



İSTANBUL MEDENİYET ÜNİVERSİTESİ
LİSANSÜSTÜ EĞİTİM ENSTİTÜSÜ
UYGULAMALI MATEMATİK VE HESAPLAMALI BİLİMLER
ANABİLİM DALI

**Destek Vektör Makinesi Sınıflandırma Algoritması için
Genişletilmiş Gauss Çekirdeği ve Uygulamaları**

Yüksek Lisans Tezi

Ayşenur Erdem

OCAK, 2023



İSTANBUL MEDENİYET ÜNİVERSİTESİ
LİSANSÜSTÜ EĞİTİM ENSTİTÜSÜ
UYGULAMALI MATEMATİK VE HESAPLAMALI BİLİMLER
ANABİLİM DALI

**Destek Vektör Makinesi Sınıflandırma Algoritması için
Genişletilmiş Gauss Çekirdeği ve Uygulamaları**

Yüksek Lisans Tezi

Ayşenur Erdem

Danışman

Doç. Dr. Betül Hiçdurmaz

OCAK, 2023

TEZ JÜRİSİ ONAYI

Ayşenur Erdem tarafından hazırlanan “Destek vektör makinesi sınıflandırma algoritması için genişletilmiş Gauss çekirdeği ve uygulamaları” başlıklı bu yüksek lisans tezi, Uygulamalı Matematik ve Hesaplamalı Bilimler Anabilim Dalında hazırlanmış ve jürimiz tarafından kabul edilmiştir.

JÜRİ ÜYELERİ

İMZA

Tez Danışmanı:

Doç. Dr. Betül HİÇDURMAZ
İstanbul Medeniyet Üniversitesi

Üyeler:

Prof. Dr. Rahmet Savaş
İstanbul Medeniyet Üniversitesi

Doç. Dr. Özgür Yıldırım
Yıldız Teknik Üniversitesi

Tez Savunma Tarihi: 26/01/2023

BEYANLAR

Yazım ve Kaynak Gösterme Kılavuzu Beyanı

Danışmanlığımda yazılan bu tezin APA yazım ve kaynak gösterme kılavuzunda belirtilen kurallara uygun olarak yapılandırıldığı ve bu kılavuzun metin içi kaynak gösterme standartlarının bu tezde tutarlı olarak uygulandığı tarafımdan incelenerek teyit edilmiştir.

İmza

Doç. Dr. Betül Hiçdurmaz

Etik İlkeler Sadakat Beyanı

Hazırladığım bu tezin tamamen kendi çalışmam olduğunu, akademik ve etik kuralları gözeterek çalıştığımı ve her alıntıya kaynak gösterdiğimi beyan ederim.

İmza

Ayşenur Erdem

ÖZET

Destek Vektör Makinesi Sınıflandırma Algoritması için Genişletilmiş Gauss Çekirdeği ve Uygulamaları

Erdem, Ayşenur

Yüksek Lisans Tezi, Uygulamalı Matematik ve Hesaplamalı Bilimler ABD

Danışman: Doç. Dr. Betül Hiçdurmaz

Ocak, 2023

Bu çalışmanın genel amacı makine destek vektör makinesi algoritmasında çekirdek fonksiyonu olarak kullanılmaya uygun ve etkili yeni çekirdek fonksiyonlarının Gauss çekirdeğinin Bernstein fonksiyonları yardımıyla genişletilmesi ile elde edilmesi ve yeni çekirdeklerin etkililiğinin ve teorik olarak çekirdek fonksiyonu olmaya uygunluğunun araştırılmasıdır.

Çalışmada, öncelikle denetimli öğrenme yöntemlerinden biri olan sınıflandırma problemleri üzerinde en etkili ve sıklıkla kullanılan destek vektör makinesi algoritması tanıtılmıştır. Sınıflandırma görevleri için kullanılan diğer tekniklerden bahsedilerek destek vektör makinelerinin avantajları ve matematiksel bir alt yapıya sahip olması bakımından öne çıktığı gösterilmiştir.

Destek vektör makinesi gibi pek çok makine öğrenmesi yönteminde matematiksel alt yapının kullanılması ile yöntemin etkililiğinin ve kapsam genişliğinin artırıldığı görülmektedir. Bu çalışma ile de destek vektör makinelerinin çekirdek fonksiyonlarının çeşitlendirilmesi için kullanılabilecek bir fonksiyonlar sınıfı olan Bernstein fonksiyonları ele alınmıştır. Bu çalışmada gösterilmiştir ki Bernstein fonksiyonları kullanılarak elde edilen tamamen monoton fonksiyonların her biri Mercer koşulu olarak isimlendirilen destek vektör makinesi için çekirdek olma şartını sağlamaktadır. Bu çalışmanın en önemli katkısı bu şekilde bir bakış açısı yardımıyla yeni çekirdek fonksiyonları oluşturmak için bir yöntem vermesidir.

Daha önce literatürde klasik tipte çekirdek fonksiyonları olan polinom, lineer, sigmoid çekirdek fonksiyonları sıklıkla kullanılmıştır. Ancak Gauss fonksiyonu çekirdekler içerisinde en sık kullanılan ve genel olarak veri setleri üzerinde en iyi sonuç veren bir fonksiyondur. Dolayısıyla çalışmanın sınırları üstel fonksiyonun üssü konumunda bulunan Bernstein fonksiyonları yardımıyla oluşturulan genişletilmiş Gauss çekirdeklerini kapsayacak şekilde çizilmiştir.

Bu çalışmada sunulan Bernstein fonksiyonları ile elde edilmiş genişletilmiş Gauss çekirdekleri Mercer koşulunu sağlama bakımından analiz edilmiştir. Sonrasında,

teorik sonuçlar uygulama veri setleri üzerinde yazılan Python kodları yardımıyla sayısal olarak analiz edilmiş ve yeni çekirdeklerin klasik Gauss çekirdeği ile karşılaştırarak etkililiği gösterilmiştir.

Anahtar Kelimeler: Sınıflandırma, Destek Vektör Makinesi, Denetimli Öğrenme, Gauss Çekirdeği, Bernstein Fonksiyonları

ABSTRACT

Extended Gauss Kernel and Its Applications for Support Vector Machine Classification Algorithm

Erdem, Ayşenur

Master' s Thesis, Applied Mathematics and Computational Sciences Department

Supervisor: Assoc. Dr. Betül Hiçdurmaz

January, 2023

The general aim of this study is to obtain effective new kernel functions suitable for use as kernel function in support vector machine algorithm by expanding the Gaussian kernel with the help of Bernstein functions and to investigate the effectiveness of new kernels and their suitability as kernel functions theoretically.

In the study, first of all, the most effective and frequently used Support Vector Machine algorithm on classification problems from supervised learning methods is introduced. By mentioning other techniques used for classification tasks, it has been shown that Support Vector Machines stand out in terms of their advantages and having a mathematical background.

In many machine learning methods such as Support Vector Machine, it is seen that the effectiveness and scope of the method are increased by the use of mathematical infrastructure. In this study, Bernstein functions, a class of functions that can be used to diversify the kernel functions of Support Vector Machines, are discussed. In this work, it is shown that each of the completely monotonic functions obtained using Bernstein functions satisfies the kernel condition for the support vector machine called the Mercer condition. The most important contribution of this study is that it provides a method for constructing new kernel functions with the help of such a perspective.

In the literature, classical type kernel functions such as polynomial, linear, sigmoid kernel functions have been used frequently. However, the Gaussian function is the most frequently used among kernels and generally gives the best results on datasets. Therefore, the boundaries of the study are drawn to cover the expanded Gaussian kernels created with the help of Bernstein functions, which are in the exponential position of the exponential function.

Extended Gaussian kernels obtained with the Bernstein functions presented in this study are analyzed in terms of meeting the Mercer condition. Afterwards, the theoretical results are analyzed numerically with the help of the Python codes on the classical datasets and the effectiveness of the new kernels was demonstrated.

Keywords: Classification, Support Vector Machine, Supervised Learning, Gaussian Kernel, Bernstein Functions

İÇİNDEKİLER

İÇİNDEKİLER.....	X
TABLolar DİZİNİ	XII
ŞEKİLLER DİZİNİ	XIII
1. GİRİŞ.....	1
1.1. ARAŞTIRMANIN AMACI VE ÖNEMİ.....	1
1.2. LİTERATÜR TARAMASI	5
2. MAKİNE ÖĞRENMESİ	9
2.1. DENETİMLİ (GÖZETİMLİ) ÖĞRENME	9
2.1.1. Regresyon	10
2.1.2. Sınıflandırma	11
2.2. DENETİMSİZ (GÖZETİMSİZ) ÖĞRENME.....	11
2.2.1. Kümeleme	12
2.3. SINIFLANDIRMA YÖNTEMLERİ.....	12
2.3.1. Karar Ağaçları	13
2.3.2. Naive Bayes	14
2.3.3. Lojistik Regresyon	15
2.3.4. Doğrusal Diskriminant Analizi.....	15
3. DESTEK VEKTÖR MAKİNESİ.....	17
4. ÇEKİRDEK (KERNEL) FONKSİYONU	22
5. GENİŞLETİLMİŞ GAUSS ÇEKİRDEĞİ FONKSİYONLARI.....	27
5.1 GAUSS ÇEKİRDEĞİ.....	27

5.2. BERNSTEİN FONKSİYONLARINDAN TÜRETİLEN ÇEKİRDEK FONKSİYONLARI	27
5.3. UYGULAMALAR.....	29
6. PYHTON İLE SINIFLANDIRMA UYGULAMALARI.....	36
7. SONUÇ VE TARTIŞMA.....	54
KAYNAKÇA	55
EKLER.....	62
EK 1- MATLAB KODLARI.....	62
EK-2- WİNE VE DİĞİTS VERİ SETİ İÇİN PYHTON KODLAR	64
ÖZGEÇMİŞ	65

TABLÖLAR DİZİNİ

Tablo 1 Geleneksel Çekirdekler Tablosu.....	23
Tablo 2 Yeni Çekirdekler Tablosu	24
Tablo 3 Fonk.I, Wine Veri Seti için ACC Tablosu	38
Tablo 4 Fonk.I, Wine Veri Seti için F1 Tablosu	40
Tablo 5 Fonk.I, Digits Veri Seti için ACC Tablosu	42
Tablo 6 Fonk.I, Digits Veri Seti için F1 Tablosu	43
Tablo 7 Fonk.II, Wine Veri Seti için ACC Tablosu	44
Tablo 8 Fonk.II, Wine Veri Seti için F1 Tablosu.....	45
Tablo 9 Fonk.II, Digits Veri Seti için ACC Tablosu.....	46
Tablo 10 Fonk.II, Digits Veri Seti için F1 Tablosu	47
Tablo 11 Fonk.III, Wine Veri Seti için ACC Tablosu	50
Tablo 12 Fonk.III, Wine Veri Seti için F1 Tablosu	51
Tablo 13 Fonk.III, Digits Veri Seti için ACC Tablosu	52
Tablo 14 Fonk.III, Digits Veri Seti için F1 Tablosu	53

ŞEKİLLER DİZİNİ

Şekil 1 Kümeleme Yönteminin Görselleştirilmesi.....	12
Şekil 2 Karar Ağacı Modeli	14
Şekil 3 Destek Vektörleri	21
Şekil 4 Fonk.I, Sıfır Civarında İncelenmesi.....	30
Şekil 5 Fonk.I, Sonsuz Civarında İncelenmesi	31
Şekil 6 Fonk.II, Sıfır Civarında İncelenmesi	32
Şekil 7 Fonk.II, Sonsuz Civarında İncelenmesi	32
Şekil 8 Fonk.III, Sıfır Civarında İncelenmesi	33
Şekil 9 Fonk.III, Sonsuz Civarında İncelenmesi.....	34
Şekil 10 Wine Veri Seti için Modelin Uygulanmasının Ekran Görüntüsü	41
Şekil 11 Digits Veri Seti için Modelin Uygulanmasının Ekran Görüntüsü	48

1. Giriş

1.1. Araştırmanın Amacı ve Önemi

Günümüz dünyasında teknolojinin gelişmesi özellikle de internet kullanımının artması nedeniyle saklanması gereken veri boyutları oldukça artmaktadır. Bu ise verinin çok daha büyük boyutlarda ulaşılabilir ve depolanabilir olma ihtiyacını doğurur. Amaç ise artan bu verilerden kullanışlı ve anlamlı olan bilgiyi elde etmektir. Bu verilerden elde edilen bilgiler, müşteri memnuniyetinin ya da yönelimlerinin anlaşılması, doğru yerde doğru reklamın yapılması, finansal tahminler, robotik vb. alanlarda sıklıkla kullanılmaktadır. Bu uygulamaların geliştirilmesi verinin fazlalaştırılması ile gerçekleşebildiği halde pek çok durumda verinin oluşturulan modelin iyileştirilmesine etkisi optimum bir seviyeye ulaştıktan sonra verinin arttırılması faydalı olmamaktadır. Büyük veri setlerinde ortaya çıkan diğer bir problem ise verinin güvenilirliğidir. Güvenilir verilerin seçilmesi bir çözüm olmakla birlikte bu durum model seçiminin önemini arttırmaktadır. Çünkü veri miktarı görece az ise bu durumda modelin tahminleme ve kararlılığın iyi olması büyük önem taşır (Özekeş, 2003). Bu tezde, bu açık probleme katkı olarak, büyük boyutlu olmayan veri setlerinde en sık kullanılan algoritma olan destek vektör makinelerinin iyileştirilmesi amaçlanmaktadır.

Veri biliminde, veriler işlenmeden önce kullanılan modeller matematiksel ve istatistiksel modelleri baz alarak oluşturulmaktaydı. Dolayısıyla modelin iyileştirilmesi ile elde edilen algoritmalar kullanılarak daha iyi ve daha hızlı sonuçlar elde edilmesi günümüzdeki aktif araştırma konularından biridir. Çoğunlukla oluşturulan algoritmanın iyileştirilmesi için matematiksel bir bakış açısı gerekir. Zira öncelikle modelin teorik yapısını anlamak ve modelde iyileştirme yapılabilecek yerleri tespit etmek için matematik bilgisi gerekir. “Veri bilimi için Matematik” başlığı altında incelenmekte olan veri bilimi modellerinin matematiğinin araştırılması ve iyileştirilmesi konuları şu an için matematik

araştırmacıları tarafından yeterince yaygın bir şekilde çalışılmamaktadır. Ancak konu hızla hak ettiği değeri kazanmaya doğru popüler bir yükseliş halindedir.

Makine öğrenmesi veri biliminin bir alt dalı olup, veri bilimi konularının önemli bir kısmı makine öğrenmesi kapsamına girmektedir. Makine öğrenmesi, bir bilgisayarın açıkça programlanmadan girdi verileriyle öğrenmesini sağlayan algoritmaları kullanıp kendi çıktısını seçmesidir (Öndeş, 2011). Bununla birlikte derin öğrenme de sıklıkla karşımıza çıkmaya başlayan bir terimdir (Gökalp, 2022). Ancak derin öğrenme algoritmaları makine öğrenmesinin bir alt başlığıdır ve bu tez kapsamında ele alınması planlanmamaktadır.

Makine Öğrenmesi problemleri için önerilen yaklaşımların bir kısmı tahminleme bir kısmı da sınıflandırma yapmaktadır. Makine öğreniminde, sınıflandırma problemleri için günümüzde birçok çalışma yapılmakta ve bunlar önemli bir yer tutmaktadır. Sınıflandırma problemlerine örnek olarak bankaların müşteri sınıflandırması ve tahmini, tıpta insanlara hangi hastalığın tanısının konulacağı, biyolojide yaprak görüntülerine göre bitki türlerinin belirlenmesi, sosyal medyada spam e-maillerin tanınması, endüstriyel üretimde kusurlu ürünlerin belirlenmesi sayılabilir (Makine Öğrenimi, t.y.).

Sınıflandırma problemi, basitçe eldeki verilere göre oluşturulan modelin ele alınan herhangi benzer örnek için bir sınıflandırma yapmasını sağlamak olarak tanımlanabilir. Makine Öğrenmesi eldeki verinin yapısına göre denetimli (gözetimli) ve denetimsiz (gözetimsiz) öğrenme olarak ikiye ayrılır. Bu çalışmada denetimli öğrenme problemleri üzerinde çalışarak çekirdeğin etkililiğinin araştırılması planlanmaktadır. Sınıflandırma için farklı yöntemler kullanılabilir. Bunlardan bazıları; karar ağaçları, yapay sinir ağları, Bayes sınıflandırıcılar, ilişki tabanlı sınıflandırıcılar, k-en yakın komşu yöntemi, destek vektör makineleri (DVM), genetik algoritmalar olarak sıralanabilir (Konular Veri Madenciliği, t.y.).

Bu tez kapsamında destek vektör makineleri metodu üzerinde iyileştirmeler araştırılacaktır.

Destek vektör makinesinde çekirdek fonksiyonunu kullanarak veriyi bir dönüşüme uğratmak ve böylece sınıflandırmayı daha doğru yapabilmek en genel mantıktır. Dolayısıyla bahsedilen dönüşümü gerçekleştiren çekirdek fonksiyonun seçimi önem taşımaktadır. Şimdiye kadar kullanılan standart çekirdek fonksiyonları polinom çekirdek fonksiyonu, lineer, sigmoid ve radyal tabanlı çekirdek fonksiyonlarıdır. Mevcut çalışmalarda radyal tabanlı fonksiyonlar en iyi sonuçları veren fonksiyonlardır (Cortes ve Vapnik, 1995).

Radyal tabanlı fonksiyonlar matematiksel olarak geniş bir fonksiyonlar sınıfını ifade etmektedir. Destek vektör makinesinde kullanılabilecek radyal tabanlı fonksiyonların bazı teorik şartları sağlaması gerekmektedir (Saatçioğlu, 2013).

Bu tez kapsamında standart olarak kullanılan radyal tabanlı fonksiyon olan Gauss fonksiyonu yerine genişletilmiş Gauss fonksiyonu önerilecektir. Genişletilmiş Gauss fonksiyonunun gerekli teorik şartları sağladığı gösterilecektir. Sonrasında teorik olarak inşa edilen genişletilmiş Gauss fonksiyonu tanımı üç farklı çekirdek fonksiyon elde edilmesi için uygulanacaktır. Bu çekirdek fonksiyonları hem teorik şartları sağlamakta hem de sayısal olarak klasik Gauss çekirdeği ile yarışabilecek kadar iyi performans göstermektedir. Oluşturulan genişletilmiş Gauss çekirdekleri standart Gauss fonksiyonu ile bazı klasik gerçek veri setleri üzerinde sayısal performans açısından karşılaştırılacaktır.

Kısaca bu tezde bir “Veri bilimi için Matematik” konusu olan sınıflandırma probleminde sıklıkla kullanılan destek vektör makineleri tekniğinin matematiksel bir bakış açısıyla açıklanması ve destek vektör makinelerinin klasik seçilen çekirdeklerinden farklı bir çekirdek olan genişletilmiş Gauss çekirdeğinin uygulanması ve etkililiğinin incelenmesi amaçlanmaktadır.

Tezin planı şu şekilde özetlenebilir: İlk bölüm giriş, destek vektör makinesi ve çekirdek fonksiyonları hakkında yapılan literatür taramasından oluşmaktadır. İkinci bölümde makine öğrenimi ve yöntemleri detaylı bir şekilde ele alınmıştır. Üçüncü bölümde bu tezdeki çalışmada kullanılacak olan destek vektör makinesi ile ilgili bilgiler verilmiştir. Dördüncü bölümde çekirdek fonksiyonu çeşitleri ve bağlantılı teoremler ele alınmıştır. Beşinci bölümde genişletilmiş Gauss çekirdekleri tanıtılmış ve uygulamaları yapılmış olup altıncı bölümde klasik veri setleri üzerinde Python uygulamaları yapılmıştır. Son başlık ise sonuçlar ve tartışma kısmıdır.

1.2. Literatür Taraması

Literatürdeki destek vektör makinesi uygulamaları; gen analizi, sinyal analizi, ses tanıma, görüntü sınıflandırma, yazı sınıflandırma gibi çok farklı alanları kapsamaktadır. Gao ve arkadaşları (2017), tarafından gerçekleştirilen çalışmada yüksek boyutta çalışılması, küçük örneklem boyutu ve gürültü fazlalığı nedeniyle, tüm gen seti ile yeterli kanser sınıflandırma doğruluğu elde etmenin zor olduğundan bahsedilmiştir. Bu nedenle yaptıkları çalışmada hibrit gen seçim yöntemi olan bilgi kazanç destek vektör makinesi önerilmiştir. Çalışmanın sonucu olarak üç gene dayalı olarak doğru bir şekilde sınıflandırılması zor olan kolon kanseri için %90,32'lik bir sınıflandırma doğruluğu elde edilmiştir.

Kuzu ve Yakut (2020), tarafından gerçekleştirilen çalışmada teknolojiyi farklı yoğunluklarda kullanan firmaların başarısız olmalarına sebep olan değişkeni tespit etmek için destek vektör makinelerinden faydalanılmıştır.

Kul ve Sayar (2021), tarafından yapılan çalışmada trafik akışını düzgün sağlayabilmek için son yıllarda sıklıkla kullanılan video kamera ve birtakım sensörlerin trafik yönetim ve gözetimindeki uygulamaları dikkate alınmıştır. Araçlar boyutlarına göre üç kategoriye ayrılması için kullanılan sürekli trafik görüntülerinde yapay sinir ağı, destek vektör makinesi ve adaboost sınıflandırıcıları ile veri eğitilmiş ve bunların performansları karşılaştırılmıştır.

Ayhan ve Erdoğan (2014), tarafından yapılan çalışmada çekirdek fonksiyonu seçiminin ve parametre optimizasyonunun destek vektör makinesi uygulamasında önemli bir yerinin olduğu belirtilmiştir. Çekirdek fonksiyonu seçerken tasarım olarak seçim süreci rassal blok deney temeline oturtulmuştur. Tek değişkenli varyans analizinden yola çıkılarak çekirdek fonksiyonunun seçimi gerçekleştirilmiştir. Analiz sonucunda rassal çekirdek fonksiyonunun diğer çekirdek fonksiyonları arasından en başarılı olduğu görülmüştür.

Eray (2008), tarafından yapılan çalışmada ses sistemleri incelenip, kullanılan yöntemler araştırılmıştır. 200'ü test 200'ü eğitim aşamasında kullanılmak üzere 400 sözcükten oluşan çalışmada iki çeşit destek vektör makinesi tasarlanmıştır. Birincisi soft marjin destek vektör makinesi, ikincisi least square destek vektör makinesi sınıflandırıcısıdır. Soft marjin sınıflandırıcısı %91 doğru tanıma başarısı elde ederken least square destek vektör makinesi %71 başarı elde etmiştir.

Kavzoglu ve Çölkesen (2010), tarafından yapılan çalışmada sınıflandırma yöntemlerinden olduğu bilinen destek vektör makinesi uydudan elde edilen görüntülerin sınıflandırılmasında çekirdek fonksiyonlarının etkisi incelenmiştir. Yapılan bu çalışmada çekirdek fonksiyonu ve parametre değerlerine bağlı değişimin destek vektör makinesinin sınıflandırılması üzerinden detaylı analizi yapılmıştır. Yapılan analizde en yaygın kullanılan dört çekirdek fonksiyonunun sınıflandırma doğruluğu üzerinde etkisi araştırılmış olup en çok benzerlik yöntemiyle en doğru sonucu destek vektör makinesinin verdiği görülmüştür.

Demirci (2007), çalışmasında karakter tanıma ile ilgili bir problemde destek vektör makinesinin başarısı gösterilmiştir. Bunun haricinde temel ve bağımsız bileşen analizi hakkında bilgi verilmiştir.

Küçüksille ve Ateş (2016), yapılan çalışmada sınıflandırma ve eğri uydurma problemlerinde yıllardır başarılı sonuç veren destek vektör makinesini kullanarak istenmeyen elektronik postaların filtrelenmesi işlemini yapmıştır.

Yakut ve ark. (2014), yapılan çalışmada destek vektör makinesi ve yapay sinir ağları kullanılarak Borsa İstanbul (BIST) endeksi ile ilgili etkili bir tahminin nasıl yapılacağını araştırmışlardır. Uygulamada BIST endeksinin en az üç gün öncesine ait değerleri ve diğer birkaç ülke borsasının 2005 ile 2012 seneleri arasındaki borsa verileri kullanılarak BIST endeks değeri destek vektör makinesi ve yapay sinir ağlarıyla tahmin edilmiştir.

Güran ve ark. (2014), alınan veya kullanılan ürünler üzerinden kişilerin günümüzde çok kullanılan Instagram, Twitter gibi sosyal iletişim araçları ile yaptıkları yorumların duygu ve düşünce analizini, bir sınıflandırma problemi haline getirmiştir. Bu çalışmada çeşitli veri setleri üzerinden destek vektör makinesi yöntemi kullanılarak duygu analizi gerçekleştirilmiştir. Farklı deneylerle parametrelerdeki değişimin sınıflandırma üzerindeki etkileri de incelenmiştir.

Zanbıl (2022), yaptığı çalışmada çok katmanlı algılayıcı ve destek vektör makinesi için parametre optimizasyonu yapmayı amaçlamıştır. Rastgele orman, Bayesian arama ve optuna gibi sınıflandırma çeşitlerine en iyileme yöntemlerinin katkısı araştırılmıştır. Üç veri seti kullanılmış sonucunda hem destek vektör makinesi hem de çok katmanlı için rastgele orman ve Bayesian arama sınıflandırma metotları, optuna yönteminden daha iyi bir performans sergilemişlerdir.

Köse (2019), tarafından yapılan çalışmada zeki optimizasyon tabanlı destek vektör makinesi kullanılarak diyabet hastalığının teşhisini gerçekleştirmek amaçlanmıştır. Bunun için beş tane zeki optimizasyon algoritması Gaussian -RBF çekirdek fonksiyonunu kullanan bir lineer olmayan destek vektör makinesini en iyileme amacıyla kullanılmıştır. Ortaya çıkan bilgilerle çeşitli algoritmalar kullanan hibrit şekilde oluşturulmuş sistemlerin, değişik düzeylerde başarılar gösterdiği görülmüştür. Zeki optimizasyon ile destek vektör makinesi diyabet hastalığı teşhisinde uygulandığında tutarlı ve yüksek başarılı sonuçlar elde edilmiştir. Bu yaklaşım ile diyabet teşhisinde yapay zekânın da kullanılacağı ve bunun yüksek potansiyele sahip bir konu olduğu teyit edilmiştir.

Kılıç ve arkadaşları (2018), yaptıkları çalışmada birçok kurumun kişisel tahminlerle yemek şirketlerinden talep ettikleri miktarın israfa yol açtığını tespit etmiş ve bir üniversite yemekhanesinin verileri üzerinde açık kaynak kodlu bir yazılımla yapay sinir ağları, destek

6vektör makinesi ve regresyon gibi analiz metotlarının kullanılması ile günlük talep edilen miktarın tahminini gerçekleştirmiştir. İlgili metotların performansları da karşılaştırılmıştır.

Okwuashi ve ark. (2020), tarafından yapılan çalışmada hiper spektral görüntü sınıflaması uygulaması için derin destek vektör makinesi yöntemi kullanılmıştır. Bahsedilen çalışmada kullanılan çekirdek fonksiyonları şunlardır; üstel, radyal ve Gauss temelli fonksiyon, nöral ve polinom fonksiyon. Derin destek vektör makinesi, destek vektör makinesi, derin sinir ağı, en yakın komşu ve k-ortalamlar yöntemlerinin karşılaştırılması sonucunda derin destek vektör makinesi yöntemi (%98,86; %98,17) en yüksek başarı elde eden sınıflandırma yöntemi olmuştur.

Shao ve arkadaşları (2020), yaptıkları çalışmada destek vektör makinesi ile bir model oluşturarak otel binalarının enerji tüketimlerini incelemiş ve analiz etmiştir. Destek vektör makinesinin çekirdek fonksiyonu olarak Gauss fonksiyonu seçilmiştir. Çekirdek fonksiyonu optimize edilerek model tahmininin doğruluğu arttırılmıştır.

Apaydın ve Aptikan (2021), tarafından yapılan çalışmada amaç Sentinel-2 uydusundan gelen görüntüleri kullanarak fındık bahçelerinin yerlerinin belirlenmesidir. Trabzon'un Beylikdüzü ilçesi merkezinde yapılan bu çalışmada piksel tabanlı görüntülerle araziler tespit edilmeye çalışılmıştır. Bunun için rastgele orman, destek vektör makinesi ve k-en yakın komşu sınıflandırma yöntemleri kullanılmıştır. Çalışma sonucunda destek vektör makinesinin diğer yöntemlere göre daha iyi çalıştığı görülmüştür.

2. Makine Öğrenmesi

Makine öğrenmesi, bilgisayar ve istatistik alanlarından faydalanan ve günümüzde oldukça kullanılan yapay zekanın bir alt dalıdır (Ercan, 2020). Makine öğrenmesi, eğitim verisi denilen çeşitli örnek veriyi ve algoritmaları kullanarak bir problemin çözümüne ilişkin öngöründe bulunmak veya karar vermek için matematiksel model oluşturan bir sistemdir (Küçük, 2019). Bu sistemin geleceğe yönelik tahminde bulunması beklenir. Makine öğrenimi terimi 1959 yılında telaffuz eden ilk kişi Amerikan bilgisayar bilimcisi Arthur Samuel'dir (Ercan, 2020).

Makine Öğrenmesi uygulamalarının gelişmesi günümüzde çeşitli öğrenme tekniklerini ortaya çıkarmıştır. Makine öğrenmesi denetimli öğrenme ve denetimsiz öğrenme şeklinde sınıflandırılabilir (Erdinç, 2019).

2.1. Denetimli (Gözetimli) Öğrenme

Denetimli öğrenme, sisteme yüklenen eğitim ve test verilerini kullanarak, sonuçta bir fonksiyon üretmeye çalışan bir makine öğrenmesi tekniğidir (Ercan,2020). Diğer bir ifadeyle denetimli öğrenme tekniğinde sisteme girilen verilerle sistemdeki çıktılar arasında bir ilişki kuran, eşleştirme yapabilen bir fonksiyon üretilir (Uzun, 2022). Makineye sunulan girdi bilgileriyle makinelerin ürettiği çıkış bilgileri istenen değerle karşılaştırılarak ara işlemde kullanılacak bilgi elde edilir (Taya, 2022).

Denetimli öğrenmede modelin parametreleri ele alınan veriye göre en az hatayı verecek şekilde güncellenerek belirlenir. Denetimli öğrenme, spam e-maillerin filtrelendirilmesi örneğinde olduğu gibi gerçek hayatta karşılaşılan problemlere çözümler sunmayı hedefler (Makine Öğrenmesi-IBM, t.y.).

Uygulama alanı çok geniş olan makine öğrenmesinde, tahmin ve kestirim gibi sonuçlara odaklanıp daha tutarlı sonuçlar vermesinden dolayı denetimli öğrenme algoritması daha sık kullanılmaktadır.

Denetimli öğrenme regresyon ve sınıflandırma olarak iki alt başlıkta ele alınır (Taya, 2022).

2.1.1. Regresyon

Regresyon, birden fazla değişkenin aralarındaki ilişkiyi belirleyen denetimli öğrenme yöntemlerinden biridir (Küçük, 2019). Başka bir deyişle kullanılmak istenen değişkenlerden matematiksel ifade tanımlamak için bir fonksiyon oluşturan analiz yöntemidir (Ercan, 2020).

Tek değişkenin olduğu regresyon tek değişkenli regresyon olarak adlandırılırken, çok değişken varsa çok değişkenli regresyon olarak adlandırılır. Değişkenlerden birinin bağımlı olup diğerinin bağımsız olarak adlandırıldığı regresyon tek değişkenli regresyondur. Diğer değişkenlerden etkilenen, bağımlı değişken olarak tanımlanabilir (Küçük, 2019).

Matematiksel olarak regresyon aşağıdaki gibi ifade edilir:

- Basit doğrusal regresyon modeli;

$$Y_i = \beta_0 + \beta_1 X_i + u_i, i=1,2,3 \dots n. \quad (1)$$

Basit doğrusal regresyon modelinde yer alan Y_i değişkeni bağımlı değişkendir. X_i değişkeni ise bağımsız değişken olarak adlandırılır. u_i bir hata terimi olarak adlandırılır. β_0 parametresi sabit parametre olup, doğrunun Y eksenini kestiği noktayı gösterir. β_1 parametresi eğim parametresidir. Bağımsız değişkendeki bir birimlik değişimin bağımlı değişkene olan etkisini belirtir. i her bir gözlemi ifade eden alt indeksleridir (Taya, 2022).

- Çoklu regresyon modeli:

$$Y_i = \beta_0 + \beta_1 X_{1i} + \beta_2 X_{2i} + \dots + \beta_p X_{pi} + u_i, \quad i=1,2,3,\dots,n. \quad (2)$$

Y_i değişkeni bağımlı değişkendir. $X_{1i}, X_{2i}, X_{3i}, X_{4i}, \dots, X_{pi}$, değişkenleri bağımsız değişkenlerdir. β_0 , bağımsız değişken değerleri sıfır iken bağımlı değişkenin hangi değeri alacağını gösteren sabit terimdir. $\beta_1, \beta_2, \dots, \beta_p$, kısmi regresyon katsayılarıdır. β_p , ise modelde p tane parametrenin ve $p-1$ tane açıklayıcı değişken bulunduğunu göstermektedir. u_i hata terimidir. i her bir gözlemi ifade eden alt indeksleridir (Taya, 2022).

2.1.2. Sınıflandırma

Sınıflandırma farklı veri sınıflarını gruplandırabilen bir denetimli öğrenme algoritmasıdır. Veriler sınıflandırılırken genellikle küçük veri boyutu kullanarak uygulamaya başlamak, modeli kurduktan sonra büyük verilerin daha iyi anlaşılıp çözümlenmesine yardımcı olmaktadır (Pekel, 2018). Denetimli öğrenme algoritmalarından ve istatistiksel öğrenme yöntemlerinden biri olan sınıflandırma, nitel verilere değerler atayarak işleme alan yöntemdir. Sınıflandırma başlığı adı altında birçok sınıflandırma yöntemi vardır. Karar ağaçları, Naive Bayes, diskriminant analizi, lojistik regresyon ve destek vektör makinesi temel sınıflandırma teknikleri olarak sayılabilir (Taya, 2022). Sınıflandırma yöntemlerinden ve destek vektör makinesi konusundan ileriki bölümlerde detaylı olarak bahsedilecektir.

2.2. Denetimsiz (Gözetimsiz) Öğrenme

Denetimsiz öğrenme, verilerdeki başlıca benzerlik, farklılık ya da uzaklıkları belirlemeye yarayan bir makine öğrenmesi yöntemidir (Taya, 2022). Makine çıktı verisi olmayan bu öğrenme yalnızca girdi verileriyle insan müdahalesi olmadan makine

öğrenmesinin algoritma çözümlemesini yapmak ve etiketsiz verileri kümelemek için kullanır (Makine Öğrenmesi-IBM, t.y.).

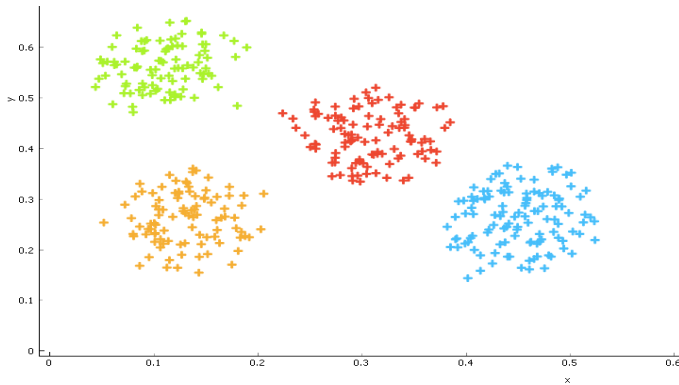
Denetimsiz öğrenme algoritmasını denetimli öğrenme algoritmasıyla kıyaslayacak olursak denetimsiz öğrenme daha karmaşık olan modelleri tanıyıp doğru gruplara ayırmaya çalışır.

2.2.1. Kümeleme

Kümeleme, benzer özellikteki ve farklı kategorideki verileri kendilerine en benzer olan kümelere ekleyerek birbirinden ayırmaktır. Düzenli verilerin oluşturulması ve modellerin tahmin edilmesinde kümeleme yöntemi veri işlemede bir ön aşama olarak da kullanılabilir. Kümeleme analizi denetimsiz öğrenme algoritmalarının temel başlığı olup, k-ortalımalı kümeleme, hiyerarşik kümeleme, olasılıklı kümeleme gibi birkaç alt başlıkta incelenebilmektedir (Taya, 2022).

Şekil 1

Kümeleme Yönteminin Görselleştirilmesi



2.3. Sınıflandırma Yöntemleri

Denetimli öğrenme yöntemlerinden olan sınıflandırma için literatür incelendiğinde birçok yöntemin mevcut olduğu görülür. Ancak destek vektör makinesinin diğer sınıflandırma yöntemlerine göre kullanımı daha kolaydır ve diğer metotlara nazaran başarı

oranı daha yüksektir. Ancak karşılaştırmaya imkan vermek açısından bu başlık altında diğer sınıflandırma yöntemlerinden kısaca bahsedilecektir.

2.3.1. Karar Ağaçları

Karar ağacı yöntemi, bir sınıflandırma yöntemi olup veri madenciliğinde yaygın olarak kullanılan tekniklerinden biridir. Karar ağaçlarının yapısının basit olmasının yanında oluşturulan sınıflandırma modelinin kolay anlaşılabilir olması, farklı sınıflandırma yöntemlerinden daha hızlı bir şekilde oluşturulması, karar ağaçlarının sınıflandırmada tercih edilmesinin sebebidir (Gehrke, 2003).

Bu yapının dezavantajı ise kullanılan veri setinin çok fazla sınıftan oluşması durumunda ortaya çıkar. Veri sayısının fazla olması durumunda, yapı enine doğru genişlemekte ve düğüm sayısı artmaktadır. Düğüm sayısının artması ise sınıf bilgisini azaltmakta ve bunun sonucunda sistemin güvenilirliğini düşürmektedir (Metlek & Kayaalp, 2020).

Karar ağacı, girdi ve çıktı verilerini kullanarak bir problemin çözümü için tümevarım yöntemiyle ağaç diyagramı oluşturan bir yapıdır. Karar ağacı kök, dal, yaprak ve bunlar arasındaki karar düğümlerinden oluşmaktadır (Metlek & Kayaalp, 2020; Tüminçin & Sadab, 2019). Ağaç yapısındaki en üstteki soruların sorulmaya başlandığı, girdisi olmayan ilk düğüm kök düğümüdür. Bu yapıda kendinden sonra çıkan birden fazla olan düğümler iç düğümleridir. Kendinden sonra dal oluşmamış düğümlere yaprak/sonuç düğümleri denir (Sayıcı, 2003).

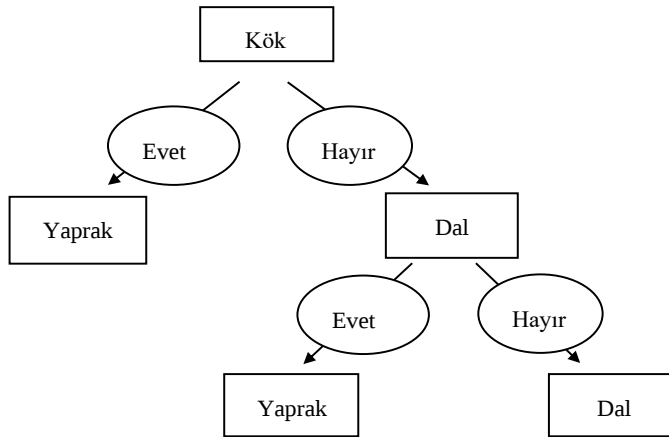
Karar ağacı ilk düğüm olan kök düğümünden başlayarak üstten alta doğru yapraklara ulaşmak için birbirini takip eden art arda gelen düğümleri takip eder. Oluşturulan ağacın dalları sınıflama işlemini tamamlamak için adaydır. Sınıflama işlemi bir dalın ucunda gerçekleşmezse, sonuçta bir karar düğümü oluşur. Fakat dalın ucunda bir sınıf oluşuyorsa, o

dalın ucunda yaprak oluşur. Oluşan bu yaprak, verinin belirlenen sınıflardan biridir (Özekeş, 2003).

Verilecek olan kararların sınıflarının her birini ağaçtaki yapraklar tarafından temsil edilir. Karar ağacında, kök düğümünden belli bir rota izlenerek yapraklara doğru götüren kararlar ve durumlar (rassal olaylar) dizisi yol (path) olarak adlandırılır ve her yaprağa (son noktaya) giden sadece tek bir yol bulunur. Bir yolun oluşumu sorulan sorulara verilen yanıtlara bağlıdır. Bu sayede kuralların oluşumu gerçekleşir. Ağaçta birden fazla yaprak aynı sınıflandırmaya sahip olabilir, lakin daha önce belirtildiği gibi takip edilen her yol tek bir yaprağa özgüdür (Sayıcı, 2003).

Şekil 2

Karar Ağacı Modeli



2.3.2. Naive Bayes

Naive Bayes, sınıflandırılması gereken sınıfların (kümelerin) ve örnek verilerin hangi sınıfa ait olduğunu olasılıksal olarak belirleyen bir sınıflandırma yöntemidir (Mitchell,1997). Büyük veriler için oldukça kullanışlı bir yöntemdir (Küçük, 2019). Bayes teoremlerini kullanarak sınıflandırma yapan Naive Bayes yöntemi, elde edilen olasılık

değerlerine göre değişkenin ait olduğu sınıfa karar verir (Pekel, 2018). Naive Bayes, modelleme için kullanıldığında herhangi bir Bayes yöntemine başvurmadan, parametre tahmini için en üst düzey olasılık tahminini kullanan hızlı ve kolay eğitilebilen bir yöntemdir.

$$P\left(\frac{C_j}{x}\right) = \frac{P\left(\frac{x}{C_j}\right) \cdot P(C_j)}{P(x)} \quad (3)$$

$P(C_j)$: C_j 'nin olma olasılığıdır.

$P\left(\frac{x}{C_j}\right)$: Bir olayın x olma olasılığıdır.

$P\left(\frac{C_j}{x}\right)$: x olan bir örneğin sınıf j 'den gelme olasılığıdır.

$P(C_j)$: Sınıf j 'nin ilk olasılığını temsil eder.

$P(x)$: Herhangi bir örneğin x olma olasılığını temsil eder (Küçük, 2019).

2.3.3. Lojistik Regresyon

Lojistik regresyon genellikle kategorik veri sınıflandırmasında kullanılır. Veri sınıflandırmasında bağımlı değişken ile bağımsız değişken arasındaki bağlantıyı açıklamak için kullanılır (Raj, 2020). Lineer regresyonun en küçük kareleri kullanmasının zıttı olarak maksimum olabilirlik prensibini kullanır. İsminde regresyon geçmesine rağmen aşağıda verilen sigmoid fonksiyonu yardımıyla sınıflandırma yapmak için kullanılır (Gök, 2017).

$$f(s) = \frac{1}{1+e^{-s}} \quad (4)$$

2.3.4. Doğrusal Diskriminant Analizi

Çok değişkenli istatistiksel yöntemlerden olan sınıflandırma da çok eskiden beri kullanılan bir yöntem olup güvenilirliğini hala korumaktadır. Diskriminant analizinin ilk amacı verileri her bir nitelik için ayrı ayrı işlem yaparak sınıflandırma ve ayırmadır.

Diskriminant analizi sınıflandırma yöntemleri arasında güçlü bir teorik yapıya sahip olduğu için çok kullanılan yöntemlerden biridir. Destek vektör makinesinden farkı doğrusal olmayan veriler için yüksek performanslı çözümler üretememesidir (Taya, 2022).

3. Destek Vektör Makinesi

Destek vektör makinesi; sınıflandırma ile regresyon analizlerinde sıklıkla kullanılan denetimli öğrenme yöntemlerinden biridir. Destek vektör makinesi diğer sınıflandırma metotlarına göre daha çok tercih edilir. Destek vektör makinesinin sınıflandırma problemlerinde diğer makine öğrenmesi yöntemlerine göre en önemli avantajı sınıflandırma problemini bir optimizasyon problemine dönüştürüp çözmesidir. Bir diğer avantajı düzensiz verilerle de iyi çalışmasıdır. Uygun bir çekirdek seçimi ile karmaşık verileri sınıflandırabilir, yüksek boyutlu verilerin nispeten iyi ölçeklenmesini sağlar.

Destek vektör makinesi, iki sınıfa ait verileri uygun olacak daha yüksek boyuta dönüştürmek için doğrusal olmayan bir haritalama kullanır. Haritalama ile verilerin dönüştürüldüğü yeni uzayda verileri en doğru şekilde ayırabilmek için hiper düzlem kullanılır. Başka bir ifadeyle veri kümelerini birbirinden ayırmak için karar sınırları belirlenir. Çizilen bu hiper düzlem ile destek vektör makineleri arasındaki marjin (uzaklık) maksimum olacak şekilde hiper düzlemi kullanarak sınıflandırma görevini gerçekleştirmeye çalışır. Diğer bir ifadeyle iki grubun üyelerine de sınır en uzak mesafede olmalıdır. Destek vektör makinesi, el yazısı tanıma, nesne tanıma ve konuşmacı kimliği saptama gibi birçok uygulama alanına sahiptir. Destek vektör makinesi bir sınıflandırıcı yöntem olarak verilerin doğrusal olarak ayrılabilmediği ve ayrılamadığı durumlarda etkili bir şekilde çalışır (Gunn, 1998; Dabanoğlu, 2018).

Verilerin doğrusal olarak ayrılamadığı durumlarda bir çekirdek fonksiyonu yardımıyla verilerin yüksek boyutlu bir uzaya dönüşümü gerçekleştirilerek yeni uzayda doğrusal olarak ayırmak amaçlanır. Burada çözüm için uygun çekirdek seçmek zordur ve her veriye göre en uygun çözümü veren bir en iyi çekirdekten bahsedilemez. Bununla birlikte büyük veri kümeleri kullanıldığında eğitim süresi uzun sürebilir.

Destek vektör makinesinin arkasındaki ana fikir, modelin genelleme yeteneğini en üst düzeye çıkaran bir yaklaşımla eğitim verilerini bir ayırma yüzeyi ile sınıflandırmaktır.

Burada, teorik tanımları basit tutmak için destek vektör makinesi ile ikili sınıflandırmaya odaklanacağız. Eğitim verilerimizin, N giriş vektörü, x_i ve ilişkili etiketler y_i olmak üzere, $i = 1 \cdots N$ olduğunu varsayalım.

$$x_i \in R^d, d \in \mathbb{Z}^+, y_i \in \{+1, -1\} \text{ şeklindedir.} \quad (5)$$

İlk adımda verilerimizin lineer olarak ayrılabilir olduğunu yani herhangi bir kesişim noktası olmaksızın lineer bir yüzeye sahip etiketlere göre ayrılabilceğini varsayıyoruz.

Başka bir varsayım, basit bir geometrik gösterim için $x_i \in R^2$ olacaktır.

Destek vektör makinesi, verileri maksimum marjla ayıran optimum ayırma hiper düzlemini bulmayı amaçlar. Ayırma hiper düzlemi H 'nin denklemi aşağıdaki gibi verilsin:

$$w^T x_i + b = 0. \quad (6)$$

H 'ye eşit uzaklıkta olan iki paralel hiper düzlem olacaktır. Bu düzlemlerin denklemleri aşağıda verilmiştir:

$$w \cdot x^+ + b = 1, w \cdot x^- + b = -1. \quad (7)$$

Burada x^+ ile y etiketi +1 olan x örnekleri, x^- ile de y etiketi -1 olan x örnekleri ifade edilmektedir. Ayrıca w , H 'yi belirleyen ağırlıkların vektörüdür, b sapmayı temsil eder. Denklem (7)'den aşağıdaki eşitsizliğe ulaşırız.

$$y_i (w \cdot x_i + b) \geq 1, \forall i. \quad (8)$$

Yani, aşağıdaki optimizasyon probleminin çözümü araştırılır:

$$\begin{aligned} \min_{w,b} & \|w\|^2 \\ \text{s. t. } & y_i (w \cdot x_i + b) \geq 1, \forall i. \end{aligned} \quad (9)$$

Bunun sonucunda, maksimum marjin $\gamma = \frac{1}{\|w\|}$ olarak bulunabilir. Sonuç olarak, geometrik marjinin maksimize edilmesi, w ağırlık vektörünün uzunluğunun minimize edilmesini gerektirir. Kenar boşluğu maksimize edildiğinde, bazı noktalar H_1 ve H_2 de olacaktır. Bu noktalar “destek vektörleri” olarak adlandırılır ve destek vektör makinesi modelinin kritik unsurlarıdır. Öte yandan, birçok gerçek hayat verisinin doğrusal olarak ayrılamaz olduğu bilinmektedir. Bu nedenle, verilerin bazı kesişme noktalarına sahip olmasına izin vermenin daha esnek bir yolunu bulmalıyız. Bu nedenle, tümü negatif olmayan gevşek değişkenlerden η_i , $i = 1 \dots N$ tanıtıyoruz ve destek vektör makinesi modelimiz için "yumuşak marjlar" elde ediyoruz. Böylece, optimizasyon problemi (9)'un durumu değiştirilecektir.

$$y_i(w \cdot x_i + b) \geq 1 - \eta_i, \quad i = 1 \dots N. \quad (10)$$

Ek olarak, gevşek değişkenleri kontrol etmek ve hatayı düzenlemek için bir C parametresi eklenir. C deneysel olarak kullanıcı tarafından seçilir. Böylece, aşağıdaki optimizasyon problemini elde ederiz:

$$\begin{aligned} \min_{w,b,\eta} \quad & \frac{1}{2} \|w\|^2 + C \sum_{i=1}^N \eta_i \\ \text{s.t.} \quad & y_i(w \cdot x_i + b) \geq 1 - \eta_i, \quad \eta_i \geq 0, \quad i = 1 \dots N. \end{aligned} \quad (11)$$

İşlemimizin ikinci adımı olarak lineer hiper düzlem ile ayrıştırılamayan ve lineer olmayan ve ayırmaya ihtiyaç duyulan bir veri setini ele alıyoruz. Bu nedenle, yeni özellik uzayında doğrusal diskriminant analizi yapabilmek için verilerin yeni bir yüksek boyutlu özellik uzayına dönüştürülmesini kullanıyoruz. Doğrusal olmayan dönüşüm şu şekilde temsil edilebilir:

$$\begin{aligned} \phi: R^d &\mapsto R^n \\ x \in R^d &\mapsto \phi(x) = [\phi_1(x), \dots, \phi_n(x)]^T \in R^n. \end{aligned} \quad (12)$$

Optimizasyon problemimiz şu şekilde değişir:

$$\begin{aligned} & \omega, b, \eta \frac{1}{2} \|w\|^2 + C \sum_{i=1}^N \eta_i \\ & s. t. y_i(w^T \psi(x_i) + b) \geq 1 - \eta_i, \eta_i \geq 0, i = 1 \dots N. \end{aligned} \quad (13)$$

Destek vektör makinesi uygulamasında Lagrange optimizasyon problemi, Lagrange çarpanlarının hesaplanmasının kolaylığı ve nokta çarpıma bağlı daha pratik bir denklem biçimi kullanma olasılığı nedeniyle, destek vektör makinesinin orijinal optimizasyon probleminin bir duali inşa edilir. Dolayısıyla, optimizasyon problemi (13)'ün dual formu olarak basit gösterim için yumuşak kenar boşlukları olmadan şu şekilde yazılabilir:

$$\begin{aligned} & \max_{\alpha} \frac{1}{2} \alpha^T Q \alpha - e^T \alpha \\ & s. t. y^T \alpha = 0 \\ & 0 \leq \alpha_i \leq C, i = 1, \dots, N. \end{aligned} \quad (14)$$

Burada e, birlerden oluşan bir vektördür. Q , $Q_{ij} = y_i y_j K(x_i, x_j)$, $i, j = 1 \dots N$ ile tanımlanan bir $N \times N$ pozitif yarı-belirli matristir. α_i Lagrange çarpanlarıdır ve sıfır olmayan α_i destek vektörlerine karşılık gelir. Burada $K(x_i, x_j)$ bu durum için $K(x_i, x_j) = \psi(x_i)^T \psi(x_j)$ olan çekirdek fonksiyonunu temsil eder. Çekirdek fonksiyonu, dönüştürme fonksiyonu değerleri $\psi(x)$ 'in hesaplanmasına gerek kalmadan verilerin yüksek boyutlu bir uzaya dönüştürülmesini ve doğrusal olarak ayrılmasını sağlar. Sınıflandırıcı fonksiyonu olarak aşağıdaki fonksiyon kullanılır:

$$f(x) = \text{sign}\left(\sum_{i=1}^N \alpha_i y_i K(x_i, x) + b\right). \quad (15)$$

Farklı $\psi(x)$ haritalamaları farklı çekirdek fonksiyonları ortaya çıkaracaktır. Kısaca, haritalama fonksiyonlarının değerlerini hesaplamamız gerekmediğinden, çekirdek

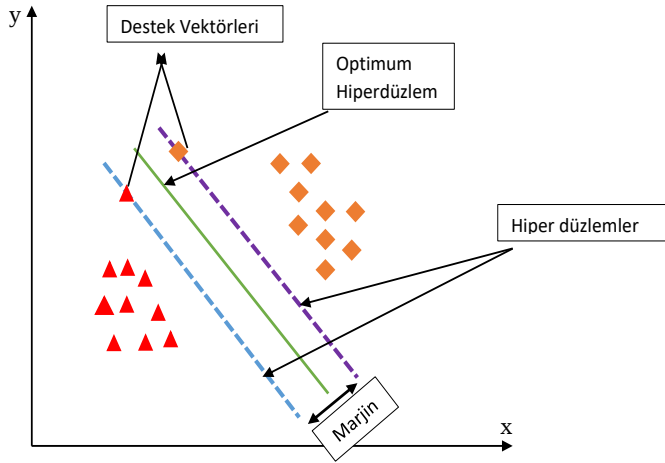
fonksiyonlarına ve özelliklerine odaklanmak daha faydalıdır. Bu durum, çekirdek fonksiyonu olmaya uygun farklı fonksiyonları burada aynı adımlarla kullanmamıza izin verecektir (Hiçdurmaz vd., 2023).

Destek vektör makinesi ve uygulaması hakkında kapsamlı bir çalışma için (Cervantes vd., 2020)'ye başvurulabilir. Çekirdek seçimi (Cervantes vd., 2020)'de önemli ve açık bir problem olarak belirtilmiştir.

Mevcut çalışmanın odak noktası, matematiksel analiz araçları kullanılarak yeni çekirdek fonksiyonlarının tasarımı ve analizidir. Bir sonraki bölüm, destek vektör makinesi için çekirdek seçimi ve çekirdek tasarımı hakkındadır.

Şekil 3

Destek Vektörleri



4. Çekirdek (Kernel) Fonksiyonu

Regresyon ve sınıflandırma problemlerinde çekirdek fonksiyonları ihtiyaca göre seçilmektedir. Hangi çekirdeğin kullanılacağı problemin eğitim verisinde farklı sınıflara ait verilerin birbirlerine göre geometrik durumu, karmaşıklığı, verinin büyüklüğü gibi farklı etkenlere dikkat edilerek belirlenmektedir. Örneğin doğrusal ayrılabilen bir problem için doğrusal çekirdek, hızlı ve etkin bir şekilde çalışırken, gerçek hayatta karşılaştığımız pek çok veri daha karmaşık bir yapıya sahiptir ve Gauss çekirdeği öncelikli olmak üzere farklı radyal tabanlı fonksiyonlar ile dönüşüm gerektirmektedir (Taya, 2022).

Çekirdek fonksiyonları destek vektör makinesi yöntemine esneklik kazandırır. Dolayısıyla doğru çekirdek seçimi ile performans açısından çok yüksek sonuçlar elde edilebilir (Güldoğan, 2017). Doğrusal olmayan iki sınıf arasındaki hiper düzlemi bulmak için kullanılan haritalama yönteminin fonksiyonunu bulmak genellikle gereksiz ve zordur. Bu fonksiyonların yerine haritalama matrisinin elemanlarını oluşturan çekirdek fonksiyonları kullanılır. Bu çalışmada radyal tabanlı çekirdek fonksiyonlarından bazıları tanıtılacak ve incelenecektir. Destek vektör makinesi, seçilen çekirdek fonksiyonunu kullanarak sınıflandırılmak istenen doğrusal olmayan verileri daha yüksek boyutlu nitelik uzayına taşır (Küçüksille ve Ateş, 2013; Ercan, 2020). Doğrusal çekirdek fonksiyonu, polinom çekirdek ve sigmoid çekirdek fonksiyonu klasik çekirdek fonksiyonlarından (Ercan, 2020). En sık kullanılan çekirdek fonksiyonu ise bu çalışma kapsamında genelleştirilmiş versiyonlarını araştırdığımız Gauss çekirdek fonksiyonudur.

Destek vektör makinesi için yeni bir çekirdek tasarımı, literatürde pek çok yazar tarafından çalışılmıştır (bkz. Ayat vd., 2001; Zhang vd., 2013). Yeni çekirdek tasarımı veya çekirdek karşılaştırması ile destek vektör makinesi üzerine yapılan çalışmaların kapsamlı bir incelemesi için bakınız (Cervantes vd., 2020).

Literatürdeki çekirdek fonksiyonları ya doğrusal, sigmoid veya polinom çekirdekler gibi nokta çarpım fonksiyonları $K(x, y) = K(x \cdot y)$ veya radyal tabanlı fonksiyonlar dediğimiz uzaklık tabanlı $K(x, y) = K(\|x - y\|)$ fonksiyonlarıdır. Burada radyal tabanlı fonksiyon, değeri yalnızca girdi ile bazı sabit noktalar arasındaki mesafeye bağlı olan gerçek değerli fonksiyonların genel adıdır. Bu tez kapsamında çekirdek olarak yaygın kullanılan ve son zamanlarda tanıtılan radyal tabanlı fonksiyonlara odaklanacağız.

Tablo 1

Geleneksel Çekirdekler Tablosu

İsim	Çekirdek Fonksiyonları
Lineer çekirdek	$K(P_i, P_j) = P_i \cdot P_j$
Polinom çekirdek	$K(P_i, P_j) = (P_i \cdot P_j + 1)^P, P \in \mathbb{Z}^+$
Gauss çekirdek (RBF)	$K(P_i, P_j) = \exp(-\gamma \ P_i - P_j\ ^2), (\gamma > 0)$
Sigmoid Çekirdek (Hiperbolik Teğet)	$K(P_i, P_j) = \tanh(\eta P_i P_j + v), \eta > 0, v < 0$

Veri setine bağlı olarak daha iyi sonuçlara ulaşılabilen bazı yeni tanıtılan çekirdek fonksiyonları da Tablo 2’de sunulmuştur. Tablo 1 ve Tablo 2’de çekirdek fonksiyonları normalleştirme parametresi gösterimi olmadan verilmiştir. Bu çekirdeklerin bazı lineer kombinasyonları da incelenmiş ve literatürde karşılaştırılmıştır (Shukla, 2010).

Tablo 2*Yeni Çekirdekler Tablosu*

İsim	Çekirdek Fonksiyonları
Üstel RBF (ERBF)	$K(P_i, P_j) = \exp(-\gamma \ P_i - P_j\) \quad (\gamma > 0)$
Orta derecede azalan çekirdek (KMOD)	$K(P_i, P_j) = \left[\exp\left(\frac{\gamma}{\ P_i - P_j\ ^2 + \sigma^2}\right) - 1 \right] \quad (\sigma \neq 0, \gamma > 0)$
Lorentz Çekirdeği (UKF)	$K(P_i, P_j) = \frac{1}{\ P_i - P_j\ ^2 + \sigma^2} \quad (a > 0, \sigma \neq 0)$
Genelleştirilmiş UKF	$K(P_i, P_j) = \frac{1}{(\ P_i - P_j\ ^2 + \sigma^2)^a} \quad (a > 0, \sigma \neq 0)$

Destek vektör makinesine herhangi bir fonksiyonunun çekirdek olarak seçilebilmesi için bir takım şartlar vardır. Orijinal uzayı artırılmış öznitelik uzayına haritalamada ortaya çıkan bir fonksiyonun varlığı, Mercer teoremi ile sağlanır.

Teorem 1

Bir $K(x, y)$ fonksiyonunun geçerli bir Mercer çekirdeği olması için aşağıdaki integral her bir sonlu $g(x)$ fonksiyonu için non negatif olmalıdır.

$$\iint K(x, y)g(x)g(y)dxdy \geq 0 \quad (16)$$

Mercer teoremi, çekirdeklerin simetrik, sürekli ve pozitif tanımlı olması gerektiğini gösterir. Ayrıca çekirdek matrisleri tümü negatif olmayan özdeğere sahiptir. Böyle bir

çekirdeğin kullanılması, optimizasyon probleminin konveks olmasını ve tek çözüme sahip olmasını sağlar.

Tanım 1

Tamamen Monotonluk

f fonksiyonunun $[0, \infty)$ aralığında sürekli, $(0, \infty)$ aralığında sonsuz kere türevlenebilir olması aşağıdaki koşulla birlikte tamamen monoton olması için gerek ve yeter şarttır.

$$(-1)^k f^{(k)} \geq 0, k=0,1,\dots,N. \quad (17)$$

Teorem 2

$g(t) \neq C$, $g(t)$ tamamen monoton ve $[0, \infty)$ aralığında tanımlı olsun. $Q(r) = g(r^2)$ ise $Q(r)$ pozitif tanımlı olur.

$K(x, y)$ fonksiyonu $r = \|x - y\|$ değişken dönüşümü ile $f(r) = K(x, y)$ olarak yazılabilen radyal tabanlı bir fonksiyon olsun. Eğer $f(r) = g(r^2)$ ise $f(r)$ 'nin sürekli ve pozitif tanımlı olduğunu göstermemiz yeterlidir. Bunun içinde Teorem 2'den $g(r)$ fonksiyonun tamamen monoton olmasının gerekli ve yeterli olduğunu görüyoruz. Şimdiye kadar ki çalışmalarda yeni çekirdek seçimi ve dizaynı ile ilgili tamamen monoton fonksiyonlara ve Bernstein fonksiyonlarına yer verilmediğini gözlemledik.

Bu noktada bahsedilmesi gereken bir diğer konu da çekirdek fonksiyonu olması için gerekli olan yumuşak Mercer koşuludur. Bu koşula göre bir radyal tabanlı fonksiyonun Mercer çekirdek fonksiyonu olması için sıfıra yakın bölgede hızla azalması ve sonsuza giderken tamamen sıfır olmayıp pozitif kalmayı uzun süre başarmasıdır. Sonsuza giderken sıfırlanmaması fonksiyonun veriyi yeni uzaya dönüştürdüğü zaman farklı sınıflar arasındaki uzaklığı koruması anlamına gelmektedir.

Bir sonraki bölümde tamamen monotonluk tanımından faydalanarak genelleştirilmiş yeni Gauss çekirdekleri oluşturulacak ve hem teorik Mercer koşulu hem de yumuşak Mercer koşulu açısından incelenecektir.

5.Genişletilmiş Gauss Çekirdeği Fonksiyonları

5.1 Gauss Çekirdeği

Gauss çekirdeği, Alman matematikçi Carl Friedrich Gauss'un (1777-1855) adını almıştır (Gaussiankernel.Nb, t.y.). Doğrusal olmayan veri setinde kullanılan Gauss çekirdeği en çok tercih edilen yöntemlerden biridir (Taya, 2022).

$$K(x, y) = e^{-\gamma \|x-y\|^2}, \quad 0 < \gamma. \quad (18)$$

Bu tezde klasik Gauss fonksiyonunun kullanım alanını genişletmek ve çalışma performansını artırmak, daha net ve daha hızlı sonuçlar elde etmek için Gauss çekirdeğinin genişletilmiş versiyonlarını araştırıyoruz.

5.2. Bernstein Fonksiyonlarından Türetilen Çekirdek Fonksiyonları

Tanım 2

Bernstein fonksiyonları

Bernstein fonksiyonları, türevi tamamen monoton olan fonksiyonlardır. Öyle ki aşağıdaki eşitsizlik sağlanır:

$$(-1)^{(n-1)} b^{(n)}(x) \geq 0, \quad n = 1, 2, \dots \quad (19)$$

Bernstein fonksiyonunun özellikleri:

- $b_1(x), b_2(x)$ bir Bernstein fonksiyonu ise $a_1 b_1(x) + a_2 b_2(x), (a_1, a_2 \geq 0)$ şeklindeki lineer kombinasyon da bir Bernstein fonksiyonudur.
- Bernstein fonksiyonlarının bileşkesi de $(b_1(b_2(s)))$ bir Bernstein fonksiyonudur.
- Yakınsak bir Bernstein fonksiyonları serisinin noktasal limiti, bir Bernstein fonksiyonudur.
- Tamamen monoton fonksiyon $m(x)$ ve Bernstein fonksiyonu $b(x)$ 'in bir bileşkesi $m(b(x))$ fonksiyonu da tamamen monoton bir fonksiyondur.

- Eğer $b(x)$ bir Bernstein fonksiyonu ise $m(x) = \frac{b(x)}{x}$ fonksiyonu da tamamen monotondur.
- $b_1(x), b_2(x)$ Bernstein fonksiyonları ve $a_1 a_2 \in (0,1)$, $a_1 + a_2 \leq 1$ olmak üzere $b_1(x^{a_1})b_2(x^{a_2})$ fonksiyonu da bir Bernstein fonksiyonudur (Bernstein Functions, t.y.).

Bu kısımda aşağıdaki teoremi kullanarak genişletilmiş Gauss çekirdeklerini araştıracağız.

Teorem 3

$g(r)$ bir Bernstein fonksiyonu olmak üzere $r = \|x - y\|$ alındığında radyal tabanlı simetrik ve sürekli $K(x, y)$ fonksiyonları $e^{-g(r^2)}$ şeklinde ifade edilebiliyorsa bir Mercer çekirdek fonksiyonu olur.

İspat

$f(x) = e^{-x}$ fonksiyonu tamamen monoton bir fonksiyon olduğu aşağıda gösterilmiştir.

$$f(x) = e^{-x}$$

$$(-1)^k f^{(k)} = (-1)^k (-1)^k e^{-x} \geq 0, k=0,1,\dots,N. \quad (20)$$

Bu nedenle, $g(x)$ bir Bernstein fonksiyonu olmak üzere Bernstein fonksiyonlarının özelliklerinden $f(g(x)) = e^{-g(x)}$ şeklinde elde edebildiğimiz fonksiyonlar da tamamen monoton fonksiyonlar olacaktır. Bu durumda Teorem 2'den $f(g(r^2)) = e^{-g(r^2)}$ bir pozitif tanımlı fonksiyon olur. Teorem 1'den $r = \|x - y\|$ değişken dönüşümü ile elde edilebilen $K(x, y)$ fonksiyonları Mercer çekirdeği olur.

Bu tip çekirdeklere odaklanmamızın bir sebebi de $0 < e^{-g(x)} \leq 1$ olmasıdır. Yani normalizasyon gerektirmemesidir. Bir sonraki bölümde teorik olarak oluşturduğumuz genişletilmiş Gauss fonksiyonunun üç farklı örneği gösterilecektir. Bu örnekler elde edilirken literatürdeki mevcut pek çok Bernstein fonksiyonu genişletilmiş Gauss çekirdeği haline getirilmiş, sayısal olarak farklı veri setlerinde denenmiştir. Bu fonksiyonlar içerisinden klasik Gauss fonksiyonuna göre öne çıkan en iyi örnekler seçilmiştir. Bu çalışmanın asıl katkısı Bernstein fonksiyonlarından makine öğrenmesinde faydalı olabilecek başka çekirdekler elde edilebileceğini göstermesidir. Teorik olarak Mercer koşulunu sağlayan her çekirdek fonksiyonu yüksek performansla çalışmayacaktır. Çekirdek seçiminde sayısal deneylerle ele alınan çekirdek fonksiyonunun performansının klasik çekirdeklerle yarışabilir olmasına dikkat edilmelidir.

5.3. Uygulamalar

Teorem 4

$K(x, y) = e^{-\|x-y\|^\beta}$, $0 < \beta < 2$ Mercer koşulunu sağlayan bir çekirdek fonksiyondur.

İspat

$g(x) = x^{\frac{\beta}{2}}$ bir Bernstein fonksiyonu, $f(x) = e^{-x}$ tamamen monoton olmak üzere $f(g(r^2)) = e^{-r^\beta}$ (Fonk. I) pozitif tanımlı bir fonksiyondur. $r = \|x - y\|$ değişken dönüşümü yapılarak elde edilen $K(x, y)$ 'nin sürekli ve simetrik olduğu dikkate alınarak ispat tamamlanır.

g fonksiyonunun Bernstein fonksiyonu olduğu açıkça hesaplanabilir. Bunun için g' fonksiyonunun tamamen monoton olduğu gösterilmesi yeterli olacaktır:

$$g = x^{\frac{\beta}{2}}, \quad 0 < \beta < 2,$$

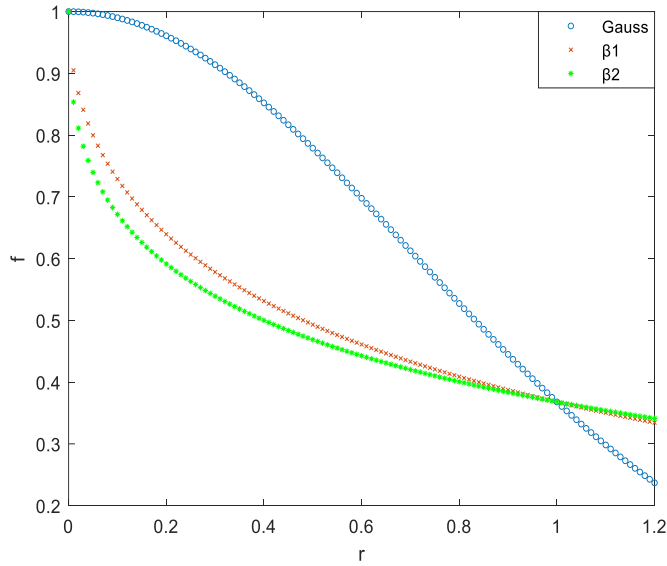
$$g' = \frac{\beta}{2} x^{\frac{\beta}{2}-1} \geq 0, \quad x > 0,$$

$$g'' = \frac{\beta}{2} \left(\frac{\beta}{2} - 1 \right) x^{\frac{\beta}{2}-2}, \quad g''(-1) \geq 0. \quad (21)$$

Burada $0 < \frac{\beta}{2} < 1$ ve $-1 < \frac{\beta}{2} < 0$ olduğu için $(-1)^k g^{(k+1)} \geq 0$ sağlanır.

Şekil 4

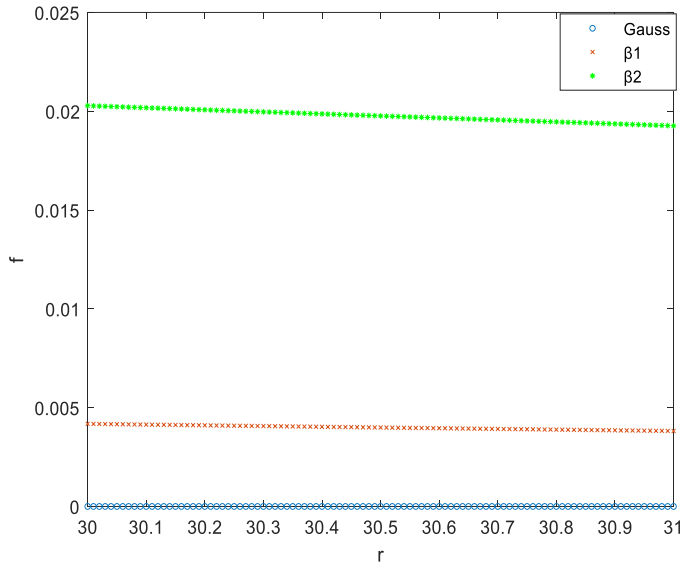
Fonk.I, Sıfır Civarında İncelenmesi



Şekil 4’de görüldüğü gibi $f(g(r^2)) = e^{-r^\beta}$ fonksiyonu, Mercer koşulu sağlaması için gerekli olduğu şekilde sıfıra yakın bölgede hızlı bir şekilde azalmıştır. Burada $\beta_1 = 0.5$ ve $\beta_2 = 0.4$ alınmıştır. Ayrıca Gauss fonksiyonu ve Fonk.I için $\gamma=1$ alınmıştır.

Şekil 5

Fonk.I, Sonsuz Civarında İncelenmesi



Şekil 5'te görüldüğü gibi $f(g(r^2)) = e^{-r^\beta}$ fonksiyonu yumuşak Mercer koşulunu sağlamış, sonsuza giderken Gauss fonksiyonunun üstünde kalıp hemen sıfırlanmamıştır. Burada $\beta_1 = 0.5$ ve $\beta_2 = 0.4$ alınmıştır. Ayrıca Gauss fonksiyonu ve Fonk.I için $\gamma=1$ alınmıştır.

Teorem 5

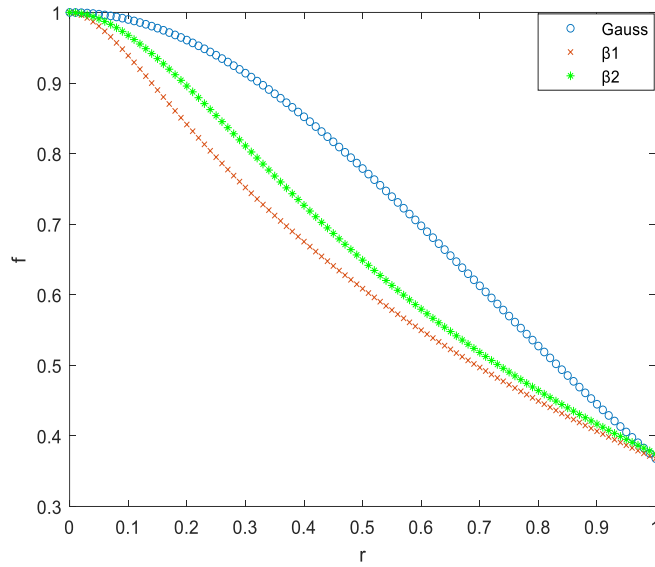
$K(x, y) = e^{-\|x-y\|(1-e^{-2\beta\|x-y\|})}$, $0 < \beta$ Mercer koşulunu sağlayan bir çekirdek fonksiyondur.

İspat

$g(x) = x^{1/2} (1 - e^{-2\beta x^{1/2}})$, $\beta > 0$ bir Bernstein fonksiyonudur. (Bkz. (Bernstein Functions.(t.y.) s. 228) $f(x) = e^{-x}$ tamamen monoton olmak üzere $f(g(r^2)) = e^{-r(1-e^{-2\beta r})}$ (Fonk.II) pozitif tanımlı bir fonksiyondur. $r = \|x - y\|$ değişken dönüşümü yapılarak elde edilen $K(x, y)$ 'nin sürekli ve simetrik olduğu dikkate alınarak ispat tamamlanır.

Şekil 6

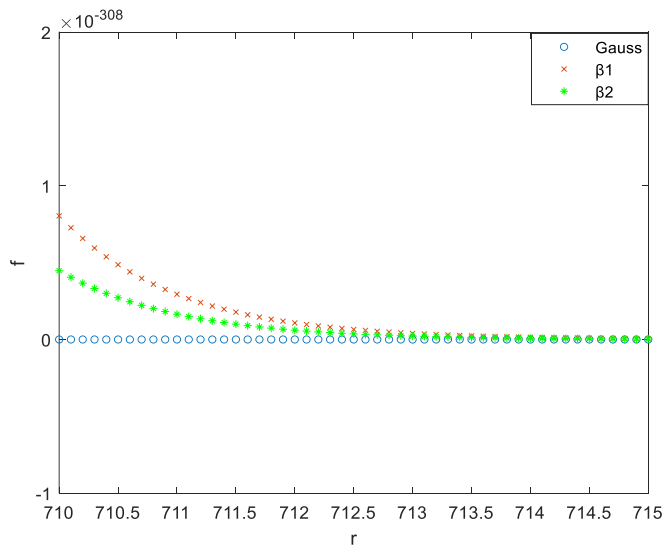
Fonk.II, Sıfır Civarında İncelenmesi



Şekil 6'da görüldüğü gibi $f(g(r^2)) = e^{-r(1-e^{-2\beta r})}$ fonksiyonu yumuşak Mercer koşulu sağlayacak şekilde sıfıra yakın bölgede hızlı bir şekilde azalmıştır. Burada $\beta_1 = 5$ ve $\beta_2 = 2$ alınmıştır. Ayrıca Gauss fonksiyonu için $\gamma=1$ alınmıştır.

Şekil 7

Fonk.II, Sonsuz Civarında İncelenmesi



Şekil 7’de görüldüğü gibi $f(g(r^2)) = e^{-r(1-e^{-2\beta r})}$ fonksiyonu yumuşak Mercer koşulu sağlayacak şekilde sonsuza giderken Gauss fonksiyonunun üstünde kalıp hemen sıfırlanmamıştır. Burada $\beta 1 = 0.005$ ve $\beta 2 = 1000$ alınmıştır. Ayrıca Gauss fonksiyonu için $\gamma=1$ alınmıştır.

Teorem 6

$K(x, y) = e^{-\|x-y\|(1+e^{-2\beta\|x-y\|})}$, $0 < \beta$, Mercer koşulunu sağlayan bir çekirdek fonksiyondur.

İspat

$g(x) = x^{1/2} (1 + e^{-2\beta x^{1/2}})$, $\beta > 0$ bir Bernstein fonksiyonudur (Bkz. *Bernstein*

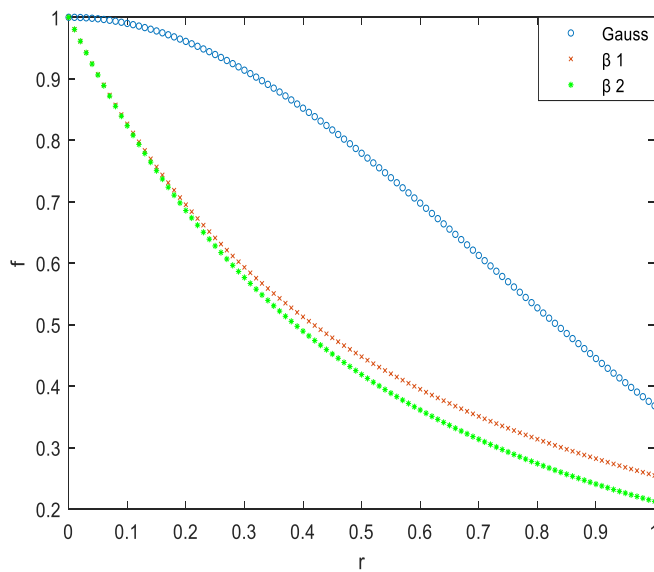
Functions, (t.y.) s. 228). $f(x) = e^{-x}$ tamamen monoton olmak üzere $f(g(r^2)) = e^{-r(1+e^{-2\beta r})}$

(*Fonk.III*) pozitif tanımlı bir fonksiyondur. $r = \|x - y\|$ değişken dönüşümü yapılarak elde edilen

$K(x, y)$ ’nin sürekli ve simetrik olduğu dikkate alınarak ispat tamamlanır.

Şekil 8

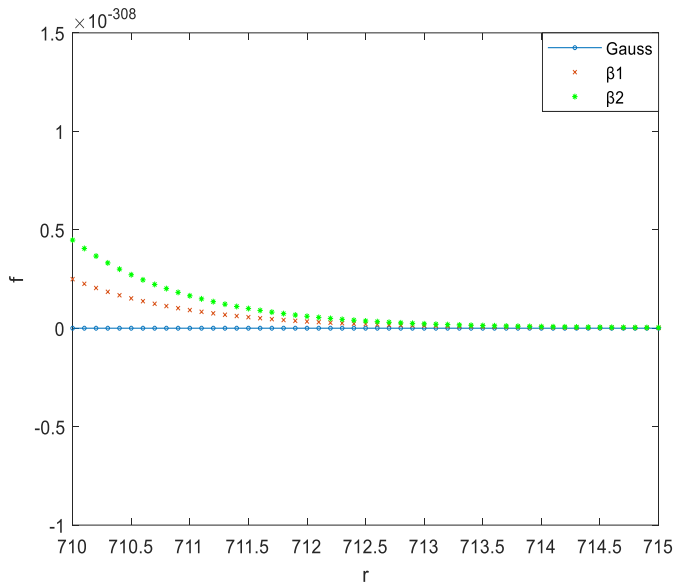
Fonk.III, Sıfır Civarında İncelenmesi



Şekil 8’de görüldüğü gibi $f(g(r^2)) = e^{-r(1+e^{-2\beta r})}$ fonksiyonu yumuşak Mercer koşulu sağlayacak şekilde sifıra yakın bölgede hızlı bir şekilde azalmıştır. Burada $\beta_1 = 0.5$ ve $\beta_2 = 0.3$ alınmıştır. Ayrıca Gauss fonksiyonu için $\gamma=1$ alınmıştır.

Şekil 9

Fonk.III, Sonsuz Civarında İncelenmesi



Şekil 9’da görüldüğü gibi $f(g(r^2)) = e^{-r(1+e^{-2\beta r})}$ fonksiyonu Mercer koşulu sağlayacak şekilde sonsuza giderken Gauss fonksiyonunun üstünde kalıp hızlı bir şekilde sıfırlanmamıştır. Burada $\beta_1 = 0.005$ ve $\beta_2 = 1000$ alınmıştır. Ayrıca Gauss fonksiyonu ve için $\gamma=1$ alınmıştır.

Burada bahsedilen uygulamalara ek olarak bir yöntem daha verilebilir. Mevcut bilinen çekirdek fonksiyonu ve bu tez kapsamında tanıtılan genişletilmiş gauss çekirdekleri ve aşağıdaki özellikler kullanılarak yeni çekirdekler elde edilebilir. Böylece veri setine daha uygun bir çekirdek fonksiyonu bulma fırsatı oluşur.

$k_1(x, x')$ ve $k_2(x, x')$ şeklinde geçerli Mercer çekirdekleri verildiğinde bunlardan elde edilen aşağıdaki çekirdekler de geçerli Mercer çekirdeği olacaktır.

- $k(x, x') = ck_1(x, x')$
- $k(x, x') = f(x)k(x, x')f(x')$
- $k(x, x') = q(k_1(x, x'))$
- $k(x, x') = \exp(k(x, x'))$
- $k(x, x') = k_1(x, x') + k_2(x, x')$
- $k(x, x') = k_1(x, x') \cdot k_2(x, x')$
- $k(x, x') = k_3(\phi(x), \phi(x'))$
- $k(x, x') = x^T A x'$
- $k(x, x') = k_a(x_a, x'_a) + k_b(x_b, x'_b)$
- $k(x, x') = k_a(x_a, x'_a) \cdot k_b(x_b, x'_b)$

Burada;

- $c > 0$ bir sabittir.
- $f(\cdot)$ herhangi bir fonksiyondur, $q(\cdot)$ negatif olmayan katsayılı bir polinomdur.
- $\phi(x)$, x 'den R^M ye bir haritalama fonksiyonudur.
- $k_3(\cdot, \cdot)$, R^M de geçerli bir çekirdektir.
- A simetrik pozitif yarı-belirli bir matristir.
- k_a ve k_b , tanımlı olduğu aralıklarda geçerli çekirdek fonksiyonudur.

6. Python İle Sınıflandırma Uygulamaları

Önerilen yöntemin sınıflandırma başarısını ölçmek için ACC (Doğruluk /Accuracy) ve F1 performans göstergeleri kullanılmıştır.

ACC, bir modelin sınıflandırma başarısını ölçmek için yaygın olarak kullanılan bir metriktir. ACC değeri, bir sınıflandırıcının doğru tahminlerinin bir veri kümesindeki toplam örnek sayısına oranına eşittir. Ancak dengesiz veri setlerinde sadece model doğruluğu ACC değerinin yüksek olması için yeterli değildir. ACC'nin hesaplama formülü aşağıdaki gibidir:

$$ACC = \frac{TP + TN}{TP + TN + FP + FN}.$$

F1 puanı, kesinlik ve geri çağırma değerlerinin harmonik ortalaması hesaplanarak elde edilir. Dengesiz veri kümelerinde yanlış sınıflandırıcı modeli seçmemek için ACC yerine F1 göstergesinin kullanılması daha uygundur (Taya, 2022).

$$F1 = \frac{2}{TPR^{-1} + PPV^{-1}}.$$

Bu tez kapsamında yeni çekirdek fonksiyonlarının modelin çalışma performansında etkisi araştırıldığı için ACC ve F1 göstergelerinin ikisi de kullanılmıştır (Elen vd., 2022). ACC ve F1 değerleri 1 değerine ne kadar yakınsa model veri setinde o kadar iyi çalışmaktadır. Burada kullanılan wine ve digits veri setleri Python açık yazılımından scikit learn kütüphanesi kullanılarak indirilmiştir.

Wine veri seti Portekiz'de üretilen “Vinho Verde” şarabının beyaz ve kırmızı renkleriyle ilgili olan ve veri setinin 13 farklı değerine bakılarak şarapların kalitelerine göre sınıflandıran bir veri setidir. Bu veri setinde üzüm türleri, şarap markası, şarap satış fiyatı gibi değişkenler gizlilik ve lojistik konularından dolayı yoktur. Bunların yerine sadece fizikokimyasal (girişler) ve duyuşal (çıkı) değişkenler mevcuttur. Giriş değişkenleri

fizikokimyasal testlere dayanarak oluşturulmuştur. Bunlar sabit asit, uçucu asitliği, sitrik asit, klorürler, artık şeker, serbest sülfür dioksit, toplam sülfür dioksit, yoğunluk, pH, sülfatlar, alkol olarak sıralanabilir. Çıkış değişkenleri ise duyuşal verilere göre oluşturulmuş kalitesi 0 ile 10 arasında olacak şekilde var olan kalite puanları için beş ayrı seviye belirlenmiştir. Bu çalışmada kullandığımız versiyonu scikit-learn kütüphanesinden indirilen temizlenmiş wine veri setidir. Bu veri setinde şarabın kalitesini ifade etmek için kullanılan etiketler 0, 1 ve 2'den oluşmaktadır. Wine veri bu çalışmada çoklu sınıflandırma görevi için kullanılmıştır (*UCI Makine Öğrenimi Deposu*, t.y.).

Digits veri seti el yazısı rakamlardan oluşur. Her sınıf bir rakamı ifade ettiği 10 sınıf vardır. Digits MNIST veri tabanı uluslararası alanda yaygınca kullanılmaktadır. Bu veri tabanında 60.000 eğitim verisi vardır. Bunun yanında 10.000 test verisi içerir. Rakamların her biri 28x28 piksel ölçüsünde bulunan çerçevede 20x20 piksel ölçüsündedir. Tahmin algoritmalarında ilk adım olan bir veri setidir. Bu çalışmada kullandığımız versiyonu scikit-learn kütüphanesinden indirilen temizlenmiş veri setidir. El yazısı fotoğrafları 8x8 piksel ölçüsündedir. Mevcut çalışmada fotoğraflardan 0-9 arasındaki rakamları tahmin etme üzerine yapılan çoklu sınıflandırma görevi için kullanılmıştır (Karakaya, 2020).

Tablo 3*Fonk.I, Wine Veri Seti için ACC Tablosu*

$\beta \backslash \gamma$	.08	.1	.25	.5	1	1.1	1.15	1.5	1.9	2	Ortalama
.08	.66666	.70370	.74074	.72222	.68518	.68518	.68518	.68518	.74074	.74074	.70555
.09	.70370	.74074	.74074	.68518	.72222	.68518	.68518	.68518	.74074	.74074	.71296
.1	.74074	.74074	.74074	.70370	.72222	.70370	.70370	.68518	.70370	.74074	.71852
.25	.74074	.77777	.77777	.77777	.70370	.70370	.68518	.68518	.61111	.57407	.70367
.5	.75925	.75925	.77777	.68518	.44444	.42592	.40740	.38888	.38888	.38888	.54259
.9	.74074	.74074	.62962	.40740	.38888	.38888	.38888	.38888	.38888	.38888	.48518
1	.74074	.68518	.51851	.40740	.38888	.38888	.38888	.38888	.38888	.38888	.46851
1.1	.66666	.64814	.46296	.38888	.38888	.38888	.38888	.38888	.38888	.38888	.44999
1.15	.64814	.62962	.44444	.38888	.38888	.38888	.38888	.38888	.38888	.38888	.44443
1.5	.50000	.48148	.40740	.38888	.38888	.38888	.38888	.38888	.38888	.38888	.41110
2	.40740	.40740	.38888	.38888	.38888	.38888	.38888	.38888	.38888	.38888	.39259

Tablo 3’de farklı γ , β değerleri için Fonk.I çekirdeği ile destek vektör makinesi modeli kullanılarak wine veri seti üzerinde elde edilen ACC değerleri sunulmuştur. Burada $\beta = 2$ durumu klasik Gauss fonksiyonuna ve $\beta = 1$ durumu Laplace çekirdek fonksiyonuna karşılık gelmektedir. Dolayısıyla $\beta = 2$ satırı, $\beta = 1$ satırı ve diğer β değerlerine karşılık gelen ACC değerleri karşılaştırılmalıdır. Bu karşılaştırma için her bir satırın sonunda her bir β değeri için farklı γ değerleri ile elde edilen ACC değerlerinin ortalaması da verilmiştir. γ

farklılaştırılmasına rağmen ortalama ACC değerinin en yüksek olduğu β değeri $\beta = 0.1$ olarak elde edilmiştir. Tablo 3’de görüldüğü üzere Gauss ve Laplace çekirdek fonksiyonlarının wine veri seti üzerindeki performansı Fonk.I’ın performansına ulaşamamaktadır.

Tablo 4*Fonk.I, Wine Veri Seti için F1 Tablosu*

$\beta \backslash \gamma$	.08	.1	.25	.5	1	1.1	1.15	1.5	1.9	2	Ortalama
.08	.52745	.55619	.57592	.55799	.62584	.62584	.62584	.60256	.57592	.57592	.58495
.09	.55619	.57933	.57592	.62207	.68539	.62584	.62584	.62207	.57592	.57592	.60445
.1	.58333	.57592	.57592	.65811	.68255	.65505	.65505	.62584	.55253	.57592	.61402
.25	.71119	.76410	.76228	.76228	.65505	.65505	.62584	.57189	.48028	.44526	.64332
.5	.74444	.74444	.76880	.63838	.28535	.25527	.22252	.18666	.18666	.18666	.42192
.9	.72605	.72605	.59745	.22252	.18666	.18666	.18666	.18666	.18666	.18666	.33920
1	.71862	.66182	.44791	.22252	.18666	.18666	.18666	.18666	.18666	.18666	.31708
1.1	.63871	.61267	.36170	.18666	.18666	.18666	.18666	.18666	.18666	.18666	.29197
1.15	.62337	.59651	.32187	.18666	.18666	.18666	.18666	.18666	.18666	.18666	.28484
1.5	.42635	.40345	.22252	.18666	.18666	.18666	.18666	.18666	.18666	.18666	.23589
2	.22252	.22252	.18666	.18666	.18666	.18666	.18666	.18666	.18666	.18666	.19383

Tablo 4’de farklı γ, β değerleri için Fonk.I çekirdeği ile destek vektör makinesi modeli kullanılarak wine veri seti üzerinde elde edilen F1 değerleri sunulmuştur. ortalama F1 değerinin en yüksek olduğu β değeri $\beta = 0.25$ olarak elde edilmiştir.

Bu tablolarda elde edilen performans değerlerinden faydalanarak $\beta = 0.25, \gamma = 0.1$ değerleri ile elde edilen Fonk.I kullanan destek vektör makinesi modeli test verisinden rastgele bir seçim üzerinde aşağıdaki ekran görüntüsünde görüldüğü gibi uygulanmıştır.

Şekil 10

Wine Veri Seti için Modelin Uygulanmasının Ekran Görüntüsü

```

Test verisinin 3. satırı:
[1.373e+01 1.500e+00 2.700e+00 2.250e+01 1.010e+02 3.000e+00 3.250e+00
2.900e-01 2.380e+00 5.700e+00 1.190e+00 2.710e+00 1.285e+03]

Test verisinin 3.satırına karşılık gelen gerçek sınıf etiketi:
0

Test verisinin 3.satırına karşılık gelen klasik Gauss fonk SVM ile tahmin edilen sınıf etiketi:
[1]

Test verisinin 3.satırına karşılık gelen 1. Fonk SVM ile tahmin edilen sınıf etiketi:
[0]

```

Şekil 10'dan da görüldüğü gibi gerçek sınıf etiketi ile Fonk.I'in aynı çalıştığı görülmüştür. Bu ise Fonk.I'in destek vektör makinesi ile wine veri seti üzerinde klasik Gauss'tan daha iyi çalışabildiğinin somut bir uygulama ile gösterimidir.

Tablo 5*Fonk.I, Digits Veri Seti için ACC Tablosu*

$\beta \backslash \gamma$	.01	.05	.08	.2	.4	.5	.6	.7	.8	.9	Ortalama
.01	.09259	.09259	.09259	.09259	.09259	.09259	.09259	.09629	.13148	.15925	.10352
.05	.09259	.09259	.09259	.33703	.87592	.91111	.92407	.92962	.92962	.93333	.61185
.08	.09259	.09259	.17037	.84814	.93518	.94074	.94444	.94629	.94629	.94814	.68648
.1	.09259	.09259	.32592	.92037	.94444	.95000	.95185	.95555	.95555	.95740	.71462
.3	.20370	.93888	.95925	.97222	.97592	.97592	.97592	.97407	.97222	.96851	.89166
.5	.92037	.97407	.97962	.98333	.97777	.97777	.97222	.96296	.95370	.95000	.96518
.6	.94629	.98148	.98333	.98518	.97777	.96666	.95744	.94259	.90000	.77777	.94185
.7	.96296	.98518	.99074	.98518	.96296	.94444	.87777	.64259	.43888	.32962	.81203
.9	.97777	.99074	.98703	.97407	.48518	.26851	.17777	.09259	.09259	.09259	.51388
1	.98518	.98888	.98703	.89074	.21111	.09259	.09259	.09259	.09259	.09259	.45259
1.8	.95555	.13518	.09259	.09259	.09259	.09259	.09259	.09259	.09259	.09259	.18315
2	.09259	.09259	.09259	.09259	.09259	.09259	.09259	.09259	.09259	.09259	.09259

Tablo 5’de farklı γ , β değerleri için Fonk.I çekirdeği ile destek vektör makinesi modeli kullanılarak digits veri seti üzerinde elde edilen ACC değerleri sunulmuştur. Ortalama ACC değerinin en yüksek olduğu β değeri $\beta = 0.6$ olarak belirlenmiştir. Yine $\beta = 1$ (Laplace çekirdeği) ve $\beta = 2$ (Klasik Gauss çekirdeği) değerlerinin Fonk.I ile karşılaştırıldığında iyi bir performans göstermediği görülmektedir.

Tablo 6*Fonk.I, Digits Veri Seti için F1 Tablosu*

$\beta \backslash y$	.01	.05	.08	.2	.4	.5	.6	.7	.8	.9	Ortalama
.01	.01694	.01694	.01694	.01694	.01694	.01694	.01694	.02427	.07433	.10129	.03184
.05	.01694	.01694	.01694	.31379	.87544	.91168	.92411	.92968	.92968	.93303	.58682
.08	.01694	.01694	.11154	.84674	.93470	.94021	.94371	.94558	.94558	.94765	.66496
.1	.01694	.01694	.29927	.92052	.94373	.94918	.95118	.95544	.95544	.95749	.69661
.3	.15482	.93832	.95841	.97295	.97679	.97679	.97679	.97475	.97268	.96928	.88715
.5	.91966	.97476	.98034	.98420	.97869	.97869	.97222	.96330	.95394	.95012	.96559
.6	.94512	.98212	.98420	.98598	.97869	.96722	.95740	.94307	.90767	.80699	.94584
.7	.96246	.98580	.99098	.98576	.96368	.94520	.89081	.68640	.47334	.35321	.82376
.9	.97827	.99098	.98769	.97438	.54117	.29906	.13357	.01694	.01694	.01694	.49559
1	.98518	.98888	.98148	.44250	.21111	.09259	.09259	.09259	.09259	.09259	.61381
1.8	.95561	.11793	.01694	.01694	.01694	.01694	.01694	.01694	.01694	.01694	.12090
2	.01694	.01694	.01694	.01694	.01694	.01694	.01699	.01694	.01694	.01694	.01694

Tablo 6’da en yüksek ortalama F1 değerlerine $\beta = 0.7$ civarında ulaşıldığı görülmektedir. $\beta = 1$ (Laplace çekirdeği) ve $\beta = 2$ (Klasik Gauss çekirdeği) için değerler oldukça düşük gözlemlenmektedir.

Tablo 7*Fonk.II, Wine Veri Seti için ACC Tablosu*

γ β	.005	.01	.07	.1	Ortalama
.001	.77777	.77777	.75925	.75925	.76851
.005	.77777	.75925	.72222	.74074	.74999
.06	.75925	.75925	.72222	.68518	.73147
.1	.75925	.75925	.72222	.68518	.73147
1.5	.75925	.75925	.72222	.68518	.72684
2	.75925	.75925	.72222	.68518	.72684
50	.75925	.75925	.72222	.68518	.72684
600	.75925	.75925	.72222	.68518	.72684
1000	.75925	.75925	.72222	.68518	.72684
RBF	.66666	.64814	.40740	.40740	.53240

Tablo 7’de verilen değerler wine veri seti için Fonk.II ile elde edilen ACC değerleridir. Kısaltma amaçlı kullanılan RBF (Radial Basis Function) klasik Gauss fonksiyonu için literatürde kullanılan bir isimdir. Tablo 7’de de son RBF satırı klasik Gauss fonksiyonu için elde edilen ACC değerlerinden oluşmaktadır. Bundan sonraki bütün sayısal deneylerde RBF için , $\gamma = 1$ alınmıştır. Tablo 7 ve 8’e göre wine veri seti üzerinde Fonk. II ile oluşturulan destek vektör makinesi algoritması klasik Gauss fonksiyonu ile oluşturulandan çok daha iyi çalışmaktadır.

Tablo 8*Fonk.II, Wine Veri Seti için F1 Tablosu*

$\beta \backslash \gamma$	.005	.01	.07	.1	Ortalama
.001	.76410	.76410	.74317	.74317	.56784
.005	.76410	.74317	.70935	.72605	.73566
.06	.74444	.74317	.70300	.66370	.71357
.1	.74444	.74317	.70300	.66370	.71357
1.5	.74444	.74317	.70300	.66182	.71310
2	.74444	.74317	.70300	.66182	.71310
50	.74444	.75925	.70300	.66182	.71712
600	.74444	.75925	.70300	.66182	.71712
1000	.74444	.75925	.70300	.66182	.71712
RBF	.63806	.63601	.26137	.22252	.43949

Tablo 9*Fonk.II, Digits Veri Seti için ACC Tablosu*

$\beta \backslash \gamma$	.005	.01	.07	.1	Ortalama
.001	.89074	.92407	.97592	.97962	.94258
.005	.95185	.97962	.99074	.99074	.97823
.06	.97777	.98518	.98703	.98518	.98379
.1	.97592	.98518	.98703	.98148	.98240
1.5	.97592	.98518	.98703	.98148	.98240
2	.97592	.98518	.98703	.98148	.98240
50	.97592	.98518	.98703	.98148	.98240
600	.97592	.98518	.98703	.98148	.98240
1000	.97592	.98518	.98703	.98148	.98240
RBF	.95555	.70555	.09259	.09259	.46157

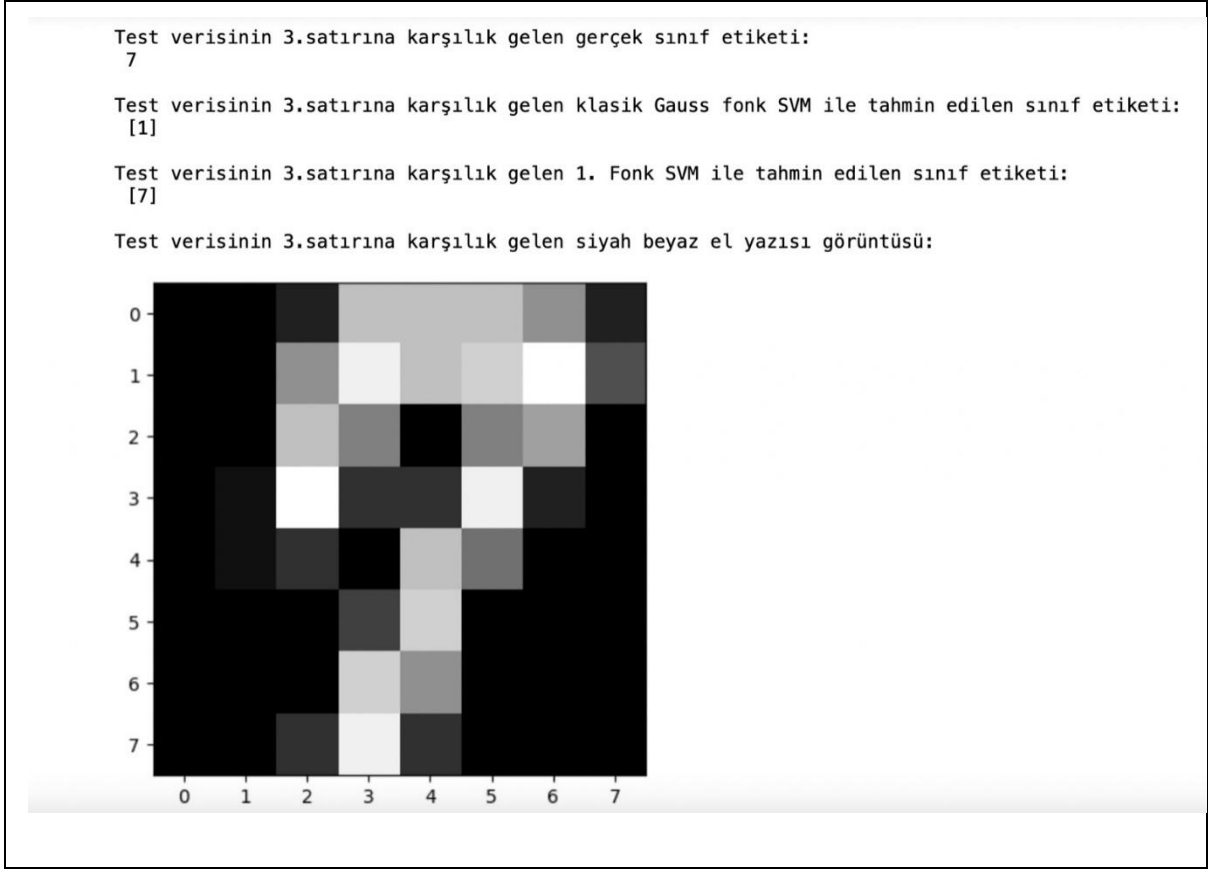
Tablo 10*Fonk.II, Digits Veri Seti için F1 Tablosu*

$\gamma \backslash \beta$	.005	.01	.07	.1	Ortalama
.001	.89013	.92286	.97560	.97933	.94198
.005	.95024	.97962	.99060	.99091	.97784
.06	.97827	.98580	.98769	.98534	.98427
.1	.97637	.98580	.98769	.98216	.98300
1.5	.97637	.98580	.98769	.98252	.98309
2	.97637	.98580	.98769	.98252	.98309
50	.97637	.98580	.98769	.98252	.98309
600	.97637	.98580	.98769	.98252	.98309
1000	.97637	.98580	.98769	.98252	.98309
RBF	.95575	.76397	.01694	.01694	.43840

Tablo 9 ve Tablo 10’da verilen değerler digits veri seti için Fonk.II ile elde edilen ACC ve F1 değerleridir. Tablo 9 ve Tablo 10 Fonk.II’nin digits veri seti üzerinde klasik Gauss fonksiyonuna göre destek vektör makinesi algoritmasında çok daha etkili bir şekilde çalışan bir çekirdek olduğunu göstermektedir.

Şekil 11

Digits Veri Seti için Modelin Uygulanmasının Ekran Görüntüsü



Bu tablolarda elde edilen performans değerlerinden faydalanarak $\beta = 0.06$, $\gamma = 0.07$ değerleri ile elde edilen Fonk.II kullanan destek vektör makinesi modeli test verisinden rastgele bir seçim üzerinde aşağıdaki ekran görüntüsünde görüldüğü gibi uygulanmıştır.

Şekil 11’de digits veri setinin test verisinin 3. satırının değerleri verilmiştir. Şekil 11’deki uygulama Fonk.II’in digits veri seti üzerinde klasik Gauss’tan daha iyi çalışabildiğinin somut bir örnek veri üzerinde bir uygulamasıdır.

Bu çalışmada teorik alt yapısı elde edilen fonksiyonların her birinin wine ve digits veri setleri üzerinde yapılan uygulamalarda bu örneklerde verildiği gibi daha iyi çalıştığı

görülmektedir. Uygulama yapılabilmesi için kodlar tezin diğer kodları ile birlikte <https://github.com/ayseur933/pyhton-code-.git> adresine yüklenmiştir.

Tablo 11*Fonk.III, Wine Veri Seti için ACC Tablosu*

$\beta \backslash \gamma$	.005	.01	.07	.1	Ortalama
.001	.75925	.75925	.64814	.59259	.68980
.005	.75925	.75925	.62962	.62962	.69443
.06	.75925	.75925	.74074	.64814	.72684
.1	.75925	.75925	.74074	.66666	.73148
1.5	.75925	.75925	.72222	.68518	.73148
2	.75925	.75925	.72222	.68518	.73148
50	.75925	.75925	.72222	.68518	.73148
600	.75925	.75925	.72222	.68518	.73148
1000	.75925	.75925	.72222	.68518	.73148
RBF	.66666	.64814	.40740	.40740	.53240

Tablo 12*Fonk.III, Wine Veri Seti için F1 Tablosu*

$\beta \backslash \gamma$	.005	.01	.07	.1	Ortalama
0.001	.74317	.75467	.61267	.55020	.66518
0.005	.74444	.74907	.59414	.59745	.67128
0.06	.74444	.74317	.72605	.60920	.70572
.1	.74444	.74317	.72605	.63625	.71248
1.5	.74444	.74317	.70300	.66182	.71311
2	.74444	.74317	.70300	.66182	.71311
50	.74444	.74317	.70300	.66182	.71311
600	.74444	.74317	.70300	.66182	.71311
1000	.74444	.74317	.70300	.66182	.71311
RBF	.63806	.63601	.26137	.22252	.43949

Tablo 11 ve Tablo 12’de verilen değerler wine veri seti için Fonk.III ile elde edilen ACC ve F1 değerleridir. Tablo 11 ve Tablo 12’de Fonk.III’ün wine veri seti üzerinde klasik Gauss fonksiyonuna göre çok daha etkili bir şekilde çalışan bir çekirdek olduğunu göstermektedir.

Tablo 13*Fonk.III, Digits Veri Seti için ACC Tablosu*

$\beta \backslash \gamma$	.005	.01	.07	.1	Ortalama
.001	.98518	.99074	.97592	.90740	.96481
.005	.98333	.98888	.97777	.94444	.97360
.06	.97407	.98518	.98703	.98333	.98240
.1	.97592	.98518	.98703	.98333	.98286
1.5	.97592	.98518	.98703	.98148	.98240
2	.97592	.98518	.98703	.98148	.98240
50	.97592	.98518	.98703	.98148	.98240
600	.97592	.98518	.98703	.98148	.98240
1000	.97632	.98518	.98703	.98148	.98240
RBF	.95555	.70555	.09259	.09259	.46157

Tablo 14*Fonk.III, Digits Veri Seti için F1 Tablosu*

$\beta \backslash \gamma$	.005	.01	.07	.1	Ortalama
.001	.98580	.99098	.97610	.91464	.96688
.005	.98373	.98928	.97810	.94548	.97414
.06	.97431	.98580	.98769	.98427	.98301
.1	.97637	.98580	.98769	.98427	.98353
1.5	.97637	.98580	.98769	.98252	.98309
2	.97637	.98580	.98703	.98252	.98293
50	.97637	.98580	.98769	.98252	.98309
600	.97637	.98580	.98769	.98252	.98309
1000	.97690	.98580	.98769	.98252	.98322
RBF	.95575	.76397	.01694	.01694	.43840

Tablo 13 ve Tablo 14’de verilen değerler digits veri seti için Fonk. III ile elde edilen ACC ve F1 değerleridir. Tablo 13 ve Tablo 14’de Fonk.III’ün digits veri seti üzerinde klasik Gauss fonksiyonuna göre çok daha etkili bir şekilde çalışan bir çekirdek olduğunu göstermektedir.

7. Sonuç ve Tartışma

Destek vektör makinesinde çekirdek fonksiyonlarının kullanımı son yılların popüler araştırmalarından biri olup, yeni çekirdek fonksiyonlarını geliştirme çabaları giriş bölümünde detaylı olarak verilmiştir.

Temel çekirdek fonksiyonlarından biri olan Gauss çekirdek fonksiyonu, lineer olmayan ayrılabilir veri kümeleri üzerinde belirgindir. Bu çekirdek fonksiyonu tanıtılmış ve destek vektör makinesindeki kullanımı açıklanmıştır.

Bu çalışmada, etkili bir çekirdek fonksiyonu olan Gauss çekirdeği kullanılarak üç farklı tipte Mercer çekirdek fonksiyonu oluşturulmuştur. Bu yaklaşım genelleştirilmiş Gauss fonksiyonları olarak adlandırdığımız bir tamamen monoton fonksiyonlar kümesi oluşturmaktadır.

Elde edilen Mercer çekirdek fonksiyonları klasik veri setleri üzerinde destek vektör makinesine konularak çalıştırıldığında destek vektör makinesinin performansında üst düzey iyileştirmelere sebep olduğu gözlemlenmiştir.

Bu çalışmada sunulan yaklaşımla oluşturulup etkili çalışabilecek başka çekirdek fonksiyonların araştırılması, mevcut Mercer fonksiyonlarının farklı kombinasyonları ile elde edilen yeni Mercer çekirdeklerinin performans analizi ve farklı veri setleri üzerinde yeni oluşturulan çekirdeklerin denenmesi ilk akla gelen açık problemler olarak görülmektedir.

Kaynakça

- Apaydin, C., & Abdikan, S. (2021). Fındık Bahçelerinin Sentinel-2 Verileri Kullanılarak Piksel Tabanlı Sınıflandırma Yöntemleriyle Belirlenmesi. *Geomatik*, 6(2), Art. 2. <https://doi.org/10.29128/geomatik.705988>
- Ayat, N. E., Cheriet, M., Remaki, L., & Suen, C. Y. (2001). KMOD - a new support vector machine kernel with moderate decreasing for pattern recognition. Application to digit image recognition. *Proceedings of Sixth International Conference on Document Analysis and Recognition*, 1215-1219. <https://doi.org/10.1109/ICDAR.2001.953976>
- Ayhan, S., & Erdoğan, Ş. (2014). Destek Vektör Makineleriyle Sınıflandırma Problemlerinin Çözümü İçin Çekirdek Fonksiyonu Seçimi. *Eskişehir Osmangazi Üniversitesi İktisadi ve İdari Bilimler Dergisi*, 9(1), Art. 1. <https://dergipark.org.tr/en/pub/oguiibf/issue/5712/76473>
- Cortes, C., & Vapnik, V. (1995). Support-vector networks. *Machine Learning*, 20(3), 273-297. <https://doi.org/10.1007/BF00994018>
- Cervantes, J., Garcia-Lamont, F., Rodríguez-Mazahua, L., & Lopez, A. (2020). A comprehensive survey on support vector machine classification: Applications, challenges and trends. *Neurocomputing*, 408, 189-215. <https://doi.org/10.1016/j.neucom.2019.10.118>
- Dabanoğlu, M. B. (2018). Makine öğrenmesi ile Gerçek Zamanlı Videodan Yüz İfadesi Analizi. *Yüksek Lisans Tezi .Karadeniz Teknik Üniversitesi,Fen Bilimleri Enstitüsü*.
- Demirci, D. A. (2007). Destek vektör makineleri ile karakter tanıma. Yüksek lisans tezi .*Yıldız Teknik Üniversitesi,Fen Bilimleri Enstitüsü*

Elen, A., Baş, S., & Közkurt, C. (2022). An Adaptive Gaussian Kernel for Support Vector Machine. *Arabian Journal for Science and Engineering*, 47(8), 10579-10588.

<https://doi.org/10.1007/s13369-022-06654-3>

Erdinç , U. (2019, Kasım 5). 02. supervised ve unsupervised learning.

https://erdincuzun.com/makine_ogrenmesi/hangisini-secmeliyim-supervised-ve-unsupervised-learning/ adresinden alındı.[Erişim: 11.11. 2022]

Eray, O. (2008). *Destek vektör makineleri ile ses tanıma uygulaması* [MasterThesis, Pamukkale Üniversitesi Fen Bilimleri Enstitüsü].

<http://acikerisim.pau.edu.tr:8080/xmlui/handle/11499/1501>

Ercan, S. (2020). Destek Vektör Makinelerini Kullanan Teknik Seçmeli Öneri Sistemi.

Yüksek Lisans Tezi .*Sakarya Üniversitesi,Fen Bilimleri Enstitüsü.*

Gao, L., Ye, M., Lu, X., & Huang, D. (2017). Hybrid Method Based on Information Gain and

Support Vector Machine for Gene Selection in Cancer Classification. *Genomics,*

Proteomics & Bioinformatics, 15(6), 389-395.

<https://doi.org/10.1016/j.gpb.2017.08.002>

03Gaussiankernel.nb.(t.y.).13.[https://pages.stat.wisc.edu/~mchung/teaching/MIA/reading/diff](https://pages.stat.wisc.edu/~mchung/teaching/MIA/reading/diffusion.gaussian.kernel.pdf)

[usion.gaussian.kernel.pdf](https://pages.stat.wisc.edu/~mchung/teaching/MIA/reading/diffusion.gaussian.kernel.pdf) adresinden alındı.[Erişim : 22.11.2022]

Gehrke, J. (2003) “Decision Trees”, The Handbook of Data Mining, Editör: Nong Ye,

Lawrence Erlbaum Associates Publishers, London, 149-175.

Güran, A., Uysal, M., & Doğrusöz, Ö. (2014). Destek Vektör Makineleri Parametre

Optimizasyonunun Duygu Analizi Üzerindeki Etkisi. Dokuz Eylül Üniversitesi

Mühendislik Fakültesi Fen ve Mühendislik Dergisi, 16(48), Art. 48.

<https://dergipark.org.tr/en/pub/deumffmd/issue/40797/492168>

Gök, M. (2017). Makine öğrenmesi yöntemleri ile akademik başarının tahmin edilmesi. *Gazi University Journal of Science Part C: Design and Technology*, 5(3), 139-148.

Gökalp, Ö. (2022). *Makine Öğrenmesi—Machine Learning* (s. 14).

Gunn, S. (1998). Support Vector Machines for Classification and Regression, Image Speech and Intelligent Systems Technical Report. Bildiriler Kitabı: 230–67.

Güldoğan, E. (2017). Çeşitli çekirdek fonksiyonları ile oluşturulan destek vektör makinesi modellerinin performanslarının incelenmesi: Bir klinik uygulama. Biyoistatistik ve Tıp Bilişimi Ana Bilim Dalı, Doktora Tezi. *İnönü Üniversitesi, Sağlık Bilimleri Enstitüsü*,

Hiçdurmaz, B. ,Çalık, N.& Üstebay, S.(2023) .Gauss-like Logarithmic Kernel Function to Improve the Performance of Kernel Machines on the Small Datasets, Pattern Recognition Letters, submitted

Karakaya, R. (2020). *Makine öğrenmesi yöntemleriyle el yazısı tanıma* .Yüksek lisans Tezi Sakarya Üniversitesi Fen Bilimleri enstitüsü

Kavzoglu, T., & Colkesen, I. (2010). Destek Vektör Makineleri İle Uydu Görüntülerinin Sınıflandırılmasında Kernel Fonksiyonlarının Etkilerinin İncelenmesi. *Harita Dergisi*, 76, 73-82.

Kılıç, F., Akkaya, M. R., & Memili, N. (2018). Yemekhane için Yapay Zeka Teknikleri Kullanımı ile Günlük Talep Tahmini. *Avrupa Bilim ve Teknoloji Dergisi*, 13, Art. 13. <https://doi.org/10.31590/ejosat.397549>

Konular Veri Madenciliği - Anasayfa. (t.y.). Erişim 22 aralık 2022, <https://web.itu.edu.tr/~sgunduz/courses/verimaden/slides/d3.pdf>

Kul, S., & Sayar, A. (2021). Yapay Sinir Ağları, Destek Vektör Makineleri ve AdaBoost Algoritması ile Araç Sınıflandırmasının Değerlendirilmesi. *Avrupa Bilim ve Teknoloji Dergisi*, 29, Art. 29. <https://doi.org/10.31590/ejosat.1023889>

Köse, U. (2019). Zeki Optimizasyon Tabanlı Destek Vektör Makineleri ile Diyabet Teşhisi. *Politeknik Dergisi*, 22(3), Art. 3. <https://doi.org/10.2339/politeknik.418851>

Kuzu, B. S., & Yakut, S. G. (2020). DESTEK VEKTÖR MAKİNELERİ YARDIMIYLA İMALAT SANAYİSİNDE MALİ BAŞARISIZLIK TAHMİNLERİNİN TEKNOLOJİ YOĞUNLUĞUNA GÖRE İNCELENMESİ. *Osmaniye Korkut Ata Üniversitesi İktisadi ve İdari Bilimler Fakültesi Dergisi*, 4(2), Art. 2. <https://dergipark.org.tr/en/pub/oskauiibfd/issue/59226/812031>

Küçükşille, E. U., & Ateş, N. (2016). Destek Vektör Makineleri ile Yaramaz Elektronik Postaların Filtrelenmesi. *Türkiye Bilişim Vakfı Bilgisayar Bilimleri ve Mühendisliği Dergisi*, 6(1), Art. 1. <https://dergipark.org.tr/en/pub/tbbmd/issue/22246/238807>

Küçük , R. G. (2019, Ağustos). Makine Öğrenmesi Yöntemleri ile Parkinson Hastalığının Teşhis Edilmesi. Yüksek Lisans Tezi . *İstanbul Aydın Üniversitesi Fen bilimleri Enstitüsü* .

Makine Öğrenmesi nedir? - Türkiye | IBM. (t.y.). Geliş tarihi 22 Aralık 2022, gönderen <https://www.ibm.com/tr-tr/cloud/learn/machine-learning>

Makine Öğrenimi Nedir? (t.y.). Geliş tarihi 21 Aralık 2022, gönderen

<https://www.endustri40.com/makine-ogrenimi-nedir/>

Metlek, S., & Kayaalp, K. (2020). Derin Öğrenme ve Destek Vektör Makineleri İle Görüntüden Cinsiyet Tahmini. *Düzce Üniversitesi Bilim ve Teknoloji Dergisi*, 8(3), 2208-2228.

Mitchell, T. M., & Mitchell, T. M. (1997). *Machine learning* (Vol. 1, No. 9). New York: McGraw-hill.

Okwuashi, O., & Ndehedehe, C. E. (2020). Deep support vector machine for hyperspectral image classification. *Pattern Recognition*, 103, 107298.
<https://doi.org/10.1016/j.patcog.2020.107298>

Özekes, S.(2003). Veri Madenciliği Modelleri ve Uygulama Alanları. *İstanbul Ticaret Üniversitesi Fen Bilim. Derg.* 2,(3), 65–82.

Öndeş,E.(2021).“Makine Öğrenmesi Nedir? (1.Bölüm)”.

<https://tr.linkedin.com/pulse/makine-%C3%B6%C4%9Frenmesi-nedir-1-b%C3%B6l%C3%BCm-enes-%C3%B6nde%C5%9F> [Erişim 11 Aralık 2022].

Pekel, E. (2018). Farklı Makine Öğrenmesi Algoritmalarının Karşılaştırılması . Yüksek Lisans Tezi .*Ondokuz Mayıs Üniversitesi Fen Bilimleri Enstitüsü*.

Saatçioğlu ,K.T.(21 Kasım 2013).*Radyal Temelli Fonksiyon Ağları*. prezi.com. (t.y.). Erişim 28.12.2022

Raj, A. (2021, Ocak 5). *The Perfect Recipe for Classification Using Logistic Regression*. Medium.<https://towardsdatascience.com/the-perfect-recipe-for-classification-using-logistic-regression-f8648e267592> adresinden alındı.[Erişim :13.11.2022]

- Rokach, L. (2005). *Data mining and knowledge discovery handbook*. Springer Science+ Business Media, Incorporated.
- Sayıcı, G. (2013). Karar Ağaçları, Bayes Ağları ve Etki Diyagramları Aracılığı ile Bilgi Keşfi ve Karar Verme. *Yayınlanmamış Yüksek Lisans Tezi*
- Shao, M., Wang, X., Bu, Z., Chen, X., & Wang, Y. (2020). Prediction of energy consumption in hotel buildings via support vector machines. *Sustainable Cities and Society*, 57, 102128. <https://doi.org/10.1016/j.scs.2020.102128>
- Schilling, R. L., Song, R., & Vondracek, Z. (1 ekim 2012). *Bernstein functions*. De Gruyter. Erişim .21 kasım 2022, <https://www.degruyter.com/document/doi/10.1515/9783110269338/html>
- Tamamen Monoton, Bernstein ve Stieltjes Fonksiyonları .(21.11.2022). <https://link.springer.com/content/pdf/bbm:978-3-030-29614-8/1.pdf>
- Taya, Ö. (2022). Destek Vektör Makineleri ve Lojistik Regresyon Yöntemi ile sınıflandırma :Türkiye'nin İthalat ve İhracat Örneği .Yüksek Lisans Tezi .*İnönü Üniversitesi Sosyal Bilimler Enstitüsü*.
- Tüminçin, P. D. F., & Baran, K. (2019). SADAB-2019 sadabsempozyum.org. 478.
- UCI Makine Öğrenimi Deposu. (t.y.). Geliş tarihi 06 Ocak 2023, gönderen <https://archive.ics.uci.edu/ml/index.php>
- Yakut, Y. B. E. T. Y., Yakut, E., Elmas, D. dr bekir, & Yavuz, Y. D. D. S. (2014). YAPAY SİNİR AĞLARI VE DESTEK VEKTÖR MAKİNELERİ YÖNTEMLERİYLE BORSA ENDEKSİ TAHMİNİ. *Süleyman Demirel Üniversitesi İktisadi ve İdari Bilimler Fakültesi Dergisi*, 19(1), Art. 1. <https://dergipark.org.tr/en/pub/sduibfd/issue/20816/222712>

Zanbıl, W. S. A. (2022). İkili sınıflandırma problemlerinde çok katmanlı algılayıcı ve destek vektör makineleri sınıflandırıcılarının hiperparametrelerinin en iyilenmesi [Thesis, Kastamonu Üniversitesi].

<http://earsiv.kastamonu.edu.tr:8080/http://earsiv.kastamonu.edu.tr:8080/xmlui/handle/20.500.12597/1937>

Zhang, X., Liu, X., & Wang, Z. J. (2013). Evaluation of a set of new ORF kernel functions of SVM for speech recognition. *Engineering Applications of Artificial Intelligence*, 26(10), 2574-2580. <https://doi.org/10.1016/j.engappai.2013.04.008>

Ekler

Ek 1- Matlab Kodları

Şekil 4- Fonk.I, Sıfır Civarında Matlab Kodu

```
t=0:.01:1.2;
gamma=1;
y_classic=exp(-gamma*t.^2);

beta=.5;
y_beta_1=exp(-gamma*t.^beta);

beta=.4;
y_beta_2=exp(-gamma*t.^beta);

plot(t,y_classic,'o',t,y_beta_1,'x',t,y_beta_2,'g*')
```

Şekil 5 -Fonk.I, Sonsuz Civarında Matlab Kodu

```
function gen_gauss_inf
t=50:.01:60;

gamma=1;
y_classic=exp(-gamma*t.^2);

beta=0.5;
y_beta_1=exp(-gamma*t.^beta);

beta=0.4;
y_beta_2=exp(-gamma*t.^beta);

plot(t,y_classic,t, y_beta_1,'o',t,y_beta_2,'*')

%ylim([-10^(-25) 10^(-5)])
```

Şekil 6 -Fonk.II, Sıfır Civarında Matlab Kodu

```
t=0:.01:1;

gamma=1
y_classic=exp(-gamma*t.^2);

beta=5;
y_beta_1=exp(-gamma*t.*(1-exp(-2*beta*t)));

beta=2;
y_beta_2=exp(-gamma*t.*(1-exp(-2*beta*t)));

%beta=1.5;
%y_beta_3=exp(-gamma*t.*(1-exp(-2*beta*t)));

plot(t,y_classic,'o',t,y_beta_1,'x',t,y_beta_2,'g*')
```

Şekil 7- Fonk.II, Sonsuz Civarında Matlab Kodu

```
function gen_gauss_inf_2_2
t=710:.1:715;

gamma=1;
y_classic=exp(-gamma*t.^2);

beta=0.005;
y_beta_1=exp(-gamma*t.*(1-exp(-2*beta*t)));

beta=1000;
y_beta_2=exp(-gamma*t.*(1-exp(-2*beta*t)));

plot(t,y_classic,'o',t,y_beta_1,'x',t,y_beta_2,'g*')
%ylim([-10^(-25) 10^(-5)])
```

Şekil 8- Fonk.III, Sıfır Civarında Matlab Kodu

```
t=0:.01:1;

gamma=1;
y_classic=exp(-gamma*t.^2);

beta=0.5;
y_beta_1=exp(-gamma*t.*(1+exp(-2*beta*t)));

beta=0.3;
y_beta_2=exp(-gamma*t.*(1+exp(-2*beta*t)));

%beta=1.5;
%y_beta_3=exp(-gamma*t.*(1+exp(-2*beta*t)));

plot(t,y_classic,'o',t,y_beta_1,'x',t,y_beta_2,'g*')
```

Şekil 9- Fonk.III, Sonsuz Civarında Matlab Kodu

```
function gen_gauss_inf_3
t=710:.1:715;

gamma=1;
y_classic=exp(-gamma*t.^2);

beta=0.005;
y_beta_1=exp(-gamma*t.*(1+exp(-2*beta*t)));

beta=1000;
y_beta_2=exp(-gamma*t.*(1+exp(-2*beta*t)));

plot(t,y_classic,'o',t,y_beta_1,'x',t,y_beta_2,'g*')
%ylim([-10^(-25) 10^(-5)])
```

Ek-2 Wine ve Digits Veri Seti İçin Pyhton Kodları

<https://github.com/ayseur933/pyhton-code-.git>

Özgeçmiş

Kişisel Bilgiler

Adı Soyadı: Ayşenur Erdem

Mail: aysenurerdm@gmail.com

Eğitim

Lisans	Gazi Üniversitesi-Matematik (Farabi)	2013-2014
Lisans	Amasya Üniversitesi-Matematik	2011-2016
Yüksek Lisans	İMÜ UMHB Tezli YL	2019-

Sertifikalar

Pedagojik Formasyon Sertifikası- Amasya Üniversitesi (2016)

Yayınlar

Hiçdurmaz, B., Erdem, A., Destek Vektör Makinesi Sınıflandırma Algoritması için Genişletilmiş Gauss Çekirdeği ve Uygulamaları, Abstract Book of the Conference ICAAM 2022, 85, 31.10.2023-6.11.2023, Antalya, Turkey.