03/02/2025

CURRENCY PREDICTION REPORT

PREPARED BY:

BENSON NYOTA

1. ABSTRACT:

The uncertainty in future currency fluctuations poses significant challenges for businesses, particularly in international trade, where accurate forecasting of currency exchange rates is crucial for making informed decisions. Developing a predictive model that can estimate future currency prices is essential for businesses to plan effectively. This study presents a model trained on historical currency data, incorporating engineered features to capture underlying patterns and correlations that influence exchange rate movements. The results demonstrate the model's potential to provide valuable insights for businesses, especially in facilitating decision-making related to import and export activities. This predictive approach offers promising prospects for improving future business strategies in a volatile economic environment.

2. Introduction

2.1 Problem Statement

There is a significant demand for a predictive model that can forecast future currency exchange rates with minimal input, reducing the reliance on large amounts of complex or detailed data. In the real world, obtaining accurate and timely data about economic indicators, market sentiment, and geopolitical events can be challenging, especially for businesses with limited resources or access to such information. A model capable of predicting currency rates without requiring extensive data would be highly beneficial, as it could provide more accessible, real-time predictions, enabling businesses to make better decisions without the need for elaborate datasets or specialized knowledge. Developing such a model would be valuable not only for improving efficiency in decision-making but also for reducing the costs and complexities associated with gathering and analyzing large quantities of financial data.

3.0 Objectives

- **To develop a predictive model that can achieve high accuracy while requiring minimal input data.** The goal is to create a model that performs effectively with reduced information, simplifying the data collection process and making the model accessible for businesses with limited data resources. By optimizing the model to extract key insights from fewer inputs, it can be both efficient and practical for real-world applications.
- **To design a model capable of making reliable currency exchange rate predictions for the next 6-12 months.** The model aims to provide actionable insights for businesses by forecasting currency rates over a medium-term horizon, enabling better financial planning, risk assessment and strategic decision-making.
- **To ensure that the model can adapt to various market conditions and account for long-term trends.** The model will incorporate features that capture the underlying factors influencing exchange rate fluctuations, allowing it to provide consistent predictions even as market dynamics evolve.

4.0   Scope

The model is built by training on data sourced from various reliable external providers. It focuses on identifying and incorporating features that are strongly correlated with currency price fluctuations. While challenges may arise, such as obtaining high-quality data and performing effective feature engineering, these will be addressed throughout the process. An important assumption in the project is that each feature included is initially assumed relevant, with irrelevant features being filtered out during the evaluation phase.

5.0   Methodology

5.1 Time Series Overview

The dataset utilized for this analysis was sourced from Central Bank Of Kenya: https://www.centralbank.go.ke/rates/forex-exchange-rates/ about Key CBK Indicative Exchange Rates, spanning from October 2016 to January 2024, consisting of approximately 1,803 entries.

Given the time-series nature of the data, an Augmented Dickey-Fuller (ADF) test was conducted to assess whether the series was stationary. The ADF test results indicated non-stationarity, with a p-value greater than 0.05, leading to the **failure** to reject the **null hypothesis**.

To address the non-stationarity, the data was transformed through differencing. Additionally, feature engineering was performed to create new features, including Month, Day, Week of the Year, Quarter, and Day of the Week. The transformed data was then prepared for regression modeling.

Various regression algorithms were trained on the dataset, but none performed exceptionally well. Ridge regression emerged as the best model, achieving an accuracy of 3.58%. Attempts to apply SARIMAX on the differenced data, as well as ARIMA and SARIMA models on the original data, also yielded poor performance.

**5.2 Regression and Feature Importance**

Further feature engineering was conducted to enhance the dataset, incorporating additional variables such as Day of Year, Election Year, and whether it was a U.S election year. After training a variety of regression models—including *Ridge, Lasso, RandomForest, KNeighborsRegressor, SVR, XGBRegressor, ElasticNet, GradientBoostingRegressor, AdaBoostRegressor, LGBMRegressor,* and *MLPRegressor*—the accuracy improved to 28%, though it remained below the 50% threshold.

To improve the model's predictive capability, data from November 2024 to January 2025 was added, along with the incorporation of **US Interest Rate** data sourced from Federal Reserve Bank (https://fred.stlouisfed.org/series/DGS1).

The US Interest Rate exhibited a positive correlation with USD Exchange Rates over time, as illustrated in the graph below.

Currency Price(ksh) and US Interest rate over time
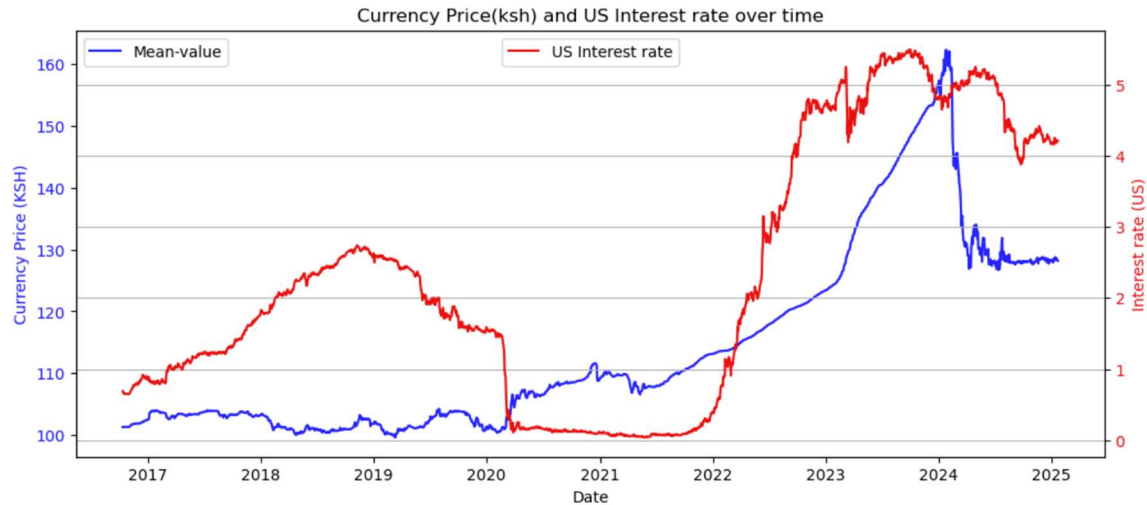
Additionally, Recursive Feature Elimination (RFE) was applied to identify the most significant features influencing the target variable. Both the LGBMRegressor and RandomForestRegressor models selected similar features but **ranked** them differently in terms of importance. The final selected features were: [US Interest Rate, Day of Year, Week of Year, Month, Election Year].
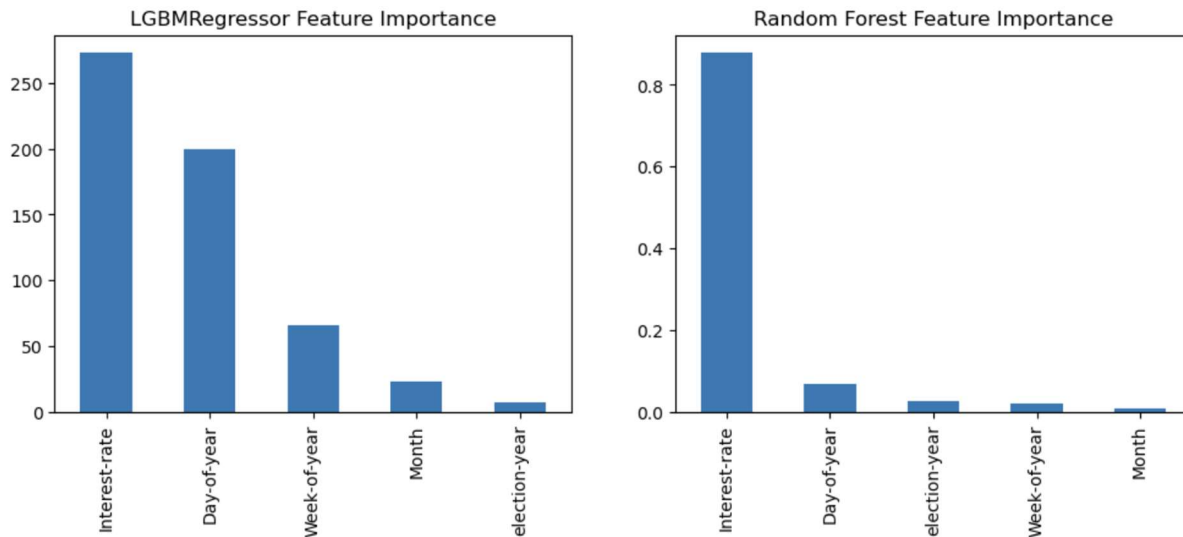

After training the model with the selected features, the accuracy increased to 87%, demonstrating a significant improvement and indicating that the US Interest Rate feature played a crucial role in determining currency prices.

5.3 Model Training and Prediction Results

At this stage, the models were trained using unscaled data and all available features in the dataset. Interestingly, this approach yielded the most accurate results, with the model achieving an impressive 96% accuracy after training.

Subsequent predictions were made using the X_test data, and these predictions were plotted alongside the actual values to visually assess the model's performance. The comparison clearly demonstrated how well the Random Forest model was able to predict currency exchange rates, reinforcing its reliability and effectiveness for this task.
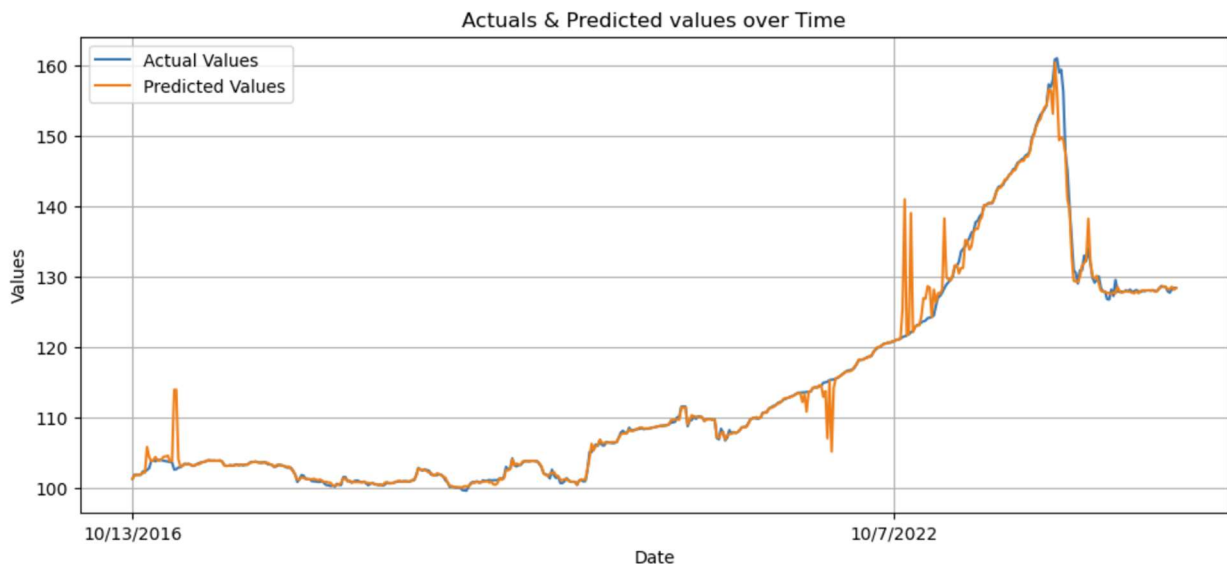
6.0   Discussion

In this project, the features selected were all deemed relevant to predicting currency exchange rates, and the models that performed the best were the **RandomForestRegressor** and **LightGBMRegressor**. These models demonstrated strong predictive capabilities and offered the highest accuracy.

Additionally, other models such as **XGBoost**, **GradientBoostRegressor**, and **AdaBoost** also showed promising results, with accuracies exceeding 90%, indicating that they effectively captured patterns in the data. While these models performed well, further optimization could potentially improve their performance.

For future work, the performance of these models can be revisited to explore whether additional tuning, feature engineering or different modeling techniques can lead to even better results. Furthermore, to enhance the accuracy and make the model more robust, sentiment analysis could be incorporated as an additional feature. By analyzing market sentiment or news articles in line with currency fluctuations, this new feature could provide deeper insights, further improving prediction accuracy.

7.0   Conclusion

This project successfully addressed the challenge of predicting currency exchange rates, a critical task for businesses operating in international markets, where currency volatility significantly impacts decision-making. By developing a model that could forecast future exchange rates with minimal input data, the

project aimed to simplify the prediction process. After performing data preprocessing, feature engineering, and training multiple regression models, it was clear that certain features—such as Interest Rates, time-related factors (e.g., Day of Year, Month) and Election Year played a crucial role in improving prediction accuracy.

The positive correlation between US Interest Rates and USD Exchange Rates highlighted the value of incorporating macroeconomic indicators into the models. Future work could explore additional factors, such as sentiment analysis or external economic indicators, to enhance the robustness of the predictions.

Overall, this project demonstrates the potential of machine learning to aid businesses in navigating the uncertainties of currency fluctuations, ultimately improving strategic decision-making in a dynamic global economy.