Journal of Cyber-Physical Security and Robotics
Volume 02, Issue 1, 11-20, 2026

Research Article

Intelligent Mortgage Rate Forecasting Using AI: Enhancing Accuracy, Governance, and Financial System Resilience

Varun Bitkuri
Masters in Software Engineering, Stratford University, USA

Received: 18th January, 2026 Accepted: 19th March, 2026

Abstract

The mortgage rate forecasting is a difficult task because financial markets are subject to complex interaction of macroeconomic indicators and market dynamics. In this paper, an intelligent AI-based prediction system in mortgage rates prediction is proposed based on Multisource Stock Market Trends dataset that involves stock prices, economic variables, and sentiment variables. The processing of the data is done via the treatment of the missing values, encoder of the sentiment, feature engineering (lag, rolling, and interaction features), StandardScaler scaling, and the time order division of the data. In order to achieve better predictive performance, a hybrid ensemble model is resorted to that is trained with XGBoost and LightGBM as well as a meta-learner which is stacking-based. The experimental findings which show that the hybrid model is superior to the base models since it offers a superior accuracy with a higher value of R2 of 0.9799, lower value of RMSE of 0.1920, lower value of MAE of 0.1462 and less value of M APE of 2.69%. Moreover, the study employs explainable AI (SHAP), stress testing, and stability testing with the aim of ensuring transparency and strength. The findings reveal how the proposed framework can be useful in improving the precision of the prediction and facilitating the implementation of the model governance and resilience evaluation AI-based mortgage rate forecasting.

Keywords: financial system, multisource stock market trends dataset, mortgage forecasting, hybrid model, machine learning, governance framework.
Doi: 10.64235/sw1r1761

Introduction

There is a strong relationship between stability within the financial system and also on the dynamics within the mortgage markets because any change of the mortgage rates directly influences the demand of a housing, credit growth and overall economic performance [1] [2]. Mortgage rate forecasting is therefore significant in ensuring that the financial institutions, the policy makers and investors forecast on the direction of the market to prevent the risk of systemic [3]negative equity has emerged as a significant concern for homeowners, lenders, and policymakers alike. This phenomenon, characterized by homeowners owing more on their mortgages than the current value of their homes, can have far-reaching economic and social implications. The main goal of this research project was to develop machine learning models that can effectively predict negative equity trends in U.S. housing markets. This involved a multi-faceted approach that encompasses data collection, model development, and validation to ensure the accuracy and reliability of predictions. The historical housing market data used for this research covers various regions across the United States, from urban to suburban and rural, to provide diversified dynamics in the markets. The dataset utilized for this analysis comprises a comprehensive collection of variables relevant to understanding negative equity trends in the U.S. housing market. It includes historical housing prices, which reflect property values across various regions, mortgage rates that provide insights into borrowing costs, and key economic indicators such as employment rates, inflation, and consumer confidence indices. The data has been sourced from reputable platforms, including public records from county assessors, real estate platforms like Zillow and Redfin for transaction data, and government databases such as the Federal Housing Finance Agency (FHFA[4]. The concentration of the dynamics of mortgage rates that is extended by the macroeconomic variables such as inflation, monetary policy decisions and international financial shocks in whichever way introduce enormous difficulties to the traditional forecasting models [5][6]. The interest-rate forecast is a central aspect of macroeconomic and financial analysis whose outputs are used by central banks to make monetary-policy decisions, institutional investors to make portfolio decisions, and corporate treasuries to hedge the interest-rate risk. Conventional econometric models classical ARIMA, GARCH, and conditional-mean and conditional-heteroskedasticity models have not been very flexible, and sometimes cannot adapt to the historical speed of structural changes in the global economy and financial markets that have been promoted by technological advancement and the changing behavior of market participants. Their interest-rate forecasting ability is an unexplored field even though interest rate information is considered to be a new asset class and vital resource of effective contracting and risk delegation of both corporate community and capital markets [7][8]. Mortgages have a special interest since it has two characteristics, which are huge loan amounts, and long term. These features are equal to huge losses in financial terms in case of default.

The new emerging technology that has transformed the game of mortgage rate forecasting is the Artificial Intelligence (AI) which offers advanced capabilities of estimating the complex, nonlinear relationship and process large financial data. Machine learning models that have been found to have superior performance in determining the temporal patterns and latent structures of financial time series are gradient boosting, random forests, and deep learning (LSTM networks) models [9][10]data governance is critical for ensuring regulatory compliance, data quality, and user trust. Traditional rule-based systems are often rigid, unable to cope with the dynamic and heterogeneous nature of modern data ecosystems. This paper presents an AI-driven data governance framework designed to automate policy enforcement, detect anomalies, and ensure continuous compliance across complex infrastructures. Leveraging machine learning and natural language processing, the system can adapt to evolving regulatory requirements, perform real-time data classification, and recommend corrective actions. Our proposed solution demonstrates significant improvements in compliance assurance, data quality scores, and governance efficiency. Experimental results across multi-cloud datasets reveal a 92% accuracy in detecting policy violations and a 38% reduction in manual auditing tasks, illustrating the transformative potential of AI in governance landscape s.”,”author”:[{“dropping-particle”:””,”family”:”Singamsetty”,”given”:”Sudheer”,”non-dropping-particle”:””,”parse-names”:false,”suffix”:””}],”container-title”:”International Journal of Computational Mathematical Ideas”,”id”:”ITEM-1”,”issue”:”03”,”issued”:{“date-parts”:[[“2021”]]},”page”:”1007-1017”,”title”:”AI-Based Data Governance: Empowering Trust and Compliance in Complex Data Ecosystems”,”type”:”article-journal”,”volume”:”13”},”uris”:[“http://www.mendeley.com/documents/?uuid=d02ccede-cf55-461b-8c22-252c95dd3b02”]}],”mendeley”:{“formattedCitation”:”[10]”,”plainTextFormattedCitation”:”[10]”,”previouslyFormattedCitation”:”[10]”},”properties”:{“noteIndex”:0},”schema”:”https://github.com/citation-style-language/schema/raw/master/csl-citation.json”}. In addition to that, AI enables the incorporation of alternative sources of data including textual data, market sentiment, and macroeconomic indicators, which enhances the power and credibility of the forecasting. This goes a long way in creating smart forecasting systems that would contain the disadvantages of the traditional approaches. Ways of governing AI-based financial systems are explainable models, rigorous validation procedures, and regulatory compliance to mitigate the risk of bias, overfitting, and model opaquity. Coupled with a blend of better accuracy and good governance.

Motivation and Contributions of the Study

The proper forecasting of mortgage rates is of paramount importance to financial institutions, policy-makers and investors to control the risk and maintain financial stability. The traditional statistical models do not tend to deal with the non-linear complex relations and dynamic interactions between macroeconomic and financial variables. With the increase in the instability of the market and economic uncertainties, there is a demand to find intelligent forecasting approaches that may enhance the quality of prediction besides ensuring transparency, strength and resiliency. This has resulted in the consideration of advanced AI and ensemble models in order to generate a justifiable and explainable mortgage rate prediction model. The following are the important contributions of this paper

- Used the Multisource Stock Market Trends dataset that entails stock, economic, and sentiment data to analyze them thoroughly.
- Systematic preprocessing that is applied such as the handling of missing values, encoding, and feature engineering, feature scaling using StandardScaler.
- Suggested a composite version of XGBoost and LightGBM but with a Ridge meta-learner to have better predictions.
- Evaluated the performance of the model through various measures, such as R^2 score, RMSE, MAE, and MAPE, residual analysis and confidence intervals.
- Transparency and reliability Applied SHAP explainability, stress testing and stability checks.
- Monitored the soundness of the model in diverse economical conditions to facilitate the stability of the financial system.

Justification and Novelty

This study is justified by the fact that while predicting mortgage rates with accuracy and reliability is required in the case of intricate and dynamic interactions among financial markets and macroeconomic variables, conventional models are inefficient in many cases.

The work is novel, as it combines a hybrid ensemble model that uses XGBoost and LightGBM and stacking-based meta-learner, deep data preprocessing, feature engineering, and explainable artificial intelligence. Also, the presence of governance mechanisms in the form of SHAP-based interpretability, stress testing and stability analysis makes this framework more realistic in the real world financial application as it does not only enhance predictive capability but also provides transparency and resilience of financial systems.

Organization of the paper

The structure of the paper is the following: Section II provides a literature review of AI-based mortgages rate forecasting and financial modeling; Section III explores the dataset and methodology; Section IV presents the experimental results and discussion; and the conclusion with the final remarks about the significant findings and future research possibilities, is given in the Section V.

Literature Review

In this section, the author will provide an extensive discussion concerning the modern situation in the field of mortgage rate forecasting and the usage of the machine learning methods to enhance the accuracy of prediction, governance of the model, and the financial system.

Poonia et al. (2025) the capacity to foresee and lessen the impact of financial crises is crucial to the resilience of the global economy. Prepare the data for analysis by preprocessing it; small and medium-sized enterprises (SMEs) in business-to-business and consumer-focused sectors are particularly susceptible to market changes. To identify crises, Principal Component Analysis (PCA) is used to extract the most important information. The NRFO-SANN model does not just perform better. Having a remarkable 96% success rate [11]especially for small and medium-sized enterprises (SMEs. Krasoytskyi and Stavtskyy (2024) outline a number of methods that financial institutions can use to predict mortgage defaults. These include decision trees, linear discriminant analysis, support vector machines with linear kernels, neural networks, and real GDP growth and debt-service-to-income ratio (DSTI). Striking a balance between actual default precision and minimizing false defaults is essential. Additionally useful are GDP projections and accurate evaluations of the borrower's DSTI based on loan details like age and interest rate. the likelihood of default is influenced by the remaining term of mortgage loans [12]while regulators monitor systemic risk, which this sector may possess. This research is focused on predicting defaults on a one-year horizon using data from the Ukrainian credit registry applying machine-learning methods. This research is useful for not only academia but also policymakers since it helps to assess the need for implementation of macroprudential instruments. We tested two data balancing techniques: weighting the original sample and synthetic minority oversampling technique and compared the results. It was found that random forest and extreme gradient-boosting decision trees are better classifiers regarding both accuracy and precision. These models provided an essential balance between actual default precision and minimizing false defaults. We also tested neural networks, linear discriminant analysis, support vector machines with linear kernels, and decision trees, but they showed similar results to logistic regression. The result suggested that real gross domestic product (GDP. Kündig, Purohit, and Verma (2024) propose machine learning methods for predicting mortgage meltdown based on borrower, property, market, and loan factors. A large dataset of locked mortgage applications from a prominent U.S. lender was examined. In terms of both predicted efficacy and stability, the random forest model was clearly the best. Using empirical research with logistic regression and the SHAP approach, identified key predictors of mortgage fallout [13]. Barbaglia, Manzan and Tosetti (2023) examine loan default behavior in multiple European countries. The study examines the occurrence of default as a function of borrower characteristics, loan-specific variables, and local economic conditions. A set of machine learning algorithms outperforms logistic regression in explaining loan defaults, with interest rates and local economic characteristics proving the most compelling factors. Compared to the American mortgage market, the European market is characterized by lower Loan-to-worth (LTV) ratios, with most lenders limiting loans to 80% of the property's worth [14]. Tewar (2022) suggested a learning-driven mortgage pricing framework that harnesses advanced data engineering and artificial intelligence to enable real-time, personalized lending decisions. Taking advantage of feature engineering, time-sensitive data organization, and live streaming. These frameworks self-tune complicated nonlinear profiles linking risk and cost sensitivity, to optimize mortgage rates whilst staying by regulatory criteria and risk controls AI-based mortgage pricing to reinvent conventional lending, minimize bias, and enhance financial inclusion using responsive and data-driven decision mechanisms [15].

Research gap

The current research in mortgage rate prediction is mainly based on the conventional statistical models or independent machine learning methods, which are not as effective as complex nonlinear relations and dynamic interactions between financial and macroeconomic variables. In addition, minimal focus has been put in model governance, interpretability and stability in

Table 1: Summary of Related Studies on AI-Driven Mortgage Rate Prediction and Financial System Stability

Author	*Dataset*	*Methodology*	*Key Findings*	*Challenges*	*Future Work*
Poonia et al. (2025)	Financial crisis & SME-related datasets	PCA + NRFO-SANN model	Achieved high prediction accuracy (~96%) for financial crisis detection; PCA improved feature extraction	Sensitivity to market fluctuations; limited generalization across sectors	Improve robustness across diverse economic environments
Krasoytskyi & Stavtskyy (2024)	Mortgage loan data (GDP, DSTI, borrower info)	Neural Networks, LDA, SVM, Decision Trees	GDP growth and DSTI are strong predictors of mortgage default; borrower characteristics improve prediction	Balancing accuracy vs false defaults; data dependency	Enhance macroeconomic forecasting integration
Purohit & Verma (2024)	Large U.S. mortgage dataset	Random Forest, Logistic Regression, SHAP analysis	Random Forest showed better performance; SHAP identified key influencing factors	Model interpretability; dependence on large datasets	Improve explainability and real-time prediction models
Barbaglia et al. (2023)	European mortgage datasets	ML models vs Logistic Regression	ML models outperform traditional methods interest rate & economic conditions are key factors	Variability across regions; limited standardization	Expand cross-country comparative models
Tewar (2022)	Mortgage pricing and lending datasets	AI-driven pricing models, feature engineering, real-time analytics	Enables personalized lending, reduces bias, improves pricing efficiency	Regulatory constraints; real-time data complexity	Develop adaptive and explainable AI pricing systems

different economic situations. The integrated frameworks that unite advanced feature engineering, ensemble learning, and explainable AI are also lacking. This study fills these gaps with the suggestion of a hybrid, interpretable, and resilient AI-based forecasting model. The research summary comparison is provided in Table I based on data, results, and limitations.

Methodology

The illustrated methodology is a structured pipeline of mortgage rate forecasting as depicted in Fig. 1.

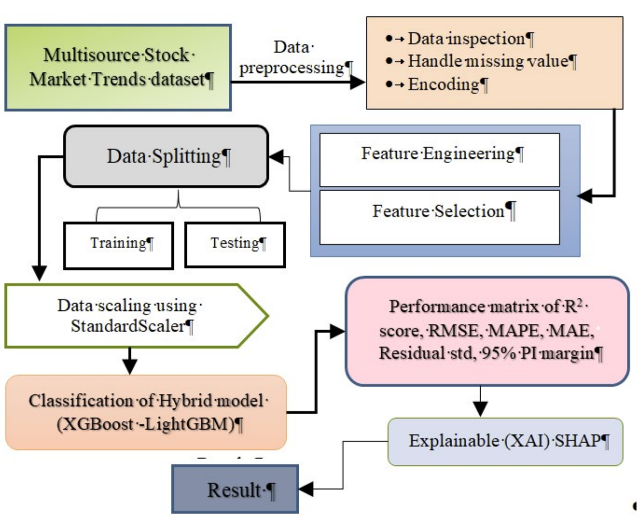

Figure 1: Flowchart of the hybrid mortgage rate forecasting model.

The Multisource stock market trends data is initially preprocessed by the inspection of data, data missing and encoding of categorical data. The process of feature engineering is used to create lag, rolling and interaction features, and the selection of features is then done through correlation and importance. The data is then chronologically divided into training (80) and testing (20) and scaled using StandardScaler. A hybrid ensemble model of an XGBoost and LightGBM is trained. R^2, RMSE, MAE, MAPE, the residual standard deviation of the results, and the 95% prediction intervals were used to measure model performance, and SHAP was applied to interpret the model performance.

Data Collection

The dataset used in this study is obtained from Kaggle and is based on the Multisource Stock Market Trends dataset, which combines financial market data and macroeconomic indicators. It includes stock price features

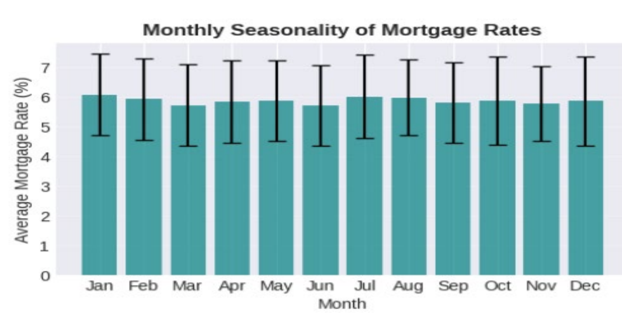

Figure 2: Monthly seasonal variation in mortage rates

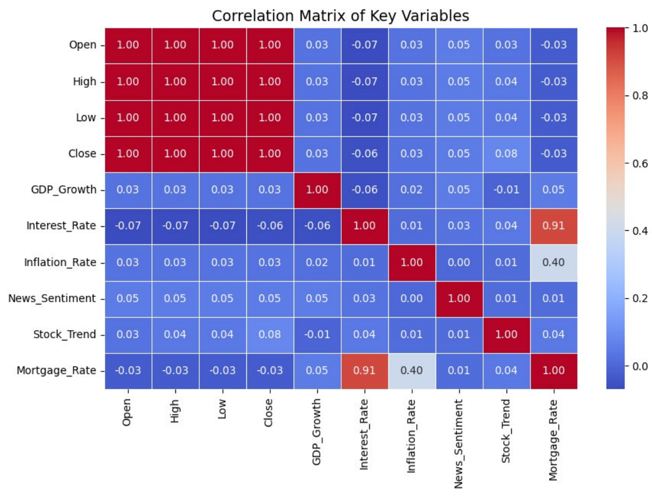

Figure 3: Correlation Heatmap of different variables

(Open, High, Low, Close), economic variables such as GDP growth, interest rate, and inflation rate, along with derived features like news sentiment and stock trends. The wide variety of data allows to thoroughly study the behavior of the market and make predictions of mortgage rates correctly as it reflects both the economic and financial processes.

The monthly seasonality of the mortgage rates and the average rates are also relatively steady with slight variations in all months in Fig. 2. Error bars are used to indicate variability, which implies that the mortgage rates show tendencies and are consistent within the year, without any major trends in the season.

Correlation matrix demonstrates the correlation between the critical variables in Fig. 3 where the correlation between the stock price features (Open, High, Low and Close) is both strongly and positively correlated and the correlation between the Interest rate and Mortgage rate (≈0.91). Mortgage rate is averagely correlated with Inflation rate (correlation= about 0.40) and weak or non-significant correlations with other variables.

Data Preprocessing

Data Preprocessing Module cleans the data and handles missing values, encodes values, and scales values. The next steps involve feature engineering and data splitting that will enhance the performance of the model.

Data inspection

The first thing to do with the dataset is to investigate its structure, types of features, and summary statistics to be able to learn about the data quality and distribution.

Handle Missing value

The treatment of missing values is undertaken through the forward fill whereby missing values are filled with the last valid observation to maintain the temporal continuity. This method makes time-series data consistent and will not lose valuable sequential data during model training.

Encoding

The News_Sentiment feature is encoded, i.e. the categorical values are transformed into numerical ones, positive is set to 1, neutral to 0.5, and negative to 0.

Feature engineering

The feature engineering can be described as transforming unprocessed data to useful inputs to improve a model new features are engineered with lag variables (last mortgage rates), rolling statistics (mean and standard deviation) and interaction terms, such as interest rate and inflation, to assist the model to replicate the temporal patterns and the complex relationship in the mortgage rate forecasting.

Feature selection

The process of feature selection presents the most pertinent variables that matter a lot in mortgage rate prediction leaving out the redundant or less informative variables like interest rate, inflation, GDP growth, stock indicators, and engineered variables are then chosen after conducting correlation analysis and importance measures to optimize the accuracy of the model and make the model less complex.

Data splitting

The dataset will be divided according to the chronological structure in order to retain the time-series structure, 80% will be used in the training and 20% in the testing. This gives a size (776, 26) training set and size (195, 26) test set, which makes sure that the model is tested on untested future observations to make sound mortgage rates predictions.

Data scaling using standardscaler

StandardScaler is used to scale the data and normalize the input features to zero mean and unit variance. This is particularly to ensure that it is consistent across models including XGBoost and LightGBM, which would mean that training is stable, and predictive is better.

Proposes Models

In this section it gives the theoretical explanation about the DL and ML methods that were used in this study.
Hybrid model (XGboost+LightGBM)

XGBoost

XGBoost (Extreme Gradient Boosting) is a powerful ensemble ML algorithm or a solution that is based on gradient boosting and whereby multiple decision trees are developed successively to enhance the accuracy of the predictions. The new trees aim at correcting the previous one errors by minimizing a loss function.

LightGBM

LightGBM (Light Gradient Boosting Machine) is an effective gradient boosting algorithm that constructs decision trees with a leaf-wise growth method that enables LightGBM to have more rapid training and accuracy than traditional methods. LightGBM is efficient, fast, and has the capacity to learn and predict complicated trends, which makes it suitable when it comes to financial prediction duties like that of mortgage rates.

The hybrid ensemble model is proposed based on the complementary features of XGBoost and LightGBM to identify the dynamics of the complex nonlinear relationship of mortgage rate dynamics. The XGBoost model is configured with n_estimators = 500, max_depth = 10, learning_rate = 0.03, subsample = 0.85, colsample_bytree = 0.85, reg_alpha = 0.1, reg_lambda = 1.0, and min_child_weight = 3, enabling it to model deep feature interactions and reduce bias. In contrast, the LightGBM model employs n_estimators = 500, max_depth = 8, learning_rate = 0.03, num_leaves = 31, subsample = 0.85, colsample_bytree = 0.85, reg_alpha = 0.1, reg_lambda = 1.0, and min_child_samples = 20, allowing efficient leaf-wise learning and improved variance control. The predictions are merged with optimized weighted averaging (w_GB=0.55,w_GB=0.45 and further refined with a Ridge regression meta-learner), which also learns coefficients (β1=0.52, β2=0.48) to improve the generalization to show better accuracy, stability and robustness in mortgage rates forecasting.

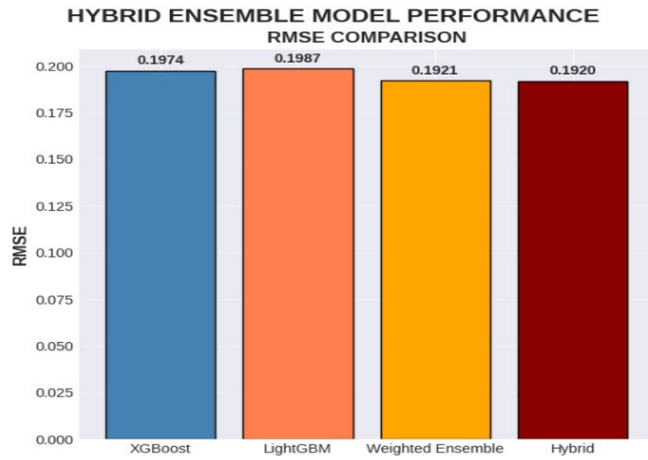

Figure 4: RMSE Analysis of individual and ensemble model

Table 2: Experiment Results of Proposed Models for mortgage rate forecasting

Model	*XGBoost*	*LightGBM*	*Weighted ensemble*	*Hybrid model*
R^2	0.978805	0.978529	0.979943	0.979952
RMSE	0.197446	0.198727	0.192074	0.192031
MAPE	2.837254	2.771865	2.720117	2.69692
MAE	0.153741	0.149132	0.147394	0.14624
Residual std	0.214165	0.194732	0.190538	0.190262
95% PI margin	0.418685	0.380695	0.372496	0.371957

Table 3: Residual Diagnostic Results for AI-Based Mortgage Rate Models

Ljung-Box	
Statistical value	P-value
11.1701	0.3444
11.9692	0.2871
10.0120	0.4394
9.9283	0.4468

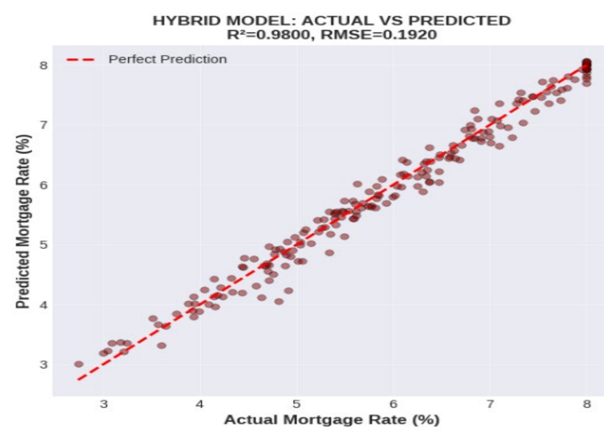

Figure 5: Actual vs predicted in hybrid model

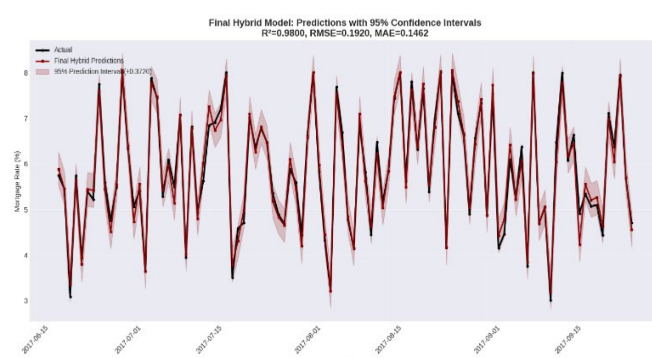

Figure 6: Performance of hybrid model for mortgage rate forecasting

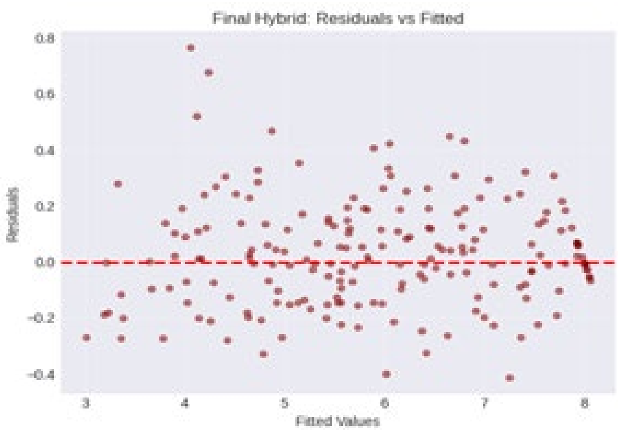

Figure 7: Residual vs Fitted in hybrid model

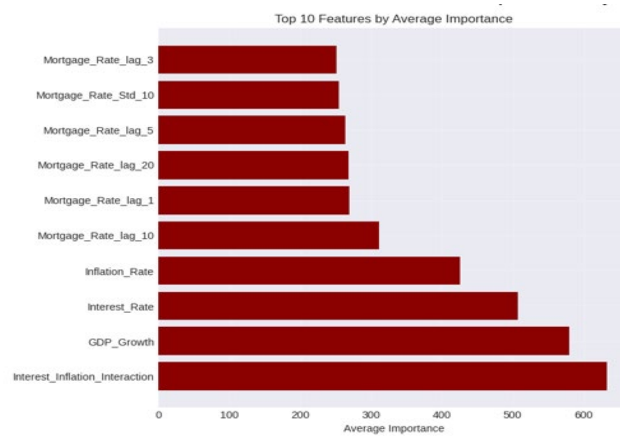

Figure 9: Top 10 features by average importance

Model Evaluation

The performance of the proposed model is evaluated using regression metric such as R^2 score, RMSE, MAE, MAPE and residual std and 95% PI margin, which effectively measure prediction accuracy and error in mortgage rate prediction (Equations (1) to (6)).

$$R^2 = 1 - \frac{\sum_{i=1}^{n}(actual - predicted)^2}{\sum_{i=1}^{n}(actual - mean\ of\ actual)^2} \quad (1)$$

$$MAE = \frac{1}{n}\sum_{i=1}^{n}|actual - predicted| \quad (2)$$

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2} \quad (3)$$

$$MAPE = \frac{1}{n}\sum_{i=1}^{n}\left|\frac{y_i - \hat{y}_i}{y_i}\right| \times 100 \quad (4)$$

$$\sigma_r = \sqrt{\frac{1}{n-p}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2}. \quad (5)$$

$$PI\ Margin = 1.96.\sigma_r \quad (6)$$

The overall performance is evaluated using multiple metrics. These measures collectively assess accuracy, error consistency, and prediction reliability of the forecasting model.

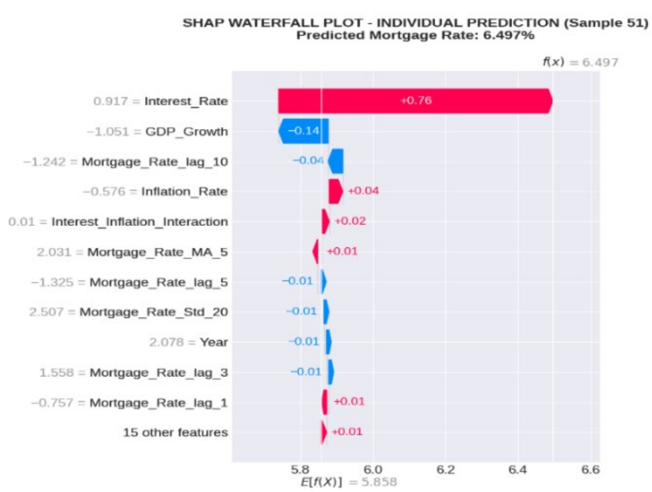

Figure 8: SHAP waterfall plot individual mortage rate prediction

Result Analysis and Discussion

The experimental evaluation of the proposed AI-based mortgage rate forecasting models was conducted using the Multisource Stock Market Trends Dataset, which integrates diverse financial It is implemented on a local computer with an Intel i5 (9th generation) CPU, 8 GB RAM, and NVIDIA GTX 1650 graphics card, and running on Windows 10. The Python language was used to build the models in the Google Colab and Visual Studio Code and the libraries used included Pandas, NumPy, Scikit-learn and Seaborn, among others. As seen in Table II, there is high predictive accuracy in all the models. Among the models, the hybrid approach is the best model since it promises the lowest values of error (MAE = 0.1462, RMSE = 0.1920, MAPE = 2.69) and highest R^2 (0.9799) value of residual standard deviation (0.1903) and a smaller confidence interval (approximately = 0.372) means that the forecasting model has the highest capabilities. The weighted ensemble model demonstrates a competitive performance as well, whereas separate models like XGBoost and LightGBM have a bit higher error rates.

RMSE values across models is shown in Fig. 4, showing

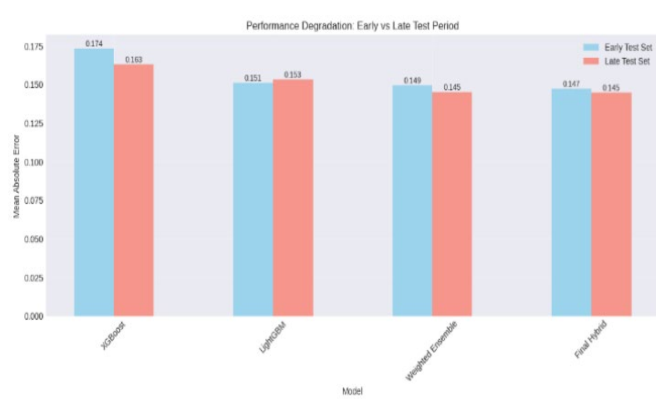

Figure 10: Performance Degradation of early and late test period

Table 4: Model Governance and Resilience Evaluation for AI-Based Mortgage Rate Forecasting

Scenario	*XGBoost Prediction*	*LightGBM Prediction*	*Hybrid Model Prediction*	*Deviation (%)*	*Stress Level*
Normal	5.810340	5.741426	5.779329	0.000000	Low
Financial Crisis (2008)	5.857510	5.743713	5.806301	0.466701	Low
High Inflation (1970s)	7.346004	7.303944	7.327077	26.780764	High
Tech Bubble Burst	5.846281	5.758421	5.806744	0.474364	Low
COVID-19 Pandemic	5.577096	5.377357	5.487214	-5.054480	Low
Rapid Recovery	6.060014	6.140923	6.096423	5.486692	Low
Stagflation	6.818634	6.770546	6.796994	17.608714	Medium

Table 5: Governance Framework and Model Validation Summary for AI-Based Mortgage Forecasting

Component	*Implementation Status*	*Priority*	*Method*	*Key Findings*
SHAP Explainability	Implemented / Completed	MUST	SHAP values computed for XGBoost and Linear models	Key features: Interest Rate, Inflation Interaction
Stress Testing	Implemented / Completed	MUST	Tested across 7 stress scenarios including crisis conditions	Max deviation: 26.8% under extreme scenarios
Scenario Analysis	Implemented / Completed	MUST	Sensitivity analysis across key macroeconomic variables	Mortgage rate highly sensitive to Interest Rate
Stability Check	Implemented / Completed	Important	PSI calculated for 26 features over time	No feature drift detected

that the hybrid and ensemble models achieve lower errors than individual models, indicating improved accuracy and more reliable mortgage rate prediction

The scatter plot compares actual and predicted mortgage rates using the hybrid model is shown in Fig. 5, showing points closely aligned along the perfect prediction line. This indicates high accuracy and strong model performance, supported by a high R^2 value and low RMSE.

The performance of the final hybrid model in forecasting mortgage rates, comparing actual values with predicted values along with a 95% confidence interval in Fig. 6. The forecast values are similar to the real trend throughout the period showing a high level of accuracy of the model with good performance indices (R^2 = 0.9800, RMSE = 0.1920, MAE = 0.1462). The small confidence intervals also show the reliability and consistency of the model.

The residuals versus fitted values plot for the hybrid model shows randomly distributed residuals around zero, indicating no clear pattern or heteroscedasticity in Fig. 7. This suggests that the model provides an adequate fit and maintains stability across predicted mortgage rate values.

The SHAP waterfall plot demonstrates the prediction of an individual mortgage rate indicating the relationship between each feature and the final value in Fig. 8. Interest Rate is the most significant positive research variable, and the growth of GDP and lag variables slightly decrease the prediction.

The top 10 features based on average importance in the model is shown in Fig. 9, that Interest–Inflation interaction, GDP growth, and Interest Rate are the most influential predictors, while lagged mortgage rate features contribute moderately to forecasting performance.

RMSE results of models in early and late test periods with minor differences in error in Fig. 10. The hybrid and ensemble models have relatively steady performances, which means that they are robust and slow to degrade with time.

The stress test analysis is used to test the strength of AI-driven mortgage rate forecasting models in Table III. across different economic conditions. The findings have shown that under normal and moderate stressing conditions, the models are stable and there are minor deviations that can be attributed to the reliability of the model in real-world financial forecasting.

Table 6: Conclusion and Future Scope

Model	*R2 score*	*RMSE*	*MAPE*
LSTM[16]	0.80	0.3482	5.59
CNN[17]	74.45	0.3026	7.52
SVR[18]	o.81	0.3495	27.5
Hybrid model	0.9799	0.1920	2.696

Table IV shows the results of the diagnostic of the residuals of the evaluated models. Residuals of most models are normally distributed except LightGBM and the hybrid model. All the models reveal, which proves the consistency and stability of the forecasting framework.

The governance framework merges explainability, robustness testing and stability testing to enhance reliability of AI-based mortgage rate forecasting. SHAP is used to identify the main drivers of the model, stress testing, and analysis of drift being used to maintain the stability of the model in different conditions. Table V results demonstrate better model governance and resilience.

Comparative Analysis

The comparison of various models in predicting the mortgage rate in terms of key performance measures in Table VI. The output consists of LSTM, CNN, SVR and the hybrid model proposed. The LSTM has a moderate performance with R^2 = 0.80 and RMSE = 0.3482 when compared to CNN, where the RMSE is lower (0.3026) but the MAPE is greater, which is the variability in the predictions. The SVR model shows a comparatively poorer performance when it has large error values. However, the proposed hybrid model has much higher R^2 score (0.9799), worse RMSE (0.1920), and MAPE (2.696) scores, which are displayed in better accuracy, stability, and effectiveness in forecasting mortgage rates.

Comparative Analysis of Models Using Performance Metrics on the Multisource Dataset

The proposed model is a step forward in predicting the mortgage rates using the state-of-the-art machine learning algorithms, such as a hybrid XGBoost/LightGBM, which ensures high accuracy and flexibility. It effectively models a complex relationship between financial and macroeconomic variables, has higher predictive accuracy through feature engineering, merges explainable AI and SHAP. The form of governance included within the framework is stress testing and stability testing which helps in the assurance of the forecasting being strong, clear and fit to be resilient to the actual world of financial decision making.

Mortgage rate forecasting is a difficult task because financial markets are driven by the complicated interactions of market dynamics and macroeconomic indicators. This paper outlines a detailed prediction model that uses machine learning and hybrid ensemble-based techniques in order to make a prediction more accurate and a financial system more stable. The findings indicate that the proposed hybrid model has a significant level of performance in comparison to the baseline models as it is highly accurate, has lesser prediction error, and is more stable. Further benefits of feature engineering and explainable AI (SHAP) make model interpretability more acceptable, whereas governance mechanisms, such as stress testing and stability analysis, ensure a robust performance of any ensemble model under different economic conditions. Further development will involve the incorporation of more effective deep learning methods, including LSTM and Transformer-based models, to enhance the significance of temporal relationships. Moreover, the use of real-time economic indicators, better uncertainty estimation and development of real-time forecasting systems will also increase the practical applicability and reliability of mortgage rate prediction models.

References

M. Bhatt and T. Kodinariya, "Mortgage rate forecasting using Box-Jenkins approach," in *Mathematical and Computational models*, 2019.

H. P. Kothandapani, "Advanced Artificial Intelligence Models for Real-Time Monitoring and Prediction of Macroprudential Risks in the Housing Finance Sector: Addressing Interest Rate Shocks and Housing Price Volatility to Support Proactive Decision-Making by Federal Agencies," *J. Artif. Intell. Res. https// thesciencebrigade. com/JAIR/article/view/507*, vol. 2, no. 1, pp. 391–416, 2022.

A. H. Jui *et al.*, "Understanding Negative Equity Trends in U.S. Housing Markets: A Machine Learning Approach to Predictive Analysis," *J. Econ. Financ. Account. Stud.*, vol. 5, no. 6, pp. 99–120, Nov. 2023, doi: 10.32996/ jefas.2023.5.6.10.

I. Nica, D. B. Alexandru, S. L. P. Crăciunescu, and Ștefan Ionescu, "Automated Valuation Modelling: Analysing Mortgage Behavioural Life Profile Models Using Machine Learning Techniques," *Sustainability*, vol. 13, no. 9, 2021, doi: 10.3390/su13095162.

M. A. Faheem, M. Aslam, and S. Kakolu, "Enhancing Financial Forecasting Accuracy Through AI- Driven Predictive Analytics Models," vol. 4, no. 12, pp. 322–328, 2021.

[6]S. B. Shah, "Hybrid Fake News Detection Using Machine Learning and Deep Learning," *Res Mil.*, vol. 13, no. 1, pp. 2858–2870, 2023.

S. G. Pillai, J. Woodbury, N. Dikshit, A. Leider, and C. C. Tappert, "Machine Learning Analysis of Mortgage Credit Risk," in *Proceedings of the Future Technologies Conference (FTC) 2019*, Cham: Springer International Publishing, 2020, pp. 107–123.

N. Kolli, J. W. Sajja, and A. Nerella, "Building Secure AI Agents for Autonomous Data Access in Compliance/ Regulatory-Critical Environments," *Comput. Fraud Secur.*, vol. 2024, no. 9, pp. 363–373, 2024, doi: 10.2139/ ssrn.5528763.

A. Sadhwani, K. Giesecke, and J. Sirignano, "Deep Learning for Mortgage Risk," *J. Financ. Econom.*, vol. 19, no. 2, pp. 313–368, 2020, doi: 10.1093/jjfinec/nbaa025.

S. Singamsetty, "AI-Based Data Governance: Empowering

Trust and Compliance in Complex Data Ecosystems," *Int. J. Comput. Math. Ideas*, vol. 13, no. 03, pp. 1007–1017, 2021, doi: 10.70153/IJCMI/2021.13301.

R. C. Poonia, P. Prabu, A. Maheshwari, A. Malhotra, and V. Gupta, "AI Enhanced Global Economic Resilience: Predicting and Mitigating Financial Crises," in *Proceedings of International Conference on Communication and Computational Technologies*, S. Kumar, S. Hiranwal, R. Garg, and S. D. Purohit, Eds., Singapore: Springer Nature Singapore, 2025, pp. 267–279. doi: 10.1007/978-981-97-7426-5_20.

Nadella, N. V. M. (2025a). Evolution of fault management in telecommunications: From reactive response to AI-driven predictive analytics. World Journal of Advanced Engineering Technology and Sciences, 15(2), 1338– 1344. https://doi.org/10.30574/wjaets.2025.15.2.0671

D. Krasovytskyi and A. Stavytskyy, "Predicting Mortgage Loan Defaults Using Machine Learning Techniques," *Ekonomika*, vol. 103, no. 2, pp. 140–160, 2024, doi: 10.15388/Ekon.2024.103.2.8.

Nadella, V. M. (2025). The backbone of modern society: How telecom technologies drive societal transformation. *World Journal of Advanced Research and Reviews*, *26*(2), 1425-1434.

P. Purohit and A. Verma, "Explainable Machine Learning for Fallout Prediction in the Mortgage Pipeline," *J. Risk Financ. Manag.*, vol. 17, no. 10, 2024, doi: 10.3390/jrfm17100431.

L. Barbaglia, S. Manzan, and E. Tosetti, "Forecasting Loan Default in Europe with Machine Learning," *J. Financ. Econom.*, vol. 21, no. 2, pp. 569–596, Mar. 2023, doi: 10.1093/jjfinec/nbab010.

Nadella, V. M. (2024). Al-Native 6G Network Management. *American International Journal of Computer Science and Technology*, *6*(1), 23-37.

R. Tewar, "Learning-Driven Mortgage Pricing: Data Engineering and AI for Real-Time Lending Decisions," *Int. J. Commun. Networks Inf. Secur.*, vol. 14, no. 3, pp. 1419–1437, 2022.

M. Bansal, A. Goyal, and A. Choudhary, "Stock market prediction with high accuracy using machine learning techniques," *Procedia Comput. Sci.*, vol. 215, pp. 247–265, 2022.

Nadella, V. M. (2025). AI/ML-Driven Service Assurance: 2024 breakthroughs transforming telecom operations. *Journal of Computer Science and Technology Studies*, *7*(4), 01-07.

M. Barua, T. Kumar, K. Raj, and A. M. Roy, "Comparative Analysis of Deep Learning Models for Stock Price Prediction in the Indian Market," *FinTech*, vol. 3, no. 4, pp. 551–568, Nov. 2024, doi: 10.3390/fintech3040029.

Sannapureddy, R., Nadella, V. M., & Nelavelli, S. (2024). Edge–Cloud Continuums for Latency-Sensitive Tasks. *International Journal of AI, BigData, Computational and Management Studies*, *5*(4), 189-201.

S. K. Lakshminarayanan and J. McCrae, "A comparative study of SVM and LSTM deep learning algorithms for stock market prediction," *CEUR Workshop Proc.*, vol. 2563, pp. 446–457, 2019.

Madhav Nadella, V. (2025). under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 (CC BY-NC-ND 4.0) International License Real-Time Monitoring Systems (5G/6G & Beyond): A Technical Review. *Sarcouncil Journal of Engineering and Computer Sciences*, *4*(08).