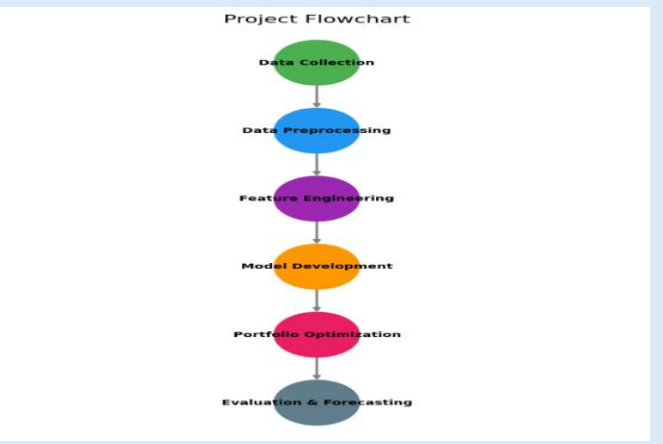
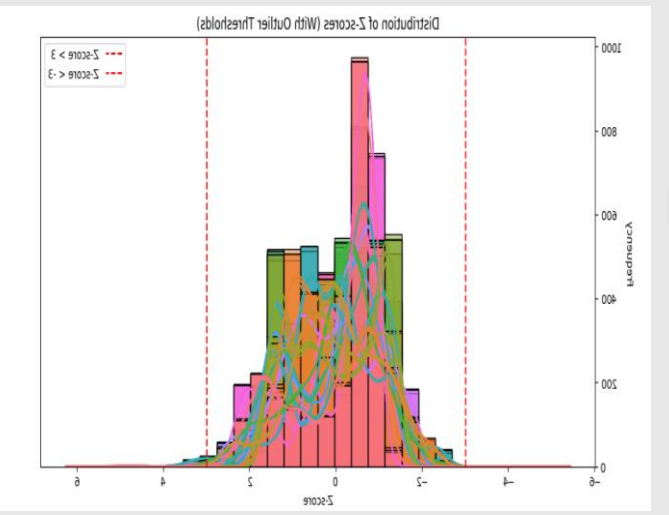
Machine Learning-Driven Forecasting of Multi-Currency Forex Rates: A Comparative Analysis of Regression, Tree-Based Models, and Neural Networks with Portfolio Optimization.

Author –Ahaan Tagare
Email ID - ataga001@gold.ac.uk
Department of Computing, Goldsmiths, University of London

Workflow:

- Data collection (Yahoo Finance API)
- Preprocessing (outlier removal, interpolation)
- Feature engineering (RSI, momentum signals, volatility)
- Model development (Ridge, RF, XG Boost, NN/LSTM)
- Portfolio optimization (equal-weighted predictions)
- Comparison with other Work
- Conclusion



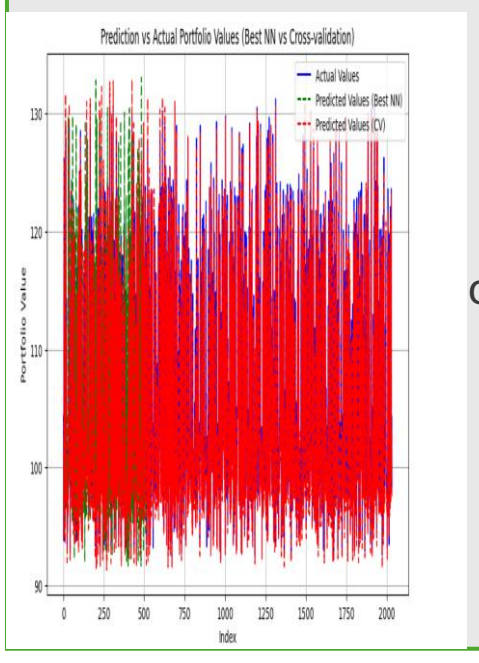


Forex Data Preprocessing and Analysis Techniques

- Linear interpolation fills missing forex data.
- Z-score removes outliers for reliable analysis.
- RSI enhances prediction of price trends.
- Preprocesses data for robust model training.
- Visualizations provide insights for better decisions.

Advanced Forex Prediction Regression Techniques

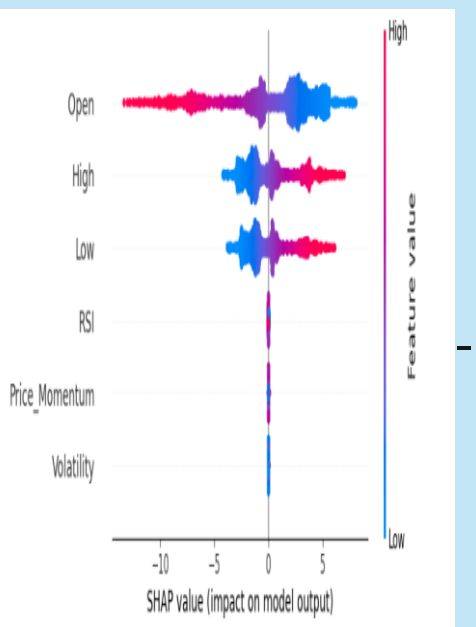
-Baseline model sets performance benchmark.
-Ridge Regression prevents overfitting effectively.
- Random Forest reduces prediction errors.



- XG Boost enhances accuracy with boosting.

- Neural Networks that capture complex patterns.

- LSTM models handle sequential forex data.

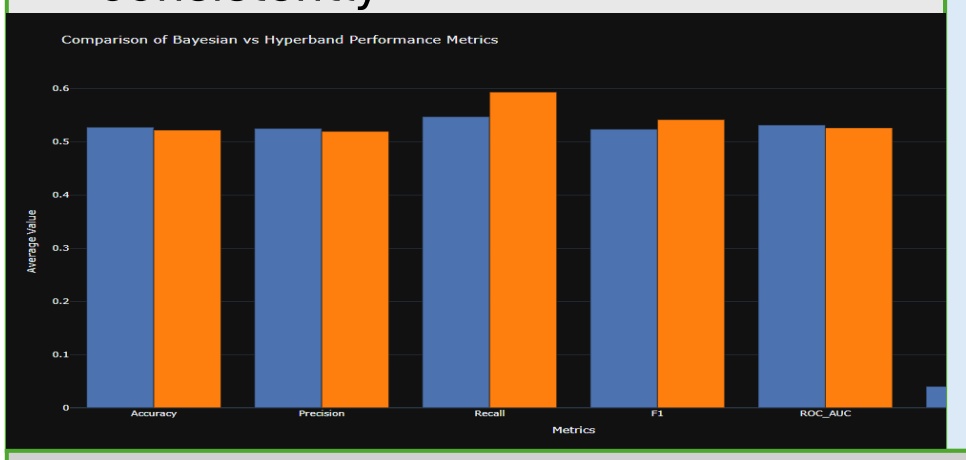


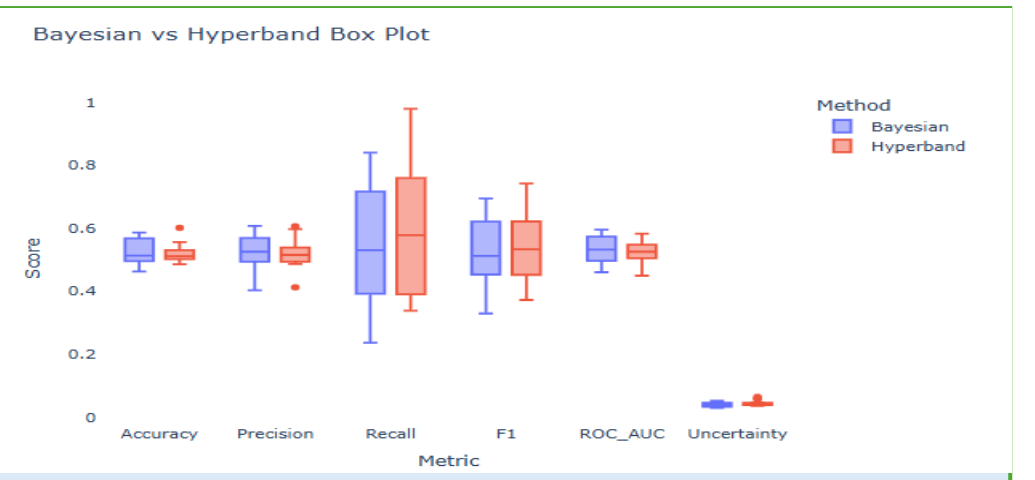
Performance Metrics

- Ridge: MSE 0.000077, R^2 0.999992
- Random Forest: MSE 0.043305, R^2 0.999801
- Neural Network: MSE 3.419810, R^2 0.961488
- XG Boost: MSE 5.458150, R^2 0.962730
- LSTM: MSE 89289.860000, R^2 -0.010933
- Cross-validation: MSE 2.065630, R^2 0.978199

Forex Price Movement Prediction (Classification)

- Bayesian model balances accuracy, uncertainty well
- Hyperband model slightly outperforms Bayesian
- Baseline model shows random guessing performance
- RSI, momentum, volatility shape model predictions
- Hyperband excels in balanced metric scores
- Both models surpass baseline consistently





Results Explanation

Bayesian: 0.527 accuracy, 0.547 recall

Hyperband: 0.522 accuracy, 0.593 recall

Baseline: 0.498 accuracy, 0.517 F1-score

Hyperband excels: 0.743 F1-score USDTRY

Bayesian strong: 0.8408 recall USDJPY

Retrained Hyperband: 75.98% validation accuracy

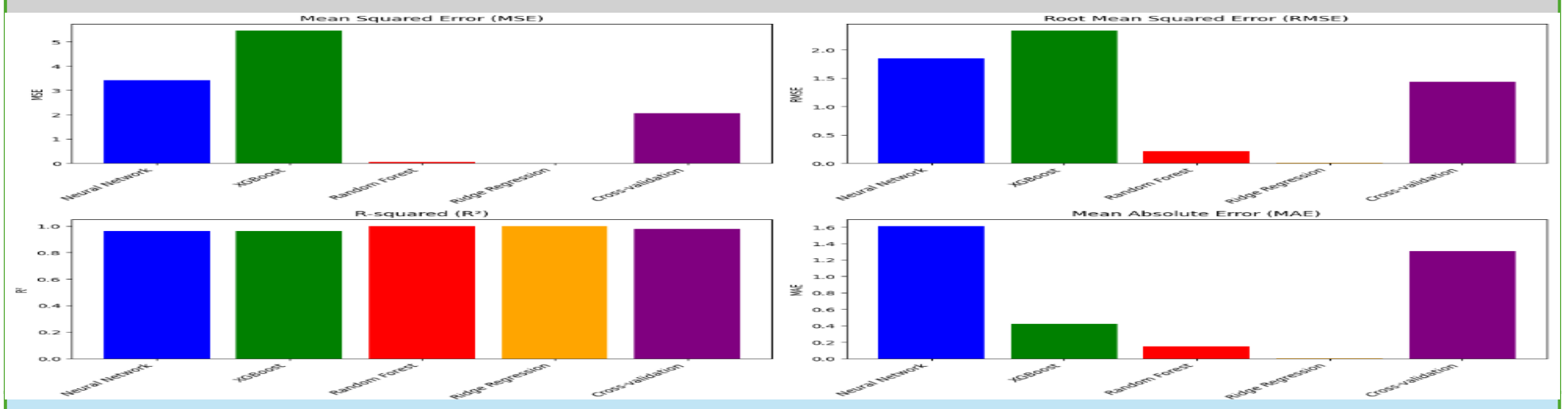
Conclusion

Key Achievement: Successfully developed machine learning models (Ridge Regression, Random Forest, XG Boost, Neural Networks) to forecast multi-currency forex rates.

Best Performing Model: Ridge Regression demonstrated near-perfect R^2 scores, highlighting its effectiveness for regression tasks.

Portfolio Optimization: Combined predictions using equal-weight allocation improved overall forecast stability and reduced variance.

Uncertainty Quantification: Bayesian Neural Networks provided probabilistic predictions, offering valuable insights into model confidence.



Comparison to other work and future recommendations

- Other papers mostly focus on broader, long term or macro economic FX, forecasting out work focuses on portfolio level predictions, for short term analysis
- Future work could macroeconomic indicators (e.g., interest rates, GDP)