**PREDICTING MARKET VOLATILITY AND DIRECTION USING ASPECT-BASED SENTIMENT ANALYSIS AND RECURRENT NEURAL NETWORKS.**

NAME: SAMUEL YEBOAH MARFO

SUPERVISOR: DR TEMITAYO

**ABSTRACT**

This project aims to develop a predictive model that utilizes Aspect-Based Sentiment Analysis (ABSA) combined with Recurrent Neural Networks (RNNs) to forecast market volatility and direction immediately following the release of economic indicators. The model will leverage sentiment analysis from various aspects of economic news to understand how different components influence market movements and apply this understanding to predict market behavior in real-time.

**INTRODUCTION**

Financial markets are exceedingly responsive to macroeconomic indicators, with entities such as the Non-Farm Payroll (NFP), Consumer Price Index (CPI), and Federal Open Market Committee (FOMC) announcements often causing significant fluctuations. Traditional analysis methods have struggled to predict these reactions with high accuracy due to the complex and multi-faceted nature of market responses. The advent of sophisticated Natural Language Processing (NLP) techniques, particularly ABSA, provides new avenues for extracting nuanced sentiment data from economic reports. When combined with the sequential data processing capabilities of RNNs, this sentiment data can be effectively used to predict market volatility and trends.

**LITERATURE REVIEW**

A review of recent literature highlights several approaches to sentiment analysis in financial contexts, with a focus on aspect-based and multimodal sentiment analysis. Key studies include:

1. FinXABSA: Explainable Finance through Aspect-Based Sentiment Analysis (2022) - This paper emphasizes the role of ABSA in enhancing explainability in financial analysis, presenting a model that correlates sentiment with stock prices using advanced statistical techniques.
2. Financial Sentiment Analysis on News and Reports Using Large Language Models (2023) - This study explores the effectiveness of large language models like FinBERT in sentiment analysis, showcasing improved sentiment detection capabilities in financial news.
3. Enhancing Aspect-Based Financial Sentiment Analysis (2021) - Focuses on identifying complex emotions in financial texts using ABSA, improving understanding of nuanced market sentiment.
4. Financial Aspect-Based Sentiment Analysis using Deep Learning (2020) - Discusses the use of deep learning techniques, especially in the context of the FiQA challenge, providing insights into the practical applications of ABSA in financial sentiment analysis.
5. Fine-Grained, Aspect-Based Sentiment Analysis on Economic and Financial Data (2021) - Proposes a methodology to perform fine-grained ABSA on financial texts, assigning real-valued polarity scores to specific topics within documents, which enhances the granularity of sentiment analysis.

These studies, while insightful, generally do not integrate sentiment analysis directly with predictive modeling of market reactions using RNNs, particularly in the context of predicting immediate market volatility and direction. This project aims to fill that gap.

**METHODOLOGY**

**Data Collection**

* **Sources**: Economic data will be sourced from financial platforms such as Forex Factory, which publishes detailed economic calendars showing scheduled releases of economic indicators including NFP, CPI, Interest rates, and FOMC outcomes. Link to source: [Economic Calendar (tradingeconomics.com)](https://tradingeconomics.com/calendar)
* **Data Points**: The data to be collected includes actual figures, previous figures, consensus estimates, and forecasts for various economic indicators.
* **Aspect Identification**: Using NLP techniques, each economic indicator will be identified and tagged as a separate aspect within the economic reports (e.g., inflation rates, unemployment figures, consumer sentiment indices).

**Data Preprocessing**

* **Sentiment Extraction**: Apply Aspect-Based Sentiment Analysis (ABSA) to extract sentiment scores associated with each aspect. This involves developing or adapting sentiment analysis models that can interpret the significance of deviations between actual data and forecasts, and categorize them as positive, negative, or neutral sentiments.
* **Feature Engineering**: Create a feature set for the predictive model, which includes:
  + **Sentiment Scores**: Quantitative scores reflecting the sentiment of each aspect.
  + **Historical Volatility**: Historical market data prior to the news release to establish a baseline of market behavior.
  + **Time Features**: Time and date of the release to capture temporal effects on market reactions.

**Model Development**

* **Choice of Model**: Given the sequential nature of the data, Recurrent Neural Networks (RNN) including Long Short-Term Memory (LSTM) and Gated Recurrent Units (GRU) will be used. These models are effective in capturing dependencies in time-series data, crucial for predicting outcomes that depend on historical context.
* **Architecture**:
  + **Input Layer**: Receives the feature set including sentiment scores and historical volatility.
  + **RNN Layers**: One or more LSTM/GRU layers to process the sequential data.
  + **Output Layer**: A dense layer with outputs corresponding to predicted market volatility and direction immediately after the news release.

**Training the Model**

* **Data Splitting**: The dataset will be divided into training, validation, and test sets. The typical split ratio will be 70% training, 15% validation, and 15% test sets to ensure sufficient data is available for learning and evaluation.
* **Training Process**: The model will be trained on the training set with validation performed at the end of each epoch to monitor for overfitting and to tune the hyperparameters effectively.
* **Optimization**: Use adaptive learning rate methods like Adam or RMSprop, which are generally more effective for RNNs due to their ability to handle sparse gradients.

**Model Evaluation**

* **Metrics**:
  + **Regression Metrics**: Since the output involves predicting the magnitude of volatility, metrics like Mean Squared Error (MSE) and Root Mean Squared Error (RMSE) will be used.
  + **Classification Metrics**: For predicting the direction of market movement (up or down), metrics like Accuracy, Precision, Recall, and F1-Score will be utilized.
* **Backtesting**: The model will be tested against historical data where the actual market reactions post-news release are known. This will provide insights into the predictive accuracy of the model under real-world conditions.

**EXPECTED OUTCOMES**

* **Model Performance**: The model is expected to accurately forecast the direction and magnitude of market volatility following economic news releases, leveraging both sentiment analysis and historical data.
* **Insights into Market Behavior**: The project aims to reveal deeper insights into how different aspects of economic reports influence market behavior, potentially guiding traders and financial analysts in their decision-making processes.
* **Contribution to Financial Forecasting**: By integrating advanced NLP techniques with RNNs, the project seeks to advance the field of financial forecasting, providing a more nuanced tool for understanding and predicting market dynamics.

**REFERENCES**

K. Ong, W. van der Heever, R. Satapathy, E. Cambria and G. Mengaldo, "FinXABSA: Explainable Finance through Aspect-Based Sentiment Analysis," 2023 IEEE International Conference on Data Mining Workshops (ICDMW). Available at: <https://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=10411537&isnumber=10411508>

Bekaert, G. and Hoerova, M. (2014) ‘The VIX, the variance premium and stock market volatility’, *Journal of Econometrics*, 183(2), pp. 181–192. Available at: <https://doi.org/10.1016/j.jeconom.2014.05.008>.

Boyd, J.H., Hu, J. and Jagannathan, R. (2005) ‘The Stock Market’s Reaction to Unemployment News: Why Bad News Is Usually Good for Stocks’, *The Journal of Finance*, 60(2), pp. 649–672. Available at: <https://doi.org/10.1111/j.1540-6261.2005.00742.x>.

Luan, Y. and Lin, S. (2019) ‘Research on Text Classification Based on CNN and LSTM’, in *2019 IEEE International Conference on Artificial Intelligence and Computer Applications (ICAICA)*. IEEE, pp. 352–355. Available at: <https://doi.org/10.1109/ICAICA.2019.8873454>.

Yang, S., Yu, X. and Zhou, Y. (2020) ‘LSTM and GRU Neural Network Performance Comparison Study: Taking Yelp Review Dataset as an Example’, in *2020 International Workshop on Electronic Communication and Artificial Intelligence (IWECAI)*. IEEE, pp. 98–101. Available at: <https://doi.org/10.1109/IWECAI50956.2020.00027>.

Yu, Y. *et al.* (2019) ‘A Review of Recurrent Neural Networks: LSTM Cells and Network Architectures’, *Neural Computation*, 31(7), pp. 1235–1270. Available at: <https://doi.org/10.1162/neco_a_01199>.

Fazlija, B. and Harder, P., 2022. Using Financial News Sentiment for Stock Price Direction Prediction. Mathematics, 10(13), p.2156. Available at: <https://doi.org/10.3390/math10132156>.

Weiss, G., Goldberg, Y. and Yahav, E., 2018. On the practical computational power of finite precision RNNs for language recognition. arXiv preprint arXiv:1805.04908. Available at: <https://arxiv.org/abs/1805.04908>.