UniBe CAS ADS Module 5

Peer Review of Lukazs Macias M3 Project: Predicting TESLA stocks using sentiments analysis

11 December 2024

Marcel Grojean
marcel.grosjean@hepl.ch
Haute Ecole Pédagogique Vaud

I. Introduction

This project focuses on predicting Tesla (TSLA) stock price movements using a combination of sentiment analysis and machine learning techniques. Tesla was selected due to its highly dynamic stock price behavior, characterized by frequent "bull runs" and sharp price drops, and the significant amount of investor sentiment data available through platforms such as Yahoo Finance. These characteristics made Tesla an ideal case study for exploring the interplay between market sentiment and stock price movements. Specifically, the project aimed to utilize investor sentiment expressed in Yahoo Finance comments alongside other financial indicators to forecast stock prices and assess the potential for profitable trading strategies.

To achieve this, the student collected a substantial dataset of 648,000 comments spanning over three years and five months. These comments were analyzed using a hybrid sentiment scoring approach. For 44.4% of the comments, sentiment scores were assigned based on explicit tags provided by users, such as "bullish," "bearish," or "neutral." For the remaining comments, the VADER sentiment analyzer was employed, a lexicon-based tool designed for social media text. This hybrid method ensured comprehensive sentiment coverage while leveraging automation for scalability. Key features, such as five-day running averages of sentiment scores, closing prices, daily volume, and daily volatility, were engineered for use in predictive modeling.

Three machine learning models were implemented: Logistic Regression for binary classification tasks, Random Forest for regression tasks, and Long Short-Term Memory (LSTM) networks for sequence-based predictions. The models were evaluated using metrics such as Mean Squared Error (MSE), Mean Absolute Error (MAE), and R-squared. Finally, the predictions were integrated into a trading simulation to evaluate the practical application of the models, yielding results that slightly outperformed a benchmark market return.

II. Good Solutions and Aspects

We're now going to dive into the good solutions and aspects if this project.

Lukasz undertook extensive data collection by scraping 648,000 comments from Yahoo Finance. This substantial dataset allowed the student to capture diverse investor sentiments across different market conditions. The long duration of the dataset, covering over three years, ensured a comprehensive analysis of sentiment trends over time, reflecting varying market dynamics.

He has an effective sentiment scoring approach. The combination of explicit sentiment tagging and VADER-based scoring enabled the student to analyze both labeled and unlabeled data effectively. This hybrid approach ensured greater coverage and provided valuable insights into market sentiment while leveraging automation for scalability. By incorporating both user-labeled sentiments and automated analysis, the methodology balanced accuracy and efficiency.

Lukasz also make a thorough EDA. The project includes a detailed EDA that highlights relationships between sentiment scores and Tesla's closing prices. For instance, the analysis revealed how positive sentiment correlates with price increases, while negative sentiment often precedes price drops. This foundational step validated the hypothesis that investor sentiment significantly influences stock prices, setting the stage for model development and adding depth to the study.

Diverse ML models were used. The three distinct machine learning approaches demonstrates a solid understanding of model selection and experimentation. Logistic Regression effectively addressed binary classification tasks, Random Forest was utilized for regression tasks, and LSTM models leveraged sequential data for time-series predictions. This comprehensive analysis not only highlighted the strengths and limitations of each approach but also allowed for a meaningful comparison of model performances.

The use of five-day running averages for sentiment scores, closing prices, daily volume, and volatility effectively captured temporal trends in market behavior. These features provided the models with valuable contextual information, enabling them to learn from historical data and predict future trends with greater accuracy. This preprocessing step was critical in aligning the dataset with the requirements of the machine learning models.

By incorporating model predictions into a trading simulation, Lukasz demonstrated the practical utility of the project's findings. The simulation's design, which included a defined investment period and decision-making criteria based on model predictions, bridged the gap between theoretical predictions and real-world applications. The positive returns achieved in the simulation underscored the potential of machine learning in informing investment strategies.

The use of MSE, MAE, and R-squared for regression models and classification accuracy for Logistic Regression ensured a rigorous assessment of model performance. These metrics provided a clear understanding of the models' predictive accuracy, highlighting their strengths and identifying areas for improvement.

The trading simulation's results, which outperformed the market benchmark by $1,243, indicate the potential value of integrating machine learning predictions into investment strategies. This outcome adds credibility to the project and underscores its relevance in financial contexts, making it a strong case for applying machine learning in stock price prediction.

III. Suggestions for Improvements

While the project is well-executed, several areas could benefit from further refinement to enhance its accuracy, robustness, and generalizability.

The VADER sentiment analyzer, while effective for general text, may lack the sophistication required for financial contexts. Implementing advanced natural language processing (NLP) models, such as BERT or FinBERT, could improve sentiment scoring accuracy by capturing domain-specific nuances. Additionally, clustering comments by topics (e.g., market trends, earnings reports) and filtering for relevance could enhance the quality of sentiment analysis, ensuring that only the most impactful comments are considered.

The current feature set, while robust, could be expanded to include macroeconomic indicators (e.g., interest rates, inflation) and external factors, such as Tesla's product launches or Elon Musk's public statements. These variables could provide additional context, enabling the models to account for broader market influences and company-specific events that impact stock prices.

The project assumes a direct relationship between sentiment and stock price movements without accounting for potential delays. Introducing lagged sentiment features or conducting time-lag correlation analysis could better capture the latency between sentiment changes and market responses. This adjustment would refine the predictive model by aligning it more closely with real-world market dynamics.

Systematic hyperparameter tuning, such as grid search or Bayesian optimization, could enhance the performance of the Random Forest and LSTM models. For example, exploring different numbers of decision trees, sequence lengths, or hidden layer configurations could yield improved results. These optimizations would ensure that each model operates at its maximum potential.

The LSTM model was trained for only 10 epochs, which may not fully exploit its ability to learn from sequence data. Increasing the number of epochs, implementing early stopping to prevent overfitting, and incorporating dropout layers to improve generalization could significantly enhance the model's performance. These adjustments would enable the LSTM to capture complex patterns in time-series data more effectively.

Cross-validation techniques could provide a more robust evaluation of model performance. Splitting the data into multiple folds and training/test sets would ensure that predictions generalize well across different subsets of data, thereby improving the reliability of the models.

The trading simulation could be made more realistic by incorporating transaction costs, risk-adjusted metrics (e.g., Sharpe ratio), and alternative strategies, such as diversified portfolios or dynamic holding periods. These additions would better reflect the complexities of real-world trading and provide a more accurate assessment of the profitability of the model's predictions.

Applying the methodology to other stocks or sectors would test its robustness and identify whether the results are specific to Tesla or generalizable to broader markets. This step would also demonstrate the scalability of the approach, potentially uncovering sector-specific patterns that could inform investment strategies.

By addressing these suggestions, the project could achieve greater accuracy, applicability, and impact. These improvements would enhance its value as a comprehensive study in financial machine learning, contributing valuable insights to both academic research and practical applications in stock market analysis.

IV. Conclusion

The project represents a commendable effort to integrate sentiment analysis and machine learning for financial prediction, addressing a complex and dynamic domain with a structured and methodical approach. By leveraging a substantial dataset and exploring diverse machine learning models, the student has effectively demonstrated the interplay between investor sentiment and stock price movements. The thoughtful implementation of feature engineering and evaluation metrics further underscores the project's rigor.

Despite its strengths, there are clear opportunities for improvement, particularly in refining sentiment analysis methods, expanding the feature set, and addressing real-world complexities such as latency and trading costs. By incorporating these enhancements, the project could achieve even greater accuracy and applicability, providing a robust framework for financial forecasting and investment strategy development. The methodology also holds promise for broader application across different stocks and market conditions, showcasing its potential as a versatile tool in the financial domain.