HYBRID INFORMATION MIXING MODULE FOR STOCK MOVEMENT PREDICTION

1) Santhosh Kumar K – M. Tech- Computer Science and Engineering, Dept of CSE, SCET (Shadan College of Engineering and Technology) (Autonomous), Hyderabad.

2) Dr. Subramanian K. M Associate Professor, Dept of CSE, SCET (Shadan College of Engineering and Technology) (Autonomous), Hyderabad.

Abstract: A lot of research is being done on the area where DL and predicting the financial markets meet. This is because the stock market is very complex and is affected by many factors. To make good predictions, you have to think about how volatile the market is, since stock prices quickly reflect all the information that is available. This project suggests a new way to predict how stock prices will move by using both stock and news data. The main goal is to create a strong predictive model that can understand the complex connection between stock prices and textual news data. Because stock markets work so well, the goal is to make predictions more accurate by using a variety of data sources and taking market instability into account. Our method involves making a hybrid information mixing module that uses two mapblocks to make the interaction between price and text data features work well. This module pulls out the multimodal relationships between price series data over time and semantic features from news data. Then, a multilayer perceptron-based model is used to guess how the stock price will move. Using the proposed method in experiments that took place in a very volatile stock market showed promising outcomes. We saw big improvements in the accuracy of our predictions when we combined price and text data features and used the hybrid information mixing module. This ALSO used ARIMA and a mix of LSTM, GRU, and Bidirectional models, which made it much better at making predictions. With real-time data, this variety of approaches led to a stronger and more accurate final prediction.

"Index terms - *Stock movement prediction, time-series forecasting, bidirectional encoder representations from transformer (BERT), gated recurrent units (GRU), multilayer perceptron (MLP)".*

1. INTRODUCTION

DL is still being studied, but DL technology has now been used in the financial field. As the COVID-19 pandemic has made the stock market more volatile, it has become harder to accurately predict how stock prices will move. This has made it harder to do study on stock market forecasting. In NLP and the financial business, studies that try to guess stock prices are becoming more and more important. The stock market is very unstable and is affected by knowledge about companies and stock price indicators. Because of this, researchers are always looking for new ways to predict how stock prices will move by using different variables. First, two major studies conducted time-based studies to predict how stock prices would be deferred. Stock price data used and other textual data such as shared news and Twitter. [1], [8]. The openings, high, low, low prices, and end prices and end volumes of technical indicators can usually be inferred how stock data is used and how stock prices are moved. To figure out how variable time series data will be, people have come up with ways to learn about them using a "convolutional neural network (CNN) or a recurrent neural network (RNN) [1]". One problem with technical analyses that use stock price data is that they can't show trends that cause stock prices to change [2].

The relationships between companies also affect stock market volatility along with price and textual statistics. Studies on stock price prediction were performed using price, text, and business relation data to build and examine attention-based models to be influenced by stocks [6, 8]. Research is conducted to infer how stock prices move by examining financial data, social media and connections between stocks in a structured manner using a hierarchical graph of attention [6]. This is because different types of information can affect stock prices.

To display stock market information in text data and predict changes in sharing, the context information is determined by inserting the context word into the "Twisted Encoder Display (BERT) of the Transformer." Indications of the multimodal market in price and text in the multimodal era affect stocks [7]. At the time of semantic functions of price data and text data, series assignments have been eliminated and held together to make information and mixed duties from different modes. The pricing and text data function may work better together if you combine the functions of the mixture function through the hybrid information mixture modules.

It is considerably easier to put together the hybrid information mixing module in this article than the transformer-based variant. This is because other MLP-based models have MLP blocking [8, 9, 10, 11, 13, 14]. We employ an MLP-based model in this work to estimate what will happen to the stock prices in the hybrid information mixing module. This can help you make more accurate predictions about your inventory. This study uses Stocknet Data Record [1].

2. LITERATURE SURVEY

It's hard to tell how stocks will move because the market is very random and we use chaotic data to make estimates that depend on time [1]. We address these three issues and present an innovative, comprehensive approach that employs both textual and value signals to accomplish this task. Our model is distinct from discriminatory or subject modeling because it uses continuous, recurrent latent variables. This makes it better at dealing with uncertainty and neural variable estimates, which is how we tackle the problem of the rear entry that just won't work. We also give you a hybrid purpose with a temporary attachment that lets you record the future's dependence in diverse methods. We show that our proposed model is better at predicting stock movements than any other model when we use a new batch of data.

The trend of a time series shows how the series moves between going up and going down. It's useful to be able to learn and predict the trend in time series data for a wide range of real-world purposes, such as allocating resources in data centers and planning work loads in smart grids. We are writing about TreNet [2], a new "end-to-end hybrid neural network that learns local and global contextual features to predict the direction of time series". This idea comes from the recent successes of neural networks. TreNet uses CNNs [24, 27] to pull out important features from time series data that is stored locally. TreNet, on the other hand, uses a LSTM [34] to understand the long-range dependencies in the order of past trends. After that, a feature fusion layer will learn how to use joint representation to guess the trend. On real datasets, TreNet does better than CNN, LSTM, the combination of CNN and LSTM, a Hidden Markov Model-based method, and a number of kernel-based baselines.

The stock index is used to judge investments and to see how the economy of a country is doing. This is why study into how to predict the stock index is still going on. There are technical, basic, and psychological factors that go into predicting the stock price index. For accurate predictions, you also need to think about complicated factors. Therefore, models need to be addressed to predict stock indexes By selecting and articulating the technical and ancillary elements influencing the methodology grounded in share value. Most of the research done on this subject predicts that studies would either employ news or a lot of economic indicators that affect market fluctuations, or they will simply look at a few indicators. In this study [3], we aim to demonstrate that news mood analysis and several macroeconomic factors are employed to accurately predict the USDO Jones index. Mood analysis was done on more than 93,000 New York Times commercial stories two years later. Burt and NLTK are two examples of natural language processing technologies that are new. We also looked at five macroeconomic indicators, gold and oil prices, and five exchange rates that have an impact on the US economy. Prediction Methods LSTM [19, 20] are known to combine optimal numbers and text, and were used in combination. “The DJI index forecast with the NLTK, BERT, OIL, GOLD, and EURUSD” combinations gave the smallest MSE value after trying out different ones.

If you want to make smart trades in the stock market, you need to be able to guess stock prices. In recent years, short-term predictions based on financial news stories have shown a lot of promise. In this study [4], we suggest a new DNN called DP-LSTM for predicting stock prices. The network uses various privacy mechanisms to use news articles as hidden information and different news sources. Based on the "Author Split Average Model (ARMA)," a sentiment arm is created by adding information from the financial message. Then, a DNN based on LSTM is constructed. The LSTM, the Vader model, and the "Differential Privacy (DP)" mechanism make up this network. The suggested DP-LSTM approach cuts down on mistakes in predictions and makes the system more reliable. We did a number of tests on the S&P 500 strains and discovered that the suggested DP-LSTM makes the intermediate products better. MPA “by 0.32% and (ii) it improves the MSE of the predictions for the S&P 500 market index by up to 127%”.

Recently, a lot of people have been interested in the stock market. As the inflation rate has gone up, people have had to invest their money in the stock market, commodities, and other places instead of saving it. Also, DL models have been shown over and over again to be able to make predictions based on time series data. When traders and investors look at the stock market, they most often use technical analysis with the help of technical indicators. Another part is sentiment analysis, which looks at how investors feel and how willing they are to invest. People all over the world have used basic ML [45] and Neural Networks in a number of different ways. From the most basic linear regression to the most complex NN People have tried every method they can think of to try to guess what will happen in the stock market. Recent events show that news and stories have an impact on the cryptocurrency and stock markets. This paper [5] suggests a group of cutting-edge ways to guess what stock prices will do. First, a version of BERT, a pre-trained Google NLP transformer model, is used to look at how people feel about the news and stories about Apple Inc., "which is listed on the NASDAQ. After that, a Generative Adversarial Network (GAN)" guesses what Apple Inc. stock price will be by looking at sentiment scores, technical indicators, market indexes from different countries, some commodities, and past prices. "Baseline models, such as the Long Short Term Memory (LSTM), Gated Recurrent Units (GRU) [19, 20, 22], plain GAN, and the Auto-Regressive Integrated Moving Average (ARIMA)" model, are used to compare.

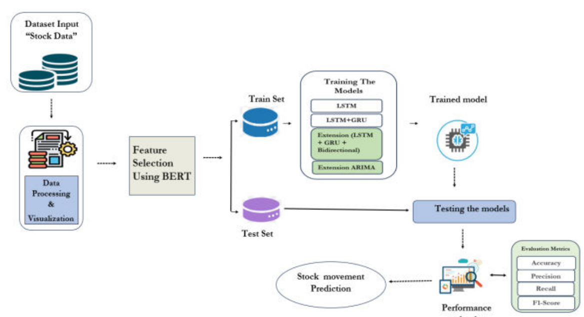
3. METHODOLOGY

i) Proposed Work:

"Hybrid Information Mixing Module," the suggested system, takes stock and news data and uses DL models like LSTM, GRU, [19, 20, 22], and BERT to get time-series and semantic features. These features are combined into mixed features that show how different data sources interact to help us better understand market signals. It has a mixed information mixing module that uses two MLP blocks to effectively integrate features. This makes predictions more accurate while using less computing power than transformer-based models. Adding ARIMA and a mixed method that combines LSTM, GRU, [19], and Bidirectional models also made it much easier to make predictions. When applied to real-time data, this variety of approaches led to a stronger and more accurate end prediction. Also, to make testing and interacting with users easier, a frontend interface with login features was built using Flask. This made it easy for users to get involved and check how well the prediction models worked in real-life situations.

ii) System Architecture:

This study proposes a hybrid mix of information that can be used to infer how stock prices move. "The hybrid information mixing module can be divided into three parts as seen in Figure 1". These are hybrid information mixing modules, binary classifiers. The mixed information module consists of the ability to mix MLPs and the interaction to mix MLPs [8, 9, 10, 11, 13, 14]. Embedded characteristics include embedding price and text embedding. The binary classifier tells you whether the predicted change of the stock price is up or down.



"Fig 1 Proposed architecture"

iii) Dataset collection:

StockNet data record [1], which consists of historical price records and Twitter data records, is used to test "the learning and performance of the hybrid information mix module in this work". The historical price dataset has data on stocks that were moved a lot in the "New York Stock Exchange (NYSE) and

NASDAQ markets that are part of the Standard and Poor's (S&P) 500 index. NASDAQ ticker symbols, like Google's $GOOG", are used in regex searches to pull out a Twitter dataset.

	Date	Open	High	Low	Close	Adj Close	Volume	Label
0	2012-09-04	95.108574	96.448570	94.928574	96.424286	87.121140	91973000	1
1	2012-09-05	96.510002	96.621429	95.657143	95.747147	86.509338	84093800	0
2	2012-09-06	96.167145	96.898575	95.828575	96.610001	87.288956	97799100	1
3	2012-09-07	96.864288	97.497147	96.538574	97.205711	87.827171	82416600	1
4	2012-09-10	97.207146	97.612854	94.585716	94.677139	85.542564	121999500	0
...	...	...	...	...	...	...	...	...
1253	2017-08-28	160.139999	162.000000	159.929993	161.470001	161.470001	25966000	1
1254	2017-08-29	160.100006	163.119995	160.000000	162.910004	162.910004	29516900	1
1255	2017-08-30	163.800003	163.889999	162.610001	163.350006	163.350006	27269600	1
1256	2017-08-31	163.639999	164.520004	163.479996	164.000000	164.000000	26785100	1
1257	2017-09-01	164.800003	164.940002	163.630005	164.050003	164.050003	16552800	1

1258 rows × 8 columns

Fig 2 dataset

iv) Data Processing:

DP is the act of turning useless data into useful information for the business. You are in charge of computer researchers most of the time. This means that it is collected, sorted, cleansed, looked into, analyzed, and turned into formats that are easy to read, such visuals and papers. There are three techniques to treat data mechanically or electrically by hand. The purpose is to spread knowledge that is valuable and helps people make decisions. This helps businesses run more smoothly and make smart judgments about their strategies. Automatic data processing tools, such computer programming, make this mostly achievable. It helps change big data and other kinds of data into valuable information that may be used for quality control and making decisions.

v) Feature selection:

Features selection is the process in which the most consistent, most convenient, and less redundant features can be used to create models. With the number and record types increasing, it is important to reduce the size in a planned manner. One of the main objectives of feature selection is to reduce the better capabilities and computing power of the predictive model.

One of the most important parts of functional engineering is the selection of characteristics, which is the process of selecting the most important features for feeding with ML algorithms. The distinctive selection method removes unnecessary or useless features and retains only the features that are most important to the ML model. This reduces the number of input variables. Instead of having your ML model do this, selecting which functions are important in advance will have the main advantage here.

vi) Algorithms:

"LSTM: Long Short-Term Memory" is a better form of the RNN. LSTM is great at recording long-term dependencies and is great at predicting sequences. It can be used for jobs that involve time series and sequences. The best thing about LSTM is that it can understand order dependence, which is important for handling hard problems like speech recognition and machine translation [4, 20].

LSTMs have a memory cell with three gates: input, forget, and output gates. This is different from regular RNNs, which only have one hidden state. These gates manage the adding, removing, and sending of data from the memory cell. This lets the network choose which data to keep and which to delete. Because of this, LSTMs can be used for jobs like translating languages, recognizing speech, and making predictions based on time series. LSTMs can also be mixed with other designs, like CNNs, to make them more useful for tasks like analyzing images and videos.

Most likely, LSTM is used in this project because it is good at modeling sequential data, like historical stock prices or text from news stories. The model can find temporal patterns and dependencies in the raw data by using LSTM. These are very important for predicting stock prices. LSTM can also remember things for a long time, which makes it good for finding trends and patterns in time-series data.

"Gated Recurrent Unit (GRU)": The LSTM architecture is put together with the GRU architecture, which is another type of gated RNN. "Recurrent neural network (RNN) is the type of network it is. For easier use instead of Long Short-Term Memory (LSTM)" networks. In the same way that LSTM can, GRU can handle sequential data like speech, text, and time lines. The main idea behind GRU is to use gates to update only some of the network's secret states for each time step. Gating systems manage data flows both inside and outside the network. The reset gate and the update target are two goals that make up the GRU. This gate determines how many of the previous secret states to be removed. This gate also determines whether to use the new input to change the hidden state. The GRU decides what to do based on the latest hidden information [19, 20].

The project probably uses this mix of LSTM and GRU to get the most out of the best parts of both designs. LSTM is great at recording long-term dependencies, and GRU is good at making computations faster. By putting them together, the model might be able to do a better job of predicting stock prices, both in terms of accuracy and speed.

4. EXPERIMENTAL RESULTS

Precision: The accuracy is the percentage of tests or situations that are correctly categorized as positive matches. So this is how we find out if something is accurate:

"Precision = True positives/ (True positives + False positives) = TP/(TP + FP)"

$$\text{Precision} = \frac{True\ Positive}{True\ Positive + False\ Positive}$$

Recall: In ML, a return is a number that shows how successfully a model can locate all the significant examples of a certain class. This shows how well a specific version of the model works. To get this number, you divide the number of approximate positive comments by the actual number of positive comments.

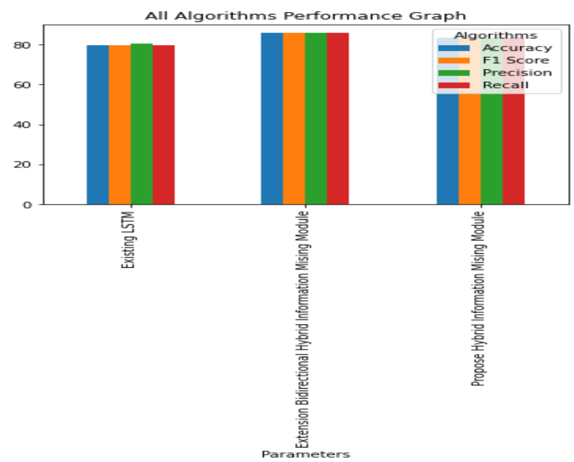
$$Recall = \frac{TP}{TP + FN}$$

"Accuracy: Accuracy is the percentage of correct predictions in a classification task". This indicates how accurate the forecasts of the right models are.

$$Accuracy = \frac{TP + TN}{TP + FP + TN + FN}$$

"F1 Score: The F1 score is a good way to measure both accuracy and memory". This is a good solution that works with both false positives and false negatives, and it can be used with data records that aren't balanced.

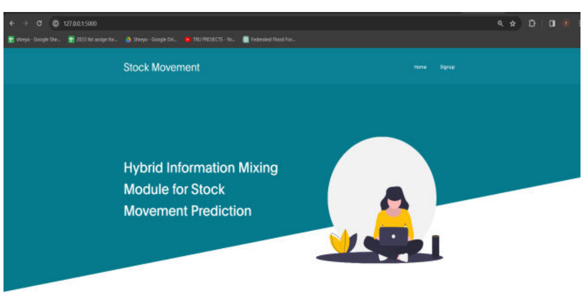
$$\text{F1 Score} = 2 * \frac{\text{Recall} \times \text{Precision}}{\text{Recall} + \text{Precision}} * 100$$



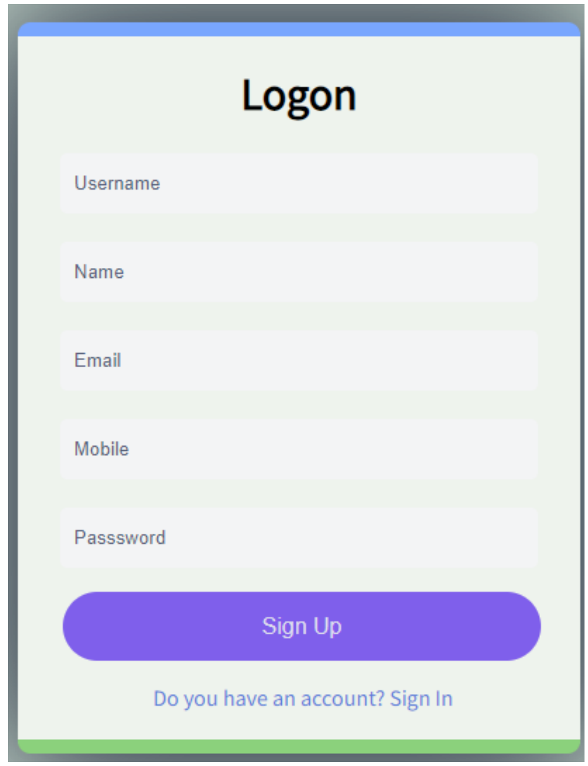
"Fig 3 Comparison graph"

Algorithm Name	Precison	Recall	FScore	Accuracy
Existing LSTM	80.300896	79.814173	79.689945	79.761905
Propose Hybrid Information Missing Module	83.414418	83.313386	83.316520	83.333333
Extension Bidirectional Hybrid Information Mis...	85.846970	85.738583	85.706183	85.714286

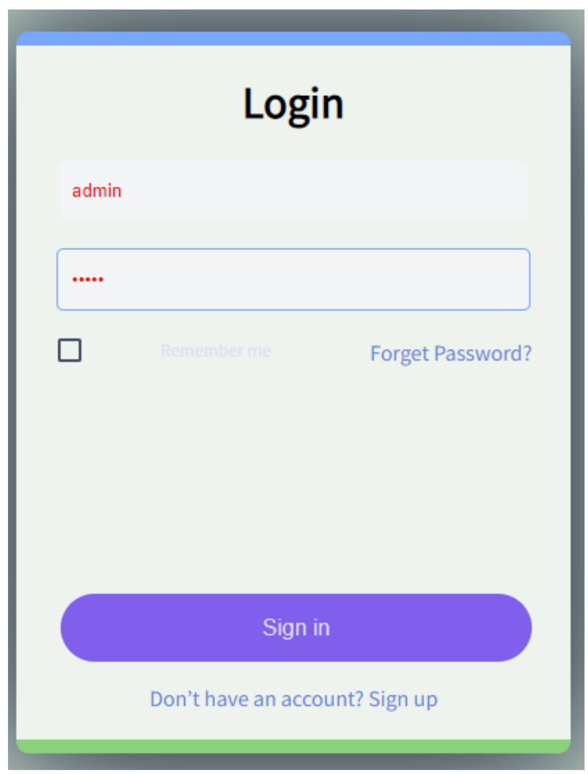
"Fig 4 Performance Evaluation"



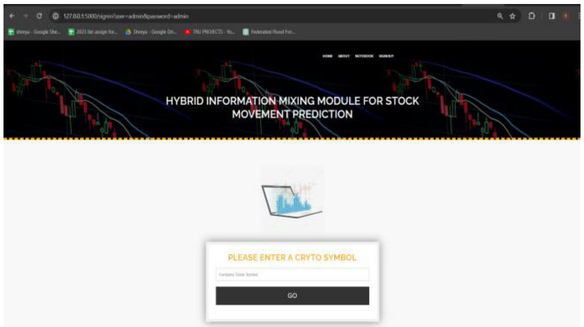
"Fig 5 Home page"



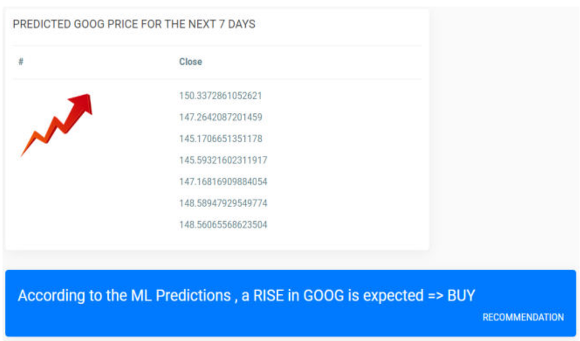
"Fig 6 Signin page"



"Fig 7 Login page"



"Fig 8 User input"



"Fig 9o Predict result for given input"

5. CONCLUSION

The new fusion method that combines features from price and text data sources makes the model better at making predictions by taking into account a wider range of market signs and information sources for more accurate predictions. By combining LSTM and GRU [19, 20] designs, the project is better at predicting how stock prices will move. This is because it uses the power of recurrent neural networks to find patterns over time in both numerical and textual data. Using LSTM and LSTM + GRU together makes it possible to model sequential data like historical stock prices and news stories well. This lets the model see complex dependencies and trends that are necessary for making accurate predictions. Using the Flask framework and SQLite database, the project provides an easy-to-use interface for users to register, log in, and upload company ticker symbols. This makes it easier for users to connect with and use the stock prediction system. The ARIMA algorithm was very good at predicting the future; it made accurate predictions about how stock prices would move with very few mistakes. Testing with real stock data on the front end proved that it was reliable and strong, giving users useful information for making smart decisions. The project gives users accurate predictions of how stock prices will move based on historical data and real-time news. This gives investors and traders the power to make smart choices, which improves their ability to navigate the constantly changing financial markets.

6. FUTURE SCOPE

By including additional data that influences stock market volatility in future research and enhancing hybrid information mixing modules, we may analyze the impact of many variables on stock markets. You can utilize the company's connection data as an extra data source to see how all the companies are linked to each other and how they affect each other. You can make multimodal information that contains three types of stock market information by putting together price data, lesson data, and business link data. Better hybrid information mixture modules [8, 9, 10, 11, 13, 14] let you blend different kinds of information. You can do a survey to see how the stock market is changing by looking at the relationship between three types of multimodal information sharing and dynamic markets.

REFERENCES

[1] Y. Xu and S. B. Cohen, ''Stock movement prediction from tweets and historical prices,'' in Proc. 56th Annu. Meeting Assoc. Comput. Linguistics, Melbourne, VIC, Australia, 2018, pp. 1970–1979.

[2] T. Lin, T. Guo, and K. Aberer, ''Hybrid neural networks for learning the trend in time series,'' in Proc. 26th Int. Joint Conf. Artif. Intell., Melbourne, VIC, Australia, Aug. 2017, pp. 2273–2279

[3] E. Jang, H. R. Choi, and H. C. Lee, ''Stock prediction using combination of BERT sentiment analysis and macro economy index,'' J. Korea Soc. Comput. Inf., vol. 25, no. 5, pp. 47–56, 2020.

[4] X. Li, Y. Li, H. Yang, L. Yang, and X.-Y. Liu, ''DP-LSTM: Differential privacy-inspired LSTM for stock prediction using financial news,'' 2019, arXiv:1912.10806.

[5] P. Sonkiya, V. Bajpai, and A. Bansal, ''Stock price prediction using BERT and GAN,'' 2021, arXiv:2107.09055.

[6] R. Sawhney, S. Agarwal, A. Wadhwa, and R. R. Shah, ''Deep attentive learning for stock movement prediction from social media text and company correlations,'' in Proc. Conf. Empirical Methods Natural Lang. Process. (EMNLP), 2020, pp. 8415–8426. [Online]. Available: https://aclanthology.org/2020.emnlp-main.676

[7] Y. Zhao, H. Du, Y. Liu, S. Wei, X. Chen, F. Zhuang, Q. Li, J. Liu, and G. Kou, ''Stock movement prediction based on bi-typed hybridrelational market knowledge graph via dual attention networks,'' 2022, arXiv:2201.04965.

[8] I. Tolstikhin, N. Houlsby, A. Kolesnikov, L. Beyer, X. Zhai, T. Unterthiner, J. Yung, D. Keysers, J. Uszkoreit, M. Lucic, and A. Dosovitskiy, ''MLPmixer: An all-MLP architecture for vision,'' in Proc. Adv. Neural Inf. Process. Sys., vol. 34, 2021, pp. 24261–24272. [Online]. Available: https://proceedings.neurips.cc/paper/2021/hash/cba0a 4ee5ccd02fda0fe3f 9a3e7b89fe-Abstract.html

[9] H. Liu, Z. Dai, D. So, and Q. V. Le, ''Pay attention to MLPs,'' in Proc. Int. Conf. Neural Inf. Process. Syst., vol. 34, 2021, pp. 9204–9215. [Online]. Available: https://openreview.net/pdf?id=KBnXrODoBW

[10] M. Li, X. Zhao, C. Lyu, M. Zhao, R. Wu, and R. Guo, ''MLP4Rec: A pure MLP architecture for sequential recommendations,'' in Proc. 31st Int. Joint Conf. Artif. Intell., Vienna, Austria, Jul. 2022, pp. 2138–2144.

[11] H. Touvron, P. Bojanowski, M. Caron, M. Cord, A. El-Nouby, E. Grave, G. Izacard, A. Joulin, G. Synnaeve, J. Verbeek, and H. Jégou, ''ResMLP: Feedforward networks for image classification with data-efficient training,'' 2021, arXiv:2105.03404.

[12] F. Mai, A. Pannatier, F. Fehr, H. Chen, F. Marelli, F. Fleuret, and J. Henderson, ''HyperMixer: An MLP-based green AI alternative to transformers,'' 2022, arXiv:2203.03691.

[13] R. Kim, C. H. So, M. Jeong, S. Lee, J. Kim, and J. Kang, ''HATS: A hierarchical graph attention network for stock movement prediction,'' 2019, arXiv:1908.07999.

[14] X. Ding, Y. Zhang, T. Liu, and J. Duan, ''Deep learning for event-driven stock prediction,'' in Proc. 24th Int. Joint Conf. Artif. Intell., Buenos Aires, Argentina, 2015, pp. 2327–2333.

[15] Z. Hu, W. Liu, J. Bian, X. Liu, and T.-Y. Liu, ''Listening to chaotic whispers: A deep learning framework for news-oriented stock trend prediction,'' in Proc. 11th ACM Int. Conf. Web Search Data

Mining, New York, NY, USA, Feb. 2018, pp. 261–269.

[16] Q. Liu, X. Cheng, S. Su, and S. Zhu, ''Hierarchical complementary attention network for predicting stock price movements with news,'' in Proc. 27th ACM Int. Conf. Inf. Knowl. Manage., Italy, Oct. 2018, pp. 1603–1606.

[17] X. Du and K. Tanaka-Ishii, ''Stock embeddings acquired from news articles and price history, and an application to portfolio optimization,'' in Proc. 58th Annu. Meeting Assoc. Comput. Linguistics, Seattle, WA, USA, 2020, pp. 3353–3363.

[18] L. Zhang, C. Aggarwal, and G.-J. Qi, ''Stock price prediction via discovering multi-frequency trading patterns,'' in Proc. 23rd ACM SIGKDD Int. Conf. Knowl. Discovery Data Mining, Halifax, NS, Canada, Aug. 2017, pp. 2141–2149.

[19] T. H. Nguyen and K. Shirai, ''Topic modeling based sentiment analysis on social media for stock market prediction,'' in Proc. 53rd Annu. Meeting Assoc. Comput. Linguistics 7th Int. Joint Conf. Natural Lang. Process., Beijing, China, vol. 1, 2015, pp. 1354–1364.

[20] X. Li, H. Xie, L. Chen, J. Wang, and X. Deng, ''News impact on stock price return via sentiment analysis,'' Knowl.-Based Syst., vol. 69, pp. 14–23, Oct. 2014.

[21] C. Rui and Q. Li, ''Modeling the momentum spillover effect for stock prediction via attribute-driven graph attention networks,'' in Proc. AAAI Conf. Artif. Intell., vol. 35 no. 1, 2021, pp. 55–62. [Online]. Available: https://ojs.aaai.org/index.php/AAAI/article/view/16077

[22] K. Stankeviciute, A. M. Alaa, and M. van der Schaar, ''Conformal timeseries forecasting,'' in Proc. Adv. Neural Inf. Process. Sys., vol. 34, 2021, pp. 6216–6228. [Online]. Available: https://proceedings.neurips.cc/paper/2021/hash/312f1ba2a72318edaaa995a67835fad5-Abstract.html

[23] C. Li, D. Song, and D. Tao, ''Multi-task recurrent neural networks and higher-order Markov random fields for stock price movement prediction: Multi-task RNN and higer-order MRFs for stock price classification,'' in Proc. 25th ACM SIGKDD Int. Conf. Knowl. Discovery Data Mining, Anchorage, AK, USA, Jul. 2019, pp. 1141–1151.

[24] D. M. Q. Nelson, A. C. M. Pereira, and R. A. de Oliveira, ''Stock market's price movement prediction with LSTM neural networks,'' in Proc. Int. Joint Conf. Neural Netw. (IJCNN), Anchorage, AK, USA, May 2017, pp. 1419–1426.

[25] V. S. Pagolu, K. N. Reddy, G. Panda, and B. Majhi, ''Sentiment analysis of Twitter data for predicting stock market movements,'' in Proc. Int. Conf. Signal Process., Commun., Power Embedded Syst. (SCOPES), Paralakhemundi, Odisha, India, Oct. 2016, pp. 1345–1350.

[26] X. Ding, Y. Zhang, T. Liu, and J. Duan, ''Using structured events to predict stock price movement: An empirical investigation,'' in Proc. Conf. Empirical Methods Natural Lang. Process. (EMNLP), Doha, Qatar, 2014, pp. 1415–1425.

[27] Z. Zhou, L. Ma, and H. Liu, ''Trade the event: Corporate events detection for news-based event-driven trading,'' 2021, arXiv:2105.12825.

[28] K. Cho, B. van Merrienboer, C. Gulcehre, D. Bahdanau, F. Bougares, H. Schwenk, and Y. Bengio, ''Learning phrase representations using RNN encoder–decoder for statistical machine translation,'' in Proc. Conf. Empirical Methods Natural Lang. Process. (EMNLP), Doha, Qatar, 2014, pp. 1724–1734.

[29] P. Murugan, ''Learning the sequential temporal information with recurrent neural networks,'' 2018, arXiv:1807.02857.

[30] G. P. C. Fung, J. X. Yu, and W. Lam, ''News sensitive stock trend prediction,'' in Proc. Pacific-Asia Conf. Knowl. Discovery Data Mining. Berlin, Germany: Springer, 2002, pp. 481–493.

[31] Z. Lu, D. Pan, and J.-Y. Nie, ''VGCN-BERT: Augmenting BERT with graph embedding for text classification,'' in Proc. 42nd Eur. Conf. Inf. Retrieval., Lisbon, Portugal: Springer, 2020, pp. 369–382.

[32] F. J. Provost, T. Fawcett, and R. Kohavi, ''The case against accuracy estimation for comparing induction algorithms,'' in Proc. ICML, vol. 98, 1998, pp. 445–453.

[33] K. Bandara, C. Bergmeir, and S. Smyl, ''Forecasting across time series databases using recurrent neural networks on groups of similar series: A clustering approach,'' Exp. Syst. Appl., vol. 140, Feb. 2020, Art. no. 112896.

[34] K. A. Althelaya, E.-S.-M. El-Alfy, and S. Mohammed, ''Stock market forecast using multivariate analysis with bidirectional and stacked (LSTM, GRU),'' in Proc. 21st Saudi Comput. Soc. Nat. Comput. Conf. (NCC), Riyadh, Saudi Arabia, Apr. 2018, pp. 1–7.

[35] S. Hochreiter and J. Schmidhuber, ''Long short-term memory,'' Neural Comput., vol. 9, no. 8, pp. 1735–1780, 1997.

[36] A. Vaswani, N. Shazeer, N. Parmar, J. Uszkoreit, L. Jones, A. N. Gomez, L. Kaiser, and I. Polosukhin, ''Attention is all you need,'' in Proc. Adv. Neural Inf. Process. Sys., Long Beach, CA, USA, 2017, pp. 5998–6008.

[37] Y. Su, Y. Zhao, C. Niu, R. Liu, W. Sun, and D. Pei, ''Robust anomaly detection for multivariate time series through stochastic recurrent neural network,'' in Proc. 25th ACM SIGKDD Int. Conf. Knowl. Discovery Data Mining, Anchorage, AK, USA, Jul. 2019, pp. 2828–2837.

[38] X. Wu, ''Event-driven learning of systematic behaviours in stock markets,'' in Proc. Findings Assoc. Comput. Linguistics. (EMNLP), 2020, pp. 2434–2444. [Online]. Available: https://aclanthology.org/2020.findings-emnlp.220

[39] H. Wang, T. Wang, and Y. Li, ''Incorporating expert-based investment opinion signals in stock prediction: A deep learning framework,'' in Proc. AAAI Conf. Artif. Intell., New York, NY, USA, vol. 34, no. 1, 2020, pp. 971–978.

[40] R. Sawhney, S. Agarwal, A. Wadhwa, T. Derr, and R. R. Shah, ''Stock selection via spatiotemporal hypergraph attention network: A learning to rank approach,'' in Proc. AAAI Conf. Artif. Intell., New York, NY, USA, vol. 35, no. 1, 2021, pp. 497–504.

[41] H. Wu, W. Zhang, W. Shen, and J. Wang, ‘‘Hybrid deep sequential modeling for social text-driven stock prediction,’’ in Proc. 27th ACM Int. Conf. Inf. Knowl. Manage., Italy, Oct. 2018, pp. 1627–1630.

[42] J. Yoo, Y. Soun, Y.-C. Park, and U. Kang, ‘‘Accurate multivariate stock movement prediction via data-axis transformer with multi-level contexts,’’ in Proc. 27th ACM SIGKDD Conf. Knowl. Discovery Data Mining, Singapore, Aug. 2021, pp. 2037–2045.

[43] G. Wang, L. Cao, H. Zhao, Q. Liu, and E. Chen, ‘‘Coupling macro-sector-micro financial indicators for learning stock representations with less uncertainty,’’ in Proc. AAAI Conf. Artif. Intell., vol. 35, no. 5, 2021, pp. 4418–4426. [Online]. Available: https://ojs.
aaai.org/index.php/AAAI/article/view/16568

[44] S. Mukherjee, B. Sadhukhan, N. Sarkar, D. Roy, and S. De, ‘‘Stock market prediction using deep learning algorithms,’’ in Proc. CAAI Trans. Intell. Technol., 2021, pp. 1–13.

[45] S. Mehtab and J. Sen, ‘‘A time series analysis-based stock price prediction using machine learning and deep learning models,’’ 2020, arXiv:2004.11697