ISSN No:-2456-2165

Business Failure Prediction through Deep Learning

[1.]Ketan Bagade
Computer Department KJ College of Engineering Management Research, Pune,

[2.]Dr. Nikita Kulkarni,
Computer Dept., KJ College of Engineering Management and Research, Pune

[3.]Dr. Vajid Khan,
Computer Dept., KJ College of Engineering Management and Research, Pune

Abstract:- During the course of carrying out company operations, complications might often arise as a result of turbulent business operating circumstances and unforeseen abnormalities. In most cases, a number of difficulties combine to cause a lengthy decrease in the project's perceived usefulness or collapse owing to a depletion of financial resources. Preemptive evaluation of a company's failure may help anticipate potential challenges and mitigate the negative effects of such challenges by methodically planning, preparing, and carrying out a business failure prediction. For an accurate forecast of the collapse of a company, it is important to do a prediction analysis of the activities of the firm in order to detect potential problems. Methods of machine learning or deep learning that can be used for the goal of generating an accurate forecast of the collapse of a firm may effectively be used to identify these issues, and they can be used to do so successfully. This methodology will be realized by the successful use of the method of K-nearest Neighbor Clustering as well as entropy estimation, in conjunction with Long-Short Term Memory and Decision Making.

Keywords:- *Business Failure, K Nearest Neighbors, Entropy Estimation, Long Short Term Memory, Decision Making.*

I. INTRODUCTION

As the economy of the world evolves, insolvency forecasting, the practice whose goal it is to examine a company's current financial status and potential for growth through the company's own accounting transactions, is playing a more significant role in the economic project lifecycle. Insolvency forecasting's objective is to examine a company's present financial status and potential for growth through the company's own accounting transactions. Despite the fact that it has been demonstrated that using ensemble approaches is an effective method for reducing premature business failure prediction inaccuracy, the vast majority of early business failure prediction algorithms ignore the severely unbalanced distribution of such observations in business failure databases. This is despite the fact that it has been shown that it is an effective method for reducing premature business failure prediction inaccuracy.

In this day and age, where innovation is the norm, having accurate business information to assist a firm in making choices about its future initiatives is essential. The processes and ideas that have a positive impact on the business decisions made by an organization are collectively referred to as predictive analytics. Predictive analytics is defined as the use of fact-based technology to the aid of decision-making. The processes of technology and design may turn raw, unpredictable data into coherent data that can be used to its full potential. By making use of this insightful facts, the company has a better chance of developing new goals, achieving organizational success, gaining an analytical knowledge, and making choices that are in the best interest of the organization moving forward.

Early warning of impending corporate collapse is a crucial component of financial risk avoidance management. It is possible for the management of a company to get early signals from an efficient proactive detection system for financial concerns, which might prevent the company from filing for bankruptcy. The implementation of financial risk warnings is vital for boosting the efficacy of investment plan and ensuring the economic viability of the organization. The number of foreclosures filed by corporations is one factor that has a substantial bearing on the economic health of a country and is also one factor that may be used to anticipate the onset of a financial crisis. As a result of the significant correlations between some of the phenomena that a lot of businesses are having financial problems and increased productivity, financial analysts are more aware of the significance of controlling and preventing the risk of bankruptcy. This is because of the significant correlations between some of the phenomena that many businesses are having financial problems.

Predictive analytics are positioned to play a large part in virtually all company kinds, both now and in the foreseeable future. This opportunity exists in both the present and the foreseeable future. Predictive analytics requires solid decision-making based on statistics, which is a vital component of predictive analytics for all sorts of businesses in all sectors. Not only does it boost the efficiency and output of business organizations, but it also reduces costs and lessens the risk of legal exposure. In addition to a plethora of other significant benefits, it improves customer retention and acquisition, and it drives up revenue. The potential movements of the market are anticipated by using predictive analytics. In order to put a concept about predictive analytics into practice by using predictive maintenance for a particular company, one instrument that is used is machine learning, and another instrument that is used is methodologies.

According to Lauren N. Singelmann et al. [1] machine learning may give simultaneously appropriate and relevant knowledge when combined with the infrastructure for development participation and the corporate strategy in the education. A computerized intelligence-based classification

ISSN No:-2456-2165

was created so that it could respond to the main investigation. Students' projects were classified using this system to place them in the appropriate proposed theoretical groups. The idea that a machine learning categorization might increase the task's homogeneity should have been developed in order to maintain stronger consistency than a bunch of human evaluators. Multiple linguistic and quantitative categorization were developed and evaluated both internally and against one another to assess the degree to which the proposed technique may lead to improvements.

As stated by Pedro Rico et al. [2] this paper's principal objective is to provide a technique for anticipating the pending sequence of actions and their descriptors in a dynamic and expandable setting related to a particular implementation of continuous improvement. For the purposes of this article, this is the essential point. The first part of this structure serves to foresee the second part's sequence of actions, while the part 2 provides time details for both parts. Each of them is a "major element," the basic building block. Both stages and the procedure for improving the forecasting models have already been detailed for adaptation to large-scale data systems. This was performed so that the interpretation would go more smoothly. Additionally, two case studies have been used to evaluate the development of this application's Apache Flink-based architecture. As a result of using this method, it has been discovered that the architecture can administer and enhance systematic design prediction.

For the purpose of gauging the efficiency of genuine outcome forecasting preventative maintenance techniques, Jongchan Kim et al. [3] reported a technique they created. The effectiveness integrity compared to the threshold criteria for segmentation, the generation uniformity across descriptor categories, and the thermal efficiency across the online component are the three key measurements included. The software's goal is to supplement the standard focus on total performance standards acquired from either a classification algorithm with a more nuanced approach to assessing the efficacy of predicting techniques. Among the metrics used to judge the success of this approach may be recall or precision. The research recommends using the notion of Optimal solutions and a ranking mechanism for situations to compare several categorization and bucketing methodology variations before applying the approach to practical systems.

This research article's literature review is found in the second part. Section 3 describes the suggested strategy, while Section 4 thoroughly evaluates the findings obtained. This study article is finalized in the section 5 including the extent of the future improvements.

II. LITERATURE SURVEY

Qingwen Li [4] et al. narrates the local government may get technology help with macroeconomic evaluation and development thanks to the Economic forecasting method. Indicators of both local and international economic expansion in addition to properly performed improvement may be found in GDP, thus this is crucial. According to the results of this research, a thermally conductive GDP prediction system is relying on deep reinforcement training should be developed. The approach of GDP projection given in this paper offers technical support for the development of reasonable economic planning and the upkeep of a prosperous development. By using signature modeling techniques like distinctive separation and categorization decision, the generator's modeling capacity might be enhanced even more. Possible benefits include improved constituent reliability and optimized computational information.

Among the main focuses of study in the area of knowledge discovery over the last two decades has been the development of new methods for measuring the accuracy of predictions, as stated by Alfonso E. Márquez Chamorro et al. [5]. This seems to be because all of these resources are fundamental for prescriptive and interpretative knowledge extraction, the goals of whom are to maintain guidance and encouragement during the implementation of the program.. In plenty of other respects, improving methods of anticipating evaluation has been a major focus of studies into knowledge discovery. Such recommendations have focused on creating tools that help improve generalization ability, whether by using a more effective teaching technique or by giving capabilities to incorporate underlying information in combination to timing, location, and behavior characteristics. Calculations may be made more confidently by using one of these methods. Nevertheless, there is no approach in the existing literature on predictive observation that supports the consumer to locate the proper metadata for the projection of a certain part of quality assurance. This could be since there wasn't enough effort has been put into the study of the issue.

Yu Liu et al. [6] introduced the LSTM-based resource forecasting for banking branches, and the resulting forecasting method was satisfactorily implemented in a laboratory setting. With this prediction method, a bank branch's everyday liquidity ratio is equivalent to its monthly buffer need. In order to estimate the future monthly buffer demand, they built a long short-term memory (LSTM) system containing 5 convolutional nodes and used it to retrieve the deposit account rules of the banking branches. Based on the results of the experiments with actual data sets, they found that the LSTM forecast technique outperformed both its forerunner, the ARIMA probabilistic model, and its main competitor, the average prognosis technique. This proposed methodology may be improved upon, despite the reality that perhaps the date attribute represents the most crucial consideration.

Qiao Li et al. [7] state that the absence of inter data management (which includes gathering) and the shortage of powerful computing tools are the three main obstacles to tackling broader societal computational concerns. Each of these issues complicate efforts to develop answers to widespread issues in social computations. The IoT is an enormous number that has shown its ability to effectively handle information as well as to its connection. Furthermore, it has been discovered that supervised learning may be employed efficiently as a virtual environment for a broad variety of classic research areas. Putting the two together

ISSN No:-2456-2165

seems like it may be a practical way to enhance many established norms in many different businesses. Inspired by this perspective, this study presents the interdisciplinary approach of IoT and creates a unique architecture to solve the first problem. The predicament may be addressed with the incorporation of a hybrid GNN/CNN artificial neural method in with this framework. Collectively, these neurons could form the cerebral system.

The targeted pricing approach increases efficiency of operations whilst decreasing redundancies and process failures, as shown by the research of Muhammad Adnan Khan et al. [8]. This means that the corporation is lacking the product lines that it obtained goods for owing to projections. Extremely precise forecasting aids in the formulation of a robust advertising strategy, the upping of utilized investment, the decreasing of operating costs related to the supply chain, and the raising of customer delight. The study did look at the benefits of rule-based programming and the capacity to forecast moving average. Forecasted calculations have demonstrated that DeepAR techniques are effective with a high degree of precision and are competitive with one another. That is why DeepAR programs can make such a high proportion of on-target recommendations. This is because the estimated proportion differences are low. As additional data is added, the model's predictions approach closer to the mark. The new research may have consequences for tracking inventory if its findings hold water. Therefore, the stock value rationalization's current status may be seen as a fresh beginning.

Concerned with the challenges of forecasting team effort and team effectiveness, Catherine Sandoval et al. [9] turned to subjective linguistic categorization in their study. Incorporating relevant post knowledge in the prediction method was thought to possess the ability to enhance classifier performance for the entire project, despite the fact that it required to enhance the functionality of the supervised learning or the level of sophistication of the classifications. The study's theories were put to the examination via a series of trials. In these tests, they integrated team activity and quality assessments into a multilayered decision-making architecture or an unified Convolutional Neural Network construction employing twofold data augmentation. In each of these examples, the hypothesis was supported by an improvement in classification results between phases of team effectiveness and group burden.

Jin Eun Yoo et al. [10] suggest using batch normalization as a suitable Machine Learning approach for authentic evaluation. There is a lack of citations for this section. Whenever combined with the gathered information, standardization provides a chance to explore undiscovered connections in between numerous aspects that are associated to the development of learners in virtual classrooms. This study's normalization not only provided an easily digestible prediction system, but also showed predicting performance on level with Random Forest. Multiple opportunities to make use of big data in schooling are emerging as development on normalization in classroom assessment advances. It is possible that standardisation and other Machine Learning methods are unlikely to provide comparable outcomes. This is largely due to the fact that svm classifier may be compromised by data fragmentation that occurs during the testing and calibration processes.

Dinh Lam Pham et al. [11] states that Engagement in corporate virtual communities and precise incident modeling are both required for fully comprehending a daily finances and controlling its personnel successfully. LSTM represents one of the strongest effective strategies for achieving this goal at the present time, there are however others. This study addresses the challenge of the phase after the acquisition of the predicted future occurrences in operational processes by compiling a number of strategies for forecasting the next important information connected to business confidence forecasting observation. This study set out to answer a question about the next step in corporate operations after the acquisition of information pertaining to anticipated future occurrences. They provided an overview of the approach and strategy for forecasting multiple corporate online social changes in along with details about the forthcoming event that used a multidimensional, multi-step LSTM network. These measures were taken in addition to disseminating details about the function. Towards the greatest of their knowledge,, the approach used in this research is the very first to deal with the prediction of an institution's social media channel using data gathered from either a query processing log or a documentary evidence of past predictions.

For predicting a company's performance, Mohamed Gihan Ali et al. [12] suggested a hybrid recommendation technique. Researchers first use Mutual Information Guided Morphological Operations to determine the much more significant features that may be used to distinguish among failed and lucrative Initial coin sales throughout two imbalanced datasets. The number of qualities chosen while retaining the same level of accuracy decreases when knowledge acquisition is utilized to promote adaptation in segmentation. Researchers compared the suggested categorization method to many different types of criteria that were not evenly weighted to show that it was successful. The classification outcomes showed that the suggested method outperformed other classifications by a large margin.

III. PROPOSED SYSTEM

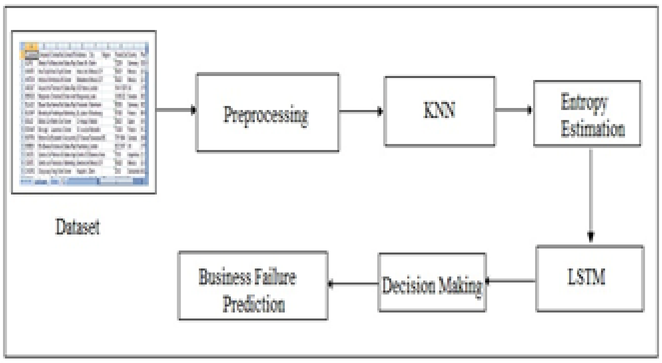

Fig 1: System Overview

The proposed methodology for business failure prediction has been realized using Long Short Term Memory that has been depicted in the figure above.

Step 1: Dataset Collection – The presented approach initiates with the dataset being provided as an input. The dataset for business failure prediction is downloaded from the URL https://www.kaggle.com/datasets/fedesoriano/company-bankruptcy-prediction. This dataset is collected for a number of different businesses along with the various parameters for the same. The attributes of the dataset are extensive and provide an in-depth insight into the functioning of the businesses. This is useful in determining the factors that can be crucial in the success and the failure of the companies.

The attributes include, class label, Operating Gross Margin, Realized Sales Gross Margin, Operating Profit Rate, Operating Expense Rate, Revenue Per Share, Operating Profit Growth Rate, Net Value Growth Rate, Cash Reinvestment Ratio, Debt ratio, Total Asset Turnover, Degree of Financial Leverage, Liability to Equity, etc. These attributes are useful in determining the state of the company and provide excellent information that can be useful in predicting the business failure with reasonable accuracy. The collected data is in the form of a spreadsheet that will be read in the next step of the procedure.

Step 2: Dataset Reading **-** Once the spreadsheet for the data is created, it can now be utilized as an input by the system. The spreadsheet is read by the system through the use of the pandas tool. The data in the spreadsheet is read in the string format, which cannot be utilized effectively in this approach, therefore, the data is converted into float type. This data is also plotted graphically to visualize the pattern of the recorded data through the code. The dataset is now in a format that can be interfaced by the system to perform the various functions and machine learning implementations in the subsequent steps.

Step 3: Dataset Preprocessing **-** The Dataset needs to be preprocessed before being utilized in the methodology, as preprocessing removes the inconsistent and improper data from the dataset which in turn improves the performance of the system. The Correlation is used to achieve the preprocessing of the dataset as depicted in fig. 20. The dataframe_corr() function is used to find the correlation of the columns in in the dataframe in a pairwise manner.

Step 4: Dataset Segmentation **-** The preprocessed dataset is provided as an input to this step, where the dataset is segmented into training and testing samples. This is done to ensure that the ample data is provided to the machine learning implementations as the amount of the data is directly proportional to the accuracy of the model deployed on the said data. Therefore, a large amount of data is useful accessory in improving the accuracy of the presented approach.

Step 5: Scaling and Dataset Reshaping **-** The segmented dataset is provided as an input to this step of the implementation. In this step the dataset is scaled or normalized using the MinMaxScaler function of the sklearn library. This function is utilized to rescale the data in a manner that the resultant values are in the range of [0, 1]. The data is reshaped after rescaling it according to the machine learning algorithm that is being utilized.

For the purpose of implementing the LSTM approach, the training and testing data is reshaped in to three dimensions. The three dimensions of the LSTM input are the features, time steps and sample size. The features are 10, the timestep is 1 and the sample size is 6706 rows.

Step 6: K-nearest Neighbor classification – At this stage of the procedure, the approach takes in both the user input and the dataset that has already been preprocessed from earlier phases. The separation among the input data as well as the individual cells of the preprocessed collection is being estimated with the help of these data. The equation 1 that is provided below would be used to get the distance.

$$ED = \sqrt{\sum(ATi - ATj)^2} \quad \text{\_\_\_\_\_\_\_\_} (1)$$

Where,
ED=Euclidian Distance
A_{Ti}=Attribute at index i
A_{Tj}= Attribute at index j

With the bubble sort method, the complete list is essentially categorized into the increasing order after the distances between every one of the rows has already been computed and added to the relevant rows. This brings the list to its final destination. The k value is going to be set to 2, which will culminate in two different clusters after this function is finished. The very first cluster, which has the relevant information, is passed on to the subsequent stage in order to estimate the entropy making use of Shannon's information gain.

Step 7: Entropy Estimation – It will be necessary to do an analysis on the information gain scores of the clusters and the properties of the produced dataset. As an input for this step of the procedure, the clusters that were produced in the stage before this one are used. The Shannon information gain method is used in the computation of entropy for attribute qualities.

An accurate score may be determined by the examination of the clusters that were completed in the past. After then, the Shannon information gain calculation, which was introduced previously, is applied to this score in order to determine the information gain. When the entropy estimates have been acquired, they are then written down in the form of a list and passed on to the subsequent step so that they may be evaluated further.

Step 8: LSTM – Long Short Term Memory – The LSTM approach is utilized through the implementation of the Keras LSTM module for our approach. The Keras library allows for the realization of the machine learning concepts with greater ease and effective control over the layers. The LSTM approach is initialized as sequential and a single LSTM layer is added with the 10 kernels and the resized shape achieved in the previous steps.

The feed the unit's value is designated as 10 for the execution of our approach. The activation function is the tanh function along with the adam optimizer that is utilized to compile the LSTM module. The batch size for the execution is designated as 30 with 150 epochs to fit the network. The Dense layer is also added which utilizes tanh as the activation function. The model is compiled with the adam optimizer and the mean squared error is being used for the metrics as accuracy.

The mathematical model for the proposed methodology has been depicted below.

Mathematical Model

S= { } be as system Business Failure Prediction System
Identify Input as D={ D_1, D_2, D_3….. D_n}
Where D = Dataset
S={D }
Identify B_P as Output i.e. Business Prediction
S= { I, B_P }
Identify Process P
S= { I, P, B_P }
P= { P,K_{NN}, L_{STM} ,D_M}
Where
P=Preprocessing
K_{NN} =K-Nearest Neighbor
L_{STM} = Long Short Term Memory
D_M = Decision Making
So complete system for Business Prediction System can be given as
S = { I, P,K_{NN}, L_{STM} ,D_M, B_P }

IV. RESULTS AND DISCUSSIONS

The research framework for Business Failure Prediction using machine learning has been developed along with the implementation of the Spyder IDE. For the purpose of implementing this strategy, the Python programming language was chosen to serve as the primary language for the software. The laptop used for the development has a standard configuration, consisting of a 1 terabyte (TB) hard drive, an Intel i5 processor, and 16 gigabytes (GB) of RAM.

This strategy has been put through a rigorous evaluation in order to ensure that an accurate assessment of the functioning of the suggested technique has been made. The concepts of precision and recall were used to the investigation of the assessment criteria.

➢ *Performance Evaluation based on Precision and Recall*

Precision and recall are two incredibly helpful techniques to measure how exactly a given module in the paradigm is implemented. Both of these metrics can be found in the context of our methodology. The component's precision, which encompasses its dependability throughout a broad range, is what defines the component's relative validity.

The ratio of the number of accurate Business Failure Predictions collected to the total number of trials that were carried out was used to determine this method's precision parameter. On the other hand, the recall criteria are a complement to the precision measurement and help in evaluating the total reliability of the LSTM constituent. This is because the precision measurement is not sufficient by itself.

In this approach, the recall is calculated by comparing the number of accurate Business Failure Forecasts to the total number of inaccurate Business Failure Predictions. The following formulae provide a quantitative expansion of this point.

Precision and Recall can be depicted as below:

- ✓ A = The number of accurate Business Failure Predictions
- ✓ B = The number of inaccurate Business Failure Predictions
- ✓ C = The number of accurate Business Failure Predictions not done

So, precision can be defined as

Precision = (A / (A+ B)) *100
Recall = (B / (B+ C)) *100

The experimental results that were obtained by utilizing the aforementioned formula are shown in a particular manner below in Table 1. These statistical measurements are applied in order to give a graphical depiction, which is shown in figure 2.

No. of Trials	Accurate Business Failure Predictions (A)	Inaccurate Business Failure Predictions (B)	Accurate Business Failure Predictions not done (C)	Precision	Recall
1	1	0	0	100	100
4	3	0	1	100	75
7	6	1	0	85.71429	100
10	8	2	0	80	100
13	10	1	2	90.90909	83.33333
16	13	2	1	86.66667	92.85714
19	17	1	1	94.44444	94.44444
22	17	3	2	85	89.47368
25	20	3	2	86.95652	90.90909
28	24	2	2	92.30769	92.30769

Table 1: Precision and Recall Measurement Table

ISSN No:-2456-2165

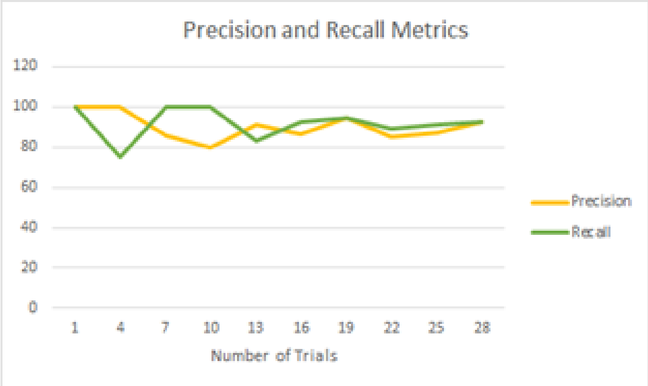

Fig 2: Comparison of Precision and Recall

The functionality of the LSTM and its ability to make accurate predictions based on the input data are shown on the graph over a wide range of trial counts. The remarkable dependability of the method is shown by its precision and recall rates, which come in at 90.19 percent and 91.83 percent, correspondingly. These statistics are remarkably substantial for the first execution of such a procedure, and the accomplishment that has been achieved as a consequence of this is satisfying.

V. CONCLUSION AND FUTURE SCOPE

The research framework for Business Failure Prediction using machine learning has been elaborated in detail in this research paper. The proposed approach initiates with the business information dataset that is collected from a number of different companies. This data consists of various attributes that are useful in the realization of the business failure prediction. The dataset is first effectively preprocessed to eliminate any redundant aspects in the dataset which need to be removed to increase the effectiveness of the approach. The preprocessed dataset is then subjected the KNN model to achieve the clusters of the data. The clusters are then utilized for the entropy estimation using the Shannon information gain. Once the entropy values are achieved, the clusters along with the entropy values are provided to the Long Short Term Memory module that performs the prediction based on the input values. The LSTM outputs an effective probability score for the business failure prediction that needs to be classified before being presented to the user. The decision making approach classifies the predictions and displays the output of the business failure prediction to the user. The approach has been tested extensively for its performance through the use of precision and recall which achieved 90.19 percent and 91.83 percent, precision and recall respectively.

For the future development purpose the propose model can be implement in banking domain, and also in stock market trading to analyze the possibilities of business failures.

REFERENCES

[1]. L. N. Singelmann and D. L. Ewert, "Leveraging the Innovation-Based Learning Framework to Predict and Understand Student Success in Innovation," in IEEE Access, vol. 10, pp. 36123-36139, 2022, doi: 10.1109/ACCESS.2022.3163744.

[2]. P. Rico, F. Cuadrado, J. C. Dueñas, J. Andión and H. A. Parada G., "Business Process Event Prediction Through Scalable Online Learning," in IEEE Access, vol. 9, pp. 136313-136333, 2021, doi: 10.1109/ACCESS.2021.3117147.

[3]. J. Kim and M. Comuzzi, "Stability Metrics for Enhancing the Evaluation of Outcome-Based Business Process Predictive Monitoring," in IEEE Access, vol. 9, pp. 133461-133471, 2021, doi: 10.1109/ACCESS.2021.3115759.

[4]. Q. Li, C. Yu and G. Yan, "A New Multipredictor Ensemble Decision Framework Based on Deep Reinforcement Learning for Regional GDP Prediction," in IEEE Access, vol. 10, pp. 45266-45279, 2022, doi: 10.1109/ACCESS.2022.3170905.

[5]. A. E. Márquez-Chamorro, K. Revoredo, M. Resinas, A. Del-Río-Ortega, F. M. Santoro and A. Ruiz-Cortés, "Context-Aware Process Performance Indicator Prediction," in IEEE Access, vol. 8, pp. 222050-222063, 2020, doi: 10.1109/ACCESS.2020.3044670.

[6]. Y. Liu, S. Dong, M. Lu and J. Wang, "LSTM based reserve prediction for bank outlets," in Tsinghua Science and Technology, vol. 24, no. 1, pp. 77-85, Feb. 2019, doi: 10.26599/TST.2018.9010007.

[7]. Q. Li, Y. Song, B. Du, Y. Shen and Y. Tian, "Deep Neural Network-Embedded Internet of Social Computing Things for Sustainability Prediction," in IEEE Access, vol. 8, pp. 60737-60746, 2020, doi: 10.1109/ACCESS.2020.2982986.

[8]. M. A. Khan et al., "Effective Demand Forecasting Model Using Business Intelligence Empowered With Machine Learning," in IEEE Access, vol. 8, pp. 116013-116023, 2020, doi: 10.1109/ACCESS.2020.3003790.

[9]. C. Sandoval, M. N. Stolar, S. G. Hosking, D. Jia and M. Lech, "Real-Time Team Performance and Workload Prediction From Voice Communications," in IEEE Access, vol. 10, pp. 78484-78492, 2022, doi: 10.1109/ACCESS.2022.3193694.

[10]. J. E. Yoo, M. Rho and Y. Lee, "Online Students' Learning Behaviors and Academic Success: An Analysis of LMS Log Data From Flipped Classrooms via Regularization," in IEEE Access, vol. 10, pp. 10740-10753, 2022, doi: 10.1109/ACCESS.2022.3144625.

[11]. D. -L. Pham, H. Ahn, K. -S. Kim and K. P. Kim, "Process-Aware Enterprise Social Network Prediction and Experiment Using LSTM Neural Network Models," in IEEE Access, vol. 9, pp. 57922-57940, 2021, doi: 10.1109/ACCESS.2021.3071789.

[12]. M. G. Ali, I. I. Gomaa and S. M. Darwish, "An Intelligent Model for Success Prediction of Initial Coin Offerings," in IEEE Access, vol. 10, pp. 58589-58602, 2022, doi: 10.1109/ACCESS.2022.3178369.