STOCK MARKET ANALYSIS AND PREDICTION

Saurabh Kumar Mishra
(1812010095)
Institute of technology and Management Gida Gorakhpur

Mohd. Aqdas
(1812010065)
Institute of technology and Management Gida Gorakhpur

Ajay Kumar
(Assistant Professor)

ABSTRACT

In Stock Market Prediction, the goal is to are expecting the destiny fee of the monetary shares of an organization. The latest fashion in inventory marketplace prediction technology is the usage of system gaining knowledge of which makes predictions primarily based totally at the values of modern inventory marketplace indices with the aid of using schooling on their preceding values In Stock Market Prediction, the goal is to are expecting the destiny fee of the monetary shares of an organization. The latest fashion in inventory marketplace prediction technology is the usage of system gaining knowledge of which makes predictions primarily based totally at the values of modern inventory marketplace indices with the aid of using schooling on their preceding values. Machine gaining knowledge of itself employs distinctive fashions to make prediction less difficult and authentic. The paper makes a specialty of the usage of Regression and LSTM primarily based totally Machine gaining knowledge of expecting inventory values. Factors taken into consideration are open, close, low, excessive, and volume.

1. INTRODUCTION

Stock marketplace prediction is the act of looking to decide the destiny fee of an organization inventory or different monetary instrument traded on an exchange. The hit prediction of an inventory's destiny rate may want to yield huge profit. The green-marketplace hypothesis indicates that inventory costs mirror all presently to be had facts and any rate modifications that aren't primarily based totally on newly found out facts accordingly are inherently unpredictable. Others disagree and people with this point of view own myriad techniques and technology which purportedly permit them to advantage destiny rate facts, The green marketplace hypothesis posits that inventory costs are a characteristic of facts and rational expectations and that newly found out facts approximately an organization’s potentialities are nearly right now pondered withinside the modern inventory rate. This could suggest that everyone publicly regarded facts approximately an organization, which manifestly consists of its rate history, could already be pondered withinside the modern rate of the inventory. Accordingly, modifications withinside the inventory rate mirror launches of recent facts, modifications withinside the marketplace generally, or random moves across the fee that displays the prevailing the facts set. Burton Malkiel, in his influential 1973 paintings A Random Walk Down Wall Street, claimed that inventory costs may want to consequently now no longer be correctly expected with the aid of using searching at rate history. As a result, Malkiel argued, inventory costs are fine defined with the aid of using a statistical technique referred to as a "random walk" which means every day's deviations from the important fee are random and unpredictable. This led Malky to finish that paying monetary offerings folks to are expecting the marketplace clearly hurt, as opposed to helping, internet portfolio goes back. A range of empirical exams and the belief that the concept applies generally, as maximum portfolios controlled with the aid of using expert inventory predictors do now no longer outperform the marketplace common go back after accounting for the managers' fees. Fundamental evaluation is constructed at the perception that human society desires capital to make development and if an organization operates properly, it has to be rewarded with extra capital and bring about a surge in inventory rate. Fundamental evaluation is broadly utilized by fund managers as it's miles the maximum reasonable, goal and crafted from publicly to be had facts like monetary assertion evaluation. Another means of essential evaluation is beyond bottom-up organization evaluation, it refers to top-down evaluation from first studying the worldwide economy, observed with the aid of using. evaluation after which region evaluation, and subsequently the organization stage evaluation. The risk posed with the aid of using cyberterrorism has grabbed the eye of the mass media, the safety community, and the facts era (IT) industry. Journalists, politicians, and professionals in numerous fields have popularized a state of affairs wherein state-of-the-art cyberterrorists electronically damage into computer systems that manage dams or air visitors manage structures, wreaking havoc and endangering now no longer best tens of thousands and thousands of lives however countrywide safety itself. And yet, notwithstanding all of the gloomy predictions of a cybergenerated doomsday, no unmarried example of actual cyberterrorism has been recorded. Just how actual is the risk that cyberterrorism poses? Because maximum crucial infrastructure in Western societies is networked via computer systems, the ability risk from cyberterrorism is, to be sure, very alarming. Hackers, even though now no longer prompted with the aid of using the equal dreams that encourage terrorists, have verified that people can advantage get admission to touchy facts and to the operation of vital offerings. Terrorists, as a minimum in concept, may want to accordingly comply with the hackers' lead after which, having damaged into authorities and personal laptop structures, cripple or as a minimum disable the military, monetary, and carrier sectors of superior economies. The developing dependence of our societies on facts era has created a brand-new shape of vulnerability, giving terrorists the threat to technique objectives that might in any other case be fully unassailable, along with countrywide protection structures and air visitors manage structures. The extra technologically evolved a rustic is, the extra inclined it turns into to cyberattacks towards its infrastructure. Concern approximately the ability hazard posed with the aid of using cyberterrorism is accordingly properly founded. That does now no longer mean, however, that every-one the fears which have been voiced withinside the media, in Congress, and indifferent public boards are rational and reasonable. Some fears are clearly unjustified, at the same time as others are distinctly exaggerated. In addition, the difference between the ability and the real harm inflicted with the aid of using cyberterrorists has too frequently been ignored, and the extraordinarily benign sports of maximum hackers had been conflated with the threat of natural cyberterrorism. This record examines the fact of the cyberterrorism risk, gift, and destiny. It starts with the aid of using outlining why cyberterrorism angst has gripped such a lot of people, defines what qualifies as "cyberterrorism" and what does now no longer, and charts cyberterrorism's attraction for terrorists. The record then appears on the proof each for and towards Western society's vulnerability to cyberattacks, drawing on a selection of latest research and guides to demonstrate the types of fears which have been expressed and to evaluate whether or not we want to be so concerned. The end appears to the destiny and argues that we ought to continue to be alert to actual risks at the same time as now no longer turning into sufferers of overblown fears. Psychological, political, and monetary forces have mixed to sell the concern of cyberterrorism. From a mental perspective, the finest fears of the current time are mixed withinside the time period "cyberterrorism." The worry of random, violent victimization blends properly with the mistrust and outright worry of the laptop era. An unknown risk is perceived as extra threatening than a regarded risk.

2. RELATED WORK

2.1. Stock Movement Prediction

Owing to the potentially large profit, stock movement prediction has been recognized as an interesting problem in the research area. According to the core concept of predicting future trends in the stock market, the related works can be further divided into two categories, which are technical analysis and news-oriented analysis. Technical analysis takes historical prices of the stock as a feature to forecast its future movement. Tao Lin, Tian Guo, and Karl Aberer [12] introduce Trent to extract features from raw time series and construct the dependency across multiple time steps. Zhang et al. [5] inspired by Discrete Fourier Transform, propose a modified version of long short term recurrent neural network (LSTM) which called State Frequency Memory (SFM) recurrent network to capture the multi-frequency trading patterns from past price data to make a future prediction in both short and long term. All of them are also evidence to prove that deep learning method are beneficial to mine effective patterns and capture nonlinearity from historical prices of stock. The news-oriented analysis takes historical prices as features, but they also extract features from the natural language processing (NLP) method as additional input, simulating the social impact on the stock market. For instance, Hu et al. [7] imitate the learning process of human beings facing such chaotic online news to predict future trends based on recent related news. Yuma Xu and Shay B. Cohen [8] also treat the stock market as a stochastic environment, jointly exploiting text and price signals for the prediction task. The above studies show that related features of the stock itself are essential in stock movement prediction. However, few consider inter-market or inter-company relations in the current financial market environment. Liu et al. [10] comprehensively use the news sentiment as the correlation between stock companies, accompanied by stock price data, to predict the future trend. Utilizing the knowledge graph and its embedding model, they find that their method can improve significantly over baseline. Matsunaga et al. [9] utilize the knowledge graph as additional input, proposing a graph convolutional neural network combining prices and knowledge graph data. Their result also proves that knowledge graphs and historical prices in financial markets hold a strong promise in creating a stable and practical stock market predictor. Deng et al. [11] extract knowledge graphs from the online news of each company. Treat all entities and edges as Trans E input to train a knowledge-driven event embedding model for each company. In combination with TCN and Trans E models, their framework can significantly outperform deep models and be explainable over prediction results. Li et al. [13] assume the stocks in the market are not dependent, modeling the connection among the whole market with a correlation matrix. The further result also shows that their structure enables a news-oriented model not only directly associated with news but also the whole market. Although they use the knowledge graph as additional input, extracting it from external resources makes it static and unreliable. It may cause model performance to fluctuate according to the quality of external resources. As compared to our method, we propose a general plugin module utilizing a graph extracted from raw data to construct a graph embedding model that can improve the accuracy of the time series backbone. Our proposed framework can be easily adapted to arbitrary deep learning models of stock movement prediction.

2.2. Lead-lag Relationships

The lead-lag relationship has been recognized as a critical stylized fact no matter on high frequency as intra-day prices or low frequency as inter-day prices. Specifically, some stocks tend to follow the other stocks on its price movement at later times. This phenomenon can be aroused for some reasons such as supply-chain, information diffusion, policy change, event driven trading, and asynchronous trading [14] [15]. It also has considerable evidence to point out that the lead lag effect should be more prevalent from the same industry, while small firms will follow big firms due to the related post announcement drift [14]. As our approach is to extract the relational graph automatically from raw time series, we use dynamic time warping as a distance function to generate graphs from individual stock and the other delayed stocks movement. Try to capture the lead-lag effect to help forecast from the whole stock market.

2.3. Graph Embedding

The graph has a natural connection with a wide diversity of real-world scenarios, user relationships in social media, transactions and users in shopping websites or entities in financial markets. Graph embedding provides an effective way to convert each element of the knowledge graph into a lower dimension, while the graph structure is preserved. Trans series models have been widely researched and applied to learning knowledge graph representation recently. This type of model represents relationships by interpreting them as a translation on the entities' low-dimensional embedding. Compared with the traditional method, the Trans series is easy to train but not liable to overfitting. Borders et al. proposed Trans E model. It makes the sum of the head vector and relation vector as close as possible with the tail vector, which treats triplet (head, relation, tail) as a transition from head to tail. That is h + r ≈ t. Wang et al. took Trans E as a reference, they proposed Trans H which models a relation as a hyperplane. Trans E and Trans H model both assume that entity and relation are vectors in the semantic space so that similar entities will close to each other. To address this issue, Lin et al. let entities and relations in two distinct spaces and perform the translation in the corresponding relation space. The relation specific projection can make the head/tail entities that actually hold the relation close with each other, and also get far away from those that do not hold the relation.

3. PORPOSED MODELLING

In this section, we provide the problem definition of stock trend prediction. Then the following part will be the detail of our proposed RGStockNet module.

3.1 Problem Definition

We use bold-face capital letters (e.g., X) for matrices and bold-face lower-case letters (e.g., x) for vectors. Moreover, normal lower-case letter (e.g., x) represent scalars and Greek letters (e.g., Θ) are used to represent hyper-parameters. We formulate stock trend prediction as a time series classification problem, learning a mapping function $y\hat{}s\ t+1 = f(Xs\ t\ , \Theta)$ from a series of sequential features in the lag of past T time step to the subsequent trend of next time step t $+ 1. Xs\ t = [xs\ t - T + 1, xs\ t - T + 2, \ldots, xs\ t\ ]$ ∈ RT ×C is the input window of stock s in time t as matrix form, which contains the sequential features from time t−T +1 to t, where C is the dimension of input features. Θ is the parameters of mapping function f and yˆs t+1 is the predicting trend at next time step t+1. Trend calculation is defined by the difference percentage of closing price at time t and closing price at time t + 1. And ys t+1 is the ground truth of future trend at time t+1. By accessing a long history of each stock, we can construct many training examples for mapping function by moving the lag along with the whole history. We then can formulate our problem as follows: Input: A set of training sequences: $\{(Xs\ t\ ), ys\ t+1\}$. Output: A prediction mapping $function\ f(Xs\ t\ , \Theta)$ that predicts future trend movement in the following time step.

3.2 Proposed Framework

Figure 1 shows the proposed RGStockNet framework that consists of two stages. In stage 1, we extract relation triplets from raw time series data of each pair of stocks as knowledge graphs, and then train a graph embedding model as a pretrained knowledge embedding (KE) model for stage 2. In stage 2, we extract a time period with lag time T from an arbitrary company. Company id will pass through the pretrained model in stage 1, generating company embedding vector, while time period data will pass through the time series encoder to generate context vector. These two vectors will be concatenated to make trend predictions through the last fully connected layer. 1) Relational Graph Generation: The relational graph generation block extract relation triplets from raw time series of each pair of stocks. By applying inter-times-series computing method to normalized time series of each pair of stocks, adjacency matrix of whole stock market is generated. Then triplets can be generated by the decision threshold of graph edge τ .

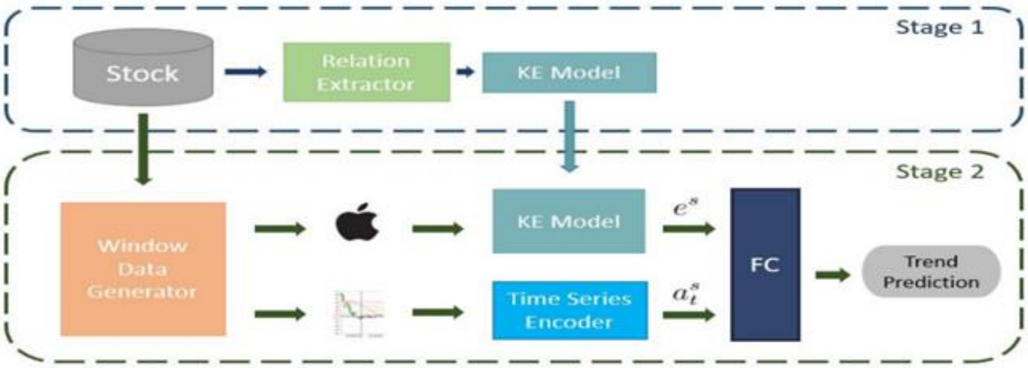

Fig:1-Framework of proposed RGStockNet module

3.3 Time Series Normalization.

Since our goal is to capture relation from each pair of raw time series data, in order to make time-series comparable to each other, we first transform the time series of every stock into the difference of two consecutive time steps along with the whole history.

3.4 Baseline

Our goal in this work is to test whether the RGStockNet module can enhance the performance of the baseline model or not. We will compare the baseline models with the RGStockNet module using the following baselines.

• **LSTM** is a long-short term memory model that belongs to the case of recurrent neural network . It contains one input layer, one LSTM layer with hidden size 128, and a fully-connected decision layer .

• **ALSTM** is the Attentive LSTM which is an encoder decoder structure. By assigning attention weighting of each time step in the encoder, the decoder decides part of the source data sequence to pay attention to.

• **Res Net** uses residual block and convolutional layer to construct classification model. It is considered a strong baseline for time series classification.

4. RESULTS AND DISCUSSIONS

4.1 Performance Comparison

Our research aims to propose a plugin RGStockNet module that is flexible enough to adapt to an arbitrary deep model, enhancing performance on the vanilla baseline. As we can see in Table I, the vanilla baseline with the RGStocknet module mostly outperforms the vanilla baseline on profit-score and accuracy, no matter the pre-trained model weight is freeze or non-freeze. Compared with the pre-trained knowledge embedding model under freeze and non-freeze settings, the RGStocknet module with non-freeze training may have higher accuracy over freeze settings and otherwise lower profit-score. There exists a trade-off between accuracy and profit-score due to the subtle imbalance distribution of our dataset. Further fine-tuned knowledge embedding model in stage 2 will make the prediction model tend to flat class, which is the majority class in our dataset. Since profit-score is a more significant metric than the accuracy of the stock movement prediction problem, RGStocknet with an unfreeze setting will be the better choice as future investment tools.

4.2 Trading Simulation

Although the classification result of RGStockNet can outperform the vanilla baseline, whether our framework can be applied in a real-life investment environment is another concern. Therefore, we make a trading simulation as back testing for the entire testing dataset. The trading strategy is followed below. If the model predicts Rise when there is no stock in hand, the buying signal will be triggered at the closing price on that day. Flat class is considered as nonsense. Once the turning point occurred,

Table 1: Trading performance of entire dataset.

Model	Return [%]	SP	PS
B&H	4.42	0.156	
LSTM	4.17	0.241	32.90%
RG-LSTM	**5.80**	**0.352**	32.53%
ALSTM	4.81	0.324	32.25%
RG-ALSTM	**4.91**	**0.335**	33.05%
Res Net	4.41	0.283	33.84%
RG-Res Net	**7.38**	**0.394**	36.22%

which is Drop class that means the closing price of the next day may fall below the close price today, we will leave current trade as profit-taking. It is noted that most well-known online brokers are commission-free for stock trading, such as TD Ameritrade, First Trade, etc. The commission will be ignored in the following trading simulation phase. Moreover, freeze setting will be used to our prediction framework

Table 2: Trading Performance of Individual Companies.

	AIV			BRK.B			COST			HPQ			VZ		
Model	RET [%]	PP [%]	HBP [%]	RET [%]	PP [%]	HBP [%]	RET [%]	PP [%]	HBP [%]	RET [%]	PP [%]	HBP [%]	RET [%]	PP [%]	HBP [%]
BH	2.72	-	-	18.38	-	-	3.26	-	-	30.43	-	-	-8.66	-	-
LSTM	-2.55	33.3	1.70	0.29	50.0	0.82	5.27	40.0	7.50	**30.43**	**100**	**100**	**-0.63**	60.00	57.8
RG-LSTM	**5.77**	**73.3**	**40.6**	**11.51**	**100**	**27.87**	**12.96**	**66.6**	**42.5**	30.21	66.6	59.20	-2.29	**66.67**	**96.33**
ALSTM	-2.41	33.3	2.54	0.45	66.6	1.64	5.25	100	7.50	20.43	**100**	**100**	**-2.2**3	**66.67**	**96.33**
RG-ALSTM	**4.54**	**64.7**	**36.44**	**5.90**	**50.0**	**18.03**	**4.47**	**50.0**	**50.83**	**23.19**	64.2	42.40	-4.45	0.00	92.66
Res Net	-7.45	57.6	38.14	3.31	44.2	34.43	15.9	56.8	35.00	4.49	56.2	23.20	5.09	**53.66**	40.37
RG-Res Net	**5.08**	**62.7**	**48.31**	**8.59**	**51.8**	**51.64**	**32.96**	**59.3**	**50.83**	**27.37**	**66.5**	**54.40**	**6.61**	47.47	**43.12**

5. EXPERIMENTS RESULTS

5.1 Performance Comparison

Our research aims to propose a plugin RGStockNet module that is flexible enough to adapt to an arbitrary deep model, enhancing performance on the vanilla baseline. As we can see in Table I, the vanilla baseline with the RGStocknet module mostly outperforms the vanilla baseline on profit-score and accuracy, no matter the pre-trained model weight is freeze or non-freeze. Compared with the pre-trained knowledge embedding model under freeze and non-freeze settings, the RGStocknet module with non-freeze training may have higher accuracy over freeze settings and otherwise lower profit-score. There exists a trade-off between accuracy and profit-score due to the subtle imbalance distribution of our dataset. Further fine-tuned knowledge embedding model in stage 2 will make the prediction model tend to flat class, which is the majority class in our dataset. Since profit-score is a more significant metric than the accuracy of the stock movement prediction problem, RGStocknet with an unfreeze setting will be the better choice as future investment tools.

5.2 Trading Simulation

Although the classification result of RGStockNet can out-perform the vanilla baseline, whether our framework can be applied in a real-life investment environment is another concern. Therefore, we make a trading simulation as back-testing for the entire testing dataset. The trading strategy is followed below. If the model predicts Rise when there is no stock in hand, the buying signal will

International Journal of Scientific Research in Engineering and Management (IJSREM)
Volume: 06 Issue: 03 | March - 2022 ISSN: 2582-3930

be triggered at the closing price on that day. Flat class is considered as nonsense. Once the turning point occurred, which is Drop class that means the closing price of the next day may fall below the close price today, we will leave current trade as profit-taking. It is noted that most well-known online brokers are commission-free for stock trading, such as TD Ameritrade, First Trade, etc. The commission will be ignored in the following trading simulation phase. Moreover, freeze setting will be used to our prediction framework.

1) Entire Dataset: However, to compare the RG-based model with the vanilla baseline, we also introduce buy and hold (B&H) strategy as the baseline, which buys at the first time point and sells out at the last trading day of the testing period. The simulation result is shown in Table II. All vanilla baseline with RGStocknet module beat B&H and outperform their vanilla baseline for cumulative return by 39.1%, 2.1%, 67.3%, and Sharpe ratio (SP) by 46.1%, 3.4%, 39.2% respect-tively. Noted that profit-score for 3-class stock classification has a positive correlation to the final return percentage and Sharpe ratio. This indicates that Rise and Drop class are more critical than Flat class prediction in the real-world investment scenario.

2) Case Study: In this section, we will inspect on simulation performance of individual stocks, including Apartment Investment and Management Co. (AIV), Berkshire Hathaway Inc. Class B (BRK.B), Costco Wholesale Corporation (COST), HP Inc. (HPQ), and Verizon Communications Inc. (VZ). Table III shows the trading performance of the companies mentioned before. Figure 2 shows the commutative return percentage plot thorough the testing period of Costco, while closing difference on two consecutive days and the prediction result is presented. Prediction results of each day are denoted as red, yellow, and green color, which are Rise, Flat, and Drop, respectively. Our model can mostly achieve better performance than the baseline on all companies. We also observe the following phenomena: i) Recurrent neural networks are eager to make a sequence of the same prediction in a row, while the RGStock-net module can ease this situation and make a better profit than the vanilla baseline. ii) The RGStocknet module is also possible to be useless. In addition to typical cases, we also consider the cyclic and non-cyclic company in our case study. The terms cyclical and non-cyclical refer to how closely correlated a company's price is to the fluctuations of the economy. We extract company with maximum and minimum standard deviation on consecutive closing difference during the testing period, HPQ and BRK.B are extracted concerning non-stable and stable stock.

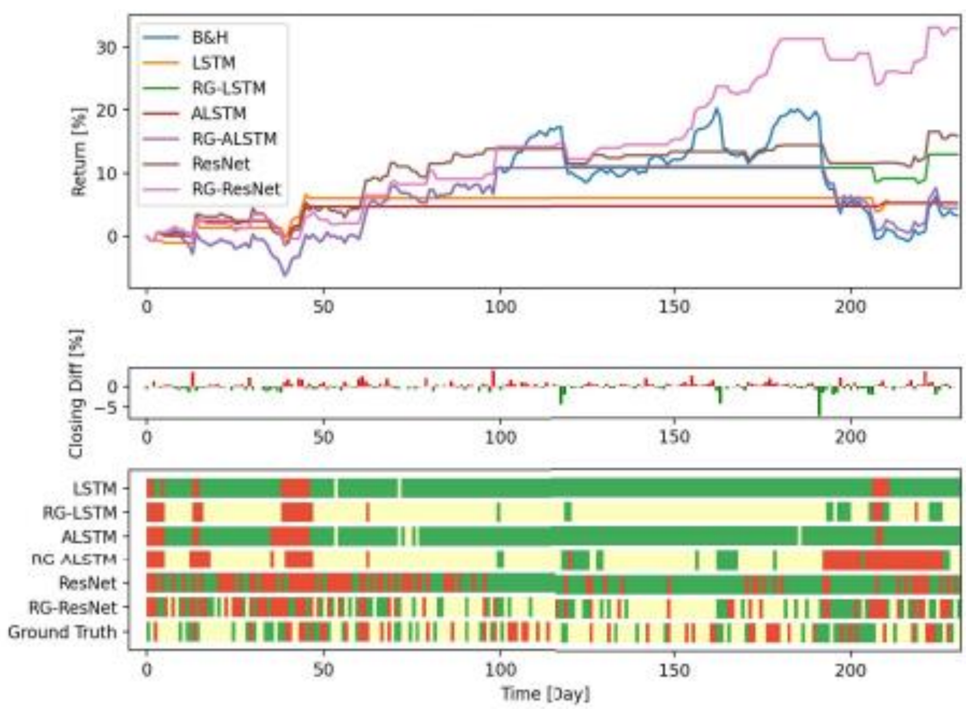

Fig 2. Trading Simulation of Cost

Companies. HPQ as non-stable stock, RG-LSTM, and RGALSTM are struggling to make correct predictions but still try to make transactions as LSTM and ALSTM act like buy and hold strategy, which holds stock in hand until the end. However, RG-ResNet can provide better prediction results and outperform vanilla ResNet on the final return. Our model can make a good performance in a stable company like BRK.B, which is a well-known holding company investing in a diverse field of industries. The case can also prove whether the stock is stable or not; our module can also effectively enhance model performance on stock movement prediction.

6.CONCLUSION

The project lays the foundation for democratizing machine learning technologies for retail investors, connecting predictions made by machine learning models to retail investors through a mobile application. It helps investors navigate through the stock markets with additional analysis and help them make more informed decisions. The findings demonstrated that the application provides significance in trend prediction. When compared to the baseline, the prediction shows useful trend tendency with the real stock trend. Through the application interface, the user can easily compare the predictions and model scores from different machine learning models, then choosing the one that fits their preference. The models used in the application will continue to improve itself by searching for a better model topology, structure and hyperparameters through evolution algorithm. The findings concluded the usefulness of the evolution algorithm in lowering the mean squared error when predicting stock prices, which is helpful for improving the trend prediction for retail investors

REFERENCES

[1] J. Patel, S. Shah, P. Thakkar, and K. Kotecha, "Predicting stock and stock price index movement using Trend Deterministic Data Preparation and machine learning techniques," Expert Systems with Applications: An International Journal, Vol. 42, Jan. 2015, pp. 259-268

[2] L. Ramchand and R. Susmel, "Volatility and cross correlation across major stock markets," Journal of Empirical Finance, vol. 5, no. 4, pp. 397–416, 1998 D. Mandic and J. Chambers, Recurrent Neural Networks for Prediction Wiley, 2001

[3] R. Williams and D. Zipser, "Gradient-based learning algorithms for recurrent networks and their computational complexity", in Backpropagation: Theory, Architectures and Applications, Hillsdale, NJ: Erlbaum, 1992, pp. 433 – 486

[4] Finance, vol. 33, pp. 178–203, 2015. [4] K. Chan, "A further analysis of the lead–lag relationship between the cash market and stock index futures market," The Review of Financial Studies, vol. 5, no. 1, pp. 123–152, 1992.

[5] L. Zhang, C. Aggarwal, and G.-J. Qi, "Stock price prediction via discovering multi-frequency trading patterns," in Proceedings of the 23rd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, 2017, pp. 2141–2149.

[6] F. Feng, H. Chen, X. He, J. Ding, M. Sun, and T.-S. Chua, "Enhancing stock movement prediction with adversarial training," arXiv preprint arXiv:1810.09936, 2018.pandas: powerful Python data analysis toolkit, Pandas. Available:

[7] 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 Proceedings of the Eleventh ACM International Conference on Web Search and Data Mining, 2018, pp. 261–269.

[8] Y. Xu and S. B. Cohen, "Stock movement prediction from tweets and historical prices," in Proceedings of the 56th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers), 2018, pp. 1970–1979.

[9] D. Matsunaga, T. Suzumura, and T. Takahashi, "Exploring graph neural networks for stock market predictions with rolling window analysis," arXiv preprint arXiv:1909.10660, 2019.

[10] J. Liu, Z. Lu, and W. Du, "Combining enterprise knowledge graph and news sentiment analysis for stock price prediction," in Proceedings of the 52nd Hawaii International Conference on System Sciences, 2019.

[11] S. Deng, N. Zhang, W. Zhang, J. Chen, J. Z. Pan, and H. Chen, "Knowledge-driven stock trend prediction and explanation via temporal convolutional network," in Companion Proceedings of the 2019 World Wide Web Conference, 2019, pp. 678–685.

[12] T. Lin, T. Guo, and K. Aberer, "Hybrid neural networks for learning the trend in time series," in Proceedings of the Twenty-Sixth International Joint Conference on Artificial Intelligence, no. CONF, 2017, pp. 2273– 2279.

[13] W. Li, R. Bao, K. Harimoto, D. Chen, J. Xu, and Q. Su, "Modeling the stock relation with graph network for overnight stock movement prediction," no. CONF, pp. 4541–4547, 2020.

[14] K. Hou, "Industry information diffusion and the lead-lag effect in stock returns," The Review of Financial Studies, vol. 20, no. 4, pp. 1113–1138, 2007.

[15] H. Shahrur, Y. L. Becker, and D. Rosenfeld, "Return predictability along the supply chain: the international evidence," Financial Analysts Journal, vol. 66, no. 3, pp. 60–77, 2010.