

DESTEK VEKTÖR MAKİNELERİ YARDIMIYLA İMALAT SANAYİSİNDE MALİ BAŞARISIZLIK TAHMİNLERİNİN TEKNOLOJİ YOĞUNLUĞUNA GÖRE İNCELENMESİ

Geliş Tarihi: 17.10.2020

Bilge SERDARER KUZU¹

Kabul Tarihi: 02.12.2020

Doç. Dr. Selay GİRAY YAKUT²

Makale Türü: Araştırma Makalesi

Özet

Bu çalışmanın amacı, farklı yoğunluklarda teknoloji kullanımına sahip imalat firmalarının başarısızlıklarını Destek Vektör Makineleri algoritmaları ile tahmin ederek çekirdek fonksiyonlarının performanslarını karşılaştırmak ve farklı teknoloji yoğunluğu kullanıma sahip imalat firmalarının finansal açıdan başarısız olmalarını etkileyen önemli değişkenlerin neler olduğunu incelemektir. Çalışmanın uygulamasında Borsa İstanbul'a kayıtlı, imalat sektöründe faaliyet gösteren 153 firmanın 2012-2015 dönemine ait 11 mali oranı ve üretimde kullandıkları teknoloji yoğunluğunu gösteren teknoloji sınıfı kategorik değişkeni kullanılmıştır. Yapılan analizler sonucunda finansal başarısızlığın tahmini uygulamalarında Destek Vektör Makinelerinin ve tüm çekirdek fonksiyonlarının kullanılabilmesi görülmüştür. Firmaların başarısız olmasını etkileyen önemli değişkenlerin tespit edilmesinde Karar Ağaçları algoritmalarından J48 kullanılmıştır. Tüm teknoloji sınıflarında finansal başarısızlık üzerinde en çok etkisi olan değişken grubunun kârlılık oranları olduğu görülmüştür.

Anahtar Kelimeler: Mali Başarısızlık, Teknoloji Yoğunluğu, Destek Vektör Makineleri

JEL Kodları: C45, C51

EXAMINATION OF FINANCIAL FAILURE ESTIMATES ACCORDING TO TECHNOLOGICAL DENSITY WITH THE HELP OF SUPPORT VECTOR MACHINES

Abstract

The aim of this study, by estimating the failures of manufacturing firms using a different density of technology with Support Vector Machines algorithms, is to compare core function performances and examine the variables that affect manufacturing firms with different technological density to fail financially. During the execution of the study, 11 financial ratios of 2012-2015 and technology class categorical variables that show the density of technology used by the companies in manufacturing from 153 companies registered in Borsa Istanbul and active in the manufacturing industry are used. As a result of the analysis made, it was found that Support Vector Machines and all core functions can be used in the estimation of financial failure. During the determination of the important variables that affect the failure of the firms, the Decision Tree algorithm J48 is used. It was found that profitability ratios are the group that is the most effective on financial failure in all technology classes.

Keywords: Financial Failure, Technological Density, Support Vector Machines

JEL Codes: C45, C51

¹ Marmara Üniversitesi, Sosyal Bilimler Enstitüsü, Dr. Öğrencisi, bilgeserdarer@gmail.com, ORCID: 0000-0002-7064-8174

² Marmara Üniversitesi, İktisat Fakültesi, Ekonometri Bölümü, selaygiray@marmara.edu.tr, ORCID: 0000-0003-4002-7956

1. GİRİŞ

Veriden öğrenme, sınırlı sayıda eğitim verisinden yola çıkılarak tahmin fonksiyonunun belirlenmesini sağlayan öğrenme makinesinin veya algoritmanın oluşturulması olarak tanımlanır. Öğrenme işlemi sınıflandırma, regresyon, kümeleme ve nitelik çıkarımı gibi işlemleri kapsar. Makine öğrenmesi teknikleri istatistiksel olarak dağılıma bağlı çalışan teknikler ve dağılımdan bağımsız çalışan teknikler olmak üzere ikiye ayrılır. Bu teknikler ayrıca kendi içlerinde denetimli ve denetimsiz öğrenen teknikler olarak da ikiye ayrılır (Çomak 2008:19). Denetimli öğrenmede veri seti, etiketleme adı verilen bir işleme tabi tutulur. Etiketleme işleminde girdilere karşılık gelen çıktı değişkenleri de veri seti olarak sisteme girilir. Denetimli öğrenme algoritmalarının amacı, etiketlenmiş eğitim verilerini analiz ederek girdiler ve çıktılar arasındaki ilişkiyi ifade eden bir fonksiyon ortaya koymaktır. Daha sonra bu fonksiyon etiketlenmemiş verilerin analizinde kullanılır. Etiketlenmiş verilerin kullanılmasının avantajı çıktıların doğruluğunun denetlenebilir olmasıdır. Denetimsiz öğrenme algoritmalarında veri seti etiketlenmemiş olarak işlenir. Algoritma veriyi analiz ederek bir yapı ortaya koymaya çalışır (Gürsakaç 2018:94). Klasik istatistiksel yöntemlerin çoğunun verilerin dağılımına ilişkin katı varsayımları vardır. Bu varsayımların pratikte sağlanması oldukça güçtür. Bu yüzden özellikle de verinin olasılıksal dağılımının bilinmediği durumlarda dağılımdan bağımsız tekniklere ihtiyaç duyulur. Destek Vektör Makineleri; dağılımdan bağımsız olarak uygulanabilen; denetimli ya da yarı denetimli olarak sınıflandırma ve regresyon problemlerine çözüm sunabilen bir makine öğrenmesi tekniğidir (Çomak 2008:19).

2. LİTERATÜR TARAMASI

Altman (1968), iflas riski olan işletmeleri tahmin etmek için Diskriminant Analizi kullanmıştır. 66 işletmenin 1946-1965 yılları arasındaki verilerinin kullanıldığı çalışmada ilk olarak 22 değişken incelenmiş ve modelde iflas riski üzerinde etkisi olduğu görülen 5 değişken kullanılmıştır. İflas riskini belirlemede etkili bulunan oranlar şunlardır: Döner sermaye/toplam aktifler, geçmiş yılların kârı/toplam aktifler, faiz ve vergi öncesi kazanç/toplam aktifler, sermayenin piyasa değeri/toplam borcun defter değeri, satışlar/toplam aktifler. Altman'ın bu çalışmasından elde ettiği modele bağlı hesaplanan değer literatürde "Altman Z skoru" olarak bilinmektedir.

Shin vd. (2005), firmaların iflas etme durumlarını Destek Vektör Makineleri ve Geri Yayılımlı Yapay Sinir Ağları ile tahmin etmek istedikleri çalışmalarında, Kore Kredi Garanti Fonu'ndan aldıkları 2320 firmaya ait 52 özellik kullanmışlardır. Yapılan analizler sonucunda Destek Vektör Makinelerinin, Geri Yayılımlı Yapay Sinir Ağlarına göre özellikle de küçük veri setlerinde çok daha başarılı olduğunu belirtmişlerdir.

Min ve Lee (2005), Kore'de küçük ve orta işletmelerin büyümesini teşvik eden bir kredi kuruluşundan aldıkları verilerle firmaların iflas risklerini tahmin etmeyi amaçlamışlardır. Kullandıkları veri seti 1888 gözlemden ve 38 finansal değişkenden oluşmaktadır. Çalışmada Destek Vektör Makineleri, Çoklu Diskriminant Analizi, Lojistik Regresyon ve Yapay Sinir Ağları kullanılarak her birinin sınıflandırma performansı karşılaştırılmıştır. Yapılan incelemeler sonucunda en iyi tahmin performansının Destek Vektör Makineleri ile elde edildiği belirtilmiştir.

Chandra vd.(2009) çalışmalarında, Wharton Research Data Services'ten elde ettikleri 240 dotcom firmasına ait 24 finansal özelliği kullanarak firmaların başarısızlıklarını tahmin etmeyi amaçlamışlardır. Bu çalışmada Çok Katmanlı Algılayıcı, Rastgele Orman, Lojistik Regresyon, Destek Vektör Makineleri ve Sınıflandırma ve Regresyon Ağaçları algoritmaları kullanılmıştır. Tüm özelliklerin kullanıldığı analizlerde doğru sınıflandırma yüzdelerine göre

algoritmalar şu şekilde sıralanmıştır: Rastgele Orman, Sınıflandırma ve Regresyon Ağaçları, Destek Vektör Makineleri, Çok Katmanlı Algılayıcı ve Lojistik Regresyon.

Li vd. (2010), Şangay Borsası ve Shenzhen Borsasında yer alan 153 firmanın 4 adet finansal oranını kullanarak firmaların başarısızlıklarını tahmin etmeyi amaçlamışlardır. Çalışmalarında; Çoklu Diskriminant Analizi, Lojistik Regresyon, Destek Vektör Makineleri, k-En Yakın Komşu, Sınıflandırma ve Regresyon Ağaçları yöntemlerini kullanmışlardır. Çalışmalarının sonucunda en iyi performansı gösteren yöntemin Sınıflandırma ve Regresyon Ağaçları olduğunu belirtmişlerdir.

Ertan ve Ersan (2018), imalat sektöründe finansal başarısızlığı modellemek için 2000-2014 yılları arasında Borsa İstanbul'da işlem gören 208 işletmeye ait veri setini kullanmışlardır. Finansal başarısızlığı, işletmelerin ulusal pazardan Yakın İzleme Pazarı'na geçişleri olarak tanımlamışlardır. Tahmin yöntemleri olarak Cox orantılı riskler, log-logistic, panel probit, logit, tamamlayıcı log-log ve rassal etkiler modelleri kullanılmıştır. Bu çalışmada, finansal başarısızlığı tahmin etmede kullanılan bağımsız değişkenler şu ana başlıklar altında incelenmiştir: Likidite, faaliyet, finansal kaldıraç, karlılık, kurumsal yönetim. Yapılan analizler sonucunda en başarılı tahmini yapan yöntemin Cox orantılı riskler yöntemi olduğu tespit edilmiştir.

Aksoy ve Boztosun (2019), makine öğrenmesi sınıflama algoritmaları kullanarak imalat sanayi sektöründeki firmaları finansal başarısızlıklarına göre sınıflandırmayı amaçlamışlardır. Veri seti olarak Borsa İstanbul imalat sanayi sektörüne kayıtlı 86 firmaya ait 2010-2012 dönemi verileri kullanılmış ve finansal başarısızlıktan 1 ve 2 yıl öncesi için sınıflandırma tahminleri yapılmıştır. Çalışmada kullanılan algoritmalar; Yapay Sinir Ağları, CART Sınıflandırma ve Regresyon ağaçları, Destek Vektör Makineleri ve k-En Yakın Komşu Algoritmasıdır. İşletmelerin başarılı-başarısız olarak sınıflandırılmasında hem finansal tablolara dayanan hem de finansal tablolara dayanmayan özel durum açıklaması ile belirtilen göstergeler kullanılmıştır. Sınıfların tahmin edilmesinde ise 14 özellik kullanılmıştır. Elde edilen bulgulara göre finansal başarısızlıktan hem 2 yıl öncesinin hem de 1 yıl öncesinin verileri ile en iyi tahmini yapan algoritmanın Yapay Sinir Ağları olduğu tespit edilmiştir.

Yürük ve Ekşi (2019), Borsa İstanbul'da işlem gören imalat sanayi sektöründeki 181 işletmenin 2008-2016 yılları arasındaki verilerini kullanarak firma başarısızlığını tahmin etmeyi amaçlamışlardır. Sınıflandırma algoritması olarak Yapay Sinir Ağları ve Destek Vektör Makineleri kullanılmıştır. Çalışmada likidite, finansal yapı, faaliyet ve karlılık oranları başlıkları altında toplanabilen 26 değişken kullanılmıştır. Başarısızlık yılından sırayla 1, 2 ve 3 yıl geriye gidilerek yapılan tahminlerin performansları karşılaştırıldığında Yapay Sinir Ağlarının daha başarılı olduğu görülmüştür.

Arslantürk Çöllü vd. (2020), Borsa İstanbul'da listelenen dokuma, giyim eşyası ve deri sektöründeki 20 şirketin, 2016-2018 dönemine ait finansal oranlarını kullanarak veri madenciliği algoritmalarının finansal başarısızlığı tahmin etmedeki başarılarını tespit etmeyi amaçlamışlardır. Çalışmalarında 22 adet finansal oran kullanarak CAHID, Exh-CHAID, CART ve QUEST algoritmalarının tahmin başarılarını karşılaştırmışlardır. Çalışma sonucunda kullanılan algoritmalarından en başarılı olanın CART olduğu; çalışmaya konu olan şirketlerin başarı ve başarısızlık durumları üzerinde en fazla etkisi olan değişkenin ise özsermaye karlılığı olduğu sonucuna ulaşmışlardır.

3. VERİ SETİ VE DEĞİŞKENLER

Bu çalışmanın amacı; DVM algoritmasının çekirdek fonksiyonlarının farklı teknoloji yoğunluğuna sahip imalat sanayi işletme gruplarının mali başarısızlık tahminindeki

performanslarını karşılaştırmak ve farklı teknoloji yoğunluğuna sahip imalat sanayi işletme gruplarında mali başarısızlığı en çok etkileyen mali oranın ne olduğunu tespit etmektir. Bunun için Borsa İstanbul'a kayıtlı, imalat sektöründe faaliyet gösteren 164 firmanın 2012-2015 dönemine ait bilanço ve gelir tablolarından elde edilen 11 adet mali oran kullanılmıştır. Ayrıca çalışmada, firmanın üretimde kullandığı teknolojinin yoğunluğunu gösteren 1 adet kategorik değişken de kullanılmıştır. Yapılan incelemeler sonucunda, eksik verisi bulunan firmalar veri setinden çıkarılarak 153 firma ile analizler yapılmıştır. Firmaların mali başarısızlığı hedef değişkeni; 3 yıl üst üste zarar etme (Aktaş, 1993) ve 2 yıl üst üste zarar etme (Güriş vd., 2017) olmak üzere iki farklı şekilde ele alınmıştır. Mali başarısızlığın farklı tanımlarından, bu ikisi baz alınarak bu çalışma gerçekleştirilmiştir. Ayrıca iki farklı hedef değişken kullanılarak, çalışma daha detaylı bir hale getirilmiştir. Çalışmada kullanılan mali oranlar; finansal başarısızlık literatüründe de sıklıkla kullanılan karlılık oranları, faaliyet oranları, finansal oranlar ve likidite oranlarıdır. Adı geçen bu mali oranlar ile yapılan finansal başarısızlık çalışmalarına şu örnekler verilebilir: İslam (2020), Giray Yakut ve Bacaksız (2019), Uyar (2019), Ertan ve Ersan (2018), Güriş vd. (2017), Selimoğlu ve Orhan (2015), Yakut ve Elmas (2013). Kullanılan mali oranların detaylı açıklamaları tablo 1'de yer almaktadır.

Tablo 1: Çalışmada kullanılan değişkenler

Hedef Değişkenler	3 yıl üst üste zarar etme (2013-2015)
	2 yıl üst üste zarar etme (2014-2015)
Üretimde Kullanılan Teknoloji Yoğunluğu	Düşük teknoloji sınıfı
	Orta-düşük teknoloji sınıfı
	Orta-ileri teknoloji sınıfı
Kârlılık Oranları	Net kâr/Toplam aktifler (kar1)
	Brüt kâr/Net satışlar (kar2)
	Net kâr/Net satışlar (kar3)
Faaliyet Oranları	Net satışlar/Toplam özkaynaklar (faal1)
	Net satışlar/Duran varlıklar (faal2)
	Net satışlar/Toplam varlıklar (faal3)
Finansal Oranlar	Toplam borçlar/Toplam varlıklar (fin1)
	Toplam borçlar/Toplam özkaynaklar (fin2)
	Kısa vadeli borçlar/Toplam varlıklar (fin3)
Likidite Oranları	Dönen varlıklar/Kısa vadeli borçlar (cari oran) (lik1)
	(Dönen varlıklar-stoklar)/Kısa vadeli borçlar (asit-test oranı) (lik2)

4. YÖNTEM

4.1. Destek Vektör Makineleri

Destek vektör makineleri (DVM), Vapnik-Chervonenkis teorisine dayanan, güçlü temelleri olan denetimli bir makine öğrenme algoritmasıdır. DVM; sinir ağları ve radyal tabanlı yapay sinir ağlarına benzerlik gösterse de genellikle bu algoritmalarından daha iyi performans gösterir. DVM'ler pazarlama, metin tanıma ve görüntü sınıflandırma gibi birçok gerçek hayat probleminde yaygın olarak kullanılmaktadır. Bu yaygın kullanımın sebebi; diğer yöntemlere göre hesaplama kolaylığı, ölçeklenebilirlik ve aykırı değerlere karşı dayanıklılık konusunda önemli gelişmeler kaydedilmesidir. DVM, az sayıda eğitim verisi ve çok sayıda özellik olduğunda bile sınıflandırma ve regresyon problemlerinde iyi performans gösterir. Ayrıca kullanılan veri sayısının bir üst sınırı yoktur; ultra büyük veri setleri için de uygundur (Zigic ve Kecman 2014).

DVM, verilerin doğrusal olarak ayrılabilmesine göre Doğrusal DVM ve Doğrusal Olmayan DVM olarak ikiye ayrılmaktadır.

4.1.1. Doğrusal Destek Vektör Makineleri

X ve Y, R^d 'nin alt kümeleri ve d özellik sayısı olmak üzere, X ve Y'nin elemanlarını farklı taraflara ayırabilen bir hiperdüzlem mevcut ise, X ve Y birbirinden doğrusal olarak ayrılabilir (Elizondo 2006). Doğrusal olarak ayrılabilen durumlar için kullanılan Hard-Marjın ve Soft-Marjın olmak üzere iki tür DVM mevcuttur.

4.1.2. Hard-Marjın Destek Vektör Makineleri

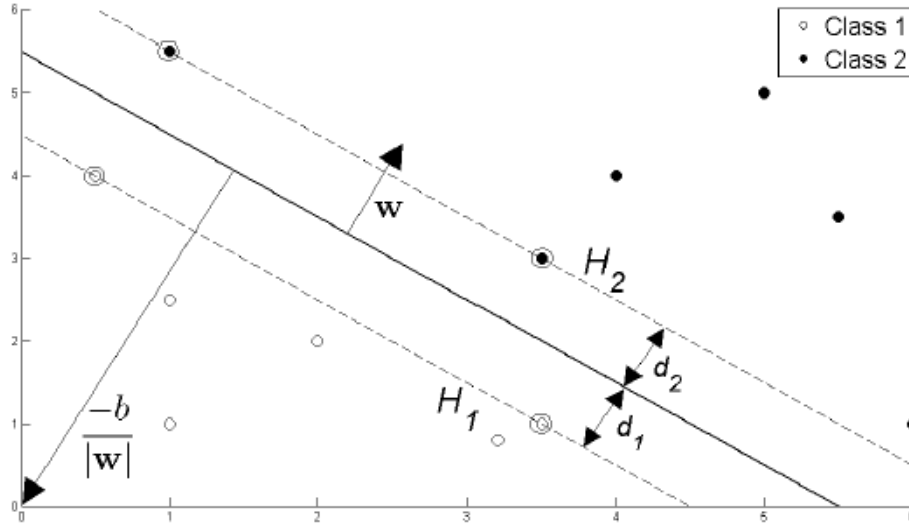
Eğitim verileri $\{x_i, y_i\}$, $i = 1, 2, \dots, L$, $y_i \in \{-1, 1\}$, $x_i \in R^d$ olacak şekilde hedef değişken 2 sıklı girilmiş olsun. Bu eğitim verilerinden pozitif ve negatif olanları birbirinden doğrusal olarak ayırabilen hiperdüzlemin denklemi (1)'deki gibidir. Bu denklemdeki w hiperdüzlemin normali, $\frac{|b|}{\|w\|}$ hiperdüzlemden orijine dik uzaklık, $\|w\|$ w'nin Öklid normudur (Burges 1998).

$$w \cdot x + b = 0 \quad (1)$$

DVM'lerin amacı; farklı sınıflara ait verileri ayıran hiperdüzlemi, tüm sınıfların birbirine en yakın noktalarına en uzak biçimde tutmaktır. DVM aslında w ve b'nin seçimi olarak özetlenebilir. Bu durumda, veri seti denklem (2)'deki gibi tanımlanabilir ve denklem (3)'teki gibi özetlenebilir: (Kartal ve Balaban 2019)

$$x_i \cdot w + b \begin{cases} \geq +1, & y_i = +1 \text{ için} \\ \leq -1, & y_i = -1 \text{ için} \end{cases} \quad (2)$$

$$\forall i \text{ için } y_i(x_i \cdot w + b) - 1 \geq 0 \quad (3)$$



Şekil 1: Doğrusal olarak sınıflandırılabilen durumlar (Hard-Marjın DVM) (Fletcher 2009)

Şekil 1'de H_1 ve H_2 ile gösterilen düzlemler destek düzlemleridir. Bu düzlemler üzerinde kalan birbirlerine en yakın farklı sınıf üyeleri de destek vektörleridir. Ayırıcı düzlem ise destek düzlemlerin tam ortasından geçmektedir. d_1 ve d_2 destek düzlemlerin ayırıcı düzleme uzaklığıdır ve bu iki uzaklık birbirine eşittir. Ayrıca bu uzaklıkların toplamı marjindir. Ayırıcı düzlemin destek vektörlerinden olabildiğince uzakta olması için bu marjının maksimize edilmesi gerekir. Marjın vektör geometrisine dayanarak $\frac{1}{\|w\|}$ olarak gösterilir ve marjının

maksimize edilmesi için $\|w\|$ 'nin minimize edilmesi gerekir. Bu minimizasyon problemi için Lagrange Çarpanları Yöntemi'nden faydalanılır. $\forall i$ için $a_i \geq 0$ olmak üzere;

$$L_p \equiv \frac{1}{2} \|w\|^2 - \sum_{i=1}^L a_i [y_i(x_i \cdot w + b) - 1] \quad (4)$$

$$L_p \equiv \frac{1}{2} \|w\|^2 - \sum_{i=1}^L a_i y_i(x_i \cdot w + b) + \sum_{i=1}^L a_i \quad (5)$$

Bu noktadan sonra amaç, denklem (5)'i minimize eden w ve b ile maksimize eden a 'nın bulunmasıdır. Bunun için ise denklem (5)'in sırasıyla w ve b 'ye göre kısmi türevleri alınarak sıfıra eşitlenir ve bulunan eşitlikler (6) ve (7) denklem (5)'te yerine yazılarak denklem (8) elde edilir.

$$\frac{\partial L_p}{\partial w} = 0 \rightarrow w = \sum_{i=1}^L a_i y_i x_i \quad (6)$$

$$\frac{\partial L_p}{\partial b} = 0 \rightarrow \sum_{i=1}^L a_i y_i = 0 \quad (7)$$

$$\forall i \text{ için } \sum_{i=1}^L a_i y_i = 0 \text{ ve } a_i \geq 0 \text{ olmak üzere,}$$

$$L_D \equiv \sum_{i=1}^L a_i - \frac{1}{2} \left(\sum_{i,j=1}^L a_i a_j y_i y_j x_i \cdot x_j \right) \quad (8)$$

Denklem (8) ile verilen bir maksimizasyon problemidir ve Dual Problem olarak bilinir. Dual Form sadece eğitim veri setindeki gözlemlerin iç çarpımını içerir ve bu durum doğrusal olarak ayrılamayan veri setleri üzerindeki işlemleri kolaylaştırmaktadır. Bu sebeple L_p 'nin minimizasyonu yerine L_D 'nin maksimizasyonu tercih edilir (Kartal ve Balaban 2019:212).

$$H_{ij} = y_i y_j x_i \cdot x_j \quad \forall i \text{ için } \sum_{i=1}^L a_i y_i = 0 \text{ ve } a_i \geq 0 \text{ olmak üzere,}$$

$$\max_a \left[\sum_{i=1}^L a_i - \frac{1}{2} a^T H a \right] \quad (9)$$

Denklem (9)'daki a_i 'lerin çoğu sıfır değerini alır. Sıfırdan farklı olanlar ise destek vektörlerine (x_s) aittir ve x_s 'ler denklem (10)'u sağlar. Buradan b , denklem (11)'daki gibi bulunur. Hesaplamaların daha kesin olması için tüm destek vektörleri üzerinden ortalama alınır ve b denklem (12)'daki şekilde elde edilir.

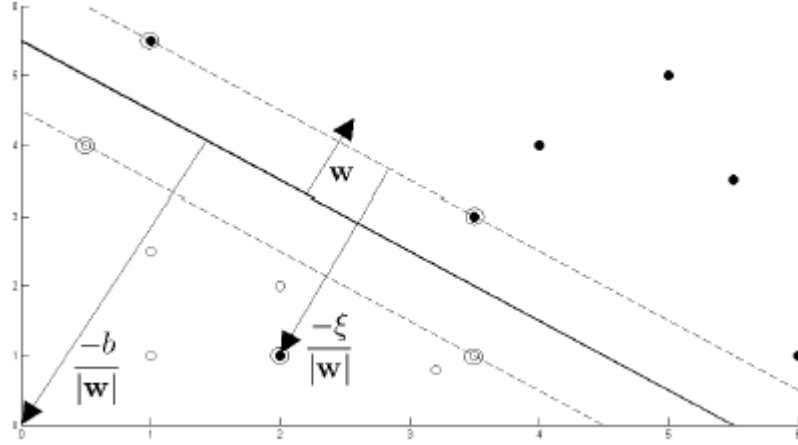
$$y_s(x_s \cdot w + b) = 1 \quad (10)$$

$$b = y_s - \sum_{m \in S} a_m y_m x_m \cdot x_s \quad (11)$$

$$b = \frac{1}{N_s} - \sum_{s \in S} (a_m y_m x_m \cdot x_s) \quad (12)$$

4.1.3. Soft-Marjin Destek Vektör Makineleri

Soft-Marjin DVM, eğitim verilerinin hatasız olarak sınıflandırılmadığı durumlarda kullanılır. Sınıflandırma işleminin en az hata ile yapılabilmesi için hatalı sınıflandırılan veriler eğitim veri setinden çıkarılır (Cortes ve Vapnik 1995).



Şekil 2: Belirli bir hata ile doğrusal olarak sınıflandırılabilen durumlar (Soft-Marjin DVM) (Fletcher 2009)

Soft-Marjin DVM'lerde denklem (3)'ü ifade edebilmek için negatif olmayan bir aylak değişken tanımlanmaktadır. Bu durumda $\xi_i \geq 0$ olmak üzere denklem (3) şu şekilde genelleştirilebilir:

$$\forall i \text{ için } y_i(x_i \cdot w + b) - 1 + \xi_i \geq 0 \quad (13)$$

ξ_i 'nin seçimi birçok durumu bir arada düşünmeyi gerektirir ve seçimler marjini etkiler. ξ_i ile marjin arasındaki dengeyi sağlamak için bir C parametresi kullanılır. C büyük seçilirse daha az hatalı sınıflandırılan veri olur. Fakat aynı zamanda $w^T w$ çarpımının büyük olmasına ve marjinin küçük olmasına sebep olur. Hard-Marjin'de eşitlik (3)'teki gibi tanımlanan Primal Problem, soft-marjin için eşitlik (14)'teki gibi ifade edilir. Daha sonrasında dual form olan eşitlik (15)'in elde edilmesi, b ve w 'nin bulunması süreçleri hard-marjin destek vektör makinelerine benzer şekilde ilerler (Kartal ve Balaban 2019:214).

$$\forall i \text{ için } y_i(x_i \cdot w + b) - 1 + \xi_i \geq 0 \text{ olmak üzere;}$$

$$\min \frac{1}{2} \|w\|^2 + C \sum_{i=1}^L \xi_i \quad (14)$$

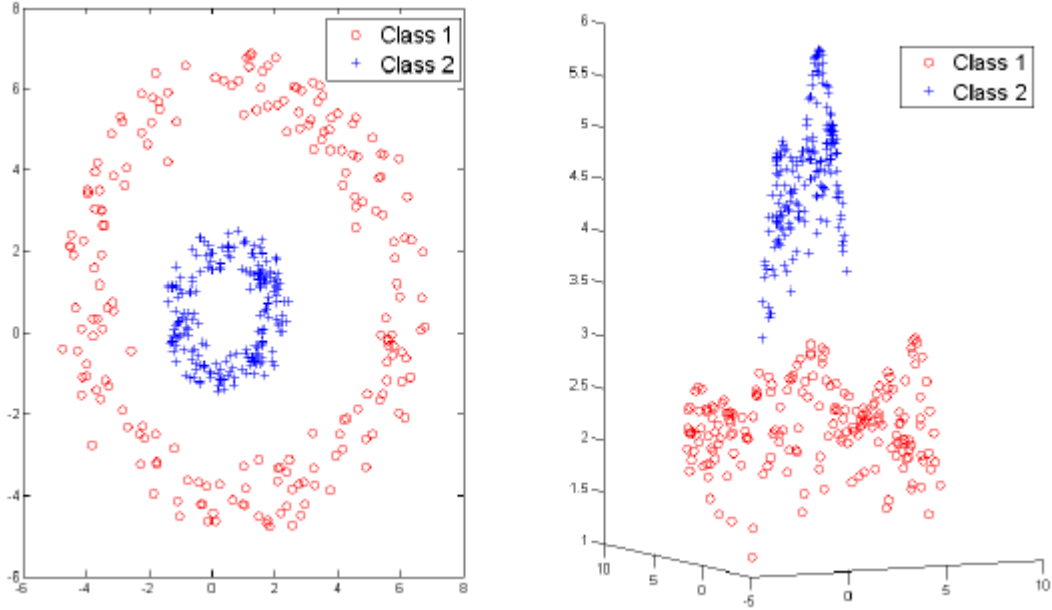
$$\forall i \text{ için } \sum_{i=1}^L a_i y_i = 0, 0 \leq a_i \leq C \text{ olmak üzere,}$$

$$L_D \equiv \sum_{i=1}^L a_i - \frac{1}{2} \left(\sum_{i,j=1}^L a_i a_j y_i y_j x_i \cdot x_j \right) \quad (15)$$

4.1.4. Doğrusal Olarak Ayrılmama Durumu

Gerçek yaşam problemlerinin çoğu doğrusal bir hiperdüzlem ile ayrılmaya uygun değildir. Bu durumda DVM'ler, girdi uzayını daha yüksek boyutlu bir uzaya haritalar. Böylelikle, sınıflar arasında doğrusal bir karar sınırı oluşturulabilir. Bunun için, $x \in R^n$ gözlem vektörü daha yüksek dereceden bir uzayda $z \in R^F$ vektörüne dönüştürülür. $R^n \rightarrow R^F$ eşlemesini yapmak için ϕ fonksiyonu $z = \phi(x)$ şeklinde ifade edilir (Yakut 2012:49).

$$x \in R^n \rightarrow z(x) = [a_1, \phi_1(x), \dots, a_n, \phi_n(x)]^T \in R^F \quad (16)$$



Şekil 3: Doğrusal olarak sınıflandırılmayan durumlarda kernel fonksiyonlarının etkisi (Fletcher 2009:14)

Doğrusal olarak ayrılmayan durumlarda yapılan bu boyut değişikliğini sağlayan haritalama fonksiyonunun bilinmemesi ve yüksek boyutlarda işlem yapmanın zor olması nedeniyle çekirdek düzenlemesi (kernel trick) denilen düzenlemeler yapılır. Böylelikle dönüştürülmüş uzaydaki haritalama fonksiyonu yerine, doğrudan girdi uzayındaki verilerin kullanımı sağlayan çekirdek fonksiyonları işleme dâhil edilmiş olur. Literatürde çok sayıda çekirdek fonksiyonu olmakla birlikte; en sık kullanılanları doğrusal fonksiyon, polinomial fonksiyon, sigmoid fonksiyon ve radyal tabanlı fonksiyonlardır (Aydoğan Culha 2015:15). Vapnik, bu çekirdek fonksiyonlarının performanslarının deneysel olarak büyük bir fark yaratmadığını belirtmiştir. Önemli olan seçilen çekirdek fonksiyonun parametrelerinin belirlenmesidir (Erasto 2001:59). Sık kullanılan çekirdek fonksiyonlarının formülasyonları şöyledir:

$$\text{Doğrusal Fonksiyon: } K(x_i, x_j) = x_i^T x_j \quad (17)$$

$$\text{Polinomial Fonksiyon: } K(x_i, x_j) = (x_i, x_j)^d \quad (18)$$

$$\text{Sigmoid Fonksiyon: } K(x_i, x_j) = \tanh(kx_i x_j - \delta) \quad (19)$$

$$\text{Radyal Tabanlı Fonksiyon: } K(x_i, x_j) = \exp(-\gamma \|x_i - x_j\|^2), \gamma > 0 \quad (20)$$

4.2 Karar Ağaçları

Veri madenciliğinde sınıflama algoritmalarından ağaç modellerinin temel prensibi, özniteliklerin belirli eşik değerlerine göre veri setini homojen alt gruplara ayırmaktır. Böylelikle, hedef değişkende hangi gruba ait olduğu bilinmeyen yeni bir verinin öznitelik değerleri, düğümlerin eşik değerleriyle karşılaştırılarak uygun dallar seçilir ve veri sınıflandırılmış olur (Hand, Mannila, ve Smyth 2001:205). Karar ağaçları denetimli öğrenme algoritmalarıdır. Bu yüzden sınıf etiketi önceden bilinen bir eğitim veri setine ihtiyaç duyulur. Ayrıca hedef değişkenin sınıfları ayrık olmalıdır. Karar ağaçlarında ilk düğüm kök düğüm, diğer düğümler yaprak düğüm ve son düğüm karar düğümü olarak adlandırılır. Bir karar ağacı, kök düğümden başlayarak yukarıdan aşağıya doğru karar düğümüne gelene kadar birbirlerine dallarla bağlanmış yaprak düğümlerden oluşur (Larose, 2005, s. 107,109). Karar ağaçlarında sıklıkla kullanılan bazı algoritmalar şunlardır: ID3, C4.5, CART, AID, CHAID, Exhaustive CHAID ve QUEST (Bilen vd. 2011).

Bu çalışmada karar ağacı algoritmalarından C4.5 kullanılmıştır. C4.5 algoritması ilk karar ağacı algoritması olan ID3'ün geliştirilmiş halidir. ID3 sadece kategorik değişkenler ile çalışabilirken C4.5 hem kategorik hem de sürekli değişkenlerle çalışabilmektedir. Ayrıca ID3 algoritması düğümleri belirlerken bilgi kazancını (eşitlik 22) kullanırken, C4.5 algoritması bilgi kazancı oranını (eşitlik 23) kullanmaktadır. Kategori sayısı fazla olan değişkenlerde, bilgi kazancı fazla çıkmakta ve yanıltıcı olabilmektedir. Bilgi kazancı oranı ile bu dezavantajın önüne geçilmiştir. Bilgi kazancı oranının hesabı entropiye dayanmaktadır. Entropi, bir değişkendeki belirsizliği ortaya koymak için kullanılır. Entropi, en düşük değeri olan sifra değişkenin tüm değerleri birbirine eşit olduğunda ulaşır. k tane düzeyi olan bir Y değişkenine ait entropi eşitlik 21'deki gibi hesaplanır. $p_j, j.$ düzeyin ortaya çıkma olasılığını gösterir ve b değeri değişken iki düzeyli olduğunda 2, ikiden fazla düzeyli olduğunda ise 10 alınır (Altunkaynak 2019:50; Ersöz 2019:74).

$$H(Y) = H(p_1, p_2, \dots, p_k) = \sum_{j=1}^k (p_j \log_b(\frac{1}{p_j})) \quad (21)$$

$$Gain(X_i) = H(Y) - \sum_{j=1}^{k_i} P(X_{ij})H\left(\frac{Y}{X_{ij}}\right); \quad i = 1, 2, \dots, m \quad (22)$$

$$GainRatio(X_i) = \frac{Gain(X_i)}{H(X_i)}; \quad i = 1, 2, \dots, m \quad (23)$$

4.3. Sınıflandırma Performans Ölçütleri

Sınıflandırma algoritmalarının başarılarını belirlemek veya birbirleriyle karşılaştırmak için kullanılan bazı ölçütler vardır. Bu ölçütler, gerçek sınıf değerlerini ve algoritmanın atadığı sınıf değerlerini bir arada gösteren karmaşıklık matrisine göre hesaplanır.

Tablo 2: Karmaşıklık matrisi

	Sınıf Bilgisi	Tahmini Sınıflar	
		Pozitif	Negatif
Gerçek Sınıflar	Pozitif	Doğru pozitif (DP)	Yanlış negatif (YN)
	Negatif	Yanlış pozitif (YP)	Doğru negatif (DN)

Bu çalışmada kullanılan performans ölçütleri ve formülleri aşağıda yer almaktadır (Han vd., 2011, s. 365):

$$Doğru Sınıflama Oranı (DSO) = \frac{DP + DN}{DP + DN + YP + YN}$$

$$\text{Hassasiyet} = \frac{DP}{DP + YP}$$

$$\text{Duyarluluk} = \frac{DP}{DP + YP}$$

$$\text{Belirleyicilik} = \frac{DN}{DN + YP}$$

$$F \text{ Ölçütü} = \frac{2xDP}{2xDP + YP + YN}$$

Kullanılan algoritmaların sınıflandırma performanslarını belirlemek için yukarıdaki ölçütlere ek olarak bir de ROC eğrisi altında kalan alanlarına bakılmıştır. ROC eğrisi; dikey ekseninde duyarlılık, yatay ekseninde 1-belirleyicilik değerleri olacak şekilde çizilir. Eğrinin altında kalan alanın 0,5 olması verilerin tamamen şansa bağlı olarak sınıflandırıldığını; 1 olması ise gerçeğe tam uyumlu olarak sınıflandırıldığını gösterir. Bu sebeple iyi bir sınıflandırma performansı için bu alanın 1'e yakın olması istenir (Alpar 2017:627).

5. UYGULAMA

5.1. Destek Vektör Makineleri Uygulama Sonuçları

Bu çalışmada, Destek Vektör Makineleri yardımıyla firmaların başarısızlık durumlarına göre sınıflandırılması amaçlanmıştır. Başarısızlık durumu olarak iki farklı değişken belirlenmiş ve analizler tekrarlanmıştır. İlk olarak üç yıl üst üste zarar eden firmalar başarısız kabul edilmiş ve daha sonra iki yıl üst üste zarar eden firmalar başarısız kabul edilmiştir. Destek Vektör Makinelerinde, seçilen çekirdek fonksiyonu ve fonksiyonun belirlenen hiper parametre değerleri algoritmanın performansında belirleyici rol oynamaktadır. Optimal hiper parametrelerin bulunması için ızgara (grid) arama yöntemi kullanılmıştır. Bu yöntemde öncelikle her bir parametre için alt sınır, üst sınır ve bir aralık değeri belirlenir. Yöntem, sınırlar dâhilinde tüm parametre kombinasyonlarını, belirlenen aralık değeri kadar atlayarak dener ve en iyi sınıflandırma performansını veren parametre değerlerini önerir (Ayhan ve Erdoğan 2014). Model performanslarının değerlendirilmesinde k-katlı çapraz geçiş (k-fold cross validation) yöntemi kullanılmıştır. Bu yöntemde, veri belirlenen k değeri kadar parçaya bölünür ve her bir parça bir kez test verisi, kalan k-1 parça da eğitim verisi olarak kullanılır. Uygulama sonunda elde edilen k tane model performansının ortalaması alınarak algoritmanın nihai performansı belirlenir. k değerinin en sık kullanılan değeri 10'dur (Balaban ve Kartal 2018:38). Parametre tahminlerinde ve DVM algoritmasının uygulanmasında DTREG programı kullanılmıştır.

5.1.1. Düşük Teknoloji Sınıfında Olan İşletmeler İçin Yapılan Tahminler

Düşük düzeyde teknoloji yoğunluğuna sahip imalat işletmelerinin, başarı durumlarına göre sınıflandırılmasında hedef değişken 3 yıl üst üste zarar etme seçildiğinde en iyi tahmin performansını gösteren Kernel Fonksiyonu, radyal tabanlı fonksiyon olmuştur. Hedef değişken 2 yıl üst üste zarar etme olduğunda ise polinomial fonksiyonun daha iyi sonuçlar verdiği tespit edilmiştir. Yapılan uygulamalar sonucunda elde edilen sınıflandırma performansları tablo 3'te yer almaktadır.

Tablo 3: Düşük teknoloji sınıfındaki işletmeler için elde edilen sınıflandırma performans ölçütleri

Başarısızlık Durumu	DSO (%)	Hassasiyet	Duyarlılık	Belirleyicilik	F Ölçütü	ROC
3 yıl	92,59	75,00	85,71	96,92	0,80	0,948
2 yıl	92,59	85,71	85,71	95,00	0,857	0,975

5.1.2. Orta-Düşük Teknoloji Sınıfında Olan İşletmeler İçin Yapılan Tahminler

İmalat teknolojisinde orta-düşük sınıfta yer alan işletmeler ile uygulanan DVM algoritmalarında, hedef değişken 3 yıl üst üste zarar etme olduğunda en iyi performansı veren fonksiyonun sigmoid fonksiyon olduğu görülmüştür. Hedef değişkenin 2 yıl üst üste zarar etme seçilmesi halinde ise sigmoid ve polinomial fonksiyonların aynı doğru sınıflandırma oranını verdikleri tespit edilmiştir. Bu sebeple seçim yapabilmek için ROC eğrisi altında kalan alanlarına bakılmış ve polinomial fonksiyonun daha iyi sonuçlar verdiği görülmüştür. Elde edilen sınıflandırma performansları tablo 4'te yer almaktadır.

Tablo 4: Orta-düşük teknoloji sınıfındaki işletmeler için elde edilen sınıflandırma performans ölçütleri

Başarısızlık Durumu	DSO (%)	Hassasiyet	Duyarlılık	Belirleyicilik	F Ölçütü	ROC
3 yıl	87,80	33,33	66,67	97,14	0,444	0,610
2 yıl	85,37	37,50	75,00	60,30	0,500	0,871

5.1.3. Orta-İleri Teknoloji Sınıfında Olan İşletmeler İçin Yapılan Tahminler

Teknoloji sınıfları orta-ileri olan işletmelerin başarısızlık durumunun tahmini için uygulanan DVM algoritmalarında, tüm Kernel fonksiyonlarının aynı doğru sınıflandırma oranına sahip olduğu görülmüştür. Bu sebeple ROC eğrisi altında kalan alanlar karşılaştırılmış ve her iki hedef değişken için de en iyi performansı gösteren Kernel fonksiyonunun, polinomial fonksiyon olduğu tespit edilmiştir. Yapılan uygulamalar ile elde edilen sınıflandırma performans ölçütleri tablo 5'te bulunmaktadır.

Tablo 5: Orta-ileri teknoloji sınıfındaki işletmeler için elde edilen sınıflandırma performans ölçütleri

Başarısızlık Durumu	DSO (%)	Hassasiyet	Duyarlılık	Belirleyicilik	F Ölçütü	ROC
3 yıl	93,55	0,00	0,00	100,00	0,00	0,241
2 yıl	90,32	0,00	0,00	100,00	0,00	0,179

5.2. Karar Ağaçları Uygulama Sonuçları

DVM ile aynı teknoloji sınıfına ait işletmelerin, çeşitli göstergeler ile başarısızlık durumlarına göre sınıflandırılabilmesi görülmüştür. Karar Ağaçları ile de yapılan sınıflandırmalarda en etkili olan değişkenlerin belirlenmesi amaçlanmıştır. Bunun için WEKA programının C4.5 uygulaması olan J.48 algoritması kullanılmıştır. C4.5 algoritması, hem nicel hem de kategorik değişkenlerle çalışabildiği için araştırmalarda sıklıkla kullanılmaktadır. Modellerin performansların değerlendirilmesinde 10 katlı çapraz geçerlilik yöntemi kullanılmıştır.

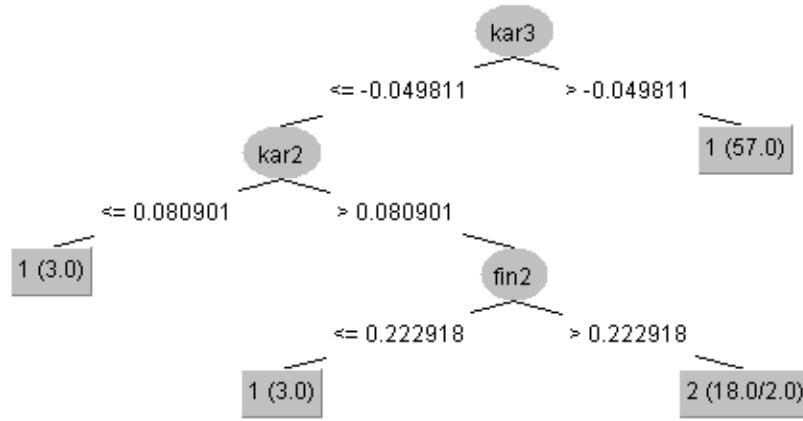
5.2.1. Düşük Teknoloji Sınıfına Ait İşletmeler İçin Karar Ağacı Sonuçları

Düşük teknoloji yoğunluğuna sahip işletmeler için iki farklı Karar Ağacı uygulaması yapılmıştır. Birinci uygulamada hedef değişken 3 yıl üst üste zarar etme durumu iken ikinci uygulamada 2 yıl üst üste zarar etme durumu seçilmiştir. Uygulama sonucunda elde edilen performans ölçüleri tablo 6'da yer almaktadır.

Tablo 6: Düşük teknoloji sınıfından olan işletmeler ile yapılan karar ağacı sınıflandırma sonuçları

Başarısızlık Durumu	DSO (%)	Hassasiyet	Duyarlılık	Belirleyicilik	F Ölçütü	ROC
3 yıl	90,12	0,91	0,97	0,88	0,937	0,889
2 yıl	86,42	0,90	0,92	0,76	0,908	0,813

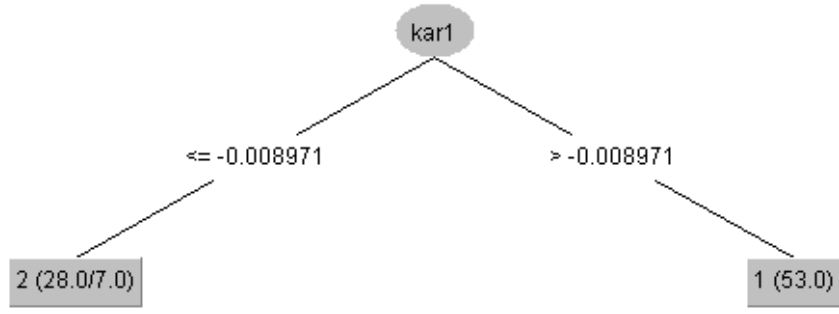
Ayrıca Karar Ağacı uygulanmasındaki asıl amaç, önemli değişkenleri tespit etmek olduğundan elde edilen ağaç diyagramları da incelenecektir. 3 yıl üst üste zarar etme durumunun başarısızlık olarak kabul edildiği durum için elde edilen karar ağacı şekil 4a'da görülmektedir.



Şekil 4a: Düşük teknoloji sınıfından olan işletmeler için karar ağacı

Düşük teknoloji sınıfına ait işletmelerin, 3 yıl üst üste zarar ederek başarısız olma durumlarını tahmin etmede en önemli değişkenler sırasıyla şöyledir: Net kâr/Net satışlar (kar3), Brüt kâr/Net Satışlar (kar2), Toplam borçlar/Toplam özkaynaklar (fin2). Elde edilen ağaç diyagramı detaylı bir şekilde incelendiğinde, ilk olarak bir işletmenin kar3 değerinin -0,049811'den büyük olması durumunda başarılı olacağı tahmin edilmektedir. Kar3 değeri bu eşik değere eşit veya bu değerden küçük olan işletmelerin ise kar2 değeri 0,080901'e eşit veya bu değerden küçük olanların başarılı olacağı öngörülmektedir. Kar2 değeri eşik değerden büyük olan işletmelerin başarı tahminleri için ise fin2 değerlerine bakılmaktadır. Fin2 değeri 0,222918'e eşit veya bu değerden küçük olanların başarılı olacağı tahmin edilmektedir.

2 yıl üst üste zarar etme durumunun başarısızlık olarak kabul edildiği durum için elde edilen karar ağacı şekil 4b'de yer almaktadır.



Şekil 4b: Düşük teknoloji sınıfından olan işletmeler için karar ağacı

Düşük teknoloji yoğunluğuna sahip işletmelerin 2 yıl üst üste zarar ederek başarısız olma durumlarını tahmin etmede önemli değişkenin sadece Net kâr/Toplam aktifler (kar1) olduğu tespit edilmiştir. Kar1 değeri -0,008971'den büyük olan işletmelerin başarılı; kar1 değeri -0,008971'den büyük olmayanların ise başarısız olacakları tahmin edilmektedir.

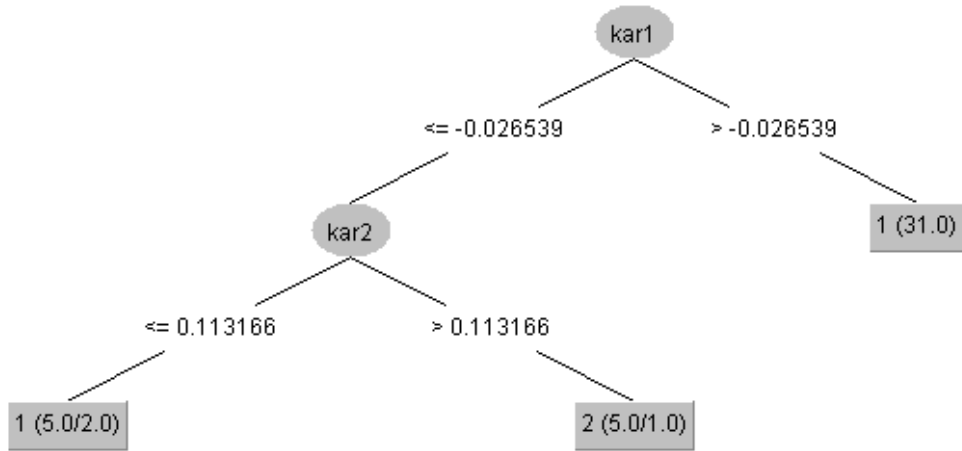
5.2.2. Orta-Düşük Teknoloji Sınıfına Ait İşletmeler İçin Karar Ağacı Sonuçları

Orta-düşük teknoloji yoğunluğuna sahip işletmeler için de üst üste zarar etme sürelerine göre iki adet karar ağacı uygulaması yapılmıştır. Yapılan uygulamalar sonucunda elde edilen sınıflama performansları tablo 7'de yer almaktadır.

Tablo 7: Orta-düşük teknoloji sınıfından olan işletmeler ile yapılan karar ağacı sınıflandırma sonuçları

Başarısızlık Durumu	DSO (%)	Hassasiyet	Duyarlılık	Belirleyicilik	F Ölçütü	ROC
3 yıl	87,80	0,89	0,97	0,83	0,925	0,821
2 yıl	90,24	0,91	0,97	0,88	0,937	0,860

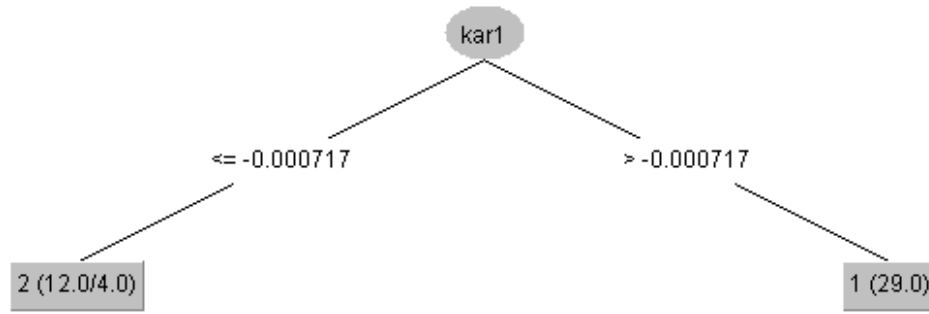
Orta-düşük teknoloji sınıfındaki işletmelerin, 3 yıl üst üste zarar ederek başarısız olma durumlarındaki önemli değişkenleri ve değişkenlerin eşik değerlerini gösteren karar ağacı şekil 5a'da görülmektedir.



Şekil 5a: Orta-düşük teknoloji sınıfından olan işletmeler için karar ağacı

Orta-düşük teknoloji yoğunluğuna sahip işletmelerin, 3 yıl üst üste zarar ederek başarısız olma durumlarının tahmininde önemli değişkenlerin Net kâr/Toplam aktifler (kar1) ve Brüt kâr/Net Satışlar (kar2) olduğu tespit edilmiştir. Orta-düşük teknoloji sınıfındaki bir işletmenin kar1 değeri -0,026539'dan büyük ise başarılı olacağı tahmin edilmektedir. Eğer işletmenin kar1 değeri -0,026539'dan büyük değil ise bu durumda kar2 değişkenine bakılarak tahmin yapılabilmektedir. Kar2 değeri, 0,113166'dan küçük veya 0,113166'ya eşit olan işletmelerin başarılı; 0,113166'dan büyük olan işletmelerin ise başarısız olacağı tahmin edilmektedir.

Başarısızlığın 2 yıl üst üste zarar etme olarak tanımlandığı orta-düşük teknoloji sınıfındaki işletmeler için elde edilen karar ağacı şekil 5b'de görülmektedir.



Şekil 5b: Orta-düşük teknoloji sınıfından olan işletmeler için karar Ağacı

Orta-düşük teknoloji yoğunluğuna sahip işletmelerin, 2 yıl üst üste zarar ederek başarısız olma durumlarının tahmininde önemli değişkenin sadece Net kâr/Toplam aktifler (kar1) olduğu tespit edilmiştir. Kar1 değeri -0,000717'den büyük olan işletmelerin başarılı, -0,000717'ye eşit veya -0,000717'den küçük olan işletmelerin ise başarısız olacakları tahmin edilmiştir.

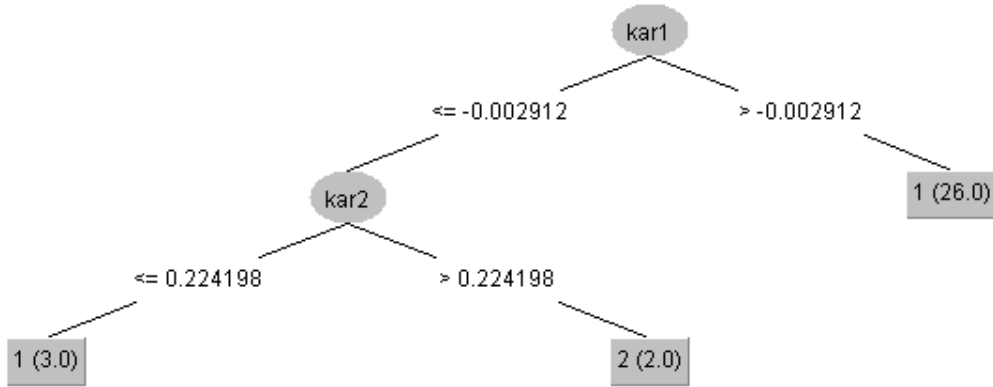
5.2.3. Orta-İleri Teknoloji Sınıfına Ait İşletmeler İçin Karar Ağacı Sonuçları

Orta-ileri teknoloji yoğunluğuna sahip işletmeler için de üst üste zarar etme sürelerine göre iki adet karar ağacı uygulaması yapılmıştır. Yapılan uygulamalar sonucunda elde edilen sınıflama performansları tablo 8'de yer almaktadır.

Tablo 8: Orta-ileri teknoloji sınıfından olan işletmeler ile yapılan karar ağacı sınıflandırma sonuçları

Başarısızlık Durumu	DSO (%)	Hassasiyet	Duyarlılık	Belirleyicilik	F Ölçütü	ROC
3 yıl	93,55	1,00	0,94	0,00	0,967	0,905
2 yıl	87,10	0,93	0,93	0,33	0,929	0,554

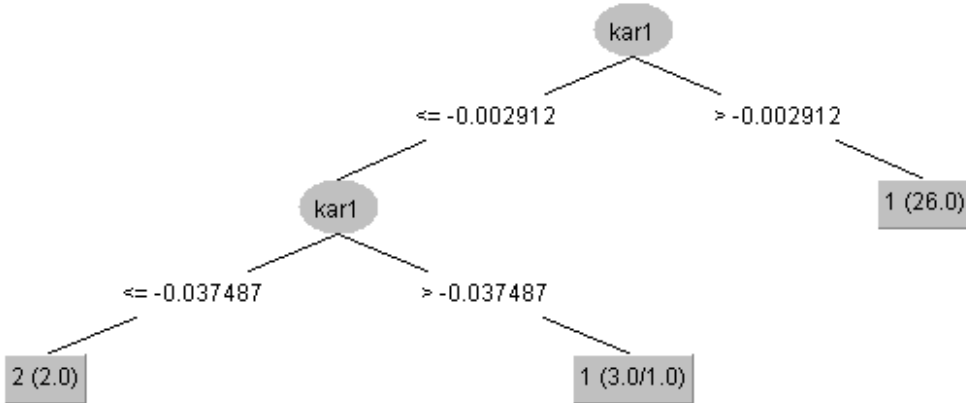
Orta-ileri teknoloji sınıfındaki işletmelerin, 3 yıl üst üste zarar ederek başarısız olma durumlarındaki önemli değişkenleri ve değişkenlerin eşik değerlerini gösteren karar ağacı şekil 6a'da görülmektedir.



Şekil 6a: Orta-ileri teknoloji sınıfından olan işletmeler için karar ağacı

Orta-ileri teknoloji sınıfına ait işletmelerin, 3 yıl üst üste zarar ederek başarısız olma durumlarını tahmin etmede en önemli değişkenlerin Net kâr/Toplam aktifler (kar1) ve Brüt kâr/Net Satışlar (kar2) olduğu belirlenmiştir. İlk olarak kar1 değeri -0,002912'den büyük olan işletmelerin başarılı olacakları yönünde tahminde bulunulabilir. Daha sonra kar1 değeri -0,002912'den büyük olmayan işletmelerden kar2 değeri 0,224198'den büyük olanların başarılı, kar2 değeri 0,224198'den büyük olmayanların ise başarısız olacakları söylenebilir.

Orta-ileri teknoloji sınıfına ait işletmelerin başarısızlığının, 2 yıl üst üste zarar etme olarak tanımlandığı durum için elde edilen karar ağacı şekil 6b'de görülmektedir.



Şekil 6b: Orta-ileri teknoloji sınıfından olan işletmeler için karar Ağacı

2 yıl üst üste zarar ederek başarısız sayılan, orta-ileri düzeydeki işletmelerin başarısızlığını belirlemede en önemli değişkenin Net kâr/Toplam aktifler (kar1) olduğu tespit edilmiştir. İlk önce kar1 değeri -0,002912'den büyük işletmelerin başarılı olacağı tahmin edilmektedir. Daha sonra kar1 değeri -0,002912'den büyük olmayanlardan ise kar1 değeri -0,037487'den büyük olanların başarılı; kar1 değeri -0,037487'den büyük olmayanların ise başarısız olacakları tahmin edilmektedir.

6. SONUÇ

Finansal başarısızlığın öngörülebilmesi, işletmelerin önlem alması ve buna göre plan yaparak başarısızlığa sebep olan etmenleri düzeltmeye çalışması açısından çok önemlidir. Bu nedenle çalışmada, olası bir finansal başarısızlığın tahmin edilebilmesi ve bu başarısızlığa sebep olan en önemli değişkenlerin belirlenmesi amaçlanmıştır.

Bu çalışmada; Borsa İstanbul'a kayıtlı, imalat sektöründe faaliyet gösteren işletmelerin 2012-2015 dönemlerine ait mali oranları ve üretimde kullandıkları teknolojinin yoğunluğunu belirten teknoloji sınıfı değişkeni kullanılmıştır. Yapılan literatür taraması sonucunda; finansal başarısızlığın incelendiği çalışmalarda sıklıkla kullanılan değişkenlerin karlılık oranları, faaliyet oranları, finansal oranlar ve likidite oranları olduğu görülmüştür. Bu sebeple çalışmada bu değişkenlere yer verilmiştir. Ayrıca imalat sektöründeki işletmelerin, üretimde teknolojiyi hangi yoğunlukta kullandıklarının da finansal başarısızlık üzerinde etkisi olduğunu gösteren az sayıda çalışma bulunmaktadır. Bu sebeple de işletmelerin ait oldukları teknoloji sınıfı değişkeni de çalışmada kullanılmış ve yapılan analizler her teknoloji sınıfı için tekrarlanmıştır. Finansal başarısızlığın göstergesi olarak da yine literatürde sıklıkla kullanılan üç yıl üst üste zarar etme ve iki yıl üst üste zarar etme durumları ayrı ayrı kullanılmıştır.

Üretimde kullanılan teknoloji yoğunluğuna göre farklı sınıflarda bulunan işletmelerin, finansal başarısızlık durumlarının tahmini için bu çalışmada Destek Vektör Makineleri (DVM) kullanılmıştır. DVM, hem sayısal hem kategorik değişkenlerle çalışabilen, küçük örneklerde ve aykırı değer olma durumunda bile başarılı sonuçlar üretebilen bir sınıflandırma algoritması olması sebebiyle tercih edilmiştir. DVM uygulamalarında, doğrusal olarak ayrılamayan durumlarda veriyi doğrusal ayrılabilen duruma getiren kernel fonksiyonları kullanılır. Bu fonksiyonların ve hiper parametrelerinin seçimi önem taşımaktadır. Bu yüzden en iyi sonucu veren kernel fonksiyonunun ve hiper parametrelerinin seçimi için ızgara (grid) yöntemi kullanılmıştır. DVM algoritmasının sonuçları aşağıda özetlenmiştir.

Finansal başarısızlığın göstergesi olarak üç yıl üst üste zarar etme durumu hedef değişken olarak kabul edildiğinde;-, en iyi doğru sınıflama oranını veren kernel fonksiyonları şöyledir: düşük teknoloji yoğunluğuna sahip işletmelerde %92,59 ile radyal tabanlı fonksiyon, orta-düşük teknoloji yoğunluğu olanlarda %87,80 ile sigmoid fonksiyonu ve orta-ileri teknoloji yoğunluğu grubunda ise tüm fonksiyonlardan aynı sonuç elde edilerek doğru sınıflama oranının %93,55 olduğu tespit edilmiştir. Hedef değişkenin iki yıl üst üste zarar etme seçildiği durumlarda ise düşük teknoloji yoğunluğuna sahip işletmelerde en iyi sınıflandırmayı yapan kernel fonksiyonunun %92,59 ile polinomial fonksiyon, orta-düşük teknoloji yoğunluğu grubunda %85,37 ile hem sigmoid hem de polinomial fonksiyon olduğu, orta-ileri teknoloji yoğunluğu grubunda ise tüm kernel fonksiyonlarının %90,32 doğru sınıflandırma oranını verdiği tespit edilmiştir.

Çalışmanın ikinci amacı, farklı teknoloji yoğunluğuna sahip işletmelerde finansal başarısızlık üzerinde etkisi olan değişkenleri tespit etmektir. Bunun için hem etkili değişkenleri hem de etkili değişkenlerin eşik değerlerini tespit edebilmek adına Karar Ağaçları uygulanmıştır. Hem sayısal hem de kategorik değişkenler ile çalışabilen C4.5 algoritmasının bir uygulaması olan J48 algoritması kullanılmıştır. Karar Ağaçları ile elde edilen sonuçlar aşağıda özetlenmiştir.

Finansal başarısızlığın üç yıl üst üste zarar etme kabul edildiği durumda, farklı teknoloji yoğunluğu grubundaki işletmeler için başarısızlığı en çok etkileyen değişkenler sırasıyla şöyledir:

- Düşük teknoloji yoğunluğuna sahip işletmeler için: Net kâr/Net satışlar, Brüt kâr/Net satışlar, Toplam borçlar/Toplam özkaynaklar
- Orta-düşük teknoloji yoğunluğuna sahip işletmeler için: Net kâr/Toplam aktifler, Brüt kâr/Net satışlar
- Orta-ileri teknoloji yoğunluğuna sahip işletmeler için: Net kâr/Toplam aktifler, Brüt kâr/Net satışlar

Finansal başarısızlığın iki yıl üst üste zarar etme kabul edildiği durumda, tüm teknoloji yoğunluğu gruplarında başarısızlığı en çok etkileyen değişkenin Net kâr/Toplam aktifler olduğu görülmüştür.

Literatürde finansal başarısızlığın tahmininde daha çok klasik istatistiksel yöntemlerin kullanıldığı, makine öğrenmesi yöntemlerini kullanan çalışmalarda ise Destek Vektör Makinelerinin diğer yöntemlere göre nispeten az kullanıldığı görülmüştür. Çalışmanın bu açıdan literatüre katkı sağlaması amaçlanmıştır. Ayrıca tüm teknoloji sınıfı gruplarında finansal başarısızlık üzerinde en etkili değişkenler incelendiğinde, kârlılık oranları değişken grubunun ön planda olması bu çalışmanın önemli bulgularındandır. Bu çalışmadan elde edilen bulgular; bilgileri kullanılan firmalar, ele alınan zaman aralığı ve kullanılan değişkenler ile kısıtlıdır.

KAYNAKÇA

- Aksoy, B. ve Boztosun, D. (2019). *İmalat işletmelerinde makine öğrenmesi yöntemleri kullanılarak finansal başarısızlık tahmini ve sınıflandırma performansının karşılaştırılması: borsa istanbul örneği*. 2. Uluslararası Bankacılık Kongresi'nde sunulan bildiri, Çorum. Erişim adresi: <http://earsiv.hitit.edu.tr/xmlui/bitstream/handle/11491/5403/bankac%C4%B1%C4%B1k-kongresi2019.pdf?sequence=1&isAllowed=y#page=11>
- Aktaş, R. (1993). *Endüstri İşletmeleri İçin Mali Başarısızlık Tahmini*. Ankara: Türkiye İş Bankası Kültür Yayınları.
- Alpar, R. (2017). *Uygulamalı Çok Değişkenli İstatistiksel Yöntemler*. Ankara: Detay Yayıncılık.
- Altman, E. I. (1968). Financial ratios, discriminant analysis and the prediction of corporate bankruptcy. *The Journal of Finance*, 23(4): 589-609.
- Altunkaynak, B. (2019). *Veri Madenciliği Yöntemleri ve R Uygulamaları*. Ankara: Seçkin Yayıncılık.
- Arslantürk Çöllü, D., Akgün, L. ve Eydurun, E. (2019). Karar ağacı algoritmalarıyla finansal başarısızlık tahmini: Dokuma, giyim eşyası ve deri sektörü uygulaması. *Uluslararası Ekonomi ve Yenilik Dergisi*, 6 (2):225-246.
- Aydoğan Culha, Ü. (2015). *Yüksek boyut düşük örneklem genişliği durumunda sınıflama algoritmalarının performanslarının karşılaştırılması*. (Yayınlanmamış Doktora Tezi). Hacettepe Üniversitesi, Sağlık Bilimleri Enstitüsü, Ankara.
- Ayhan, S. ve Erdoğmuş, Ş. (2014). Destek vektör makineleriyle sınıflandırma problemlerinin çözümü için çekirdek fonksiyonu seçimi. *Eskişehir Osmangazi Üniversitesi İktisadi ve İdari Bilimler Dergisi*, 9(1):175-201.
- Balaban, M. E. ve Kartal, E.(2018). *Veri Madenciliği ve Makine Öğrenmesi Temel Algoritmaları ve R Dili İle Uygulamaları*. Çağlayan Kitabevi.
- Bilen, Ö., Hotaman, D., Aşkin, Ö.E., ve Büyüklü, A.H. (2011). LYS başarılarına göre okul performanslarının eğitsel veri madenciliği teknikleriyle incelenmesi: 2011 İstanbul örneği. *Eğitim ve Bilim*, 139(172):78-94.
- Burges, C.J. (1998). A tutorial on support vector machines for pattern recognition. *Data Mining and knowledge discovery*, 2(2),121-167.
- Chandra, D. Karthik, Ravi, V. ve Bose, I. (2009). Failure prediction of dotcom companies using hybrid intelligent techniques. *Expert Systems with Applications*, 36(3):4830-37.

- Cortes, C., ve Vapnik, V. (1995). Support-vector networks. *Machine Learning*, 20(3):273-97.
- Çomak, E. (2008). *Destek vektör makinelerinin etkin eğitimi için yeni yaklaşımlar*. Yayınlanmamış Doktora Tezi, Selçuk Üniversitesi, Fen Bilimleri Enstitüsü, Elektrik Elektronik Mühendisliği Anabilim Dalı, Konya.
- Elizondo, D. (2006). The linear separability problem: some testing methods. *IEEE Transactions on Neural Networks*, 17(2):330-44.
- Erasto, P. (2001). *Support vector machines - backgrounds and practice*. Yayınlanmamış Yüksek Lisans Tezi, Rolf Nevanlinna Institute, Helsinki.
- Ersöz, F. (2019). *Veri Madenciliği Teknikler ve Uygulamaları*. Ankara: Seçkin Yayıncılık.
- Ertan, A. S., ve Ersan, Ö. (2018). Finansal başarısızlığı belirleyen etkenler: Türkiye imalat sektörü örneği. *Marmara Üniversitesi İktisadi ve İdari Bilimler Dergisi*, 40(2):181-207.
- Fletcher, T. (2009). *Support vector machines explained*. Erişim adresi: <http://sutikno.blog.undip.ac.id/files/2011/11/SVM-Explained.pdf>.
- Giray Yakut, S. ve Bacaksız, N.E. (2019). Teknoloji yoğunluğuna göre finansal başarısızlık tahmin modelleri değişir mi? İmalat sanayi sektörü üzerine bir uygulama. *Marmara Üniversitesi İktisadi ve İdari Bilimler Dergisi*, 40(2):536-555.
- Güriş, S., Çağlayan Akay, E., Ün, T., Kızıllarlan, Ş. (2017). Multivariate probit modeli ile finansal başarısızlığın yeniden incelenmesi: Borsa İstanbul örneği. *Social Sciences Research Journal*, 6(3):199-210.
- Gürsakal, N. (2018). *Makine Öğrenmesi*. Bursa: Dora Yayınları.
- Han, J., Kamber. M., ve Pei J. (2011). *Data Mining: Concepts and Techniques*. Third Edition. Elsevier.
- Hand, D. J., Mannila, H., ve Smyth, P. (2001). *Principles of Data Mining*. Cambridge, Mass: MIT Press.
- Islam, M.S. (2020). Predictive capability of financial ratios for forecasting of corporate bankruptcy. *Journal of Business and Management*, 22(6):13-57.
- Kartal, E., ve Balaban M. E. (2019). Destek Vektör Makineleri: Teori ve R Dili ile Bir Uygulama. M. E. Balaban ve E. Kartal (Ed.), *Veri Madenciliği ve Makine Öğrenmesi* içinde (s. 207-239). İstanbul: Çağlayan Kitabevi.
- Larose, D. T. (2005). *Discovering Knowledge in Data: An Introduction to Data Mining*. Hoboken, N.J: Wiley-Interscience.
- Li, H., Sun, J. ve Wu, J. (2010). Predicting business failure using classification and regression tree: An empirical comparison with popular classical statistical methods and top classification mining methods. *Expert Systems with Applications*, 37 (8): 5895-5904.
- Min, J., ve Lee Y. (2005). Bankruptcy prediction using support vector machine with optimal choice of kernel function parameters. *Expert Systems with Applications*, 28(4):603-14.
- Selimoğlu, S. ve Orhan, A. (2015). Finansal başarısızlığın oran analizi ve diskriminant analizi kullanılarak ölçülmesi: BİST'de işlem gören dokuma, giyim eşyası ve deri işletmeleri üzerine bir araştırma. *The Journal of Accounting and Finance*.
- Shin, K., Lee, T., ve Kim, H. (2005). An application of support vector machines in bankruptcy prediction model. *Expert Systems with Applications*, 28(1):127-35.

- Uyar, S. (2019). *Finansal rasyolar yardımıyla mali başarısızlık tahmininde alternatif tekniklerin karşılaştırılması*. Yayınlanmamış Yüksek Lisans Tezi, Marmara Üniversitesi, Sosyal Bilimler Enstitüsü, Ekonometri Anabilim Dalı, İstanbul.
- Yakut, E. (2012). *Veri madenciliği tekniklerinden c5.0 algoritması ve destek vektör makineleri ile yapay sinir ağlarının sınıflandırma başarılarının karşılaştırılması: imalat sektöründe bir uygulama*. Yayınlanmamış Doktora Tezi, Atatürk Üniversitesi, Sosyal Bilimler Enstitüsü, İşletme Anabilim Dalı, Ankara.
- Yakut, E. ve Elmas, B. (2013). İşletmelerin finansal başarısızlığının veri madenciliği ve diskriminant analizi modelleri ile tahmin edilmesi. *Afyon Kocatepe Üniversitesi İİBF Dergisi*, 15(1):237-254.
- Yürük, M. F. ve Ekşi, İ.H. (2019). Yapay zeka yöntemleri ile işletmelerin finansal başarısızlığının tahmin edilmesi: bist imalat sektörü uygulaması. *Mukaddime* 10(1):393-422.
- Zigic, L. ve Kecman, V. (2014). *Variants and performances of novel direct learning algorithms for L2 support vector machines*. International Conference on Artificial Intelligence and Soft Computing'de sunulan bildiri, Cham. Erişim adresi: https://www.researchgate.net/profile/Vojislav_Kecman/publication/295103176_Variants_and_Performances_of_Novel_Direct_Learning_Algorithms_for_L2_Support_Vector_Machines/links/573fad9508ae298602e8f713/Variants-and-Performances-of-Novel-Direct-Learning-Algorithms-for-L2-Support-Vector-Machines.pdf