

T.C.
İSTANBUL ÜNİVERSİTESİ
SOSYAL BİLİMLER ENSTİTÜSÜ
EKONOMETRİ ANABİLİM DALI
İSTATİSTİK BİLİM DALI

YÜKSEK LİSANS TEZİ

TÜRK DEMİR ÇELİK ENDÜSTRİSİ
DEĞERLENDİRMESİNDE RANDOM FOREST
ALGORİTMASI KULLANILARAK BİR MAKİNE
ÖĞRENMESİ UYGULAMASI

Mhd Refeat AL HALLAK

2501191346

TEZ DANIŞMANI

DR. ÖĞR. ÜYESİ Leyla İŞBİLEN YÜCEL

İSTANBUL – 2023

ÖZ

**TÜRK DEMİR ÇELİK ENDÜSTRİSİ
DEĞERLENDİRMESİNDE RANDOM FOREST
ALGORİTMASI KULLANILARAK BİR MAKİNE
ÖĞRENMESİ UYGULAMASI**

Mhd Refeat AL HALLAK

Demir çelik endüstrisinde yapılacak yatırımların, harcamaların ve alınacak her türlü kararın doğru ve etkin bir biçimde gerçekleştirilmesi ülke ekonomisi için kritik bir öneme sahiptir. Bu hassasiyetten doğan bakış açısıyla hazırlanan bu çalışma, Türk Demir Çelik endüstrisinde faaliyette bulunan 12 firmanın performanslarının sınıflandırılması amacıyla makine öğrenmesi yöntemlerinden Random Forest tekniği ile modellenmesini konu almaktadır. Random Forest makine öğrenmesine hedef değişken olarak tasarlanan etkinlik skorları, veri zarflama analizinin CCR ve BCC modelleriyle hesaplanmıştır. Girdi ve çıktı değişkenleri ise Shannon entropi yöntemiyle elde edilmiştir. Toplamda 7 adet girdi ve çıktı kullanılmıştır. Bu çalışmada ayrıca, Random Forest algoritmasıyla gerçekleştirilen sınıflandırmada en yüksek doğruluk düzeyine sahip bulunan BCC modelin etkin bulmadığı firmalar için etkinleştirme önerileri de sunulmuştur. Çalışmada kullanılan bilgisayar programları şu şekildedir: Shannon entropi uygulamaları için Excel, makine öğrenmesi Random Forest algoritması için R, etkinlik skorlarının veri zarflama analizi ile hesaplanması için Frontier Analyst'tir.

Anahtar kelimeler: Makine öğrenmesi, Random Forest Tekniği, Shannon Entropi, VZA, Süper Etkinlik.

ABSTRACT

A MACHINE LEARNING APPLICATION USING RANDOM FOREST ALGORITHM IN EVALUATION TURKISH IRON AND STEEL INDUSTRY

Mhd Refeat AL HALLAK

Every investment, expenditures and every decision going to be taken should be accurate and effective as it plays a critical role in the national economy development. This research has arisen from this sensitive point of view, aims to investigate in classifying the efficiency of 12 firms working in the iron steel sector using machine learning method random forest classifier. The efficiency scores from data envelopment analysis CCR, BCC and Super efficiency models has been calculated to be used as target variables for the random forest classifier. The inputs and outputs have been processed throw Shannon entropy, and in total 7 variables have been used. In this study, recommendations are also presented for firms that are inefficient according to the BCC model, which has the highest level of accuracy in the classification performed with the random forest algorithm. The programs that have been used in this study are; Microsoft excel for Shannon Entropy, R language for random forest classifier algorithm, and for data envelopment analysis it used Frontier analyst.

Keywords: Machine Learning, Random Forest Technique, Shannon Entropy, DEA, Super Efficiency.

ÖNSÖZ

Çağımızın karmaşık ve dinamik iş ortamında endüstriler, hem ulusal hem de küresel ölçekte rekabetçi kalabilmek için üretkenliklerini ve verimliliklerini en üst düzeye çıkarmak için sürekli bir baskı altında bulunmaktadır. Üretimdeki büyüme, firmanın en yüksek verimlilikte çalıştığı anlamına gelmemekte, bu da sürekli değerlendirme çalışmalarına ihtiyaç duymaktadır. Demir çelik sektörü, ekonomik gelişme ve ilerlemenin mihenk taşı konumunda olup, birçok sektör ile doğrudan veya dolaylı etkileri sebebiyle hayati bir rol oynamaktadır.

Bu araştırma, temelde makine öğrenme yöntemlerinden Random Forest tekniği ile sektörün bir modelini elde etmenin yanında, algoritmaya hizmet eden etkinlik skorlarının elde edilmesi sürecinde firmaların etkinliklerinin de veri zarflama analizi ile hesaplanmış olması nedeniyle, her bir firma için etkinlik ve süper etkinlik araştırmalarının da yapılmasını sağlamıştır.

Makine öğrenmesi Random Forest algoritmasına giden yolda, öncelikle sistemi tanımlayan, yani Türk Demir-Çelik sektöründeki üretim sürecini en kapsamlı bir şekilde ortaya kayan girdiler ve çıktılar belirlenmiştir. İkinci adımda; girdi ve çıktılar Shannon entropi yöntemi ile kümelенmiştir. 27 adet olan girdi-çıkıtı sayısı 7'ye indirgenmiştir. Üçüncü adımda; girdi-çıkıtılar üzerinden CCR ve BCC veri zarflama analizi yöntemleri kullanılarak firmaların etkinlik skorları hesaplanmıştır. Son adımda, girdiler-çıkıtılar ve etkinlik skorları üzerinden (etkinlik skorları hedef değer alınarak) Random Forest algoritması uygulanmış sistem bu şekilde modellenmeye çalışılmıştır.

Tez çalışmamın başlangıcından itibaren sürekli sağladığı sarsılmaz teşvik, rehberlik ve destek için tez danışmanım Dr. Öğr. Üyesi Leyla İşbilen Yücel'e en içten ve derin şükranlarımı sunmak isterim. Çabaları sayesinde, yol boyunca karşılaştığım zorlukların üstesinden gelebildim.

Aynı zamanda, programın ilk yılında bana eğitim veren hocalarıma da teşekkür etmek isterim. Onların özverisi, yeni bir eğitim diliyle uyum sağlamama yardımcı oldu.

Yeteneklerime inandıkları ve her engelde benden esirgemedikleri sürekli destek ve motivasyonları için aileme, arkadaşlarıma ve meslektaşlarıma yürekten teşekkürlerimi sunarım.

MHD REFEAT AL HALLAK

İstanbul, 2023

İÇİNDEKİLER

ÖZ	II
ABSTRACT	III
ÖNSÖZ	IV
TABLolar LİSTESİ	IX
ŞEKİLLER LİSTESİ	XI
KISALTMALAR	XII
SİMGELER.....	XIII
GİRİŞ	1

BİRİNCİ BÖLÜM

TÜRKİYE’DE DEMİR ÇELİK ENDÜSTRİSİ

1.1. Türkiye’de Son Dönemde Demir-Çelik İhracatı.....	7
---	---

İKİNCİ BÖLÜM:

ARAŞTIRMA METODOLOJİSİ

2. 1. Makine Öğrenmesi.....	11
2.1. 1. Çapraz Doğrulama Yöntemi.....	14
2.1.2. Karışıklık Matrisi Yöntemi (Confusion Matrix Method).....	15
2.1.3. Random Forest Yöntemi.....	17
2.1.4. Makine Öğrenmesinde Yanlılık ve Varyans.....	19
2.1.5. ROC ve AUC Grafikleri.....	19

2.2. Entropi.....	22
2.2.1. Karar Vermede Entropi.....	26
2.2.2. Entropi Yönteminin Adımları.....	27
2.2.3. Huffman Algoritması.....	27
2.3. Veri Zarflama Analizi.....	29
2.3.1. VZA'da Etkinlik Kavramı.....	30
2.3.2. CCR Model.....	32
2.3.3. BCC Model.....	36
2. 4. Süper Etkinlik.....	39
2. 4.1. Radyal Süper Etkinlik Modelleri.....	39
2.4.2. Radyal Süper Etkinlik VRS Modeli.....	40
2.4.3. Süper Radyal I-V.....	40

ÜÇÜMCÜ BÖLÜM

UYGULAMA

3.1. Veri Kümesi.....	41
3.2. Değişkenler.....	41
3.3. Uygulama adımları	44
3.3.1. Random Forest Algortiması İçin Değişken Kümelerinin Shannon Entropi İle Oluşturulması (Korelasyonların Hesaplanması)	46
3.3.2. Shannon Entropi ve Ağırlıkların Hesaplanması.....	47
3.4. Random Forest Algortiması İçin Shannon Entropi Yöntemiyle	

Elde Edilen Girdi-Çıktılar Üzerinden Hedef Değerlerin Veri	
Zarflama Analizi İle Elde Edilmesi.....	49
3.5. Random Forest Makine Öğrenmesinin Uygulanması.....	52
3.6.BCC Modelin Etkinleştirme Önerileri.....	58
SONUÇ.....	60
KAYNAKÇA.....	62
EKLER.....	69

TABLÖLAR LİSTESİ

Tablo 1. Karışıklık Matrisinin Anlatımı İçin Oluşturulmuş Kurgusal Veriler.....	16
Tablo 2. Karışıklık Matrisi.....	16
Tablo 3. Firmaların Listesi.....	41
Tablo 4. Daha Önce Yapılan Araştırmalarda Kullanılan Değişkenlere Örnekler.....	42
Tablo 5. Değişken Listesi.....	44
Tablo 6. Random Forest Makine Öğrenmesi Tekniğinin Adımları.....	45
Tablo 7. Birinci Girdi Değişken Grubu; Shannon Entropi ve Ağırlık Değerleri.....	48
Tablo 8. İkinci Girdi Değişken Grubu; Shannon Entropi ve Ağırlık Değerleri.....	48
Tablo 9. Birinci Çıktı Değişken Grubu; Shannon Entropi ve Ağırlık Değerleri.....	49
Tablo 10. İkinci Çıktı Değişken Grubu; Shannon Entropi ve Ağırlık Değerleri.....	49
Tablo 11. Girdi ve Çıktı VZA Değişkenleri.....	50
Tablo 12. Uygulamada Kullanılan Veri Kümesi.....	50
Tablo 13. CCR, BCC ve Ölçek Etkinlik Skorları	51

Tablo 14. Süper Etkinlik Skorları.....	52
Tablo 15. “532623” Tohum ile CCR Etkinlik Skorlarını Kullanan Random Forest Algoritmasının R Kodları.....	53
Tablo 16. “532623” Tohum ile BCC Etkinlik Skorlarını Kullanan Random Forest Algoritmasının R Kodları.....	54
Tablo 17. “532623” Tohum ile Süper Etkinlik Skorlarını Kullanan Random Forest Algoritmasının R Kodları.....	55
Tablo 18. “532623” Tohum ile Random Forest Algoritması Modellerinin Sonuçları.....	56
Tablo 19. Random Forest Algoritması Modellerin Farklı Tohumlarla Sonuçları.....	56
Tablo 20. Firma12 BCC Etkinlik Analizi.....	58
Tablo 21. Firma3 BCC Etkinlik Analizi.....	59

ŞEKİLLER LİSTESİ

Şekil 1. Türkiye Çelik Haritası.....	7
Şekil 2. Türk Demir Çelik Endüstrisi.....	8
Şekil 3. Türkiye'nin Sektörlere Göre İhracatı.....	9
Şekil 4 . 2021 Yılında Türkiye'nin Demir Çelik İhracatının Dağılımı.....	10
Şekil 5. Karar Ağacı Örneği.....	12
Şekil 6. Eğitim Verisinin Doğrusal Model İle Tahmin Edilmesinin Gösterimi.....	13
Şekil 7. Orijinal Verinin Eğrisel Modelle Tahmin Edilmesinin Gösterimi.....	13
Şekil 8. Test Verisinin Gösterimi.....	14
Şekil 9. Çapraz Doğrulama Yönteminde Test Verisinin Belirlenmesi.....	15
Şekil 10. Lojistik Regresyon Eğrisinin 0.5'lik Eşik Değer Gösterimi.....	20
Şekil 11. ROC Grafiği.....	21
Şekil 12. AUC Grafiği.....	22
Şekil 13. Entropide Sürpriz ve Olasılık Arasındaki Bağntı.....	23
Şekil 14. Karar Ağacı.....	28
Şekil 15. İki Girdi ve Bir Çıktı İçin Etkin Sınır.....	31
Şekil 16. CCR ve BCC Etkin Sınırları.....	37
Şekil 17. Çıktı Değişkenlerinde Korelasyon.....	46
Şekil 18. Girdi Değişkenlerinde Korelasyon.....	46

KISALTMALAR LİSTESİ

RF	: Random Forest
VZA	: Veri Zarflama Analizi
CCR	: Banker-Charnes-Cooper
BCC	: Chranes-Cooper-Rhodes
VRS	: Variable Return To Scale
CRS	: Constant Return To Scale
PCA-DEA	: Principal Component Analysis – Data Envelopment Analysis
DEA-ANN	: Data Envelopment Analysis – Artifactual Neural Network
LP	: Linear Program
FP	: Fractional Program

SİMGELER LİSTESİ

S	: Slack / gevşek
θ	: Etkinlik Değeri
$\hat{x}_0$	: O Karar Birimi İçin X Girdi Değişkeninin Optimum Değeri
$\hat{y}_0$	: O Karar Birimi İçin Y Çıktı Değişkeninin Optimum Değeri
u_r	: r. Çıktı İçin Tahsis Edilen Ağırlık
v_I	: i. Girdi İçin Tahsis Edilen Ağırlık
$P(x_{ij})$	: x_{ij} . Değişken İçin Olasılık
$E(X)$	: X Değişken İçin Beklenen Değer
H	: Shannon Entropi
x_{ij}	: i. Sütün ve j. Kolonda Karar Matrisindeki Değişken
e_j	: Karar Matrisindeki J Değişken İçin Shannon Entropi
w_j	: Karar Matrisindeki J Değişken İçin Ağırlık
Firma_j	: j. firma (j=1....12)

GİRİŞ

Türk demir çelik endüstrisinde 2021 yılı verileriyle yapılan bu çalışmada; borsaya kote olan, hem üretim kapasitesi hem de piyasa etkisi bakımından önemli bir yer tutan başlıca 12 firma¹ için Random Forest makine öğrenmesi algoritması yöntemi kullanılarak firmaların etkinlik düzeylerini en iyi öğrenebilen ve sınıflandırabilen makine öğrenmesi algoritması modellenmeye çalışılmıştır.

Yapılan okumalar ve araştırmalar sonucunda Demir-çelik sektöründeki üretim sürecini en iyi şekilde tanımladığı düşünülen 27 değişken üzerinden Shannon entropi yöntemi kullanılarak, çok sayıdaki değişkenin aralarındaki yüksek korelasyonlar, entropi puanları ve ağırlıklarına bakarak, değişken sayısı 7'ye düşürülmüştür. Verimliliğin artması için teorik olarak azaltılması gereken faktörler girdi olarak, artırılması gereken faktörler ise çıktı olarak kabul edilmiştir. Söz konusu girdi-çıktı kümesi üzerinden CCR, BCC ve süper etkinlik modelleri ile etkinlik ölçümleri yapılmış, firmaların “etkin” ve “etkin olmayan” şeklindeki skorları Random Forest makine öğrenmesi algoritması için hedef değişken olarak belirlenmiştir.

Çalışmanın birincil amacı firmaların etkinliğini ölçmek değildir, etkinlik skorları Random Forest algoritmasının sınıflandırmadaki doğruluk düzeyinin ölçülmesine hizmet etmektedir. Fakat bununla birlikte, firmaların etkinlik analizleri yapılmışken, Random Forest algoritması için önemli bulunan sonuçlar için etkinleştirme önerilerine metin içinde, görece daha az önemli olan etkinleştirme önerilerine ise tezin sonunda “Ekler” kısmında yer verilmiştir.

Bu çalışmada kullanılan veriler 2021 yılına aittir ve İstanbul Menkul Kıymetler Piyasası'nda işlem gören 12 firmanın finansal ve operasyonel verileridir.² Çalışmaya konu

¹ ALMAD, AYES, MSCH, BURCE, CELHA, CEMTS, EREGL, ISDMR, IZMDC, KOZA, KRDM, YKSLN.

² Bu veriler borsadan veri sağlayan kamuya açık internet sitelerinden sağlanmıştır, doğrudan BIST'ten alınmamıştır. Çalışmanın bulguları deneysel olarak kabul edilmelidir ve bahsi geçen firmaların gerçek etkinlik durumlarına, başarı performanslarına atfedilmemelidir. Hesaplanan etkinlik skorları gerçek veriler üzerinden hesaplanmış olmasına rağmen, çalışma boyunca firmaların etkinlik skorları, “Firma'nın etkinlik skoru” şeklinde kodlanarak verilmiştir.

olan tüm girdi ve çıktıları aynı anda birlikte içeren bir tek kaynak bulunamadığı için birkaç farklı kaynak kullanılmıştır.³ Verilerin elde edilmesi esnasında, her bir girdi ve(ya) çıktı için bir tek kaynak kullanılmasına özen gösterilmiştir. Çalışmada, değişkenler arasındaki korelasyonları hesaplamak ve korelasyon grafikleri için RStudio, Shannon entropisini ve ağırlıklarını hesaplamak için Excel, Random Forest algoritması için R, Veri zarflama analizi için de Frontier Analyst programı kullanılmıştır.



³ <https://tr.tradingview.com/markets/stocks-turkey/sectorandindustry-industry/steel/>
<https://tr.investing.com/equities/>
<https://fintables.com/>
<https://www.cemtas.com.tr/sayfa/yatirimci-iliskileri-detay/1170/ozet-mali-tablolar>
<https://halkyatirim.com.tr/skorkart/>

BİRİNCİ BÖLÜM

TÜRKİYE’DE DEMİR ÇELİK ENDÜSTRİSİ

Demir çelik endüstrisi, hem eski devletler ve imparatorlukların hem de modern ülkelerin gelişmesinde kritik bir rol oynamış, modern yaşamımızı şekillendirmiş ve insan uygarlığının çeşitli yönlerinde ve alanlarında devrim yaratan tarihsel bir rol oynamıştır. Eski zamanlarda savunma ve savaşta demir ve çeliğin; kılıç, mızrak ve zırh gibi silahlarda keşfedilmesi ve kullanılması, askeri kuvvetlerde halen bronz veya taş silahlar kullanmakta olan rakiplerine karşı önemli bir avantaj sağlamıştır. Modern yaşamda olduğu gibi demir çelik, askeri uygulamalarda kritik bir malzeme olmaya devam etmektedir. Tanklar, uçak gemileri, denizaltılar ve diğer modern askeri teçhizatlarının önemli bir kısmı halen demir çeliğin gücüne, dayanıklılığına ve korumasına güvenmektedir.

Makine ve imalatta, eski toplumlarda gelişmiş demir ve demir çelik aletlerin kullanılması, daha etkin tarım uygulamalarına ve gelişmiş zanaatkarlığa izin vererek medeniyetlerin büyümesine yol açmıştır. Modern makineler ve imalat için çeliğin ağır endüstriyel ekipmanların yapımından hassas aletlere ve robotiklere kadar temel bir malzeme olmasının yanı sıra, çeliğin gücü ve çok yönlülüğü, imalat süreçlerinde ilerlemelere ve üretkenliğin artmasına olanak sağlamıştır.

Demir çelik endüstrisi, insanların ve malların yerel veya global olarak hareketini sağlayan ulaşım sistemlerinin desteklenmesinde de önemli bir rol oynamaktadır. Otomobil, tren, uçak, gemi ve demiryollarının yapımında kullanılan demir çelik, güvenli ve etkin ulaşım için gerekli gücü ve dayanıklılığı sağlamaktadır.

Demir çelik, enerji ve güç altyapısının geliştirilmesi için de gereklidir. Enerji santralleri, iletim hatları, elektrikli teçhizat yapımında, nükleer, termik ve yenilenebilir enerji tesislerinde ve birçok yerlerde kullanılmaktadır. Türkiye'nin demir çelik endüstrisi, ülkenin endüstriyel ve ekonomik kalkınmasında önemli bir rol oynamıştır. Türkiye, demir çelik üretimi için temel hammaddeler olan demir cevheri ve kömürün zengin rezervleriyle

tanınmaktadır. Bu kaynaklara erişim, ülkenin demir çelik endüstrisi için önemli bir avantaj olmuştur.

Gelişmiş teknoloji ile donatılmış demir çelik fabrikaları ile Türkiye, demir çelik üretimi alanında büyük bir ilerleme kaydetmiştir, ileri teknolojilerin kullanılması, demir çelik üretim süreçlerinin etkinliğini ve üretkenliğini arttırmış, daha yüksek kaliteli ürünlere ve düşük üretim maliyetlerine olanak sağlamıştır.

Dünyanın önde gelen demir çelik ihracatçılarından biri haline gelen Türkiye, kıtalar genelinde çeşitli ülkelere demir çelik ürünleri tedarik ederek hem dış ticaret gelirlerini arttırmış hem de global pazardaki konumunu güçlendirmiştir.

Demir çelik endüstrisindeki büyüme, ekonominin diğer sektörleri üzerinde dalgalanma etkisi yaratmıştır. Endüstriyel faaliyetleri canlandırmış ve istihdam olanakları sağlayarak ekonomik büyüme ve kalkınmayı hızlandırmış, inşaat ve altyapı sektörleri yüksek kaliteli çeliğin mevcudiyetinden yararlanarak modern binaların, köprülerin ve ulaşım ağlarının geliştirilmesini sağlamıştır.

Türkiye'deki sanayi patlamasının sürdürülmesi ve desteklenmesi, ekonomik büyümenin sürdürülmesi ve global pazarda rekabet avantajı yaratılması açısından büyük önem taşımaktadır. Araştırma ve geliştirmeye sürekli yatırım yaparak ve yenilikçi teknolojileri benimseyen Türkiye, demir çelik üretim kapasitesini daha da geliştirebilme, ürün kalitesini iyileştirebilme ve yeni pazarlar keşfedebilme kapasitesine sahiptir.

Türkiye'deki demir çelik endüstrisi, 1937'de ilk fabrikasının kurulmasına kadar uzanan zengin ve köklü bir geçmişe sahiptir. Ekonomik kalkınma yolunda demir çelik sektörü, modernleşme ve kentleşme için gerekli yapı taşlarını sağlayarak ilerlemenin mihenk taşı olarak ortaya çıkmıştır. Türk demir çelik endüstrisi, bu sağlam malzemenin muazzam potansiyelini erkenden fark ederek dönüştürücü bir büyüme ve genişleme yolculuğuna zemin hazırlamıştır.

1939'da Türkiye, artan demir çelik üretiminin temelini oluşturan, adeta çığır açan bir gelişme olan yüksek fırının ilk kullanımına tanık olmuştur. Bir yıl sonra, 1940 yılında,

ülke, endüstriyel kapasitesini artıran ilk demir çelik fabrikasının ve haddehanenin⁴ kurulmasıyla bir başka dönüm noktasına daha ulaşmıştır.

1960'lı yıllar, demir çelik üretiminde devrim niteliğinde bir yöntem olan bazik oksijen fırınının kullanılmaya başlanmasıyla Türk demir çelik endüstrisi için yeni bir dönem başlatmıştır. Bu yenilikçi yaklaşım, daha yüksek üretim oranlarını mümkün kılmış ve demir çelik üretiminin kalitesini yükselterek, Türkiye'yi demir çelik sektöründe global öneme sahip bir konuma taşımıştır.

1970'ler boyunca sektör daha da büyümüştür. 1971'de Türkiye'nin ilk demir çelik ihracatıyla zirveye ulaşmıştır. Türk çeliğinin uluslararası pazarlarda satılmaya başlanması, dünya çapında erişimini ve etkisini artırmasıyla bir dönüm noktası olmuştur.

1980'ler, Türkiye'de demir çelik üretiminde kayda değer bir artışa tanık olmuştur. Ülkenin global demir çelik endüstrisinde önemli bir oyuncu olarak yükselmesine yol açmıştır. 1987'de Türkiye, yılda 7.0 milyon metrik ton üretimle dünyada 20. en büyük demir çelik üreticisi haline gelmiştir. İhracata yönelik bir yaklaşımın benimsenmesiyle, Türk demir çelik sektörünün büyüme rotası yoğunlaşmış, 1993 yılında üretim kapasitesi 11,5 milyon metrik tona ulaşarak dünyanın en büyük 15 demir çelik üreticisi arasına girmiştir.

2000'li yılların başları, stratejik yatırımlar ve kapasite artırımlarıyla karakterize edilen Türk demir çelik endüstrisi için istisnai bir büyüme dönemi olmuştur. İhracat hızla artarak dünyanın en büyük 10 demir çelik ihracatçısı arasında yerini almıştır.

Türkiye'nin global demir çelik pazarında öne çıkması hız kesmeden devam etmiştir. 25,3 milyon metrik tonluk şaşırtıcı üretim kapasitesiyle 2009 yılında dünyanın en büyük 7. demir çelik ihracatçısı olmuştur. 2013 yılına gelindiğinde Türkiye, 34,7

⁴ HADDEHANE birimi, ÇELİKHANE 'de hazırlanmış kütüğün, müşterinin istediği ebatlara toleranslar dahilinde haddelendiği birimdir. Vasıflı çelik üretiminde, üretilen kütüklerin modern teknolojiyi yakalamış ve verimli bir haddehanede haddelenmesi gerekmektedir. (<https://www.cemtas.com.tr/sayfa/uretim/3/haddehane#:~:text=HADDEHANE%20birimi%2C%20%C3%87EL%C4%B0KHANE'de%20haz%C4%B1rlanm%C4%B1%C5%9F,verimli%20bir%20haddehanede%20haddelenmesi%20gerekmektedir.>)

milyon metrik ton üretim kapasitesiyle dünyanın 8. büyük demir çelik üreticisi ve Avrupa'nın 2. büyük ihracatçısı konumunu sağlamlaştırmıştır.

Global ekonomik dalgalanmaların ve talepteki kaymaların yarattığı zorluklara rağmen, Türk demir çelik sektörü kayda değer bir dayanıklılık sergilemiş, 2015 yılında 31,5 milyon metrik ton üretim kapasitesi ile dünyanın en büyük 9. demir çelik üreticisi olmuştur.

2018 yılında Türkiye, global demir çelik arenasındaki konumunu güçlendirmeye devam ederek dünyanın 8. büyük demir çelik üreticisi ve Avrupa'nın 2. büyük ihracatçısı unvanını elde etmiştir.

2021 yılında 40,4 milyon metrik tona ulaşarak dünyanın 7. büyük demir çelik üreticisi ve Avrupa'nın en büyük demir çelik üreticisi olmuştur. Ancak 2022 yılında dünya demir çelik talebi gerileyerek üretimde marjinal bir düşüşe yol açmış ve Türkiye 35,1 milyon metrik tonla dünyanın 8., Avrupa'nın 2. büyük üreticisi olmuştur.

Türk demir çelik endüstrisi, ülkenin ekonomik büyümesinin ve kalkınmasının arkasındaki itici güç olarak ortaya çıkmıştır. Sektörün başarısı, diğerlerinin yanı sıra inşaat, otomotiv ve makine gibi çeşitli imalat sektörlerine sağladığı önemli katkıya da yansımaktadır.

Demir çelik üretiminin %72'sini demir çelik hurdasından sağlayan Türk demir çelik endüstrisi, global ve yerel çevre standartlarını karşılayan ve üretim sürecinde karbon emisyonlarını azaltan sürdürülebilirlik uygulamalarını benimsemiştir.

Ayrıca, Türkiye'nin en büyük inşaat demiri ihracatçısı ve dünyanın en büyük hurda metal ithalatçısı konumu, Türkiye'nin demir çelik pazarındaki global önemini bir kez daha vurgulamaktadır.

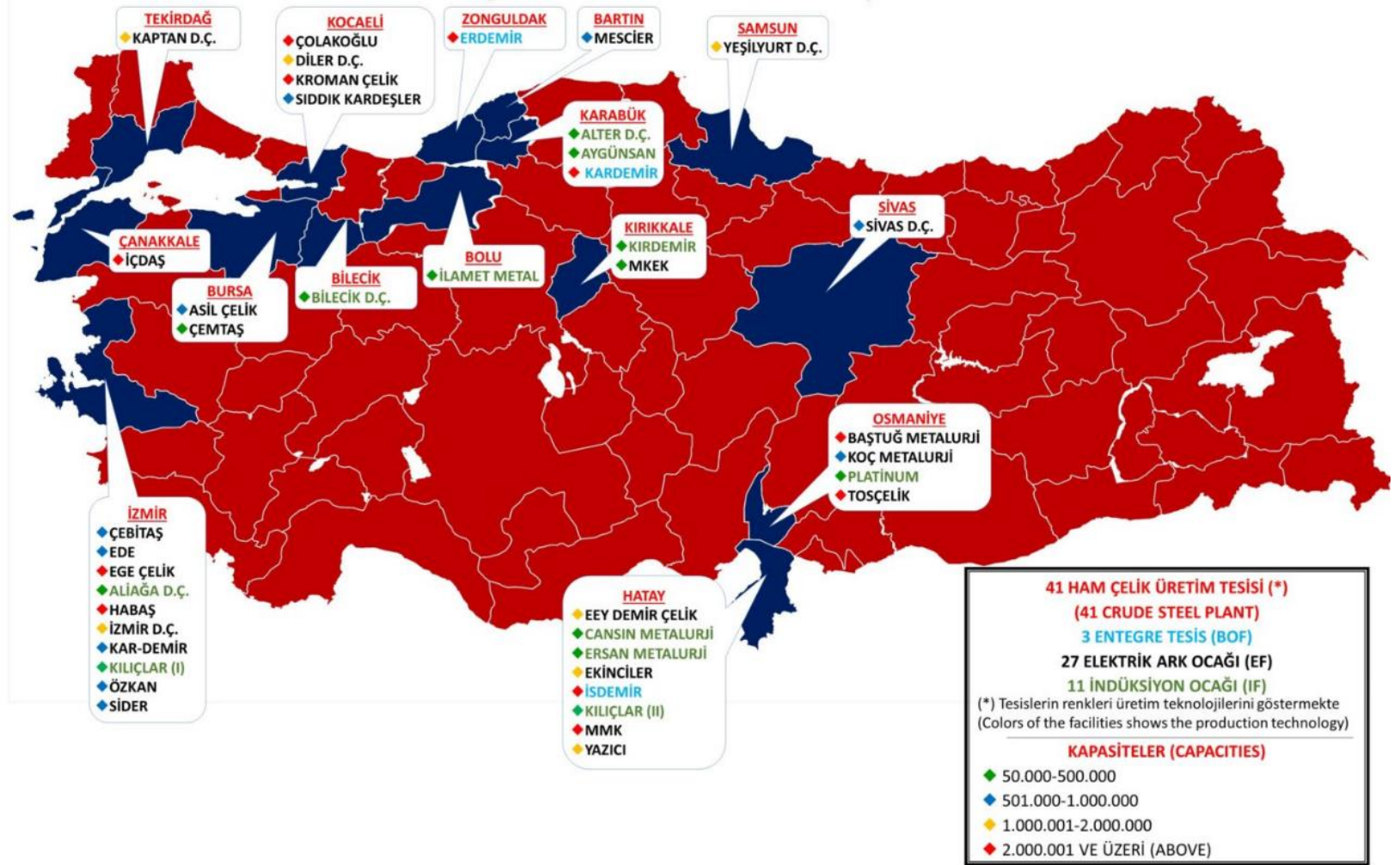
Yüksek kaliteli çeliği dünya çapında prestijli inşaat projelerinde kullanıldığından, Türk demir çelik endüstrisinin etkisi sınırlarının çok ötesine uzanmaktadır. Dikkate değer örnekler arasında Dubai, BAE'deki Burj Khalifa, BAKU-TİBİLİS-CEYHAN (BTC) ham

petrol boru hattı projesi, Hong Kong Uluslararası Havaalanı, Londra, Birleşik Krallık'taki Heathrow Havaalanı ve çok sayıda diğer ikonik yapılar yer almaktadır.

1.1. Türkiye’de Son Dönemde Demir-Çelik İhracatı

2021 yılında demir çelik sektörü, uluslararası ihracatta kayda değer bir artışa tanık olmuş ve bir önceki yıla göre %52'lik etkileyici bir artış kaydetmiştir. Sektörün ihracat değeri 2020'de 330 milyar dolardan 2021'de 502 milyar dolara yükselmiştir. Türkiye'nin genel imalat sanayiine önemli düzeydeki katkısını göstermektedir. Aynı dönemde tüm imalat sanayi ihracatta %31,3'lük bir büyüme yaşarken, tek başına demir çelik sektörü %77'lik olağanüstü bir büyüme oranı sergilemiştir. Bu durum, demir çelik sektörünü %79,5'lik bir büyüme kaydeden kuyumculuk sektörünün hemen ardından en çok büyüyen ikinci sektör haline getirmişti. Türkiye İhracatçılar Meclisi 2022 İhracat Raporu'na göre demir çelik sektörü, 22 milyar 351 milyar dolarlık toplam ihracat değeriyle Türkiye'nin önde gelen ihracat sektörleri arasında üçüncü sırayı almıştır.

Şekil 1. Türkiye Çelik Haritası⁵



⁵ Türkiye Çelik Üreticileri Derneği,

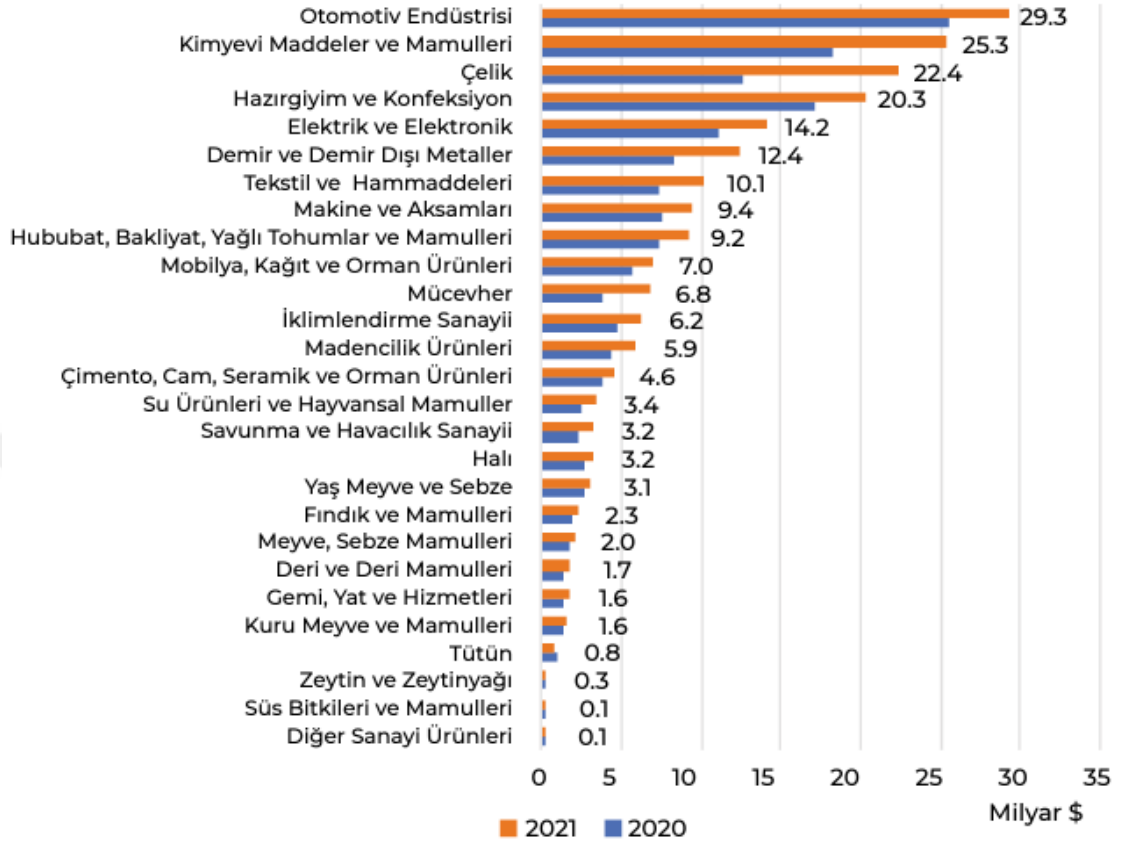
Şekil 2. Türk Demir Çelik Endüstrisi⁶



Türk demir çelik sektörü yıllar içinde dalgalanmalar göstermiştir. Geçtiğimiz yıllarda hem iç talebi hem de uluslararası ticaret dinamiklerini yansıtan önemli bir büyüme ve değişime tanık olmuştur. Üretimde, Türk demir çelik endüstri yıllar içinde istikrarlı bir artış göstererek 40,4 milyon ton ile 2021 yılında en yüksek noktasına ulaştı. Ancak, çelik uluslararası taleplerindeki değişiklik ve diğer farklı nedenlerle üretim seviyelerinde bazı dalgalanmalar olmuştur. Çelik tüketimi de dalgalanmalar yaşadı, ancak genel olarak 33,4 milyon tona ulaştığı 2021 yılına kadar artış eğilimi göstermiştir. Türk çelik ihracatı, ülkenin küresel pazardaki rekabet gücünü yansıtan istikrarlı bir büyüme kaydetti. En yüksek ihracat hacmi 23,9 milyon tona ulaşarak 2021 yılı olmuştur. Çelik ithalatı da iç talep ve uluslararası piyasa koşullarından etkilenerek dalgalanmalar gösterdi. 20,6 milyon ton 2016 yılında ithalatta zirvesi olarak görülmüştür.

⁶ CIB demir çelik ihracatçıları birliği 2023 raporu, Sayfa.8,

Şekil 3. Türkiye'nin Sektörlere Göre İhracatı⁷



2020 ve 2021 yılları arasında Türk ihracat sektörleri, birçok sektörün ihracat değerlerinde önemli artışlara tanık olmasıyla dikkat çekici bir büyüme yaşadı. Etkileyici bir şekilde çelik sektörü, bir önceki yılda 12,6 milyar dolar olan ihracatın 2021'de 22,4 milyar dolara ulaşmasıyla kayda değer bir ilerleme 9,8 milyar dolarlıkla önemli büyüme gösterdi, Türk çelik sektörünün küresel pazardaki dayanıklılığını ve rekabet gücünü yansıtıyor. Çelik sektörü ile bağlantılı olan otomotiv sektörü de etkileyici bir performans sergileyerek 2021 yılında 2020 yılına göre kayda değer bir artışla 29,3 milyar dolarlık ihracat gerçekleştirdi.

⁷ TİM Türkiye ihracat meclisi 2022 raporu, Sayfa.46 ,

Şekil 4 . 2021 Yılında Türkiye'nin Demir Çelik İhracatının Dağılımı⁸



2021 yılında Türkiye'nin demir çelik ihracatının dağılımı, uluslararası ticarete dengeli ve çeşitlendirilmiş bir yaklaşım sergilemektedir. Avrupa Birliği, bu bölgeden gelen güçlü ekonomik bağları ve talebi yansıtan, ana ihracatın %31'ini oluşturan önemli bir ihracat pazarı olmaya devam ediyor. Türkiye'nin Avrupa pazarına verdiği önem açıktır. Türkiye diğer bölgelerdeki varlığını genişletme konusunda önemli adımlar attı. Latin Amerika %12,8 ve Afrika %13,1 ülkenin gelişmekte olan pazarlara artan ilgisini göstermektedir. Ortadoğu ayrıca, komşu ülkelerle ticari ilişkileri geliştirmek için Türkiye'nin coğrafi yakınlığından ve kültürel bağlarından yararlanarak toplam demir çelik ihracatın %17,8'ini temsil eden önemli bir rol oynamaktadır. Ayrıca, Türkiye'nin demir çelik ihracatı, potansiyel büyüme fayda elde etmek amacıyla Asya ve Okyanusya bölgelere uzanmaktadır.

⁸ CIB Demir Çelik İhracatçıları Birliği 2022 Raporu, S:10 ,

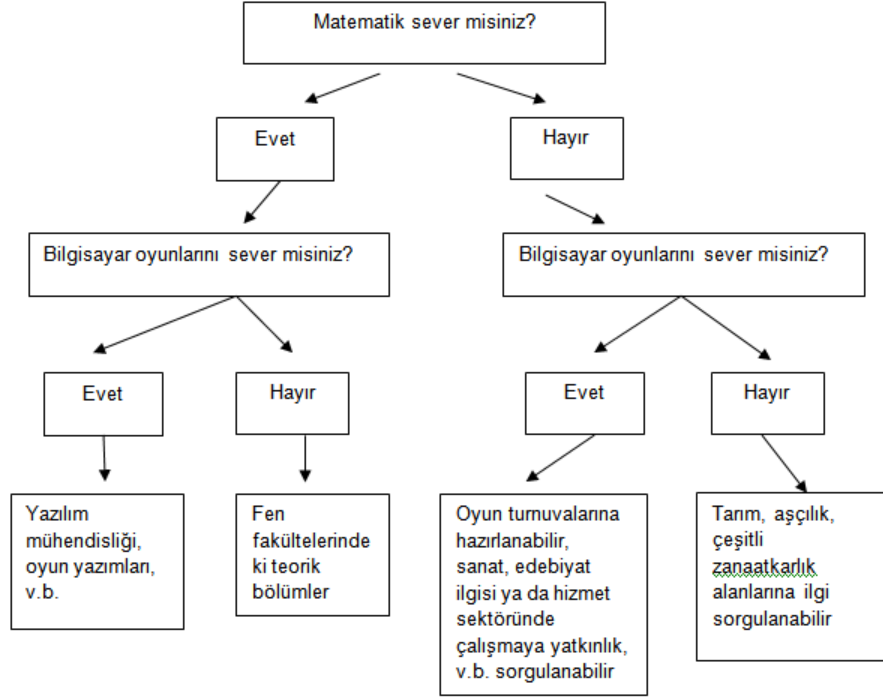
İKİNCİ BÖLÜM

ARAŞTIRMA METODOLOJİSİ

2. 1. Makine Öğrenmesi

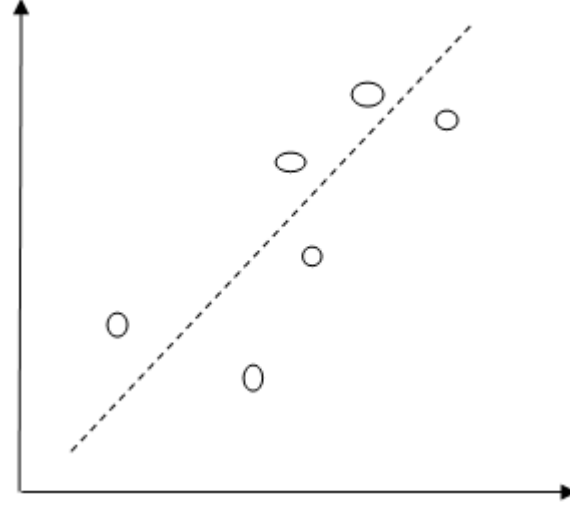
Makine öğrenmesi temelde sınıflandırma ve tahmin yapmayı sağlayan yöntemlerdir. Bu yöntemlere her yıl onlarca yenisi eklenmekte ve sürekli gelişim göstermektedir. Makine öğrenmesinde sınıflandırma işlemi genellikle karar ağaçları üzerinden yapılmaktadır. Karar ağaçları insan beyninin düşünme sistemine çok yakın olması, karar alma süreçlerinde insan beyninin izlediğine yola benzer bir yol izlemesi ve varsayım gerektirmemesi nedeniyle parametrik yöntemlere kıyasla daha pratik çözümler üretebilmektedir. Karar ağacının oluşturulması ve sınıflandırma özelliği bir örnek üzerinden açıklanmaya çalışılacaktır. Şekilde görülen karar ağacı, “Matematik sever misiniz” sorusuyla başlamaktadır. Bu kısım karar ağacının kök düğümüdür. Kök düğüm (root node), sınıflandırmanın başlangıcıdır ve sınıflandırmanın en iyi şekilde yapılabilmesi için en ayrıştırıcı özelliği temsil etmelidir. Diğer basamaklarda verilen cevaba bağlı olarak tıpkı bir ağaç gibi dallanıp budaklanmaktadır. Şekilde “evet” ve “hayır” cevaplarıyla basit sorular üzerinden karar ağaçları yoluyla sınıflandırma yapılarak gençlerin üniversitede alan seçimleri yapmalarına bir örnek verilmiştir. Alan seçiminde matematik bilimine olan ilgi ve bilgisayar oyunlarına yatkınlık temel bir sınıflayıcı gibi düşünülmüştür. Elbette bu örneğin bilimsel ya da sosyolojik bir yanı yoktur, sadece, basit sorular üzerinden karar ağaçları kullanıldığında sınıflandırmanın kolaylığını temsil etmek amacıyla tasarlanmıştır.

Şekil 5. Karar Ağacı Örneği



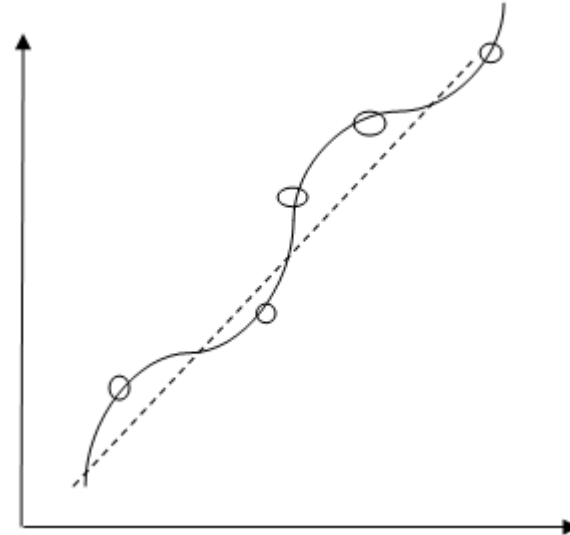
Makine öğrenmesinin tahmin özeliği ise klasik regresyon analizine benzer niteliktedir. Örneğin 100 tane 15-18 yaş arası erkeğin günlük aldıkları kalsiyum miktarı ile boy uzunlukları arasındaki ilişki modellendiğinde, bu 100 erkeğin arasında olmayan bir gencin günlük aldığı kalsiyum miktarını öğrenerek boy uzunluğu tahmin edilebilir. Makine öğrenmesinde orijinal veri kümesi, eğitim verisi (training data) olarak adlandırılır.

Şekil 6. Eğitim Verisinin Doğrusal Model İle Tahmin Edilmesinin Gösterimi



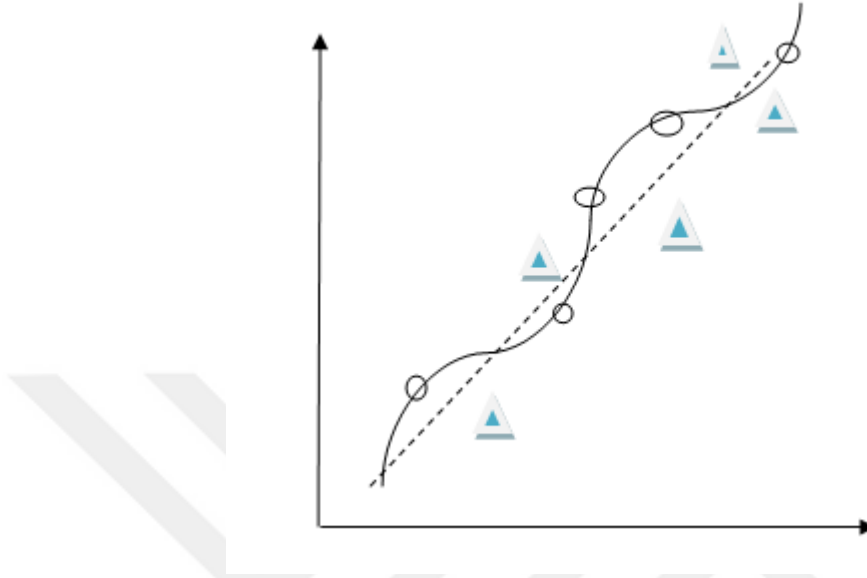
Gözlemleri ortalayan kesikli çizgili doğru parçası, eğitim verisini modellemektedir. Fakat bu gözlemleri daha iyi modelleyebilecek eğrisel bir model de olabilir, bkz. Şekil.7.

Şekil 7. Orijinal Verinin Eğrisel Modelle Tahmin Edilmesinin Gösterimi



Eğrisel modelin orijinal veriye daha uygun olduğu görülmektedir. Fakat makine öğrenmesinin amacı tahmin yapmaktır ve iki modelden hangisinin daha uygun olduğunun belirlenmesi gerekmektedir. Bunun anlaşılabilmesi için test verisine ihtiyaç vardır. Şekil.8’de üçgenlerle gösterilen veriler, test verisidir.

Şekil 8. Test Verisinin Gösterimi



Test verisi doğrusal ve eğrisel modelle yapılan tahminlerin kıyaslanmasını sağlar. Eğrisel model orijinal veriye daha uygun tahminler yapabiliyor gibi görünse de, bu her zaman bu şekilde gerçekleşmeyebilir, doğrusal model ile yapılan tahminler orijinal veriye daha yakın bulunabilir. O halde makine öğrenmesi yöntemlerinden en iyisi, test verisine en fazla uyum gösterendir denebilir.⁹

Makine öğrenmesinde test verisinin belirlenmesi tamamen rastgele yapılabildiği gibi, çeşitli yöntemlerle de gerçekleştirilmektedir. Çapraz doğrulama bunlarda biridir:

2.1. 1. Çapraz Doğrulama Yöntemi

Çapraz doğrulama yöntemi, makine öğrenme yöntemlerinden hangi yöntemin daha uygun olacağının belirlenmesinde kullanılmaktadır. Verinin test etmek üzere ayrılan kısmının, tahmin modelinin en yüksek doğrulukla çalışıp çalışmadığını kontrol etmektedir. verinin hangi kısmının test için ayrılacağı keyfi bir durumdur, eğitim verisi yan, orijinal veri dört parçaya bölünerek bu işlem gerçekleştirilirse dört katmanlı

⁹ Karar ağacı oluşturma sürecinde tüm gözlemlerin kapsandığı veri kümesine eğitim verisi, bir kısmının alındığı veri kümesine ise test verisi denmektedir.

çapraz doğrulama (four fold cross validation), on parçaya bölünerek yapılırsa on katmanlı (ten fold cross validation) yapılmış olur, v.b.

Şekil 9. Çapraz Doğrulama Yönteminde Test Verisinin Belirlenmesi

I. Durum: Veri dört parçaya bölünmüş, son parça test verisi olmuştur.



II. Durum: Veri dört parçaya bölünmüş, üçüncü parça test verisi olmuştur.



III. Durum: Veri dört parçaya bölünmüş, ikinci parça test verisi olmuştur.



IV. Durum: Veri dört parçaya bölünmüş, ilk parça test verisi olmuştur.



Şekil.9’da görülen blokların hangi kısmının test verisi olarak seçilmesi hususu, çapraz doğrulama yöntemi tarafından giderilmektedir. Çünkü çapraz doğrulama yöntemi tüm sıralamaları eşanlı olarak almakta ve test etmektedir. Böylece test verisine en iyi uyum sağlayan yöntemi tespit etmiş olmaktadır.

2.1.2. Karışıklık Matrisi Yöntemi (Confusion Matrix Method)

Karışıklık matrisi, sınıflandırıcının zayıf ve güçlü performans gösterdiği alanları belirlemeye yardımcı olmaktadır.¹⁰ Karışıklık matrisi karesel bir matristir ve boyutu, tahmin edilmek istenen durumların sayısı kadardır. Bunu bir örnek üzerinden açıklayabiliriz. Tablo.1’de 100 hastaya ait kurgusal veriler görülmektedir:

¹⁰ Sönmezler, G., Gündüz,İ. O., (2021). Covid-19 Pandemi Sürecinin BİST-30 Hisse Senetlerine Etkilerinin Karışıklık Matrisi ile Analizi, *Maliye ve Finans Yazıları* (115), s:58.

Tablo 1. Karışıklık Matrisinin Anlatımı İçin Oluşturulmuş Kurgusal Veriler

Yaş	Boy Uzunluğu	Kilo	Bel Çevresi	Obezite Durumu
59	172	95	103	Evet
65	180	70	78	Hayır
75	167	98	115	Evet
70	165	60	80	Hayır
.....				
.....				

Yaş, boy uzunluğu, kilo ve bel çevresi verilerine dayanarak, bir insanın obeziteye sahip olup olmayacağını makine öğrenmesiyle tahmin edilmek istensin. Veri kümesinin bir kısmı eğitim verisi, diğer kısmı test verisi olarak ayrılmalıdır. Daha sonra lojistik regresyon, random forest, v.b. yöntemleri hem eğitim verisine hem de test verisine uygulanır. Sonuçlar karışıklık matrisinde gösterilir:

Tablo 2. Karışıklık Matrisi

		Gerçek	
Tahmin		Obez	Obez Değil
	Obez	True +	False +
	Obez Değil	False -	True -

Karışıklık matrisinin satırları makine öğrenmesi algoritmasının tahminleridir. Sütunları ise gerçek değerlerdir.

“True +”, obeziteye sahip olanların algoritma tarafından doğru sınıflandırıldığı anlamına gelmektedir.

“True -“, obez olmayanların algoritma tarafından doğru sınıflandırıldığı anlamına gelmektedir.

“False –“, gerçekte obez olanları, algoritma tarafından yanlış sınıflandırıldığı anlamına gelmektedir.

“False +“, gerçekte obez olmayanların algoritma tarafından yanlış sınıflandırıldığı anlamına gelmektedir.

Uygulanan çeşitli makine öğrenmesi algoritmalarının karışıklık matrislerine bakılarak, hangi yöntemde “True +” ve “True –“ sayıları fazla ise, o yöntemin daha uygun olduğu söylenir. Çünkü bu yöntem diğerlerine göre daha doğru bir sınıflandırma gerçekleştirmiş demektir.

2.1.3. Random Forest Tekniği

Orijinal veri kümesinden tekrarlı örnekleme izin vererek elde edilen ve bootstrapped veri olarak adlandırılan veri kümeleri kullanılarak, sınıflandırmayı sağlamak amacıyla her defasında farklı değişken grupları göz önüne alınarak oluşturulan karar ağaçlarının meydana getirdiği yapıya, Random Forest denir.

Bootstrapped veriler oluşturulurken tekrarlı örnekleme izin verildiği için, bazı veriler kapsam dışı kalabilirler. Bu gözlemler, OOB (out of bag data) olarak adlandırılır. OOB ler kapsam dışı kaldıkları için karar ağaçları oluşumunda yer almazlar. OOB ler hatanın tahmininde kullanılırlar.¹¹ OOB’ler karar ağaçlarında işleme sokulduklarında ortaya çıkan yanlış sınıflandırılma oranına, OOB hatası denir. Genellikle verinin üçte ikisi karar ağaçlarını oluşturmak, üçte biri de OOB hatasını hesaplamak için kullanılır.¹² Her defasında farklı sayıda değişkenlerle çok sayıda karar ağacı elde edilir. Karar ağacı oluşum sürecinde N adet değişkenden seçilecek n tane değişken için başlangıç değeri olarak;

$$\sqrt{N} \quad [1]$$

ya da;

¹¹ Azar, A. T., Elshazly, H. I., Hassanien, A. E., Elkorany, A. M. (2014). A random forest classifier for lymph diseases, *Computer Methods and Programs in Biomedicine*, Volume 113, Issue 2, p:467.

¹² Dogru, N., Subasi, A. (2018). "Traffic accident detection using random forest classifier," *2018 15th Learning and Technology Conference (L&T)*, Jeddah, Saudi Arabia, 2018, p:43.

$n \ll N \log 2(N) + 1$ bağıntısı göz önünde bulundurularak minimum OOB hatası elde edene kadar denenir.¹³

Hem sınıflandırma hem de regresyon problemlerinde, dahili bir düğümü bölmek için seçilen tahminci değişkenlerin alt kümesi önceden belirlenmiş ayırma kriterlerine bağlıdır. Söz konusu kriterler optimizasyon problemi olarak formüle edilmiştir. Bölmede yaygın olarak kullanılan bir yöntem, Shannon'un entropisidir.¹⁴

$$E = -\sum_{i=1}^c p_i \log(p_i) \quad [2]$$

Bu değer karar ağacının her bölünmesinde en fazla bilgi elde etmek üzere maksimize edilmektedir.¹⁵

Sınıflandırıcı algoritmalarının performanslarını değerlendirirken, karışıklık matrisinin doğruluk oranına bakılabilir. Doğruluk oranı:¹⁶

$$\text{Doğruluk} = \frac{\text{True}_+ + \text{True}_-}{\text{True}_+ + \text{True}_- + \text{False}_+ + \text{False}_-} \quad [3]$$

Hata oranı, kesinlik ve duyarlılık:¹⁷

$$\text{HataOranı} = \frac{\text{False}_+ + \text{False}_-}{\text{True}_+ + \text{True}_- + \text{False}_+ + \text{False}_-} \quad [4]$$

$$\text{Kesinlik} = \frac{\text{True}_+}{\text{True}_+ + \text{False}_+}$$

$$\text{Duyarlılık} = \frac{\text{True}_+}{\text{True}_+ + \text{False}_-}$$

¹³ A.g.e., p:43.

¹⁴ Schonlau, M., Zou, R. Y. (2020). The random forest algorithm for statistical learning, *The Stata Journal* 20, Number 1, p.5.

¹⁵ A.g.e., p:5

¹⁶ Aydın, C. (2018). Makine Öğrenmesi Algoritmaları Kullanılarak İtfaiye İstasyonu İhtiyacının Sınıflandırılması, *Avrupa Bilim ve Teknoloji Dergisi* Sayı 14, S.173.

¹⁷ Bilgin, M. (2017). Gerçek Veri Setlerinde Klasik Makine Öğrenmesi Yöntemlerinin Performans Analizi, *Akademik Bilişim*, Aksaray, Turkey, s:4.

Kesinlik ve duyarlılık ölçütleri tek başına anlamlı bir karşılaştırma sağlamayabilir, bu durumda F ölçütü kullanılabilir:¹⁸

$$F = \frac{2DK}{D + K} \quad [5]$$

F ölçütü, [0,1] aralığında değer alır, 1'e yakın bir değer alması sınıflandırmanın başarısını gösterir.¹⁹

2.1.4. Makine Öğrenmesinde Yanlılık ve Varyans

Yanlılık istatistiksel olarak gerçek değer ile tahmin edilen değer arasındaki fark olarak tanımlanır ve gözlemlerin doğru temsil edildiği modellerde yanlılık düşük bulunur. Şayet model uygun bir model değilse yanlılık artacaktır. Makine öğrenmesi bakımından ise yanlılığın artması, algoritmanın daha az öğrenmesi demektir. Yanlılığı fazla olan bir algoritma, yanlış sonuçlar üretecektir. Bunların birlikte sadece yanlılığa odaklanmak yanlış olur. Algoritmanın öğrenme düzeyi varyansa da bağlıdır. Sistemin okuduğu veri arttıkça varyans da artma eğilimine girecektir. Bu durum algoritmanın daha iyi öğrenmesi gibi görünse de, varyansın büyümesi aşırı öğrenme sorununa yol açabilir. *“Model yüksek varyansa ve yüksek önyargıya sahip olmamalıdır. Yüksek tahmin başarısı her zaman veri setinde iyi sonuçlar vermez. Önemli olan veri setinin algoritmasının doğru oluşturulması ve genelleştirilebilir modelin yakalanmasıdır.”*²⁰

2.1.5. ROC ve AUC Grafikleri

Lojistik regresyon analizi makine öğrenmesinde sıklıkla kullanılan bir sınıflandırma yöntemidir. Lojistik regresyonda y-ekseni olasılıklara karşılık gelir. ROC ve AUC grafikleri, bir önceki bölümde bahsi geçen 100 hastaya ait kurgusal veri üzerinden açıklanmaya çalışılacaktır. Hastaların bel çevresi ile obez olup olmama durumları lojistik regresyon analizi ile modellenmiş olsun. Her bir hastanın bel çevresine karşılık gelen bir olasılık olacaktır. Fakat bu olasılıkların hangisi obez

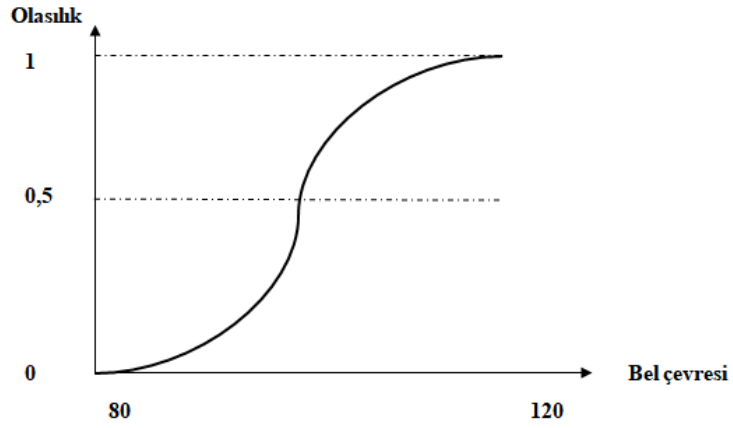
¹⁸ Nizam, H., Akın, S.S. (2014). Sosyal Medyada Makine Öğrenmesi ile Duygu Analizinde Dengeli ve Dengesiz Veri Setlerinin Performanslarının Karşılaştırılması, *XIX. Türkiye'de İnternet Konferansı, inet-tr.org.tr*, s:4.

¹⁹ Onan, A. (2014). Şirket İflaslarının Tahmin Edilmesinde Karar Ağacı Algoritmalarının Karşılaştırmalı Başarım Analizi, *Bilişim Teknolojileri Dergisi*, CİLT: 8, SAYI: 1, s:14.

²⁰ Nacar, E. N., Erdebilli, B. (2021). Makine Öğrenmesi Algoritmaları İle Satış Tahmini, *Journal Of Industrial Engineering* 32(2), s:315.

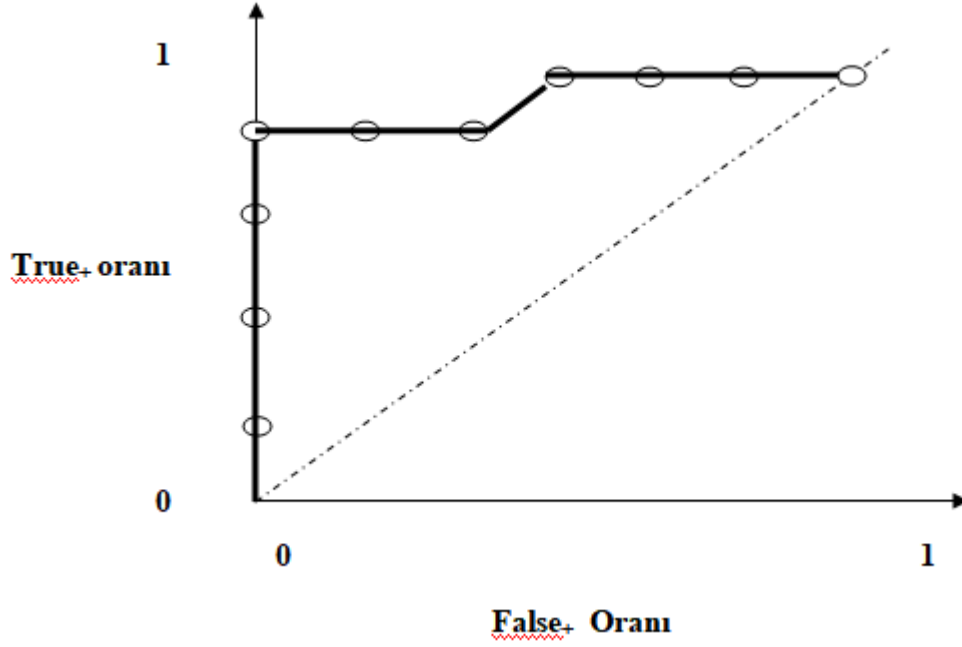
olmak için bir eşik teşkil edecektir? Aşağıda şekilde bu eşik 0.5 olarak gösterilmektedir. Yani bir hastanın bel çevresine karşılık gelen olasılık 0.5'ten büyük veya eşitse bu hasta obez olarak sınıflandırılacaktır, eğer 0.5'ten küçükse obez değildir şeklinde sınıflandırılacaktır.

Şekil 10. Lojistik Regresyon Eğrisinin 0.5'lik Eşik Değer Gösterimi



Farklı eşik değerleri için oluşturulan sınıflandırmaların her biri için bir karışıklık matrisi oluşacaktır. Her eşik düzeyinde doğru ve yanlış sınıflandırmalar olacaktır. Obez olduğu halde obez değil şeklinde yanlış sınıflandırmalar olabilir. Her bir karışıklık matrisinin True+ ve False+ oranları kullanılarak ROC grafiği elde edilir.

Şekil 11. ROC Grafiği²¹



True₊, karışıklık matrisinde obez olup obez olarak sınıflandırılanların sayısıdır. False₋, karışıklık matrisinde obez olmadığı halde obez olarak sınıflandırılanların sayısıdır. True₊ oranı;²²

$$True_{+}Oranı = \frac{True_{+}}{True_{+} + False_{-}} \quad [6]$$

False₊, karışıklık matrisinde obez olduğu halde obez değil şeklinde sınıflandırılanların sayısıdır. True₋, karışıklık matrisinde obez olmayanların obez değildir şeklinde sınıflandırılanların sayısıdır.²³

$$False_{+}Oranı = \frac{False_{+}}{False_{+} + True_{-}} \quad [7]$$

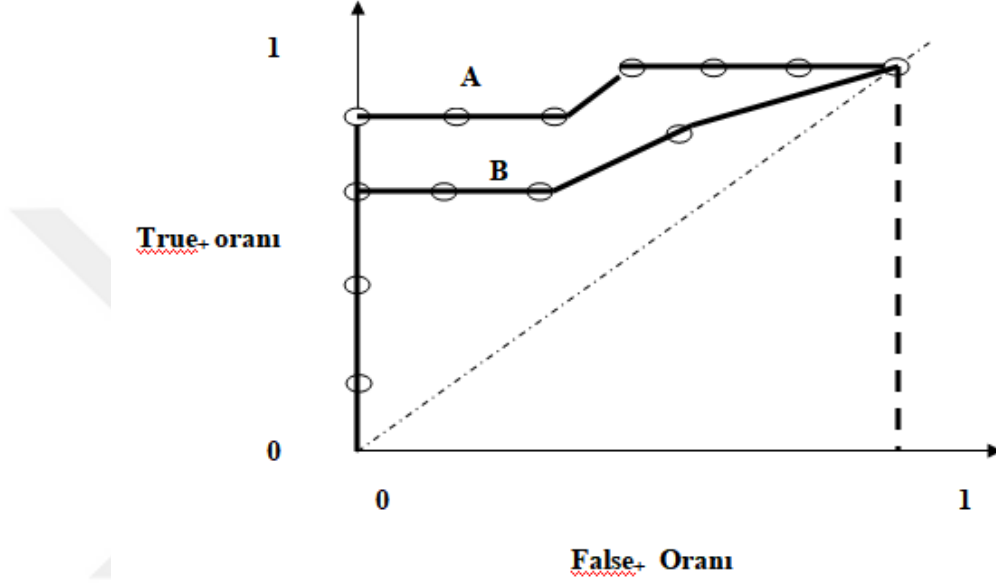
²¹ <https://statquest.gumroad.com/l/wvtmc>

²² A.g.e.

²³ A.g.e.

AUC (area under curve) grafiği, ROC grafiği altında kalan alandır. Alanı büyük olan algoritmanın sınıflandırma başarısı daha yüksektir. Şekilde A algoritması, B algoritmasına göre daha iyi bir sınıflandırma gerçekleştirmiştir.

Şekil 12. AUC Grafiği²⁴



2.2. Entropi

Entropi, veri biliminde sınıflandırma, benzerlik ve farklılıkların ölçülmesinde, boyut indirgeme v.b. konularda kullanılan bir yöntemdir.

Entropi sözcüğü, Oxford sözlüğünde tanımlandığı şekliyle, “bir sistemde var olan düzen eksikliğini ölçmenin bir yolu” dur. Bu kelime, değişmek veya dönüşmek anlamına gelen Yunanca “entropia” kökünden gelmektedir. Daha sonra Alman fizikçi Rudolph Clausius tarafından kullanılmıştır.

İstatistikte Entropi, 1948'de bilgi teorisi ve ergodik teori²⁵ üzerinde önemli bir etkisi olan Claude Shannon tarafından başlatılmıştır. Entropi, rastgele süreçlerin uzun

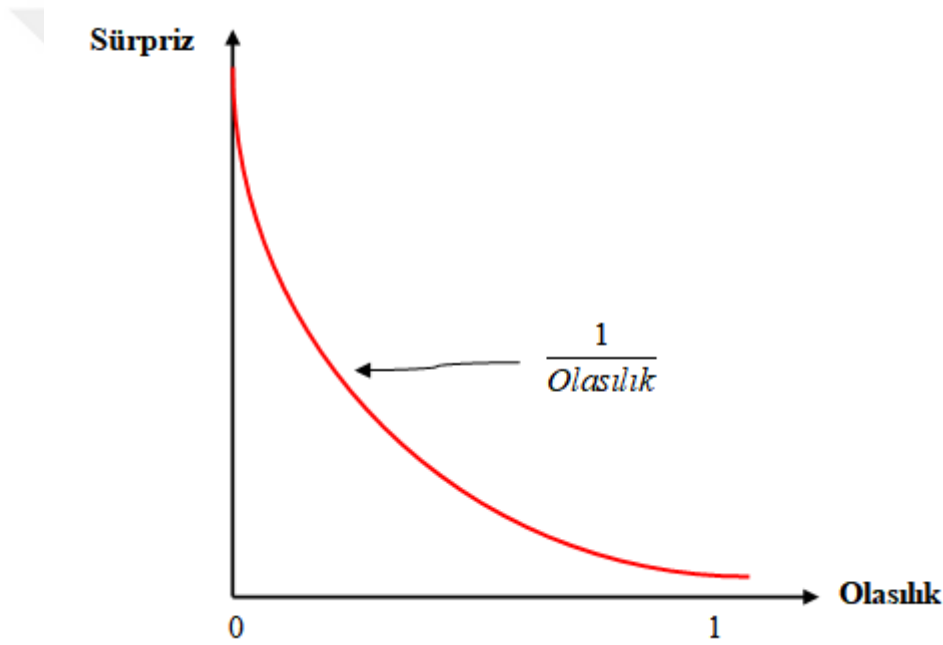
²⁴ <https://statquest.gumroad.com/l/wvtmc>

²⁵ Ergodik teori, matematik, fizik, istatistik ve diğer alanlarda kullanılan bir kavramdır. Temel olarak, bir dinamik sistemdeki uzun vadeli özelliklerin, sistemin farklı noktalarındaki zaman serileri veya uzayda farklı yerlerdeki zaman serileri boyunca karşılaştırılması ve analizini içeren bir yaklaşımı ifade eder. Ergodik teori, bir sistemin zaman içindeki davranışını inceler ve sıcaklık, basınç, moleküler hareketler gibi istatistiksel özelliklerle ilgilidir. Bir sistemdeki genel özellikleri ve dinamiği öğrenmek için kullanılır. Özellikle karmaşık sistemlerin incelenmesinde önemli bir araçtır. Ergodik teorisi, bir

vadeli davranışını anlamada çok önemli bir rol oynar ve bilgi teorisinin kodlama teoremleri için gereklidir. Shannon'ın entropi kavramı, fiziksel bir sistemin rastgelelik düzensizliğini ölçen fizikten termodinamik entropi fikrinden esinlenmiştir. Ancak Shannon, bu kavramı, entropinin bilginin belirsizliğini veya öngörülemezliğini temsil ettiği bilgi teorisi alanına uygulamıştır.

Bilgi teorisinde entropi, bir olay veya mesajda yer alan bilgi veya "sürpriz" miktarını ölçer.

Şekil 13. Entropide Sürpriz ve Olasılık Arasındaki Bağntı



Şekil 13'te görüldüğü üzere, olasılık sıfıra yaklaştıkça, sürpriz miktarı artar. Tersine, olasılık 1'e yaklaştıkça, sürpriz miktarı sıfıra yaklaşır. Aralarında ters yönlü bir ilişki vardır.

Belirli bir bilgi kaynağındaki farklı sonuçların veya sembollerin olasılıklarına göre hesaplanır. Entropi ne kadar yüksek olursa, bilginin belirsizliği veya değişkenliği de o kadar büyük olur

sistemi belirli bir zaman dilimi boyunca gözlemleyerek, o sistemin uzun vadeli davranışını çözmeye çalışır. Temel olarak, bir sistemin zaman içindeki hareketi sırasında gösterdiği tüm olası davranışları tanımlamaya çalışır. Bu, istatistik ve olasılık teorilerinde kullanılır. (<https://www.nedemek.page/kavramlar/ergodic%20teorisi>)

$$H_j = -k \sum_{i=1}^n p(x_{ij}) \log(p(x_{ij})) \quad (j = 1, 2, \dots, m) \quad 26 \quad [8]$$

n : değişken sayısı

H_j : x_j değişkeninin entropisidir

$P(x_{ij})$: x_{ij} 'in olasılığı

k : pozitif bir değer , $k = \frac{1}{\log(m)}$

Shannon'ın entropi üzerine çalışması, iletişim teorisi, veri sıkıştırma, kriptografi ve istatistiksel çıkarım dahil olmak üzere çeşitli alanlar üzerinde derin bir etkiye sahip olmuştur. Kantitatif bir bilgi ölçüsü sağlanırsa ve bilgi teorisi ve ilgili disiplinlerde daha fazla gelişme için zemin hazırlar.

Shannon'ın orijinal formülasyonundan bu yana, belirli senaryoları ve uygulamaları ele almak için koşullu entropi, ortak entropi ve diferansiyel entropi gibi entropinin çeşitli varyasyonları ve uzantıları tanıtıldı. Bu varyasyonlar, bilginin kritik bir rol oynadığı çeşitli alanlarda mantıksal entropinin anlaşılmasını ve uygulanmasını daha da zenginleştirir.

Entropinin nasıl hesaplandığı konusunun daha iyi anlaşılması için şu basit örnek üzerinden gidilebilir:

Homojen olmayan (hileli) bir para olsun. Bu paranın “yazı” gelme olasılığı 0.3, “tura” gelme olasılığı ise 0.7 olsun. İki çıktısı olan bir sistemde (yazı ve tura), sürpriz:

$$\text{sürpriz} = \log_2\left(\frac{1}{\text{Olasılık}}\right) \text{ ile hesaplanır.}^{27} \quad [9]$$

Bu durumda, yazı gelmesindeki sürpriz:

$$\text{sürpriz}_{\text{yazı}} = \log_2\left(\frac{1}{0.3}\right) = 1.736966 \text{ 'dir.}$$

²⁶ Shannon, C. E. (1948). A mathematical theory of communication, Harvard mathematics department, Reprinted with corrections from The Bell System Technical Journal, Vol. 27, pp. 379–423, 623–656, July, October 1948, Sayfa 11,

²⁷ Üç çıktılı bir deney olsaydı, $\text{sürpriz} = \log_3\left(\frac{1}{\text{Olasılık}}\right)$ ile hesaplanırdı.

Tura gelmesindeki sürpriz:

$$sürpriz_{tura} = \log_2\left(\frac{1}{0.7}\right) = 0.514573 \text{ 'dir.}$$

Görüldüğü üzere, daha olası sonuçların sürprizi, görece daha az olası olan sonuçların sürpriz miktarından düşüktür. Hileli paranın tura gelmesi yazı gelmesinden daha yüksek bir olasılık olduğu için, para atıldığında tura gelmesi çok da şaşırtıcı olmayacaktır.

Söz konusu parayı 5 defa attığımızı düşünelim ve 4 defa yazı 1 defa tura gelmiş olsun. Bu sıralamanın gerçekleşme olasılığı;

$$YYYYT = 0.3*0.3*0.3*0.3*0.7 = 0.00567 \text{ 'dir.}$$

Dört yazı bir tura gelmesindeki sürprizi hesaplamak istersek,

$$sürpriz = \log_2\left(\frac{1}{0.3*0.3*0.3*0.3*0.7}\right) = 7.462436$$

elde ederiz. Bu değer, her bir atışın tekil sürprizlerinin toplamıdır;

$$(sürpriz_{yazı} = 1.736966) + (sürpriz_{yazı} = 1.736966) + (sürpriz_{yazı} = 1.736966) + (sürpriz_{yazı} = 1.736966) + (sürpriz_{tura} = 0.514573) = 7.462437.$$

Yukarıda bahsedilenlerden entropiye geçerseniz, basitçe şu örneği verebiliriz: Hileli parayı 10 defa attığımızı düşünelim, bu durumda her atış için toplam sürpriz miktarı;

$$\frac{0.3*10*1.736966 + 0.7*10*0.514573}{10} = 0.881291 \text{ elde edilir.}$$

Entropi; paranın her atılışında beklenen sürpriz miktarıdır;

$$Entropi = E(sürpriz)$$

$$Entropi = E(sürpriz) = \frac{0.3*10*1.736966 + 0.7*10*0.514573}{10} = 0.881291$$

'dir.

2.2.1. Karar Vermede Entropi

Karar vermede entropinin kullanımı, bir karar problemindeki çeşitli faktörler veya niteliklerle ilişkili bilgi içeriğinin çeşitliliğini, belirsizliğini ölçmek için bir ölçü olarak ifade edilir.

Entropi, bilgi teorisinden türetildiği için, bir sistemdeki rastgelelik veya düzensizlik miktarının sayısal bir temsilini sağlar.

Entropi çeşitli şekillerde karar verme sürecini desteklemek için kullanılır:

Nitelik seçimi: Entropi, karar verme problemleri için en bilgilendirici veya önemli nitelikleri belirlemeye yardımcı olur. Karar vericiler, her özelliğin entropisini hesaplayarak, özellik içindeki değerlerin değişkenliğini veya dağılımını değerlendirebilir. Daha yüksek entropi değerlerine sahip nitelikler, daha büyük bir değer çeşitliliğine işaret eder ve genellikle karar verme için daha bilgilendirici olarak kabul edilir.

Ağırlık belirleme: Entropi, bilgi içeriği veya ayırmacı güç üzerindeki niteliklere ağırlık atamak için kullanılabilir. Bu, niteliklere daha nesnel ve veri yönelimli bir şekilde önem verilmesine yardımcı olur.

Kriter değerlendirmesi: Entropi, farklı kriterlerin performansını veya kalitesini değerlendirmede yardımcı olabilir. Karar vericiler, kriterlerin entropisini hesaplayarak, her bir kriter içindeki puanların veya değerlerin çeşitliliğini veya değişkenliğini değerlendirebilir. Daha yüksek entropi değerleri, puanların daha fazla dağıldığını gösterir ve kriterin önemi veya uygunluğu konusunda karar vericiler arasında fikir birliği veya anlaşma olmadığını gösterebilir.

Duyarlılık Analizi: Entropi, özellik değerlerindeki değişikliklerin genel karar üzerindeki etkisini değerlendirmek için duyarlılık analizinde kullanılabilir. Karar vericiler, belirli değişikliklerden önce ve sonra niteliklerin entropisini ölçerek, bu değişikliklerin genel sistemin belirsizliği veya çeşitliliği üzerindeki etkisini anlayabilirler.

2.2.2. Entropi Yönteminin Adımları

1. Bir karar matrisi oluşturma: karara dâhil olan alternatifi ve nitelikleri temsil eden bir matris oluşturur. Her satır bir alternatife karşılık gelir ve her sütun bir özelliği temsil eder. Matrise, her bir alternatif ve özellikteki performansı yansıtan veriler sağlar.

2. Karar matrisini normalleştirme: Nitelikler arasında karşılaştırılabilirliği sağlamak için. Bu, verilerin ortak ölçeğe dönüştürülmesini veya verilerin standartlaştırılmasını içerir.

3. Her özellik için entropiyi hesaplama: Her bir özelliğin belirsizlik veya rastgelelik derecesini ölçmek için. Daha yüksek entropiye sahip nitelikler, daha çeşitli veya tahmin edilemez değerlere sahiptir.

4. Her özelliğin ağırlığını hesaplama: Entropi değerine bağlı olarak, ağırlık entropi ile ters orantılıdır, yani daha düşük entropi değerlerine sahip nitelikler daha yüksek ağırlık alır. Bu, daha az belirsiz veya daha bilgilendirici niteliklere daha fazla önem verilmesi gerektiği fikrini yansıtır.

2.2.3. Huffman Algoritması

Kayıpsız veri sıkıştırma için yaygın olarak kullanılan bir kodlama algoritmasıdır. Adını, tekniği 1952'de tanıtan David A. Huffman'dan almıştır. Huffman kodlamasının arkasındaki ana fikir, belirli bir veri setinde daha sık görülen sembollere daha kısa kodlar atamak ve böylece kodlanmış verilerin genel boyutunu azaltmaktır.

Shannon entropi, optimal bir kodlama şemasında her bir sembolü temsil etmek için gereken minimum ikili soru biti sayısının bir tahminini sağlar.

Huffman Algoritmasına Bir Örnek

A, B, C ve D mesaj türleri ve olasılıkları $p(A) = 0.4$, $p(B) = 0.3$, $p(C) = 0.2$ ve $p(D) = 0.1$ mesaj türlerini sınıflandırmak için gereken minimum ikili soru sayısı nedir²⁸

İlk sorunun olasılığı $p(q_1) = 0.4$

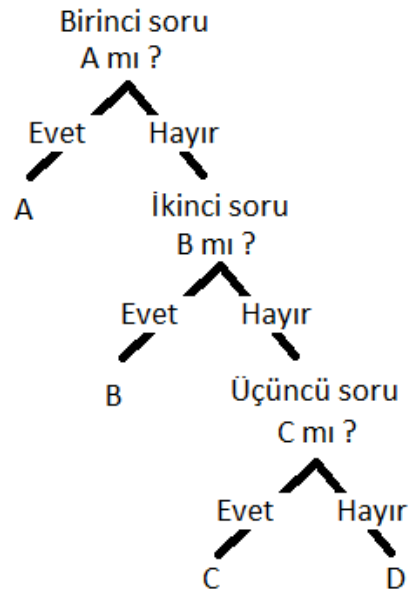
²⁸ Golomb, S. W., Peile, R. E., Scholtz, R. A., (1994) Basic Concepts in Information Theory and Coding , *The Adventures of Secret Agent* 00111 p:9.

İkinci sorunun olasılığı $p(q_2) = 0.3$

Üçüncü sorunun olasılığı $p(q_3) = 0.2$

Buradaki durumu aşağıdaki şekil 2.4 ağaç diyagramı ile açıklandı

Şekil 14. Karar Ağacı



Mesajı sınıflandırmak için gereken ortalama beklenti soru sayısı şu şekilde olacaktır:

$$E(X) = 1 P(0.4) + 2P(0.3) + 3P(0.2) = 0.4 + 0.6 + 0.6 = 1.6$$

Ve Shannon entropi değeri şu şekilde olacaktır:

$$\begin{aligned} H &= 0.4 \log (1/0.4) + 0.3 \log(1/0.3) + 0.2 \log(1/0.2) \\ &= 0.528771 + 0.52109 + 1.514247 = 1.514247 \approx 1.6 \end{aligned}$$

2.3. Veri Zarflama Analizi

Veri Zarflama Analizi (VZA) ilk kez 1978 yılında Charnes, Cooper ve Rhodes tarafından ortaya atıldığından günümüze kadar, farklı alanlardaki araştırmacılara performans değerlendirmek için pratik bir yaklaşım olarak kabul görmektedir. Bunun en önemli sebeplerinden biri, VZA'nın parametrik olmayan bir yöntem olması ve bu nedenle de çok fazla varsayım sorgulamamasıdır. VZA'yı öne çıkaran bir diğer husus da, karar verme birimlerinin etkinliğini ölçmekle kalmayıp, etkinsiz karar birimleri için etkin olabilmeleri adına her bir karar birimine ayrı ayrı önerilerde bulunmasıdır.

VZA bir doğrusal programlama tekniği olduğundan, girdi ve çıktı sayısını sınırlama dan çok sayıda değişkeni ve ilişkiyi işleyebilir. Fakat elbette VZA'nın etkin olan karar birimlerini etkinsiz olanlardan doğru bir şekilde ayrıştırabilmesi için, karar birimi sayısının, toplam girdi çıktı sayısından en az bir fazla, Golany ve Roll'a göre hatta en az iki katı²⁹, Banker, Cooper, Friedman ve Sinuany-Stern'e göre en az üç katı olmalıdır.³⁰

VZA, girdiler ile çıktılar arasındaki ilişkilerin fonksiyonel formlarının bilinmediği durumlarda bile uygulanabilir, bu da birçok kullanım olanağı sağlamaktadır. Cooper v.d. (2007) VZA metodu hakkında bunu anlattı: “en başarılı firmalara VZA ile etkinlik ölçümü uygulandığımızda bazıları büyük bir verimsizlik tespit edilmiştir, geliştirmek için yeni potansiyeller gösterdi ve eski etkinlik çalışmaları için yeniden değerlendirmeler önerdi”³¹.

Regresyon gibi diğer yöntemlerden farklı olarak, VZA merkezi eğilimler üzerinde çalışmak yerine sınırlar üzerinde çalışır, VZA verileri etkinlik sınırıyla çalışır: etkin birimler bağlayıcı etkinliği sınırında yer alırken, etkin olmayan birimler etkin sınırla karşılaştırılarak etkinlikleri hesaplanır. Regresyon yöntemi ve benzeri teknikler ise veriler üzerinden bir regresyon düzlemi uydurmaya çalışır.

²⁹ Golany, B., Roll, Y. (1989). An application procedure for DEA, p:239.

³⁰ Khezrimotlagh, D. (2014). How to deal with numbers of decision making units and variables in data envelopment analysis. Department of Applied Statistics, Faculty of Economics and Administration, University of Malaya, Kuala Lumpur. P:3,

³¹Cooper, Seiford, Tone (2007): Data envelopment analysis a comprehensive text with models, Applications, Referanss, And DEA-Solver Software, p:29.

VZA yöntemleri, diğer tüm varlıkları gözlemleyerek her birim için en iyi tahmini etkinlik skoru elde etmeye çalışır, çıktı ve girdilerdeki verimsizliği tespit eder ve etkinliğin ölçümünü sağlar. Optimal çözümden girdilerdeki fazlalıklar ve çıktılarıdaki eksiklikler gevşeklikler olarak bilinir.

VZA yöntemlerinin geliştirilmesi ve VZA'nın kullanım alanlarının genişletilmesi ile etkinliğin hesaplanmasında kazandığı önemi yansıtan çok sayıda çalışma yapılmaktadır. VZA kullanılarak yeni teknikler geliştirilmiştir, bunlardan bazılarını sayabiliriz:

PCA-DEA, VZA da meydana gelen boyutsallığın azaltılması için geliştirilmiştir.

2.3.1. VZA'da Etkinlik Kavramı

Etkinlik, bir firmanın üretkenliğini değerlendirmeye çalışırken kullanılan bir terimdir. Yani etkinlik, bir firmanın hedeflerine en az çaba, enerji ve atık miktarları ile ulaşması, üretim sürecinde kullanılan para, insan kaynakları, ekipman ve zamanı içeren üretim araçlarından elde edilen faydanın optimize edilmesi, israf edilmeden kullanılması anlamına gelir.

VZA'daki etkinliği tanımlamak için Genişletilmiş Pareto-Koopmans tanımı vardır, bu da: "Tam (% 100) etkinliğe, herhangi bir karar birimi tarafından ancak ve ancak, diğer girdi veya çıktıların bazılarını kötüleştirmeden girdi veya çıktılarından hiçbiri iyileştirilemiyorsa ulaşılır."³²

Bu tanım daha sonra deneysel bulgulara kullanım için değiştirilir ve “Göreceli Etkinlik” tanımını oluşturulmuştur:

"Bir karar birimi, ancak ve ancak diğer karar birimlerin performansları, bazı girdi veya çıktıların, diğer girdi veya çıktılarından bazılarını kötüleştirmeden iyileştirilebileceğini göstermemesi durumunda tam (%100) etkin olarak derecelendirilmelidir."³³

³² Cooper, W.W. , Seiford, L. M., Zhu, J. (2011) Handbook on Data Envelopment Analysis p:3.

³³ A.g.e., p:3.

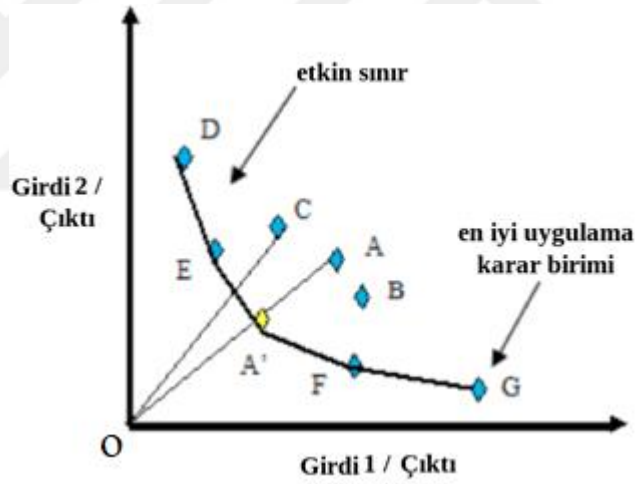
Bir çıktı ve bir girdi ile etkinliği hesaplamının basit bir örneği şu şekilde ifade edilir:

$$\frac{\text{Çıktı}}{\text{Girdi}} \quad [10]$$

"işçi saati başına çıktı" veya "çalıştırılan işçi başına çıktı" olarak da bilinmektedir. Bu oranı, aynı firmada aynı alanda veya farklı bölümlerde çalışan farklı firmalar arasında karşılaştırarak, her firma için etkinliği değerlendirebiliriz.

Etkin sınır, üretim olanakları kümesini belirlemektedir yani etkin sınır etkin karar birimlerinden oluşmaktadır.

Şekil 15. İki Girdi ve Bir Çıktı İçin Etkin Sınır³⁴



Karar birimi A, etkin sınırdan yer almadığından etkin değildir. Karar birimleri D, E, F ve G etkin sınırdadır ve etkinler, çünkü girdilerini değiştirmeden çıktıyı iyileştirmek mümkün değildir. A'ya en yakın etkin iki karar birimleri, E ve F'dir, bunlara A için referans seti denir.

³⁴ Vukic, L., Jugovic, T. P., Guidi, G., Oblak, R. (2020). Model Of Determining The Optimal, Green Transport Route Among Alternatives: Data Envelopment Analysis Settings”, *MDPI Journal Of Marine Science And Engineering* ISSN 2504-4990 Switzerland,p:9.

Karar birimi A'yı etkin hale getirmek için (A'nın bulunduğu yerde) hem girdi 1 hem de girdi 2 azaltılmalıdır. girdi 1 ve girdi 2'de değişmesi gereken miktarlara gevşeklik denir. Ve $\overline{AA'}$ segmentine iyileştirme denir.

Verimsiz karar birimi A için etkinlik değerini hesaplamak için, başlangıç noktasından A ya bir çizgi çizeriz, bu çizgi etkinlik sınırını A' noktasında keser, etkinlik değeri şu şekilde hesaplanır:

$$\frac{OA'}{OA} \quad [11]$$

2.3.2. CCR Model

CCR modeli, Charnes Cooper ve Rhodes (1978) yaklaşımlarında Veri zarflama analizi kullanarak karar birimlerin etkinliğini değerlendirme tanıtılmıştır, önceden belirlenmiş ağırlıklar gerektirmeden tek çıktı tek girdi klasik oranı; çoklu çıktı ve girdi modele genelleştirdiler.

j karar birimi, m girdi ve s çıktı için adlandırdığımız:

Karar birimi 1 , Karar birimi 2, , Karar birimi n

Girdiler $X = x_i$; $i = 1, \dots, m$

Çıktılar $Y = y_r$; $r = 1, \dots, s$

Girdi Yönelimli CCR Model

Model, girdilerin sabit bir getirisi veya bilindiği gibi “ölçeğe göre sabit getiriye (CRS)” sahip olduğunu varsayar. Model kesir olarak bu şekilde gösterilir³⁵:

$$\max \theta_0 = \frac{\sum_{r=1}^s u_r y_r}{\sum_{i=1}^m v_i x_i} \quad [12]$$

³⁵ Banker, R. D., Charnes, A., Cooper, W. W. (1984). Some Models For Estimation Technical And Scale Inefficiencies In Data Envelopment Analysis. *Management Science*, USA, P:1078.

$$1 \geq \frac{\sum_{r=1}^s u_r y_{rj}}{\sum_{i=1}^m v_i x_{ij}}, \quad j = 1, 2, \dots, n$$

$$u_r \geq \varepsilon, r = 1, 2, \dots, s$$

$$v_i \geq \varepsilon, i = 1, 2, \dots, m$$

u_r : r çıktı tahsis edilen ağırlık

v_i : i girişine tahsis edilen ağırlık

θ_0 değeri aralıktadır: $0 < \theta_0 \leq 1$

Bu problemi, kesirli programdan (FP) doğrusal programlamaya (LP) dönüştürülebilir:

36

$$\max_{\mu, v} \theta = \mu_1 y_{1o} + \dots + \mu_s y_{so} \quad [13]$$

$$v_1 x_{1o} + \dots + v_m x_{mo} = 1$$

$$\mu_1 y_{1o} + \dots + \mu_m y_{mo} \leq v_1 x_{1o} + \dots + v_m x_{mo}, \quad (j = 1, \dots, n)$$

$$v_1, v_2, \dots, v_m \geq 0$$

$$\mu_1, \mu_2, \dots, \mu_m \geq 0$$

Lineer programdaki düallite özelliğinden³⁷:

$$\theta^* = \min \theta \quad [14]$$

$$\sum_{j=1}^n x_{ij} \lambda_j \leq \theta x_{io} \quad i = 1, 2, \dots, m$$

$$\sum_{j=1}^n y_{rj} \lambda_j \geq y_{ro} \quad r = 1, 2, \dots, s$$

$$\lambda_j \geq 0 \quad j = 1, 2, \dots, n$$

³⁶ Cooper, W. W., Seiford, L. M., Tone, K. (2007). Data Envelopment Analysis A Comprehensive Text With Models, Applications, DEA-Solver Software p:23.

³⁷ Cooper, W. W., Seiford, L. M., Zhu, J. (2011) Handbook on Data Envelopment Analysis, p:9.

Sınırdan bazı noktalar etkin olabilir ancak sıfır gevşekliği yoktur; onlara “zayıf etkin” diyoruz. Bu sorun hepsinde değil bazı noktalarda olabilir, bu nedenle gevşek değişkenleri dikkate almak için θ^* değerine herhangi bir etkisi olmadan gevşek değerlerini maksimize eden doğrusal programı tanımlarız³⁸:

$$\max \quad \sum_{j=1}^m s_i^- + \sum_{r=1}^s s_r^+ \quad [14]$$

$$\sum_{j=1}^n x_{ij} \lambda_j + s_i^- = \theta^* x_{io} \quad i = 1, 2, \dots, m$$

$$\sum_{j=1}^n y_{rj} \lambda_j - s_r^+ = y_{ro} \quad r = 1, 2, \dots, s$$

$$\lambda_j, s_i^-, s_r^+ \geq 0 \quad \forall i, j, r$$

CCR, girdi-çıkıtı oranlarını veri kümesindeki en etkin karar birimleri ile karşılaştırarak karar verme birimlerinin performansını değerlendirmektedir. CCR etkinlik skoru tüm karar birimleri için $[0,1]$ aralığındadır. Sanal çıktının sanal girdiye henüz bilinmeyen (v_i^*, u_r^*) ile (etkinlik oranı) maksimize ederek karar birimi için optimum ağırlığı (v_i^*, u_r^*) belirler.

$$\text{Sanal girdi} = v_1 x_{1o} + \dots + v_m x_{mo} \quad [15]$$

$$\text{Sanal çıktı} = u_1 y_{1o} + \dots + u_s y_{so} \quad 39$$

Karar birimi O CCR-verimsiz ise⁴⁰:

- $\theta^* < 1$ veya:
- $\theta^* = 1$ ama en azından CCR lineer programında her optimal çözüm için (u^*, v^*) 'den sıfıra eşit bir eleman vardır.

³⁸ A.g.e Sayfa.10.

Cooper, W. W., Seiford, L. M., Tone, K. (2007). Data Envelopment Analysis A Comprehensive Text With Models, Applications, DEA-Solver Software, p:21.

⁴⁰ A.g.e Sayfa.25.

Karar birimi O CCR-verimsiz ve $\theta^* < 1$ olduğu durumunda; CCR-etkin olan ve girdilerini ve çıktıları dengeleyen optimal ağırlıklara sahip en az bir karar birimi olacaktır, aksi takdirde θ^* etkin olacaktır. Böylece tanımlayacağız⁴¹:

$$E'_o = \{ j : \sum_{r=1}^S u_r^* y_{rj} = \sum_{i=1}^m v_i^* x_{ij} \} \quad \text{ve} \quad j \in \{1, \dots, n\} \quad [16]$$

E'_o dan CCR-etkin karar birimlerden oluşan E_o altkümesine, Karar birimi O'nun referans kümesi veya eş grubu denir. Girdilerde ve çıktılarda birbirine yakın değerlere sahip olan bu etkin karar birim seti, Karar birim O'yu verimsiz hale getirir. Dolayısıyla bu kümeye karar birim O'nun etkin sınırı da denir.

CCR doğrusal bir programı halde ve etkinlik oranını θ^* maksimize eden optimal ağırlıkları (u^*, v^*) bulmaya çalışır:

Optimum θ_0^* , $u_r^*, v_i^* > 0$ için optimal değerlerle $0 < \theta_0^* \leq 1$ i karşılar⁴².

Bir karar birimi O için optimal θ^* eşit⁴³:

$$\theta^* = \frac{\sum_{r=1}^S u_r^* y_{ro}}{\sum_{i=1}^m v_i^* x_{io}} \quad [17]$$

Ve doğrusal programdan:

$$\sum_{i=1}^m v_i^* x_{io} = 1 \quad [18]$$

Bu, θ^* ı şuna eşit yapar:

$$\theta^* = \sum_{r=1}^S u_r^* y_{ro} \quad [19]$$

Bir girdi i ögesi için v_i^* ağırlıklarının değeri, o ögenin süreçteki önemini ifade eder. Bir çıktı ögesi r için ağırlık u_r^* ile aynıdır. $u_r^* y_{ro}$ 'nun değeri, O'nun değerine y_{ro} 'nin katkısını gösterir.

⁴¹ A.g.e Sayfa.25.

⁴² Banker, Charnes, Cooper (1984). Some Models For Estimation Technical And Scale Inefficiencies In Data Envelopment Analysis, *Management Science*, USA, p:1079.

⁴³ Cooper, W. W., Seiford, L. M., Tone, K. (2007). Data Envelopment Analysis A Comprehensive Text With Models, Applications, DEA-Solver Software. p.25.

2.3.3. BCC Model

Bu model Banker, Charnes & Cooper tarafından 1984 yılında CCR modelinden genişletilerek geliştirilmiştir, kriterler CCR ile aynıdır. CCR modelinden farklı olarak BCC, girdi fazlalıklarını ve çıktı eksikliklerini de ölçer. Bu modeldeki sınırlar, onu ölçeğe göre değişken getiri (VRS) modeli yapan parçalı doğrusal ve içbükey karakterizasyonlardır. BCC modeli, CCR modelinden $\sum_{k=1}^K \lambda_k = 1$ ekleyerek türetilen bir modeldir (Cooper v.d. 2011).

Girdi yönelimli BCC modeli, belirli bir karar verme biriminin (karar birimi O) etkinliğini bir dizi karar biriminde ($o = 1, \dots, n$ olarak endeksleri) değerlendirir. Bu değerlendirme, doğrusal bir programı zarf biçiminde çözerek yapılır⁴⁴.

$$\begin{aligned} \min \quad & \theta \\ \sum_{j=1}^n \lambda_j x_{ij} & \leq \theta x_{io} \quad , \quad i = 1, 2, \dots, m \\ \sum_{j=1}^n \lambda_j y_{rj} & \geq y_{ro} \quad , \quad r = 1, 2, \dots, s \\ \sum_{j=1}^n \lambda_j & = 1 \\ \lambda_j & \geq 0 \quad , \quad j = 1, 2, \dots, n \end{aligned} \quad [20]$$

Son modelden, optimal değerleri $\theta^* = 1$, λ^* olarak hesapladığımızda, verimsiz karar birimi için gevşek “slacks” değerleri şu şekilde tanımlayabiliriz⁴⁵:

$$\begin{aligned} S_i^- &= \theta^* x_{io} - \sum_{j=1}^n \lambda_j x_{ij} \quad i = 1, 2, \dots, m \\ S_r^+ &= \sum_{j=1}^n \lambda_j y_{rj} - y_{ro} \quad r = 1, 2, \dots, s \end{aligned} \quad [21]$$

Ölçeğe göre değişken getiri şu şekilde ifade edilir; eğer girdiler iki katına çıktığında çıktılar da iki katına çıkıyorsa, bu durumda ölçeğe göre sabit getiri söz konusudur. Ölçeği büyütmek ya da küçültmek verimliliği değiştirmemektedir. Eğer

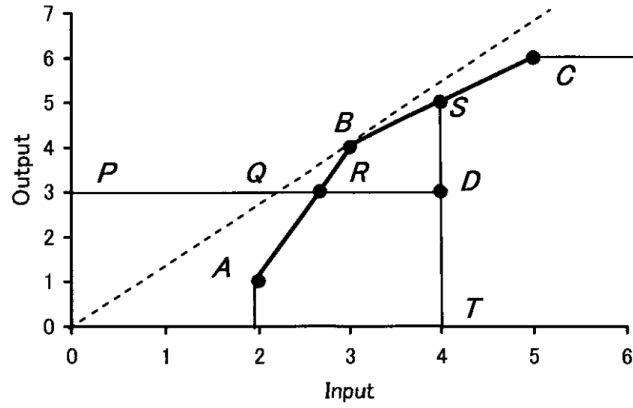
⁴⁴ Zhu, J. (2003). Quantitative Models For Performance Evaluation And Benchmarking Data Envelopment Analysis With Spreadsheets And DEA Excel Solver Chapter 1 Basic DEA Models”, *Springer Science+Business* p:5.

⁴⁵ A.g.e., p:7.

girdileri iki katına çıkarmak, çıktıları iki katından daha fazla veya daha az değiştiriyorsa, bu durumda ölçeğe göre değişken getiri söz konusudur. Ölçeği büyütmek veya küçültmek verimliliği etkilemektedir anlamına gelir.

Ölçeğe göre sabit getiri varsayımına dayanan CCR modelin etkin sınırı düz doğru şeklinde iken, ölçeğe göre değişken getiri varsayımına dayanan BCC modelin etkin sınırı parçalı doğrusal formdadır. Bu durum Şekil 16'da açıklanmıştır:

Şekil 16. CCR ve BCC Etkin Sınırları⁴⁶



Karar birimleri A, B ve C, sınır oluşturdıkları BCC etkinlik sınırında bulunurlar ve bu karar birimler BCC etkinken, CCR etkinlik çizgisi orijin ve karar birimi B ile oluşturulur ve sadece karar birimi B CCR etkindir. Karar birimi D, BCC verimsiz ve CCR verimsizdir.

CCR etkinlik sınırının BCC etkinlik sınırının üzerinde yer aldığı ve çizgilerin yalnızca B noktasında kesiştiği dikkat çekicidir; etkin değer, etkinlik sınır çizgisinden uzaklığın oranı olduğundan; CCR etkinlik değeri her zaman BCC etkinlik değerinden küçük veya ona eşit olacaktır.

⁴⁶ Cooper, W. W., Seiford, L. M., Tone, K. (2007). Data Envelopment Analysis A Comprehensive Text With Models, Applications, *DEA-Solver Software*. P:90.

Karar birimi D için CCR etkinlik değeri, $\frac{PQ}{PD}$ oranı iken, Karar birimi D için BCC etkinlik değeri $\frac{PR}{PD}$ dir.

PR değeri PQ değeri dan büyük olduğu için, BCC etkinliğini CCR etkinliğinden daha yüksek yapar. Karar birimi B için etkinlik değeri CCR ve BCC etkinlik değerleri ile aynı ve 1 olacaktır, Çıktı yönelimli modelde, $\frac{ST}{SD}$ oranı, BCC etkinliğine ulaşmak için gereken artırma seviyesini gösterir.

Banker v.d (1984) üretim olasılığını şu şekilde tanımlamışlardır:

$$P_B = \{ (x,y) \mid x \geq X\lambda, y \leq y\lambda, e\lambda = 1, \lambda \geq 0 \} \quad [22]$$

Halbuki ; $X=(x_i) \in R^{m \times n}$, $y=(y_i) \in R^{s \times n}$, $\lambda \in R^n$ elemanları l'e eşit olan bir ham vektördür.

Bir karar birimi O, yalnızca (a) $\theta^* = 1$ ve (b) $s^{-*} = 0$, $s^{+*} = 0$ ise, tam etkindir (Pareto-efficient), aksi halde karar birimi O BCC-verimsizdir. s^{-*} ve s^{+*} gevşek değişkenlerdir⁴⁷.

BCC yönteminin daha önce de bahsedildiği gibi avantajlarından biri de eksikliklerin ve fazlalıkların hesaplanmasıdır.

Gevşek değişkenleri kullanarak, birden fazla optimal çözüm olması durumunda bunlardan birini seçip iyileştirme formülü hesaplanır⁴⁸:

$$\theta^* x_0 = \sum_{j \in E_0} \theta_j^* x_0 + S^{-*} \quad [23]$$

$$y_0 = \sum_{j \in E_0} \lambda_j^* y_j - S^{+*}$$

Dolayısıyla, BCC modeli için iyileştirme formülü şöyle olacaktır⁴⁹ :

⁴⁷ Thanassoulis, E. (2001). Introduction To The Theory And Application Of Data Envelopment Analysis", *Springer Science+Business Media*, New York, 2001, <https://link.springer.com/book/10.1007/978-1-4615-1407-7> .p:66.

⁴⁸ Cooper, W. W., Seiford, L. M., Tone, K., (2007). Data Envelopment Analysis A Comprehensive Text With Models, Applications, DEA-Solver Software ,p:92.

⁴⁹ A.g.e. Sayfa.92.

$$\begin{aligned}\hat{x}_0 &\Leftarrow \theta^* x_0 - s^{-*} \\ \hat{y}_0 &\Leftarrow y_0 + s^{+*}\end{aligned}\quad [24]$$

2. 4. Süper Etkinlik

Süper etkinlik kavramı ilk kez Andersen ve Petresen 1993 da tanıtıldı, etkinlik skoru karar birim O için bağlı değişkenlerin değerleri kaldırarak hesaplanır, çözüm seti kullanarak değerlendirir. Süper etkinlik modeller kullandığı teknikler yüzünde bazı problemler yaratıyor; ölçümlerde değişmezlik eksikliğini⁵⁰ ve BCC modelde bazı karar birimleri için çözümsüz sonuçlanabilir, bu problemler radyal olmayan modeller kullanarak ortadan kaldırılabilir.

2. 4. 1. Radyal Süper Etkinlik Modelleri

Andersen ve Petersen, BCC modelinde ortaya çıkartabilecek çözümsüzlük⁵¹ ihtimalleri ortadan kaldırmak için CCR yapı kullanarak modellerini oluşturmaktadırlar.

Süper radyal I-C modeli⁵²

$$\begin{aligned}\theta^* &= \min_{\theta, \lambda, s^-, s^+} \theta - \varepsilon s^+ \\ \theta x_o &= \sum_{j=1, \neq 0}^n \lambda_j x_j + s^-\end{aligned}\quad [25]$$

⁵⁰ Ölçümler değişmezliği eksikliği farklı gruplarla ya da farklı zamanda model farklı bir yapı sahip olması demektir; bu halde model farklı grupla ya da farklı zamanda yorumlanamayan bir hale getiriyor (Putnick and Bornstein, 2016),

VZA de CCR ve BCC modelleri etkinlik skorları değişmezlik sahiptir ama gevşekler değişkenlerde değişmezlik eksiktir (Lovell 1995) sayfa 147-148,

⁵¹ Süper etkin karar birimleri yüksek performans görünümü nedeniyle bazı etkin birimler aykırı değeri olarak değerlendirme yol açar, bunu gerçekleştirdiği zaman süper etkinlik modeli sadece aykırı değeri değerlendirdiği karar birimleri için etkinlik skoru hesaplama başarısız olur ve etkinlik skoru çözümsüz denir, Banker ve Chang (2006), The super-efficiency procedure for outlier identification, not for ranking efficient units, sayfa 1311-1320,

⁵² Cooper, Seiford, Tone (2007). Data Envelopment Analysis A Comprehensive Text With Models, Applications, refe DEA-Solver software p:310.

$$y_o = \sum_{j=1, \neq 0}^n \lambda_j y_j - s^+$$

$$\lambda, s^-, s^+, \varepsilon > 0$$

2.4.2. Radyal Süper Etkinlik VRS Modeli

En son modeli aşağıdaki ilişkiyi birleştirerek ölçeğe göre değişken getiri bir modele uzatabilir⁵³:

$$\sum_{j=1}^n \lambda_j = 1, \lambda_j \geq 0, \forall j \quad [26]$$

2.4.3. Süper Radyal I-V

$$\theta^* = \min_{\theta, \lambda, s^-, s^+} \theta \quad [27]$$

$$\theta x_{io} \geq \sum_{j=1, \neq 0}^n \lambda_j x_{ij} \quad (i = 1, \dots, n)$$

$$y_{ro} \leq \sum_{j=1, \neq 0}^n \lambda_j y_{rj} \quad (r = 1, \dots, s)$$

$$\sum_{j=1}^n \lambda_j = 1, \lambda_j \geq 0, \forall j$$

⁵³ A.g.e., p:317.

ÜÇÜNCÜ BÖLÜM

UYGULAMA

3.1. Veri Kümesi

Veriler, İstanbul borsasına kote edilmiş 12 demir-çelik firmasının 2021 yılı verileridir. Çalışmaya konu olan firmaların listesi Tablo.3'teki gibidir:

Tablo 3. Firmaların Listesi

Kodlanmış adı	Açık adı
ALMAD	Altınyâğ Madencilik Ve Enerji Yatırımları Sanayi Ve Ticaret A.Ş.
AYES	Ayes Çelik Hasır ve Çit Sanayi A.Ş.
BMSCH	BMS Çelik Hasır Sanayi Ve Ticaret A.Ş.
BURCE	Burçelik Bursa Çelik Döküm Sanayi A.Ş.
CELHA	Çelik Halat ve Tel Sanayi A.Ş.
CEMTS	Çemtaş Çelik Makina Sanayi ve Ticaret A.Ş.
EREGL	Ereğli Demir ve Çelik Fabrikaları T.A.Ş.
ISDMR	İskenderun Demir ve Çelik A.Ş.
IZMDC	Izmir Demir çelik sanayi A.Ş.
KOZA	Koza Anadolu Metal Madencilik İşletmeleri A.Ş.
KRDMD	Kardemir Karabük Demir Çelik Sanayi ve Ticaret AS
YKSLN	Yükselen Çelik A.Ş.

3.2. Değişkenler

Etkinlik değerlendirmesinde daha önce yapılan araştırmaların kullanıldığı kertya değişkenler Tablo.4'te gösteriliyor.

Tablo 4. Daha Önce Yapılan Araştırmalarda Kullanılan Değişkenlere Örnekler

Yazar	Eser	Değişkenler
Yavuz Selim Akbulut	Gayrimenkul Yatırım Ortaklığı Şirketlerinin Finansal Etkinliklerinin Vza Yöntemi İle Ölçülmesi. Bist’ de Bir Uygulama, Y1 Tez, 2021	• Finansman Gideri • Öz Sermaye • Faaliyet Gideri • Satışların Maliyeti • Net Dönem Kârı • Net Dönem Kârı • Esas Faaliyet Kârı
Çağatay Biçen	Otomotiv Sektöründe Veri Zarflama Analizi İle Finansal Etkinlik Ölçümü, Y1 Tez, 2010	• Özkaynak • Toplam Aktif • İstihdam • Ciro • Net Kar • İhracat
Tuba Özkan	Türk Bankacılık Sisteminde Veri Zarflama Tekniği İle Banka Etkinliğinin Ölçülmesi, Araştırma Makalesi, 2019	• Mevduat • Faiz Giderleri • Diğer Faaliyet Giderleri • Kredi Ve Alacaklar • Faiz Gelirleri • Diğer Faaliyet Gelirleri • Net Kar
Nazlı Gamze Özel, İbrahim Erem Şahin, Ramazan Göral	Türk Bankacılık Sektöründe Etkinlik Ve Verimlilik Analizinin Veri Zarflama Yöntemi İle İncelenmesi: 2013-2015 Dönemi Uygulaması, 2017	• Toplam Mevduat/Toplam Aktifler • Faiz Giderleri/Toplam Aktifler • Diğer Faaliyet Giderleri/Toplam Aktifler • Toplam Krediler Ve Alacaklar/Toplam Aktifler • Faiz Gelirleri/Toplam Aktifler
Gökhan Karaman	Borsa İstanbul’da İşlem Gören Metal Ana Sanayi Firmalarının Ekonomik Etkinlik Analizi, Y1 Tez 2019	• Toplam Borç / Toplam Varlık • Dönen Varlıklar / Kısa Vadeli Yabancı Kaynaklar • Dönen Varlık – Stoklar / Kısa Vadeli Yabancı Kaynaklar • Maddi Duran Varlık / Öz Sermaye

		• Net Kâr / Toplam Varlık • Net Kâr / Net Satışlar
--	--	---

Tablo 4... Devam

Yazan Adı	Makale Adı	Kullanıldı Değişkenler
Pejman Peykan, Fateme Sadat Seyed Esmaeili, Mirpouya Mirmozaffari, Armin Jabbarzadeh, Mohammad Khamechian.	Input/Output Variables Selection In Data Envelopment Analysis: A Shannon Entropy Approach, 2022.	• Alacak Devir Oranı • Ödeme Gücü Oranı-Ii • Kâr Marjı Artış Oranı • Özkaynak Kârlılığı • Nakit Oranı
Seda Sütçü Asar	Using Principal Component Analysis In Data Envelopment Analysis, Y1 Tez 2014.	• Personel Sayısı • Acente Sayısı • Özsermaye • Prim Üretimi • Teknik Kar/Zarar • Dönem Net Karı/Zararı
Emine Üstündağ	Vergî Zarflama Analîzgî Ğle Vergîmlîğîğğîğn Değerlendîrgîğmesîğ: Çğmento Sektörû Üzerîğne Uygulama, Y1 Tez 2009.	• Toplam Dönen Varlıklar / Toplam Kısa Vadeli Borçlar • Toplam Yabancı Kaynak/ Toplam Aktif • Kısa Vadeli Yabancı Kaynak/ Toplam Pasif • Net Kar/ Net Satışlar • Net Kar/ Özsermaye • Net Kar/ Toplam Aktif
Esmâ Canhasî Kasemî	Vza Ve Vzâhp Bütünleşîk Metodu İle Verîmlîlîk Analîzî: Kosova Gîda Sektörû Uygulamasî, Doktora Tez, 2021.	• Öz Kaynaklar • Aktif Toplam • İşçi Giderleri • Kar • Net Satış

Demir-çelik sektöründeki üretimi tanımladığı düşünülen 27 adet değişken aşağıda Tablo.4'teki gibidir:

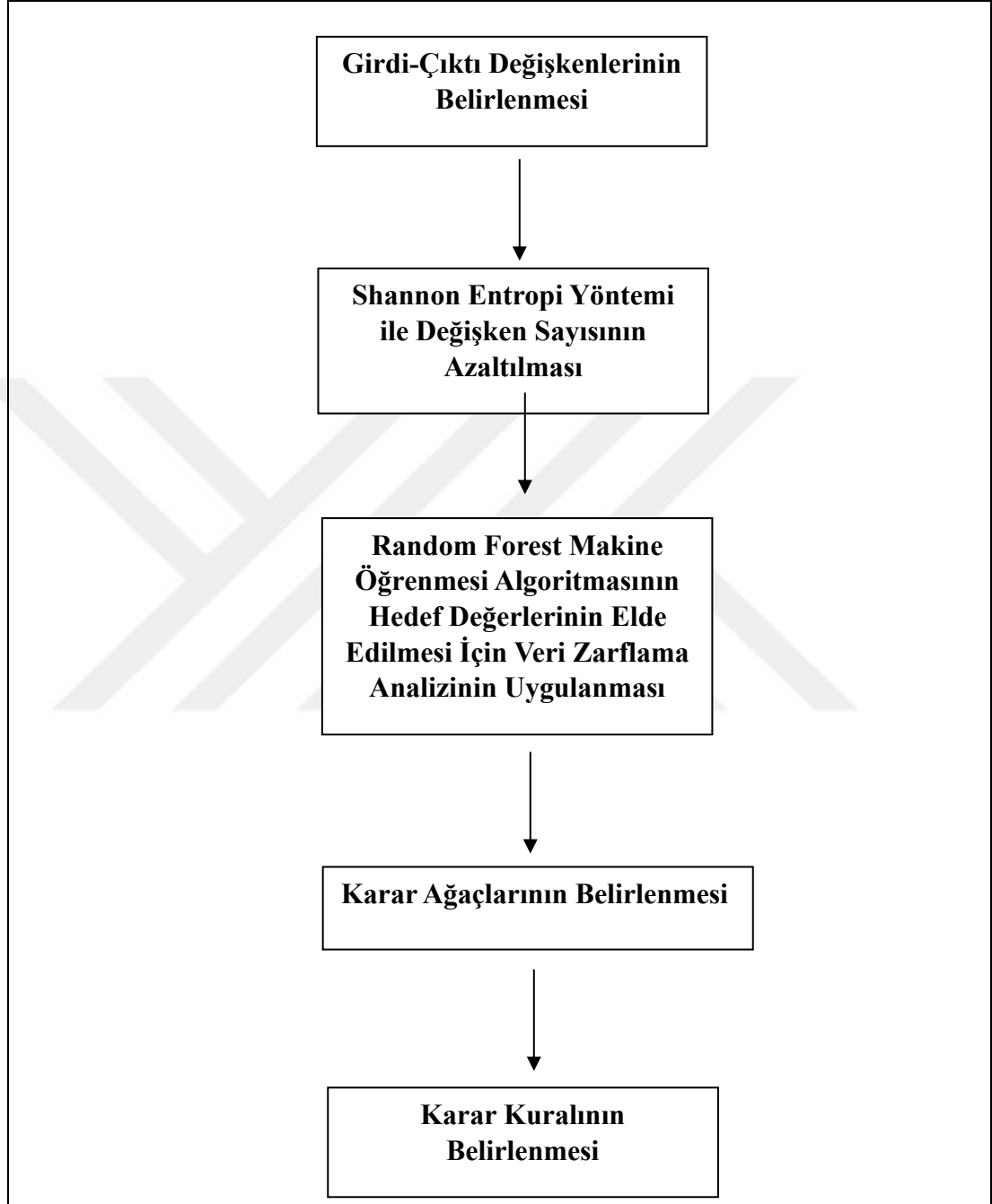
Tablo 5. Değişken Listesi

Kodlanmış Adı	Açık adı	Kodlanmış Adı	Açık adı
O-HS	Hasılat	O-NKP	Net Kar/Personel
O-BK	Brüt kar	O-LO	Likidite Oranı MRQ
O-FK	Faaliyet kar	O-CO	Cari Oran
O-VK	Sürdürülen Faaliyetler vergi öncesi karı	I-DUV	Duran Varlıklar
O-NK	Net dönen Kar	I-KVY	Kısa Vadeli Yükümlülükler
O-FVK	FAVÖK	I-UVY	Uzun Vadeli Yükümlülükler
O-DOV	Dönen Varlıklar	I-TEN	Toplam Envanter
O-NS	Net satışlar	I-SEG	Sermaye Giderleri
O-5TV	5 Yıllık Temettü Verimi Ortalaması 5YA	I-YFN	Yatırım Faaliyetlerinden Sağlanan Nakit
O-KDO	Kar Dağıtım Oranı TTM	I-FFN	Finansman Faaliyetlerinden Sağlanan Nakit
O-OZG	Özkaynak Getirisi TTM	I-TBV	Toplam Borç/Varlık
O-ADH	Aktif Devir Hızı TTM	I-BAV	Borç / aktif varlık oranı Mevcut
O-SDH	Stok Devir Hızı TTM	I-BOZS	Borç/öz sermaye oranı Mevcut
O-CP	Ciro/Personel		

3.3. Uygulama Adımları

Uygulamanın kavramsal çerçevesi Tablo.6'daki gibidir:

Tablo 6. Random Forest Makine Öğrenmesi Tekniğinin Adımları⁵⁴

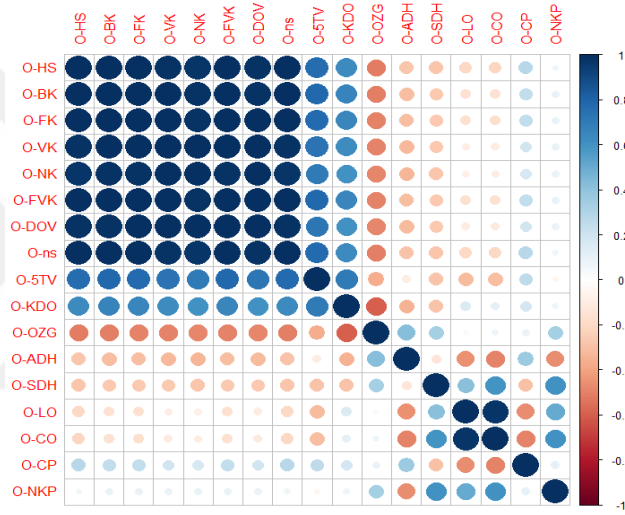


⁵⁴ Çınaroğlu, S. Onur Başer, O. (2017). Kamu Hastane Birliklerinin Teknik Verimliliklerinin Veri Zarflama Analizi ve Makine Öğrenmesi Teknikleri Kullanılarak İncelenmesi, *Ekonomik Yaklaşım*, s:102. (adı geçen kaynaktan esinlenilmiştir)

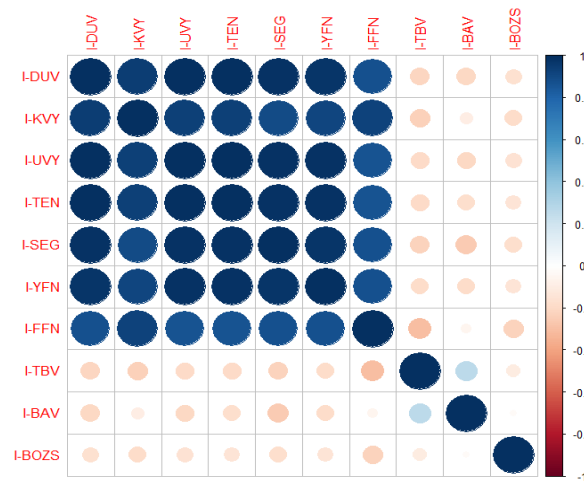
3. 3. 1. Random Forest Algoritması İçin Değişken Kümelerinin Shannon Entropi İle Oluşturulması (Korelasyonların Hesaplanması)

Yüksek oranda ilişkili değişkenler genellikle gereksiz bilgiler sağlar. Bir modele güçlü bir şekilde ilişkili birden çok değişkenin dâhil edilmesi, fazla uyumlamaya yol açabilir ve modeli daha az yorumlanabilir hale getirebilir. Korelasyonu inceleyerek, paylaşılan bilgileri yakalayan yalnızca bir temsili değişkeni tutarak gereksiz değişkenleri belirleyebilir ve kaldırabiliriz.

Şekil 17. Çıktı Değişkenlerinde Korelasyon



Şekil 18. Girdi Değişkenlerinde Korelasyon



Korelasyon değerlerine göre çıktı değişkenleri ve girdi değişkenleri kendi aralarında yüksek korelasyon gösteren daha küçük alt gruplara (karar matrisi grupları) bölünecektir. Bu gruplandırma, kendi kategorileri içinde birbirleriyle güçlü bir şekilde ilişkili olan değişkenlerin alt kümelerini tanımlamamıza olanak tanır. Değişkenleri bu daha küçük gruplar halinde düzenleyerek, Shannon entropi yöntemini uygulamak ve karar grubunda en yüksek niteliğe sahip değişkenleri VZA değerlendirme etkinliği prosedürüne dahil etmek için belirlemek mümkündür.

Giriş ve çıkış grupları, her biri iki alt gruba bölünecektir. Girdi grubunun ilk alt grubunda, yüksek korelasyonlu değişkenlerden bir değişken seçilecektir. Giriş grubunun ikinci alt grubunda iki değişken seçilecektir. Benzer şekilde çıkış grubunun ilk alt grubundan bir değişken seçilecek, fakat ikinci alt grubundan üç değişken seçilecektir. Seçilen girdi değişkenler ve çıktı değişkenler aralarındaki korelasyon değeri %40 dan az olması önem verilecektir.

3.3.2. Shannon Entropi ve Ağırlıkların Hesaplanması

Aşağıdaki denklemleri⁵⁵ kullanarak entropi ve ağırlıkları dört adımda hesaplanmıştır:

Birinci adım: her değişken grubu için karar matrisi oluşturma

$$\begin{pmatrix} x_{11} & \cdots & x_{1j} \\ \vdots & \ddots & \vdots \\ x_{i1} & \cdots & x_{ij} \end{pmatrix} \quad [28]$$

İkinci adım: Karar matrisi normalleştirmektir:

$$p_{ij} = \frac{x_{ij}}{\sum_{i=1}^m x_{ij}} \quad [29]$$

Üçüncü adım: Entropinin hesaplanması

⁵⁵ Alinezhad, A., Khalili, J. (2019). New Methods And Applications In Multiple Attribute Decision Making (MADM), Springer Science+Business Media, New York <https://link.springer.com/book/10.1007/978-3-030-15009-9>, p:134.

$$e_j = -h \sum_{i=1}^m p_{ij} \ln p_{ij} \quad [30]$$

$$h = \frac{1}{\ln(m)}$$

$$j = 1, 2, \dots, n, \quad m = 12$$

Dördüncü adım: Ağırlıkların hesaplanması

$$w_j = \frac{d_j}{\sum_{j=1}^n d_j} \quad [31]$$

$$d_j = 1 - e_j$$

Tablo 7. Birinci Girdi Değişken Grubu; Shannon Entropi ve Ağırlık Değerleri

	I-DUV	I-KVY	I-UVY	I-YFN	I-FFN	I-TEN	I-SEG	
Ej	0.198786	0.250575	0.180354	0.174028	0.302908	0.198785	0.210717	SUM
1-Ej	0.801214	0.749425	0.819646	0.825972	0.697092	0.801215	0.789283	5.48385
Wj	0.146104	0.136660	0.149466	0.150619	0.127117	0.146104	0.143929	1

Ej: Shannon entropi değeri, Wj: entropi ağırlık değeri

Tablo 8. İkinci Girdi Değişken Grubu; Shannon Entropi ve Ağırlık Değerleri

	I-TBV	I-BAV	I-BOZS	
Ej	0.271651	0.376253	0.260203	SUM
1-Ej	0.728349	0.623747	0.739797	2.091893
Wj	0.348177	0.298174	0.353650	1

Ej: Shannon entropi değeri, Wj: entropi ağırlık değeri

Tablo 9. Birinci Çıktı Değişken Grubu; Shannon Entropi ve Ağırlık Değerleri

	O- HS	O- BK	O- FK	O- VK	O- NK	O- FVK	O- DOV	O- NS	O- 5TV	O- KDO	O- OZG	
Ej	0.24 64	0.22 21	0.21 76	0.22 77	0.23 54	0.21 61	0.23 38	0.24 25	0.26 46	0.25 15	0.41 74	SUM
1- Ej	0.75 36	0.77 79	0.78 24	0.77 23	0.76 46	0.78 39	0.76 62	0.75 75	0.73 54	0.74 85	0.58 26	8.22 49
Wj	0.09 16	0.09 46	0.09 51	0.09 39	0.09 30	0.09 50	0.09 32	0.09 21	0.08 94	0.09 10	0.07 08	1

Ej: Shannon entropi değeri, **Wj:** entropi ağırlık değeri

Tablo 10. İkinci Çıktı Değişken Grubu; Shannon Entropi ve Ağırlık Değerleri

	O-ADH	O-SDH	O-CP	O-NKP	O-LO	O-CO	
Ej	0.388668	0.340056	0.390423	0.369464	0.338340	0.305544	SUM
1-Ej	0.611332	0.659944	0.609577	0.630536	0.661660	0.694456	2.51139
Wj	0.243424	0.262780	0.242725	0.251071	0.263464	0.276523	1

Ej: Shannon entropi değeri, **Wj:** entropi ağırlık değeri

3.4. Random Forest Algortiması İçin Shannon Entropi Yöntemiyle Elde Edilen Girdi-Çıktılar Üzerinden Hedef Değerlerin Veri Zarflama Analizi İle Elde Edilmesi

Bu bölümde öncelikle Shannon Entropi yöntemi uygulanarak, performansı en fazla etkilediği tespit edilen 3 girdi ve 4 çıktı ile yapılan uygulamaya yer verilmiştir. Uygulamaya konu olan girdiler ve çıktılar Tablo 10'daki gibidir:

Tablo 11. Girdi ve Çıktı VZA Değişkenleri

I1	Toplam Borç/Varlık
I2	Borç/öz sermaye oranı
I3	Yatırım Faaliyetlerinden Sağlanan Nakit
O1	Faaliyet karı
O2	Stok Devir Hızı TTM
O3	Net Kar/Personel TTM
O4	Cari Oran

Tablo 12. Uygulamada Kullanılan Veri Kümesi

	I-						
	TBV	I-BOZS	I-YFN	O-FK	O-SDH	O-NKP	O-CO
	I1	I2	I3	O1	O2	O3	O4
Firma4	6.39	0.05	26,800,000	5832709	46910000	3830000	8740000
Firma5	121.98	1.86	7,722,490	98171573	9710000	465720	380000
Firma9	12.56	0.31	5,800,000	32990518	13700000	578400	830000
Firma1	71.05	0.55	4,640,000	7825951	3250000	91670	640000
Firma12	541.68	4.1	27,590,000	31523969	4950000	68910	430000
Firma2	17.62	0.06	67,200,000	481691327	3350000	2090000	2290000
Firma11	24.9	0.25	7,520,856,000	24871472000	2580000	1600000	880000
Firma8	21.83	0.22	2,449,550,000	13009462000	2920000	1560000	780000
Firma3	21.83	7.62	297,090,000	1173279088	7310000	228630	340000
Firma7	0.53	0.01	594,470,000	1995840000	3830000	1370000	8930000
Firma10	19.3	0.19	155,180,000	4797420845	3760000	878090	480000
Firma6	152.61	0.57	22,310,000	160015506	4760000	2520000	780000

Tablo 13. CCR, BCC ve Ölçek Etkinlik Skorları^{56, 57}

Karar Birimleri	CCR Etkinlik Skorları	BCC Etkinlik Skorları	Ölçek Etkinliği
Firma ₄	100 %	100 %	100%
Firma ₅	87.5 %	100 %	88%
Firma ₉	100 %	100 %	100%
Firma ₁	47.4 %	100 %	47%
Firma ₁₂	10.7 %	19.7 %	54%
Firma ₂	75.9 %	100 %	76%
Firma ₁₁	87.1 %	100 %	87%
Firma ₈	100 %	100 %	100%
Firma ₃	23.7 %	38.2 %	62%
Firma ₇	100 %	100 %	100%
Firma ₁₀	100 %	100 %	100%
Firma ₆	100 %	100 %	100%
Ortalama	78%	88%	85%

Ölçeğe göre sabit getirili CCR modele göre etkin bulunan firma sayısı 6'tır ve etkinlik ortalaması %78'tür. Ölçeğe göre değişken getirili BCC modeline göre ise 9 firma etkin bulunmuştur ve ortalama etkinlik %88'e yükselmiştir. Bunun nedeni, CCR modelin etkin sınırının BCC modelin etkin sınırına kıyasla daha kuzeybatıda olması ve ulaşılmasının daha zor olmasıdır. Oysaki BCC modelin etkin sınırına ulaşmak ve etkin olmak daha kolaydır. Ölçek etkinliklerine bakıldığında, ölçek etkinliği ortalaması %85'dur fakat optimal ölçekte faaliyet gösteren firma oranı %50'dir.

⁵⁶ Firma etkinlik sıralaması 2021 yılına ait kesit verilere dayandığı ve araştırmacının entropi yöntemiyle elde ettiği girdi-çıkıtı kümesi üzerinden hesaplandığı için deneysel kabul edilmelidir. Bir başka dönemde, farklı girdi-çıkıtı kümeleri üzerinden hesaplanmış olsa, başka sonuçlar çıkması kaçınılmazdır. Bahse konu etkinlik skorları firmalar için mutlak bir başarı sıralaması anlamına gelmemesine rağmen, yine de, firma ismi verilmeden kodlar kullanılarak sunulmuştur.

⁵⁷ CCR model CRS, yani, ölçeğe göre sabit getirili modeldir. BCC model VRS, yani, ölçeğe göre değişken getirili modeldir. Ölçek etkinliği ise, CCR/BCC etkinliklerinin bölünmesiyle elde edilmektedir.

Tablo 14. Süper Etkinlik Skorları

Karar Birimi	Süper I-C Modeli	Süper I-V Modeli
Firma ₄	1000.00%	1000.00%
Firma ₅	87.50%	100.40%
Firma ₉	152.20%	294.10%
Firma ₁	47.40%	125.00%
Firma ₁₂	10.70%	19.70%
Firma ₂	75.9%	105.4%
Firma ₁₁	87.10%	1000.00%
Firma ₈	105.70%	137.60%
Firma ₃	23.70%	38.20%
Firma ₇	1000.00%	1000.00%
Firma ₁₀	578.70%	577.90%
Firma ₆	100.80%	102.10%

3.5. Random Forest Makine Öğrenmesinin Uygulanması

Tohum “seed”: Yazılımda tohum, rasgele sayıları üretimde bir başlangıç değeri görevi görür. Böylece rastgele sayı dizileri üretmek için algoritmaları takip eder. Bu, aynı tohumla her başladıklarında, aynı sayı dizisini ürettikleri anlamına gelir. Tohumlar, rastgele sayıların işlenmesinde çok önemli bir rol oynar, çünkü aynı tohumun kullanılması, aynı rastgele sayılar dizisinin yeniden oluşturulmasını garanti eder. Bu tekrarlanabilirlik, hata ayıklama ve deneyleri çoğaltmak için hayati önem taşır. Bu deneyimde başta farklı tohumları deneyerek tüm modellerde eğitim setinde en çok çeşitliliği sağlayabildiği tohum “532623” olduğu için seçilmiştir.

Aşağıda sırasıyla CCR, BCC ve Süper etkinlik modellerinden elde edilen etkinlik skorlarını kullanan Random Forest algoritmasının bulguları görülmektedir. Yaygın bölme şekillerden biri olarak (Brownlee 2020⁵⁸): Eğitim kümesi %50 ve test kümesi %50 olarak tanıtıldı.

⁵⁸ Brownlee, J. (2020) Train-Test Split for Evaluating Machine Learning Algorithms, Machine Learning Mastery
<https://machinelearningmastery.com/train-test-split-for-evaluating-machine-learning-algorithms/>

Tablo 15. “532623” Tohum ile CCR Etkinlik Skorlarını Kullanan Random Forest Algoritmasının R Kodları

```

Console Terminal Background Jobs
R 4.3.1 ~ /
> set.seed(532623)
> dadt02 <- read_excel("C:/Users/PC/Desktop/Data22DT.xlsx")
New names:
* `` -> `...11`
> data1<-dadt02[1:7]          ### Degiskenler almak
> data1["y"]<- dadt02$'CCR-Eff' ### VZA dan CCR etkinlik degisken
> data1$y<- as.factor(data1$y) ### Etkinlik degisken faktor tanima
> test<- sample(1:nrow(data1),6) ### Egitim kumesi tanima
> data<- data1[test,]        ### veryidan test KVBler almak
> ### Modeli ulusturma
> model <- randomForest(y ~ ., data=data, proximity=TRUE, ntree=1000)
> model                      ### Modeli cagirma

Call:
randomForest(formula = y ~ ., data = data, proximity = TRUE, ntree = 1000)
Type of random forest: classification
Number of trees: 1000
No. of variables tried at each split: 2

OOB estimate of error rate: 16.67%
Confusion matrix:
      Etkin EtkinDegil class.error
Etkin      4          0          0.0
EtkinDegil  1          1          0.5
> ### Test seti icin model ile tahmin
> predictions2 <- predict(model, data1[-test,], type="class")
> predictions2
      1      2      3      4      5      6
EtkinDegil Etkin Etkin EtkinDegil Etkin Etkin
Levels: Etkin EtkinDegil

```

Tablo 15...devamı

```

> ptestdata<-data1[-test,]    ### test veryi tanima
> mean(ptestdata$y==predictions2)### Dogruluk orani hesaplama
[1] 0.6666667
> ### Tum veri icin model ile tahmin etme
> predictions3 <- predict(model, data1, type="class")
> predictions3                ### Dogruluk orani hesaplama
      1      2      3      4      5      6      7
Etkin EtkinDegil Etkin EtkinDegil EtkinDegil Etkin Etkin
      8      9     10     11     12
Etkin EtkinDegil Etkin Etkin Etkin
Levels: Etkin EtkinDegil
> mean(data1$y==predictions3)
[1] 0.8333333
>

```

CCR etkinlik skorları hedef değeri alındığında Random forest algoritması %83 oranında doğru bir sınıflandırma yapmıştır.

Tablo 16. “532623” Tohum ile BCC Etkinlik Skorlarını Kullanan Random Forest Algoritmasının R Kodları

```

R 4.3.1 ~ /
> set.seed(532623)
> dadt02 <- read_excel("C:/Users/PC/Desktop/Data22DT.xlsx")
New names:
  ... -> '...11'
> data1<-dadt02[1:7]          ### Degiskenler almak
> data1["y"]<- dadt02$'BCC-Eff' ### VZA dan BCC etkinlink degisken
> data1$y<- as.factor(data1$y) ### Etkinlik degisken fakor tanima
> test<- sample(1:nrow(data1),6) ### Egitim kumesi tanima
> data<- data1[test,]         ### veryidan test kvBler almak
> ### Modeli ulusturma
> model <- randomForest(y ~ ., data=data, proximity=TRUE, ntree=1000)
> model                       ### Modeli cagirma

Call:
randomForest(formula = y ~ ., data = data, proximity = TRUE,      ntree = 1000)
Type of random forest: classification
Number of trees: 1000
No. of variables tried at each split: 2

OOB estimate of error rate: 0%
Confusion matrix:
      Etkin EtkinDegil class.error
Etkin      5          0          0
EtkinDegil 0          0         NaN
> ### Test seti icin model ile tahmin
> predictions2 <- predict(model, data1[-test,], type="class")
> predictions2
      1      2      3      4      5      6
      Etkin Etkin Etkin EtkinDegil Etkin Etkin
Levels: Etkin EtkinDegil
> ptestdata<-data1[-test,]      ### test veryi tanima
> mean(ptestdata$y==predictions2)### Dogruluk orani hesaplama
[1] 0.8333333

```

Tablo 16...devamı

```

> ### Tum veri icin model ile tahmin etme
> predictions3 <- predict(model, data1, type="class")
> predictions3      ### Dogruluk orani hesaplama
      1      2      3      4      5      6      7
      Etkin Etkin Etkin Etkin EtkinDegil Etkin Etkin
Etkin EtkinDegil
      10      11      12
      Etkin Etkin Etkin
Levels: Etkin EtkinDegil
> mean(data1$y==predictions3)
[1] 0.9166667
>

```

BCC etkinlik skorları hedef değeri alındığında Random forest algoritması %92 oranında doğru bir sınıflandırma yapmıştır.

Tablo 17. “532623” Tohum ile Süper Etkinlik Skorlarını Kullanan Random Forest Algoritmasının R Kodları

```

R 4.3.1 ~ /
> set.seed(532623)
> dadt02 <- read_excel("C:/Users/PC/Desktop/Data22DT.xlsx")
New names:
* `` -> `...11`
> data1<-dadt02[1:7]          ### Degiskenler almak
> data1["y"]<- dadt02$Sup     ### VZA dan Super etkinlink degisken
> data1$y<- as.factor(data1$y) ### Etkinlik degisken faktor tanima
> test<- sample(1:nrow(data1),6) ### Egitim kumesi tanima
> data<- data1[test,]         ### Veryidan test KVBlar almak
> ### Modeli ulusturma
> model <- randomForest(y ~ ., data=data, proximity=TRUE, ntree=1000)
> model                       ### Modeli cagirma

Call:
randomForest(formula = y ~ ., data = data, proximity = TRUE,      ntree = 1000)
Type of random forest: classification
Number of trees: 1000
No. of variables tried at each split: 2

OOB estimate of error rate: 0%
Confusion matrix:
      Super SuperDegil class.error
Super      0          0         NaN
SuperDegil 0          5          0
> ### Test seti icin model ile tahmin
> predictions2 <- predict(model, data1[-test,], type="class")
> predictions2
      1      2      3      4      5      6
SuperDegil SuperDegil SuperDegil SuperDegil SuperDegil SuperDegil
Levels: Super SuperDegil
> ptestdata<-data1[-test,]      ### test veryi tanima
> mean(ptestdata$y==predictions2)### Dogruluk orani hesaplama
[1] 0.6666667

```

Tablo 17...devamı

```

> ### Tum veri icin model ile tahmin etme
> predictions3 <- predict(model, data1, type="class")
> predictions3          ### Dogruluk orani hesaplama
      1      2      3      4      5      6      7
Super SuperDegil SuperDegil SuperDegil SuperDegil SuperDegil SuperDegil
      8      9     10     11     12
SuperDegil SuperDegil SuperDegil SuperDegil SuperDegil
Levels: Super SuperDegil
> mean(data1$y==predictions3)
[1] 0.8333333
>

```

Süper etkinlik skorları hedef değeri alındığında Random forest algoritması %83 oranında doğru bir sınıflandırma yapmıştır.

Tablo 18. “532623” Tohum ile Random Forest Algoritması Modellerinin Sonuçları

Model	Eğitim Kümesi Doğruluk Oranı	Test Kümesi Doğruluk Oranı	Tüm Veri İçin Tahmin Edilen Doğruluk Oranı
CCR Model	83.33%	66.67%	83.33%
BCC Model	100%	83,33%	91,67%
Süper Etkinlik	100%	66,67%	83,33%
Ortalama	94,44%	72,22%	86,11%

Tohumları değiştirerek, verilerin makine öğrenimi alanındaki eğitim ve test kümelerine bölünmesini etkiler. Farklı veri düzenlemeleri, farklı sonuçlar üretebilir, bu da model performansının değerlendirmesini etkileyebilir.

Buradaki amacımız, kategorilerinde dengeli bir eğitim seti oluşturacak her model için uygun bir tohum seçmektir. Örneğin, yalnızca verimli karar birimlerden oluşan bir eğitim seti, modelin eğitilmesinde çok büyük eksiklik olur ve zayıf model oluşumuna yol açabilir.

Aşağıdaki Tablo.19, farklı tohumlar kullanarak farklı modellerin sonuçları göstermektedir:

Tablo.19. Random Forest Algoritması Modellerin Farklı tohumlarla Sonuçları

Model	Tohum	OOB Hata	Test Kümesi Doğruluk Oranı	Tüm Veri İçin Tahmin Edilen Doğruluk Oranı
CCR Model	532623	16.67%	66.67%	83.33%
CCR Model	7159	0%	83.33%	91.67%
BCC Model	532623	0%	83,33%	91,67%
BCC Model	8429	83.33%	100%	91.67%
Süper Etkinlik	532623	0%	66,67%	83,33%
Süper Etkinlik	127	50%	100%	100%

Her model için uygun tohumlar kullanarak CCR modelinde eğitim seti için doğruluk oranı %83,33 ve tüm veri için %91,67 bulunmuştur. BCC modelinde eğitim seti için doğruluk oranı eğitim seti için %100 ve tüm veri için %91,67'ye ulaşmıştır.

Süper etkinlik modeli için hem eğitim setinde hem de test setinde tamamında %100'e ulaşılmıştır.

Modeller ve çıkarıldığı sonuçlar belirli bir tohum için benzersiz değildir, eğitim setinin oluşturulmasında benzer sonuçlar nedeniyle aynı modeller ve sonuçlar verecek birçok farklı tohum bulmak mümkündür.

OOB hatası ve Test seti hatası modelin performansını ölçer, fakat OOB hatası eğitim setinde kullanılmamıştır. Birimlerin tahmin hatası hesaplanmıştır. Test seti hatası, tamamen bağımsız bir veri seti içerdiğinden genellikle sonuçları daha güvenilir kabul edilmektedir.

Eğitim ve test setlerini ayrı ayrı kullanmak, modelin ne kadar iyi çalıştığını daha net ve gerçekçi bir şekilde göstermeye yardımcı olmaktadır. Baştan tüm veriyi kullanarak yalnızca OOB hatasına bakmak yerine, ayrı ayrı setler kullanmak daha doğru sonuçlar vermektedir. Tüm veriyi kullanmak daha iyimser sonuçlar elde etmemize neden olabilir. Buna rağmen ayrı setler kullanıldığında, tüm veriyi test için kullanarak yapılan değerlendirme, eğitim setinde olabilecek hataları da belirlemede yardımcı olabilmektedir.

Makine öğrenmesi Random Forest algoritmasında eğitim kümesi (testing set) için en yüksek doğruluk düzeyi BCC model ile Süper etkinlik modelin etkinlik skorları kullanıldığında elde edilmiştir. CCR modelin etkinlik skorlarıyla yapılan sınıflandırmada eğitim kümesinin doğruluk düzeyi %83.33 olarak gerçekleşmiştir. Üç modelin eğitim kümesi için ortalama doğruluk düzeyi %94.44'tür.

Tüm veri için doğruluk düzeyine bakıldığında, süper etkinlik model %100'lük en yüksek doğruluk oranı ile ilk sırada yer almaktadır. CCR modeli ve BCC modeli doğruluk düzeyleri aynı oranda %91.67 düzeyinde bulunmuştur. Üç modelin ortalama doğruluk düzeyi ise %96 olarak gerçekleşmiştir.

Random Forest algoritmasıyla elde edilen doğruluk düzeyleri yani birimlerin doğru sınıflandırılma düzeylerinin ölçeğe göre değişken getirili BCC modelinde en yüksek düzeyde bulunması, firmaların kaynak kullanımlarının optimal düzeyde olmadığını işaret etmektedir. Firmaların kaynak dağılımlarını optimal seviyelere çıkarabilmeleri ise VZA ile elde edilen etkinleştirme önerilerine dikkate almalarıyla

sağlanabilir. Bu nedenle bir sonraki bölümde veri zarflama analizinin BCC model bulgularından elde edilen etkinleştirme önerilerine yer verilmiştir. CCR ve Süper etkinlik analiz bulgularına çalışmanın sonunda Ekler kısmında yer verilmiştir.

3.6. BCC Modelin Etkinleştirme Önerileri

BCC modele göre Firma₂ %74.3, Firma₃ %33.2, Firma₅ %73.8, Firma₆ %25.8, Firma₈ %73, Firma₁₂ %17.7 oranında etkin bulunmuşlardır, yani etkin değildirler. Aşağıda bu firmalardan sembolik olarak sadece ikisi için etkinleştirme önerilerine yer verilmiştir. Diğerleri için hesaplanan etkinlik önerileri Ekler kısmında verilmiştir.

Tablo 20. FİRMA₁₂ BCC Etkinlik Analizi

17.68% Firma12			
Potential Improvements			
Variable	Actual	Target	Potential Improvement
I1	541.68	61.50	-88.65 %
I2	4.10	0.51	-87.55 %
I3	27590000.00	4878090.31	-82.32 %
O1	31523969.00	31523969.00	0.00 %
O2	4950000.00	4950000.00	0.00 %
O3	68910.00	172064.76	149.69 %
O4	430000.00	671107.92	56.07 %
Peer Contributions			
Firma1	I1		3.33 %
Firma1	I2		9.89 %
Firma1	I3		17.31 %
Firma1	O1		17.00 %
Firma1	O2		44.97 %
Firma1	O3		54.02 %
Firma1	O4		20.12 %
Firma11	I1		96.54 %
Firma11	I2		89.96 %
Firma11	I3		71.02 %
Firma11	O1		20.69 %
Firma11	O2		54.72 %
Firma11	O3		43.91 %
Firma11	O4		79.59 %
Firma9	I1		0.13 %
Firma9	I2		0.15 %
Firma9	I3		11.67 %
Firma9	O1		62.31 %
Firma9	O2		0.31 %
Firma9	O3		2.07 %
Firma9	O4		0.29 %

Analiz Yorumu: BCC modeline göre Firma₁₂ etkin değildir. Etkin olabilmesi için I1'i hedef değeri 61,33 , I2'yi hedef değeri 0,51 ve I3'ü 5444808,72 değeri ulaşmamaya

çalışmalıdır. Örnek alması gereken etkin firmaların konveks kombinasyonu, I1 için %3,33 oranında Firma9, %96,54 oranında Firma1, %0,13 oranında Firma10'dur. I2 için %9,89 Firma9, %89,96 oranında Firma1, %0,15 Firma10'dur. I3 için %17,31 Firma9, %71,02 oranında Firma1, %11,67 oranında Firma10'dur.

Tablo 21. FİRMA3 BCC Etkinlik Analizi

38.16% Firma3				Peers: 3
Potential Improvements				References: 0
Variable	Actual	Target	Potential Improvement	
I1	21.83	8.33	-61.84 %	
I2	7.62	0.07	-99.03 %	
I3	297090000.00	113357485.82	-61.84 %	
O1	1173279088.00	1173279088.00	0.00 %	
O2	7310000.00	33692808.94	360.91 %	
O3	228630.00	2978158.08	1202.61 %	
O4	340000.00	7116592.23	1993.12 %	
Peer Contributions				
Firma11	I1		53.20 %	
Firma11	I2		47.14 %	
Firma11	I3		16.40 %	
Firma11	O1		0.34 %	
Firma11	O2		96.56 %	
Firma11	O3		89.19 %	
Firma11	O4		85.17 %	
Firma4	I1		0.68 %	
Firma4	I2		1.46 %	
Firma4	I3		56.36 %	
Firma4	O1		18.28 %	
Firma4	O2		1.22 %	
Firma4	O3		4.94 %	
Firma4	O4		13.49 %	
Firma7	I1		46.11 %	
Firma7	I2		51.40 %	
Firma7	I3		27.24 %	
Firma7	O1		81.37 %	
Firma7	O2		2.22 %	
Firma7	O3		5.87 %	
Firma7	O4		1.34 %	

Analiz Yorumu: Firma3 etkin değildir. Etkin olabilmesi için I1'i hedef değeri 8.33 değeri, I2'y hedef değeri 0,07 değeri ve I3'ü 113357485,82 hedef değeri ulaşmaya çalışmalıdır. Örnek alması gereken etkin firmaların konveks kombinasyonu, I1 için %53,20 oranında Firma4, %0,68 oranında Firma7, %46,11 oranında Firma10'dur. I2 için %47,14 Firma4, %1,46 oranında Firma7, %51,40 oranında Firma10'dur. I3 için %16,40 Firma4, %56,36 oranında Firma7, %27,24 oranında Firma10'dur.

SONUÇ

Türk demir-çelik endüstrisinde İstanbul borsasına kote edilmiş 12 demir-çelik firmasının 2021 yılı verileri kullanılarak Random Forest makine öğrenmesi algoritması uygulanmıştır. Demir-çelik endüstrisindeki üretim sürecini tanımladığı düşünülen 27 değişken, Shannon entropi yöntemiyle boyut indirgenmesine tabi tutularak 7'ye düşürülmüştür. Bu değişkenlerden, kavramsal olarak azaltılması gerekenler girdileri, arttırılması gerekenler de çıktıları temsil etmek üzere, bu girdi ve çıktılarıyla veri zarflama analizi uygulanarak Random Forest makine öğrenmesi algoritması için hedef değerler (etkinlik skorları sayısal değer olarak değil, etkin veya etkin değil şeklinde alınmıştır) belirlenmiştir. Çalışmada kullanılan veriler gerçek veriler olmasına rağmen, deneysel olarak entropi ile elde edilen değişken kümesi üzerinden yapılan bir kesit analize dayandığı için, VZA ile elde edilen etkinlik skorları; firmalar için mutlak bir başarı sıralaması anlamına gelmemektedir. Buna rağmen, yanlış bilgilendirmeye sebebiyet vermemek amacıyla, çalışmaya dâhil edilen firma isimleri verilmiştir, fakat etkinlik skorları firma isimleri verilmeden, kodlanarak gizlenmiştir. Tekrar belirtmek gerekir ki söz konusu bulgular mutlak değil, deneyseldir.

Makine öğrenmesi Random forest algoritması ile elde edilen bulgulara göre, en yüksek doğruluk düzeyine, ölçeğe göre değişken getirili BCC model ve Süper etkinlik modellerinde ulaşılmıştır. Ölçeğe göre sabit getirili CCR model kullanılarak yapılan Random Forest uygulamalarının görece daha düşük bir doğruluk düzeyine ulaştığı görülmüştür. BCC etkinlik skorları hedef değer alındığında Random Forest algoritmasının %100 oranında, Süper etkinlik skorları hedef değer alındığında %100 oranında, CCR etkinlik skorları hedef değer alındığında %83.33 oranında doğru bir sınıflandırma yaptığı görülmüştür. Söz konusu bulgulara bakıldığında, demir-çelik sektöründe bir ölçek sorunu olduğu söylenebilir. Firmaların yarısından daha azı (yaklaşık %42'si) optimal ölçekte üretim yapmaktadırlar.

Random Forest makine öğrenmesi algoritması ile gerçekleştirilen bu çalışmada karşılaşılan birtakım hususlara dikkat çekmenin faydalı olacağı düşünülmektedir. Örneğin, tohum (seed) kullanmada her defasında farklı modeller ve farklı sonuçlarla

karşılaşmanın olası olduğu görülmüştür. Veri kümesinin yeterince büyük olmaması, modelin öğrenme düzeyini olumsuz etkilemektedir. Bu sorunun önüne geçebilmek adına başta sınıflandırma değişkeni için (CCR etkinlik, BCC etkinlik ya da Süper etkinlik) eğitim veri kümesinde (training data) çok sayıda deneme yapılarak en çok çeşitliliği sağlayan tohum “532623” seçilmiştir. Daha sonra her model için çeşitliliği daha çok sağlayabilecek tohumlar denenmiştir. R’da oluşturulan standart ağaç sayısı 500 olup, bu sayının arttırılması sadece CCR modelde bir miktar doğruluk düzeyi artışı sağlamış, BCC ve Süper etkinlik modellerinin doğruluk düzeylerinde herhangi bir artış olmamıştır.



KAYNAKÇA

- Shannon,C.E.: 1948 "A Mathematical Theory Of Communication Reprinted With Corrections From", **The Bell System Technical Journal** Vol. 27 , Pp. 379–423, 623–656.
<https://people.math.harvard.edu/~ctm/home/text/others/shannon/entropy/entropy.pdf>
- Banker,R.D.Charnes,A.& "Some Models For Estimation Technical And Scale Inefficiencies In Data Envelopment Analysis", **Management Science Journal**, USA, Pp. 1078-1092.
Cooper,W.W.: 1984 <http://www.jstor.org/stable/2631725>
- Golomb, S. W., Peile, R. **Basic Concepts In Information Theory And Coding**
E., Scholtz, R. A.: 1994 **The Adventures Of Secret Agent 00111**, Springer Science+Business Media, New York,
<https://link.springer.com/book/10.1007/978-1-4757-2319-9>
- C.A. Knox Lovell ve T. "Units invariant and translation invariant DEA models", **Elsevier Science B.V**,
Pastor: 1995 [https://doi.org/10.1016/0167-6377\(95\)00044-5](https://doi.org/10.1016/0167-6377(95)00044-5)
- Thanassoulis, E.: 2001 **Introduction To The Theory And Application Of Data Envelopment Analysis**, Springer Science+Business Media, New York,
<https://link.springer.com/book/10.1007/978-1-4615-1407-7>
- Zhu, J.: 2003 **Quantitative Models For Performance Evaluation And**

Benchmarking Data Envelopment Analysis With Spreadsheets And DEA Excel Solver Chapter 1 Basic DEA Models, Springer Science+Business Media, New York.

<https://link.springer.com/book/10.1007/978-1-4757-4246-6>

Banker ve Chang: 2006 "The super-efficiency procedure for outlier identification, not for ranking efficient units",
European Journal of Operational Research 175, Pp 1311- 1320,
<https://doi.org/10.1016/j.ejor.2005.06.028>

Cooper, W. W., Seiford, L. M., Tone, K.: 2007 **Data Envelopment Analysis A Comprehensive Text With Models, Applications, DEA-Solver Software**, Springer Science+Business Media, New York,
<https://link.springer.com/book/10.1007/978-0-387-45283-8>

Golany, B., Roll, Y.: 1989 "An Application Procedure for DEA", **OMEGA Int. Journal. of Mgmt Sci.**, Vol. 17, No. 3, pp. 237-250. [https://doi.org/10.1016/0305-0483\(89\)90029-7](https://doi.org/10.1016/0305-0483(89)90029-7)

Khezrimotlagh, D.: 2014 "How To Deal With Numbers Of Decision Making Units And Variables In Data Envelopment Analysis". Department Of Applied Statistics, Faculty Of Economics And Administration, University Of Malaya, Kuala Lumpur, pp. 1-11.

Cooper, W. W., Seiford, L. M., Zhu, J.: 2011 **Handbook On Data Envelopment Analysis**, Springer Science+Business Media, New York,

<https://link.springer.com/book/10.1007/978-1-4419-6151-8>

- Azar, A. T., Elshazly, H. I., Hassanien, A. E., Elkorany, A. M.: 2014 "A Random Forest Classifier For Lymph Diseases", **Computer Methods And Programs In Biomedicine journal**, Volume 113, Issue 2, February 2014, Pages 465-473.
- Nizam, H., Akın, S. S.: 2014 "Sosyal Medyada Makine Öğrenmesi İle Duygu Analizinde Dengeli ve Dengesiz Veri Setlerinin Performanslarının Karşılaştırılması", **XIX. Türkiye’de İnternet Konferansı**, inet-tr.org.tr
- Onan, A.: 2014 "Şirket İflaslarının Tahmin Edilmesinde Karar Ağacı Algoritmalarının Karşılaştırmalı Başarım Analizi, **Bilişim Teknolojileri Dergisi**, Cilt: 8, Sayı: 1, ss:9-19.
- Putnick DL, Bornstein MH.: 2016 "Measurement Invariance Conventions and Reporting: The State of the Art and Future Directions for Psychological Research". **Development Review** Sep;V 41:71-90
<https://doi.org/10.1016/j.dr.2016.06.004>
- Bilgin, M.: 2017 "Gerçek Veri Setlerinde Klasik Makine Öğrenmesi Yöntemlerinin Performans Analizi", Akademik Bilişim, Aksaray, Turkey.
<https://ab.org.tr/ab17/bildiri/101.pdf>
- Çınaroğlu, S., Başer, O.: 2017 "Kamu Hastane Birliklerinin Teknik Verimliliklerinin Veri Zarflama Analizi ve Makine Öğrenmesi

- Teknikleri Kullanılarak İncelenmesi", **Ekonomik Yaklaşım derneği**, 28(104): 81-120.
<http://dx.doi.org/10.5455/ey.36122>
- Aydın, C.: 2018 "Makine Öğrenmesi Algoritmaları Kullanılarak İtfaiye İstasyonu İhtiyacının Sınıflandırılması", **Avrupa Bilim Ve Teknoloji Dergisi** Sayı 14, S.169-175.
- Dogru, N., Subasi, A.: 2018 "Traffic Accident Detection Using Random Forest Classifier", **15th Learning And Technology Conference (L&T)**, Jeddah, Saudi Arabia, 2018, Pp. 40-45, Doi: 10.1109/LT.2018.8368509.
- Alinezhad, A. ve Khalili, J.: 2019 **New methods and applications in multiple attribute decision making (MADM)**, Springer Science+Business Media, New York, <https://link.springer.com/book/10.1007/978-3-030-15009-9>
- Schonlau, M., Zou, R. Y.: 2020 "The Random Forest Algorithm For Statistical Learning", **The Stata Journal** 20, Number 1, Pp. 3–29.
- Vukic, L., Jugovic, T. P., Guidi, G., Oblak, R.: 2020 "Model of Determining the Optimal, Green Transport Route among Alternatives: Data Envelopment Analysis Settings", **MDPI Journal of Marine Science and Engineering** ISSN 2504-4990 Switzerland, <https://doi.org/10.3390/jmse8100735>
- Brownlee, J.: 2020 "Train-Test Split for Evaluating Machine Learning

Algorithms", **Machine Learning Mastery**

<https://machinelearningmastery.com/train-test-split-for-evaluating-machine-learning-algorithms/>

Sönmezler, G., Gündüz,
İ. O.: 2021

"Covid-19 Pandemi Sürecinin BİST-30
Hisse Senetlerine Etkilerinin Karışıklık Matrisi İle
Analizi", **Maliye Ve Finans Yazıları** - 2021 - (115), 51
– 70.

Nacar, E. N., Erdebili,
B.: 2021

"Makine Öğrenmesi Algoritmaları İle Satış
Tahmini", **Journal Of Industrial Engineering** 32(2),
307-320.

Turkish Steel Catalogue:
2022

"CIB çelik ihracatçıları birliği 2022 katalog
https://www.cib.org.tr/files/Doc/files/KATALOG2022_ENG.pdf Erişim 26.7.2023

İhracat 2022 Raporu

"TİM Türkiye ihracat meclisi, İstanbul,
https://tim.org.tr/files/downloads/Strateji_Raporlari/TIM%20I%CC%87hracat%20Raporu%202022.pdf Erişim 26.7.2023

Türkiye Çelik Haritası

"Türkiye Çelik Üreticileri Derneği, Ankara,
<https://celik.org.tr/harita/> Erişim 26.7.2023

Turkish Steel Catalogue:
2022

"CIB çelik ihracatçıları birliği 2022 katalog
https://www.cib.org.tr/files/Doc/files/KATALOG2022_ENG.pdf Erişim 26.7.2023.

Akbulut, Y. S.: 2021

**Gayrimenkul Yatırım Ortaklığı Şirketlerinin
Finansal**

Etkinliklerinin Vza Yöntemi İle Ölçülmesi. Bist' De Bir Uygulama, Aydın Adnan Menderes Üniversitesi, Sosyal Bilimler Enstitüsü, İşletme Ana Bilim Dalı, Muhasebe Finansman Bilim Dalı, Y1 Tezi.

Biçen, Ç.: 2010

Otomotiv Sektöründe Veri Zarflama Analizi İle Finansal Etkinlik Ölçümü, Yıldız Teknik Üniversitesi, Sosyal Bilimler Enstitüsü, İşletme Bölümü, İşletme Ana Bilim Dalı, İşletme Yönetimi Bilim Dalı, Y1 Tezi.

Özkan, T.: 2019

"Türk Bankacılık Sisteminde Veri Zarflama Tekniği İle Banka Etkinliğinin Ölçülmesi", Araştırma Makalesi, **Üçüncü Sektör Sosyal Ekonomi Dergisi**,54(3) 2019 1511-1529.

Özel, N. G., Şahin, İ. E. & Göröl, R.: 2017

"Türk Bankacılık Sektöründe Etkinlik Ve Verimlilik Analizinin Veri Zarflama Yöntemi İle İncelenmesi: 2013-2015 Dönemi Uygulaması". **Sosyal Ekonomik Araştırmalar Dergisi**, 30. Yıl Özel Sayısı, 85-100 . DOI: 10.30976/susead.348135.

Karaman, G.: 2019

Borsa İstanbul'da İşlem Gören Metal Ana Sanayi Firmalarının Ekonomik Etkinlik Analizi, Nevşehir Hacı Bektaş Veli Üniversitesi Sosyal Bilimler, YL Tez.

Peykani, P., Seyed Esmaceli, F.S., Mirmozaffari, M., Jabbarzadeh, A., Khamechian, M.:2022

"Input/Output Variables Selection in Data Envelopment Analysis: A Shannon Entropy Approach". **Machine Learning Knowledge extraction journal** 4, 688-699. <https://doi.org/10.3390/make4030032>

Asar, S. S.: 2014

Using Principal Component Analysis In Data Envelopment Analysis, Fen Bilimleri Enstitüsü İstatistik Bölümü Hacettepe Üniversitesi ,YL Tez.

Üstündağ, E.: 2009

Veri Zarflama Analizi İle Verimliliğin Değerlendirilmesi: Çimento Sektörü Üzerine Bir Uygulama, Selçuk Üniversitesi, Fen Bilimleri Enstitüsü, İstatistik Ana Bilim Dalı, YL Tez.

Kasemi, E. C.: 2021

VZA ve VZAHP Bütünleşik Metodu İle Verimlilik Analizi: Kosova Gıda Sektörü Uygulaması, Balıkesir Üniversitesi, Fen Bilimleri Enstitüsü Doktora Tezi.

Veri Kaynakları

<https://tr.tradingview.com/markets/stocks-turkey/sectorandindustry-industry/steel/>

<https://tr.investing.com/equities/>

<https://fintables.com/>

<https://www.cemtas.com.tr/sayfa/yatirimci-iliskileri-detay/1170/ozet-mali-tablolar>

<https://halkyatirim.com.tr/skorkart/>

<https://halkyatirim.com.tr/skorkart/>

EKLER

EK1: Girdi Değişkenleri Grup 1

	O-HS	O-BK	O-FK	O-VK	O-NK	O-FVK	O-DOV	O-ns	O-5TV	O-KDO	O-OZG
Firma4	17,000,334	12,331,070	5,832,709	48,460,052	44,498,987	6,820,000	196,719,393	17,000,000	0.00	0.00	62.22
Firma5	1,340,579,006	128,098,914	98,171,573	46,524,636	35,889,748	98,387,336	281,600,944	1,341,630,874	2.36	4.05	76.88
Firma9	447,792,698	47,339,745	32,990,518	6,970,970	24,187,748	33,790,000	169,347,244	24,187,748	1.01	11.47	33.38
Firma1	119,009,315	24,120,000	7,825,951	2,751,919	2,265,525	12,900,000	63,594,561	119,010,000	0.00	0.00	23.63
Firma12	686,839,428	123,180,878	31,523,969	10,202,725	7,572,680	63,930,000	480,915,117	686,840,000	8.93	16.31	49.73
Firma2	1,510,820,977	444,725,171	481,691,327	501,165,338	403,327,293	383,660,000	932,676,482	1,510,820,000	1.54	2.94	66.69
Firma11	68,227,404,000	25,516,411,000	24,871,472,000	23,916,713,000	16,078,558,000	26,885,849,000	65,693,526,000	68,227,400,000	12.89	82.74	23.92
Firma8	36,499,516,000	13,718,772,000	13,009,462,000	10,157,036,000	6,468,354,000	13,920,852,000	25,998,080,000	36,499,520,000	13.22	178.28	18.23
Firma3	10,823,179,851	1,128,257,646	1,173,279,088	1,209,234,578	782,513,004	1,366,704,711	4,317,512,549	9,694,922,205	0.00	0.00	28.50
Firma7	4,149,686,000	2,510,962,000	1,995,840,000	3,949,170,000	2,909,256,000	2,164,860,000	9,753,790,000	4,149,690,000	0.00	92.90	28.82
Firma10	14,251,188,133	4,782,356,123	4,797,420,845	4,467,032,440	3,852,707,219	4,908,895,714	12,337,672,778	14,764,791,145	0.00	10.97	49.09
Firma6	678,717,588	180,269,251	160,015,506	146,232,953	109,141,188	160,230,000	502,734,732	678,720,000	0.71	0.00	61.68

EK2: Girdi Değişkenleri Grup 2

	O-ADH	O-SDH	O-CP	O-NKP	O-LO	O-CO
Firma4	0.23	46.91	1620000.00	3830000.00	4.82	8.74
Firma5	5.55	9.71	11360000.00	465720.00	0.51	0.38
Firma9	3.66	13.70	8440000.00	578400.00	0.87	0.83
Firma1	0.99	3.25	955950.00	91670.00	0.46	0.64
Firma12	2.16	4.95	2940000.00	68910.00	0.43	0.43
Firma2	1.96	3.35	6280000.00	2090000.00	2.29	2.29
Firma11	0.98	2.58	9910000.00	1600000.00	0.82	0.88
Firma8	1.07	2.92	13660000.00	1560000.00	0.62	0.78
Firma3	2.43	7.31	10910000.00	228630.00	0.63	0.34
Firma7	0.48	3.83	2130000.00	1370000.00	7.00	8.93
Firma10	1.19	3.76	5210000.00	878090.00	0.52	0.48
Firma6	1.80	4.76	19200000.00	2520000.00	0.74	0.78

EK3: Girdi Değişkenleri Grup 1

	I-DUV	I-KVY	I-UVY	I-TEN	I-SEG	I-YFN	I-FFN
Firma4	130663102	17684163	4583238	0.00	2038.00	26,800,000	170,180,000
Firma5	36881813	2295840	6073955	241.43	3.49	7,722,490	9,430,000
Firma9	66057939	47368065	8537092	88.47	5.67	5,800,000	59,280,000
Firma1	109092112	67419204	28206345	24.86	4.88	4,640,000	9,370,000
Firma12	96797650	494935687	31315297	214.72	23.09	27,590,000	130,020,000
Firma2	257367553	194299654	20253497	491.50	50.52	67,200,000	26,670,000
Firma11	60748771000	23477310000	18417383000	28960.20	386205.00	7,520,856,000	3,425,760,000
Firma8	29757191000	8592839000	8005263000	13919.00	226098.00	2,449,550,000	1,498,130,000
Firma3	3433596246	6287039528	663610156	1812.21	10622.00	297,090,000	1,634,700,000
Firma7	2527775000	1194037000	424009000	450.32	61586.00	594,470,000	1,403,280,000
Firma10	9477296747	9590469397	2262931458	3196.54	29912.00	155,180,000	1,004,980,000
Firma6	37235714	234473563	83239976	132.46	2248.00	22,310,000	129,540,000

EK4: Girdi Değişkenleri Grup 2

	I-TBV	I-BAV	I-BOZS
Firma4	6.39	0.04	0.05
Firma5	121.98	0.50	1.86
Firma9	12.56	0.24	0.31
Firma1	71.05	0.23	0.55
Firma12	541.68	0.36	4.10
Firma2	17.62	0.18	0.06
Firma11	24.90	0.16	0.25
Firma8	21.83	0.14	0.22
Firma3	21.83	0.77	7.62
Firma7	0.53	0.00	0.01
Firma10	19.30	0.08	0.19
Firma6	152.61	0.28	0.57

EK5: Seçilen Değişkenlerle VZA İçin Veri Kümesi

	I-YFN	I-TBV	I-BOZS	O-FK	O-SDH	O-NKP	O-CO
Firma4	26,800,000	6.39	0.05	5832709	4691000 0	383000 0	874000 0
Firma5	7,722,490	121.9 8	1.86	98171573	9710000	465720	380000
Firma9	5,800,000	12.56	0.31	32990518	1370000 0	578400	830000
Firma1	4,640,000	71.05	0.55	7825951	3250000	91670	640000
Firma1 2	27,590,000	541.6 8	4.10	31523969	4950000	68910	430000
Firma2	67,200,000	17.62	0.06	481691327	3350000	209000 0	229000 0
Firma1 1	7,520,856,0 00	24.9	0.25	248714720 00	2580000	160000 0	880000
Firma8	2,449,550,0 00	21.83	0.22	130094620 00	2920000	156000 0	780000
Firma3	297,090,000	21.83	7.62	117327908 8	7310000	228630	340000
Firma7	594,470,000	0.53	0.01	199584000 0	3830000	137000 0	893000 0
Firma1 0	155,180,000	19.3	0.19	479742084 5	3760000	878090	480000
Firma6	22,310,000	152.6 1	0.57	160015506	4760000	252000 0	780000

EK 6: Girdi Yönelimli CCR Model Bulguları

100.00% Firma 4				Peers: 0 References: 6
<i>Potential Improvements</i>				
Variable	Actual	Target	Potential Improvement	
I1	6.39	6.39	0.00 %	
I2	0.05	0.05	0.00 %	
I3	26800000.00	26800000.00	0.00 %	
O1	5832709.00	5832709.00	0.00 %	
O2	46910000.00	46910000.00	0.00 %	
O3	3830000.00	3830000.00	0.00 %	
O4	8740000.00	8740000.00	0.00 %	

Firma4 etkindir, girdilerinden herhangi bir azaltmaya ihtiyaç bulunmamaktadır.

87.50% Firma 5				Peers: 3 References: 0
<i>Potential Improvements</i>				
Variable	Actual	Target	Potential Improvement	
I1	121.98	8.28	-93.21 %	
I2	1.86	0.20	-89.39 %	
I3	7722490.00	6757411.16	-12.50 %	
O1	98171573.00	98171573.00	0.00 %	
O2	9710000.00	9710000.00	0.00 %	
O3	465720.00	465720.00	0.00 %	
O4	380000.00	733149.54	92.93 %	

Firma5 etkin değildir. Etkin olabilmesi için I1'i %93,21 oranında, I2'yi %89,39 oranında, I3'ü %12.60 oranında azaltmalıdır.

100.00% Firma 9				Peers: 0 References: 3
<i>Potential Improvements</i>				
Variable	Actual	Target	Potential Improvement	
I1	12.56	12.56	0.00 %	
I2	0.31	0.31	0.00 %	
I3	5800000.00	5800000.00	0.00 %	
O1	32990518.00	32990518.00	0.00 %	
O2	13700000.00	13700000.00	0.00 %	
O3	578400.00	578400.00	0.00 %	
O4	830000.00	830000.00	0.00 %	

Firma9 etkindir, girdilerinden herhangi bir azaltmaya ihtiyaç bulunmamaktadır.

47.40% Firma 1				Peers: 2 References: 0
Potential Improvements Variable	Actual	Target	Potential Improvement	
I1	71.05	0.50	-99.30 %	
I2	0.55	0.00	-99.28 %	
I3	4640000.00	2199544.24	-52.60 %	
O1	7825951.00	7825951.00	0.00 %	
O2	3250000.00	3436882.91	5.75 %	
O3	91670.00	281487.57	207.07 %	
O4	640000.00	640000.00	0.00 %	

Firma1 etkin değildir. Etkin olabilmesi için I1'i %99,30 oranında, I2'yi %99,28 oranında, I3'ü %52,60 oranında azaltmalıdır.

10.70% Firma 12				Peers: 3 References: 0
Potential Improvements Variable	Actual	Target	Potential Improvement	
I1	541.68	3.81	-99.30 %	
I2	4.10	0.09	-97.80 %	
I3	27590000.00	2951296.48	-89.30 %	
O1	31523969.00	31523969.00	0.00 %	
O2	4950000.00	4950000.00	0.00 %	
O3	68910.00	252718.38	266.74 %	
O4	430000.00	430000.00	0.00 %	

Firma12 etkin değildir. Etkin olabilmesi için I1'i %99,30 oranında, I2'yi %97,80 oranında, I3'ü %89,30 oranında azaltmalıdır.

75.86% Firma 2				Peers: 3 References: 0
<i>Potential Improvements</i>				
Variable	Actual	Target	Potential Improvement	
I1	17.62	4.89	-72.25 %	
I2	0.06	0.04	-24.14 %	
I3	67200000.00	50979126.48	-24.14 %	
O1	481691327.00	481691327.00	0.00 %	
O2	3350000.00	24487934.26	630.98 %	
O3	2090000.00	2090000.00	0.00 %	
O4	2290000.00	4881795.52	113.18 %	

Firma2 etkin değildir. Etkin olabilmesi için I1'i %72,25 oranında, I2'yi %24,14 oranında, I3'ü %24.14 oranında azaltmalıdır.

87.10% Firma 11				Peers: 2 References: 0
<i>Potential Improvements</i>				
Variable	Actual	Target	Potential Improvement	
I1	24.90	17.66	-29.08 %	
I2	0.25	0.22	-12.90 %	
I3	7520856000.00	6550627288.52	-12.90 %	
O1	24871472000.00	24871472000.00	0.00 %	
O2	2580000.00	34466670.19	1235.92 %	
O3	1600000.00	12638919.88	689.93 %	
O4	880000.00	76735866.81	8619.98 %	

Firma11 etkin değildir. Etkin olabilmesi için I1'i %29,08 oranında, I2'yi %12,90 oranında, I3'ü %12,90 oranında azaltmalıdır.

100.00% Firma 8				Peers: 0 References: 2
<i>Potential Improvements</i>				
Variable	Actual	Target	Potential Improvement	
I1	21.83	21.83	0.00 %	
I2	0.22	0.22	0.00 %	
I3	2449550000.00	2449550000.00	0.00 %	
O1	13009462000.00	13009462000.00	0.00 %	
O2	2920000.00	2920000.00	0.00 %	
O3	1560000.00	1560000.00	0.00 %	
O4	780000.00	780000.00	0.00 %	

Firma8 etkindir, girdilerinden herhangi bir azaltmaya ihtiyaç bulunmamaktadır.

23.65% Firma 3				Peers: 3 References: 0
<i>Potential Improvements</i>				
Variable	Actual	Target	Potential Improvement	
I1	21.83	5.16	-76.35 %	
I2	7.62	0.05	-99.35 %	
I3	297090000.00	70273858.31	-76.35 %	
O1	1173279088.00	1173279088.00	0.00 %	
O2	7310000.00	7310000.00	0.00 %	
O3	228630.00	780906.84	241.56 %	
O4	340000.00	1759069.95	417.37 %	

Firma3 etkin değildir. Etkin olabilmesi için I1'i %76,35 oranında, I2'yi %99,35 oranında, I3'ü %76.35 oranında azaltmalıdır.

100.00% Firma 7				Peers: 0 References: 4
<i>Potential Improvements</i>				
Variable	Actual	Target	Potential Improvement	
I1	0.53	0.53	0.00 %	
I2	0.01	0.01	0.00 %	
I3	594470000.00	594470000.00	0.00 %	
O1	1995840000.00	1995840000.00	0.00 %	
O2	3830000.00	3830000.00	0.00 %	
O3	1370000.00	1370000.00	0.00 %	
O4	8930000.00	8930000.00	0.00 %	

Firma7 etkindir, girdilerinden herhangi bir azaltmaya ihtiyaç bulunmamaktadır.

100.00% Firma 10				Peers: 0 References: 6
<i>Potential Improvements</i>				
Variable	Actual	Target	Potential Improvement	
I1	19.30	19.30	0.00 %	
I2	0.19	0.19	0.00 %	
I3	155180000.00	155180000.00	0.00 %	
O1	4797420845.00	4797420845.00	0.00 %	
O2	3760000.00	3760000.00	0.00 %	
O3	878090.00	878090.00	0.00 %	
O4	480000.00	480000.00	0.00 %	

Firma10 etkindir, girdilerinden herhangi bir azaltmaya ihtiyaç bulunmamaktadır.

100.00% Firma 6				Peers: 0
				References: 1
<i>Potential Improvements</i>				
Variable	Actual	Target	Potential Improvement	
I1	152.61	152.61	0.00 %	
I2	0.57	0.57	0.00 %	
I3	22310000.00	22310000.00	0.00 %	
O1	160015506.00	160015506.00	0.00 %	
O2	4760000.00	4760000.00	0.00 %	
O3	2520000.00	2520000.00	0.00 %	
O4	780000.00	780000.00	0.00 %	

Firma6 etkindir, girdilerinden herhangi bir azaltmaya ihtiyaç bulunmamaktadır.

EK 7: Girdi Yönelimli BCC Model Bulguları

100.00% Firma4				Peers: 0 References: 5
<i>Potential Improvements</i>				
Variable	Actual	Target	Potential Improvement	
I1	6.39	6.39	0.00 %	
I2	0.05	0.05	0.00 %	
I3	26800000.00	26800000.00	0.00 %	
O1	5832709.00	5832709.00	0.00 %	
O2	46910000.00	46910000.00	0.00 %	
O3	3830000.00	3830000.00	0.00 %	
O4	8740000.00	8740000.00	0.00 %	

Firma4 etkindir, girdilerinden herhangi bir azaltmaya ihtiyaç bulunmamaktadır.

100.00% Firma 5				Peers: 0 References: 1
<i>Potential Improvements</i>				
Variable	Actual	Target	Potential Improvement	
I1	121.98	121.98	0.00 %	
I2	1.86	1.86	0.00 %	
I3	7722490.00	7722490.00	0.00 %	
O1	98171573.00	98171573.00	0.00 %	
O2	9710000.00	9710000.00	0.00 %	
O3	465720.00	465720.00	0.00 %	
O4	380000.00	380000.00	0.00 %	

Firma5 etkindir, girdilerinden herhangi bir azaltmaya ihtiyaç bulunmamaktadır.

100.00% Firma9				Peers: 0 References: 4
<i>Potential Improvements</i>				
Variable	Actual	Target	Potential Improvement	
I1	12.56	12.56	0.00 %	
I2	0.31	0.31	0.00 %	
I3	5800000.00	5800000.00	0.00 %	
O1	32990518.00	32990518.00	0.00 %	
O2	13700000.00	13700000.00	0.00 %	
O3	578400.00	578400.00	0.00 %	
O4	830000.00	830000.00	0.00 %	

Firma9 etkindir, girdilerinden herhangi bir azaltmaya ihtiyaç bulunmamaktadır.

100.00% Firma1				Peers: 0
				References: 3
<i>Potential Improvements</i>				
Variable	Actual	Target	Potential Improvement	
I1	71.05	71.05	0.00 %	
I2	0.55	0.55	0.00 %	
I3	4640000.00	4640000.00	0.00 %	
O1	7825951.00	7825951.00	0.00 %	
O2	3250000.00	3250000.00	0.00 %	
O3	91670.00	91670.00	0.00 %	
O4	640000.00	640000.00	0.00 %	

Firma1 etkindir, girdilerinden herhangi bir azaltmaya ihtiyaç bulunmamaktadır.

19.73% Firma 12				Peers: 3
				References: 0
<i>Potential Improvements</i>				
Variable	Actual	Target	Potential Improvement	
I1	541.68	61.33	-88.68 %	
I2	4.10	0.51	-87.57 %	
I3	27590000.00	5444808.72	-80.27 %	
O1	31523969.00	31523969.00	0.00 %	
O2	4950000.00	4950000.00	0.00 %	
O3	68910.00	173973.42	152.46 %	
O4	430000.00	670216.06	55.86 %	

Firma12 etkin değildir. Etkin olabilmesi için I1'i %88,68 oranında, I2'yi %87,57 oranında, I3'ü %80,27 oranında azaltmalıdır.

100.00% Firma 2				Peers: 0
				References: 1
<i>Potential Improvements</i>				
Variable	Actual	Target	Potential Improvement	
I1	17.62	17.62	0.00 %	
I2	0.06	0.05	0.00 %	
I3	67200000.00	67200000.00	0.00 %	
O1	481691327.00	481691327.00	0.00 %	
O2	3350000.00	3350000.00	0.00 %	
O3	2090000.00	2090000.00	0.00 %	
O4	2290000.00	2290000.00	0.00 %	

Firma11 etkindir, girdilerinden herhangi bir azaltmaya ihtiyaç bulunmamaktadır.

100.00% Firma 11				Peers: 0 References: 1
<i>Potential Improvements</i>				
Variable	Actual	Target	Potential Improvement	
I1	24.90	24.90	0.00 %	
I2	0.25	0.25	0.00 %	
I3	7520856000.00	7520856000.00	0.00 %	
O1	24871472000.00	24871472000.00	0.00 %	
O2	2580000.00	2580000.00	0.00 %	
O3	1600000.00	1600000.00	0.00 %	
O4	880000.00	880000.00	0.00 %	

Firma11 etkindir, girdilerinden herhangi bir azaltmaya ihtiyaç bulunmamaktadır.

100.00% Firma 8				Peers: 0 References: 1
<i>Potential Improvements</i>				
Variable	Actual	Target	Potential Improvement	
I1	21.83	21.83	0.00 %	
I2	0.22	0.22	0.00 %	
I3	2449550000.00	2449550000.00	0.00 %	
O1	13009462000.00	13009462000.00	0.00 %	
O2	2920000.00	2920000.00	0.00 %	
O3	1560000.00	1560000.00	0.00 %	
O4	780000.00	780000.00	0.00 %	

Firma8 etkindir, girdilerinden herhangi bir azaltmaya ihtiyaç bulunmamaktadır.

38.16% Firma 3				Peers: 3 References: 0
<i>Potential Improvements</i>				
Variable	Actual	Target	Potential Improvement	
I1	21.83	8.33	-61.84 %	
I2	7.62	0.07	-99.03 %	
I3	297090000.00	113357485.82	-61.84 %	
O1	1173279088.00	1173279088.00	0.00 %	
O2	7310000.00	33692808.94	360.91 %	
O3	228630.00	2978158.08	1202.61 %	
O4	340000.00	7116592.23	1993.12 %	

Firma3 etkin değildir. Etkin olabilmesi için I1'i %61,84 oranında, I2'yi %99,03 oranında, I3'ü %61.84 oranında azaltmalıdır.

100.00% Firma 7				Peers: 0
				References: 2
Potential Improvements				
Variable	Actual	Target	Potential Improvement	
I1	0.53	0.53	0.00 %	
I2	0.01	0.01	0.00 %	
I3	594470000.00	594470000.00	0.00 %	
O1	1995840000.00	1995840000.00	0.00 %	
O2	3830000.00	3830000.00	0.00 %	
O3	1370000.00	1370000.00	0.00 %	
O4	8930000.00	8930000.00	0.00 %	

Firma7 etkindir, girdilerinden herhangi bir azaltmaya ihtiyaç bulunmamaktadır.

100.00% Firma 10				Peers: 0
				References: 3
Potential Improvements				
Variable	Actual	Target	Potential Improvement	
I1	19.30	19.30	0.00 %	
I2	0.19	0.19	0.00 %	
I3	155180000.00	155180000.00	0.00 %	
O1	4797420845.00	4797420845.00	0.00 %	
O2	3760000.00	3760000.00	0.00 %	
O3	878090.00	878090.00	0.00 %	
O4	480000.00	480000.00	0.00 %	

Firma10 etkindir, girdilerinden herhangi bir azaltmaya ihtiyaç bulunmamaktadır.

100.00% Firma 6				Peers: 0
				References: 1
<i>Potential Improvements</i>				
Variable	Actual	Target	Potential Improvement	
I1	152.61	152.61	0.00 %	
I2	0.57	0.57	0.00 %	
I3	22310000.00	22310000.00	0.00 %	
O1	160015506.00	160015506.00	0.00 %	
O2	4760000.00	4760000.00	0.00 %	
O3	2520000.00	2520000.00	0.00 %	
O4	780000.00	780000.00	0.00 %	

Firma6 etkindir, girdilerinden herhangi bir azaltmaya ihtiyaç bulunmamaktadır.