



**ÇEŞİTLİ ÇEKİRDEK FONKSİYONLARI İLE
OLUŞTURULAN DESTEK VEKTÖR MAKİNESİ
MODELLERİNİN PERFORMANSLARININ İNCELENMESİ:
BİR KLİNİK UYGULAMA**

Emek GÜLDOĞAN
İnönü Üniversitesi ve Mersin Üniversitesi
Biyoistatistik ve Tıp Bilişimi Anabilim Dalı
Ortak Doktora Programı

Tez Danışmanı: Prof. Dr. Cemil ÇOLAK
Ortak Tez Danışmanı: Prof. Dr. Bahar TAŞDELEN

Doktora Tezi - 2017

**T.C.
İNÖNÜ ÜNİVERSİTESİ
SAĞLIK BİLİMLERİ ENSTİTÜSÜ**

**ÇEŞİTLİ ÇEKİRDEK FONKSİYONLARI İLE OLUŞTURULAN DESTEK
VEKTÖR MAKİNESİ MODELLERİNİN PERFORMANSLARININ
İNCELENMESİ: BİR KLİNİK UYGULAMA**

Emek GÜLDOĞAN

**İnönü Üniversitesi ve Mersin Üniversitesi
Biyoistatistik Ve Tıp Bilişimi Anabilim Dalı
Ortak Doktora Tezi**

**Tez Danışmanı
Prof. Dr. Cemil ÇOLAK**

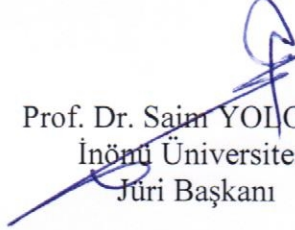
**Ortak Tez Danışmanı
Prof. Dr. Bahar TAŞDELEN**


**MALATYA
2017**

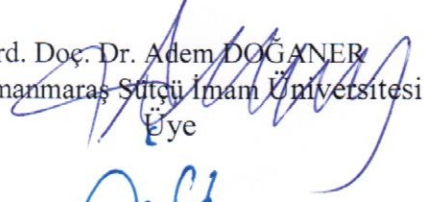
KABUL VE ONAY SAYFASI

İnönü Üniversitesi ile Mersin Üniversitesi Sağlık Bilimleri Enstitüleri Biyoistatistik ve Tıp Bilişimi Anabilim Dalı Ortak Doktora Programı çerçevesinde yürütülmüş olan; **Emek GÜLDOĞAN'ın** “**Çeşitli Çekirdek Fonksiyonları ile Oluşturulan Destek Vektör Makinesi Modellerinin Performanslarının İncelenmesi: Bir Klinik Uygulama**” konulu bu çalışması, aşağıdaki jüri tarafından Doktora tezi olarak kabul edilmiştir.

Tez Savunma Tarihi: 23/05/2017


Prof. Dr. Saim YOLOĞLU
İnönü Üniversitesi
Jüri Başkanı


Prof. Dr. Cemil ÇOLAK
İnönü Üniversitesi
Tez Danışmanı
Üye


Yrd. Doç. Dr. Adem DOĞANER
Kahramanmaraş Sütcü İmam Üniversitesi
Üye


Yrd. Doç. Dr. Harika G. GÖZÜKARA BAĞ
İnönü Üniversitesi
Üye


Yrd. Doç. Dr. Hüsamettin KAYA
Fırat Üniversitesi
Üye

ONAY

Bu tez, İnönü Üniversitesi Lisansüstü Eğitim-Öğretim Yönetmeliği'nin ilgili maddeleri uyarınca yukarıdaki jüri üyeleri tarafından kabul edilmiş ve Enstitü Yönetim Kurulu'nun/...../2017 tarih ve 2017/..... sayılı Kararıyla da uygun görülmüştür.

Prof. Dr. Yusuf TÜRKÖZ
Enstitü Müdürü

İÇİNDEKİLER

ÖZET	vi
ABSTRACT.....	vii
SİMGELER VE KISALTMALAR DİZİNİ.....	viii
ŞEKİLLER DİZİNİ.....	ix
TABLOLAR DİZİNİ.....	xi
1. GİRİŞ.....	1
2. GENEL BİLGİLER.....	3
2.1. Veri Tabanlarında Bilgi Keşfi Süreci.....	3
2.1.1. Veri Seçimi.....	3
2.1.2. Veri Önışleme.....	3
2.1.3. Veri Dönüşümü ve Veri İndirgeme.....	4
2.1.4. Veri Madenciliği.....	4
2.1.5. Deęerlendirme ve Yorumlama.....	9
2.2. Destek Vektör Makinesi.....	13
2.2.1. Doğrusal DVM.....	13
2.2.2. Doğrusal Olmayan DVM.....	14
2.2.3. Çekirdek Fonksiyonları.....	15
2.3 Araştırma Konusu ile İlgili Benzer Çalışmalar.....	18
3. MATERYAL VE METOT.....	22
3.1. Örneklem Büyüklüğü.....	23
3.2. Geliştirilen Web Tabanlı Yazılım.....	26
3.3. Oluşturulan Destek Vektör Makinesi Modelleri.....	34
3.4. Biyoistatistiksel Analiz.....	39
4. BULGULAR.....	41
5. TARTIŞMA.....	76
6. SONUÇ VE ÖNERİLER.....	80

KAYNAKLAR	81
EKLER	87
EK-1. Özgeçmiş	87
EK-2. Etik Kurul Onay Formu	88



TEŐEKKÜR

Tez alıőmam sűresince bilgi, birikim ve deneyimleri ile bana yol gűsterici ve destek olan deęerli birinci danıőman hocam Sayın Prof. Dr. Cemil OLAK'a, deęerli ikinci danıőman hocam Sayın Prof. Dr. Bahar TAŐDELEN'e, eęitimim boyunca desteęini esirgemeyen ve űnerileriyle bana ıőık tutan Sayın Prof. Dr. Saim YOLOęLU'na teőekkűr ve saygılarımı sunarım. alıőmalarım boyunca yardımını hi esirgemeyen deęerli arkadaőım Arő. Gűr. Ahmet Kadir ARSLAN'a teőekkűrű bir bor bilirim. Akademik eęitimim boyunca destek ve katkılarıyla beni hibir zaman yalnız bırakmayan eőime ve aileme sonsuz teőekkűr ederim.

Okt. Emek GŪLDOęAN

ÖZET

Çeşitli Çekirdek Fonksiyonları ile Oluşturulan Destek Vektör Makinesi Modellerinin Performanslarının İncelenmesi: Bir Klinik Uygulama

Amaç: Bu araştırmanın birincil amacı; çeşitli çekirdek fonksiyonları ile oluşturulan destek vektör makinesi modellerinin, akut koroner sendromlu hastalarda diabetes mellitus'u sınıflandırma performanslarının incelenmesi ve karşılaştırılmasıdır. Bu araştırmanın ikincil amacı ise, destek vektör makinesi modeli oluşturulurken kullanılan çeşitli çekirdek fonksiyonlarının parametrelerinin optimize edilerek en iyi sınıflandırma performansını elde etmeye çalışmaktır.

Materyal ve Metot: Bu çalışmada incelenen veriler, İnönü Üniversitesi Turgut Özal Tıp Merkezi Kardiyoloji Anabilim Dalı için geliştirilen veritabanından geriye yönelik (retrospektif) olarak seçilmiştir. Çalışmadaki söz konusu veriler akut koroner sendromlu hastalarda tip 2 diabetes mellitus ile değişik demografik ve klinik değişkenleri içermektedir. Akut koroner sendromlu hastalarda tip 2 diabetes mellitus'un sınıflandırılması için Destek Vektör Makinesi modelleri kullanılmıştır. İlgili modeller, ANOVA radyal tabanlı fonksiyon, bessel, doğrusal, Gaussian radyal tabanlı fonksiyon, laplace, polinomiyal ve sigmoid çekirdekleri ile oluşturulmuştur.

Bulgular: Laplace çekirdek fonksiyonu ile oluşturulan en iyi sınıflama performansına sahip destek vektör makinesi modeline ilişkin doğruluk, ROC eğrisi altında kalan alan, duyarlılık ve özgüllük [seçicilik] ölçütleri ile % 95 güven aralığı değerleri sırasıyla; 0.9804 (0.9716 - 0.987), 0.9332 (0.9096 - 0.9567), 0.9999 (0.9791 – 1.000) ve 0.9776 (0.9675 – 0.9852) olarak elde edilmiştir.

Sonuç: İncelenen değişik çekirdek fonksiyonları ile oluşturulan modeller arasında söz konusu performans ölçütleri dikkate alındığında, en iyi sınıflama performansı laplace Destek Vektör Makinesi modelinden elde edilmiştir. İlerleyen çalışmalarda, farklı klinik verilerde değişik çekirdek fonksiyonlu Destek Vektör Makinesi modelleri ile diğer makine öğrenmesi ya da veri madenciliği algoritmalarının kullanılması hastalıkların sınıflandırma başarısını artırabilecektir.

Anahtar Kelimeler: Destek vektör makinesi, çekirdek fonksiyonları, akut koroner sendromu, Tip II diabetes mellitus.

ABSTRACT

The Performance Exploration of Support Vector Machines Models Constructed with Various Kernel Functions: A Clinical Application

Aim: The primary aim of this study is to examine and compare the classification performance of support vector machine models generated by various core functions used to classify diabetes mellitus in acute coronary syndrome patients. The secondary aim is to optimize the parameters of the various kernel functions which are used for constructing the support vector machine model and to achieve the best classification performance.

Material and Methods: The data examined in this study were selected retrospectively from the database developed for Inonu University Turgut Ozal Medical Center Cardiology Department. The study included type 2 diabetes mellitus and various demographic and clinical variables in acute coronary syndrome patients. The Support Vector Machine model was used to classify type 2 diabetes mellitus in acute coronary syndrome patients. The related models are constructed by ANOVA radial basis function, bessel, linear, Gaussian radial basis function, laplace, polynomial and sigmoid kernel functions.

Results: The best classification performance was obtained by Support Vector Machine model constructed by laplace kernel function based on the results of performance metrics. The accuracy, area under ROC curve, sensitivity and specificity metrics with 95% CI were calculated as; 0.9804 (0.9716 - 0.987), 0.9332 (0.9096 - 0.9567), 0.9999 (0.9791 – 1.000) and 0.9776 (0.9675 – 0.9852), respectively.

Conclusion: When the performance metrics were taken into account, the best classification performance was achieved from the laplace Support Vector Machine model. In subsequent studies, the use of Support Vector Machine models with different kernel functions and other machine learning or data mining algorithms in different clinical trials may improve the classification success of the diseases.

Key words: Support vector machine, kernel functions, acute coronary syndrome, type II diabetes mellitus.

SİMGELER VE KISALTMALAR DİZİNİ

ACE	: Anjiyotensin Dönüştürücü Enzim
ARB	: Anjiotensin Reseptör Blokerleri
APA	: Amerikan Fizyoloji Derneği
ASA	: Aspirin
AKS	: Akut Koroner Sendrom
ANOVA	: Varyans Analizi
AUC	: Eğri altında kalan alan
CABG	: Koroner Arter Bypass
DKB	: Diastolik Kan Basıncı
DM	: Diabetes mellitus
DVM	: Destek Vektör Makinesi
GB	: Gigabayt
HDL	: Yüksek Yoğunluklu Lipoprotein
HTML	: Hyper Text Markup Language
KAH	: Koroner Arter Hastalığı
LDL	: Düşük Yoğunluklu Lipoprotein
LOF	: Lokal Aykırı Faktörü
ME	: Miyokardial Enfarktüs
PAH	: Periferik Arter Hastalığı
PHP	: Personel Home Page
RBF	: Radyal Tabanlı Fonksiyon
ROC	: Alıcı İşlem Karakteristiği
SKB	: Sistolik Kan Basıncı
VKİ	: Vücut Kitle İndeksi
VTBK	: Veri Tabanlarında Bilgi Keşfi
VTYS	: Veri Tabanı Yönetim Sistemi
WHO	: Dünya Sağlık Örgütü

ŞEKİLLER DİZİNİ

Şekil No	Sayfa No
Şekil 2.1: VTBK sürecinin grafiksel gösterimi (6)	3
Şekil 2.2: Veri madenciliğinin kapsadığı ve ilişkili olduğu disiplinler	5
Şekil 2.3: Veri madenciliği uygulama alanları	6
Şekil 2.4: Veri madenciliği yöntemlerinin denetimli-denetimsiz olarak sınıflandırılması	7
Şekil 2.5: Veri madenciliğinin uygulamasında kullanılacak bir yöntem bilimsel yaklaşım	8
Şekil 2.6: Veri madenciliği yöntemlerinin tahmini ve tanımlayıcı olarak sınıflandırılması	9
Şekil 2.7: 5-katlı çapraz geçerlilik yöntemine ilişkin grafiksel gösterim	10
Şekil 2.8: Bootstrap örnekleme tekniğine ilişkin grafiksel gösterim	11
Şekil 2.9: ROC eğrisi altında kalan 4 farklı alanın değişik değerlerine sahip dört ROC eğrisi	12
Şekil 2.10: Doğrusal DVM modeline ilişkin grafiksel gösterim	14
Şekil 2.11: Doğrusal olarak sınıflandırılmayan girdi uzayının bir üst boyuta çekirdek fonksiyonu ile haritalanması	15
Şekil 3.1: Güç analizi hesabına ilişkin gösterim	24
Şekil 3.2: Web tabanlı kardiyoloji veri giriş sisteminin giriş ekranı	27
Şekil 3.3: Hasta ile ilgili demografik verilerin kaydedildiği ara yüz	28
Şekil 3.4: Hastaneye başvuran hastanın tansiyon, nabız, laboratuvar parametreleri gibi verilerinin kaydedildiği ekran	29
Şekil 3.5: “hasta_bilgi_demografik” tablosunun yapısı	30
Şekil 3.6: “hasta_bilgi_basvuru” tablosunun yapısı	32
Şekil 3.7: “kurumlar” tablosunun yapısı	33
Şekil 3.8: Web tabanlı programı kullanan hastaneler	33
Şekil 3.9: Kullanıcıların şifreleri ile giriş yapabilmeleri için gerekli bilgilerin kaydedildiği tablo	34
Şekil 3.10: Doğrusal DVM'ye ilişkin destek vektörler, en uygun hiperdüzlem ve en geniş sınırlar	35
Şekil 3.11: DVM'nin yapısı	36
Şekil 4.1: Maliyet, derece ve ölçek parametrelerine göre doğruluk metriğinin değişimi (Çekirdek fonksiyon=Polinomial)	57

Şekil 4.2: Maliyet parametresine göre doğruluk metriğinin değişimi (Çekirdek fonksiyon= Gaussian RBF).....	58
Şekil 4.3: Maliyet parametresine göre doğruluk metriğinin değişimi (Çekirdek fonksiyon=Doğrusal)	59
Şekil 4.4: Maliyet parametresine göre doğruluk metriğinin değişimi (Çekirdek fonksiyon=Laplace)	60
Şekil 4.5: Maliyet ve derece parametrelerine göre doğruluk metriğinin değişimi (Çekirdek fonksiyon=Bessel).....	61
Şekil 4.6: Maliyet ve derece parametrelerine göre doğruluk metriğinin değişimi (Çekirdek fonksiyon= ANOVA RBF)	62
Şekil 4.7: Maliyet ve ölçek parametrelerine göre doğruluk metriğinin değişimi (Çekirdek fonksiyon= hiperbolik tanjant (sigmoid)).....	63
Şekil 4.8: DVM modelinde Laplace çekirdek fonksiyonu için ROC eğrisi grafiği	64
Şekil 4.9: DVM modelinde Gaussian RBF çekirdek fonksiyonu için ROC eğrisi grafiği.....	65
Şekil 4.10: DVM modelinde Polinomiyal çekirdek fonksiyonu için ROC eğrisi grafiği.....	66
Şekil 4.11: DVM modelinde Doğrusal çekirdek fonksiyonu için ROC eğrisi grafiği.....	67
Şekil 4.12: DVM modelinde Bessel çekirdek fonksiyonu için ROC eğrisi grafiği	68
Şekil 4.13: DVM modelinde Hiperbolik Tanjant (Sigmoid)çekirdek fonksiyonu için ROC eğrisi grafiği.....	69
Şekil 4.14: DVM modelinde ANOVA RBF çekirdek fonksiyonu için ROC eğrisi grafiği.....	70
Şekil 4.15: DVM modelinde kullanılan tüm çekirdek fonksiyonları için ROC eğrisi grafiği.....	71
Şekil 4.16: DVM'nin tüm çekirdek fonksiyonlarının çeşitli performans metriklerine göre sınıflandırma performansları (% 95 güven aralığı ile)	73

TABLULAR DİZİNİ

Tablo No	Sayfa No
Tablo 2.1: Bazı performans metriklerinin hesaplamasına yönelik karmaşıklık matrisi	11
Tablo 3.1: Çalışmada kullanılan değişkenlerin tanımları	22
Tablo 3.2: Çekirdek tipleri ve fonksiyonları	37
Tablo 3.3: Çekirdek fonksiyonlarının optimizasyonları parametrelerine ilişkin tanımlayıcı bilgiler	38
Tablo 3.4: Karmaşıklık matrisi	39
Tablo 4.1: DM değişkenin dağılımı	41
Tablo 4.2: AKS tipi değişkeni açısından DM varlığı ve yokluğuna ilişkin dağılım.....	41
Tablo 4.3: Cinsiyet değişkeni açısından DM varlığı ve yokluğuna ilişkin dağılım.....	42
Tablo 4.4: ACE inhibitör değişkeni açısından DM varlığı ve yokluğuna ilişkin dağılım	42
Tablo 4.5: ARB diğer değişkeni açısından DM varlığı ve yokluğuna ilişkin dağılım ..	43
Tablo 4.6: ASA değişkeni açısından DM varlığı ve yokluğuna ilişkin dağılım	43
Tablo 4.7: Beta bloker değişkeni açısından DM varlığı ve yokluğuna ilişkin dağılım .	44
Tablo 4.8: Diüretik değişkeni açısından DM varlığı ve yokluğuna ilişkin dağılım.....	44
Tablo 4.9: Statin değişkeni açısından DM varlığı ve yokluğuna ilişkin dağılım.....	45
Tablo 4.10: HT değişkeni açısından DM varlığı ve yokluğuna ilişkin dağılım.....	45
Tablo 4.11: Kalp yetmezliği değişkeni açısından DM varlığı ve yokluğuna ilişkin dağılım	45
Tablo 4.12: ME değişkeni açısından DM varlığı ve yokluğuna ilişkin dağılım	46
Tablo 4.13: Renal yetmezlik değişkeni açısından DM varlığı ve yokluğuna ilişkin dağılım	46
Tablo 4.14: Sigara değişkeni açısından DM varlığı ve yokluğuna ilişkin dağılım.....	47
Tablo 4.15: Kalsiyum kanal bloker değişkeni açısından DM varlığı ve yokluğuna ilişkin dağılım	47
Tablo 4.16: Clopidogrel değişkeni açısından DM varlığı ve yokluğuna ilişkin dağılım	48
Tablo 4.17: AKS aile öyküsü değişkeni açısından DM varlığı ve yokluğuna ilişkin dağılım	48
Tablo 4.18: Koroner arter by-pass değişkeni açısından DM varlığı ve yokluğuna ilişkin dağılım	49

Tablo 4.19: Hiperlipidemi deęişkeni açısından DM varlığı ve yokluęuna ilişkin dağılım	49
Tablo 4.20: Koroner arter hastalığı deęişkeni açısından DM varlığı ve yokluęuna ilişkin dağılım	50
Tablo 4.21: Malignite deęişkeni açısından DM varlığı ve yokluęuna ilişkin dağılım ..	50
Tablo 4.22: PAH deęişkeni açısından DM varlığı ve yokluęuna ilişkin dağılım	51
Tablo 4.23: İnme deęişkeni açısından DM varlığı ve yokluęuna ilişkin dağılım.....	51
Tablo 4.24: Nicel deęişkenlerin DM'nin varlığına ve yokluęuna göre tanımlayıcı istatistikleri.....	52
Tablo 4.25: Gaussian RBF çekirdek fonksiyonlu DVM modeli için tahminlere ilişkin performans ölçütleri ve %95 güven aralığı deęerleri.....	53
Tablo 4.26: Hyperbolic tangent (Sigmoid) çekirdek fonksiyonlu DVM modeli için tahminlere ilişkin performans ölçütleri ve %95 güven aralığı deęerleri	53
Tablo 4.27: Polinomiyal çekirdek fonksiyonlu DVM modeli için tahminlere ilişkin performans ölçütleri ve %95 güven aralığı deęerleri.....	54
Tablo 4.28: Doğrusal çekirdek fonksiyonlu DVM modeli için tahminlere ilişkin performans ölçütleri ve %95 güven aralığı deęerleri.....	54
Tablo 4.29: Laplace çekirdek fonksiyonlu DVM modeli için tahminlere ilişkin performans ölçütleri ve %95 güven aralığı deęerleri.....	55
Tablo 4.30: ANOVA RBF çekirdek fonksiyonlu DVM modeli için tahminlere ilişkin performans ölçütleri ve %95 güven aralığı deęerleri.....	55
Tablo 4.31: Bessel çekirdek fonksiyonlu DVM modeli için tahminlere ilişkin performans ölçütleri ve %95 güven aralığı deęerleri.....	56
Tablo 4.32: DVM'nin tüm çekirdek fonksiyonlarının çeşitli performans metriklerine göre sınıflandırma performansları (%95 güven aralığı ile) ...	72
Tablo 4.33: Çalışmada kullanılan deęişkenlerin laplace çekirdek fonksiyonuyla oluşturulan DVM modeline göre önemlilik dereceleri	74

1. GİRİŞ

Akut koroner sendrom, acil servise başvuruların ve hastanelerin koroner yoğun bakım birimleri başta olmak üzere hastaların yatırılma sebeplerinin başında yer almaktadır. Akut koroner sendromun neden olduğu iş gücü kaybı, morbidite ve mortaliteye ek olarak yüksek hastane maliyetlerine sebep olması açısından günümüzde en önemli sağlık problemlerinden biri olarak kabul edilmektedir (1). Akut koroner sendromu, kalp kasının bir kısmının düzgün çalışmaması veya işlevinin kaybolması gibi koroner atardamarlarda azalmış kan akışına bağlı olarak ortaya çıkan bir sendromdur. En sık rastlanan belirtilerinden bazıları çoğunlukla sağ omuza yayılan göğüs ağrısı ile bulantı ve terlemedir. Akut koroner sendromu, genellikle ST yükselmeli miyokard enfarktüsü (STEMI, % 30), ST yükselmesi olmayan miyokard enfarktüsü (NSTEMI, % 25) veya kararsız angina (% 38) olmak üzere üç nedenden ortaya çıkmaktadır. Bu tipler, EKG / EKG görünümüne göre ST segment yükselmesi olmayan miyokard infarktüsü ve ST segment yükselmesi miyokard enfarktüsü olarak adlandırılır (2, 3).

Diabetes mellitus, kandaki glikoz düzeyinin yükselmesiyle neticelenen, çoğunlukla kalıtsal ve çevre ile ilgili faktörlerin bir araya gelmesi ile meydana gelen bir hastalıktır. Dünya Sağlık Örgütü (WHO) verilerine göre, diyabetli insanların sayısı 1980'de 108 milyon iken, 2014 'de bu sayı 422 milyona; hastalığın 18 yaş üstü erişkinlerde genel prevalansı ise 1980'de % 4.7 iken, 2014'de % 8.5'e yükselmiştir. Diyabetin 2030'da tüm dünyadaki ölüm nedenleri arasında 7. sırada olacağı öngörülmektedir. Orta ve düşük gelirli ülkelerde, diyabet prevalansı hızla artmaktadır. Ayrıca, diyabet; körlük, böbrek yetmezliği, kalp krizi, inme ve alt ekstremitte amputasyonunun ana nedenlerindedir (4).

Diabetes mellitus, pankreasın yeterli insülin salgılamaması ya da üretilen insüline normal olarak tepki vermeyen vücut hücrelerinden kaynaklanmaktadır. Diabetes mellitusun üç ana türü vardır:

Tip 1 DM, pankreasın yeterli insülin salgılamamasından kaynaklanmaktadır. Tip 1 DM daha önce "insüline bağımlı diabetes mellitus olarak adlandırılmıştır (5).

Tip 2 DM ise, insülin direnci ile başlayan ve hücrelerin insüline normal olarak tepki vermemesi durumudur. Tip 2 DM'nin ilerlemesi durumunda, insülin eksikliği de

meydana gelebilir. Tip 2 DM, daha önce "insüline bağımlı olmayan diabetes mellitus" olarak adlandırılmıştır. Tip 2 DM'nin en yaygın nedenleri arasında, aşırı vücut ağırlığı ve yeterince egzersiz yapmama yer almaktadır (5). Tip 2 diyabetin gelişmesini önlemede, sağlıklı beslenme, düzenli fiziksel aktivite, normal vücut ağırlığı ve sigarayı bırakma vb. faktörler önemli rol oynamaktadır (4).

Gestasyonel DM ise, daha önceden diyabet öyküsü olmayan gebe kadınlarda, yüksek kan şekeri düzeyi gelişimiyle ortaya çıkan üçüncü tip DM türüdür. Risk faktörleri arasında, obezite, önceden gestasyonel diyabet geçirme hikâyesi, aile öyküsü, tip 2 diyabet öyküsü ve polikistik over sendromu bulunabilmektedir (5).

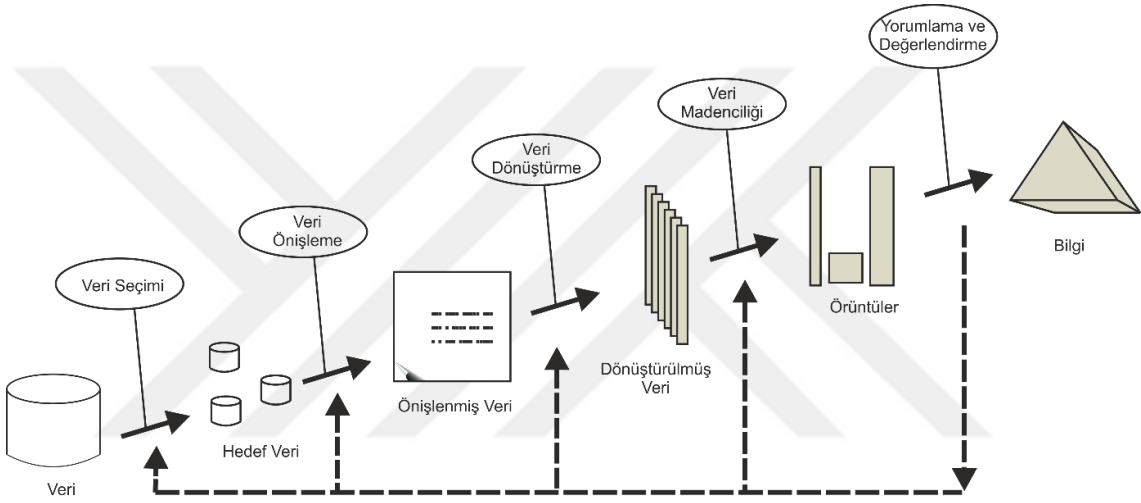
Destek vektör makinesi modelleri son zamanlarda popüler bir algoritma haline gelmektedir. Destek vektör makineleri, sınıflandırma veya regresyon problemleri için kullanılabilen denetimli/danışmanlı bir makine öğrenme/veri madenciliği algoritmasıdır. Bu süreçte verileri dönüştürmek için çekirdek çözümü (kernel trick) adı verilen bir teknik kullanır. Bu çekirdek çözümü, veri dönüşümlerine dayanarak olası çıktılar arasında en uygun sınırı bulur. Başka bir ifadeyle, çekirdek çözümü son derece karmaşık veri dönüşümleri yapar ve daha sonra bu verilerin, tanımlanan etiketler veya çıktılara dayalı olarak nasıl ayrıldığı belirlenir. Destek vektör makineleri, optimal bir hiperdüzlem oluşturmak için hata fonksiyonunu en aza indirmek amacıyla kullanılan yinelemeli bir eğitim algoritması kullanır. Bu sayılan esnek ve önemli özellikleri nedeniyle, bu çalışmada ele alınan sınıflama probleminin çözümünde değişik çekirdek fonksiyonlarının tahmin performanslarını incelemek için destek vektör makineleri seçilmiştir.

Bu araştırmanın birincil amacı; çeşitli çekirdek fonksiyonları ile oluşturulan destek vektör makinesi modellerinin, akut koroner sendromlu hastalarda diabetes mellitus'u sınıflandırma performanslarının incelenmesi ve karşılaştırılmasıdır. Bu araştırmanın ikincil amacı ise, destek vektör makinesi modeli oluşturulurken kullanılan çeşitli çekirdek fonksiyonlarının parametrelerini optimize ederek en iyi sınıflandırma performansını elde etmeye çalışmaktır.

2. GENEL BİLGİLER

2.1. Veri Tabanlarında Bilgi Keşfi Süreci

Veri tabanlarında bilgi keşfi (VTBK), üretme ve saklama kapasitesine sahip olunan büyük miktarda veride, gizli olarak bulunan bilgiyi ortaya çıkarma sürecidir. VTBK süreci, veritabanlarından yararlanma suretiyle, veri seçimi, veri önışleme, veri dönüşümü ve veri indirgemesi, veri madenciliği ve veri madenciliği çıktılarının değerlendirilmesi ile yorumlanması adımlarından oluşur (6, 7). VTBK sürecini oluşturan basamaklar Şekil 1’de verilmiştir (6).



Şekil 2.1: VTBK sürecinin grafiksel gösterimi (6)

2.1.1. Veri Seçimi

Bu aşamada, VTBK sürecinde kullanılacak olan ham veri ilgili veritabanlarından elde edilir.

2.1.2. Veri Önışleme

Bu aşamada kayıp değer analizleri yapılır ve eğer veri setinde kayıp değerler mevcut ise bu kayıp değerlere çeşitli yöntemler kullanılarak uygun değerler atanır. Ayrıca veri setinde aşırı/aykırı değerler çeşitli yaklaşımlar yapılarak tespit edilir ve bu değerler veri setinden çıkarılır.

2.1.3. Veri Dönüşümü ve Veri İndirgeme

Bu adımda, normalizasyon, standardizasyon gibi çeşitli dönüşüm teknikleri uygulanır. Ayrıca, istatistik tabanlı boyut indirgeme ve makine öğrenmesi tabanlı özellik seçimi yöntemleri kullanılarak veri boyutunda indirgeme yapılır. Verinin boyutunu indirmede amaç, veri madenciliği sürecinde sonuca katkısı olmayacak verilerin devre dışı bırakılmasıdır.

2.1.4. Veri Madenciliği

Veri madenciliği, yığın veri setlerindeki ilişki ve örüntüleri çeşitli istatistik ve makine öğrenmesi yöntemleri kullanarak ortaya çıkaran ve bunlardan tutarlı tahminler yapmak için işletilen bir süreç olarak tanımlanabilir (8, 9). Şekil 2.2’de de belirtildiği üzere, Veri madenciliği; istatistik, yapay zekâ, yönetim bilişim sistemleri, örüntü tanıma, matematiksel modelleme ve veritabanı faaliyetlerini kapsayan disiplinler arası bir disiplindir (10).



Şekil 2.2: Veri madenciliğinin kapsadığı ve ilişkili olduğu disiplinler

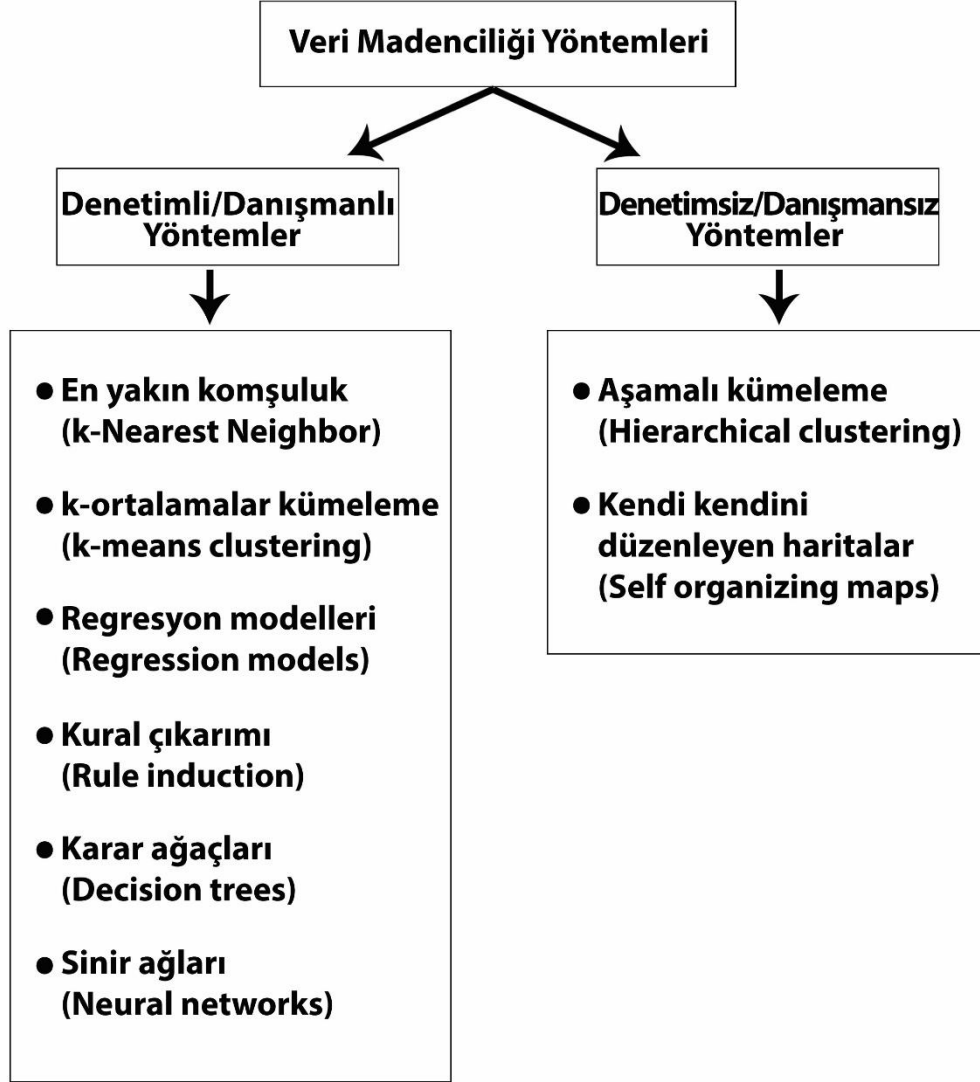
Veri madenciliği birçok alanda uygulama alanına sahiptir. Başlıca uygulama alanları Şekil 2.3’de verilmiştir (11).

Uygulama alanı	Uygulama amacı
Pazarlama	Pazar segmentasyonu, müşteri değerlendirme ve çapraz satış analizleri.
Bankacılık	Risk analizi, usulsüzlük tespiti, müşteri kazanma ve mevcut müşterileri elde tutma analizleri, çapraz satış.
Sigortacılık	Müşteri kaybı sebeplerinin belirlenmesi, usulsüzlüklerin önlenmesi, ana giderlerin azaltılması, poliçe fiyatlarının belirlenmesi.
Perakendecilik	Satış noktası veri analizleri, alışveriş sepeti analizleri, tedarik ve mağaza yerleşim optimizasyonları.
Borsa	Hisse senedi fiyat tahmini, genel piyasa analizleri, alım-satım stratejilerinin optimizasyonu.
Telekomünikasyon	Kalite iyileştirme, hile tespiti, hataların yoğunluk tahmini, müşteri kazanma ve elde tutma analizleri.
İlaç	Test sonuçlarının tahmini, ürün geliştirme
Sağlık	Tıbbi teşhis ve tanı, uygun tedavi sürecinin belirlenmesi, sınıflandırma.
Endüstri	Kalite kontrol, lojistik, üretim süreçlerinin optimizasyonu.
Bilim ve Mühendislik	Ampirik veriler üzerinde modeller kurularak bilimsel ve teknik problemlerin çözülmesi.

Şekil 2.3: Veri madenciliği uygulama alanları

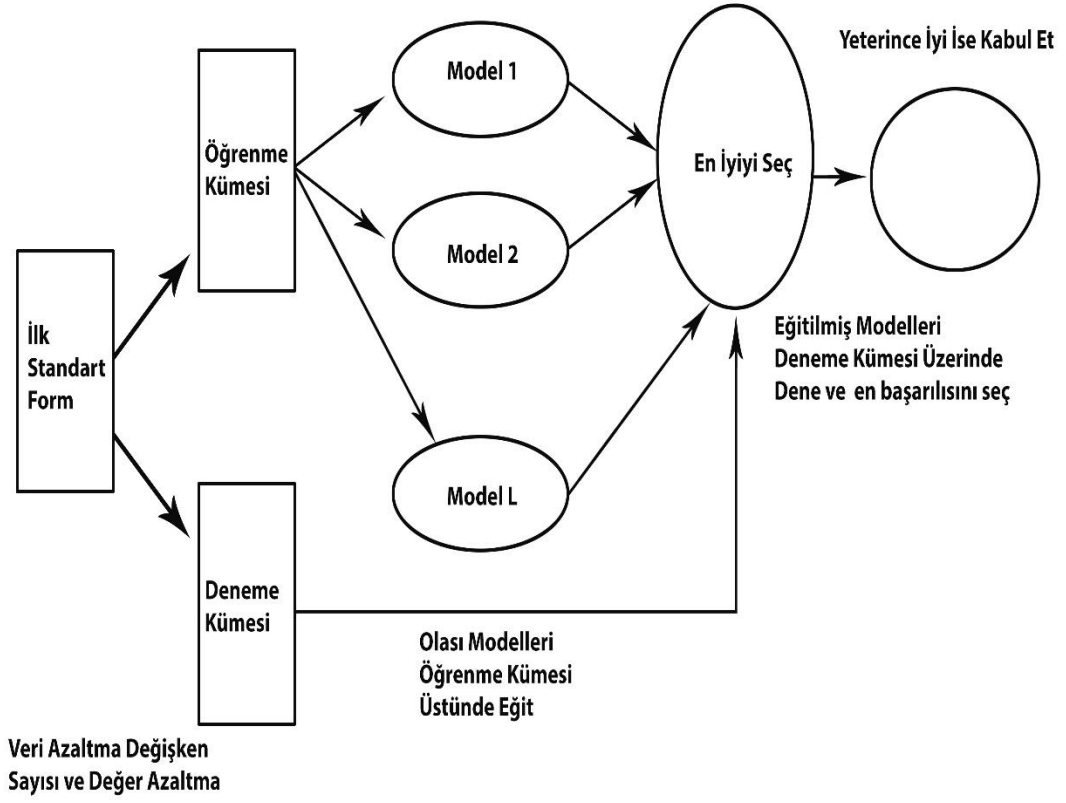
Veri madenciliği yöntemleri, denetimli/danışmanlı (supervised learning) ve denetimsiz/danışmansız (unsupervised learning) öğrenme olmak üzere genel olarak ikiye ayrılır. Denetimli öğrenmede, girdi verilerine göre çıktı değerleri tahmin edilmektedir. Denetimli öğrenmede tahmin edilecek çıktı değerleri önceden belirlenmiştir (12). Örneğin, bir hastalığa ilişkin risk faktörleri ve demografik verilerden yola çıkarak bireylerin hasta ve kontrol olarak sınıflandırılması denetimli öğrenmeye örnektir. Denetimsiz öğrenmede ise, çıktı değerleri olmadığı ve belirlenmediği durumlarda sadece

girdi değerlerinden yola çıkarak girdi verilerinin gruplandırılması hedeflenir (12). Veri madenciliği yöntemlerinin denetimli-denetimsiz olarak sınıflandırılması Şekil 2.4’de verilmiştir (13).



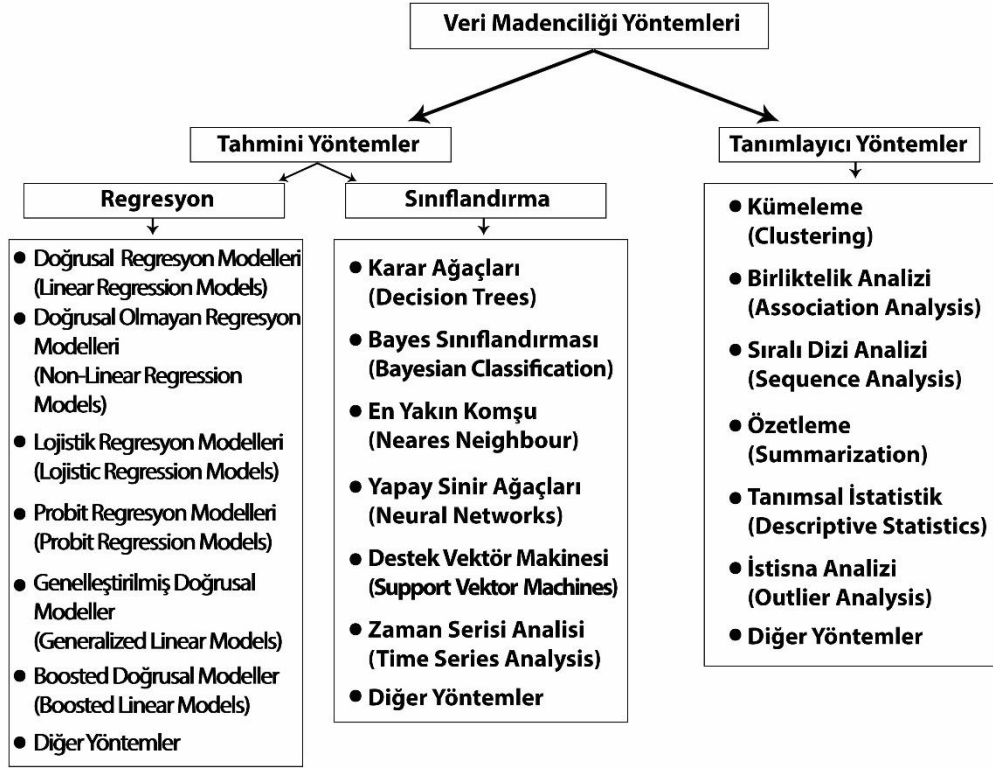
Şekil 2.4: Veri madenciliği yöntemlerinin denetimli-denetimsiz olarak sınıflandırılması

Veri madenciliğinin uygulamasında kullanılacak yöntem bilimsel yaklaşımlardan birisi Şekil 2.5’de gösterilmiştir (14). Bu yaklaşımda, öncelikle verideki değişken sayısı, özellik seçimi (Feature Selection) teknikleri uygulanarak azaltılır. Sonrasında ise girdi verisi eğitim ve test verisi olarak ikiye ayrılır. Eğitim verisi ilgili veri madenciliği algoritmalarını eğitmede kullanılır. Algoritmanın eğitim verisinden öğrendikleri ise test verisi üzerinde uygulanır. Sonucu aşamada ise çeşitli uyum iyiliği ölçütleri kullanılarak modelin performansı belirlenir.



Şekil 2.5: Veri madenciliğinin uygulamasında kullanılabilir bir yöntem bilimsel yaklaşım

Şekil 2.6’da görüldüğü gibi, veri madenciliği yöntemleri aynı zamanda tahmini yöntemler ve tanımlayıcı yöntemler olarak iki grupta incelenebilir (13).



Şekil 2.6: Veri madenciliği yöntemlerinin tahmini ve tanımlayıcı olarak sınıflandırılması

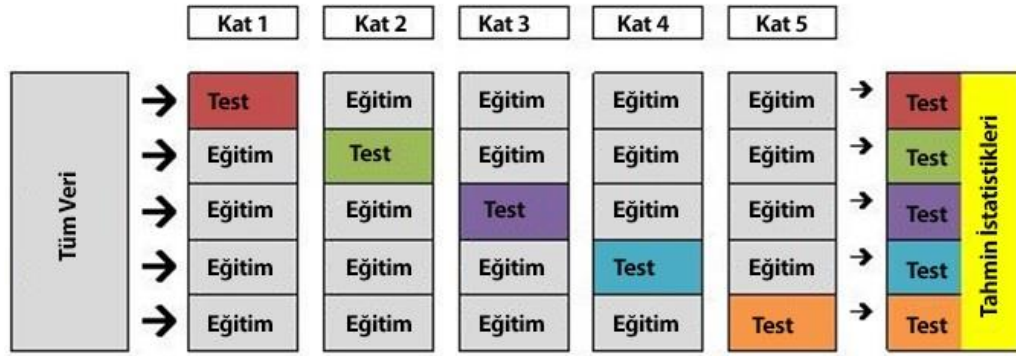
2.1.5. Değerlendirme ve Yorumlama

Değerlendirme aşamasında çeşitli yeniden örnekleme teknikleri kullanılarak, modelin öğrenme performansı değerlendirilir. Yeniden örnekleme teknikleri, eğitilmiş modelin tahmin performansının yanlılığını ve şansa bağlı olarak elde edilme olasılığını azaltır. Yeniden örnekleme yöntemlerinden en bilinen ve en çok kullanılanları aşağıda verilmiştir:

- a. **Basit Çapraz Geçerlilik (Holdout):** Bu yöntemde, kullanılan veriseti eğitim ve test veriseti olmak üzere ikiye bölünür. Eğitim veriseti kullanılarak veri eğitilirken, test veriseti kullanılarak eğitilmiş modelin tahmin performansı ölçülür. Modelin optimizasyon parametreleri mevcut ise, bu parametrelerin uygun hale getirilmesi eğitim verisetinde yapılır. Test verisetinin tahmin edilmesinde en uygun hale getirilen parametreler kullanılır. Veriseti, eğitim ve

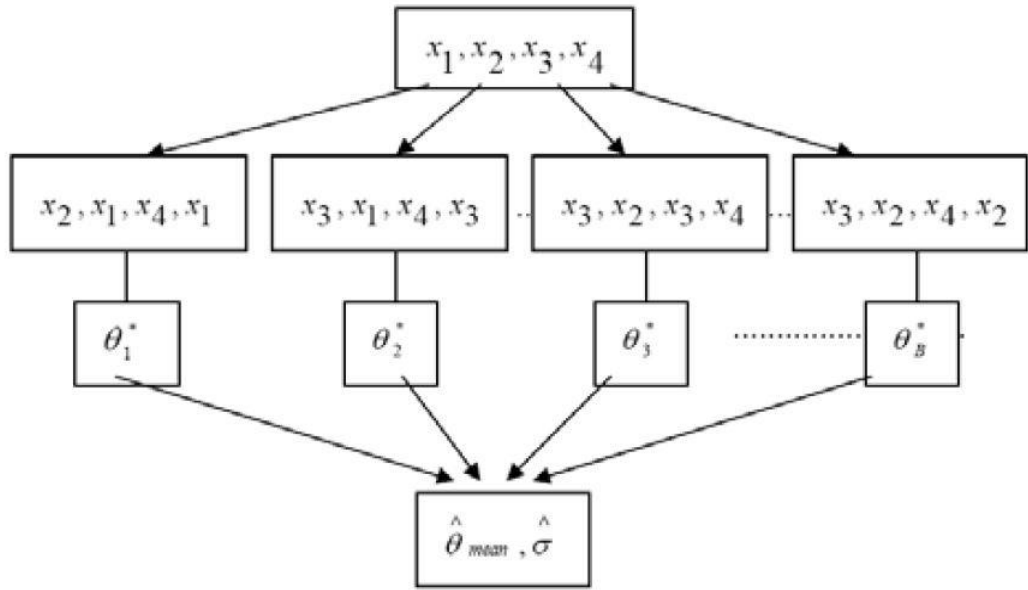
test verisetleri olarak ikiye ayrılırken, eğitim verisetinin test verisetinden fazla seçilmesi önerilebilir.

- b. k-katlı Çapraz Geçerlilik:** Bu teknikte veriseti $k > 2$ olmak üzere k parçaya ayrılır. k parçaya ayrılan verinin $(k-1)$ tanesi eğitim veriseti, 1 tanesi ise test veriseti olarak kullanılır. Bu süreç k defa tekrarlanarak ve her defasında farklı eğitim ve test verisetleri kullanılır. Süreç sonucunda elde edilen k adet doğruluk değerinin ortalaması alınır ve nihai doğruluk değeri modelin doğruluk değeri olarak kabul edilir. Bu yöntemde basit çapraz geçerlilik yönteminin aksine tüm veriseti kullanılır. Bu bağlamda, k -katlı çapraz geçerlilikte, basit çapraz geçerlilik yöntemine göre daha tutarlı sonuçlar elde edilir. 5-katlı çapraz geçerlilik yöntemine ilişkin grafiksel gösterim Şekil 2.7'de gösterilmiştir (15).



Şekil 2.7: 5-katlı çapraz geçerlilik yöntemine ilişkin grafiksel gösterim

- c. Bootstrap örnekleme:** Bootstrap yöntemi, standart sapma ve güven aralığı vb. istatistikler ve parametrik olmayan kestirim problemlerinde yaygınlaştıran bir yaklaşımdır. Bootstrap yönteminin temel çıkış noktası, mevcut veri kümesindeki gözlem değerlerinin rastgele yer değiştirmesi ile yeniden örneklemlerin oluşturulmasıdır. Böylelikle, veri kümesinden daha büyük miktarda bilgi çıkarımı yapılabilmektedir. Bootstrap yönteminde, öncelikle asıl örneklemden yerine koyma tekniği ile örnek gözlemler alınarak bootstrap örnekleme elde edilmektedir. Sonrasında ise benzer biçimde yeterli miktarda bootstrap örneklemleri elde edilerek, bu örneklemler ile tahminleyicinin güven aralığı ve testin istatistiksel önemliliği belirlenmektedir (16). Bootstrap örnekleme tekniğine ilişkin grafiksel gösterim Şekil 2.8'de verilmiştir (16).



Şekil 2.8: Bootstrap örnekleme tekniğine ilişkin grafiksel gösterim

Yorumlama aşamasında, çeşitli performans metrikleri ve istatistiksel uyum iyiliği testleri kullanılarak, kullanılan veri madenciliği modelinin hedef değişkeni tahminleme performansı ölçülür ve yorumlanır. Doğruluk (accuracy), duyarlılık (sensitivity), özgüllük [seçicilik] (specificity), ROC (Receiver Operating Characteristics) eğrisi altında kalan alan, vb. en çok kullanılan performans ölçüm metriklerindedir. Bu performans metriklerinin hesaplamaları aşağıdaki karmaşıklık matrisine göre yapılabilmektedir.

Tablo 2.1: Bazı performans metriklerinin hesaplamasına yönelik karmaşıklık matrisi

		Gerçek Durum		
		Pozitif	Negatif	Toplam
Tahmin	Pozitif	Doğru pozitif (DP)	Yanlış negatif (YN)	DP+YN
	Negatif	Yanlış pozitif (YP)	Doğru negatif (DN)	YP+DN
	Toplam	DP+YP	YN+DN	DP+DN+YP+YN

$$\text{Doğruluk} = (DP+DN)/(DP+DN+YP+YN)$$

$$\text{Duyarlılık} = DP/(DP+YP)$$

$$\text{Özgüllük [Seçicilik]} = DN/(DN+YN)$$

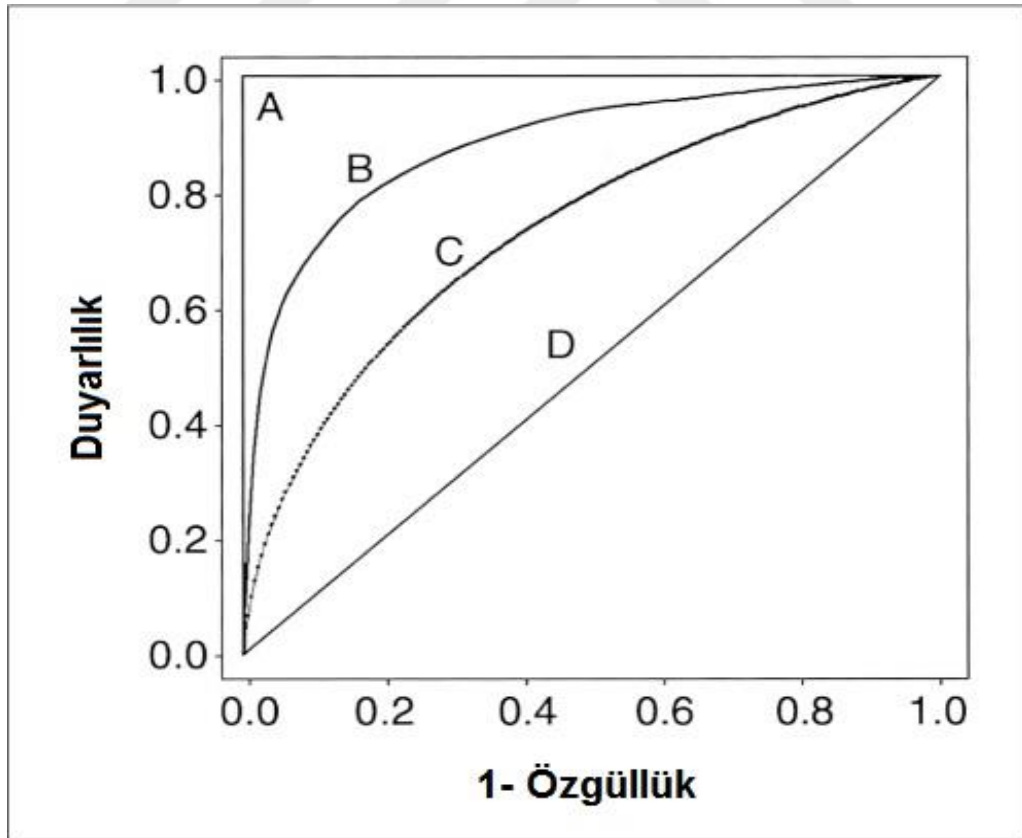
$$\text{Pozitif tahmin değeri} = DP/(DP+YN)$$

$$\text{Negatif tahmin değeri} = DN/(DN+YP)$$

$$G_{\text{ortalama}} = \sqrt{(\text{Duyarlılık} \times \text{Özgüllük})}$$

Burada, DP: doğru pozitif sayısını, DN: doğru negatif sayısını, YP: yanlış pozitif sayısını ve YN: yanlış negatif sayısını göstermektedir.

İşlem karakteristik eğrisi (ROC eğrisi), bir ikili sınıflandırıcı yöntemin değişik kesim noktalarındaki tahmin performansını gösteren grafiksel bir yöntemdir. ROC eğrisi, doğru pozitif (DP) oranının (duyarlılık) yanlış pozitif (YP) oranına (1-özgüllük) karşı değişik kesim noktalarında çizilmesi ile oluşturulur. Doğru pozitif oranı duyarlılık ve yanlış pozitif oranı ise özgüllük [seçicilik] ya da seçicilik olarak adlandırılır.



Şekil 2.9: ROC eğrisi altında kalan 4 farklı alanın değişik değerlerine sahip dört ROC eğrisi.

Şekil 2.9’da görüldüğü gibi, ROC eğrisi altında alanın farklı değerlerine sahip dört ROC eğrisi söz konusu olabilir. Eğer bir test/yöntem sınıfları tam olarak ayırabiliyorsa, ROC eğrisi altında kalan alan 1’e eşittir (A). Eğer ROC eğrisi diyagonal çizgi üzerindeyse (D, (0, 0) ile (1, 1) noktaları arasındaki çizgi), ROC eğrisinin altındaki alan 0.5’e eşittir. Bu ise testin sınıfları ayırmasama gücünün olmadığını ifade etmektedir. Eğer testlere ilişkin ROC eğrileri B ve C’deki gibi bir durumda ise testin ayırmasama gücü bulunmaktadır. Bu durumda, B eğrisinin altında kalan alan C eğrisinin altında kalan alana göre daha fazla olduğu için ROC eğrisi B olan test/yöntem, ROC eğrisi C olan test/yönteme göre daha iyi ayırmasama performansına sahiptir.

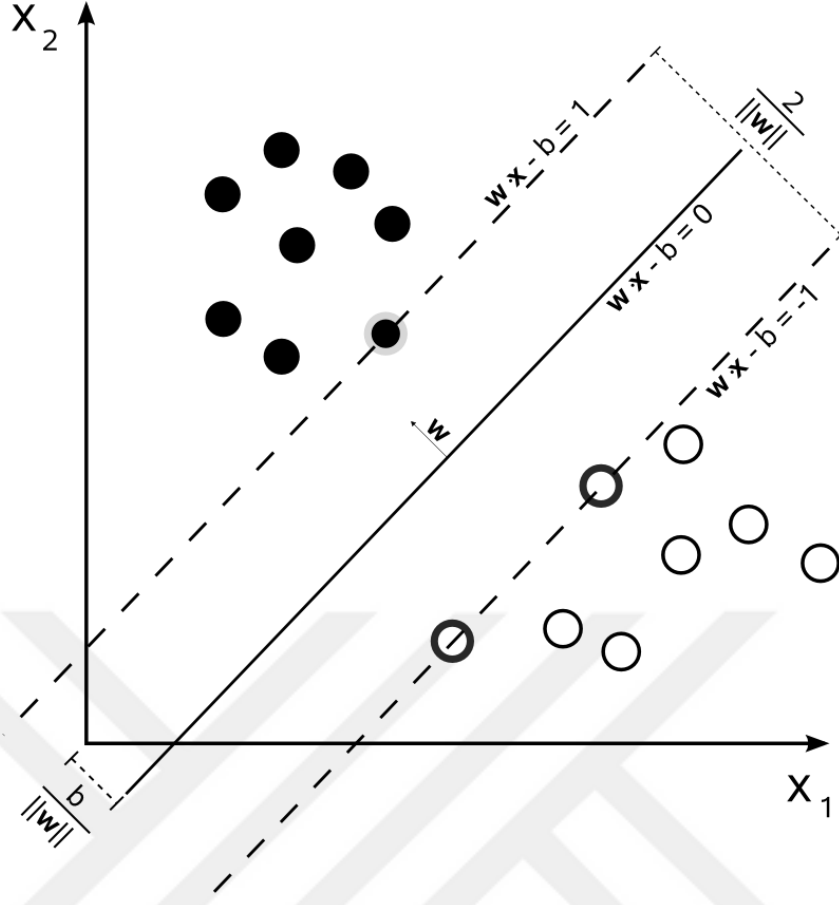
2.2. Destek Vektör Makinesi

DVM, Vapnik - Chervonenkis tarafından geliştirilen, sınıflandırma başta olmak üzere kümeleme ve regresyon problemlerinde kullanılan bir makine öğrenmesi modelidir. DVM modelinin amacı temel olarak, hedef değişkene ait sınıfları birbirinden en uygun şekilde ayıracak hiperdüzlemi tespit etmektir (17). Bu çalışmada, DVM modeli 2-sınıflı hedef değişkeni (DM hastalığı) sınıflandırmak için kullanılmıştır.

DVM, verinin doğrusal ya da doğrusal olmayan şekilde sınıflandırılmasına göre iki durumda incelenir (18).

2.2.1. Doğrusal DVM

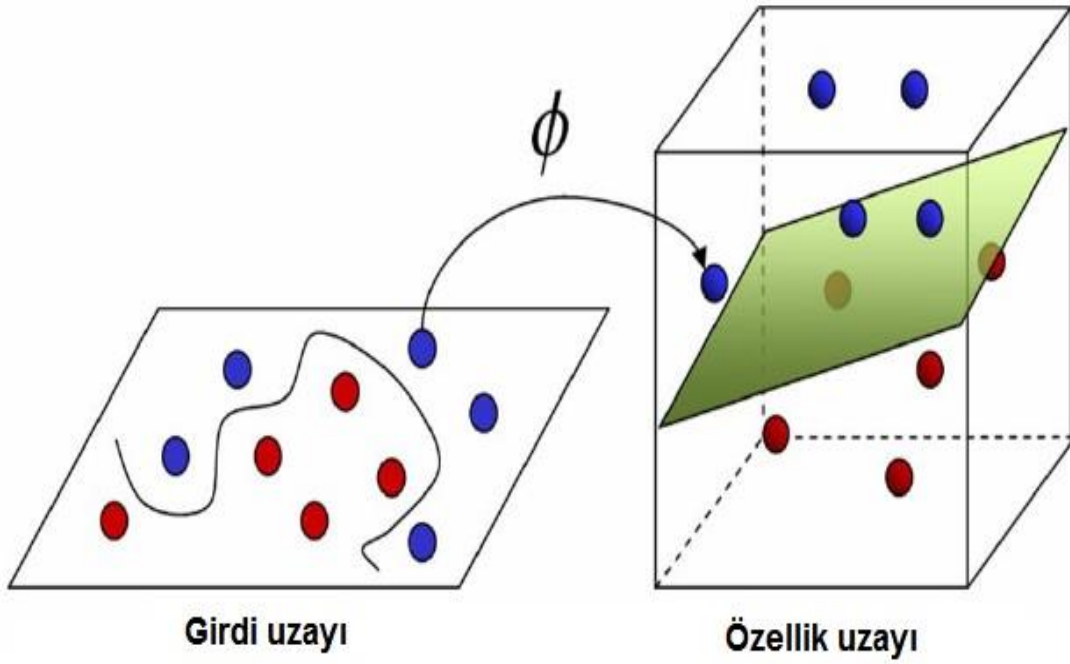
$(\vec{x}_1, y_1), (\vec{x}_2, y_2), \dots, (\vec{x}_n, y_n)$ bir grup gözlemden oluşan bir veri kümesi olsun. y_i değerlerinin her biri -1 ya da +1 sayısını alması ile her bir x_i değerine karşılık gelen bir sınıfı ifade etmektedir. Her x_i değeri p boyutlu bir vektördür. Maksimum aralıklı hiperdüzlemi, $y_i=1$ değeri için x_i gözlem noktalarını $y_i=-1$ için incelenen gözlem noktalarından en uygun biçimde ayırmaktadır. Bu hiperdüzlemde, her iki grubun en yakın noktası arasındaki mesafenin en üst düzeye çıkarılması amaçlanmaktadır. Herhangi bir hiperdüzlem, $\vec{w} \cdot \vec{x} = 0$ eşitliğini sağlayan \vec{x} noktalar kümesi olarak yazılabilir. Burada \vec{w} ; hiperdüzleme doğru bir ağırlık vektörünü göstermektedir. $\frac{b}{\|\vec{w}\|}$ parametresi; \vec{w} normal vektörü boyunca orijinden geçen hiperdüzlemi, b; eğilim değerlerini ve $\frac{2}{\|\vec{w}\|}$ ise; $w^T x^{(i)} + b = -1$ ile $w^T x^{(i)} + b = +1$ arası uzaklığı ifade etmektedir. Bu yapı, aşağıdaki şekilde betimlenmektedir (19). Şekil 2.10’da doğrusal DVM modeline ilişkin grafiksel gösterim verilmiştir.



Şekil 2.10: Doğrusal DVM modeline ilişkin grafiksel gösterim

2.2.2. Doğrusal Olmayan DVM

Doğrusal olmayan DVM yönteminde, değişik yapıdaki çekirdek fonksiyonları maksimum aralıklı hiperdüzleme uygulanarak doğrusal olmayan sınıflandırıcılar elde edilmektedir. Ortaya çıkan algoritma, doğrusal DVM ile benzerdir; ancak, her iç çarpımı doğrusal olmayan bir çekirdek fonksiyonu ile değiştirilmiştir. Böylelikle algoritma, maksimum aralıklı hiperdüzlemin dönüştürülmüş örnek uzayına yerleşmesine izin vermektedir. Söz konusu dönüşüm, doğrusal yapıda olmayabilir ve dönüştürülmüş örnek uzayı ise yüksek boyutlu olabilir. İncelenen sınıflandırıcı dönüştürülmüş örnek uzayında bir hiperdüzlem olmasına rağmen, orijinal girdi uzayında doğrusal olmayabilir (20). Doğrusal olarak sınıflandırılmayan girdi uzayının bir üst boyuta çekirdek fonksiyonu ile haritalanarak doğrusal olarak sınıflandırılması Şekil 2.11’de gösterilmektedir (21).



Şekil 2.11: Doğrusal olarak sınıflandırılmayan girdi uzayının bir üst boyuta çekirdek fonksiyonu ile haritalanması

2.2.3. Çekirdek Fonksiyonları

Bu araştırmada değişik destek vektör makinesi modellerini oluşturmak için aşağıda tanımlanan ANOVA radyal tabanlı fonksiyon, bessel, doğrusal, Gaussian radyal tabanlı fonksiyon, laplace, polinomiyal ve sigmoid çekirdek fonksiyonları kullanılmıştır. Çekirdek fonksiyonları sınıflama ve regresyon problemlerinde bir benzerlik fonksiyonu olarak görev yapmaktadır. Bu fonksiyonlar makine öğrenmesi algoritmalarına çeşitli esnek özellikler kazandırır. Böylelikle sınıflama ve regresyon problemlerinde bu esnek özellikleri nedeni ile daha yüksek performans sonuçları elde edilebilmektedir. Bir örnek ile çekirdek fonksiyonu açıklanacak olursa;

$$X = (x_1, x_2), Y = (y_1, y_2) \text{ ve } k(x, y) = x^T y + c$$

çekirdek fonksiyonu olduğunda,

$$k(x, y) = x^T y + c = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} [y_1 \ y_2] + c = (x_1 y_1 + x_2 y_2) + c = \phi(x)^T \phi(y)$$

olarak elde edilir.

a. Doğrusal çekirdek fonksiyonu

Doğrusal çekirdek, en basit çekirdek fonksiyonudur. $\langle x, y \rangle$ iç çarpımı ve bir c sabiti ile belirtilir. İlgili çekirdek fonksiyonu aşağıdaki formül ile ifade edilir:

$$k(x, y) = x^T y + c$$

b. Polinomial çekirdek fonksiyonu

Polinomial çekirdek fonksiyonu durağan olmayan bir çekirdektir. Polinomial çekirdeği, tüm eğitim verilerinin normalleştirildiği problemler için çok uygundur. İlgili çekirdek fonksiyonu aşağıdaki formül ile ifade edilir:

$$k(x, y) = (\alpha x^T y + c)^d$$

Ayarlanabilir parametrelerden, alfa eğimi, c sabit terimi ve d polinom derecesini ifade etmektedir.

c. Gaussian radyal tabanlı fonksiyon çekirdeği

Gauss çekirdeği radyal taban fonksiyonu çekirdeğine bir örnektir. İlgili çekirdek fonksiyonu aşağıdaki formül ile ifade edilir:

$$k(x, y) = \exp\left(-\frac{\|x-y\|^2}{2\sigma^2}\right)$$

Alternatif olarak,

$$k(x, y) = \exp(-\gamma\|x - y\|^2)$$

formülü ile ifade edilir.

Ayarlanabilir parametre sigma, çekirdeğin performansında önemli bir rol oynar ve eldeki soruna göre dikkatle seçilmelidir. Aşırı tahmin söz konusu olduğunda, üstel ifade yaklaşık olarak doğrusallaşır ve yüksek boyutlu projeksiyon doğrusal olmayan yapısını kaybetmeye başlar. Öte yandan, eğer gerçek değerden daha düşük tahmin gerçekleşiyorsa, fonksiyon düzgülendirmeyecek ve karar sınırı ise eğitim verisindeki gürültü değerlere karşı son derece duyarlı olacaktır.

d. Laplace çekirdek fonksiyonu

Laplace çekirdeği bütünüyle üstel çekirdek fonksiyonuna eşdeğerdir; ancak sigma parametresindeki olası değişikliklere karşı daha az bir duyarlılık söz konusudur. Ayrıca Laplace çekirdek fonksiyonu, radyal tabanlı fonksiyon çekirdeği ile eşdeğerdir. İlgili çekirdek fonksiyonu aşağıdaki formül ile ifade edilir:

$$k(x, y) = \exp\left(-\frac{\|x-y\|}{\sigma}\right)$$

Gauss çekirdeği için sigma parametresi hakkında yapılan gözlemlerin üstel ve Laplace çekirdeği için de geçerli olduğu gözlemlenmiştir.

e. ANOVA radyal tabanlı fonksiyon çekirdek fonksiyonu

ANOVA radyal tabanlı fonksiyon çekirdek fonksiyonu, Gauss ve Laplace çekirdeği gibi radyal tabanlı fonksiyon çekirdeğidir. Çok boyutlu regresyon problemlerinde iyi performans gösterdiği bildirilmektedir (22). İlgili çekirdek fonksiyonu aşağıdaki formül ile ifade edilir:

$$k(x, y) = \sum_{k=1}^n \exp(-\sigma(x^k - y^k)^2)^d$$

f. Hiperbolik Tanjant (Sigmoid) çekirdek fonksiyonu

Hiperbolik Tanjant (Sigmoid) Çekirdeği, Sigmoid Çekirdeği ve Çok Katmanlı Algılayıcı (MLP) çekirdeği olarak da bilinir. Sigmoid Çekirdeği, yapay sinir ağlarında aktivasyon fonksiyonu olarak da kullanılır. İlgili çekirdek fonksiyonu aşağıdaki formül ile ifade edilir:

$$k(x, y) = \tanh(\alpha x^T y + c)$$

Sigmoid çekirdek fonksiyonunu kullanan bir DVM modelinin, iki katmanlı algılayıcı sinir ağıyla eşdeğerdir. Bu çekirdek fonksiyonunun pratikte iyi performans gösterdiği tespit edilmiştir.

Sigmoid çekirdeğinde alfa eğim ve c kesme sabiti olmak üzere iki ayarlanabilir parametre vardır. Alfa için ortak bir değer $1 / N$ 'dir, burada N veri boyutudur. Sigmoid çekirdekler hakkında daha ayrıntılı bilgi Hsuan-Tien ve Chih-Jen'in yayınından (23) elde edilebilir.

g. Bessel çekirdek fonksiyonu

Bessel çekirdeği, aşağıdaki denklem ile ifade edilir:

$$k(x, y) = \frac{J_{v+1}(\sigma \|x-y\|)}{\|x-y\|^{-n(v+1)}}$$

Bu denklemde J birinci tip Bessel fonksiyonudur. Bununla birlikte, kernlab for R kaynağında (24), Bessel çekirdeğinin şu şekilde gösterildiği bildirilmektedir:

$$k(x, x') = -Bessel_{(nu+1)}^n(\sigma |x - x'|^2)$$

2.3 Araştırma Konusu ile İlgili Benzer Çalışmalar

Bir çalışmada (25), akut koroner sendrom tanısı için yapay sinir ağı modellerini geliştirmek ve optimize etmek için farklı merkezlerden ileriye dönük olarak toplanan veriler üzerinde bu modellerin performansları incelenmiş. Çalışmada, akut göğüs ağrısı şikâyeti ile 3 hastaneye başvuran 3.147 hastadan elde edilen 40 değişkene ait veriler kullanılmıştır. İncelenen veri seti, lojistik regresyon ve yapay sinir ağı modellerinin oluşturulması için kullanılmıştır. Modellerinin performansı, işlem karakteristik eğrisi (ROC) analizi kullanılarak değerlendirilmiştir. Çalışmanın sonuçları, yapay sinir ağlarının tanı amaçlı algoritmaları geliştirmede yararlı bir yaklaşım olabileceğini doğrulamış ve yapay sinir ağı modellerinin performansının, uygun şekilde optimize edilmiş lojistik regresyon modellerinden anlamlı ölçüde farklı bulunmadığı bildirilmiştir (25).

Başka bir çalışmada (26), TIMI-DISPERSE2 klinik araştırmasındaki hastalardan toplanan sürekli elektrokardiyografik verilerin zaman serileri analizlerinden biyolojik belirteçleri türetmek amacıyla istenilen bilgileri çıkarmak için tasarlanmış makine öğrenimi ve veri madenciliği yöntemleri kullanılmıştır. Bu biyolojik belirteçleri, akut koroner sendromlu 4500'den fazla hasta üzerinde 15'den fazla bağımsız değişken değerlendirilmiştir. Sonuç olarak, söz konusu biyolojik belirteçlerin, akut koroner sendromlu hastalardan rutin olarak toplanan verilerden elde edilebileceği ve daha doğru risk sınıflaması ile potansiyel olarak daha iyi hasta tedavisine imkân vereceği bildirilmiştir (26).

Diğer bir çalışmada (27), veri madenciliği teknikleri yardımıyla akut koroner sendromunu tahmin edebilen modeller üretmek için farklı hasta verileri analiz edilmiştir. Adı geçen çalışmada, akut koroner sendromu veritabanında yer alan 422 hastanın, kronik koşulları, risk faktörleri, laboratuvar sonuçları gibi 28 değişkene ilişkin veriler üzerinde özellik seçimi ve öğrenme algoritmalarını uygulanmıştır. Çalışmanın deneysel sonuçları, oluşturulan modellerin akut koroner sendromunun tahmini için kullanılabileceğini göstermektedir (27).

Koroner arter hastalığını tahmin edebilmek için değişik sekiz öğrenme algoritması ile farklı yapay sinir ağı modelleri oluşturulan bir çalışmada çok katmanlı “perseptron” yapay sinir ağları, toplam 237 kayıt üzerinde (171’i eğitimde ve 66’sı teste) kullanılarak oluşturulmuştur. Modellerin başarısının değerlendirilmesinde; duyarlılık, özgüllük

[seçicilik] ve doğruluk değerleri kullanılmış ve bu değerler %71.0'den daha yüksek olarak elde edilmiştir. Sonuçta; geriye yayılım algoritması dışındaki diğer öğrenme yöntemleri ile daha büyük örnek büyüklüklerinde elde edilebilecek sınıflama başarısının daha yüksek olabileceği ve incelenen modellerin KAH'ın tahmininde oldukça iyi kestirim sonuçları verdiği bildirilmiştir (28).

Başka bir çalışmada (29), koroner arter hastalığı tanısı için bulanık kurallara dayalı karar destek sistemi oluşturma ve değerlendirme amacıyla kullanılan veri seti, demografik ve geçmiş verilerin yanı sıra laboratuvar tetkikleri de dahil olmak üzere 19 özellik ile karakterize edilen 199 kişiden oluşmaktadır. On kat çapraz doğrulama kullanılarak, karar ağacından çıkarılan kurallar seti yardımıyla elde edilen ortalama duyarlılık ve özgüllük [seçicilik] değerleri sırasıyla; % 62 ve % 54 iken, bulanık mantık ve optimizasyon aşamaları kullanıldığında ise ortalama duyarlılık ve özgüllük [seçicilik] sırasıyla % 80 ve %65'e yükseldiği belirtilmiştir. Önerilen sistemin, koroner arter hastalığı tanısı için faydalı olacağı ve bu tür hastalıkların tahmin edilmesinde kullanılabileceği önerilmiştir (29).

Diğer bir araştırmada (30), akut koroner sendrom semptomları ile acil servise gelen hastalardan alınan triyaj bilgilerinin bir veri kümesini temsil etmek için bir Bayes Ağının oluşturulması amaçlanmıştır. Bu amaçla, triyaj veri seti kullanılarak eğitilen değişik Bayes Ağları, veri setinde temsil edilen hastalar için 30 günlük advers sonuçların tahmininde doğruluk oranları açısından karşılaştırılmıştır. ROC eğrileri kullanılarak yapılan karşılaştırmada, işlenmiş Bayes Ağının (the mined Bayesian Network) diğer oluşturulan ağ modelinin (the domain expert generated network) biraz üstünde performans gösterdiği bildirilmiştir (30).

Başka bir çalışmada (31), 303 koroner arter hastası ve 54 değişkeni olan bir veri seti üzerinde veri madenciliği yöntemleri uygulanmıştır. İncelenen değişkenlerin/özelliklerin, koroner arter hastalığı üzerindeki etkinliğini belirlemek için Bilgi Kazancı ve güven (information gain and confidence) teknikleri kullanılmıştır. İncelenen bu çalışmada, veri madenciliği yöntemleri ve özellik oluşturma algoritmasını kullanılması ile literatürdeki bilinen yaklaşımlardan daha yüksek olan % 94.08 doğruluk oranı elde edilmiştir (31).

2004-2009 yılları arasında AKS'nin Koreli nüfus üzerindeki ekonomik etkisini inceleyen bir çalışmada (32), Kore Ulusal Sağlık Sigortası Kurumunun talep veritabanı,

Güney Koreli nüfusun yıllık sağlık kullanımı ve harcamaları hakkında veri edinmek için kullanılmıştır. Veri madenciliği tekniği kullanılarak, sağlık talepleri AKS ile ilgili hastalık kodlarıyla tespit edilmiş ve masraflar makro maliyetleme yöntemi kullanılarak hesaplanmıştır. Söz konusu çalışmanın sonuçları, AKS'nin Kore toplumuna toplam maliyetinin yüksek olduğunu ve AKS ile ilişkili mortalite ve morbiditenin azaltılması için bu sendromun erken ve etkili bir biçimde yönetilmesi gerektiğini göstermektedir (32).

Yapılan başka bir araştırmada (33), veri madenciliği teknikleri yardımıyla akut koroner sendrom riskini artırmaya önemli katkıda bulunan faktörleri incelemek amaçlanmıştır. Bu kapsamda, Pakistan'ın Karachi kentindeki iki farklı kalp hastanesinden alınan veri setinde bağımlı değişken olan akut koroner sendromu (hastalığın varlığını veya yokluğunu göstermekte) ve bu sendromu etkileyebilecek 15 bağımsız değişkenden önemli faktörlerin belirlenmesi için ikili lojistik regresyon (binary logistic regression) analizi uygulanmıştır. Sonuç olarak, temel bileşenler analizi ile yapılan veri indirgeme sonuçlarına dayanarak, bu faktörlerden 14'ünün akut koroner sendrom riskini artırmaya önemli katkıda bulunduğu bildirilmiştir (33).

Koroner arter hastalığının tahmin edilmesi için lojistik regresyon modeli seçim yöntemlerinin kıyaslanması amaçlanan bir araştırmada (34), KAH verileri Kardiyoloji bölümüne başvuran ardışık 237 bireyin değişik klinik ve demografik bağımsız değişkenlerini içermektedir. Söz konusu araştırmada, her bir model seçim tekniğine ilişkin duyarlılık, özgüllük ve doğruluk değerleri % 91.9'dan daha yüksek olarak elde edilmiştir. Sonuç olarak, Lojistik regresyon modeli seçim teknikleri KAH'ın tahmininde iyi performans göstermiş olup, olabilirlik oran testi bulgularına göre incelenen adimsal tekniklerin Giriş (Enter) yaklaşımından daha iyi kestirimler verdiği belirlenmiştir. Ele alınan bağımsız değişkenler arasından yaş, DM, HT, aile öyküsü, sigara, LDL, trigliserid, stres ve obezite faktörlerinin KAH'ı etkileyen önemli risk faktörlerinden olduğu bildirilmiştir (34).

Başka bir araştırmada (35), hastanın yaşı, cinsiyeti, risk faktörleri ve kardiyak enzimleri gibi hasta verileri yardımıyla AKS'yi teşhis etmek ve makine öğrenme yöntemlerinde biri olan DVM'yi kullanarak hastanın yatışı veya taburculuğu konularında doktora yardımcı olma durumu amaçlanmıştır. Göğüs ağrısı şikayetiyle acil servise başvuruda bulunan 228 hastanın laboratuvar, radyolojik görüntüleme ve klinik verileri incelenmiş ve bu verilere yapay sinir ağları, Naive Bayes, Lojistik Regresyon ve DVM

yöntemleri uygulanmıştır. Doğruluk oranının en yüksek olduğu SVM sonuçları ve sınıflandırmadaki başarısının % 99.13 olduğu raporlanmıştır. Bu çalışmanın sonuçları, makine öğrenmesi teknikleri kullanılarak hızlıca üretilen verilerden çıkarılan bilginin acil servis çalışanlarının karar verme sürecinde yardımcı olabileceğini göstermiştir (35).

Yukarıda sözü edilen çalışmalar, akut koroner sendromu ya da koroner arter hastalığının değişik boyutlarda incelenmesinde veri madenciliği ve makine öğrenmesi teknikleri ele almaktadır. Ancak, yapılan yayın taraması bulgularına göre, akut koroner sendromlu hastalardan oluşan bireylerde tip 2 diyabet mellitusun varlığı ya da yokluğunu tıbbi bilgi keşfi süreci ile inceleyen bir çalışmaya rastlanmamıştır. Söz konusu araştırma bu yönüyle ilgili literatüre önemli bir katkı sağlayacaktır.



3. MATERYAL VE METOT

Bu çalışma, Malatya Klinik Araştırmalar Etik Kurulu'nun 2016/161 protokol numaralı izni ile onaylanmıştır (Ek-1). Bu çalışmada, anabilim dalımız tarafından Kardiyoloji Anabilim Dalı için geliştirilen PHP tabanlı veri giriş, sorgulama, silme, güncelleme, vb. işlemleri yapan yazılım kullanıldı. Bu kapsamda akut koroner sendromlu hastalarda DM'nin varlığı ve yokluğu, aşağıda belirtilen değişkenlere dayalı olarak sınıflandı. Bu değişkenlere ait tanımlayıcı bilgiler Tablo 3.1'de verilmiştir.

Tablo 3.1: Çalışmada kullanılan değişkenlerin tanımları

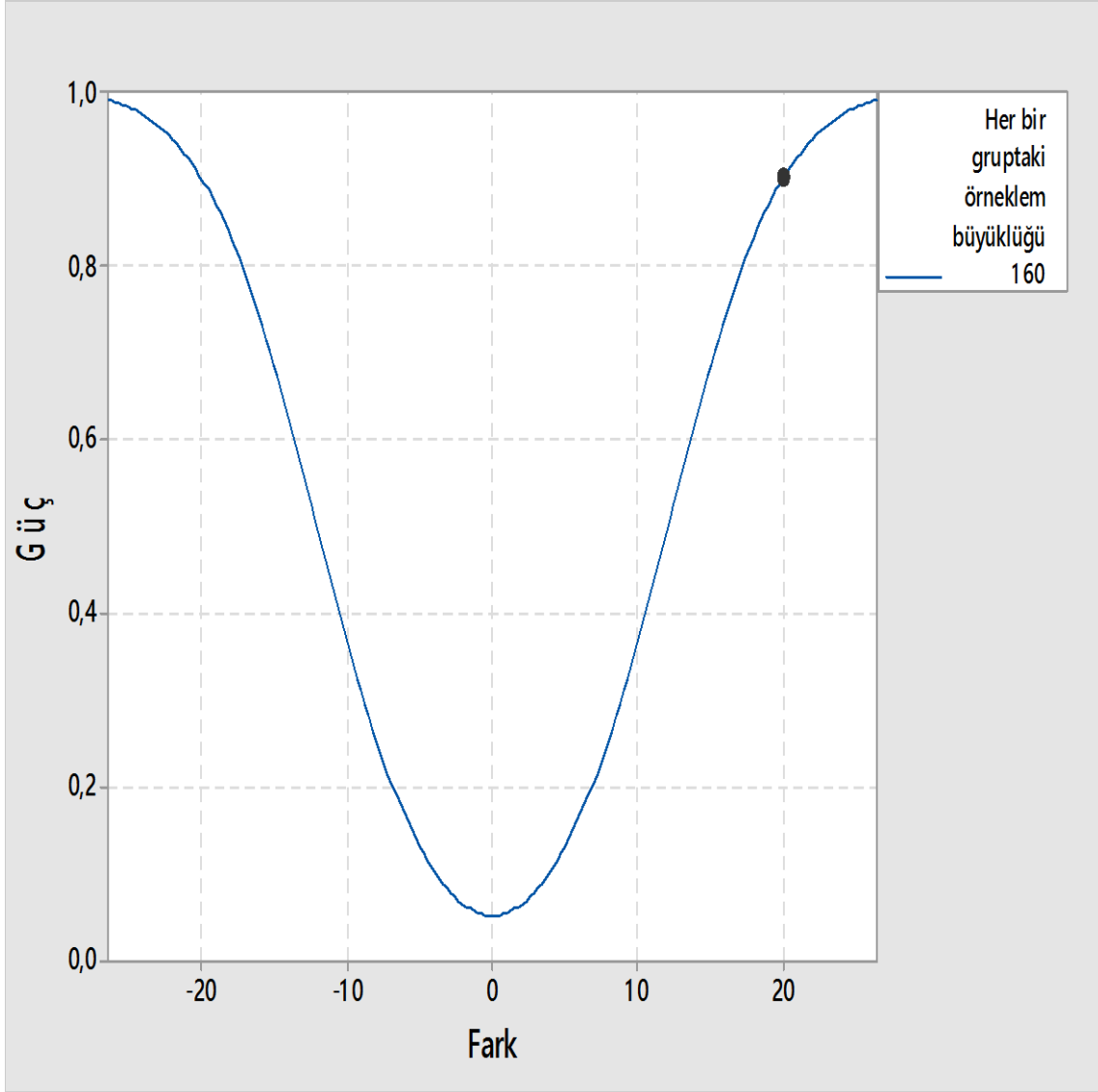
	Değişkenler	Değişken tipi	Değişken Açıklaması	Değişken Rolü
Demografik	Diabetes Mellitus (DM)	Kategorik	Var/Yok	Bağımlı/Hedef
	Akut Koroner Sendrom (AKS) tipi	Kategorik	Unstable Angina/STEMI/NSTEMI	Bağımsız/Tahminleyici
	Yaş	Nümerik	Doğal sayı	Bağımsız/Tahminleyici
	Cinsiyet	Kategorik	Kadın/Erkek	Bağımsız/Tahminleyici
	Vücut Kitle İndeksi (VKİ)	Nümerik	Pozitif reel sayı	Bağımsız/Tahminleyici
Hastalık öyküsü	Hipertansiyon (HT)	Kategorik	Var/Yok	Bağımsız/Tahminleyici
	Sigara içme durumu	Kategorik	Var/Yok	Bağımsız/Tahminleyici
	Renal yetmezlik öyküsü	Kategorik	Var/Yok	Bağımsız/Tahminleyici
	Miyokardiyal Enfarktüs (ME) öyküsü	Kategorik	Var/Yok	Bağımsız/Tahminleyici
	AKS Aile öyküsü	Kategorik	Var/Yok	Bağımsız/Tahminleyici
	Malignite öyküsü	Kategorik	Var/Yok	Bağımsız/Tahminleyici
	Hiperlipidemi öyküsü	Kategorik	Var/Yok	Bağımsız/Tahminleyici
	Periferik arter öyküsü	Kategorik	Var/Yok	Bağımsız/Tahminleyici
	Koroner Arter By-pass öyküsü	Kategorik	Var/Yok	Bağımsız/Tahminleyici
	İnme hastalığı öyküsü	Kategorik	Var/Yok	Bağımsız/Tahminleyici
	Kalp yetmezliği öyküsü	Kategorik	Var/Yok	Bağımsız/Tahminleyici

İlaç öyküsü	Aspirin	Kategorik	Var/Yok	Bağımsız/Tahminleyici	
	Clopidogrel	Kategorik	Var/Yok	Bağımsız/Tahminleyici	
	Beta Bloker	Kategorik	Var/Yok	Bağımsız/Tahminleyici	
	Kalsiyum Kanal Bloker	Kategorik	Var/Yok	Bağımsız/Tahminleyici	
	Statin	Kategorik	Var/Yok	Bağımsız/Tahminleyici	
	Digoksin	Kategorik	Var/Yok	Bağımsız/Tahminleyici	
	Anjiyotensin Dönüştürücü Enzim (ACE) inhibitör	Kategorik	Var/Yok	Bağımsız/Tahminleyici	
	Anjiyotensin Reseptör Bloker (ARB) diğer	Kategorik	Var/Yok	Bağımsız/Tahminleyici	
	Laboratuvar	Kreatinin	Nümerik	Pozitif reel sayı	Bağımsız/Tahminleyici
		BUN	Nümerik	Pozitif reel sayı	Bağımsız/Tahminleyici
Kolesterol		Nümerik	Pozitif reel sayı	Bağımsız/Tahminleyici	
Trigliserid		Nümerik	Pozitif reel sayı	Bağımsız/Tahminleyici	
LDL		Nümerik	Pozitif reel sayı	Bağımsız/Tahminleyici	
HDL		Nümerik	Pozitif reel sayı	Bağımsız/Tahminleyici	
Sistolik Kan Basıncı		Nümerik	Pozitif reel sayı	Bağımsız/Tahminleyici	
Diastolik Kan Basıncı		Nümerik	Pozitif reel sayı	Bağımsız/Tahminleyici	
Diüretik		Kategorik	Var/Yok	Bağımsız/Tahminleyici	
Glikoz		Nümerik	Pozitif reel sayı	Bağımsız/Tahminleyici	

3.1. Örneklem Büyüklüğü

İki grup arası tahmini glikoz düzeyi farkı 20, varsayılan ortak standart sapması 55, tip I hata (alfa) 0.05 ve tip II hata (beta) 0.10 olduğunda, her grupta en az 160 olmak üzere toplamda en az 320 birey gerektiği güç analizi ile hesaplandı (36). Güç analizi hesabına ilişkin gösterim Şekil 3.1’de verilmiştir. Ayrıca, çok değişkenli istatistiksel yöntemlerde önerilen örneklem büyüklüğü hesabında; gözlem sayısı (n), bağımsız/tahminleyici değişken/özellik sayısının en 5 katı ve ideal olanın ise bağımsız/tahminleyici değişken/özellik sayısının 20 katı kadar olması gerektiği bildirilmektedir. Kestirim amacıyla yapılacak modelleme çalışmalarında ise bu oranın daha da fazla olması (örneğin 40 kat) önerilmektedir. Başka bir kaynakta, çok değişkenli bir istatistik model denkleminde bağımsız/tahminleyici değişken/özellik sayısı 6 ya da daha büyük olduğunda, her bağımsız/tahminleyici değişken/özellik için 10 birey gerektiği ve $n > 104 + k$ (k: bağımsız/tahminleyici değişken/özellik sayısı) eşitliğinin kullanılabileceği bildirilmektedir (37, 38). Bu çalışmada 33 bağımsız/tahminleyici değişken/özellik

dikkate alındığında, yukarıda sözü edilen arařtırmalarda önerilen olası örneklem büyüklükleri; 330 (33x10), 660 (33x20), 1320 (33x40) ya da 137 (104 + 33) olarak hesaplanmaktadır. Bu planlanan arařtırma kapsamına 1378 bireye ilişkin veriler alınmıştır.



Şekil 3.1: Güç analizi hesabına ilişkin gösterim

Bu çalışmada uygulanan VTBK süreci aşamaları aşağıda verilmiştir:

- Veri seçimi:** Veritabanlarından elde edilen ve bu arařtırmada bağımlı/hedef deęişken olarak DM, bağımsız/tahminleyici deęişkenler olarak ise yukarıdaki tabloda detaylıca açıklanan faktörler veri seçimi kapsamında incelenmiştir.
- Veri önışleme:** Veri setindeki kayıp deęer bulunan sınıflar verisetinden çıkarılmıştır. Aşırı/aykırı deęer tespiti lokal aykırı faktörü (LOF) analizi ile yapılmış ve tespit edilen aşırı/aykırı deęerler verisetinden çıkarılmıştır.

Aşırı/aykırı değer tespitinde, yerel aykırı faktör (LOF) algoritması (39) kullanıldı. LOF, yakın zamanda geliştirilmiş olan yoğunluğa dayalı aşırı/aykırı gözlem tespitinde kullanılan yöntemlerden biridir. LOF, diğer aşırı/aykırı gözlem saptama algoritmalarıyla karşılaştırıldığında birçok avantaja sahiptir. LOF, bir gözlemin diğer komşu nesnelere olan uzaklığını bir sayısal ölçüğe dönüştürerek, yerel aşırı/aykırı gözlemlerin tespit edilmesini sağlar (40). Bu algoritma aşırı/aykırı tespitinde yaygın olarak kullanılan güçlü bir tekniktir. LOF algoritmasında, veri setini oluşturan tüm kayıtlara, ilgili kaydın aşırı/aykırı değer derecesini tayin bir faktör değeri atanır ve bu faktör değeri ile aşırı/aykırı değer tespiti yapılır.

- c. Veri dönüşümü ve veri indirgeme:** Veri setindeki sayısal değişkenler için standardizasyon (Z-dönüşümü) uygulanmıştır. Standardizasyona ilişkin denklem aşağıda verilmiştir:

$$x_{yeni} = \frac{x - \bar{x}}{s}$$

Burada \bar{x} ; örneklem aritmetik ortalamasını ve s ; örneklem standart sapmasını ifade etmektedir.

Ayrıca, değişken ölçeklendirmede kullanılabilen diğer bir yöntem ise normalizasyon tekniğidir. Bu yöntem aşağıdaki denklem ile tanımlanabilir.

$$x_{yeni} = \frac{x - x_{en\ küçük}}{x_{en\ büyük} - x_{en\ küçük}}$$

Bu tekniklerin her ikisinin de belirli dezavantajları vardır. Eğer veri kümesinde aykırı değerleriniz varsa, veriler normalize edilerek çok küçük bir aralıkla ölçeklendirilebilir. Standardizasyon yöntemi kullanırken, yeni veriler için bir sınır söz konusu değildir (41).

- d. Veri madenciliği:** Çeşitli çekirdek fonksiyonları kullanılarak oluşturulan DVM modelleri kullanılarak veri setinden ilişki, örüntüler çıkarma ve tahminler yapılmıştır. Bu adıma ilişkin detaylı bilgiler, “Oluşturulan Destek Vektör Makinesi Modelleri” alt başlığında verilmiştir.
- e. Değerlendirme ve yorumlama:** Çeşitli çekirdek fonksiyonları kullanılarak oluşturulan DVM modellerinin tahmin performansları; 10-katlı çapraz geçerlilik tekniği ile değerlendirilmiştir. Ayrıca, doğruluk, duyarlılık, özgüllük [seçicilik], ROC eğrisi altında kalan alan gibi performans metrikleri kullanılarak ilgili model çıktıları yorumlanmıştır. Bu metriklerle ilişkin ayrıntılı formül tanımlamaları önceki bölümlerde bahsedilmiştir.

3.2. Geliştirilen Web Tabanlı Yazılım

Bu yazılımda, PHP (Personal Home Page) programlama dili kullanılmıştır PHP, sunucu üzerinde çalışan ve HTML kodları içine yerleştirme temelli betik (script) bir dildir. PHP kodlarının çalışması için bir derleyici gereksinimi yoktur. Kodların yazılması için istenilen yazı editörü tercih edilebilir. Web sayfalarında hazırladığımız HTML kodları özel olarak derlenmediği sürece sabit çıktılar verirler. Bu yüzden düz bir HTML koduyla yapılamayacak şeyler vardır. HTML kodlarını kullanarak bir web sunucusunda bulunan yazı dosyası üzerinde okuma ve yazma işlemini yapamayız, herhangi bir veritabanı yönetim sistemi ile bağlantı kuramayız. Bu gibi işlemler için HTML kodları içerisine yerleştireceğimiz betiklere gereksinim vardır (42, 43).

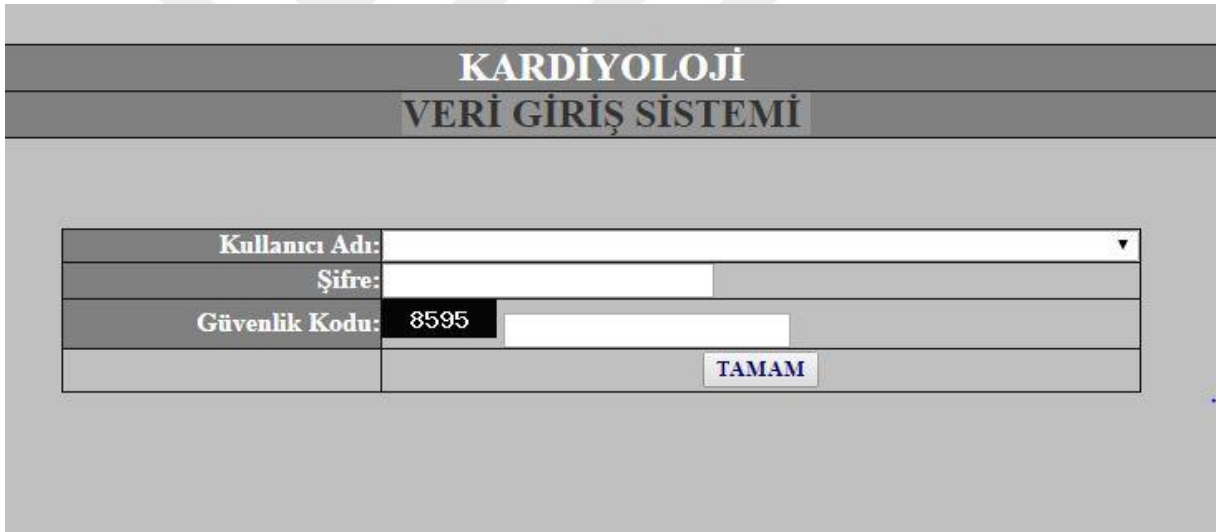
PHP'yi HTML kodları arasına yerleştirdiğimiz program betikleri olarak düşünebiliriz. Şayet web sunucumuzda bu program parçacıklarını anlayıp yorumlama yeteneği kazandırırsak bu sayede HTML dosyalarımızı istenildiği şekilde programlayabiliriz. PHP desteği sağlayacak programı web sunucularımıza kurarak, PHP dosyalarının algılanmasını ve yorumlanmasını sağlarız. PHP'nin diğer betik dillerine göre en önemli farklılığı sunucu taraflı yorumlanıyor olmasıdır. Bu sayede web sunucusu tarafından PHP kodları yorumlanıp, düzenlenip istemcinin tarayıcısına gönderilir (42).

PHP kolaylıkla öğrenilebilen ve kodlanabilen bir dildir. Java, C ve Perl dillerine benzerlik göstermesinin yanında bu dillerden farklı birtakım özellikler katılarak oluşturulmuştur. Bu web arayüzü geliştirilirken PHP tercih etmemizin sebeplerini şöyle sıralayabiliriz; hızlı çalışması, dilin rahatça öğrenilmesi, kolay ve hızlı kodlanması, kararlı oluşu ve platform bağımsız çalışması, MySQL veritabanı yönetim sistemi ile son derece başarılı ve yüksek performanslı sonuçlar elde edilmesidir (42, 43).

Ayrıca bu web tabanlı yazılımda MySQL Veri Tabanı Yönetim Sistemi (VTYS) kullanılmıştır. Çok kanallı ve kullanıcı, yüksek hızlı ve güvenilir bir yapıya sahiptir. MySQL PHP, Python, Java vb. programlama dilleriyle erişilebilir. APACHE sunucu programı ve PHP beraber web-veritabanı uygulamalarında sıklıkla kullanılmaktadır. Birden çok işlemci ile rahatça çalışabilmektedir. Veritabanı tablolarının denetimi, optimizasyonu ve onarımı hızlı bir şekilde yapılabilir. MySQL VTYS'yi geliştirenler 10.000 tabloluk veritabanlar ile yaklaşık olarak 100 GB'lık veriyi MySQL'de saklayabilmekteler. Bilhassa internet ortamında değer kazanan güçlü bir kullanıcı erişim ve yetkilendirme yapısına sahiptir. MySQL VTYS çok erişimli sistemlerde ortaya çıkan

yetkilendirme sorununu çözmek için yanı sıra stored procedure ve trigger özelliklerini desteklemektedir. Görünüm (view) özelliği olmamasına rağmen MySQL VTYS'nin gelişmiş erişim yetkilendirme sistemi olduğundan bu özelliğe ihtiyaç duyulmamaktadır. Diğer taraftan görsel bütünlük oluşturma işi yazılımcının tercihine bırakılmıştır. Fakat bunun bir dezavantaj olduğu düşünülmemelidir. Çünkü çoğu veritabanı yazılımcısı VTYS'deki görsel bütünlüğün zorlayıcı ve esnek olmadığını düşünmektedir. MySQL VTYS çeşitli tablo biçimleri ve işleme çeşitliliğiyle esnek bir yapı sağlar. Hızın ve kolaylığın öne çıktığı projelerde, MySQL VTYS uygun bir tercihtir. Fakat tablo sayısı çok ve karmaşıksa geleneksel VTYS sunucularındaki ileri özelliklere talep duyulacaktır (43-45).

Web tabanlı kardiyoloji veri giriş sisteminin giriş ekranı Şekil 3.2 deki gibi tasarlanmıştır. Bu ekran, verisi girilecek hastanenin kullanıcılarının kimlik doğrulama yaptığı ara yüzdür.



KARDİYOLOJİ VERİ GİRİŞ SİSTEMİ	
Kullanıcı Adı:	<input type="text"/>
Şifre:	<input type="password"/>
Güvenlik Kodu:	8595 <input type="text"/>
<input type="button" value="TAMAM"/>	

Şekil 3.2: Web tabanlı kardiyoloji veri giriş sisteminin giriş ekranı

Hasta ile ilgili demografik verilerin kaydedildiği ara yüz Şekil 3.3'de yer almaktadır. Hastaneye başvuru tarihi, adı, soyadı, TC kimlik numarası, telefonu, cinsiyeti, doğum yılı, sosyo ekonomik düzeyi, okuma yazma durumu, mesleği, boyu, kilosu, tıbbi öyküleri ve ilaç öyküleri ile ilgili gerekli bilgiler bu ekran kullanılarak kayıt altına alınır.

DEMOGRAFİK	
Başvuru Tarihi	▼ ▼ ▼
Ad	
Soyad	
Tc Kimlik No	
Telefon	
Cinsiyet	Kız <input type="radio"/> Erkek <input type="radio"/>
Doğum Yılı	▼
Sosyo Ekonomik Düzey	▼
Okuma Yazma Durumu	▼
Meslek	▼
Boy	▼
Kilo	▼
Tıbbi Öykü	DM <input type="checkbox"/>
	HT <input type="checkbox"/>
	Sigara <input type="checkbox"/>
	Renal Yetmezlik <input type="checkbox"/>
	MI Öyküsü <input type="checkbox"/>
	PAH <input type="checkbox"/>
	CABG öyküsü <input type="checkbox"/>
	Hiperlipidemi <input type="checkbox"/>
	Stroke öyküsü <input type="checkbox"/>
	Maliğnite <input type="checkbox"/>
	Kalp yetmezliği öyküsü <input type="checkbox"/>
	KAH öyküsü (\geq % 50 darlık) <input type="checkbox"/>
	\geq 24 saat 2 den fazla anjinal atak <input type="checkbox"/>
	Aile Öyküsü <input type="checkbox"/>
İlaç Öyküsü	ASA <input type="checkbox"/>
	Clopidogrel <input type="checkbox"/>
	Beta Bloker <input type="checkbox"/>
	Kalsiyum Kanal Bloker <input type="checkbox"/>
	Statin <input type="checkbox"/>
	Digoksin <input type="checkbox"/>
	ACE inh <input type="checkbox"/>
	ARB Diğer <input type="checkbox"/>
	Diüretik <input type="checkbox"/>

Şekil 3.3: Hasta ile ilgili demografik verilerin kaydedildiği ara yüz

Hastaneye başvuran hastanın tansiyon, nabız, laboratuvar parametreleri gibi verilerinin kaydedildiği ekran Şekil 3.4’de gösterilmiştir.

BAŞVURU		
Tansiyon	Sistolik <input type="text"/> Diastolik <input type="text"/>	
Nabız	<input type="text"/>	
Killip Sınıf	<input type="text"/>	
Akut Koroner Sendrom Tipi	<input type="text"/>	
EKG	<input type="text"/>	
Kardiyak Enzimler	Total CK	Baseline <input type="text"/> Maksimum <input type="text"/>
	CK MB	Baseline <input type="text"/> Maksimum <input type="text"/>
	Troponin	Baseline <input type="text"/> Maksimum <input type="text"/>
	Glukoz	<input type="text"/>
	Kreatinin	<input type="text"/>
	BUN	<input type="text"/>
Laboratuvar Parametreleri	T.Kolesterol	<input type="text"/>
	Trigliserid	<input type="text"/>
	LDL	<input type="text"/>
	HDL	<input type="text"/>

Şekil 3.4: Hastaneye başvuran hastanın tansiyon, nabız, laboratuvar parametreleri gibi verilerinin kaydedildiği ekran

MySQL VTYS’de kardiyoloji adında veritabanı tanımlanmış ve “*hasta_bilgi_demografik*”, “*hasta_bilgi_basvuru*”, “*kurumlar*” ve “*kullanıcılar*” tabloları oluşturulmuştur.

“*hasta_bilgi_demografik*” tablosuna hasta ile ilgili demografik verilerin kaydedilmesi sağlanmıştır. Tablolardan “*hasta_bilgi_demografik*” e yönelik yapı Şekil 3.5’de verilmiştir.

Sunucu: localhost Veritabanı: kardiyoloji Tablo: hasta_bilgi_demografik

Gözet Yapı SQL Ara Ekle Dışarı Aktar İçeri Aktar İşlemler Boşalt Kaldır

Alan	Türü	Karşılaştırma	Öznitelikler	Boş	Varsayılan	Ekstra	Eylem
<input type="checkbox"/> hasta_id	int(11)			Hayır	None		
<input type="checkbox"/> il_id	tinyint(4)			Hayır	None		
<input type="checkbox"/> kurum_id	tinyint(4)			Hayır	None		
<input type="checkbox"/> kullanıcı_id	smallint(6)			Hayır	None		
<input type="checkbox"/> hasta_basvuru_tarihi	date			Evet	NULL		
<input type="checkbox"/> hasta_ad	varchar(30)	utf8_turkish_ci		Evet	NULL		
<input type="checkbox"/> hasta_soyad	varchar(30)	utf8_turkish_ci		Evet	NULL		
<input type="checkbox"/> hasta_tc_kimlik_no	varchar(11)	utf8_turkish_ci		Evet	NULL		
<input type="checkbox"/> hasta_telefon	varchar(30)	utf8_turkish_ci		Evet	NULL		
<input type="checkbox"/> hasta_cinsiyet	char(1)	utf8_turkish_ci		Evet	NULL		
<input type="checkbox"/> hasta_dogum_yili	char(4)	utf8_turkish_ci		Evet	NULL		
<input type="checkbox"/> hasta_sosyo_ekonomik_duzey	char(1)	utf8_turkish_ci		Evet	NULL		
<input type="checkbox"/> hasta_okuma_yazma_durumu	char(1)	utf8_turkish_ci		Evet	NULL		
<input type="checkbox"/> hasta_meslek	char(1)	utf8_turkish_ci		Evet	NULL		
<input type="checkbox"/> hasta_boy	char(3)	utf8_turkish_ci		Evet	NULL		
<input type="checkbox"/> hasta_kilo	char(3)	utf8_turkish_ci		Evet	NULL		
<input type="checkbox"/> tibbi_oyku_dm	char(1)	utf8_turkish_ci		Evet	NULL		
<input type="checkbox"/> tibbi_oyku_insulin	char(1)	utf8_turkish_ci		Evet	NULL		
<input type="checkbox"/> tibbi_oyku_ht	char(1)	utf8_turkish_ci		Evet	NULL		
<input type="checkbox"/> tibbi_oyku_sigara	char(1)	utf8_turkish_ci		Evet	NULL		
<input type="checkbox"/> tibbi_oyku_renal_yetmezlik	char(1)	utf8_turkish_ci		Evet	NULL		
<input type="checkbox"/> tibbi_oyku_diyaliz	char(1)	utf8_turkish_ci		Evet	NULL		
<input type="checkbox"/> tibbi_oyku_mi_oykusu	char(1)	utf8_turkish_ci		Evet	NULL		
<input type="checkbox"/> tibbi_oyku_pah	char(1)	utf8_turkish_ci		Evet	NULL		
<input type="checkbox"/> tibbi_oyku_cabg_oykusu	char(1)	utf8_turkish_ci		Evet	NULL		
<input type="checkbox"/> tibbi_oyku_hiperlipidemi	char(1)	utf8_turkish_ci		Evet	NULL		
<input type="checkbox"/> tibbi_oyku_stroke_oykusu	char(1)	utf8_turkish_ci		Evet	NULL		
<input type="checkbox"/> tibbi_oyku_malignite	char(1)	utf8_turkish_ci		Evet	NULL		
<input type="checkbox"/> tibbi_oyku_kalp_yetmezligi	char(1)	utf8_turkish_ci		Hayır	None		
<input type="checkbox"/> tibbi_oyku_kah	char(1)	utf8_turkish_ci		Hayır	None		
<input type="checkbox"/> tibbi_oyku_kalp_anjinal_atak	char(1)	utf8_turkish_ci		Hayır	None		
<input type="checkbox"/> tibbi_oyku_aile	char(1)	utf8_turkish_ci		Hayır	None		
<input type="checkbox"/> ilac_oykusu_asa	char(1)	utf8_turkish_ci		Evet	NULL		
<input type="checkbox"/> ilac_oykusu_asa_son_yedi_gun	char(1)	utf8_turkish_ci		Hayır	None		
<input type="checkbox"/> ilac_oykusu_clopiyogrel	char(1)	utf8_turkish_ci		Evet	NULL		
<input type="checkbox"/> ilac_oykusu_B_blk	char(1)	utf8_turkish_ci		Evet	NULL		
<input type="checkbox"/> ilac_oykusu_ca_kblk	char(1)	utf8_turkish_ci		Evet	NULL		
<input type="checkbox"/> ilac_oykusu_statin	char(1)	utf8_turkish_ci		Evet	NULL		
<input type="checkbox"/> ilac_oykusu_statin_turu	char(1)	utf8_turkish_ci		Hayır	None		
<input type="checkbox"/> ilac_oykusu_digoksin	char(1)	utf8_turkish_ci		Evet	NULL		
<input type="checkbox"/> ilac_oykusu_ace_inh	char(1)	utf8_turkish_ci		Evet	NULL		
<input type="checkbox"/> ilac_oykusu_arb_diger	char(1)	utf8_turkish_ci		Evet	NULL		
<input type="checkbox"/> ilac_oykusu_diuretik	char(1)	utf8_turkish_ci		Evet	NULL		

Şekil 3.5: "hasta_bilgi_demografik" tablosunun yapısı

Oluşturulan "hasta_bilgi_demografik" tablosundaki "hasta_id" alanı hasta numarası; "il_id" alanı hastanın programa kaydedildiği il plakası; "kurum_id" alanı programı kullanan hastanenin numarası; "kullanıcı_id" alanı hasta numarası; "hasta_basvuru_tarihi" alanı hastanın hastaneye başvuru tarihi; "hasta_ad" alanı hastanın ad; "hasta_soyad" alanı hastanın soyad; "hasta_tc_kimlik_no" alanı hastanın TC kimlik numarası; "hasta_telefon" alanı hastanın telefonu; "hasta_cinsiyet" alanı hastanın cinsiyet (Kadın, Erkek); "hasta_dogum_yili" alanı hastanın doğduğu yıl; "hasta_sosyo_ekonomik_duzey" alanı hastanın ekonomik düzey (Kötü, Orta, İyi);

"hasta_okuma_yazma_durumu" alanı hastanın okuma yazma durumu (Yok, İlköğretim, Lise, Lisans, Yüksek Lisans); "hasta_meslek" alanı hastanın meslek (İssiz, İşçi, Memur, Serbest Meslek, Emekli, Ev Hanımı); "hasta_boy" alanı hastanın boy uzunluğu; "hasta_kilo" alanı hastanın kilo; "tibbi_oyku_dm" alanı hastanın diabetüs melitüs öyküsü olup olmadığı; "tibbi_oyku_insulin" alanı hastanın insülin öyküsü olup olmadığı; "tibbi_oyku_ht" alanı hastanın hipertansiyon öyküsü olup olmadığı; "tibbi_oyku_sigara" alanı hastanın sigara kullanıp kullanmadığı; "tibbi_oyku_renal_yetmezlik" alanı hastanın renal yetmezlik öyküsü olup olmadığı; "tibbi_oyku_diyaliz" alanı hastanın diyaliz öyküsü olup olmadığı; "tibbi_oyku_mi_oykusu" alanı hastanın miyokard infarktüs öyküsü olup olmadığı; "tibbi_oyku_pah" alanı hastanın periferik damar hastalığı öyküsü olup olmadığı; "tibbi_oyku_cabg_oykusu" alanı hastanın koroner by-pass öyküsü olup olmadığı; "tibbi_oyku_hiperlipidemi" alanı hastanın hiperlipidemi öyküsü olup olmadığı; "tibbi_oyku_stroke_oykusu" alanı hastanın inme öyküsü olup olmadığı; "tibbi_oyku_malignite" alanı hastanın kanser öyküsü olup olmadığı; "tibbi_oyku_kalp_yetmezligi" alanı hastanın kalp yetmezliği öyküsü olup olmadığı; "tibbi_oyku_kah" alanı hastanın koroner arter öyküsü olup olmadığı; "tibbi_oyku_kalp_anjinal_atak" alanı hastanın kalp anjinal atak öyküsü olup olmadığı; "tibbi_oyku_aile" alanı hastanın aile öyküsü olup olmadığı; "ilac_oykusu_asa" alanı hastanın aspirin kullanıp kullanmadığı; "ilac_oykusu_asa_son_yedi_gun" alanı hastanın son 7 günde aspirin kullanıp kullanmadığı; "ilac_oykusu_clopidogrel" alanı hastanın clopidogrel kullanıp kullanmadığı; "ilac_oykusu_B_blk" alanı hastanın beta bloker kullanıp kullanmadığı; "ilac_oykusu_ca_kblk" alanı hastanın kalsiyum kanal bloker kullanıp kullanmadığı; "ilac_oykusu_statin" alanı hastanın statin kullanıp kullanmadığı; "ilac_oykusu_statin_turu" alanı hastanın kullandığı statin türünü (Atorvastatin, Rosuvastatin, Fluvastatin, Pravastatin); "ilac_oykusu_digoksin" alanı hastanın digoksin kullanıp kullanmadığı; "ilac_oykusu_ace_inh" alanı hastanın ace inhibitörü kullanıp kullanmadığı; "ilac_oykusu_arb_diger" alanı hastanın anjiotensin reseptör bloker diğer kullanıp kullanmadığı; "ilac_oykusu_diuretik" alanı hastanın diüretik kullanıp kullanmadığı bilgisini tutar.

Hastanın hastaneye başvurusunda alınan bilgiler "*hasta_bilgi_basvuru*" tablosuna kaydedilmiştir. "*hasta_bilgi_basvuru*" tablosunun yapısı Şekil 3.6'da verilmiştir.



The screenshot shows the MySQL database interface for the 'hasta_bilgi_basvuru' table. The table structure is as follows:

Field	Type	Collation	Attributes	Null	Default	Extra	Action
hasta_id	int(11)			No	None		
il_id	tinyint(4)			No	None		
kurumu_id	tinyint(11)			No	None		
kullanici_id	smallint(6)			No	None		
basvuru_tansiyon_sistolik	char(3)	utf8_turkish_ci		Yes	NULL		
basvuru_tansiyon_diastolik	char(3)	utf8_turkish_ci		Yes	NULL		
basvuru_nabiz	char(3)	utf8_turkish_ci		Yes	NULL		
basvuru_cillip_sinif	char(1)	utf8_turkish_ci		Yes	NULL		
basvuru_aks_tipi	char(1)	utf8_turkish_ci		Yes	NULL		
basvuru_ekg	char(1)	utf8_turkish_ci		Yes	NULL		
basvuru_total_ck_baseline	varchar(6)	utf8_turkish_ci		Yes	NULL		
basvuru_total_ck_maksimum	varchar(6)	utf8_turkish_ci		Yes	NULL		
basvuru_ck_mb_baseline	varchar(6)	utf8_turkish_ci		Yes	NULL		
basvuru_ck_mb_maksimum	varchar(6)	utf8_turkish_ci		Yes	NULL		
basvuru_troponin_baseline	varchar(6)	utf8_turkish_ci		Yes	NULL		
basvuru_troponin_maksimum	varchar(6)	utf8_turkish_ci		Yes	NULL		
basvuru_laboratuvar_glukoz	varchar(6)	utf8_turkish_ci		Yes	NULL		
basvuru_laboratuvar_kreatinin	varchar(6)	utf8_turkish_ci		Yes	NULL		
basvuru_laboratuvar_bun	varchar(6)	utf8_turkish_ci		Yes	NULL		
basvuru_laboratuvar_kolesterol	varchar(6)	utf8_turkish_ci		Yes	NULL		
basvuru_laboratuvar_trigliserid	varchar(6)	utf8_turkish_ci		Yes	NULL		
basvuru_laboratuvar_idf	varchar(6)	utf8_turkish_ci		Yes	NULL		
basvuru_laboratuvar_hdl	varchar(6)	utf8_turkish_ci		Yes	NULL		

Şekil 3.6: "*hasta_bilgi_basvuru*" tablosunun yapısı

"*hasta_bilgi_basvuru*" tablosundaki "hasta_id" alanı hasta numarası; "il_id" alanı hastanın programa kaydedildiği il plakası; "kurumu_id" alanı programı kullanan hastanenin numarası; "kullanici_id" alanı hasta numarası, "basvuru_tansiyon_sistolik" alanı hastanın hastaneye başvurduğu zamanki sistolik tansiyon; "basvuru_tansiyon_diastolik" alanı hastanın hastaneye başvurduğu zamanki diastolik tansiyon; "basvuru_nabiz" alanı hastanın hastaneye başvurduğu zamanki nabız; "basvuru_cillip_sinif" alanı hastanın hastaneye başvurduğu zamanki killip sınıf tipi (1,2,3,4); "basvuru_aks_tipi" alanı hastanın hastaneye başvurduğu zamanki akut korner sendrom tipi (Unstabl Anjina, STEMI, NSTEMI); "basvuru_ekg" alanı hastanın hastaneye başvurduğu zamanki ekg türü (ST segment evelasyonu, ST segment depresyonu, T negatifliği, LBBB, Tanımlanamayan, ST-T değişiklik yok); "basvuru_total_ck_baseline" alanı hastanın hastaneye başvurduğu zamanki Total CK kardiyak enzim baseline; "basvuru_total_ck_maksimum" alanı hastanın hastaneye başvurduğu zamanki Total Ck kardiyak enzim maksimum; "basvuru_ck_mb_baseline" alanı hastanın hastaneye başvurduğu zamanki CK MB kardiyak enzim baseline; "basvuru_ck_mb_maksimum" alanı hastanın hastaneye başvurduğu zamanki CK MB kardiyak enzim baseline; "basvuru_troponin_baseline" alanı hastanın hastaneye başvurduğu zamanki troponin baseline; "basvuru_troponin_maksimum" alanı hastanın

hastaneye başvurduğu zamanki troponin maksimum; "basvuru_laboratuvar_glikoz" alanı hastanın hastaneye başvurduğu zamanki glikoz; "basvuru_laboratuvar_kreatinin" alanı hastanın hastaneye başvurduğu zamanki kreatinin; "basvuru_laboratuvar_bun" alanı hastanın hastaneye başvurduğu zamanki bun; "basvuru_laboratuvar_kolesterol" alanı hastanın hastaneye başvurduğu zamanki kolesterol; "basvuru_laboratuvar_trigliserid" alanı hastanın hastaneye başvurduğu zamanki trigliserid; "basvuru_laboratuvar_ldl" alanı hastanın hastaneye başvurduğu zamanki LDL; "basvuru_laboratuvar_hdl" alanı hastanın hastaneye başvurduğu zamanki HDL bilgisini tutar.

Bu web tabanlı programın diğer illerdeki hastanelerde de kullanılması için “kurumlar” adlı tablo oluşturulmuştur. “kurumlar” tablosunun yapısı Şekil 3.7’de verilmiştir. “kurumlar” tablosunda yer alan hastaneler Şekil 3.8’de verilmiştir.

Alan	Türü	Karşılaştırma	Öznitelikler	Boş	Varsayılan	Ekstra	Eylem
<input type="checkbox"/> kurum_id	smallint(6)			Hayır	None	auto_increment	
<input type="checkbox"/> kurum_adi	varchar(255)	utf8_turkish_ci		Hayır	None		
<input type="checkbox"/> kurum_il	tinyint(4)			Hayır	None		

Şekil 3.7: “kurumlar” tablosunun yapısı

	kurum_id	kurum_adi	kurum_il
<input type="checkbox"/>	1	Turgut Özal Tıp Merkezi	44
<input type="checkbox"/>	2	Devlet Hastanesi	45
<input type="checkbox"/>	3	Beydağı Devlet Hastanesi	44
<input type="checkbox"/>	4	Elazığ Eğitim ve Araştırma	23
<input type="checkbox"/>	5	Fırat Üniversitesi Tıp Fakültesi	23
<input type="checkbox"/>	6	Harran Üniversitesi Tıp Fakültesi	63
<input type="checkbox"/>	7	Dicle Üniversitesi Hastanesi	21
<input type="checkbox"/>	8	Başkent Üniversitesi Ankara Hastanesi	6
<input type="checkbox"/>	9	Atatürk Üniversitesi Tıp Fakültesi	25
<input type="checkbox"/>	10	Başkent Hastanesi	1
<input type="checkbox"/>	11	Trabzon Ahi Evren Gögüs Kalp ve Damar Cerrahisi Ha...	61
<input type="checkbox"/>	12	Karadeniz Teknik Üniversitesi	61
<input type="checkbox"/>	13	Eğitim ve Araştırma Hastanesi	61
<input type="checkbox"/>	14	Ondokuz Mayıs Üniversitesi	55
<input type="checkbox"/>	15	Koşuyolu Hastanesi	34
<input type="checkbox"/>	16	Mehmet Akif Ersoy	34
<input type="checkbox"/>	17	Sütçü İmam Üniversitesi	46
<input type="checkbox"/>	18	Hacettepe Üniversitesi	6
<input type="checkbox"/>	19	Erciyes Üniversitesi	38
<input type="checkbox"/>	20	Adana Eğitim Hastanesi	1
<input type="checkbox"/>	21	Başkent Hastanesi	42
<input type="checkbox"/>	22	Selçuk Meram Üniversitesi	42
<input type="checkbox"/>	23	Cumhuriyet Üniversitesi	58
<input type="checkbox"/>	24	Çukurova Üniversitesi	1
<input type="checkbox"/>	25	Adana Numune Eğitim Araştırma Hastanesi	1
<input type="checkbox"/>	26	Uludağ Üniversitesi	16
<input type="checkbox"/>	27	Bursa Yük. İht. EAH 1.Kard. Kliniği	16
<input type="checkbox"/>	28	Bursa Yük. İht. EAH 2.Kard. Kliniği	16
<input type="checkbox"/>	29	Denizli Devlet Hastanesi	20
<input type="checkbox"/>	30	Diyarbakır Eğitim Araştırma Hastanesi	21

Şekil 3.8: Web tabanlı programı kullanan hastaneler

Bu web tabanlı programa kullanıcıların şifreleri ile girişlerinin sağlanması için gerekli bilgiler “*kullanici*lar” tablosuna kaydedilmiştir. “*kullanici*lar” tablosunun yapısı Şekil 3.9’da verilmiştir.

Alan	Türü	Karşılaştırma	Öznitelikler	Boş	Varsayılan	Ekstra	Eylem
<input type="checkbox"/> kullanici_id	smallint(6)			Hayır	None	auto_increment	[Icons]
<input type="checkbox"/> kurum_id	smallint(6)			Evet	NULL		[Icons]
<input type="checkbox"/> ad	varchar(50)	utf8_turkish_ci		Hayır	None		[Icons]
<input type="checkbox"/> soyad	varchar(50)	utf8_turkish_ci		Hayır	None		[Icons]
<input type="checkbox"/> kullanici_adi	varchar(50)	utf8_turkish_ci		Evet	NULL		[Icons]
<input type="checkbox"/> sifre	varchar(50)	utf8_turkish_ci		Evet	NULL		[Icons]
<input type="checkbox"/> sifre_geceriligi	smallint(6)			Evet	NULL		[Icons]
<input type="checkbox"/> kullanici_statu	smallint(6)			Evet	NULL		[Icons]

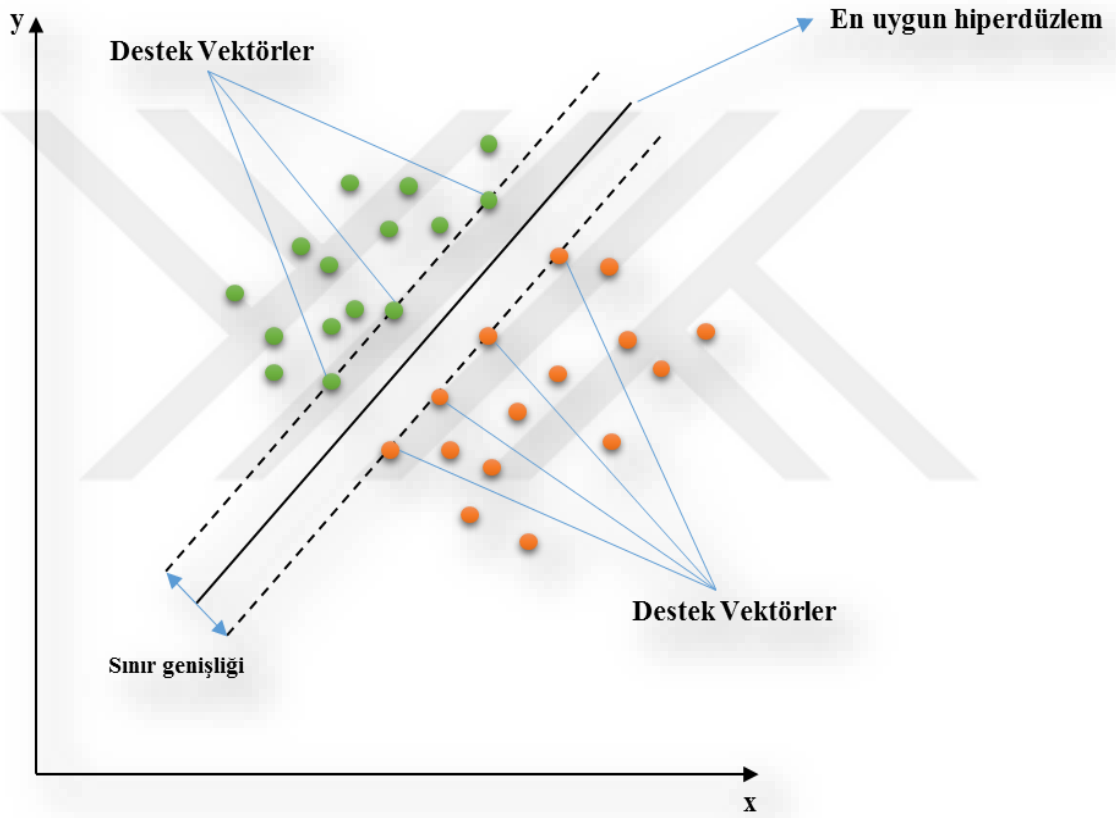
Şekil 3.9: Kullanıcıların şifreleri ile giriş yapabilmeleri için gerekli bilgilerin kaydedildiği tablo

3.3. Oluşturulan Destek Vektör Makinesi Modelleri

Değişik uygulamaları yapılan DVM son yıllarda sıklıkla kullanılan sınıflandırma yöntemlerinden biridir. Bu yöntem, hastalıkların tanısı [2, 36], farklı mühendislik uygulamaları vb. farklı alanlarda kullanılmaktadır. Cortes ve Vapnik tarafından önerilen DVM yöntemi yapısal olarak çalışan riski minimize etme ilkesini kullanmaktadır. Bu teknikte, iki sınıfın/grubun birbirlerine en yakın gözlemlere ilişkin uzaklıklarını en üst düzeye çıkarıldığı bir düzlem incelenir. Doğrusal yöntemlerle sınıflandırılmayan veriler, DVM yöntemiyle giriş vektörü daha büyük boyutlu bir uzaya lineer olmayan bir işlev aracılığıyla eşleştirilir. DVM’nin eğitilmesinde 2. dereceden bir optimizasyon denklemi kullanılabilir (46).

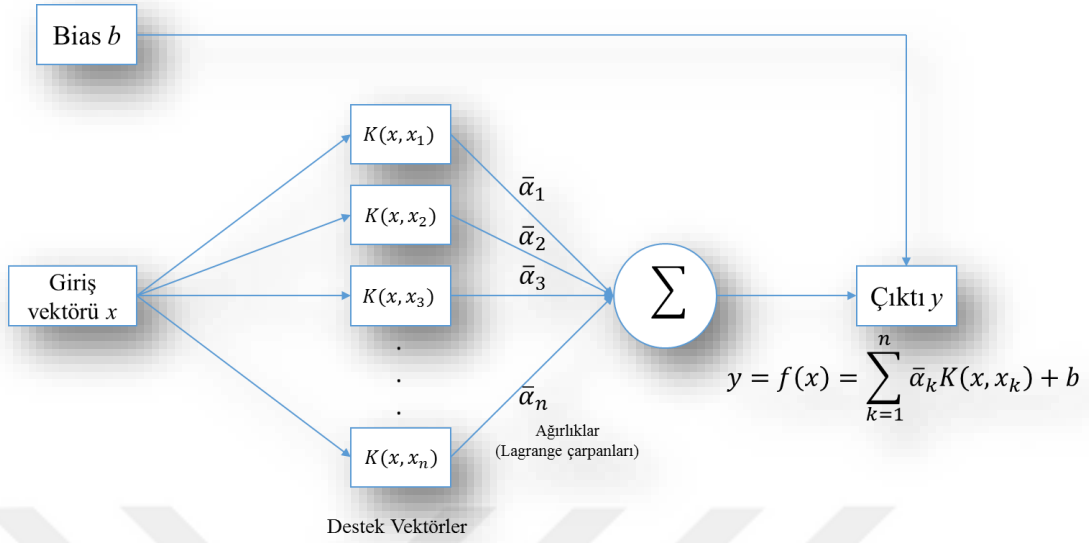
Aynı zamanda, makine öğrenme yöntemlerinde, destek vektör makineleri (destek vektör ağları olarak da adlandırılır) sınıflandırma ve regresyon analizi için kullanılan öğrenme algoritmalarıyla denetlenen modellerdendir. Her biri, her iki kategoriden birine ya da diğerine ait olarak işaretlenmiş bir dizi eğitim örneği verildiğinde, bir DVM eğitim algoritması, ikili bir sınıflayıcıdan yararlanarak incelenen kategorilerden birisine ya da diğerine yeni örnekler atayan bir model oluşturur. Bir DVM modelinde, örnekler uzaydaki noktalar olarak gösterilir ve ayrı kategorilerin örnekleri mümkün olduğunca açık bir boşluk ile bölünür. Yeni örnekler daha sonra aynı alana atanır ve yeni örneklerin boşluğun hangi tarafına dayanan bir kategoriye ait olduğu öngörülür. Doğrusal sınıflandırmaya ek olarak, DVM’ler doğrusal olmayan sınıflandırma problemlerini de

başarılı bir şekilde gerçekleştirebilir ve girişlerini yüksek boyutlu özellik alanlarına örtülü olarak eşler. Verilerin atanabilecekleri olası gruplar belli olmadığında, denetlemeli/danışmanlı öğrenme mümkün değildir. Bu durumda verilerin gruplara doğal olarak kümelenmesini ve daha sonra bu gruplara yeni verilerin atanmasını gerçekleştiren denetimsiz/danışmansız bir öğrenme yaklaşımı gereklidir (47). Şekil 3.10, doğrusal DVM'ye ilişkin destek vektörler, en uygun hiperdüzlem ve en geniş sınır aralığını göstermektedir



Şekil 3.10: Doğrusal DVM'ye ilişkin destek vektörler, en uygun hiperdüzlem ve en geniş sınır aralığı

DVM'nin yapısı Şekil 3.11'de verilmiştir (48).



Şekil 3.11: DVM'nin yapısı

Şekil 3.11'de $K(x, x_1), K(x, x_2), K(x, x_3), \dots, K(x, x_n)$ ile gösterilen fonksiyonlar çekirdek fonksiyonlarını; $\bar{\alpha}_1, \bar{\alpha}_2, \bar{\alpha}_3, \dots, \bar{\alpha}_n$ ile ifade edilen Lagrange çarpanları ise ağırlık katsayılarıdır.

Bu çalışmada çeşitli DVM modelleri, Tablo 3.1'de ayrıntılı şekilde tanımlanan değişik çekirdek fonksiyonları ile oluşturulmuş ve incelenen veri setine uygulanmıştır. Hiperparametre optimizasyonu için literatürde en yaygın ve en çok bilinen yöntem, bir parametre uzayı tanımlamak ve bu aralıktaki uzayda ızgara araması (grid search) tekniğini kullanmaktadır. Bu teknikte, hiperparametreler tespit edilen sabit bir adım büyüklüğü ile kullanılabilir olan en büyük aralıkta uygulanır ve her bir sıralanmış biçimi (kombinasyon) için performans ölçütlerine göre değerlendirilir (49). Izgara Arama tekniği daha önceden tanımlanmış model kapsamında incelenen uzayın bütünü bir ızgara aracılığı ile küçük parçalara bölerek tarama temeline dayanmaktadır. Bu yöntem, uygulanan hiperparametrelerin sayısının artırılması ve/veya belirlenen aralıklarının geniş değerli olması halinde çok yoğun hesaplamalar gerektirmesine karşın tam çözüm verme niteliğinin olması ve uygulama kolaylığı gibi nedenlerden dolayı yaygın olarak kullanılmaktadır (49). Ayrıca ızgara arama (grid search) yöntemi ya da evrimsel optimizasyon algoritması ile bulunan en uygun çekirdek fonksiyonunun parametreleri optimize edilmiştir. Izgara arama yöntemi, iki değişkenin değeri ile ızgara oluşturulan ve

ızgaranın her noktasındaki deęişken deęerleri ile arama yapılan bir tekniktir (50). Evrimsel algoritma, biyolojik evrimden esinlenerek üreme, mutasyon, rekombinasyon ve doğal seçilime benzer mekanizmalar kullanır. Optimizasyon problemlerinin aday çözümleri bir popülasyondaki bireyleri temsil eder ve seçim deęeri fonksiyonları çözümlerin içinde "yaşadığı" çevreyi belirler (51, 52).

Tablo 3.2: Çekirdek tipleri ve fonksiyonları

Çekirdek tipi	Fonksiyon
Doęrusal	$k(x, y) = x^T y + c$
Polinomial	$k(x, y) = (\alpha x^T y + c)^d$
Hiperbolik Tanjant (Sigmoid)	$k(x, y) = \tanh(\alpha x^T y + c)$
Bessel	$k(x, x') = -Bessel_{(nu+1)}^n(\sigma x - x' ^2)$
ANOVA RBF	$k(x, y) = \sum_{k=1}^n \exp(-\sigma(x^k - y^k)^2)^d$
Gaussian RBF	$k(x, y) = \exp(-\gamma\ x - y\ ^2)$
Laplace	$k(x, y) = \exp\left(-\frac{\ x - y\ }{\sigma}\right)$

Tablo 3.3'de, DVM modelinin oluşturulmasında kullanılan çekirdek fonksiyonlarının optimizasyon parametreleri, aralıkları ve kombinasyon sayıları verilmiştir.

Tablo 3.3: Çekirdek fonksiyonlarının optimizasyonları parametrelerine ilişkin tanımlayıcı bilgiler

Çekirdek Fonksiyonu	Optimizasyon Parametreleri	Parametre Aralığı	Kombinasyon Sayısı
Gaussian RBF	C, sigma	$(2^{-2}-2^5)$, (0.02-0.20)	24
Hiperbolik Tanjant (Sigmoid)	C, ölçek	$(2^{-2}-2^5)$, $(10^{-3}-10^4)$	64
Polinomial	C, derece, ölçek	$(2^{-2}-2^5)$, (1-3), $(10^{-3}-10^4)$	192
Doğrusal	C	$(2^{-2}-2^5)$	8
Laplace	C, sigma	$(2^{-2}-2^5)$, (0.02-0.20)	24
ANOVA RBF	C, sigma, derece	$(2^{-2}-2^5)$, (0.02-0.20), (1-3)	72
Bessel	C, sigma, derece	$(2^{-2}-2^5)$, (0.02-0.20), (1-3)	72

C (cost, maliyet) parametresi, ayırıcı hiperdüzlemin düzgünlüğü ile eğitim verilerinin yanlış sınıflandırılması arasındaki dengeyi kontrol eder. Maliyet parametresinin nispeten küçük seçildiği durumda, bazı eğitim verilerini yanlış sınıflandırılması pahasına geniş aralıklı ve düzgün bir hiperdüzlem seçilecektir. Maliyet parametresinin nispeten büyük seçildiği değerlerde ise eğitim verilerinin doğru bir şekilde sınıflandırabilmek için daha küçük aralıklı ve daha karmaşık bir hiperdüzlem seçilir (53).

Maliyet parametresinin aksine σ (sigma), scale (ölçek) ve degree (derece) parametreleri çekirdek fonksiyonu parametreleridir. Sigma; Gaussian RBF, Laplace, ANOVA RBF ve Bessel çekirdek fonksiyonlarının parametresi, ölçek; Hiperbolik Tanjant (Sigmoid) ve Polinomial çekirdek fonksiyonlarının parametresi, derece; Polinomial, ANOVA RBF ve Bessel çekirdek fonksiyonlarının parametresidir. Sigma, ölçek ve derece parametreleri sınıflandırma performansına doğrudan etki eden parametreler olup değer aralıkları dikkatle seçilmelidir.

Çalışmada; doğruluk, ROC eğrisi altında kalan alan (AUC), duyarlılık ve özgüllük [seçicilik] performans ölçütleri kullanılmıştır. Bu ölçütlerden doğruluk, duyarlılık ve özgüllük [seçicilik] değerleri oluşturulan karışıklık matrisi ile elde edilmiştir. Çalışmada oluşturulan karışıklık (confusion) matrisi şu şekildedir:

Tablo 3.4: Karmaşıklık matrisi

Diabetes mellitus hastalığı	Tahmin edilen durum		Toplam
	Var	Yok	
Gerçek durum	Var	A B	A+B
	Yok	C D	C+D
Toplam		A+C B+D	A+B+C+D

Bu matriste sayısı A ile gösterilen kutucukta, gerçek durumda DM hastası olup, ilgili model tarafından da DM hastası olarak tahmin edilen kişi sayısı gösterilmiştir. Sayısı B ile gösterilen kutucukta gerçek durumda DM hastası olup, model tarafından DM hastası olmayan olarak tahmin edilen kişi sayısı belirtilmiştir. Sayısı C ile gösterilen kutucukta gerçek durumda DM hastası olmayıp, model tarafından DM hastası olarak tahmin edilen kişi sayısı belirtilmiştir. Sayısı D ile gösterilen kutucukta gerçek durumda DM hastası olmayıp, yine model tarafından da DM hastası olmayan olarak tahmin edilen kişi sayısı belirtilmiştir. Modelde kullanılan doğruluk, duyarlılık ve özgüllük [seçicilik] ölçütleri yukarıdaki karmaşıklık matrisi kullanılarak şöyle hesaplanır:

$$\text{Doğruluk} = \frac{A + D}{A + B + C + D}$$

$$\text{Duyarlılık} = \frac{A}{A + B}$$

$$\text{Özgüllük} = \frac{D}{C + D}$$

Bu çalışmadaki veritabanlarında bilgi keşfi sürecinin uygulanan bütün adımlarda, optimizasyon ve değişik çekirdek fonksiyonları ile oluşturulan destek vektör makinesi modellerinin gerçekleştirilmesinde; istatistiksel hesaplamada R projesi [The R project in statistical computing] (54) ve caret paketi [caret package] (55) kullanılmıştır.

3.4. Biyoistatistiksel Analiz

Nicel veriler ortalama ve standart sapma, nitel veriler ise sayı ve yüzde özetlendi. Verilerin normal dağılıma uygunluğu Kolmogorov Smirnov testi ve varyansların homojenlik kontrolü ise Levene testi ile değerlendirildi. Verilerin analizinde bağımsız örneklerde t-testi, Pearson Ki-kare testi, Yates Düzeltilmiş Ki-kare testi ve Fisher Kesin

Ki-kare testi uygun olan yerlerde kesin (exact) yaklaşıma dayalı olarak kullanıldı. Analizlerde IBM SPSS Statistics version 24.0 for Windows paket programı (56) kullanıldı. $p < 0.05$ değeri istatistiksel olarak önemli kabul edildi.



4. BULGULAR

Aşırı/aykırı değer analizi sonucunda 2 gözlem DM olmayan bireylerin oluşturduğu veri setinden çıkarılmıştır. Bu durumda dağılım Tablo 4.1’de gösterilmiştir.

Tablo 4.1: DM değişkenin dağılımı

DM	Sayı	Yüzde (%)
Yok	1176	85.3
Var	202	14.7
Toplam	1378	100.0

Tablo 4.2’de, AKS tipi değişkeni açısından DM varlığı ve yokluğuna ilişkin dağılım verilmiştir. AKS tipi değişkeni, Unstabl Anjina, STEMI ve NSTEMI kategorilerinden oluşmaktadır. DM’nin varlığı AKS tipi Unstabl Anjina olan 25 (%12.4), STEMI olan 153 (%75.7) ve NSTEMI olan 24 (%11.9) kişide görülürken, DM’nin yokluğu AKS tipi Unstabl Anjina olan 130 (%11.1), STEMI olan 962 (%81.8) ve NSTEMI olan 84 (%7.1) kişide gözlemlenmiştir. AKS tipi değişkeni açısından incelediğinde, DM varlığı ve yokluğu arasında istatistiksel olarak fark vardır ($p=0.049$). Bu farklılığı oluşturan alt kategorilerin STEMI ve NSTEMI olduğu görülmektedir. Unstabl Anjina kategorisi açısından DM varlığı ve yokluğu arasında istatistiksel olarak fark bulunmamaktadır.

Tablo 4.2: AKS tipi değişkeni açısından DM varlığı ve yokluğuna ilişkin dağılım

Değişken adı/Kategorileri	Diabetes Mellitus				<i>p</i>	
	Yok (<i>n</i> =1176)		Var (<i>n</i> =202)			
	Sayı	Yüzde (Sütun)	Sayı	Yüzde (Sütun)		
AKS tipi	Unstabl Anjina	130 ^a	11.1	25 ^a	12.4	0.049
	STEMI	962 ^a	81.8	153 ^b	75.7	
	NSTEMI	84 ^a	7.1	24 ^b	11.9	

APA stiline göre her satırda farklı harfleri taşıyan değerler, istatistiksel olarak birbirinden farklıdır ($p<0.05$).

Tablo 4.3’de, cinsiyet değişkeni açısından DM varlığı ve yokluğuna ilişkin dağılım verilmiştir. DM’nin varlığı 69 (%34.2) kadında ve 133 (%65.8) erkekte görülürken, DM’nin yokluğu 224 (%19.0) kadında ve 952 (%81.0) erkekte gözlemlenmiştir. Cinsiyet değişkeni açısından DM varlığı ve yokluğu arasında istatistiksel olarak fark vardır ($p<0.001$). Cinsiyet değişkeninin alt kategorileri incelendiğinden bu farklılığın hem erkeklerde hem de kadınlarda olduğu görülmektedir.

Tablo 4.3: Cinsiyet değişkeni açısından DM varlığı ve yokluğuna ilişkin dağılım

Değişken adı/Kategorileri		Diabetes Mellitus				<i>p</i>
		Yok (<i>n</i> =1176)		Var (<i>n</i> =202)		
		Sayı	Yüzde (Sütun)	Sayı	Yüzde (Sütun)	
Cinsiyet	Kadın	224 ^a	19.0	69 ^b	34.2	<0.001
	Erkek	952 ^a	81.0	133 ^b	65.8	

APA stiline göre her satırda farklı harfleri taşıyan değerler, istatistiksel olarak birbirinden farklıdır ($p<0.05$).

Tablo 4.4’de, ACE inhibitör değişkeni açısından DM varlığı ve yokluğuna ilişkin dağılım verilmiştir. DM’nin varlığı ACE inhibitörü kullanmayan 172 (%85.1) ve kullanan 30 (%14.9) kişide görülürken, DM’nin yokluğu ACE inhibitörü kullanmayan 1116 (%94.9) ve kullanan 60 (%5.1) kişide gözlemlenmiştir. ACE inhibitör değişkeni açısından DM varlığı ve yokluğu arasında istatistiksel olarak fark vardır ($p<0.001$).

Tablo 4.4: ACE inhibitör değişkeni açısından DM varlığı ve yokluğuna ilişkin dağılım

Değişken adı/Kategorileri		Diabetes Mellitus				<i>p</i>
		Yok (<i>n</i> =1176)		Var (<i>n</i> =202)		
		Sayı	Yüzde (Sütun)	Sayı	Yüzde (Sütun)	
ACE inhibitör	Yok	1116 ^a	94.9	172 ^b	85.1	<0.001
	Var	60 ^a	5.1	30 ^b	14.9	

APA stiline göre her satırda farklı harfleri taşıyan değerler, istatistiksel olarak birbirinden farklıdır ($p<0.05$).

Tablo 4.5’de ARB diğer değişkeni açısından DM varlığı ve yokluğuna ilişkin dağılım verilmiştir. DM’nin varlığı ARB diğer kullanmayan 183 (%90.6) ve kullanan 19 (%9.4) kişide görülürken, DM’nin yokluğu ARB diğer kullanmayan 1130 (%96.1) ve kullanan 46 (%3.9) kişide gözlemlenmiştir. ARB diğer değişkeni açısından DM varlığı ve yokluğu arasında istatistiksel olarak fark vardır ($p=0.001$).

Tablo 4.5: ARB diğer deęişkeni açısından DM varlığı ve yokluęuna ilişkin dağılım

Deęişken adı/Kategorileri	Diabetes Mellitus				<i>p</i>	
	Yok (<i>n</i> =1176)		Var (<i>n</i> =202)			
	Sayı	Yüzde (Sütun)	Sayı	Yüzde (Sütun)		
ARB diğer	Yok	1130 ^a	96.1	183 ^b	90.6	0.001
	Var	46 ^a	3.9	19 ^b	9.4	

APA stiline göre her satırda farklı harfleri taşıyan deęerler, istatistiksel olarak birbirinden farklıdır ($p<0.05$).

Tablo 4.6’da ASA deęişkeni açısından DM varlığı ve yokluęuna ilişkin dağılım verilmiştir. DM’nin varlığı ASA kullanmayan 144 (%71.3) ve kullanan 58 (%28.7) kişide görülürken, DM’nin yokluęu ASA kullanmayan 1065 (%90.6) ve kullanan 111 (%9.4) kişide gözlemlenmiştir. ASA deęişkeni açısından DM varlığı ve yokluęu arasında istatistiksel olarak fark vardır ($p<0.001$).

Tablo 4.6: ASA deęişkeni açısından DM varlığı ve yokluęuna ilişkin dağılım

Deęişken adı/Kategorileri	Diabetes Mellitus				<i>p</i>	
	Yok (<i>n</i> =1176)		Var (<i>n</i> =202)			
	Sayı	Yüzde (Sütun)	Sayı	Yüzde (Sütun)		
ASA	Yok	1065 ^a	90.6	144 ^b	71.3	<0.001
	Var	111 ^a	9.4	58 ^b	28.7	

APA stiline göre her satırda farklı harfleri taşıyan deęerler, istatistiksel olarak birbirinden farklıdır ($p<0.05$).

Tablo 4.7’de Beta bloker deęişkeni açısından DM varlığı ve yokluęuna ilişkin dağılım verilmiştir. DM’nin varlığı Beta bloker kullanmayan 173 (%85.6) ve kullanan 29 (%14.4) kişide görülürken, DM’nin yokluęu Beta bloker kullanmayan 1084 (%92.2) ve kullanan 92 (%7.8) kişide gözlemlenmiştir. Beta bloker deęişkeni açısından DM varlığı ve yokluęu arasında istatistiksel olarak fark vardır ($p=0.004$).

Tablo 4.7: Beta bloker deęişkeni açısından DM varlığı ve yokluęuna ilişkin daęılım

Deęişken adı/Kategorileri	Diabetes Mellitus				<i>p</i>	
	Yok (<i>n</i> =1176)		Var (<i>n</i> =202)			
	Sayı	Yüzde (Sütun)	Sayı	Yüzde (Sütun)		
Beta bloker	Yok	1084 ^a	92.2	173 ^b	85.6	0.004
	Var	92 ^a	7.8	29 ^b	14.4	

APA stiline göre her satırda farklı harfleri taşıyan deęerler, istatistiksel olarak birbirinden farklıdır ($p<0.05$).

Tablo 4.8’de Diüretik deęişkeni açısından DM varlığı ve yokluęuna ilişkin daęılım verilmiştir. DM’nin varlığı Diüretik kullanmayan 168 (%83.2) ve kullanan 34 (%16.8) kişide görülürken, DM’nin yokluęu Diüretik kullanmayan 1120 (%95.2) ve kullanan 56 (%4.8) kişide gözlemlenmiştir. Diüretik deęişkeni açısından DM varlığı ve yokluęu arasında istatistiksel olarak fark vardır ($p<0.001$).

Tablo 4.8: Diüretik deęişkeni açısından DM varlığı ve yokluęuna ilişkin daęılım

Deęişken adı/Kategorileri	Diabetes Mellitus				<i>p</i>	
	Yok (<i>n</i> =1176)		Var (<i>n</i> =202)			
	Sayı	Yüzde (Sütun)	Sayı	Yüzde (Sütun)		
Diüretik	Yok	1120 ^a	95.2	168 ^b	83.2	<0.001
	Var	56 ^a	4.8	34 ^b	16.8	

APA stiline göre her satırda farklı harfleri taşıyan deęerler, istatistiksel olarak birbirinden farklıdır ($p<0.05$).

Tablo 4.9’da Statin deęişkeni açısından DM varlığı ve yokluęuna ilişkin daęılım verilmiştir. DM’nin varlığı Statin kullanmayan 187 (%92.6) ve kullanan 15 (%7.4) kişide görülürken, DM’nin yokluęu Statin kullanmayan 1146 (%97.4) ve kullanan 30 (%2.6) kişide gözlemlenmiştir. Statin deęişkeni açısından DM varlığı ve yokluęu arasında istatistiksel olarak fark vardır ($p=0.001$).

Tablo 4.9: Statin deęiřkeni aısından DM varlıęı ve yokluęuna iliřkin daęılım

Deęiřken adı/Kategorileri	Diabetes Mellitus				<i>p</i>	
	Yok (<i>n</i> =1176)		Var (<i>n</i> =202)			
	Sayı	Yüzde (Sütun)	Sayı	Yüzde (Sütun)		
Statin	Yok	1146 ^a	97.4	187 ^b	92.6	0.001
	Var	30 ^a	2.6	15 ^b	7.4	

APA stiline göre her satırda farklı harfleri taşıyan deęerler, istatistiksel olarak birbirinden farklıdır ($p<0.05$).

Tablo 4.10'da HT deęiřkeni aısından DM varlıęı ve yokluęuna iliřkin daęılım verilmiřtir. DM'nin varlıęı hiper tansiyonu olmayan 77 (%38.1) ve olan 125 (%61.9) kiřide görölürken, DM'nin yokluęu hiper tansiyonu olmayan 655 (%55.7) ve olan 521 (%44.3) kiřide gözlemlenmiřtir. HT deęiřkeni aısından DM varlıęı ve yokluęu arasında istatistiksel olarak fark vardır ($p<0.001$).

Tablo 4.10: HT deęiřkeni aısından DM varlıęı ve yokluęuna iliřkin daęılım

Deęiřken adı/Kategorileri	Diabetes Mellitus				<i>p</i>	
	Yok (<i>n</i> =1176)		Var (<i>n</i> =202)			
	Sayı	Yüzde (Sütun)	Sayı	Yüzde (Sütun)		
HT	Yok	655 ^a	55.7	77 ^b	38.1	<0.001
	Var	521 ^a	44.3	125 ^b	61.9	

APA stiline göre her satırda farklı harfleri taşıyan deęerler, istatistiksel olarak birbirinden farklıdır ($p<0.05$).

Tablo 4.11'de Kalp yetmezlięi deęiřkeni aısından DM varlıęı ve yokluęuna iliřkin daęılım verilmiřtir. Kalp yetmezlięi deęiřkeni aısından DM varlıęı ve yokluęu arasında istatistiksel olarak fark vardır ($p=0.044$).

Tablo 4.11: Kalp yetmezlięi deęiřkeni aısından DM varlıęı ve yokluęuna iliřkin daęılım

Deęiřken adı/Kategorileri	Diabetes Mellitus				<i>p</i>	
	Yok (<i>n</i> =1176)		Var (<i>n</i> =202)			
	Sayı	Yüzde (Sütun)	Sayı	Yüzde (Sütun)		
Kalp yetmezlięi	Yok	1173 ^a	99.7	199 ^b	98.5	0.044
	Var	3 ^a	0.3	3 ^b	1.5	

APA stiline göre her satırda farklı harfleri taşıyan değerler, istatistiksel olarak birbirinden farklıdır ($p<0.05$).

Tablo 4.12’de MI değişkeni açısından DM varlığı ve yokluğuna ilişkin dağılım verilmiştir. DM’nin varlığı miyokard infarktüsü olmayan 186 (%92.1) ve olan 16 (%7.9) kişide görülürken, DM’nin yokluğu miyokard infarktüsü olmayan 1133 (%96.3) ve olan 43 (%3.7) kişide gözlemlenmiştir. MI değişkeni açısından DM varlığı ve yokluğu arasında istatistiksel olarak fark vardır ($p=0.01$).

Tablo 4.12: ME değişkeni açısından DM varlığı ve yokluğuna ilişkin dağılım

Değişken adı/Kategorileri		Diabetes Mellitus				<i>p</i>
		Yok (<i>n</i> =1176)		Var (<i>n</i> =202)		
		Sayı	Yüzde (Sütun)	Sayı	Yüzde (Sütun)	
ME	Yok	1133 ^a	96.3	186 ^b	92.1	0.01
	Var	43 ^a	3.7	16 ^b	7.9	

APA stiline göre her satırda farklı harfleri taşıyan değerler, istatistiksel olarak birbirinden farklıdır ($p<0.05$).

Tablo 4.13’de Renal yetmezlik değişkeni açısından DM varlığı ve yokluğuna ilişkin dağılım verilmiştir. DM’nin varlığı renal yetmezliği olmayan 196 (%97.0) ve olan 6 (%3.0) kişide görülürken, DM’nin yokluğu renal yetmezliği olmayan 1168 (%99.3) ve olan 8 (%0.7) kişide gözlemlenmiştir. Renal yetmezlik değişkeni açısından DM varlığı ve yokluğu arasında istatistiksel olarak fark vardır ($p=0.01$).

Tablo 4.13: Renal yetmezlik değişkeni açısından DM varlığı ve yokluğuna ilişkin dağılım

Değişken adı/Kategorileri		Diabetes Mellitus				<i>p</i>
		Yok (<i>n</i> =1176)		Var (<i>n</i> =202)		
		Sayı	Yüzde (Sütun)	Sayı	Yüzde (Sütun)	
Renal yetmezlik	Yok	1168 ^a	99.3	196 ^b	97.0	0.01
	Var	8 ^a	0.7	6 ^b	3.0	

APA stiline göre her satırda farklı harfleri taşıyan değerler, istatistiksel olarak birbirinden farklıdır ($p<0.05$).

Tablo 4.14’de Sigara değişkeni açısından DM varlığı ve yokluğuna ilişkin dağılım verilmiştir. DM’nin varlığı sigara içmeyen 125 (%61.9) ve içen 77 (%38.1) kişide

görülürken, DM'nin yokluğu sigara içmeyen 463 (%39.4) ve içen 713 (%60.6) kişide gözlemlenmiştir. Sigara değişkeni açısından DM varlığı ve yokluğu arasında istatistiksel olarak fark vardır ($p<0.001$).

Tablo 4.14: Sigara değişkeni açısından DM varlığı ve yokluğuna ilişkin dağılım

Değişken adı/Kategorileri	Diabetes Mellitus				<i>p</i>	
	Yok (<i>n</i> =1176)		Var (<i>n</i> =202)			
	Sayı	Yüzde (Sütun)	Sayı	Yüzde (Sütun)		
Sigara	Yok	463 ^a	39.4	125 ^b	61.9	<0.001
	Var	713 ^a	60.6	77 ^b	38.1	

APA stiline göre her satırda farklı harfleri taşıyan değerler, istatistiksel olarak birbirinden farklıdır ($p<0.05$).

Tablo 4.15'de Kalsiyum kanal bloker değişkeni açısından DM varlığı ve yokluğuna ilişkin dağılım verilmiştir. DM'nin varlığı kalsiyum kanal bloker kullanmayan 195 (%96.5) ve kullanan 7 (%3.5) kişide görülürken, DM'nin yokluğu kalsiyum kanal bloker kullanmayan 1128 (%95.9) ve kullanan 48 (%4.1) kişide gözlemlenmiştir. Kalsiyum kanal bloker değişkeni açısından DM varlığı ve yokluğu arasında istatistiksel olarak fark yoktur ($p=0.827$).

Tablo 4.15: Kalsiyum kanal bloker değişkeni açısından DM varlığı ve yokluğuna ilişkin dağılım

Değişken adı/Kategorileri	Diabetes Mellitus				<i>p</i>	
	Yok (<i>n</i> =1176)		Var (<i>n</i> =202)			
	Sayı	Yüzde (Sütun)	Sayı	Yüzde (Sütun)		
Kalsiyum kanal bloker	Yok	1128	95.9	195	96.5	0.827
	Var	48	4.1	7	3.5	

Tablo 4.16'da Clopidogrel değişkeni açısından DM varlığı ve yokluğuna ilişkin dağılım verilmiştir. DM'nin varlığı clopidogrel kullanmayan 201 (%99.5) ve kullanan 1 (%0.5) kişide görülürken, DM'nin yokluğu clopidogrel kullanmayan 1164 (%99.0) ve kullanan 12 (%1.0) kişide gözlemlenmiştir. Clopidogrel değişkeni açısından DM varlığı ve yokluğu arasında istatistiksel olarak fark yoktur ($p=0.705$).

Tablo 4.16: Clopidogrel deęişkeni aısından DM varlıęı ve yokluęuna iliřkin daęılım

Deęişken adı/Kategorileri		Diabetes Mellitus				<i>p</i>
		Yok (<i>n</i> =1176)		Var (<i>n</i> =202)		
		Sayı	Yüzde (Sütun)	Sayı	Yüzde (Sütun)	
Clopidogrel	Yok	1164	99.0	201	99.5	0.705
	Var	12	1.0	1	0.5	

Tablo 4.17’de Aile öyküsü deęişkeni aısından DM varlıęı ve yokluęuna iliřkin daęılım verilmiřtir. DM’nin varlıęı aile öyküsü olmayan 172 (%85.1) ve olan 30 (%14.9) kiřide görölürken, DM’nin yokluęu aile öyküsü olmayan 952 (%81.0) ve olan 224 (%19.0) kiřide gözlemlenmiřtir. Aile öyküsü deęişkeni aısından DM varlıęı ve yokluęu arasında istatistiksel olarak fark yoktur ($p=0.155$).

Tablo 4.17: AKS aile öyküsü deęişkeni aısından DM varlıęı ve yokluęuna iliřkin daęılım

Deęişken adı/Kategorileri		Diabetes Mellitus				<i>p</i>
		Yok (<i>n</i> =1176)		Var (<i>n</i> =202)		
		Sayı	Yüzde (Sütun)	Sayı	Yüzde (Sütun)	
Aile öyküsü	Yok	952	81.0	172	85.1	0.155
	Var	224	19.0	30	14.9	

Tablo 4.18’de koroner arter by-pass deęişkeni aısından DM varlıęı ve yokluęuna iliřkin daęılım verilmiřtir. DM’nin varlıęı koroner arter by-pass olmayan 200 (%99.0) ve olan 2 (%1.0) kiřide görölürken, DM’nin yokluęu koroner arter by-pass olmayan 1156 (%98.3) ve olan 20 (%1.7) kiřide gözlemlenmiřtir. Koroner arter by-pass deęişkeni aısından DM varlıęı ve yokluęu arasında istatistiksel olarak fark yoktur ($p=0.155$).

Tablo 4.18: Koroner arter by-pass deęişkeni açısından DM varlığı ve yokluęuna ilişkin daęılım

Deęişken adı/Kategorileri		Diabetes Mellitus				<i>p</i>
		Yok (<i>n</i> =1176)		Var (<i>n</i> =202)		
		Sayı	Yüzde (Sütun)	Sayı	Yüzde (Sütun)	
Koroner arter bypass	Yok	1156	98.3	200	99.0	0.760
	Var	20	1.7	2	1.0	

Tablo 4.19’da Hiperlipidemi deęişkeni açısından DM varlığı ve yokluęuna ilişkin daęılım verilmiştir. DM’nin varlığı hiperlipidemi öyküsü olmayan 192 (%95.0) ve olan 10 (%5.0) kişide görülürken, DM’nin yokluęu hiperlipidemi öyküsü olmayan 1143 (%97.2) ve olan 33 (%2.8) kişide gözlemlenmiştir. Hiperlipidemi deęişkeni açısından DM varlığı ve yokluęu arasında istatistiksel olarak fark yoktur ($p=0.161$).

Tablo 4.19: Hiperlipidemi deęişkeni açısından DM varlığı ve yokluęuna ilişkin daęılım

Deęişken adı/Kategorileri		Diabetes Mellitus				<i>p</i>
		Yok (<i>n</i> =1176)		Var (<i>n</i> =202)		
		Sayı	Yüzde (Sütun)	Sayı	Yüzde (Sütun)	
Hiperlipidemi	Yok	1143	97.2	192	95.0	0.161
	Var	33	2.8	10	5.0	

Tablo 4.20’de Koroner arter hastalığı deęişkeni açısından DM varlığı ve yokluęuna ilişkin daęılım verilmiştir. DM’nin varlığı koroner arter hastalığı olmayan 187 (%92.6) ve olan 15 (%7.4) kişide görülürken, DM’nin yokluęu koroner arter hastalığı olmayan 1117 (%95.0) ve olan 59 (%5.0) kişide gözlemlenmiştir. Koroner arter hastalığı deęişkeni açısından DM varlığı ve yokluęu arasında istatistiksel olarak fark yoktur ($p=0.217$).

Tablo 4.20: Koroner arter hastalığı değişkeni açısından DM varlığı ve yokluğuna ilişkin dağılım

Değişken adı/Kategorileri		Diabetes Mellitus				<i>p</i>
		Yok (<i>n</i> =1176)		Var (<i>n</i> =202)		
		Sayı	Yüzde (Sütun)	Sayı	Yüzde (Sütun)	
Koroner arter hastalığı	Yok	1117	95.0	187	92.6	0.217
	Var	59	5.0	15	7.4	

Tablo 4.21’de Malignite değişkeni açısından DM varlığı ve yokluğuna ilişkin dağılım verilmiştir. DM’nin varlığı malignite öyküsü olmayan 201 (%99.5) ve olan 1 (%0.5) kişide görülürken, DM’nin yokluğu malignite öyküsü olmayan 1170 (%99.5) ve olan 6 (%0.5) kişide gözlemlenmiştir. Malignite değişkeni açısından DM varlığı ve yokluğu arasında istatistiksel olarak fark yoktur ($p=0.999$).

Tablo 4.21: Malignite değişkeni açısından DM varlığı ve yokluğuna ilişkin dağılım

Değişken adı/Kategorileri		Diabetes Mellitus				<i>p</i>
		Yok (<i>n</i> =1176)		Var (<i>n</i> =202)		
		Sayı	Yüzde (Sütun)	Sayı	Yüzde (Sütun)	
Malignite	Yok	1170	99.5	201	99.5	0.999
	Var	6	0.5	1	0.5	

Tablo 4.22’de PAH değişkeni açısından DM varlığı ve yokluğuna ilişkin dağılım verilmiştir. DM’nin varlığı perifrik damar hastalığı olmayan 200 (%99.0) ve olan 2 (%1) kişide görülürken, DM’nin yokluğu perifrik damar hastalığı olmayan 1172 (%99.7) ve olan 4 (%0.3) kişide gözlemlenmiştir. PAH değişkeni açısından DM varlığı ve yokluğu arasında istatistiksel olarak fark yoktur ($p=0.215$).

Tablo 4.22: PAH deęişkeni aısından DM varlıęı ve yokluęuna iliřkin daęılım

Deęişken adı/Kategorileri		Diabetes Mellitus				<i>p</i>
		Yok (<i>n</i> =1176)		Var (<i>n</i> =202)		
		Sayı	Yüzde (Sütun)	Sayı	Yüzde (Sütun)	
PAH	Yok	1172	99.7	200	99.0	0.215
	Var	4	0.3	2	1.0	

Tablo 4.23’de İnme deęişkeni aısından DM varlıęı ve yokluęuna iliřkin daęılım verilmiřtir. DM’nin varlıęı inme hastalıęı olmayan 202 (%100.0) kiřide görölürken, DM’nin yokluęu inme hastalıęı olmayan 1166 (%99.1) ve olan 10(%0.9) kiřide gözlemlenmiřtir. İnme deęişkeni aısından DM varlıęı ve yokluęu arasında istatistiksel olarak fark yoktur ($p=0.374$).

Tablo 4.23: İnme deęişkeni aısından DM varlıęı ve yokluęuna iliřkin daęılım

Deęişken adı/Kategorileri		Diabetes Mellitus				<i>p</i>
		Yok (<i>n</i> =1176)		Var (<i>n</i> =202)		
		Sayı	Yüzde (Sütun)	Sayı	Yüzde (Sütun)	
İnme	Yok	1166	99.1	202	100.0	0.374
	Var	10	0.9	0	0.0	

Nicel deęişkenlerin DM’nin varlıęına ve yokluęuna göre tanımlayıcı istatistikleri Tablo 4.24’de verilmiřtir. Tablo 4.24 incelendięinde, BUN, glikoz, HDL, VKİ deęişkenleri aısından DM varlıęı ve yokluęu arasında istatistiksel olarak fark varken ($p<0.05$), LDL, trigliserid, nabız, DKB, SKB, kolesterol deęişkenleri aısından DM varlıęı ve yokluęu arasında istatistiksel olarak farklılık yoktur ($p>0.05$).

Tablo 4.24: Nicel deęişkenlerin DM'nin varlığına ve yokluęuna göre tanımlayıcı istatistikleri

Deęişkenler	Diabetes mellitus				<i>p</i>
	Yok (<i>n</i> =1176)		Var (<i>n</i> =202)		
	Ortalama	Standart Sapma	Ortalama	Standart Sapma	
BUN	17.91 ^a	11.04	21.52 ^b	12.62	<0.001
Glikoz	129 ^a	57	210 ^b	99	<0.001
HDL	36 ^a	9	35 ^b	8	0.036
Kolesterol	180 ^a	36	182 ^a	43	0.525
Kreatinin	1.05 ^a	2.06	1.58 ^b	4.38	0.094
LDL	119.8 ^a	30.3	115.5 ^a	34.2	0.10
Trigiliserid	148 ^a	61	155 ^a	64	0.145
Nabız	73 ^a	11	72 ^a	12	0.272
DKB	80 ^a	12	80 ^a	13	0.097
SKB	132 ^a	19	132 ^a	20	0.902
VKİ	25.21 ^a	2.22	25.72 ^b	2.97	0.021

APA stiline göre her satırda farklı harfleri taşıyan deęerler, istatistiksel olarak birbirinden farklıdır ($p<0.05$).

Veritabanlarında bilgi keşfi sürecinin “Veri Madencilięi” ve “Deęerlendirme ve Yorumlama” adımları kapsamında deęişik çekirdek fonksiyonları ile oluşturulan DVM modelleri elde edilmiştir. Bu bağlamda; Gaussian RBF çekirdek fonksiyonlu DVM modeli için tahminlere ilişkin performans ölçütleri ve %95 güven aralığı deęerleri Tablo 4.25’de özetlenmiştir. Bu verilere göre; Gaussian RBF çekirdek fonksiyonlu DVM modeli için Doğruluk deęeri 0.9448 (0.9315 – 0.9563), ROC eğrisi altında kalan alan (AUC) 0.8139 (0.7805-0.8473), Duyarlılık deęeri 0.9922 (0.9572 – 0.9998) ve Özgüllük [Seçicilik] deęeri 0.9400 (0.9254 - 0.9525) olarak bulundu. Bu deęerler arasında Gaussian RBF çekirdek fonksiyonlu DVM modeli için en yüksek performans ölçütü deęeri 0.9922 ile duyarlılık metrięine aitti. Sonrasında bu performansı sırasıyla; 0.9448 deęeri ile Doğruluk, 0.9400 deęeri ile Özgüllük [Seçicilik] ve 0.8139 deęeri ile ROC eğrisi altında kalan alan (AUC) izlemektedir.

Tablo 4.25: Gaussian RBF çekirdek fonksiyonlu DVM modeli için tahminlere ilişkin performans ölçütleri ve %95 güven aralığı değerleri

Performans ölçütü	Değer (% 95 Güven Aralığı)
Doğruluk	0.9448 (0.9315 – 0.9563)
ROC eğrisi altında kalan alan (AUC)	0.8139 (0.7805-0.8473)
Duyarlılık	0.9922 (0.9572 – 0.9998)
Özgüllük [Seçicilik]	0.9400 (0.9254 - 0.9525)

Hyperbolic tangent (Sigmoid) çekirdek fonksiyonlu DVM modeli için tahminlere ilişkin performans ölçütleri ve %95 güven aralığı değerleri Tablo 4.26’da özetlenmiştir. Bu verilere göre; Hyperbolic tangent (Sigmoid) çekirdek fonksiyonlu DVM modeli için Doğruluk değeri 0.8716 (0.8527 - 0.8888), ROC eğrisi altında kalan alan (AUC) 0.5906 (0.5631 - 0.618), Duyarlılık değeri 0.7359 (0.5967 - 0.8474) ve Özgüllük [Seçicilik] değeri 0.8770 (0.8581 - 0.8942) olarak bulundu. Bu değerler arasında Hyperbolic tangent (Sigmoid) çekirdek fonksiyonlu DVM modeli için en yüksek performans ölçütü değeri 0.8770 ile Özgüllük [Seçicilik] metriğine aitti. Sonrasında bu performansı sırasıyla; 0.8716 değeri ile Doğruluk, 0.7359 değeri ile Duyarlılık ve 0.5906 değeri ile ROC eğrisi altında kalan alan (AUC) izlemektedir.

Tablo 4.26: Hyperbolic tangent (Sigmoid) çekirdek fonksiyonlu DVM modeli için tahminlere ilişkin performans ölçütleri ve %95 güven aralığı değerleri

Performans ölçütü	Değer (% 95 Güven Aralığı ile)
Doğruluk	0.8716 (0.8527 - 0.8888)
ROC eğrisi altında kalan alan (AUC)	0.5906 (0.5631 - 0.618)
Duyarlılık	0.7359 (0.5967 - 0.8474)
Özgüllük [Seçicilik]	0.8770 (0.8581 - 0.8942)

Polinomial çekirdek fonksiyonlu DVM modeli için tahminlere ilişkin performans ölçütleri ve %95 güven aralığı değerleri Tablo 4.27’de özetlenmiştir. Bu verilere göre; Polinomial çekirdek fonksiyonlu DVM modeli için Doğruluk değeri 0.8846 (0.8666 - 0.901), ROC eğrisi altında kalan alan (AUC) 0.6495 (0.6171 - 0.6819), Duyarlılık değeri 0.75294 (0.6475 - 0.8401) ve Özgüllük [Seçicilik] değeri 0.89327 (0.8752 - 0.9096) olarak bulundu. Bu değerler arasında Polinomial çekirdek fonksiyonlu DVM modeli için en yüksek performans ölçütü değeri 0.89327 ile Özgüllük [Seçicilik]

metriğine aitti. Sonrasında bu performansı sırasıyla; 0.8846 değeri ile Doğruluk, 0.75294 değeri ile Duyarlılık ve 0.6495 değeri ile ROC eğrisi altında kalan alan (AUC) izlemektedir.

Tablo 4.27: Polinomial çekirdek fonksiyonlu DVM modeli için tahminlere ilişkin performans ölçütleri ve %95 güven aralığı değerleri

Performans ölçütü	Değer (% 95 Güven Aralığı ile)
Doğruluk	0.8846 (0.8666 - 0.901)
ROC eğrisi altında kalan alan (AUC)	0.6495 (0.6171 - 0.6819)
Duyarlılık	0.75294 (0.6475 - 0.8401)
Özgüllük [Seçicilik]	0.89327 (0.8752 - 0.9096)

Doğrusal çekirdek fonksiyonlu DVM modeli için tahminlere ilişkin performans ölçütleri ve %95 güven aralığı değerleri Tablo 4.28’de özetlenmiştir. Bu verilere göre; Doğrusal çekirdek fonksiyonlu DVM modeli için Doğruluk değeri 0.8788 (0.8604 - 0.8956), ROC eğrisi altında kalan alan (AUC) 0.6256 (0.5948-0.6564), Duyarlılık değeri 0.73973 (0.6238 – 0.8355) ve Özgüllük [Seçicilik] değeri 0.88659 (0.8681 – 0.9033) olarak bulundu. Bu değerler arasında Doğrusal çekirdek fonksiyonlu DVM modeli için en yüksek performans ölçütü değeri 0.88659 ile Özgüllük [Seçicilik] metriğine aitti. Sonrasında bu performansı sırasıyla; 0.8788 değeri ile Doğruluk, 0.73973 değeri ile Duyarlılık ve 0.6256 değeri ile ROC eğrisi altında kalan alan (AUC) izlemektedir.

Tablo 4.28: Doğrusal çekirdek fonksiyonlu DVM modeli için tahminlere ilişkin performans ölçütleri ve %95 güven aralığı değerleri

Performans ölçütü	Değer (% 95 Güven Aralığı ile)
Doğruluk	0.8788 (0.8604 - 0.8956)
ROC eğrisi altında kalan alan (AUC)	0.6256 (0.5948-0.6564)
Duyarlılık	0.73973 (0.6238 – 0.8355)
Özgüllük [Seçicilik]	0.88659 (0.8681 – 0.9033)

Laplace çekirdek fonksiyonlu DVM modeli için tahminlere ilişkin performans ölçütleri ve %95 güven aralığı değerleri Tablo 4.29’da özetlenmiştir. Bu verilere göre; Laplace çekirdek fonksiyonlu DVM modeli için Doğruluk değeri 0.9804 (0.9716 - 0.987), ROC eğrisi altında kalan alan (AUC) 0.9332 (0.9096 - 0.9567), Duyarlılık değeri 0.9999 (0.9791 – 1.000) ve Özgüllük [Seçicilik] değeri 0.9776 (0.9675 – 0.9852) olarak bulundu.

Bu deęerler arasında Laplace çekirdek fonksiyonlu DVM modeli için en yüksek performans ölçütü deęeri 0.9999 ile Duyarlılık metrięine aitti. Sonrasında bu performansı sırasıyla; 0.9804 deęeri ile Doğruluk, 0.9776 deęeri ile Özgüllük [Seçicilik] ve 0.9332 deęeri ile ROC eğrisi altında kalan alan (AUC) izlemektedir.

Tablo 4.29: Laplace çekirdek fonksiyonlu DVM modeli için tahminlere ilişkin performans ölçütleri ve %95 güven aralığı deęerleri

Performans ölçütü	Deęer (% 95 Güven Aralığı ile)
Doęruluk	0.9804 (0.9716 - 0.987)
ROC eğrisi altında kalan alan (AUC)	0.9332 (0.9096 - 0.9567)
Duyarlılık	0.9999 (0.9791 – 1.000)
Özgüllük [Seçicilik]	0.9776 (0.9675 – 0.9852)

ANOVA RBF çekirdek fonksiyonlu DVM modeli için tahminlere ilişkin performans ölçütleri ve %95 güven aralığı deęerleri Tablo 4.30’da özetlenmiştir. Bu verilere göre; ANOVA RBF çekirdek fonksiyonlu DVM modeli için Doğruluk deęeri 0.8803 (0.8619 - 0.8969), ROC eğrisi altında kalan alan (AUC) 0.6182 (0.5883 - 0.6482), Duyarlılık deęeri 0.7937 (0.6730 - 0.8853) ve Özgüllük [Seçicilik] deęeri 0.8844 (0.8659 - 0.9012) olarak bulundu. Bu deęerler arasında ANOVA RBF çekirdek fonksiyonlu DVM modeli için en yüksek performans ölçütü deęeri 0.8844 ile Özgüllük [Seçicilik] metrięine aitti. Sonrasında bu performansı sırasıyla; 0.8803 deęeri ile Doğruluk, 0.7937 deęeri ile Duyarlılık ve 0.6182 deęeri ile ROC eğrisi altında kalan alan (AUC) izlemektedir.

Tablo 4.30: ANOVA RBF çekirdek fonksiyonlu DVM modeli için tahminlere ilişkin performans ölçütleri ve %95 güven aralığı deęerleri

Performans ölçütü	Deęer (% 95 Güven Aralığı ile)
Doęruluk	0.8803 (0.8619 - 0.8969)
ROC eğrisi altında kalan alan (AUC)	0.6182 (0.5883 - 0.6482)
Duyarlılık	0.7937 (0.6730 - 0.8853)
Özgüllük [Seçicilik]	0.8844 (0.8659 - 0.9012)

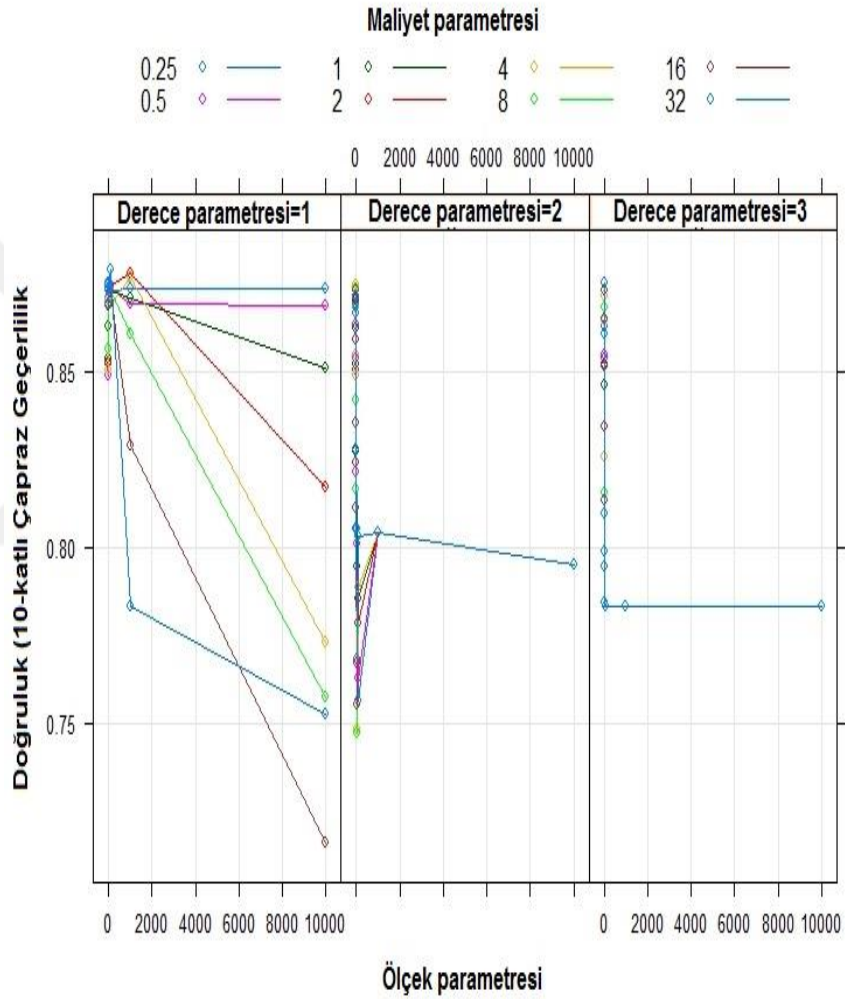
Tablo 4.31’de Bessel çekirdek fonksiyonlu DVM modeli için tahminlere ilişkin performans ölçütleri ve %95 güven aralığı deęerleri özetlenmiştir. Doğruluk deęeri 0.8832 (0.865 - 0.8997), ROC eğrisi altında kalan alan (AUC) 0.6261 (0.5955 - 0.6566),

Duyarlılık değeri 0.8154 (0.6997 - 0.9008) ve Özgüllük [Seçicilik] değeri 0.8865 (0.8681 - 0.9032) Bessel çekirdek fonksiyonlu DVM modeli için elde edilmiştir. Özgüllük [Seçicilik] metriği performans ölçütü değeri 0.8865 ile Bessel çekirdek fonksiyonlu DVM modeli için en yüksek değerli metriktir. 0.8832 değeri ile Doğruluk, 0.8154 değeri ile Duyarlılık ve 0.6261 değeri ile ROC eğrisi altında kalan alan (AUC) bu metriği takip etmektedir.

Tablo 4.31: Bessel çekirdek fonksiyonlu DVM modeli için tahminlere ilişkin performans ölçütleri ve %95 güven aralığı değerleri

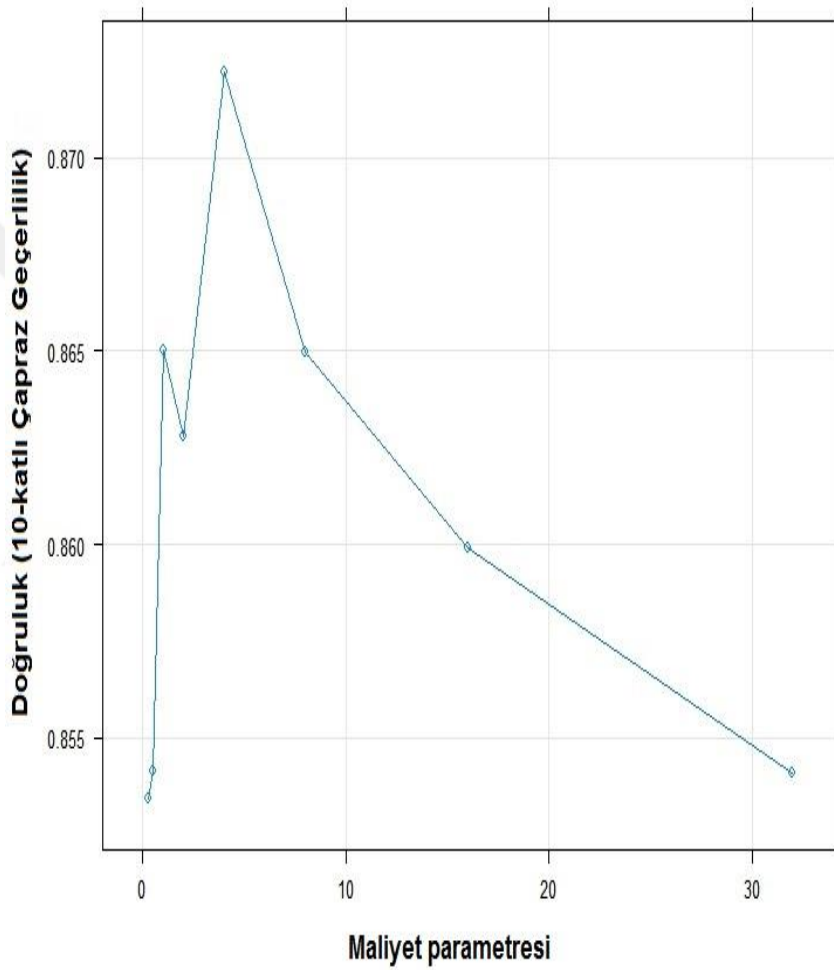
Performans ölçütü	Değer (% 95 Güven Aralığı ile)
Doğruluk	0.8832 (0.865 - 0.8997)
ROC eğrisi altında kalan alan (AUC)	0.6261 (0.5955 - 0.6566)
Duyarlılık	0.8154 (0.6997 - 0.9008)
Özgüllük [Seçicilik]	0.8865 (0.8681 - 0.9032)

DVM modeli polinomial çekirdek fonksiyonu ile oluşturulduğunda, ilgili çekirdek fonksiyonunun optimizasyon parametrelerine göre elde edilen model performansları Şekil 4.1’de gösterilmiştir. En uygun optimizasyon parametresinin hangisi olduğuna karar vermek için doğruluk metriği kullanılmıştır. Şekil 4.1 incelendiğinde, derece parametresi 1, maliyet parametresi 32, ölçek parametresi 100 olduğunda doğruluk değeri 0.8788 olarak elde edilmiştir. Bir başka deyişle, modelin en iyi öğrenme performansı yukarıda bahsedilen parametreler ile gerçekleşmiştir.



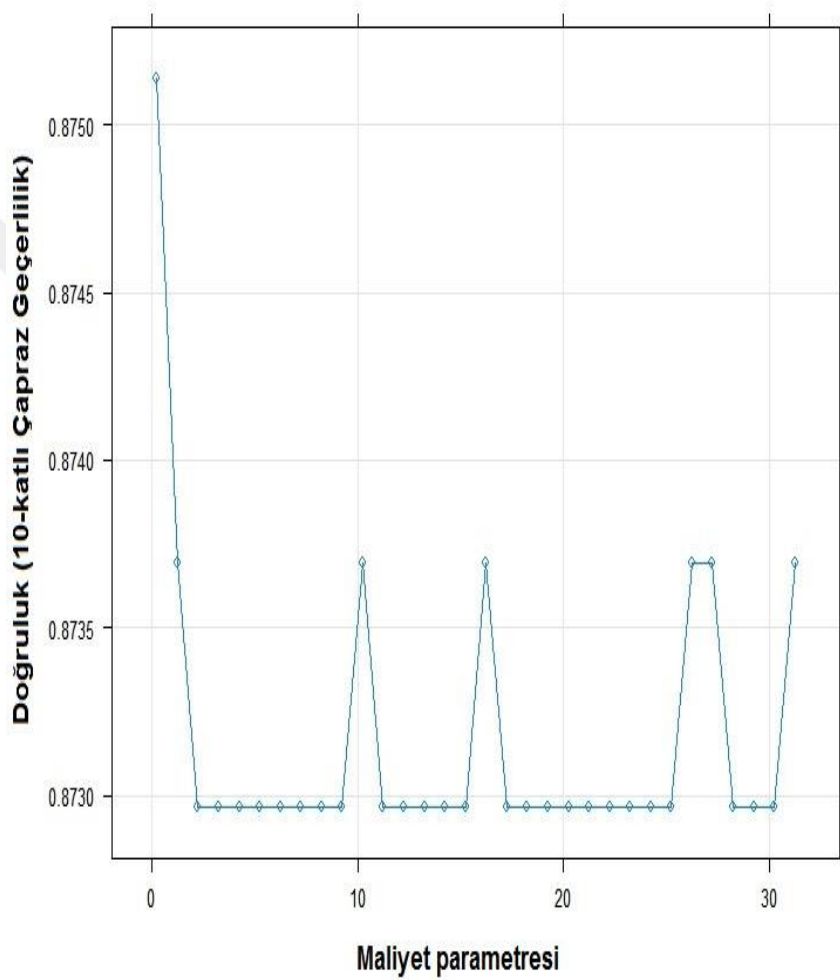
Şekil 4.1: Maliyet, derece ve ölçek parametrelerine göre doğruluk metriğinin değişimi (Çekirdek fonksiyon=Polinomial)

DVM modeli Gaussian RBF çekirdek fonksiyonu ile oluşturulduğunda, ilgili çekirdek fonksiyonunun optimizasyon parametrelerine göre elde edilen model performansları Şekil 4.2’de gösterilmiştir. En uygun optimizasyon parametresinin hangisi olduğuna karar vermek için doğruluk metriği kullanılmıştır. Şekil 4.2 incelendiğinde, maliyet parametresi 4 olduğunda doğruluk değeri 0.8722 olarak elde edilmiştir. Yani modelin en iyi öğrenme performansı maliyet parametresi 4 iken gerçekleşmiştir.



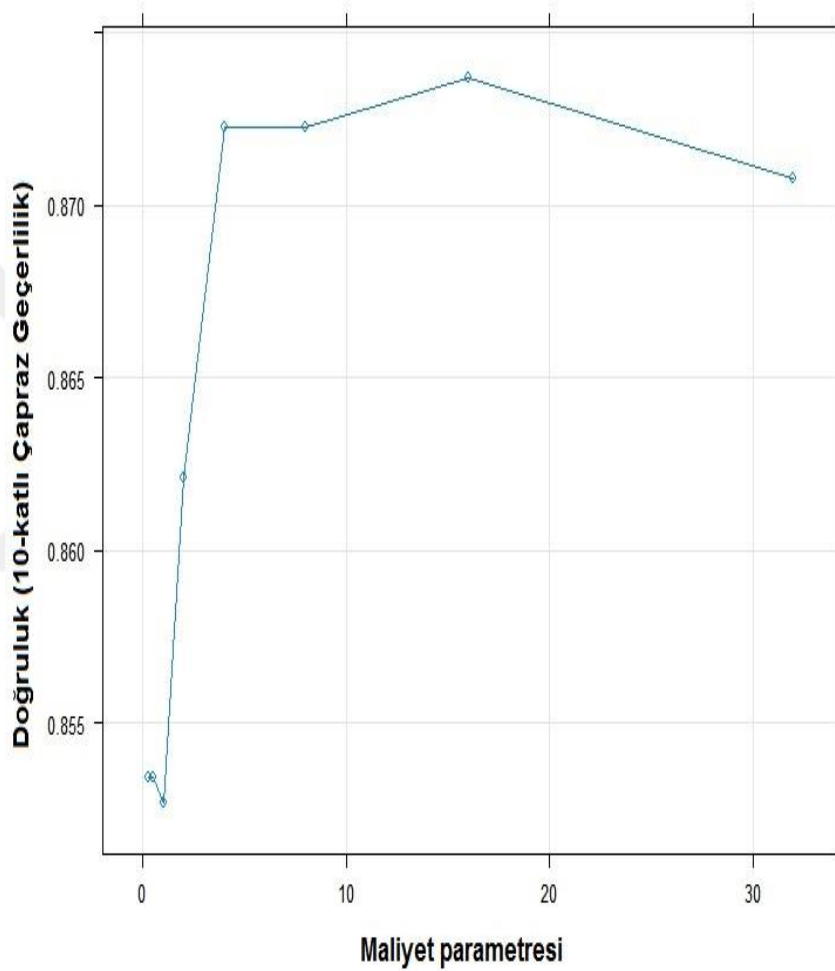
Şekil 4.2: Maliyet parametresine göre doğruluk metriğinin değişimi (Çekirdek fonksiyon= Gaussian RBF)

DVM modeli doğrusal çekirdek fonksiyonu ile oluşturulduğunda, ilgili çekirdek fonksiyonunun optimizasyon parametrelerine göre elde edilen model performansları Şekil 4.3’de gösterilmiştir. En uygun optimizasyon parametresinin hangisi olduğuna karar vermek için doğruluk metriği kullanılmıştır. Şekil 4.3 incelendiğinde, maliyet parametresi 0.25 olduğunda doğruluk değeri 0.8751 olarak elde edilmiştir. Yani modelin en iyi öğrenme performansı maliyet parametresi 0.25 iken gerçekleşmiştir.



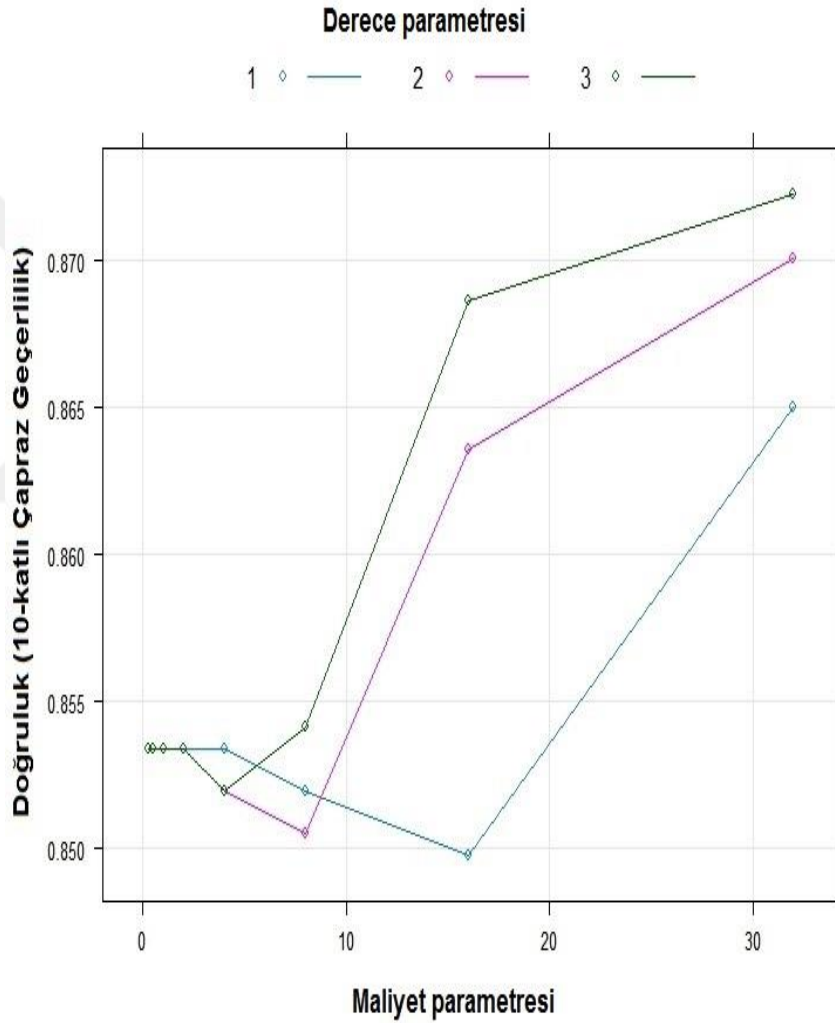
Şekil 4.3: Maliyet parametresine göre doğruluk metriğinin değişimi (Çekirdek fonksiyon=Doğrusal)

DVM modeli laplace çekirdek fonksiyonu ile oluşturulduğunda, ilgili çekirdek fonksiyonunun optimizasyon parametrelerine göre elde edilen model performansları Şekil 4.4'de gösterilmiştir. En uygun optimizasyon parametresinin hangisi olduğuna karar vermek için doğruluk metriği kullanılmıştır. Şekil 4.4 incelendiğinde, maliyet parametresi 16 olduğunda doğruluk değeri 0.8737 olarak elde edilmiştir. Yani modelin en iyi öğrenme performansı maliyet parametresi 16 iken gerçekleşmiştir.



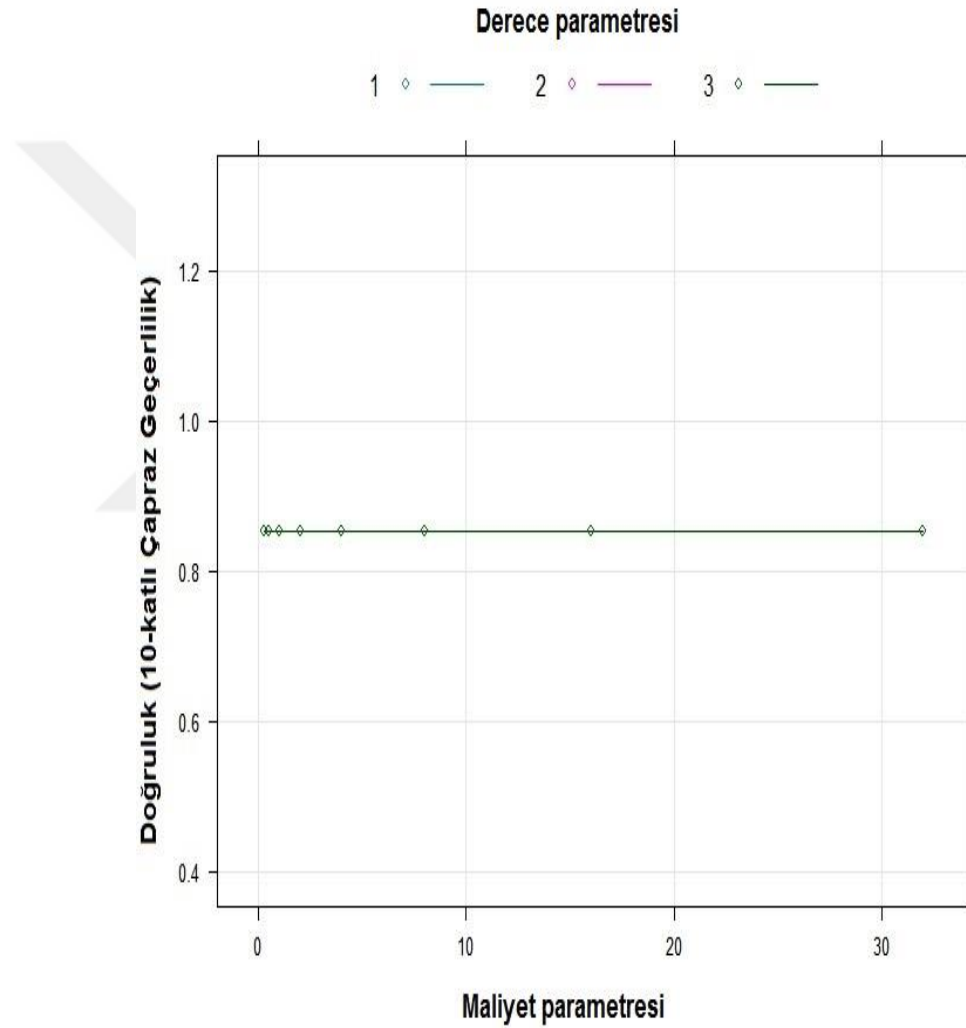
Şekil 4.4: Maliyet parametresine göre doğruluk metriğinin değişimi (Çekirdek fonksiyon=Laplace)

DVM modeli bessel çekirdek fonksiyonu ile oluşturulduğunda, ilgili çekirdek fonksiyonunun optimizasyon parametrelerine göre elde edilen model performansları Şekil 4.5’de gösterilmiştir. En uygun optimizasyon parametresinin hangisi olduğuna karar vermek için doğruluk metriği kullanılmıştır. Şekil 4.5 incelendiğinde, derece parametresi 3, maliyet parametresi 32 olduğunda doğruluk değeri 0.8723 olarak elde edilmiştir. Yani modelin en iyi öğrenme performansı derece parametresi 3, maliyet parametresi 16 iken gerçekleşmiştir. Ayrıca Şekil 4.5’de ifade edilememekle birlikte, bu en iyi öğrenme sonucu sigma parametresinin 0.0367 değeri için elde edilmiştir.



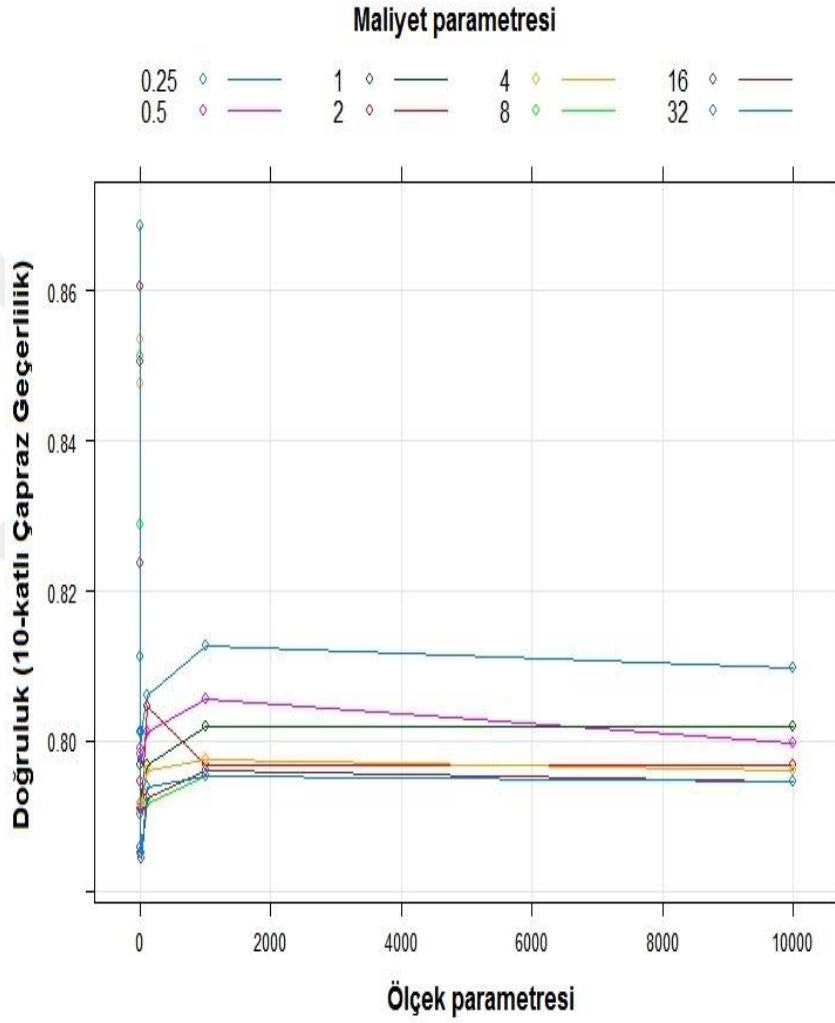
Şekil 4.5: Maliyet ve derece parametrelerine göre doğruluk metriğinin değişimi (Çekirdek fonksiyon=Bessel)

DVM modeli ANOVA RBF çekirdek fonksiyonu ile oluşturulduğunda, ilgili çekirdek fonksiyonunun optimizasyon parametrelerine göre elde edilen model performansları Şekil 4.6’da gösterilmiştir. En uygun optimizasyon parametresinin hangisi olduğuna karar vermek için doğruluk metriği kullanılmıştır. Şekil 4.6 incelendiğinde, tüm derece ve maliyet parametreleri için doğruluk değeri 0.8534 olarak elde edilmiştir. Yani modelin en iyi öğrenme performansı tüm derece ve maliyet parametreleri için gerçekleşmiştir. Ayrıca Şekil 4.6’da ifade edilememekle birlikte, bu en iyi öğrenme sonucu sigma parametresinin 0.0367 değeri için elde edilmiştir.



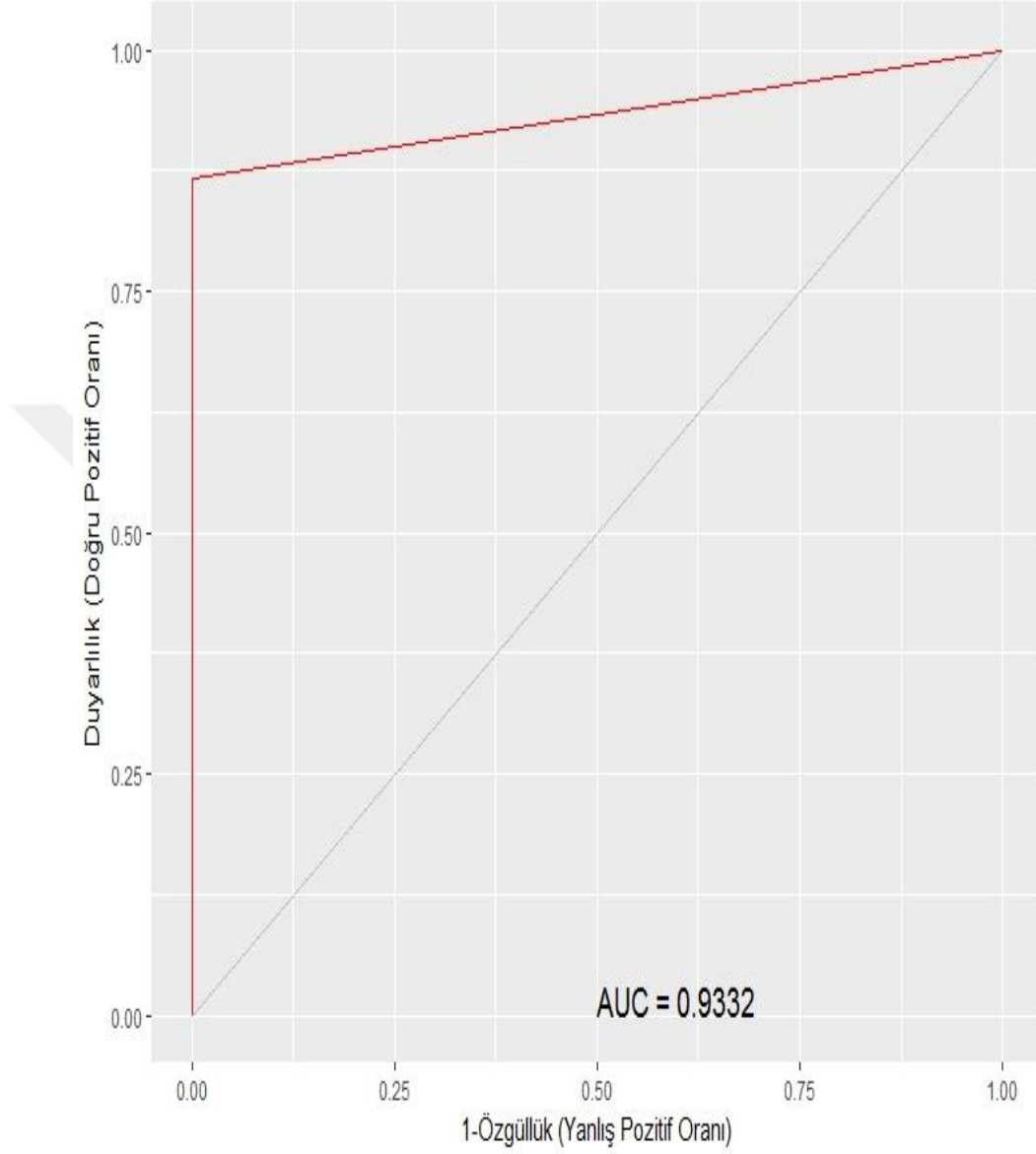
Şekil 4.6: Maliyet ve derece parametrelerine göre doğruluk metriğinin değişimi (Çekirdek fonksiyon= ANOVA RBF)

DVM modeli hiperbolik tanjant (sigmoid) çekirdek fonksiyonu ile oluşturulduğunda, ilgili çekirdek fonksiyonunun optimizasyon parametrelerine göre elde edilen model performansları Şekil 4.7’de gösterilmiştir. En uygun optimizasyon parametresinin hangisi olduğuna karar vermek için doğruluk metriği kullanılmıştır. Şekil 4.7 incelendiğinde, tüm derece ve maliyet parametreleri için doğruluk değeri 0.8686 olarak elde edilmiştir. Yani modelin en iyi öğrenme performansı maliyet parametresi 32, ölçek parametresi 0.001 için gerçekleşmiştir.



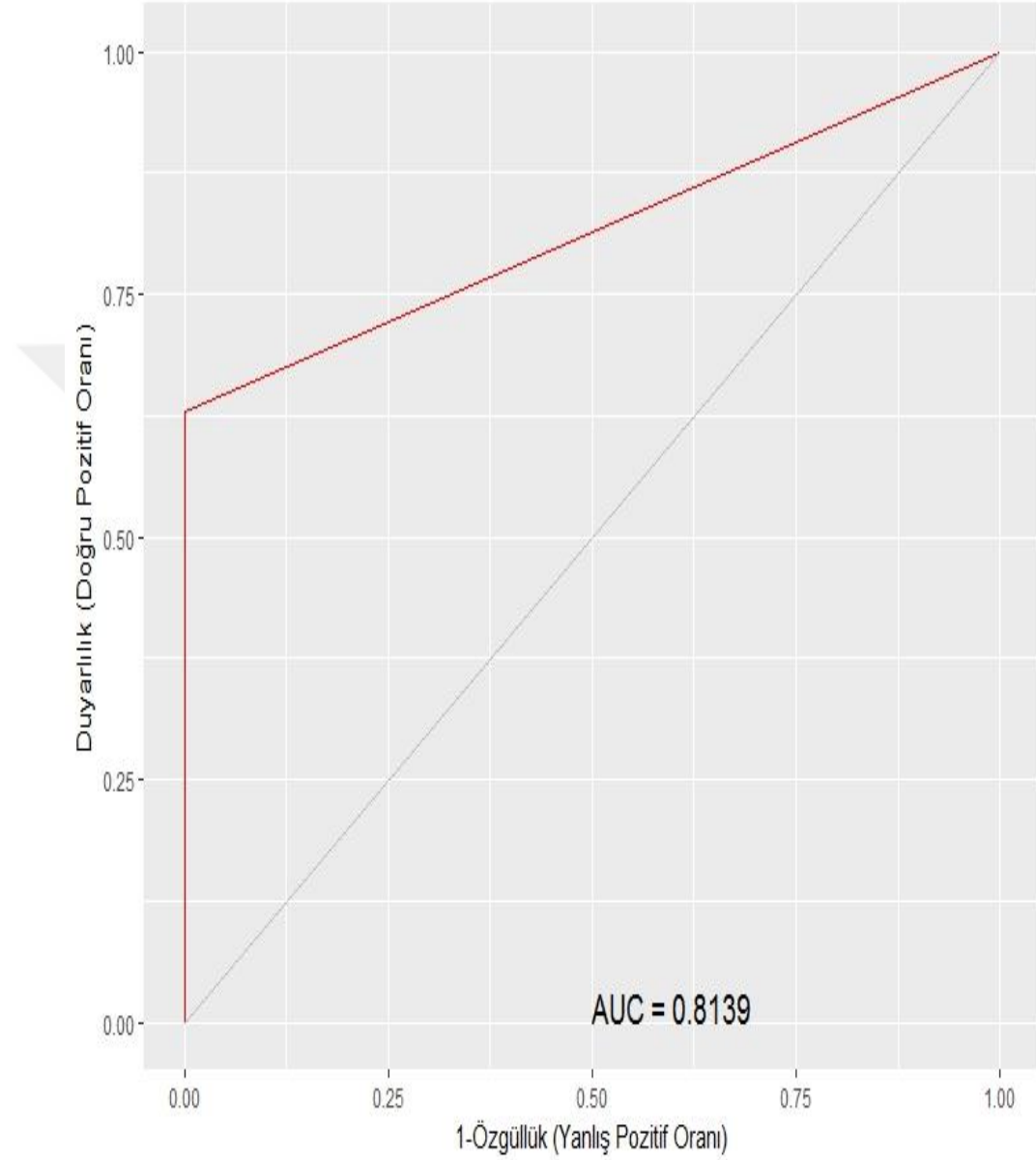
Şekil 4.7: Maliyet ve ölçek parametrelerine göre doğruluk metriğinin değişimi (Çekirdek fonksiyon= hiperbolik tanjant (sigmoid))

Şekil 4.8’de, Laplace çekirdek fonksiyonlu DVM modeli için ROC eğrisi ve ROC eğrisi altında kalan alan gösterilmiştir. Laplace çekirdek fonksiyonlu DVM modeli için ROC eğrisi altında kalan alan değeri 0.9332 (0.9096 - 0.9567) olarak hesaplanmıştır.



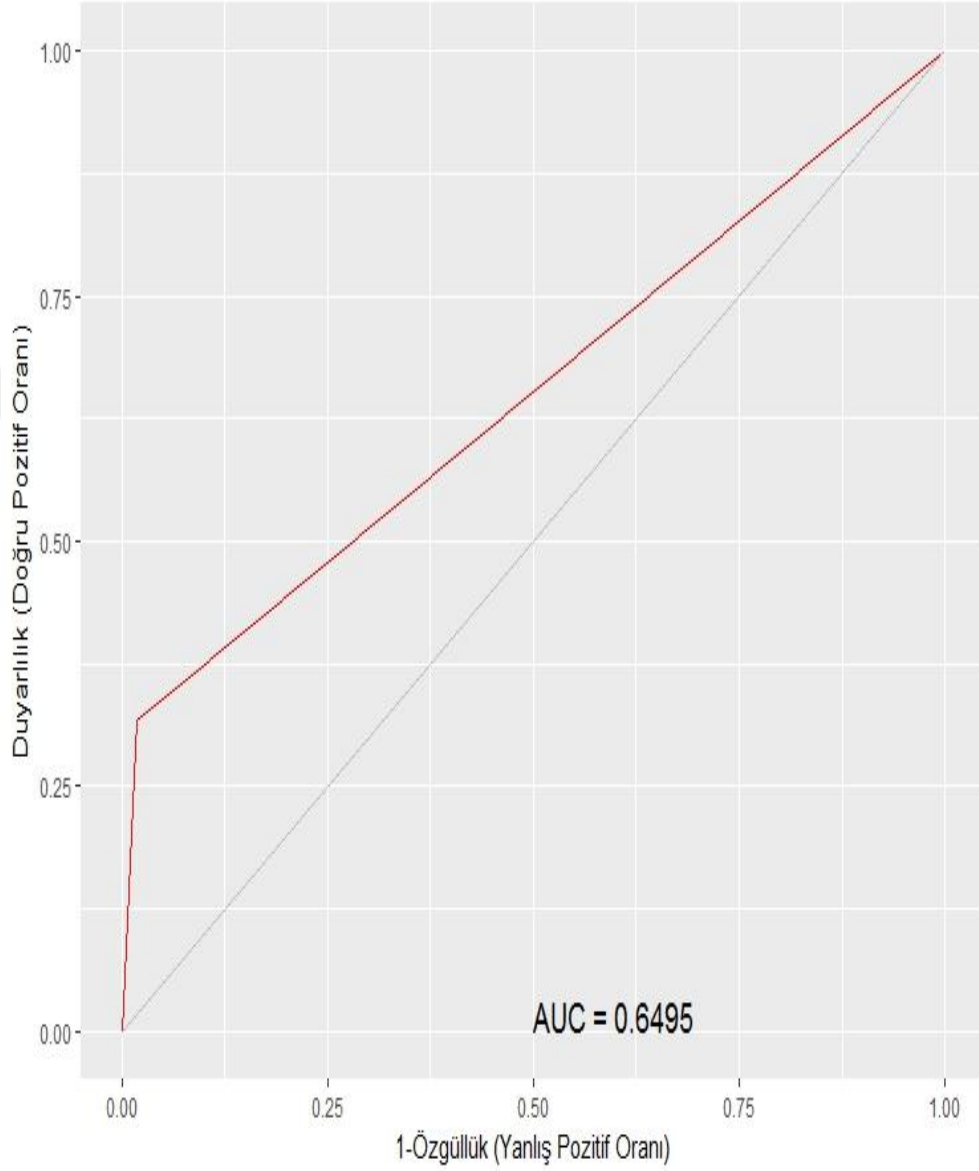
Şekil 4.8: DVM modelinde Laplace çekirdek fonksiyonu için ROC eğrisi grafiği

Şekil 4.9'da, Gaussian RBF çekirdek fonksiyonlu DVM modeli için ROC eğrisi ve ROC eğrisi altında kalan alan gösterilmiştir. Gaussian RBF çekirdek fonksiyonlu DVM modeli için ROC eğrisi altında kalan alan değeri 0.8139 (0.7805-0.8473) olarak hesaplanmıştır.



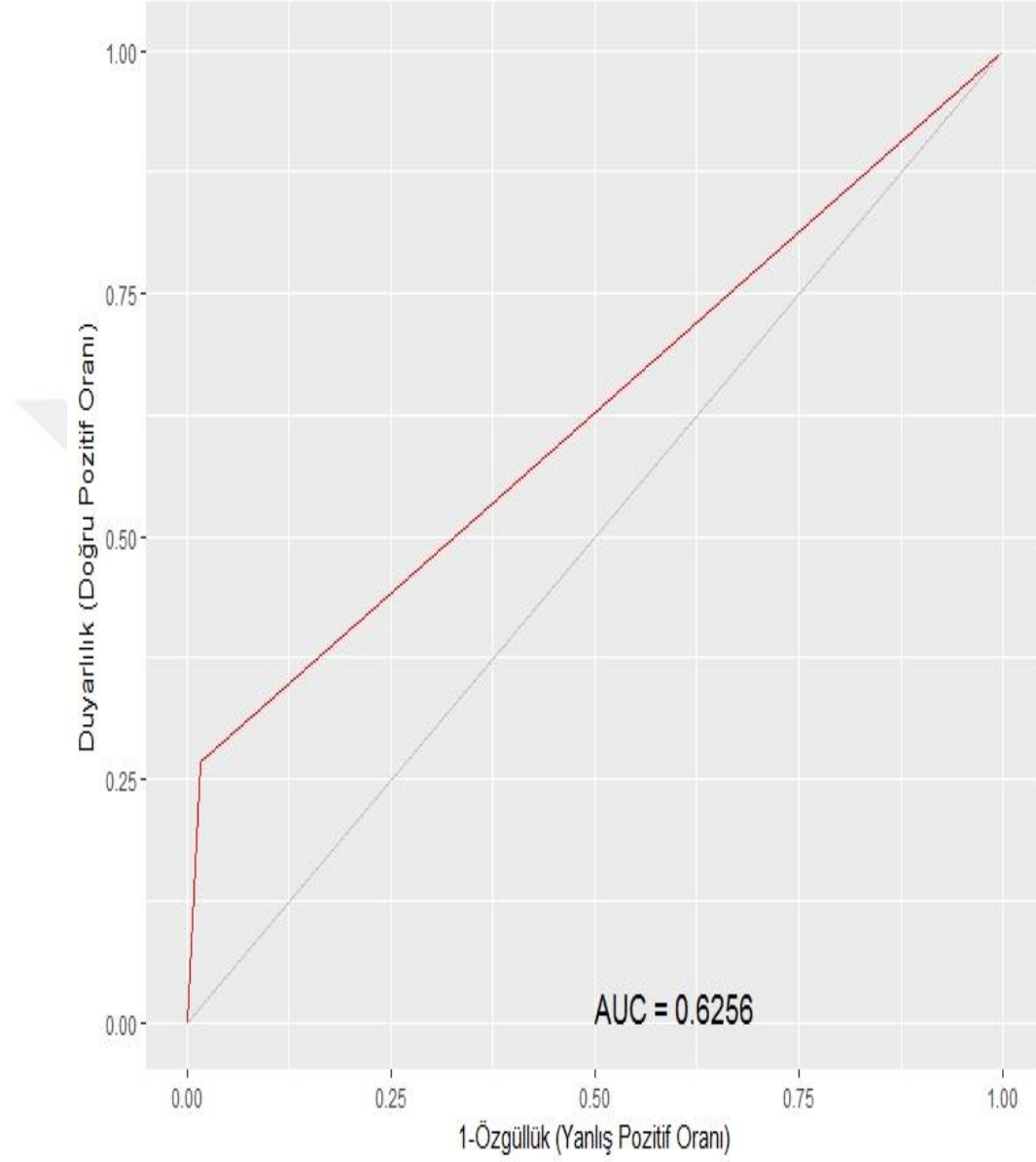
Şekil 4.9: DVM modelinde Gaussian RBF çekirdek fonksiyonu için ROC eğrisi grafiği

Şekil 4.10'da, Polinomiyal çekirdek fonksiyonlu DVM modeli için ROC eğrisi ve ROC eğrisi altında kalan alan gösterilmiştir. Polinomiyal çekirdek fonksiyonlu DVM modeli için ROC eğrisi altında kalan alan değeri 0.6495 (0.6171 - 0.6819) olarak hesaplanmıştır.



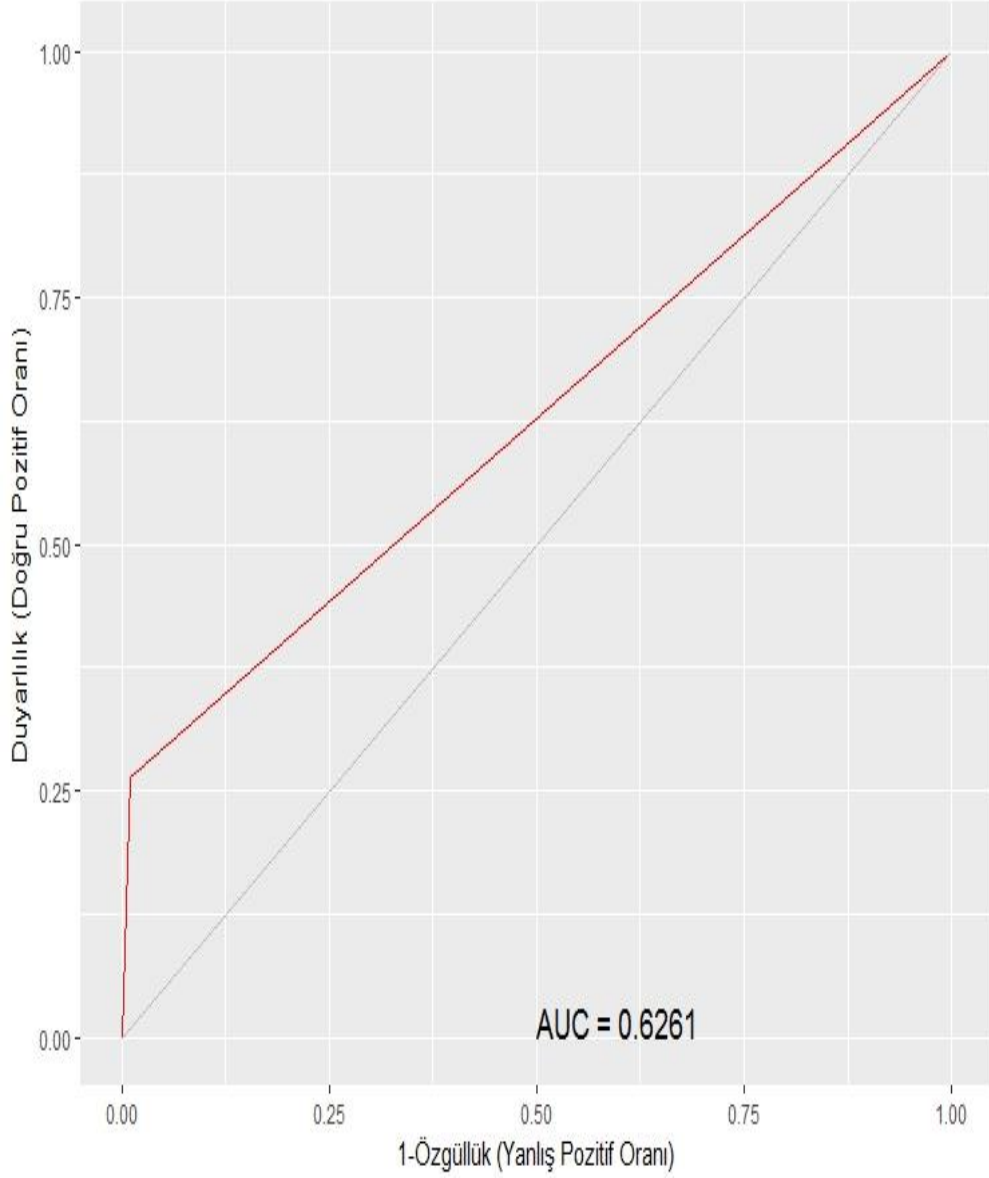
Şekil 4.10: DVM modelinde Polinomiyal çekirdek fonksiyonu için ROC eğrisi grafiği

Şekil 4.11’de, Doğrusal çekirdek fonksiyonlu DVM modeli için ROC eğrisi ve ROC eğrisi altında kalan alan gösterilmiştir. Doğrusal çekirdek fonksiyonlu DVM modeli için ROC eğrisi altında kalan alan değeri 0.6256 (0.5948-0.6564) olarak hesaplanmıştır.



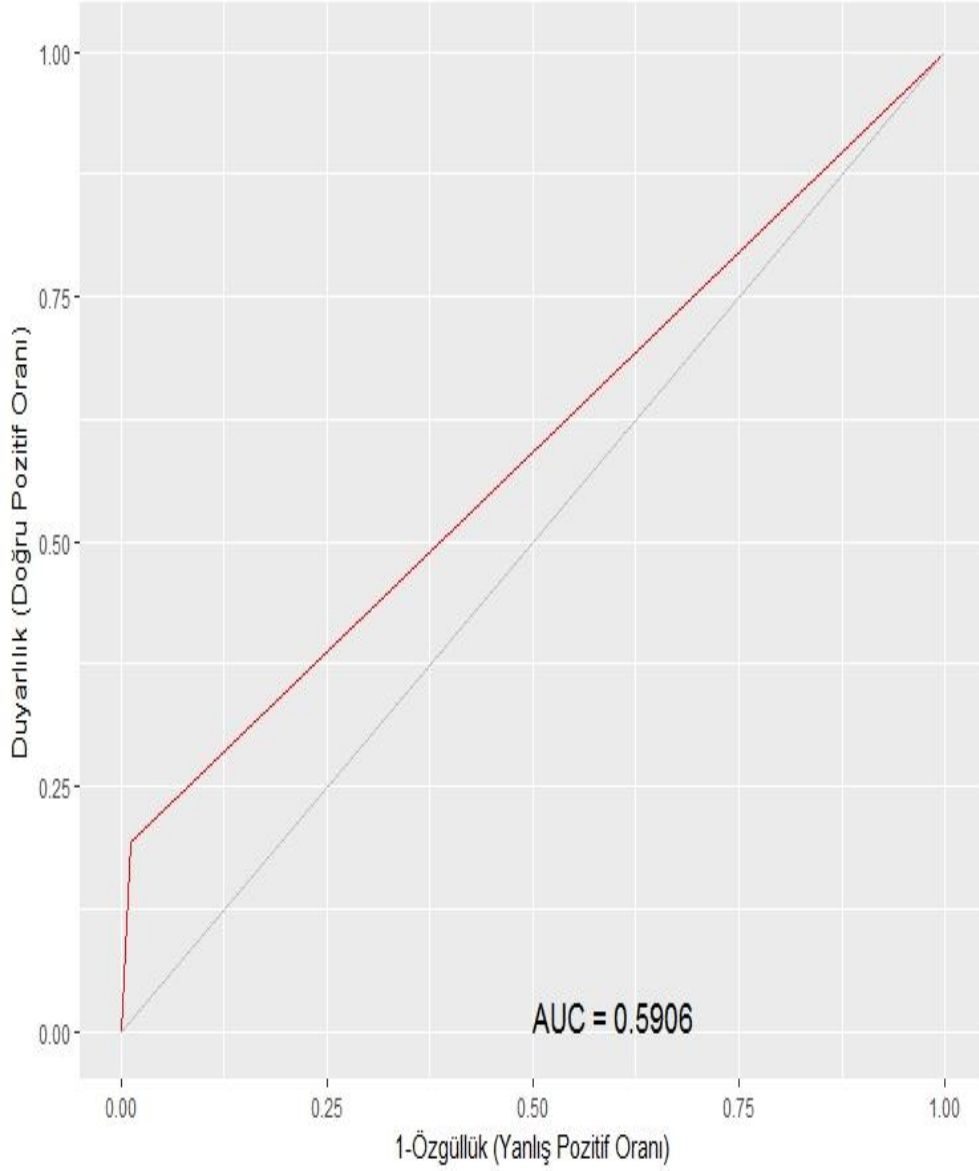
Şekil 4.11: DVM modelinde Doğrusal çekirdek fonksiyonu için ROC eğrisi grafiği

Şekil 4.12’de, Bessel çekirdek fonksiyonlu DVM modeli için ROC eğrisi ve ROC eğrisi altında kalan alan gösterilmiştir. Bessel çekirdek fonksiyonlu DVM modeli için ROC eğrisi altında kalan alan değeri 0.6261 (0.5955 - 0.6566) olarak hesaplanmıştır.



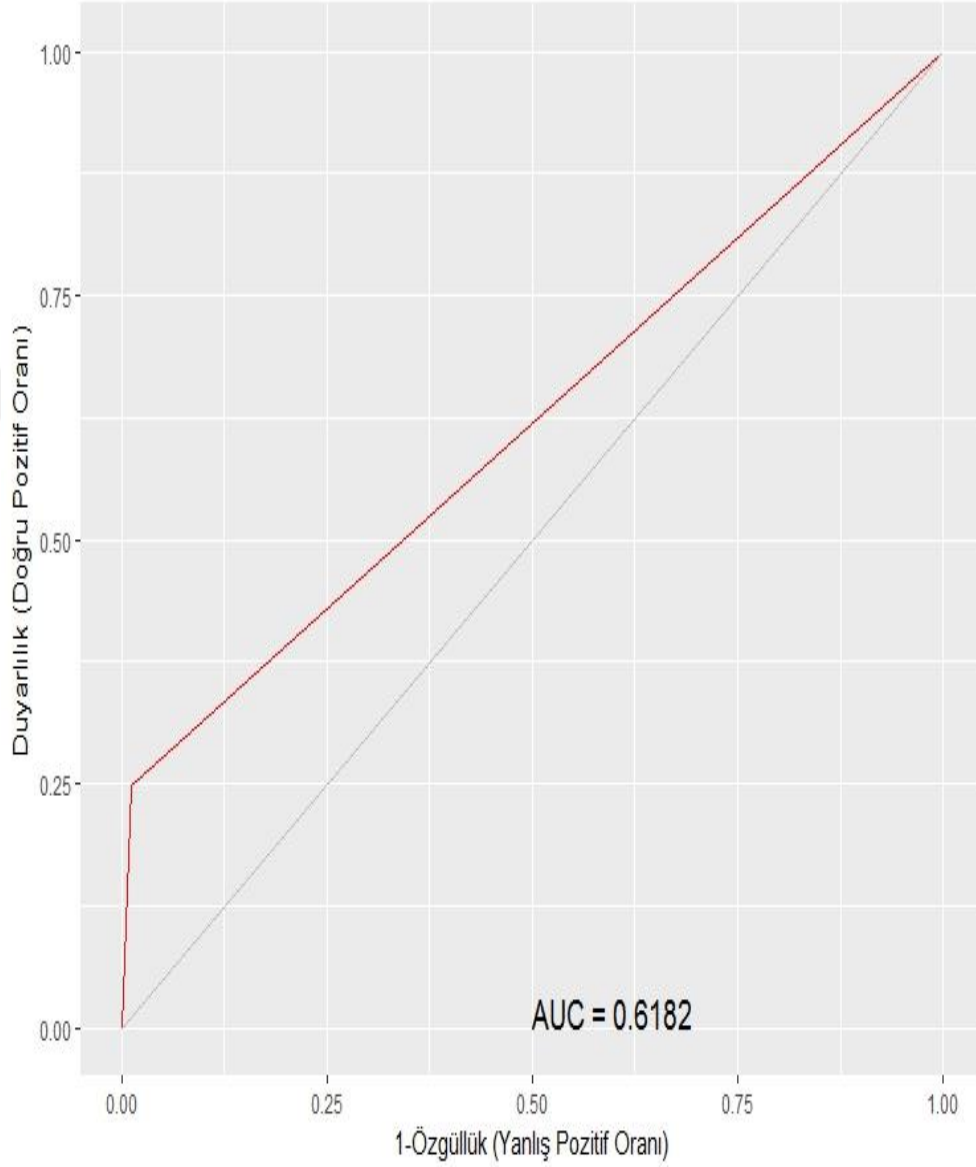
Şekil 4.12: DVM modelinde Bessel çekirdek fonksiyonu için ROC eğrisi grafiği

Şekil 4.13’de, Hiperbolik Tanjant (Sigmoid) çekirdek fonksiyonlu DVM modeli için ROC eğrisi ve ROC eğrisi altında kalan alan gösterilmiştir. Hiperbolik Tanjant (Sigmoid) çekirdek fonksiyonlu DVM modeli için ROC eğrisi altında kalan alan değeri 0.5906 (0.5631 - 0.618) olarak hesaplanmıştır.



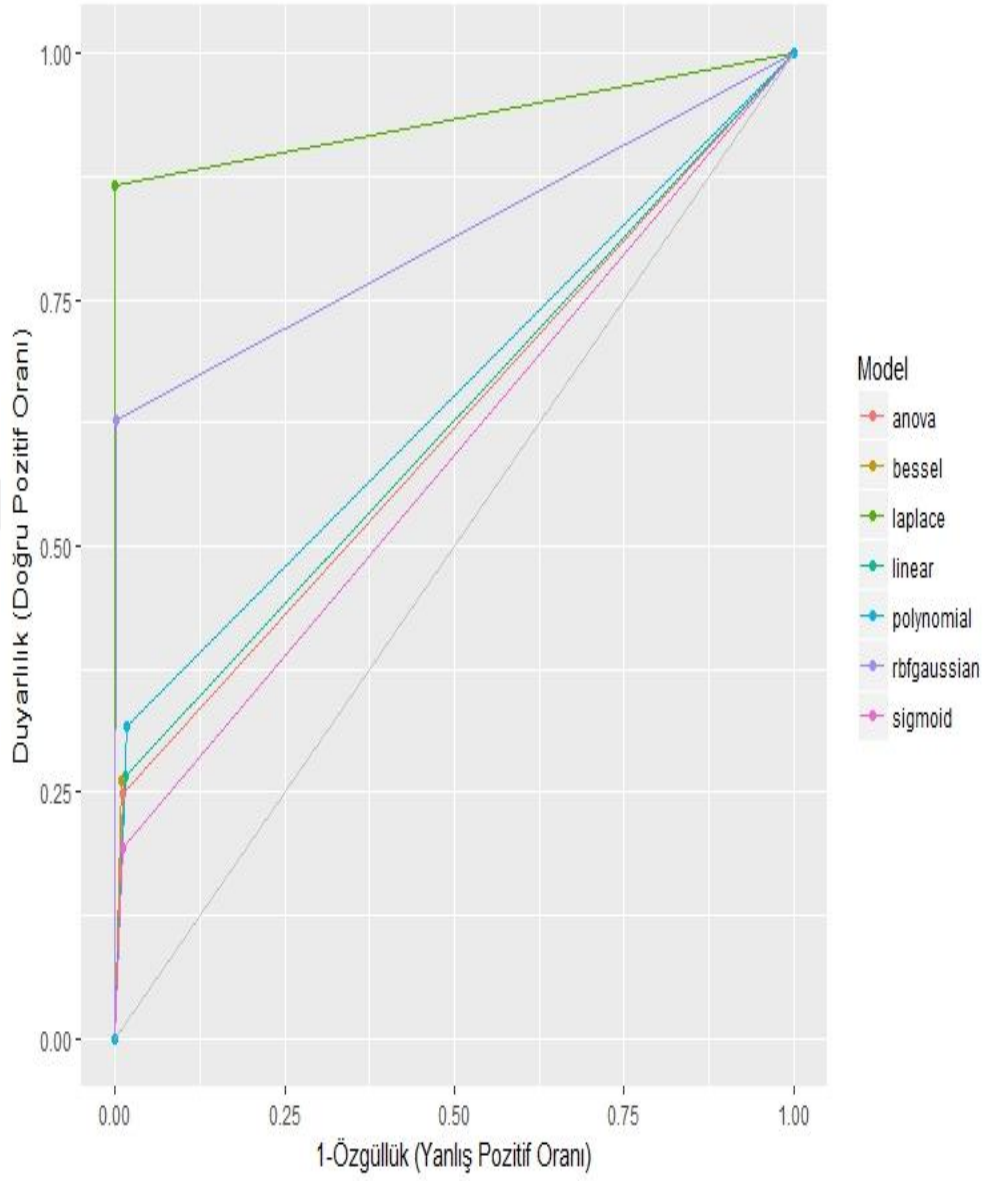
Şekil 4.13: DVM modelinde Hiperbolik Tanjant (Sigmoid) çekirdek fonksiyonu için ROC eğrisi grafiği

Şekil 4.14'de, ANOVA RBF çekirdek fonksiyonlu DVM modeli için ROC eğrisi ve ROC eğrisi altında kalan alan gösterilmiştir. ANOVA RBF çekirdek fonksiyonlu DVM modeli için ROC eğrisi altında kalan alan değeri 0.6182 (0.5883 - 0.6482) olarak hesaplanmıştır.



Şekil 4.14: DVM modelinde ANOVA RBF çekirdek fonksiyonu için ROC eğrisi grafiği

Şekil 4.15’de, tüm çekirdek fonksiyonları için karşılaştırmalı ROC eğrisi gösterilmiştir.

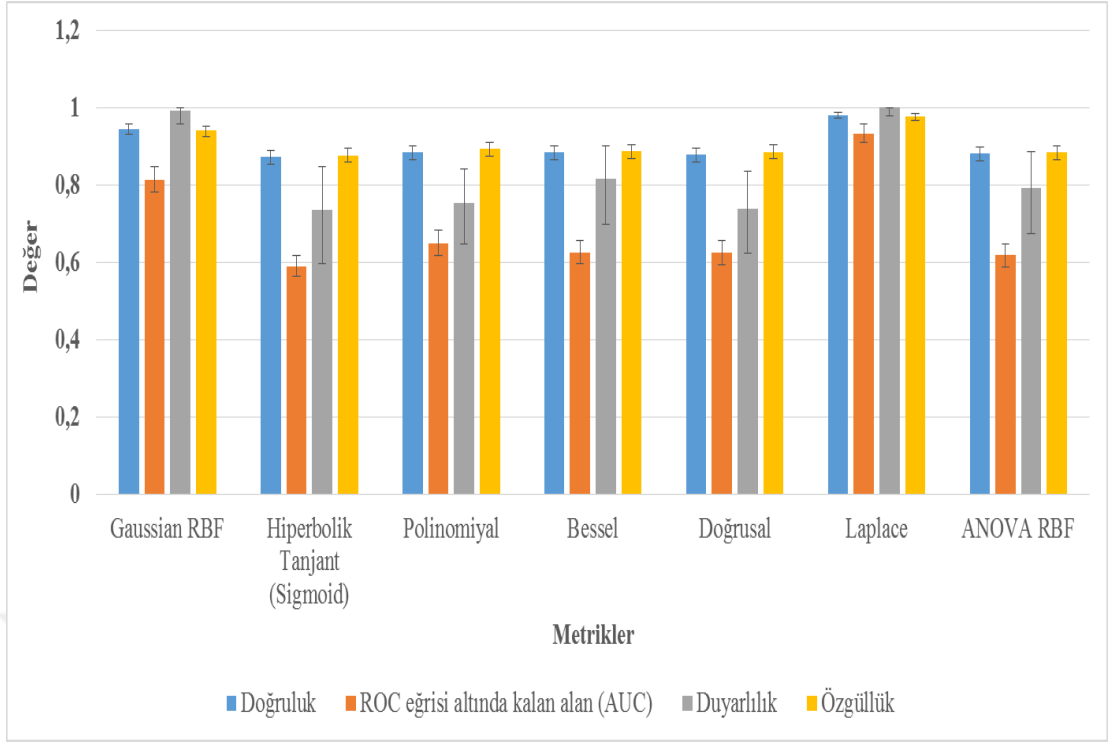


Şekil 4.15: DVM modelinde kullanılan tüm çekirdek fonksiyonları için ROC eğrisi grafiği

Şekil 4.16 ve Tablo 4.32’de, DVM’nin tüm çekirdek fonksiyonlarının çeşitli performans metriklerine göre sınıflandırma performansları (% 95 güven aralığı ile) detaylı bir şekilde verilmiştir.

Tablo 4.32: DVM'nin tüm çekirdek fonksiyonlarının çeşitli performans metriklerine göre sınıflandırma performansları (%95 güven aralığı ile)

Çekirdek Fonksiyonları	Doğruluk	ROC eğrisi altında kalan alan (AUC)	Duyarlılık	Özgüllük [Seçicilik]
Gaussian RBF	0.9448 (0.9315 – 0.9563)	0.8139 (0.7805-0.8473)	0.9922 (0.9572 – 0.9998)	0.94 (0.9254 - 0.9525)
Hyperbolic tangent (Sigmoid)	0.8716 (0.8527 - 0.8888)	0.5906 (0.5631 - 0.618)	0.7359 (0.5967 - 0.8474)	0.877 (0.8581 - 0.8942)
Polinomiyal	0.8846 (0.8666 - 0.901)	0.6495 (0.6171 - 0.6819)	0.7529 (0.6475 - 0.8401)	0.8933 (0.8752 - 0.9096)
Bessel	0.8832 (0.865 - 0.8997)	0.6261 (0.5955 - 0.6566)	0.8154 (0.6997 - 0.9008)	0.8865 (0.8681 - 0.9032)
Doğrusal	0.8788 (0.8604 - 0.8956)	0.6256 (0.5948-0.6564)	0.73973 (0.6238 – 0.8355)	0.88569 (0.8681 – 0.9033)
Laplace	0.9804 (0.9716 - 0.987)	0.9332 (0.9096 - 0.9567)	0.9999 (0.9791 – 1.000)	0.9776 (0.9675 – 0.9852)
ANOVA RBF	0.8803 (0.8619 - 0.8969)	0.6182 (0.5883 - 0.6482)	0.7937 (0.6730 - 0.8853)	0.8844 (0.8659 - 0.9012)



Şekil 4.16: DVM'nin tüm çekirdek fonksiyonlarının çeşitli performans metriklerine göre sınıflandırma performansları (% 95 güven aralığı ile)

Tablo 4.33’de çalışmada kullanılan değişkenlerin laplace çekirdek fonksiyonuyla oluşturulan DVM modeline göre önemlilik dereceleri verilmiştir. Hesaplanan önemlilik dereceleri, (0, 100) aralığında normalize edilmiştir. Laplace DVM modelinden hesaplanan önemliliği en yüksek olan değişken Glikoz için %100’dür. Diğer değişkenler için önemlilik değerleri sırasıyla; BUN için %58.41, kreatinin için %55.6, aspirin için %51.86, hipertansiyon için %49.75, diüretik için %42.91, anjiyotensin dönüştürücü enzim inhibitör için %40.04, vücut kitle indeksi için %38.08, trigliserid için %36.29, beta bloker için %36.04, anjiotensin reseptör blokerleri diğer için %34.75, statin için %33.99, ME için %33.23, akut koroner sendrom tipi için %31.64, kolesterol için %31.01, KAH için %30.93, renal yetmezlik için %30.78, hiperlipidemi için %30.60, SKB için %30.53, kalp yetmezliği için %29.46, DKB için %29.41, PAH için %28.74, malignite için %27.92, klopidogrel için %27.28, Ca KBLK için %27.17, CABG için %27.05, inme öyküsü için %26.88, aile öyküsü için %22.73, HDL için %20.44, nabız için %20.41, LDL için %20.23 ve cinsiyet için %9.18 bulunmuştur.

Tablo 4.33: Çalışmada kullanılan değişkenlerin laplace çekirdek fonksiyonuyla oluşturulan DVM modeline göre önemlilik dereceleri

Değişken ismi	Değişken Önemliliği
Glikoz	100.00
BUN	58.41
Kreatinin	55.63
ASA	51.86
Hipertansiyon	49.75
Diüretik	42.91
ACE İnhibitör	40.04
VKİ	38.08
Trigliserid	36.29
Beta bloker	36.04
ARB diğer	34.75
Statin	33.99
ME	33.23
AKS tipi	31.64
Kolesterol	31.01
KAH	30.93
Renal yetmezlik	30.78
Hiperlipidemi	30.60
SKB	30.53
Kalp yetmezliği	29.46
DKB	29.41
PAH	28.74
Malignite	27.92
Klopidogrel	27.28

Ca KBLK	27.17
CABG	27.05
İnme öyküsü	26.88
Aile öyküsü	22.73
HDL	20.44
Nabız	20.41
LDL	20.23
Cinsiyet	9.18



5. TARTIŞMA

Klinik açıdan KAH ve trombotik komplikasyonları sanayileşmiş ülkelerde morbidite ve mortalitenin önde gelen nedenlerindedir. KAH oranının önümüzdeki on yılda hızlı bir şekilde artacağı ve yaşlanmaya sebep olacağı tahmin edilmektedir. Ayrıca, önümüzdeki yıllarda obezite, tip 2 DM ve metabolik sendromun prevalansında dünya çapında bir artış beklenmekle beraber, genç insanlar arasında kardiyovasküler risk faktörlerinin de önemli derecede artış göstermesi öngörülmektedir. Özellikle tip 2 DM olan bireylerde kardiyovasküler hastalık riski, DM olmayan bireylere kıyasla 2-4 kat daha fazladır. Tip 2 DM hastalarının %75'inde mortalite sebebi KAH'tır (57). Akut koroner sendromu, KAH'ın ilerlemiş bir klinik tablosu biçiminde meydana gelmektedir. Bu nedenle akut koroner sendromu ile ilişkili risk faktörlerinin belirlenmesi ve kontrol altına alınması, kardiyovasküler hastalıklarının önlenmesi (birincil koruma) ve tanısı konulmuş hastalıkların tekrarlama durumların önlenmesi (ikincil koruma) için oldukça önemlidir. Bu bağlamda, söz konusu çalışmada, Tip 2 DM olan akut koroner sendromlu bireylerden oluşan bir hasta örneğinde, KAH gelişmesini etkilediği kabul edilen DM'nin sınıflandırılması için değişik çekirdek fonksiyonları ile geliştirilen DVM modellerinin tahmin performanslarının değerlendirilmesi amaçlanmıştır.

Diyabet, insülin sekresyonu, insülin etkisi veya her ikisindeki kusurlarla sonuçlanan hiperglisemi ile tanımlanan bir metabolik hastalık grubudur. Diyabetin kronik hiperglisemi, uzun vadeli hasar, işlev bozukluğu ve farklı organların, özellikle gözlerin, böbreklerin, sinirlerin, kalp ve damarların yetmezliğiyle ilişkilidir. Birkaç patojenik süreç diyabet gelişiminde rol oynar. Pankreas β -hücrelerinin bu otoimmün yıkım aralıkları, dolayısıyla insülin eksikliği anormalliklerle insülin etkisine direnç gösterir. Diyabette karbonhidrat, yağ ve protein metabolizması anormalliklerinin temeli, hedef dokularda insülinin eksik etkisidir. Eksik insülin etkisi, yetersiz insülin sekresyonundan ve/veya hormon eyleminin karmaşık yollarındaki bir veya daha fazla noktada insüline verilen doku yanıtlarının azalmasından kaynaklanır. Diyabetin uzun vadeli komplikasyonları; görme kaybı olasılığı olan retinopati, böbrek yetmezliğine yol açan nefropati, ayak ülseri, amputasyonlar ve charcot eklemleri riski olan periferik nöropati, gastrointestinal, genitoüriner ve kardiyovasküler semptomlara ve cinsel işlev bozukluğuna neden olan otonomik nöropatidir. Diyabetli hastalarda aterosklerotik kardiyovasküler, periferik arteriyel ve serebrovasküler hastalık insidansı artmaktadır. Diyabetiklerde hipertansiyon

ve lipoprotein metabolizması anormallikleri genellikle diyabetlilerde bulunur. Diyabet vakalarının büyük çoğunluğu iki etyopatogenetik kategoriye girmektedir. Birinci kategoride, tip 1 DM yer almakta olup, bunun nedeni mutlak insülin salgı eksikliğidir. Bu tip DM'nin gelişme riski yüksek bireyler, pankreatik adacıklarda ve otoimmün patolojide genetik işaretlerle ortaya çıkan serolojik kanıtlarla tanımlanabilir. Diğeri ise çok daha yaygın kategori olan tip 2 DM olup, nedeni insülin hareketine direnç ve yetersiz telafi edici insülin salgısı tepkisidir. Tip 2 DM'de, bazı dokularda patolojik ve fonksiyonel değişikliklere neden olan bir miktar hipergliseminin varlığı yeterlidir (58). Bu çalışmada, koroner arter hastalarında tip 2 DM'nin sınıflandırılması için çeşitli çekirdek fonksiyonları ile oluşturulan destek vektör makinesi modellerinin performanslarının incelenmesi ve karşılaştırılması amaçlanmıştır. Bu kapsamda, özellikle koroner arter hastalarında tip 2 DM'nin tahmin edilmesi, koruyucu hekimlik uygulamaları açısından oldukça önemlidir (59).

Tip 2 DM insidansı dünya çapında artmaktadır. Tip 2 DM, genetik bir yatkınlık ile davranışsal ve çevresel risk faktörleri arasındaki etkileşim sonucu ortaya çıkmaktadır. Tip 2 diyabetin genetik temeli henüz kesin olarak belirlenmemiştir; ancak obezite ve fiziksel hareketsizlik gibi değiştirilebilir risk faktörlerinin temel nedenleri arasında olduğuna dair güçlü kanıtlar vardır (28).

Akut koroner sendromlu hastalarda, cinsiyet değişkeni açısından DM varlığı ve yokluğu karşılaştırıldığında, DM varlığı ve yokluğu arasında istatistiksel olarak farklılık vardı. AKS tipi değişkeni, Unstabl Anjina, STEMI ve NSTEMI kategorilerinden oluşmaktadır. AKS tipi değişkeni açısından incelediğinde, DM varlığı ve yokluğu arasında istatistiksel olarak farklılık bulunmaktaydı. Bu farklılığı oluşturan alt kategorilerin STEMI ve NSTEMI olduğu görülmektedir. Unstabl Anjina kategorisi açısından DM varlığı ve yokluğu arasında fark bulunmamaktaydı. Bu değişkenler dışında kalan diğer nitel değişkenlerin alt kategorileri var/yok şeklindedir. Bu değişkenlerden, ACE inhibitör, ARB Diğeri, ASA, Beta Bloker, Diüretik, Statin, HT, Kalp yetmezliği öyküsü, MI Öyküsü, Renal Yetmezlik, Sigara açısından DM varlığı ve yokluğu arasında istatistiksel olarak farklılık vardı. Kalsiyum Kanal Bloker, Clopidogrel, Aile Öyküsü, CABG öyküsü, Hiperlipidemi, KAH öyküsü, Malignite, PAH değişkenleri açısından DM varlığı ve yokluğu arasında istatistiksel olarak farklılık yoktu. Bu çalışma kapsamında incelenen değişkenler, tip 2 DM için literatürde belirtilen risk faktörleri ile benzerlik göstermektedir (60-66).

Bu arařtırmada, performans ölçütleri kullanılarak, sınıflama başarısı en yüksek Laplace DVM modelinden elde edilen deęişken önemlilikleri sırasıyla; glikoz (%100.00), kan üre nitrojen (BUN) (%58.41), kreatinin (%55.63), ASA (%51.86) olmak üzere en önemli ilk dört faktör arasında yer almıřtır.

Bu arařtırmanın temel amacı; çeřitli çekirdek fonksiyonları ile oluşturulan destek vektör makinesi modellerinin akut koroner sendromu olan bireylerde tip 2 DM'nin sınıflama performanslarının incelenmesi ve karşılaştırılmasıdır. Bu amaçla, ANOVA RBF, Bessel, doğrusal, Gaussian RBF, Laplace, polinomial ve sigmoid çekirdek fonksiyonları ile geliştirilen DVM modelleri elde edilmiřtir. Bu DVM modelleri arasında, DM'yi sınıflamada en yüksek doğruluk deęeri Laplace çekirdek fonksiyonu (0.9804) ile hesaplandı. Doğruluk deęeri dikkate alındığında, DM'yi sınıflamada en yüksek deęerler sırasıyla; Gaussian RBF (0.9448), Polinomial (0.8846), Bessel (0.8832), ANOVA RBF (0.8803), Hyperbolic tangent (Sigmoid) (0.8716) ve doğrusal (0.8788) olarak elde edilmiřtir. ROC eğrisi altında kalan alan açısından tahminlenen deęişik DVM modelleri deęerlendirildiğinde, çekirdek fonksiyonlarına iliřkin en yüksek deęerler sırasıyla; Laplace (0.9804), Gaussian RBF (0.8139), Polinomial (0.6495), Bessel (0.6261), doğrusal (0.6256), ANOVA RBF (0.6182) ve Sigmoid (0.5906) olarak bulunmuřtur. Performans ölçütleri arasından duyarlılık deęerleri büyükten küçüğe sıralandığında; Laplace (0.9999), Gaussian RBF (0.9922), Bessel (0.8154), ANOVA RBF (0.7937), Polinomial (0.7529), doğrusal (0.7397) ve Sigmoid (0.7359) olarak elde edilmiřtir. Ayrıca, oluşturulan modeller özgüllük [seçicilik] deęerleri açısından incelendiğinde, en yüksek özgüllük [seçicilik] deęerleri sırasıyla; Laplace (0.9776), Gaussian RBF (0.94), Polinomial (0.8933), Bessel (0.8865), ANOVA RBF (0.8844), doğrusal (0.8857) ve Sigmoid (0.877) olarak bulunmuřtur.

Sonuç olarak; incelenen deęişik çekirdek fonksiyonları ile oluşturulan modeller arasında söz konusu performans ölçütleri dikkate alındığında, en iyi sınıflama performansı Laplace Destek Vektör Makinesi modelinden elde edilmiřtir. İlerleyen çalışmalarda, farklı klinik verilerde deęişik çekirdek fonksiyonlu Destek Vektör Makinesi modelleri ile dięer makine öğrenmesi ya da veri madencilięi algoritmalarının kullanılması hastalıkların sınıflandırma başarısını artırabilecektir.

Yukarıda sözü edilen çalışmalar, akut koroner sendromu ya da koroner arter hastalıęının deęişik boyutlarda incelenmesinde veri madencilięi ve makine öğrenmesi

teknikleri ele almaktadır. Ancak, yapılan yayın taraması bulgularına göre, koroner arter hastalığının özel bir türü olan akut koroner sendromlu hastalardan oluşan bireylerde tip 2 diyabet mellitusun varlığı ya da yokluğunu tıbbi bilgi keşfi süreci ile inceleyen bir çalışmaya rastlanmamıştır. Söz konusu araştırma bu yönüyle ilgili literatüre önemli bir katkı sağlayacaktır.



6. SONUÇ VE ÖNERİLER

Sonuç olarak; incelenen değişik çekirdek fonksiyonları ile oluşturulan modeller arasında söz konusu performans ölçütlerinden; doğruluk, duyarlılık, özgüllük [seçicilik] ve ROC eğrisi altında kalan alan değerleri dikkate alındığında, en iyi sınıflama performansı Laplace Destek Vektör Makinesi modelinden elde edilmiştir. İlerleyen çalışmalarda, farklı klinik verilerde değişik çekirdek fonksiyonlu Destek Vektör Makinesi modelleri ile diğer makine öğrenmesi ya da veri madenciliği algoritmalarının kullanılması hastalıkların sınıflandırma başarısını artırabilecektir.

Bu çalışmanın ilerleyen aşamalarında, bu araştırma kapsamında incelenmeyen (değişik çekirdek fonksiyonları ile oluşturulmuş Destek Vektör Makinesi modelleri haricinde kalan) istatistiksel öğrenme, makine öğrenmesi ve veri madenciliği alanlarında sıklıkla kullanılan ve mevcut çalışmada ele alınan verisetinin yapısına uygun olan:

- Bayes temelli yaklaşımlardan; Naive Bayes, Gaussian Naive Bayes, Bayes İnanç Ağları (Bayesian Belief Network), Bayes Ağları,
- Yapay sinir ağları temelli yaklaşımlardan; Çok katmanlı algılayıcı, geriye yayımlı sinir ağları, radyal tabanlı fonksiyon ağı,
- Karar ağacı temelli yaklaşımlardan; Sınıflandırma ve Regresyon Ağaçları (CART), Tekrarlı İkilikçi Ağaç (Iterative Dichotomiser 3), C4.5, C5.0, Ki-Kare Otomatik İlişki Tarayıcısı (CHAID), M5, Koşullu Karar Ağaçları,
- Derin öğrenme temelli yaklaşımlardan; Derin Boltzmann Makinesi, Derin İnanç Ağları, Evrimci Sinir Ağları,
- Toplu öğrenme temelli yaklaşımlardan; Boosting, Bagging, AdaBoost, Gradyent Boosting Makineleri, Gradyent Boosted Regresyon Ağaçları, Rastgele Orman,
- Regresyon temelli yaklaşımlardan; Lojistik Regresyon modelleri, Probit Regresyon modelleri, Boosted Genelleştirilmiş Doğrusal Model, Boosted Doğrusal Model

yöntemleri uygulanarak sınıflama ve tahmin başarısı daha da artırılabilmesi açısından önerilebilir.

KAYNAKLAR

1. Sezgin AT, Yıldırım A, Müderrisoğlu H. Akut Koroner Sendromlar III: Kararsız Angina Pektoris/ST Yükselmesi Miyokard İnfarktüsü.
2. Amsterdam EA, Wenger NK, Brindis RG, Casey Jr DE, Ganiats TG, Holmes Jr DR, Jaffe AS, Jneid H, Kelly RF, Kontos MC. 2014 AHA/ACC Guideline for the management of patients with non-ST-elevation acute coronary syndromes: executive summary: a report of the American College of Cardiology/American Heart Association Task Force on Practice Guidelines. *Circulation* 2014, 130(25): 2354.
3. Grech ED, Ramsdale DR. Acute coronary syndrome: unstable angina and non-ST segment elevation myocardial infarction. *BMJ* 2003, 326(7401): 1259-61.
4. Organization WH. Global report on diabetes. 2016.
5. Organization WH. Diabetes Fact sheet N 312. 2012. *World Health Organization* 2010.
6. Fayyad U, Piatetsky-Shapiro G, Smyth P. From data mining to knowledge discovery in databases. *AI magazine* 1996, 17(3): 37.
7. Mahalakshmi M, Deepak R, Nedunchelian R. Enhanced Personalized Mobile Search Engine towards Secured Data Analysis. *Int J Innov Res Sci Eng Technol* 2014, 3(5): 12209-14.
8. Koyuncugil AS, Özgülbaş N. Veri Madenciliği: Tıp ve Sağlık Hizmetlerinde Kullanımı ve Uygulamaları. *Int J of Inf Tech* 2009, 2(2).
9. Colak C, Karaman E, Turtay MG. Application of knowledge discovery process on the prediction of stroke. *Comput Methods Programs Biomed* 2015, 119(3): 181-5.
10. Babič M, Jerman-Blažič B, editors. New cybercrime taxonomy of visualization of data mining process. Information and Communication Technology, Electronics and Microelectronics (MIPRO); 2016: IEEE.
11. Köktürk F, Ankaralı H, Sümbüloğlu V. Veri Madenciliği Yöntemlerine Genel Bakış. *Turkiye Klinikleri J Biostat* 2009, 1(1): 20-5.
12. Balaban ME, E. K. *Veri Madenciliği ve Makine Öğrenmesi Temel Algoritmaları ve R Dili ile Uygulamaları*, Çağlayan Kitabevi 2015.
13. <http://visualdatamining.blogspot.com.tr/>. 25 Nisan 2017

14. <http://enm.blogcu.com/veri-madenciligi-nedir-2/2793556>. 25 Nisan 2017
15. <http://blog.goldenhelix.com/goldenadmin/cross-validation-for-genomic-prediction-in-svs/>. 27 Nisan 2017
16. Temel GO, Erdoğan S, Ankaralı H. Sınıflama Modelinin Performansını Geliştirmede Yeniden Örnekleme Yöntemlerinin Kullanımı. *Int J Inf Tech* 2012, 5(3): 1-8.
17. Vapnik V. *The Nature of Statistical Learning Theory*, Springer Science & Business Media 2013.
18. Ayhan S, Erdoğan Ş. Destek vektör makineleriyle sınıflandırma problemlerinin çözümü için çekirdek fonksiyonu seçimi. *Eskişehir Osmangazi Üniversitesi İktisadi ve İdari Bilimler Dergisi* 2014, 9(1).
19. Boser BE, Guyon IM, Vapnik VN, editors. A training algorithm for optimal margin classifiers. Proceedings of the fifth annual workshop on Computational learning theory; 1992: ACM.
20. Aizerman M, Braverman EM, Rozonoer L. Theoretical foundations of potential function method in pattern recognition. *Autom Remote Control* 1964, 25917-36.
21. Yu W, Liu T, Valdez R, Gwinn M, Khoury MJ. Application of support vector machine modeling for prediction of common diseases: the case of diabetes and pre-diabetes. *BMC Med Inform Decis Mak* 2010, 10(1): 16.
22. Hofmann T, Scholkopf B, Smola AJ. Kernel Methods in Machine Learning. 2008, 1171-220.
23. Lin H-T, Lin C-J. A study on sigmoid kernels for SVM and the training of non-PSD kernels by SMO-type methods. *Neural Comput* 2003, 1-32.
24. Zeileis A, Hornik K, Smola A, Karatzoglou A. kernlab-an S4 package for kernel methods in R. *J Stat Softw* 2004, 11(9): 1-20.
25. Harrison RF, Kennedy RL. Artificial neural network models for prediction of acute coronary syndromes using clinical data from the time of presentation. *Ann Emerg Med* 2005, 46(5): 431-9.
26. Syed Z, Stultz CM, Scirica BM, Gutttag JV. Computationally generated cardiac biomarkers for risk stratification after acute coronary syndrome. *Sci Trans Med* 2011, 3(102): 102ra95-ra95.
27. Lavesson N, Halling A, Freitag M, Odeberg J, Odeberg H, Davidsson P, editors. Classifying the severity of an acute coronary syndrome by mining patient data.

The Swedish AI Society Workshop May 27-28; 2009 IDA; Linköping University; 2009: Linköping University Electronic Press.

28. Çolak MC, Çolak C, Kocatürk H, Sağıroğlu Ş, Barutçu İ. Predicting coronary artery disease using different artificial neural network models. *Anatol J Cardiol* 2008, 8(4): 249-54.
29. Tsipouras MG, Exarchos TP, Fotiadis DI, Kotsia AP, Vakalis KV, Naka KK, Michalis LK. Automated diagnosis of coronary artery disease based on data mining and fuzzy modeling. *IEEE J Biomed Health Inform* 2008, 12(4): 447-58.
30. Novobilski A, Fesmire FM, Sonnemaker D, editors. Mining Bayesian Networks to Forecast Adverse Outcomes Related to Acute Coronary Syndrome. FLAIRS Conference; 2004.
31. Alizadehsani R, Habibi J, Hosseini MJ, Mashayekhi H, Boghrati R, Ghandeharioun A, Bahadorian B, Sani ZA. A data mining approach for diagnosis of coronary artery disease. *Comput Methods Programs Biomed* 2013, 111(1): 52-61.
32. Kim J, Lee E, Lee T, Sohn A. Economic burden of acute coronary syndrome in South Korea: a national survey. *BMC Cardiovasc Disord* 2013, 13(1): 55.
33. Jilani TA, Yasin H, Yasin M, Ardil C. Acute coronary syndrome prediction using data mining techniques-an application. *World Acad Sci Eng Technol* 2009, 59(4): 295-9.
34. Çolak C, Cengiz Çolak M, N. Orman M. The comparison of logistic regression model selection methods for the prediction of coronary artery disease. *Anatol J Cardiol* 2007, 7(1): 6-11.
35. Berikol GB, Yildiz O, Özcan İT. Diagnosis of Acute Coronary Syndrome with a Support Vector Machine. *J Med Syst* 2016, 40(4): 84.
36. Minitab I. MINITAB statistical software. *Minitab Release* 2015, 16.
37. Alpar R. *Uygulamalı istatistik ve geçerlik-güvenirlik: spor, sağlık ve eğitim bilimlerinden örneklerle*, Detay Yayıncılık 2016.
38. Sümbüloğlu V, Sümbüloğlu K. *Klinik saha araştırmalarında örnekleme yöntemleri ve örneklem büyüklüğü*, Vildan Sümbüloğlu 2005.
39. Breunig MM, Kriegel H-P, Ng RT, Sander J, editors. LOF: identifying density-based local outliers. *ACM Sigmod Record*; 2000: ACM.

40. Lee J, Kang B, Kang S-H. Integrating independent component analysis and local outlier factor for plant-wide process monitoring. *J Process Control* 2011, 21(7): 1011-21.
41. Pathak J, Bailey KR, Beebe CE, Bethard S, Carrell DS, Chen PJ, Dligach D, Endle CM, Hart LA, Haug PJ. Normalization and standardization of electronic health records for high-throughput phenotyping: the SHARPN consortium. *J Am Med Inform Assoc* 2013, 20(e2): e341-e8.
42. Çaycı Ö. *PHP ve MySQL*, Seçkin Yayıncılık 2003: 368.
43. Nixon R. *Learning PHP, MySQL, JavaScript, and CSS: A step-by-step guide to creating dynamic websites*, O'Reilly Media, Inc. 2012.
44. Aslan E, Durmaz F. Rehabilitasyon Amaçlı Bilgisayar Veri Tabanı Yardımıyla Bölgesel Engelli Kişi Haritasının Oluşturulması. *CBÜ Soma Meslek Yüksekokulu Teknik Bilimler Dergisi* 2011, 1(15): 64-73.
45. Nixon R. *Learning PHP, MySQL, JavaScript, CSS & HTML5: A Step-by-Step Guide to Creating Dynamic Websites*, O'Reilly Media, Inc. 2014.
46. Eskidere Ö. A Comparison of Feature Selection Methods for Diagnosis of Parkinson 's disease From Vocal Measurements. *Sigma* 2012, 30402-14.
47. Ben-Hur A, Horn D, Siegelmann HT, Vapnik V. Support vector clustering. *J Mach Learn Res* 2001, 2(Dec): 125-37.
48. Chen S, Yu P. Support vector networks puning on fllood forecasting and data mining. *J Hydro* 2007, 34767-78.
49. Eminoğlu U. Matlab GUI Kullanılarak Rüzgâr Türbini Sistemleri İçin Tasarım Optimizasyonu Paketinin Geliştirilmesi Developing of Design Optimization Package for Wind Turbine Systems By Using Matlab GUI.
50. Şeker AA, Hoccoğlu MH. Dağıtım Sistemine Yerel Elektrik Santrallerinin Yerleştirilmesi. 2014.
51. Clune J, Beckmann BE, Ofria C, Pennock RT, editors. Evolving coordinated quadruped gaits with the HyperNEAT generative encoding. 2009 IEEE Congress on Evolutionary Computation; 2009 18-21 May 2009.
52. Hornby GS, Pollack JB. Creating high-level components with a generative representation for body-brain evolution. *Artif Life* 2002, 8(3): 223-46.
53. Cortes C, Vapnik V. Support-vector networks. *Machine learning* 1995, 20(3): 273-97.

54. Team RC. The R project for statistical computing. *Available at www R-project org/ Accessed May 2017, 012017.*
55. Kuhn M. Caret package. *J Stat Softw* 2008, 28(5): 1-26.
56. Corp I. IBM SPSS statistics for windows, version 24.0. IBM Corp Armonk, NY; 2017.
57. Aladağ A, Ergene M, Koru M, Yılmaz MS. Tip 2 Diyabetik Hastalarda Koroner Kalp Hastalığı ile İlişkili Faktörler.
58. Association AD. Diagnosis and classification of diabetes mellitus. *Diabetes Care* 2014, 37(Supplement 1): S81-S90.
59. Yeşilbursa D, Yuvaç U, Serdar A, Odabaşı A, Cordan J. Akut koroner sendrom ve kararlı anginada trombosit hacim ve sayısındaki değişiklikler. *Turkiye Klinikleri J Cardiol* 2003, 16(2): 93-6.
60. White WB, Cannon CP, Heller SR, Nissen SE, Bergenstal RM, Bakris GL, Perez AT, Fleck PR, Mehta CR, Kupfer S, Wilson C, Cushman WC, Zannad F. Alogliptin after Acute Coronary Syndrome in Patients with Type 2 Diabetes. *N Engl J Med* 2013, 369(14): 1327-35.
61. Gress TW, Nieto FJ, Shahar E, Wofford MR, Brancati FL. Hypertension and Antihypertensive Therapy as Risk Factors for Type 2 Diabetes Mellitus. *N Engl J Med* 2000, 342(13): 905-12.
62. Bonora E, Kiechl S, Willeit J, Oberhollenzer F, Egger G, Meigs JB, Bonadonna RC, Muggeo M. Population-based incidence rates and risk factors for type 2 diabetes in white individuals. *Diabetes* 2004, 53(7): 1782-9.
63. Kempf K, Herder C, Erlund I, Kolb H, Martin S, Carstensen M, Koenig W, Sundvall J, Bidel S, Kuha S. Effects of coffee consumption on subclinical inflammation and other risk factors for type 2 diabetes: a clinical trial. *The American journal of clinical nutrition* 2010, 91(4): 950-7.
64. Rich-Edwards JW, Colditz GA, Stampfer MJ, Willett WC, Gillman MW, Hennekens CH, Speizer FE, Manson JE. Birthweight and the risk for type 2 diabetes mellitus in adult women. *Ann Int Med* 1999, 130(4_Part_1): 278-84.
65. Baptiste-Roberts K, Barone BB, Gary TL, Golden SH, Wilson LM, Bass EB, Nicholson WK. Risk factors for type 2 diabetes among women with gestational diabetes: a systematic review. *The American journal of medicine* 2009, 122(3): 207-14. e4.

66. Wannamethee SG, Shaper AG, Perry IJ. Smoking as a modifiable risk factor for type 2 diabetes in middle-aged men. *Diabetes Care* 2001, 24(9): 1590-5.



EKLER

EK-1. Özgeçmiş

26.04.1978 tarihinde Malatya'da doğdum. İlkokulu Mustafa Kemal Atatürk İlköğretim-Derme İlköğretim okullarında, ortaokulu Malatya Anadolu Lisesi'nde ve lise eğitimimi Malatya Fen Lisesi-Malatya Lisesi'nde tamamladım. 2001 yılında Orta Doğu Teknik Üniversitesi Bilgisayar Mühendisliği Bölümünden mezun oldum. 2001 yılında İnönü Üniversitesi Enformatik Bölümünde bilgisayar okutmanı olarak göreve başladım. 2005 yılında İnönü Üniversitesi Sağlık Bilimleri Enstitüsü, Biyoistatistik ve Tıp Bilişimi Anabilim Dalında yüksek lisansımı tamamladım. 2009 yılında beri İnönü Üniversitesi Turgut Özal Tıp Merkezi Bilgi İşlem Müdürlüğünde çalışmaktayım. Evli ve 2 çocuk babasıyım. İyi derecede İngilizce bilmekteyim.

EK-2. Etik Kurul Onay Formu

KLİNİK ARAŞTIRMALAR ETİK KURULU KARAR FORMU

ARAŞTIRMANIN AÇIK ADI	Çeşitli Çekirdek Fonksiyonları ile Oluşturulan Destek Vektör Makinesi Modellerinin Performanslarının İncelenmesi: Bir Klinik Uygulama.
VARSA ARAŞTIRMANIN PROTOKOL KODU	2016/161

ETİK KURUL BİLGİLERİ	ETİK KURULUN ADI	MALATYA KLİNİK ARAŞTIRMALAR ETİK KURULU
	AÇIK ADRESİ:	İnönü Üniversitesi Merkez Kampüsü, 44280, Malatya, Türkiye
	TELEFON	+90 422 341 06 60 / 1219
	FAKS	+90 422 341 00 36
	E-POSTA	inu.dhek@inonu.edu.tr

BAŞVURU BİLGİLERİ	KOORDİNATÖR/SORUMLU ARAŞTIRMACI UNVANI/ADI/SOYADI	Doç. Dr. Cemil Çolak			
	KOORDİNATÖR/SORUMLU ARAŞTIRMACININ UZMANLIK ALANI	İnönü Üniversitesi Tıp Fakültesi Biyoistatistik ve Tıp Bilişimi AD			
	KOORDİNATÖR/SORUMLU ARAŞTIRMACININ BULUNDUĞU MERKEZ	MALATYA			
	VARSA İDARİ SORUMLU UNVANI/ADI/SOYADI				
	DESTEKLEYİCİ				
	PROJE YÜRÜTÜCÜSÜ UNVANI/ADI/SOYADI (TÜBİTAK vb. gibi kaynaklardan destek alanlar için)				
	DESTEKLEYİCİNİN YASAL TEMSİLCİSİ				
	ARAŞTIRMANIN FAZİ VE TÜRÜ	FAZ 1	<input type="checkbox"/>		
		FAZ 2	<input type="checkbox"/>		
		FAZ 3	<input type="checkbox"/>		
FAZ 4		<input type="checkbox"/>			
Gözlemsel ilaç çalışması		<input type="checkbox"/>			
Tıbbi cihaz klinik araştırması		<input type="checkbox"/>			
İn vitro tıbbi tanı cihazları ile yapılan performans değerlendirme çalışmaları		<input type="checkbox"/>			
İlaç dışı klinik araştırma	<input type="checkbox"/>				
DİĞER İSE BELİRTİNİZ					
ARAŞTIRMAYA KATILAN MERKEZLER	TEK MERKEZ <input type="checkbox"/>	ÇOK MERKEZLİ <input type="checkbox"/>	ULUSAL <input type="checkbox"/>	ULUSLARARASI <input type="checkbox"/>	

Etik Kurul Başkanının
Unvanı/Adı/Soyadı: Prof. Dr. Rıfat KARLIDAĞ
İmza:

Not: Etik kurul başkanının her sayfada imzasının olması gerekmektedir.

KLİNİK ARAŞTIRMALAR ETİK KURULU KARAR FORMU

ARAŞTIRMANIN AÇIK ADI	Çeşitli Çekirdek Fonksiyonları ile Oluşturulan Destek Vektör Makinesi Modellerinin Performanslarının İncelenmesi: Bir Klinik Uygulama.
VARSA ARAŞTIRMANIN PROTOKOL KODU	2016/161

DEĞERLENDİRİLEN BELGELER	Belge Adı	Tarihi	Versiyon Numarası	Dili
	ARAŞTIRMA PROTOKOLÜ			Türkçe <input type="checkbox"/> İngilizce <input type="checkbox"/> Diğer <input type="checkbox"/>
	BİLGİLENDİRİLMİŞ GÖNÜLLÜ OLUR FORMU			Türkçe <input type="checkbox"/> İngilizce <input type="checkbox"/> Diğer <input type="checkbox"/>
	OLGU RAPOR FORMU			Türkçe <input type="checkbox"/> İngilizce <input type="checkbox"/> Diğer <input type="checkbox"/>
	ARAŞTIRMA BROŞÜRÜ			Türkçe <input type="checkbox"/> İngilizce <input type="checkbox"/> Diğer <input type="checkbox"/>
DEĞERLENDİRİLEN DİĞER BELGELER	Belge Adı	Açıklama		
	SIGORTA	<input type="checkbox"/>		
	ARAŞTIRMA BÜTÇESİ	<input type="checkbox"/>		
	BIYOLOJİK MATERYEL TRANSFER FORMU	<input type="checkbox"/>		
	İLAN	<input type="checkbox"/>		
	YILLIK BİLDİRİM	<input type="checkbox"/>		
	SONUÇ RAPORU	<input type="checkbox"/>		
	GÜVENLİLİK BİLDİRİMLERİ	<input type="checkbox"/>		
DİĞER:	<input type="checkbox"/>			
KARAR BİLGİLERİ	Karar No:2016/161	Tarih:27.07.2016		
	Yukarıda bilgileri verilen başvuru dosyası ile ilgili belgeler araştırmacı/çalışmanın gereke, amaç, yaklaşım ve yöntemleri dikkate alınarak incelenmiş ve uygun bulunmuş olup araştırmacı/çalışmanın başvuru dosyasında belirtilen merkezlerde gerçekleştirilmesinde etik ve bilimsel sakınca bulunmadığına toplantıya katılan etik kurul üye tam sayısının salt çoğunluğu ile karar verilmiştir. İlaç ve Biyolojik Ürünlerin Klinik Araştırmaları Hakkında Yönetmelik kapsamında yer alan araştırmalar/çalışmalar için Türkiye İlaç ve Tıbbi Cihaz Kurumu'ndan izin alınması gerekmektedir.			
KLİNİK ARAŞTIRMALAR ETİK KURULU				
ETİK KURULUN ÇALIŞMA ESASI	İlaç ve Biyolojik Ürünlerin Klinik Araştırmaları Hakkında Yönetmelik, İyi Klinik Uygulamaları Kılavuzu			
BAŞKANIN UNVANI / ADI / SOYADI:	Prof. Dr. Rifat KARLIDAĞ			

Unvanı/Adı/Soyadı	Uzmanlık Alanı	Kurumu	Cinsiyet	Araştırma ile ilişki	Katılım *	İmza
Prof. Dr. Rifat KARLIDAĞ	Psikiyatri	İnönü Üniversitesi Tıp Fakültesi	E <input checked="" type="checkbox"/> K <input type="checkbox"/>	E <input type="checkbox"/> H <input checked="" type="checkbox"/>	E <input checked="" type="checkbox"/> H <input type="checkbox"/>	
Prof. Dr. Metin GENÇ	Halk Sağlığı	İnönü Üniversitesi Tıp Fakültesi	E <input checked="" type="checkbox"/> K <input type="checkbox"/>	E <input type="checkbox"/> H <input checked="" type="checkbox"/>	E <input checked="" type="checkbox"/> H <input type="checkbox"/>	
Prof. Dr. Saim YOĞLU	Biyoistatistik	İnönü Üniversitesi Tıp Fakültesi	E <input checked="" type="checkbox"/> K <input type="checkbox"/>	E <input type="checkbox"/> H <input type="checkbox"/>	E <input type="checkbox"/> H <input type="checkbox"/>	
Prof. Dr. Türkan TOĞAL	Anesteziyoloji ve Rea.	İnönü Üniversitesi Tıp Fakültesi	E <input type="checkbox"/> K <input checked="" type="checkbox"/>	E <input type="checkbox"/> H <input checked="" type="checkbox"/>	E <input checked="" type="checkbox"/> H <input type="checkbox"/>	
Prof. Dr. İbrahim ŞAHİN	İç Hastalıkları	İnönü Üniversitesi Tıp Fakültesi	E <input checked="" type="checkbox"/> K <input type="checkbox"/>	E <input type="checkbox"/> H <input checked="" type="checkbox"/>	E <input checked="" type="checkbox"/> H <input type="checkbox"/>	
Prof. Dr. Sedat YILDIZ	Fizyoloji	İnönü Üniversitesi Tıp Fakültesi	E <input checked="" type="checkbox"/> K <input type="checkbox"/>	E <input type="checkbox"/> H <input type="checkbox"/>	E <input type="checkbox"/> H <input type="checkbox"/>	
Doç. Dr. Seda TAŞDEMİR	Tıbbi Farmakoloji	İnönü Üniversitesi Tıp Fakültesi	E <input type="checkbox"/> K <input checked="" type="checkbox"/>	E <input type="checkbox"/> H <input checked="" type="checkbox"/>	E <input type="checkbox"/> H <input type="checkbox"/>	

Etik Kurul Başkanının
Unvanı/Adı/Soyadı: Prof. Dr. Rifat KARLIDAĞ
İmza:

Not: Etik kurul başkanının her sayfada imzasının olması gerekmektedir.

KLİNİK ARAŞTIRMALAR ETİK KURULU KARAR FORMU

ARAŞTIRMANIN AÇIK ADI		Çeşitli Çekirdek Fonksiyonları ile Oluşturulan Destek Vektör Makinesi Modellerinin Performanslarının İncelenmesi: Bir Klinik Uygulama.							
VARSA ARAŞTIRMANIN PROTOKOL KODU		2016/161							
Doç. Dr. Derya DOĞAN	Çocuk Sağlığı ve Hast.	İnönü Üniversitesi Tıp Fakültesi	E <input type="checkbox"/>	K <input checked="" type="checkbox"/>	E <input type="checkbox"/>	H <input type="checkbox"/>	E <input type="checkbox"/>	H <input type="checkbox"/>	
Doç. Dr. Özden KAMIŞLI	Nöroloji	İnönü Üniversitesi Tıp Fakültesi	E <input type="checkbox"/>	K <input checked="" type="checkbox"/>	E <input type="checkbox"/>	H <input checked="" type="checkbox"/>	E <input checked="" type="checkbox"/>	H <input type="checkbox"/>	<i>[Signature]</i>
Doç. Dr. Hakan HARPUTLUOĞLU	Onkoloji	İnönü Üniversitesi Tıp Fakültesi	E <input checked="" type="checkbox"/>	K <input type="checkbox"/>	E <input type="checkbox"/>	H <input checked="" type="checkbox"/>	E <input checked="" type="checkbox"/>	H <input type="checkbox"/>	<i>[Signature]</i>
Yrd. Doç. Dr. Mehmet KARATAŞ	Tıp Tarihi ve Etik	İnönü Üniversitesi Tıp Fakültesi	E <input checked="" type="checkbox"/>	K <input type="checkbox"/>	E <input type="checkbox"/>	H <input type="checkbox"/>	E <input type="checkbox"/>	H <input type="checkbox"/>	<i>[Signature]</i>
Dr. Mahmut Barkın AKGÜL	Tıp Doktoru	Halk Sağlığı Müdürlüğü	E <input checked="" type="checkbox"/>	K <input type="checkbox"/>	E <input type="checkbox"/>	H <input checked="" type="checkbox"/>	E <input checked="" type="checkbox"/>	H <input type="checkbox"/>	<i>[Signature]</i>
Metin TAY	Eczacı	Serbest Eczacı	E <input checked="" type="checkbox"/>	K <input type="checkbox"/>	E <input type="checkbox"/>	H <input type="checkbox"/>	E <input type="checkbox"/>	H <input type="checkbox"/>	<i>[Signature]</i>
Zafer ERGÜZEL	Hukuk	İnönü Üniversitesi	E <input checked="" type="checkbox"/>	K <input type="checkbox"/>	E <input type="checkbox"/>	H <input type="checkbox"/>	E <input type="checkbox"/>	H <input type="checkbox"/>	<i>[Signature]</i>
Hasan KONAN	Sivil Üye	MSD Ltd. Şti.	E <input checked="" type="checkbox"/>	K <input type="checkbox"/>	E <input type="checkbox"/>	H <input checked="" type="checkbox"/>	E <input checked="" type="checkbox"/>	H <input type="checkbox"/>	<i>[Signature]</i>

Etik Kurul Başkanının
Unvanı/Adı/Soyadı: Prof. Dr. Rifat KARLIDAĞ
İmza:

Not: Etik kurul başkanının her sayfada imzasının olması gerekmektedir.