A Comparative Study of Sentiment Analysis Models in Financial Markets

Pranav Agarwal
Parth Maheshwari

Indian Institute of Technology Guwahati

November 25, 2023

Introduction

- Historical quest for prediction: The study is driven by a deep-rooted human endeavor to forecast future outcomes using data analysis, a practice that dates back through history.
- Complexity in financial forecasting: Focuses on the intricacies of financial time series forecasting, a field marked by substantial complexity and significant implications.
- Sentiment analysis in forecasting: Explores the integration of sentiment analysis, particularly from social media like Twitter, into stock market forecasting models, examining its impact on predictive accuracy.

Aim

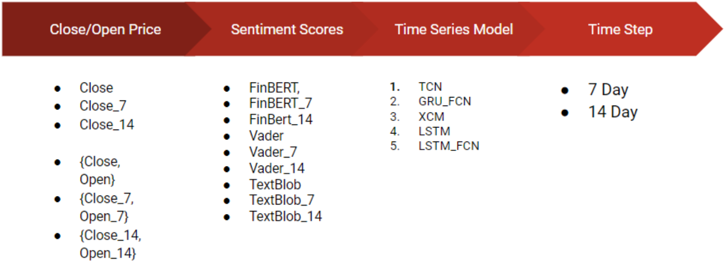
- Comprehensive model evaluation: Undertakes a thorough comparison of various advanced forecasting models in financial time series, emphasizing deep learning and machine learning techniques.
- Role of open stock prices: Investigates how including open stock prices as an input parameter affects model accuracy.
- Sentiment analysis and temporal scales: Assesses the impact of sentiment analysis and studies model performance across different temporal scales to understand forecasting reliability and accuracy.

Why Open Price?

- Reflection of external influences:
  - Open prices capture market reactions to overnight or weekend events, which are not reflected in the previous day's closing price.
  - Essential for understanding market responses to news, global events, and other developments happening when the market is closed.
- Comprehensive market analysis:
  - Including open prices provides a more complete picture, filling the information gap between market closure and the next opening.
  - Particularly crucial for analyzing market trends after extended closures like weekends or holidays.

Input Feature Combinations

- Analysis of TSLA stock data: The study meticulously analyzes TSLA stock data, forming the basis of the input features for predictive models.
- Creation of diverse feature sets: 54 unique combinations of features were developed, half including only closing prices, and half combining opening and closing prices, to thoroughly investigate the data's predictive capacity.



Algorithms Employed

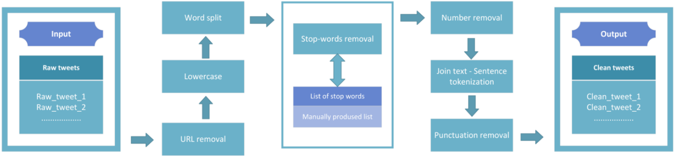
- Selection of state-of-the-art algorithms: The study employs five advanced forecasting algorithms, chosen for their effectiveness in complex data analysis and relevance in financial forecasting.
- Contextual effectiveness: These algorithms were specifically selected for their proven capability in similar financial contexts and their ability to handle intricate data structures.

Time Steps

- Diverse temporal analysis: Experiments were conducted over two distinct time steps to capture both short-term and long-term market trends, reflecting the stock data's predictive patterns.
- Comprehensive model assessment: This approach enabled a thorough evaluation of the models' performance across varying temporal contexts, ensuring a more complete understanding of their predictive capabilities.

Data Preprocessing

- Sourcing and preparing data: Stock-related data for one year was sourced from Kaggle, with a focus on daily stock information essential for analysis.
- Detailed preprocessing steps: Involved removing hyperlinks from tweets, converting text to lowercase, and breaking it into words. Extraneous phrases and numerical strings were removed, and the text was restructured into its original form with tokenization.



Sentiment Analysis

- Crucial role of sentiment analysis: Utilized VADER, FinBERT, and TextBlob for sentiment analysis, applying these tools to TSLA-related Twitter data to derive sentiment scores.
- Depth with rolling averages: Incorporated 7-day and 14-day rolling averages to capture short-term and medium-term sentiment trends, providing a richer analysis of market sentiment.
- Multi-dimensional market view: The integration of these sentiment scores into forecasting models offered a more nuanced perspective, enhancing the accuracy of stock market predictions.

Algorithms (Detail)

- Overview of employed algorithms: Detailed presentation of the algorithms used, including TCN, GRU FCN, XCM, LSTM, and LSTM FCN, each offering unique strengths in time series analysis.
- Utilization of tsAI library: The tsAI library, a robust Python toolkit for deep learning in time series analysis, was instrumental in implementing these algorithms efficiently.

Metrics

- Mean Absolute Error (MAE):
  - Formula: $\text{MAE} = \frac{1}{n}\sum_{i=1}^{n} |y_i - \hat{y}_i|$
  - Represents the average magnitude of the errors in a set of predictions, without considering their direction.
- Root Mean Squared Error (RMSE):
  - Formula: $\text{RMSE} = \sqrt{\frac{1}{n}\sum_{i=1}^{n} (y_i - \hat{y}_i)^2}$
  - Emphasizes larger errors, providing a measure of the magnitude of error in prediction.

Key Findings

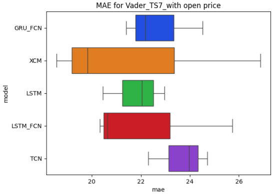
- Timestep impact: Observations show that longer timesteps generally lead to less accurate predictions, as evidenced by higher MAE and RMSE.
- Open price influence: Inclusion of open price data in models affects error metrics variably, indicating its conditional relevance.
- Sentiment analysis tool comparison: Discrepancies across different sentiment analysis tools in terms of error metrics, with Vader showing better overall performance.
- Variability in model performance: Notable fluctuations in maximum and minimum error values across configurations highlight the need for careful model interpretation.

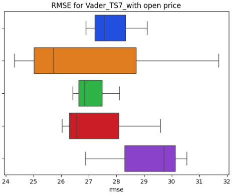
Comparative Analysis of Models

- Comparative study of models: Analyzing the performance of GRU FCN, LSTM FCN, TCN, and XCM, with a focus on MAE and RMSE metrics.
- Insights into model accuracy: TCN shows the most accurate average predictions, while other models exhibit comparable levels of accuracy, as seen in their mean MAE and RMSE.

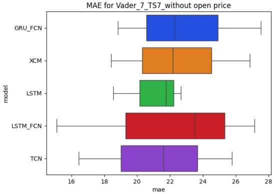
Model Performance Graphs

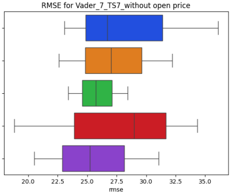
Vader TS7 with Open Price





Vader 7 TS7 without Open Price





Conclusion

- Model performance insights: Highlights the consistency and average performance of various models, with LSTM showing the greatest consistency and TCN the best average performance.
- General observations: Discusses the sensitivity of model accuracy to timestep length, variable impact of open price inclusion, and comparative performance of different models.
- Recommendations for future research: Suggests exploring advanced data preprocessing techniques and investigating high variability in model performance for more targeted and effective prediction models.