

## Research Questionnaire

**Note:**

- Fill up **Table 2** for each paper. (COPY TABLE 2 AND PASTE AT THE END OF THIS FILE FOR NEXT PAPER)
- **Green** – Write few or more lines of required in your own words.
- **Red** – write down the list of what is required and description about each in the list

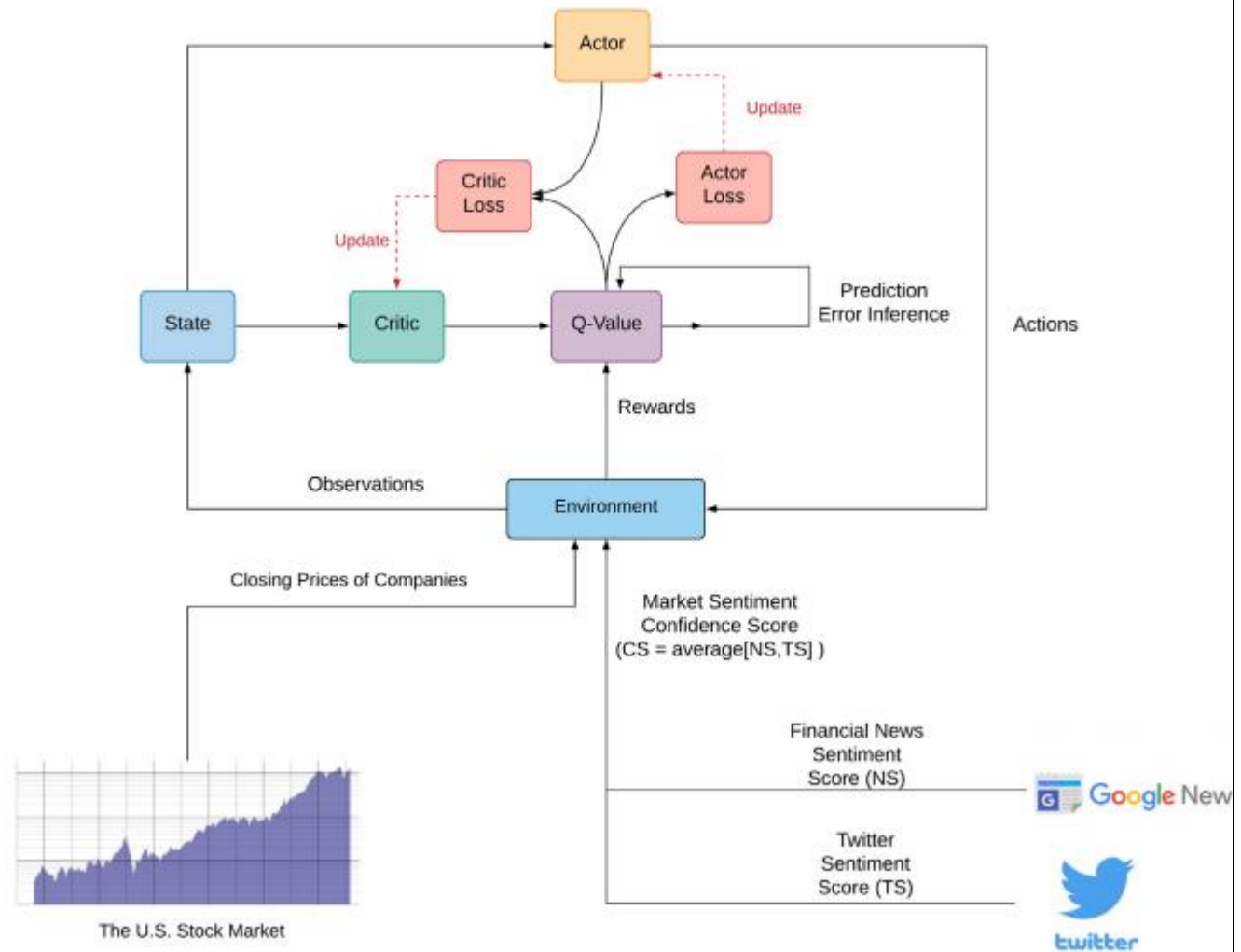
TABLE 1	
<b>Reg. No. &amp; Name</b>	19BCE0959 Devang Gupta
<b>Team No.</b>	7
<b>Paper Title</b>	Market sentiment-aware deep reinforcement learning approach for stock portfolio allocation  Stock Prediction by using NLP and Deep Learning Approach  An Efficient Word Embedding and Deep Learning Based Model to Forecast the Direction of Stock Exchange Market Using Twitter and Financial News Sites: A Case of Istanbul Stock Exchange (BIST 100)
<b>Citation (APA style)</b>	Koratomaddi, P., Wadhwani, K., Gupta, M., & Sanjeevi, S. G. (2021). Market sentiment-aware deep reinforcement learning approach for stock portfolio allocation. <i>Engineering Science and Technology, an International Journal</i> , 24(4), 848-859.  Deshmukh, R. (2021). Stock Prediction by using NLP and Deep Learning Approach. <i>Turkish Journal of Computer and Mathematics Education (TURCOMAT)</i> , 12(1S), 202-211.  Kilimci, Z. H., & Duvar, R. (2020). An Efficient Word Embedding and Deep Learning Based Model to Forecast the Direction of Stock Exchange Market Using Twitter and Financial News Sites: A Case of Istanbul Stock Exchange (BIST 100). <i>IEEE Access</i> , 8, 188186-188198.

**TABLE 2**

<b>Problem answered in this paper.</b> (1-2 lines)	To solve stock portfolio allocation and maintain it to get maximum return with minimum risk involved.
<b>Detailed description about the problem</b> (5-8 lines)	One of the most difficult and fascinating challenges in modern finance is portfolio allocation. This is due to the stock market's complexity, which is characterized by a web of interconnected return effects that necessitates a significant amount of computational work to decipher and predict return regularities. Furthermore, because the stock market is continuously changing, anticipating stock price fluctuations is not an easy process.
<b>Why that problem is chosen in this paper? Scope of the problem and solution</b> ( <i>Refer Introduction</i> ) (5-8 lines)	People are confused and find themselves in difficult situation when it comes to maintain their stock portfolio. They have to maintain the stocks in their portfolio such that then not only get maximum returns but also with minimum risk involved. To help people from this situation this paper proposes an algorithm that will analyze the news and tweets to help find a stock that can give maximum return with minimum risk involved.
<b>History of the problem.</b> ( <i>Refer Introduction</i> ) (8-10 lines)	Some of the earlier efforts on portfolio allocation were based on mathematical models that used quadratic programming, stochastic calculus, numerical analysis, and other techniques. Statistical learning approaches were initially utilized for simple numerical analysis problems – for example, the Newton-Raphson algorithm was used to solve logistic regression. However, when supervised machine learning methods such as artificial neural networks became more prominent in the 1990s, multiple deep learning algorithms for diverse stock market applications were proposed. The ability of neural networks to learn complex nonlinear functions is why they outperform other machine learning approaches in predicting market returns.
<b>List of the related/similar problems</b> ( <i>Refer Related work</i> ) – Describe each with proposed solutions	
<b>Related problem 1 –</b> Describe (3-4 lines)	To solve the financial problem of portfolio selection.
<b>Paper in APA style</b>	Koyano, S., & Ikeda, K. (2017, November). Online portfolio selection based on the posts of winners and losers in stock microblogs. In <i>2017 IEEE Symposium Series on Computational Intelligence (SSCI)</i> (pp. 1-4). IEEE.
<b>Related problem 2 –</b> Describe (3-4 lines)	To predict one-month-ahead stock returns in the cross-section in the Japanese stock market and investigates the performance of the method
<b>Paper in APA style</b>	Abe, M., & Nakayama, H. (2018, June). Deep learning for forecasting stock returns in the cross-section. In <i>Pacific-Asia conference on knowledge discovery and data mining</i> (pp. 273-284). Springer, Cham.

<b>Related problem 3 –</b> Describe (3-4 lines)	To find an appropriate modeling system that can incorporate the complexities of the stock market and generate practical trading strategies.
<b>Paper in APA style</b>	Bao, W., & Liu, X. Y. (2019). Multi-agent deep reinforcement learning for liquidation strategy analysis. <i>arXiv preprint arXiv:1906.11046</i> .
<b>Related problem 4 –</b> Describe (3-4 lines)	To investigate the role of market sentiment in an asset allocation problem.
<b>Paper in APA style</b>	Xing, F. Z., Cambria, E., & Welsch, R. E. (2018). Intelligent asset allocation via market sentiment views. <i>IEEE Computational Intelligence Magazine</i> , 13(4), 25-34.
<b>Related problem 5 –</b> Describe (3-4 lines)	Dynamic portfolio optimization is the process of sequentially allocating wealth to a collection of assets in some consecutive trading periods, based on investors' return-risk profile. Automating this process with machine learning remains a challenging problem.
<b>Paper in APA style</b>	Yu, P., Lee, J. S., Kulyatin, I., Shi, Z., & Dasgupta, S. (2019). Model-based deep reinforcement learning for dynamic portfolio optimization. <i>arXiv preprint arXiv:1901.08740</i> .
<b>What is the proposed solution in this paper for the problem chosen?</b> (Refer Proposed work) (5-8 lines)	In this they proposed an adaptive, sentiment-aware deep deterministic policy gradients approach to solve portfolio allocation that not only learns from historical stock price trends, but also from market sentiment – which is an influential environment input that captures the overall mood of investors. They've consolidated an extensive dataset of Google News and Twitter tweets that reflect the sentiment of the 30 Dow Jones companies. It also provided the methodology and mathematical definitions used to calculate market sentiment, and enable the adaptive DDPG algorithm to leverage it sufficiently.

**Architecture of the proposed solution.**  
*(Refer proposed work)*  
**Diagram**



**Name of the approach as stated by the authors (if not, you try to give a name based on the concepts used)**

Sentiment-aware deep reinforcement learning approach

<p><b>List of existing algorithms used by the authors to complete the proposed work.</b> (1-2 lines for each algorithm)</p>	<p>(DDPG) Deep Deterministic Policy Gradients algorithm Deep Deterministic Policy Gradients (DDPG) is a deep reinforcement learning algorithm that concurrently learns a Q-function (learnt by the critic network) and a policy (learnt by the actor network). The algorithm is a model-free, off-policy actor-critic algorithm using deep function approximators that can learn policies in high-dimensional, continuous action spaces Adaptive Deep Deterministic Policy Gradients (Adaptive DDPG) As an improvement to the DDPG algorithm that allowed for learning the bullish and bearish aspects of the market actively</p>
<p><b>List of datasets used.</b> (Refer experimental evaluation/result discussion) (3-4 lines)</p>	<p>Google News Twitter</p>
<p><b>References/links to each of the dataset used in this paper (in APA style)</b></p>	<p>1. <a href="https://www.kaggle.com/getting-started/8223">https://www.kaggle.com/getting-started/8223</a> 2. <a href="https://www.kaggle.com/kazanova/sentiment140">https://www.kaggle.com/kazanova/sentiment140</a> ...</p>
<p><b>Why the above dataset(s) used?</b> (Refer experimental evaluation/result discussion) (3-4 lines)</p>	<p>Twitter and Google News are used as the data set in this paper as, Twitter tweets was found useful to predict the stock prices because twitter is a platform where they share their opinions and in a smaller number of words which can be easily analyzed. Google news is a reliable news platform and using its news and twitter tweets the predicted price of the stock was expected to be more precise.</p>
<p><b>List of equations that are very well applied in this problem domain</b></p>	<p>Equation 1:</p> $NS_{(c,d)} = \frac{\sum_{i=1}^N PS(a_{(i,c,d)})}{N}$ <p>Description: Sentiment analysis was performed to obtain a google news sentiment score.</p>

	<p>Equation 2:</p> $TS_{(c,d)} = \frac{\sum_{i=1}^N w(t_{(i,c,d)})PS(t_{(i,c,d)})}{N}$ <p>Description: Sentiment analysis was performed to obtain a twitter sentiment score</p> <p>Equation 3:</p> $CS_{(c,d)} = \frac{NS_{(c,d)} + TS_{(c,d)}}{2}$ <p>Description: The confidence score is computed using the financial news sentiment score (NS) and twitter sentiment score (TS)</p>
<p><b>List of method(s)/metrics used to evaluate the proposed approach.</b> (Refer experimental evaluation/result discussion) (5-8 lines)</p>	<p>Sharpe Ratio Annualized Return Annualized Std. Error Final Portfolio Value</p>
<p><b>List of supporting tools/concepts</b> (3-4 lines)</p>	<p>VADER (Valence Aware Dictionary and sentiment Reasoner). VADER is a lexicon and rule-based sentiment analysis tool implemented in Natural Language Toolkit.</p>
<p><b>What are the similar approaches with which the proposed approach is compared?</b> (Refer experimental evaluation/result discussion)</p>	<p>Approach/method 1: Supervised machine learning tools such as artificial neural networks. Neural networks approach was more successful as compared to other machine learning approaches in predicting stock returns is attributed to their ability to learn complex nonlinear functions.</p> <p>Approach/method 2: Deep learning approaches, LSTMs are very powerful in sequence prediction problems because they're able to store past information. This is important in our case because the previous price of a stock is crucial in predicting its future price.</p>

<b>Explain each of these approach (3-4 lines)</b>	
<b>How the results of proposed approach are compared with other similar approaches?</b> <i>(Refer experimental evaluation/result discussion)</i>	It is noteworthy that the sentiment aware approach has significant improvements across all the considered metrics compared to the baselines. With an initial investment of 10,000 dollars, the final portfolio value of our approach reaches 25,051 dollars which is much higher than 21,881 dollars by Adaptive DDPG and 18,156 dollars by DDPG. This approach achieves an annualized rate of return of 22.05% as compared to 18.85% and 14.7% returns by Adaptive and DDPG respectively. The risk taken by this model is also lesser compared to the other baselines as shown by obtaining the least annualized standard deviation error of 0.096. With a 2.07 Sharpe ratio value, our sentiment-aware approach is shown to be more robust and effective in balancing return and risk compared to 1.49, 0.93 for the adaptive DDPG and DDPG respectively.
<b>Advantages/merits of proposed solution in your view.</b> <i>(Refer conclusion / result discussion / experimental evaluation)</i>	<b>Deep Deterministic Policy Gradient (DDPG)</b> is a model-free off-policy algorithm for learning continuous actions. It combines ideas from DPG (Deterministic Policy Gradient) and DQN (Deep Q-Network). It uses Experience Replay and slow-learning target networks from DQN, and it is based on DPG, which can operate over continuous action spaces.
<b>Disadvantages/limitations of proposed solution in your view.</b> <i>(Refer conclusion / result discussion / experimental evaluation)</i>	In environments with continuous state and action spaces, Deep Deterministic Policy Gradient (DDPG) algorithms can solve very complex problems, yet can also fail in environments that seem trivial, but the reason for such failures is still poorly understood.
<b>Future work as stated by authors</b> <i>(Refer conclusion / result discussion / experimental evaluation)</i>	Future work could focus on acquiring more tweets per day, expanding acquisition to get insights from different sources such as stock market-specific news websites (CNBC, Business Standard, etc.), and processing photos, as most tweets and news online are now provided as image snippets. For this case, multi-agent reinforcement learning algorithms can be investigated. We're also excited to address the presence of several exogenous restraints on retail and institutional traders, such as transaction fees, trading restrictions, cash holding restrictions, and liquidity shortages. Furthermore, due to the extensive use of metaphors, sarcasm, domain specific terminology, and other indirect linguistic references in common language, especially in material that expresses an opinion, natural language processing on financial data is a non-trivial effort. Being able to grasp such language might aid in more accurately predicting market sentiment.

**Your one page write-up about this paper**

The stock market is still one of the most challenging financial systems to model. As a result, stock portfolio allocation is a difficult problem to solve, as an optimal investing strategy for a curated group of companies must be developed that efficiently maximizes return while minimizing risk. When used to automate portfolio allocation by training an intelligent agent on historical stock prices, deep reinforcement learning algorithms have demonstrated promising outcomes. Modern investors, on the other hand, are using digital platforms like social media and online news websites to better understand and analyses their portfolios. Market sentiment refers to the overall attitude that investors have about a certain stock or financial market. Existing techniques ignore market mood, which has been proved to influence investor decisions empirically. They suggested a revolutionary deep reinforcement learning strategy for efficiently training an intelligent automated trader that not only uses historical stock price data but also detects market sentiment for a Dow Jones stock portfolio. They showed that their method is more robust than existing baselines in terms of standardized metrics like the Sharpe ratio and annualized investment return.

**Your findings: (possible alternate for the solution proposed)**

- ARIMA is a very popular statistical method for time series forecasting. ARIMA models take into account the past values to predict the future values.

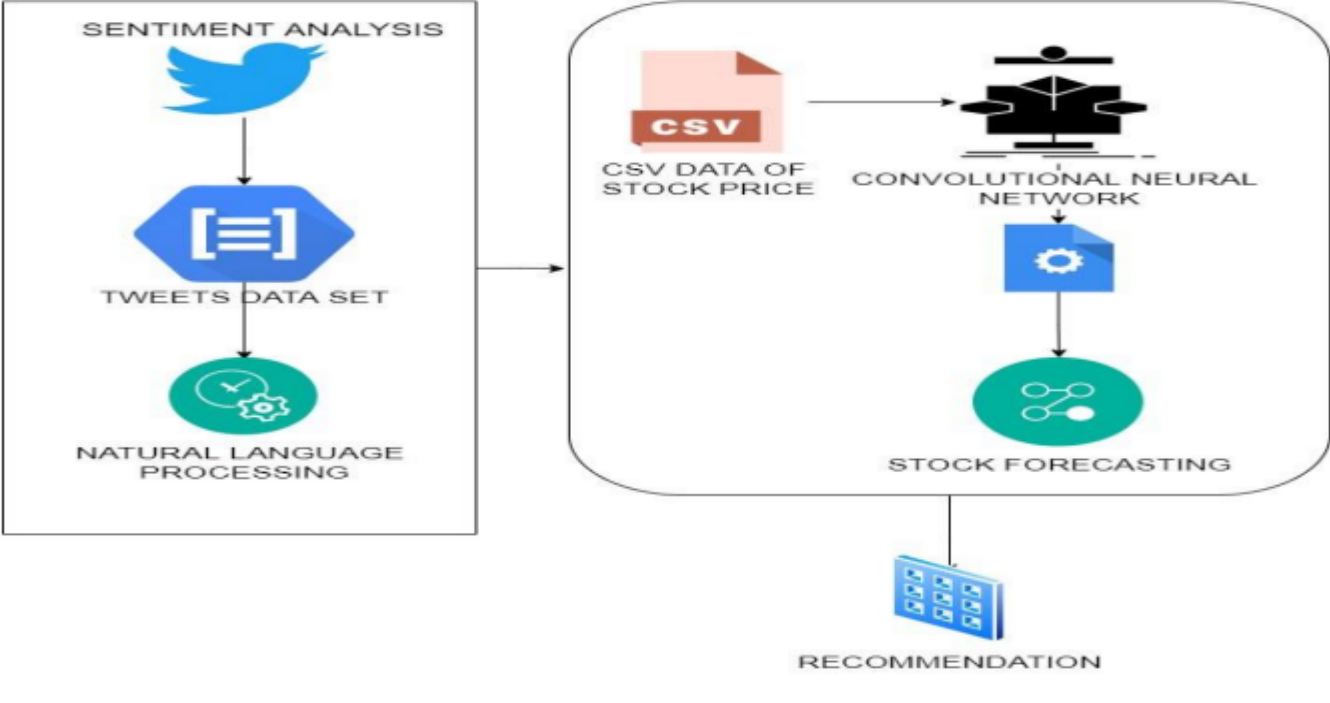
**TABLE 2**

<b>Problem answered in this paper.</b> (1-2 lines)	To properly recognize which shares to promote with a purpose to get more profits.
<b>Detailed description about the problem</b> (5-8 lines)	People have a tendency to analyze existing strategies and so planned new strategies for inventory prediction. They predict the share price based on random assumption. People are not able to keep track of the stock prices whether it is going up or down. Financial analysts investing in stocks usually, but they are not aware about the inventory market place conduct. They usually go through the problem of trading as they do not properly recognize which shares to shop or which shares to promote with a purpose to get greater profits.
<b>Why that problem is chosen in this paper? Scope of the problem and solution</b> (Refer Introduction)	In today's world, all the information pertaining to inventory market is available. Analyzing all these records in my opinion or manually is pretty much difficult. As such, automation of the method is required. This is where Data mining techniques help. Understanding that analysis of numerical time series offers close results, wise traders use system learning techniques in predicting the inventory market conduct. This will allow financial



(5-8 lines)	analysts to foresee the conduct of the inventory that they may be interested by and consequently act accordingly.
<b>History of the problem.</b> (Refer Introduction) (8-10 lines)	They had used one system to find out the opening value of stocks in the financial market. However, their developed system was self-learner so that they were able to predict the opening value of the market. They had given the stocks data to their developed system to find out the forecasted value. Last, they developed another network system and compared both the system with each other to predict the starting day value of the stock. Used ANN algorithm for the share market prediction. On the basis of their studies of ANN they had claimed that they had achieved the high percentage of accuracy while predicting the values in the stock market.
<b>List of the related/similar problems</b> (Refer Related work) – Describe each with proposed solutions	
<b>Related problem 1 –</b> Describe (3-4 lines)	They employed a sentiment analysis technique for stock tweets that were related to a different type of Apple product; for this, she had extracted stock-related tweets from various social networking sources for an eight-year period. These individuals had elected to use stocks-related data from the Yahoo Finance source for that time period in addition to share data. To determine the polarity of those tweets, they applied the SVM approach.
<b>Paper in APA style</b>	Batra, R., & Daudpota, S. M. (2018, March). Integrating StockTwits with sentiment analysis for better prediction of stock price movement. In <i>2018 International Conference on Computing, Mathematics and Engineering Technologies (iCoMET)</i> (pp. 1-5). IEEE.
<b>Related problem 2 –</b> Describe (3-4 lines)	The focus of this article was on the market's share price swings. They employed additional important information in addition to mining approaches for stock predictions. Their findings revealed that they calculated the polarity of the stocks in order to better anticipate the stock's price.
<b>Paper in APA style</b>	Wang, Y., & Wang, Y. (2016, March). Using social media mining technology to assist in price prediction of stock market. In <i>2016 IEEE International conference on big data analysis (ICBDA)</i> (pp. 1-4). IEEE.
<b>Related problem 3 –</b> Describe (3-4 lines)	They had utilized a single system to determine the stock market's opening value. However, their built system was self-learner, allowing them to estimate the market's opening value. They had fed the stock data into their custom-built algorithm, which calculated the anticipated value.
<b>Paper in APA style</b>	Zhang, Z., Shen, Y., Zhang, G., Song, Y., & Zhu, Y. (2017, November). Short-term prediction for opening price of stock market based on self-adapting variant PSO-Elman neural network. In <i>2017 8th IEEE International Conference on Software Engineering and Service Science (ICSESS)</i> (pp. 225-228). IEEE.
<b>Related problem 4 –</b> Describe (3-4 lines)	They employed the ANN algorithm to predict the stock market. They stated that their research of ANN had resulted in a high percentage of accuracy in predicting stock market prices. They researched several approaches for this, and after doing so, they discovered accurate and proper findings.
<b>Paper in APA style</b>	Firdaus, M., Pratiwi, S. E., Kowanda, D., & Kowanda, A. (2018, October). Literature review on artificial neural networks techniques application for stock market prediction and as decision support tools. In <i>2018 Third International Conference on Informatics and Computing (ICIC)</i> (pp. 1-4). IEEE.

<b>Related problem 5 –</b> Describe (3-4 lines)	They were primarily interested in predicting share values using historical data. They used decade data from the NSE and BSE, two well-known stock exchanges. They used the SVM method to create a model for this.
<b>Paper in APA style</b>	Sharma, N., & Juneja, A. (2017, April). Combining of random forest estimates using LSboost for stock market index prediction. In <i>2017 2nd International conference for convergence in technology (I2CT)</i> (pp. 1199-1202). IEEE.
<b>What is the proposed solution in this paper for the problem chosen?</b> (Refer Proposed work) (5-8 lines)	In this paper they tend to area unit planned System that is working with Improved level of recommendation. System is developed with Natural Language Processing (NLP) technique of computer science and Convolutional Neural Network (CNN) of Deep Learning. Natural Language Processing technology is used facilitate system to search out companies with excellent news in terms of live performance in market. That helped to facilitate to create selection of best performer in market. NLP is used to classify news in positive and negative sets and to provide performance graph of selected organization. Supported to that we got to know the best performing company. Natural Language Processing provides to system NLP (Natural Language Processing) that worked on our twits for detection merchandise and unhealthy of its impact.
<b>Architecture of the proposed solution. (Refer proposed work) Diagram</b>	

	
<p><b>Name of the approach as stated by the authors (if not, you try to give a name based on the concepts used)</b></p>	<p>Prediction of stock prices using CNN and VADAR Sentiment analysis algorithm.</p>
<p><b>List of existing algorithms used by the authors to complete the proposed work.</b>  <b>(1-2 lines for each algorithm)</b></p>	<p>CNN Algorithm - A Convolutional Neural Network (ConvNet/CNN) is a Deep Learning algorithm which can take in an input image, assign importance (learnable weights and biases) to various aspects/objects in the image and be able to differentiate one from the other.</p> <p>VADER Algorithm - VADER (Valence Aware Dictionary for Sentiment Reasoning) is a model used for text sentiment analysis that is sensitive to both polarity (positive/negative) and intensity (strength) of emotion. VADER sentimental analysis relies on a dictionary that maps lexical features to emotion intensities known as sentiment scores.</p>

<b>List of datasets used.</b> <i>(Refer experimental evaluation/result discussion)</i> <b>(3-4 lines)</b>	Twitter Tweets top 5 performing companies' data in IT sector
<b>References/links to each of the dataset used in this paper (in APA style)</b>	1. <a href="https://www.kaggle.com/kazanova/sentiment140">https://www.kaggle.com/kazanova/sentiment140</a> 2. ...
<b>Why the above dataset(s) used?</b> <i>(Refer experimental evaluation/result discussion)</i> <b>(3-4 lines)</b>	Twitter tweets was found more useful to predict the stock prices as twitter is a platform where they share their opinions and in a smaller number of words which can be easily analyzed.
<b>List of equations that are very well applied in this problem domain</b>	Equation 1: Description: Equation 2: Description: Equation 3: Description:
<b>List of method(s)/metrics used to evaluate the proposed approach.</b> <i>(Refer experimental evaluation/result discussion)</i> <b>(5-8 lines)</b>	The real data of the stock prices were used to evaluate the CNN model for prediction and forecasting. Comparing both the data they measured the prediction accuracy. The model first analyzed the previous data then compared the future data which was then compared with the stock price data. And using this the model achieved a polarity and based on this polarity they measured the top 10 well performing companies in the IT sector.
<b>List of supporting tools/concepts</b> <b>(3-4 lines)</b>	VADER sentiment analysis
<b>What are the similar approaches with which</b>	Approach/method 1: ANN Algorithm to predict the stock prices.

<b>the proposed approach is compared?</b> (Refer experimental evaluation/result discussion) <b>Explain each of these approach (3-4 lines)</b>	They used ANN algorithm for the share market prediction. On the basis of their studies of ANN they had claimed that they had achieved the high percentage of accuracy while predicting the values in the stock market. For this they had studied different methods and after studying that they had find out the accurate and proper results.
<b>How the results of proposed approach are compared with other similar approaches?</b> (Refer experimental evaluation/result discussion)	CNN has the characteristic of paying attention to the most obvious features in the line of sight, so it is widely used in feature engineering. With comparison to other approach i.e., ANN algorithm CNN performs much better.
<b>Advantages/merits of proposed solution in your view.</b> (Refer conclusion / result discussion / experimental evaluation)	The proposed CNN achieved relatively higher prediction accuracy of 84.6%, while the ANN, SVM, and KNN algorithms obtained prediction accuracies of 73.5%, 67.9%, and 65.9% using 11 HRV features, respectively.
<b>Disadvantages/limitations of proposed solution in your view.</b> (Refer conclusion / result discussion / experimental evaluation)	CNN do not encode the position and orientation of object. Lack of ability to be spatially invariant to the input data. Lots of training data is required. If CNN-LSTM was used it can provide a reliable stock price forecasting with the highest prediction accuracy.
<b>Future work as stated by authors</b> (Refer conclusion / result discussion / experimental evaluation)	In future they will attempt to execute more calculations and all the newer methods planning to give live proposal to securities exchange financial specialists. Additionally, their emphasis will be on entire securities exchange for forecasting.
<b>Your one-page write-up about this paper</b>	

Early stock prediction research relied entirely on random walks, machine learning, numerical prediction, and support vector machines, but with the introduction of behavioral finance, people's market literacy was taken into account when predicting stock movement. They have proposed a model that will use sentiment analysis on tweets associated with special sectors such as Information Technology, Banking, Pharmaceuticals, Automobiles, and Infrastructure that are extracted from twitter in order to exploit the benefits of sentiment analysis on enterprise associated inventory. These tweets were taken from Twitter and used to calculate polarity. To make it more efficient, they applied NLP technology to analyses the sentiment of Stock Tweets. They put the theory into action by gathering sentiment and stock price data and creating a CNN model for prediction and forecasting, as well as measuring prediction accuracy. The results showed that we had attained polarity, and we measured the top 10 performing companies in each category using this polarity.

**Your findings: (possible alternate for the solution proposed)**

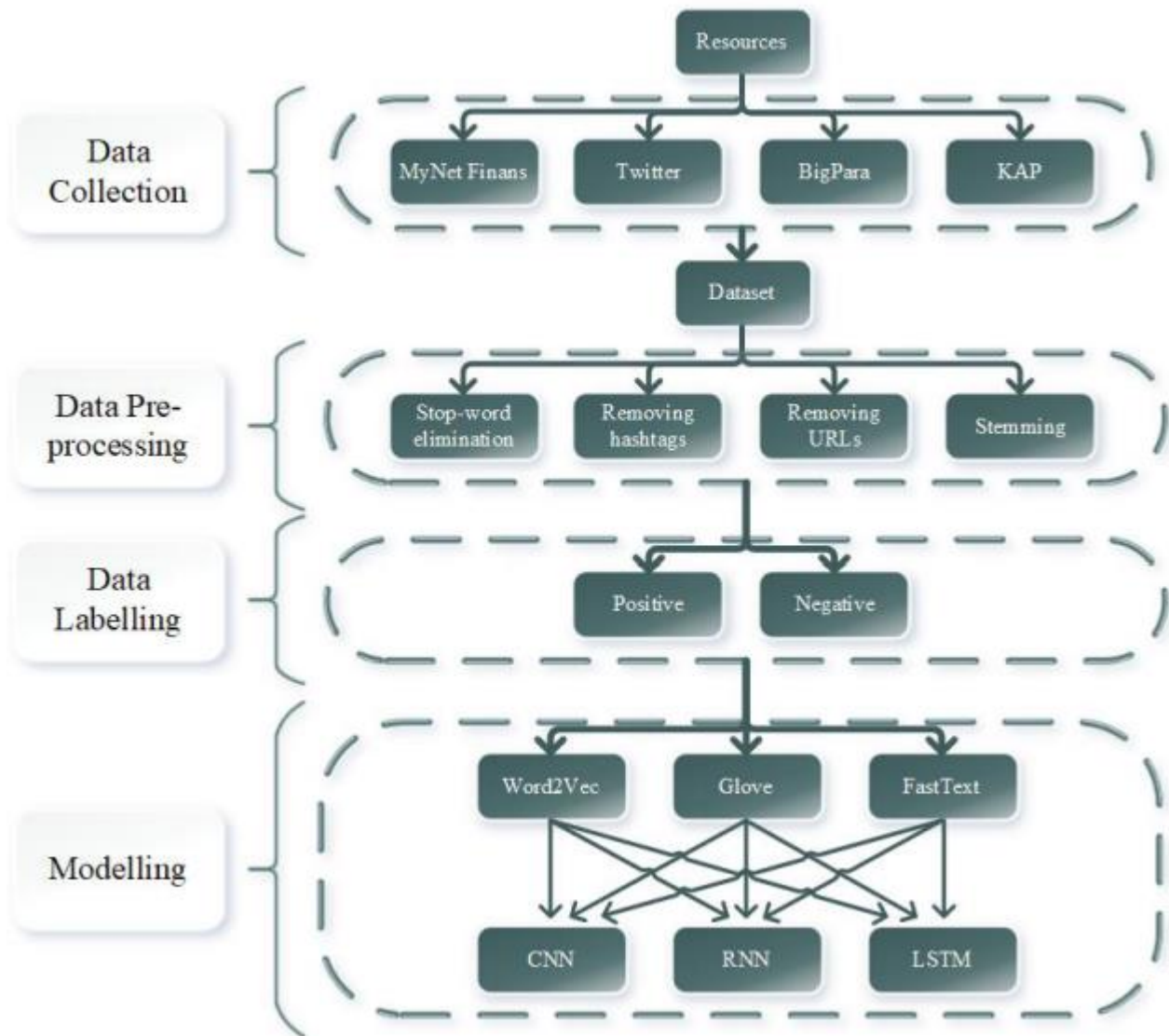
- Convolution Neural Network (CNN) and Long–Short-Term Memory Neural Network (LSTM). This new method is aptly named stock sequence array convolutional LSTM (SACLSTM).

**TABLE 2**

<b>Problem answered in this paper.</b> (1-2 lines)	Prediction of stock price in Istanbul Stock Exchange
<b>Detailed description about the problem</b> (5-8 lines)	In this study, unlike the recent researches on forecasting the stock market direction, they focus on financial sentiment analysis using the Turkish data sets collected from both a social media platform and websites including technical analysis and news to analyze the stock market direction by evaluating high volume stocks.
<b>Why that problem is chosen in this paper? Scope of the problem and solution</b> (Refer Introduction) (5-8 lines)	To forecast the movement directions of stocks, exchange rates, and stock markets are significant and an active research area for investors, analysts, and researchers. So, in this paper word embedding and deep learning-based direction prediction of Istanbul Stock Exchange (BIST 100) is proposed by analyzing nine banking stocks with high volume in BIST 100.
<b>History of the problem.</b> (Refer Introduction) (8-10 lines)	In some previous researches the combination of technical and fundamental analysts approaches applied to market trend forecasting through the use of conventional machine learning techniques applied to time series prediction and sentiment analysis on the same data but the results were not as expected.

<b>List of the related/similar problems (Refer Related work) – Describe each with proposed solutions</b>	
<b>Related problem 1 – Describe</b> (3-4 lines)	Equity investments are one of the most important asset classes. Equity investments have high return yield however also high risk due to the variability of share prices. Therefore, precise share price modeling is essential.
<b>Paper in APA style</b>	Tekin, S., & Çanakoğlu, E. (2019, April). Analysis of price models in istanbul stock exchange. In <i>2019 27th Signal Processing and Communications Applications Conference (SIU)</i> (pp. 1-4). IEEE.
<b>Related problem 2 – Describe</b> (3-4 lines)	Sentiment Analysis on turkish web comments using random forest classifier on Turkish Web comments.
<b>Paper in APA style</b>	PERVAN, N., & KELEŞ, H. Y. (2017). Sentiment analysis using a random forest classifier on turkish web comments. <i>Communications Faculty of Sciences University of Ankara Series A2-A3 Physical Sciences and Engineering</i> , 59(2), 69-79.
<b>Related problem 3 – Describe</b> (3-4 lines)	To estimate the direction of Borsa Istanbul 100 Index by using financial sentiment analysis and to enrich the datasets with various techniques from a semantic perspective and improve the classification performance of system by blending ensemble learning approach with deep learning algorithms.
<b>Paper in APA style</b>	Kilimci, Z. H. (2020). Financial sentiment analysis with Deep Ensemble Models (DEMs) for stock market prediction. <i>Journal of the Faculty of Engineering and Architecture of Gazi University</i> , 35(2), 635-650.
<b>Related problem 4 – Describe</b> (3-4 lines)	To forecast the direction of stocks is a significant for investors, analysts, and researchers. In this study, we propose to predict the direction of stocks in Turkish stock market (BIST100) by employing Turkish texts such as social media platforms. For this purpose, different deep learning methodologies
<b>Paper in APA style</b>	Othan, D., Kilimci, Z. H., & Uysal, M. (2019, December). Financial sentiment analysis for predicting direction of stocks using bidirectional encoder representations from transformers (BERT) and deep learning models. In <i>Proc. Int. Conf. Innov. Intell. Technol.</i> (Vol. 2019, pp. 30-35).
<b>Related problem 5 – Describe</b> (3-4 lines)	The stock market parameter forecasting is an important research subject both for financial professionals and the machine learning experts due to the challenges and opportunities it possess. Despite the difficulties in financial data, interest in this research area is growing rapidly.
<b>Paper in APA style</b>	Tekin, S., & Çanakoğlu, E. (2018, May). Prediction of stock returns in Istanbul stock exchange using machine learning methods. In <i>2018 26th Signal Processing and Communications Applications Conference (SIU)</i> (pp. 1-4). IEEE.
<b>What is the proposed solution in this paper for the problem chosen? (Refer Proposed work)</b> (5-8 lines)	In this paper long short-term memory networks, recurrent neural networks, convolutional neural networks as deep learning algorithms and Word2Vec, GloVe, and FastText as word embedding models are evaluated. To demonstrate the effectiveness of proposed model, four different sources of Turkish news are collected. The news articles about stocks from Public Disclosure Platform (KAP), text-based technical analysis of each stock from Bigpara, user comments from both Twitter and Mynet Finans platforms are gathered.

**Architecture of the proposed solution.** (Refer proposed work) **Diagram**





<b>Name of the approach as stated by the authors (if not, you try to give a name based on the concepts used)</b>	Word embedding techniques and deep learning approaches
<b>List of existing algorithms used by the authors to complete the proposed work.</b> (1-2 lines for each algorithm)	Word embedding models namely, FastText, GloVe, Word2Vec, deep learning techniques such as CNN, LSTM, RNN as well as combinations of these methods are approaches that are evaluated in this study. For the deep learning models and word embedding algorithms, Keras library is utilized.
<b>List of datasets used.</b> (Refer experimental evaluation/result discussion) (3-4 lines)	Twitter as a social media platform KAP as public disclosure platform Mynet Finans as news platform Bigpara as a technical analysis platform
<b>References/links to each of the dataset used in this paper (in APA style)</b>	1. <a href="https://www.kaggle.com/kazanova/sentiment140">https://www.kaggle.com/kazanova/sentiment140</a> 2. ...
<b>Why the above dataset(s) used?</b> (Refer experimental evaluation/result discussion) (3-4 lines)	Though English news articles have been employed for forecasting of market direction previously, to the best of our knowledge, Turkish news articles and user comments from social media and different platforms have not been utilized with the combination of deep learning techniques and word embedding methods to predict the direction of Turkish stocks and market
<b>List of equations that are very well applied in this problem domain</b>	<p>Equation 1: <math display="block">J_{\theta} = \frac{1}{T} \sum_{t=1}^T \sum_{-n \leq j \leq n} \log p(w_{t+j}   w_t)</math></p> <p>Description: Given a center word and sequence of training words <math>w_1, w_2, w_3, \dots, w_t</math> skip-gram model maximizes the average log probability of <math>n</math> surrounding words of the center word <math>w_t</math>, <math>n</math> denotes the size of training context.</p> <p>Equation 2: <math display="block">J_{\theta} = \sum_{i,j=1}^V f(X_{ij}) \left( w_i^T w_j + b_i + b_j - \log X_{ij} \right)^2</math></p>

	<p>Description: Glove method first contracts a word co-occurrence matrix X. Each element of <math>X_{ij}</math> shows the number of times word i appears in the context word j. The Glove model utilizes (3) to calculate cost.</p> <p>Equation 3: <math>s(w, c) = \sum_{g \in G_w} z_g^T v_c</math></p> <p>Description: FastText uses the skip-gram model with negative sampling proposed for Word2Vec with a modified skip-gram loss function. Let <math>G_w \supset \{1, \dots, G\}</math> be the set of n-grams appearing in a word w, the score of the word is calculated by the sum of the vector representations of its n-grams:</p>
<p><b>List of method(s)/metrics used to evaluate the proposed approach.</b> (Refer experimental evaluation/result discussion) (5-8 lines)</p>	<p>The success of proposed model is appraised with evaluation metrics namely, F-criterion, and accuracy. F-criterion is most often used when comparing statistical models that have been fitted to a data set, in order to identify the model that best fits the population from which the data were sampled.</p>
<p><b>List of supporting tools/concepts</b> (3-4 lines)</p>	<p><b>Word2Vec</b> is accepted as a pioneer word embedding method that starts a new trend in natural language processing. Word2Vec tries to express words in a vector space and it is a prediction-based and unsupervised model.</p> <p><b>Global Vectors (GloVe)</b> is another popular word embedding algorithm.</p> <p><b>FastText</b> is an artificial neural network library developed for text classification. Converts text or words into continuous vectors that can be used in any language, such as a speech-related task.</p> <p><b>Bidirectional Encoder Representations from Transformers (BERT)</b> is a machine learning framework which is designed for natural language processing. In this every output element is connected to every input element and it is designed to pre-train deep bidirectional representations from unlabeled text by jointly conditioning on both left and right context in all layers</p>
<p><b>What are the similar approaches with which the proposed approach is compared?</b> (Refer experimental evaluation/result discussion) Explain each of these approach</p>	<p>Approach/method 1: Evolutionary Neural Network - In this work, the prediction of status of stocks in BIST 100 are proposed by using ESA. The classification performance obtained by ESA is evaluated to be higher than those obtained with the chi-square feature selection and logistic regression classifier.</p> <p>Approach/method 2: LSTM, multi-layer perceptron (MLP), and random forest classifier (RFC) are employed. an approach is proposed to compute the asset-level market sentiment from social media data stream, and integrate it to the state-of-the-art asset allocation method using market views.</p>

(3-4 lines)	
<b>How the results of proposed approach are compared with other similar approaches?</b> <i>(Refer experimental evaluation/result discussion)</i>	The combination of RNN and LSTM models with word embedding models exhibit remarkable experiment results for all data sets. When data sets are evaluated, it is observed that either Word2Vec or FastText word embedding models present more successful results. Mynet Finans and Twitter data sets contain user comments which have much shorter text content compared to the Bigpara and KAP data sets.
<b>Advantages/merits of proposed solution in your view.</b> <i>(Refer conclusion / result discussion / experimental evaluation)</i>	The main Advantage is that since the model uses RNN, LSTM, Machine Learning and Deep Learning models the prediction of stock prices will be more accurate. And also, in the model it can predict the future 30 days Stock Prices and it can show it in a graph. Also, the main feature is that the model can show an output of the Individual Predicted Close prices of the Predicted 30 days.
<b>Disadvantages/limitations of proposed solution in your view.</b> <i>(Refer conclusion / result discussion / experimental evaluation)</i>	LSTM require 4 linear layer (MLP layer) per cell to run at and for each sequence time-step. Linear layers require large amounts of memory bandwidth to be computed, in fact they cannot use many computes unit often because the system has not enough memory bandwidth to feed the computational units.
<b>Future work as stated by authors</b> <i>(Refer conclusion / result discussion / experimental evaluation)</i>	In future work authors may try to increase the data set and also tune the parameters to predict more accurate value of stock prices.
<b>Your one-page write-up about this paper</b>	
For investors, analysts, and researchers, forecasting the movement directions of stocks, exchange rates, and stock markets is a key and active research topic. By evaluating nine banking stocks with large volume on the Istanbul Stock Exchange (BIST 100), this paper proposes word embedding and deep learning-based direction prediction. Although English news articles have previously been used to forecast market direction, to our knowledge, Turkish news articles and user comments from social media and other platforms have not been combined with deep learning techniques and word embedding methods to forecast the direction of Turkish stocks and markets. t. For this objective, long short-term memory networks, recurrent neural networks, convolutional neural networks as deep learning algorithms and Word2Vec, GloVe, and FastText as word embedding models are evaluated. Four distinct sources of Turkish news are gathered to demonstrate the efficiency of the suggested model. Public Disclosure Platform (KAP) news articles about stocks, Bigpara text-based technical analysis of each stock, and user comments from both Twitter and Mynet Finans platforms are aggregated.	

**Your findings: (possible alternate for the solution proposed)**

- Instead of word embedding approach text mining approach offers good prediction and great profit.