**MAKİNE ÖĞRENMESİ ALGORİTMALARI İLE EKONOMİK RESESYON ÖNGÖRÜSÜ**

**ÖZET**

*Çalışmanın amacı çeşitli makine öğrenmesi algoritmaları ile Amerikan ekonomisindeki iktisadi resesyonların öngörülmesidir. Ekonomik durgunlukların gerçekleşme sıklıklarının düşük olması, dengesiz sınıf probleminin ortaya çıkmasına neden olmaktadır. Bu problemin üstesinden gelmek için üç farklı örnekleme yöntemi (eksik örnekleme, aşırı/sentetik örnekleme, normal örnekleme) kullanılarak dengesiz sınıf problemini ortadan kaldıracak uygun örnekleme yöntemi belirlenmiştir. Eğri altı alan (AUC) kriterine göre uygun örnekleme yöntemi eksik örnekleme olarak tespit edilmiştir. Analizin ikinci aşamasında öngörü başarısı yüksek algortimayı belirlemek için öğrenme verisi olarak Ocak 1982-Haziran 2021 dönemi için Amerikan Merkez Bankası’nın veri tabanından aylık resesyon verisi ile finansal ve makroekonomik zaman serileri (işsizlik, üretim endeksi, S&P 500 endeksi, federal faiz oranı, getiri eğrisi, reel para stoğu) elde edilmiştir. Üretim endeksi ve işsizlik değişkenleri reel sektördeki değişimleri; reel para stoğu, federal faiz oranı ve S&P 500 endeksindeki değişimler para politikasındaki beklenen ve beklenmeyen etkileri temsilen resesyon üzerinde etkili olabilecek değişkenler olarak belirlenmiştir. Bu verilere dayalı olarak belirli bir aydaki durgunluk olasılığı, öngörü başarısı en yüksek olan makine öğrenmesi algoritması ile tahmin edilmiştir. Algoritmalar arasındaki öngörü performans karşılaştırması da AUC kriterine göre yapılmıştır. Bu kritere göre rassal orman algoritması öngörü başarısı en yüksek algoritma olarak belirlenmiştir. Rassal orman algoritmasına dayalı resesyon olasılık grafiği, algoritmanın sıkılaştırılmış makroekonomik ortam sonucu 2019’un son aylarında artan durgunluk olasılığı ile COVID-19 pandemisine bağlı yüksek durgunluk olasılıklarını öngörmede başarılı olduğunu göstermektedir. Rassal orman algoritması ile tahmin sonrası permütasyona dayalı yaklaşım ile değişkenlerin önem sıralaması belirlenmiştir. Buna göre, incelenen dönem için Amerikan resesyonlarında en etkili değişken işsizlik değişkeni, ikinci en etkili değişken federal faiz oranı ve üçüncü en etkili değişken üretim endeksi olarak belirlenmiştir. Ayrıca, analiz bulguları resesyon olasılığı ile açıklayıcı değişkenler arasında doğrusal olmayan ilişkilerin olduğunu göstermektedir. Değişkenlerin belirli eşik değerlerinin üzerinde ve altında resesyon olasılığı üzerindeki etkilerinin farklılaştığı gözlemlenmiştir.*

***Anahtar Kelimeler****: Makine Öğrenmesi, Rassal Orman Algoritması, Ekonomik Durgunluk*

**RECESSION FORECASTING WITH MACHINE LEARNING ALGORTIHMS**

**ABSTRACT**

*The aim of the study is to predict economic recessions in the US economy with various machine learning algorithms. The low-frequency occurrence of economic recessions causes the imbalanced class problem to emerge. To overcome this problem, three different sampling methods (under sampling, over/synthetic sampling, normal sampling) were used and an appropriate sampling method was determined to eliminate the imbalanced class problem. According to the area under the curve (AUC) criterion, the appropriate sampling method was determined as under sampling. In the second stage of the analysis, monthly recession, and financial and macroeconomic time series data (unemployment rate, industrial production index, S&P 500 index, federal funds rate, slope of the yield curve, real money) were obtained from the database of the US Federal Reserve for the period January 1982-June 2021 as training data to determine the algorithm with high predictive performance.*

*Changes in industrial production index and unemployment rate were used to represent changes in the real sector while changes in real money stock, federal funds rate and the S&P 500 index were used to capture the expected and unexpected effects of monetary policy. These were determined as variables that may have an impact on the US recession. Based on these data, the probability of recession in a given month was estimated by the machine learning algorithm with the highest predictive success. The prediction performance comparison among the algorithms was also made according to the AUC criterion. According to this criterion, the random forest algorithm was determined as the algorithm with the highest predictive success. The recession probability graph based on the random forest algorithm shows that the algorithm is successful in predicting the increased recession probability in the last months of 2019 as a result of the tightened macroeconomic environment, and high recession probabilities due to the COVID-19 pandemic. The order of importance of the variables was determined with the permutation-based approach after the estimation with the random forest algorithm. Accordingly, the most influential variable was unemployment, the second most influential variable was the federal interest rate, and the third most influential variable was the industrial production index for the analyzed period. Furthermore, the findings show that there are nonlinear relationships between the probability of recession and explanatory variables. It has been observed that the effects of the variables on the probability of recession vary above and below certain threshold values.*

***Keywords****: Machine Learning, Random Forest Algorithm, Economic Recession*