T.C.
SAKARYA ÜNİVERSİTESİ
FEN BİLİMLERİ ENSTİTÜSÜ

BANKACILIK ALANINDA MAKİNE ÖĞRENME ALGORİTMALARI İLE İŞ YÜKÜ TAHMİNLEMESİ

YÜKSEK LİSANS TEZİ

Yunus Emre BALCI

Endüstri Mühendisliği Anabilim Dalı

Endüstri Mühendisliği Bilim Dalı

TEMMUZ 2024

T.C.
SAKARYA ÜNİVERSİTESİ
FEN BİLİMLERİ ENSTİTÜSÜ

BANKACILIK ALANINDA MAKİNE ÖĞRENME ALGORİTMALARI İLE İŞ YÜKÜ TAHMİNLEMESİ

YÜKSEK LİSANS TEZİ

Yunus Emre BALCI

Endüstri Mühendisliği Anabilim Dalı

Endüstri Mühendisliği Bilim Dalı

Tez Danışmanı: Alparslan Serhat DEMİR

TEMMUZ 2024

Yunus Emre BALCI tarafından hazırlanan "Bankacılık Alanında Makine Öğrenme Algoritmaları İle İş Yükü Tahminlemesi" adlı tez çalışması 08.07.2024 tarihinde aşağıdaki jüri tarafından oy birliği ile Sakarya Üniversitesi Fen Bilimleri Enstitüsü Endüstri Mühendisliği Anabilim Dalı **Endüstri Mühendisliği** Bilim Dalı'nda Yüksek Lisans tezi olarak kabul edilmiştir.

Tez Jürisi

Jüri Başkanı :	**Doç. Dr. Alparslan Serhat DEMİR** Sakarya Üniversitesi	.....
Jüri Üyesi :	**Doç.Dr. Berrin DENİZHAN** Sakarya Üniversitesi	.....
Jüri Üyesi :	**Dr. Öğr. Üyesi Celal ÖZKALE** Kocaeli Üniversitesi	.....

ETİK İLKE VE KURALLARA UYGUNLUK BEYANNAMESİ

Sakarya Üniversitesi Fen Bilimleri Enstitüsü Lisansüstü Eğitim-Öğretim Yönetmeliğine ve Yükseköğretim Kurumları Bilimsel Araştırma ve Yayın Etiği Yönergesine uygun olarak hazırlamış olduğum "Bankacılık Alanında Makine Öğrenme Algoritmaları İle İş Yükü Tahminlemesi" başlıklı tezin bana ait, özgün bir çalışma olduğunu; çalışmamın tüm aşamalarında yukarıda belirtilen yönetmelik ve yönergeye uygun davrandığımı, tezin içerdiği yenilik ve sonuçları başka bir yerden almadığımı, tezde kullandığım eserleri usulüne göre kaynak olarak gösterdiğimi, bu tezi başka bir bilim kuruluna akademik amaç ve unvan almak amacıyla vermediğimi ve 20.04.2016 tarihli Resmi Gazete'de yayımlanan Lisansüstü Eğitim ve Öğretim Yönetmeliğinin 9/2 ve 22/2 maddeleri gereğince Sakarya Üniversitesi'nin abonesi olduğu intihal yazılım programı kullanılarak Enstitü tarafından belirlenmiş ölçütlere uygun rapor alındığını, çalışmamla ilgili yaptığım bu beyana aykırı bir durumun ortaya çıkması halinde doğabilecek her türlü hukuki sorumluluğu kabul ettiğimi beyan ederim.

(……/……/20…..).

(imza)

Öğrencinin Adı Soyadı

Eşime ve aileme …

TEŞEKKÜR

Çalışmamın her aşamasında bana destek olan, bilgi ve deneyimleri ile yol gösteren danışman hocam Doç. Dr. Alparslan Serhat Demir'e, yoğun çalışmalarım sırasında sabır gösterdiği ve beni desteklediği için eşim Mehtap'a, yüksek lisans sürecimde destek gösteren yöneticim Şule Kutlu'ya ve çalışmam sırasında küçük veya büyük desteklerini esirgemeyen herkese teşekkür ederim.

Yunus Emre Balcı

İÇİNDEKİLER

Sayfa

KISALTMALAR

Adam :Adaptive Moment Estimation (Adaptif Moment Tahmini)

A.Ş. : Anonim şirket

ANN : Artificial neural networks (Yapay sinir ağları)

AR : Autocorrelation (Otoregresyon)

ARIMA : Autoregressive Integrated Moving Average (Otokorelatif entegre hareketli ortalama)

ARMA : Autoregressive Moving Average(Otokorelatif hareketli ortalama ve otokorelatif)

BIST : Borsa İstanbul

CNN : Convolutional neural networks (Evrişimli sinir ağları)

DFNN : Deeper variant of a feed-forward neural network (İleri beslemeli sinir ağlarının daha derin varyantı)

EKF : Extended kalman filter (Genişletilmiş Kalman Filtresi)

EYT : Emeklilikte yaşa takılanlar

FNN : Feed-forward neural network (İleri beslemeli sinir ağı)

GS : Grid search (Izgara arama)

I : Entegrasyon

KAP : Kamuyu aydınlatma platformu

KNN : K-En yakın komşu

LIGHTGBM : Light gradient boosting machine (Hafif gradyan artırma makinesi)

LSTM : Long short-term memory (Uzun kısa vadeli bellek)

MA : Moving Average (Hareketli ortalama)

MAE : Mean absolute error (Ortalama Mutlak Hata)

MAPE : Mean absolute percentage error (Ortalama Mutlak Yüzde Hata)

MLP : Multi-layer perceptrons (Çok katmanlı algılayıcılar)

MSE : Mean squared error (Ortalama Kare Hata)

RNN : Recurrent neural network (Yinelemeli sinir ağı)

RMSprop : Root mean square propagation (Kök ortalama kare yayılma)

RS : Random search (Rastgele arama)

SARIMA : Seasonal Autoregressive Integrated Moving Average (Mevsimsel otoregresif entegre hareketli ortalama)

SGD : Stochastic Gradient Descent (Stokastik gradyan inişi)

seq2seq : sequence-to-sequence (Dizi-dizilim)

SVM : Support vector machines (Destek vektör makineleri)

SVR : Destek vektör regresyonu (Destek vektör regresyonu)

TCN : Temporal convolutional network (Zamansal evrişimli ağ)

UKF : Unscented kalman filter (Kokusuz Kalman Filtresi)

XGBoost : Extreme gradient boosting (Aşırı gradyan artırma)

YSA : Yapay sinir ağı (Artificial neural network)

SİMGELER

q	: Derece
α	: Düzleştirme Katsayısı (Smoothing Coefficient)
φ	: Hata Terimlerinin Katsayıları (Coefficients of Error Terms)
ε	: Hata Terimleri (Error Terms)
p	: Otoregresif Katsayısı (Autoregressive Coefficient)
m	: Mevsimsel Periyodu (Seasonal Period)
k	: Komşuluk Sayısı
n	: Veri Sayısı

TABLO LİSTESİ

Sayfa

ŞEKİL LİSTESİ

Sayfa

BANKACILIK ALANINDA MAKİNE ÖĞRENME ALGORİTMALARI İLE İŞ YÜKÜ TAHMİNLEMESİ

ÖZET

Dijitalleşen dünya, özellikle finans sektöründe köklü değişikliklere neden olmaktadır. Son yıllarda dijital bankaların sayısındaki artış, geleneksel şubeli bankalar için hem bir tehdit hem de dijitalleşme yolunda bir uyarıcı olarak karşımıza çıkmaktadır. Bu yeni dönemde, şubeli bankaların personel giderleri gibi büyük maliyet kalemlerini yönetebilmesi, dijital bankalar ile rekabette avantaj sağlamaları için büyük önem taşımaktadır. Dijital bankalar, düşük operasyonel maliyetleri ve hızlı hizmet sunumları ile müşteri çekme konusunda avantajlıdır. Bu nedenle, şubeli bankaların sektördeki paylarını koruyabilmeleri ve sürdürülebilir büyüme sağlayabilmeleri adına dijitalleşmeyi benimsemeleri ve kadro planlamalarını etkin bir şekilde gerçekleştirmeleri gerekmektedir. Kadro sayısının optimizasyonu, hem çalışan memnuniyetini hem de müşteri deneyimini doğrudan etkileyebilecek kritik bir faktördür. Kadro sayısının azaltılması, çalışanların iş yükünü artırarak memnuniyetlerini azaltabilir ve müşteri bekleme sürelerinin uzamasına yol açabilir. Fazla kadro belirlenmesi ise kullanılmayan kapasiteyle maliyetleri yükseltebilir ve olası yer değişiklikleri sebebiyle çalışanların sosyal yaşamlarını olumsuz etkileyebilir. Bu nedenle, bankaların optimal kadro sayısını belirleyerek hem maliyetleri düşürmesi hem de hizmet kalitesini artırması gerekmektedir. Bu çalışma, finans sektöründe faaliyet gösteren bir banka şubesinin kadro planlaması ve kaynak yönetimi süreçlerinde yapay zekâ tekniklerinin potansiyel avantajlarını keşfetmeyi amaçlamaktadır. Farklı metotlar ile, belirli bir unvan grubu ele alınarak geçmiş performans verileri, müşteri talepleri ve diğer etkenler detaylı bir şekilde analiz edilmiştir. Gelecekteki iş yüklerini doğru bir şekilde öngörmek için bu veriler kullanılmıştır. Çalışma kapsamında, elde edilen 0,99 ve 0,91 yüksek korelasyon değerleri ile Extrem Gradyan Arttırma (eXtreme Gradient Boosting, XGBoost) yönteminin en etkili sonuçlara sahip olduğu görülmüştür. Bu yöntem, bankaların daha etkin kadro planlaması yaparak hem maliyetleri optimize etmelerini hem de müşteri ve çalışan memnuniyetini maksimize etmelerini mümkün kılmaktadır. Sonuç olarak, bu araştırma, yapay zekanın kadro planlamasındaki rolünü derinlemesine inceleyerek, bankacılık sektöründe kaynak yönetimi ve operasyonel verimliliği artırmaya yönelik yenilikçi bir yaklaşım sunmaktadır. Bankaların, dijitalleşen dünyada rekabet edebilmek için teknolojileri kullanarak daha verimli ve etkili stratejiler geliştirmesi gerektiği vurgulanmaktadır. Bu bağlamda, yapay zekâ tabanlı kadro planlama modellerinin, bankaların hem maliyetlerini azaltma hem de hizmet kalitesini artırma konusunda önemli bir araç olduğu ifade edilebilir.

WORKLOAD ESTİMATİON WİTH MACHİNE LEARNİNG ALGORİTHMS İN BANKİNG

SUMMARY

The financial sector has entered a radical transformation process with the rapid advancement of digital technologies. Innovations such as the ability to perform financial transactions over the internet, remote account opening, digital fund management, mobile signature use and instant payment systems offer great convenience to customers. This transformation has led to radical changes in customer behavior and increased competition in the sector. Especially in Turkey, the entry of digital banks into the market and the digitalization investments of existing banks have caused a decrease of 702 in the total number of branches. However, the changes observed in the number of employees are complex; there have been increases in some periods and decreases in some periods. This situation shows that the number of employees is not only under the influence of digital processes, but also affected by external dynamics such as economic factors and changes in interest rates.

The digitalization process has created significant challenges for banks. Digital banks and the digitalization investments of existing banks have required banks with branches to rapidly implement digital transformation strategies in order to increase operational efficiency and maintain customer satisfaction. In this process, personnel employment and management emerge as a significant challenge. Personnel expenses are one of the largest cost items that directly affect the profitability of banks. Insufficient staff can increase the pressure on existing staff and cause processing times to extend. At the same time, excessive staff estimation can increase costs and negatively affect the social life of the staff. This can lead to both employee and customer dissatisfaction. Determining the number of employees correctly will both increase efficiency and maintain customer satisfaction. With this transformation in the financial sector, the use of estimation methods has increased. As traditional methods began to fall short, machine learning (ML) methods have gained importance. The development of computer technologies and the popularization of artificial intelligence have led researchers to turn to more easily applicable and understandable machine learning techniques. Machine learning methods such as support vector machines (SVM) provide effective results in the field of financial estimation. Methods such as XGBoost and long-short-term memory (LSTM) stand out with their ability to make fast and effective estimations in large data sets. While XGBoost can automatically learn the interactions and important features between variables with its tree-based structure, LSTM stands out with its capacity to process long-term dependencies.

Hyperparameter optimization is applied to optimize the performance of machine learning methods. Methods such as Grid Search, Random Search and Bayesian Optimization test various parameter combinations to obtain the best results. While Grid Search performs a complete scan over a predetermined set of values, Random Search makes random selections from the specified parameter ranges. Bayesian Optimization determines which hyperparameters should be tested by learning from

previous experiences. These methods contribute to models obtaining more accurate and reliable results.

In the financial sector, digital transformation and changes in customer expectations create significant challenges for banks. Banks need to consider the digitalization process from a strategic perspective and support it with accurate data analysis and workforce management practices. Human resources management plays a critical role in this process as well as technological developments. Correct staff planning and workforce management are becoming a fundamental pillar for the success of banks. The use of machine learning methods contributes to banks obtaining more accurate and reliable results in their forecasting and workforce management processes. When comparing statistical methods, machine learning and deep learning methods, it is seen that machine learning methods are generally more powerful in capturing complex data structures and relationships. In particular, XGBoost's tree-based structure can automatically learn the interactions and important features between variables. This feature shows that it is useful in time series forecasting. The best performance of XGBoost indicates that the complexity and features of your data set are suitable for XGBoost's strengths. It has been seen that the Grid Search method maximizes model performance in hyperparameter optimization. This has increased the effectiveness of the forecast model.

The developed forecast model enables the bank's teller staff to be managed more accurately and efficiently. In the old system, the staff requirement was determined based on the last 6 months of workload data. However, this method was insufficient to provide optimum planning due to variable trends and factors in the workload. For example, giving too much standard staff to a branch led to idle capacity, and this situation negatively affected the social life of the employee as a result of being transferred to a different city where he was needed. On the other hand, the inadequate determination of the norm staff increased the workload of other employees within the branch, which led to employee dissatisfaction and extended waiting times for customers.Accurately predicting the future workload with the forecast model and using this data as input in the staff determination process will ensure that workloads are distributed more optimally and efficiently. The workload estimates obtained will contribute to the bank's decision-making processes in the correct planning of resources such as staff numbers and business models. In this way, staff planning can be made more efficiently and job descriptions for titles can be updated more healthily.

First, the study was conducted with ARIMA, Prophet, XGBoost, LSTM and LightGBM models, then the results were tried to be optimized using the hyperparameter method for XGBoost, LSTM and LightGBM, and these hyperparameter methods were compared with each other. The XGBoost-GS model shows the best performance among the XGBoost models. The correlation value of this model is at the highest level with 0.91, and the MSE and MAE values are also quite low. The XGBoost-RS and XGBoost-Bayes models are similar to each other in terms of correlation value (0.87), and the MSE and MAE values are slightly higher than the XGBoost-GS model. When the LSTM models are evaluated, it is seen that they generally exhibit low performance. The correlation values of the LSTM-GS and LSTM-RS models are quite low (0.16 and 0.17, respectively) and the MSE and MAE values are high. The LSTM-Bayes model performs better than other LSTM models, and its correlation value is 0.7, but this value is still low compared to the XGBoost and LightGBM models.Among the LightGBM models, the LightGBM-GS model performs best. The correlation value of this model is 0.85, and the MSE and MAE

values are at a reasonable level. The LightGBM-RS and LightGBM-Bayes models perform similarly, with correlation values of 0.81. The MSE and MAE values of these models are slightly higher than the LightGBM-GS model.When evaluated in general, the XGBoost-GS model has the highest correlation value and shows the best performance. The LightGBM-GS model stands out as a good alternative.

When statistical methods, machine learning and deep learning methods are compared, it is seen that machine learning methods are generally quite powerful in capturing complex data structures and relationships. In addition, XGBoost's tree-based structure can automatically learn the interactions and important features between variables, which is seen in the study to be useful in time series prediction. The fact that XGBoost gives the best performance may mean that the complexity and features of your data set are in line with XGBoost's strengths. At the same time, these results show that the GS method maximizes model performance in hyperparameter methods. With this model, the bank's teller staff will be managed more accurately and efficiently. In the old system, the staff is determined based on the workload average by looking at the last 6 months of data. However, this method has been insufficient in providing optimum planning due to the workload being affected by different factors and variable trends in the time series. For example, giving too much standard staff to a branch caused idle capacity, and this situation negatively affected the social life of the employee as a result of being shifted to a different city where he was needed. On the other hand, the inadequate determination of the norm staff has increased the workload of other employees in the branch, led to employee dissatisfaction, and increased customer dissatisfaction due to the extended waiting times of customers. In this context, the accurate prediction of the future workload with the developed forecast model and the use of this data as input in the staff determination process will ensure that workloads are distributed more optimally and efficiently. The workload estimates obtained will contribute to the bank's decision-making processes in the correct planning of resources such as the number of staff and business models. In this way, staff planning can be done more efficiently and job descriptions for titles can be updated more healthily.

Since the workforce is calculated in man-day (ft) type when making staff plans, fractions are rounded up according to whether they are above or below 0.5. Accordingly, in the staff planning for 2023, an average workload of 12.47 ft was formed in the past 6-month workload data, and since the fractional value of 0.47 is below 0.5, a staff plan of 12 people was provided according to the old calculation method. The workload average obtained with the estimation model was calculated as 13.48 for 2023, and a norm staff of 13 people will be determined for this workload. Since the absolute deviations on the actual workload values for these two staff numbers were compared, a 23.6% decrease was observed in the estimation model. At the same time, when the cost assessment is made in the study, it is assumed that the workload cost of 1 person will be one third compared to situations such as customer dissatisfaction, employee dissatisfaction and opportunity costs in unrealized transactions that will occur in case of a shortage of staff. In this case, if we were to make a cost calculation based on the staff value determined in the calculation, it would be correct to state it as follows; For example; If the determined staff value is 12 people and the workload for a sample day in 2023 is 11 ft, in this case, the cost for that day is equal to 1 person's idle cash flow.In another scenario, let's consider the workload of one day as 13.5 ft. Since the given staff is again 12, there will be an excess workload of 1.5 ft. In order to convert the cost here to man-day (ft) based on the above explanations, we will need to multiply it by 3 and the cost of 1.5 ft excess workload

will be 4.5 ft.Based on this, in a cost calculation made for 2023, when the staff number is 12 in the old model, the cost incurred is 1537.95 ft, while in the new estimation model, when the staff number is 13 people, the cost incurred will be 1030.86 ft. The difference of 507.10 ft between the old and new models is equivalent to the income of 2 personnel based on 245 working days of 1 person. When calculating the total income for a year, one personnel cost should be deducted since one additional personnel is given in the new estimation model. As a result, the annual cost can be calculated as 1 personnel.

This situation promises the model to determine more optimum staffing rather than lower staffing. In future studies, the problem can be remodeled by adding different variables. When adding these variables, ensemble approaches can be used to increase the accuracy of the estimates. For example, using variables such as customer behavior, economic indicators, seasonal effects and local events in the model can increase the accuracy of workload estimates. It is thought that this study makes significant contributions to the operational efficiency of the bank and customer satisfaction. In addition, similar models can be developed for other sectors and used in workforce management. Such models allow for more effective and efficient use of resources and increase the competitiveness of businesses.

1. GİRİŞ

1.1. Tezin Kapsamı

Finans sektörü dijital teknolojinin ilerlemesiyle birlikte önemli bir dönüşüm sürecinden geçmektedir. Artık müşterilerin finansal işlemlerini gerçekleştirmek için fiziksel banka şubelerine gitmelerine gerek kalmadan uzaktan hesap açılışından dijital fon yönetimine, mobil imza kullanımından anında ödeme sistemlerine kadar birçok işlemi internet üzerinden yapabilmeleri mümkün hale gelmiştir. Bu dönüşüm, müşteri davranışlarında köklü değişikliklere yol açarken, finans sektöründe de yeni oyuncuların piyasaya girişini hızlandırmış ve rekabeti artırmıştır. 2019-2024 yılları arasında yeni dijital bankaların sahneye çıkışı ve mevcut bankaların dijitalleşme yatırımları 10 yeni bankanın piyasaya girmesiyle sonuçlanmış ve bu süreçte toplam şube sayısında 702 adetlik bir azalma yaşanmıştır. Ancak şube sayısındaki bu düşüş, çalışan sayısında karışık bir tablo ortaya koymuş; bazı dönemlerde artışlar bazı dönemlerde ise azalışlar gözlemlenmiştir. Bu durum, çalışan sayısının yalnızca dijital süreçlerin etkisi altında olmadığını aynı zamanda ekonomik faktörler, faiz oranlarındaki değişiklikler gibi çeşitli dış dinamiklerden de etkilendiğini göstermektedir. Dijitalleşme ve müşteri beklentilerindeki değişim, şubeli bankalar için zorlu bir adaptasyon sürecini beraberinde getirmiştir. Dijital bankalara kıyasla daha ağır bir yük altında olan bu bankalar hem operasyonel verimliliği artırmak hem de müşteri memnuniyetini korumak için dijital dönüşüm stratejilerini hızla hayata geçirmek zorundadır. Bu süreçte, personel istihdamı ve yönetimi önemli bir meydan okuma olarak öne çıkmaktadır. Personel giderleri, bankaların karlılığını doğrudan etkileyen en büyük maliyet kalemlerinden biri olup bu alanda yapılan her türlü iyileştirme doğrudan finansal sonuçlara yansımaktadır. Banka şubelerinde farklı pozisyonlar ve unvanlar altında çalışan personelin her biri çok çeşitli görev ve sorumluluklar üstlenmektedir. Bu çeşitlilik, iş yüklerinin değişken faktörlere göre farklılık gösterdiğini ve dolayısıyla kadro planlamasının büyük bir dikkatle yapılması gerektiğini ortaya koymaktadır. Kadro sayısının eksik olması mevcut personelin üzerindeki baskıyı artırarak işlem sürelerinin uzamasına ve sonuç olarak hem müşteri hem de çalışan memnuniyetsizliğine yol açabilir. Diğer yandan kadro sayısının fazla

tahmin edilmesi durumunda ise kullanılmayan iş gücü maliyetleri artırırken, personelin farklı lokasyonlara aktarılması onların sosyal yaşamını olumsuz etkileyebilir ve bu da çalışan memnuniyeti ve bağlılığını düşürebilir. Bu bağlamda, finans sektöründe dijital dönüşümün getirdiği fırsatlar kadar zorluklar da bulunmaktadır. Bankaların rekabetçi bir piyasada ayakta kalabilmek ve müşteri beklentilerini karşılayabilmek için dijitalleşme sürecini stratejik bir perspektiften ele alması, doğru veri analizi ve iş gücü yönetimi uygulamaları ile desteklenmesi gerekmektedir. Bu süreçte teknolojik gelişmeler kadar insan kaynakları yönetimi de kritik bir rol oynamakta, bankaların başarısı için temel sütun haline gelmektedir.

1.2. Tezin Amacı

Bu çalışma, dijitalleşen finans sektöründe şubeli bankaların karşılaştığı personel maliyeti yönetimi zorluklarına çözüm bulmak ve rekabetçi avantaj elde etmelerine yardımcı olmaktır. Bu bağlamda, yapay zeka teknolojilerinin, kadro planlaması ve kaynak yönetimi süreçlerindeki potansiyel faydalarını araştırmak hedeflenmektedir. Çalışma, geçmiş performans verileri ve müşteri talepleri gibi çeşitli faktörleri analiz ederek, gelecekteki iş yüklerini doğru bir şekilde öngörebilme kapasitesini ortaya koymayı amaçlamaktadır. Bu tez, bankaların kaynak yönetiminde yapay zekâ kullanımının etkinliğini detaylı bir şekilde inceleyerek, operasyonel verimliliği artırma yolunda yenilikçi bir yaklaşım sunmayı hedeflemektedir.

2. TEMEL KAVRAMLAR VE YÖNTEMLER

2.1. Temel Kavramlar ve İstatistiksel Veriler

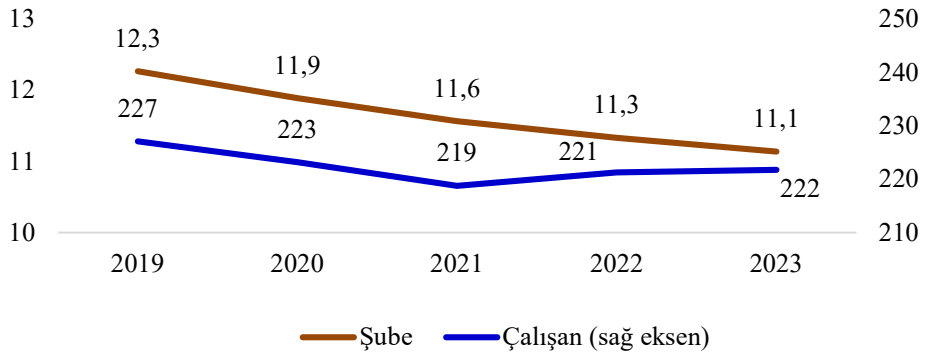
Bankacılık alanında faaliyet gösteren kurumların, çalışma kapsamında bazı istatistiksek verileri, iş yükü hesaplama tekniği ve unvan görev tanımları ile ilgili bilgi verilmiştir.

2.1.1. Banka sayısı

Bankacılık sisteminde 58 banka faaliyet göstermektedir. Mevduat bankaları 35, kalkınma ve yatırım bankaları 17, katılım bankaları 6 tanedir. Misyon Yatırım Bankası A.Ş. 14 Haziran 2023 tarihinde faaliyete geçmiştir. Q Yatırım Bankası A.Ş. ve Tera Yatırım Bankası A.Ş. faaliyet izinlerini almışlardır. Türk Ticaret Bankası A.Ş. özel sermayeli mevduat bankaları grubuna geçmiştir (Türkiye Bankalar Birliği, 2023).

2.1.2. Çalışan sayısı

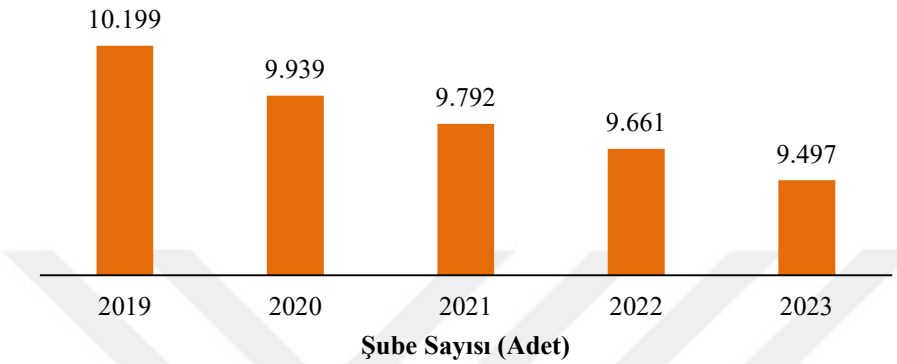
Çalışan sayısı, Haziran 2023 itibariyle mevduat bankaları ile kalkınma ve yatırım bankalarında 188.850 kişi olmuştur. Çalışan sayısı bir önceki çeyreğe göre 2.359 kişi azalırken, bir önceki yılın aynı dönemine göre ise 3.735 kişi artmıştır. Haziran ayında çalışan sayısındaki değişiminde EYT düzenlemesi etkili olmuştur. Çalışan sayısı geçen yılın aynı dönemine göre mevduat bankalarında 2.768 kişi, kalkınma ve yatırım bankalarında ise 967 kişi artmıştır (Türkiye Bankalar Birliği, 2023).



Şekil 2.1. 100.000 kişiye düşen çalışan ve şube sayısı

2.1.3. Şube sayısı

Haziran 2023 yılında mevduat bankaları ile kalkınma ve yatırım bankalarında şube sayısı 9.639'dir. Şube sayısı, bir önceki yıl aynı çeyreğe göre 114 adet azalmış, 2022 yılı sonuna göre ise 22 adet artmıştır. Şube Sayısı (adet) Haziran 2023 itibariyle mevduat bankalarında banka başına ortalama 274 adet olmuştur. (Türkiye Bankalar Birliği, 2023).



Şekil 2.2. Yıllara göre şube sayıları

Şube dışı alternatif dağıtım kanallarının gelişmesi, mobil ve dijital bankacılık hizmetlerine olan talebin artması ve bazı hizmetlerin destek hizmeti kuruluşlarından temin edilmesi; şube ve çalışan sayısını etkilemektedir.

2.1.4. Şube türleleri ve organizasyon yapısı

Bankalar, geniş ve çeşitlendirilmiş organizasyon yapılarıyla faaliyet göstermektedir. Bu yapılar genellikle merkeziyetçi veya dağıtık olabilir. Merkezi yapıda, kararlar genel merkez tarafından alınırken, dağıtık yapıda bölgesel şubeler daha fazla yetkiye sahip olabilmektedir. Türk bankacılık sektörü, hem yerel hem de uluslararası piyasalarda rekabet edebilmek için çeşitli şube türlerine sahiptir. Bu şube türlerinden bazıları aşağıda paylaşılmıştır.

Ticari Şubeler: Kurumsal bankacılık hizmetleri sunan şubelerdir. Bu şubeler, işletmelerin günlük bankacılık işlemleri, kredi talepleri ve diğer finansal danışmanlık hizmetlerini sağlar.

Bireysel Şubeler: Genelde bireysel müşterilere hizmet veren ve perakende bankacılık ürünleri sunan şubelerdir. Bu tür şubeler, kredi kartları, bireysel krediler ve tasarruf hesapları gibi ürünlerle hizmet verir.

Özel Bankacılık Şubeleri: Yüksek net değere sahip bireyler için özelleştirilmiş hizmetler sunan şubelerdir. Bu şubeler, müşteri mahremiyetine ve kişiselleştirilmiş finansal çözümlere büyük önem verir.

KOBİ Şubeleri: Küçük ve orta ölçekli işletmelerle (KOBİ) ve düşük gelirli bireylerle çalışan, onlara uygun finansal ürünler sunan şubelerdir. Bu şubeler genellikle mikro krediler ve KOBİ kredileri gibi hizmetler üzerine yoğunlaşmıştır.

Dijital Şubeler: Fiziksel bir mekan olmadan, tamamen çevrimiçi hizmet veren şubelerdir. Bu yeni şube türü, dijital bankacılık hizmetlerini kullanarak müşterilere daha hızlı ve daha uygun maliyetli hizmet sunmayı amaçlar.

Türk bankalarının bu çeşitli şube türleri, farklı müşteri segmentlerine özelleştirilmiş hizmetler sunarak, sektördeki rekabetçiliklerini artırmayı hedeflemektedir. Organizasyon yapıları ise, genel stratejik hedeflere göre şekillenmekte ve bankacılık düzenlemeleriyle uyum içinde evrilmektedir.

2.1.5. Unvan çeşitleri ve görev tanımları

Türk bankalarındaki şube unvan yapıları ve görev tanımları, bankanın genel işleyişini destekleyen ve müşteri hizmetlerinin verimli bir şekilde yürütülmesini sağlayan bir dizi rol ve sorumluluk içerir. Şube içerisinde yer alan bazı temel pozisyonlar ve bunların görev tanımları şu şekildedir:

Şube Müdürü: Şubenin en üst yöneticisi olan şube müdürü, şubenin genel performansından sorumludur. Tüm şube operasyonlarını yönetir, çalışanları denetler ve müşteri ilişkilerini yönlendirir. Ayrıca, şubenin finansal hedeflerine ulaşması için stratejik planlamalar yapar.

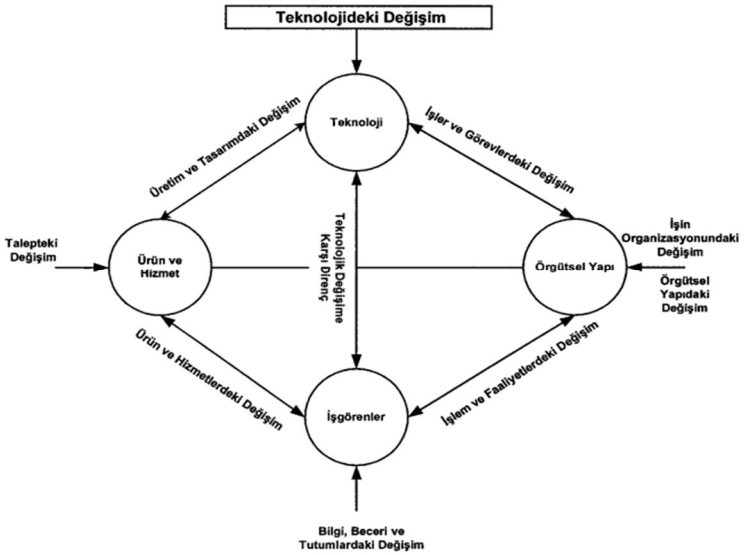
Gişe Çalışanı: Gişe Çalışanları, müşterilerin günlük bankacılık işlemlerini yürütmekten sorumludur. Bu pozisyonda çalışanlar, hesap açma, para transferi ve ödeme işlemleri gibi konularda müşterilere yardımcı olur ve ürünler hakkında bilgi verir.

Operasyon Görevlisi: Bankacılık işlemlerinin arka planında yer alan operasyon görevlileri, işlemlerin doğru ve zamanında gerçekleşmesini sağlar. Para transferleri, hesap işlemleri ve diğer tüm operasyonel süreçler bu pozisyondaki kişiler tarafından yönetilir.

Kobi Satış, Bireysel Satış, Ticari Satış: Bu pozisyondaki çalışanlar, bankacılık ürün ve hizmetlerinin pazarlanması ve satışı konusunda görev alır. Müşteri ihtiyaçlarını belirleyerek, uygun bankacılık ürünlerini tanıtarak ve satışını yaparak şubenin gelir hedeflerine ulaşmasına katkı sağlar.

2.1.6. Teknolojik değişimlerin etkileri

İşletmenin teknolojisinin sık sık adaptasyon faaliyeti gösteren içsel bir değişken olduğu vurgulanmaktadır. Herhangi bir pazar talebi değişikliği, işletmenin ürünlerini etkileyerek onları yeniden tasarlamaya zorlar ve rekabetçi kalmak için yeni teknolojileri kullanmaya zorlar. Şekil 2.3.'teki çerçeve, teknoloji, örgütsel yapı, çalışanlar ve ürünler arasındaki ilişkileri açıklar ve bu değişkenler arasındaki etkileşimi gösterir. Çalışanlar, değişime uygun yetenek ve davranışları geliştiremeyeceklerine inandıklarında direnç gösterebilirler. İşletmenin teknolojisi, faaliyetleri ve yönetim yapılarını etkileyerek değişim üzerinde önemli bir etkiye sahiptir. Teknolojik değişim, tasarım ve üretim faaliyetlerini etkileyerek değişimin gerekli olduğunu ortaya çıkarır ve görevleri yeniden tasarlanmalıdır. Sonuç olarak, işletme başarılı bir performans için teknoloji, yapı ve çalışanlar arasında en iyi uyumu sağlamalıdır.



Şekil 2.3. Teknolojik değişim ve karşılıklı ilişkilerin kavramsal çerçevesi (Ghani, 2000)

2.1.7. İş yükü hesaplamaları ve önemi

İş yükü hesaplamaları, iş süreçlerinin etkinliğini, çalışanların performansını ve iş yoğunluklarını ölçmek amacı ile kullanılan tekniktir. Çalışmada ilk adım,

hesaplamanın yapılacağı iş süreçlerindeki tüm görevleri ve faaliyetleri tanımlamaktır. Bu, hangi işlerin yapılması gerektiğini ve bu işlerin kapsamını belirler. Daha sonra her bir görev için gereken zaman belirlenir. Bu süreler, iş etüdü, geçmiş tecrübeler ve işlem arka planında oluşan veriler yardımı ile belirlenebilmektedir. Çalışanlara veya ekiplere görevler atanırken, mevcut iş yükleri ve performans kapasiteleri dikkate alınır. İş yükünün verimli bir şekilde dağıtılması önem arz eder. Son olarak gerçekleşen işlem adetleri ve işlem süreleri üzerinden kişi başına düşen günlük, aylık ve dönemlik iş yükleri hesaplanır. İş yükü hesaplamaları dinamiktir ve sürekli olarak güncellenmesi gerekebilir. Proje ilerledikçe görevler, süreler ve kaynak ihtiyaçları değişebilir. Bu nedenle sürekli izleme ve gerekirse ayarlamalar yapmak önemlidir.

İş yükleri doğru bir şekilde tahmin edilmesi, müşteri ve çalışan memnuniyeti için büyük önem taşır. Müşteriler açısından iş yükünün doğru yönetilmesi bekleme sürelerini önemli ölçüde azaltabilir. Bu da işlemlerin hızlı ve etkin bir şekilde tamamlanmasını sağlayarak müşteri memnuniyetini artırır. Ayrıca personelin yeterli olması ve işlerini etkin bir şekilde yapabilmesi, müşterilere sunulan hizmet kalitesini doğrudan etkiler. Müşteriler, işlemlerinin sorunsuz bir şekilde yönetildiğini gördüklerinde bankaya olan güvenleri artar, bu da uzun vadeli marka sadakati ve bağlılığını pekiştirir. Çalışanlar açısından ise, iş yükünün doğru bir şekilde dengelenmesi iş stresini azaltır. Çalışanların üzerindeki baskı makul düzeylerde olduğunda, hatalar azalır ve genel performans artar. Ayrıca iş yükünün adil ve dengeli dağıtılması çalışanların iş tatminini ve motivasyonunu yükseltir. Bu da genel iş verimliliğini olumlu yönde etkiler. Dengeli bir iş yükü ayrıca çalışan devir hızını düşürerek işe alım ve eğitim maliyetlerinde tasarruf sağlar.

Özetle, bir banka iş yükünü doğru bir şekilde tahmin edip yönettiğinde, hem müşteri memnuniyetini hem de çalışanların iş tatmini ve performansını artırmış olur. Bu, bankanın operasyonel verimliliğini ve pazar konumunu güçlendirirken, uzun vadeli başarısını da destekler.

2.2. Zaman Serisi Kavramı ve Tahmin Yöntemleri

Zaman serisi verileri, sayısal değerlerin dönemsel olarak birbiri ardına sıralanmış gözlemlerini içerir. Gözlemlenen verilerin mutlaka ardışık olması şart değilse de, analizin doğruluğu için verilerin düzenli zaman dilimlerinde incelenmesi kritik önem taşır. Bu düzenlilik, veri dizisinin zaman içindeki gelişimini daha iyi takip etmeyi

sağlar (Seddighi vd., 2000). Zaman serisi verileri, günümüzde pek çok disiplinde yaygın olarak kullanılmaktadır. Politika, ekonomi, psikoloji, sosyoloji gibi sosyal bilimlerden biyomedikal istatistik ve meteoroloji gibi fen bilimlerine kadar geniş bir yelpazede, bu tür veriler analiz edilerek geleceğe dair öngörülerde bulunulmaktadır. Zaman serilerinin analizi, geçmişteki verilerden yola çıkarak sistemli tahminler yapılmasını sağlar, böylece karar verme süreçleri daha bilgiye dayalı ve stratejik hale gelir. Bu da özellikle hızlı değişim gösteren ve belirsizlik içeren alanlarda büyük bir avantaj sağlar.

2.2.1. Geleneksel yöntemler

Geleneksel zaman serisi yöntemleri, geçmiş verileri temel alarak gelecekteki eğilimleri ve desenleri tahmin etmek için hareketli ortalamalar, üssel düzleştirme, ARIMA ve SARIMA gibi teknikler kullanır. Bu yöntemler genellikle belirli bir istatistiksel modele dayanarak veri analizi yaparlar.

2.2.1.1. Hareketli ortalamalar (MA)

MA modeller içerdikleri geçmiş dönem hata terimi sayısına göre isimlendirilmektedir. q derece bir haraketli ortalama sürecinde her X_t gözlem değeri, q periyot geriye giden rasgele hataların ağırlıklı ortalaması olarak üretilmektedir. Bu proses MA(q) ile temsil edilir ve denklem aşağıdaki 2.2'deki gibi ifade edilmektedir (Box ve Jenkins, 1976).

$$Y_t = \emptyset_0 Y_t + \emptyset_1 Y_{t-1} + \emptyset_2 Y_{t-2} + \cdots + \emptyset_q Y_{t-q} \tag{2.1}$$

2.2.1.2. Üstel düzleştirme (Exponential smoothing)

Üstel düzeltme metotları, eşit olmayan ağırlık setini verilere uygular. Üstel düzeltme metodunda, ağırlıklar en yeni gözlemden en uzaktaki gözleme doğru üstel şekilde azaltılır. Üstel düzeltme metotları, basitliği ve doğruluğu bakımından yoğun bir şekilde kullanılmıştır (Makridakis vd., 1982).Üstel düzeltme metodunun bazen Box-Jenkins yöntemi kadar iyi sonuçlar verebileceğini belirtmişlerdir.

$$\hat{y}_{t+1} = \alpha y_t + (1-\alpha)\hat{y}_t \tag{2.2}$$

Burada $\hat{y}_{t+1}$ zaman serisinin $t+1$ anındaki tahmini, y_t gerçek gözlem, $\hat{y}_t$ ise t anındaki tahminidir. α düzleştirme katsayısıdır ve 0 ile 1 arasında değer alır. Diğer

düzeltme metotları: çift üstel düzeltme, yüksek dereceli üstel düzeltme, Winter metodu ve çift hareketli ortalama metotlarıdır.

2.2.1.3. ARIMA (Otokorelatif entegre hareketli ortalama)

Otoregresif Bütünleşik Hareketli Ortalamalar modeli, zaman serileri temelli sayısal bir model olup; tek veya çok değişkenli olabilmektedir. Talep tahmini çalışmalarında çoğunlukla tek değişkenli Otoregresif Bütünleşik Hareketli Ortalamalar yöntemi kullanılmaktadır (Hu, 2002). Modelin temel işlevi otoregresif süreçler içermesidir. Yani, geçmiş verilerden yararlanarak gelecekle ilgili istatistiksel tahminler yapabilmesi, süreçleri birleştirmesi ve hareketli ortalama süreçlerinin hepsini birlikte içermesi olarak bilinmektedir. ARIMA modellerinin genel matematiksel denklemi 2.3'deki gibi verilebilir (Hipel vd., 1977).

$$Y_t = \alpha + \beta_1 Y_{t-1} + \beta_2 Y_{t-2} + \cdots + \beta_p Y_{t-p} + \varphi_1 \varepsilon_{t-1} + \varphi_2 \varepsilon_{t-2} + \cdots + \varphi_q \varepsilon_{q-1} \quad (2.3)$$

Burada, α sabit değeri, β gözlem değerleri için katsayılarını, Y_t t anındaki zaman serisinin değerini, φ hata terimlerinin katsayılarını, ε hata terimlerini, p oto regresif katsayısını yani gecikme değerlerini göstermektedir. q değeri ise hareketli ortalama parametresidir. Dolayısıyla ARIMA (p, d, q) ifadesi ile otoregresif sıralama(p), fark alma işleminin kaç sefer yapılacağı (d) ve hareketli ortalama parametresi(q) bilgileri ortaya konulur.

2.2.1.4. Mevsimsel ARIMA (SARIMA)

Mevsimsel Otoregresif Entegre Hareketli Ortalama modeli, ARIMA'nın genişletilmiş bir formudur ve özellikle mevsimsel etkileri olan zaman serileri verileri için tasarlanmıştır. Bu model, ARIMA'nın sunduğu entegrasyon (I), otoregresyon (AR) ve hareketli ortalamalar (MA) kavramlarını mevsimsellikle birleştirerek daha kapsamlı ve etkili tahminler yapılmasını sağlar. Zaman serisinin zamana bağlı veya korelogram şeklindeki grafiksel incelemeleri farklı tipte mevsimsel durağan-dışılığın bulunmasını ve tanımlanmasını sağlamaktadır (Yafee ve McGee, 2000). Model genellikle *SARIMA*(*p*,*d*,*q*)(*P*,*D*,*Q*)[*m*] şeklinde ifade edilir. Burada: *p*,*d*,*q* sırasıyla non-mevsimsel AR, I ve MA terimlerinin sıralarını, *P*,*D*,*Q* ise mevsimsel AR, I ve MA terimlerinin sıralarını,*m* ise mevsimsel periyodu belirtir.

2.2.1.5. Box-Jenkins metodu

Box-Jenkins yöntemi, zaman serisi analizinde yaygın olarak kullanılan ve özellikle kısa dönem tahminlerde etkili sonuçlar sunan bir modelleme tekniğidir. Bu metod, özünde otoregresif entegre hareketli ortalama (ARIMA) modellerinin belirli bir sırayla ve disiplinli bir yaklaşımla geliştirilmesine dayanır. Box-Jenkins yöntemi, serinin durağan olduğu ve eşit zaman aralıklarıyla düzenli gözlemler içerdiği varsayımına dayanır. Bu, serinin trend veya mevsimsellik gibi zamanla değişen bileşenlerden arındırılmış olması gerektiği anlamına gelir.Yöntemin temel amacı, gözlemlenen veri setini en iyi açıklayacak, ancak en az sayıda parametreyi içeren modeli tespit etmektir. Bu süreç, modelin tanımlanması, tahmin edilmesi ve doğrulanmasını içeren üç aşamalı bir yaklaşımı takip eder. İlk aşama, uygun AR ve MA terimlerinin sıralarını belirlemek için veri setinin dikkatli bir şekilde incelenmesini içerir. İkinci aşama, seçilen model yapısına göre parametrelerin tahmin edilmesidir. Son aşama ise modelin yeterliliğinin çeşitli istatistiksel testlerle değerlendirilmesidir. Box-Jenkins metodolojisi, otomatik model seçimi süreçleri sunan modern yazılımlarla birlikte daha erişilebilir hale gelmiş olmasına rağmen, bu sürecin tamamının otomatik olarak gerçekleşmediği önemli bir noktadır. Modelin başlangıç aşamalarında analistin müdahalesi gerekebilir, özellikle model parametrelerinin belirlenmesi ve sonraki doğrulama testlerinin yorumlanmasında analitik beceriler ön plandadır. Diğer zaman serisi tahmin yöntemlerine göre daha karmaşık olan Box-Jenkins, derinlemesine istatistiksel bilgi gerektirir ve analistin serinin özelliklerini ve yapısal bileşenlerini iyi anlamasını şart koşar. Karmaşıklığına rağmen, Box-Jenkins yöntemi, özellikle finansal piyasalar, ekonomik veriler ve endüstriyel uygulamalar gibi alanlarda güçlü tahminler yapma kapasitesi nedeniyle tercih edilir. Yöntemin, durağan olmayan serileri modelleme yeteneği, ARIMA'nın farklılaştırılmasıyla sağlanırken, mevsimsel etkileri içeren seriler için SARIMA modeli gibi genişletilmiş formları kullanılabilir. Bu geniş uygulama yelpazesi, Box-Jenkins yöntemini zaman serisi tahmininde vazgeçilmez kılar (Çağlar, 2007).

2.2.2. Makine öğrenimi tabanlı yöntemler

Öğrenme süreçleri genellikle iki temel aşamadan oluşur. İlk aşama, veri kümesi içindeki bilinmeyen ilişkileri tahmin etmektir. İkinci aşama ise bu tahmin edilen ilişkilere dayanarak yeni çıktılar üretmektir. Bu süreçler, Denetimli Öğrenme ve Denetimsiz Öğrenme olmak üzere iki ana öğrenme metodunu tanımlar. Makine

öğrenimi tabanlı yöntemler, zaman serisi verilerini modellemek ve tahmin etmek için çeşitli algoritmalar kullanır. Örneğin, XGBoost, Rastgele Ormanlar ve Destek Vektör Makineleri (SVM) gibi teknikler, büyük veri setleriyle etkili bir şekilde başa çıkabilir ve genellikle karmaşık tahmin modelleri oluşturabilir. Bu yaklaşımlar, veriler arasındaki ilişkileri keşfetmek ve gelecekteki değerleri doğru bir şekilde tahmin etmek için gelişmiş algoritmalar ve matematiksel modellerden yararlanır.

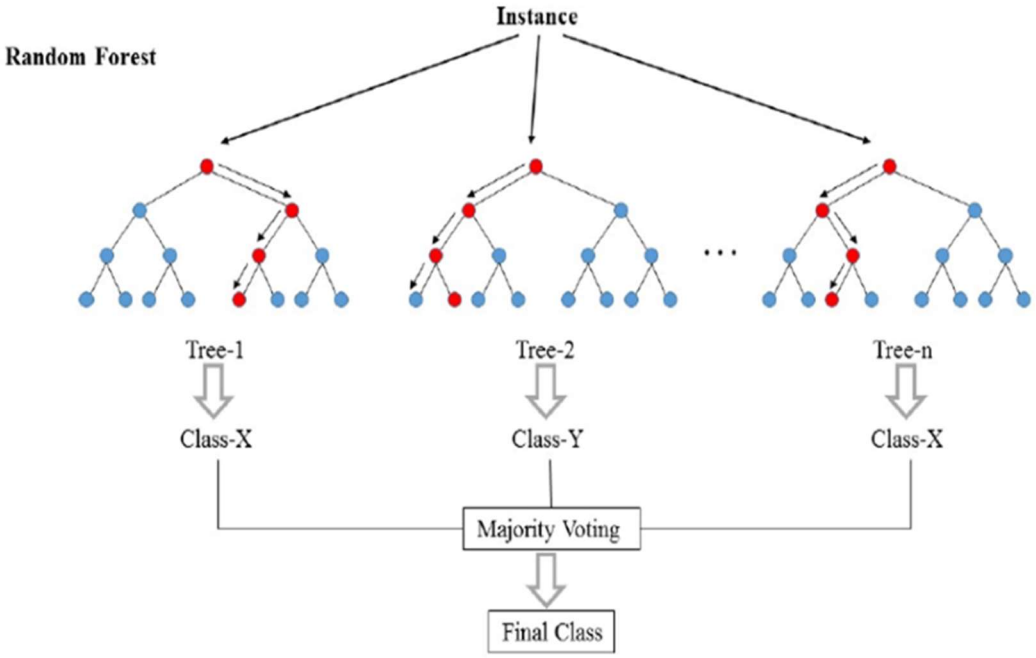
2.2.2.1. Aşırı gradyan arttırma (XGBoost)

XGBoost (eXtreme Gradient Boosting), Gradyan Artırma yöntemlerinin gelişmiş bir uygulaması olarak bilinir. Bu algoritma, hem sınıflandırma hem de regresyon problemlerinde kullanılan karar ağaçları temelinde çalışır. Algoritma, bir dizi zayıf öğreniciyi (genellikle regresyon ağaçları) bir araya getirerek, her bir iterasyonda elde edilen hataları minimize edecek şekilde modeli sürekli olarak güçlendirir. XGBoost, modelin karmaşıklığını kontrol altında tutmak için regülarizasyon teknikleri uygular ve bu da onu diğer Gradyan Artırma yöntemlerinden daha dayanıklı hale getirir.(Chen ve Guestrin, 2016) XGBoost'un başarıya ulaşmasındaki en önemli faktörlerden biri, hız ve ölçeklenebilirlik özellikleridir. Büyük veri kümeleri üzerinde çalışabilme kapasitesi sayesinde, veri bilimciler arasında popüler bir seçim olmuştur. Algoritmanın ayrıca eksik verilerle etkin bir şekilde başa çıkabilme yeteneği de önemli bir avantajdır. XGBoost, eksik veri noktalarını otomatik olarak işleyerek, bu noktalarda en iyi bölünme yollarını bulma konusunda ağacı yönlendirir. Bu özellik, veri ön işleme süreçlerini basitleştirir ve veri kaybını azaltır. XGBoost'un sahip olduğu ölçeklendirilebilirlik ve esneklik, farklı platformlarda ve çok çeşitli uygulama senaryolarında etkin bir şekilde kullanılmasını sağlar. Geliştiriciler, bu algoritmayı birden fazla çekirdek veya farklı makinelerde paralel olarak çalıştırarak, büyük ölçekli veri setleri üzerinde hızlı ve etkili sonuçlar elde etmektediler.

2.2.2.2. Rastgele ormanlar (Random forests)

Rastgele Orman (Random Forest) modeli, denetimli öğrenme algoritmaları içinde yer alan ve birçok karar ağacının oluşturduğu bir topluluk öğrenme yöntemidir. Leo Breiman tarafından 1997 yılında geliştirilen bu yöntem, daha önceki Amit ve Geman'ın çalışmalarına dayanarak şekillendirilmiştir. Rastgele Orman, hem sınıflandırma hem de regresyon problemleri için uygundur ve bu alanda yaygın olarak kullanılır.Rastgele Orman modelinin çalışma prensibi, bir dizi karar ağacını rastgele seçilmiş veri alt kümeleri üzerinde bağımsız olarak eğitmek ve bu ağaçların

tahminlerini birleştirerek nihai bir tahmin elde etmektir. Bu yöntem, her bir karar ağacının veri setinin farklı bir rastgele alt kümesi üzerinde eğitilmesiyle gerçekleştirilir. Bu işlem, modelin genel performansını artırırken aynı zamanda aşırı uyuma (overfitting) karşı dayanıklılığını da artırır. (Breiman, 2001). Rastgele Orman, bu özellikleri sayesinde çeşitli sınıflandırma ve regresyon problemlerinde güçlü ve güvenilir sonuçlar sunar. Modelin işleyişi, Şekil 2.4.'teki gibi görselleştirilerek daha iyi anlaşılabilir. Bu görseller, modelin veri üzerinde nasıl bir yapı kurduğunu ve tahminlerin nasıl elde edildiğini somut bir şekilde gösterir.



Şekil 2.4. Rastgele orman regresyonu işleyişi (Nomidl Machine Learning, 2022).

2.2.2.3. Destek vektör makineleri (Support vector machines, SVM)

İlk defa Vapnik ve Cortes (1995) tarafından önerilen SVM, sıklıkla regresyon ve sınıflama problemlerinde tercih edilen bir makine öğrenmesi yöntemidir. Düşük boyuta sahip veri setlerinin doğrusal olarak ayrıştırılamadığı durumlarda SVM, bu veri setini yüksek boyuta sahip uzaya çıkararak doğrusal ayrıştırılabilir hale getirmektedir (Tay ve Cao, 2001). SVM, regresyon problemlerinde kullanıldığında Destek Vektör Regresyonu (SVR) ismini almaktadır. Bir eğitim setinde xi ve yi sırasıyla girdi ve çıktı değişkenleri olmak üzere $T = \{(x1, y1), \ldots, (xn, yn)\}$ olduğu varsayılsın. SVR, eğitim verisindeki gerçek değer yi 'den en fazla ε kadar uzakta doğrusal bir fonksiyon bulabilmeyi amaçlar (Smola ve Scholkopf, 2004).

2.2.2.4. K-En yakın komşu (K-nearest neighbors, KNN)

K-En Yakın Komşu (KNN) tahmin yöntemi parametrik olmayan sınıflandırma ve regresyon algoritmasıdır. Uygulanabilme kolaylığı, basit bir matematiksel temele sahip olması pek çok farklı alanda tahmin modelleri kullanımını yaygınlaştırmıştır (Kermani vd., 2018). Bir vakanın sonucu kendisine en yakın komşu vakalarının sonucu ile aynı olur fikrini esas alır. Geçmiş gözlemlere dayalı olan eğitim seti vasıtasıyla, veri setinin her elemanının (vakanın) sonucu olan bağımlı değişkenler belirlenir. İleriye dönük tahminler, mevcut vakaların sonuçlarının eğitim veri setindeki en yakın elemanların sonuçlarının ortalamasına eşit olacaktır. Genellikle, en yakın gözlemler, göz önünde bulundurulan veri noktasına en küçük Öklid mesafesine sahip olanlar olarak tanımlanır. Gözlemler arasındaki öklid mesafesi 2 boyutlu çözüm kümesi örneğinde x düzlemindeki doğrusal uzaklık ix 'ye ve y düzlemindeki doğrusal uzaklık yi 'ye bağlı olarak bulunabilir. Dikkate alınacak komşu sayısını gösteren k'nın belirlenmesi için model optimizasyonu gerekir. Komşuluk sayısı k değeri arttıkça hesaplama adımı artsa da eğitim setindeki gürültüye karşı kırılganlığı azalacak ve test verisine uyum artacaktır.

2.2.2.5. Hafif gradyan artırma (LightGBM)

LightGBM (Light Gradient Boosting Machine), Microsoft tarafından geliştirilen ve özellikle hız ve performans açısından optimize edilmiş bir makine öğrenmesi algoritmasıdır. Büyük veri kümeleri ve yüksek boyutlu veriler üzerinde etkili sonuçlar elde etmek için tasarlanan LightGBM, gradient boosting algoritmaları arasında öne çıkmaktadır. Algoritmanın performansını ve doğruluğunu artırırken eğitim süresini önemli ölçüde kısaltan birçok yenilikçi özelliği bulunmaktadır. LightGBM, geleneksel seviye-temelli büyüme stratejisi yerine yaprak-temelli bir strateji kullanarak belirli bir seviyedeki en büyük kaybı azaltan yaprağın genişletilmesini sağlar. Bu yaklaşım, eğitim süresini azaltırken doğruluğu artırır. Ayrıca, veriyi histogramlar şeklinde bölerek, bölme işlemlerini hızlandırır ve bellek kullanımını optimize eder. Gölge testi (out-of-bag error estimation) ile modelin aşırı uyum göstermesini (overfitting) önler ve daha doğru tahminler yapılmasını sağlar. LightGBM, kategorik değişkenleri doğrudan destekler ve bu değişkenlerin en uygun şekilde işlenmesini sağlar (Ju ve diğerleri, 2019).

LightGBM, finansal analiz ve kredi skorlama, pazarlama analitiği, müşteri davranış tahminleri, sağlık hizmetleri veri analitiği ve zaman serisi tahminleri gibi birçok farklı

uygulama alanında kullanılabilir. Performans açısından diğer gradient boosting algoritmalarına kıyasla belirgin avantajlar sunar. Büyük veri kümeleri ve yüksek boyutlu veri setlerinde, LightGBM'nin eğitim süresi ve doğruluğu, diğer yöntemlere kıyasla genellikle daha üstündür. Özellikle, XGBoost ile yapılan karşılaştırmalarda, LightGBM'nin hem hız hem de doğruluk açısından avantajlı olduğu birçok çalışmada gösterilmiştir. LightGBM, modern veri analitiği ve makine öğrenmesi uygulamalarında güçlü ve verimli bir araçtır. Büyük ve karmaşık veri setleri üzerinde hızlı ve doğru sonuçlar elde etmek isteyen araştırmacılar ve veri bilimciler için ideal bir seçimdir. Sağladığı hız, doğruluk ve esneklik, onu günümüzün veri analitiği dünyasında vazgeçilmez bir araç haline getirmektedir.

2.2.3. Derin öğrenme yöntemleri

Derin öğrenme tabanlı zaman serisi yöntemleri, karmaşık desenleri tanımak ve uzun vadeli eğilimleri tahmin etmek için LSTM ve CNN gibi algoritmaları kullanır. Bu yaklaşımlar, çok katmanlı yapıları sayesinde verilerin karmaşıklığını daha iyi ele alabilir.

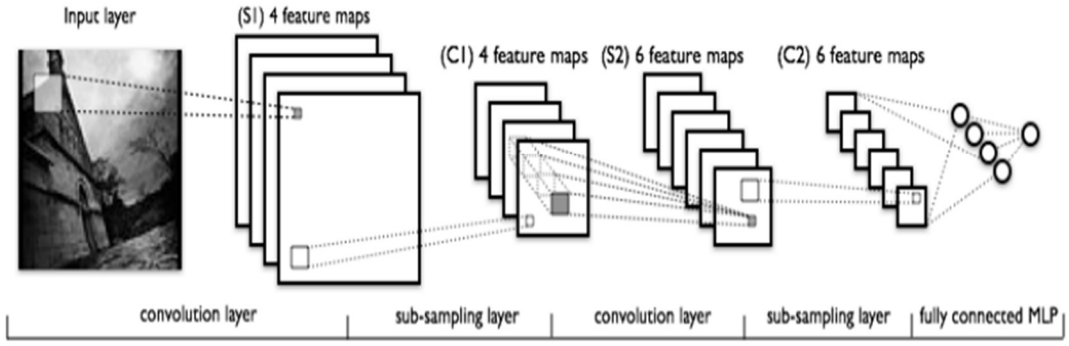
2.2.3.1. Çok katmanlı algılayıcılar (Multi-layer perceptrons, MLP)

Çok Katmanlı Algılayıcılar (Multi-Layer Perceptrons, MLP), yapay sinir ağlarının bir türü olup, çeşitli veri işleme ve tahminleme görevlerinde kullanılan güçlü bir makine öğrenme modelidir. Bu tür ağlar, bir giriş katmanı, bir veya daha fazla gizli katman ve bir çıkış katmanı içerir. Giriş katmanı, modelin veri aldığı yerdir ve bu verinin özelliklerini temsil eder. Gizli katmanlar, giriş verilerini işler ve modelin karmaşık ilişkileri öğrenmesine yardımcı olur. Çıkış katmanı ise modelin nihai tahminlerini veya sonuçlarını verir. MLP'ler, katmanlar arasındaki bağlantılarda ağırlıklar ve bias terimleri kullanır. Bu parametreler modelin öğrenme sürecinde güncellenir ve böylece model, eğitim verisindeki desenleri ve kalıpları öğrenebilir. MLP'lerde yaygın olarak kullanılan aktivasyon fonksiyonları, modelin doğrusal olmayan ilişkileri öğrenmesini sağlar. Geri yayılım, MLP'lerin en önemli bileşenlerinden biridir. Bu algoritma, modelin hata bilgisini kullanarak ağırlık ve bias terimlerini günceller, böylece modelin tahminlerinin doğruluğu artar. Gizli katman sayısı en az bir tane olmak üzere birden fazla da olabilmektedir. Gizli katmandaki her işlem elemanı, doğrusal olmayan bir transfer fonksiyonuna sahiptir ve bu fonksiyonlar aracılığıyla elde edilen sonuçlar bir sonraki işlem elemanlarına girdi sağlamaktadır. MLP'lerde en çok kullanılan transfer

fonksiyonları sigmoid fonksiyonu ve hiperbolik tanjant fonksiyonudur (Bayramoğlu, 2007).

2.2.3.2. Konvolüsyonel sinir ağları (Convolutional neural networks, CNN)

Konvolüsyonel Sinir Ağları (Convolutional Neural Network - CNN) bir çeşit ileri beslemeli yapay sinir agı olup, katmanlar arası bağlantı modeli, görsel korteksin organizasyonuna benzetilmektedir. Genel olarak iki çeşit katmandan oluşur: konvolüsyonel katman (convolutional layer), örnekleme katmanı (sub-sampling layer). CNN yapısal olarak, arka arkaya gelen konvolüsyonel ve örnekleme katmanından oluşmaktadır. Son olarak katmanlar tamamen baglı çok katmanlı algılayıcıya (fully connected MLP) bağlanır. CNN'ler görüntü ve video tanıma, tavsiye sistemleri ve dogal dil işleme alanlarında geniş uygulamalara sahiptirler. Genel CNN yapısı, ¸Sekil 2.5.'te gösterilmektedir. Şekilde de gösterildigi gibi girdi verisi arka arkaya konvolüsyonel ve örnekleme katmanlarından geçer, son olarak MLP'ye baglanır.



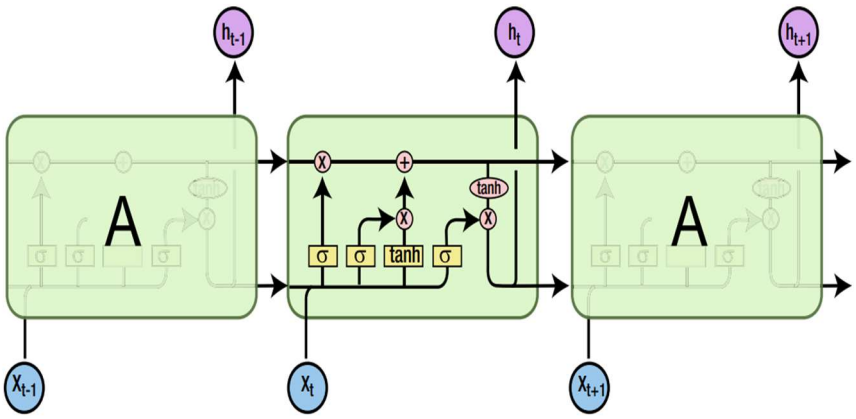
Şekil 2.5. Konvolüsyonel sinir ağları (LeCun vd., 1989).

Literatürde, CNN genelde resim ve video tanıma alanında kullanılmıştır. Ayrıca CNN zaman serisi verilerinin analizinde kullanılan bir yöntemdir.

2.2.3.3. Uzun kısa süreli bellek (Long short-term memory, LSTM)

Uzun-kısa süreli bellek (LSTM), geleneksel RNN'lere göre daha doğru sonuçlar sağlayabilmesiyle bilinir ve uzun menzilli bağımlılıkları işleyebilme yeteneğine sahiptir. Temelde, LSTM birimleri farklı zaman adımlarında çıktıları hatırlayabilen ve bunları açıkça iletebilen bir bellek hücresi kullanır. Bu bellek hücresi, zamansal bağlamları hatırlamak için hücre durumlarını kullanır ve unutma, giriş ve çıkış geçitlerine sahiptir.

RNN'lerdeki öğrenme sürecinde karşılaşılan zorluk, kaybolan gradyanlar sorunudur. Ancak LSTM unutma, giriş ve çıkış geçitlerinden oluşan bir yapı kullanarak bu dezavantajı aşar. Bu geçitler, uzun süreli belleğin düzenlenmesine yardımcı olur. LSTM hücre bloğu, bilgileri unutma veya hafızaya alma, bir sonraki hücreye ne kadar bilgi aktarılacağına karar verme gibi işlevlere sahiptir. Böylece uzun vadeli bağımlılıkları işleyebilme kabiliyetini sağlar. Yapılan çalışmalarda, LSTM'nin üç kapısının uzun vadeli belleğin düzenlenmesini kolaylaştırdığı bulunmuştur. LSTM'nin yapısı unutma, giriş ve çıkış geçitlerinin etkin kullanımıyla dikkat çeker. Şekil 2.6.'da, LSTM yapısının görselleştirilmiş hali bulunmaktadır (Yang ve Chang, 2020).



Şekil 2.6. LSTM modeli etkileşimli dört katman (Beysolow, 2018).

LSTM'de kullanılan unutma kapısı, geçmiş bilgilerin ne kadarının unutulacağını belirler. Bu kapı, sigmoid aktivasyon fonksiyonu kullanarak çalışır. Önceki hücre durumu ve mevcut giriş bilgileriyle birlikte ağırlıklı bir işlem yapar. Sonuç olarak, 0 ile 1 arasında bir değer üretir; 0'a yakın değerler geçmiş bilgileri unutmayı, 1'e yakın değerler ise unutmamayı ifade eder. Giriş kapısı, yeni bilgilerin ne kadarının hafızaya ekleneceğini belirler. İki önemli bileşenden oluşur: aday durum hesaplama ve giriş kapısı seçilen bilgilerin eklenmesi. Aday durum hesaplanırken tanh aktivasyon fonksiyonu kullanılır ve yeni bilgileri içeren bir aday hücre durumu oluşturulur.

Giriş kapısı, sigmoid aktivasyon fonksiyonu kullanarak yeni bilgilerin ne kadarının aday duruma ekleneceğini belirler. Hücre durumu unutma ve giriş kapıları tarafından hesaplanan değerlerin birleştirilmesiyle güncellenir. Unutma kapısının çıktısı önceki hücre durumunu ne kadarının korunacağını belirler. Giriş kapısının çıktısı aday hücre durumunun ne kadarının gerçek hücre durumuna ekleneceğini belirler. Bu iki değer toplanarak güncellenmiş hücre durumu elde edilir. Çıkış kapısı, hücre durumunun ne

kadarının ağın çıktısına aktarılacağını belirler. Hücre durumu üzerinden geçirilerek sigmoid aktivasyon fonksiyonu kullanılır. Sonuç, ağın çıktısı olan gizli durumu oluşturmak için kullanılır. Gizli durum hücre durumunu kullanarak ağın çıktısını üretir. Çıkış kapısının çıktısı hücre durumu üzerinden aktivasyon fonksiyonu kullanılarak çarpılır. Bu, ağın mevcut gizli durumunu oluşturur.

2.2.3.4. Gradyan tabanlı optimizasyon algoritmaları

Gradyan tabanlı optimizasyon algoritmaları, zaman serisi verileriyle çalışan makine öğrenimi modellerinin eğitiminde kritik bir rol oynar. Bu algoritmalar, model parametrelerini güncellemek için kayıp fonksiyonunun model parametrelerine göre türevini hesaplar ve böylece modelin eğitimi sırasında maliyeti minimize etmeye çalışır. Zaman serisi verileri, zaman içindeki trendler, döngüler ve mevsimsellik gibi benzersiz özellikler içerir. Modelin bu karmaşık yapıları yakalaması ve anlaması gerekmektedir.Zaman serisi verileriyle çalışırken, uygun model seçimi hayati önem taşır. Sinir ağları ve lineer modeller gibi yaklaşımlar, zaman serisi verilerinin karmaşık özelliklerini başarılı bir şekilde modelleyebilir. Gradyan tabanlı optimizasyon, modelin eğitim sürecinde parametrelerini sürekli güncelleyerek öğrenmeyi yönlendirir. Öğrenme oranının doğru bir şekilde ayarlanması, modelin etkili bir şekilde öğrenmesini sağlar; çok yüksek bir öğrenme oranı modelin dengesizleşmesine neden olabilirken, düşük bir oran ise öğrenmeyi yavaşlatabilir

Stokastik Gradyan İnişi (SGD), Adam ve RMSprop gibi yaygın olarak kullanılan optimizasyon algoritmaları, modelin farklı zaman serisi özelliklerini ve karmaşıklıklarını öğrenmesi için parametrelerini uygun şekilde güncelleyerek eğitimi etkili bir şekilde yönlendirir. Zaman serisi verileri, öngörülemez ve değişkenliği yüksek olabilir; bu nedenle modelin doğru tahminler yapabilmesi için yeterli veri ve özenle seçilmiş özniteliklere ihtiyaç vardır. Optimizasyon algoritmaları, modelin bu karmaşıklıkların üstesinden gelerek güvenilir ve doğru tahminlerde bulunmasına yardımcı olur.

2.2.4. Prophet ve diğer ileri zaman serisi yöntemleri

Prophet ve diğer ileri zaman serisi yöntemleri, mevsimsellik, tatil etkileri ve eğilimler gibi unsurları dikkate alarak doğru tahminler yapmak için tasarlanmıştır. Holt-Winters ve Kalman Filtreleri gibi teknikler de bu kategoride yer alarak karmaşık veri setlerini modellemek için kullanılır.

2.2.4.1. Prophet

"Prophet modeli, Facebook veri bilimi ekibi üyeleri tarafından duyurulan, zaman serileri tahmininde kullanılan, Python ve R destekli, açık kaynak kodlu bir tahmin uygulamasıdır (Taylor ve Letham, 2017). Bu çalışma Python ile gerçekleştirilmiştir. Facebook Prophet, doğrusal olmayan verilerde yıllık, aylık, günlük tahminler yapabilen ve belirtilen tatil günlerini tahmin hesabına katabilen bir algoritmadır. Prophet'in sezgisel işleyişi, verinin detaylarına boğulmadan etkili tahminler yapılmasın (Taylor ve Letham, 2018). Ayrıca, hem genelleştirilmiş doğrusal modellerden (linear model) hem de eklemeli modellerden (additive model) özellikler taşır (Mata, 2020). Zaman serisini tanımlayan genel formül şu şekildedir:

$$y(t) = g(t) + s(t) + h(t) + \varepsilon(t) \tag{2.4}$$

g (t), zaman serisi değerlerindeki periyodik olmayan değişkenleri modelleme fonksiyonudur. s (t) yıllık, aylık ve haftalık periyodik değişimleri temsil eder. h (t), tatillerin genellikle düzensiz olarak bir ya da daha çok gün için ortaya çıkardığı etkilerini temsil eder. Genellikle normal dağılım olarak modellenen hata terimi ε (t), modelde bulunmayan her türlü kendine has değişiklikleri ifade eder. Prophet ile diğer istatistiksel yöntemler arasındaki temel fark, döngüdeki analist yaklaşımıdır. Bu yaklaşım, analistin veri hakkında alan bilgilerini kullanarak tahmin algoritması uygulamasına olanak tanır, böylece içeriden çalışan istatistiksel yöntemler hakkında bilgi sahibi olmadan tahmin yapılabilir. Bu nedenle, Prophet hem istatistiksel tahmin hem de insan uzmanların kararlarına dayanan tahmin yöntemlerini birleştirmeyi amaçlar (Mata, 2020).

2.2.4.2. Kalman filtreleri

Kalman filtreleri, zaman serisi analizi ve sistem izleme alanlarında yaygın olarak kullanılan güçlü bir araçtır. Son yıllarda yapılan araştırmalar, Kalman filtrelerinin çeşitli alanlarda kullanımını genişletmiş ve yeni yöntemler geliştirmiştir. Özellikle, doğrusal olmayan sistemler ve gürültülü verilerle başa çıkma yetenekleri üzerine yapılan çalışmalar dikkat çekicidir. Doğrusal olmayan sistemler için genişletilmiş Kalman filtreleri (EKF) ve unscented Kalman filtreleri (UKF) gibi yöntemler, literatürde yoğun olarak incelenmektedir (Julier ve Uhlmann, 2004).

Ayrıca, Kalman filtrelerinin makine öğrenimi ve yapay zeka alanlarıyla birleştirilerek kullanımı da artmaktadır. Derin öğrenme teknikleri ile birleştirilmiş Kalman filtreleri, karmaşık sistemlerin izlenmesi ve tahmin edilmesi için yeni olanaklar sunmaktadır (Chiappa ve Calandra, 2019). Finansal piyasalarda Kalman filtrelerinin kullanımı da dikkat çekicidir. Özellikle, portföy yönetimi ve risk analizi gibi alanlarda Kalman filtreleri, volatilite tahmini ve fiyat hareketlerinin modellenmesi için yaygın olarak kullanılmaktadır (Bauwens ve Veredas, 2004).

Sonuç olarak, Kalman filtreleri zaman serisi analizi ve sistem izleme alanlarında önemli bir araç olarak kalmaya devam etmektedir. Yeni gelişmeler ve araştırmalar, Kalman filtrelerinin uygulama alanlarını genişletmekte ve performansını artırmaktadır.

2.2.4.3. Gradyan artışı

Gradyan Artışı (Gradient Boosting), son yıllarda zaman serisi analizi alanında önemli bir popülerlik kazanmış ve birçok uygulama alanında başarıyla kullanılmıştır. Gradyan Artışı, zayıf tahmin edicilerin bir araya gelerek güçlü bir tahmin edici oluşturduğu bir toplu öğrenme tekniğidir. Gradyan Artışı yöntemi, başlangıçta Friedman (2001) tarafından önerilmiş ve ardından birçok türetilmiş algoritma geliştirilmiştir. Bu algoritmalar arasında XGBoost (Chen ve Guestrin, 2016), LightGBM (Ke vd., 2017) ve CatBoost (Dorogush vd., 2018) gibi öne çıkanlar bulunmaktadır.

Bu algoritmalar, zaman serisi verilerinin özelliklerini dikkate alarak tahmin modelleri oluşturabilir. Özellikle, Gradyan Artışı'nın esnek parametre ayarları ve yüksek performansı, finansal piyasalardan sağlık sektörüne kadar birçok alanda başarıyla kullanılmasını sağlamıştır. Gradyan Artışı'nın zaman serisi analizindeki kullanımı, model güçlendirme ve tahmin performansının artırılması açısından önemlidir. Ayrıca, Gradyan Artışı'nın paralel hesaplama yetenekleri ve büyük ölçekli veri kümeleriyle etkili bir şekilde başa çıkabilmesi, büyük veri analizinde de tercih edilen bir yöntem haline gelmiştir. Sonuç olarak, Gradyan Artışı, zaman serisi analizi alanında etkili bir tahmin aracı olarak kabul edilir. Yeni gelişmeler ve araştırmalar, Gradyan Artışı'nın zaman serisi verilerinin analizindeki rolünü daha da genişletebilir.

2.3. Hiper Parametre Yöntemleri

Hiperparametreler, bir makine öğrenimi modelinin performansını ve davranışını kontrol eden ayarlanabilir parametrelerdir. Bu parametreler, modelin yapılandırılmasında kullanılır ve modelin eğitim sürecinde sabitlenmeyen, elle ayarlanması gereken ve deneme yanılma yöntemiyle optimize edilmesi gereken değerlerdir. Hiperparametreler, modelin karmaşıklığını, öğrenme sürecini ve genelleme yeteneğini etkiler. Hiperparametre optimizasyonu, doğru hiperparametre değerlerinin seçilmesini sağlamak için kullanılan bir süreçtir. Bu optimizasyon, genellikle çapraz doğrulama veya otomatik hiperparametre arama gibi tekniklerle gerçekleştirilir.

2.3.1. Grid search

Grid Search, bir makine öğrenimi modelinin hiperparametrelerini optimize etmek için kullanılan bir yöntemdir. Belirli bir hiperparametre aralığında önceden belirlenmiş değerlerin kombinasyonlarını sistemli bir şekilde deneyerek en iyi performansı sağlayan hiperparametre setini bulmaya çalışır.

2.3.2. Random search

Grid Search yönteminden farklı olarak, belirli bir hiperparametre aralığında rastgele seçilen değerlerle deneme yapar. Bu yöntem, tüm olası hiperparametre kombinasyonlarını aramak yerine belirli bir bölgedeki rastgele noktalarda deneme yaparak daha geniş bir hiperparametre uzayını keşfeder. Bu sayede, Grid Search yöntemine kıyasla daha verimli bir arama yapabilir ve daha az hesaplama maliyetiyle daha iyi sonuçlar elde edebilir.

2.3.3. Bayesian optimizasyon algoritması

Bu yöntem, bir modelin performansını ölçerek, en iyi hiperparametre değerlerini belirlemek için bir olasılık modeli kullanır. Önceki denemelerden elde edilen bilgilere dayanarak, bir sonraki hiperparametre değerlerini belirler ve bu değerlerle modeli eğitir. Bu süreç, modelin performansını ölçerek en iyi hiperparametre değerlerini bulana kadar tekrarlanır. Bayesian Optimization, rastgele arama ve grid search gibi yöntemlere kıyasla daha az deneme ile daha iyi sonuçlar elde edebilir, özellikle çok boyutlu ve karmaşık hiperparametre uzaylarında etkilidir.

2.4. Performans Ölçüm Yöntemleri

2.4.1. MAE (Mean absolute error - Ortalama mutlak hata)

Ortalama mutlak hata (MAE), regresyon ve zaman serisi problemlerinde sıklıkla kullanılan bir performans metriğidir ve gerçek ile tahmin edilen değerler arasındaki mutlak farkların ortalamasını hesaplamaktadır. Ortalama mutlak hata, her bir veri noktasının model performansına eşit ağırlık verdiği bir ölçüttür. Bu nedenle, küçük hataların da büyük hatalar kadar önemli olduğu durumlarda kullanılmaktadır. Ortalama mutlak hata değeri sıfıra yaklaştıkça, modelin tahminleri gerçek değerlere daha yakın hale gelir ve modelin performansı artar. Ortalama mutlak hata, kolay yorumlanabilir bir metrik olduğu için özellikle endüstriyel ve ticari uygulamalarda tercih edilir.

$$\frac{1}{n}\sum_{i=1}^{n}|x_i - x| \tag{2.5}$$

2.4.2. MAPE (Mean absolute percentage error - Ortalama mutlak hata yüzdesi)

MAPE (Mean Absolute Percentage Error - Ortalama Mutlak Hata Yüzdesi), bir tahmin modelinin doğruluğunu değerlendirmek için kullanılan bir ölçüdür. MAPE, tahminlerin gerçek değerlerden ne kadar sapma gösterdiğini yüzde olarak ifade eder. MAPE'nin avantajı, tahminlerin yüzde cinsinden hatalarını doğrudan ölçmesidir, bu da farklı ölçeklerdeki verileri karşılaştırmak için kullanışlıdır. Ancak, MAPE'nin sıfıra bölünme hatası oluşabileceği ve bazı durumlarda yanıltıcı sonuçlar verebileceği önemli bir dezavantajı vardır.

$$\frac{1}{n}\sum_{i=1}^{n}\frac{|A_i - F_i|}{A_i} \tag{2.6}$$

2.4.3. MSE (Mean squared error)

Ortalama kare hata, özellikle regresyon modellerinde sıklıkla kullanılan bir performans metriğidir. MSE, gerçek ve tahmin edilen değerler arasındaki farkların karelerinin ortalamasıdır. Ortalama kare hata daha yüksek hataların daha fazla ağırlığına sahip olduğu bir ölçüttür. Bu nedenle, büyük hataların model performansını daha fazla düşüreceği durumlarda kullanılır. Ortalama kare hata değeri sıfıra yaklaştıkça, modelin tahminleri gerçek değerlere daha yakın hale gelir ve modelin

performansı artar. Ancak, ortalama kare hatanın yorumlanması için gerçek değerlerin aralığına ve veri kümesinin özelliklerine göre değerlendirme yapılması gerekmektedir.

$$\frac{1}{N}\sum_{i=1}^{N}(y_i - \hat{y}_i)^2 \tag{2.7}$$

2.4.4. Korelasyon katsayısı

Korelasyon katsayısı, iki değişken arasındaki ilişkinin gücünü ve yönünü ölçen bir istatistiksel ölçüdür. Korelasyon katsayısı, genellikle -1 ile +1 arasında bir değer alır.

+1: Mükemmel pozitif korelasyon, yani iki değişken arasında tam bir doğrusal ilişki olduğunu gösterir. Bir değişken artarken diğeri de artar.

0: İlişki yoktur veya tamamen rastgeledir.

-1: Mükemmel negatif korelasyon, yani iki değişken arasında tam bir ters doğrusal ilişki olduğunu gösterir. Bir değişken artarken diğeri azalır.

$$r = \frac{\sum_{i=1}^{n}(X_i - \bar{X})(Y_i - \bar{Y})}{\sqrt{\sum_{i=1}^{n}(X_i - \bar{X})^2} \times \sqrt{\sum_{i=1}^{n}(Y_i - \bar{Y})^2}} \tag{2.8}$$

Burada:

- *Xi* ve *Yi*, sırasıyla iki değişkenin değerleridir.
- $\bar{X}$ve $\bar{Y}$, sırasıyla *X* ve *Y* değişkenlerinin ortalamalarıdır.
- *n*, veri noktalarının toplam sayısıdır.

3. LİTERATÜR TARAMASI

3.1. İş Gücü ve Diğer Zaman Serisi Tahminlemeleri

Bu alanda yapılmış birçok çalışma bulunmaktadır. Bu çalışmaları incelediğimizde bazı örnekler aşağıda yer almaktadır. Chung ve arkadaşlarının çalışmasında gerçek dünya verilerindeki belirsizlikleri dikkate alarak kanser bakımı için hemşire ihtiyacını tahmin ederler. Bu sayede, sağlık politikalarının etkin şekilde planlanmasına yardımcı olurlar (Chung vd., 2021). Bu yaklaşım, personel eksikliğinin önceden tespit edilip giderilmesini sağlar.

Babu ve ekibinin çalışmasında, ARIMA-ANN hibrit modeli zaman serisi tahmininde kullanılmıştır. Araştırmada, "sunspotdataset" yanı sıra elektrik fiyatı zaman serisi verileri ve L&T hisse senedi fiyat verileri gibi dört farklı model denenmiştir. Bulgular, ARIMA ve ANN modellerinin tek başlarına zaman serisi tahmininde yetersiz kaldığını, ancak hibrit modellerin her iki yöntemi de aşan bir başarı elde ettiğini göstermiştir. Babu ve Reddy'nin çalışması, zaman serisi verilerinin genellikle doğrusal ve doğrusal olmayan varyasyonlar içerdiğini, tek başına doğrusal ARIMA modelleri ve doğrusal olmayan YSA modellerinin bu tür verileri doğru bir şekilde modelleyemediğini vurgulamıştır. Sonuç olarak, ARIMA ve ANN modellerinin güçlü yanlarını birleştiren hibrit modellerin, her iki modelin avantajlarını aynı anda kullanarak bireysel modellere göre daha etkili olduğu sonucuna varılmıştır (Babu ve Reddy, 2014).

Ayoobi ve ekibi (2021), COVID-19 vakalarını ve ölümlerini tahmin etmek için zaman serisi analizini derin öğrenme yöntemleriyle birleştirmiştir. Talkhi ve arkadaşları (2021), İran'daki COVID-19 vakalarını ve ölümlerini tahmin etmeye çalışmış ve bu amaçla ARIMA gibi geleneksel yöntemlerle yapay sinir ağlarını bir araya getirmişlerdir. Bi ve diğerleri (2021), turizm verileri üzerinde derin öğrenme ile zaman serisi analizi yaparak gelecekteki turist sayısını tahmin etmişlerdir. Kim ve Kim (2021), elektrikli araç kullanımının artmasıyla birlikte elektrikli şarj istasyonlarına olan talebi tahmin etmek için ARIMA ve yapay sinir ağlarını kullanmışlardır. Cheng ve meslektaşları (2021), hastanelerdeki acil servisler için saatlik verileri ele alarak

SARIMA modelini kullanmışlardır. Alduailij ve ekibi (2021), akıllı binalar için enerji ihtiyacını zaman serisi analiziyle belirlemiş ve bu amaçla ARIMA modelini kullanmışlardır. Bu çalışmaların yanı sıra, çeşitli alanlarda benzer birçok araştırma bulunmaktadır. Wang ve diğerleri tıp lisansüstü öğrencilerinin arz ve talebini tahmin etmek için doğrusal regresyon ve gri tahmin modelini kullanmıştır (Wang vd., 2003). Tsai ve Shiue ters aktarımlı sinir ağlarıyla Tayvan işgücü piyasasındaki insan gücü talebini öngörmüşlerdir (Tsai ve Shiue, 2005). Wong ve diğerleri inşaat sektöründeki insan gücü talebini değerlendirmek için vektör düzeltme modelini tercih etmişlerdir (Wong vd., 2007). Enerji sektörüne odaklanan Merikull ve arkadaşları, işgücü talebini tahmin etmek için komut dosyası oluşturma yöntemini kullanmışlardır (Merikull vd., 2012). Bakım hizmetleri için Lin zaman serisi ve üstel düzeltme yöntemini kullanarak çalışma yapmıştır (Lin, 2015). Singh ise bulanık-nöro-entropi ve uzman sistemlerini kullanarak zaman serisi tahminleri yapmıştır (Singh, 2017). Liu ve diğerleri ve Edwards ve arkadaşları sırasıyla sağlık ve cerrahi alanlarında sistem dinamikleri ve simülasyon modelleri kullanarak arz ve talep tahminleri yapmışlardır (Liu vd., 2014; Edwards vd., 2014). İnsan kaynakları sistemlerinin optimizasyonu için Mutingi ve Mbohwa bulanık sistem dinamiğini, Wu ve Xu ise sistem dinamiği ve bulanık çok amaçlı programlama yöntemlerini kullanmıştır (Mutingi ve Mbohwa, 2012; Wu ve Xu, 2013).

3.2. Finans Alanında Tahminleme

x ve ekibinin çalışması, ARMA veya ARIMA (aynı zamanda Box-Jenkins yöntemi olarak da bilinir) gibi zaman serisi analizi tekniklerinin hisse senedi verileri üzerindeki performansını araştırmıştır. Bu araştırma, ARIMA modelinin hisse senedi verileri üzerindeki etkisini ölçmek amacıyla gerçekleştirilmiştir (Mondal vd., 2014).

Livieris ve ekibinin çalışmasında, LSTM-CNN hibrit modeli altın fiyatları zaman serisi üzerinde kullanılmıştır. Yazarlar, bu çalışmada CNN'in bilgi çıkarma yeteneğinden ve LSTM ağının uzun ve kısa vadeli etkileri tanımlama kapasitesinden faydalanmışlardır. Sonuç olarak, hibrit modelin yalnızca LSTM kullanan modellere kıyasla daha iyi performans gösterdiği belirlenmiştir. Araştırmada, kısa ve uzun vadeli bağımlılıkların zaman serisi tahmininde doğruluğu arttırabileceği sonucuna varılmıştır (Livieris vd., 2020). Vergil ve Ozkan (2007) Türkiye'nin döviz kurlarını tahmin etmek amacıyla ARIMA modellerinin tahmin gücünü araştırmışlardır. Sonuç olarak, ARIMA

modellerinin satin alma paritesi teorisine göre daha iyi tahmin performansı sergilediği tespit edilmiştir. Finansal endekslerdeki zaman serisi tahminleri için Singh ve Huang hibrit bir yöntem geliştirmişlerdir (Singh ve Huang, 2019).

Qiu ve diğerleri, birleşik derin öğrenme ağları (ensemble deep learning belief network) yapısını kullanarak zaman serisi modeli geliştirmiştir. Çalışmada elektrik yükleri talebi veri seti üzerinde destek vektör makineleri, yapay sinir ağları ve derin sinir ağları karşılaştırması yapılarak en iyi sonucun derin öğrenme yöntemleri ile elde edildiği belirtilmiştir (Qiu vd., 2014).

Gasparin ve diğerleri, elektrik tahmini problemini derin öğrenme zaman serisi modelleri ile çözmüşlerdir. Yöntemler arasında ileri beslemeli sinir ağı (feed-forward neural network-FNN), daha derin ileri beslemeli sinir ağı varyantı (deeper variant of a feed-forward neural network-DFNN), geçici evrişimli ağ (temporal convolutional network-TCN), LSTM, GRU ve seq2seq modelleri bulunmaktadır. En iyi performansın ise LSTM modellerinde kaydedildiği belirtilmiştir (Gasparin vd., 2022). Santur, 2006-2018 aralığındaki günlük verileri kullanarak geliştirilen derin öğrenme modelinde %82 oranında doğru trend tahmini gerçekleştirmiştir. Çalışmada trend yönü tahmini için çok katmanlı LSTM hücreleri kullanılmış ve performans ölçütü olarak doğruluk oranları tercih edilmiştir. Çalışma sonucunda 694 işlem gününe ait portföy büyüklüğü %39 oranında artırılarak başarılı sonuçlar elde edilmiştir (Santur, 2020).

Hasan ve diğerleri, teknik indikatörlere dayalı bir tahmin modelini BİST-100 endeksi üzerinde gerçekleştirmişlerdir. Endeksin tahmini için derin öğrenme, destek vektör makineleri, rassal orman ve lojistik regresyon modelleri kurulmuş ve performanslar karışıklık matrisi, geri dönüş oranı gibi ölçütler ile değerlendirilmiştir (Hasan vd., 2020). Gündüz ve diğerleri, BİST 100 endeksinde işlem gören 9 banka hissesinin tahmini için LSTM modeli geliştirmişlerdir. Girdi özellikleri arasında finansal haberlerden elde edilen nitelikler de bulunmaktadır. Geliştirilen model, rastgele ve Naive yöntemleri ile kıyaslanmış ve sonuçların benzer performanslara sahip olduğu belirlenmiştir (Gündüz vd., 2018).

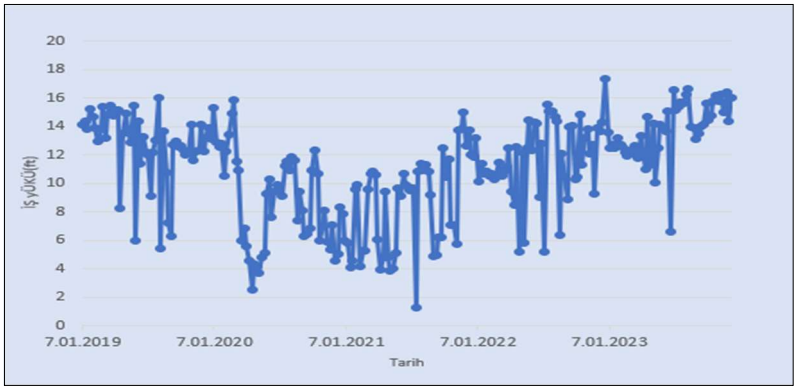
Aslan ve Erdur, kamu aydınlatma platformu verileri ile üç spor kulübü (Galatasaray, Fenerbahçe ve Beşiktaş) hisselerini tahmin etmişlerdir. Çalışmada tahmin için LSTM modeli geliştirilmiş ve deneysel sonuçlara göre KAP verilerinin hisse senetleri üzerindeki etkileri vurgulanmıştır (Aslan ve Erdur, 2020).

4. UYGULAMA

Bankalar, şubelerinde görev yapan personelin rollerini ve sorumluluklarını belirlemek için çeşitli unvanlar kullanır. Bu unvanlar, genellikle pazarlama ve operasyon olmak üzere iki ana kategori altında gruplanır. Pazarlama departmanındaki unvanlar genellikle satış ve müşteri ilişkileri üzerine odaklanırken, operasyon departmanındaki unvanlar genellikle işlemlerin yönetimi ve operasyonel süreçlerin yürütülmesiyle ilgilenir. Pazarlama tarafında, iş yükünü belirlemek genellikle zor olabilir. Çünkü pazarlama faaliyetleri, doğrudan satışa dönüşmeyebilir ve bu nedenle pazarlama ekibinin katkısını ölçmek karmaşık hale gelebilir. Operasyon tarafında ise, iş yükünü belirlemek daha doğrudan olabilir. Çünkü operasyonel işlemler genellikle belirli bir süreci veya prosedürü takip eder ve her bir işlem genellikle bir veri girişi veya bir kayıt olarak tutulur. Bu nedenle, operasyonel unvanların iş yükünü değerlendirmek daha sistematik bir şekilde yapılabilir. Bu sebeple çalışmada hedef unvan olarak operasyon başlığı altına alınabilecek gişe unvanı seçilmiştir.

4.1. Veri Toplama

Bu çalışmada, bir bankasının Sakarya ilinde faaliyet gösteren şubelerinin 7 Ocak 2019-11 Aralık 2023 yılları arasında elde edilen gişe unvanına ait iş yükleri adam-gün(ft) biriminden kullanılmıştır. Veriler haftalık olarak değerlendirilmiş olup toplam 258 haftalık veri kullanılmış ve modellerde tarih girdisi olarak hafta başları alınmıştır. Veri setine Şekil 4.1'de yer verilmiştir.



Şekil 4.1. Haftalık iş yükü (ft)

İş yükleri, bankacılık sisteminde çalışanların yapmış oldukları işlem adetleri üzerinden hesaplanan verilerdir. Arka planda her işlem için bir veri satırı oluşturulur ve o işlem hakkında çeşitli bilgiler burada depolanır. Her işlem; işlem adı, süresi, şube ve çalışan bilgisi gibi veriler içerir. Bu veriler üzerinden süreler toplanarak çalışanın günlük çalışma süresine bölünür ve iş yükü hesabı sağlanır.

Çalışmada 400'den fazla farklı işlem göz önüne alındığı için işlem adlarına göre iş yükü yoğunlukları incelenmiştir. Bu incelemenin amacı, ana işlem kalemlerini belirlemek ve bu kalem üzerinden değişkenler hakkında çıkarımlar yapmaktır. İnceleme sonucunda, ana işlem yükünün fiziksel nakit işlem süreçlerinden oluştuğu görülmüştür. Bu nedenle, enflasyon ve dolar kuru gibi fiziki para hareketliliğine sebep olabilecek veriler, makine öğrenimi ve derin öğrenim modellerinin içine dahil edilmiştir. Model değişkenleri Tablo 4.1'de yer almaktadır.

Tablo 4.1. Değişkenler ve tanımlar

Değişken	Tanım
Tarih	İlgili hafta başı tarihtir.
Ay	Tarih verisinden oluşturan ay verisidir. 1-12 arası değer alır.
Yıl	Tarih verisinden oluşturan yıl verisidir. 2019-2023 arası değer alır.
Çeyrek	Tarih verisinden oluşturan çeyrek verisidir. 1-4 arası değer alır
Tatil	1 Hafta içinde çalışma günü sayısı 5 alınarak hesaplanmış tatil sayısıdır. (Tüm tatil günlerini içerir Bayram vb.)
Arife	Hafta içerisinde yarım gün çalışma durumunu ele alır.
Yılın Haftası	Tarih verisinden oluşturan haftası verisidir. 1-52 arası değer alır.
Enflasyon Birikimli	Birikimli enflasyon değerini ele alır. (Türkiye Cumhuriyet Merkez Bankası, t.y.)
Gişe Kapasite	1 Hafta içindeki iş yükü kapasitesini ele alır. Çalışan sayısının iş günü sayısı ile çarpımından bulunur.
Dolar kuru	Haftalık dolar kuru ortalamasıdır. (Türkiye Cumhuriyet Merkez Bankası, t.y.)

4.2. Veri İşleme ve İlk Uygulama

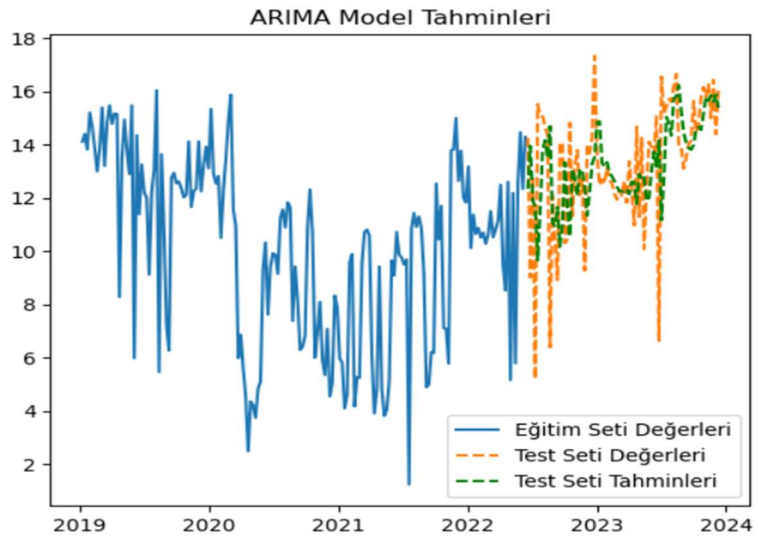
Veri seti %70 eğitim ve %30 test seti olmak üzere iki kısma ayrılmış, çalışmada 4 farklı metot üzerinden uygulamalar gerçekleştirilmiştir. ARIMA ve PROPHET metotlarında girdi olarak sadece zaman değişkeni kullanılırken LSTM, XGBOOST ve LightGBM metotlarından girdi alanına tahmin gücünü arttırmak amacı ile Tablo 4.2.'de yer alan değişkenler eklenmiştir.

	Tarih	A	Ay	Yıl	Çeyrek	Tatil	Arife	Yılın Haftası	Enflasyon Birikimli	Gişe/BST	Dolar kuru
0	2019-01-07	14.110000	1	2019	1	0.0	NaN	2	398.07	16	5.41956
1	2019-01-14	14.400000	1	2019	1	0.0	NaN	3	398.07	16	5.40686
2	2019-01-21	13.820000	1	2019	1	0.0	NaN	4	398.07	14	5.30768
3	2019-01-28	15.200000	1	2019	1	0.0	NaN	5	398.07	15	5.25536
4	2019-02-04	14.660000	2	2019	1	0.0	NaN	6	398.71	14	5.22162
...	...	...	...	...	...	...	...	...	...	...	...
253	2023-11-13	16.255427	11	2023	4	0.0	NaN	46	1806.50	15	28.63978
254	2023-11-20	14.995427	11	2023	4	0.0	NaN	47	1806.50	15	28.79392
255	2023-11-27	16.432159	11	2023	4	0.0	NaN	48	1806.50	15	28.90894
256	2023-12-04	14.390728	12	2023	4	0.0	NaN	49	1859.38	15	28.91696
257	2023-12-11	16.030775	12	2023	4	0.0	NaN	50	1859.38	15	28.98828

258 rows × 11 columns

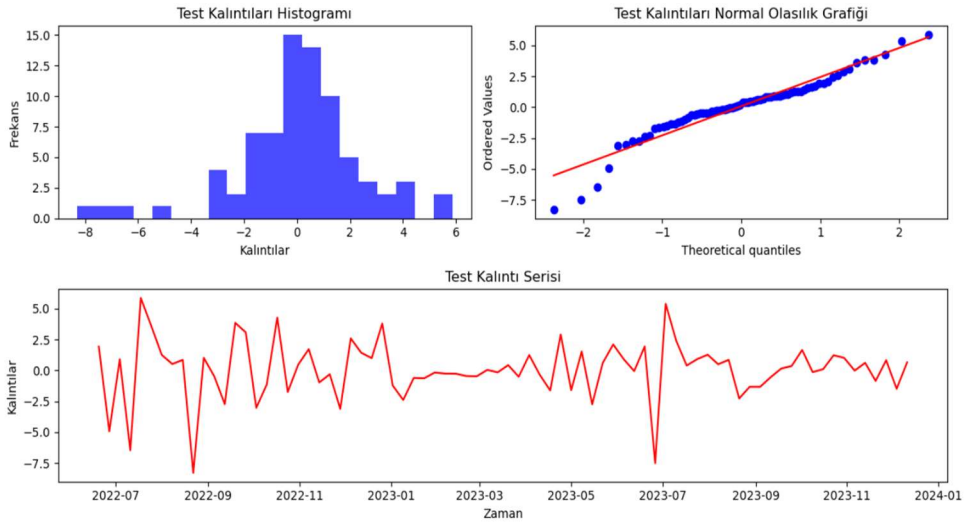
Şekil 4.2. Veri tablosu görünümü

Kullanılan ilk model arıma modeli olmakla beraber bu modelde sadece tarih verisi üzerinden tahminleme sağlanmaktadır. Arıma modeline ait çıktılar Şekil 4.3.'te yer almaktadır.



Şekil 4.3. Arıma model çıktıları

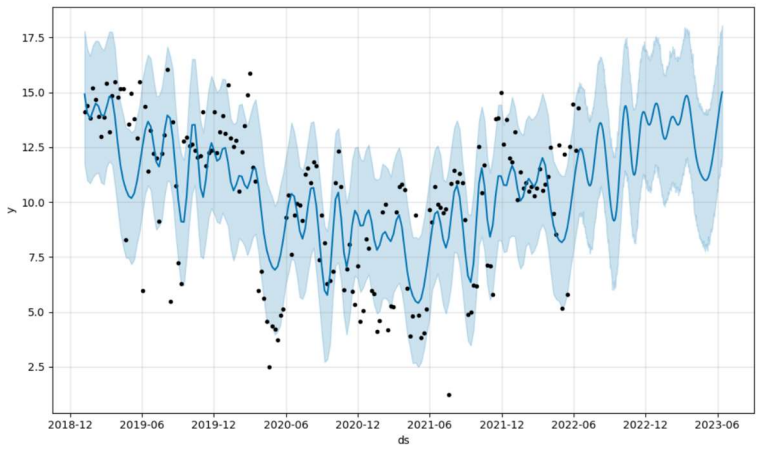
Arıma modeli üzerinde parametre değerleri (p, d, q) =(4, 1, 0) olarak kullanılmış ve model çıktıları bu paramete değeri üzerinden elde edilmiştir. İlgili model Python uygulamasında “statsmodels.tsa.arima.model” kütüpühanesi altında yer ARIMA modeli kullanılarak gerçekleştirilmiştir.



Şekil 4.4. Arıma model performans grafikleri

Şekil 4.4.’te yer alan ARIMA modeli performans metriklerine göre kalıntılar normal dağılım özelliği gösterse de Histogram grafiği simetrik özellik taşımaması, uç değerlerin olması ve modelin bazı noktalarda hatalı tahminler yaptığı sonucuna ulaşılabilir.

Prophet modeli yine Arıma modeli gibi sadece zaman değişkeninin girdi olarak kullanılabileceği bir modeldir. Bu model Python uygulamasında “prophet” kütüphanesi altında yer alan prophet modeli kullanılarak gerçekleştirilmiştir.



Şekil 4.5. Prophet gelecekteki değer tahminleri

Şekil 4.5.'te geçmişteki ve gelecekteki tarihleri içeren bir veri seti oluşturulmuştur. Bu genişletilmiş zaman çerçevesi üzerinde, model tahminler yapar. Bu grafik, tahmin edilen gelecekteki değerleri ve bu tahminlerin belirsizlik aralıklarını göstermektedir.

XGBoost, LSTM ve LightGBM modellerinde tarih verisi haricinde tahmin doğruluğunu arttırmak amacıyla eklenen değişkenler ile üç model varsayılan parametre değerleri üzerinden çalıştırılmıştır. Bu değerlere Tablo 4.2.'de yer verilmiştir. İlgili modeller için Python uygulamasında "tensorflow.keras.layers" ,"xgboost" ve "lightgbm" kütüpühaneleri altında yer alan XGBoost, LSTM, LGB modelleri kullanılmıştır.

Tablo 4.2. XGBoost, LSTM ve LightGBM modelleri için varsayılan parametre değeri

Algoritma	Hiper Parametre	Varsayılan Değer
XGBoost	Colsample_bytree	1.0
	Learning_rate	0.3
	N_estimators	100
	Subsample	1.0
	Max_depth	6
LSTM	Units	32
	Batch_size	1
	Epochs	10
LightGBM	Colsample_bytree	1.0
	Learning_rate	0.1
	N_estimators	100
	Subsample	1.0
	Max_depth	-1

Tablo 4.3. Varsayılan parameteler ile model çıktıları

Modeller	Eğitim Seti Korelasyon	Test Seti Korelasyon	Test Seti MSE	Test Seti MAE
ARIMA	0,61	0,32	5,65	1,85
PROPHET	0,36	0,05	6,04	1,97
LSTM	0,63	0,17	7,83	2,17
XGBOOST	0,99	0,75	2,55	1,27
LightGBM	0,97	0,81	2,48	1,29

Veri seti üzerinde dört farklı model ile elde edilen ilk sonuçlar Tablo 4.3.'te yer almaktadır. ARIMA, Prophet, LSTM, XGBoost ve LightGBM modellerinin performansını karşılaştıran bir değerlendirme yapılmaktadır. ARIMA modeli, Prophet ve LSTM modellerine kıyasla eğitim ve test seti korelasyonu açısından daha iyi performans göstermektedir. Ancak, XGBoost ve LightGBM modellerine kıyasla test seti korelasyonu daha düşük ve test seti MSE ve MAE değerleri daha yüksektir.

Prophet modelinin eğitim ve test seti korelasyonları oldukça düşüktür ve test seti MSE ve MAE değerleri de yüksektir, bu da Prophet modelinin diğer modellere kıyasla daha düşük performans sergilediğini gösterir. LSTM modelinin eğitim seti korelasyonu yüksek olmasına rağmen, test seti korelasyonu oldukça düşüktür ve test seti MSE ve MAE değerleri yüksektir, bu da modelin düşük performans gösterdiğini belirtir. XGBoost modeli, hem eğitim hem de test seti korelasyonu açısından oldukça yüksek bir performans sergilemektedir. Ayrıca, test seti MSE ve MAE değerleri diğer modellere kıyasla oldukça düşüktür, bu da XGBoost modelinin genel olarak en iyi performansı gösterdiğini ifade eder. LightGBM modeli de yüksek performans sergilemektedir; eğitim seti korelasyonu ve test seti korelasyonu oldukça yüksek olup, test seti MSE ve MAE değerleri düşüktür. LightGBM modeli, özellikle test seti korelasyonu açısından en iyi sonuçları vermektedir. Genel olarak, XGBoost ve LightGBM modelleri en iyi performansı gösterirken, Prophet ve LSTM modelleri daha düşük performans sergilemektedir. ARIMA modeli ise Prophet ve LSTM'ye göre daha iyi ancak XGBoost ve LightGBM'ye göre daha düşük performans göstermektedir.

4.3. Hiper Parametre Çalışması

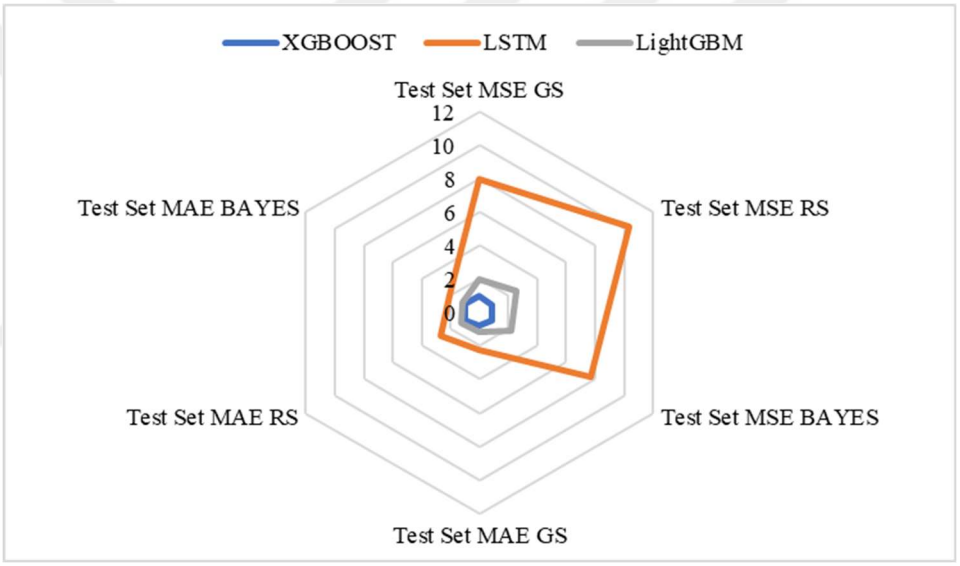
XGBoost, LSTM ve LightGBM modelleri için hiper parametre yöntemi kullanılarak sonuçlar optimize edilmeye çalışılmış ve bu hiper parametre yöntemleri birbirleri ile kıyaslanmıştır. Modellerde kullanılan hiper parametre değer aralıkları Tablo 4.4.'te gösterilmiştir.

Tablo 4.4. Hiper parametre değerleri

Algoritma	Hiper Parametre	Aralık
XGBoost	Colsample_bytree	0.5-1.0
	Learning_rate	0.01-1.0
	N_estimators	100-1000
	Subsample	0.5-1.0
	Max_depth	1-20
LSTM	Units	50-150
	Batch_size	1-32
	Epochs	5-15
LightGBM	Colsample_bytree	0.5-0.8
	Learning_rate	0.01-0.1
	N_estimators	200-1000
	Subsample	0.5-0.8
	Max_depth	(-1)-7

XGBoost modelleri arasında en iyi performansı XGBoost-GS modeli göstermektedir. Bu modelin korelasyon değeri 0,91 ile en yüksek seviyede olup, aynı zamanda MSE ve MAE değerleri de oldukça düşüktür. XGBoost-RS ve XGBoost-Bayes modelleri ise korelasyon değeri açısından birbirine benzer olup (0,87), MSE ve MAE değerleri XGBoost-GS modeline göre biraz daha yüksektir.

LSTM modelleri değerlendirildiğinde, genel olarak düşük performans sergiledikleri görülmektedir. LSTM-GS ve LSTM-RS modellerinin korelasyon değerleri oldukça düşük (sırasıyla 0,16 ve 0,17) ve MSE ile MAE değerleri yüksektir. LSTM-Bayes modeli ise diğer LSTM modellerine göre daha iyi performans göstermekte olup, korelasyon değeri 0,7'dir ancak bu değer yine de XGBoost ve LightGBM modellerine kıyasla düşüktür.



Şekil 4.6. Hiper parametre sonuçları

LightGBM modelleri arasında, LightGBM-GS modeli en iyi performansı sergilemektedir. Bu modelin korelasyon değeri 0,85 olup, MSE ve MAE değerleri makul düzeydedir. LightGBM-RS ve LightGBM-Bayes modelleri ise benzer performans sergilemekte olup, korelasyon değerleri 0,81'dir. Bu modellerin MSE ve MAE değerleri ise LightGBM-GS modeline göre biraz daha yüksektir.

Genel olarak değerlendirildiğinde, XGBoost-GS modeli en yüksek korelasyon değerine sahip olup, en iyi performansı göstermektedir. LightGBM-GS modeli ise iyi bir alternatif olarak öne çıkmaktadır. LSTM modelleri ise düşük korelasyon değerleri ve yüksek hata metrikleri nedeniyle tercih edilmemelidir. Bu çalışmada elde edilen performans sonuçlarına Şekil 4.6.'da ve Tablo 4.5.'te yer verilmiştir.

Tablo 4.5. Hiper parametre sonuçları

Model ve Yöntem	Test Set MSE	Test Set MAE	Test Set Correlation
XGBoost- GS	0,99	0,77	0,91
XGBoost- RS	0,87	0,96	0,87
XGBoost- Bayes	0,87	1,01	0,87
LSTM- GS	7,95	2,23	0,16
LSTM- RS	10,31	2,72	0,17
LSTM- Bayes	7,65	2,09	0,7
LightGBM-GS	1,96	1,14	0,85
LightGBM-RS	2,53	1,29	0,81
LightGBM-BAYES	2,23	1,18	0,81

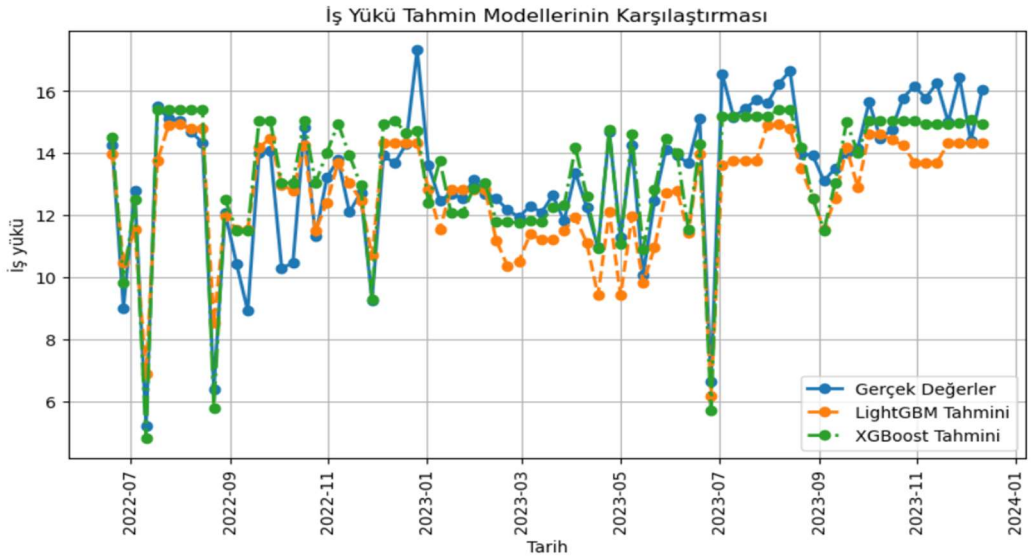
4.4. Sonuçların Değerlendirilmesi

Uygulama sonrasında Hiper Parametre uygulanmış en iyi performans gösteren iki algoritmanın çıktıları Tablo 4.6.'da yer almaktadır. Bu veri seti üzerinde XGBOOST modelinin en iyi performans metriklerine sahip olduğu ve daha sonrasında en iyi performans metriğinin LightGBM modeline ait olduğu görülmüştür.

Tablo 4.6. Modellerin performans metriği değerleri

Metot	MSE	MAE	Korelasyon Eğitim seti	Korelasyon Test Seti
LightGBM-GS	1,96	1,14	0.97	0.85
XGBOOST-GS	0.99	0.77	0.99	0.91

Aynı zamanda XGBoost ve LightGBM modellerinin test setlerinin tahmin değerleri ile görsel olarak karşılaştırılması Şekil 4.7.'de yer verilmiştir. Görsel üzerinde de tahmin performansları hakkında çıkarımlar yapılabilmektedir.



Şekil 4.7. XGBoost ve LightGBM model çıktıları

Problemin çözümü için hata oranları ve test setindeki 0.91 korelasyon değeri ile XGBOOST modeli önerilmiştir. XGBoost modeli için belirlenen hiper parametre değerleri Tablo 4.7.'de görülmektedir ve bu parametre değerleri Grid search metodu ile elde edilmiştir.

Tablo 4.7. XGBoost modeli hiper parametre değerleri

Metrik	Açıklama	Değer
N_estimators	Ağaç sayısı	500
Learning_rate	Öğrenme oranı	1
Max_depth	Maksimum ağaç derinliği	7
Subsample	Alt örnekleme oranı	0.8
Colsample_bytree	Ağaç için sütun örnekleme oranı	0.8

4.5. Çalışmada Kullanılan Uygulamalar

Çalışmada Python uygulaması üzerinde Tablo 4.8'de yer alan çeşitli açık kaynak kütüphaneler ve fonksiyonlar kullanılmıştır.

Tablo 4.8. Phyton kütüphaneler

Kullanım Amaçları	Kütüphane ve Fonksiyonlar
Veri Manipülasyonu ve Analizi	pandas: pd.read_csv(), pd.DataFrame()
Bilimsel Hesaplama	numpy: np.array(), np.mean()
Görselleştirme	matplotlib: plt.plot(), plt.show()
Makine Öğrenimi ve Modelleme	xgboost: xgb.train(), xgb.XGBClassifier(), statsmodels: ARIMA(),lightgbm
Derin Öğrenme	tensorflow.keras: Sequential(), LSTM(), Dense()

Tablo 4.8. (Devamı) Phyton kütüphaneler

Kullanım Amaçları	Kütüphane ve Fonksiyonlar
Model Seçimi ve Hiperparametre Optimizasyonu	sklearn.model_selection: train_test_split(), GridSearchCV(), RandomizedSearchCV() bayes_opt: BayesianOptimization() skopt: BayesSearchCV(), Real(), Integer(), Categorical()
İstatistiksel Fonksiyonlar	scipy.stats: uniform(), randint()
Yardımcı Araçlar	itertools: product()

5. SONUÇLAR VE ÖNERGELER

5.1. Sonuç ve Öneriler

İstatistiksel yöntemler, makine öğrenimi ve derin öğrenme yöntemleri karşılaştırıldığında makine öğrenimi yöntemlerinin genellikle karmaşık veri yapılarını ve ilişkilerini yakalama konusunda oldukça güçlü olduğu görülmektedir. Ayrıca XGBoost'un ağaç tabanlı yapısı, değişkenler arasındaki etkileşimleri ve önemli özellikleri otomatik olarak öğrenebilir, bu da zaman serisi tahmininde faydalı olduğu çalışmada gözükmektedir. XGBoost'un en iyi performansı vermesi veri setinizin karmaşıklığının ve özelliklerinin XGBoost'un güçlü yönlerine uygun olduğu anlamına gelebilir. Aynı zamanda bu sonuçlar, Hiper parametre yöntemlerinde GS yönteminin model performansını en üst düzeye çıkardığını göstermektedir.

Bu model ile bankanın gişe kadrosu çalışanlarının daha doğru ve verimli bir şekilde yönetilebilmesi sağlanacaktır. Eski sistemde, son 6 ay verilere bakılarak iş yükü ortalaması üzerinden kadro belirlenmektedir. Ancak bu yöntem, iş yükünün farklı etkenlerden etkilenmesi ve zaman serisindeki değişken trendler nedeniyle optimum planlama sağlamada yetersiz kalmıştır. Örneğin, bir şubeye fazla norm kadro verilmesi, atıl kapasiteye sebep olmuş ve bu durum çalışanın ihtiyaç duyulan farklı bir şehre kaydırılması sonucu sosyal hayatını olumsuz etkilemiştir. Diğer yandan, norm kadronun eksik belirlenmesi, şube içindeki diğer çalışanların iş yükünü artırmış, çalışan memnuniyetsizliğine yol açmış ve müşterilerin bekleme sürelerinin uzaması nedeniyle müşteri memnuniyetsizliğini de artırmıştır.

Bu bağlamda geliştirilmiş olan tahmin modeli ile gelecekte oluşacak iş yükünün doğru bir şekilde öngörülmesi ve kadro belirleme sürecinde bu verinin girdi olarak kullanılması, iş yüklerinin daha optimum ve verimli bir şekilde dağıtılmasını sağlayacaktır. Elde edilen iş yükü tahminleri, kadro sayısı ve iş modelleri gibi kaynakların doğru planlanmasında bankanın karar verme süreçlerine katkı sağlayacaktır. Bu sayede kadro planlamaları daha verimli bir şekilde yapılabilecek ve unvanlara ait görev tanımları daha sağlıklı bir şekilde güncellenebilecektir.

Bu çalışmaya konu olan Sakarya ilinde çıktılar üzerinden yapılan bir hesaplama Tablo 5.1.'de yer almaktadır. Kadro planlamaları yapılırken iş gücü adam-gün(ft) türünden hesaplandığı için küsüratlar 0.5'ten aşağı veya yukarı olma durumlarına göre yuvarlanmaktadır. Buna göre 2023 yılı kadro planlamasında geçmiş 6 aylık iş yükü verisinde ortalama 12.47 ft'lik bir iş yükü oluşmuştur, küsürat olan 0.47 değeri 0.5 değerinden aşağı olduğu için eski hesaplama yöntemine göre 12 kişilik bir kadro planlaması sağlanmıştır. Tahminleme modeli ile elde edilen iş yükü ortalaması ise 2023 yılı için 13.48 olarak hesaplanmıştır bu iş yükü için 13 kişilik bir norm kadro belirlenmiş olacaktır. Bu iki kadro sayısı için gerçek oluşan iş yükü değerleri üzerinde mutlak sapmaları karşılaştırıldığından tahminleme modelinde %23,6 oranında bir azalma görülmüştür.

Aynı zamanda çalışmada maliyet değerlendirmesi yapıldığında 1 kişinin iş yükü maliyetinin, kadro sayısının eksik olması durumunda oluşacak müşteri memnuniyetsizliği, çalışan memnunyetsizliği ve gerçekleşmeyen işlemlerdeki fırsat maliyetleri gibi durumlarla karşılaştırıldığında üçte biri kadar olacağı varsayılmaktadır. Bu durumda hesaplamada belirlenen kadro değeri üzerinden bir maliyet hesaplaması yapacak olursak bunu şu şekilde belirtmek doğru olacaktır;

Örneğin; belirlenen kadro değeri 12 kişi ve 2023 yılında örnek bir gün için oluşan iş yükü 11 ft bu durumda o gün için 1 kişilik atıl kasapasite kadar maliyet oluşmaktadır.

Diğer bir senaryoda bir gün oluşan iş yükünü 13.5 ft olarak ele alalım. Verilen kadro yine 12 olduğu için 1.5 ft değerinde fazla iş yükü durumu söz konusu olacaktır. Buradaki maliyeti yukarıda anlatılanlara istinaden adam-gün(ft) olarak çevirmemiz için ise 3 ile çarmamız gerekecektir ve 1.5 ft fazla iş yükünün maliyeti 4.5 ft olacaktır.

Buna istinaden 2023 yılı için yapılan bir maliyet hesabında eski modelde kadro sayısı 12 kişi olduğunda oluşan maliyet 1537,95 ft iken yeni tahmin modelinde kadro sayısı 13 kişi olduğunda oluşan maliyet 1030,86 ft olacaktır. Eski ve yeni model arasındaki 507,10 ft'lik fark 1 kişinin 245 iş günü çalışma süresini baz alarak 2 personel kazancına eşdeğerdir. Bir yıllık toplam kazanç hesabı yapılırken yeni tahmin modelinde bir personel ekstra verildiği için bir personel maliyeti düşülmelidir. Bunun sonucunda da 1 yıllık maliyet 1 personel olarak hesaplanabilmektedir.

Tablo 5.1. Model kazanımları

Değerlendirme Metriği	Eski Hesaplama Modeli (2023)	Tahmin Modeli ile Hesaplama (2023)
Ortalama (Ft)	12,47	13,48
Belirlenen Kadro (Kişi)	12	13
Mutlak Sapma (Ft)	1040,76	722,93
Maliyet (Ft)	1537,95	1030,86

Bu durum, modelin daha düşük kadro belirlemeden ziyade daha optimum kadro belirlemesini taahhüt etmektedir. Gelecek çalışmalarda, problem farklı değişkenler eklenerek tekrar modellenebilir. Bu değişkenleri eklerken tahmin doğruluğunu artırmak amacıyla topluluk yaklaşımları kullanılabilir. Örneğin, müşteri davranışları, ekonomik göstergeler, mevsimsel etkiler ve yerel etkinlikler gibi değişkenlerin modelde kullanılması, iş yükü tahminlerinin doğruluğunu artırabilir. Bu çalışmanın bankanın operasyonel verimliliğine ve müşteri memnuniyetine önemli katkılar sağladığı düşünülmektedir. Ayrıca, diğer sektörler için de benzer modeller geliştirilebilir ve iş gücü yönetiminde kullanılabilir. Bu tür modeller, kaynakların daha etkin ve verimli bir şekilde kullanılmasına olanak tanır ve işletmelerin rekabet gücünü artırır.

KAYNAKLAR

Alduailij, M. A., Petri, I., Rana, O., Alduailij, M. A., & Aldawood, A. S. (2021). Forecasting peak energy demand for smart buildings. The Journal of Supercomputing, 77(6), 6356-6380. https://doi.org/10.1007/s11227-020-03540-3

Aslan, B., & Erdur, R. C. (2020). Stock Market Prediction with Deep Learning Using Public Disclosure Platform Data. 2020 Innovations in Intelligent Systems and Applications Conference (ASYU), 1-5.

Ayoobi, N., Sharifrazi, D., Alizadehsani, R., Shoeibi, A., Gorriz, J. M., Moosaei, H., Khosravi, A., Nahavandi, S., Gholamzadeh Chofreh, A., Goni, F. A., Klemeš, J. J., & Mosavi, A. (2021). Time series forecasting of new cases and new deaths rate for COVID-19 using deep learning methods. Results in Physics, 27, 104495. https://doi.org/10.1016/j.rinp.2021.104495

Babu, C. N., & Reddy, B. E. (2014). A moving-average filter-based hybrid ARIMA–ANN model for forecasting time series data. Applied Soft Computing, 23, 27–38. doi: 10.1016/j.asoc.2014.05.028

Bauwens, L., & Veredas, D. (2004). "The stochastic conditional duration model: a latent variable model for the analysis of financial transaction data". Journal of Econometrics, 119(2), 381-412.

Bayramoğlu, M. F. (2007). Finansal endekslerin öngörüsünde yapay sinir ağı modellerinin kullanılması: İMKB ulusal 100 endeksinin gün içi en yüksek ve en düşük değerlerinin öngörüsü üzerine bir uygulama. Yüksek Lisans Tezi, Karaelmas Üniversitesi, Sosyal Bilimler Enstitüsü.

Beysolow, T. (2018). Applied natural language processing with Python. Apress.

Bi, J.-W., Li, H., & Fan, Z.-P. (2021). Tourism demand forecasting with time series imaging: A deep learning model. Annals of Tourism Research, 90, 103255. https://doi.org/10.1016/j.annals.2021.103255

Box, G. E. P., & Jenkins, G. M. (1976). Time Series Analysis: Forecasting and Control. Holden-Day.

Breiman, L. (2001). Random Forests. Machine Learning, 45, 5–32.

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.

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.

Cheng, Q., Argon, N. T., Evans, C. S., Liu, Y., Platts-Mills, T. F., & Ziya, S. (2021). Forecasting emergency department hourly occupancy using time series analysis. The American Journal of Emergency Medicine, 48, 177-182. https://doi.org/10.1016/j.ajem.2021.04.075

Chiappa, S., & Calandra, R. (2019). "Uncertainty-aware Learning in Continuous-time Dynamical Systems". Advances in Neural Information Processing Systems, 32.

Chung, M. H., Hung, K. C., Chiou, J. F., Fang, H. F., & Chiu, C. H. (2021). Nursing Manpower Forecast For Cancer Patients. Computer Methods and Programs in Biomedicine, 201, 05967. https://5cef0b7b43e7a42fd019406d30fe2e79ab6b0d1f.vetisonline.com/science/article/abs/pii/S0169260721000420

Çağlar, T. (2007). Talep Tahmininde Kullanılan Yöntemler ve Fens Teli Üretimi Yapan Bir İşletmede Uygulanması. Yüksek Lisans Tezi, Kırıkkale Üniversitesi, Fen Bilimleri Enstitüsü.

Dorogush, A. V., Ershov, V., & Gulin, A. (2018). CatBoost: gradient boosting with categorical features support. Proceedings of the 32nd International Conference on Neural Information Processing Systems.

Edwards, J. P., Datta, I., Hunt, J. D., Stefan, K., Ball, C. G., Dixon, E., & Grondin, S. C. (2014). A novel approach for the accurate prediction of thoracic surgery workforce requirements in Canada. J. Thorac. Cardiovasc. Surg., 148, 7–12. doi: 10.1016/j.jtcvs.2014.03.031

Friedman, J. H. (2001). "Greedy Function Approximation: A Gradient Boosting Machine". Annals of Statistics, 29(5), 1189-1232.

Gasparin, A., Lukovic, S., & Alippi, C. (2022). Deep learning for time series forecasting: The electric load case. CAAI Transactions on Intelligence Technology, 7(1), 1-25. https://doi.org/10.1049/cit2.12060

Ghani, K. A. (2000). Advanced manufacturing technology and planned organizational change. The Journal of High Technology Management Research, 11(1), 1-10.

Gündüz, H., Yaslan, Y., & Çataltepe, Z. (2018). Stock market prediction with deep learning using financial news. 2018 26th Signal Processing and Communications Applications Conference (SIU), 1-4. https://doi.org/10.1109/SIU.2018.8404616

Hasan, A., Kalıpsız, O., & Akyokuş, S. (2020). Modeling traders' behavior with deep learning and machine learning methods: Evidence from BIST 100 index. Complexity, 2020. https://doi.org/10.1155/2020/8285149

Hipel, K. W., McLeod, A. I., & Lennox, W. C. (1977). Advances in Box-Jenkins modeling: 1. Model construction. Water Resources Research, 13(3), 567-575. Erişim adresi: https://doi.org/10.1029/WR013i003p00567

Hu, Clark (2002). Advanced Tourism Demand Forecasting: ANN and Box-Jenkins Modelling. Doktora Tezi, Purdue University, MI, USA.

Ju, Y., Sun, G., Chen, Q., Zhang, M., Zhu, H., Rehman, M.U., A model combining convolutional neural network and LightGBM algorithm for ultra-short-term wind power forecasting, Ieee Access, 7, 28309-28318, 2019.

Julier, S. J., & Uhlmann, J. K. (2004). "Unscented Filtering and Nonlinear Estimation". Proceedings of the IEEE, 92(3), 401-422.

Ke, G., Meng, Q., Finley, T., Wang, T., Chen, W., Ma, W., Ye, Q., & Liu, T.-Y. (2017). "LightGBM: A Highly Efficient Gradient Boosting Decision Tree". Advances in Neural Information Processing Systems, 30.

Kermani, E. F., Barani, G. A., & Hessaroeyeh, M. G. (2018). Cavitation damage prediction on dam spillways using Fuzzy-KNN modeling. J Appl Fluid Mech, 11(2), 323-329. doi:10.29252/JAFM.11.02.28356

Kim, Y., & Kim, S. (2021). Forecasting Charging Demand of Electric Vehicles Using Time-Series Models. Energies, 14(5), Art. 5. https://doi.org/10.3390/en14051487

LeCun, Y., & et al. (1989). Handwritten digit recognition: Applications of neural network chips and automatic learning. IEEE Communications Magazine, 27(11), 41-46.

Lin, C. Y. (2015). Developing a need assessment model for manpower of home care worker. Journal for Social Development Study, 67–94. doi:10.6687/JSDS.2015.16.3

Liu, P. L., Chen, C. H., & Pan, Y. W. (2014). Application of system dynamics on supplydemand of military physician manpower in R.O.C. J. Health. Manag., 15, 73–88. doi:10.6174/JHM2014.15(1).73

Livieris, I. E., Pintelas, E., & Pintelas, P. (2020). A CNN–LSTM model for gold price time-series forecasting. Neural Computing and Applications, 32(23), 17351–17360. doi:10.1007/s00521-020-04867-

Makridakis, S., Anderson, A., Carbone, R., Fildes, R., Hibdon, M., Lewandowski, R., Newton, J., Parzen, E., & Winkler, R. (1982). The Accuracy of Extrapolation (Time Series) Methods: Results of A Forecasting Competition. Journal of Forecasting, 1(2), 111-153.

Mata, A. G. (2020). A Comparison Between LSTM and Facebook Prophet Models: A Financial Forecasting Case Study. Universitat Oberta de Catalunya.

Merikull, J., Eamets, R., Humal, K., & Espenberg, K. (2012). Power without manpower: Forecasting labour demand for Estonian energy sector. Energy Policy, 49, 740–750. doi: 10.1016/j.enpol.2012.07.018

Mondal, P., Shit, L., & Goswami, S. (2014). Study of effectiveness of time series modeling (ARIMA) in forecasting stock prices. International Journal of Computer Science, Engineering and Applications, 4, 13–29.

Mutingi, M., & Mbohwa, C. (2012). Fuzzy system dynamics and optimization with application to manpower systems. Int. J. Ind. Eng. Comput., 3, 873–886. doi: 10.5267/j.ijiec.2012.05.004

Nomidl Machine Learning. (2023). "Random Forest algorithm for Machine Learning". Erişim adresi: https://www.nomidl.com/machine-learning/random-forest-algorithm-for-machine-learning/.

Qiu, X., Zhang, L., Ren, Y., Suganthan, P. N., & Amaratunga, G. (2014). Ensemble deep learning for regression and time series forecasting. 2014 IEEE Symposium on Computational Intelligence in Ensemble Learning (CIEL), 1-6. https://doi.org/10.1109/CIEL.2014.7015739

Santur, Y. (2020). Deep learning based regression approach for algorithmic stock trading: A case study of the Bist30. Gümüşhane Üniversitesi Fen Bilimleri Dergisi, 10(4), 1195-1211.

Seddighi, H. R., Lawyer, K. A., & Katos, A. V. (2000). Econometrics: A Practical Approach. Routledge Taylor and Francis Group.

Singh, P. (2017). High-order fuzzy-neuro-entropy integration based expert system for time series forecasting. Neural Comput. & Applic., 28, 3851–3868. doi:10.1007/s00521-016-2261-4

Singh, P., Huang, Y. P. (2019). A High-order neutrosophic-neuro-gradient descent algorithm-based expert system for time series forecasting. International Journal of Fuzzy Systems, 21, 2245–2257. doi:10.1007/s40815-019-00690-2

Smola, A. J., & Scholkopf, B. (2004). A Tutorial on Support Vector Regression. Stat. Comput. 14, 199–222.

Talkhi, N., Akhavan Fatemi, N., Ataei, Z., & Jabbari Nooghabi, M. (2021). Modeling and forecasting number of confirmed and death caused COVID-19 in IRAN: A comparison of time series forecasting methods. Biomedical Signal Processing and Control, 66, 102494. https://doi.org/10.1016/j.bspc.2021.102494

Tay, F. E., & Cao, L. (2001). Application of Support Vector Machines In Financial Time Series Forecasting. Omega, 29(4), 309-317.

Taylor, S. J., & Letham, B. (2017). Prophet: Forecasting at Scale Facebook. Erişim adresi: https://facebook.github.io/prophet/.

Taylor, S. J., & Letham, B. (2018). Forecasting at Scale. The American Statistician, 72(1), 37-45. doi:10.1080/00031305.2017.1380080

Tsai, C. Y., & Shiue, Y. C. (2005). Application of back-propagation neural network for predicting the manpower demands of labor market in Taiwan. Chung Hua Journal of Management, 6, 1–14. doi:10.30053/CHJM.200506.0001

Türkiye Bankalar Birliği. (2023). Bankacılık sistemde banka, çalışan ve şube sayıları, Eylül 2023 (Rapor Kodu: DT13). Erişim adresi: https://www.tbb.org.tr/Content/Upload/istatistikiraporlar/ekler/4196/Banka_Calisan_ve_Sube_Sayilari-Eylul_2023.pdf

Vapnik, V., & Cortes, C. (1995). Support Vector Networks. Machine Learning, 20(3), 273–297.

Vergil, H., & Ozkan, F. (2007). Purchasing Power Parity and ARIMA Models in Forecasting Exchange Rates: The Case of Turkey. Istanbul Stock Exchange Review, 9(35), 37-50.

Wang, Z. H., Zhao, Z. G., Huang, Q., Wu, S. L., & Lu, Z. X. (2003). Forecasting study on demand and supply of medical postgraduates. J. Huazhong Univ. Sci. Technol. Med. Sci. 23, 94–96. doi:10.1007/BF02829476

Wong, J. M., Chan, A. P., & Chiang, Y. H. (2007). Forecasting construction manpower demand: A vector error correction model. Build. Environ., 42, 3030–3041. doi: 10.1016/j.buildenv.2006.07.024

Wu, Z., & Xu, J. (2013). Predicting and optimization of energy consumption using system dynamics-fuzzy multiple objective programming in world heritage areas. Energy, 49, 19–31. doi: 10.1016/j.energy.2012.10.030

Yafee, R., & Mcgee, M. (2000). Introduction to Time Series Analysis and Forecasting with Applications of SAS and SPSS. Academic Press.

Yang, C., & Chang, P. (2020). Forecasting the demand for container throughput using a mixed-precision neural architecture based on CNN–LSTM. Mathematics, 8, 1-17.

EKLER

EK A. 40 Haftalık veri seti örneği

EK A

Tarih	A	A_SAF	Ay	Yıl	Çeyrek	Tatil	Arife	Yılın Haftası	Enflasyon Birikim	Gişe/BST	Dolar kuru
7.01.2019	14,11	14,11	1	2019	1	0		2	398,07	16	5,41956
14.01.2019	14,4	14,4	1	2019	1	0		3	398,07	16	5,40686
21.01.2019	13,82	13,82	1	2019	1	0		4	398,07	14	5,30768
28.01.2019	15,2	15,2	1	2019	1	0		5	398,07	15	5,25536
4.02.2019	14,66	14,66	2	2019	1	0		6	398,71	14	5,22162
11.02.2019	13,88	13,88	2	2019	1	0		7	398,71	15	5,26716
18.02.2019	13	13	2	2019	1	0		8	398,71	15	5,30066
25.02.2019	13,85	13,85	2	2019	1	0		9	398,71	15	5,315
4.03.2019	15,39	15,39	3	2019	1	0		10	402,81	15	5,40742
11.03.2019	13,2	13,2	3	2019	1	0		11	402,81	14	5,45008
18.03.2019	14,83	14,83	3	2019	1	0		12	402,81	16	5,45616
25.03.2019	15,47	15,47	3	2019	1	0		13	402,81	16	5,51324
1.04.2019	14,78	14,78	4	2019	2	0		14	409,63	16	5,57308
8.04.2019	15,16	15,16	4	2019	2	0		15	409,63	15	5,68532
15.04.2019	15,15	15,15	4	2019	2	0		16	409,63	15	5,7782
22.04.2019	8,28	8,28	4	2019	2	1		17	409,63	8	5,85416
29.04.2019	13,53	13,53	4	2019	2	1		18	409,63	11	5,95824
6.05.2019	14,94	14,94	5	2019	2	0		19	413,52	16	6,11452
13.05.2019	13,79	13,79	5	2019	2	0		20	413,52	16	6,02946
20.05.2019	12,9	12,9	5	2019	2	0		21	413,52	15	6,06886
27.05.2019	15,47	15,47	5	2019	2	0		22	413,52	16	6,01488
3.06.2019	5,98	5,98	6	2019	2	3,5	1	23	413,63	7	5,79424
10.06.2019	14,35	14,35	6	2019	2	0		24	413,63	16	5,81678
17.06.2019	11,39	11,39	6	2019	2	0		25	413,63	12	5,83354
24.06.2019	13,25	13,25	6	2019	2	0		26	413,63	12	5,79526
1.07.2019	12,21	12,21	7	2019	3	0		27	419,24	11	5,66692
8.07.2019	12	12	7	2019	3	0		28	419,24	10	5,69148
15.07.2019	9,12	9,12	7	2019	3	1		29	419,24	9	5,69028
22.07.2019	12,2	12,2	7	2019	3	0		30	419,24	11	5,69036
29.07.2019	13,06	13,06	7	2019	3	0		31	419,24	14	5,60644
5.08.2019	16,04	16,04	8	2019	3	0		32	422,84	16	5,52436
12.08.2019	5,46	5,46	8	2019	3	3		33	422,84	5	5,5612
19.08.2019	13,64	13,64	8	2019	3	0		34	422,84	10	5,69414
26.08.2019	10,75	10,75	8	2019	3	1		35	422,84	9	5,81534
2.09.2019	7,24	7,24	9	2019	3	0		36	427,04	5	5,7463
9.09.2019	6,27	6,27	9	2019	3	0		37	427,04	5	5,72944
16.09.2019	12,76	12,76	9	2019	3	0		38	427,04	10	5,7037
23.09.2019	12,94	12,94	9	2019	3	0		39	427,04	11	5,6988
30.09.2019	12,55	12,55	9	2019	3	0		40	427,04	10	5,68778
7.10.2019	12,62	12,62	10	2019	4	0		41	435,59	11	5,81506

Şekil A.1. 40 Haftalık veri seti örneği

ÖZGEÇMİŞ

Ad-Soyad : Yunus Emre BALCI

ÖĞRENİM DURUMU:

- **Lisans :** 2019, Sakarya Üniversitesi, Mühendislik Fakültesi, Endüstri Mühendisliği
- **Yükseklisans :** Devam Ediyor, Sakarya Üniversitesi, Fen Bilimleri Enstitüsü, Endüstri

MESLEKİ DENEYİM VE ÖDÜLLER:

- 2021 yılından bu yana Kuveyttürk Katılım Bankası'nda iş geliştirme bölümünde çalışmaktadır.