2020 IEEE/WIC/ACM International Joint Conference on Web Intelligence and Intelligent Agent Technology (WI-IAT) | 978-1-6654-1924-6/20/$31.00 ©2020 IEEE | DOI: 10.1109/WIIAT50758.2020.00117

Spot Gold Price Prediction Using Financial News Sentiment Analysis

Zhou Junjie
Department of Journalism
Hong Kong Baptist University
Hong Kong, China
19447515@life.hkbu.edu.hk

Paolo Mengoni, *IEEE Member*
Department of Journalism
Hong Kong Baptist University
Hong Kong, China
pmengoni@hkbu.edu.hk

***Abstract*—Data analytic helps investors to make prediction of the financial market. Financial news or other related information is an invaluable asset for investors' accurate prediction and efficient decision. Most researchers focus on stock price prediction. Commodity market, and especially Spot Gold, has not been investigated deeply in literature. Spot Gold price prediction is proved to be an intractable task due to the complexity of market information and a huge amount of capital pool. This paper using financial sentiment analysis to transform text data into numerical data and using MLP model to predict Spot Gold price. The results show that using news and history price to predict Spot Gold prices is feasible and trustable.**

***Index Terms*—Financial sentiment analysis, Spot Gold, MLP, Machine learning, Financial news**

I. Introduction

Financial market prediction has been always a hot topic for researchers. It's like a Pandora box which attracts researchers to predict the trend and support investment decisions, especially the stock price. Econometrics is an important methodology for researchers to build up their models. There is no doubt that time series prediction is one of the most widely used method. The Internet makes the World tightly linked and information is spreading fast, which provides various sources for economists or investors to do analysis and prediction efficiently.

The prediction of stock prices using financial news is one of the hot topic in recent years. Support Vector Machine (SVM), Natural Language Processing (NLP), Support Vector Regression (SVR) and other textual analysis approaches can transform text data into numerical data. Scientists would like to use these new numerical data to predict the stock price using various different models.

However, most of the researchers focus on the stock market while the commodity market still has much space to be explored, especially the Spot Gold. Gold is considered a safe-haven asset as its price dynamics are mainly regulated by market demand and supply. When stock prices fluctuate, the Gold price relatively remains stable. Spot Gold's price is more stable than most of the stock prices. On the contrary, the Gold's capital pool is much larger than any stocks. Therefore, none of the single trader or investment company can control the Gold market.

However, stable does not mean its price will keep in a certain interval. The basic function of Gold is preservation. It can offset inflation and avoid the social risk which some unpredictable news or events which will hurt the financial market.

In this work we investigated how the Spot Gold price is influenced by financial news sentiment. Moreover, we explored the possibility to predict accurately the Spot Gold price using machine learning. Furthermore, we demonstrated how to apply intelligent computational methods, basing on news and historical Spot Gold price, to predict the trend of Gold future's price.

The remainder of this paper is organized as follows: literature review about financial market prediction is conducted in Section 2. In Section 3 we describe the methodology of our approach to predict the Spot Gold price using financial news sentiment. In Section 4 we analyze the model performance and the correlation between sentiment and Spot Gold price. The conclusions and future works are discussed in Section 5.

II. Literature Review

A. Approach of analysis

Financial forecasting has been investigated for a long period in the research area. The fundamental analysis started in 1928 with an important stock market investor, Benjamin Graham, which stated that investors need to study some fundamental attributes of a company before investing, such as the size of the firm, capitalization, and price-earnings ratio [33]. Altman and Beaver [1] were the earliest scholars that use daily information to do financial prediction. Before the year 2000, most of researchers used some fundamental analysis or technical approaches to do a long-term financial prediction. [26].

In that period, many useful and creative models were introduced by many scholars. These models include time series regression, exponential smoothing, ARIMA [39], and its variations, GARCH [17], among others (Lin et al. [22]; Wang [34]). All these models dominated the financial prediction area before year 2000 and even nowadays still many scholars combine different models to achieve higher accuracy in prediction. Therefore, Kroner, Kneafsey, and Claessens [21] approached the problem by studying the volatility of the daily price of different kinds of products (e.g. oil, Spot Gold, and

978-1-6654-1924-6/20/$31.00 ©2020 IEEE
DOI 10.1109/WIIAT50758.2020.00117

other commodities) in the long term by using a new model which combines GARCH model and ISD (Implied Standard Deviation) models. Besides, Tully and Lucy [32] applied an AP–GARCH model to analyze the effect of macroeconomic on Gold using Spot price, the future price of Gold, and macroeconomic variables spanning a time frame of 20 years from 1983 to 2003. Making use of new financial knowledge, such as the Hurst exponent, Qian el al. [29] found that it could be used to elicit the correlation between Dow Jones daily returns and its historical data. More complex models are on demand to fit the unpredictable financial data to achieve more accurate predictions. Moreover, information spread so quickly so that the financial time series model is influenced by many economic factors such as investor's psychology, and expectations, movement of other stock markets, and political events [19].

With the computational power grows after year 2000, more techniques have been used in financial market analytics. Schumaker [30] applied SVM to analyze the breaking news. They built a model to predict the stock price and the S&P 500. This research could predict the stock price after the news released for 20 minutes and predict shorten trends. It filled the gap that the time series prediction method could not predict the shorter trends. Combining news and technical indicators also perform well to predict the daily stock price [38]. They also applied SVM to classify the history price and news. Their model achieved a high accuracy of stock price prediction. Summarization Model is a new model that uses financial news to analyze the impact of the stock price [22]. After analysing and summarizing the news, it selects the 30 most important sentences into SVM to classify the trend.

The introduction of new technologies like text mining, NLP, Sentiment analysis, and other useful techniques were of great help for prediction. An extensive survey can be found in Cavalcante et al. [6]. Financial Sentiment Analysis (FSA) survey showed that most studies applied their FSA result to stock or Forex market and a few use them on commodities markets such as Spot Gold market and Oil market. Besides, they found there was relationship between industry news and sentiment impact [24]. However, it still has space to improve by applying novel techniques to those financial prediction models.

B. Algorithms

In the last decade, many algorithms have been used to analyze the financial market. Nikfarjam [27] applied text mining to extract qualitative information of companies and use this information to predict the future behavior of stock prices based on positive or negative company news. Schumaker [30] improved the performance of time series regression using SVM and financial news articles to predict the stock price.

Approaches based on a single algorithm presented some drawbacks, such as local optima, overfitting, the difficulty in selecting between a large number of parameters, which directly affect forecasting accuracy [15]. Liang [23] proposed an hybrid methodology, composed by parametric and non-parametric approaches, to effectively forecast option prices. The nonparametric methods investigated include linear neural network, MLP, and SVR while the parametric methods include the Binomial tree, finite difference, and Monte Carlo methods. This hybrid approach, originally tested on stock price prediction, has been extended to predict commodities prices.

Yazdani and Chamzini [37] introduced a new model based on the combination of ANFIS (Adaptive Network Fuzzy Inference System) used to capture the change of Gold price and Artificial Neural Network (ANN) to test the model results. Their results show that the combination of ANFIS and ANN applied to Gold price prediction can achieve better results than the ARIMA model. Moreover, the combination ANN and GARCH improves the performance of a single GARCH model. Results show that the ANN–GARCH model increased the prediction accuracy of Gold Spot price volatility by 25% and Gold futures price volatility by 38% [20].

Various algorithms such as MLP, ANN, SVM, ANFIS, Text mining, auto-regression and etc. have been applied by different scholars for stock and commodity prices prediction [6]. Furthermore, financial news and social media information have been investigated and found to play an important role. Emotion and sentiment information extracted from these data have been used to improve the prediction of stock and commodities prices [24].

C. Data Source

The data sources used in the various approaches in literature may include textual data and stock/commodity price historical information. The textual data can be provided in the form of required shareholder reports, government-mandated forms, or news reports about the company's prospects. An unexpected report can cause significant changes in security prices [30]. Social media data is gaining increasing attention. The quick development of social media introduced a new channel for people to receive the latest information [35]. However, gathering data from social media, such as Twitter, require particular attention to data quality assessment as poor data quality will lead to a bad performing model [4], [9], [28].

Wall Street Journal, Financial Times, Dow Jones News Services, Thomson Reuters, Bloomberg, and Forbes have special authority in finance. However, access to these data collection is difficult for commercial reasons [36]. Chen [8] used their proposed model to verify using experimental datasets from ChinaTimes.com.cn, YES.com, Yahoo! stock market news, and Google stock market news over 18 months. The result provided huge support to their proposed model.

High-frequency financial price data can help to produce better model result. However, most of the datasets can only provide daily data, where hourly data can provide more detailed information, as the market reaction can attenuate very fast after rounds of adaptation [36]. Moreover, Gidofalvi and Elkan [16] found that the fluctuations of stock prices between 20 minutes before the start of news coverage and 20 minutes after the end of news coverage have high predictive power.

D. News influence

Baur and Lucey [2] and Baur and McDermott [3] claim that Gold is a safe haven in times of economic turmoil. In 2020, the COVID-19 outbreak is detrimental to the whole world economy. Negative financial news are became dominant in the early stage of the pandemic. However, some researchers believe that investors may overreact to unexpected and dramatic news or events which can lead to the potential for profitable trading opportunities [5], [10], [12].

III. Methodology

A. Data Source

The data source for the Spot Gold price prediction will contain first-hand data. In this paper we will analyse news information and Spot Gold price quotations.

The professional financial news used in our work has been gathered from sources that have special authority in financial areas. However, some of the financial websites like Financial Times and the New York Times don't provide full-text access to the information as the majority of their content is behind a paywall. Thus, our selection of free access authoritative platforms includes full-text news from Yahoo! Finance, South China Morning Post (SCMP), and Etnet. All of these websites focus or have specific sections for financial news and post information which may influence the economics and the financial markets.

After selecting the data sources, we need to select the topic to focus on. This will help to reduce the data noise. For SCMP, the "Global economy" and "Banking-Finance" channels have been selected. For Yahoo! Finance, we focus on "Global News" and "stock market" news. The "latest news" channel has been selected for Etnet as it covers the whole day's financial news messages. Although the information is coming from different websites and topics/channels, we still can find some overlapping news among the sources. The news data has been gathered from February 17, 2020 to April 25, 2020.

Table I shows the detailed news data from each website. The news format of each Each website is different. Etnet much more focus on some urgent messages while SCMP more focus on some event's influence, hence the number of words in each SCMP news is much more than the other websites we considered. By contrast, the number of news in SCMP is lower as their news reports are longer and more complex. Yahoo! finance present news that are a combination of the previous two. It supports different types of information in the form of articles, short massages, or financial reports. The number of words in each piece may vary between 250 and 4500.

The Spot Gold data (Code: XAU/USD) is provided by Sina, Nasdaq, and Investing. After investigation of the characteristics and frequency of data update, we select Investing real-time Gold data as source for investigation. The dataset includes two months' data from February 17, 2020 to April 25, 2020.

For all of the news and Gold data, their timestamps are based on the Zoom 8 (UTC/GMT+08:00).

TABLE I
DETAIL OF NEWS RECORDS IN THE DATASET

News source	News format	Count	Date period	Words
SCMP	News/Report	232	17/02/2020 - 25/04/2020	1000-6000
Yahoo! Finance	News/Message /Report	952	20/02/2020 - 24/04-2020	250-4500
Etnet	Message	6749	19/02/2020 - 24/04/2020	30-150

B. Sentiment analysis

The textual information needs to be transformed into numerical data for machine learning algorithms. There are many methods of text mining which include Bag of Words, Term Frequency–Inverse Document Frequency (TF-IDF), classification, and other basic NLP techniques. In this work, for a deeper understanding of the news text we apply sentiment analysis to decode the semantic information [7], [25].

Sentiment analysis is a set of NLP techniques that help to extract the polarization of sentiments from text, i.e. it helps to evaluate the feeling that a human reader will get from reading a specific text. In literature, different techniques can be found to do sentiment analysis that include Linguistic Inquiry and Word Count, SentiWordNet, and other machine learning approaches such as Naive Bayes, Maximum Entropy, and Support Vector Machine (SVM) [11], [13], [14], [24]. To minimize the training time and the training resources, Hutto [18] created VADER (Valence Aware Dictionary for sEntiment Reasoning) a simple rule-based model for general sentiment analysis. VADER works well on social media text and can be generalized to fit multiple domains. Its corpora includes four types of information: social media text (Tweets), movie reviews (rotten.tomatoes.com), technical product review (Amazon), and opinion news article (NYT). VADER performed very well in different contexts, especially in the social media and opinion news contexts. VADER outperforms individual human rates in a social media context and can reach nearly 50% accuracy about opinion news articles while other models only can achieve 20% accuracy. The combination of these characteristics renders it very suitable for our text data which contains short messages and opinion news reports.

C. Data preparation and algorithms selection

After using VADER to transform financial news data into numerical data, we create a new dataset that combines with financial sentiment data and Spot Gold data. Gold data have been gathered every minute from Investing.com. However, this information needs to be transformed to the daily average to be combined with financial sentiment data. The daily average price is our model label and the value to be predicted. Besides, the timeliness of news is different in various areas. Smales [31] also use news sentiment to do some regression analysis to test the significant value between news and Gold future return. He

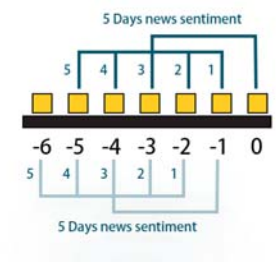

Fig. 1. 5-days news sentiment

found most of the trader will overreact even in some cases and it will be mitigated within a week. To comply these findings we create two different datasets: the first one is using 5-days news sentiment before the price day (Figure 1), while the other is using 1-day news sentiments before the price day (Figure 2). The day at 0 time point is the day we want to predict, -1 means one day before the predicted day. Similarly, -5 means 5 days before the predicted day. The notation concept of Figure 1 also applies to Figure 2.

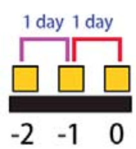

Fig. 2. 1-day news sentiment

After creating the dataset, the appropriate algorithms will be selected for model building to train and predict the Spot Gold price. From previous research, sentiment analysis was not widely used for financial prediction. The most common approaches apply time series models or SVM/SVR text mining approaches [6]. In our work, we apply a more novel approach than SVM. Machine Learning is the main approach we applied. Considering the features of the dataset, Multi-layer Perceptron (MLP) has been selected. After careful parameter selection, our MLP model has been fixed to 3 hidden layers, 6 features input, 256 neurons for each hidden layer and a single neuron output layer (i.e. representing the predicted price of Spot Gold). The loss function selected for evaluation of the training improvement is Least Squares Error or $L2$ loss. The MLP hidden layer weights have been optimized using the AdaGrad Optimizer. The training has been conducted using K-fold cross-validation.

IV. Results

The first part of our analysis is dedicated to figure out whether news sentiment is related to the Spot Gold real price. Table II shows that 5-days news sentiment significant P-value is 0.027 and 1-day news sentiment significant P-value is 0.047. The results show that both 5-days news sentiment and 1-day news sentiment are correlated with the Spot Gold real price. Besides, their Pearson correlation is lower than 0 which means that when the positive news sentiment increases, the gold price decreases. This represents one of the basic attributes of the Spot Gold. It can resist against financial events such as inflation or some unexpected events such as stock market crash, war or other negative events. However, Spot Gold is not a good commodity to invest for its characteristics and its return rate is low when compared with other financial products.

TABLE II
Data source statistical test

Types	Significance (P-value)	Person Correlation
5-days news sentiment(pos)	0.027	-0.306
1-day news sentiment(pos)	0.047	-0.271

In Figure 3 and Figure 4 are represented the positive sentiment and the predicted price trends. Figure 3 is using 1-day news sentiment and Figure 4 is using 5-days news sentiment. 1-day news sentiment dataset is more volatile than 5-days news sentiment dataset. Besides, 1-day news sentiments dataset has more extreme values such as 0.116 and 0.059 while the 5-days dataset only has 0.094 and 0.06.

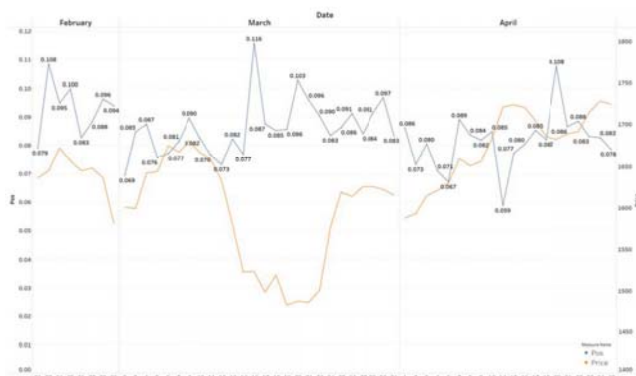

Fig. 3. 1-day news sentiment: Predicted price vs. Positive sentiment

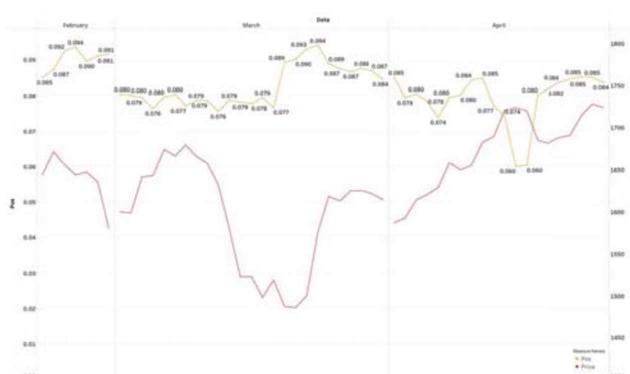

Fig. 4. 5-day news sentiment: Predicted price vs. Positive sentiment

After testing the correlation between news sentiment and Spot Gold price, we analyse the performance of Spot Gold price prediction. The results in Table 3 show that the prediction of 5-days news sentiment dataset demonstrate higher accuracy than 1-day news sentiment. The accuracy at 5-days can reach 78.43% of the real Spot Gold value. Furthermore, 5-days news sentiment predicted price brings to a lower L2 loss, which means the prediction is more stable and more precise.

TABLE III
PREDICTION RESULTS

Types	Exact prediction trend	Real trend	Accuracy	L2 loss
5-days news sentimnet	40	51	78.43%	0.0044
1-day news sentiment	38	53	71.70%	0.0082

Figure 5 and Figure 6, respectively for 5-days news Spot Gold price prediction and 1-day news Spot Gold price prediction, show the predicted trend compared with the real price. The 1-day news Spot Gold price prediction is more volatile.

However, in some days the predicted price differs slightly from the real price especially on February 25, February 27, April 18 to 20. 1-day news prediction price has 20$ to 30$ difference when compared with the real price. In these dates the real price had higher volatility. 5-days news prediction perform better than 1-day news prediction on those dates as the difference between the predicted value and real value are smaller. Compare with the trend of 1-day news prediction price and the trend of real price, the trend of 5-days news prediction price is more stable and smooth. Its training performance is better than 1-day news sentiment dataset.

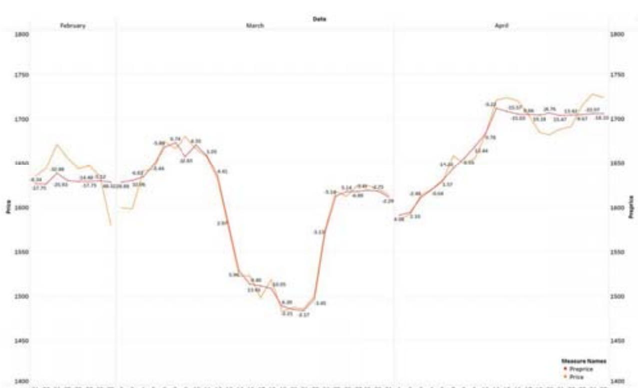

Fig. 5. 1-day news sentiment: Predicted price vs. Real price

V. CONCLUSIONS AND FUTURE WORKS

Financial market prediction has been always a hot topic for researchers. With the ongoing development of technologies, information and news flow even faster. Financial product prices are influenced by many factors such as news, events or some financial reports.

In this paper, financial news from several real World trustworthy platforms such as Yahoo! Finance, SCMP and Etnet, have been used for Spot Gold price prediction. VADER has been applied to do sentiment analysis, which transformed text data into numerical data. Instead of time series model, machine learning model (MLP) has been applied to train the Spot Gold price prediction model. The experimental results showed that by using our methods of financial news sentiment analysis to predict Spot Gold price, the prediction accuracy rate reached more than 78% and the prediction L2 loss value as low as 0.0044.

In future works we plan to apply the Recurrent Neural Network (RNN) model and especially Long-term Short-Term Memory (RNN-LSTM) to predict the Spot Gold price based on time-series data. Moreover, instead of using daily average price, the price of every minute will be used, so that the price difference of before-after the news post can create a new corpus to train the sentiment model.

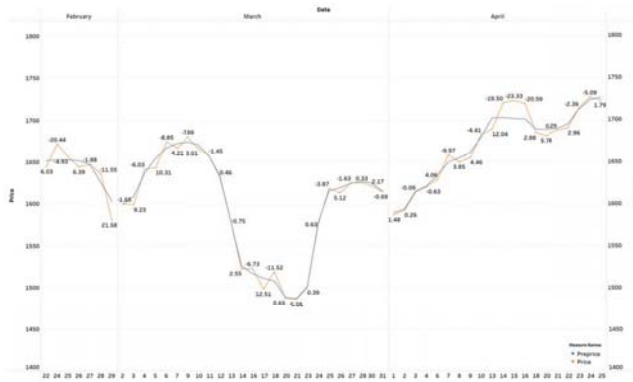

Fig. 6. 5-day news sentiment: Predicted price vs. Real price

ACKNOWLEDGMENT

This work was supported by the "Teaching Development Grant" - Department of Journalism, School of Communication, Hong Kong Baptist University, Kowloon Tong, Hong Kong, China

REFERENCES

[1] E. I. Altman. Financial ratios, discriminant analysis and the prediction of corporate bankruptcy. *The journal of finance*, 23(4):589–609, 1968.

[2] D. G. Baur and B. M. Lucey. Is gold a hedge or a safe haven? an analysis of stocks, bonds and gold. *Financial Review*, 45(2):217–229, 2010.

[3] D. G. Baur and T. K. McDermott. Is gold a safe haven? international evidence. *Journal of Banking & Finance*, 34(8):1886–1898, 2010.

[4] J. Bollen, H. Mao, and X. Zeng. Twitter mood predicts the stock market. *Journal of computational science*, 2(1):1–8, 2011.

[5] R. M. Brooks, A. Patel, and T. Su. How the equity market responds to unanticipated events. *The Journal of Business*, 76(1):109–133, 2003.

[6] R. C. Cavalcante, R. C. Brasileiro, V. L. Souza, J. P. Nobrega, and A. L. Oliveira. Computational intelligence and financial markets: A survey and future directions. *Expert Systems with Applications*, 55:194–211, 2016.

[7] S. W. Chan, V. Franzoni, P. Mengoni, and A. Milani. Context-based image semantic similarity for prosthetic knowledge. In *2018 IEEE First International Conference on Artificial Intelligence and Knowledge Engineering (AIKE)*, pages 254–258, 2018.

[8] M.-Y. Chen and T.-H. Chen. Modeling public mood and emotion: Blog and news sentiment and socio-economic phenomena. *Future Generation Computer Systems*, 96:692–699, 2019.

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

[10] W. F. De Bondt and R. Thaler. Does the stock market overreact? *The Journal of finance*, 40(3):793–805, 1985.

[11] J. J. Deng, C. H. C. Leung, P. Mengoni, and Y. Li. Emotion recognition from human behaviors using attention model. In *2018 IEEE First International Conference on Artificial Intelligence and Knowledge Engineering (AIKE)*, pages 249–253, 2018.

[12] L. H. Ederington and J. H. Lee. The short-run dynamics of the price adjustment to new information. *Journal of Financial and Quantitative Analysis*, pages 117–134, 1995.

[13] V. Franzoni, Y. Li, and P. Mengoni. A path-based model for emotion abstraction on facebook using sentiment analysis and taxonomy knowledge. In *Proceedings of the International Conference on Web Intelligence*, WI '17, page 947–952, New York, NY, USA, 2017. Association for Computing Machinery.

[14] V. Franzoni, Y. Li, P. Mengoni, and A. Milani. Clustering facebook for biased context extraction. In *Computational Science and Its Applications – ICCSA 2017*, pages 717–729, Cham, 2017. Springer International Publishing.

[15] I. A. Gheyas and L. S. Smith. A novel neural network ensemble architecture for time series forecasting. *Neurocomputing*, 74(18):3855–3864, 2011.

[16] 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.

[17] A. Grané and H. Veiga. Outliers, garch-type models and risk measures: A comparison of several approaches. *Journal of Empirical Finance*, 26:26–40, 2014.

[18] C. J. Hutto and E. Gilbert. Vader: A parsimonious rule-based model for sentiment analysis of social media text. In *Eighth international AAAI conference on weblogs and social media*, 2014.

[19] Y. Kara, M. A. Boyacioglu, and Ö. K. Baykan. Predicting direction of stock price index movement using artificial neural networks and support vector machines: The sample of the istanbul stock exchange. *Expert systems with Applications*, 38(5):5311–5319, 2011.

[20] W. Kristjanpoller and M. C. Minutolo. Gold price volatility: A forecasting approach using the artificial neural network–garch model. *Expert Systems with Applications*, 42(20):7245–7251, 2015.

[21] K. F. Kroner, K. P. Kneafsey, and S. Claessens. Forecasting volatility in commodity markets. *Journal of Forecasting*, 14(2):77–95, 1995.

[22] X. Li, H. Xie, Y. Song, S. Zhu, Q. Li, and F. L. Wang. Does summarization help stock prediction? a news impact analysis. *IEEE intelligent systems*, 30(3):26–34, 2015.

[23] X. Liang, R.-C. Chen, Y. He, and Y. Chen. Associating stock prices with web financial information time series based on support vector regression. *Neurocomputing*, 115:142–149, 2013.

[24] X. Man, T. Luo, and J. Lin. Financial sentiment analysis (fsa): A survey. In *2019 IEEE International Conference on Industrial Cyber Physical Systems (ICPS)*, pages 617–622. IEEE, 2019.

[25] P. Mengoni, A. Milani, and Y. Li. Multi-term semantic context elicitation from collaborative networks. In *2018 IEEE First International Conference on Artificial Intelligence and Knowledge Engineering (AIKE)*, pages 234–238, 2018.

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

[27] A. Nikfarjam, E. Emadzadeh, and S. Muthaiyah. Text mining approaches for stock market prediction. In *2010 The 2nd international conference on computer and automation engineering (ICCAE)*, volume 4, pages 256–260. IEEE, 2010.

[28] M. Nofer and O. Hinz. Using twitter to predict the stock market. *Business & Information Systems Engineering*, 57(4):229–242, 2015.

[29] B. Qian and K. Rasheed. Hurst exponent and financial market predictability. In *IASTED conference on Financial Engineering and Applications*, pages 203–209, 2004.

[30] R. Schumaker and H. Chen. Textual analysis of stock market prediction using financial news articles. *AMCIS 2006 Proceedings*, page 185, 2006.

[31] L. A. Smales. News sentiment in the gold futures market. *Journal of Banking & Finance*, 49:275–286, 2014.

[32] E. Tully and B. M. Lucey. A power garch examination of the gold market. *Research in International Business and Finance*, 21(2):316–325, 2007.

[33] B. Vanstone and G. Finnie. An empirical methodology for developing stockmarket trading systems using artificial neural networks. *Expert systems with Applications*, 36(3):6668–6680, 2009.

[34] B. Wang, H. Huang, and X. Wang. A novel text mining approach to financial time series forecasting. *Neurocomputing*, 83:136–145, 2012.

[35] D. Westerman, P. R. Spence, and B. Van Der Heide. Social media as information source: Recency of updates and credibility of information. *Journal of computer-mediated communication*, 19(2):171–183, 2014.

[36] F. Z. Xing, E. Cambria, and R. E. Welsch. Natural language based financial forecasting: a survey. *Artificial Intelligence Review*, 50(1):49–73, 2018.

[37] A. Yazdani-Chamzini, S. H. Yakhchali, D. Volungevičienė, and E. K. Zavadskas. Forecasting gold price changes by using adaptive network fuzzy inference system. *Journal of Business Economics and Management*, 13(5):994–1010, 2012.

[38] Y. Zhai, A. Hsu, and S. K. Halgamuge. Combining news and technical indicators in daily stock price trends prediction. In *International symposium on neural networks*, pages 1087–1096. Springer, 2007.

[39] B. Zhu and Y. Wei. Carbon price forecasting with a novel hybrid arima and least squares support vector machines methodology. *Omega*, 41(3):517–524, 2013.