



2025, 14 (2), 869-886 | Araştırma Makalesi

Makina Öğrenmesi Yoluyla İşsizlik Oranlarının Tahmini: Türkiye İçin Bir Uygulama

Gülferah Ertürkmen¹

Ali Öter²

Öz

Bu çalışma, Türkiye’de işsizlik oranlarının tahmini için geleneksel istatistiksel yöntemler ile makine öğrenmesi tekniklerini karşılaştırarak analiz etmektedir. İşsizlik oranları; ekonomik büyüme, enflasyon, nüfus artışı, göç hareketleri ve eğitim harcamaları gibi makroekonomik değişkenlerden etkilenmektedir. Bu nedenle, TÜİK verilerine dayalı olarak Rastgele Ormanlar (RF), Gradyan Artırma (GB) ve Çok Katmanlı Algılayıcı (MLP) gibi makine öğrenmesi algoritmaları kullanılarak işsizlik oranları tahmin edilmiş ve modellerin performansları karşılaştırılmıştır. Çalışmada, makine öğrenmesi modelleri arasında MLP modeli en düşük hata oranları ile en iyi tahmin performansını (MAE: 1,945, RMSE: 2,235) göstermiştir. RF ve GB modelleri belirli bir doğruluk seviyesine ulaşsa da, hata oranları MLP modeline kıyasla daha yüksek çıkmıştır. Bulgular, makine öğrenmesi tekniklerinin işsizlik tahmininde geleneksel istatistiksel modellere göre daha başarılı olduğunu göstermektedir. Özellikle MLP modeli, doğrusal olmayan ilişkileri öğrenebilme kapasitesi sayesinde diğer modellere göre daha hassas tahminler sunmuştur. Ayrıca, yapılan korelasyon analizleri, işsizlik oranlarının enflasyon, ekonomik büyüme ve göç hareketleri ile anlamlı ilişkiler içinde olduğunu ortaya koymuştur. Ekonomik büyümenin işsizlik üzerindeki negatif etkisi belirgin şekilde gözlemlenirken, göç hareketlerinin işsizlik oranlarını artırabileceği tespit edilmiştir. Özellikle enflasyon ve eğitim harcamaları arasındaki negatif korelasyon, ekonomik istikrarsızlık dönemlerinde eğitim yatırımlarının azaldığını göstermektedir. Bu çalışmada, makine öğrenmesi teknikleri kullanılarak 2025, 2026 ve 2027 yıllarına ait işsizlik oranları tahmin edilmiştir. Yapılan analizler sonucunda, 2025 yılında işsizlik oranının %9,2 ile %11,5 arasında değişmesi beklenirken, 2026 yılında bu oranın %8,8 ile %11,0 seviyelerinde olacağı öngörülmüştür. 2027 yılı tahminlerine göre ise işsizlik oranının %8,5 ile %10,7 seviyelerine gerileyeceği hesaplanmıştır. Analiz sonuçları, MLP modelinin TÜİK verilerine en yakın tahminleri sunduğunu göstermektedir. RF ve GB modelleri ise işsizlik oranlarını %9-14 aralığında tahmin ederek daha geniş bir hata payına sahip olmuştur. Bu tahminler, ekonomik politika yapıcıları ve işgücü piyasası analistleri için önemli bir rehber niteliğinde olup, işsizlik oranlarının gelecekteki seyri hakkında öngörüler sunmaktadır.

Anahtar Kelimeler: İşsizlik, Makine Öğrenmesi, İşsizlik Oranı Tahmini, Rastgele Ormanlar, Gradyan Artırma, Çok Katmanlı Algılayıcı

Ertürkmen, G., & Öter, A. (2025). Makina Öğrenmesi Yoluyla İşsizlik Oranlarının Tahmini: Türkiye İçin Bir Uygulama. İnsan Ve Toplum Bilimleri Araştırmaları Dergisi, 14(2), 869-886. <https://doi.org/10.15869/itobiad.1629420>

Geliş Tarihi	29.01.2025
Kabul Tarihi	03.06.2025
Yayın Tarihi	30.06.2025
*Bu CC BY-NC lisansı altında açık erişimli bir makaledir.	

1 Dr.Öğr. Üyesi, Kahramanmaraş Sütçü İmam Üniversitesi, Göksun Uygulamalı Bilimler Yüksekokulu, Finans ve Bankacılık Bölümü, Kahramanmaraş, Türkiye, gbozkaya@ksu.edu.tr, ORCID: 0000-0003-2239-0241

2 Dr.Öğr. Üyesi, Kahramanmaraş Sütçü İmam Üniversitesi, Teknik Bilimler Meslek Yüksekokulu, Elektronik ve Otomasyon Bölümü, Kahramanmaraş, Türkiye, alioter@ksu.edu.tr, ORCID: 0000-0002-9546-0602



2025, 14 (2), 869-886 | Research Article Estimation of Unemployment Rates through Machine Learning: An Application for Türkiye

Gülferah Ertürkmen ¹

Ali Öter ²

Abstract

This research compares and analyzes traditional statistical methods and machine learning techniques for forecasting unemployment rates in Türkiye. Unemployment rates are affected by macroeconomic variables such as economic growth, inflation, population growth, migration movements and education expenditures. Therefore, unemployment rates are estimated using machine learning algorithms such as Random Forests (RF), Gradient Boosting (GB) and Multilayer Perceptron (MLP) based on TurkStat data and the performances of the models are compared. In the research, among the machine learning models, the MLP model showed the best forecasting performance (MAE: 1,945; RMSE: 2,235) with the lowest error rates. Although the RF and GB models achieved a certain level of accuracy, their error rates were higher compared to the MLP model. The findings suggest that machine learning techniques are more successful in unemployment forecasting than traditional statistical models. In particular, the MLP model provides more accurate forecasts than other models thanks to its capacity to learn nonlinear relationships. Moreover, correlation analyses reveal that unemployment rates are significantly correlated with inflation, economic growth and migration flows. While the negative effect of economic growth on unemployment is clearly observed, migration movements are found to increase unemployment rates. In particular, the negative correlation between inflation and education expenditures suggests that education investments decrease during periods of economic instability. In this research, unemployment rates for the years 2025, 2026 and 2027 are estimated using machine learning techniques. As a result of the analysis, the unemployment rate is expected to vary between 9.2% and 11.5% in 2025, while this rate is expected to be between 8.8% and 11.0% in 2026. According to the 2027 forecasts, the unemployment rate is expected to decline to between 8.5% and 10.7%. The results of the analysis show that the MLP model provides the closest forecasts to the TurkStat data. The RF and GB models, on the other hand, have a wider margin of error, predicting unemployment rates in the range of 9-14%. These forecasts are an important guide for economic policy makers and labor market analysts and provide predictions about the future course of unemployment rates.

Keywords: Unemployment, Machine Learning, Unemployment Rate Prediction, Random Forests, Gradient Boosting, Multilayer Perceptron

Ertürkmen, G., & Öter, A. (2025). Estimation of Unemployment Rates through Machine Learning: An Application for Türkiye. *Journal of Human and Social Sciences Research*, 14(2), 869-886. <https://doi.org/10.15869/itobiad.1629420>

Date of Submission	29.01.2025
Date of Acceptance	03.06.2025
Date of Publication	30.06.2025
*This is an open access article under the CC BY-NC license.	

¹ Asst . Prof. Dr., Kahramanmaraş Sütçü İmam University, Göksun School of Applied Sciences, Department of Finance and Banking, Kahramanmaraş, Türkiye, gbozkaya@ksu.edu.tr, ORCID: 0000-0003-2239-0241

² Asst. Prof. Dr., Kahramanmaraş Sütçü İmam University, Vocational School of Technical Sciences, Department of Electronics and Automation, Kahramanmaraş, Türkiye, alioter@ksu.edu.tr, ORCID: 0000-0002-9546-0602

1. Giriş

İşsizlik, ekonomik kalkınma, toplumsal refah ve bireysel yaşam standartları üzerinde doğrudan etkisi olan temel makroekonomik göstergelerden biridir. İşsizlik oranlarının yükselmesi, gelir dağılımındaki adaletsizlikleri artırırken, sosyal huzursuzluklara ve ekonomik durgunluğa yol açabilir. Özellikle ekonomik krizler, küresel salgınlar ve teknolojik dönüşümler gibi dışsal şoklar, işgücü piyasasında ani değişimlere neden olarak işsizlik oranlarını öngörmeyi daha da zorlaştırmaktadır (Olmedo, 2014, s. 189). Bu nedenle, işsizlik oranlarının güvenilir tahminlerle öngörülmesi, ekonomi yönetiminin önlem alabilmesi açısından kritik bir önem taşımaktadır.

Geleneksel ekonomik tahmin yöntemleri, genellikle zaman serisi analizlerine veya istatistiksel regresyon modellerine dayanmaktadır. Örneğin, Otokorelasyon Fonksiyonu (ACF) ve Gecikmeli Kısmi Otokorelasyon (PACF) analizleri kullanılarak oluşturulan ARIMA (Autoregressive Integrated Moving Average) modelleri, geçmiş işsizlik verilerinin istatistiksel özelliklerini kullanarak gelecekteki işsizlik oranlarını tahmin etmeye çalışmaktadır (Tüzemen & Yıldız, 2017, s.90). Bunun yanı sıra, GARCH (Generalized Autoregressive Conditional Heteroskedasticity) modelleri gibi volatilité analizine dayalı yöntemler, işsizlik oranlarının zaman içindeki değişkenliğini modellemekte ve tahminlerde kullanılmaktadır (Gil-Alina vd., 2019, s.222). Ancak, işsizlik oranları sadece geçmiş verilere bağlı kalmadan, ekonomik büyüme, enflasyon, nüfus artışı, göç hareketleri ve eğitim harcamaları gibi birçok makroekonomik faktörden etkilenmektedir. Bu noktada, geleneksel istatistiksel yöntemler, doğrusal olmayan ve karmaşık ilişkileri yakalamakta yetersiz kalabilmekte ve alternatif tahmin tekniklerine ihtiyaç duyulmaktadır (Katris, 2020, s.674).

Son yıllarda, makine öğrenmesi (ML) ve yapay zekâ tabanlı modeller işsizlik oranlarının tahmininde giderek daha fazla kullanılmaya başlanmıştır. Makine öğrenmesi, büyük veri kümeleri üzerinde örüntüleri ve ilişkileri öğrenerek, istihdam piyasasında ani değişimlere karşı daha duyarlı ve esnek tahminler yapabilmektedir. Örneğin, Yapay Sinir Ağları (ANN - Artificial Neural Networks) ve Uzun Kısa Süreli Bellek Ağları (LSTM - Long Short-Term Memory Networks) gibi derin öğrenme modelleri, doğrusal olmayan ilişkileri öğrenerek işsizlik oranlarını uzun vadeli olarak tahmin edebilme kapasitesine sahiptir (Mulaudzi & Ajoodha, 2020). Benzer şekilde, Destek Vektör Makineleri (SVM - Support Vector Machines) ve Gradyan Artırma (GB - Gradient Boosting) yöntemleri, farklı ekonomik göstergeler arasındaki karmaşık ilişkileri öğrenerek tahmin doğruluğunu artırabilmektedir (Ahmad vd., 2021, s.529).

Bu çalışmada, Türkiye’de işsizlik oranlarının tahmini amacıyla makine öğrenmesi teknikleri kullanılarak geleneksel yöntemlerle karşılaştırmalı bir analiz gerçekleştirilmiştir. RF, GB ve MLP gibi popüler makine öğrenmesi algoritmaları uygulanarak, farklı model yapılandırmalarının tahmin performansı ölçülmüştür. Kullanılan veri seti, TÜİK (Türkiye İstatistik Kurumu) tarafından sağlanan aylık işsizlik oranlarını içermekte olup, tahmin modellerine enflasyon, eğitim harcamaları, ekonomik büyüme, nüfus artışı ve göç gibi ekonomik değişkenler de dahil edilmiştir. Çalışmada ele alınan temel araştırma soruları şunlardır:

- Makine öğrenmesi yöntemleri, geleneksel istatistiksel modellere kıyasla işsizlik oranlarını daha başarılı bir şekilde tahmin edebilir mi?

- Hangi makine öğrenmesi yöntemi en düşük hata oranı ile en doğru tahminleri sunmaktadır?
- İşsizlik oranlarını belirleyen en önemli makroekonomik faktörler nelerdir ve bu faktörlerin tahmin modeline etkisi nasıl değişmektedir?

Bu araştırmanın temel amacı, işsizlik oranlarının tahmininde geleneksel yöntemler ile makine öğrenmesi yaklaşımlarını karşılaştırmalı olarak analiz ederek en etkili tahmin yöntemini belirlemektir. Elde edilen bulguların, ekonomik politika yapımcılar ve işgücü piyasası analistleri için yol gösterici olması hedeflenmektedir. Makine öğrenmesi modellerinin sağladığı yüksek doğruluk sayesinde, mevcut tahmin yöntemlerinin geliştirilmesi ve gelecekteki tahmin süreçlerinin daha güvenilir hale getirilmesi amaçlanmaktadır.

Çalışmada, literatürde genellikle sadece geçmiş işsizlik oranlarına odaklanan araştırmalardan farklı olarak; enflasyon, ekonomik büyüme, eğitim harcamaları, nüfus artışı ve göç gibi çeşitli makroekonomik değişkenler bir arada ele alınmış ve bu değişkenlerin işsizlik tahmini üzerindeki etkileri incelenmiştir. Ayrıca, farklı makine öğrenmesi algoritmaları (Random Forest, Gradient Boosting, Multi-Layer Perceptron) aynı veri seti üzerinde karşılaştırmalı olarak test edilmiş, algoritmaların performansları ayrıntılı bir şekilde değerlendirilmiştir. Çalışmanın özgün yönlerinden biri, geliştirilen modeller aracılığıyla Türkiye için gelecekteki işsizlik oranlarının tahmin edilmesi ve olası eğilimlerin ortaya konmasıdır. Böylece araştırma, hem kullanılan değişken çeşitliliği hem de algoritmaların sistematik karşılaştırması ve ileriye dönük tahmin sunması açısından literatüre önemli bir katkı sağlamaktadır. Çalışmanın devamında sırasıyla literatür özeti, araştırmada kullanılan veri seti ve yöntem, elde edilen bulguların tartışılması ve son bölümde sonuçlar ile politika önerileri sunulmaktadır.

2. Literatür

İşsizlik oranlarının tahmini, ekonomi literatüründe önemli bir araştırma alanı olarak dikkat çekmektedir. Artan işsizlik oranı tüm dünyada, özellikle de ülkemizde en acil sorunlardan biridir. Sosyal ve politik önemi göz önüne alındığında, işsizlik oranlarını tahmin etmek, politika yapımcıların karar alma süreçlerine yardımcı olmak için özellikle önemlidir. Bu konuda yapılan çalışmalar, hem geleneksel yöntemler hem de makine öğrenimi ve hibrit modeller gibi çağdaş yaklaşımları kapsamaktadır. Aşağıda, bu alanda yapılan bazı önemli çalışmalar özetlenmiştir.

Olmedo (2014) çalışmasında, İspanya'da işsizlik oranlarının tahmin edilmesinde Yakın Komşu ve Sinir Ağı Teknikleri kullanılmıştır. Araştırmanın amacı, yeniden yapılandırma ve öğrenme yaklaşımlarını vurgulayarak, parametrik olmayan tahmin yöntemlerini karşılaştırmaktır. İlk yöntemde tek adımlı tahmin yapılırken, ikinci yöntemde daha uzun bir zaman dilimi kullanılarak tahmin yapılmıştır. Yöntemlerin etkinliği, farklı Avrupa ülkelerinin işsizlik zaman serilerine uygulanarak değerlendirilmiştir. Sonuçlar, uzun vadeli tahminlerin daha başarılı sonuçlar verdiğini göstermektedir.

Çiftçi (2016) çalışmasında, Türkiye'deki işsizlik oranının tahmini için Markov zincirleri yöntemi kullanılmıştır. Bu yöntemle, işsizlik oranlarının gelecekteki değerleri belirlenmeye çalışılmıştır. Çalışma, işsizlik oranının zaman içindeki geçiş olasılıklarını analiz ederek, gelecekteki eğilimleri öngörmeyi amaçlamaktadır. Elde edilen sonuçlar, Markov zincirleri yönteminin işsizlik oranı tahmininde etkili bir araç olabileceğini

göstermektedir.

Yolcu ve Baş (2016) çalışmasında, işsizlik oranı ve işgücüne katılım oranının tahmini üzerine odaklanılmış; Polonya ve Türkiye örneklerinde çeşitli bulanık zaman serisi yöntemleri uygulanmıştır. Analizler sonucunda, bulanık zaman serisi modellerinin bu tür ekonomik göstergelerin tahmininde yüksek doğruluk sağladığı ve etkili bir tahmin aracı olarak kullanılabileceği ortaya konmuştur. Ayrıca, bu yöntemlerin geliştirilmesi durumunda tahmin başarılarının daha da artırılabilceği belirtilmiştir.

Tüzemen ve Yıldız (2017), Türkiye'nin 01/2005-10/2016 dönemi işsizlik oranlarını tahmin etmek için Toplamsal ve Çarpımsal-Mevsimsel Holt-Winters yöntemlerini kullanmıştır. Analizde yer alan Ortalama Kare Hatası (MSE), gerçek ve tahmin değerleri arasındaki farkların karelerinin ortalamasını; Ortalama Mutlak Yüzde Hatası (MAPE) ise bu farkların gerçek değerlere oranının mutlak değerlerinin ortalamasını yüzde cinsinden ifade etmektedir. Performans analizlerinde, Toplamsal-Mevsimsel HW yöntemi daha düşük hata oranlarıyla (MSE = 0,073711, MAPE = 0,021695) öne çıkarken, Çarpımsal-Mevsimsel HW yöntemi daha yüksek hata değerleri (MSE = 0,8618929, MAPE = 0,02217202) vermiştir. Çalışmada, Toplamsal-Mevsimsel HW yöntemiyle işsizlik oranlarının 2017 yılında artmaya devam edeceği tahmin edilmiştir.

Kızılkaya(2017) çalışmasında, Türkiye'nin enflasyon ve işsizlik oranlarının tahmini için Yapay Sinir Ağları (YSA) ve ARIMA modelleri kullanılmıştır. 1923–2014 arası işsizlik ve 1969–2014 arası enflasyon verileri ile modelleme yapılmıştır. Tahminlerde, YSA'nın yaygın bir türü olan Çok Katmanlı Algılayıcı (ÇKA) ile Box-Jenkins yöntemi temel alınarak uygun ARIMA süreci belirlenmiştir. 2015–2020 dönemi için elde edilen tahmin sonuçları, resmi tahminlere göre daha başarılı bulunmuştur.

Yıldırım ve Başgeçmez (2017), Türkiye'deki işsizlik oranının analizini ve tahminini gerçekleştirmek amacıyla zaman serisi analiz yöntemlerini kullanmıştır. Çalışmada, geçmiş dönem işsizlik verileri incelenerek mevcut eğilimler değerlendirilmiş ve gelecekteki işsizlik oranları öngörülmüştür. Araştırma, işsizlik oranının zaman içindeki davranışını anlamak ve gelecekteki değişimleri tahmin etmek için istatistiksel modelleme tekniklerinin etkinliğini ortaya koymaktadır.

Gil-Alina vd. (2019), Türkiye işsizlik oranını incelemek için uzun hafızalı modeller ve kısmi bütünlük yöntemler kullanmışlardır. 1988-2013 verilerine dayanan çalışma, işsizliğin Türkiye'de büyük ölçüde kalıcı olduğunu ve entegrasyon derecesinin genelde 1 veya üzerinde olduğunu göstermektedir. Çalışma, işgücü piyasası ve makroekonomik politikaların işsizlik üzerinde uzun vadeli etkiler yaratabileceğini vurgulamıştır.

Kütük ve Güloğlu (2019) Türkiye'de işsiz bireylerin istihdama geçiş olasılıklarını tahmin etmek için lojistik regresyon, yapay sinir ağları ve makine öğrenmesi yöntemlerini karşılaştırmıştır. 2004–2016 verileriyle yapılan analizde, makine öğrenmesi yöntemleri, özellikle XGBoost ve RF, %67 doğrulukla en iyi performansı göstermiştir; bu da işgücü piyasası analizinde bu tekniklerin etkinliğini ortaya koymuştur.

Katris (2020) tarafından yapılan çalışma, işsizlik tahmininde kullanılan yöntemlerin başarı oranlarının tahmin dönemi uzunluğuna göre değiştiğini ortaya koymuştur. Çalışmada, Fractional ARIMA (FARIMA) modelinin kısa vadeli (1 adımlı) tahminlerde en iyi performansı sunduğu belirlenmiştir. Ancak, tahmin dönemi uzadığında (örneğin 12 adımlı tahminlerde), sinir ağı modelleri (neural network models) daha başarılı

sonuçlar sağlamıştır. Bu durum, farklı yöntemlerin tahmin ufku uzunluğuna göre seçilmesi gerektiğini vurgulamaktadır. Çalışma, tahmin süresi ve kullanılan yöntem arasındaki uyumun tahmin doğruluğunu önemli ölçüde etkilediğini göstermektedir.

Mulaudzi ve Ajoodha (2020) çalışmasında, Güney Afrika'nın işsizlik oranını tahmin etmek için altı geleneksel istatistiksel model ve yedi makine öğrenimi modelini karşılaştırarak, tek değişkenli makine öğrenimi yöntemlerini değerlendirmişlerdir. Çalışmanın sonuçlarına göre, Çok Katmanlı Perceptron en düşük hata oranını sağlarken, Ridge Regresyon modeli en yüksek R^2 değerini elde etmiştir. ARIMA, LASSO ve Elastik Ağ modelleri ise bu yöntemlerin hemen ardından gelmiş ve genel bulgular, makine öğrenimi modellerinin geleneksel yöntemlerden daha doğru tahminler sunduğunu ortaya koymuştur.

Zhao (2020), uzun süreli işsizlik riskini tahmin etmek için Avrupa'daki kamu istihdam verilerini kullanarak XGBoost gibi makine öğrenmesi algoritmalarını uygulamış ve %81,2 doğruluk oranı elde etmiştir. Çalışmada SHAP yöntemiyle model kararları açıklanmış, özellikle yaşlılar ve göçmenler gibi gruplar için önyargı riskleri belirlenmiştir; sonuçlar, kamu istihdam politikalarında şeffaf ve adil yapay zeka kullanımının önemini vurgulamaktadır.

Bağcı (2020), Türkiye'de işsizlik oranını tahmin etmek amacıyla önce GM(1,1) modelini, ardından bu modele Markov zincirlerini entegre ederek Gri Markov modelini kullanmıştır. İlk olarak GM(1,1) modeliyle işsizlik oranlarına ilişkin tahminler yapılmış ve hata oranları hesaplanmıştır. Daha sonra hata oranları ve orijinal verilerle durum geçiş matrisi oluşturularak Markov analizi gerçekleştirilmiştir. Gri Markov modeli ile yapılan tahminlerin, GM(1,1) modeline göre daha düşük hata oranına sahip olduğu ve daha isabetli sonuçlar ürettiği belirlenmiştir.

Lai ve diğerleri (2021), Asya'daki beş gelişmiş ve beş gelişmekte olan ülkenin işsizlik oranlarını tahmin etmek amacıyla bir çalışma gerçekleştirmiştir. Araştırmada, ARIMA-ARNN, ARIMA-ANN ve ARIMA-SVM gibi hibrit modeller kullanılmıştır. Analiz sonuçları, Asya'nın gelişmiş ülkelerinde hibrit modeller arasında en iyi performansı ARIMA-ARNN modelinin sergilediğini ortaya koymuştur.

Ahmad vd. (2021) çalışması, Avrupa'nın bazı ülkelerinde, özellikle Fransa, İspanya, Belçika, Türkiye, İtalya ve Almanya'da işsizlik oranlarını daha doğru bir şekilde tahmin etmek amacıyla doğrusal ve doğrusal olmayan hibrit tahmin modellerini incelemiştir. Araştırmanın bulgularına göre, Fransa, Belçika, Türkiye ve Almanya'da hibrit ARIMA-ARNN modelinin başarılı bir şekilde tahmin yaptığı, İspanya ve İtalya'da ise hibrit ARIMA-SVM modelinin daha etkili olduğu belirlenmiştir. Çalışma, COVID-19 pandemisinin işsizlik oranlarını artıracak ve bu etkinin söz konusu ülkelerde en az 5 yıl süreceğini göstermektedir.

Davidescu vd. (2021), Romanya'nın işsizlik oranını tahmin etmek için hem tek değişkenli (univariate) hem de çok değişkenli (multivariate) yöntemleri karşılaştırmalı olarak değerlendirmiştir. Çalışmada kullanılan modellerden SETAR (Threshold AutoRegressive) ve VECM (Vector Error Correction Model), doğruluk açısından oldukça yakın sonuçlar vermiştir. Özellikle pandemi öncesi dönemi yansıtan tahminlerde, her iki modelin de güvenilir sonuçlar sunduğu belirlenmiştir. Araştırmanın bulguları, karmaşık yapıda olmayan dönemler için bu modellerin uygun olabileceğini ve işsizlik oranlarını

doğru bir şekilde öngörebildiğini ortaya koymaktadır .

Chakraborty vd. (2021), yedi farklı ülkeye ait işsizlik oranı verilerini kullanarak, hibrit modellemelerin geleneksel yöntemlere kıyasla daha üstün performans sunduğunu ortaya koymuştur. Araştırmada kullanılan hibrit yöntemler, tahmin doğruluğunu artırarak işsizlik oranlarının daha iyi bir şekilde öngörülmesini sağlamıştır.

Sen ve diğerleri (2022) çalışmasında, Hindistan'daki istihdam fırsatlarını artırmaya ihtiyaç duyan eyaletleri belirlemek için Lojistik Regresyon, Destek Vektör Makineleri, K-en yakın komşu (kNN) algoritması ve Karar Ağaçları gibi farklı popüler denetimli makine öğrenimi algoritmalarını kullanmışlardır. Çalışma, Karar Ağaçları algoritmasının daha üstün bir performans sergilediğini ortaya koymuştur..

Karahan ve Çetintaş (2022), Türkiye'nin 2005-2018 yıllarına ait aylık ekonomik verilerini kullanarak, işsizlik oranının gelecekteki dönemlere ilişkin tahminini gerçekleştirmeyi hedeflemiştir. Araştırmada toplam altı temel ekonomik gösterge dikkate alınmıştır. Çalışmanın sonuçları, geleneksel yöntemlerin hatalı veya aşırı sapmalı sonuçlar üretme riskinin yüksek olduğu durumlarda, kaotik problemlerin çözümünde çağdaş yaklaşımlardan biri olan Yapay Sinir Ağları (YSA) yönteminin güvenilir ve gerçeğe yakın tahminler sağladığını ortaya koymuştur.

Simionescu ve Cifuentes-Faura (2022), İspanya'daki genç işsizlik oranını tahmin etmek için Google Trend verilerini kullanarak yenilikçi bir yaklaşım benimsemiştir. Araştırmada, hem bölgesel hem de ulusal düzeyde farklı modelleme teknikleri kullanılmıştır. Bölgesel analizlerde, Bayes panel veri modelleri ve sabit etkiler modelleri tercih edilmiştir. Ulusal düzeyde ise Bayesyen vektör otoregresif (BVAR) modelleri ve hata düzeltme modelleri (VECM) uygulanmıştır. Sonuçlar, Bayes yöntemleri ile elde edilen tahminlerin dokuz farklı bölge için en başarılı ve doğru sonuçları sunduğunu göstermiştir. Bu yaklaşım, işsizlik oranlarını tahmin etmede Google Trend gibi alternatif veri kaynaklarının kullanılabilirliğini ve farklı ölçeklerde (bölgesel veya ulusal) etkili sonuçlar elde edilebileceğini vurgulamaktadır.

Celbiş (2022) çalışması, kırsal bölgelerde işsizlikle ilişkili temel bireysel faktörleri belirlemek ve politika geliştirmeye katkı sağlamak amacıyla ağaç tabanlı sınıflandırma modelleri kullanmıştır. Sınıflandırma ağacı, bootstrap agregasyonu, rastgele orman ve gradyan artırma algoritmalarıyla yapılan analizlerde, eğitim, cinsiyet, yaş ve ebeveyn eğitimi gibi faktörlerin iş gücü piyasasındaki eşitsizlikleri azaltmada kritik olduğu bulunmuştur. SHAP analiziyle desteklenen bulgular, eğitim programlarına erişimin önemini ve bazı alt gruplarda eşitsizliklerin daha belirgin olduğunu vurgulamaktadır.

Yıldız (2022) çalışmasında , TÜİK'in İllerde Yaşam Endeksi ve diğer veri setlerini kullanarak Türkiye'de işsizlik oranını tahmin etmek için veri madenciliği ve makine öğrenmesi yöntemleri uygulanmıştır. Farklı algoritmaların karşılaştırıldığı analizde, Genel Regresyon Sinir Ağı (GRNN) yönteminin SVM ve Gen İfadesi Programlama (GEP) yöntemlerinden daha başarılı olduğu bulunmuştur. Ayrıca, eğitimle ilgili verilerin işsizlik tahmininde en etkili değişken grubu olduğu belirlenmiştir.

Wei vd. (2023) çalışması, üniversite mezunlarının “yavaş istihdam” riskini tahmin etmek için geliştirilmiş yarasa algoritması (GEBA) ve destek vektör makinelerini (SVM) birleştirerek bGEBA-SVM adında bir makine öğrenmesi modeli sunmuştur. 1.694 öğrenci verisiyle test edilen model, %93,86 doğruluk oranıyla diğer yöntemlerden daha iyi sonuç

vermiştir. Çalışma, bireysel ve ailesel faktörlerin istihdam olasılığı üzerindeki etkisini vurgulamış ve eğitim kurumları için stratejik öneriler sunmuştur.

Yamaçlı ve Yamaçlı (2023) çalışmalarında Türkiye'deki işsizlik oranının tahmininde ARIMA ve Yapay Sinir Ağları (YSA) yöntemlerinin performansı karşılaştırılmıştır. 2008-2022 dönemi aylık verilerle yapılan analizde, ARMA (2,1) modelinin Root Mean Squared Error – Kök Ortalama Kare Hatası (RMSE), Mean Absolute Error – Ortalama Mutlak Hata (MAE), Mean Absolute Percentage Error – Ortalama Mutlak Yüzde Hatası (MAPE) ve Belirleme Katsayısı (R^2) kriterlerine göre en uygun model olduğu bulunmuştur. Ancak, Covid-19 pandemisi gibi ekonomik belirsizlik dönemlerinde YSA'nın daha düşük tahmin hataları sunduğu gözlemlenmiştir. Sonuçlar, ARIMA'nın işsizlik oranındaki kronik eğilimleri daha iyi yakalarken, YSA'nın belirsizlik dönemlerinde avantajlı olduğunu göstermektedir.

Yurtsever (2023) çalışmasında, Amerika Birleşik Devletleri, Birleşik Krallık, Fransa ve İtalya'nın işsizlik oranlarını tahmin etmek amacıyla, Ocak 1983 ile Mayıs 2022 arasındaki aylık işsizlik verilerini kullanarak, LSTM ve GRU (Geçitli Yinelemeli Birim) katmanlarını birleştiren bir model geliştirmiştir. Sonuçlar, hibrit modelin İtalya dışında kalan dört ülke için daha başarılı tahminler sağladığını göstermektedir.

Vinaya vd. (2023) çalışmalarında, işsizlik oranlarının tahmininde en iyi yöntemi belirlemek için farklı modelleri karşılaştırmıştır. Spiral tabanlı sinir ağı ve öğrenme vektör kantizasyonu birlikte kullanıldığında tahmin doğruluğunun arttığı tespit edilmiştir. Ayrıca, doğrusal regresyon, destek vektör regresyonu, SVM, Rastgele Orman ve Gradyan Artırma gibi yöntemlerle tahmin performansı daha da iyileştirilmiştir. Elde edilen modeller, gelecekteki işsizlik oranlarını doğru bir şekilde tahmin edebilmek için bir eğitim modeli oluşturulmasını sağlamaktadır.

Aliyev (2024) çalışmasında, Türkiye'deki işsizlik oranlarının tahmin edilmesinde yapay zekâ, makine öğrenimi ve klasik yöntemlerin karşılaştırılmasını amaçlamaktadır. TÜİK verileriyle 2005-2023 dönemi işsizlik oranı verileri kullanılmıştır. ARIMA, SARIMA, Rastgele Orman, XGBoost, LSTM ve GRU modelleri uygulanmıştır. Sonuçlar, SARIMA modelinin performansının ARIMA modeline göre daha iyi olduğunu göstermektedir. Makine öğrenimi modelleri arasında ise, en uygun tahmin sonuçlarını GRU modeli vermektedir, çünkü MAE ve RMSE değerleri en düşük, R^2 değeri ise en yüksektir. Rastgele Orman modeli ise buna en yakın performansı sergilemiştir.

Güler vd. (2024), Türkiye'de 2005–2023 verileriyle işsizlik oranlarını tahmin etmek için ANN, SVM ve XGBoost algoritmalarını karşılaştırmıştır. Yalnızca geçmiş verilerle ve ek ekonomik göstergelerle iki model geliştirilmiş, performansları MAE, RMSE ve MAPE ile değerlendirilmiştir. XGBoost'un en yüksek doğruluğu sağladığı belirlenmiş; sonuçlar, makine öğrenmesi yöntemlerinin işsizlik tahmininde etkili olduğunu ve politika yapıcılara destek sunabileceğini göstermiştir.

Maigur (2024) çalışmasında, 2017–2021 verilerini kullanarak Rusya'daki işsiz bireylerin iş bulma süresini tahmin etmek için Cox orantılı tehlike modeli ve sinir ağı tabanlı Cox-Time modelini karşılaştırmıştır. Çalışmada, Cox-Time modeli %64 doğruluk oranıyla en iyi sonucu vermiş ve iş aranan bölgenin işsizlik süresi üzerinde en etkili faktör olduğu belirlenmiştir.

3. Veri Seti ve Yöntem

3.1. Veri Seti

Çalışmada kullanılan veri seti 2005-2024 yılları arasındaki TÜİK verilerinden elde edilmiştir. Tüm değişkenler için veri bütünlüğünün ve karşılaştırılabilirliğinin sağlanabilmesi amacıyla, veri seti 2005 yılından başlatılmıştır. İşsizlik, Enflasyon, Eğitim Harcamaları, Ekonomik Büyüme ve Nüfus Artış oranı yüzde (%) oran şeklinde değerlendirilirken, Göç verisi olarak Türkiye’de geçici koruma dışında kalan mülteciler ele alınırken; göç verisi sayısal değer olarak kaydedilmiştir. Tablo 1’de 20 yıllık verilerin değişkenlere göre istatistiksel dağılımını gösterilmektedir.

Tablo 1. Değişkenlerin İstatistiksel Özellikleri

	Maks.	Min.	Ortalama	Std	Çarpıklık	Basıklık
İşsizlik	15.1	7.0	10.38	1.83	0.56	-0.4
Enflasyon	72.45	5.93	16.85	17.63	2.1	2.87
Eğitim	12.14	8.2	10.31	1.17	-0.18	-1.11
Harcamaları						
Ekonomik	22.3	-14.5	5.18	5.01	-0.94	4.22
Büyüme						
Nüfus Artış	1.5	0.16	1.1	0.33	-1.46	1.71
Göç	15.14	7.78	12.61	2.87	-0.57	-1.48

Tablo 1’de yer alan tanımlayıcı istatistikler, işsizlik, enflasyon, eğitim harcamaları, ekonomik büyüme, nüfus artışı ve göç oranı gibi temel makroekonomik değişkenlerin dağılım özelliklerine dair genel bir çerçeve sunmaktadır. Veriler incelendiğinde, enflasyon ve ekonomik büyümenin görece yüksek varyans sergilediği ve zaman içinde önemli dalgalanmalar gösterdiği görülmektedir. Buna karşılık, eğitim harcamaları ve nüfus artış oranı daha sınırlı değişkenliğe sahip olup, daha istikrarlı bir yapı arz etmektedir. İşsizlik ve göç oranları ise orta düzeyde dalgalanma göstermekte ve dağılımları genel olarak dengeli bir görünüm sunmaktadır. Bu sonuçlar, ilgili ekonomik göstergelerin dönemsel dinamiklerini ortaya koymakta ve sonraki ekonometrik analizler için önemli bir zemin hazırlamaktadır.

2.2. Yöntem

Giriş verileri olarak enflasyon, eğitim harcamaları, ekonomik büyüme, nüfus artışı ve göç değişkenleri kullanılmıştır. Veriler üzerinde minimum-maksimum normalizasyonu uygulanarak ölçeklendirme işlemi gerçekleştirilmiştir. Göç verileri ise logaritması alınarak dönüştürülmüş ve ardından karesi hesaplanarak (enerji değeri) makine öğrenmesi modellerine giriş değişkeni olarak kullanılmıştır.

2.2.1. Rastgele Ormanlar (RF)

Rastgele Ormanlar (RF), birden fazla Karar Ağacı (DT)’nı birleştirerek daha sağlam ve geliştirilebilir bir model oluşturan bir makine öğrenimi tekniğidir (Breiman, 2001; Öter, 2024). Aşırı uyum, makine öğreniminde önemli bir zorluktur ve bu zorluğun üstesinden gelmek ve sağlamlık sağlamak için RF yöntemi, rastgele özellik seçimi ile örnekleme ve rastgeleleştirme kullanır. Bu nedenle, her ağaç farklı bir alt küme üzerinde eğitilirken, her düğümde bölmek için rastgele bir özellik alt kümesi kullanılır ve bu da modelin genelleme yeteneğini artırır (Hastie vd, 2009,s.64). Ormandaki her ağaç bir sınıf

tahmini verir ve nihai tahmin bu bireysel tahminlerin ortalaması olarak hesaplar ve şu şekilde verilir:

$$y = \frac{1}{n} \sum_{i=1}^n y_i \quad (1)$$

Burada y tahmin edilen değer ve m ormandaki toplam ağaç sayısıdır.

RF yöntemi, veri kümeleri kapsamlı ve karmaşık olduğunda yüksek doğruluk sağlayarak sınıflandırma ve regresyon uygulamalarında performansı artırır. Ayrıca RF, modele girdi olan özniteliklerin ağırlıklarının anlaşılmasını ve açıklanmasını sağlar. Birçok çalışma, RF'nin çeşitli veri kümeleri üzerinde geleneksel karar ağaçları ve diğer topluluk yöntemlerinden daha iyi performans gösterdiğini ve finansal tahmin, görüntü işleme, tıbbi teşhis ve biyoinformatik gibi alanlarda uygulanabilir olduğunu göstermektedir. Özetle, RF'de uygulanan topluluk öğrenme ilkeleri ile rastgeleliği birleştirmek, sağlam ve uyarlanabilir bir tahmin modeli oluşturur.

2.2.2. Gradient Boosting (GB)

Gradyan Artırma (Gradient Boosting - GB) olarak bilinen topluluk tekniği, zayıf öğrenenleri, özellikle karar ağaçlarını art arda eğiterek bir tahmin sorununu ele almak için güçlü bir model oluşturur. Bu yaklaşımla, her yeni öğrenen önceki model tarafından yapılan hataları en aza indirmek için eğitilir. Gözlemsel ve öngörü tutarsızlıkları, modelleme sürecini yönlendiren kayıp fonksiyonunun gradyanını oluşturur (Öter, 2024). Bu iteratif süreç, hataları minimize ederek her bir ağacın performansını artırmayı hedefler. GB'nin en önemli avantajı, büyük ölçekli veri setleri ve karmaşık özellikler üzerinde çalışırken yüksek doğruluk oranları elde etme yeteneğidir (Chen ve Guestrin, 2016). Bu özellikleri, onu özellikle zorlu problemlerde etkili bir yöntem haline getirmektedir.

2.2.3. Çok Katmanlı Algılayıcı (MLP)

İnsan nöron ağı ve sinir sisteminden esinlenen Yapay Sinir Ağları (YSA) ve bunların bir çeşidi olan Çok Katmanlı Algılayıcı (MLP), karmaşık veri setlerinden öğrenme ve gürültülü veriler içerse bile karmaşık doğrusal/doğrusal olmayan ilişkileri tahmin etme yeteneğine sahiptir. İnsan nöron ağından ve sinir sisteminden esinlenen yapı, nöronlar ve katmanlardan oluşan dögümsel bir ağ yapısına sahiptir (Öter 2024; Öter vd. 2024). Birkaç nöron bir araya geldiğinde katmanlar oluşur ve katmanlar arasında aktivasyon fonksiyonu uygulanarak ve her nörona ağırlıklar atanarak nöronlar arasında bilgi iletilir. YSA ağı eğitim sırasında bilgiyi öğrenir ve genelleştirir, test aşamasında ise yeni verileri sınıflandırır. Ağın ağırlıklarını ayarlayarak birden fazla model oluşturabilen YSA yöntemi, veri boyutuna bağlı olarak hızlı çalışabilir, gürültülü verileri tolere edebilir ve derin öğrenme de dahil olmak üzere geniş bir uygulama alanına sahiptir. YSA modelinin birçok avantajına rağmen, hesaplama talebine bağlı olarak güçlü donanım gerektirebilir. Çözümün detayları elde edilemediği için kara kutu tekniği olarak da anılmaktadır. Ancak son yıllarda açıklanabilir yapay zeka teknikleri ile kara kutu yöntemleri anlaşılabilir ve yorumlanabilir hale gelmiştir (Haeffele ve Vidal, 2017)

Nöronlardan oluşan klasik bir YSA modelinin çıkışı Net, giriş sayısı n ve ağırlık w_i , giriş x_i ve bias b ile aktivasyon fonksiyonu f ile aşağıdaki gibi hesaplanır.

$$Net = \sum_{i=1}^n f(w_i x_i + b_i) \quad (2)$$

2.3. Değerlendirme Kriterleri

Model performansını değerlendirmek için değerlendirme metrikleri kullanmak, yeni veriler üzerinde üretime geçmeden önce bir modelin tahmin gücünü anlamamızı ve geliştirmemizi sağlar. Bu çalışmada, RMSE, ve MAE, gibi bir regresyon modelinin performansını değerlendirmek ve raporlamak için sıklıkla kullanılan üç hata ölçüsü kullanılmıştır. RMSE ve MAE, bir modelin tahmin edilen ve gerçek değerleri arasındaki sapmayı hesaplar (Öter vd. 2024; Sezer vd. 2024). Bu çalışmada, deneysel sonuçlarımıza en yakın tahminsel modellemeyi bulmayı, deneysel değerlerden sapmaları görmeyi ve bağımlı değişken ile bağımsız değişken arasındaki ilişkileri yakalamayı amaçlıyoruz. Yapılan tahmin çalışmalarında MSE ve MAE metrikleri kullanılmıştır.

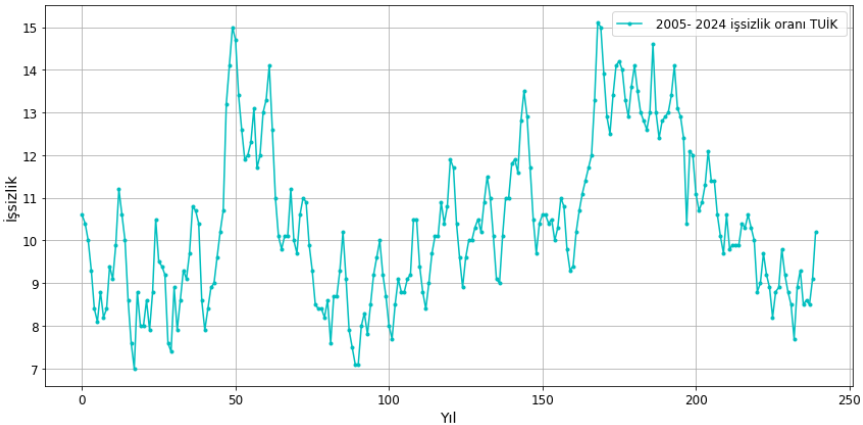
$$RMSE = \sqrt{\frac{1}{N} \sum_{i=1}^N (Y_i - Y_{predict(i)})^2} \quad (3)$$

$$MAE = \frac{1}{n} \sum_{i=1}^n |Y_i - Y_{mean}| \quad (4)$$

n : veri değeri sayısı; Y_i : kümedeki veri değerleri; Y_{mean} : veri kümesinin ortalama değeri $Y_{predict}$: veri kümesinin tahmin edilen değeri göstermektedir.

3. Bulgular ve Tartışma

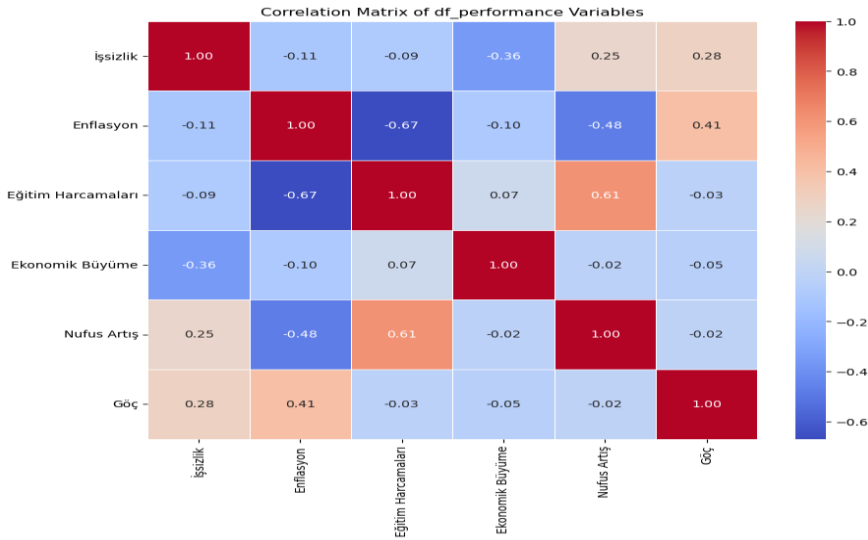
Çalışmada, 2005-2024 yılları arasında TÜİK verilerinden elde edilen veri seti kullanılmış ve bu veri setindeki değişkenler arasındaki ikili ilişkiler analiz edilmiştir. Şekil 1 de İşsizlik oranının 20 yıllık dağılımı gösterilmiştir. En yüksek oran 15.00 ile 49 ay (2009 Ocak) ve 15.8 ile 168. Ay (2016 Aralık) olarak kaydedilmiştir. Çalışmada Enflasyon, Eğitim Harcamaları, Ekonomik Büyüme, Nüfus Artış ve Göç giriş değişkeni olarak, İşsizlik ise çıkış değişkeni olarak kullanılmıştır. Eğitim verileri olarak 2005-2022 yılları arasındaki veriler kullanılmıştır. Test verileri için 2022-2024yıllarındaki veriler kullanılmıştır. Göç değerlerinin kullanımında logaritması alınarak optimizasyon işlemi yapılmış elde edilen değerler önerilen modellerde kullanılmıştır.



Şekil 1. 20 Yıllık (240 Aylık) İşsizlik Oranlarının Dağılımı

2005-2024 yılları arasındaki 240 aylık işsizlik oranlarının değişimini göstermektedir. İşsizlik oranı genel olarak 7% ile 15% arasında dalgalanmaktadır ve belirli dönemlerde keskin artışlar ve düşüşler görülmektedir. Özellikle 2008-2009 yıllarında küresel ekonomik krizle birlikte işsizlik oranının hızla yükseldiği ve 15% seviyelerine ulaştığı dikkat çekmektedir. Benzer şekilde, 2018-2020 yılları arasında ekonomik kriz ve COVID-19 pandemisinin etkisiyle işsizlik oranında keskin bir artış yaşanmıştır. 2020 sonrasında ise işsizlik oranında bir düşüş eğilimi gözlemlense de dalgalanmalar devam etmektedir. 2024 yılına doğru işsizlik oranı 8%-10% seviyelerinde seyretmektedir, ancak gelecekteki ekonomik ve politik gelişmelere bağlı olarak bu oranın nasıl şekilleneceği belirsizdir. Genel olarak, grafik Türkiye’de işsizlik oranlarının ekonomik krizler, enflasyon, büyüme oranları ve pandemiler gibi çeşitli faktörlerden etkilendiğini ortaya koymaktadır.

Değişkenler arasındaki ilişkilerin detaylı görselleştirilmesi Şekil 2’de sunulmuştur.



Şekil 2. Değişkenler Arası İkili İlişkinin Dağılımı

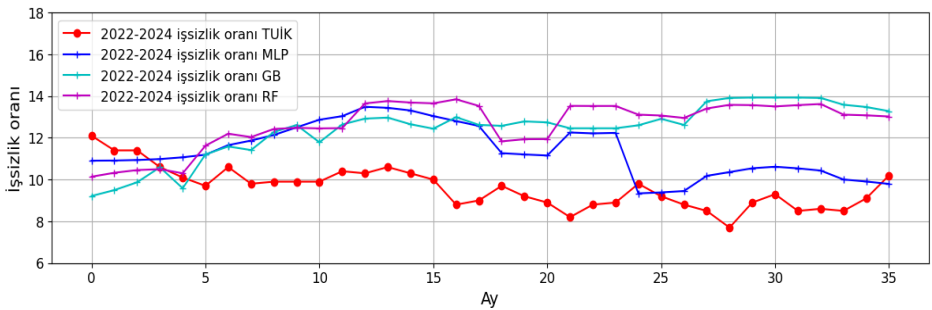
Şekil 2’deki korelasyon matrisi incelendiğinde, bazı değişkenler arasında anlamlı ve dikkat çekici ilişkiler olduğu görülmektedir. Enflasyon ile eğitim harcamaları arasında gözlenen güçlü negatif ilişki, enflasyon oranı yükseldikçe eğitim harcamalarının görece azaldığını göstermektedir. Ekonomik büyüme ile işsizlik oranı arasında ise negatif bir ilişki söz konusudur; bu durum, büyümenin hızlandığı dönemlerde işsizliğin azaldığına işaret etmektedir. Ayrıca nüfus artışı ile eğitim harcamaları arasında pozitif yönlü bir ilişki görülmekte, bu da artan nüfusun kamu eğitim yatırımları üzerinde artırıcı bir etkisi olabileceğini düşündürmektedir. Göç oranı ise hem enflasyon hem de işsizlik ile pozitif yönde ilişkilidir; bu durum, özellikle ekonomik belirsizlik dönemlerinde göç hareketlerinin arttığını göstermektedir. Bu ilişkiler, değişkenler arasındaki temel etkileşimleri anlamak ve ileri analizlerde nedensel bağlantılar kurmak açısından önemli bir temel sunmaktadır.

Çalışmada RF, GB, MLP 3 farklı ML yöntemi kullanılarak gerçekleştirilmiştir. Her ML modeli Tablo 2’de verilen eğitim parametreleri kullanılarak en iyi parametreler belirlenmiştir.

Tablo 2 . ML Modellerinin Parametreleri

Modeller	Eğitimlerde Kullanılan Parametreler	En İyi Modelin Parametreleri
RF	n_estimators: [50, 100, 150], max_depth: [10, 20, 30], min_samples_split: [2, 5, 10], min_samples_leaf: [1, 2, 4]	n_estimators: 150 max_depth: 20, min_samples_leaf: 1, min_samples_split: 2,
GB	n_estimator: 100, 200, 300 max_depth: 3, 5, 7 min_samples_split: 2, 5, 10 min_samples_leaf: 1, 2, 4 subsample: 0.8, 1.0 learning_rate: 0.01, 0.05, 0.1 random_state:42	n_estimators:200 learning_rate: 0.2 max_depth:7 min_samples_split:2 min_samples_leaf:2 subsample:0.8 random_state=42,
MLP	hidden_layer_sizes: [(50,),(100,),(50, 50), (100, 50)], activation: ['relu', 'tanh', 'logistic'], solver: ['adam', 'sgd'], alpha: [0.0001, 0.001, 0.01], learning_rate: ['constant', 'adaptive'], max_iter: [200, 500, 1000]	hidden_layer_sizes: (50, 50), activation: 'relu', solver: 'adam' alpha: 0.001, learning_rate: 'constant', max_iter: 1000,

Tablo 2’de ML modellerinde kullanılan en iyi parametreler belirlenmesi ile bu modellerler kullanılarak 2022-2024 yılları arasındaki oluşturulan veriler ile Test işlemleri gerçekleştirilmiştir. Şekil 3 ‘te her model için tahmin değerler verilmiştir.



Şekil 3 2022-2024 Arasındaki Veriler ile ML Yöntemlerinin Karşılaştırılması

Şekil 3’te 2022-2024 arasındaki 36 aylık enflasyon, eğitim harcamaları, ekonomik büyüme, nüfus artış oranları ve göç verileri kullanılarak 3 farklı ML yöntemi ile Tahmin edilen İşsizlik verileri ile TÜİK verilerinin 36 aylık değişimi verilmiştir. TÜİK verileri 36 aylık dönemde % 8-12 arasında değişmiştir. ML yöntemleri ile yapılan tahmin çalışmalarında ise TÜİK en yakın değerler MLP yöntemiyle elde edilmiştir. RF ve GB yöntemleri Benzer sonuçlar vererek %9-14 arasında değişmektedir. Göç oranları logaritmik olarak kullanıldığında aşağıda verilen şekilde ise MLP yönteminin TÜİK verilerine daha yakın

olduğu gözlemlenmiştir.

Tablo 3 yıllık test verisi

Model	MAE	RMSE
RF	3,277	3,589
GB	3,174	3,488
MLP	1,945	2,235

Tablo 3 'te, işsizlik oranlarının tahmini için kullanılan RF, GB ve MLP modellerinin performansları karşılaştırılmıştır. Modellerin başarısı MAE ve RMSE değerleri üzerinden değerlendirilmiştir. Sonuçlara göre, en düşük hata oranlarına sahip olan MLP modeli (MAE: 1,945 ; RMSE: 2,235) en başarılı tahmin performansını göstermiştir. GB modeli, RF modeline kıyasla daha iyi bir performans sergilese de MLP kadar düşük hata oranına ulaşamamıştır. RF modeli, en yüksek hata oranına sahip olduğu için en düşük tahmin doğruluğunu göstermiştir. Bu bulgular, işsizlik oranlarının tahmininde MLP modelinin diğer modellere kıyasla daha başarılı olduğunu ortaya koymaktadır. Gelecekte tahmin performansını artırmak için MLP modelinin daha fazla veriyle eğitilmesi veya farklı mimarilerin denenmesi önerilmektedir.

4. Sonuç ve Öneriler

Bu çalışma, Türkiye'de 2005-2024 dönemine ait işsizlik oranlarının tahmininde makine öğrenmesi algoritmalarının performansını değerlendirmek ve geleneksel istatistiksel yöntemlerle karşılaştırmak amacıyla gerçekleştirilmiştir. Çalışmada, işsizlik oranlarının ekonomik büyüme, enflasyon, eğitim harcamaları, nüfus artışı ve göç gibi çeşitli makroekonomik değişkenlerden etkilendiği varsayımından hareketle, tahmin modellerine bu değişkenler dahil edilmiştir.

Analizlerin ilk aşamasında, bağımlı ve bağımsız değişkenler arasındaki ilişkiler korelasyon analizi aracılığıyla incelenmiştir. Bulgular, ekonomik büyüme ve eğitim harcamalarının işsizlik oranı ile negatif; enflasyon ve göçün ise pozitif yönlü anlamlı ilişkiler gösterdiğini ortaya koymuştur. Bu ilişkiler, işgücü piyasasının çok boyutlu yapısını yansıtmakta ve tahmin sürecinde dikkate alınması gereken temel faktörleri işaret etmektedir.

Korelasyon analizinin ardından, tahmin sürecinde Rastgele Ormanlar (RF), Gradyan Artırma (GB) ve Çok Katmanlı Algılayıcı (MLP) algoritmaları uygulanmış; modellerin performansları Ortalama Mutlak Hata (MAE) ve Kök Ortalama Kare Hata (RMSE) kriterlerine göre karşılaştırılmıştır. Elde edilen bulgular, en düşük hata oranlarının MLP modeliyle (MAE: 1.945, RMSE: 2.235) elde edildiğini, dolayısıyla bu modelin en başarılı tahmin performansını sergilediğini göstermiştir. GB ve RF modelleri de makul doğruluk seviyelerine ulaşmakla birlikte, hata oranları MLP modeline kıyasla daha yüksek bulunmuştur. Özellikle, MLP modelinin doğrusal olmayan ilişkileri öğrenebilme kapasitesi, işsizlik oranlarının tahmininde üstün performans göstermesinin temel nedenlerinden biri olarak değerlendirilmektedir. Ayrıca, ileriye dönük tahminlerde MLP modelinin, 2025-2027 yılları için işsizlik oranlarına dair en tutarlı ve TÜİK verilerine en yakın tahminleri sunduğu belirlenmiştir.

Çalışmadan elde edilen bulgular, işgücü piyasası için çeşitli politika önerileri geliştirilmesine zemin oluşturmıştır. Bu kapsamda; dijital dönüşümün yarattığı yeni çalışma biçimlerine uyum sağlanması, esnek çalışma modellerinin teşvik edilmesi, kayıt

dışı istihdamın azaltılmasına yönelik vergi ve sosyal güvenlik reformlarının hayata geçirilmesi öncelikli stratejiler arasında önerilmektedir. Ayrıca, işgücü piyasasının ihtiyaç duyduğu niteliklerin karşılanabilmesi amacıyla STEM alanlarına yönelik eğitim yatırımlarının artırılması ve yaşam boyu öğrenme fırsatlarının yaygınlaştırılması önem arz etmektedir. İstihdamın artırılmasına yönelik olarak, küçük ve orta ölçekli işletmelerin desteklenmesi, girişimcilik ekosisteminin güçlendirilmesi ve yeşil ekonomi ile dijital ekonomi gibi yenilikçi sektörlere yatırım yapılması önerilmektedir. Bunun yanı sıra, bölgesel kalkınma politikaları kapsamında altyapı yatırımlarının artırılması ve yerel ekonomilerin desteklenmesi, işsizlik oranlarının düşürülmesine katkı sağlayacak diğer stratejiler arasında yer almaktadır.

Sonuç olarak, bu çalışma, makine öğrenmesi algoritmalarının Türkiye’de işsizlik oranlarının tahmininde yüksek doğruluk sunduğunu ve özellikle MLP modelinin üstün performans sergilediğini ortaya koymuştur. Çalışma bulguları, işsizlik gibi karmaşık ve çok değişkenli ekonomik bir göstergenin tahmininde makine öğrenmesi yöntemlerinin etkili bir araç olabileceğini göstermesi bakımından literatüre katkı sağlamaktadır. Gelecekte, daha geniş veri setleri ve alternatif algoritmalarla gerçekleştirilecek çalışmalarla model performanslarının daha da artırılabilirliği değerlendirilmektedir.

Değerlendirme	İki Dış Hakem / Çift Taraflı Körleme
Etik Beyan	Bu makale etik kurul kararı gerektirmemektedir ve daha önce herhangi bir yerde sözlü bildiri olarak sunulmamış ve bir tez çalışmasından üretilmemiştir. Tüm etik ilkeler uygun biçimde araştırmada göz önünde bulundurulmuştur.
Benzerlik Taraması	Yapıldı – Ithenticate
Etik Bildirim	itobiad@itobiad.com
Çıkar Çatışması	Çıkar çatışması beyan edilmemiştir.
Finansman	Bu araştırmayı desteklemek için dış fon kullanılmamıştır.
Yazar Katkıları	Çalışmanın Tasarlanması: 1. Yazar (%50), 2. Yazar (%50) Veri Toplanması: 1. Yazar (%50), 2. Yazar (%50) Veri Analizi: 1. Yazar (%50), 2. Yazar (%50) Makalenin Yazımı: 1. Yazar (%50), 2. Yazar (%50) Makale Gönderimi ve Revizyonu: 1. Yazar (%50), 2. Yazar (%50)
Peer-Review	Double anonymized - Two External
Ethical Statement	This article does not require ethics committee approval and has not been previously presented as an oral presentation anywhere and is not derived from a thesis. All ethical principles have been appropriately considered in the research.
Plagiarism Checks	Yes - Ithenticate
Conflicts of Interest	The author(s) has no conflict of interest to declare.
Complaints	itobiad@itobiad.com
Grant Support	The author(s) acknowledge that they received no external funding in support of this research.
Author Contributions	Design of Study: 1. Author (%50), 2. Author (%50) Data Acquisition: 1. Author (%50), 2. Author (%50) Data Analysis: 1. Author (%50), 2. Author (%50) Writing up: 1. Author (%50), 2. Author (%50) Submission and Revision: 1. Author (%50), 2. Author (%50)

Kaynakça | References

- Ahmad, M., Khan, Y. A., Jiang, C., Kazmi, S. J. H., & Abbas, S. Z. (2023). The impact of COVID-19 on unemployment rate: An intelligent based unemployment rate prediction in selected countries of Europe. *International Journal of Finance & Economics*, 28(1), 528-543.
- Aliyev, P. (2023). İşsizlik oranı öngörülerinde makine öğrenimi yaklaşımları: Türkiye üzerine bir uygulama. *Iğdir University Journal of Faculty of Economics and Administrative Sciences*, 11, 1-15.
- Breiman, L. (2001). Random forests. *Machine Learning*, 45, 5-32.
- Celbiş, M. G. (2023). Unemployment in rural Europe: A machine learning perspective. *Applied Spatial Analysis and Policy*, 16(3), 1071-1095.
- Chakraborty, T., Chakraborty, A. K., Biswas, M., Banerjee, S., & Bhattacharya, S. (2021). Unemployment rate forecasting: A hybrid approach. *Computational Economics*, 57(1), 183-201.
- Çiftçi, C. (2016). Forecasting of unemployment rate for Turkey: Markov chains approach. *Eurasian Academy of Sciences Eurasian Business & Economics Journal*, S2, 657-665. <http://dx.doi.org/10.17740/eas.econ.2016-MSEMP-140>
- Davidescu, A. A., Apostu, S.-A., & Marin, A. (2021). Forecasting the Romanian unemployment rate in time of health crisis—a univariate vs. multivariate time series approach. *International Journal of Environmental Research and Public Health*, 18(21), 11165.
- Gil-Alana, L. A., Özdemir, Z. A., & Tansel, A. (2019). Long memory in Turkish unemployment rates. *Emerging Markets Finance and Trade*, 55(1), 201-217.
- Güler, M., Kabakçı, A., Koç, Ö., Eraslan, E., Derin, K. H., Güler, M., Ünlü, R., Türkan, Y. S., & Namlı, E. (2024). Forecasting of the unemployment rate in Turkey: Comparison of the machine learning models. *Sustainability*, 16(15), 6509. <https://doi.org/10.3390/su16156509>
- Karahan, M., & Çetintaş, F. (2022). Forecasting of Turkey's unemployment rate for future periods with artificial neural networks. *Erciyes Üniversitesi İktisadi ve İdari Bilimler Fakültesi Dergisi*, 62, 163-184.
- Katris, C. (2020). Prediction of unemployment rates with time series and machine learning techniques. *Computational Economics*, 55(2), 673-706.
- Kızılkaya, O. (2017). Türkiye'nin enflasyon ve işsizlik oranının yapay sinir ağları ve Box-Jenkins yöntemiyle tahmini. *Social Sciences Studies Journal*, 3(12), 2197-2207.
- Kütük, Y., & Güloğlu, B. (2019). Prediction of transition probabilities from unemployment to employment for Turkey via machine learning and econometrics: A comparative study. *İktisat Araştırmaları Dergisi*, 3(1), 58-75. <https://doi.org/10.24954/jore.2019.29>
- Lai, H., Khan, Y. A., Thaljaoui, A., Chammam, W., & Abbas, S. Z. (2023). Retracted article: COVID-19 pandemic and unemployment rate: A hybrid unemployment rate prediction approach for developed and developing countries of Asia. *Soft Computing*, 27(1), 615-615.
- Maigur, A. A. (2024). Machine learning algorithms for predicting unemployment duration in Russia. *Russian Journal of Economics*, 10(4), 365-384.

<https://doi.org/10.32609/j.ruje.10.128611>

Mulaudzi, R., & Ajoodha, R. (2020). An exploration of machine learning models to forecast the unemployment rate of South Africa: A univariate approach. *2020 2nd International Multidisciplinary Information Technology and Engineering Conference (IMITEC)*.

Olmedo, E. (2014). Forecasting Spanish unemployment using near neighbour and neural net techniques. *Computational Economics*, 43, 183-197.

Öter, A. (2024). Automatic detection of epileptic seizures from EEG signals using artificial intelligence methods. *Gazi University Journal of Science Part C: Design and Technology*, 1, 257-266.

Öter, A., Ersöz, B., Berktaş, Z., Bülbül, H. İ., Orhan, E., & Sağiroğlu, Ş. (2024). An artificial intelligence model estimation for functionalized graphene quantum dot-based diode characteristics. *Physica Scripta*, 99(5), 056001. <https://doi.org/10.1088/1402-4896/ad3515>

Sen, M., Basu, S., Chatterjee, A., Banerjee, A., Pal, S., Mukhopadhyay, P. K., Dutta, S., & Tarafdar, A. (2022). Prediction of unemployment using machine learning approach. *2022 OITS International Conference on Information Technology (OCIT)*.

Sezer, S., Öter, A., Ersöz, B., Topcuoğlu, C., Bülbül, H. İ., Sağiroğlu, Ş., Akın, M., & Yılmaz, G. (2024). Explainable artificial intelligence for LDL cholesterol prediction and classification. *Clinical Biochemistry*, 110791. <https://doi.org/10.1016/j.clinbiochem.2024.110791>

Simionescu, M., & Cifuentes-Faura, J. (2022). Forecasting national and regional youth unemployment in Spain using Google Trends. *Social Indicators Research*, 164(3), 1187-1216.

TÜİK. İstihdam, işsizlik, ücret. <https://data.tuik.gov.tr/Kategori/GetKategori?p=istihdam-issizlik-ve-ucret-108&dil=1>

Tüzemen, D., & Yıldız, H. (2017). İşsizlik oranlarının tahmininde zaman serisi analizleri: Türkiye örneği. *İktisat ve Finans Dergisi*, 32(4), 89-112.

Wei, Y., Rao, X., Fu, Y., Song, L., Chen, H., & Li, J. (2023). Machine learning prediction model based on enhanced bat algorithm and support vector machine for slow employment prediction. *PLOS ONE*, 18(11), e0294114. <https://doi.org/10.1371/journal.pone.0294114>

Yamaçlı, D. S., & Yamaçlı, S. (2023). Estimation of the unemployment rate in Turkey: A comparison of the ARIMA and machine learning models including Covid-19 pandemic periods. *Heliyon*, 9(1), e12796. <https://doi.org/10.1016/j.heliyon.2023.e12796>

Yıldırım, H., & Başeğmez, H. (2017). Analysis and forecast of Turkey unemployment rate. *Global Journal of Mathematical Analysis*, 1(1), 11-15. <https://doi.org/10.14419/gjma.v5i1.6841>

Yıldız, İ. (2022). Veri madenciliği: Makine öğrenme algoritmaları ile Türkiye'nin işsizlik oranı tahminini etkileyen faktörlerin tespit edilmesi. *International Journal of Management Information Systems and Computer Science*, 6(2), 78-91. <https://doi.org/10.33461/uybisbbd.1129013>

Yolcu, U., & Baş, E. (2016). The forecasting of labour force participation and the unemployment rate in Poland and Turkey using fuzzy time series methods. *Comparative*

Economic Research, 19(2), 5–25. <https://doi.org/10.1515/cer-2016-0010>

Yurtsever, M. (2023). Unemployment rate forecasting: LSTM-GRU hybrid approach. *Journal for Labour Market Research*, 57(1), 18.

Zhao, L. F. (2020). Data-driven approach for predicting and explaining the risk of long-term unemployment. *E3S Web of Conferences*, 214, 01023. <https://doi.org/10.1051/e3sconf/202021401023>