

Comparative Analysis of Sentiment Classification Models for Algorithmic Trading

Perera R.D.M.^{1*}, Ishanka U.A.P.²

¹Department of Software Engineering, Faculty of Computing,
Sabaragamuwa University of Sri Lanka, Sri Lanka

²Department of Data Science, Faculty of Computing,
Sabaragamuwa University of Sri Lanka, Sri Lanka

*dulanya2001@gmail.com

Market sentiment, as driven by news and online discussion sites is increasingly impacting algorithmic trading systems. In this work, we analyze financial news headlines from Yahoo Finance, enhanced with related discussion context from Reddit, to identify market sentiment for Bitcoin. We trained various machine learning and deep learning sentiment classifiers with a dataset of 11,293 Bitcoin-related text samples primarily extracted from Yahoo Finance news headlines and aligned with relevant Reddit discussion content. The dataset had three classes of sentiment which were Positive, Neutral and Negative. The extracted data was fetched from news headlines ranging for five years. Multiple classification models were tested, the classical machine learning models, deep learning architectures and then transformer based language models, which enabled a systematic comparison across different models. In baseline models, Logistic Regression achieved the best performance (accuracy 71.7%, F1=0.704), and among neural network-based ones, BiLSTM does so with accuracy of 70.5%. Transformer-based models were subsequently fine-tuned and evaluated for comparative performance. The model FinBERT achieved improved classification performance (accuracy of 83.7%, F1=0.835), which was further confirmed statically by McNemar's statistical test as compared with the other models. FinBERT showed the effectiveness of domain specific contextual pretraining for the financial sentiment because it was able to outperform all the other models by a large margin. This study highlights the importance of incorporating sentiment aware modeling into the financial prediction workflows and also provide insights that would be needed for future development of reliable trading strategies.

Keywords: *Algorithmic Trading, Bitcoin, Cryptocurrency, Machine Learning, Market Sentiment*