IEEE Access
Multidisciplinary ⋮ Rapid Review ⋮ Open Access Journal

Received 12 December 2022, accepted 27 December 2022, date of publication 30 December 2022, date of current version 5 January 2023.

Digital Object Identifier 10.1109/ACCESS.2022.3233529

RESEARCH ARTICLE

Performance, Efficiency, and Target Setting for Bank Branches: Time Series With Automated Machine Learning

ILKER MET[1], AYFER ERKOÇ[2], AND SADI EVREN SEKER[3]
[1]Ziraat Bank, 34398 Istanbul, Turkey
[2]Ziraat Bank, 06050 Ankara, Turkey
[3]Department of Computer Science and Engineering, Antalya Bilim University, 07190 Antalya, Turkey

Corresponding author: Ilker Met (imet@ziraatbank.com.tr)

ABSTRACT Setting targets for the bank branches and distribution of annual targets to the branches and portfolio managers, quarterly, is a crucial process for decision making and strategic planning in the banking industry. Performance of the bank branches and portfolio managers are also evaluated by the quarterly divided targets to the branches and portfolio managers. In this study, the focus is on performance prediction by using state-of-art machine learning algorithms. A novel automated machine learning approach with combined algorithm selection and hyperparameter optimization is also applied for each of the branches since all the branches might have different customer segmentation and behavior. Moreover, the postconditions are executed to finalize the target calculation and distribution over the performance predictions. The study shows the success of the methodology with a successful application of 98% accuracy in the prediction and majority of branch target calculations. An end-to-end solution found to the seasonality and periodicity problem, which is the biggest problem faced by branches while achieving their goals. Also, the novel approach increases the success of branch targets by 10% in overall. The most significant innovation this study provides to the literature and practitioners is that, unlike classical studies, it solves the seasonality and periodicity problem through multiple time series modeling. The target setting procedure was employed by the largest financial institution in Turkey, Ziraat Bank, to evaluate the operating performance of its branches. The empirical study demonstrates the applicability of the proposed model in the banking sector. The outputs of the study are implemented in real life for all retail branches of Ziraat Bank. In addition, the study awarded the most innovative use of AI/ML, the most innovative project for in-house implementation related to the innovative aspect of the work, by the Global FinTech Innovation Awards 2022.

INDEX TERMS Banking, performance prediction, strategic planning, decision support systems, time series, machine learning, AutoML, artificial intelligence.

I. INTRODUCTION

Banks are the most important actors for the continuity of the financial system. The sustainability of a bank is based on profitability like for every commercial business. Therefore, performance evaluation and measurement and also effective target setting prediction are required to manage profitability and to coach bank branches. Although, there are many financial prediction studies on the literature, this study is one of the first studies applying target setting for the branches and portfolio managers in the banking industry.

The associate editor coordinating the review of this manuscript and approving it for publication was Faissal El Bouanani.

The main purpose in the present study is to conduct target setting calculations, target distribution, and performance prediction for the banking sector via analytical methods with innovative ways. The purpose of producing solutions with machine learning algorithms (AutoML, clustering methods, correlations, pre- and postcondition rules) is to ensure that the model includes the effects of variables that cannot be included in the system manually. In this way, a performance measurement system that produces a more effective and faster solution for banking is targeted. For effective performance measurement to be made, the targets set must be accurate and achievable. Undoubtedly, the heart of the target work is periodicity and seasonal effects. If these effects are not considered in the criteria for target distribution, the realization of the distributed targets becomes impossible, especially for units that are exposed to periodic effects, or growth far exceeding the target is encountered. Even though the realization of the target seems to be positive in perception, it is an indication that the process is not managed well and can disrupt the bank's general strategic targets. In the analyses performed within the scope of the targeting study, it was observed that the units that cannot achieve their targets or that exceed 100% are due to periodic effects. To find a solution to this problem, time series analysis and machine learning methods appear to be good solutions.

For many years, in the banking sector, the performance evaluation of bank branches has been a crucial metric for strategic planning and decision making [1]. The very first implementation of targeting procedures in the literature started in the 1970s with the central bank's implementations [2]. Although there are some classical approaches to performance evaluation [3], some researchers in the literature applied statistical methods like data envelopment analysis (DEA) [4]. DEA is a linear programming-based method for assessing the performance of homogeneous organizational units and is additionally used for identifying efficiency. It is a non-parametric method that utilizes linear programming to measure the level of efficiency of comparable decision-making units by employing multiple inputs and outputs. The main application areas for DEA in banking and branch performance analysis include the following: efficiency ranking, resource allocation, and efficiency trends investigation [5]. In addition to this mainstream analysis field, evaluation of the ability the branches to meet the targets established by bank management, which is studied with DEA, showed that targets can be substantially reduced without significant loss or distortion of information to bank management [6]. When this method is used with financial ratio analyses, it provides financial institutions with versatile decision-making support in performance evaluation [7]. The purpose of this joint usage is to minimize the disadvantages of the method, i.e., ignoring the effect of exogenous variables on the operation, ignoring the statistical error, not indicating a way to improve efficiency, difficulty in performing statistical tests with the results, computational difficulties depending on the number of variables, and when used for target setting not always feasible and bounded [8]. Conventional models assume non-negative inputs and outputs. On the other hand, data can take also negative values in the real-world application of target setting in the financial sector, for instance, annual growth variable. To overcome this issue, semi-oriented radial measure (SORM) approach modeling can be used with parametric or nonparametric models [9]. Another method used in efficiency measurement is financial ratio analysis (FRA). FRA is the structural and logical way to present the overall financial performance of a financial institution. Financial ratios are tools used to measure efficiency proportionally with important metrics such as profitability, liquidity, and solvency performance [10]. The method ensures mostly two variables to be compared with each other and can show the adequacy or evolution of the calculated ratios over the periods, thus the claimed single-perspective evaluation. However, the analysis is not sufficient alone for target setting or distribution problems. Financial ratios can be used as input in other parametric/nonparametric approaches [11].

Besides the nonparametric methods, parametric approaches are used for efficiency forecasting and measurement. The most common of these approaches are stochastic frontier analysis (SFA), the thick frontier approach (TFA), distribution-free approach (DFA), and economic frontier approach (EFA), in descending order. Theoretically, a production function gives the maximum potential output with a given set of inputs. This is different from its common regression counterpart, which specifies the conditional mean of output. The production function defines a limitation or "frontier", deviations from which can be interpreted as inefficiency. The econometrics of stochastic frontier analysis provides procedures for modeling the frontier notion within a regression framework so that inefficiency can be estimated.

Although the assumptions and preconditions of the methods are different, in literature reviews it is seen that the findings support each other. With frontier analyses, findings are revealed for bank managers to increase profitability and reduce costs for branches that perform efficiently and inefficiently. This type of analysis not only guides bank managers in terms of performance management but also sheds light on revealing the branching [12], [13].

As data-driven systems grow, the banking target system inevitably has to keep up with the data world. Efficiency measurement correspondingly has to keep up with this world as well. This harmony emerges with the modification of classical methods or the study of new methods. According to the new categorization of the efficiency measurement factors introduced with the growing needs and data and also regarding some previous reviews' classical methods, models can be modified to include the desired properties like slacks based measure network data envelopment analysis (SBM-NDEA). SBM-NDEA is a linear programming technique, an advanced type of linear programming analysis, and it considers the interconnections of inputs and outputs with all variables. By applying the modified model, branch efficiency scores and efficient targets can be obtained via an innovative perspective [14]. In addition to systematic approaches [15], algorithmic and mathematical methods are also encountered in the literature. These methods are general multi-criteria decision-making (MCDM) approaches [16]. MCDM techniques are appropriate tools to prioritize in a sophisticated environment and are able to rank alternatives in decision problems with conflicting criteria. MCDM's different approaches include the collection of requirements for a "good" selection approach [17].

In banking, target setting and estimation include numerous criteria due to the nature of the business, such as a loan growth target, a deposit growth target, and a credit card sales target. In this case, both the selection of the sub-variables affecting the criteria and the efficiency measurement of the criteria are critical for the correct execution of the targeting process. MCDM analyzes the criteria to determine whether each criterion is a favorable or unfavorable choice for a particular application [18]. It also attempts to compare this criterion, based on the selected criteria, against every other available option in an attempt to assist the decision maker in selecting an option with minimal compromise and maximum advantages. The criteria used in the analyses of these criteria can be either qualitative or quantitative [19], [20].

In more recent years, algorithm-based approaches have taken their place on the stage as an indispensable part of performance management, as in many other automation processes. Strategizing with artificial intelligence (AI) and machine learning is as valuable as discovering and exploiting strategic opportunities. AI and machine learning techniques can help determine which results in performance management processes, how to measure, how to distribute the target, and what the potential of the measured unit is.

Time series analysis techniques are quite new to rely on the machine learning domain for banking applications. In addition to classical time-series estimation methods, artificial intelligence and machine learning methods have also been used in the financial field. A comparison of time-series prediction using traditional and machine-learning methods revealed that machine-learning methods yield better results than conventional prediction methods [21]. Model selection strategies play an important role in time series prediction because the number of harmonic or random components required to account for the complexity of seasonality can be very large [22]. Automatic machine learning models come to the fore in solving this problem recent years. Financial time series prediction is a hot topic in the machine learning field and can be used for seasonal-trend decomposition based on loess as a preprocessing technology [23]. An algorithm suitable for a time series can provide strong predictive results depending on both the performance of the algorithm and features of the time series. To use time-series algorithms in terms of application, examining the time-series features together will increase the predictive power of the studies [24].

In target setting, predictability is as much an important part of banking strategy as efficiency measurement is. Furthermore, guidance on how to improve branch performance is critical. In a study on branch efficiency measurement, artificial neural networks (ANNs) were used. The results of the two methodologies (DEA via ANNs) were comparable. ANNs are also applicable to short-term efficiency prediction [25].

In recent years, artificial intelligence-based fuzzy clustering, the adaptive neuro-fuzzy inference system (ANFIS) and ordered weighted averaging (OWA) methods have been frequently encountered in the literature for solving problems that require complex prediction, especially for issues such as targeting and efficiency in banking. The ANFIS model has the advantage of being both numerical and grammatical; it is more transparent and causes fewer memorization errors [26]. Fuzzy clustering method, on the other hand, is very convenient for clustering customers efficiently and especially for preparing offers [27]. OWA is a method based on a new aggregation technique used to determine the weights of the inputs used in estimation methods [28]. Especially The nonlinear relationships among selection criteria greatly impact the decision-making process. OWA and probabilistic OWA is encountered as an effective method for criterion selection and weighting [29].OWA also has different usage areas in the financial field and can be used frequently, particularly in decision-making processes [30]. It is also possible to come across the use fuzzy methods in applications as a risk-factor-reducing effect in the selection of intermittent numbers and more complex risky variables [31].

Machine learning algorithms can be used alone or together with classical approaches. One criterion that banks should consider in target-setting problems is non-performing loans (NPLs). NPLs are a critical constituent that impacts the operational performance of banks. A rising level of risk leads to poor operational performance. Accordingly, a hybrid approach combining with ANN is developed to measure and predict the operational efficiency scores of banks [32].

Efficiency measurement at branch level is critical both to reveal potential areas and provide data to model target criteria. One of the major difficulties in the banking domain is that each branch has its own customer behavior and it is almost impossible to apply one size fits all solutions to all branches. The solution proposed here in is to use an automated machine learning (AutoML) approach to solve the combined algorithms selection and hyperparameter optimization (CASH) problem [33]. In addition, each branch has its own unique algorithm with optimized hyperparameters and the best data preprocessing applications.

The novelty of the study is based on three development at the same time. 1) The first-time successful application of a target setting problem in the banking industry with 98%. 2) with the novelty of the data set, which includes thousands of branches and portfolio managers. 3) with AutoML application for different customer behavior and segmentations. Also, the methodology is applied for 3 quarters and replaced the classical applications in target setting and distribution of annual target to the branches and portfolio managers. We believe, above three novelty points will bring a new baseline in the literature by this study. In addition, the study won the IBS Intelligence Global FinTech Innovation Awards 2022 for the most innovative use of AI/ML: the most innovative project for in-house implementation–an international award program related to the innovative aspect of the work.

II. PERFORMANCE PREDICTION AND TARGET CALCULATION FOR BANK BRANCHES (PROBLEM DEFINITION)

Organizations commonly set targets with definitive and quantitative reference points for their branches that reflect top management aspirations for these units [34]. To obtain measurable success, it is vital to establish authentic targeting systematically. These determined objectives must be realistic and data-based. Performance measurement techniques should contain data that can both monitor previous performance and plan future performance. It is aimed to reveal the effective units and to involve the less efficient units in the system by considering their potential. Currently, performance management systems need to be rearranged in parallel with technological evolutions and changes in economic conditions. It emerges that these regulations should also include issues such as creating high added value, flexibility and adapting to the market, process management in harmony with the time, preventing obliteration, and increasing the motivation and performance of employees. High targets that do not reflect the potential and are not compatible with productivity measurements cause the branches to withdraw from the race at the beginning of the target period with the belief that they will not be able to achieve the target. The need for regeneration of the system originated mostly to solve this problem.

The fact that the system is accurate and based on data analysis increases the motivation of the employees and the possibility of realizing the targets. To get the best benefit the system structure should be objective and explainable. The second major problem is that the system needs to be renewed and the method should reflect seasonality and trends. Attempts have been made to solve this problem by the integration of time series analysis. For the targets to be realistic, the periods in which periodic payments and activities such as agricultural harvest periods, tourism periods, official holidays, or the dates that cause an increase or decrease due to seasonal effects should be considered in the target modeling system.

Ziraat Bank has 1,639 retail branches and over 8,000 portfolio managers. Targets are set for branches and portfolio managers over a total of 18 criteria in quarterly periods, and measured at the end of the term according to the success in achieving the target. Considering the size of the bank and the number of target criteria, it has been determined that classical methods are insufficient in terms of time planning for processing big data. Goals are distributed for 3-month periods, immediately after the completion of each term. Therefore, it is critical for sales and marketing planning to be communicated to the branches and portfolio managers about the next term' targets in the first week of the new period. With these requirements, the existing classical system was reviewed and attempts were made to produce an effective resolution via a new approach to the banking target distribution system.

The present study mainly focuses on calculation of the performance predictions for each branch of the bank. The performance of previous months is already stored in the data warehouse and the historical data show time series behavior. The performance predictions are important for the bank for the quarter of a year, but the data are collected on a monthly basis. To incorporate the branches' location potentials in the modeling study, external data such as monthly banking data from the Banking Regulation and Supervision Agency (BDDK), insurance data from Turkey's Pension Monitoring Center (EGM), and growth data from the Turkish Statistical Institute (TUİK) are used as well as the banking data.

After modeling, the postcondition rules are used to align the forecasts with the bank's strategies. By combining external and internal data and alignment with conditions, targets are created transparently and completely data-based.

After the models were tested in bank branches for a year (4 performance periods) and their realization rates of success were observed, system automation studies were initiated.

III. DATASET

The historical dataset of all branches and portfolios is collected for 5 years with monthly resolution. To see seasonality and trends, at least 5 years of data have been studied. In addition to the data produced by the branches and portfolios at the bank, non-banking data sources are also used to include the location potential in the modeling. Ensuring compatibility with the bank data, the periodically generated external resource data were converted into monthly frequency by the interpolation method. Variables produced from internal bank data and external source data are given in Table 1. External data: income per capita gained from the TUİK [35], sector ratio for loan and debits and banking products gained from the BDDK [36], location protentional data gained from the Banks Association of Turkey (TBB) [37] and Credit Record Bureau (KKB) [38], and private pension insurance data gained from the EGM [39] respectively.

IV. METHODOLOGY

The flow of the process in the present study follows the cross-industry standard processing for data mining (CRISP-DM) [40] steps. Besides the business understanding and data understanding steps already discussed in the previous sections, this section mainly focuses on the data preprocessing and modeling steps. The steps of CRISP-DM are shown in Figure 1.

The journey of the data starts with the data connection layer. During the data connection, two types of data are gathered for the calculations. External data are mainly gathered from national institutes about the bank branches' performances. Internal data, on the other hand, are gathered from the data warehouse and databases of the bank.

After the data connection layer, the preconditions are executed on the dataset. Dirty/noisy data, missing data, or any data which does not obey the predefined rules are filtered, imputed, or interpolated during this step. For example, some of the critical data for some branches do not exist for 5 years, because the branch has moved or opened less than 5 years ago. This problem is solved by using linear interpolation [41] or k-nearest neighbor [42] approaches with $k = 5$ and the distance metrics are built on time series features extracted from the TSFEL library with Euclidean distance.

Figure 2 demonstrates the overall flow of data from internal datasets on branch DWH (Data Warehouse) to time series analysis. Although, the major data are flowing from branches DWH, time series analysis also includes the external data

TABLE 1. Data and data properties.

Variable	Explanation	Measurement Level
growth rate	growth data for branches and portfolios	monthly, internal data, continuous
annual growth	annual growth, calculated with rolling mean approaches for relevant months	monthly, internal data, continuous
last term growth	at least 63-month relevant performance term growth	periodic, internal data, continuous
concentrated customer extension	customer trade with large balances extension balances in relevant month	monthly, internal data, continuous monthly, internal data, continuous
seasonality	seasonality effect in relevant term, calculated with autoTS	monthly, internal data, continuous
scale effect	size status according to the scale group in which branch or portfolio located	monthly, internal data, continuous
average balance	average balance growth amount for units	monthly, internal data, continuous
expectation	unit's own expectation for increase or decrease for next performance term	periodic, internal data, categorical ("low", "middle", "high")
waived expense amount	discounted amounts (bank charges) that affect the profitability of the branches and portfolios	monthly, internal data, continuous
GDP	income per capita	periodic, external data, continuous
loan customer	ratio of customers using loans in the province where the branch is located, gained from BRSA's FINTURK data and from Turkey Credit Bureau' database GEOMİS.	periodic, external data, continuous
deposit customer	ratio of customers that have deposited in the province where the branch is located, gained from BRSA's FINTURK data and from Turkey Credit Bureau' application GEOMİS.	periodic, external data, continuous
cc customer	number of customers using credit card	monthly, internal data, continuous
ncc customer	number of customers not having credit card	monthly, internal data, continuous
product customer	customer or potential customers using the banking product in the relevant target criteria	monthly, internal data, continuous
financial transaction	number of financial transactions	monthly, internal data, continuous
private pension insurance	number of participating customers	periodic, external data, continuous
private pension insurance amount	number of private pension participants' province share	periodic, external data, continuous
gross non-performing loans	non-performing loan amount of the relevant unit	monthly, internal data, continuous

TABLE 1. *(Continued.)* Data and data properties.

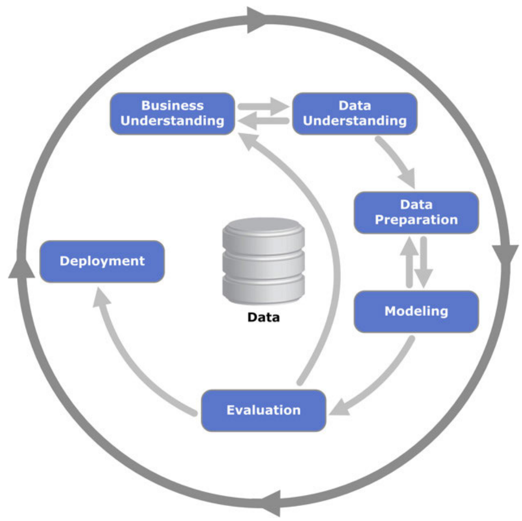

FIGURE 1. Cross-industry standard processing for data mining (CRISP-DM) [40].

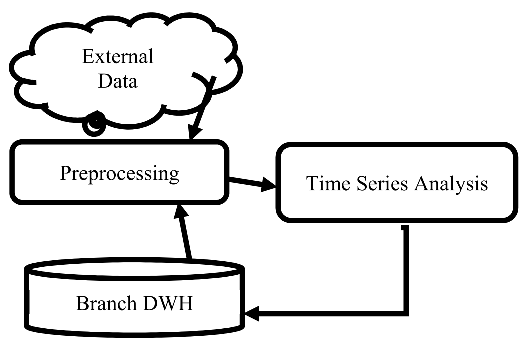

FIGURE 2. Flow of the data process and three major steps of the methodology.

for the higher success rates. The outcome of time series analysis is also saved back to the branch DWH for the further processes and other data analytics operations in the bank can benefit from these time series analysis outputs.

The time series analysis in Figure 3 is a compound step with 3 major sub-steps. These sub-steps can be demonstrated as in Figure 3.

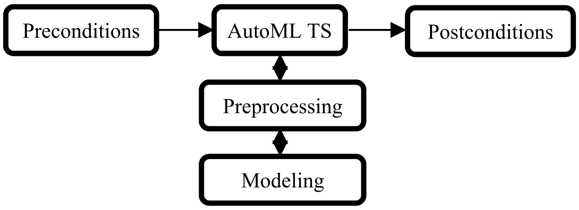

FIGURE 3. **Overview flow of the data process and three major steps of the methodology.**

Figure 3 demonstrates the overview of the data processing system created for the solution. The methodology starts with checking the preconditions and flow of the data to the AutoML time series process. The AutoML time-series process includes two sub-processes, the data preprocessing step for data manipulation and feature engineering, and the modeling phase for the combination and selection of best algorithm alternatives. After the completion of the AutoML time series phase, the postconditions are executed as demonstrated in Figure 3. The data processing system starts with execution of preconditions like rarity, volatility, or missing value checks. If any problem is detected for the rules in the preconditions, this branch is labeled a problem branch. The list of problem branches is handled at the final phase of postconditions but until the postconditions a branch can be labeled a problem branch during the precondition or time series analysis. Briefly, the precondition phase is for detecting the problematic branches with some predefined rules on the missing value, rarity, or volatility checks. For example, most of the problems might occur because some of the branches are newly launched or hold some customers transferred from other branches. Another example for the external data is the expert view, which is a data collection from external sources and predefined rules and limits for the branches are executed in this step. In the preconditions phase, the rarity, uniqueness, missing value imputation, outlier detection, and cleaning operations are handled.

During the AutoML TS process in Figure 3, the extraction of time series features, solving the CASH problem, and finding the best combined algorithm with the best hyperparameters for the time series are handled. Finally, in the postconditions step in Figure 3, the predefined rules from business analysis are applied to the data collected from the outcome of the time series.

A. TIME SERIES PREDICTION WITH AUTOML

Historical data for 5 years with the monthly resolution are represented in Figure 4.

In Figure 4, the average of all branches is decomposed into seasonal, trend, and residue data. During the decomposition Figure 4 is calculated by STL (seasonal and trend decomposition using Loess) method [43], and Figure 5 is calculated by the naive method. For both approaches, the additive parameters are analyzed.

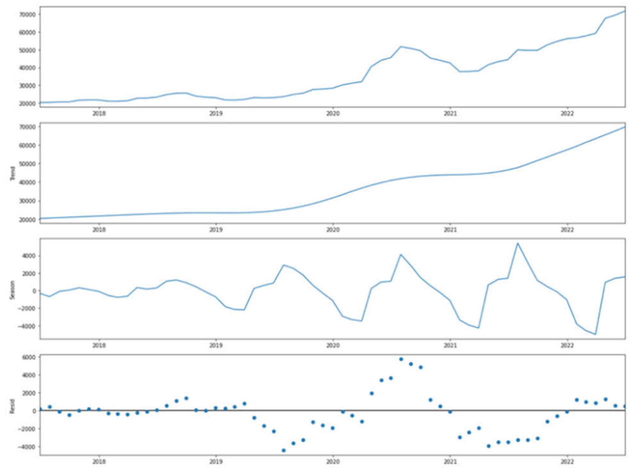

FIGURE 4. **Time series feature decomposition into seasonal, trend, and residue for 5 years on average of branches. The STL method.**

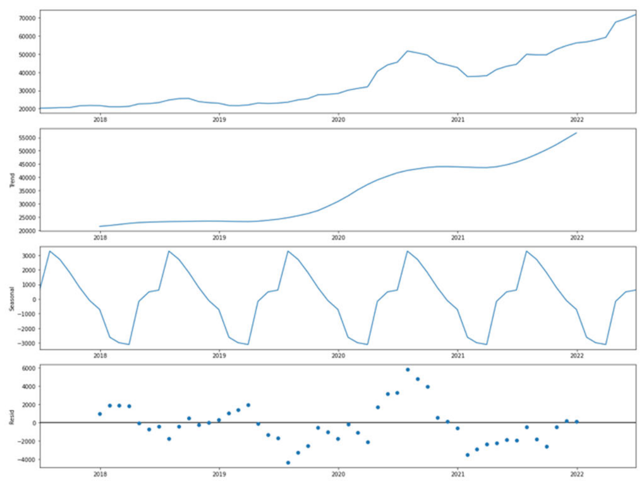

FIGURE 5. **Time series feature decomposition into seasonal, trend, and residue for 5 years on average of branches. The naive method.**

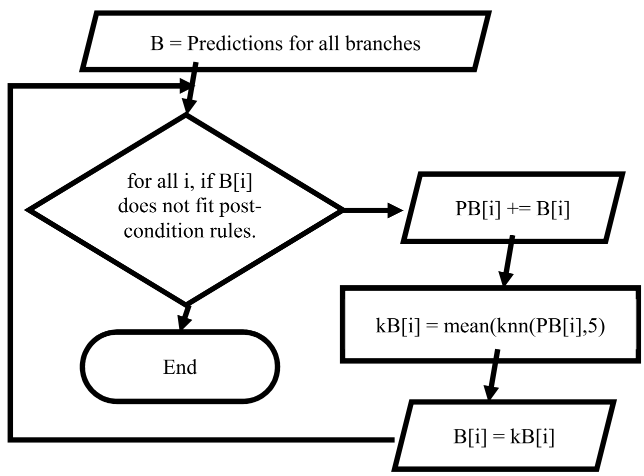

FIGURE 6. **Flow chart of Algorithm 1: KNN solution for the branches with problems.**

The dataset of the past performance for any branch is a time series with data points indexed in time order on a monthly basis. Although the monthly frequency has some problems

TABLE 2. Time series feature extraction.

Feature Domain	Feature	Meaning
Temporal	Autocorrelation (L)	Autocorrelation of time series with the lag L.
Temporal	Centroid	Gravity Center of the Time Series
Temporal	Mean Absolute Difference	mean($\|\Delta s\|$)
Temporal	Mean Difference	mean(Δs)
Temporal	Sum of Absolute Distances	sum($\|\Delta s\|$)
Temporal	Entropy	sum($P(x)\log_2 P(x)$)
Temporal	Peak to peak distance	$\|$ max(s) – min(s) $\|$
Temporal	AUC	Area Under Curve
Temporal	Maximum Peaks	count number of maximum peaks
Temporal	Slope / Trend	Linear Regression with Ordinary Least Squares
Statistical	Interquartile Range	Q3 – Q1, where Q3 is third and Q1 is the first quartile.
Statistical	RMS	Root Mean Square, root of mean of square of each data point in time series.
Statistical	Kurtosis	
Statistical	Skewness	
Statistical	ECDF	Empirical Cumulative Distribution Function

s = time series data set, s_i = i^{th} member of the series; Q = Quartile

with non-equal spaced points in time, time-series features such as seasonality, trend, force, or torque are extracted and evaluated before the AutoML process.

During the present study, two time series extraction libraries are implemented on the data: TSFEL [44] and TSFRESH [45].

The extracted time series are given in Table 2.

B. AUTOMATED MACHINE LEARNING

During the present study, the AutoML approach was applied to find the best algorithm combinations and also the pre-processing of the dataset. The application of AutoML for the performance prediction is a special time, which can be considered as a time series analysis using the machine learning algorithms in the temporal domain. The combination of AutoML with time series is called automated time series analysis in the literature and there are some already implemented libraries for the purpose, such as project AutoTS [46], and project FEDOT from the NSS Team at ITMO University [47]. Automated machine learning steps include the solution for the CASH problem, so the algorithms listed below are selected, combined, and hyperparameter optimized for each branch of the bank. The result is a combination of the best algorithms with the best parameters for each branch. The list of algorithms for this study is given in Table 3.

The 18 algorithms given in Table 3 are included in the project from 7 different libraries and research projects. From the AutoTS library, the algorithms tested are also added from stats models [48], Facebook Prophet [49], TensorFlow [50], Greykite from LinkedIn [51], and Gluon time series

TABLE 3. List of algorithms in AutoML phase.

Name of Algorithm	Library	Detail
Last Value Naive	AutoTS	Naive forecasting predicting a dataframe of the last series value
Seasonal Naive	AutoTS	Naive forecasting predicting a dataframe with seasonal (lag) forecasts.
Average Value Naive	AutoTS	Naive forecasting predicting a dataframe of the series' median values
Zeroes Naive	AutoTS	Naive forecasting predicting a dataframe of zeroes (0's)
GLS	Stats Models	Simple linear regression from Stats Models
ETS	Stats Models	Exponential Smoothing from Stats Models
GLM	Stats Models	Simple linear regression from Stats Models
ARIMA	Stats Models	ARIMA (AutoRegressive Integrated Moving Average)
Unobserved Components	Stats Models	Unobserved Components from Stats Models.
Datepart Regression	AutoTS	Regression not on series but datetime
Rolling Regression	Sci-Kit Learn	General regression-framed approach to forecasting using sklearn.
Motif Simulation	AutoTS	More dark magic created by the evil mastermind of this project. Basically a highly-customized KNN
FBProphet	Facebook	Facebook's Prophet
TensorFlow STS	TensorFlow	STS from TensorFlow Probability.
TFP Regression	TensorFlow	TensorFlow Probability regression.
Greykite	LinkedIn Greykite	Greykite
GluonTS	Gluon Time Series	Gluon Time Series (GluonTS) is the Gluon toolkit for probabilistic time series modeling, focusing on deep learning-based models.
VECM	Stats Models	VECM from Stats Models

library [52] and a different statistical modeling approach for time series is also tried for comparing results [53].

The algorithms have different advantages and, in this project, all the algorithms from different libraries are tested, fine-tuned, and ensembled. The ensemble process also includes more than 1 algorithm at a time and the statistical details of successful algorithms are given in Table 4.

Although the AutoTS library includes multiple ensemble techniques such as MOSAIC (MultiObjective Genetic Sampling for Imbalanced Classification) [54] ensembling or ensembling per series [55], our study is limited to the BestN ensemble technique. This technique obtains the average of the best algorithm predictions.

C. POSTCONDITIONS

Execution of machine learning algorithms, even with the AutoML approach, involves some risks that might be

TABLE 4. Statistical details of algorithms and ensembles.

Successful Algorithms	Number of Occurrences
Ensemble	1.045
Seasonal Naive	111
Last Value Naive	87
ARIMA	86
Unobserved Components	56
Average Value Naive	56
Zeroes Naive	55
Rolling Regression	41
Date Part Regression	32
GLM	32
ETS	22
GLS	16

catastrophic for the banking industry. Although all the safety checks are preliminarily executed before the machine learning phase, there is always a risk of outputting an impossible target for the branch or target calculation for underperformance. The results of time series modeling can produce negative results. Since a negative size cannot be distributed to the units as a target, identifying units with similar trends at these points where the estimations do not produce the correct result and including the results of clustering and correlation analysis in the business rules eliminate these errors.

Moreover, manual target calculation of a banking branch is somewhat complicated, including qualitative prediction such as the Delphi method [56], which covers the opinions of banking authorities and the banking limitations. For example, from a pure time series prediction, a branch target can also be an increase of 200%, which is possible statistically but impossible from the business analysis perspective. The application of the business analysis perspective is added to this final step of postconditions. The one major parameter is limiting the upper and lower increase and decrease on the targets to 10%. Further, some of the branches are not suitable for time-series predictions, and the majority of the algorithms return errors. The reason for errors is sometimes missing data or sometimes the period/cycles depending on the algorithms. The solution for these branches is handling the error both from a business perspective and statistical perspective by the KNN approach. In the AutoML phase, we try to find the predictions by automated machine learning and we handle the errors, if any, in the postcondition phase.

Error handling is solved by the KNN algorithm. The KNN algorithm is built on the distance calculation and to find the closest branches to the branches with errors we use time-series features. The time-series features are first converted to numerical values and normalized for distance calculation; then the KNN algorithm is utilized to calculate the closest branches to the branches with errors. The error handling returns the average of k-nearest neighbor predictions for each branch. In addition, the average value of branches is meaningless, and we convert the percentage of increase or decrease for the average.

Algorithm 1 Handling of Time Series Predictions by KNN

Input: List of predictions and time-series features from the branches with problems as *probBranch*, List of predictions and time-series features from the branches without problems as *predBranches,* Number of neighbors as *k*

Output: List of corrected predictions for the branches with problems as *resultBranches*

1: Convert branch predictions to increase/decrease ratios by: predBranches[‘prediction’] = predBranches[‘prediction’] / predBranches[‘prediction’

2: for i in *probBranch*:

3: listofNeighbors = KNN(*k*, *i, predBranches*)

4: t = average(listofNeighbors[‘predictions’])

5: *resultBranches*.add(i,t)

6: return *resultBranches*

The major challenge in the present study is to be able to access the right data, especially in large-scale banks or businesses, and to be able to model these data correctly. For example, to estimate the target of 1,639 branches and approximately 8,000 portfolios at the Ziraat bank scale:

- To obtain 5 years of historical (month frequency) data for each unit,
- Adding external data that affect each criterion and determining these datasets (16 different external datasets),
- Running models for each criterion and unit (16 criteria in total),
- Making time plans so that the target works can be estimated in a certain time interval,
- Comparing forecast results with realization results and measuring success and employee satisfaction are both challenges faced and critical success factors.

Another major challenge in the project is the number of branches and the dynamic structure of the banking sector. Calculation of target branches and portfolios requires a data science expert power to handle all the branches one by one and updating the models for each quarter increases the required number of analyses. As a solution, one major impact is utilizing the AutoML approaches.

V. RESULTS

The present study demonstrates the automation of data analytics and data science processes for automatically setting performance targets for branches. In the last step before the development, the project outcomes are evaluated and the success of the project is investigated. The outcomes are found satisfactory by the business units of the bank.

Table 5 shows the machine learning outcomes for the 1,639 branches of the bank.

In the present study, a solution was produced by calculating the periodic effects for 1,639 branches of Ziraat Bank and combining them with other criteria, again with machine learning methods. The purpose of using 5-year data is to capture the behavior of the branch for at least 5 periods when evaluated on a quarterly basis.

TABLE 5. Autots trend results.

Trend	Number of Occurrences
increasing trend	1,539
decreasing trend	83
stable	17

TABLE 6. Example of trend Data for two branches.

DATE	BRANCH A	BRANCH B
30.06.2017	69,029.66	32,218.87
31.07.2017	*70,434.54*	*32,476.98*
31.08.2017	*71,603.06*	*32,575.40*
30.09.2017	*72,571.40*	*32,543.15*
30.06.2018	79,280.43	34,993.77
31.07.2018	*84,727.19*	*35,062.88*
31.08.2018	*89,299.59*	*35,143.01*
30.09.2018	*89,800.98*	*35,587.89*
30.06.2019	111,786.22	43,581.19
31.07.2019	*114,384.17*	*42,788.42*
31.08.2019	*115,806.58*	*42,418.10*
30.09.2019	*118,131.73*	*40,610.54*
30.06.2020	118,142.02	49,387.79
31.07.2020	*122,685.95*	*52,130.07*
31.08.2020	*123,948.53*	*51,955.36*
30.09.2020	*125,107.36*	*50,036.77*
30.06.2021	130,970.80	50,324.71
31.07.2021	*132,459.82*	*50,735.23*
31.08.2021	*134,495.94*	*50,417.95*
30.09.2021	*134,183.50*	*50,736.28*

TABLE 7. Example of target data for two branches.

	Branch A		Branch B	
DATE	Target	Realization	Target	Realization
30.09.2017	*18%*	*5%*	*10%*	*1%*
31.09.2018	*15%*	*15%*	*9%*	*1%*
31.09.2019	*22%*	*5%*	*10%*	*2%*
30.09.2020	***55%***	***5%***	***5%***	***5%***
30.09.2021	***5%***	***5%***	***5%***	***7%***
31.09.2022	***11%***	***12%***	***3%***	***1%***

For example, when the increase trends in the 3rd quarter periods of the two branches whose five-year data are monitored in Table 6, it is obvious that there is a seasonality effect in the B branch. While the growth rate of the A branch is higher, the B branch exhibits more stable behavior.

The targets of branches A and B are shown in Table 7 by periods.

For the deposit target criteria of Branches A and B, manual target estimation until 2020 and estimations made by machine learning methods after 2020 are given in the table. Forecasting studies conducted after 2020 include both unit-based modeling and seasonality and periodicity effects. Branch B is normally located where agricultural production is conducted, and it is a branch with a very low growth expectation in the deposit criteria in the 3rd quarter. In Branch B, the need for loans increases in the 3rd quarter of the year, but the growth of deposits remains limited. When target distribution is performed for this branch without considering the seasonal effects, it becomes impossible to meet its targets and the general strategy of the bank cannot be achieved. However, when the results of the modeling studies are examined, it is seen that more accurate estimates are made and the success is higher.

AutoML results suggest an increase of around 10% for Branch A and a decreasing trend of around 2% for Branch B.

When the results of machine learning are combined with the sector share of the branch, its annual development, potential data, the opinion of the Regional Presidency, the recent growth, and scale effect, the growth target of Branch A in the predicted quarter is 15% and that of Branch B is 3% (due to potential effects). It has been observed that 87% of the targets deployed with machine learning are compatible with branch growth.

Using time series for periodicity and seasonality detection and working with more than one model on each unit is a subject that has not been studied before.

The method was tried for the first time and 98% success was achieved. Model diversity has been ensured and each branch has been enabled to move in its potential pattern. The study is easily applicable to the banking sector. The basis of the study consists of historical data and business rules. The datasets and the rules determined by the banks during target distribution can be easily integrated into the modeling work and used in the financial sector.

VI. CONCLUSION

Machine learning methods are used to obtain consistent results from the data and to reveal the periodicity results. To be able to make accurate target estimations of the units, it is determined which units were affected by seasonality in which periods utilizing autoTS estimations by taking the annual data on a monthly frequency in the 2017-2022 period. For a total of 16 different criteria, 13 different time series models are run from the autoTS library and the success percentages of the models are calculated. When the model results are examined for 1 year, it is observed that 20% of the branches in each period are under the influence of periodicity and seasonality. For each criterion, the best estimator from 13 time series models, growth data, external data, and business unit rules are added and approximately 10 variables are used as inputs per criterion, and quarterly target estimates are obtained with a 3-stage modeling structure. With our analytical study, the manual operational work lasting approximately 15 days with Excel in the target working departments was transferred to the automated machine learning system, and it was ensured that it worked with big data that could not be calculated manually, because of limited human resources, sensitivity of data, and the limited time frame for the results. The positive feedback received from the targeted branches increased the belief that the system makes a fair distribution. Thanks to the system established, data-based decision systems and business processes have contributed to the harmonization of the future. Besides these benefits, owing to the machine learning algorithm's multiple model selection methods, the selection of the model that best predicts the periodicity is performed automatically. The issue of periodicity, which should be considered in target distribution projects, has been studied for the first time with multiple time series models. The fact that periodicity has been tried via a new approach with machine learning methods and tested in the real environment will guide both researchers and practitioners for future studies.

REFERENCES

[1] I. Met, G. Tunali, A. Erkoç, S. Tanrikulu, and M. Ö. Dolgun, "Branch efficiency and location forecasting: Application of Ziraat bank," *J. Appl. Finance Banking*, vol. 7, no. 4, pp. 1–13, Jun. 2017.

[2] G. H. Sellon, "Monetary targets and inflation: The Canadian experience," *Federal Reserve Bank Kansas City Econ. Rev.*, vol. 67, pp. 16–31, Jan. 1982.

[3] A. Shaout and M. K. Yousif, "Performance evaluation—Methods and techniques survey," *Int. J. Comput. Inf. Technol.*, vol. 3, no. 5, pp. 966–979, Sep. 2014.

[4] S. Kaffash and M. Marra, "Data envelopment analysis in financial services: A citations network analysis of banks, insurance companies and money market funds," *Ann. Oper. Res.*, vol. 253, no. 1, pp. 307–344, Jun. 2017.

[5] J. C. Paradi, Z. Yang, and H. Zhu, "Assessing bank and bank branch performance," in *Handbook on Data Envelopment Analysis* (International Series in Operations Research & Management Science), vol. 164. Boston, MA, USA: Springer, 2011, pp. 315–361.

[6] C. A. K. Lovell and J. T. Pastor, "Target setting: An application to a bank branch network," *Eur. J. Oper. Res.*, vol. 98, no. 2, pp. 290–299, Apr. 1997.

[7] Q.-J. Yeh, "The application of data envelopment analysis in conjunction with financial ratios for bank performance evaluation," *J. Oper. Res. Soc.*, vol. 47, no. 8, p. 980, Aug. 1996.

[8] I. A. Ali and C. S. Lerme, "Comparative advantage and disadvantage in DEA," *Ann. Oper. Res.*, vol. 73, pp. 215–232, Oct. 1997.

[9] R. K. Matin, G. R. Amin, and A. Emrouznejad, "A modified semi-oriented radial measure for target setting with negative data," *Measurement*, vol. 54, pp. 152–158, Aug. 2014.

[10] I. Aminul, "An analysis of the financial performance of national bank limited using financial ratio," *J. Behavioural Econ., Finance, Entrepreneurship, Accounting Transp.*, vol. 2, no.5, pp. 121–129, Nov. 2014.

[11] N. K. Avkiran, "Association of DEA super-efficiency estimates with financial ratios: Investigating the case for Chinese banks," *Omega*, vol. 39, no. 3, pp. 323–334, Jun. 2011.

[12] A. N. Berger, J. H. Leusner, and J. J. Mingo, "The efficiency of bank branches," *J. Monetary Econ.*, vol. 40, no. 1, pp. 141–162, Sep. 1997.

[13] D. I. Giokas, "Assessing the efficiency in operations of a large Greek bank branch network adopting different economic behaviors," *Econ. Model.*, vol. 25, no. 3, pp. 559–574, May 2008.

[14] F. Boloori and J. Pourmahmoud, "A modified SBM-NDEA approach for the efficiency measurement in bank branches," *Oper. Res.*, vol. 16, no. 2, pp. 301–326, Jul. 2016.

[15] N. Ahmad, A. Naveed, S. Ahmad, and I. Butt, "Banking sector performance, profitability, and efficiency: A citation-based systematic literature review," *J. Econ. Surveys*, vol. 34, no. 1, pp. 185–218, Feb. 2020, doi: 10.1111/JOES.12346.

[16] A. Mardani, A. Jusoh, K. M. D. Nor, Z. Khalifah, N. Zakwan, and A. Valipour, "Multiple criteria decision-making techniques and their applications—A review of the literature from 2000 to 2014," *Econ. Res.*, vol. 28, no. 1, pp. 516–571, Sep. 2015.

[17] E. Kornyshova and C. Salinesi, "MCDM techniques selection approaches: State of the art," in *Proc. IEEE Symp. Comput. Intell. Multi-Criteria Decis.-Making*, Apr. 2007, pp. 22–29, doi: 10.1109/MCDM.2007.369412.

[18] G. Kou, Y. Lu, Y. Peng, and Y. Shi, "Evaluation of classification algorithms using MCDM and rank correlation," *Int. J. Inf. Technol. Decis. Making*, vol. 11, no. 1, pp. 197–225, Jan. 2012, doi: 10.1142/S0219622012500095.

[19] S. Opricovic and G. H. Tzeng, "Comparing DEA and MCDM method," in *Multi-Objective Programming and Goal Programming*. Berlin, Germany: Springer, 2013, doi: 10.1007/978-3-540-36510-5_32.

[20] A. Yousefi and A. Hadi-Vencheh, "Selecting six sigma projects: MCDM or DEA?" *J. Model. Manage.*, vol. 11, no. 1, pp. 309–325, Feb. 2016, doi: 10.1108/JM2-05-2014-0036.

[21] W. Hussain, J. Merigó, M. Raza, and H. Gao, "A new QoS prediction model using hybrid IOWA-ANFIS with fuzzy C-means, subtractive clustering and grid partitioning," *Inf. Sci.*, vol. 584, pp. 280–300, Jan. 2022.

[22] T. Proietti and D. J. Pedregal, "Seasonality in high frequency time series," *Econometrics Statist.*, vol. 2022, pp. 1–33, Feb. 2022.

[23] H. He, S. Gao, T. Jin, S. Sato, and X. Zhang, "A seasonal-trend decomposition-based dendritic neuron model for financial time series prediction," *Appl. Soft Comput.*, vol. 108, Sep. 2021, Art. no. 107488.

[24] G. Petelin, G. Cenikj, and T. Eftimov, "Towards understanding the importance of time-series features in automated algorithm performance prediction," *Expert Syst. Appl.*, vol. 213, Mar. 2023, Art. no. 119023.

[25] D. Wu, Z. Yang, and L. Liang, "Using DEA-neural network approach to evaluate branch efficiency of a large Canadian bank," *Expert Syst. Appl.*, vol. 31, no. 1, pp. 108–115, Jul. 2006.

[26] J.-S. Wang and C.-X. Ning, "ANFIS based time series prediction method of bank cash flow optimized by adaptive population activity PSO algorithm," *Information*, vol. 6, no. 3, pp. 300–313, Jun. 2015.

[27] P. Alam, D. Booth, K. Lee, and T. Thordarson, "The use of fuzzy clustering algorithm and self-organizing neural networks for identifying potentially failing banks: An experimental study," *Expert Syst. Appl.*, vol. 18, no. 3, pp. 185–199, Apr. 2000.

[28] J. Kacprzyk, H. Nurmi, and S. Zadrożny, "The role of the OWA operators as a unification tool for the representation of collective choice sets," in *Recent Developments in the Ordered Weighted Averaging Operators: Theory and Practice*. Berlin, Germany: Springer, 2011, pp. 149–166.

[29] W. Hussain, J. M. Merigo, H. Gao, A. M. Alkalbani, and F. A. Rabhi, "Integrated AHP-IOWA, POWA framework for ideal cloud provider selection and optimum resource management," *IEEE Trans. Services Comput.*, early access, Nov. 4, 2021, doi: 10.1109/TSC.2021.3124885.

[30] E. Leon-Castro, F. Blanco-Mesa, W. Hussain, M. Flores-Sosa, and L. A. Perez-Arellano, "Tax revenue measurement using OWA operators," *Cybern. Syst.*, vol. 2022, pp. 1–15, Aug. 2022.

[31] W. Hussain, M. R. Raza, M. A. Jan, J. M. Merigo, and H. Gao, "Cloud risk management with OWA-LSTM and fuzzy linguistic decision making," *IEEE Trans. Fuzzy Syst.*, vol. 30, no. 11, pp. 4657–4666, Nov. 2022, doi: 10.1109/TFUZZ.2022.3157951.

[32] S. Roy, "Application of hybrid approach in banking system: An undesirable operational performance modelling," *Global Bus. Rev.*, vol. 2021, Aug. 2021, Art. no. 097215092110267, doi: 10.1177/09721509211026789.

[33] L. Kotthoff, C. Thornton, H. H. Hoos, F. Hutter, and K. Leyton-Brown, "Auto-WEKA: Automatic model selection and hyperparameter optimization in WEKA," in *Automated Machine Learning*. Cham, Switzerland: Springer, 2019, pp. 81–95.

[34] C. Aranda, J. Arellano, and A. Davila, "Organizational learning in target setting," *Acad. Manage. J.*, vol. 60, no. 3, pp. 1189–1211, Jun. 2017.

[35] (2022). *Turkish Statistical Ýnstitute*. Accessed: May 2, 2022. [Online]. Available: https://data.tuik.gov.tr/Search/Search?text=GDP&dil=2

[36] (2022). *Banking Regulation and Supervision Agency*. Accessed: May 2, 2022. [Online]. Available: https://www.bddk.org.tr/bultenfinturk

[37] (2022). *Banks Association of Turkey*. Accessed: May 2, 2022. [Online]. Available: https://verisistemi.tbb.org.tr/

[38] (2022). *Credit Record Bureau*. Accessed: May 2, 2022. [Online]. Available: https://www.kkb.com.tr/en

[39] (2022). *Pension Monitoring Center*. Accessed: May 2, 2022. [Online]. Available: https://www.egm.org.tr/data-center/statistics/ips-statistics/summary-ips-data/

[40] W. Rudiger and J. Hipp, "CRISP-DM: Towards a standard process model for data mining," in *Proc. 4th Int. Conf. Practical Appl. Knowl. Discovery Data Mining* vol. 1, 2000, pp. 29–39.

[41] P. J. Davis, *Interpolation and Approximation*. Chelmsford, MA, USA: Courier Corp., 1975.

[42] H. Samet, "K-nearest neighbor finding using MaxNearestDist," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 30, no. 2, pp. 243–252, Feb. 2008.

[43] R. B. Cleveland, W. S. Cleveland, J. E. McRae, and I. Terpenning, "STL: A seasonal-trend decomposition," *J. Off. Stat*, vol. 6, no. 1, pp. 3–72, 1990.

[44] M. Barandas, D. Folgado, L. Fernandes, S. Santos, M. Abreu, P. Bota, H. Liu, T. Schultz, and H. Gamboa, "TSFEL: Time series feature extraction library," *SoftwareX*, vol. 11, Jan. 2020, Art. no. 100456.
[45] M. Christ, N. Braun, J. Neuffer, and A. W. Kempa-Liehr, "Time series FeatuRe extraction on basis of scalable hypothesis tests (tsfresh—A Python package)," *Neurocomputing*, vol. 307, pp. 72–77, Sep. 2018.
[46] D. Khider, F. Zhu, and Y. Gil, "autoTS: Automated machine learning for time series analysis," in *Proc. AGU Fall Meeting Abstr.*, vol. 2019, 2019, pp. 43D–1637.
[47] I. V. Barabanova, P. Vychuzhanin, and N. O. Nikitin, "Sensitivity analysis of the composite data-driven pipelines in the automated machine learning," *Proc. Comput. Sci.*, vol. 193, pp. 484–493, Jan. 2021, doi: 10.1016/J.PROCS.2021.10.050.
[48] S. Seabold and J. Perktold, "Statsmodels: Econometric and statistical modeling with Python," in *Proc. 9th Python Sci. Conf.*, vol. 57, no. 61, 2022, pp. 25010–25080.
[49] S. Mahmud, "Bangladesh COVID-19 daily cases time series analysis using Facebook prophet model," *SSRN Electron. J.*, vol. 2020, pp. 1–8, Jul. 2020, doi: 10.2139/SSRN.3660368.
[50] Y. J. Mo et al., "Performance of deep learning computation with TensorFlow software library in GPU-capable multi-core computing platforms," in *Proc. 9th Int. Conf. Ubiquitous Future Netw. (ICUFN)*, Jul. 2017, pp. 240–242.
[51] J. Herzen, "Darts: User-friendly modern machine learning for time series," *J. Mach. Learn. Res.*, vol. 23, no. 124, pp. 1–6, 2022.
[52] A. Alexandrov, K. Benidis, M. Bohlke-Schneider, V. Flunkert, J. Gasthaus, T. Januschowski, D. C. Maddix, S. Rangapuram, D. Salinas, J. Schulz, and L. Stella, "GluonTS: Probabilistic and neural time series modeling in Python," *J. Mach. Learn. Res.*, vol. 21, no. 116, pp. 1–6, 2020.
[53] İ. Met, L. Özbek, H. Aksoy, and A. Erkoç, "A modeling study on the estimation of COVID-19 daily and weekly cases and reproduction number using the adaptive Kalman filter: The example of Ziraat bank, Turkey," *J. Stat. Econ. Methods*, vol. 2021, pp. 15–27, Aug. 2021, doi: 10.47260/JSEM/1032.
[54] E. R. Q. Fernandes, A. C. P. L. F. de Carvalho, and X. Yao, "Ensemble of classifiers based on multiobjective genetic sampling for imbalanced data," *IEEE Trans. Knowl. Data Eng.*, vol. 32, no. 6, pp. 1104–1115, Jun. 2020, doi: 10.1109/TKDE.2019.2898861.
[55] R. Godahewa, K. Bandara, G. I. Webb, S. Smyl, and C. Bergmeir, "Ensembles of localised models for time series forecasting," *Knowl.-Based Syst.*, vol. 233, Dec. 2021, Art. no. 107518, doi: 10.1016/J.KNOSYS.2021.107518.
[56] S. E. Seker, "Computerized argument Delphi technique," *IEEE Access*, vol. 3, pp. 368–380, 2015, doi: 10.1109/ACCESS.2015.2424703.

ILKER MET received the degree (Hons.) and master's degrees from the Department of Geological Engineering, Middle East Technical University, in 1999, the M.B.A. degree from Sakarya University, and the Ph.D. degree (Hons.) from the Military Academy in Technology Management. Starting his career at Middle East Technical University, in 1996, he joined Ziraat Bank, in 2000, as an Assistant Specialist. He worked as a Specialist, the Director, the Project Management Department Manager, the Senior Vice President, and the Head of Enterprise Architecture Group, consecutively. He has been the Head of the Human Resources Group, since December 2019. Some of his awards and honors include the Brandon Hall Group Award, the Global Banking and Finance Award, the Financial Innovation Award, the Stevie Award, the Globee Award, the Efma-Accenture Banking Innovation Award, and the Fast Company Award (Most Innovative HR Leader).

AYFER ERKOÇ received the B.S. and M.S. degrees in statistics, in 2005 and 2008, respectively. She has been working at Ziraat Bank, since 2008. She worked at the retail marketing, enterprise architecture, human resources, and performance management departments, consecutively. She has international awards and papers, especially about data usage and statistical modeling, due to projects worked on. Her research interests include data-based projects, statistical modeling, big data, and analyses.

SADI EVREN SEKER was born in Istanbul, Turkey, in 1979. He received the B.S., M.S., and Ph.D. degrees in computer science and engineering, in 2000, 2003, and 2010, respectively.

From 2000 to 2004, he was a Research Assistant with the Department of Computer Science, Yeditepe University. After 2004, parallel to his academic career, he also lectured on computer science and management information systems in 18 different universities in six different countries. From 2012 to 2015, he was a Postdoctoral Researcher and conducted research projects with the Department of Computer Science on Big Data Analytics and Streaming Data Mining, UT Dallas. The author also leads four technology startups in USA and Turkey, in the artificial intelligence domains for the banking, finance, telecom, and retail sectors. His company focuses on the democratization of artificial intelligence, especially for small and medium enterprises. Besides the AI industry, he has been a full-time Professor with Antalya Bilim University, since 2019. He has three PCT patents and one national patent, 11 books internationally published, and more than 100 journals and conference papers. His research interests include automated machine learning, responsible artificial intelligence, explainable artificial intelligence, and big data analytics.