Stock and Currency Price Prediction Using ML Methods

07.19.2023

Mohammad Javad Maheronnaghsh
Mohammad Mahdi Gheidi
Sharif University of Technology

Overview

This proposal outlines a research study aimed at predicting stock and currency prices using state-of-the-art AI and ML techniques, including **Transformers**, **Ichimoku Cloud**, **LSTM**, **Simple RNN**, **NHits**, and **NBeats**. The primary objective of this research is to explore the efficacy of these models in forecasting financial markets' price movements and comparing their performance in terms of accuracy, efficiency, and robustness. The proposed research aims to contribute to the existing body of knowledge in the field of financial forecasting and provide valuable insights for investors, financial analysts, and researchers. This article/project will detail the methodology, data collection, model implementation, evaluation metrics, and potential applications of the research findings.

The financial markets are characterized by complex and dynamic price movements, making accurate predictions challenging but crucial for informed decision-making. Recent advances in AI and ML, especially Transformers, LSTM, Simple RNN, NHits, and NBeats, have shown promising results in various time-series forecasting tasks. This research proposal seeks to investigate and compare the performance of these techniques in predicting stock and currency prices.

Goals (Objectives)

1. To explore the **applicability** of Transformers, LSTM, Simple RNN, NHits, and NBeats for stock and currency price prediction.
2. To assess the predictive **accuracy** of each model under different market conditions and time horizons.
3. To identify the **strengths and weaknesses** of each model and their suitability for real-world financial forecasting applications.

Specifications (Methodology)

Data Collection

Historical stock and currency price data will be collected from reliable financial sources, such as financial market databases, APIs, or financial institutions.

Data Preprocessing

The collected data will undergo preprocessing to handle missing values, outliers, and other data anomalies. Standardization or normalization may be applied to ensure consistent scaling across different features.

Data Partitioning

Divide the historical data into training, validation, and test sets to assess the models' performance accurately. The training set will be used for model parameter learning, the validation set for hyperparameter tuning, and the test set for unbiased evaluation.

Model Definition (Model Architecture)

Here we are going to define 6 models:

a. **Transformers**: Implement the Transformer architecture using libraries such as TensorFlow or PyTorch, with attention mechanisms and positional encoding for sequence modeling.

b. **Ichimoku Cloud**: Implementing the Ichimoku Cloud in Python involves calculating the various components of the cloud, such as the Tenkan-sen, Kijun-sen, Senkou Span A, Senkou Span B, and Chikou Span, based on historical price data.

c. **LSTM**: Set up the Long Short-Term Memory (LSTM) model using the chosen deep learning framework, configuring the number of LSTM layers, hidden units, and dropout rates as necessary.

d. **Simple RNN**: Implement the Simple RNN model, specifying the number of recurrent units and other hyperparameters.

e. **NHits**: Implement the NHits ensemble forecasting model, incorporating various time-series forecasting methods to generate combined predictions.

f. **NBeats**: Implement the NBeats deep learning architecture, designing the necessary components for effective time-series forecasting.

Model Training, Validation and Evaluation

a. **Training**: Train each model using the training data and appropriate optimization algorithms (e.g., Adam, RMSprop) with suitable learning rates.

b. **Validation**: Utilize the validation set to tune hyperparameters, such as learning rates, batch sizes, and the number of epochs, to optimize model performance.

c. **Evaluation Metrics**: Assess the models' predictive accuracy using common evaluation metrics such as Mean Squared Error (MSE), Root Mean Squared Error (RMSE), and Mean Absolute Error (MAE).

Comparative Analysis

a. **Performance Metrics**: Quantitatively compare the models' performance based on the evaluation metrics, identifying the most accurate and reliable model.
b. **Statistical Analysis**: Perform statistical tests (e.g., t-tests or ANOVA) to determine if any observed differences in performance are statistically significant.

Discussion of Results

Present the findings of the comparative analysis, providing insights into the strengths and weaknesses of each model. Discuss the models' abilities to capture complex price patterns, their robustness in different market conditions, and their suitability for stock and currency price prediction.

Potential Limitations

Address any limitations or challenges encountered during the research, such as data quality issues, model interpretability, or computational resources.

Conclusion

Summarize the key findings of the research, emphasizing the model that demonstrated superior performance in stock and currency price prediction. Highlight the implications of the research and suggest avenues for future work, such as exploring hybrid models or incorporating external market indicators to enhance predictions.

Milestones

I. Milestone 1: Data Collection and Preprocessing (Week 1)

- Collect historical stock and currency price data from reliable financial sources.
- Preprocess the data to handle missing values, outliers, and ensure consistent scaling across features.

II. Milestone 2: Data Partitioning and Model Definition (Week 2)

- Divide the preprocessed data into training, validation, and test sets.
- Define the six models to be used: Transformers, Ichimoku Cloud, LSTM, Simple RNN, NHits, and NBeats.

III. Milestone 3: Model Training and Validation (Week 3)

- Train each model using the training data and appropriate optimization algorithms (e.g., Adam, RMSprop) with suitable learning rates.
- Utilize the validation set to fine-tune hyperparameters for optimal model performance.

IV. Milestone 4: Model Evaluation and Comparative Analysis (Week 4)

- Evaluate the predictive accuracy of each model using evaluation metrics such as MSE, RMSE, and MAE.
- Conduct a comprehensive comparative analysis to identify the most accurate and reliable model.

V. Milestone 5: Discussion of Results and Conclusion (Week 5)

- Present the findings of the comparative analysis, highlighting the strengths and weaknesses of each model.
- Discuss the models' suitability for stock and currency price prediction in different market conditions.
- Address potential limitations and challenges encountered during the research.
- Summarize key findings and emphasize the model that demonstrated superior performance.
- Draw conclusions, discuss implications, and suggest future research directions.

Some Other Notes

- The final objective of this project is to publish it in journals, conferences, or - at least - Arxiv.
- A proper title for the article: Leveraging AI and ML Techniques for Stock and Currency Price Prediction: A Comprehensive Comparative Study

- The supervisor of this project is Mr. Abolfazl Younesi and the supervising professor is Professor M. Amin Fazli.
- It has to be further checked whether we add TinyML and also implement these on Raspberry Pi or not.
- In the end, Auto Trader Bot is a possible solution to end-up our project (It depends on the progress of the project).

Thanks for your attention