Uluslararası İleri Doğa Bilimleri ve Mühendislik Araştırmaları Dergisi
Sayı 8, S. 10-20, 2, 2024

Araştırma Makalesi

International Journal of Advanced Natural Sciences and Engineering Researches
Volume 8, pp. 10-20, 2, 2024

Research Article

https://as-proceeding.com/index.php/ijanser
ISSN: 2980-0811

Data-Driven Approaches to Optimize Branch and Team-Based Targeting in Banking

Mert Güneş[1], Metin Samet Korkmaz [2]

[1] *Türkiye Vakıflar Bankası T.A.O, Vakıfbank R&D Center Istanbul, Turkey*
[2] *Türkiye Vakıflar Bankası T.A.O, Vakıfbank R&D Center Istanbul, Turkey*

(Received: 22 February 2024, Accepted: 07 March 2024)

(4th International Artificial Intelligence and Data Science Congress ICADA 2024, March 14-15, 2024)

ATIF/REFERENCE: Güneş, M. & Korkmaz, M. S. (2024). Data-Driven Approaches to Optimize Branch and Team-Based Targeting in Banking. *International Journal of Advanced Natural Sciences and Engineering Researches, 8(2), 10-20.*

Abstract – The utilization of machine learning techniques to enhance strategic and operational decision-making within the banking sector is explored in this research. The second-largest state bank in Turkey conducted the studies, focusing on performance target prediction for two fundamental SME banking products: non-cash loans and demand deposits. Given the complex influencing factors such as volatile market conditions, customer creditworthiness, macro and microeconomic indicators, and team-specific variables, accurate performance prediction remains a significant challenge. The aim was to develop robust machine learning models capable of accurately predicting performance targets, thereby enabling efficient resource allocation and performance management. Techniques ranging from data mining and data preprocessing to feature selection and predictive modeling were applied in the studies. The effectiveness of the Orthogonal Matching Pursuit CV algorithm for branch targeting of non-cash loans and Stacked Regression algorithm for a dynamic team-based targeting process of demand deposits was revealed in the findings. The transformative potential of data analytics in banking and the importance of refining these models to cater to evolving industry needs are underlined by these insights.

Keywords – Target Prediction, Machine Learning, Banking, Data Science

I. INTRODUCTION

In the midst of an unprecedented digital revolution, the banking sector finds itself at a crossroads. Embracing cutting-edge technologies like machine learning is becoming more evident than ever as a significant edge in navigating the increasingly complex financial landscape. Our institution, as the second-largest state bank in Turkey, stands at the vanguard of this transformation, leveraging machine learning to inform strategic and operational decisions. A comprehensive exploration of two pioneering studies that employ machine learning techniques to predict performance targets associated with two fundamental SME (Small and Medium Enterprises) banking products: non-cash loans and demand deposits, is provided in this article.

II. PROBLEM DEFINITION

A. Non-Cash Loan Sales Performance Prediction

Our initial study undertakes the intricate task of performance target prediction for non-cash loans, a crucial offering for our SME customers across various branches. Non-cash loans, inclusive of a variety of instruments such as letters of credit and bank guarantees, serve as a cornerstone of our SME product portfolio. Accurately predicting performance targets for these products, however, is a daunting challenge given the myriad of influencing factors. These factors include volatile market conditions, the creditworthiness of SMEs, a range of macro and microeconomic indicators, and an array of branch-specific factors such as geographical location, staff expertise, and local competition.

B. Demand Deposit Performance Prediction

The second study turns its attention to demand deposits, another essential product for our SME customers. Demand deposits, due to their liquidity, are vital to the smooth functioning of SMEs, allowing them to manage their cash flow efficiently and conduct business transactions without a hitch. However, accurately predicting performance targets for demand deposits at the team level brings its own set of complexities. Influencing factors range from local economic conditions, which can impact SMEs' cash flow and hence their demand for deposits, to intricate customer behaviors, team-specific variables such as team size and skill set, and even seasonal trends in deposit activity [1].

III. PURPOSE OF STUDY

The objectives underlying these studies are multifaceted.

Foremost, our aim is to develop robust machine learning models capable of predicting performance targets for non-cash loans and demand deposits with high precision. Accurate predictions can function as a valuable compass in navigating the often turbulent waters of strategic planning, resource allocation, and performance management at both the branch and team levels. They can assist in ensuring that resources are directed where they are most needed and that potential issues are identified and tackled before they can impact performance.

Simultaneously, the aim to unearth the most influential predictors of these targets is pursued. This endeavor is not viewed as a mere academic exercise, but a crucial step towards achieving a deeper understanding of the factors that drive the business. Light is shed on the drivers of customer behavior and market trends, with the potential to shape product development and marketing strategies. The goal is to enable better service to SME customers and stay ahead of the competition.

The power of machine learning is harnessed with the aim of elevating decision-making capabilities, optimizing financial performance, and creating a banking experience that aligns closely with the needs and expectations of SME customers.

IV. LITERATURE REVIEW

The importance of predictive analytics in the banking industry's daily operations, especially in setting goals for individual branches or teams, is well-recognized. A broad array of studies explores this theme, delving into topics like the application of machine learning, use of econometric models, principles of management, and aspects of behavioral economics. This literature review aims to provide a comprehensive analysis of existing research, with an emphasis on the methods used, the prediction precision of various models, and the real-world impact of these predictions on team dynamics and institutional performance. The aim is to identify topics that existing studies might have missed and point out possible directions for further academic investigation. Firstly, it can be said that the widespread use of machine learning in banking is reflected in the literature. Machine learning is frequently employed in credit risk management, reaching the right customers, and artificial intelligence applications for financial competition. For instance, Aphale & Shinde (2020) demonstrate in their study that the approval of a customer's loan request can be predicted with 76-80% accuracy using machine learning. They also show that creating the final model by building important features from the database and training it by testing many different models is the correct method. [2] Similar literature can be considered when determining the main method.

Moreover, noteworthy instances come up when the research is restricted to Turkey. High success rates in predicting customer numbers using machine learning techniques have been shown in studies by Yetiz, Terzioglu, and Kayakus (2021). The benefits of artificial neural networks and SVM, with an accuracy over 94%, were revealed. The importance of evaluating R squared, MSE, RMSE, and MAPE together is emphasized in the results evaluation section, with each being explained separately. [3]

Another significant aspect to consider is that the studies we will conduct need to account for different customer types and products. Therefore, examining literature that provides this breakdown is essential. There are studies that investigate the topic among different customer types. For example, in a study published in 2019 for SME customers, Zhu, Zhou, and Xie et al. emphasized that by considering various distinct characteristics of SMEs (credit risk, profit margin, rate of return on assets, etc.), machine learning models can achieve predictions with an accuracy rate close to 85 percent. [4]

On the other hand, the burgeoning incorporation of artificial intelligence in banking, with a focus on deep learning and data mining, has been profound. Despite this trend, our literature review highlights that the practical application of such advanced AI techniques does not always yield a superior performance. The efficacy of machine learning is exemplified by the logistic regression-based model referenced in the work of A. Shinde et al., which achieved an 82 percent accuracy rate in its application, a metric that stands out for its efficiency and reliability . [5]In contrast, deep learning, as extensively reviewed by Hassani et al. in their work on its implementations in banking, did not show a distinct advantage in our project . [6] Consequently, our project did not necessitate the use of deep learning, underscoring that the choice of AI methodology must be dictated by its relevance and effectiveness to the task at hand rather than the allure of its complexity.

Furthermore, the critical issue of loan defaults is rigorously examined through a study that utilizes logistic regression to predict loan delinquency. This approach endeavors to minimize non-performing assets and amplify profits. The study produced a machine learning algorithm that predicts loan eligibility with commendable accuracy, based on a comprehensive set of 13 variables, with the model achieving an apex accuracy score of 0.785 . [7]

In conclusion, the trajectory of this research points to future improvements in the algorithm's reliability, efficiency, and robustness, as well as its integration with automated banking systems a step that could revolutionize risk mitigation and loan approval processes in the banking industry.

V. CASE STUDY

A. NON-CASH LOAN SALES PERFORMANCE PREDICTION

This case study explores the application of predictive analytics in improving branch targeting for non-cash loans at Vakıfbank. As banking sectors strive for increased efficiency and customer satisfaction, understanding the right strategy to employ in branch targeting becomes paramount. This investigation intends to provide insights into how data-driven decision-making can enhance financial services in Vakıfbank.

I. DATA PREPARATION

The initial phase of this project was marked by the retrieval of relevant data from the bank's databases. The intricate relationship between various bank operations necessitated the extraction of data from multiple tables, resulting in the use of complex queries. The primary language for these database queries was PL/SQL, a tool appreciated for its efficiency and security, designed specifically for seamless data manipulation and processing in Oracle databases. The Jupyter Notebook environment served as the arena for the execution and testing of these queries. The value of this interactive platform was exhibited in the real-time testing, modification, and validation of the PL/SQL scripts. The highly optimized queries, resulting from meticulous design and rigorous testing, performed excellently despite the vastness of the data involved. Over a thousand rows of data, encompassing numerous facets of Vakıfbank's operations, were fetched by the executed queries. The use of unions in assembling this rich dataset from multiple tables allowed for a comprehensive and detailed investigation.

A solid foundation for the subsequent stages of the study was provided by this data acquisition process, which ensured accurate, insightful, and effective analytics. The robustness and comprehensiveness of the retrieved data were attributed to the PL/SQL querying process and the versatility of the Jupyter Notebook platform.

The original six-month dataset, prepared in line with the performance targets set by business units, was restructured into three-month intervals to align with changes in the performance management system. This adaptation contributed to the overall quantity and diversity of the data. Throughout the project, these aggregated three-month datasets were utilized in the models to enhance their accuracy and robustness.

In the next phase of the analysis, the necessary libraries that facilitated data processing and modeling were imported. Key libraries utilized included pandas for data manipulation and analysis, sklearn for machine learning, and seaborn for data visualization. At this stage, the data prepared using PL/SQL queries was also brought into the Python environment.

Upon loading the required libraries and data, the target variable and the set of features to be used for its prediction were designated. The loan stock at the end of the term served as the dependent variable, while the remaining variables, acting as predictors in the model, were the independent variables. A train-test split was implemented to partition the data into a training set and a testing set, a standard practice in machine learning to evaluate how well a model generalizes to unseen data. An 80-20 split was employed, with 80% of the data used for training and 20% reserved for testing.

Missing values within the dataset were handled through the fillna method, a strategy that replaced all NaN values with zero. This approach resulted from an intensive process of exploration and testing, which also evaluated various other methods. A comparative study of multiple methods revealed that substituting the missing values with zero yielded the most reliable and accurate results for this specific dataset, providing the least distortion and maintaining data integrity.

Feature selection, a crucial step in building effective machine learning models, followed. The feature selection employed RandomForestRegressor, LightGBM, and XGBoost, all recognized for their excellent performance in feature importance estimation. Each of these algorithms possesses a 'feature_importances_' attribute, providing a score indicating the usefulness of each feature in the prediction task.

Notably, the 'tree_method' parameter was set to 'gpu_hist' in the use of XGBoost. The 'gpu_hist' method is a histogram-based algorithm allowing the model to utilize the Graphics Processing Unit (GPU) instead of the Central Processing Unit (CPU) during training. This change significantly enhances computational speed, as GPUs excel in handling parallel processes, an aspect heavily employed in boosting algorithms like XGBoost. As a result, model training becomes much faster, allowing for more efficient iteration through different models and hyperparameters.

Table 1: Most Important Branch Based Features according to feature importances of three different models

Feature Name	Description
SB_GUNCEL_DONEM_GNK_ACILAN_DIGER	Number of customers opened new non-cash loan
SB_GUNCEL_DONEM_GNK_STOK	Total non-cash loan amount
SB_GUNCEL_DONEM_GNK_STOK_DIGER	Total non-cash loan amount of non-big customers
SB_GUNCEL_DONEM_GNK_STOK_KETLI	Total non-cash loan of upper high customers
VAKIFBANK_LIMIT	Total limit of customers in the bank

VI. MODEL SELECTION

The computation of feature importance scores by each of the three machine learning models (Random Forest, XGBoost, and LightGBM) led to the selection of the top 20 features from each. These selected features were combined and only those that appeared in the top 20 list of all three models were retained. This intersection of features, deemed crucial by all three models, formed the final feature set for modeling. The first four most important branch-based features were displayed in Table 1. The crucial step of model selection saw the utilization of the LazyPredict Python library, a tool facilitating quick fitting and evaluation of a variety of machine learning models. A comprehensive overview of the performance of many models on the data was provided by LazyPredict. However, to ensure a more thorough evaluation, the most promising models underwent additional hyperparameter tuning. This process enhanced their predictive power and allowed exploration of the models' performance under different settings. The outcomes of the fine-tuned models were summarized and evaluated. OrthogonalMatchingPursuitCV stood out as the best performer among all the tested models in terms of R-squared and Mean Absolute Percentage Error (MAPE). Additionally, its relatively low training time rendered it an attractive option.

OrthogonalMatchingPursuitCV, an algorithm utilized for linear regression problems, fits a linear model through the iterative selection of predictors. In each iteration, a predictor, most highly correlated with the target variable given the residual of the previous model, is identified by the algorithm. The "CV" in the algorithm's name, standing for "cross-validation," signifies a robust method deployed to evaluate the generalization of a model's results to an independent data set. This is accomplished by partitioning the original sample into a training set, where the model is fit, and a test set, where it is evaluated. The OrthogonalMatchingPursuitCV algorithm, serving as an efficient tool for variable selection and regularization, demonstrates exceptional performance when the problem involves multicollinearity or when a limitation on the number of features in a high-dimensional dataset is necessary [8].

Table 2: Model Performance List - Non-Cash Loan

Model Name	R_Squared	MAPE
OrthogonalMatchingPursuit	0.84	0.54
LGBMRegressor	0.73	0.62
ExtraTreesRegressor	0.73	0.69
HistGradientBoostingRegressor	0.73	0.75
RandomForestRegressor	0.71	1.5

VII. RESULTS

The final step involved the selected OrthogonalMatchingPursuitCV model being utilized to make predictions for each bank branch. The model was put to work, with loan stocks being forecasted for each individual branch. Following this, coefficients for each branch's prediction were computed by dividing the specific branch's prediction by the total sum of all predictions. These coefficients represented the proportion of the total predicted loan stock for which each branch was responsible.

This valuable information was then relayed to the Small and Medium Enterprises (SME) department. These results were used by them as a guide in their branch targeting process. The predictions of the model, represented by the coefficients, served as a tool for the SME department to prioritize their resources and efforts. It allowed them to concentrate on branches that were expected to contribute more significantly to the bank's loan portfolio. Looking at the initial results, the feedback from the SME department was very positive, especially in terms of speeding up the targeting process. As a result, a model that acts as a pioneer in spreading the performance process to machine learning in all products with more advanced models by providing full automation in the long term has been created.

B. DEMAND DEPOSIT PERFORMANCE PREDICTION

In this second case study, certain aspects of the previous work were revised. The data, previously prepared based on the branches of the bank, was prepared based on teams this time. Features found useful in the previous study were brought to the main focus in this study, thus time in the data preparation process was saved. The aim was to create a more dynamic targeting process by working with 3-month data. Following the success of the study in credit products, the focus this time was shifted to deposit products. In this context, the model created for SME customers in terms of Turkish Lira for the "non-term deposit" product, which is thought to have achieved successful results, was included.

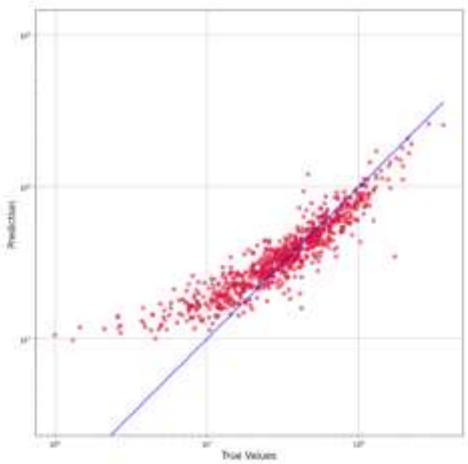

Figure 1: Selected Model True Values & Predictions Graph - Non-Cash Loan

I. DATA PREPARATION

Firstly, at the data preparation stage, collaboration occurred with the necessary teams to examine the factors affecting the sale of the deposit product. The queries prepared earlier were expanded this time by revising them to team breakdowns instead of branches in the same databases. The 6-month data was converted into 3-month data, thereby the quantity and diversity of the data was increased. In addition, external data from banking institutions in Turkey was added to the data on a regional basis to improve prediction quality.

Although the study was prepared by separately evaluating term and non-term deposit products for SME customers under four main headings in terms of foreign currency and Turkish Lira, the data largely parallels. Here, the work carried out while preparing the data of the non-term deposit, which yielded the most successful results, is included:

1. The breakdown of the number of customers who have a non-term deposit product by each customer segment
2. Total sales volumes of teams in the last 3-6 months
3. General characteristics of teams
4. External Data (Especially Data about General Banking Metrics of Location)

When all these are examined, in the ranking made using the prioritization of the "XGBoost" algorithm, the first 5 features can be seen in Table 3.

Another different perspective is the inclusion of multiple "target" data during data preparation. With the idea that if success is achieved, the best target will be beneficial in all kinds of targeting, both average and stock deposit amounts have been prepared on a team basis.

II. MODEL SELECTION

In model selection, LazyPredict was used similarly to the previous study, but since a more detailed study was needed, the most efficient models were subjected to parameter tuning as an alternative, and trials were conducted. The results of the tuned models can be found in Table 4 (sorted by decreased R_Squared). R_Squared means squared R score of model and MAPE means Mean Absolute Percentage Error. In general, R_Squared increased and MAPE decreased except one percent differences.

Table 3: Most Important Team Based Features according to XGBoost Model

Feature Name	Description
VDSZ_MUSTERI_SAYISI	Number of customers have demand deposit
MEVD_SON_3_AY_ORT	Last 3 months average demand deposit
MEVD_SON_1_AY_ORT	Last month average demand deposit
DIS_TICARET	Change in the number of customers of the branch
POS_UC_AYLIK_CIRO	Last month average demand deposit

Table 4: Model Performance List- Deposit

Model Name	R_Squared	MAPE
RandomForestRegressor	0.86	0.35
GradientBoostingRegressor	0.84	0.34
ExtraTreesRegressor	0.84	0.35
XGBRegressor	0.83	0.36
OrthogonalMatchingPursuit	0.79	0.37

Additionally, in this study, the StackedRegressor algorithm from the Sklearn library was, for the first time, included in the trials. StackedRegressor, also known as Stacking Regression, is recognized as a type of ensemble learning method where multiple regression models are trained to predict the same output variable, and their predictions are, in some way, combined to make the final prediction. The basic idea underlying stacking is to leverage the strengths of different models to enhance the overall performance. Individual models are trained based on the complete training set; then, a meta-model is fitted based on the outputs, or the so-called out-of-sample predictions, of the individual models to generate a final prediction. This approach aids the model in correcting the mistakes of individual models, thereby improving the model’s performance. The list of models used in StackedRegressor and the results are included in Table 5 [9].

Table 5: StackedRegressor Model List

Stacked Model List	Final Estimator	R_Squared
RandomForestRegressor GradientBoostingRegressor OrthogonalMatchingPursuit XGBRegressor	LinearRegression	0.81

Upon examination of all models, the StackedRegressor model, with an R-squared score of 0.81 and a MAPE score of 0.38, was not the highest-scoring in terms of metrics. However, it has been appraised for introducing an innovative perspective and for its ability to leverage multiple models to produce competitively close results.

The Stacked Regressor model, selected for its robustness, yielded the outcomes detailed in Table 6, showcasing a variety of metrics. The graphical representation in Figure 2 allows for a visual assessment of the model’s predictions in comparison to the actual test values. The results indicate a commendable degree of accuracy, with the model’s predictions closely mirroring the real figures.

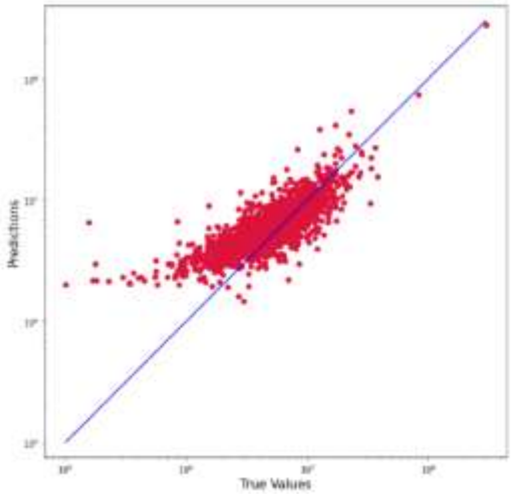

Figure 2: Selected Model True Values & Predictions Graph - Deposit

The model has been designed to forecast the average non-term deposit amounts expected at the end of a quarter, utilizing data from the preceding three months and adjusting for team-specific variables. Initial observations have indicated a transition within the bank, moving away from rule-based systems towards a greater dependency on machine learning models for decision support.

While the on-paper metrics suggest that demand deposit predictions outperform those of non-cash loans, the real-world application tells a different story. Contrary to what the numbers might imply, the non-cash loan predictions have proven to be more valuable in practice. They have been particularly effective and have received positive feedback from the bank's SME team, who have noted the practical utility and accuracy of these predictions in their operations.

The initial slower adaptation rate of the model can be attributed to the novelty of the system and the relative immaturity of the dataset. Yet, the expectation is that with time, as the system matures and the dataset grows, full automation will be feasible, and the model will provide even more precise targets. This, in turn, is anticipated to enhance the bank's operational efficiency and profitability, evidencing the pragmatic superiority of the non-cash loan performance predictions in the banking environment.

Table 6: Selected Model Performance List

Model Name	R_Squared	MAPE
StackedRegressor	0.81	0.38

VIII. DISCUSSION

In the discussion section of this study, the strategic value and potential implications of our findings are examined, along with the shifts in methodology implemented in the second case study. The objective of these studies was to fine-tune the targeting process for Vakıfbank's non-cash loans and non-term deposit products, focusing on branch-level and team-level analysis respectively. The bank's aspiration to optimize their resource allocation and elevate the efficiency of their targeting strategies underpinned this exploratory effort.

The first case study revolved around branch targeting for non-cash loans, utilizing a combination of SQL queries, data pre-processing, feature selection, and predictive modeling techniques. The employment of the OrthogonalMatchingPursuitCV algorithm, distinguished by its performance in terms of Mean Absolute Percentage Error (MAPE), R-squared value, and training time, marked a significant advancement towards achieving more accurate targeting. The SME department was able to identify branches projected to contribute more significantly to the bank's loan portfolio through the predictions generated by the algorithm, thereby enhancing the overall targeting process.

In the second case study, a transition from the branch-based approach to a team-based tactic was made, focusing on the performance prediction for demand deposits. This change in perspective not only introduced a new dimension to the bank's targeting process but also utilized the valuable insights gleaned from the first study. Moreover, the modification from six-month data to three-month data facilitated a more dynamic targeting process.

This second study represented a strategic initiative to diversify the bank's product focus. After the successful incorporation of predictive modeling in credit products, attention was shifted towards deposit products, particularly non-term deposit product for SME customers. In alignment with the first study, this model was developed to guide the bank in resource allocation and strategy formulation.

In conclusion, both studies provide substantial contributions towards the refinement of Vakıfbank's targeting process. The employment of predictive modeling, in conjunction with a dynamic and targeted approach, allows the bank to allocate their resources and efforts more effectively. As the banking sector becomes progressively competitive, our studies highlight the significance of leveraging data analytics for strategic decision-making. Future research could build upon these findings by exploring the applicability of our models to different products, customer segments, or even geographical areas, potentially presenting a wider array of innovative solutions to the banking industry.

IX. CONCLUSION & FUTURE WORK

The enhancement of the targeting process for Vakıfbank's non-cash loans and non-term deposit products was established as the fundamental objective of these studies. A systematic approach was adopted, with initial focus on refining branch targeting for non-cash loans through the application of techniques ranging from SQL queries and data pre-processing to feature selection and predictive modeling. Subsequently, focus was shifted to the performance prediction for demand deposits at a team level.

From the conducted research, the value of predictive modeling in the banking sector, particularly in relation to product targeting, was a prominent takeaway. The effectiveness of the OrthogonalMatchingPursuitCV algorithm in identifying branches likely to contribute more significantly to the bank's loan portfolio was disclosed in the first study. A similar strategy was employed in the second study for team-based targeting for deposit products, culminating in a more dynamic targeting process.

The cumulative significance of these studies is realized in their contribution towards the enhancement of the efficiency and precision of resource allocation within Vakıfbank. The underscored power of data analytics in strategic decision-making, as highlighted by the findings, paves the way for more focused, timely, and impactful strategies.

At its core, the conducted research is a testament to the transformative potential of data analytics in banking. As the industry undergoes evolution, the insights derived from such studies are set to become integral to banks' strategic planning, resource allocation, and overall performance. Therefore, the exploration and refinement of these models should be continuously undertaken, extending their applicability to other products, customer segments, and geographical regions to offer broader, more innovative solutions to the banking industry. [10]

REFERENCES

[1] I. Met, A. Erkoc， , and S. E. Seker, "Performance, efficiency, and target setting for bank branches: Time series with automated machine learning," IEEE Access, 2022.

[2] A. S. Aphale and S. R. Shinde, "Predict loan approval in banking system machine learning approach for cooperative banks loan approval," 2020.

[3] F. Yetiz, M. Terzioˇglu, and M. Kayakus， , “An analysis of turkish deposit banks’ customer forecasting with machine learning methods,” Sosyoekonomi, vol. 29, no. 50, pp. 413–432, 2021. doi: 10 . 17233 / sosyoekonomi.2021.04.19.

[4] L. Wang, F. Jia, L. Chen, and Q. Xu, “Forecasting smes’ credit risk in supply chain finance with a sampling strategy based on machine learning techniques,” Annals of Operations Research, 2022. doi: 10.1007/s10479-022-04518-5.

[5] A. Shinde, Y. Patil, I. Kotian, A. Shinde, and R. Gulwani, “Loan prediction system using machine learning,” in ITM Web Conf., vol. 44, 2022, p. 03 019. doi: 10.1051/itmconf/20224403019.

[6] H. Hassani, X. Huang, E. Silva, and M. Ghodsi, “Deep learning and implementations in banking,” Annals of Data Science, vol. 7, pp. 433– 446, 2020.

[7] S. Dosalwar, K. Kinkar, R. Sannat, and N. Pise, “Analysis of loan availability using machine learning techniques,” International Journal of Advanced Research in Science, Communication and Technology (IJARSCT), vol. 9, no. 1, pp. 15–20, 2021.

[8] J.Wang, S. Kwon, and B. Shim, “Generalized orthogonal matching pursuit,” IEEE Transactions on signal processing, vol. 60, no. 12, pp. 6202– 6216, 2012.

[9] R. Polikar, “Ensemble learning,” in Ensemble machine learning: Methods and applications, 2012, pp. 1–34.

[10] M. Amakobe, “The impact of big data analytics on the banking industry,” Ph.D. dissertation, Colorado Technical University, 2015.