2021 IEEE International Conference on Systems, Man, and Cybernetics (SMC)
17-20 October, 2021. Melbourne, Australia

Stock Price Prediction Using Sentiment Analysis

Thendo Sidogi
School of Electrical Engineering
University of Johannesburg
Johannesburg, South Africa
thendos@gmail.com

Rendani Mbuvha
School of Statistics and Actuarial Science
University of Witwatersrand
Johannesburg, South Africa
rendani.mbuvha@wits.ac.za

Tshilidzi Marwala
School of Electrical Engineering
University of Johannesburg
Johannesburg, South Africa
tmarwala@uj.ac.za

2021 IEEE International Conference on Systems, Man, and Cybernetics (SMC) | 978-1-6654-4207-7/21/$31.00 ©2021 IEEE | DOI: 10.1109/SMC52423.2021.9659283

***Abstract*—We investigate the influence of financial news headline sentiment on the predictability of stock prices using Long Term Short Term Memory (LSTM) networks. The investigation is performed on intraday data with specific lag-times between published article headlines and realised stock prices. FinBERT, a natural language processing model which is fine-tuned specifically for financial news is used to perform sentiment analysis on the company related news headlines. Two base models, one with only historical stock price data as inputs and the other with both historical stock price data and sentiment data from the original BERT model is tested. An alternative model with have both historical stock price data and sentiment data from the fine tuned FinBERT model as additional features. A comparison is performed on both the base and alternative models using Root Mean Square Error (RMSE) and mean absolute error (MAE) as performance metrics. The results suggest that the use of news headline sentiment features from FinBERT significantly improve the predictive performance of LSTM networks in intraday stock price prediction. FinBERT features are also found to outperform features based BERT model trained on a general corpus, illustrating the positive effect of domain specific fine tuning for Large Language models.**

***Index Terms*—FinBERT, language model, sentiment analysis, prediction, LSTM**

I. Introduction

Stock price prediction is a problem that practitioners and researchers have attempted to understand since the early days of the stock exchange [1]. Stock price prediction is a largely non-trivial task since it contradicts the well established theory of efficient-market hypothesis (EMH) [2], [3]. It is largely accepted that the form of EMH exhibited within financial markets has been identified as semi-strong EMH. This form of EMH states that prices of stocks or listed financial instruments reflect all past and current available public information [4]. In other words, it is impossible to make correct future predictions consistently on a risk-adjusted basis since prices should react to new information in near real-time.

It is generally assumed that markets take some time to discount market information [5]. Information published in news articles are a relatively reliable source of insights and can assist in stock price prediction [6]. News headlines provide a new dimension for the stock price prediction problem. Sentiment from the news articles becomes a key factor in predicting the rise and fall of stock prices [7]. It is important to analyse the information as soon as possible so it can help predict stock prices [5]. The task of analysing the vast number of articles manually and predicting the stock prices can be very tedious [8]. This means that an automated system can potentially assist to extract, process, structure and predict. It is therefore ideal for news articles to be extracted from websites, processed into a structured format (sentiment score) and stored for prediction.

This work aims to investigate the use of financial language models to performs sentiment analysis on financial news headlines and to further investigate whether this analysis can assist in our stock price prediction problem using an LSTM network. Depending on the quality of our news source data, this can be seen as an aid to fundamental analysis of the market [9]. Intraday financial news headlines sentiment scores and the respective historical intraday price data are used as inputs to the problem. These input data is then fed into an LSTM network to predict the stock price post publishing of the respective financial news headlines. The paper starts by looking at different methodologies used in stock price prediction in Section II. Then describing the methodology undertaken for the approach in Section IV and the discussion of the corresponding results in SectionV. The paper concludes by drawing associations between the results of this research and prior work in Section VI.

II. Background to Stock Price Prediction

In the literature we find popular traditional methods to stock price prediction that fall in three broad categories fundamental analysis, technical analysis and computational methods.

A. Fundamental Analysis

Fundamental analysis is a technique used to assess or predict the price of a stock using publicly available information such as financial statements or economic indicators [10]. It is a method used to calculate the intrinsic value of an a financial instrument by forecasting future cash flows and discounting them by the expected discount rate. Other company information is taken into consideration like management views on the operational performance and the competitiveness of the operations [9]. This can come with a lot of subjectivity and may not be scalable to fundamentally analysing thousands of companies listed on different exchanges around the world. These methodologies are popular in small stock exchanges where estimating earnings per share using company fundamentals is important [11].

B. Technical Analysis

Technical analysis involves using charts, trading data such as price and volume as input features with very little subjectivity on the analysis to evaluate investment opportunities [12]. This is done by analysing the statistical trends of the collected data. In contrast to fundamental analysis, technical analysis does not use any company fundamentals like financial statements data. The goal is to determine future price movement of a stock based solely on statistical trends of the past price or volume data with the belief that historical prices tend to repeat themselves [13].

C. Computational methods

Computational methods such as using artificial neural networks to predict stock prices were used in [14]. The advent of digital computers and recent notable development of computational intelligence theory and big data has seen quantitative investment strategies adopting more of these new technologies. Specifically, practitioners have looked at other data sources to assist in predicting stock price movements [15]. These include satellite images, financial news articles, twitter or social media data, google trends [5], [7] etc. Researchers have proposed a number of methods aimed at using language features to predict market variables. In [16], the authors investigated the importance of text analysis for stock price prediction in response to financial events in 8-K fillings and reported a prediction accuracy in excess of 10% over a strong baseline models that are using fundamental features.

With the rise in the number of news articles being published, the use of text mining together with machine learning algorithms has received more attention in recent years [17]. Through analysing these text data from online news websites and other sources, the authors in [18] also investigate the effects of sentiment on stock price returns.

There is increasing evidence that computational methods for stock price prediction can be a combination of both technical and fundamental analysis [19]. In this work, we explore this idea through testing whether sentiment from financial news headlines can be used as additional features with historical price movements to improve in stock price prediction.

III. BERT and FinBERT: Sentiment Analysis

Sentiment analysis is a branch of Natural Language Processing (NLP) that analyses pieces of text to infer an authors attitude as negative, positive or neutral about a subject [20]. State-of-the-art techniques include using language models employing machine learning techniques and model architecture such as transformers [21], [22], to get statistical and probabilistic properties of words in a sentence as represented by trained neural networks.

Applications of sentiment analysis include product recommendations, healthcare, politics and online surveillance [23]. We aim to apply sentiment analysis to financial news headlines. Financial news data is often consumed by traders who participate in the stock market. The effects of financial news as a source of high frequency data is very important in gauging the aggregated opinions of investors or traders [5]. The task of analysing the vast number of articles manually by traders can prove to be manually tedious using human capability. We propose to use sentiment scores of financial news headlines to predict stock prices. Financial news article headlines will be analysed to generate a sentiment score confirming whether the financial news sentiment has a predictive influence on stock price movement.

A. BERT (Bidirectional Encoder Representations from Transformers)

With a shortage of data to train the language model, we make use of pre-trained language models. These models are trained on large amounts of text or corpus in order to understand a specific language through multi-task self-supervised training techniques.

- A pre-trained model can be fine-tuned on NLP tasks including sentiment analysis instead of training the models from scratch. To do sentiment analysis, we make use of the Bidirectional Encoder Representations from Transformers (BERT) model.
- BERT is a deep bidirectional model pre-trained on a plain unlabelled text corpus (in this case, Wikipedia and book corpus) to get a statistical property representation of the English language [22]. Attention mechanisms in [21] make use of an alignment score function to quantify the relevance of each item in a list to another item. This is how context of a word in a sentence is established. Transformers like BERT utilise a scaled dot-product attention mechanism within both masked language model prediction and next sentence prediction.
- Prior to BERT, sequence modelling was used to establish language understanding. This often involved finding relations between words through tokenzation of each word as model inputs to the self-supervised task. BERT removes these constrain by using the complete tokenized sequence of words as model input and hence establish word relation in a non-directional way.

B. FinBERT: BERT with financial language model

We introduce a further pre-trained BERT model on financial domain language called FinBERT by [24]. Similarly, they use a financial corpus from Reuters which consist of 1.8 million news articles that were published by Reuters[1] between 2008 and 2010. The articles are filtered for financially relevant words resulting in 20 million words and 300 thousand sentences. Using the constructor provided in the BERT python library, the data was reprocessed so that it matches the BERT tokenization format and be able to be trained on new financial corpus.

[1]The corpus can be obtained for research purposes here: https://trec.nist.gov/data/reuters/reuters.html

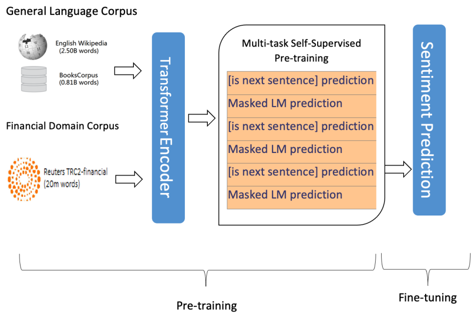

Fig. 1. An illustration of the architecture for FinBERT [24].

Figure 1 shows how the BERT language model was simultaneously trained on a general corpus and financial domain corpus. During fine-tuning to adapt to language understanding tasks, financial sentiment classification training is conducted by adding a dense layer after the last hidden state of the token. The classifier network or the dense layer is trained on the labelled sentiment data set. The labelled sentiment data set used was the FinancialPhrase Bank [25]. Figure 2 is an example of news article headlines and how their sentiment would be classified in the dataset for the supervised learning task[2].

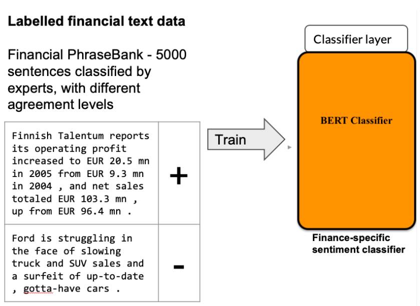

Fig. 2. An illustration of labelling of financial text used to train the classifier layer of FinBERT

We divide the combined data set into a training set and a testing set with an 80:20 split. The model achieves 82% validation accuracy on respective test data this result is similar to that of [24]. Once training and validation is complete on the financial sentiment model, we can then test for sentiment as a signal to our trading or stock price prediction model.

C. FinBERT Sentiment classifier: An example

We now illustrate the FinBERT sentiment classifier using an example on MTN (a telecommunication company listed on the Johannesburg Stock Exchange) and its related news articles to verify if we can use sentiment scoring as signal for systematic trading. The fine-tuned FinBERT model was used to get sentiment score of news article headlines on the day MTN was fined by the Nigerian authorities. On the 4^{th} of September 2018, news broke that Nigeria has handed MTN a \$2 billion tax bill. Testing articles were used to see the performance of the model. The model also generates the magnitude or probabilistic output of the classification as sentiment score between 0 and 1. Figure 3 shows the news article headlines and the sentiment score from the FinBERT model.

[2]https://medium.com/prosus-ai-tech-blog/finbert-financial-sentiment-analysis-with-bert-b277a3607101

IV. Experiment Setup

We make use of a deep-learning neural network architecture in the form of an LSTM network [26] for stock price prediction using historical time series data and financial news headline sentiment. LSTMs are powerful networks which not only have the capability of carrying information from one time step to the other but can also store or establish short-term dependencies and translate them into long-term dependencies. They have internal mechanisms called gates that can regulate the flow of information. This information will then go through gates to inform whether it will be sufficiently important to recall or not.

The ability of LSTM networks to model time series data makes them ideal for the stock price prediction problem. Their short term memory/cache ability helps them learn both the long term and short term dependencies of stock price movement in relation to its historical movement. Therefore, looking at the realised stock price history, LSTMs can predict the next (or next few) points by learning the short and long term patterns.

Figure 4 shows an LSTM network made of a cell state, and different gates. The cell state acts a trend builder or information highway that transfers relative information through the time series sequence. Unlike recurrent neural networks, LSTM have gated cell states which has capacity to carry relevant information throughout the processing of a time series. Using gates, information can be carried to later time steps, and hence introducing long term effects. The gates are neural networks that use a sigmoid function to filter for relevant information to keep or forget.

The LSTM network models were implemented using the Keras deep learning package in python [28]. A single layer of LSTM was used and hyper parameters such as the learning rate, dropout, activation function, optimization, bias, epoch and batch size are chosen to determine the performance of each model. The Adam optimizer [29] is used in our experiments.

A. Data preparation

As shown in figure 5, a web crawler or scrapper is built in python to collect news articles from the web. For this research, 5 months of news articles relating to AMAZON, an American multinational technology company listed on the NASDAQ, was collected between May 2020 and September 2020. The data was a sum total of 2621 AMAZON related news articles headlines to be investigated. These news articles are filtered for only their headlines using the 'BeautifulSoup' python library [30] and parse data from the $FinViz$ news website. For stock

Source	HeadLine	publ_time	stock_price	sent_class	sent_score
MoneyWeb	In Latest class action, Nigeria hands MTN $2bn tax bill	2018/09/04 13:12	84.67	**negative**	0.699
Reuters Africa	In Latest class action, Nigeria hands South Africa's MTN $2bn tax bill	2018/09/04 13:20	85.215	**negative**	0.72
MybroadBand	More huge problems for MTN - Nigeria demands $2 billion in back taxes	2018/09/04 13:50	84.345	**negative**	0.918
IOL	Nigeria slaps South Africa's MTN with $2bn tax nill, shares dip	2018/09/04 13:57	84.415	**negative**	0.876
TimeLive	Nigeria slaps MTN with R30- billion tax bill	2018/09/04 14:01	84.415	**negative**	0.959
BusinessLive	MTN slides on Nigeria's $2bn claim for taxes	2018/09/04 14:24	84.46	**negative**	0.949
Fin24	MTN says Nigeria is seeking to recover $2bn in back taxes	2018/09/04 14:26	84.505	**positive**	0.819
News24	Nigeria hits MTN with new liability, $2bn tax bill	2018/09/04 16:07	81.54	**negative**	0.637
Mail & Guardian	Nigeria hits MTN with new liability, $2bn tax bill	2018/09/04 16:28	79.615	**negative**	0.637
PremiumTimesNigeria	More trouble for MTN as telecom giant faces fresh $2 billion tax	2018/09/04 18:14	79.485	**negative**	0.921
TechCentral	MTN's Nigeria nightmare deepens	2018/09/04 18:35	79.485	**negative**	0.906

Fig. 3. An example of how FinBERT classified MTN news article headlines on 4th September 2018.

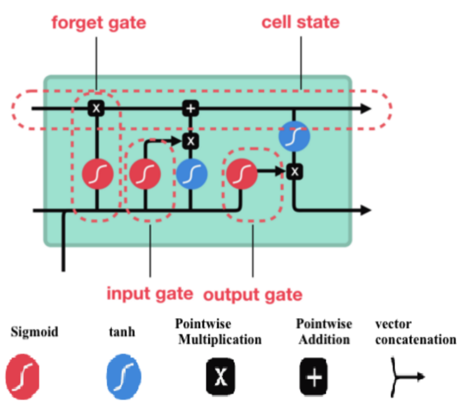

Fig. 4. LSTM network showing the regulation data of using a gating system [27].

data, we used Bloomberg as the source of tick data where price action for every 5 minutes is collected.

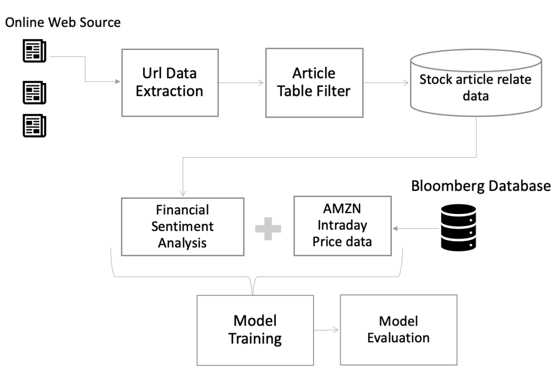

Fig. 5. Data preparation for sentimental analysis and stock price prediction

Sentiment classification on each article headline is either positive, neutral or negative.For model comparison, we generates sentiment scores from both BERT and FinBERT model. The historical intraday data was from 8 May 2020 to 18 August 2020 and 19 August 2020 to 11 September 2020 for training and testing respectively. Each sample of data consist a X of window size and Y of forecast size. As shown in Table II and III sentiment becomes an extra feature to model inputs.

X	Y
Data within input window	Target price data

TABLE I
MODEL INPUT DATA TO LSTM NETWORK FOR MODEL A

X	BERT Sentiment	Y
Data within input window	Headline Sentiment Scores	Target price data

TABLE II
MODEL INPUT DATA TO LSTM NETWORK FOR MODEL B

X	FinBERT Sentiment	Y
Data within input window	Headline Sentiment Scores	Target price data

TABLE III
MODEL INPUT DATA TO LSTM NETWORK FOR MODEL C

B. Model Design

Model A can be described as LSTM model to predict stock price S as S_{t+1} , S_{t+2} , S_{t+3} $\dots S_{t+f}$ using S_t , S_{t-1} , S_{t-2} $\dots S_{t-w}$ where f is the forecast size and w is the window size. A sliding window technique is used. Data normalization is applied to get the stock prices to the same scale. Data is transformed to a range between -1 to 1.

Model B is constructed similarly but as a multivariate model with sentiment scores for every stock price. The sentiment score here is generated from the original BERT model. Model C is similar to model B but with sentiment scores generated from FinBERT model.

The time taken to discount market information varies according to the type of market or the asset under consideration. A look back window size of $w = 7$ is chosen with a 25 minute lag between realised price and published article. The forecast size is $f = 1$. Prior studies on stock market sentiment analysis find that new market information takes around 20 minutes to reflect in the stock prices [31].

V. EVALUATION AND RESULTS

The root mean square error (RMSE) and mean absolute error (MAE) are used to assess the model prediction errors. For the testing period, both figure 10 and 11 shows that model B and model C predictions appears visually to be closer to realised values than with model A in figure 9 while it was the opposite during early periods of model training. This is more evident in figure 11.

A possible reason for the improvement in prediction performance is because of the sentiment signal/trend being

Fig. 6. Model A training and testing loss over each epochs

Fig. 7. Model B training and testing loss over each epochs

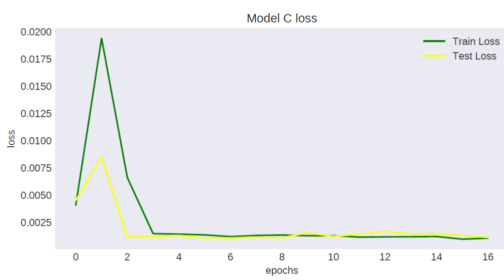

Fig. 8. Model C training and testing loss over each epochs

Fig. 9. Amazon shock price forecast using only historical stock price as input to LSTM network

available to model B and model C versus model A. It can also be seen that model C improves prediction performance over time by establishing long/short term relations between historical price and sentiment. All indicators in Table IV show that the LSTM model with the news sentiment is the better performer. We notice that model B using the

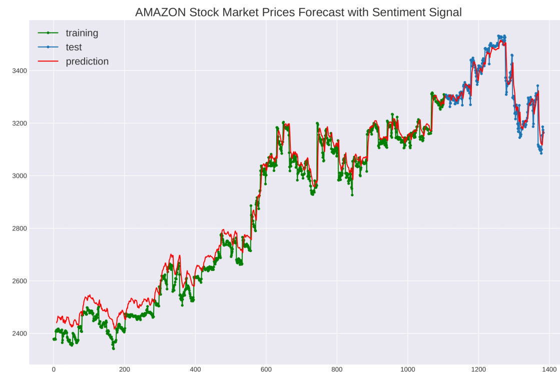

Fig. 10. Amazon shock price forecast using both historical stock price and respective BERT sentiment signal as inputs to LSTM network

Fig. 11. Amazon shock price forecast using both historical stock price and respective FinBERT sentiment signal as inputs to LSTM network

original BERT sentiment scores underperforms the FinBERT model C but outperforms model A without any sentiment features highlighting the importance of pre-training on financial corpus. The LSTM without news sentiment displays inferior performance compared to when incorporating the news sentiment with historical stock price features in the input data. In other words, the impact of the financial news sentiment on stock market prediction is found to be influential.

This strongly suggests that the semi-strong form of efficient market hypothesis does not always hold within intraday stock price actions. The increase in financial news flow and the ability to timeously process these news by market participants maybe a reason for the delayed reaction in price action. A state-of-the-art financial language model (FinBERT) has presented an opportunity to process large amounts of financial news headlines timeously and with this study, has shown the ability to improve on stock price predictability using LSTM networks.

Metric	Model A (Historical Prices Only)	Model B (BERT)	Model C (FinBERT)
MAE	32.23	27.37	22.33
RMSE	41.23	39.82	34.08

TABLE IV
EVALUATION METRICS FOR STOCK PRICE PREDICTION USING LSTM NETWORKS

VI. Conclusion and Future Work

It has become increasingly evident that using historical prices only for stock price prediction may not be sufficient and other alternative data needs to be explored. We have performed a principled comparison on the performance of LSTM models for stock market forecasting under the same conditions but with an objective assessment of the significance of incorporating financial news sentiment as inputs to the model. This work can be extended in a multitude of ways. Drawing from domain knowledge and experience we can test several other theories that can enhance the performance of our model as future work including the following;

- Sentiment Analysis and trading works better with small cap (with low number of floating shares) companies. Through anecdotal evidence, we have seen that small caps react more strongly to news sentiment as compared to mid cap or large caps. Negative news attracts more attention from journalists giving it more media coverage.
- Low volatility, quality stocks have the strongest reaction to positive sentiment, while high volatility stocks have the strongest reaction to negative sentiment due to confirmation biases within investors. Specific sectors are more sensitive to news sentiment than others. Sectors like utilities, pharmaceuticals and banking.
- In order to draw further inferences on the performances of our models, we can perform statistical significance tests on the testing RMSEs of the models. This is done by adopting the hypothesis testing technique used by [32] and [33], in testing for differences between models using the Friedman test and the post-hoc Nemenyi test.
- Bayesian inference methods for the LSTM parameters including automatic relevance determination priors for feature selection, similar to the approach in [34] is also a potential future direction of interest.

References

[1] J. L. Livermore, *How to trade in stocks: The livermore formula for combining time element and price*. Laurus-Lexecon Kft., 1940.

[2] B. G. Malkiel and E. F. Fama, "Efficient capital markets: A review of theory and empirical work," *The journal of Finance*, vol. 25, no. 2, pp. 383–417, 1970.

[3] E. F. Fama, "The behavior of stock-market prices," *The journal of Business*, vol. 38, no. 1, pp. 34–105, 1965.

[4] B. G. Malkiel, "The efficient market hypothesis and its critics," *Journal of economic perspectives*, vol. 17, no. 1, pp. 59–82, 2003.

[5] V. S. Pagolu, K. N. Reddy, G. Panda, and B. Majhi, "Sentiment analysis of twitter data for predicting stock market movements," in *2016 international conference on signal processing, communication, power and embedded system (SCOPES)*. IEEE, 2016, pp. 1345–1350.

[6] J. Li, H. Bu, and J. Wu, "Sentiment-aware stock market prediction: A deep learning method," in *2017 international conference on service systems and service management*. IEEE, 2017, pp. 1–6.

[7] H. Choi and H. Varian, "Predicting the present with google trends," *Economic record*, vol. 88, pp. 2–9, 2012.

[8] A. Yadav and D. K. Vishwakarma, "Sentiment analysis using deep learning architectures: a review," *Artificial Intelligence Review*, vol. 53, no. 6, pp. 4335–4385, 2020.

[9] R. Palat, *Fundamental analysis for investors*. Vision Books, 2016.

[10] D. Dodd and B. Graham, "Security analysis," *New York: Whittlesey house,*, p. 185, 1934.

[11] A. Solomon, F. Memba, and W. Muturi, "Earnings per share and equity share investment in companies listed on nigerian stock exchange," *European Journal of Business, Economics and Accountancy*, vol. 4, no. 1, pp. 68–78, 2016.

[12] C. D. Kirkpatrick II and J. A. Dahlquist, *Technical analysis: the complete resource for financial market technicians*. FT press, 2010.

[13] J. J. Murphy, *Technical analysis of the financial markets: A comprehensive guide to trading methods and applications*. Penguin, 1999.

[14] X. Wang, P. K. H. Phua, and W. Lin, "Stock market prediction using neural networks: Does trading volume help in short-term prediction?" in *Proceedings of the International Joint Conference on Neural Networks, 2003.*, vol. 4. IEEE, 2003, pp. 2438–2442.

[15] A. Denev and S. Amen, *The Book of Alternative Data: A Guide for Investors, Traders and Risk Managers*. John Wiley & Sons, 2020.

[16] H. Lee, M. Surdeanu, B. MacCartney, and D. Jurafsky, "On the importance of text analysis for stock price prediction." in *LREC*, vol. 2014, 2014, pp. 1170–1175.

[17] G. Birz and J. R. Lott Jr, "The effect of macroeconomic news on stock returns: New evidence from newspaper coverage," *Journal of Banking & Finance*, vol. 35, no. 11, pp. 2791–2800, 2011.

[18] S. Y. K. Mo, A. Liu, and S. Y. Yang, "News sentiment to market impact and its feedback effect," *Environment Systems and Decisions*, vol. 36, no. 2, pp. 158–166, 2016.

[19] I. K. Nti, A. F. Adekoya, and B. A. Weyori, "A systematic review of fundamental and technical analysis of stock market predictions," *Artificial Intelligence Review*, pp. 1–51, 2019.

[20] C. Manning and H. Schutze, *Foundations of statistical natural language processing*. MIT press, 1999.

[21] A. Vaswani, N. Shazeer, N. Parmar, J. Uszkoreit, L. Jones, A. N. Gomez, L. Kaiser, and I. Polosukhin, "Attention is all you need," *arXiv preprint arXiv:1706.03762*, 2017.

[22] J. Devlin, M.-W. Chang, K. Lee, and K. Toutanova, "Bert: Pre-training of deep bidirectional transformers for language understanding," *arXiv preprint arXiv:1810.04805*, 2018.

[23] B. Liu and L. Zhang, "A survey of opinion mining and sentiment analysis," in *Mining text data*. Springer, 2012, pp. 415–463.

[24] D. Araci, "Finbert: Financial sentiment analysis with pre-trained language models," *arXiv preprint arXiv:1908.10063*, 2019.

[25] P. Malo, A. Sinha, P. Korhonen, J. Wallenius, and P. Takala, "Good debt or bad debt: Detecting semantic orientations in economic texts," *Journal of the Association for Information Science and Technology*, vol. 65, no. 4, pp. 782–796, 2014.

[26] S. Hochreiter and J. Schmidhuber, "Long short-term memory," *Neural computation*, vol. 9, no. 8, pp. 1735–1780, 1997.

[27] M. Phi, "Illustrated guide to lstm's and gru's: A step by step explanation," Jun 2020. [Online]. Available: https://towardsdatascience.com/illustrated-guide-to-lstms-and-gru-s-a-step-by-step-explanation

[28] F. Chollet *et al.*, "Keras: Deep learning library for theano and tensorflow," *URL: https://keras. io/k*, vol. 7, no. 8, p. T1, 2015.

[29] D. P. Kingma and J. Ba, "Adam: A method for stochastic optimization," *arXiv preprint arXiv:1412.6980*, 2014.

[30] L. Richardson, "Beautiful soup documentation," *Dosegljivo: https://www. crummy. com/software/BeautifulSoup/bs4/doc/.[Dostopano: 7. 7. 2018]*, 2007.

[31] G. Gidofalvi and C. Elkan, "Using news articles to predict stock price movements," *Department of Computer Science and Engineering, University of California, San Diego*, 2001.

[32] R. Mbuvha, M. Jonsson, N. Ehn, and P. Herman, "Bayesian neural networks for one-hour ahead wind power forecasting," in *2017 IEEE 6th International Conference on Renewable Energy Research and Applications (ICRERA)*. IEEE, 2017, pp. 591–596.

[33] R. Mbuvha, I. Boulkaibet, T. Marwala, and F. B. de Lima Neto, "A hybrid ga-pso adaptive neuro-fuzzy inference system for short-term wind power prediction," in *International Conference on Swarm Intelligence*. Springer, 2018, pp. 498–506.

[34] R. Mbuvha, I. Boulkaibet, and T. Marwala, "Bayesian automatic relevance determination for feature selection in credit default modelling," in *Artificial Neural Networks and Machine Learning – ICANN 2019: Workshop and Special Sessions*, I. V. Tetko, V. Kůrková, P. Karpov, and F. Theis, Eds. Cham: Springer International Publishing, 2019, pp. 420–425.