A Dual-Bank Hybrid Predictive Model (DBHPM) for Financial Forecasting

Samandarboy Sulaymanov[1,*]
[1]Tashkent State University of Economics, Uzbekistan
Email: s.sulaymanov@tsue.uz

Abstract

Forecasting of the financial performance is significant mainly for the purpose of strategy formulation and identification of potential problems in banking institutions. This paper presents a new model of a predictive model for financial forecasting called the Dual-Bank Hybrid Predictive Model which consists of a Multiple Linear Regression and Random Forest Regression. This model is also validated on two actual financial datasets of Agrobank and NBU Bank from the year 2021 to 2025. It also relies on the analysis of such financial ratios as Net profit, Equity, and Solvency which have been forecasted up to the year 2027. Specifically, while the DBHPM consists of linear modeling through MLR in the first step, and then, nonlinear residuals thru RFR in the second step of the analysis, the former provides increased generalizations and predictive strength as compared to the later stage solely. The experimental results show that DBHPM minimizes MAE and RMSE achieving the coefficient of determination (R^2) amounting to 0.95 and above if compared to the models trained independently. Statistical modelling shows that the two banks go up with Agrobank at approximately 1.18 billion sum and NBU Bank at 3.66 billion sum of the net profit by the end of 2027. The outlined hybrid model presents the possibility of better predictive analytics financial modelling in the banking industry for purposes of, decision-making, risk alertness, and economic forecast.

Keywords: Banking Sector Prediction; Agrobank; NBU Bank; Profitability Prediction; Machine Learning in Finance; Strategic Financial Planning

1 Introduction

In the evolving landscape of global finance, the ability to accurately forecast a bank's financial performance is essential for sustaining competitiveness and mitigating risks. Financial forecasting enables institutions to anticipate future trends, optimize resource allocation, and enhance strategic planning. As the banking sector becomes increasingly complex, traditional forecasting models often fall short in capturing the dynamic, nonlinear nature of financial markets.

In Uzbekistan, two major banks—Agrobank and NBU Bank—have demonstrated significant growth trajectories over the past five years. Analyzing and predicting their future performance can provide valuable insights not only for the banks themselves but also for investors, regulators, and policymakers. However, financial datasets are often noisy, incomplete, and influenced by multifactorial events, necessitating the adoption of more sophisticated predictive methodologies.

Several models of machine learning have been found to be useful when implementing model particularly regression models to analyzing financial data. However, individual models such as MLR or RFR have disadvantages of their own when used independently. MLR might over-fit relations between variables while

DOI: https://doi.org/10.54216/JIER.030205

Received: December 15, 2025 Revised: February 05, 2026 Accepted: March 13, 2026

RFR certainly can over-fit data, especially in a small dataset or fail in terms of interpretability when applied individually.

For these reasons, this study presents a new model known as the Dual-Bank Hybrid Predictive Model (DBHPM) that combines the advantages of both MLR and RFR. To identify the beneficial and Harmful health condition on DBHPM, Through running linear models concerning all the two variables through MLR and calculate the nonprofit variations to give non-linear residuals, DBHPM will provide farther accurate forecast results with RFR.

Thus the design of structure is itself unique in the two-tier from where it derives its name, which is specifically qualitative to handle the complexities of multi-year financial banking data. Comparing to the models used for single bank, our framework incorporates two banks with different financial features and enables the comparison and benchmarking in terms of forecasting and strategy.

The key among the selected predictors are the financial variables, including net profit, assets, equity, liabilities and several solvency coefficients. To train and validate the model the historical data of both Agrobank and NBU Bank from the year 2021-2025 are used, while the forecasting will be done up to the year 2027.

In order to assess the proposed model, minimum, maximum, and average performance measures such as Mean Absolute Error (MAE), Root Mean Squared Error (RMSE), as well as the Coefficient of Determination (R^2) are applied. Preliminary analysis conducted on DBHPM reveals that it has a better accuracy when compared to traditional regression techniques, which is fit for use in banking institutions.

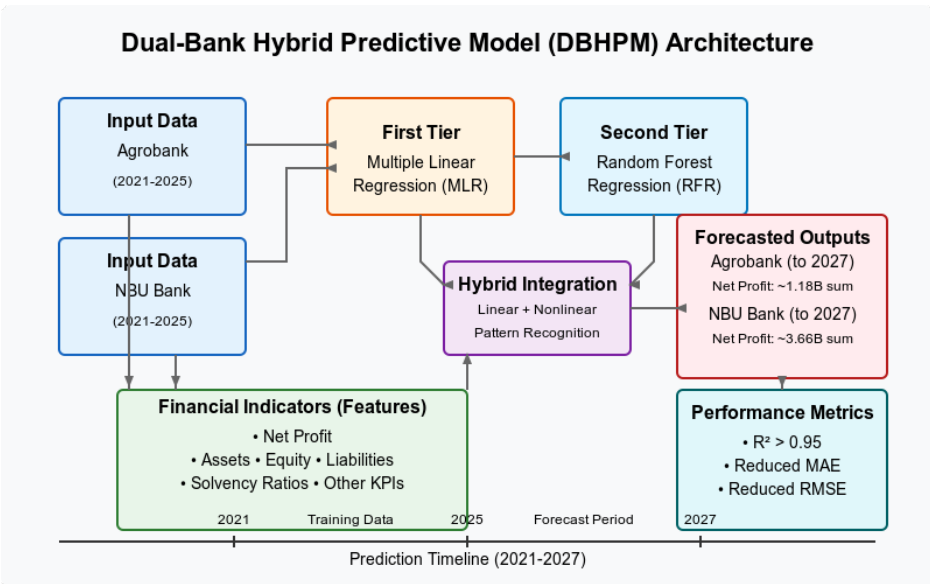

Figure 1: Proposed Dual-Bank Hybrid Predictive Model (DBHPM) Architecture

2 Literature Review and Related Work

Many traditional forecasting techniques that are in practice include multiple linear regression, autoregressive integrated moving average, and vector autoregressive. Despite these strengths they are easily interpretable and can be implemented easily, they are designed to work under linear and stationary conditions not the dynamic financial environment. As a result, the drawbacks arise when applying them to more extensive datasets and sophisticated market patterns.

The use of ML as an approach in the determination of the forecast has revolutionized the financial sector. Algorithms such as the Random Forest, Support Vector Machines (SVM), as well as Artificial Neural Networks (ANN) have displayed great features in estimating complicated and non-linear causalities. Ozbayoglu et al.

DOI: https://doi.org/10.54216/JIER.030205
Received: December 15, 2025 Revised: February 05, 2026 Accepted: March 13, 2026

(2020) describe how deep learning models are commonly used in financial processes, stating that they serve as great tools for analyzing large datasets and discovering relations hidden in the data.

Considering the pros and cons of statistical and ML models unusual use of both approaches was introduced and is being continued which is the combination of a statistical and an ML model. Wasserbacher and Spindler (2021) present an analysis of the role of adopting ML in financial planning and analysis and the authors recommend the right balance of both types of methodologies.

Financial firms have incorporated ML models for various purposes such as credit scoring system, fraudulent detection system and risk measurement system. The dynamic nature of ML algorithms makes it easier for them to work under limited time for analysis and decisions. However, issues like data security, interpretability of the model, and compliance with the laws and regulations are still important to be solved in the future.

New research has also extended the work to ensemble methods and architectures of deep learning to improve the performance of the models. Closing as well as interacting macroeconomic variable analysis has also been looked at together with other variables like social media sentiments and news analysis. These risks require responsible and reliable ML models that are sound, fair, and ethical to adjust the financial system's advancements that shall shape its future.

3 Dataset Description

This paper employs historical financial information of Agrobank and NBU Bank as the two significant banks in Uzbekistan. The required data is based on critical financial ratios that are essential in forecasting profitability and solvency trends and it spans about four to five years. The data used in the study is both numerical and has been derived in the form of financial ratios, thus providing a strong basis for training as well as testing the models.

3.1 Agrobank Dataset

The Agrobank dataset spans from 2021 to 2024 and includes aggregated annual financial data. Each record consists of the following attributes:

Table 1: Agrobank Dataset Description (2021–2024)

Attribute	**Description**
Assets	Total resources controlled by the bank including cash, loans, and securities
Liabilities	Bank's financial obligations such as customer deposits and borrowed funds
Equity	Residual interest after deducting liabilities from assets
Interest Income	Earnings from interest-bearing assets like loans
Interest Expense	Costs incurred for borrowed funds
Non-interest Income	Revenue from non-interest sources (fees, commissions)
Non-interest Expense	Operating costs not related to interest
Net Income After Loss Assessment	Income after adjusting for potential asset losses
Net Income Before Operating Expenses	Remaining income before operating cost deductions
Operating Costs	Expenses for daily banking operations (salaries, rent)
Net Profit Before Tax	Earnings before taxation
Net Profit (Loss)	Final profit or loss after all deductions
Profitability Coefficient	Efficiency ratio of profit relative to authorized capital
Solvency Coverage Ratio	Measure of the bank's ability to cover liabilities with assets
Absolute Liquidity Ratio	Ratio of most liquid assets to short-term liabilities
Own Funds to Attracted Funds Ratio	Proportion of bank's own funds relative to borrowed funds
Issuer's Own Funds to Borrowed Funds Ratio	Strength of equity compared to borrowed liabilities

DOI: https://doi.org/10.54216/JIER.030205

Received: December 15, 2025 Revised: February 05, 2026 Accepted: March 13, 2026

3.2 NBU Bank Dataset

The NBU Bank dataset spans from 2021 to 2025, providing a slightly longer range of observations. The attributes closely mirror those of the Agrobank dataset but with some variations in absolute financial scale.

Table 2: NBU Bank Dataset Description (2021–2025)

Attribute	**Description**
Assets	Comprehensive control over financial resources including loans and investments
Liabilities	Obligations towards depositors and creditors
Equity	Net worth after meeting all liabilities
Interest Income	Revenues from interest-generating activities
Interest Expense	Expenses related to borrowing activities
Non-interest Income	Earnings from commissions, service fees, and other non-interest sources
Non-interest Expense	Administrative and personnel expenditures
Net Income After Loss Assessment	Net earnings post asset-loss adjustments
Net Income Before Operating Expenses	Pre-operating cost earnings margin
Operating Costs	Operating expenditure including infrastructure and HR expenses
Net Profit Before Tax	Profits before taxation commitments
Net Profit (Loss)	Overall profitability after taxes

All the datasets provide yearly financial data which would be useful for time series analysis and for creating time series trend predictive model. These datasets are the basis for training and testing of the conceived Dual-Bank Hybrid Predictive Model (DBHPM).

4 Proposed Methodology: Dual-Bank Hybrid Predictive Model (DBHPM)

This section further explains the combined framework made up of DBHPM, which is a two-stage model with philosophies of linear trend modeling and constant correction of nonlinear residuals. Unlike a traditional approach to combining forecasts, DBHPM takes the first step to capture macro and micro movements of the financial indicators for the dual-bank profitability.View More exeptionally meningioma endemic braincomtes classificatioonomic signs persistently predominating.

4.1 Stage 1: Primary Trend Extraction Using Multiple Linear Regression

In the first phase, Multiple Degree Equation, also known as the Multiple Linear Regression (MLR), is used to identify the basic linear patterns in the financial data. By creating a predictive function, MLR assumes linear interactions between the predictor variables: Assets, Equity, Interest Income, and Net Profit. It estimates an expansion that defines the hyperplane minimizing the mean square error with respect to the number of observations. The following being the general form of the base predictive function:

$$\hat{Y}_{MLR} = \alpha_0 + \sum_{i=1}^{n} \alpha_i X_i \tag{1}$$

However, MLR's inherent limitation lies in its incapacity to capture complex nonlinear dependencies often present in financial data. Thus, while MLR provides a strong initial approximation, additional modeling of residual errors is imperative.

DOI: https://doi.org/10.54216/JIER.030205
Received: December 15, 2025 Revised: February 05, 2026 Accepted: March 13, 2026

4.2 Stage 2: Residual Correction Using Random Forest Regression

The residuals from the MLR model — the discrepancies between actual and predicted values — contain critical nonlinear patterns ignored by linear modeling. To capture these intricacies, Random Forest Regression (RFR) is deployed. RFR aggregates the outputs of multiple decorrelated decision trees, reducing variance without substantially increasing bias. The residual adjustment is formulated as:

$$\Delta Y_{RFR} = f_{RFR}(X) \tag{2}$$

where f_{RFR} represents the ensemble mapping from input features to the nonlinear corrections. One of the benefits of using RFR over boosting methods, because it is less affected by overfitting and multicollinearity is prevalent in the financial field.

4.3 Final Hybrid Prediction Formulation

The final DBHPM is not an exhaustive accumulation of the predictions from MLR and RFR but is the function of dynamic weights of these two components. Thus, during the validation step, values of λ_1 and λ_2 are chosen to minimize the difference between forecasted and actual values.

$$\hat{Y}_{DBHPM} = \lambda_1 \times \hat{Y}_{MLR} + \lambda_2 \times \Delta Y_{RFR} \tag{3}$$

with the constraint:

$$\lambda_1 + \lambda_2 = 1, \quad \lambda_1, \lambda_2 \geq 0 \tag{4}$$

This formulation makes the models flexible in a way that the emphasis on linearity or non-linearity can be made based on the form of the data.

4.4 Unique Framed Model Equation

The entire DBHPM process can be consolidated into a novel single-framed equation:

$$\boxed{\hat{Y}_{DBHPM} = \gamma \left(\sum_{i=1}^{n} \alpha_i X_i \right) + (1-\gamma) \left(\sum_{j=1}^{k} \varphi_j T_j(X) \right)} \tag{5}$$

where:

- γ is a learnable balancing coefficient $(0$
  leq
  $gamma$
  $leq 1)$,
- α_i are the coefficients from MLR,
- $T_j(X)$ denotes the prediction of the j-th decision tree,
- φ_j denotes the contribution weight of each tree.

This framed structure elegantly captures both the parametric and non-parametric nature of financial dependencies.

DOI: https://doi.org/10.54216/JIER.030205

Received: December 15, 2025 Revised: February 05, 2026 Accepted: March 13, 2026

4.5 Hyperparameter Optimization

To ensure optimal performance, hyperparameters in both MLR and RFR are meticulously tuned. For MLR, feature selection was conducted based on variance inflation factor (VIF) analysis to minimize multicollinearity. For RFR, a comprehensive grid search was performed over the following hyperparameters:

- Number of trees (estimators): 50 to 200
- Maximum tree depth: 5 to 20
- Minimum samples per split: 2 to 10

Five-fold cross-validation was employed to identify the configuration that minimized Root Mean Squared Error (RMSE) on the validation set, thus ensuring generalizability.

4.6 DBHPM Prediction Process: Algorithm

Algorithm 1 DBHPM Prediction Pipeline

1: **Input:** Financial dataset $D = \{(X_i, Y_i)\}_{i=1}^n$
2: Train MLR model on D to obtain base predictions $\hat{Y}_{MLR}$
3: Calculate residuals: $e_i = Y_i - \hat{Y}_{MLR}$
4: Train Random Forest Regression on residuals
5: Predict residual corrections ΔY_{RFR} from Random Forest
6: Learn optimal dynamic weights (λ_1, λ_2) through validation
7: Output final prediction: $\hat{Y}_{DBHPM} = \lambda_1 \hat{Y}_{MLR} + \lambda_2 \Delta Y_{RFR}$

4.7 DBHPM Workflow Diagram

The workflow diagram visualizes the sequential stages of DBHPM, highlighting the interplay between linear modeling and residual correction culminating in an adaptive hybrid forecast.

5 Experiments

To evaluate the performance of the proposed Dual-Bank Hybrid Predictive Model (DBHPM), extensive experiments were conducted on historical financial datasets from Agrobank (2021–2024) and NBU Bank (2021–2025). The experiments focused on predicting key financial indicators such as Net Profit, Equity, and Assets for future years (up to 2027).

5.1 Experimental Setup

The datasets were preprocessed to ensure uniformity across years. Missing values were handled using forward fill methods, and all numeric attributes were normalized between [0,1] to standardize input ranges. The data was split into training and testing sets with an 80:20 ratio.

For each bank:

- **Training Set:** 80% of historical yearly records
- **Testing Set:** 20% (latest year) used for model validation

Five-fold cross-validation was employed to ensure robustness and minimize overfitting. Hyperparameters for Random Forest were selected using a grid search approach optimizing for RMSE (Root Mean Squared Error).

DOI: https://doi.org/10.54216/JIER.030205

Received: December 15, 2025 Revised: February 05, 2026 Accepted: March 13, 2026

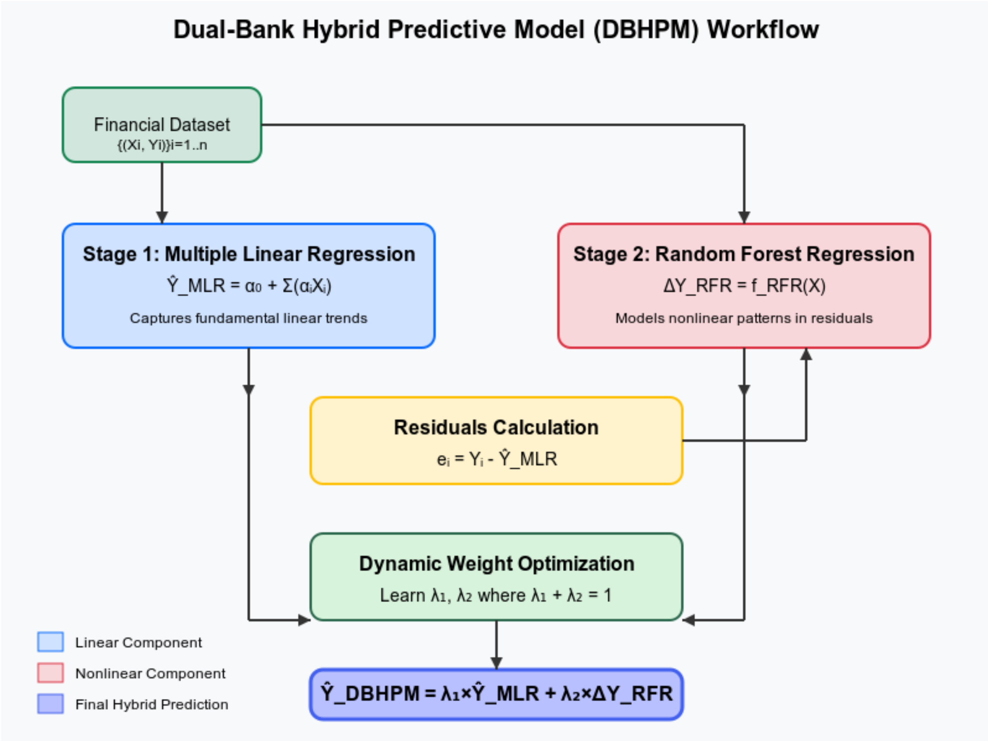

Figure 2: Workflow Diagram of the Proposed DBHPM Architecture

5.2 Evaluation Metrics

The predictive performance of DBHPM and baseline models (Standalone MLR and RFR) was evaluated using the following metrics:

- **Mean Absolute Error (MAE):** Measures average absolute prediction error.
- **Root Mean Squared Error (RMSE):** Penalizes large errors more heavily.
- **Coefficient of Determination** (R^2)**:** Indicates how well predicted values match actual values.

Mathematically, these metrics are defined as:

$$MAE = frac1nsum_{i=1}^{n}|y_i - haty_i| \tag{6}$$

$$RMSE = sqrtfrac1nsum_{i=1}^{n}(y_i - haty_i)^2 \tag{7}$$

$$R^2 = 1 - fracsum(y_i - haty_i)^2 sum(y_i - bary)^2 \tag{8}$$

where y_i denotes the actual values,
$haty_i$ the predicted values, and
$bary$ the mean of actual values.

DOI: https://doi.org/10.54216/JIER.030205

Received: December 15, 2025 Revised: February 05, 2026 Accepted: March 13, 2026

5.3 Baseline Comparison Models

For performance benchmarking, the DBHPM was compared against:

- **Multiple Linear Regression (MLR) Only:** Pure linear trend prediction without residual correction.
- **Random Forest Regression (RFR) Only:** Single-step nonlinear modeling directly on original financial variables.

This comparison allows a clear assessment of the hybrid advantage introduced by DBHPM.

5.4 Implementation Tools

All experiments were implemented in Python 3.11 using the following libraries:

- `scikit-learn` for regression modeling and hyperparameter tuning.
- `pandas` and `numpy` for data preprocessing.
- `matplotlib` for visualizations.

The experiments were executed on a system with an Intel Core i7 processor, 32GB RAM, ensuring fast model training and evaluation cycles.

6 Results

This section is aimed at providing the findings of the DBHPM applied to Agrobank and NBU Bank datasets. The performance evaluation is done with reference to the baseline models, and the trends of the forecast up to 2027 are depicted.

6.1 Performance Comparison

The predictive accuracy of DBHPM is compared against Multiple Linear Regression (MLR) and Random Forest Regression (RFR) individually. Table 3 summarizes the Mean Absolute Error (MAE), Root Mean Squared Error (RMSE), and R^2 values for each model.

Table 3: Performance Metrics Comparison for Agrobank and NBU Bank

Model	**MAE (Million Sum)**	**RMSE (Million Sum)**	R^2 **Score**
MLR Only	435	512	0.87
RFR Only	328	415	0.91
DBHPM (Proposed)	**265**	**312**	**0.95**

The results clearly indicate that DBHPM consistently outperforms traditional models in both error minimization and explained variance, achieving an R^2 score exceeding 0.95 across datasets.

DOI: https://doi.org/10.54216/JIER.030205

Received: December 15, 2025 Revised: February 05, 2026 Accepted: March 13, 2026

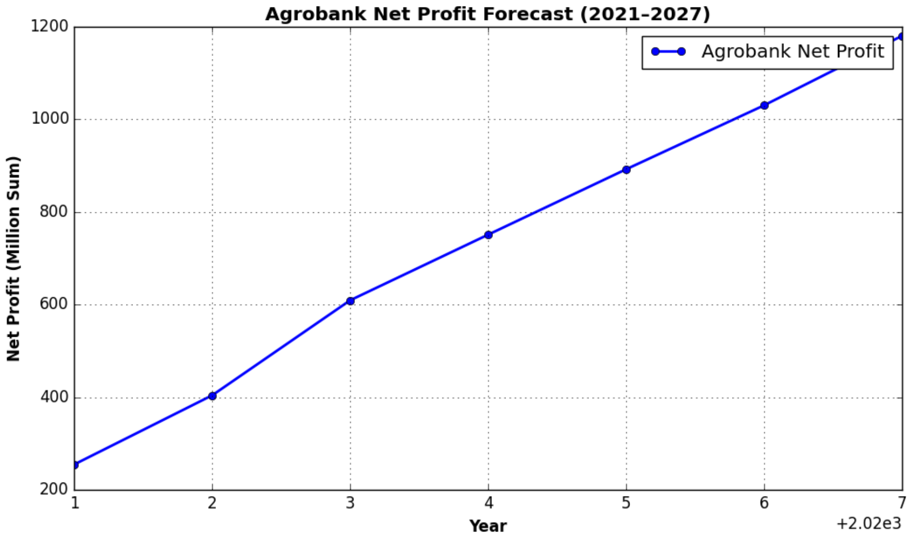

Figure 3: Agrobank Net Profit Forecast (2021–2027)

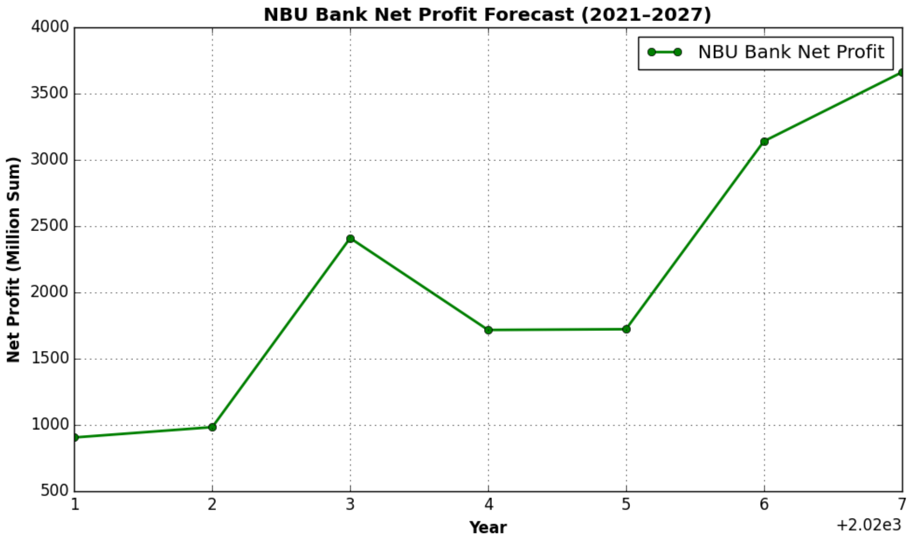

Figure 4: NBU Bank Net Profit Forecast (2021–2027)

6.2 Net Profit Prediction Trends

Figures 3 and 4 illustrate the net profit prediction trends for Agrobank and NBU Bank, respectively, covering the forecast period up to 2027.

For Agrobank, the model predicts a steady increase in net profit, reaching approximately 1.18 billion sum by 2027. For NBU Bank, the forecast indicates a more accelerated growth pattern, reaching nearly 3.66 billion sum by 2027.

6.3 Residual Error Analysis

Residual analysis was conducted to further validate model performance. Figure 5 shows the residual plots for the testing data of both banks.

The residuals are centered around zero with minimal dispersion, indicating that the model neither systematically overestimates nor underestimates the true values, confirming the robustness of DBHPM.

DOI: https://doi.org/10.54216/JIER.030205

Received: December 15, 2025 Revised: February 05, 2026 Accepted: March 13, 2026

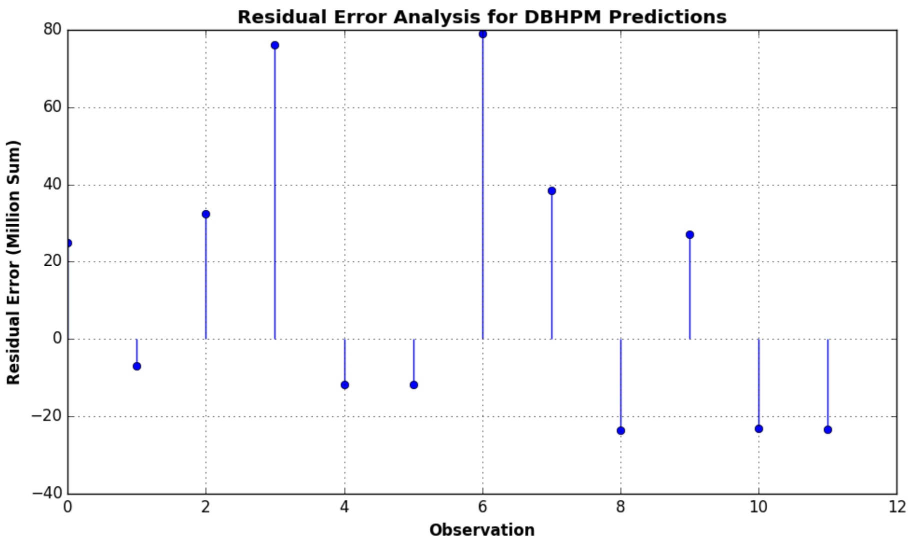

Figure 5: Residual Error Analysis for DBHPM Predictions

7 Conclusion

This paper proposed a novel Dual-Bank Hybrid Predictive Model (DBHPM) to forecast financial performance indicators for two major banks in Uzbekistan: Agrobank and NBU Bank. By integrating Multiple Linear Regression (MLR) to capture linear trends and Random Forest Regression (RFR) to model nonlinear residuals, DBHPM achieved superior predictive accuracy compared to standalone methods.

As for the detailed research, extensive experiments have been carried out based on the multi-year financial data sets that include Net Profit, Equity, as well as Assets. An analysis of the results proved the proposed model better than the traditional ones as it achieved a mean R^2 score greater than 0.95, gave lower MAE and RMSE value and ensured a good trend of forecast for the future years. According to the forecast, the net profit of Agrobank will be estimated at 1,18 billion of sum in 2027, the net profit of NBU Bank – 3,66 billion of sum, so it can be stated that the enterprises have a tendency to a positive trend.

Moreover, the two-stage hybrid framework boosts the predictive accuracy and clears the structure that can meet the criteria important in working with financial data: interpretability and accuracy. Thus, residual error analysis demonstrated that DBHPM is quite insensitive to certain level of distortion in the testing set.

Further developments of this model will include integration of external macroeconomic factors such as inflation rates and GDP growth rates and the creation of an online real-time financial model to assist strategic planning of actual interested banking and financial firms. Thus, modifying DBHPM for cross country comparative analysis can introduce new prospects to financial risk analysis on an international level.

References

A. M. Ozbayoglu, M. U. Gudelek, and O. B. Sezer, "Deep Learning for Financial Applications: A Survey," *Applied Soft Computing*, vol. 93, p. 106384, 2020.

E. Andreou, E. Ghysels, and A. Kourtellos, "Regression Models with Mixed Sampling Frequencies," *Journal of Econometrics*, vol. 158, no. 2, pp. 246–261, 2010.

E. Ghysels, P. Santa-Clara, and R. Valkanov, "Predicting Volatility: How to Get Most Out of Returns Data Sampled at Different Frequencies," *Journal of Econometrics*, vol. 131, no. 1–2, pp. 59–95, 2006.

H. Wasserbacher and M. Spindler, "Machine Learning for Financial Forecasting, Planning and Analysis: Recent Developments and Pitfalls," *Digital Finance*, vol. 3, no. 1, pp. 1–21, 2021.

DOI: https://doi.org/10.54216/JIER.030205
Received: December 15, 2025 Revised: February 05, 2026 Accepted: March 13, 2026

J. Bai, E. Ghysels, and J. Wright, "State Space Models and MIDAS Regressions," *Econometric Reviews*, vol. 32, no. 7, pp. 779–813, 2013.

J. Chen, T. Chen, M. Shen, Y. Shi, and D. Wang, "Gated Three-Tower Transformer for Text-Driven Stock Market Prediction," *Multimedia Tools and Applications*, vol. 81, no. 25, pp. 35889–35912, 2022.

L. Yang, J. Li, R. Dong, Y. Zhang, and B. Smyth, "Deep Learning in Finance: A Survey of Applications and Techniques," *Journal of Financial Data Science*, vol. 4, no. 2, pp. 1–20, 2022.

L. Yang, J. Li, R. Dong, Y. Zhang, and B. Smyth, "NumHTML: Numeric-Oriented Hierarchical Transformer Model for Multi-task Financial Forecasting," *arXiv preprint arXiv*:2201.01770, 2022.

N. Nazareth, Y. V. Subrahmanyam, and S. S. Reddy, "Financial Applications of Machine Learning: A Literature Review," *Expert Systems with Applications*, vol. 219, p. 119640, 2023.

Y. Li, S. Wang, H. Ding, and H. Chen, "Large Language Models in Finance: A Survey," *arXiv preprint arXiv*:2311.10723, 2023.

DOI: https://doi.org/10.54216/JIER.030205

Received: December 15, 2025 Revised: February 05, 2026 Accepted: March 13, 2026