	**MKK**	**PRC Alanı**	**EAKA**
Çok Katmanlı Algılayıcı	**Yaşıyor**	0,947	0,945	0,942	0,731	0,986	0,953
	Ölü	0,779	0,786	0,794	0,731	0,865	0,953
	Genel	0,912	0,912	0,911	0,731	0,953	0,953
Destek Vektör Makinesi	**Yaşıyor**	0,962	0,950	0,938	0,750	0,933	0,861
	Ölü	0,760	0,798	0,840	0,750	0,688	0,861
	Genel	0,920	0,920	0,918	0,750	0,882	0,861
Lojistik Regresyon	**Yaşıyor**	0,960	0,949	0,938	0,744	0,985	0,951
	Ölü	0,760	0,794	0,832	0,744	0,857	0,951
	Genel	0,918	0,917	0,916	0,744	0,958	0,951
J48	**Yaşıyor**	0,949	0,934	0,919	0,665	0,916	0,846
	Ölü	0,683	0,728	0,780	0,665	0,657	0,846
	Genel	0,894	0,891	0,890	0,665	0,862	0,846
Random Forest	**Yaşıyor**	0,972	0,947	0,923	0,725	0,987	0,957
	Ölü	0,692	0,770	0,867	0,725	0,880	0,957
	Genel	0,914	0,910	0,912	0,725	0,965	0,957
Hibrit Model	**Yaşıyor**	0,990	0,984	0,978	0,920	0,998	0,952
	Ölü	0,913	0,936	0,960	0,920	0,948	0,952
	Genel	0,974	0,974	0,974	0,920	0,988	0,952

Özet olarak gerçek veri setinde ve 250, 500 örnekleme sahip veri setlerinde performans ölçütleri hesaplanıp karşılaştırmalar yapıldığında örneklem büyüklüğü arttıkça performans ölçütlerinin de arttığı görülmüştür. Gerçek veri seti için MKK değeri 0,730 olarak bulunurken, bu değer 250 örnekleme sahip veri setinde 0,798 ve 500 örnekleme sahip veri setinde 0,920 olarak bulunmuştur.

TARTIŞMA

Algoritmalar ve görselleştirme araçları gibi veri madenciliği algoritmalarındaki son gelişmeler, her türlü ham verinin üst düzey bilgiye dönüştürülmesini mümkün kılmıştır. Standart veri madenciliği yöntemlerine ait kısıtlılıklar hibrit veri madenciliği yaklaşımına ihtiyacı ortaya çıkarmıştır.

Ha ve Joo, göğüs ağrısı şikayetiyle tedavi gören 478 hasta üzerinde yaptıkları çalışmada, hibrit model ile Sinir Ağı ve Destek Vektör Makinası yöntemlerinin sonuçlarını karşılaştırmıştır ve hibrit modelin en iyi performansa sahip yöntem olduğunu bildirmişlerdir (4). Simsek ve ark., 240 meme kanseri hastasına ait 7399 gen üzerinde yaptıkları çalışmada Genetik Algoritma adını verdikleri bir hibrit

model oluşturmuşlardır ve hibrit modelin Yapay Sinir Ağı ve Lojistik Regresyon yöntemlerine göre daha iyi performansa sahip olduğunu bildirmişlerdir (5). Çalışmamızda literatürle uyumlu olarak hibrit modele ait performans ölçütleri vei madenciliği yöntemlerine göre daha yüksek bulunmuştur. Üretilen simüle veriler üzerindeki 27 senaryoda ve gerçek veri setleri üzerinde yapılan analizler sonucunda hibrit modelin daha iyi sonuç verdiği, performansı ciddi oranda iyileştirdiği görülmüştür.

Çalışmamızda tüm veri setlerinde hibrit model daha iyi sonuçlar vermiştir. Uygun veri madenciliği yöntemlerinin birlikte kullanımı durumunda, yapılacak çalışmalarda hibrit model yaklaşımının daha iyi modeller elde etmemizi sağlayacabileceği sonucuna ulaşılmıştır.

KAYNAKLAR

1. Kantardzic, M (2011). Data Mining: Concepts, Models, Methods, and Algorithms. John Wiley & Sons.
2. Wu X, Kotagiri R, Korb KB (2006). Research and Development in Knowledge Discovery and Data Mining: Second Pacific-Asia Conference, PAKDD'98, Melbourne, Australia. Springer.
3. Syed MR, Syed SN. (2008). Handbook of Research on Modern Systems Analysis and Design Technologies and Applications. IGI Global.
4. Ha SH, Joo SH (2010). A hybrid data mining method for the medical classification of chest pain. International Journal of Computer and Information Engineering, 4(1): 33-38.
5. Simsek S, Kursuncu U, Kibis E, Anisabdellatif M, Dag A (2020). A hybrid data mining approach for identifying the temporal effects of variables associated with breast cancer survival. Expert Systems with Applications, 139: 1-13.

Küçük Küme ve Periyotlu Boylamsal Çalışmalarda Değişken ve Model Seçimi

[1]Merve TURKEGÜN, [2]Bahar TAŞDELEN, [3]Saim YOLOĞLU

[1]Mersin Üniversitesi Biyoistatistik Bölümü, Mersin, Türkiye
[2]Mersin Üniversitesi Biyoistatistik Bölümü, Mersin, Türkiye
[3]İnönü Üniversitesi, Tıp Fakültesi, Biyoistatistik ve Tıbbi Bilişim Anabilim Dalı, Malatya

Amaç: Boylamsal verilerde, küçük küme sayısının sebep olduğu en önemli problemlerden biri bir ya da daha fazla açıklayıcı değişkenin ikili sonuç değişkenini mükemmel ya da mükemmele yakın olarak tahminlemesi ile ortaya çıkan tam ayrılma ya da aşırı uyum durumudur. Aşırı uyumun gözlendiği durumlarda adımsal değişken seçim yöntemlerinin kullanımı katsayıların standart hatalarını şişirerek yanlı tahminlere sebep olur. Aşırı uyum probleminin çözümü için tüm açıklayıcı değişkenleri aynı anda değerlendirerek, regresyon katsayı tahminlerini bir ceza fonksiyonu yardımıyla sıfıra doğru küçülten cezalandırma yöntemlerinin kullanımı önerilmektedir. Çeşitli ceza fonksiyonları ile Genelleştirilmiş Tahmin Eşitlikleri (GEE) yöntemini birleştiren Cezalı Genelleştirilmiş Tahmin Eşitlikleri (PGEE) yöntemi, aşırı uyum durumunda ve çok yüksek boyutlu boylamsal verilerde değişken seçimi için sıklıkla kullanılan bir yöntemdir. Amacımız; aşırı uyum problemine sebep olan açıklayıcı değişken sayısının küme sayısında

Gereç ve Yöntem: 500 tekrarlı simülasyon verisinin üretimi ve analizler için R programlama dili kullanılmıştır. Simülasyon çalışmasında, populasyonda nadir görülen hastalıkların araştırıldığı klinik çalışmalarda sıklıkla karşılaşılan veri yapılarına ve çalışma kısıtlılıklarına benzerlik göstermesi amaçlanmıştır. Bu benzerlikler, ölçümlerin tekrar ve küme sayıları dikkate alınarak yakalanmaya çalışılmıştır. Tekrar sayısı sayısının minimumda olduğu (k=2), küme sayısının (N); 20, 30, 50,100 ve 200 olduğu durumlar için açıklayıcı değişken sayısı (P);10,20,50 olacak şekilde ve açıklayıcı değişkenler arası korelasyonun r=0.20, 0.50 ve 0.80 olan çok değişkenli normal dağılıma uygun veri üretilmiştir. İkili yapıda sonuç değişkeni binom dağılımından üretilmiştir. Sonuç değişkenin tekrarlanan ölçümü için gözlemler arasındaki korelasyon 0.40 alınmış bu sebeple çalışan korelasyon yapısı sadece "değiştirilebilir" olarak seçilebilmiştir. GEE ve PGEE model performansları, katsayılara ait Medyan Kare Hatası (MEDSE), Medyan Mutlak Hatası (MDAE) ve Bias istatistikleri ile değerlendirilmiştir. PGEE modelinde değişken seçiminde, katsayı değeri>0,001 olanlar sıfırdan farklı kabul edilmiştir. Değişken seçim performansları ise doğru seçim (Correct) ve hatalı seçim (Incorrect), kesinlik (Precision), hassasiyet (Recall) ve F1 skor kriterleri kullanılmıştır.

Bulgular: Değişken ve küme sayısının birbirine yakın olduğu durumlarda (P=10, 20, 50; N=20, 30, 50) GEE katsayılarının MEDSE, MDAE ve Bias değerleri çok büyük hesaplanmıştır. PGEE ise tüm senaryolarda 0'a yakın MEDSE ve MDAE değerlerine sahip olmuştur.GEE etkisiz sayılabilecek tüm katsayıları (sıfıra eşit olan) sıfırdan farklı tahminlemiştir. Bu sebeple kesinlik ve F1 skoru hesaplanamamıştır. Hassasiyet değeri ise tüm senaryolarda 0 olarak hesaplanmıştır. Değişken sayısı ile küme sayısının arttırılması PGEE'nin tahmin gücünü arttırmıştır (P=50, N=200 ,F1 skor=0,78).

Sonuç: PGEE, P>n durumunda hem küme içindeki hem de değişkenler arasındaki korelasyonlu yapıyı açıklamada GEE'ye göre daha başarılı olmuştur. PGEE, hem küme içindeki hem de değişkenler arasındaki ilişkileri kontrol altında tutarak daha tutarlı sonuçlar üretmiştir. PGEE'nin, değişken seçiminden ziyade değişken ayıklama mantığıyla çalıştığı görülmüştür. Doğrudan etkili değişkenleri seçmek yerine dolaylı olarak seçerek, modele etkisi olmayan değişkenleri tespit etmede daha başarılı olduğu tespit edilmiştir.

Anahtar Kelimeler: GEE, cezalı GEE, değişken seçimi, model seçimi

1. GİRİŞ

İlişkili sonuç değişkeninin yer aldığı boylamsal verilerin modellemesinde kullanılan en popüler yaklaşım ise Genelleştirilmiş Tahmin Eşitlikleri (Generalized Estimating Equations, GEE) yöntemidir (1). GEE, küme sayısının küçük olduğu durumlarda, GEE tahmin edicilerine ait tip I hata (α) oranları artış gösterir ve yanlı tahminler üretir (2). Küçük küme sayısının GEE'de sebep olduğu en önemli problemlerden biri, açıklayıcı değişkenlerin, ikili sonuç değişkenini mükemmel ya da mükemmele yakın olarak tahminlemesi ile ortaya çıkan aşırı uyum (overfitting) durumudur (3). Örneklem ya da küme sayısının az olması, nadir görülen sonuç değişkeninin varlığı, sonuç değişkeni için olayın görülme ve görülmeme dağılımının dengesizliği, açıklayıcı değişken sayısının çok olması ve yüksek küme içi korelasyonlar aşırı uyuma neden olur (4).

Açıklayıcı değişken sayısının fazla olduğu durumlarda, regresyon modellerinde değişken seçimi için geleneksel adımsal yöntemler yerine son yıllarda cezalı regresyon yöntemlerinin kullanımı popüler hale gelmiştir. Çok sayıda açıklayıcı değişkenin yer aldığı çalışmalarda değişkenlerin doğrusal kombinasyonları aşırı uyumu ortaya çıkarır. Özellikle bu durum küçük örneklem genişliklerinde görülür. Ancak yeterli büyüklükte bir örneklem olsa bile eğer nadir görülen bir sonuç değişkeni ile çalışılıyorsa; tam ayrılma ortaya çıkabilir (5). Bu

sebeple özellikle nadir görülen ikili sonuç değişkenin varlığında aşırı uyum problemine sebep olan açıklayıcı değişken sayısının küme sayısından büyük olma (P>n) durumuyla birlikte bu durumun beraberinde gelen çoklu bağlantı (ÇB) sorununun, GEE ve cezalı GEE (PGEE) modellerinin değişken ve model seçim performanslarına olan etkisi, simülasyon çalışması ile incelenmesi amaçlanmıştır.

2. MATERYAL VE METOT

Simülasyon çalışmasında, özellikle popülasyonda nadir görülen hastalıkların araştırıldığı klinik çalışmalarda sıklıkla karşılaşılan veri yapılarına ve çalışma kısıtlılıklarına benzerlik göstermesi amaçlanmıştır. Tekrarlanan ölçüm sayısının iki olduğu durumda; küme sayısının, değişken sayısının ve değişkenler arasındaki korelasyonun; GEE ve PGEE modellerinin model performanslarına ve değişken seçimine olan etkisi incelenmiştir.

Simülasyon verisinin üretimi ve analizler için R programlama dili kullanılmıştır. Simülasyon adımları ise şöyledir;

- İkili yapıdaki bağımlı değişken binom dağılımından türetilmiştir.
- Bağımlı değişkenin tekrarlanan ölçümü için gözlemler arasındaki korelasyon 0.40 ve çalışan korelasyon yapısı "değiştirilebilir" olarak belirlenmiştir.
- Açıklayıcı değişkenler arasında tekrarlı ölçüm sayısı 2 olarak belirlenmiştir.
- Küme büyüklüğü N=20, 30, 50, 100, 200 olarak belirlenmiştir.
- Her kümede açıklayıcı değişken sayısı P=10, 20, 50 olacak şekilde alınmıştır.
- Açıklayıcı değişkenler 0 ortalama, 1 standart sapma ve aralarındaki korelasyon r=0.20,0.50 ve 0.80 olacak şekilde çok değişkenli normal dağılımdan türetilmiştir.
- Üretilen verilere GEE ve PGEE analizleri uygulanmıştır.
- PGEE'de kullanılmak üzere, başlangıç beta katsayıları sırasıyla; 0.5, 0.01, 0.5, 6.4 alınmıştır. Değişken sayısı 10 ve 20 olan senaryolarda 3 katsayı değeri (-0.5, 0.01, 0.5) değişken sayısının 50 olan senaryoda ise 4 katsayı değeri (-0.5, 0.01, 0.5, 6.4) başlangıçta belirlenmiştir. Geri kalan beta katsayı değerleri ise 0 alınmıştır.
- GEE ve PGEE modellerinin tahmin doğruluğunun değerlendirmek amacıyla katsayılar için medyan kare hatası (MEDSE), medyan mutlak hatası (MDAE) ve BIAS istatistikleri hesaplatılıp kaydedilmiştir (Eşitlik 1).

$$\mathrm{MEDSE} = \mathrm{medyan}_s[(\hat{\beta}_s - \beta_s)^2]) \qquad \mathrm{MDAE} = \mathrm{medyan}_s(|(\hat{\beta}_s - \beta_s)|) \qquad \text{Eşitlik 1}$$

- Değişken seçim performansının değerlendirilmesi için *"doğru seçim"* (Correct, C) ve *"hatalı seçim"* (Incorrect, IC), kesinlik (Precision), hassasiyet (Recall) ve F_1 skor kriterleri hesaplatılmıştır.GEE ve PGEE modellerinin belirtilen gerçek beta katsayılarını sıfır ve sıfırdan farklı tahminlemesine göre sınıflama performansları değerlendirilmiştir. Gerçek değeri sıfır iken tahminlenen değeri de sıfır olan katsayıların 500 simülasyon için ortalama sayısı doğru pozitif ya da doğru tahmin (C) değerleri ile gerçek değeri sıfırdan farklı iken tahminlenen değeri sıfır olan katsayıların 500 simülasyon için ortalama sayısı yanlış pozitif ya da yanlış tahmin (IC) değerleri hesaplanmıştır. Yöntemlerin değişken seçim performansları için kesinlik (Precision) (eşitlik 2), hassasiyet (Recall) (eşitlik 3) ve F_1 skor değerleri (eşitlik 3) ile hesaplandı.

$$\mathrm{Kesinlik} = \frac{\mathrm{DP}}{\mathrm{DP+YP}} \qquad \text{Eşitlik 2}$$

$$\mathrm{Hassiyet} = \frac{\mathrm{DP}}{\mathrm{DP+YN}} \qquad \text{Eşitlik 3}$$

$$\mathrm{F_1skoru} = 2 * \frac{\mathrm{Kesinlik*Hassasiyet}}{\mathrm{Kesinlik+Hassasiyet}} \qquad \text{Eşitlik 4}$$

2.1. Genelleştirilmiş Tahmin Eşitlikleri

Genelleştirilmiş tahmin eşitlikleri, ortalama sonuç değişkenini modelleyerek popülasyon ortalaması hakkında çıkarım yapılmasını sağlar (9). GEE'de her bir denek bir küme olarak kabul edilir. Küme içindeki gözlemlere ait sonuç değişkeni kendi arasında bağımlı olup kümeler arasında birbirinden bağımsızdır.

i. birey için sonuç vektörü $Y_i = (Y_{i1}, \ldots, Y_{ini})^T$ ve kovaryat matrisi $X_{ij}=(X_{i1}, \ldots, X_{ini})^T$ olmak üzere Marjinal regresyon yaklaşımında, GEE'de her bir sonuç değişkeni için beklenen değer vektörü; $E(Y_{ij}|X_{ij}) = g(X'_{ij}\beta) = \mu_{ij} = (\mu_{i1}, \ldots, \mu_{in})$ olarak yazılır.

İkili sonuç değişkeni $Y_{ij}(0,1)$ için lojit bağlantı fonksiyonu eşitlik 5 ile gösterilir.

$$g^{-1}(\mu_{ij}) = \log\left(\frac{\mu_{ij}}{1-\mu_{ij}}\right) = \text{logit}(\mu_{ij}) = X'_{ij}\beta$$ Eşitlik 5

Her bir sonuç değişkeni Y_{ij}'nin varyans-kovaryans matrisi $V_i = \phi A_i^{\frac{1}{2}} R_i(\alpha) A_i^{\frac{1}{2}}$'dir. *A*; köşegen elemanları sonuç değişkenin varyansı ile ortalaması arasındaki ilişkiyi açıklamaya yardımcı olan $\upsilon(\mu_{ij})$'lerden oluşan köşegen matristir. $(A = \text{Diag}\{\upsilon(\mu_{i1}), \ldots, \upsilon(\mu_{in_i})\})$. ϕ yayılım parametresidir. $R_i(\alpha)$, n_i x n_i boyutlu olup tekrarlanan gözlemler arasındaki ilişkiyi gösteren çalışan korelasyon matrisi, α ise gözlemler arası ilişkileri gösteren model ile ilişkili parametreler vektörüdür.

β'yı tahminleme aşamasında kovaryans matrisinin doğru belirlenmesi katsayıların tutarlığı için önemlidir. Sandviç (sandwich) tahmin edicileri olarak da bilinen robust tahmin edicileri, kovaryans parametresi tam belirlenemediğinde ya da çalışan korelasyon matrisi yanlış belirlediğinde tutarlı varyans tahminleri ve standart hatalar üretirler. Y_i'nin varyans-kovaryans matrisinin sandviç tahmin edicisi $\text{Var}(\hat{\beta}) = I_0^{-1} I_1 I_0^{-1}$ olup; burada, $I_0 = \left(\sum_{i=}^{K} D_i' V_i^{-1} D_i\right)$ ve $I_1 = \left(\sum_{i=}^{K} D_i' V_i^{-1} \text{Cov}(Y_i) V_i^{-1} D_i\right)$ (10,11-12).

GEE'de, çalışan korelasyon matrisleri tekrarlanan ölçümler arasındaki ilişkileri belirlemek için kullanılır. GEE'nin en büyük avantajı da bu çalışan korelasyon yapısı yanlış belirtilmiş olsa bile yansız ve tutarlı parametre tahminleri sağlamasıdır (13). Bağımsız, değiştirilebilir, yapılandırılmamış, birinci dereceden otoregresif, M-bağımlı ve sabit çalışan korelasyon yapıları olmak üzere kullanım amacına, küme sayısına ve gözlem sayısına göre değişen farklı çalışan korelasyon yapıları bulunmaktadır. Uygun korelasyon yapısı, *"Yarı-Olasılık Bilgi Kriteri"* (Quasi- Likelihood Information Criterion-QIC(R)) ile belirlenir (14).

2.2. Cezalı Genelleştirilmiş Tahmin Eşitlikleri

Cezalı genelleştirilmiş tahmin eşitlikleri yöntemi (Penalized Generalized Estimating Equations, PGEE), hem tahminin hem de değişken seçiminin aynı anda yapılmasını sağlamak amacıyla geliştirilmiştir (15). Cezalandırılmış tahmin denklemleri, küçük katsayıları sıfıra büzüştürür, böylece sıfır farklı olan katsayıların robust tahmin edicilerini üretirken aynı zamanda değişken seçimi gerçekleştirebilmektedir. Cezalı GEE modeli eşitlik 6 ile gösterilir.

$$U(\beta) = S(\beta) - q_\lambda(|\beta|) * \text{sign}(\beta)$$ Eşitlik 6

$S(\beta)$, GEE için tahmin fonksiyonudur ve eşitlik 7 ile ifade edilir.

$$S(\beta) = \sum_{i=}^{K} X_i^T A_i^{1/2} \widehat{R(\alpha)}^{-1} A_i^{-1/2} (Y_i - \mu_i) = 0$$ Eşitlik 7

Eşitlik 6'da; $q_{\lambda_n}(|\beta_n|) = \left(q_{\lambda_n}(|\beta_{n1}|), \ldots, q_{\lambda_n}(|\beta_{n_{pn}}|)\right)^T$ p_n boyutlu ceza fonksiyonunu, λ; ceza fonksiyonu için büzülme miktarını belirleyen ayar (tuning) parametresini; $\text{sign}(t) = I(t>0) - I(t<0)$ ile $\text{sign}(\beta_n) = \left(\text{sign}(\beta_{n1}), \ldots, \text{sign}(\beta_{n_{pn}})\right)^T$ olmak üzere $q_\lambda(|\beta|) * \text{sign}(\beta)$ ise iki vektörün element bazında çarpımını ifade etmektedir. Ceza fonksiyonu $q_{\lambda_n}(|\beta_j|)$'nın rolü, küçük regresyon katsayılarının tahminlerini sıfıra doğru büzmektir. Tahminleri sıfıra doğru küçültülen (büzdürülen) katsayılar final modelden çıkarılırlar (15,16).

2.2.1. Kolayca Kesilmiş Mutlak Sapma Cezası

Ceza fonksiyonunun tahmin edicilerinde olması gereken üç temel özellik yansızlık, seyreklik ve sürekliliktir. SCAD ceza fonksiyonu, bu üç özelliği de taşımakta ve dış bükey (convex) olmayan, $[0, +\infty]$ aralığında tanımlanmaktadır (17). Ceza fonksiyonu eşitlik 8 ile gösterilir.

$$q_{\lambda_n}(\theta) = \lambda_n \left\{ I(\theta \leq \lambda_n) + \frac{(\alpha\lambda_n - \theta)+}{(\alpha-1)\lambda_n} I(\theta > \lambda_n) \right\}$$ Eşitlik 8

Eşitlik 8 için; $\lambda \geq 0$, $\theta \geq 0$ ve $\alpha > 2$ c doğruysa I(c)=1, değilse 0 değerini almak üzere I(.) indikatör fonksiyonudur ve gerçek b sayısı için $b_+ = bI(b>0)$'dır ve α=3.7'dir.

Dış bükey olmayan SCAD cezası, büyük katsayıları aşırı cezalandırmadan kaçınır ve böylece tutarlı değişkenlerin seçimine olanak sağlar. Gereksiz kovaryat değişkenlere ait katsayıları tam olarak sıfıra büzüştürür (17).

Cezalı regresyon yöntemlerinde, hem elde edilen tahmin edicilerin robust ve tutarlı olması hem de önemli değişkenlerin belirlenmesi ceza fonksiyonunda yer alan ayar (tuning) parametresinin seçimine bağlıdır. Gereğinden büyük seçilen ayar parametresi veride aşırı cezalandırmaya, gereğinden küçük seçilmesi ise düşük cezalandırmaya sebep olmaktadır. SCAD ceza fonksiyonu, λ ve α olmak üzere iki tip ayar parametresi içerir. α ayar parametresini α=3.7 sabit olup; λ ayar parametresi ise optimal değeri çapraz geçerlilik (CV) kriteri kullanılarak belirlenir (15) (Eşitlik 9).

$$PE_{-k}(\lambda) = \frac{1}{|N_{-k}|} \sum_{i \in N_{-k}} \frac{1}{n_i} \sum_{t=1}^{n_i} \left(Y_{it} - g(X_{it}^T \hat{\beta})\right)^2 \qquad CV(\lambda) = \frac{1}{K} \sum_{k=1}^{K} PE_{-k}(\lambda)$$ Eşitlik 9

3. BULGULAR

3.1. Değişken Sayısı 10 iken PGEE ve GEE için Model ve Değişken Seçim Performansları

Tüm senaryolarda PGEE katsayılarının MEDSE, MDAE ve BIAS değerleri 0 ve altında elde edilmiştir. GEE, küçük kümelerde büyük kümelere göre korelasyonun artışından daha çok etkilenmiştir. Ayrıca küme sayısı arttıkça tüm korelasyonlar için MEDSE, MDAE ve BIAS düşüş göstermiştir. PGEE modeli küçük kümelerde de korelasyondan etkilenmezken PGEE, GEE'ye göre daha tutarlı sonuçlar üretmiştir (Tablo-1).

Tablo-1. 10 değişkenli PGEE ve GEE modellerinin katsayılarına ait

P=10		r=0.20			r=0.50			r=0.80		
Medyan [IQR]		MEDSE	MDAE	BIAS	MEDSE	MDAE	BIAS	MEDSE	MDAE	BIAS
N=20	PGEE	<0.0001 [<0.0001]	0.00001 [0.00000]	0.001 [0.000]	<0.0001 [<0.0001]	0.00001 [0.00002]	0.001 [0.000]	<0.0001 [<0.0001]	0.00001 [0.00003]	0.001 [0.000]
	GEE	**0.168 [0.346]**	**0.4064 [0.3593]**	**0.006 [0.157]**	**0.229 [0.474]**	**0.47639 [0.43086]**	**-0.003 [0.100]**	**0.496 [0.858]**	**0.69664 [0.54613]**	**-0.002 [0.079]**
N=30	PGEE	<0.0001 [<0.0001]	0.00001 [0.00003	0.001] [0.000]	<0.0001 [<0.0001]	0.00001 [0.00002]	0.001 [0.000]	<0.0001 [<0.0001]	0.000004 [0.00002]	0.001 [0.000]
	GEE	**0.062 [0.077]**	**0.24353 [0.14578]**	**0.003 [0.084]**	**0.100 [0.113]**	**0.31249 [0.17476]**	**-0.003 [0.063]**	**0.217 [0.260]**	**0.46098 [0.26303]**	**0.001 [0.052]**
N=50	PGEE	<0.0001 [<0.0001]	0.00001 [0.00003]	0.001 [0.000]	<0.0001 [<0.0001]	0.00001 [0.00003]	0.001 [0.000]	<0.0001 [<0.0001]	0.000003 [0.00002]	0.001 [0.000]
	GEE	**0.030 [0.033]**	**0.17104 [0.09491]**	**-0.001 [0.055]**	**0.043 [0.043]**	**0.20534 [0.10126]**	**0.000 [0.041]**	**0.098 [0.116]**	**0.31087 [0.17720]**	**-0.001 [0.034]**
N=100	PGEE	0.000 [0.013]	0.00003 [0.11322]	0.001 [0.000]	<0.0001 [<0.0001]	0.00001 [0.00004]	0.001 [0.000]	<0.0001 [<0.0001]	0.000002 [0.00002]	0.001 [0.000]
	GEE	**0.013 [0.012]**	**0.11133 [0.05004]**	**0.001 [0.040]**	**0.019 [0.020]**	**0.13696 [0.07071]**	**0.002 [0.028]**	**0.047 [0.051]**	**0.21428 [0.11554]**	**-0.002 [0.020]**
N=200	PGEE	0.000 [0.007	0.00719 [0.08194]	0.001 [0.003]	<0.0001 [<0.0001]	0.00002 [0.00003]	0.001 [0.000]	<0.0001 [<0.0001]	0.000004 [0.00003]	0.001 [0.000]
	GEE	**0.006 [0.006]**	**0.07676 [0.03477]**	**0.002 [0.027]**	**0.010 [0.009]**	**0.09762 [0.04659]**	**0.000 [0.019]**	**0.023 [0.022]**	**0.15002 [0.06910]**	**0.000 [0.014]**

Tablo-2. 10 değişkenli PGEE ve GEE yöntemlerinin değişken seçim performans ölçümleri

	r=0.20				r=0.50				r=0.80			
P=10	**PGEE**		**GEE**		**PGEE**		**GEE**		**PGEE**		**GEE**	
	C	**IC**	**C**	**IC**	**C**	**IC**	**C**	**IC**	**C**	**IC**	**C**	**IC**
N=20	492.7	493	0.0	500.0	495.4	491.67	0.0	500.0	485.1	483.33	0.0	500.0
Kesinlik	0.50		-		0.50		-		0.50		-	
Hassasiyet	0.99		0.0		0.99		0.0		0.97		0.0	
F_1 skor	**0.66**				**0.66**				**0.66**			
N=30	485.3	482	0.0	500.0	491.7	491.0	0.0	500.0	488.7	487.67	0.0	500.0
Kesinlik	0.50		-		0.50		-		0.50		-	
Hassasiyet	0.97		0.0		0.98		0.0		0.98		0.0	
F_1 skor	**0.66**				**0.66**				**0.66**			
N=50	447.6	442.33	0.0	500.0	469.1	465.33	0.0	500.0	485.0	483.33	0.0	500.0
Kesinlik	0.50		-		0.50		-		0.50		-	
Hassasiyet	0.90		0.0		0.94		0.0		0.97		0.0	
F_1 skor	**0.64**				**0.65**				**0.66**			
N=100	325.6	305.33	0.0	500.0	444.3	432.67	0.0	500.0	483.1	481	0.0	500.0
Kesinlik	0.52		-		0.51		-		0.50		-	
Hassasiyet	0.65		0.0		0.89		0.0		0.97		0.0	
F_1 skor	**0.58**				**0,65**				**0,66**			
N=200	304.9	236	0.0	500.0	453.3	444	0.0	500.0	488.9	487	0.0	500.0
Kesinlik	0.54		-		0.51		-		0.50		-	
Hassasiyet	0.61		0.0		0.91		0.0		0.98		0.0	
F_1 skor	**0,57**				**0,68**				**0,66**			

Tablo-2'ye göre, GEE yöntemi, tüm senaryolarda katsayıların hepsini sıfırdan farklı tahminlemeyerek değişken seçiminde yetersiz kalmıştır. PGEE ise tüm senaryolar için C ve IC değerleri birbirine çok yakın hesaplanmıştır. Yüksek C değerlerine karşılık yüksek IC değerleri PGEE'nin en cimri modeli oluştururken, sıfırdan farklı katsayıları da sıfır olarak tahminleme ve böylece gereksiz değişkenleri ayıklama eğiliminde olduğunu göstermektedir. Tüm senaryolar için kesinlik değerleri 0.50 olarak hesaplanmıştır. Genel olarak küme sayısı ve korelasyondaki artışlar PGEE'nin tahmin gücünü değiştirmediği gözlemlenmiştir (F_1 skor>0.60).

3.2. Değişken Sayısı 20 iken PGEE ve GEE için Model ve Değişken Seçim Performansları

Tüm senaryolarda PGEE modelinden elde edilen MEDSE, MDAE ve BIAS değerleri 0'dan küçük elde edilmiştir. PGEE modelinin düşük küme sayısı ve yüksek korelasyondan etkilenmediği görülmüştür. Değişken sayısının arttığı (P=20) buna karşılık küme sayısının küçük olduğu durumda GEE modeline ait MEDSE değerleri oldukça yüksek hesaplanmıştır. Korelasyondaki artışlar da bu değerleri etkilemiştir. Küme sayısının artmasıyla GEE için hesaplanan MEDSE değerleri de 0.05'in altına düşmüştür. MEDSE, orta ve yüksek korelasyon için sıfıra yakın hesaplanmıştır. BIAS değerlerinin ise MEDSE değerleri ile birlikte kararsız düşüşleri gözlemlenmiştir. N=100 ve 200 için MDAE değerleri 0.0-0.20 arasında hesaplanmış ve küme sayısının küçük olduğu durumlarda MEDSE değerlerinde artış gözlemlenmiştir. Değişken sayısının küme sayısına eşit ya da yakın olduğu tüm korelasyonlarda GEE modeli PGEE modeline göre zayıf kalmıştır (Tablo 3).

Tablo-3. 20 değişkenli PGEE ve GEE modellerinin katsayılarına ait performans ölçütleri

P=20		r=0.20			r=0.50			r=0.80		
Medyan [IQR]		MEDSE	MDAE	BIAS	MEDSE	MDAE	BIAS	MEDSE	MDAE	BIAS
N=20	PGEE	0.00	0.000005	0.001	0.00	0.000004	0.001	0.00	0.000003	0.001
		[0.00]	[0.00001]	[0.000]	[0.00]	[0.000006]	[0.000]	[0.00]	[0.000011]	[0.000]
	GEE	**2908.34**	**53.91689**	**0.148**	**5568.85**	**74.57746**	**0.022**	**14783.79**	**121.4735**	**-0.047**
		[50147.70]	**[197.0316]**	**[10.088]**	**[5.66E+28]**	**[2.38E+14]**	**[7.624]**	**[2.53E+29]**	**[5.01E+14]**	**[7.858]**
N=30	PGEE	0.00	0.000004	0.001	0.00	0.000003	0.001	0.00	0.000003	0.001
		[0.00]	[0.00001]	[0.000]	[0.00]	[0.00001]	[0.000]	[0.00]	[0.00001]	[0.000]
	GEE	**0.2444**	**0.49245**	**-0.005**	**0.499**	**0.70500**	**-0.003**	**1.105**	**1.05075**	**0.000**
		[3.39E+28]	**[1.84E+14]**	**[0.135]**	**[7.50E+28]**	**[2.73E+14]**	**[0.117]**	**[1.84E+29]**	**[4.26E+14]**	**[0.076]**
N=50	PGEE	0.00	0.00001	0.001	0.00	0.00001	0.001	0.00	0.000002	0.000
		[0.00]	[0.00001]	[0.000]	[0.00]	[0.00003]	[0.000	[0.00]	[0.000007]	[0.000]
	GEE	**0.041**	**0.20323**	**0.002**	**0.043**	**0.20534**	**-0.001**	**0.140**	**0.37188**	**0.000**
		[0.040]	**[0.09711]**	**[0.037]**	**[0.043]**	**[0.10126]**	**[0.024]**	**[0.123]**	**[0.15961]**	**[0.019]**
N=100	PGEE	0.00	0.00001	0.000	0.00	0.000004	0.000	0.00	0.000001	0.000
		[0.00]	[0.00001]	[0.000]	[0.00]	[0.00001]	[0.000]	[0.00]	[0.000012]	[0.000]
	GEE	**0.015**	**0.12091**	**0.000**	**0.022**	**0.14695**	**0.001**	**0.050**	**0.22400**	**0.000**
		[0.010]	**[0.04176]**	**[0.024]**	**[0.017]**	**[0.05769]**	**[0.016]**	**[0.038}**	**[0.08272]**	**[0.012]**
N=200	PGEE	0.00	0.00001	0.001	0.00	0.00001	0.000	0.00	0.000001	0.000
		[0.00]	[0.08206]	[0.000]	[0.00]	[0.00001]	[0.000]	[0.00]	[0.00001]	[0.000]
	GEE	**0.006**	**0.07902**	**0.001]**	**0.010**	**0.09931**	**0.000**	**0.023**	**0.15141**	**0.000**
		[0.004]	**[0.02801]**	**[0.015]**	**[0.007]**	**[0.03478]**	**[0.009]**	**[0.015]**	**[0.04930]**	**[0.008]**

Tablo-4'e göre; değişken seçiminde GEE yöntemi, tüm senaryolarda katsayıların hepsini sıfırdan farklı tahminlemiştir. Tüm senaryolar için IC=500 ve hassasiyeti 0'dır. Kesinlik değeri ise hesaplanamamıştır. Değişken sayısının artışıyla birlikte küçük kümelerde ve periyotta GEE değişken seçiminde yetersiz kalmıştır. PGEE, tüm senaryolar için C ve IC değerleri birbirine çok yakın hesaplamıştır. Değişen korelasyon ve küme sayısı PGEE'nin değişken seçim performansını etkilememiştir. Tüm senaryolar için kesinlik değerleri 0.50 olarak hesaplanmıştır. Hassasiyet değerleri 0.95 ve üzeri hesaplanmıştır. Genel olarak küme sayısındaki ve korelasyondaki artışlar PGEE'nin tahmin gücünü değiştirmediği gözlemlenmiştir (F_1 skor>0.60).

Tablo-4. 20 değişkenli PGEE ve GEE yöntemlerinin değişken seçim performans ölçümleri

	r=0.20				r=0.50				r=0.80			
P=20	**PGEE**		**GEE**		**PGEE**		**GEE**		**PGEE**		**GEE**	
	C	**IC**	**C**	**IC**	**C**	**IC**	**C**	**IC**	**C**	**IC**	**C**	**IC**
N=20	497.7	497.67	0.0	500.0	496.9	496.67	0.0	500.0	492.8	490.00	0.0	500.0
Kesinlik	0.50		-		0.50		-		0.50		-	
Hassasiyet	1.00		0.0		0.99		0.0		0.99		0.0	
F_1 skor	**0.67**				**0.66**				**0.66**			
N=30	496.1	496	0.0	500.0	496.3	495.33	0.0	500.0	495.0	495.0	0.0	500.0
Kesinlik	0.50		-		0.50		-		0.50		-	
Hassasiyet	0.99				0.99		0.0		0.99		0.0	
F_1 skor	**0.66**				**0.66**				**0.66**			
N=50	495.2	494.33	0.0	500.0	497.3	496.67	0.0	500.0	496.6	494.33	0.0	500.0
Kesinlik	0.50		-		0.50		-		0.50		-	
Hassasiyet	0.99				0.99		0.0		0.99		0.0	
F_1 skor	**0.66**				**0.66**				**0.66**			
N=100	466.1	461.33	0.0	500.0	487.1	483.67	0.0	500.0	498.5	498.33	0.0	500.0
Kesinlik	0.50		-		0.50		-		0.50		-	
Hassasiyet	0.93		0.0		0.97		0.0		1.00		0.0	
F_1 skor	**0.65**				**0,66**				**0.67**			
N=200	468.9	461	0.0	500.0	469.9	461.67	0.0	500.0	495.2	493.67	0.0	500.0
Kesinlik	0.50		-		0.50		-		0.50		-	
Hassasiyet	0.94		0.0		0.94		0.0		0.99		0.0	
F_1 skor	**0.65**				**0.65**				**0.66**			

3.3. Değişken Sayısı 50 iken PGEE ve GEE için Model ve Değişken Seçim Performansları

Tablo-5'e göre, tüm senaryolarda PGEE modelinden elde edilen MEDSE, MDAE ve BIAS değerleri 0'ın altında elde edilmiştir. PGEE modelinin düşük küme sayısı ve yüksek korelasyondan etkilenmediği görülmüştür. GEE 'de küçük küme sayısında değişken sayısının artmasına bağlı olarak yakınsama sorunu ortaya çıkmış yöntem çalışmadığı için hesaplama yapılamamıştır. N=200'e kadar olan tüm küme büyüklerinde GEE katsayılarına ait MEDSE ve BIAS değerleri 0'dan çok uzaklaşmışlardır. Her bir küme büyüklüğünde korelasyon artıkça MEDSE, MDAE ve BIAS değerleri de artmıştır. Küme sayısı 200'e çıktığında ise MEDSE değerleri düşüşler gözlemlenmiştir. Küme sayısının değişken sayısından küçük ya da eşit olduğu durumlarda GEE modelli tutarsız ve yanlı sonuçlar üretmiştir. Ancak Küme sayısının arttırılması ile model tahminleri bekleneni karşılamıştır. Küme sayısının düşük olduğu tüm korelasyon senaryolarında GEE modeli PGEE'nin gerisinde kalmıştır.

Tablo 6'daki bulgulara göre; GEE yöntemi, tüm senaryolarda katsayıların hepsini sıfırdan farklı tahminlemiştir. Tüm senaryolar için IC=500 ve hassasiyeti 0'dır. Kesinlik değeri ise hesaplanamamıştır. Bu bulgular GEE'nin değişken seçiminde yetersiz kaldığını göstermektedir. PGEE yönteminde ise tüm senaryolar için C ve IC değerleri birbirine çok yakındır. Yüksek C değerlerine karşılık yüksek IC değerleri PGEE'nin en cimri modeli oluştururken, sıfırdan farklı katsayıları da sıfır olarak tahminleme ve böylece gereksiz değişkenleri ayıklama eğiliminde olduğunu göstermektedir. Değişen korelasyon ve küme sayısı PGEE'nin değişken seçim performansını etkilememiştir. Tüm senaryolar için kesinlik değerleri 0.50 ve üstünde hesaplanmıştır. Hassasiyet değerleri 0.90 ve üzeri hesaplanmıştır. F1 skorları tüm senaryolarda 0,60 ve üstündedir.

Tablo 5. 50 değişkenli PGEE ve GEE modellerinin katsayılarına ait performans ölçütleri

P=50 Medyan [IQR]		r=0.20 MEDSE	r=0.20 MDAE	r=0.20 BIAS	r=0.50 MEDSE	r=0.50 MDAE	r=0.50 BIAS	r=0.80 MEDSE	r=0.80 MDAE	r=0.80 BIAS
N=20	PGEE	0.000 [0.000]	0.000001 [0.000004]	-0.097 [0.339]	0.000 [0.000]	0.0000004 [0.000001]	-0.128 [0.185]	0.000 [0.000]	0.0000004 [0.0000008]	-0.113 [0.208]
	GEE	-- --	-- --	-- --	-- --	-- --	--- --	-- --	-- --	-- --
N=30	PGEE	0.000 [0.000]	0.000003 [0.00001]	0.128 [0.307]	0.000 [0.000]	0.000001 [0.000004]	0.041 [0.273]	0.000 [0.000]	0.0000004 [0.000001]	-0.012 [0.124]
	GEE	**104.82 [128.67]**	**10.23608 [5.8800]**	**-0.813 [0.296]**	**166.64 [171.50]**	**12.90881 [6.23]**	**-0.817 [0.165]**	**415.03 [467.39]**	**20.37 [10.70]**	**-0.826 [0.107]**
N=50	PGEE	0.000 [0.000]	0.000004 [0.000006]	0.128 [0.262]	0.000 [0.000]	0.000001 [7.3325E-7]	-0.014 [0.083]	0.000 [0.000]	3.45E-07 [0.000001]	-0.002 [0.063]
	GEE	**30.16 [16.27]**	**5.48616 [1.4800]**	**-1.407 [0.482]**	**46.28 [25.54]**	**6.80 [1.84]**	**-1.411 [0.288]**	**118.52 [73.58]**	**10.88 [3.32]**	**-1.413 [0.265]**
N=100	PGEE	0.000 [0.000]	0.000001 [0.000003]	-0.030 [0.214]	0.000 [0.000]	4.34E-07 [0.0000004]	0.001 [0.034]	0.000 [0.000]	0.0000003 [0.0000003]	-0.002 [0.037]
	GEE	**17.74 [7.74]**	**4.20827 [0.91837]**	**-3.888 [1.723]**	**27.35 [12.82]**	**5.23 [1.23]**	**-3.958 [1.520]**	**69.79 [31.91]**	**8.36 [1.90494]**	**-3.991 [1.635]**
N=200	PGEE	0.000 [0.000]	0.000001 [0.000001]	-0.012 [0.030]	0.000 [0.000]	0.0000003 [0.0000004]	0.001 [0.027]	0.000 [0.000]	0.0000003 [0.0000004]	0.000 [0.024]
	GEE	**0.29 [796.92]**	**0.542537 [27.87735]**	**-0.218 [16.648]**	**0.45 [1684.91]**	**0.67 [40.62]**	**-0.210 [17.088]**	**0.86 [915.02]**	**0.93 [29.58]**	**-0.173 [8.842]**

Tablo 6. 50 değişkenli PGEE ve GEE yöntemlerinin değişken seçim performans ölçümleri

	r=0.20				r=0.50				r=0.80			
P=50	**PGEE**		**GEE**		**PGEE**		**GEE**		**PGEE**		**GEE**	
	C	**IC**	**C**	**IC**	**C**	**IC**	**C**	**IC**	**C**	**IC**	**C**	**IC**
N=20	478.0	415.00	0.0	500.0 4	68.2 3	66.25	0.0	500.0	452.1	415.00 0	.0 5	00.0
Kesinlik 0	.50		-		0.56 -		0		.56		-	
Hassasiyet 0	.96		0.0		0.94 0		.0 0		.90		0.0	
F_1 skor	**0.66**				**0.70**				**0.69**			
N=30	484.3	434.00	0.0	500.0 4	77.0 3	86 0	.0 5	00.0	488.0	388.5	0.0	500.0
Kesinlik 0	.53		-		0.55 -		0		.56		-	
Hassasiyet 0	.97		0.0		0.95 0		.0 0		.98		0.0	
F_1 skor	**0.69**				**0.70**				**0.71**			
N=50	478.1	424.75	0.0	500.0 4	91.5 3	73 0	.0 5	00.0	494.0	386.75 0	.0 5	00.0
Kesinlik 0	.53		-		0.57 -		0		.56		-	
Hassasiyet 0	.96		0.0		0.50 0		.0 0		.99		0.0	
F_1 skor	**0.69**				**0.53**				**0.72**			
N=100	452.9	331.5 0	.0	500.0 4	91.3 3	62.25	0.0	500.0	492.6	374.5	0.0	500.0
Kesinlik 0	.59		-		0.58 -		0		.57		-	
Hassasiyet 0	.91		0.0		0.98 0		.0 0		.99		0.0	
F_1 skor	**0.72**				**0.73**				**0.72**			
N=200	458.8	222.5 0	.0	500.0 4	91.8 3	48.5 0	.0 5	00.0	492.9	364.25 0	.0 5	00.0
Kesinlik 0	.67		-		0.59 -		0		.59		-	
Hassasiyet 0	.92		0.0		0.98 0		.0 0		.98		0.0	
F_1 skor	**0.78**				**0.74**				**0.74**			

4.TARTIŞMA

Açıklayıcı değişken sayısının fazla olduğu boylamsal çalışmalarda yeterli küme sayısı ile çalışmak tahminlerin tutarlığını önemli ölçüde etkilemektedir. Değişken sayısının örneklem sayısından daha büyük olduğu çok boyutlu veri setlerinde, katsayı tahminleri tutarsız ve yanlı olup; yanlış yorumları da beraberinde getirecektir. Bu durumlarda model ve değişken seçimi için cezalandırma yöntemleri oldukça başarılı sonuçlar vermiştir (18).

PGEE, p>n durumunda hem küme içindeki hem de değişkenler arasındaki ilişkileri kontrol altında tutarak daha tutarlı sonuçlar üretmiştir. PGEE'nin değişkenler arasında en azından orta düzey bir korelasyon olması durumunda iyi çalıştığını söylemek mümkündür. Değişken sayısı ile birlikte küme sayısının arttırılması ise PGEE'nin tahmin gücünü arttırmıştır. GEE, minimum periyot sayısından ve P>n durumundan olumsuz etkilenmiş, değişken ve model seçiminde yetersiz kalmıştır. PGEE, değişken seçiminden ziyade değişken ayıklama mantığıyla çalışmıştır. Doğrudan etkili değişkenleri seçmek yerine dolaylı olarak seçerek, modele etkisi olmayan değişkenleri tespit etmede daha başarılı olmuştur.

Bu durum tanı ve tarama testleri arasındaki farka benzetilmiştir. PGEE de tıpkı tarama testi gibi çalışmıştır. Pozitiflik durumu beta katsayılarının sıfır olması olarak belirtilen çalışmamızda PGEE, yüksek yanlış pozitif oranları üreterek tüm senaryolarda yüksek duyarlılık ya da hassasiyet değerlerine ulaşmıştır.

İkili yanıt değişkenin varlığında, küme sayısının en az 30, sürekli ya da kesikli yanıt değişkeni için de en az 50 olması önerilmektedir (19). Eğer, P>n varlığında bu durum sağlanamıyorsa, değişkenler arasında orta ya da yüksek korelasyonlu yapılar varsa; adımsal yöntemlerle klasik GEE uygulamak yerine etkili ve yansız sonuçlar veren cezalı GEE'nin kullanımını önermekteyiz.

KAYNAKLAR

[1] Park, P. (2021). Transition models for longitudinal data analysis with sparse hierarchical penalization. Yayımlanmış doktora tezi. McGill University, Montreal.
[2] Morel, J. G., Bokossa, M. C., Neerchal, N. K. (2003). Small sample correction for the variance of GEE estimators. Biometrical Journal. 45 (4), 395–409.
[3] Albert, A., Anderson, J. A., (1984). On the existence of maximum likelihood estimates in logistic regression models. Biometrika. 71 (1), 1-10.
[4] Heinze, G., Schemper, M. (2002). A solution to the problem of separation in logistic regression. Statistics In Medicine. 21 (16), 2409-2419.
[5] Lesaffre, E, Albert, A. (1989). Partial separation in logistic discriminati. Journal of the Royal Statistical Society. 51 (1), 109-116.
[9] Diggle, P.J., Heagerty, P., Liang, K.Y. (2002). Analysis of Longitudinal Data.Oxford: Oxford University Press [10]. Fitzmaurice, G., Davidian M., Verbeke, G., Molenberghs, G. (2009). Longitudinal Data Analysis. Longitudinal data analysis. New York: Chapman and Hall/CRC.
[11] Liu, X. (2016). Methods and Applications of Longitudinal Data Analysis.London: Elsevier.
[12] Ming, Wang. (2014). Generalized Estimating Equations in Longitudinal Data Analysis: A Review and Recent Developments, Advances in Statistics. 2014, https://doi.org/10.1155/2014/303728.
[13] Liang, K. Y., Zeger, S. L. (1986). Longitudinal data analysis using generalized linear models. Biometrika. 73 (1), 13-22.
[14] Kleinbaum, D.G., Klein, M. (2010) Logistic Regression. New York: Springer.
[15] Wang, L., Zhou, J., Qu, A. (2012). Penalized generalized estimating equations for High-Dimensional longitudinal data analysis. Biometrics. 68 (2), 353-360.
[16] Gül, İ., Wang, L. (2017). PGEE: An R package for analysis of longitudinal data with High-Dimensional covariates. The R Journal 9 (1), 393-402.
[17] Fan, J., Li, R. (2001). Variable selection via nonconcave penalized likelihood and its oracle properties. Journal of the American Statistical Association. 96(456), 1348-1360.
[18] Nadarajaha, T., Variyathb A.M., Loredo-Ostib J.C., (2021). Penalized Empirical Likelihood-Based Variable Selection for Longitudinal Data Analysis. Amerıcan Journal of Mathematıcal and Management Scıences, 40(3), https://doi.org/10.1080/01966324.2020.1837042
[19] McNeish D.M, Harring, J. R. (2017) Clustered data with small sample sizes: Comparing the performance of model-based and design-based approaches, Communications in Statistics- Simulation and Computation, 46(2), 855-869, DOI: 10.1080/03610918.2014.98364

Genom-boyu İlişki Çalışmalarında Poligenik Risk Skorunun Makine Öğrenimi ve Derin Öğrenme Yöntemleri ile Tahmin Edilmesi

[1]Ragıp Onur ÖZTORNACI, [2]Erdal COŞGUN, [3]Bahar TAŞDELEN, [3]Cemil ÇOLAK

[1]Koç Üniversitesi Translasyonel Araştırma Merkezi, Koç Universitesi, İstanbul, Türkiye
[2]Microsoft Research, 14820 NE 36th Street, Building 99, Redmond, WA, 98052, USA
[3]Mersin Üniversitesi Tıp Fakültesi Biyoistatistik ve Tıbbi Bilişim Anabilim Dalı, Mersin, Türkiye
[3]İnönü Üniversitesi Tıp Fakültesi Biyoistatistik ve Tıbbi Bilişim Anabilim Dalı, Malatya, Türkiye

Özet

Amaç: Son yıllarda genetik yatkınlığın değerlendirilmesinde Poligenik risk skorunun (PRS) kullanılması yaygın bir yaklaşımdır. PRS, aynı anda birden çok SNP kullanılarak, bir hastalık için genetik risk skoru sağlayan bir ölçüttür ve tipik bir hastalık skoru olarak hesaplanabilir. Bu çalışmada, klasik PRS hesaplama yöntemlerine alternatif olarak derin öğrenme (DL) ve makine öğrenmesi (ML) kullanılarak literatüre alternatif yöntemler katılması ve klasik yöntemin validasyonu amaçlanmıştır.

Gereç ve Yöntem: PLINK Programı ile Genom-boyu İlişki Çalışmaları (GWAS) verileri farklı allel frekanslarında, farklı örnek genişliklerinde yüz kez tekrar edilerek oluşturulmuştur. DL ve ML yöntemleri ile elde edilen yeni ağırlıklar ile PRS hesaplanmıştır. Elde edilen PRS sonuçlarının dağılımı üzerinde durularak, çevresel faktörlerle birlikte nasıl yorumlanacağı üzerine durulmuştur. Kullanılan programlar ve işletim sistemleri; R, Python ve Ubuntu Linux Operating System'dir.

Bulgular: Farklı örnek genişliklerinde ve farklı SNP sayılarında, ML ile hesaplanan PRS, klasik yöntem (PRS) ile hesaplanan PRS'e göre ortalamalar bakımından hasta-kontrol ayırımını yapmada daha tutarlı sonuçlar elde edebileceği görülmüştür. Benzer bir yaklaşımla SNP sayısını sabit tutulup örnek genişliğindeki artışın sırasında meydana gelen değişiklikler gözlemlenirse; klasik yöntem (PRS), destek vektör makineleri (SVM) ve DL için etkili olabileceği görülürken, rastgele orman (RF) yöntemi için değişim gözlemlenmemiştir. SNP sayısı sabitenken örnek genişliğindeki değişimden, klasik PRS ve DL etkilenmiştir.

Sonuç: Klasik yöntemle PRS hesaplanırken eğitim ve test seti olmak üzere iki farklı veri seti kullanılması çok sık kullanılmaktadır. Bu her zaman kolay olmamakla birlikte klasik yöntem için bir dezavantaj getirmektedir. ML ve DL yöntemleri tek bir veri seti kullanarak eğitim aşamasında yapılan çapraz doğrulama ile hem aşırı uyum sorunundan kaçınmaktadır hem de klasik yöntemin iki farklı veri seti kullanımı dezavantajına karşın iyi bir alternatif yöntem olarak kullanılabilir. Klasik PRS hesaplama yöntemine göre ML ve DL hem daha dar değişim aralığı ve daha dar güven aralıkları elde etmişlerdir. PRS hesaplamak için klasik yöntemlere alternatif olarak ML ve DL yöntemlerinin kullanılması, alternatif yöntemler olarak önerilmiştir.

Anahtar Kelimeler: Genom-boyu ilişki çalışmaları (GWAS), Poligenik risk skoru (PRS), Makine öğrenimi, derin öğrenme.

1. GİRİŞ

Genom-boyu ilişki analizi çalışmaları (Genome-wide Association Studies, GWAS) ile belirli bir hastalığa neden olan tek nükleotid polimorfizmlerinin (TNP) genom üzerindeki yerleri tespit edilerek, hastalığa yatkınlık risk oranları hesaplanabilir. Günümüzde birçok hastalığa neden olan TNP'ler belirlenmiştir. Dolasıyla artık, GWAS'ın ötesinde çalışmalara ihtiyaç duyulmaktadır. Bu çalışmalar, kişiselleştirilmiş tıp ve koruyucu tıp olarak beli bir hastalığın henüz ortaya çıkmadan önlenmesi adına önem teşkil etmektedir ve klinisyenlere oldukça önemli faydalar sağlamaktadırlar [1]. Bu çalışmalar arasında en yaygın olarak kullanılanlardan birisi de Poligenik Risk Skoru (PRS) hesaplama çalışmalarıdır. PRS, yaygın olarak kişiselleştirilmiş tıp alanında hastalığın daha ortaya çıkmadan, ilerlemeden önüne geçilmesi için oldukça kullanışlı, klinisyenlere ve topluma bilgi sağlayan bir yöntemdir [2]. PRS, aynı anda birden çok tek nükleotid polimorfizmi (TNP) kullanılarak, bir hastalık için genetik risk skoru sağlayan bir ölçüttür ve tipik bir hastalık skoru olarak hesaplanabilir. Bu çalışmanın amaçları, risk tahmini için en doğru modeli bulmak, klasik PRS hesaplama yönteminin yanı sıra makine öğrenimi (MÖ) ve derin öğrenim (DÖ) yöntemleri kullanılarak en uygun modeli belirlemek, GWAS ile hasta-sağlıklı ayırımı olarak belirlenmiştir.

GEREÇ VE YÖNTEMLER

Bu çalışmada PRS hesaplanmasında, en uygun modeli belirlemek amacıyla GWAS verisi formatında (bed, bim, fam uzantılı) 1000 genom projesi bünyesindeki gerçek verilere dayanılarak üretilen, 251 hasta ve 232 kontrol, 489805 TNP'den oluşan obezite ile ilişkili TNP'leri bulmak amacıyla oluşturulmuş açık erişimli veri setinde; klasik yöntem, makine öğrenimi algoritmaları ve derin öğrenme algoritması çalışmamızın amacı doğrultusunda karşılaştırılmıştır. Dolayısıyla, obezite eğilimi yüksek olan gruplar hasta grubu olarak tanımlanmıştır [3]. Yaptığımız simülasyonlarda, farklı odds oranlarına sahip hasta ve kontrol verileri farklı örneklem genişliklerinde üretilmişlerdir. Tüm

verilerde hasta gruplarında 0.01 oranında olacak şekilde hastalıkla ilişkili olan TNP'ler üretilmiştir. Hasta grubunun kontrol grubundan ayırt edilmesi amacıyla, hastalıkla ilişki olacak şekilde TNP'ler için (binary veriler kullanılacağından ötürü) lojistik regresyon analizi ile elde edilecek olan, p değerleri 0.05'ten küçük olacak şekilde üretilmiştir.

Poligenik Risk Skoru

Genetik olarak karmaşık hastalıkların, risk tahmininde bireysel mutasyonlara bakmaktan daha faydalıdır ve hesaplanırken, genom üzerindeki tüm hastalıkla ilişkili olan varyantların etki büyüklükleri değerlendirilerek, tüm varyantların toplamı olarak bulunur.

$$\mathrm{PRS}= \sum_{i=1}^{n} w_i * X_i \,, \tag{1}$$

Formülasyonu (1) ile hesaplanır. Burada, w_i i-nci TNP'e ait ağırlıkları temsil ederken, X_i ise ağırlıkların etki ettiği alelleri ifade eder [3].

Makine Öğrenimi Yöntemleri

Kullanılan verideki örüntüleri yakalayarak, öğrenen diğer bir ifade ile performansını geliştiren algoritma sistemlerine genel anlamda makine öğrenimi (MÖ) adı verilir [4]. Klasik istatistiksel yöntemlere göre varsayım gerektirmemesi, hızlı işlem yapabilme kabiliyeti ve problem türüne göre farklı şekilde ayar parametrelerine sahip olması nedeniyle, makine öğrenimi yöntemleri giderek cazip hale gelmektedir [5, 6]. Genetik epidemiyoloji verileri, yapıları itibariyle büyük verilerdir; dolasıyla, bu yapılar göz önünde bulundurulduğunda, kullanılan yöntemler de matematiksel olarak bu yapıları destekler nitelikte olmalıdırlar. MÖ ve DÖ yöntemleri tahmin gücü yüksek yöntemler olduğundan ötürü, büyük veri analizlerinde giderek başvurulan yöntemler haline gelmişlerdir [7].

Destek Vektör Makineleri

Vladimir Vapnik ilk olarak 1992 sınıflama problemleri için lagrange çarpanlarını kullanarak sınıfları arası maksimum ayrıma yeteneğine sahip minimum hata ile sınıflar arasına hayali bir çizgi çekmeyi hedefleyen algoritma geliştirmiştir. Eğitim süresi olarak diğer algoritmalarla kıyaslayınca uzun sürmesine rağmen, destek vektör makineleri aşırı uyum (overfitting) problemine karşı dirençlidir [8, 9].

Rastgele Orman (Random Forest)

En temel MÖ olan karar ağaçlarının bir araya getirilerek oluşturduğu hem sürekli yapıdaki değişkeneler için kullanılabilen hem de kategorik yapıdaki değişkenler için kullanılabilen RF algoritmasının, eksik verilerin tamamlanması, özellik seçimi (feature selection) ve bu çalışmanın kapsamında da kullanılacak olan modele katkı sunan değişkenlerin ağırlıklandırılması gibi birçok kullanım alanı vardır. Ağaçlar inşa edilirken Gini indexi veya Twoing kuralı uygulanır [10-13].

Derin Öğrenim Yöntemi

Makine öğreniminin alt dalı olan derin öğrenme, MÖ'den farklı olarak özellik seçimine dolasıyla da veri ön işleme sürecine gerek duymaz [14]. ML için özellik seçimi esnasında araştırmacı tarafından kullanılan korelasyon, ki-kare testi gibi veri ön işleme yöntemleri yerine, DÖ algoritmaları için gizli katmanlar (hidden layer), batch ve epoch sayıları gibi DÖ'ye ait kavramların kullanılır ve bu DÖ kavramları, DÖ algoritmalarının kendi kendine öğrenme sürecine etki ederek algoritmaların performanslarını değiştirdiği gözlemlenir [15-17]. Gizli katmanların sayısının arttırılması yanı sıra doğru şekilde ağırlıklandırma ile optimize edilerek de doğrusal olmayan sınıflama problemlerin çözüm gücü artırılabilir [17].

Poligenik Risk Skoru Tahmini İçin Yeni Geliştirilen Yöntemler

Destek vektör makineleri (DVM) için, ağırlık vektörü $SVMW_i$, DVM sınıflanma formüllerindeki w_i ağırlık vektörü değerleridir. RF yöntemi için ise RFI_i olarak tanımlanan yeni ağırlık vektörleri ise, I_i değişkenlerin tahmin denklemindeki önemlilik dereceleri (variable importance) ile hesaplanmıştır. DÖ yönteminde ise sınıflama öncesi elde edilen ağırlık matrisleri DLW_i ile ifade edilerek, her bir TNP ile çarpılarak, kişisel risk skorları elde edilmiştir. Bu durumda formüller sırasıyla,

$$PRS_{SVM} = \sum_{i=1}^{n} SVMW_i * SNP_i \tag{2}$$

$$PRS_{RF} = \sum_{i=1}^{n} RFI_i * SNP_i \tag{3}$$

$$PRS_{DL} = \sum_{i=1}^{n} DLW_i * SNP_i \tag{4}$$

Olarak hesaplanmıştır.

Simülasyon Adımları

Belirli bir Hastalık için GWAS verilerine ulaşmak oldukça maliyetli ve zordur. PLINK programında simülasyonlar ile üretilen (bed, bim, fam dosyaları) ham veriler ile yeni yöntemlerin prototipini oluşturmak, farklı yaklaşımların gücünü karşılaştırmak mümkündür. Simülasyon kriterleri ise; Hasta-Kontrol Sayıları: 500-500, 1000-1000, 2000-2000, TNP sayıları: 5.000, 10.000, 50.000, 100.000 olarak belirlenmiştir. Kullanılan Programlar ve İşletim Sistemleri; PLINK, R, Python, Linux, Microsoft Azure Platformu, STATISTICA 13'tür [18-23].

4. BULGULAR

Hasta-Kontrol grupları arasında, tüm TNP sayılarında ve tüm örnek genişliklerinde klasik yöntem ile hesaplanan PRS bakımından istatistiksel olarak anlamlı fark olduğu gözlemlenmiştir ($p<0.001$). Tüm örnek genişliklerinde hem kontrol hem de hasta grubunda en düşük sonuçlar, klasik yöntem PRS ile elde edilirken, en yüksek DÖ ile elde edildiği gözlemlenmiştir. Tüm TNP sayılarında ve tüm örnek genişliklerinde RF yönteminde daha homojen bir değişim aralığı elde edildiği gözlemlenmiştir. DVM için de benzer durum söz konusudur. Düşük TNP sayısında tüm yeni geliştirilen yöntemlerin klasik yönteme alternatif olarak kullanabilecekleri söylenebilir.

Tablo-1. Klasik Poligenik Risk Skoru (PRS) ve Diğer Modellere Göre Ağırlıklandırılmış Risk Skorlarının Hasta ve Kontrol Grubunda Karşılaştırmaları (N:500-500)

TNP Sayısı	Yöntem	500 Hasta- 500 Kontrol Poligenik Risk Skoru Tablosu		p Değeri
		Kontrol	Hasta	
		Ort. ±Std. Sapma	Ort. ±Std. Sapma	
		[Min:Maks]	[Min:Maks]	
5000 TNP	PRS	0,081±0,269	0,212±0,267	<0.001
		[-0,645:0,916]	[-0,712:0,863]	
	DVM	0,089±0,129	0,172±0,077	<0.001
		[-0,178:0,428]	[-0,125:0,389]	
	RF	0,078±0,025	0,083±0,019	<0.001
		[0,033:0,128]	[0,045:0,121]	
	DÖ	0,489±0,318	0,505±0,296	<0.001
		[-0,098:1,247]	[-0,044:1,234]	
10000 TNP	PRS	-0,513±0,451	-0,476±0,471	<0.001
		[-1,104:0,039]	[-1,119:0,079]	
	DVM	0,019±0,074	0,062±0,049	<0.001
		[-0,155:0,162]	[-0,075:0,156]	
	RF	0,039±0,012	0,042±0,009	<0.001
		[0,018:0,063]	[0,025:0,061]	
	DÖ	0,640±0,317	0,643±0,306	0.361
		[-0,016:1,441]	[0,012:1,443]	
50000 TNP	PRS	-0,001±0,113	0,130±0,163	<0.001
		[-0,234:0,354]	[-0,241:0,587]	
	DVM	0,053±0,096	0,137±0,063	<0.001
		[-0,103:0,252]	[-0,102:0,241]	
	RF	0,077±0,024	0,084±0,017	<0.001
		[0,036:0,120]	[0,050:0,119]	
	DÖ	0,616±0,351	0,618±0,348	0.642
		[-0,005:1,468]	[0,004:1,468]	
100000 TNP	PRS	0,286±0,358	0,452±0,266	<0.001
		[-0,246:0,963]	[0,045:0,986]	
	DVM	0,029±0,047	0,072±0,029	<0.001
		[-0,052:0,111]	[-0,028:0,111]	
	RF	0,038±0,013	0,042±0,008	<0.001
		[0,015:0,060]	[0,024:0,058]	
	DÖ	0,669±0,373	0,670±0,371	0.803
		[-0,004:1,339]	[0,002:1,351]	

PRS: Poligenik Risk Skoru, DVM: Destek Vektör Makineler, RF: Rastgele Orman, DÖ: Derin Öğrenme

En düşük örnek genişliği olan 500 hasta ve 500 kontrol grupları arasında 10.000 TNP için, klasik PRS hesaplama yöntemi DVM ve RF ile hesaplanan genetik risk skoru bakımından istatistiksel olarak anlamlı fark varken (p<0.001), DÖ yöntemi kullanıldığında istatistiksel olarak anlamlı fark bulunamamıştır (p=0.361). TNP sayısı 50.000 olarak belirlendiği zaman, klasik PRS hesaplama yöntemi DVM ve RF ile hesaplanan genetik risk skoru bakımından istatistiksel olarak anlamlı fark varken (p<0.001), DÖ yöntemi kullanıldığında istatistiksel olarak anlamlı fark yoktur (p=0.642). TNP sayısı 100.000 olarak belirlendiği zaman, klasik PRS hesaplama yöntemi DVM ve RF ile hesaplanan genetik risk skoru bakımından istatistiksel olarak anlamlı fark varken (p<0.001), DÖ yöntemi kullanıldığında istatistiksel olarak anlamlı fark bulunamamıştır (p=0.803). Dolayısıyla, TNP sayısındaki değişim yalnızca DÖ yöntemi için önemlidir. Düşük TNP sayısında DÖ yöntemi iyi sonuç verirken, TNP sayısı arttıkça PRS belirleme yeteneğinin azaldığını söylemek mümkündür. RF ve destek vektör makineleri tüm TNP sayılarında klasik hesaplama yöntemi yerine kullanılabilmektedirler ve TNP sayısı arttıkça bu yöntemlerin daha dar güven aralıklarına sahip olarak Hasta-kontrol ayrımını daha keskin şekilde yaptıklarını söylemek mümkündür (Tablo-1).

Tablo-2. Klasik Poligenik Risk Skoru (PRS) ve Diğer Modellere Göre Ağırlıklandırılmış Risk Skorlarının Hasta ve Kontrol Grubunda Karşılaştırmaları (N:1000-1000)

TNP Sayısı	Yöntem	1000 Hasta- 1000 Kontrol Poligenik Risk Skoru Tablosu		p Değeri
		Kontrol	Hasta	
		Ort. ±Std. Sapma	Ort. ±Std. Sapma	
		[Min:Maks]	[Min:Maks]	
5000 TNP	PRS	-0,003±0,112	0,066±0,147	<0.001
		[-0,400:0,463]	[-0,403:0,777]	
	DVM	0,195±0,151	0,282±0,071	<0.001
		[-0,139:0,540]	[-0,011:0,545]	
	RF	0,080±0,026	0,085±0,018	<0.001
		[0,027:0,136]	[0,044:0,128]	
	DÖ	0,673±0,320	0,689±0,295	<0.001
		[-0,043:1,432]	[0,022:1,457]	
10000 TNP	PRS	-0,550±0,498	-0,516±0,519	<0.001
		[-1,130:0,087]	[-1,127:0,132]	
	DVM	0,073±0,095	0,119±0,062	<0.001
		[-0,168:0,261]	[-0,069:0,243]	
	RF	0,038±0,012	0,042±0,009	<0.001
		[0,015:0,069]	[0,025:0,062]	
	DÖ	0,902±0,430	0,905±0,419	0.467
		[0,003:1,640]	[0,061:1,657]	
50000 TNP	PRS	-0,023±0,079	0,109±0,182	<0.001
		[-0,209:0,303]	[-0,164:0,815]	
	DVM	0,039±0,085	0,123±0,068	<0.001
		[-0,128:0,259]	[-0,076:0,245]	
	RF	0,077±0,024	0,084±0,017	<0.001
		[0,037:0,127]	[0,046:0,118]	
	DÖ	0,828±0,382	0,830±0,379	0.652
		[-0,002:1,551]	[0,008:1,544]	
100000 TNP	PRS	0,094±0,189	0,234±0,172	<0.001
		[-0,174:0,545]	[-0,079:0,552]	
	DVM	0,021±0,045	0,063±0,037	<0.001
		[-0,041:0,116]	[-0,043:0,113]	
	RF	0,039±0,013	0,043±0,008	<0.001
		[0,013:0,065]	[0,028:0,060]	
	DÖ	0,892±0,461	0,893±0,459	0.825
		[0,000:1,727]	[0,005:1,732]	

PRS: Poligenik Risk Skoru, DVM: Destek Vektör Makineler, RF: Rastgele Orman, DÖ: Derin Öğrenme

Örnek genişliği 2000 için, hasta ve 1000 kontrol grupları arasında; hasta ve kontrol grubunu birbirinden ayırt etme gücü en yüksek yöntem ise DVM olarak karşımıza çıkmaktadır. 5000 TNP için tüm yöntemler hasta-kontrol ayırımını yapabilmektedir. Ancak, 10.000 TNP için, klasik PRS, RF ve DVM hasta sağlıklı ayırımını ortalamalar bakımından istatistiksel olarak anlamlı düzeyde yapabilirken (p<0.001), DL yöntemi kullanıldığında istatistiksel olarak anlamlı fark bulunamamıştır (p=0.467). 50.000 TNP için, hasta-kontrol grupları arasında, klasik PRS hesaplama yöntemi DVM ve RF ile hesaplanan genetik risk skoru bakımından istatistiksel olarak anlamlı fark varken (p<0.001), DL yöntemi kullanıldığında istatistiksel olarak anlamlı fark bulunamamıştır (p=0.652). Benzer durum 100.000 TNP için, hasta-kontrol grupları arasında, klasik PRS hesaplama yöntemi DVM ve RF ile hesaplanan genetik risk skoru bakımından istatistiksel olarak anlamlı fark varken (p<0.001), DL yöntemi kullanıldığında istatistiksel olarak anlamlı fark bulunamamıştır (p=0.825) (Tablo-2).

Tablo-3. Klasik Poligenik Risk Skoru (PRS) ve Diğer Modellere Göre Ağırlıklandırılmış Risk Skorlarının Hasta ve Kontrol Grubunda Karşılaştırmaları (N:2000-2000)

TNP Sayısı	Yöntem	2000 Hasta- 2000 Kontrol Poligenik Risk Skoru Tablosu		p Değeri
		Kontrol	Hasta	
		Ort. ±Std. Sapma	Ort. ±Std. Sapma	
		[Min:Maks]	[Min:Maks]	
5000 TNP	PRS	-0,515±0,459	-0,485±0,472	<0.001
		[-1,135:0,029]	[-1,123:0,046]	
	DVM	0,254±0,248	0,354±0,182	<0.001
		[-0,384:0,757]	[-0,152:0,773]	
	RF	0,075±0,027	0,081±0,022	<0.001
		[0,018:0,132]	[0,030:0,129]	
	DÖ	1,214±0,526	1,224±0,501	0,028
		[-0,076:2,150]	[0,126:2,155]	
10000 TNP	PRS	-0,534±0,477	-0,504±0,495	<0.001
		[-1,086:0,048]	[-1,084:0,084]	
	DVM	0,072±0,125	0,117±0,096	<0.001
		[-0,205:0,342]	[-0,210:0,343]	
	RF	0,039±0,013	0,042±0,009	<0.001
		[0,014:0,068]	[0,023:0,065]	
	DÖ	1,506±0,681	1,507±0,667	0.843
		[-0,072:2,448]	[0,030:2,435]	
50000 TNP	PRS	-0,013±0,087	0,060±0,191	<0.001
		[-0,178:0,370]	[-0,177:0,639]	
	DVM	0,135±0,164	0,224±0,127	<0.001
		[-0,190:0,440]	[-0,090:0,444]	
	RF	0,077±0,023	0,084±0,017	<0.001
		[0,040:0,120]	[0,048:0,115]	
	DÖ	2,207±0,966	2,210±0,962	0.760
		[0,007:3,430]	[0,018:3,409]	
100000 TNP	PRS	1,372±1,247	1,393±1,251	<0.001
		[-0,008:2,547]	[-0,002:2,574]	
	DVM	0,019±0,042	0,111±0,040	<0.001
		[-0,095:0,117]	[-0,013:0,158]	
	RF	0,027±0,004	0,036±0,004	<0.001
		[0,018:0,044]	[0,020:0,044]	
	DÖ	1,456±0,484	1,460±0,483	0.382
		[0,004:1,940]	[0,010:1,927]	

PRS: Poligenik Risk Skoru, DVM: Destek Vektör Makineler, RF: Rastgele Orman, DÖ: Derin Öğrenme

En yüksek örnek genişliği olan 2000 için, en yüksek poligenik risk skorunu hem hasta hem de kontrol grupları için 5000 TNP için, DL yöntemi elde ederken onu DVM yöntemi [0,20:0,40] aralığında takip etmiştir. En düşük sonuçları ise klasik PRS yöntemi elde etmiştir. 10.000 TNP sayısında, hasta ve kontrol grubu ile ortalamalar bakımından birbirine en uzak sonucu DVM yöntemi elde etmiştir. Hasta-Kontrol grupları arasında, klasik PRS hesaplama yöntemi DVM ve RF ile hesaplanan genetik risk skoru bakımından istatistiksel olarak anlamlı fark varken ($p<0.001$), DÖ yöntemi kullanıldığında istatistiksel olarak anlamlı fark bulunamamıştır ($p=0.843$). TNP sayısı 50.000 olduğu zaman, klasik PRS ile elde edilen poligenik risk skorlarının diğer yöntemlere göre hem kontrol hem de hasta grubunda daha düşük sonuçlar verdiği görülmektedir. RF ve DVM yöntemleri birbirlerine yakın sonuçlar elde etmiştir. Hasta-Kontrol grupları arasında, klasik PRS hesaplama yöntemi DVM ve RF ile hesaplanan genetik risk skoru bakımından istatistiksel olarak anlamlı fark varken ($p<0.001$), DL yöntemi kullanıldığında istatistiksel olarak anlamlı fark bulunamamıştır ($p=0.760$). En yüksek TNP sayısı olan 100.000 TNP için, klasik yöntem ve ML yöntemleri birbirine yakın olarak kontrol gruplarında yaklaşık olarak [- 0,10:0,35] aralığında sonuçlar elde ederken; hasta gruplarında ise yaklaşık olarak [0,10:0,45] aralığında sonuçlar elde etmişlerdir. En yüksek poligenik risk skorlarının DÖ yöntemi ile elde ettiği görülmektedir. Hasta-Kontrol grupları arasında, klasik PRS hesaplama yöntemi DVM ve RF ile hesaplanan genetik risk skoru bakımından istatistiksel olarak anlamlı fark varken ($p<0.001$), DL yöntemi kullanıldığında istatistiksel olarak anlamlı fark bulunamamıştır ($p=0.382$) (Tablo-3).

Tablo-4. Gerçek Verileri Kullanarak Klasik Poligenik Risk Skoru (PRS) ve Diğer Modellere Göre Ağırlıklandırılmış Risk Skorlarının Hasta Grubu ve Sağlıklı Kontrol Grubu ile Karşılaştırmaları

Yöntem	251 Hasta- 232 Kontrol Poligenik Risk Skoru Tablosu		p Değeri
	489805 SNP		
	Kontrol	Hasta	
	Ort. ±Std. Sapma	Ort. ±Std. Sapma	
	[Min:Maks]	[Min:Maks]	
PRS	-0,214±0,049	0,216±0,051	<0.001
	[-0,289: -0,073]	[0,121:0,364]	
DVM	-0,073±0,046	0,077±0,043	<0.001
	[-0,101:0,026]	[-0,020:0,103]	
RF	0,481±0,154	0,495±0,137	0.290
	[0,005:0,550]	[0,005:0,553]	
DÖ	0,178±004	0,178±0,004	0.989
	[0,162:0,189]	[0,168:0,191]	

PRS: Poligenik Risk Skoru, DVM: Destek Vektör Makineler, RF: Rastgele Orman, DÖ: Derin Öğrenme

Hasta-Kontrol grupları arasında, klasik PRS hesaplama yöntemi ve SVM ile hesaplanan genetik risk skoru bakımından istatistiksel olarak anlamlı fark varken ($p<0.001$), DL yöntemi ($p=0.989$) ve RF yöntemi ($p=0.290$) kullanıldığında istatistiksel olarak anlamlı fark bulunamamıştır (Tablo 4). Hasta ve kontrol grupları için ayrı ayrı değerlendirme yapılacak olursa; kontrol grubunda, klasik PRS hesaplama yöntemi için -0,214 ortalama değeri simülasyonlar ile hesaplanan [- 0,246:0,963] değer aralığı içerisinde yer almaktadır. Benzer şekilde, DÖ yöntemi için 0,178 ortalama değeri [-0,004:1,339] aralığı içerisinde yer almaktadır. DVM yöntemi için ise simülasyon sonuçları gerçek verilerden elde edilen ortalama değerlere oldukça yakındır. Ancak, RF yöntemi için benzerlik söz konusu değildir. Hasta gruplarında ise durum; PRS, DVM ve DÖ yöntemi için, gerçek verilerden elde edilen ortalama değeri, simülasyon sonucu elde edilen [min:maks] değerleri kapsamaktadırlar. Dolasıyla, önceki yapılan benzer çalışmalarda da yalnızca kontrol grubu için risk skoru belirlenebileceği görüldüğünden ötürü, yaklaşık olarak örnek genişliğinin 500 olduğu durumda, PRS yöntemine alternatif en güçlü yöntem olarak DVM yöntemi kullanılabileceği söylenebilir [24]. Ancak, yalnızca kontrol veya yalızca hasta grubunun sahip olduğu genetik risk skorları, iki farklı popülasyon gibi düşünülerek belirlemek istenirse, diğer yöntemlerin de kullanılabileceği göz ardı edilmemelidir. Çünkü, poligenik risk skorunun yorumlanması bölümünde irdelendiği üzere, elde edilen ortalama sonuçlar hasta ve sağlıklı popülasyonları için birbirinden bağımsız ve farklı olacak şekilde normal dağılım eğrisine sahip olabilirler. Bu durumun en temel nedeni SNP'ler ile genetik varyasyonu tamamının açıklanmasının mümkün olamayacağı göz önünde bulundurulmalı ve çevre-gen etkileşiminin yanı sıra klinik bulguların göz ardı edilmemesi gerekmektedir. Simülasyonlardan farklı olarak RF yöntemini kullanarak hesaplanan poligenik risk skorlarının hasta ve kontrollerde ortalamalar bakımından en yüksek değerleri elde ettiği görülmektedir.

SONUÇ

PRS diagnostik testler ile karıştırılmamalıdır. PRS ile bireylerin hastalığa sahip olması veya sahip olacağına dair nitel bir bilgi elde edilemez. Ancak, PRS ile bireylerde ilerlemekte olan bir hastalığa karşı risk ölçümü yapılması mümkündür [25]. Kullanılacak olan yöntem ile elde edilecek ortalama risk skoru ve buna bağlı olarak elde edilecek skorların standart sapma değerlerinin küçük olmasıyla birlikte değişim aralığının (Range) da küçük olması PRS kullanımı için avantaj sağlayabilir, MÖ yöntemleri bu anlamda DÖ'den ve geleneksel yöntemden daha iyi sonuçlar verdiği söylenebilir. Klasik yöntemle PRS hesaplanırken eğitim ve test seti olmak üzere iki farklı veri seti kullanılması çok sık kullanılmaktadır. Bu her zaman kolay olmamakla birlikte klasik yöntem için bir dezavantaj getirmektedir [26]. ML ve DL yöntemleri tek bir veri seti kullanarak eğitim aşamasında yapılan çapraz doğrulama ile hem aşırı uyum sorunundan kaçınmaktadır hem de klasik yöntemin iki farklı veri seti kullanımı dezavantajına karşın iyi bir alternatif yöntem olarak kullanılabilir. Daha önce yapılan benzer nitelikteki çalışmalarda da sürekli verilerde GWAS özet istatistikleri kullanılarak, ML yöntemlerinin klasik poligenik risk skoru hesaplama yönteminden daha iyi sonuçlar elde ettiği gösterilmiştir [27-29]. Bu çalışmada ise GWAS özet istatistikleri yerine ham veriler ile (bed, bim, fam dosyaları) yapılan hasta-kontrol sınıflamaları esnasında elde edilen ağırlık vektörlerinin kullanılması literatüre yeni bir bakış açısı getirmiştir. Geliştirilen bu MÖ ve DÖ yöntemi ile elde edilen poligenik risk skorlarının klasik yönteme göre bir tek ham veri kullanımı kısıtlaması vardır. Tip-II diyabet ile ilgili yapılan gen-gen ve gen- çevre etkileşimini inceleyen benzeri bir çalışmada, PRS ile klinik risk skorunun yüksek oranda benzeştiği görülmüştür. Benzer mantıkla yapılan diğer bir çalışmada ise klinik risk teşkil eden değişkenler ile PRS kullanılmıştır [30]. PRS yalnızca tek bir yöntem ile elde edilmesi yerine, farklı yöntemlerin de farklı hastalıklarda kullanılması açısından bu çalışma önem arz etmektedir.

KAYNAKLAR

[1] Dorak, M. T. (2016). Genetic association studies: background, conduct, analysis, interpretation. Garland Science.
[2] Konuma, T., & Okada, Y. (2021). Statistical genetics and polygenic risk score for precision medicine. *Inflammation and Regeneration, 41*(1), 1-5.
[3] Choi, S. W., Mak, T. S. H., & O'Reilly, P. F. (2020). Tutorial: a guide to performing polygenic risk score analyses. *Nature Protocols, 15*(9), 2759-2772.)
[4] E. Alpaydın, Introduction to Machine Learning, The MIT Press, 2004
[5] Akpınar H., "Data Veri Madenciliği Veri Analizi", 1. Baskı, Papatya Yayıncılık, İstanbul, 2013 , ISBN 978-605-4220- 81-6
[6] Gönen, M., & Alpaydın, E. (2011). Multiple kernel learning algorithms. The Journal of Machine Learning Research, 12, 2211-2268.
[7] Köse, T., Özgür, S., Coşgun, E., Keskinoğlu, A., & Keskinoğlu, P. (2020). Effect of missing data imputation on deep learning prediction performance for vesicoureteral reflux and recurrent urinary tract infection clinical study. BioMed Research International, 2020.
[8] Jiawei H., Kamber M., Han J., Kamber M., Pei J., "Data Mining: Concepts and Techniques", San Francisco, 2012 ,ISBN 978-0-12-381479-1
[9] Pisner, D. A., & Schnyer, D. M. (2020). Support vector machine. In *Machine learning* (pp. 101-121). Academic Press.
[10] Temel G. Ö., "Sınıflama ve Regresyon Ağaçları", Yüksek Lisans Tezi, Mersin Üniversitesi Sağlık Bilimleri Enstitüsü, Mersin, 2004
[11] Temel G. Ö., Çamdeviren H., Akkuş Z., "Sınıflama Ağaçları Yardımıyla Restless Legs Syndrome (RLS) Hastalarına Tanı Koyma", İnönü Üniversitesi Tıp Fakültesi Dergisi
[12] Strobl, C., & Zeileis, A. (2008). Danger: High power!–exploring the statistical properties of a test for random forest variable importance.
[13] Liu, Y., & Zhao, H. (2017). Variable importance-weighted random forests. Quantitative Biology, 5(4), 338-351.
[14] Aminanto, E., & Kim, K. (2016). Deep learning in intrusion detection system: An overview. In 2016 International Research Conference on Engineering and Technology (2016 IRCET). Higher Education Forum
[15] Min, S., Lee, B., & Yoon, S. (2017). Deep learning in bioinformatics. Briefings in bioinformatics, 18(5), 851-869. [16] Schmidhuber, J. (2015). Deep learning in neural networks: An overview. Neural networks, 61, 85-117
[17] Vidal, R., Bruna, J., Giryes, R., & Soatto, S. (2017). Mathematics of deep learning. arXiv preprint arXiv:1712.04741.
[18] Sarumathi, S., Shanthi, N., Vidhya, S., & Ranjetha, P. (2015). Statistica software: a state of the art review. *International Journal of Computer and Information Engineering, 9*(2), 473-480.
[19] Team, R. C. (2000). R language definition. *Vienna, Austria: R foundation for statistical computing*.
[20] Copeland, M., Soh, J., Puca, A., Manning, M., & Gollob, D. (2015). Microsoft azure. *New York, NY, USA: Apress,* 3- 26.
[21] Python, W. (2021). Python. *Python Releases for Windows, 24.*
[22] Purcell, S., Neale, B., Todd-Brown, K., Thomas, L., Ferreira, M. A., Bender, D., ... & Sham, P. C. (2007). PLINK: a tool set for whole-genome association and population-based linkage analyses. *The American journal of human genetics, 81*(3), 559-575.
[23] Кофлер, M. (2014). Linux.
[24] Black, M. H., Li, S., LaDuca, H., Lo, M. T., Chen, J., Hoiness, R., ... & Xu, J. (2020). Validation of a prostate cancer polygenic risk score. The Prostate, 80(15), 1314-1321.
[25] Mamani, N. M. (2020). Machine Learning techniques and Polygenic Risk Score application to prediction genetic diseases. ADCAIJ: Advances in Distributed Computing and Artificial Intelligence Journal, 9(1), 5-14.
[26] Paré, G., Mao, S., & Deng, W. Q. (2017). A machine-learning heuristic to improve gene score prediction of polygenic traits. Scientific reports, 7(1), 1-11.
[27] Polygenic risk scores outperform machine learning methods in predicting coronary artery disease status
[28] Badré, A., Zhang, L., Muchero, W., Reynolds, J. C., & Pan, C. (2021). Deep neural network improves the estimation of polygenic risk scores for breast cancer. Journal of Human Genetics, 66(4), 359-369.
[29] He, Y., Lakhani, C. M., Rasooly, D., Manrai, A. K., Tzoulaki, I., & Patel, C. J. (2021). Comparisons of polyexposure, polygenic, and clinical risk scores in risk prediction of type 2 diabetes. Diabetes care, 44(4), 935-943.
[30] Li, J. H., Szczerbinski, L., Dawed, A. Y., Kaur, V., Todd, J. N., Pearson, E. R., & Florez, J. C. (2021). A Polygenic Score for Type 2 Diabetes Risk Is Associated With Both the Acute and Sustained Response to Sulfonylureas. Diabetes, 70(1), 293-300.

Log Multinomial Regresyon Modeline Hosmer-Lemeshow Uyum İyiliği Testinin Uyarlanması

[1]Yasemin ÖZTÜRK, [2]Erdem KARABULUT, [3]Nimet Anıl DOLGUN

[1]T.C. Sosyal Güvenlik Kurumu, Genel Sağlık Sigortası Genel Müdürlüğü, Ankara
[2] Hacettepe Üniversitesi Tıp Fakültesi, Biyoistatistik Anabilim Dalı, Ankara
[3] CSL, Melbourne, Avustralya

Amaç: İki durumlu veya multinomial lojistik regresyon modelinin uyumunun değerlendirilmesinde Hosmer ve Lemeshow tarafından önerilen uyum iyiliği testi kullanılmaktadır. İleriye yönelik çalışmalarda göreli risk kestirimi elde edilebildiğinden, iki kategorili bağımlı değişken olduğunda log-binomial ve 3 ya da daha fazla kategorili bağımlı değişken olduğunda log-multinomial regresyon modelinin kullanılması önerilmektedir. Ancak, log-multinomial regresyon modelinin veriye uyumunun değerlendirilmesine yönelik herhangi bir metot bulunmamaktadır. Bu çalışmada Hosmer Lemeshow uyum iyiliği testi log multinomial regresyon modeline uyarlanmış ve benzetim çalışması ile performansı değerlendirilmiştir.

Yöntem: Uyarlanan Hosmer Lemeshow uyum iyiliği testinin uygulanabilirliği çeşitli senaryolar altında benzetim çalışması ile incelenmiştir. Söz konusu senaryolarda bağımlı ve bağımsız değişkenlerin farklı yapıları, çeşitli örneklem genişliklerinde benzetim çalışması yardımıyla karşılaştırılmıştır. Benzetim çalışmasında, ikiden fazla kategorili bağımlı kategorik değişkenler, sürekli ya da kategorik bağımsız değişkenler ve farklı örneklem genişlikleri (200, 400, 600, 800 ve 1000) dikkate alınarak toplamda 70 farklı senaryo kullanılmıştır. Benzetim senaryolarındaki modellerde kullanılan regresyon katsayıları, kullanılan gerçek veri setlerinde yer alan değişkenler yardımı ile kestirilen log-multinomial modellerden elde edilmiştir. Her bir senaryo 1000 kez tekrar edilmiş ve her bir tekrarda uyum iyiliği testi sonuçları tip I ve tip II hata sıklıkları elde edilerek değerlendirilmiştir. Belirlenen senaryolar için benzetim çalışması, modelleme ve uyum iyiliği testinin sonuçlarının incelenmesi R yazılımında yer alan "RStata", "general hoslem" ve "MASS" paketleri ve STATA yazılımı kullanılarak gerçekleştirilmiştir.

BULGULAR

Senaryo 1: Tip I Hata Performansının Değerlendirilmesi

$$\pi_j(x)=\exp(x'\beta_j) \quad (1)$$

Uyarlanan Hosmer-Lemeshow uyum iyiliği testinin tip I hata açısından değerlendirilmesi için ": Kurulan log multinomial model veriye uyumludur" yokluk hipotezi altında bağımlı değişkenin ikiden fazla durumlu olduğu, aşağıdaki senaryolarda türetilen verilere log multinomial regresyon modeli uygulanmıştır.

Senaryo 1.a. Kategorik bağımlı değişkenin 3 durumlu olduğu, bir sürekli sayısal (x) ve bir iki durumlu kategorik bağımsız değişkenin (d) olduğu model:

(1) numaralı eşitlik ile verilen modelde, x sürekli sayısal bağımsız değişkene ilişkin parametreler N(52,16) olan normal dağılımdan türetilmiştir. d ise iki durumlu kategorik bağımsız değişkeni ifade etmektedir. İki durumlu kategorik bağımsız değişken ise parametreleri Bernoulli(0.5) olan Bernoulli dağılımından türetilmiştir. Farklı örneklem genişliklerinde söz konusu modelin veriye uyumunun test edildiği uyum iyiliği testlerinin kaç tanesinin 1000 tekrarda yanlışlıkla reddedildiği sonucu Tablo 1'de sunulmaktadır.

Tablo-1. Senaryo.1.a. Hosmer-Lemeshow uyum iyiliği testi simüle edilmiş Ho hipotezi reddedilme oranları

Örneklem Büyüklüğü	Anlamlılık Düzeyi α=0,01 iken Tip I Hata Oranları (Ho Red Oranları)	Anlamlılık Düzeyi α=0,05 iken Tip I Hata Oranları (Ho Red Oranları)
200	0,021	0,060
400	0,011	0,056
600	0,007	0,035
800	0,008	0,039
1000	0,005	0,034

Tablo-1'e göre, 3 kategorili kategorik bağımlı değişkene sahip log multinomial regresyon modelin veriye uyumlu olduğu hipotezi altında kurulan modelin simülasyon sonuçlarına göre, anlamlılık düzeyi α=0,01 iken değerlendirilen her bir örneklem büyüklüğünde tip I hata oranları 0,005 ile 0,021 arasındadır. Anlamlılık düzeyi α=0,05 iken ise tip I hata oranlarının 0,034 ile 0,06 arasında değişim gösterdiği görülmektedir. Her iki anlamlılık düzeyinde de örneklem büyüklüğü arttıkça tip I hata oranının düştüğü ve çoğunlukla nominal düzeyin altında olduğu sonucu elde edilmiştir.

Senaryo 1.b. Kategorik bağımlı değişken 4 durumlu olduğu, bir sürekli sayısal (x) ve bir iki durumlu bağımsız değişkenin (d) olduğu model:

(1) numaralı eşitlik ile verilen modelde, x sürekli sayısal bağımsız değişkene ilişkin parametreleri N(23.23,5.29) olan normal dağılımdan türetilmiştir. d ise iki durumlu kategorik bağımsız değişkeni ifade etmektedir. İki durumlu kategorik bağımsız değişkeni ise parametreleri Bernoulli(0.5) olan Bernoulli dağılımından türetilmiştir. Farklı örneklem genişliklerinde söz konusu modelin veriye uyumunun test edildiği uyum iyiliği testlerinin kaç tanesinin 1000 tekrarda yanlışlıkla reddedildiği sonucu Tablo-2'de sunulmaktadır.

Tablo-2. Senaryo 1.b. Hosmer-Lemeshow uyum iyiliği testi simüle edilmiş Ho hipotezi reddedilme oranları

Örneklem Büyüklüğü	Anlamlılık Düzeyi α=0,01 iken Tip I Hata Oranları (Ho Red Oranları)	Anlamlılık Düzeyi α=0,05 iken Tip I Hata Oranları (Ho Red Oranları)
200	0,009	0,047
400	0,009	0,032
600	0,005	0,047
800	0,006	0,04
1000	0,003	0,041

Tablo-2'ye göre, 4 kategorili kategorik bağımlı değişkene sahip log multinomial regresyon modelin veriye uyumlu olduğu hipotezi altında kurulan modelin simülasyon sonuçlarına göre, anlamlılık düzeyi α=0,01 iken değerlendirilen her bir örneklem büyüklüğünde tip I hata oranları 0,003 ile 0,009 arasındadır. Anlamlılık düzeyi α=0,05 iken ise tip I hata oranlarının 0,041 ile 0,047 arasında değişim gösterdiği görülmektedir. Her iki anlamlılık düzeyinde de örneklem büyüklüğü arttıkça tip I hata oranının düştüğü ve çoğunlukla nominal düzeyin altında olduğu sonucu elde edilmiştir.

Senaryo 1.a. ve 1.b çerçevesinde, uyarlanan Hosmer-Lemeshow uyum iyiliği testinin hipotezinin doğru olduğu varsayımı altında, log multinomial modelin veriye uyumlu olup olmadığının değerlendirildiği benzetim sonuçlarına göre iyi performans gösterdiği ifade edilebilir.

Senaryo 2: Tip II Hata Performansının Değerlendirilmesi

Uyarlanan Hosmer-Lemeshow uyum iyiliği testinin tip II hata açısından değerlendirilmesi için ": Kurulan log multinomial model veriye uyumlu değildir" alternatif hipotezi altında bağımlı değişkenin ikiden fazla durumlu olduğu, farklı senaryolarda türetilen verilere log multinomial regresyon modeli uygulanmıştır.

Senaryo 2.a. Kategorik bağımlı değişken 3 durumlu, bir sürekli sayısal (x), bir iki durumlu bağımsız değişkenin (d), sürekli ve iki durumlu bağımsız değişkenler arasında etkileşimin bulunduğu model:

(1) numaralı eşitlik ile verilen modelde, x sürekli sayısal bağımsız değişkene ilişkin parametreler N(23,5) olan normal dağılımdan türetilmiştir. d ise iki durumlu kategorik bağımsız değişken ifade etmektedir. İki durumlu kategorik bağımsız değişken ise parametreleri Bernoulli(0.5) olan Bernoulli dağılımdan türetilmiştir. Sürekli sayısal değişken ile iki durumlu değişken arasındaki etkileşim terimi modele dahil edilmiştir. Farklı örneklem genişliklerinde söz konusu modelin veriye uyumunun test edildiği uyum iyiliği testlerinin kaç tanesinin 1000 tekrarda reddedildiği sonucu Tablo-4'te sunulmaktadır.

Tablo-4. Senaryo 2.a. Hosmer-Lemeshow uyum iyiliği testi simüle edilmiş Ho hipotezi reddedilme oranları (Güç Değerleri)

Örneklem Büyüklüğü	Anlamlılık Düzeyi α=0,01 iken İstatistiksel Güç (%)	Anlamlılık Düzeyi α=0,05 iken İstatistiksel Güç (%)
200	7,3	18,6
400	24,8	46,0
600	44,7	68,7
800	63,7	81,2
1000	80,9	92,2

Tablo 4'e göre, 3 kategorili kategorik bağımlı değişkene sahip log multinomial regresyon modelin veriye uyumlu olmadığı hipotezi altında kurulan modelin simülasyon sonuçlarına göre, anlamlılık düzeyi α=0,01 iken değerlendirilen her bir örneklem büyüklüğünde tip II hata oranları %19,1 ile %92,7 arasındadır. Anlamlılık düzeyi α=0,05 iken ise tip II hata oranlarının %7,8 ile %81,4 arasında değişim gösterdiği görülmektedir. Her iki anlamlılık düzeyinde de örneklem büyüklüğü arttıkça tip II hata oranının düştüğü sonucu elde edilmiştir.

Senaryo 2.b. Kategorik bağımlı değişken 3 durumlu, bir sürekli sayısal (x), bir iki durumlu bağımsız değişkenin (d), sürekli ve iki durumlu bağımsız değişkenler arasında etkileşim terimi ve sürekli değişkene ilişkin karesel terimin olduğu model:

(1) numaralı eşitlik ile verilen modelde, x sürekli sayısal bağımsız değişkene ilişkin parametreler N(23,5) olan normal dağılımdan türetilmiştir. d ise iki durumlu kategorik bağımsız değişkeni ifade etmektedir. İki durumlu kategorik bağımsız değişken ise parametreleri Bernoulli (0.5) olan Bernoulli dağılımdan türetilmiştir. Sürekli sayısal değişken ile iki durumlu değişken arasındaki etkileşim terimi ile ikinci dereceden sürekli sayısal değişkeni ifade eden karesel terim modele dahil edilmiştir. Farklı örneklem genişliklerinde söz konusu modelin veriye uyumunun test edildiği uyum iyiliği testlerinin kaç tanesinin 1000 tekrarda reddedildiği sonucu Tablo 5'te sunulmaktadır.

Tablo-5. Senaryo 2.b. Hosmer-Lemeshow uyum iyiliği testi simüle edilmiş Ho hipotezi reddedilme oranları (Güç Değerleri)

Örneklem Büyüklüğü	Anlamlılık Düzeyi α=0,01 iken İstatistiksel Güç (%)	Anlamlılık Düzeyi α=0,05 iken İstatistiksel Güç (%)
200	17,7	33,7
400	62,9	79,9
600	86,6	94,2
800	96,2	98,7
1000	98,0	99,1

Tablo-5'e göre, 3 kategorili kategorik bağımlı değişkene sahip log multinomial regresyon modelin veriye uyumlu olmadığı hipotezi altında kurulan modelin simülasyon sonuçlarına göre, anlamlılık düzeyi α=0,01 iken değerlendirilen her bir örneklem büyüklüğünde tip II hata oranları %2 ile %82,3 arasındadır. Anlamlılık düzeyi α=0,05 iken ise tip II hata oranlarının %0,9 ile %66,3 arasında değişim gösterdiği görülmektedir. Her iki anlamlılık düzeyinde de örneklem büyüklüğü arttıkça tip II hata oranının düştüğü sonucu elde edilmiştir.

Benzer şekilde, 4 kategorili kategorik bağımlı değişkene sahip log multinomial regresyon modelin veriye uyumlu olmadığı hipotezi altında kurulan modellerin simülasyon sonuçlarına göre, her iki anlamlılık düzeyinde örneklem büyüklüğü arttıkça tip II hata oranının düştüğü sonuçları elde edilmiştir.

Sonuç: Uyarlanan Hosmer-Lemeshow uyum iyiliği testinin, sürekli sayısal bağımsız değişkene ilişkin karesel terim ve bağımsız değişkenler arasındaki etkileşim teriminin dahil edilmemesi gibi çeşitli uyumsuzlukları belirleyebildiği görülmektedir. Tüm senaryolarda örneklem büyüklüğü arttıkça tip I ve tip II hata oranının düştüğü ve çoğunlukla nominal düzeyin altında olduğu sonucu elde edilmiştir. Hosmer-Lemeshow uyum iyiliği testinin log multinomial regresyon modelinin uyum iyiliğinin tespit edilmesinde kullanılabilir yapıda olduğu görülmektedir.

Anahtar Sözcükler: Log multinomial regresyon, Hosmer Lemeshow test istatistiği, uyum iyiliği.

KAYNAKLAR

1. Blizzard, L., and D. W. Hosmer. The Log Multinomial Regression Model for Nominal Outcomes with More than Two Attributes. Biometrical Journal;2007; 49; 889-902.
2. Hilbe, J., M., Logistic Regression Models, CRC Press; 2009, 109-110.
3. Agresti, A. An introduction to categorical data analysis.; 2018, John Wiley & Sons.
4. Saraçbaşı O., Dolgun A., Lojistik Regresyon Çözümlemesi, Hacettepe Üniversitesi Yayınları; 2015.
5. Qiu, Y., Liu, L., Lai, X., & Qiu, Y. An Online Test for Goodness-of-Fit in Logistic Regression Model. IEEE Access; 2019; 7, 107179-107187.
6. Hosmer Jr, D. W., Lemeshow, S., & Sturdivant, R. X. Applied logistic regression Vol. 398. John Wiley & Sons.;2013;154.
7. Hosmer, D. W., Hosmer, T., Le Cessie, S., & Lemeshow, S. A comparison of goodness- of- fit tests for the logistic regression model. Statistics in medicine;1997; 16(9), 965-980.
8. Paul, P., Pennell, M. L., & Lemeshow, S. Standardizing the power of the Hosmer–Lemeshow goodness of fit test in large data sets. Statistics in medicine; 2013; 32(1), 67-80.
9. Fagerland, M. W., Hosmer, D. W., & Bofin, A. M. Multinomial goodness- of- fit tests for logistic regression models. Statistics in medicine; 2008; 27(21), 4238-4253.
10. Fagerland, M. W., & Hosmer, D. W. A goodness- of- fit test for the proportional odds regression model. Statistics in medicine; 2013; 32(13), 2235-2249.
11. Yu, W., Xu, W., & Zhu, L. A modified Hosmer–Lemeshow test for large data sets. Communications in Statistics-Theory and Methods; 2017; 46(23), 11813-11825.

Index

Z


ANKARA ÜNİVERSİTESİ
1946
1943
A.Ü.
TIP
FAKÜLTESİ
ANKARA
BİYOİSTATİSTİK
DERNEĞİ
1997
μ