

Makine Öğrenimi ile Ticaretin Öngörülmesi: Türkiye ve Balkan Ülkeleri Üzerine Çekim Modeli Analizi

Forecasting Trade with Machine Learning: A Gravity Model Analysis of Türkiye and Balkan Countries

Haldun SOYDAL *

Mustafa AY **

Sümeyye KOÇ ***

ÖZ

Bu çalışmada, Türkiye'nin Balkan ülkeleri ile olan ticaret hacmi çekim modeli kullanılarak analiz edilmiş ve Türkiye ile bu ülkeler arasındaki ticaret hacmini tahmin etmede makine öğrenimi yöntemlerinin performansları karşılaştırılmak istenmiştir. Bu amaçla çalışmada 2004 yılından başlayarak 2023 yılına kadar olan veriler kullanılmıştır. Ancak Sırbistan ve Karadağ'ın 2006 yılında bağımsız iki ülke hâline gelmesi, Kosova'nın da 2008 yılında Sırbistan'dan tek taraflı bağımsızlığını ilan etmesi nedeniyle verilerinin incelenmesi sonucu görülen devamsız ve sağlıksız yapı sebebi ile bu üç ülke analiz dışında bırakılmıştır. Çalışmada kullanılan veriler sayısal değişkenler olarak Türkiye ile Balkan ülkeleri arasındaki ihracat ve ithalat verileri, ülkelerin milli gelirleri, aralarındaki mesafe; kukla değişkenler olarak ise ülkelerin birbirleriyle olan sınırı, ortak dil, ülkelerin karayla çevriliilik durumu, Dünya Ticaret Örgütü (DTÖ) üyelikleri ve Avrupa gümrük birliği üyelikleridir. Bu veriler, ticaret hacmini tahmin etmek için yedi farklı makine öğrenmesi modeli ile analiz edilmiştir. Analiz dört farklı eğitim-test veri seti bölünmesi ve 10 farklı iterasyon ile çapraz doğrulama yoluyla uygulanmıştır. Uygulanan makine öğrenmesi modellerinin başarısı MAPE (Ortalama Mutlak Yüzde Hata) ve R2 değerleri üzerinden kıyaslanmıştır. Bulgular, en iyi tahmin yönteminin analize konu ülkeye ve eğitim-test ayırım oranlarına göre değişim gösterdiğini ortaya koymuştur. Sonuçlar arasındaki bu farklılık, Türkiye ve Balkan ülkeleri arasındaki ticaret ilişkilerinin daha iyi anlaşılması, bu ilişkilerin gelecekteki seyrinin tahmin edilmesi ve bölgesel ekonomik politikaların oluşturulmasına önemli katkılar sağlayacaktır. Bu sebeple çalışma, bu ülkelerle olan ticaretin gelişimine yönelik stratejilerin belirlenmesi açısından önemlidir.

ANAHTAR KELİMELELER

Çekim Modeli, Gradyan Artırma, İkili Ticaret, Karar Ağacı, Makine Öğrenmesi

ABSTRACT

This study analyzes Türkiye's trade volume with Balkan countries using the gravity model and aims to compare the performance of machine learning methods in predicting the trade volume between Türkiye and these countries. For this purpose, data from 2004 to 2023 have been utilized. However, due to the dissolution of Serbia and Montenegro into two independent countries in 2006, and Kosovo's unilateral declaration of independence from Serbia in 2008, the data for these three countries exhibit discontinuous and inconsistent patterns, leading to their exclusion from the analysis. The data used in the study includes numerical variables such as the export and import figures between Türkiye and Balkan countries, the GDP of the countries, and the distance between them. The dummy variables include whether the countries share a border, a common language, their landlocked status, and their memberships in the World Trade Organization (WTO) and the European Customs Union. These data have been analyzed using seven different machine learning models to predict the trade volume. The analysis was conducted using four different train-test data splits and cross-validation with 10 iterations. The performance of the applied machine learning models was compared based on MAPE (Mean Absolute Percentage Error) and R2 values. The findings revealed that the best prediction method varies depending on the country analyzed and the train-test split ratios. These differences in results contribute significantly to a better understanding of trade relations between Türkiye and the Balkan countries, forecasting the future course of these relations, and shaping regional economic policies. Therefore, the study is important in determining strategies for the development of trade with these countries.

KEYWORDS

Gravity Model, Gradient Boosting, Bilateral Trade, Decision Tree, Machine Learning

Makale Geliş Tarihi / Submission Date	Makale Kabul Tarihi / Date of Acceptance
09.09.2024	07.11.2024
Atf	Soydal, H., Ay, M. ve Koç, S. (2024). Makine Öğrenimi ile Ticaretin Öngörülmesi: Türkiye ve Balkan Ülkeleri Üzerine Çekim Modeli Analizi. <i>Selçuk Üniversitesi Sosyal Bilimler Meslek Yüksekokulu Dergisi</i> , 27 (2), 746-765.

* Prof. Dr., Selçuk Üniversitesi, İktisadi ve İdari Bilimler Fakültesi, hsoydal@selcuk.edu.tr, ORCID: 0000-0001-6979-0256

** Arş. Gör., Selçuk Üniversitesi, İktisadi ve İdari Bilimler Fakültesi, mustafa.ay@selcuk.edu.tr, ORCID: 0000-0001-8635-6101

*** Doktora Öğrencisi, sumeyyekoc350@gmail.com, ORCID: 0000-0003-4255-3562

GİRİŞ

Yapay zeka 1950'li yıllarda ortaya çıkmış ve günümüzde insan yaşamının hemen her alanında uygulamaları olan multidisipliner bir alan hâline gelmiştir. Makine öğrenimi, veri örnekleri veya geçmiş deneyimler arasındaki kalıpları tespit etmede belirli bir performans düzeyi elde etmek için bilgisayarların programlanmasını içerir. Bu süreçte makine öğrenimi geleceğe yönelik tahminlerde bulunmak, verilerden içgörü elde etmek veya her ikisi için de kullanılabilir (Alpaydin, 2014; Kulkarni ve More, 2016).

Ekonomide her iki amaç için de makine öğrenimi uygulamalarının birçok örneği vardır. Bunlardan biri de birçok ülke için önemli bir araç olan makroekonomik değişkenlerin tahminidir. Yapay zeka ve makine öğrenimi, makroekonomi ve uluslararası ticaret için önemli araçlar olarak kabul edilmektedir (Goldfarb ve Trefler, 2018). Bu nedenle uluslararası ticareti tahmin etmek için yapay zeka ve makine öğreniminden de yararlanılmaktadır. Bu alanda Türkiye için öncü çalışma, ana ticaret ortaklarıyla ikili ticaret akışlarını analiz eden Nuroğlu (2012) tarafından yürütülmüştür. Çalışmada 1985'ten 2010'a kadar olan yıllık veriler kullanılarak panel veri modelleri ve sinir ağları kullanılmış ve elde edilen sonuçlara göre sinir ağları, Türkiye'nin ikili ticaret akışlarının analizinde geleneksel panel veri modellerine göre üstünlük göstermiştir.

Bu çalışma, Türkiye ile Balkan ülkeleri arasındaki ikili dış ticaret akışlarının tahmin edilmesinde makine öğrenmesi algoritmalarının performansını değerlendirmeyi amaçlamaktadır. Analizdeki girdi değişkenleri ihracatçıların GSYH'si, ithalatçıların GSYH'si, mesafe ve kukla değişkenler (dil, AB gümrük birliği üyeliği, sınır, Karayla çevrili olma ve DTÖ üyeliği), çıktı değişkenleri ise ithalat ve ihracat değerleridir. Veriler, 2004-2023 dönemi için toplanmıştır ve Türkiye'nin ikili ticaret akışlarını tahmin etmede en başarılı modeli tespit etmek için kullanılmıştır.

Tahmin algoritmalarının doğruluğu, ortalama mutlak yüzde hata (MAPE) ve R2 hesaplanarak incelenmiştir. Çalışma kapsamında Türkiye-Balkan ülkeleri ikili ticaret akışlarını tahmin etmek için yedi farklı makine öğrenimi algoritması (Lineer Regresyon, Karar Ağacı, Rastgele Orman, Bagging Karar Ağacı, Extratrees, Gradyan Artırımı, Ekstra Gradyan Artırımı) kullanılarak performansları karşılaştırılmıştır.

Çalışma giriş, teorik çerçeve, Türkiye'nin ticaret akışı ve Balkan ülkeleri ile ikili ticareti, yöntemlerin tanıtılması, bu yöntemlerle yapılan analizlerin bulguları ve sonuç şeklindedir ve altı bölümden oluşmaktadır.

1. TEORİK ÇERÇEVE

Gravity Model, uluslararası ticaret akımlarını açıklamak için kullanılan en temel modellerden biridir. Newton'un kütle çekim yasasından ilham alınarak geliştirilen bu model, iki ülke arasındaki ticaret hacminin ülkelerin ekonomik büyüklükleri ile doğru orantılı, aralarındaki mesafe ile ters orantılı olduğunu öngörür. Modelin temelleri, ilk olarak Tinbergen (1962) ve Pöyhönen (1963) tarafından atılmıştır. Bu teorik çerçeve, Gravity Model'in gelişimi, temel unsurları ve çeşitli genişletmeleri hakkında ayrıntılı bir bakış sunmaktadır.

Gravity Model'e göre, iki ülke arasındaki ticaret hacmi, ülkelerin ekonomik büyüklükleri (genellikle Gayri Safi Yurtiçi Hasıla - GDP) ile doğru orantılıdır. Ekonomik büyüklük, bir ülkenin üretim kapasitesini ve ticari potansiyelini yansıtır. Daha büyük ekonomiler, daha fazla mal ve hizmet üretip tüketebilme kapasitesine sahip oldukları için, ticaret hacimlerinin de büyük olması beklenir (Bergstrand, 1985). Modelde mesafe, iki ülke arasındaki ticaret maliyetlerinin bir göstergesi olarak kullanılır. Mesafe arttıkça, ulaşım ve diğer lojistik maliyetler artar ve bu da ticaret hacmini azaltır. Mesafenin sadece fiziksel uzaklıkla değil, aynı zamanda kültürel ve kurumsal farklılıklarla da ilişkili olduğu belirtilmiştir (Linnemann, 1966).

Gravity Model'in temel matematiksel formülasyonu genellikle şu şekilde ifade edilir:

$$Ticaret_{ij} = \frac{A \times GSYİH_i \times GSYİH_j}{Mesafe_{ij}}$$

Yukarıda $GSYİH_i$ ve $GSYİH_j$ ticaret yapan iki ülkenin milli geliri, $Mesafe_{ij}$ ülkeler arası mesafeyi, A ise genel ticaretin seviyesini belirleyen bir çarpandır.

Gravity Model'in başlangıçtaki basit formülü, sadece ekonomik büyüklük ve mesafeyi içerirken, zamanla çeşitli genişletmelerle daha sofistike hâle getirilmiştir. Anderson ve van Wincoop (2003), ticaret maliyetlerini dikkate alarak modelin mikroekonomik temellerini güçlendirmiş ve ticaret maliyetlerinin tarifeler, kotalar ve diğer ticaret engellerini içerdiğini vurgulamıştır. Serbest ticaret anlaşmaları (FTA'lar) da modele dahil edilmiş ve bu anlaşmaların ticaret bariyerlerini azaltarak iki ülke arasındaki ticareti artırabileceği gösterilmiştir (Baier ve Bergstrand, 2007).

Bunun yanı sıra, ülkeler arasındaki ortak dil ve kültürel bağlar da ticaret hacmini etkileyen faktörler olarak modele eklenmiştir. Bu faktörler, iletişim maliyetlerini düşürerek ticari işlemleri kolaylaştırır (Melitz, 2008). McCallum (1995) ise fiziksel sınırların ticaret üzerindeki etkisini incelemiş ve sınırların, gümrük prosedürleri ve düzenleyici farklılıklar nedeniyle ticareti azaltabileceğini belirtmiştir. Ayrıca Helpman vd. (2008), ülkelerin

pazar potansiyeli ve ekonomik yapılarının benzerliğinin de ticaret akımlarını etkileyebileceğini, benzer ekonomik yapılar arasındaki ticaretin daha güçlü olabileceğini göstermiştir.

2. LİTERATÜR TARAMASI

Literatürde çekim modeli çerçevesinde ileriye dönük tahmine ilişkin geleneksel ve gelişmiş yöntemler ile yapılmış birçok çalışma bulunmaktadır. Bu çalışma konusu itibarıyla literatür incelemesi yapılırken makine öğrenimi yöntemlerinin kullanıldığı çalışmalara ek olarak geleneksel yöntemler ile makine öğrenimi yöntemlerinin tahmin konusundaki performanslarının kıyaslandığı çalışmalara da yer verilmiştir.

Nuroğlu (2012) tarafından 15 AB üyesi ülke için ikili ticaret tahminine yönelik yapılan çalışma literatürde özellikle yapay sinir ağlarının kullanıldığı önde gelen çalışmalardan biridir. Çalışmada AB üyesi 15 ülke için 1964-2003 yılları arasındaki veriler kullanılmıştır. Çalışmada panel veri analiz yöntemleri ve yapay sinir ağlarının ileriye yönelik ticaret tahminindeki başarısı karşılaştırılmıştır. Çalışmanın bulguları yapay sinir ağlarının, geleceğe yönelik ticaret tahmini yaparken geleneksel panel veri analizinden daha başarılı sonuçlar ürettiğini göstermiştir.

Circlaeys vd. (2017) tarafından yapılan çalışmada dünya ülkelerinin 1800'lü yıllardan 2014 yılına kadar olan 1,9 milyon verilik ikili ticaret veriseti kullanılmıştır. Çalışmada geleneksel yöntemler aracılığıyla farklı lineer regresyon uygulamalarına ek olarak gecikmeli çıktı verileri kullanarak tam bağlantılı yapay sinir ağları yöntemi olarak uygulanmıştır. Çalışma sonucunda elde edilen bulgulara göre en başarılı tahmin, yapay sinir ağları tarafından üretilmiştir.

Wohl ve Kennedy tarafından 2018 yılında ortaya konulan çalışmada 68 ülkenin 1986-2006 yılları arasındaki verileri kullanılarak geleceğe dönük ikili ticaret ilişkisi tahmin edilmek istenmiştir. Bu amaçla OLS, PPML ve yapay sinir ağları yöntemleri kullanılmıştır. Analiz sonucu ortaya konulan bulgular, ileriye dönük ikili ticaret tahmininde yapay sinir ağlarının diğer yöntemlere göre daha başarılı sonuçlar ürettiği belirtilmiştir.

2018 yılında ortaya konulan bir diğer çalışma da Quimba ve Barral tarafından 21 APEC ülkesi için gerçekleştirilen analize ilişkindir. Çalışmada kullanılan veriseti 1996-2001 yılları arasındaki verileri içermektedir. Quimba ve Barral, APEC ülkeleri için ileriye yönelik ticaret tahmininde OLS, PPML, GPML ve yapay sinir ağları yöntemlerini kullanmış ve bu analiz yöntemlerinin performanslarını kıyaslamıştır. Çalışmanın ortaya koyduğu bulgular, yapay sinir ağlarının diğer yöntemlere göre daha iyi tahmin sonuçları elde ettiğini göstermiştir.

Kottou vd. (2020) tarafından yapılan ileriye yönelik ikili ticaret ilişkisi tahminlemede GSR, Karar Ağacı, LSTM ve CNN yöntemleri kullanılmış ve bu yöntemlerin tahmin başarıları karşılaştırılmak istenmiştir. Çalışmada 1827-2014 yılları arasındaki ticaret verilerinden oluşan 2,7 milyon verilik bir veriseti kullanılmıştır. Bu çalışmanın analiz süreci sonunda elde edilen bulgulara göre, karar ağacı yönteminin geleceğe dönük ikili ticaret tahmininde diğer yöntemlere göre daha başarılı tahminler yaptığı görülmüştür.

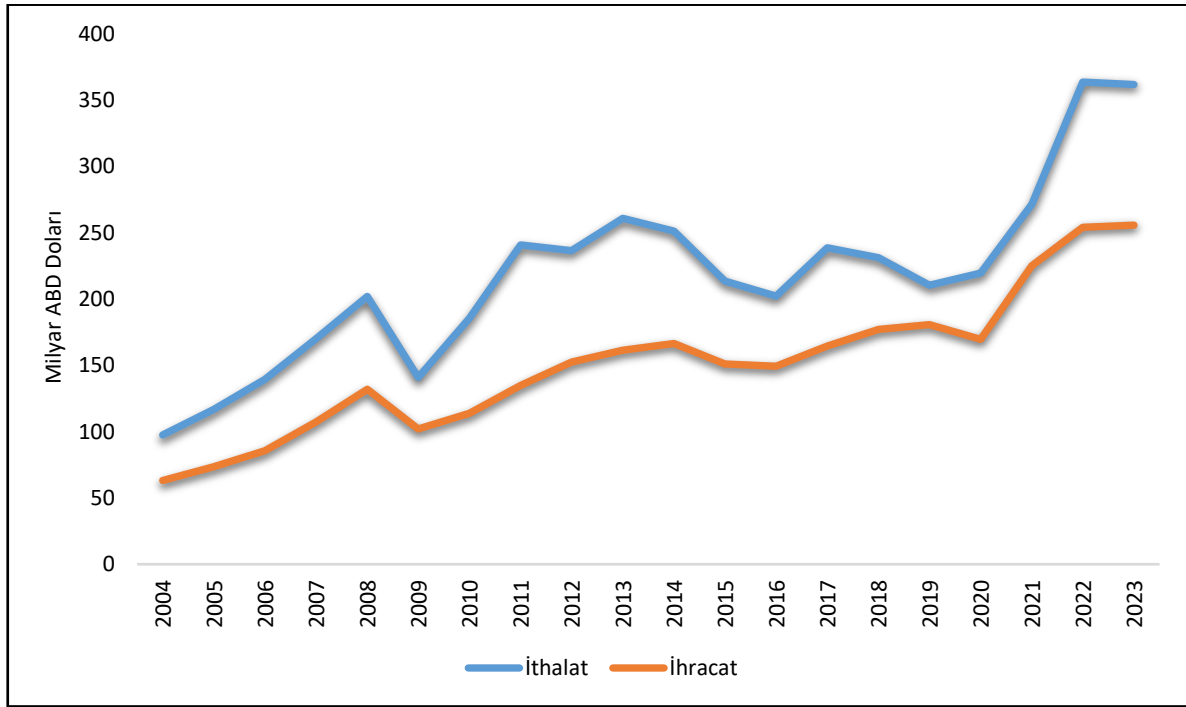
Jošić ve Žmuk (2022) tarafından yürütülen çalışmada Hırvatistan'ın 2001-2019 yılları arasındaki dış ticaret verileri ele alınmış ve ikili ticaret tahmininde lineer regresyon, GSR ve çok katmanlı algılayıcılar kullanılarak tahminler üretilmiştir. Bu çalışmadaki analiz sonuçları yöntemlerin en başarılı olduğu ülke sayısına göre gruplandırılmıştır. Gruplama sonucunda çok katmanlı algılayıcıların, uygulanan diğer tahmin yöntemlerine göre daha fazla ülke için en iyi sonuçları ortaya koyduğu görülmüştür.

3. TÜRKİYE'NİN TİCARET YAPISI VE BALKANLAR İLE İLİŞKİSİ

Türkiye, jeopolitik konumu, nüfus yoğunluğu ve büyüme potansiyeli nedeniyle uluslararası ticaret içerisinde önemli bir ülkedir. Konumu itibarıyla Avrupa ile Asya pazarları arasında bir konumda olmasının da etkisiyle Türkiye, dünyadaki stratejik ticaret merkezlerinden biri hâline gelmiştir. Türkiye'nin ticaret yapısı farklı sektörlerle dayanmaktadır. Türkiye'nin ihracatında otomotiv, tekstil, makine, gıda ürünleri ve elektronik gibi sektörler dikkat çekerken; ithalatında ise enerji, ham madde, ara mallar ve ileri teknolojiye yönelik ürünler öne çıkmaktadır. Türkiye'nin dış ticaret politikaları, küresel ekonomik entegrasyonu ve ticaret hacminin artırılmasına yönelik stratejilerle şekillenmektedir.

Türkiye'nin jeopolitik konumu ve dünya ticaretindeki yeri, serbest ticaret anlaşmaları ve bölgesel iş birliklerini de gerekli kılmıştır. Avrupa Birliği, Orta Doğu, Kuzey Afrika ve Asya ülkeleri ile olan ticari ilişkiler, Türkiye'nin ekonomik büyüme performansına ciddi katkı sunmaktadır. Bu kapsamda önce Türkiye'nin ihracat ve ithalat miktarlarının yıllar içerisindeki değişimi incelenecek (Şekil 1) ardından da Balkan ülkeleri ile ihracat (Şekil 2) ve ithalatı (Şekil 3) incelenecektir.

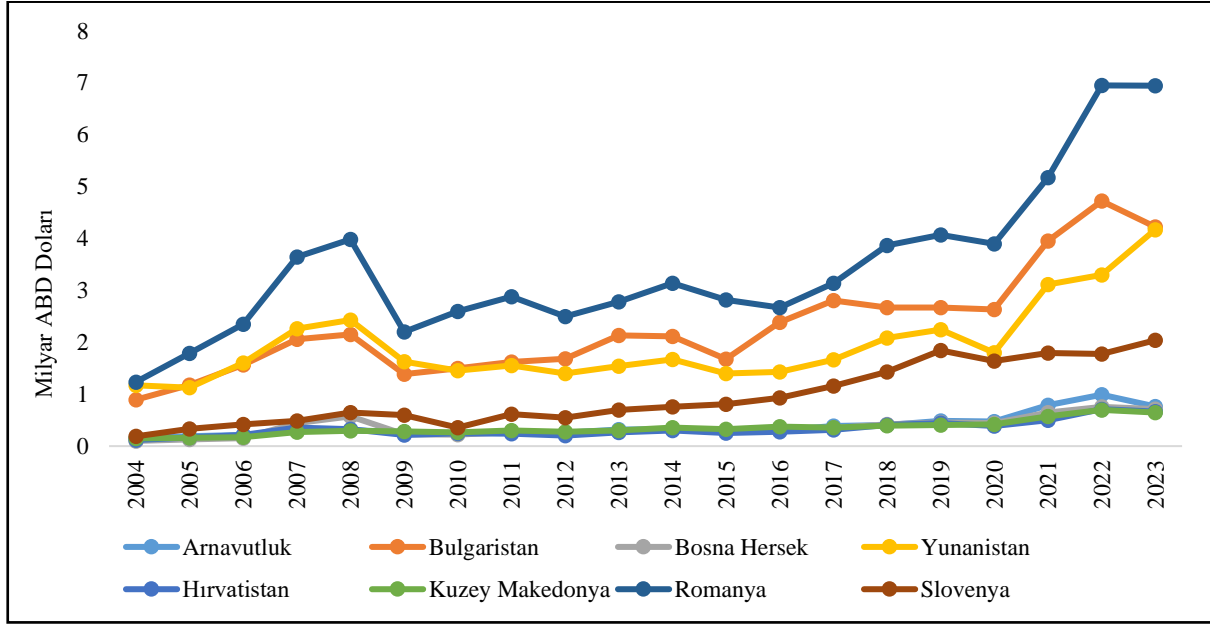
Şekil 1. Türkiye'nin İhracat ve İthalat Miktarları (1992-2023) (Milyar ABD Doları)



Kaynak: UN Comtrade (2024)

Türkiye'nin temel ticaret göstergeleri 2004-2023 yılları arasında önemli bir artış ve çeşitlenme göstermiştir. 2004 yılında 60 milyar dolar seviyesinde olan ihracat, 2023 yılında 250 milyar doları aşmıştır. Aynı dönemde ithalat ise 100 milyar dolardan yaklaşık 360 milyar dolara yükselmiştir. 2001 krizi sonrasında uygulanan yapısal reformlar sonraki dönemde hem ihracatın hem de ithalatın büyümesine önemli katkılar sağlamış ve bu büyüme 2008'e kadar devam etmiştir. 2008 küresel finans krizinin etkisiyle 2009 yılında ihracat ve ithalatta belirgin bir düşüş gerçekleşmiş, fakat sonraki yıllarda hızlı bir toparlanma sürecine girilmiştir. İhracat 2009-2019 yılları arasında 2015 ve 2016 yılları haricinde sürekli artma eğilimi göstermiştir. Buna ek olarak ithalat, hem kurdaki artış ve volatilitenin yüksekliği hem de enerji fiyatlarının dünyada düşmesinden dolayı 2014-2019 yıllarında düşme eğilimi göstermiştir. Özellikle 2014 yılına kadar ki süreçte düşük döviz kuru sebebiyle ihracattaki artış, ithalattaki artış oransal olarak yakalayamamıştır (Erdoğan ve Tek, 2019). 2020'de COVID-19 pandemisinin başlaması küresel ticarete duraklamaya yol açmış ve bundan Türkiye'nin dış ticaret göstergeleri de etkilenmiştir. Bu dönemde ihracat artarken ithalat duraksamıştır. Ancak Türkiye'nin dış ticaret göstergeleri 2021 sonrasında hızlı bir toparlanma eğilimine girmiş ve büyük bir sıçrama göstermiştir.

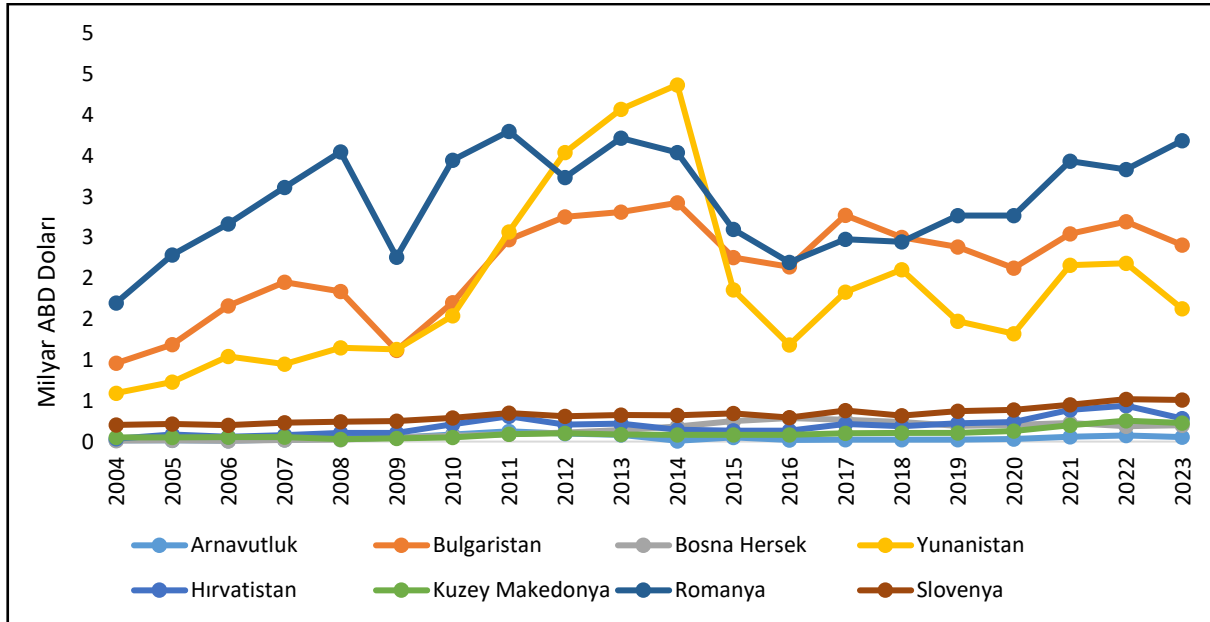
Şekil 2. Türkiye'nin Balkan Ülkelerine İhracat Miktarı (2004-2023) (Milyar ABD Doları)



Kaynak: UN Comtrade (2024)

Türkiye'nin Balkan ülkelerine yaptığı ihracat miktarlarına bakıldığında 2004 yılından 2008 yılına kadar düzenli bir artış görülmektedir. 2008 küresel krizinin etkisini gösterdiği 2009 yılında Balkan ülkelerine ihracat miktarı azalmıştır. Sonrasında ise 2020 yılına kadar ihracat miktarı yakın seviyelerde iken 2020 sonrasında bir artış yaşanmıştır. Balkan ülkeleri arasında Türkiye'nin en yüksek ihracat yaptığı üç ülke ise Romanya, Bulgaristan ve Yunanistan olurken en düşük ihracat yaptığı ülkeler ise Hırvatistan, Kuzey Makedonya ve Bosna Hersek'tir.

Şekil 3. Türkiye'nin Balkan Ülkelerinden İthalat Miktarı (2004-2023) (Milyar ABD Doları)



Kaynak: UN Comtrade (2024)

Türkiye'nin Balkan ülkelerinden yaptığı ithalat miktarı Şekil 3'te gösterilmiştir. Buna göre Türkiye'nin ihracatçı olduğu durumdaki en önemli üç ticaret ortağı, ithalatçı olduğu durumda da aynı şekilde Romanya, Bulgaristan ve Yunanistan'dır. Diğer beş Balkan ülkesinin ise ihracat miktarlarında olduğu gibi ithalat miktarında da bu üç ticaret ortağına göre oldukça geride olduğu görülmektedir.

4. EKONOMETRİK METODOLOJİ

Çalışma 2004'ten 2023'e kadar Türkiye'nin Balkan ülkeleri ile yaptığı ikili ticaret (ihracat ve ithalat) verilerini kullanarak ileriye yönelik tahmin konusunda farklı makine öğrenimi yöntemlerinin çekim modeli çerçevesinde uygulanarak tahmin başarılarını kıyaslamayı amaçlamaktadır. Modeldeki bağımsız değişkenler, ihracatçı ve ithalatçı ülkelerin GSYİH'leri, ülkeler arasındaki mesafe ve kukla değişkenlerdir (dil, sınır, karayla çevrili olma ve DTÖ üyeliği, Avrupa Gümrük Birliği üyeliği). Yukarıda belirtilen değişkenlerin birçoğu, uluslararası ticaretin çekim modeli için ana açıklayıcı değişkenlerdir. Bu nedenle, ticaretin çekim modellerinin tahmin edilmesinde sıklıkla kullanılmaktadırlar. Çalışmada bu değişkenlere ek olarak analize konu ülkelerin buldukları coğrafya sebebiyle Avrupa Gümrük Birliği üyelik durumları da makine öğrenimi algoritmaları için bağımsız değişken olarak kullanılmıştır. Bağımlı değişken olarak ise Türkiye'nin analize konu Balkan ülkeleri ile gerçekleştirdiği ihracat ve ithalat değerlerine yer verilmiştir. Kullanılan değişken, kısa açıklaması ve kullanılan verinin kaynağı Tablo 1'de gösterilmiştir.

Tablo 1. Çalışmada Kullanılan Değişkenler

Değişken	Açıklama	Veri Kaynağı
İhracat	İhracat Miktarı (Bin ABD Doları)	UN Comtrade (2024)
İthalat	İthalat Miktarı (Bin ABD Doları)	UN Comtrade (2024)
İhracatçı GSYH	İhracatçı Ülke GSYH (ABD Doları)	Dünya Bankası (2024)
İthalatçı GSYH	İthalatçı Ülke GSYH (ABD Doları)	Dünya Bankası (2024)
Dil	İhracatçı ve İthalatçı Ülke Aynı Dili Konuşuyor (0=Hayır, 1=Evet)	CEPII (2024)
Sınır	İhracatçı ve İthalatçı Ülke Sınır Komşusu (0=Hayır, 1=Evet)	CEPII (2024)
Uzaklık	İhracatçı ve İthalatçı Ülke Arasındaki Uzaklık (Km)	CEPII (2024)
Karayla Çevrililik	İhracatçı/İthalatçı Ülke Karayla Çevrili Ülke (0=Hayır, 1=Evet)	CEPII (2024)
DTÖ Üyeliği	İhracatçı ve İthalatçı Birlikte DTÖ Üyesi (0=Hayır, 1=Evet)	DTÖ (2024)
AB Gümrük Birliği Üyeliği	İhracatçı ve İthalatçı Birlikte AB Gümrük Birliği Üyesi (0=Hayır, 1=Evet)	CEPII (2024)

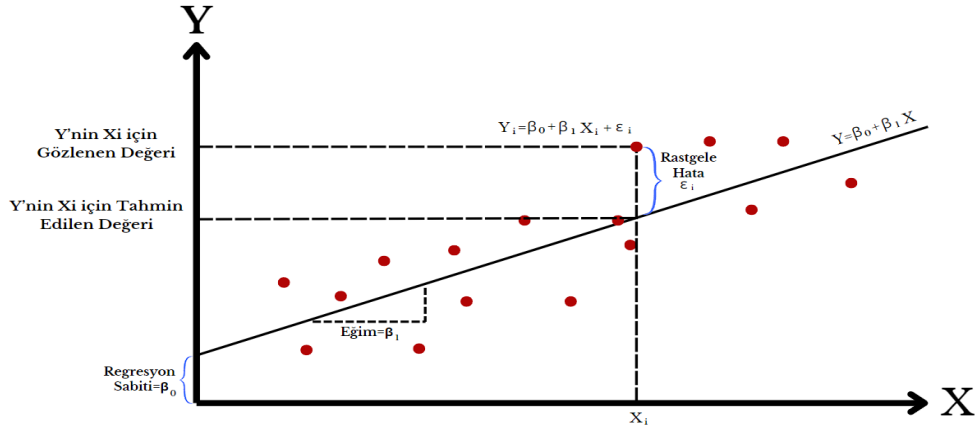
Analiz için veriler çeşitli kaynaklardan sağlanmıştır: UN Comtrade, Dünya Bankası, CEPII ve DTÖ. Çalışmada hem ihracat hem de ithalat değişkenleri bin ABD doları cinsinden ifade edilmiştir. İhracatçı GSYİH ve İthalatçı GSYİH değişkenleri ABD doları olarak analize dâhil edilmiştir. Dil değişkeni, eğer hem ihracatçı hem de ithalatçı ülkeler resmi dilleri olarak aynı dili paylaşıyorlarsa, değeri "1" (bir) olan bir kukla değişken kullanılmaktadır. Sınır değişkeni de eğer ihracatçı ve ithalatçı ülkeler komşu ülkeler ise değeri "1" (bir) olan kukla değişken kullanılmaktadır. Uzaklık, ticaret ortağı ülkelerin başkentleri arasındaki mesafeyi kilometre cinsinden ifade etmektedir. Karayla çevrililik, ortak ülkelerden birinin açık denize doğrudan erişimi yoksa kukla değişken "1" (bir) olarak alınmaktadır. İthalatçı ve ihracatçı ülkelerin her ikisi de Dünya Ticaret Örgütü üyesi ise, DTÖ kukla değişkeninin değeri "1" (bir), AB Gümrük Birliği üyesi ise EU kukla değişkeninin değeri "1" (bir) olarak alınmaktadır.

Türkiye, dünya üzerinde birçok ülke ile ticaret ortaklığına sahip olsa da bu çalışma özellikle Avrupa kıtası tarafındaki yakın komşu ülkelerini oluşturan Balkan ülkeleri ile ilişkileri üzerine bir inceleme sunmaktadır. Amaç, ikili uluslararası ticaret akışlarını kapsamlı bir şekilde tahmin edebilecek bir analiz çerçevesi sunmaktır. Türkiye'nin Balkan ülkeleri ile ikili ihracat ve ithalatının gelecekteki değerlerini tahmin etmede en başarılı yöntemi tespit etmek amacıyla, python programlama dili kullanılarak seçilen makine öğrenmesi algoritmaları (Lineer Regresyon, Karar Ağacı, Rastgele Orman, Bagging Karar Ağacı, Ekstra Ağaçlar, Gradyan Artırımı, Extra Gradyan Artırımı) uygulanmıştır.

A. Lineer Regresyon (Linear Regression)

Lineer regresyon, bağımsız değişken veya değişkenler (x_1, x_2, \dots, x_n) ile bağımlı değişkenin (y) arasındaki doğrusal ilişki yapısını anlamlandırabilmek için kullanılır. Basit lineer regresyon için değişkenler arası ilişkisel yapı aşağıdaki gibi görselleştirilir:

Şekil 4. Lineer Regresyon Yapısı



Kaynak: (Bangdiwala, 2018)

Yukarıda iki boyutlu kartezyen koordinat yapısında y bağımlı (tahmin edilmek istenen) değişkeni, x_1, x_2, \dots, x_n bağımsız (girdi) değişkenleri, β_0 regresyon sabitini, $\beta_1, \beta_2, \dots, \beta_n$ regresyon katsayılarını (eğimlerini), ϵ hata terimini ifade eder. Bu noktadan hareketle, bağımlı değişkeni açıklamak için bağımsız değişken, regresyon sabiti, regresyon katsayıları ve hata terimini içeren aşağıdaki eşitlik yapısı kullanılır:

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \dots + \beta_n x_n + \epsilon$$

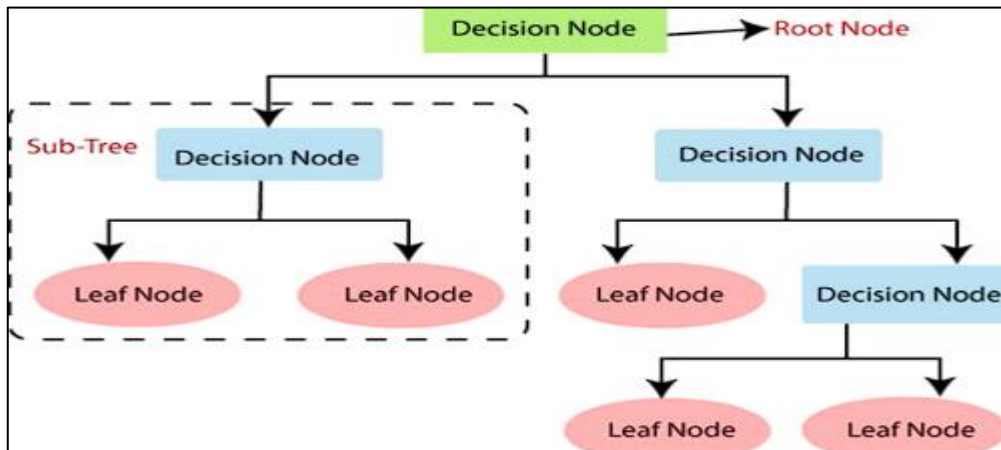
Lineer regresyonun amacı, gözlemlenen değer ile modelin tahmin ettiği değerler arasındaki farkı, diğer bir deyişle hata düzeyini (ϵ) minimize etmektir.

B. Karar Ağacı (Decision Tree)

Karar Ağacı (Decision Tree) algoritması, sınıflandırma ve regresyon analizlerinde yaygın olarak kullanılan denetimli öğrenme yöntemlerinden biridir (Quinlan, 1986). Bu algoritma, veri kümesindeki özellikleri kullanarak belirli karar noktaları üzerinden veri örneklerini sınıflara ayırır. Bu süreç, ağacın kök düğümünden başlayarak yaprak düğümlere kadar devam eder ve her düğümde bir özellik seçilerek veri alt kümelere bölünür. Karar ağaçları, sezgisel olarak anlaşılması kolay modeller sunar ve bu nedenle veri bilimi uygulamalarında yaygın bir şekilde tercih edilir (Loh, 2011).

Karar ağacı, dalların yapısını temsil eden iç düğümlerden, algoritmanın verdiği kararı temsil eden veri kümesinden ve her yaprak düğümünün bir sonucu temsil ettiği bir yapıdan oluşur. İki tür düğüm vardır: Birincisi, bir karar vermek için kullanılan ve çeşitli dallara sahip olan karar düğümü; ikincisi ise, karar düğümlerinin sonucunu temsil eden ve daha fazla dallanmaya sahip olmayan yaprak düğümüdür. Karar ağacı, şeklinden dolayı bu ismi almıştır çünkü bir ağaca benzer bir yapıdadır. Kök düğüm, çeşitli dallara genişleyen bir başlangıç noktasıdır ve bu şekilde ağaç benzeri bir yapı oluşturur. Karar ağacı, bir soruya verilen cevaba (evet veya hayır) bağlı olarak ağacı alt ağaçlara ayırır (Bansal vd., 2020).

Şekil 5. Karar Ağacı Yapısı



Kaynak: (Bansal vd., 2022)

Karar ağacının temel amacı, veri kümesini mümkün olan en iyi şekilde ayırarak sınıflandırma hatasını minimize etmektir. Bu amaçla, veri kümesi içindeki en iyi bölme noktalarını belirlemek için çeşitli ölçütler kullanılır. En yaygın kullanılan ölçütler Gini Impurity (Gini Kirliliği) ve Information Gain (Bilgi Kazancı) ölçütleridir.

Gini kirliliği, veri kümesindeki heterojenliği ölçer. Bir düğümdeki Gini kirliliği, düğümdeki sınıfların olasılıklarının karesinin toplamının birden çıkarılmasıyla hesaplanır (Breiman vd., 1984). Gini kirliliği, 0 ile 0.5 arasında bir değer alır ve değeri ne kadar düşükse, düğümdeki veri örnekleri o kadar homojendir.

$$Gini(D) = 1 - \sum_{i=1}^c p_i^2$$

Burada p_i , i . sınıfa ait olasılığı ve C sınıf sayısını temsil eder.

Bilgi kazancı, bir özelliğin veri kümesini ne kadar iyi böldüğünü ölçer. Bilgi kazancı, bölünmeden önceki ve bölünmeden sonraki entropi farkı olarak tanımlanır (Quinlan, 1986). Entropi, düğümdeki bilgi miktarını ölçer ve şu formülle hesaplanır:

$$Entropy(D) = - \sum_{i=1}^c p_i \log_2(p_i)$$

Bilgi kazancı ise şu şekilde hesaplanır:

$$Information\ Gain(D, A) = Entropy(D) - \sum_{v \in Values(A)} \frac{|D_v|}{|D|} \cdot Entropy(D_v)$$

Burada A bölme yapılacak özellik, D_v ise A özelliğinin v değeri için veri kümesidir.

Son yıllarda, karar ağaçlarının optimal bölme noktalarını bulma ve model karmaşıklığını yönetme konularında önemli ilerlemeler kaydedilmiştir. Örneğin, Gini kirliliği ve bilgi kazancı gibi geleneksel ölçütlerin yanı sıra, regularization yöntemleri ve ensemble teknikleri kullanılarak karar ağaçlarının genelleme yetenekleri artırılmaktadır (Rokach ve Maimon, 2014).

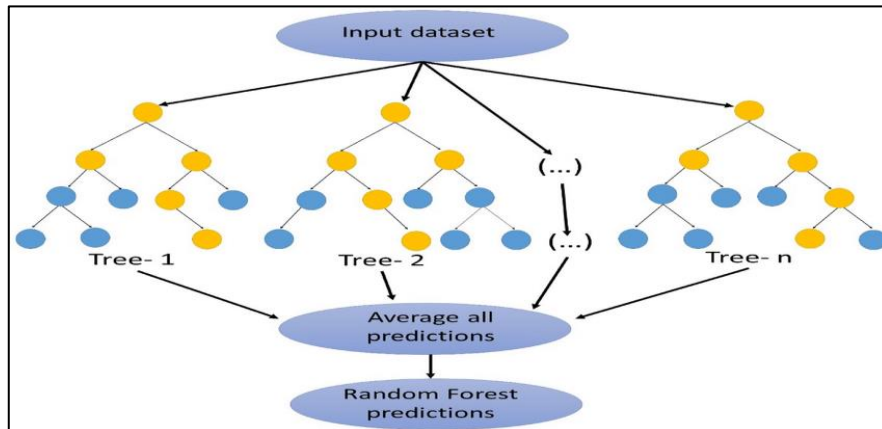
Karar ağacı algoritmasının uygulanması, veri kümesinin yapısına ve seçilen ölçütlere bağlı olarak değişiklik gösterebilir. Ancak, genel prensipler ve formülasyonlar yukarıda belirtilen adımları izler. Son zamanlarda, ağaçların aşırı uyum (overfitting) riskini azaltmak için pruning teknikleri ve ensemble yöntemleri (örneğin, Random Forest ve Gradient Boosting) kullanılmaktadır (Hastie vd., 2009).

C. Rastgele Orman (Random Forest)

Random Forest, hem sınıflandırma hem de regresyon görevlerinde kullanılan güçlü bir ensemble öğrenme yöntemidir (Breiman, 2001). Bu algoritma, birçok karar ağacının (decision trees) oluşturulması ve bu ağaçların sonuçlarının birleştirilmesiyle çalışır. Random Forest, ağaç tabanlı modellerin yüksek varyansını (overfitting) azaltarak daha sağlam ve genelleştirilebilir modeller üretir (Liaw ve Wiener, 2001).

Random Forest algoritmasının temel amacı, veri setinin farklı alt kümelerini kullanarak birçok karar ağacı oluşturmak ve bu ağaçların tahminlerini birleştirerek nihai bir tahmin yapmak ve böylece modelin genel performansını artırmaktır. Bu süreç, bootstrap aggregating (bagging) ve random feature selection tekniklerine dayanır.

Şekil 6. Rastgele Orman Yapısı



Kaynak: (Sahour vd., 2021)

Bagging, her bir karar ağacının, orijinal veri kümesinden bootstrap yöntemi ile rastgele seçilen örnekler kullanılarak eğitilmesi sürecidir (Breiman, 1996). Bootstrap örnekleme, orijinal veri kümesinden tekrarlarla örnekleme almayı içerir ve bu yöntem, modelin varyansını azaltmaya yardımcı olur.

$$f_{\text{Bagging}}(x) = \frac{1}{B} \sum_{b=1}^B f_b(x)$$

Burada $f_b(x)$ b. karar ağacının tahminini ve B toplam ağaç sayısını ifade eder.

Random Forest algoritması, her bir düğümde bölme yaparken, tüm özellikler yerine rastgele seçilen bir alt küme özelliği kullanır. Bu yöntem, ağaçların birbirinden bağımsız olmasını sağlar ve modelin bias-variance trade-off dengesini iyileştirir (Ho, 1995).

Bölme Kriteri: $\operatorname{argmax}_{a \in A'} \Delta \text{Impurity}$

Burada A' , rastgele seçilen özelliklerin alt kümesini temsil eder ve $\Delta \text{Impurity}$, Gini kirliliği veya bilgi kazancı gibi bir ölçütle hesaplanır.

D. Bagging Karar Ağacı (Bagging Decision Tree)

Bagging (Bootstrap Aggregating), istatistiksel öğrenmede tahmin modelinin varyansını azaltmak için kullanılan bir ensemble yöntemidir. Bagging, özellikle yüksek varyansa sahip modellerde (örneğin, karar ağaçları) genelleme hatasını azaltmada etkilidir (Breiman, 1996). Bu metodoloji, orijinal veri kümesinden rastgele örnekler alarak birden fazla model oluşturmayı ve bu modellerin tahminlerini birleştirmeyi içerir.

Bagging algoritmasının temel amacı, farklı örnekleme teknikleriyle oluşturulan birçok modelin tahminlerini birleştirerek daha istikrarlı ve güvenilir bir sonuç elde etmektir. Bagging, her bir modelin hatalarını ortadan kaldırarak ve farklı modellerin tahminlerini birleştirerek genelleme performansını artırır.

Bagging algoritmasının ilk adımı, orijinal veri kümesinden bootstrap yöntemi ile birçok alt küme oluşturulmasıdır. Bootstrap örnekleme, orijinal veri kümesinden tekrarlarla örnekleme almayı içerir. Bu sayede, her bir model farklı veri alt kümeleri üzerinde eğitilir.

$$D^* = \{D_1^*, D_2^*, \dots, D_B^*\}$$

Burada D_i^* , bootstrap örnekleme ile elde edilen i. veri alt kümesini temsil eder ve B, toplam bootstrap örnekleme sayısını ifade eder.

Her bir bootstrap örnekleme üzerinde bir regresyon modeli (örneğin, karar ağacı regresyon modeli) eğitilir. Eğitilen bu modellerin tahminleri birleştirilerek nihai tahmin elde edilir. Regresyon problemlerinde, bu tahminler genellikle ortalama alınarak birleştirilir.

$$f'_{\text{Bagging}}(x) = \frac{1}{B} \sum_{b=1}^B f'_b(x)$$

Burada $f'_b(x)$, b. modelin x girdisi için tahminini ve $f'_{\text{Bagging}}(x)$, bagging yönteminin nihai tahminini temsil eder.

E. Ekstra Ağaçlar (ExtraTrees)

ExtraTrees (Extremely Randomized Trees), hem sınıflandırma hem de regresyon görevlerinde kullanılan bir ensemble öğrenme yöntemidir. Bu algoritma, temel olarak karar ağaçlarıyla çalışır ancak ağaçların oluşturulması sırasında ek rastgelelik unsurları ekleyerek modelin varyansını azaltmayı hedefler (Geurts vd., 2006). ExtraTrees, Random Forest algoritmasına benzer şekilde çalışır ancak bölme noktalarının seçimi ve ağaçların oluşturulması süreçlerinde bazı farklılıklar içerir.

ExtraTrees algoritmasının temel amacı, veri setinin farklı alt kümelerini kullanarak birçok karar ağacı oluşturmak ve bu ağaçların tahminlerini birleştirerek nihai bir tahmin yapmaktır. Random Forest algoritması ile aynı şekilde bölme gerçekleştirirse de ExtraTrees daha fazla rastgelelik ekleyerek her bir düğümdeki bölme noktalarını da rastgele seçer.

Her bir düğümde, tüm özellikler yerine rastgele seçilen bir alt küme özelliği kullanılır. Bu yöntem, ağaçların birbirinden bağımsız olmasını sağlar ve modelin bias-variance trade-off dengesini iyileştirir (Geurts vd., 2006).

Bölme Kriteri: $\operatorname{argmax}_{a \in A'} \Delta \text{Impurity}$

Burada random forest algoritması ile aynı şekilde A' , rastgele seçilen özelliklerin alt kümesini temsil eder ve $\Delta \text{Impurity}$, Gini kirliliği veya bilgi kazancı gibi bir ölçütle hesaplanır.

ExtraTrees algoritmasının belirgin özelliği, her bir özelliğin bölme noktalarının rastgele seçilmesidir. Bu, ağaçların daha çeşitlilik göstermesini sağlar ve genel model performansını artırır.

Bölme Noktası: $\operatorname{random}(a_{\min}, a_{\max})$

Burada random(a_{min}, a_{max}) özellik a için minimum ve maksimum değerler arasında rastgele bir bölme noktası seçimini ifade eder.

F. Gradyan Artırımı (Gradient Boosting)

Gradient Boosting, sınıflandırma ve regresyon görevlerinde kullanılan güçlü bir ensemble öğrenme yöntemidir. Bu algoritma, zayıf öğrencileri ardışık olarak ekleyerek bir model oluşturur ve her yeni model, önceki modellerin hatalarını azaltmaya çalışır (Friedman, 2001). Gradient Boosting, ağaç tabanlı modellerle birlikte kullanıldığında, yüksek doğruluk ve genelleme performansı sağlar.

Gradient Boosting algoritmasının temel amacı, hataları azaltarak ve ardışık modeller ekleyerek güçlü bir tahmin modeli oluşturmaktır. Bu süreç, her bir yeni modelin, önceki modellerin hatalarını öğrenmesiyle gerçekleştirilir.

Gradient Boosting, kayıp fonksiyonunu minimize etmek için fonksiyonel gradient descent yöntemini kullanır. Amaç, her bir adımda kayıp fonksiyonunun gradyanına göre modeli güncellemektir.

Verilen bir veri kümesi $\{(x_i, y_i)\}_{i=1}^N$ için, ilk olarak sabit bir model $F_0(x)$ seçilir. Ardından, her bir iterasyon için;

Kayıp fonksiyonunun negatif gradyanı hesaplanır. Bu gradyanlar, mevcut modelin hatalarını temsil eder.

$$r_{im} = - \left[\frac{\partial L(y_i, F(x_i))}{\partial F(x_i)} \right]_{F=F_{m-1}}$$

Hesaplanan gradyanlara göre yeni bir zayıf öğrenci $h_m(x)$ eğitilir. Bu zayıf öğrenci, hata terimlerini tahmin etmeye çalışır.

$$h_m(x) = \underset{h}{\operatorname{argmin}} \sum_{i=1}^N (r_{im} - h(x_i))^2$$

Yeni model, önceki modele eklenir ve bir ağırlık çarpanı γ_m ile ölçeklendirilir.

$$F_m(x) = F_{m-1}(x) + \gamma_m h_m(x)$$

Burada γ_m , aşağıdaki şekilde hesaplanan bir ağırlık çarpanıdır:

$$\gamma_m = \underset{\gamma}{\operatorname{argmin}} \sum_{i=1}^N L(y_i, F_{m-1}(x_i) + \gamma h_m(x_i))$$

Bu adımlar, belirli bir iterasyon sayısına veya durdurma kriterine ulaşılan kadar tekrarlanır.

G. Ekstra Gradyan Artırımı (XGBoost)

XGBoost algoritması Gradient Boosting algoritmasının optimize edilmiş bir türüdür (Chen ve Guestrin, 2016). Önceki versiyonlara göre sağladığı avantajları XGBoost kullanımının yaygınlaşmasındaki en önemli nedendir. XGBoost, ağacı oluştururken maksimum derinlik değerini kullanır. Oluşturulan ağaç aşağı yönde aşırı ilerleme gösterirse, budama gerçekleştirilir. Aşırı öğrenmenin önüne geçilir. Gradient Boosting algoritması, kayıp fonksiyonun hesaplanmasında birinci dereceden fonksiyon kullanırken, XGBoost bu hesaplamaları ikinci dereceden fonksiyonlar kullanarak gerçekleştirir. Paralel çalışma özelliği, diğer algoritmalara göre sonuca daha kısa sürede ulaşılmasını sağlar (Kelle ve Yüce, 2022).

XGBoost algoritmasının temel amacı, hataları azaltarak ve ardışık modeller ekleyerek güçlü bir tahmin modeli oluşturmaktır. Bu süreç, her bir yeni modelin, önceki modellerin hatalarını öğrenmesiyle gerçekleştirilir.

XGBoost, her bir iterasyonda kayıp fonksiyonunu minimize etmeye çalışır. Verilen bir veri kümesi için, ilk olarak sabit bir model seçilir (Chen ve Guestrin, 2016):

$$F_m(x) = F_{m-1}(x) + h_m(x)$$

Burada $h_m(x)$, m. iterasyondaki zayıf öğrenci (genellikle karar ağacı) olarak tanımlanır.

XGBoost, kayıp fonksiyonunu hem model doğruluğunu hem de modelin karmaşıklığını göz önünde bulundurarak optimize eder. Bu, aşırı öğrenmeyi (overfitting) önlemek için L1 ve L2 düzenleme terimlerini içeren bir hedef fonksiyonuyla yapılır:

$$L(\theta) = \sum_{i=1}^n l(y_i, \hat{y}_i) + \sum_{k=1}^K \Omega(f_k)$$

Burada, l tahmin edilen değerlerle gerçek değerler arasındaki kaybı ifade eden kayıp fonksiyonu, $\Omega(f_k)$ ise ağaç karmaşıklık cezasını gösteren düzenleme terimidir:

$$\Omega(f_k) = \gamma T + \frac{1}{2} \lambda \sum_{j=1}^T w_j^2$$

Yukarıdaki denklemde T , karar ağacındaki yaprak sayısını ve w_j her yapraktaki ağırlıkları temsil eder. γ ve λ düzenleme parametreleridir.

XGBoost, kayıp fonksiyonunu optimize etmek için ikinci dereceden Taylor serisi genişlemesini kullanır. İlk ve ikinci türevler, modelin güncellenmesinde kullanılır:

$$L(\theta) \approx \sum_{i=1}^n \left[(g_i f(x_i) + \frac{1}{2} h_i f(x_i)^2) \right] + \Omega(f)$$

Burada g_i , birinci dereceden gradyanı diğer bir deyişle kayıp fonksiyonunun türevini ($\partial_{y_i} l(y_i, \hat{y}_i)$) ifade ederken, h_i ikinci dereceden türev ($\partial_{y_i}^2 l(y_i, \hat{y}_i)$) olan hessian matrisi olarak tanımlanır. Bu ikinci dereceden bilgi, her adımda modelin daha iyi optimize edilmesini sağlar.

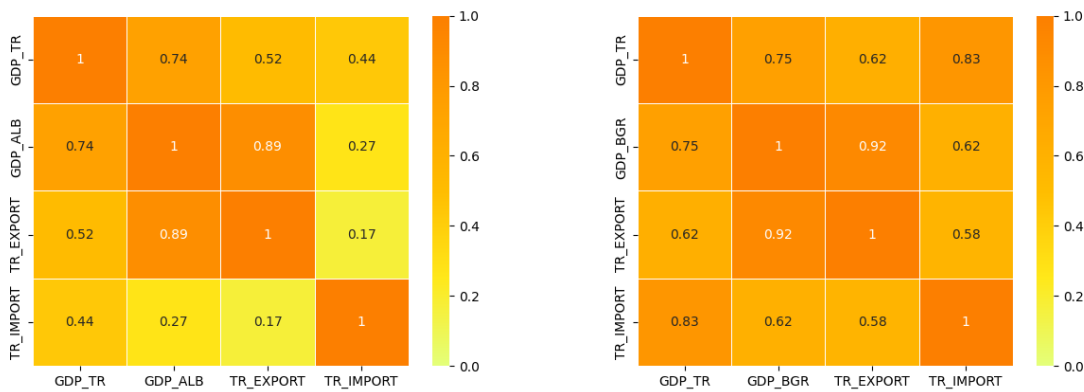
5. BULGULAR

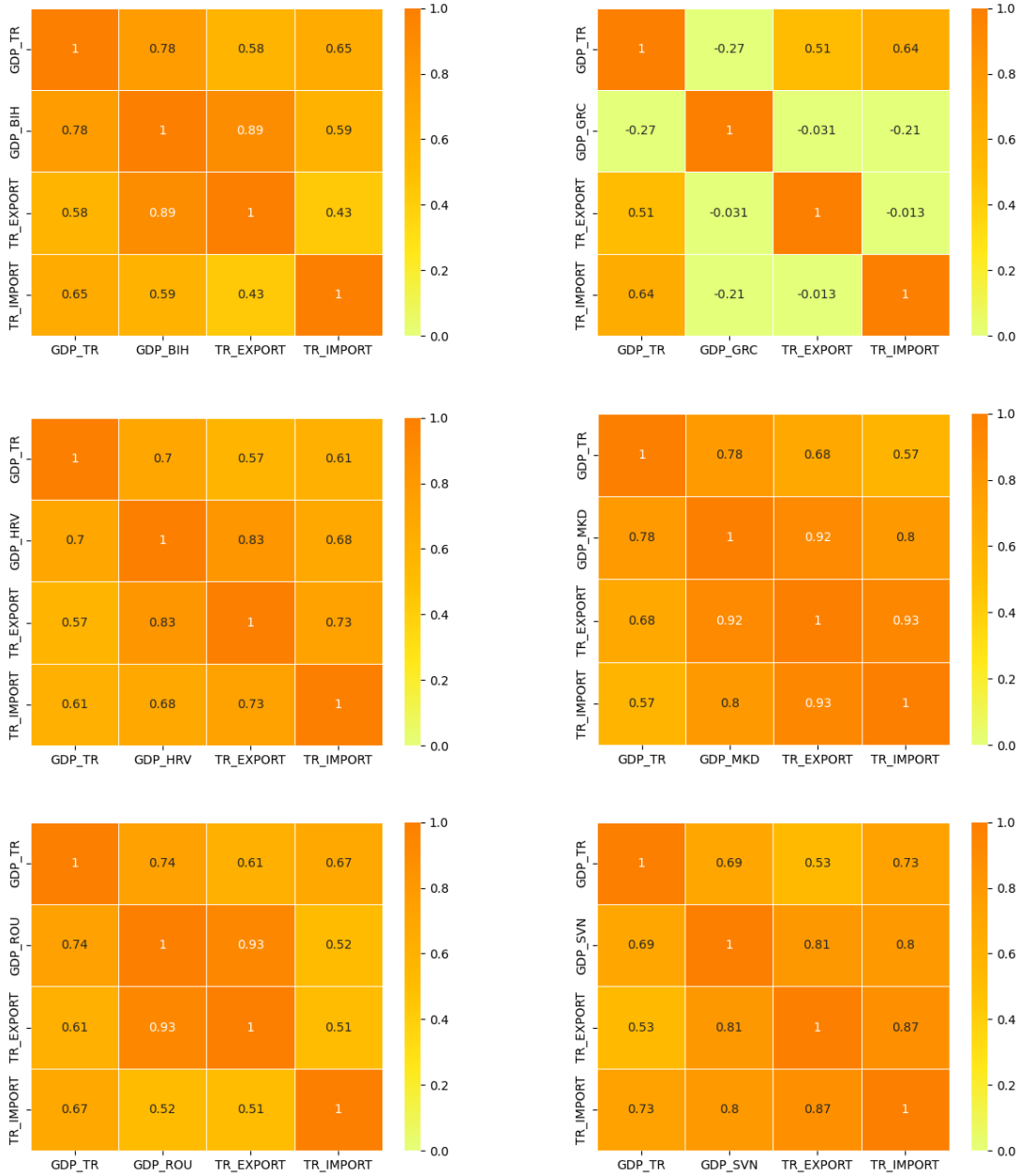
Uluslararası ticaret çekim modeli, daha önce belirtildiği gibi ticareti açıklamak için ülkelerin ekonomik büyüklükleri ve coğrafi uzaklıklarını temel almaktadır. Model, iki ülke arasındaki ticaret hacminin, ülkelerin milli gelirleriyle doğru, aralarındaki mesafe ile ters orantılı olduğunu varsayar. Bu varsayımın göre ülkelerin ekonomik büyüklükleri arttıkça ticaret potansiyeli de artarken, coğrafi uzaklık maliyetleri yükselterek ticareti sınırlar.

Bu temel ilişkiyi daha iyi anlamak için modelde yer alan değişkenler arasındaki korelasyon yapısının incelenmesi, ticareti belirleyen faktörlerin etkilerini daha net ortaya koyacaktır. Bu kapsamda öncelikle bu çalışmada değerlendirilecek ülkelerin Türkiye ile ikili ticaret miktarları ve ülkelerin milli gelirleri arasındaki ilişkiyi anlamak için oluşturulan korelasyon tabloları incelenecektir. Sonrasında ise 2023 yılı için bu ülkelerin Türkiye'ye coğrafi uzaklıkları ve ticaret büyüklükleri arasındaki ilişki değerlendirilecektir.

Şekil 7'de Türkiye ile analize konu her bir ülkenin milli gelirleri ve aralarındaki ticaret miktarları korelasyon matrisleri ile gösterilmiştir. Matris içerisinde dikey ve yatay eksene ülkelerin milli gelirleri ve ithalat-ihracat değerleri yerleştirilmiştir. Blokların içindeki sayıların her biri yatay ve dikey eksenle değişkenler arasındaki korelasyon katsayısını gösterirken, sağ tarafında bulunan renk ölçekleri ilişkisel yapının kuvvetini göstermektedir. Blok içerisindeki korelasyon katsayısı arttıkça renk tonu koyulaşırken, azaldıkça renk tonu da açılmaktadır.

Şekil 7. Türkiye'nin Balkan Ülkeleri ile İkili Ticaret Sayısal Göstergeleri İçin Korelasyon Matrisleri





Kaynak: Yazarlar tarafından oluşturulmuştur.

Şekil 7'deki korelasyon matrislerine göre ikili ticaret yapısında Türkiye'nin Arnavutluk'a ihracatı ile Arnavutluk'un milli geliri %89 pozitif, Türkiye'nin milli geliri %52 pozitif ilişkilidir. Türkiye'nin Arnavutluk için ithalatı ile Arnavutluk'un milli geliri %27 pozitif, Türkiye'nin milli geliri %44 pozitif, Arnavutluk'a ihracatı ise %17 pozitif ilişkilidir. İki ülkenin milli gelirleri ise birbirleri ile %74 pozitif ilişkilidir.

Bulgaristan'a yapılan ihracat ile Bulgaristan'ın milli geliri arasında %92, Türkiye'nin milli geliriyle ise %62 oranında pozitif bir korelasyon gözlemlenmiştir. Türkiye'nin Bulgaristan'dan ithalatı Bulgaristan'ın milli geliriyle %62, Türkiye'nin milli geliriyle %83 oranında pozitif korelasyona sahiptir. Ayrıca, Türkiye'nin Bulgaristan'a ihracatı %58 oranında Bulgaristan'dan ithalatı ile pozitif korelasyon göstermektedir. Türkiye ve Bulgaristan'ın milli gelirleri arasında %75 oranında pozitif bir ilişki bulunmaktadır.

Bosna-Hersek'e yapılan ihracat ile Bosna-Hersek'in milli geliri arasında %89, Türkiye'nin milli geliriyle %58 oranında pozitif korelasyon vardır. Türkiye'nin Bosna-Hersek'ten ithalatı Bosna-Hersek'in milli geliriyle %59, Türkiye'nin milli geliriyle %65 oranında pozitif korelasyona sahiptir. Türkiye'nin Bosna-Hersek'e ihracatı ile ithalatı %43 oranında pozitif korelasyon göstermektedir. Türkiye ve Bosna-Hersek'in milli gelirleri arasındaki korelasyon %78 oranında pozitif bir seviyededir.

Yunanistan'a yapılan ihracat ile Yunanistan'ın milli geliri arasında %3 oranında negatif, Türkiye'nin milli geliri ile %51 oranında pozitif korelasyon bulunmuştur. Türkiye'nin Yunanistan'dan ithalatı, Yunanistan'ın

milli geliriyle %21 negatif, Türkiye'nin milli geliriyle %64 oranında pozitif korelasyona sahiptir. Türkiye'nin Yunanistan'dan ithalatı ile Yunanistan'a ihracatı %1 oranında negatif zayıf bir korelasyon gösterirken, iki ülkenin milli gelirleri arasında %27 oranında negatif bir ilişki gözlemlenmiştir.

Hırvatistan'a yapılan ihracat ile Hırvatistan'ın milli geliri arasında %83, Türkiye'nin milli geliriyle %57 oranında pozitif korelasyon bulunmaktadır. Türkiye'nin Hırvatistan'dan ithalatı Hırvatistan'ın milli geliriyle %68, Türkiye'nin milli geliriyle %61 oranında pozitif korelasyon içindedir. Türkiye'nin Hırvatistan'a ihracatı ve Hırvatistan'dan ithalatı %73 oranında pozitif korelasyon göstermektedir. Türkiye ve Hırvatistan'ın milli gelirleri arasındaki korelasyon %70 oranında pozitif bir ilişkiye sahiptir.

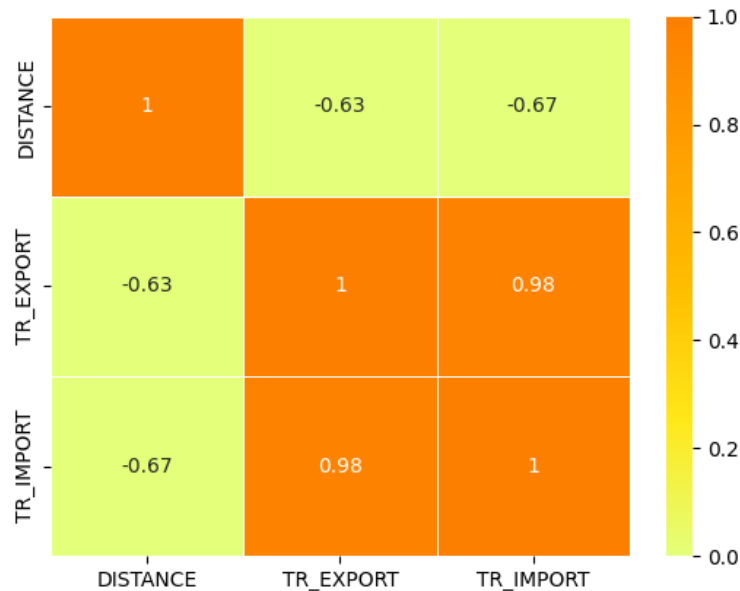
Kuzey Makedonya'ya yapılan ihracat ile Kuzey Makedonya'nın milli geliri %92 oranında, Türkiye'nin milli geliriyle %68 oranında pozitif korelasyon göstermektedir. Türkiye'nin Kuzey Makedonya'dan ithalatı Kuzey Makedonya'nın milli geliriyle %82, Türkiye'nin milli geliriyle %57 oranında pozitif korelasyon içindedir. Türkiye'nin Kuzey Makedonya'dan ithalatı ile ihracatı %93 oranında pozitif bir korelasyonla ilişkilidir. Türkiye ve Kuzey Makedonya'nın milli gelirleri arasında %78 oranında pozitif korelasyon mevcuttur.

Romanya'ya yapılan ihracat Romanya'nın milli geliriyle %93, Türkiye'nin milli geliriyle %61 oranında pozitif korelasyona sahiptir. Türkiye'nin Romanya'dan ithalatı Romanya'nın milli geliriyle %52, Türkiye'nin milli geliriyle %67 oranında pozitif bir korelasyon içindedir. Türkiye'nin Romanya'ya ihracatı ile ithalatı %51 oranında pozitif bir korelasyon gösterirken, Türkiye ve Romanya'nın milli gelirleri arasında %74 oranında pozitif bir ilişki bulunmaktadır.

Slovenya'ya yapılan ihracat ile Slovenya'nın milli geliri arasında %81, Türkiye'nin milli geliriyle %53 oranında pozitif korelasyon gözlemlenmiştir. Türkiye'nin Slovenya'dan ithalatı Slovenya'nın milli geliriyle %80, Türkiye'nin milli geliriyle %73 oranında pozitif korelasyon içindedir. Türkiye'nin Slovenya'dan ithalatı ile ihracatı %87 oranında pozitif korelasyon göstermektedir. Türkiye ve Slovenya'nın milli gelirleri arasındaki pozitif korelasyon %69 oranındadır.

Şekil 7 ve şekil 7'den hareketle yukarıda yapılan değerlendirmeye göre ikili ticarete gelir ile ticaret arasında negatif ilişkilerin görüldüğü tek ülke Yunanistan'dır. Türkiye'nin Arnavutluk ile ticareti için de zayıf korelasyon ilişkileri görülmesine rağmen pozitif yapıdadır. Bu iki ülke dışındaki ülkelerin Türkiye ile ikili ticaret yapılarında ise hem milli gelir miktarları hem de ticaret hacimleri için çekim modelini destekleyen orta-güçlü korelasyon ilişkileri gözlemlenmektedir.

Şekil 8. Türkiye'nin 2023 Yılı Balkan Ülkeleri Ticareti ile Ülkelere Uzaklıklarının Korelasyon Matrisi



Kaynak: Yazarlar tarafından oluşturulmuştur.

Şekil 8'de analize dâhil edilen ülkeler ile 2023 yılında gerçekleştirilen ticaret miktarları ve bu ülkelerin Türkiye'ye uzaklıkları arasındaki ilişkiyi gösteren korelasyon matrisi paylaşılmıştır. Buna göre ülkelerin Türkiye ile arasındaki mesafe, Türkiye'nin bu ülkelere ihracat miktarıyla %63, bu ülkelere ithalat miktarıyla ise %67 oranında negatif yönlü ilişkidir. Bu durum çekim modelinin temel varsayımlarından olan, ülkeler arasındaki mesafe ile ticaret hacmi arasındaki ters yönlü ilişkiyi de göstermektedir.

Uygulama aşamasında dört farklı (%15-85, %20-80, %25-75, %30-70) eğitim-test veri seti ayrıştırması yapılmıştır. Analize dâhil her ülke ile hem ihracat hem de ithalat yönlü olarak dört farklı eğitim-test yapısı katlı çapraz doğrulama yöntemi ile on farklı iterasyon şeklinde gerçekleştirilmiştir. Bu uygulama sonucunda elde edilen r^2 ve MAPE değerlerinde aşırı öğrenme ve eksik öğrenme durumları göz önüne alınarak her ülke için ikili ticaret tahmininde hem ihracat hem de ithalat için en iyi sonuçları üreten eğitim-test ayrımları, analize konu tüm yöntemler ile birlikte tablo 2’de gösterilmiştir.

Tablo 2. Ülkeler ve Bağımlı Değişkenlere Göre Sonuç Tabloları

Arnavutluk													
Eğitim: %75 Test: %25	Bağımlı Değişken: İthalat						Eğitim: %85 Test: %15	Bağımlı Değişken: İhracat					
Model	R2			MAPE			Model	R2			MAPE		
	Eğitim R2	Test R2	Fark R2	Eğitim MAPE	Test MAPE	Fark MAPE		Eğitim R2	Test R2	Fark R2	Eğitim MAPE	Test MAPE	Fark MAPE
ExtraTrees	1.00	0.96	0.04	0.00	17.60	17.60	ExtraTrees	1.00	0.99	0.01	0.00	7.53	7.53
Gradient Boosting	1.00	0.91	0.09	0.24	21.46	21.23	Bagging	0.97	0.99	-0.02	7.20	0.60	-6.59
Decision Tree	1.00	0.91	0.09	0.00	25.52	25.52	Random Forest	0.95	0.97	-0.03	4.66	3.77	-0.89
Bagging	0.84	0.90	-0.06	36.76	18.95	-17.81	Linear Regression	0.84	0.95	-0.11	15.95	2.42	-13.53
Random Forest	0.84	0.89	-0.05	38.91	21.40	-17.51	XGBoost	1.00	0.90	0.10	0.00	2.09	2.09
XGBoost	1.00	0.85	0.15	0.00	18.50	18.50	Gradient Boosting	1.00	0.82	0.18	0.05	17.12	17.07
Linear Regression	0.18	0.51	-0.34	83.58	40.50	-43.08	Decision Tree	1.00	0.80	0.20	0.00	7.30	7.30
Bulgaristan													
Eğitim: %85 Test: %15	Bağımlı Değişken: İthalat						Eğitim: %85 Test: %15	Bağımlı Değişken: İhracat					
Model	R2			MAPE			Model	R2			MAPE		
	Eğitim R2	Test R2	Fark R2	Eğitim MAPE	Test MAPE	Fark MAPE		Eğitim R2	Test R2	Fark R2	Eğitim MAPE	Test MAPE	Fark MAPE
ExtraTrees	1.00	0.96	0.04	0.00	10.08	10.08	Bagging	0.96	1.00	-0.04	9.36	0.84	-8.51
Decision Tree	1.00	0.95	0.05	0.00	10.22	10.22	Random Forest	0.96	0.99	-0.03	7.12	2.31	-4.81
Gradient Boosting	1.00	0.94	0.06	0.02	2.43	2.41	ExtraTrees	1.00	0.98	0.02	0.00	11.91	11.91
XGBoost	1.00	0.93	0.07	0.00	12.10	12.10	Linear Regression	0.85	0.97	-0.12	14.12	12.04	-2.08
Random Forest	0.96	0.92	0.03	5.98	5.32	-0.66	Gradient Boosting	1.00	0.96	0.04	0.03	2.26	2.23
Linear Regression	0.69	0.87	-0.18	13.41	11.85	-1.55	Decision Tree	1.00	0.96	0.04	0.00	2.29	2.29
Bagging	0.93	0.85	0.08	7.50	6.26	-1.24	XGBoost	1.00	0.61	0.39	0.00	21.76	21.76
Bosna Hersek													
Eğitim: %75 Test: %25	Bağımlı Değişken: İthalat						Eğitim: %85 Test: %15	Bağımlı Değişken: İhracat					
Model	R2			MAPE			Model	R2			MAPE		
	Eğitim R2	Test R2	Fark R2	Eğitim MAPE	Test MAPE	Fark MAPE		Eğitim R2	Test R2	Fark R2	Eğitim MAPE	Test MAPE	Fark MAPE
Random Forest	0.98	0.97	0.01	14.82	14.46	-0.35	ExtraTrees	1.00	0.99	0.01	0.00	10.49	10.49
XGBoost	1.00	0.93	0.07	0.00	21.79	21.79	Linear Regression	0.77	0.99	-0.21	20.67	14.71	-5.96
ExtraTrees	1.00	0.92	0.08	0.00	23.69	23.69	Random Forest	0.95	0.94	0.02	8.59	24.15	15.57
Gradient Boosting	1.00	0.92	0.08	0.04	22.23	22.19	Gradient Boosting	1.00	0.93	0.07	0.04	4.09	4.05
Decision Tree	1.00	0.91	0.09	0.00	22.52	22.52	Bagging	0.94	0.87	0.07	7.14	23.98	16.84
Bagging	0.97	0.91	0.06	18.69	18.66	-0.02	XGBoost	1.00	0.66	0.34	0.00	21.62	21.62
Linear Regression	0.91	0.82	0.09	56.45	127.1	70.65	Decision Tree	1.00	0.65	0.35	0.00	23.56	23.56

Yunanistan													
Eğitim: %80 Test: %20	Bağımlı Değişken: İthalat						Eğitim: %85 Test: %15	Bağımlı Değişken: İhracat					
Model	R2			MAPE			Model	R2			MAPE		
	Eğitim R2	Test R2	Fark R2	Eğitim MAPE	Test MAPE	Fark MAPE		Eğitim R2	Test R2	Fark R2	Eğitim MAPE	Test MAPE	Fark MAPE
ExtraTrees	1.00	1.00	0.00	0.00	0.34	0.34	ExtraTrees	1.00	0.78	0.22	0.00	3.97	3.97
XGBoost	1.00	0.97	0.03	0.00	4.30	4.30	Gradient Boosting	1.00	0.77	0.23	0.13	13.68	13.55
Gradient Boosting	1.00	0.97	0.03	0.06	2.55	2.50	Decision Tree	1.00	0.73	0.27	0.00	15.36	15.36
Decision Tree	1.00	0.94	0.06	0.00	6.48	6.48	Bagging	0.94	0.70	0.24	8.24	16.45	8.21
Random Forest	0.96	0.89	0.06	10.30	7.45	-2.84	Random Forest	0.94	0.65	0.30	6.96	17.58	10.61
Bagging	0.95	0.75	0.20	11.52	12.94	1.42	Linear Regression	0.69	0.54	0.15	19.19	29.00	9.81
Linear Regression	0.54	0.63	-0.09	33.45	16.20	-17.25	XGBoost	1.00	0.27	0.73	0.00	7.44	7.44
Hırvatistan													
Eğitim: %85 Test: %15	Bağımlı Değişken: İthalat						Eğitim: %85 Test: %15	Bağımlı Değişken: İhracat					
Model	R2			MAPE			Model	R2			MAPE		
	Eğitim R2	Test R2	Fark R2	Eğitim MAPE	Test MAPE	Fark MAPE		Eğitim R2	Test R2	Fark R2	Eğitim MAPE	Test MAPE	Fark MAPE
Gradient Boosting	1.00	0.92	0.08	0.06	6.62	6.55	Random Forest	0.95	0.97	-0.03	8.31	1.85	-6.46
Random Forest	0.91	0.92	-0.01	13.90	49.52	35.61	Linear Regression	0.85	0.92	-0.07	14.01	10.68	-3.33
Bagging	0.92	0.92	0.00	11.48	43.97	32.48	XGBoost	1.00	0.87	0.13	0.00	22.70	22.70
ExtraTrees	1.00	0.91	0.09	0.00	53.95	53.95	Bagging	0.92	0.85	0.07	12.98	3.79	-9.20
Decision Tree	1.00	0.90	0.10	0.00	57.29	57.29	ExtraTrees	1.00	0.85	0.15	0.00	23.19	23.19
Linear Regression	0.67	0.81	-0.13	24.88	30.72	5.84	Decision Tree	1.00	0.85	0.15	0.00	23.33	23.33
XGBoost	1.00	0.79	0.21	0.00	61.31	61.31	Bagging	0.94	0.82	0.12	7.46	31.15	23.68
Kuzey Makedonya													
Eğitim: %85 Test: %15	Bağımlı Değişken: İthalat						Eğitim: %75 Test: %25	Bağımlı Değişken: İhracat					
Model	R2			MAPE			Model	R2			MAPE		
	Eğitim R2	Test R2	Fark R2	Eğitim MAPE	Test MAPE	Fark MAPE		Eğitim R2	Test R2	Fark R2	Eğitim MAPE	Test MAPE	Fark MAPE
Random Forest	0.96	0.99	-0.03	10.64	3.34	-7.30	Random Forest	0.97	0.98	-0.01	3.98	2.09	-1.89
ExtraTrees	1.00	0.96	0.04	0.00	4.29	4.29	Gradient Boosting	1.00	0.97	0.03	0.02	4.41	4.39
Bagging	0.94	0.96	-0.02	7.75	8.01	0.27	ExtraTrees	1.00	0.96	0.04	0.00	8.46	8.46
Gradient Boosting	1.00	0.90	0.10	0.06	6.44	6.39	Decision Tree	1.00	0.95	0.05	0.00	5.13	5.13
Linear Regression	0.71	0.59	0.13	33.22	46.37	13.15	Bagging	0.97	0.92	0.05	4.42	17.81	13.39
XGBoost	1.00	0.56	0.44	0.00	11.91	11.91	Linear Regression	0.82	0.87	-0.05	11.98	24.17	12.19
Decision Tree	1.00	0.23	0.77	0.00	15.68	15.68	XGBoost	1.00	0.80	0.20	0.00	11.28	11.28

Romanya													
Eğitim: %85 Test: %15	Bağımlı Değişken: İthalat						Eğitim: %70 Test: %30	Bağımlı Değişken: İhracat					
Model	R2			MAPE			Model	R2			MAPE		
	Eğitim R2	Test R2	Fark R2	Eğitim MAPE	Test MAPE	Fark MAPE		Eğitim R2	Test R2	Fark R2	Eğitim MAPE	Test MAPE	Fark MAPE
Bagging	0.78	0.99	-0.21	8.59	0.88	-7.70	XGBoost	1.00	0.95	0.05	0.00	1.46	1.46
Random Forest	0.90	0.99	-0.09	5.88	1.66	-4.22	Linear Regression	0.90	0.93	-0.03	7.98	12.30	4.32
Decision Tree	1.00	0.96	0.04	0.00	1.47	1.47	ExtraTrees	1.00	0.74	0.26	0.00	2.33	2.33
Gradient Boosting	1.00	0.95	0.05	0.06	1.18	1.12	Gradient Boosting	1.00	0.71	0.29	0.01	32.78	32.78
Linear Regression	0.69	0.80	-0.11	10.05	6.54	-3.50	Random Forest	0.97	0.63	0.33	4.75	39.29	34.54
ExtraTrees	1.00	0.76	0.24	0.00	7.11	7.11	Decision Tree	0.96	0.58	0.38	6.84	7.82	0.98
XGBoost	1.00	0.74	0.26	0.00	7.77	7.77	Bagging	0.98	0.56	0.42	4.83	43.50	38.67

Slovenya													
Eğitim: %70 Test: %30	Bağımlı Değişken: İthalat						Eğitim: %70 Test: %30	Bağımlı Değişken: İhracat					
Model	R2			MAPE			Model	R2			MAPE		
	Eğitim R2	Test R2	Fark R2	Eğitim MAPE	Test MAPE	Fark MAPE		Eğitim R2	Test R2	Fark R2	Eğitim MAPE	Test MAPE	Fark MAPE
ExtraTrees	1.00	0.97	0.03	0.00	4.21	4.21	Linear Regression	0.96	0.98	-0.01	14.14	1.62	-12.52
Bagging	0.96	0.96	-0.01	4.56	4.93	0.37	Bagging	0.93	0.97	-0.04	16.04	1.21	-14.83
Random Forest	0.94	0.94	0.01	5.41	5.58	0.17	Gradient Boosting	1.00	0.96	0.04	0.06	0.81	0.76
Gradient Boosting	1.00	0.89	0.11	0.02	5.67	5.66	ExtraTrees	1.00	0.96	0.04	0.00	1.42	1.42
XGBoost	1.00	0.84	0.16	0.00	6.69	6.69	XGBoost	1.00	0.91	0.09	0.00	32.67	32.67
Linear Regression	0.88	0.81	0.07	8.91	0.72	-8.19	Random Forest	0.98	0.90	0.08	6.11	38.50	32.39
Decision Tree	1.00	0.52	0.48	0.00	10.93	10.93	Decision Tree	1.00	0.85	0.15	0.00	39.01	39.01

Kaynak: Yazarlar tarafından oluşturulmuştur.

Tablo 2'deki sonuçlar eğitim ve test verileri için yapılan ayrımlar içerisinde her bir ülke ve her bir bağımlı değişken için en başarılı tahminlerin olduğu ayrımları içermektedir. En iyi yöntemin belirlenmesi hususunda öncelikli olarak yöntemin tahmin başarısını işaret ettiği için test MAPE değerlerinin düşük olması dikkate alınmıştır. Sonrasında eğitim MAPE değerlerinin eksik öğrenme durumuna sebebiyet verecek şekilde yüksek olmamasına dikkat edilmiştir. Sonrasında ise modelin açıklama gücünü ifade eden eğitim ve test r^2 değerleri incelenerek en dengeli model yapıları belirlenerek tabloda altı çizili şekilde işaretlenmiştir.

Tablo 2'de gösterilen sonuçlar, en iyi tahmin üreten yöntemlerin modelde kullanılan veri setinin bölünme yapısına, ülkeler arasındaki ilişkiler ve ticari dinamiklere, aynı zamanda bağımlı değişkenlere göre farklılık gösterdiğini ortaya koymaktadır. Bu durum, tahmin modellerinin her ülke ve veri setine özgü değişkenlerle farklı performans sergileyebileceğini işaret eder. Örneğin, bazı yöntemler belirli bir bölgeye ya da ülke grubuna yönelik daha isabetli tahminler sunarken, başka bir veri seti veya bağımlı değişkenle çalışıldığında aynı yöntemler beklenen başarıyı göstermeyebilir. Bu nedenle, model seçiminde veri setinin yapısına ve bağımlı değişkenlerin karakteristiklerine dikkat edilmesi önem taşımaktadır.

Tablo 3. Ülkelere ve Bağımlı Değişkenlere Göre En Başarılı Yöntemler

Bağımlı Değişken	Ülke	Arnavutluk	Bulgaristan	Bosna-Hersek	Yunanistan	Hırvatistan	Kuzey Makedonya	Romanya	Slovenya
İthalat		ExtraTrees	Gradient Boosting	Random Forest	Gradient Boosting	Gradient Boosting	ExtraTrees	Gradient Boosting	ExtraTrees
İhracat		Bagging	Gradient Boosting	Gradient Boosting	ExtraTrees	Random Forest	Random Forest	XGBoost	Gradient Boosting

Kaynak: Yazarlar tarafından oluşturulmuştur.

Tablo 2'deki sonuçlar üzerinden seçilen en dengeli modeller tablo 3'te özetlenmiştir. Buna göre analize dâhil edilen sekiz ülke için hem ihracat hem de ithalat yönlü en fazla başarılı tahmin üreten model gradyan artırımı algoritmasıdır. Gradyan artırımı algoritmasının sekiz ülke ile ikili ticarete iki farklı bağımlı değişken için yapılan analizlerin yedisinde en başarılı yöntem olduğu görülmüştür. Analizlerin dördünde ekstra ağaçlar algoritması en başarılı tahminler üreten model olmuştur. Rastgele orman yapısının ise üç analiz için en başarılı

yöntemleri üreten model olduğu gözlenmiştir. Bagging karar ağacı ve ekstra gradyan artırımı ise birer kez en başarılı tahmin sonuçlarını üretmiştir. Bu beş yöntemin dışında kalan karar ağacı ve lineer regresyon yapıları hiçbir analiz için diğer modellere göre başarılı tahminler yapamamıştır. Ancak tablo 2 incelendiğinde de görüleceği gibi bu algoritmaların da en başarılı yönteme yakın tahminler ürettiği analizler bulunmaktadır.

SONUÇ VE DEĞERLENDİRME

Türkiye'nin Balkan ülkeleri ile olan ticaret hacmini tahmin etmek amacıyla kullanılan makine öğrenmesi yöntemleri, ülkeler arası ticari dinamiklerin karmaşıklığını başarılı bir şekilde yansıtmış ve her ülkenin farklı ekonomik yapısına göre çeşitlilik gösteren sonuçlar elde edilmiştir. Çalışmada kullanılan yedi farklı modelin performansı, dört farklı eğitim-test veri seti bölünmesi ve 10 katlı çapraz doğrulama ile değerlendirilmiştir. Bulgular, bu analizlerin her birinin kendi içinde değerlendirilmesi gerektiğini ve ticaret hacmi tahminlerinde, kullanılan modelin seçiminin büyük bir öneme sahip olduğunu göstermiştir. Özellikle, ülkeler arası ticari ilişkilerin tahmininde belirli modellerin genel performansı yüksek olsa da her ülkenin kendine özgü ticari dinamikleri nedeniyle en iyi performans gösteren model ülkeden ülkeye ve bağımlı değişkenlere göre değişiklik göstermiştir.

MAPE (Ortalama Mutlak Yüzde Hata) ve r^2 gibi değerlendirme metrikleri üzerinden yapılan karşılaştırmalar, ticaret hacmi tahminlerinde bazı modellerin belirli ülkeler için üstünlük sağladığını ortaya koymuştur. Ancak bu sonuçlar sabit bir modelin her durumda en doğru sonuçları vereceği anlamına gelmemektedir. Aksine, elde edilen veriler ışığında her ülke için farklı modellerin daha uygun olduğu anlaşılmıştır. Bu durum, ticaret hacminin tahmin edilmesi sürecinde tek bir yönteme bağlı kalmanın sınırlayıcı olabileceğini ve analizin farklı yöntemler üzerinden gerçekleştirilmesinin daha sağlıklı sonuçlar vereceğini göstermektedir. Çalışmada kullanılan modellerin başarılarının farklı ülkeler ve veri setleri üzerinde değişiklik göstermesi, ticaret ilişkilerinin dinamik yapısına işaret etmektedir.

Türkiye ile Balkan ülkeleri arasındaki ticaret hacmi incelendiğinde, her iki tarafın da ekonomik büyüklükleri, jeopolitik konumları, uluslararası anlaşmalara üyelik durumları ve coğrafi faktörleri gibi değişkenlerin bu ticaret üzerinde önemli bir etkisi olduğu görülmektedir. Özellikle ülkelerin Dünya Ticaret Örgütü (DTÖ) üyelikleri, gümrük birliği üyelikleri, karayla çevrili olma durumu ve ortak sınır gibi faktörlerin ticaret hacmini nasıl etkilediği, analizde belirleyici olmuştur. Sayısal değişkenler arasında ise ülkelerin milli gelirleri, Türkiye ile Balkan ülkeleri arasındaki mesafe, ihracat ve ithalat verileri ön plana çıkmıştır. Bu faktörlerin her biri, ticaret hacmi tahminlerinde önemli birer bileşen olmuş ve sonuçlar üzerinde doğrudan etkiye bulunmuştur.

Yapılan analizler, Türkiye ile Balkan ülkeleri arasındaki ticari ilişkilerin tahmin edilmesi ve gelecekteki ticari seyrin öngörülmesi noktasında önemli veriler sunmaktadır. Özellikle ticaret politikalarının geliştirilmesinde kullanılabilecek stratejik öneriler için bu analizler büyük bir kaynak teşkil etmektedir. Türkiye'nin Balkan ülkeleri ile olan ticari ilişkilerini daha derinlemesine anlayabilmek, bu ülkelerle yapılacak ticari müzakerelerde ve ekonomik iş birliklerinde önemli fırsatlar yaratacaktır. Ticaret hacminin daha sağlıklı ve doğru bir şekilde tahmin edilmesi, bu ülkelerle yapılacak ticaretin daha verimli hale getirilmesi ve karşılıklı ekonomik faydanın maksimuma çıkarılması açısından büyük bir önem taşımaktadır. Elde edilen sonuçlar, Türkiye'nin Balkan ülkeleri ile olan ticari ilişkilerini daha iyi yönetebilmesi ve bu ülkelerle olan ticaret hacmini artırabilmesi için stratejik kararlar alınmasına olanak tanıyacak niteliktedir. Buna ek olarak Balkan coğrafyasında bulunan ülkeler ile ticarete konu mal ve hizmetlerin hem nitelik hem de nicelik olarak derinlemesine değerlendirilmesi de bu ülkeler ile gelecek yıllardaki ticaret politikalarının şekillenmesi konusunda destekleyici bir niteliğe sahip olacaktır.

Ayrıca çalışmanın sonuçları sadece ticari tahminler açısından değil, bölgesel ekonomik politikaların oluşturulmasında da dikkate değer katkılar sağlamaktadır. Türkiye'nin bölgedeki ticari etkisini artırabilmesi için, özellikle bölgesel dinamikleri göz önünde bulundurarak, ticari politikalarını esnek ve ülkeye özgü bir şekilde şekillendirmesi gerekmektedir. Balkan ülkeleri ile ticari ilişkilerin gelecekteki potansiyelini daha doğru bir şekilde değerlendirmek ve stratejik öncelikleri belirlemek, bu çalışmada elde edilen bulguların ışığında mümkündür. Çalışmanın sunduğu bulgular hem akademik anlamda hem de ticaret politikalarının geliştirilmesi noktasında önemli bir referans noktası olma özelliği taşımaktadır.

Türkiye'nin Balkan ülkeleri ile ticaret hacminin tahmin edilmesine yönelik bu kapsamlı analiz, makine öğrenimi yöntemlerinin ticaret verilerinin tahmininde nasıl etkin bir şekilde kullanılabileceğini göstermiştir. Her ne kadar her ülke için farklı modellerin daha başarılı olduğu görülmüş olsa da bu yöntemler ticaret tahminlerinde önemli bir potansiyele sahiptir ve geliştirilebilir yapıdadır. Gelecekte bu yöntemlerin daha da geliştirilmesi, veri setlerinin daha detaylı bir şekilde işlenmesi ve ticaret ilişkilerindeki yeni dinamiklerin dikkate alınmasıyla, daha kesin ve güvenilir tahminlere ulaşılması mümkündür. Bu çalışma, hem mevcut

ticaret ilişkilerinin daha iyi anlaşılması hem de Türkiye'nin Balkan ülkeleri ile olan gelecekteki ticaret stratejilerine yön verecek önemli ipuçları sunmaktadır.

KAYNAKÇA

- Alpaydin, E. (2014). Introduction to Machine Learning, third edition. MIT Press. <https://books.google.com.tr/books?id=7f5bBAAAQBAJ>
- Anderson, J. E., & van Wincoop, E. (2003). Gravity with Gravitas: A Solution to the Border Puzzle. *American Economic Review*, 93(1), 170–192. <https://doi.org/10.1257/000282803321455214>
- Baier, S. L., & Bergstrand, J. H. (2007). Do free trade agreements actually increase members' international trade? *Journal of International Economics*, 71(1), 72–95. <https://doi.org/https://doi.org/10.1016/j.jinteco.2006.02.005>
- Bangdiwala, S. I. (2018). Regression: simple linear. *International Journal of Injury Control and Safety Promotion*, 25(1), 113–115. <https://doi.org/10.1080/17457300.2018.1426702>
- Bansal, M., Goyal, A., & Choudhary, A. (2022). A comparative analysis of K-Nearest Neighbor, Genetic, Support Vector Machine, Decision Tree, and Long Short Term Memory algorithms in machine learning. *Decision Analytics Journal*, 3, 100071. <https://doi.org/https://doi.org/10.1016/j.dajour.2022.100071>
- Bansal, M., Prince, Yadav, R., & Ujjwal, P. K. (2020). Palmistry using Machine Learning and OpenCV. 2020 Fourth International Conference on Inventive Systems and Control (ICISC), 536–539. <https://doi.org/10.1109/ICISC47916.2020.9171158>
- Bergstrand, J. (1985). The Gravity Equation In International Trade: Some Microeconomic Foundations And Empirical Evidence. *The Review of Economics and Statistics*, 67, 474–481. <https://doi.org/10.2307/1925976>
- Breiman, L. (1996). Bagging Predictors. *Machine Learning*, 24, 123–140. <https://doi.org/10.1007/BF00058655>
- Breiman, L. (2001). Random Forests. *Machine Learning*, 45, 5–32. <https://doi.org/10.1023/A:1010950718922>
- Breiman, L., Friedman, J., Stone, C. J., & Olshen, R. A. (1984). Classification and Regression Trees. Taylor & Francis. <https://books.google.com.tr/books?id=JwQx-WOmSyQC>
- Chen, T., & Guestrin, C. (2016). XGBoost: A Scalable Tree Boosting System. Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, 785–794. <https://doi.org/10.1145/2939672.2939785>
- Circlaeys, S., Kanitkar, C., & Kumazawa, D. (2017). Bilateral trade flow prediction. Unpublished Manuscript, Available for Download at <Http://Cs229.Stanford.Edu/Proj2017/Final-Reports/5240224>. Pdf.
- Erdoğan, S., & Tek, M. (2019). Türkiye Ekonomisinin Genel Görünümü. M. Alagöz & G. Akar (Der.), Sektörel Ekonomik Analiz Türkiye (2003-2018) İçinde (ss. 1–34). Gazi Kitabevi.
- Friedman, J. H. (2001). Greedy function approximation: A gradient boosting machine. *The Annals of Statistics*, 29(5), 1189–1232. <https://doi.org/10.1214/aos/1013203451>
- Geurts, P., Ernst, D., & Wehenkel, L. (2006). Extremely Randomized Trees. *Machine Learning*, 63, 3–42. <https://doi.org/10.1007/s10994-006-6226-1>
- Goldfarb, A., & Trefler, D. (2018). Artificial Intelligence, and International Trade. *The Economics of Artificial Intelligence: An Agenda*, 463–492.
- Hastie, T., Tibshirani, R., & Friedman, J. (2009). The Elements of Statistical Learning: Data Mining, Inference, and Prediction, Second Edition (Springer Series in Statistics).
- Helpman, E., Melitz, M., & Rubinstein, Y. (2008). Estimating Trade Flows: Trading Partners and Trading Volumes. *The Quarterly Journal of Economics*, 123(2), 441–487. <https://econpapers.repec.org/RePEc:oup:qjecon:v:123:y:2008:i:2:p:441-487>
- Ho, T. K. (1995). Random decision forests. Proceedings of 3rd International Conference on Document Analysis and Recognition, 1, 278–282 vol.1. <https://doi.org/10.1109/ICDAR.1995.598994>
- Jošić, H., & Žmuk, B. (2022). A Machine Learning Approach to Forecast International Trade: The Case of Croatia. *Business Systems Research*, 13(3), 144–160. <https://doi.org/10.2478/bsrj-2022-0030>
- Kelle, A. C., & Yüce, H. (2022). MQTT Trafikinde DoS Saldırılarının Makine Öğrenmesi ile Sınıflandırılması ve Modelin SHAP ile Yorumlanması TT - Classification of DoS Attacks in MQTT Network with Machine Learning and Interpretation of The Model with SHAP. *Journal of Materials and Mechatronics: A*, 3(1), 50–62. <https://doi.org/10.55546/jmm.995091>
- Kottou, E. M., Grubelich, T. A., & Wang, X. (2020). Bilateral Trade Flow Prediction Models Enhanced By Wavelet and Machine Learning Algorithms. 2020 International Conference on Computational Science and Computational Intelligence (CSCI), 1510–1516. <https://doi.org/10.1109/CSCI51800.2020.00279>
- Kulkarni, A., & More, A. (2016). Formulation of a Prediction Index with the Help of WEKA Tool for Guiding the Stock Market Investors. *Oriental Journal of Computer Science and Technology*, 9, 212–225. <https://doi.org/10.13005/ojcsct/09.03.07>
- Liaw, A., & Wiener, M. (2001). Classification and Regression by RandomForest. *Forest*, 23.
- Linnemann, H. (1966). An econometric study of international trade flows [North-Holland Pub. Co.]. <https://doi.org/LK-https://worldcat.org/title/239355>
- Loh, W.-Y. (2011). Classification and Regression Trees. *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery*, 1, 14–23. <https://doi.org/10.1002/widm.8>
- McCallum, J. (1995). National Borders Matter: Canada-U.S. Regional Trade Patterns. *The American Economic Review*, 85(3), 615–623. <http://www.jstor.org/stable/2118191>

- Melitz, J. (2008). Language and foreign trade. *European Economic Review*, 52(4), 667–699. <https://doi.org/https://doi.org/10.1016/j.euroecorev.2007.05.002>
- Nuroğlu, E. (2012). Estimating Trade Flows of Turkey Using Panel Data Analysis and Neural Networks. *Bartın Üniversitesi İİBF Dergisi*, 3, 13–34.
- Pöyhönen, P. (1963). A Tentative Model for the Volume of Trade between Countries. *Weltwirtschaftliches Archiv*, 90, 93–100. <http://www.jstor.org/stable/40436776>
- Quimba, F. M. A., & Barral, M. A. A. (2018). Exploring neural network models in understanding bilateral trade in APEC: A review of history and concepts (Issues 2018–33). *Philippine Institute for Development Studies (PIDS)*. <https://hdl.handle.net/10419/211053>
- Quinlan, J. R. (1986). Induction of decision trees. *Machine Learning*, 1(1), 81–106. <https://doi.org/10.1007/BF00116251>
- Rokach, L., & Maimon, O. (2014). *Data Mining with Decision Trees: Theory and Applications*, 2nd Edition. In *Data Mining with Decision Trees: Theory and Applications*, 2nd Edition (Vol. 81). <https://doi.org/10.1142/9097>
- Sahour, H., Gholami, V., Torkman, J., Vazifedan, M., & Saeedi, S. (2021). Random forest and extreme gradient boosting algorithms for streamflow modeling using vessel features and tree-rings. *Environmental Earth Sciences*, 80. <https://doi.org/10.1007/s12665-021-10054-5>
- Tinbergen, J. (1962). *Shaping the World Economy; Suggestions for an International Economic Policy*. <http://hdl.handle.net/1765/16826>
- Wohl, I., & Kennedy, J. (2018). *Neural network analysis of international trade*. US International Trade Commission: Washington, DC, USA.