

## Türkiye'nin İthalat ve İhracatının Veri Madenciliği Yöntemleriyle Tahmini

İlker İbrahim AVŞAR<sup>1\*</sup>, Orhan ECEMİŞ<sup>2</sup>

<sup>1</sup>Osmaniye Korkut Ata Üniversitesi, Bahçe MYO, Lojistik Bölümü, 80500, Osmaniye

<sup>2</sup>Gaziantep Üniversitesi, Teknik Bilimler MYO, Bilgisayar Teknolojileri Bölümü, 27310, Gaziantep

<sup>1</sup><https://orcid.org/0000-0003-2991-380X>

<sup>2</sup><https://orcid.org/0000-0002-8270-0857>

\*Sorumlu yazar: iibrahimavsar@osmaniye.edu.tr

### Araştırma Makalesi

### ÖZ

#### Makale Tarihiçesi:

Geliş tarihi: 18.12.2022

Kabul tarihi: 14.03.2023

Online Yayınlanma: 04.12.2023

#### Anahtar Kelimeler:

İthalat

İhracat

Veri madenciliği

Weka

Ekonomik süreçlerin hızla değiştiği günümüz dünyasında dış ticaret her geçen gün daha da önemli hale gelmekte olup ülkeler ekonomik kalkınma açısından ihracatı önemli bir araç olarak görmektedirler. İhracata yönelik çabalar ülkelerin politika yapıcılarını açısından dikkatle izlenmektedir. Yine ülkelerin vaz geçemediği eylemlerden bir tanesi de ithalattır. Ülkeler açısından hem ithalat hem de ihracat verilerinin tahmin edilebilmesi yöneticilere avantaj sağlayacaktır. Çalışmada Türkiye'nin ithalat ve ihracat verileri tahmin edilmiştir. Tahmin için Türkiye'nin 1969-2022 yılları arasındaki verileri kullanılmıştır. Tahmin çalışmasında; Lineer Regresyon, MLPRegressor, Gaus Süreç Regresyon, RBF Regressor ve Destek Vektör Regresyon yöntemleri kullanılmıştır. İhracat tahmininde test verilerine göre %8,6987 MAPE değeriyle SmoReg algoritması en iyi sonucu vermiştir. İthalat tahmininde %7,6685 MAPE değeriyle yine SmoReg algoritması en düşük sonucu vermiştir. Çalışma MAPE kriterine göre Türkiye'nin ithalat verilerinin ihracat verilerine göre daha düşük hata oranıyla tahmin edilebildiğini göstermektedir.

## Forecast of Turkey's Import and Export by Data Mining Methods

### Research Article

### ABSTRACT

#### Article History:

Received: 18.12.2022

Accepted: 14.03.2023

Published online: 04.12.2023

#### Keywords:

Imports

Export

Data mining

Weka

In today's world, where economic processes are changing rapidly, foreign trade is gaining importance day by day. Countries consider exports as an important tool for economic development. Export efforts are carefully monitored by the countries' policy makers. One of the measures that countries cannot do without is imports. Estimating import and export data for countries is useful for managers. In this study, Turkey's import and export data were estimated. Data for the years 1969-2022 in Turkey were used for the estimation. Linear regression, MLPRegressor, Gaussian Process Regression, RBF Regressor and Support Vector Regression methods were used for estimation. The SmoReg algorithm provided the lowest result with a MAPE value of 8.6987% for the export estimate. The MLPRegressor algorithm provided the lowest result with a MAPE value of 7.6685 for the imports estimate. The study shows that according to the MAPE criterion, Turkey's import data can be estimated with a lower error rate than export data.

**To Cite:** Avşar İ., Ecemiş O. Türkiye'nin İthalat ve İhracatının Veri Madenciliği Yöntemleriyle Tahmini. Osmaniye Korkut Ata Üniversitesi Fen Bilimleri Enstitüsü Dergisi 2023; 6(3): 1890-1907.

## 1. Giriş

Küreselleşmeyle birlikte ülkeler kendisini ticari anlamda dünya pazarının içerisinde bulmuştur (Bayat, 2013). Bu bağlamda Türkiye özellikle 1980 yılı sonrasında uygulanan politikalarla birlikte dünya ekonomisiyle bütünleşmeye yönelik çalışmalar yürütmektedir. Türkiye, Avrupa Birliği (AB) ve Ortadoğu gibi farklı pazarlara yönelik ticari ilişkiler içerisinde. Ülke son yıllarda ihracatını artırma ve ithalatını düşürme noktasında çaba içerisinde (Aslan ve Terzi, 2007;Tütünsatar, 2022). Kısacası genel çerçeveye göz attığımızda Türkiye’de 1980 yılından bu yana döviz gelirlerini yükseltme ve dış ticaret açığını kapamaya yönelik politikalar devam etmektedir (Subaşı Ertekin, 2001).

Dış ticaret bir ülkenin ekonomik süreçlerinin önemli parçasıdır. Gelişmekte olan ülkelerin ithalat ve ihracat süreçlerinde verimliliği arttırmaları gerekmektedir. Bu iki olgu birbiriyle ilişkilidir ve ithalatla ilgili yapısal sorunların aşılmasında ülkeler ihracata bel bağlamaktadır. İhracat aynı zamanda bir ülkenin yerli üretiminin gelişmesinde, teknolojik gelişimin sağlanmasında, üretimde kalitesinin yükselmesinde ve ekonomide dengenin sağlanmasında hayati öneme sahiptir. Bu açıdan bakıldığında ihracat rakamlarını doğru tahmin etmek, ülkenin makroekonomik planlarının geliştirilmesinde ve özellikle dış ticaret politikalarının belirlenmesinde çok önemlidir (Zhang ve ark., 2009).

Büyük oranda uluslararası ticaret teşvik edildiği günümüz dünyasında hem ithalat hem de ihracat önem taşımaktadır. Planlama bağlamında ülkeler açısından ele alındığında dış ticarete yönelik verileri doğru tahmin edebilmek büyük avantajlar sağlayacaktır (Dave ve ark., 2021). İthalat ve ihracat değerlerinde değişimin önceden tahmin edilmesi uluslararası ticarete yönelik planların doğru şekilde yapılarak konuyla ilgili doğru politika üretilmesi açısından önemlidir. Dış ticarete doğru tahminler yapılarak ülke ekonomisini akılcı yönetmek mümkün olacaktır (Yuan, 2017).

Veri madenciliğine yönelik uygulamalar her geçen gün akademik camianın daha fazla ilgisini çekmektedir (Durmuşoğlu, 2017). Konuya ilişkin literatür incelendiğinde; Weka programı kullanılarak yapılan akademik çalışmalara karbon emisyon hacminin tahmini, hastalık tedavi başarımlarının tahmini, müşteri kayıp tahmini, talep tahmini, gelir tahmini, ülkelerin gelişmişlik tahmini gibi farklı örnekler bulunmaktadır (Kılıç ve ark., 2018; Demirci ve Karaatlı, 2019;Uzun ve ark., 2019; Aydemir ve ark., 2020;Abuzir ve Abuzir, 2021;Uyar Erdem ve ark., 2021).

Dış ticaret verilerinin öneminden hareketle çalışmada Türkiye’nin ithalat ve ihracat değerlerine yönelik tahmin yapılmıştır. Tahminde Lineer Regresyon, MLPRegressor, Gaus Süreç Regresyon, RBF Regressor ve Destek Vektöre Regresyon algoritmaları kullanılmıştır. Çalışmada kullanılan veriler 1969-2022 yılları arasında kapsamaktadır. Çalışma, Türkiye’nin ithalat ve ihracat verilerinin tahmin edilebilirlik seviyesini göstermesi açısından önem taşımaktadır.

Çalışmada hata değerlendirme kriteri olarak Atsalakis ve ark. (2015), Boonyakunakorn ve ark. (2018) ve Dave ve ark. (2021) tarafından yapılan çalışmalarla uyumlu olarak MAPE, RMSE, MSE ve MAE kullanılmıştır.

## 2.Literatür Araştırması

### 2.1. Veri Madenciliği

Veri madenciliği; güncel durumda verinin işlenerek bilgiye dönüştürülmesi konusunun öneminden aldığı güçle veriden öğrenme ve veriyle tahminde bulunma odaklı gerçek dünya uygulamalarda kendine yer bulmaktadır (Albayrak ve Yılmaz, 2009; Alaeddinoğlu ve ark., 2014). Veri madenciliği güncel çalışmalarda genellikle vektör uygulamaları şeklinde olmasına karşın zaman içerisinde çok daha karmaşık veri türleri üzerinde çalışılıyor olması olasıdır (Kriegel ve ark., 2007). Sürekli gelişim içerisinde olan veri madenciliği, teknoloji odaklı dünyada büyük verileri hızlı şekilde anlamlandırmaktadır. Bir konu hakkındaki eğilimlerin ortaya konması ve gelecekte yapılacak çalışmalara yol göstermesi gibi önemli misyona sahiptir (Akgün ve Bulut Özek, 2020).

Veri madenciliği istatistiksel modellere dayanan ve büyük veri kümelerinde analiz yapmaya olanak bir alandır (Emre ve Selçukcan Erol, 2017). Veri madenciliği; bankacılık, tıp, ekonomi, iletişim teknolojileri, yazılım ve eğitim gibi farklı alanlarda uygulanabilmektedir (Özbay, 2015; Gürsoy ve Bilgin, 2016; Akgün, 2106; Özdemir ve ark., 2018; Artsın, 2019). Eğitim alanında öğrenci başarısının değerlendirilmesi, tıp alanında hastalık belirtileri odaklı şablonlar oluşturulması, ekonomiye yönelik uygulamalarda müşteri ve pazar analizine yönelik uygulamalara rastlanmaktadır. Bankacılık ve borsa odaklı çalışmalar daha çok tahmin ağırlıklı olarak ortaya çıkmaktadır (Savaş ve ark., 2012).

Veri madenciliği giderek kullanımı artan alan olmanın yanı sıra kritik bir teknoloji olmasıyla da öne çıkmaktadır. Büyük miktarlarda verinin üretildiği sektörlerde veri madenciliği stratejik öneme sahiptir (İrmak ve ark., 2012). Veri madenciliğine yönelik çok farklı alanda uygulamalar bulunmaktadır. Tablo 1’de bu uygulamalara örnekler verilmektedir.

**Tablo 1.** Veri Madenciliğine Yönelik Diğer Literatür.

<b>Yazar</b>	<b>Açıklama</b>
<b>Raza ve ark. (2009)</b>	Pakistan’ın ihracatına yönelik veri madenciliği uygulamasını 2004-2007 yılları arasındaki verileri kullanarak yapmışlardır.
<b>Özel ve Topsakal (2014)</b>	9 farklı algoritma kullanarak beton basınç dayanımını belirlemişlerdir. Kullandıkları algoritmalar: “Gaussian Processes”, “Linear Regression”, “Multilayer Perceptron”, “SMOreg”, “IBK”, “KStar”, “Additive Regression”, “Bagging” ve “M5Rules” şeklindedir. Değerlendirme kriteri olarak Korelasyon Katsayısı, Ortalama Mutlak Hata (MAE), Karekök Ortalama Hata (RMSE) ve Nisbi Mutlak Hata (RAE) kriterlerini kullanmışlardır. En iyi sonucu Kstar algoritmasından elde etmişlerdir.
<b>Sebik ve Bülbül (2018)</b>	Akciğer kanseri teşhisinin daha doğru yapılabilmesi için veri madenciliği modellerini kullanılmıştır. Çalışmadan “Naive Bayes “, “Bayes Net”, “Lojistik reg.”, “Multilayerpercep.”, “KStar”, “Bagging”, “OneR”, “ZeroR”, “J48” ve “Random Tree” algoritmaları kullanılmıştır. En iyi sonucu “Naive Bayes” algoritmasıyla elde etmişlerdir.
<b>Garcia Rodriguez ve ark. (2021)</b>	Kamu ihalelerinde fiyatı tahmin eden uygulamalarında doğrusal regresyon ve rastgele orman gibi geleneksel regresyon yöntemlerinin yanı sıra izotonik regresyon ve popüler yapay sinir ağı modelleri yani göreceli olarak daha yeni yöntemleri de kullanmışlardır.
<b>Ayanoğlu ve Kurt (2019)</b>	Yapay sinir ağı algoritması kullanarak metal sektörüne yönelik kaza tahmin modeli oluşturmuşlardır. Modelde 165 işletmeye ait çok değişkenli veri kümesi kullanılmıştır. Model %90 doğruluk oranını yakalamıştır.

<b>Aydemir (2019)</b>	Yabancı dil 2 dersini alan öğrencilerin 2017-2018 eğitim yılındaki verilerini kullanarak ders geçme notunu tahmin etmiştir. Çalışmada öğrencilerin başarı durumunu önceden öngörmek amaçlanmaktadır. Uygula Weka programında farklı 19 algoritma kullanılarak gerçekleştirilmiştir. En iyi tahmin sonucu Bagging (M5P) algoritmasıyla elde edilmiştir.
<b>Acı ve Ayyıldız Doğansoy (2022)</b>	Yerel bir dükkânın 2019-2020 ürün satış verileri üzerinde tahmin modelleri oluşturmuşlardır.
<b>Kitessa ve ark. (2021),</b>	Sürdürülebilir bir kentsel planlama için uzun vadeli su ve enerji tahminini Weka yazılımı kullanarak yapmışlardır. Çalışmada doğrusal regresyon modeli kullanılmıştır.
<b>Özarı ve Demirkale (2022)</b>	2008-2021 yılları arası BIST100, BIST50, BIST30, Avro/TL ve Dolar/TL kapanış fiyatlarını kullanarak tahmin çalışması yapmışlardır. Çalışmada elde edilen tahmin sonuçları bir sonraki gün için al-sat-bekle kararını vermede kullanılmıştır. Uygulamada k-NN algoritması tercih edilmiştir.

## 2.2. Benzer Çalışmalar

Lehmann (2020), Eurostat verilerini kullanarak 1996-2016 yıllarını ve 18 ülkeyi kapsayan Avrupa ülkeleri ihracat tahmini yapmıştır. Çalışmada, standart otoregresif dağıtılmış gecikme (ADL) modeli kullanılmıştır.

Zhang ve ark. (2009), Çin'in Hubei eyaletinin ihracat rakamlarını tahmin etmek için gri sistem teorisinden GM (1,1) modelini kullanmışlardır. Bu model elde bulunan veriye dayanarak tahmin modeli geliştirilmesini sağlamaktadır. Çalışma araştırmaya konu olan eyaletin ihracatının 2009 yılında artacağını göstermektedir.

Narayan ve ark. (2008) Fiji'nin ithalatını ve ihracatını tahmin etmişlerdir. Çalışmada otoregresif hareketli ortalama (ARMAX) modeli kullanılmıştır. Çalışma sonucunda ilgili dönemde ithalatın ihracattan daha yüksek artacağı ve bu durumun dış ticaret açığını yükselteceği ortaya konmuştur. Çalışmada model hata değerlendirme kriteri olarak MAPE kullanılmıştır.

Wong ve ark. (2010), Tayvan ihracatının tahmininde klasik zaman serisinin uygulamasına yönelik ARIMA modelini ve Vektör ARMA modelini kullanmıştır. Ayrıca Bulanık Zaman Serisi Yöntemine yönelik olarak İki faktörlü model, Sezgisel model ve Markov modeli tercih edilmiştir. Sonuç kısa zaman serilerinde bulanık modelin daha iyi olduğunu göstermektedir.

Yuan (2017), PSO (Particle Swarm Optimization Algorithm) ve GMDH (Particle swarm Method of Data Handling) modelini kullanarak Shenzhen bölgesinin ithalatının ve ihracatının tahminini yapmıştır. Tahmin önce tek başına GMDH ağıyla yapılmıştır. İkinci tahmindeyse PSO ve GMDH birlikte kullanılmıştır. Sonuçlar ikinci modelin daha iyi sonuç verdiğini göstermektedir.

Boonyakunakorn ve ark. (2018), Tayland'ın ASEAN ülkelerine ihracatını tahmin etmişlerdir. Çalışmada SETAR, LSTAR, MSAR ve Kink AR modeli olmak üzere dört doğrusal olmayan model kullanılmıştır. Bu dört modelle karşılaştırmak için AR analizi de yapılmıştır. Değerlendirme kriteri

olarak RMSE ve MAE tercih edilmiştir. Sonuçlar SETAR modelinin diğerlerinden daha iyi sonuç verdiğini göstermektedir.

Urrutia ve ark. (2019), Filipinler'in ithalat ve ihracat tahminlerinde Otoregresif Bütünleşik Hareketli Ortalama (ARIMA) ve Bayesian Yapay Sinir Ağı (BANN) kullanılmışlardır. Çalışmanın amacı bu iki modelin karşılaştırılmasıdır. Veriler 1993-2017 yılları arasını kapsamaktadır ve bu döneme ait 100 gözlem bulunmaktadır. Model hata değerlendirme kriteri olarak MSE, NMSE, MAE, RMSE ve MAPE kullanılmıştır. Çalışma BANN modelinin daha iyi sonuç verdiğini göstermektedir. Çalışma ithalata ve ihracata yönelik tahminde Filipin ekonomisine katkıda bulunma gibi bir amaca sahiptir.

Eckert ve ark. (2021), büyük bir zaman serisi veri kümesiyle desteklenen İsviçre dış ticaretine yönelik tahmin çalışmasında Bayes modelini temel almıştır. Yöntemin elde ettiği sonuç umut verici olarak değerlendirilmektedir.

Sohrabpour ve ark. (2021), yapay zekâ tekniklerinden olan genetik algoritma kullanarak bir firmanın ihracat tahmin modellemesini yapmışlardır. Tahmin 6 haftalık süreci kapsamaktadır. Elde edilen tahmin sonuçları gerçek değerlerle karşılaştırılmaktadır.

Ghauri ve ark. (2020), 2002-2019 yıllarını kapsayan ithalat ve ihracat verilerini kullanarak Pakistan dış ticaretine yönelik tahmin çalışması yapmışlardır. Box Jenkins, AR ve ARIMA kullanılan modellerdendir. Modelde hata değerlendirme kriteri olarak RMSE ve MAE kullanılmıştır.

Liu (2021), dış ticaret verilerinde tahmin odaklı çalışmasında sinir ağı ve bulanık teoriyi birleştirilerek oluşturduğu karma modeli kullanmıştır.

Dave ve ark. (2021), Endonezya'nın ihracat tahminine yönelik modelde makine öğrenmesi tekniklerini kullanmışlardır. Çalışma Endonezya hükümetlerine doğru ihracat tahmin modeli önermektedir. Analiz 1998 sonrası aylık verileri kapsamaktadır. Çalışmada LSTM ve ARIMA modelleri karma çerçevede kullanılmıştır. LSTM, verilerin doğrusal olmayan bileşenlerinde ve ARIMA verilerin doğrusal bileşenlerinde kullanılmıştır. Karma model %7.38 MAPE ve RMSE  $1.66 \times 10^{13}$  değerinde sonuç üretirek bağımsız modellerden daha iyi sonuç elde etmiştir.

Wang (2022) ithalat ve ihracat hacmi tahmininde denetimli öğrenme modeli olan RVM kullanmıştır. Ardından RVM ve PSO modellerini karma olarak kullanmıştır.

Han ve ark. (2022) zaman serisi olarak elde edilen Çin'in ihracat hacmini Destek Vektör Sinir Ağı (SVNN) algoritması kullanarak tahmin etmiştir. Modelin hata değerlendirme oranı %10 seviyesinin altında sonuç vererek çok iyi performans göstermiştir.

### **3. Materyal ve Metot**

Türkiye İhracat ve İthalat değerleri; WEKA programında yer alan algoritmalarından SMOReg, MLPRegressor, Gaus Süreç Regresyon, RBF Regressor ve Doğrusal Regresyon yöntemleriyle 1969-2022 yılları arasındaki veriler kullanılarak tahmin edilmiştir. Veri seti %80-%20 oranında eğitim ve test verileri olmak üzere 2 ye ayrılmıştır. Çünkü bahsi geçen yıllar arasındaki veri seti kullanılarak

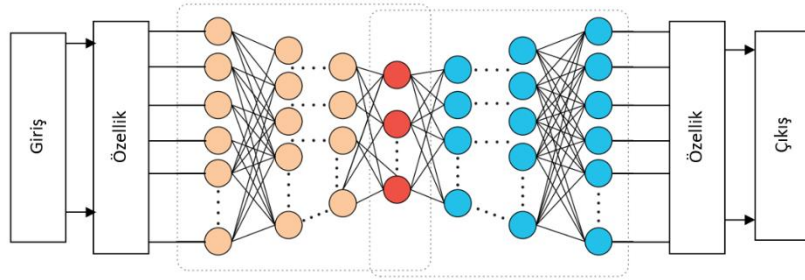
Türkiye'nin ithalat ve ihracat değerleri üzerinde tahmin çalışması yapılacaktır. Yapılan tahmin sonucunda ithalat ve ihracat verilerinin farklı algoritmalara göre performansı karşılaştırılacaktır.

### 3.1.Tahmin Modeli

Türkiye'nin ithalat ve ihracat değerlerinin veri madenciliği yöntemleriyle (SMOreg, MLPRegressor, Gaus Süreç Regresyon, RBF Regressor ve Doğrusal Regresyon) tahmin edilmesi ve tahmin sonuçlarının karşılaştırılması çalışmanın modelini oluşturmaktadır. Çalışmada veri madenciliği aracı olarak seçilen Weka yazılımı farklı algoritmaların kullanıma olanak tanımaktadır (Eibe ve ark., 2016; Purlu ve Türkay, 2021). Çalışma kapsamında oluşturulan tahmin modelleri karşılaştırılmıştır.

#### 3.1.1 Radial Basis Functions (RBF Regressor)

Şekil 1'de yer alan RBF ağı, üç katmanlı bir ileri beslemeli yapay sinir ağ modeli olarak düşünülebilir. RBF sinir ağı modelinde, giriş özellikleri  $X_k = (x_1, x_2 \dots x_n)$ ,  $k=1,2,\dots,m$  şeklinde ifade edilmektedir. Burada m-n ve m-n sırasıyla toplam örnek sayısını ve bir örneğin toplam özellik sayısını ifade etmektedir. Genel olarak giriş katmanındaki nöron sayısı ile toplam öznitelik sayısı birbirine eşittir. Girdi katmanındaki nöron sayısı, örnek öznitelikler filtrelendikten sonra genellikle özniteliklerin toplam sayısından daha azdır (Yu, 2022).



Şekil 1. Yu (2022) kaynağına göre RBF modeli

Yu (2022) tarafından verilen modelde k örneğinin çıktısı  $Y_k = (y_1, y_2 \dots y_n)$  ve n çıktı katmanındaki nöron sayısıdır. İlk olarak, girdi örnekleri, birinci gizli katmanın değerlerini verecek şekilde ağırlıklandırma katsayılarıyla ayarlanır (Eşitlik 1).

$$S_{1j} = \sum_{i=1}^n W_{1ij}X_i - \theta_{1j}, j = 1, 2, \dots, p. \quad (1)$$

Eşitlik 2'de olduğu gibi oluşturulan katmanın değerlerinin dönüştürülmesi gerekir.

$$b_{1j} = \exp\left(\frac{\left\|\sum_{i=1}^n W_{1ij}X_i - \theta_{1j} - c_j\right\|^2}{2\sigma^2}\right) \quad (2)$$

Burada  $\sigma$ , 0'dan büyük gerçekte sayıdır,  $c_j$  ise j gizli katman merkezidir. RBF sinir ağı için seçilen özellik dönüştürme fonksiyonu Gauss fonksiyonudur. Birinci gizli katman girdi olarak kullanılır ve ağırlık katsayıları ayarlandıktan sonra ikinci gizli katmana ait değerler elde edilir (Eşitlik 3).

$$S_{2j} = \left( \sum_{i=n+1}^{2n} W_{2ij} X_j + \theta_{1j} \right) + \sum_{i=n+1}^{2n} \sum_{j=1}^p W'_{ij} b_{1j} - \theta_{2j} \quad (3)$$

Eşitlik 4'te gösterildiği gibi dönüştürme işlevi daha sonra çözüm işlemini gerçekleştirmek için kullanılır.

$$b_{2j} = \exp \left( \frac{\left\| \sum_{i=n+1}^{2n} W_{2ij} X_j + \theta_{1j} + \sum_{i=n+1}^{2n} \sum_{j=1}^p W'_{ij} b_{1j} - \theta_{2j} - c_j \right\|^2}{2\sigma^2} \right) \quad (4)$$

Oluşturulan tüm katmanlardan sonra çıktı, ek ağırlıklar  $V_{ij}$  gerektirir (Eşitlik 5).

$$L_t = \sum_{j=1}^p V_{jt} b_{2j} \quad (5)$$

Tüm modelin çıktısını elde etmek için Eşitlik 6'da olduğu gibi Gauss fonksiyonu çözülür.

$$C_t = \exp \left( - \frac{\left\| \sum_{i=1}^n V_{jt} b_{2j} - c_j \right\|^2}{2\sigma^2} \right) \quad (6)$$

k-inci örnek için hata sonuçları aşağıdaki gibi gösterilir (Eşitlik 7):

$$E_k = \frac{\sum_{t=1}^q (y_t^k - C_t^k)}{2} \quad (7)$$

Tüm örnekler için hatalar Eşitlik 8'de gösterilmiştir:

$$E = \sum_{k=1}^m \sum_{j=1}^q \frac{(y_t^k - C_t^k)}{2}, \Delta v_{jt} = \frac{-\partial E_K}{\partial v_{jt}} \quad (8)$$

Örtülü ve çıktı katmanları arasındaki ağırlıklar:  $\Delta W_{jt} = \alpha d_t^k b_{j,t}$ ,  $j = 1, 2, \dots, n; k = 1, 2, \dots, m$ , şeklindedir. Burada,  $\alpha$  öğrenme hızıdır ve  $d = (y_j^k - C_j^k)t(1 - C_j^k)$  olarak gösterilir.

### 3.1.2. SMOreg

Destek Vektör Makineleri (DVM) Cortes ve Vapnik (1995), istatistiksel öğrenme teorisi üzerine yapmış oldukları çalışmalarda sınıflandırma problemlerinin çözümü için öne sürülmüştür. Destek Vektör Regresyon yöntemi ise Smola ve Scholkopf (2004) ve Shevade ve ark. (2000) tarafından DVM'lerin regresyon problemlerine çözümü için öne sürülmüştür (Witten ve Frank, 2005). Sıralı En Düşük Optimizasyon (Sequential Minimal Optimization Regression) algoritması DVM yönteminin regresyon problemlerine uyarlanmasıdır.

Regresyon analizi; tahmin probleminde değişkenler arasındaki ilişkiyi ele alan bir modeldir. SMOreg, sıradan SMO'da bulunan tek eşliğin neden olduğu kaotik ve verimsiz problemin üstesinden gelir ve ikili problemle ilgili optimizasyon yönergelerinden ipuçları alır. İki eşik parametresi kullanır ve orijinal SMO'dan daha verimlidir. SMOreg, model olarak yapısal risk minimizasyonu kısıtlamalarını kullanır. Regresyon tahmini ve doğrusal olmayan, küçük örnek verileri işleme konusunda iyidir. Bu nedenle SMOreg, zaman serisi verilerini doğru şekilde tahmin edebilir (Wang ve ark., 2012).

DVM regresyon analizinin birincil amacı, her giriş vektörü ( $v_i$ ) için en uygun hedef değerleri ( $t_i$ ) gösteren hiperdüzlem (Denklemler (9) ve (10)) için bir  $f(v)$  fonksiyonu bulmaktır. Bu, veri noktalarını içeren setin ( $v_i, t_i$ ) eğitilmesiyle yapılır (Platt, 1998):

$$f(v) = h \cdot j(v) + b \quad (9)$$

$$m = \frac{1}{\|\eta\|_2} \quad (10)$$

Eğitim için  $v_i = i$ th vektörlerinin girişi;  $t_i = i$ th hedef vektörleri,  $i$ 'inci giriş vektörlerine karşılık gelir;  $b =$ yanlılık;  $\eta =$  hiperdüzlemin normal vektörü;  $m =$ marj. Marj,  $m$ , eşitlik 11 aracılığıyla optimizasyon analiziyle maksimize edilir.

Marj, Denklem (11) aracılığıyla optimizasyon analizi yöntemiyle maksimize edilir:

$$\min_{n, b} \frac{1}{2} \|\eta\|^2 \text{ subject to } t_i (\eta \cdot v_i) \geq 1, \forall i \quad (11)$$

Optimizasyon problemini ikili forma dönüştürmek için Lagrange çarpanları kullanılarak aşağıdaki gibi hesaplanan amaç fonksiyonları Denklem (12) ve (13) ile verilmektedir.

$$\min_{\vec{\beta}} \delta(\vec{\beta}) = \min_{\vec{\beta}} \frac{1}{2} \sum_{i=1}^N \sum_{j=1}^N t_i t_j (\vec{v}_i \cdot \vec{v}_j) \beta_i \cdot \beta_j - \sum_{i=1}^N \beta_i \quad (12)$$

$$\beta_i \geq 0, \forall i \text{ ve } \sum_{i=1}^N t_i \beta_i = 0 \quad (13)$$

$N$ , eğitim örneklerinin sayısıdır. Cortes ve Vapnik (1995), Denklemlerde (14–16) verildiği gibi, hatanın öncelik değerinden küçük olduğu durumlarda hata düzeltmeleri yapmak için duyarsız kayıp fonksiyonunu kullanmışlardır.

$$\min \frac{1}{2} \|\eta\|^2 + C \sum_i X_j + X'_i \quad (14)$$

$$t_i - \eta \cdot v_i - b \leq \varepsilon + X_i \quad (15)$$

$$X_j \cdot X'_i \geq 0 \quad (16)$$

Burada  $X$  değişkenlerin kenar boşluklarını geçemez,  $C$  bir parametredir ve  $\varepsilon$  priori'dir. Primal problem olarak bilinen optimizasyon problemi, yukarıda verilen DVM formülasyonu kullanılarak çözülmüştür. Kısıtları değiştirmek için optimizasyon problemi değişkenleri içermeyen ikili forma dönüştürülür. Ardından, çıktıyı elde etmek için doğrusal olmayan SVM, Eşitlik 17'de gösterildiği gibi Lagrange çarpanları kullanılarak çözülür. Platt, 1998 tarafından verilen modelde girdi vektörü ile saklanan eğitim vektörü arasındaki mesafeyi belirlemek için  $K$  adlı kernel fonksiyonu kullanılır.

$$f = \sum_{j=1}^N t_j \beta_j K(\vec{v}_j, \vec{v}) - b \quad (17)$$



### 3.1.3.Gauss Süreci

Gauss süreci; sonlu sayıdaki değişkenlerle ilgilidir. Gauss süreci ortalama fonksiyonu  $m(x)$  ve kovaryans fonksiyonu  $k(x, x')$  tarafından belirlenir. Bu vektör veya matrisin ortalama ve kovaryansa göre Gauss dağılımının genel sıralamasını ifade etmektedir. Sürece ilişkin adımlar aşağıda verilmektedir (David, 1998;Rasmussen, 2004):

Gauss süreciyle ilgili fonksiyon Eşitlik 9'da verilmektedir.

$$f \sim GP(m, k) \quad (9)$$

Gauss sürecinde değişkenler vektördeki konumuna göre indekslenir. Dizin kümesi  $x$  ile ifade edilir. Her  $x$  girişi için rastgele  $f(x)$  değeri vardır. Eşitlik 10'da uygulanabilir bir Gauss süreci verilmektedir.

$$m(x) = \frac{1}{4}x^2, \text{ ve } k(x, x') = \exp\left(-\frac{1}{2}(x, x')\right)^2 \quad (10)$$

Eşitlik 11 düzenli Gauss sürecini vermektedir. Süreç ve dağılımı ifade edebilmek için  $m$  ve  $k$  kullanılırken ilki için  $\mu$  ve ikincisi için  $\Sigma$  kullanılmaktadır. Bu noktada  $x$ 'e karşılık gelen  $f(x)$  rasgele vektör oluşturulabilecektir.

$$\mu_i = m(x_i) = \frac{1}{4}x_i^2, i = 1, \dots, n \text{ ve } \Sigma_{ij} = k(x_i, x_j) = \exp\left(-\frac{1}{2}(x_i, x_j)^2\right), i, j = 1, \dots, n \quad (11)$$

Bu süreçlerden sonra verilerin eğitim aşamasına geçilmektedir. Veriler eğitim ve test olarak ikiye ayrılır. Kovaryans vektörü hesaplanır ve tahmin bu altyapı üzerinde yapılır.  $f$  eğitimin bilinen fonksiyonu olduğunda ve  $f^*$  test verisi olduğunda eşitlik 12 oluşacaktır

$$\begin{bmatrix} f \\ f^* \end{bmatrix} \sim N\left(\begin{bmatrix} \mu \\ \mu^* \end{bmatrix}, \begin{bmatrix} \Sigma & \Sigma^* \\ \Sigma^{*T} & \Sigma^{**} \end{bmatrix}\right) \quad (12)$$

Sonrasında test verileri bilindiği için tahminle bilinen test verisi arasındaki ilişkiye yoğunlaşmaktadır.

### 3.1.4.MLP regresyon

MLPRegressor algoritması ileri beslemeli yapay sinir ağlarının regresyon problemlerine uyarlanmasıdır (Pereira ve Aires, 2018). Weka MLP regresyon (MLPRegressor) optimizasyon sınıfı kullanılarak verilen kayıp fonksiyonu ve Newton-benzeri yöntemlerden BFGS, (BroydenFletcher-Goldfarb-Shanno) yöntemiyle ikinci dereceden cezayı optimize ederek tek bir gizli katmana sahip algılayıcıyı eğitmektedir. Ridge parametresi ağırlıkların boyutu üzerindeki cezayı belirler. Çıkış katmanında sınıflandırma için sigmoid fonksiyonu kullanılır. Giriş katmanı için yaklaşık sigmoid değeri belirlenmişse çıkış katmanı için de kullanılabilir (George-Nektarios, 2013):

- Gizli birim sayısı; varsayılan değer 2'dir.
- Sırt (Ridge) parametresi: ağırlıkların büyüklüğündeki cezayı belirlemek için kullanılır.
- Aktivasyon fonksiyonları;
  - Yaklaşık sigmoid:  $f(x) = 1/(1 + e^{(-x)})$

- Sigmodi:  $f(x) = 1/(1 + e^{-x})$
- Softplus:  $f(x) = \ln(1 + e^x)$
- Kayıp fonksiyonu;
  - Yaklaşık Mutlak Hata:  $loss(a, b) = \sqrt{(a - b)^2 + \epsilon}$
  - Karesel Hata:  $loss(a, b) = (a - b)^2$
- Delta değerleri için tolerans parametresi
- Eğitim ivmesi için eşlenik gradyan girişi

### 3.1.5. Doğrusal regresyon

İçerisinde birden fazla fonksiyon barındıran doğrusal regresyon birçok bilim dalında kullanılan ve tahmin uygulamaları yapılabilen önemli bir modeldir (Weisberg, 2005; Kılıç, 2013; Nagy, 2018).

Doğrusal regresyon, tüm nitelikler sayısal olduğunda üzerinde durulması gereken klasik bir tekniktir. Doğrusal regresyon sayısal tahmin için mükemmel modeldir. Sınıfı önceden belirlenmiş ağırlıklarla niteliklerin doğrusal kombinasyonunu ifade etmektedir. X'in sınıfı, a özellik ve w ağırlık değeri olduğu durumda Eşitlik 13 ve sonraki adımlar oluşmaktadır. Aşağıda doğrusal regresyon modelinin ana hatları verilmektedir (Witten ve ark., 2017):

$$x = w_0 + w_1 a_1 + w_2 a_2 + \dots + w_k a_k \quad (13)$$

Ağırlıklar eğitim verilerinden hesaplanır. İlk örneğin sınıfı için tahmin edilen değer Eşitlik 14'de verilmektedir.

$$w_0 a_0^1 + w_1 a_1^1 + w_2 a_2^1 + \dots + w_k a_k^1 = \sum_{j=0}^k w_j a_j^1 \quad (14)$$

Sınıfın tahmin edilen değerlerine ulaşıldıktan sonra önemli olan tahmin edilen değerle gerçek değer arasındaki farktır. En küçük kareler doğrusal yöntemi tüm eğitim örnekleri üzerinde bu farkların karelerinin toplamını en aza indirmek için  $w_j$  katsayılarını seçmektir. O zaman farkların kareler toplamı Eşitlik 15 gibi olmaktadır.

$$\sum_{i=1}^n \left( x^{(i)} - \sum_{j=0}^k w_j a_j^{(i)} \right)^2 \quad (15)$$

### 3.2. Model Değerlendirme Kriterleri

Zaman serilerinde tahmin sonuçları birçok kriterle değerlendirilebilmektedir. En sık kullanılan değerlendirme araçları RMSE ve MSE/MAPE olarak görülmektedir (Yurduseven ve Müngen, 2022). MAPE tahminin doğru şekilde ölçülmesini sağlar (Smith ve Sincich, 1988). Benzer işi yapan RMSE, model hatasının standart sapmasıdır. RMSE hesaplanırken tahmin değeriyle gerçek değer arasındaki fark hesaplanır. Sonrasında ortalama değer kökü hesaplanmadan önce fark değerinin karesi alınır. MAE, tahmin edilen ve gerçek değer arasındaki mutlak farkların ortalamasıdır (Kitesa ve ark., 2021). Eşitlik 16 Ortalama Mutlak Hata (Mean Absolute Error, MAE), eşitlik 17 Karekök Ortalama Hata (Root Mean Square Error, RMSE), eşitlik 18 Ortalama Mutlak Yüzde Hatası (Mean Absolute

Percentage Error, MAPE), eşitlik 19 Ortalama Kare Hata (Mean Square Error, MSE) değerlendirme ölçütlerinin genel çerçevesini vermektedir (Emeksiz ve ark., 2016; Namlı ve ark. 2019; Garcia Rodriguez ve ark., 2021).

$$MAE = \frac{1}{n} \sum_{i=1}^n |y - y'| \quad (16)$$

$$RMSE = \sqrt{\frac{\sum_{i=1}^n (\hat{y}_i - y_i)^2}{n}} \quad (17)$$

$$MAPE = \frac{100\%}{n} \sum_{i=1}^n \left| \frac{e_i}{y_t} \right| \quad (18)$$

$$MSE = \frac{1}{n} \sum_{i=1}^n (Y_i - X_i) \quad (19)$$

MAE, performans ölçütünün belirli bir olayın olasılığı olduğu durumlarda sistemin simülasyonu ve analizinde kullanılabilir (Mendo, 2009). Alt limiti MAE'ye sabitlenen RMSE, hata büyüklüklerinin dağılımı daha değişken hale geldikçe MAE değerine göre daha büyük olma eğilimindedir (Willmott ve Matsuura, 2005). Myttenaere ve ark. (2016), MAPE'yi göreceli hata açısından sezgisel yorumu nedeniyle pratikte sıklıkla kullanılan model olarak görmektedir. MAPE, bu özelliğiyle göreceli değerlerin ölçüldüğü finansla ilgili modellemelerde kullanılabilir. MAPE, gerçek dünya olaylarında tahmin edilen değerlerin sıfırın üzerinde olduğu durumlarda kullanılabilir uygun seçenek olarak görülmektedir.

Lewis (1982), MAPE değerinin %10 altında olması durumunda bunu yüksek doğrulukta tahmin olarak değerlendirmektedir. Sonrasında %10-%20 arası tahmin iyi, %20-%50 arası tahmin makul, %50 üzeri tahmin ise yanlış tahmin olarak görülmektedir.

### 3. Bulgular ve Tartışma

Finansal tahmin problemlerinde farklı algoritmalar kullanılabilir (Basak, 2007). Ulaşılabilen literatürde Weka kullanılarak Türkiye'nin ticaret verilerine yönelik yeterli tahmin çalışması bulunmamaktadır. Dış ticaret verilerinin öneminden hareketle çalışmada Türkiye'nin ithalat ve ihracat değerlerine yönelik tahmin yapılmıştır. Tahminde Lineer Regresyon, MLPRegressor, Gauss Süreç Regresyon, RBF Regressor ve Destek Vektöre Regresyon algoritmaları kullanılmıştır. Çalışmada kullanılan veriler 1969-2022 yılları arasını kapsamaktadır. Çalışma bu noktadaki literatürü desteklemesi açısından önem taşımaktadır.

Vaka çalışması Türkiye'nin ithalat ve ihracat verilerini kapsamaktadır. 1969 yılından başlayarak 2022 yılının ilk 7 ayını kapsayan TÜİK ithalat ve ihracat verileri çalışmanın veri setini oluşturmaktadır. Veri seti 703 adet aylık değeri kapsamaktadır. Namlı ve ark., (2019) tarafından belirtildiği gibi makine öğrenme modellerinden en iyi verimi elde edebilmek için veriler normalize edilmiştir.

Aksu ve Doğan (2019), Urfaloğlu ve Şahin (2022) tarafından yapılan çalışmada dile getirildiği gibi Weka programı veri madenciliğine yönelik akademik çalışmalarda kullanılabilir ve buradan hareketle analizlerde Weka programı tercih edilmiştir. Weka programının açık kaynak kodlu olması diğer bir tercih sebebidir.

Acı ve Ayyıldız Doğan (2022) tarafından ifade edildiği gibi makine öğrenmesi modellerinde veriler eğitim ve test verisi olarak ikiye ayrılabilir. Örneğin %80 eğitim ve %20 test verisi olarak ikiye ayrılmış veri setinde model performans ölçütü olarak test verisinde elde edilen sonuç değerlendirilmelidir (Namlı ve ark., 2019).

Türkiye ihracat tahmininde Tablo 2’de yöntemlerin eğitim aşaması kriter değerleri yer almaktadır. Eğitim aşaması kriter değerlerinin sonuçlarına göre; MAPE oranları %12,6080 değeriyle MLPRegressor algoritması en düşük değere sahiptir. %18,6803 değeriyle SmoReg algoritmasına en yüksek değere sahiptir.

**Tablo 2.** İhracat tahmini eğitim aşaması

	MAE	MAPE	RMSE	MSE
<b>Lineer Regresyon</b>	274113929,1	13,5595	548001058,1	300.305.159.672.289.000
<b>MLPRegressor</b>	244596383,2	12,6080	511000479,4	261.121.489.896.288.000
<b>Gaussian</b>	267608275,5	13,5019	540941733,8	292.617.959.339.646.000,00
<b>RBF Regressor</b>	264382881,3	14,5418	543485494,3	295.376.482.524.545.000,00
<b>SmoReg</b>	257.733.116	18,6803	551728517,9	304.404.357.495.654.000,00

Türkiye ihracat tahmininde Tablo 3’de yöntemlerin test aşaması kriter değerleri yer almaktadır. MAPE oranları %8,6987 değeriyle SmoReg algoritması en düşük değere sahiptir. %9,4072 değeriyle Lineer Regresyon en yüksek değere sahiptir.

**Tablo 3.** İhracat tahmini test aşaması

Yöntem	MAE	MAPE	RMSE	MSE
<b>Lineer Regresyon</b>	1266350696	9,4072	1622163722	2,63142E+18
<b>MLP Regressor</b>	1248316085,74	8,8731	1672998136,124	2,79892E+18
<b>Gaussian</b>	1225844318	9,0728	1571714342	2,47029E+18
<b>RBF Regressor</b>	1254911796	8,8761	1701939639	2,8966E+18
<b>SmoReg</b>	1.154.607.471	8,6987	1502685070	2,25806E+18

Türkiye ithalat tahmininde Tablo 4’te yöntemlerin eğitim aşaması kriter değerleri yer almaktadır. MAPE oranı %13,198 değeriyle MLPRegressor algoritması en düşük değere sahiptir. %13,951 değeriyle RBF Regressor algoritması en yüksek değere sahiptir.

**Tablo 4.** İthalat tahmini eğitim aşaması

	MAE	MAPE	RMSE	MSE
<b>Lineer Regresyon</b>	414989636	13,5658	886491446,9	785.867.085.406.080.000
<b>MLP Regressor</b>	380071284,2	13,198	835103580	697.397.989.316.070.000
<b>Gaussian</b>	419084027,2	13,593	893623019,2	798.562.100.525.132.000,00
<b>RBF Regressor</b>	419.660.201	13,951	925045737,8	855.709.617.113.834.000,00
<b>SmoReg</b>	400.774.093	13,6118	876086715,1	767.527.932.379.369.000,00

Türkiye ithalat tahmininde Tablo 5’te yöntemlerin test aşaması kriter değerleri yer almaktadır. MAPE oranları %7,6685 değeriyle SmoReg algoritması en düşük değere sahiptir. %7,967 değeriyle MLPRegressor en büyük değere sahiptir.

**Tablo 5.** İthalat tahmini test aşaması

	MAE	MAPE	RMSE	MSE
<b>Lineer Regresyon</b>	1.448.912.271	7,7223	1854909197	3.440.688.128.858.990.000
<b>MLPRegressor</b>	1512176271	7,967	1914187887	3.664.115.267.570.530.000
<b>Gaussian</b>	1458121534	7,7533	1862446599	3.468.707.332.918.260.000,00
<b>RBF Regressor</b>	1.473.043.170	7,799	1.898.970.415	3.606.088.638.339.180.000,00
<b>SmoReg</b>	1.421.497.791	7,6685	1835814987	3.370.216.666.640.010.000,00

#### 4. Sonuç

Türkiye ihracat tahmininde test aşaması kriter değeri MAPE oranları %8,6987 değeriyle SmoReg algoritması en düşük değeri vermiştir. Aynı şekilde Türkiye ithalat tahmininde test aşaması kriter değeri MAPE oranları %7,6685 değeriyle SmoReg algoritması vermiştir.

Çalışma MAPE kriterine göre Türkiye’nin ithalat verilerinin ihracat verilerine göre daha düşük hata oranıyla tahmin edilebildiğini göstermektedir.

#### Çıkar Çatışması Beyanı

Makale yazarları herhangi bir çıkar çatışması olmadığını beyan eder.

#### Araştırmacıların Katkı Oranı Beyan Özeti

Yazarlar makaleye eşit oranda katkı sağlamış olduğunu beyan eder.

#### Kaynakça

- Abuzir S., Abuzir Y. Data mining for CO2 emissions prediction In Italy. Mühendislik Bilimleri ve Araştırmaları Dergisi 2021; 3(1): 59-68.
- Acı M., Ayyıldız Doğansoy G. Makine öğrenmesi ve derin öğrenme yöntemleri kullanılarak e-perakende sektörüne yönelik talep tahmini. Gazi Üniversitesi Mühendislik Mimarlık Fakültesi Dergisi 2022; 37(3): 1325-1340.
- Akgün K., Bulut Özek M. Eğitsel veri madenciliği yöntemi ile ilgili yapılmış çalışmaların incelenmesi: İçerik analizi. Uluslararası Eğitim Bilim ve Teknoloji Dergisi 2020; 6(3):197-213.
- Akgün Z. Veri madenciliği ile yazılım hata tespiti. El-Cezeri 2016; 3(2). DOI: 10.31202/ecjse.264197
- Aksu G., Doğan N. Veri madenciliğinde kullanılan güncel bir analiz programı: WEKA. Journal of Measurement and Evaluation in Education and Psychology 2019; 10(1): 80-95.
- Alaeddinoğlu M., Aydın T., Dal D. Birliktelik kuralları ile mekânsal-zamansal veri madenciliği. Erzincan University Journal of Science and Technology 2014; 5(2): 191-212.

- Albayrak YS., Yılmaz ÖK. Veri madenciliği: Karar ağacı algoritmaları ve İMKB verileri üzerine bir uygulama. Süleyman Demirel Üniversitesi İktisadi ve İdari Bilimler Fakültesi Dergisi 2009; 14(1): 31-52.
- Artsın M. Veri madenciliği ve bilgi keşfi (Kitap özeti). Açıköğretim Uygulamaları ve Araştırmaları Dergisi 2019; 5(3): 174-180
- Aslan N., Terzi N. Gelişmiş ve gelişen ülkelerde uluslararası ticaret- ücret ilişkisi ve Türkiye uygulaması. Öneri Dergisi 2007; 7(27): 217-233.
- Atsalakis G., Frantzis D., Zopounidis C. Energy's exports forecasting by a neuro-fuzzy controller. Energy Syst 2015; 6: 249–267.
- Ayanoğlu CC., Kurt M. Metal Sektöründe veri madenciliği yöntemleri ile bir iş kazası tahmin modeli önerisi. Ergonomi 2019; 2(2): 78-87.
- Aydemir E. Ders geçme notlarının veri madenciliği yöntemleriyle tahmin edilmesi. Avrupa Bilim ve Teknoloji Dergisi 2019; 15: 70-76.
- Aydemir E., Kaysi F., Yavuz M. İlaç satış verileri kullanılarak ağaç algoritmaları ile elde edilen gelirin tahmin edilmesi. Computer Science 2020; 5(1): 14-21.
- Basak D., Pal S., Patranabis D. Support vector regression, neural information processing – Letters and Reviews 2007; 11(10): 203-224.
- Bayat M. Uluslararası ticarete rekabet gücü elde etmede küçük ve orta boy işletmelerin rolü ve önemi/ Mustafa Kemal Üniversitesi Sosyal Bilimler Enstitüsü Dergisi 2013; 8(16): 567-575.
- Boonyakunakorn P., Pastpipatkul P., Sriboonchitta S. Forecasting Thailand's Exports to ASEAN with non-linear models. In: Kreinovich, V., Sriboonchitta, S., Chakpitak, N. (eds) Predictive Econometrics and Big Data. TES 2018. Studies in Computational Intelligence 2018;753. doi.org/10.1007/978-3-319-70942-0\_24.
- Cortes C., Vapnik V. Support vector networks. Machine Learning 1995; 20(3): 273–297.
- Dave E., Leonardo A., Jeanice M. Forecasting Indonesia exports using a hybrid model ARIMA-LSTM. 5th International Conference on Computer Science and Computational Intelligence (ICCSCI), 2021;179: 480-487.
- David JCM. Introduction to gaussian processes. Dept. Of Physics, Cambridge University 1998. www.inference.org.uk/mackay/gpB.pdf.
- Demirci E., Karaatlı M. Ülkelerin gelişmişlik seviyelerinin tahmininde kullanılan sınıflandırma algoritmalarının karşılaştırılması. Süleyman Demirel Üniversitesi İktisadi ve İdari Bilimler Fakültesi Dergisi 2019; 24(3): 703-714.
- Durmuşoğlu A. Veri madenciliği çalışmaları üzerine bir analiz: Türkiye adresli yayınlar. Elektronik Sosyal Bilimler Dergisi 2017; 16(62): 1034-1047.
- Eckert F., Hyndman RJ., Panagiotelis A. Forecasting swiss exports using bayesian forecast reconciliation. European Journal of Operational Research 2021; 291(2): 693-710. 10.1016/j.ejor.2020.09.046.

- Eibe Frank Mark AH., Ian HW. The WEKA workbench. Online Appendix for "Data Mining: Practical Machine Learning Tools and Techniques". Morgan Kaufmann, 2016.
- Emeksiz C., Doğan Z., Gökrem L., Yavuz AH. Tokat Bölgesi rüzgâr karakteristiğinin istatistiksel yöntemler ile incelenmesi. *Politeknik Dergisi* 2016; 19(4): 481-489.
- Emre İE., Selçukcan Erol Ç. Veri analizinde istatistik mi veri madenciliği mi? *Bilişim Teknolojileri Dergisi* 2017; 10(2): 161-167.
- Garcia Rodriguez MJ., Rodriguez Montequin V., Aranguren Ubierna A., Santana Hermida R., Sierra Araujo B., Zelaia Jauregi A. Award price estimator for public procurement auctions using machine learning algorithms: Case study with tenders from Spain, *Studies in Informatics and Control* 2021; 30(4): 67-76.
- GaussianProcesses. GaussianProcesses weka.classifiers.functions fonksiyonu. (Erişim tarihi: <https://pentaho-community.atlassian.net/wiki/spaces/DATAMINING/pages/284461153/GaussianProcesses>), (Erişim tarihi: 30 Ağustos 2022).
- George-Nektarios T. Weka classifiers summary. Athens University of Economics and Business Intracom-Telecom, Athens. 2013.
- Ghauri SP., Ahmed RR., Streimikiene D., Streimikis, J. Forecasting exports and imports by using autoregressive (ar) with seasonal dummies and box-jenkins approaches: A case of Pakistan, *Inzinerine Ekonomika-Engineering Economics* 2020; 31(3): 291-301..
- Gürsoy UT., Bilgin Ş. Banka müşterilerinin internet bankacılığına ilişkin yaklaşımlarının veri madenciliği teknikleri ile incelenmesi. *Kafkas Üniversitesi İktisadi ve İdari Bilimler Fakültesi Dergisi* 2017; 7(14): 421-442.
- Han ZA., Zhu ZJ., Zhao SJY., Dai WH. Research on nonlinear forecast and influencing factors of foreign trade export based on support vector neural network, *Neural Computing & Applications*, 2022; 34(4): 2611-2622.
- Irmak S., Köksal CD., Asilkan Ö. Hastanelerin gelecekteki hasta yoğunluklarının veri madenciliği yöntemleri ile tahmin edilmesi. *Uluslararası Alanya İşletme Fakültesi Dergisi* 2012; 4(1): 101-114.
- Kılıç F., Akkaya MR., Memili N. Yemekhane için yapay zekâ teknikleri kullanımı ile günlük talep tahmini. *Avrupa Bilim ve Teknoloji Dergisi* 2018; 13: 65-71.
- Kılıç S. Doğrusal regresyon analizi. *Journal of Mood Disorders* 2013; 3(2): 90-92.
- Kitessa BD., Ayalew SM., Gebrie GS., Teferi ST. Long-term water-energy demand prediction using a regression model: a case study of Addis Ababa city. *Journal of Water and Climate Change* 2021; 12(6): 2555-2578.
- Kriegel HP., Borgwardt KM., Kroger P., Pryakhin A., Schubert M., Zimek A. Future trends in data mining, *Data Mining and Knowledge Discovery* 2007; 15(1): 87-97.
- Lehmann R. Forecasting exports across Europe: What are the superior survey indicators?. *Empirical Economics* 2020; 60(5): 2428-2459.

- Lewis CD. Industrial and business forecasting methods. Butterworths Scientific. London;1982.
- Linear Regression. Linear Regression weka.classifiers.functions fonksiyonu. (Erişim tarihi: <https://pentaho-community.atlassian.net/wiki/spaces/DATAMINING/pages/284461164/LinearRegression>). (Erişim tarihi: 30 Ağustos 2022).
- Liu Y. Foreign trade export forecast based on fuzzy neural network. Complexity. 2021. 10.1155/2021/5523222.
- Mendo L. Estimation of a probability with guaranteed normalized mean absolute error. IEEE Communications Letters 2009; 13(11): 817-819.
- MLPRegressor. Weka sınıflandırma fonksiyonları (weka.classifiers.functions, Class MLPRegressor), <https://weka.sourceforge.io/doc/packages/multiLayerPerceptrons/weka/classifiers/functions/MLPRegressor.html>, (Erişim: 3.12.2022)
- Multilayer Perceptron. Multilayer Perceptron weka.classifiers.functions , <https://pentaho-community.atlassian.net/wiki/spaces/DATAMINING/pages/284461168/MultilayerPerceptron>, (Erişim: 30.08.2022).
- Myttenaere AD., Golden B., Grand BL., Rossi F., Mean absolute percentage error for regression models. Neurocomputing 2016; 92: 38-48.
- Nagy G. Sector based linear regression, a new robust method for the multiple linear regression. Acta Cybernetica 2018; 23(4): 1017-1038.
- Namlı E., Ünlü R., Gül E. Fiyat tahminlemede makine öğrenmesi teknikleri ve doğrusal regresyon yöntemlerinin kıyaslanması; Türkiye’de satılan ikinci el araç fiyatlarının tahminlenmesine yönelik bir vaka çalışması. Konya Mühendislik Bilimleri Dergisi 2019; 7(4): 806-821.
- Narayan PK., Narayan S., Prasad BC. Forecasting Fiji's exports and imports, 2003-2020, International Journal of Social Economics 2008; 35(12): 1005.
- Özari Ç., Demirkale Ö. K-en yakın komşu algoritması ile dolar-TL ve Euro-TL kuru kullanarak borsa endeks tahmini. maliye ve finans yazıları 2022; 117: 41-62.
- Özbay Ö. Veri madenciliği kavramı ve eğitimde veri madenciliği uygulamaları. Uluslararası Eğitim Bilimleri Dergisi 2015; 5: 262-272.
- Özdemir A., Saylam R., Bilen BB. Eğitim sisteminde veri madenciliği uygulamaları ve farkındalık üzerine bir durum çalışması. Atatürk Üniversitesi Sosyal Bilimler Enstitüsü Dergisi 2018; 2159-2172.
- Özel C., Topsakal A. Veri madenciliği kullanarak beton basınç dayanımının belirlenmesi. Cumhuriyet Üniversitesi Fen Edebiyat Fakültesi Fen Bilimleri Dergisi 2014; 35(1): 1-11.
- Pereira F., Aires-de-Sousa J. Machine learning for the prediction of molecular dipole moments obtained by density functional theory. Journal of cheminformatics, 2018; 10(1): 1-11.
- Platt JC. Fast training of support vector machines using sequential minimal optimization. In Advances in Kernel Methods: Support Vector Machines, ed. B. Scholkopf, C. Burges, and A. Smola, Cambridge, MA: MIT Press;1998.



- Purlu M., Turkyay BE. Estimating the distributed generation unit sizing and its effects on the distribution system by using machine learning methods. *Elektronika Ir Elektrotehnika* 2021; 27(4): 24-32.
- Rasmussen CE. Gaussian processes in machine learning. In: Bousquet, O., von Luxburg, U., Rätsch, G. (eds) *Advanced Lectures on Machine Learning. ML 2003. Lecture Notes in Computer Science* 2004;3176. doi.org/10.1007/978-3-540-28650-9\_4.
- Raza H., Manarvi I., Ahmed J., Khan K., Rehman K. A methodology of export sectors identification through data mining. *International Conference on Computers & Industrial Engineering* 2009: 1496-1499.
- Savaş S., Topaloğlu N., Yılmaz M. Veri madenciliği ve Türkiye'deki uygulama örnekleri. *İstanbul Ticaret Üniversitesi Fen Bilimleri Dergisi* 2012; 11(21): 1-23.
- Sebik NB., Bülbül, Hİ. Veri madenciliği modellerinin akciğer kanseri veri seti üzerinde başarılarının incelenmesi. *TÜBAV Bilim Dergisi* 2018; 11(3): 1-7.
- Shevade SK., Keerthi SS., Bhattacharyya C., Murthy KKK. Improvements to the SMO Algorithm for SVM Regression. *IEEE Transactions on Neural Networks* 2000; 11(5): 1188-1193.
- Smith S., Sincich T. Stability over time in the distribution of population forecast errors. *Demography* 1988; 25: 461-474.
- Smola AJ., Scholkopf B. A tutorial on support vector regression, *Statistics and Computing* 2004; 14: 199-222.
- SMOreg. SMOreg weka.classifiers.functions fonksiyonu. (Erişim tarihi: <https://pentaho-community.atlassian.net/wiki/spaces/DATAMINING/pages/284461182/SMOreg>), (Erişim tarihi: 30 Ağustos 2022).
- Sohrabpour V., Oghazi P., Toorajipour R. Nazarpour A. Export sales forecasting using artificial intelligence. *Technological Forecasting and Social Change* 2021; 163. 10.1016/j.techfore.2020.120480.
- Subaşı Ertekin M. Türkiye'de 1980 sonrasında döviz kuru politikaları ve dış ticaret. *Afyon Kocatepe Üniversitesi İktisadi ve İdari Bilimler Fakültesi Dergisi* 2001; 3(1): 183-198.
- TUİK, Türkiye İstatistik Kurumu. <https://data.tuik.gov.tr/Kategori/GetKategori?p=dis-ticaret-104>, (Erişim: 29.08.2022).
- Tütünsatar A. Zengezur koridorunun uluslararası ticaret politikası ve lojistik açıdan önemi. *Türkiye Mesleki ve Sosyal Bilimler Dergisi* 2022; 8: 66-74.
- Urfaloğlu M., Şahin S. Veri madenciliği ile özel öğrenme güçlüğü tahmini. *Türkiye Bilimsel Araştırmalar Dergisi* 2022; 7(1): 115-122.
- Urrutia JD., Abdul AM., Atienza JBE. Forecasting Philippines imports and exports using bayesian artificial neural network and autoregressive integrated moving average. 8th SEAMS-UGM International Conference on Mathematics and Its Applications, AIP Conference Proceedings, 2019;2192.

- Uyar Erdem Z., Çalış B., Fırat SÜ. Customer Churn prediction analysis in a telecommunication company with machine learning algorithms. *Endüstri Mühendisliği* 2021; 32(3): 496-512.
- Uzun R., İşler Y., Toksan M. WEKA yazılım paketinin siğil tedavi yöntemlerinin başarısının tahmininde kullanımı. *Düzce Üniversitesi Bilim ve Teknoloji Dergisi* 2019; 7(1): 699-708.
- Wang L., Tan L., Yu C., Wu Z. Study and application of non-linear time series prediction in ground source heat pump system. *2nd International Conference on Consumer Electronics, Communications and Networks (CECNet)* 2012:3522-3525. 10.1109/CECNet.2012.6201751.
- Wang Y. Import and export trade forecasting algorithm based on blockchain security and PSO optimized hybrid RVM model, *Security and Privacy* 2022.
- Weisberg S. *Applied linear regression*, Wiley. New Jersey;2005.
- Willmott CJ., Matsuura K. Advantages of the mean absolute error (MAE) over the root mean square error (RMSE) in assessing average model performance, *Climate Research* 2005; 30(1): 79-82.
- Witten IH., Frank E. *Data mining practical machine learning tools and techniques*. Morgan Kaufmann; 2005.
- Witten IH., Frank E., Hall MA. Pal CJ. *Data mining practical machine learning tools and techniques* Morgan Kaufmann; 2017. ISBN: 978-0-12-804291-5.
- Wong HL., Tu YH., Wang CC. Application of fuzzy time series models for forecasting the amount of Taiwan export, *Expert Systems with Applications* 2010; 37(2): 1465-1470.
- Yu Y. GDP Economic forecasting model based on improved RBF neural network. *Mathematical Problems in Engineering* 2022. doi.org/10.1155/2022/7630268.
- Yuan YL. Forecasting method for import and export trade on the basis of GMDH network model, *Journal of Discrete Mathematical Sciences and Cryptography* 2017; 20(4): 755-766. 10.1080/09720529.2017.1358859.
- Yurduseven N., Müngen AA. Finansal zaman serilerini tahminlemede kullanılan yöntemlere genel bir bakış. *Bilecik Şeyh Edebali Üniversitesi Fen Bilimleri Dergisi* 2022; 9(1): 653-671.
- Zhang DB., Zhu H., Zhang JG. Forecasting of customs export based on gray theory. *2nd International Conference on Business Intelligence and Financial Engineering* Jul 24-26 Beijing China. 2009:630-633. 10.1109/BIFE.2009.148.