2023 International Conference on Applied Intelligence and Sustainable Computing (ICAISC)

Forecasting of E-Commerce System for Sale Prediction Using Deep Learning Modified Neural Networks

1st S. Neelakandan
Department of Computer Science and Engineering,
R.M.K. Engineering College,
Chennai, India,
snk.cse@rmkec.ac.in

2nd V. Prakash
Department of Computer Science and Engineering
R.M.K. Engineering College
Chennai, India
prak18227@rmkec.ac.in

3rd M. S. PranavKumar
Department of Computer Science and Engineering,
R.M.K. Engineering College
Chennai, India,
pran18228@rmkec.ac.in

4th R. Balasubramaniam
Department of Computer Science and Engineering,
R.M.K. Engineering College
Chennai, India,
raja18230@rmkec.ac.in

***Abstract*—E-commerce is the practice of purchasing and selling goods over the Internet. Customers will appreciate the simplicity of not having to make physical purchases using e-primary commerce. They will order the item online and deliver it to their home as soon as feasible. This research aimed to create deep learning algorithms capable of forecasting e-commerce sales. This project aims to develop and test a model for predicting online product sales over a wide range of online product types. It bases its decisions on the requirements for online product sales, the factors influencing online product sales in various industries, and the benefits of the deep learning algorithm. A continuous Stochastic Fractal Search (SFS) method for optimizing the parameters of a deep learning-modified neural network (DLMNN) is introduced in this paper. In e-commerce demand forecasting studies, a time series dataset is also analyzed. The DLMNN model's performance improvements across multiple product categories are illustrated using a non-deep learning model as a baseline comparison.The experiment also shows that the unsupervised pretrained DLMNN model outperforms the competition in terms of sales predictions. For root mean square error, the proposed technique obtained RMSE, Mean, and Standard Deviation.**

Keywords—E-commerce system, forecasting, deep learning modified neural networks(DLMNN), deep learning.

I. Introduction

As just a brand-new commerce model for the twenty-first period, e-commerce has a lot of potentials. Economically speaking, e-commerce growth was never higher [1]. The e-commerce industry is increasing. It is anticipated that the market will be worth $4.88 trillion. Identifying the appropriate pricing for all of the product lines on the platform regularly to maximize profits and revenues is unique of the most excellent inspiring tasks for e-commerce businesses. Demand can change significantly from day to day in the erratic E-commerce sector. Retailers and wholesalers must therefore be agile adapters to shifts in consumer demand. More than simply reacting to market changes is required because you risk running out of inventory to fulfill an increased request or missing out on a trend entirely. You must be proactive in anticipating changes to stay ahead of the competition. The two components of e-trading volume able to forecast techniques developed are the regression analysis model and the device learning model. Machine learning methods are more attractive than numerical regression models since they can fit non-linear data more effectively and have a more significant aspect of data ability to adapt [2].

One of the most widely used types of computational intelligence in applications for medical imaging is deep learning. The classification of medical images, a pattern recognition application, extensively uses deep learning techniques. These deep learning-based techniques for classifying medical images are frequently employed in automated disease diagnostic systems. The population-based stochastic diffusion search (SDS) pattern was first described in 1989. It is a member of a group of SDS, the first of a family of naturally inspired search algorithms. The tandem calling mechanism one ant species uses comparable to the direct (one-to-one) communication method SDS uses between the agents.

SDS agents carry out low-cost, incomplete evaluations of a hypothesis. They then have direct one-on-one conversations to exchange information about their views. The diffusion mechanism identifies high-quality solutions from clusters of agents sharing the same idea. The best way to comprehend how SDS works is through a straightforward analogy. The following portions of the paper are structured as surveys. In part 2, we evaluate prior research that has addressed the privacy issue and suggest alternative remedies. The proposed algorithm and the privacy analysis are described in Section 3. The experimental data are presented and discussed in detail in Section 4. The study's result is summarised in Section 5, along with potential directions for future research.

II. Literature Survey

Convolutional neural network technology was reportedly used in a study by K. Zhao et al.[3] to anticipate e-commerce sales, according to a single source. The process needs case-by-case manual feature engineering for particular scenarios, which is difficult, time-consuming, and requires a high skill level. This study's goal was to address the mentioned limitation. This study evaluated the approach's capacity to auto-extract the most beneficial features and generate sales forecasts based on such characteristics.

Ganatra et al. [4]. Deep learning is one of the subcategories of the larger category of algorithms for machine learning described in 2018. An algorithmic type is deep learning. It is

2023 International Conference on Applied Intelligence and Sustainable Computing (ICAISC) | 979-8-3503-2379-5/23/$31.00 ©2023 IEEE | DOI: 10.1109/ICAISC58445.2023.10199817

predicated on the notion that education should develop naturally over time. E. M. Hassib et al. [5] Predicting e-commerce sales, which include transactions and growth, constancy, and decline, can help you improve and recognize the lifecycle of an e-commerce stage. This can help you make better business decisions. This knowledge can also help you better understand how various elements affect short-term sales, such as discounts, prices, seasonal variations, and web rankings. Time series analysis has become a popular method for sales forecasting, according to research by Ranjith et al. [6]. This technique uses the auto-regressive function, a helpful tool for many prediction investigations. Examining machine learning algorithms for sales forecasting revealed that sales prediction has become a popular business intelligence tactic. The neural network model developed by Q. Qu et al. [7] for estimating cigarette sales is one significant addition to this subject. An improved Levenberg-Marquardt algorithm is used in the model for the backpropagation neural network. The importance of this strategy is supported by the fact that numerous studies using time series prediction approaches have also shown equivalent results.

J. Liu et al. [8] employed a deep learning technique to build a crown model that was then used to categorize online sales of agricultural products. This was accomplished by first considering all the characteristics of the data obtained from agricultural e-commerce sales and then basing the crown model on the deep learning algorithm. This model was used in their research.

Arunraj et al. [9]. Small and medium-sized e-commerce businesses can succeed by providing customized goods and services while implementing market segmentation and improving their services to fulfill the rising need for online transactions. Due to technology and resource limitations, small and medium-sized e-commerce enterprises need to give enough thought to the enormous volumes of transaction data they collect.

The main goal of time series analysis, according to Haviluddin et al. [10], is to make future predictions using historical data. A type of time series data is sales data for online goods. Theoretically, future sales might be predicted using fitting and regression techniques applied to previous data, notably time series analysis and its associated models. However, handling the sale of various products in a different orders is necessary.

According to Fouad, M. M. et al. [11], accurate demand forecasting significantly impacts business decision-making, which is vital for the success of e-commerce. Businesses need reliable sales forecasts for their e-commerce platform to manage their workforce and finances and increase operational effectiveness efficiently. Additionally, forecasting gives companies access to information about their financial status within e-commerce, allowing them to optimize their supply chain management and boost overall economic performance. A sales forecast helps manage inventory, competitive rates, and timeously marketing strategies by enabling an e-commerce framework to term profits to better accuracy and dependability

III. PROPOSED SYSTEM

As illustrated in fig. 1 sales forecasting is a crucial task for e-commerce because it significantly impacts the company's strategic choices. Understanding their financial position and managing their workforce is made possible by needing sales expectations for the e-commerce stage [12]. However, they can enhance their supply chain management and further their financial condition to better understand their financial situation by having transaction expectations for the e-commerce stage.

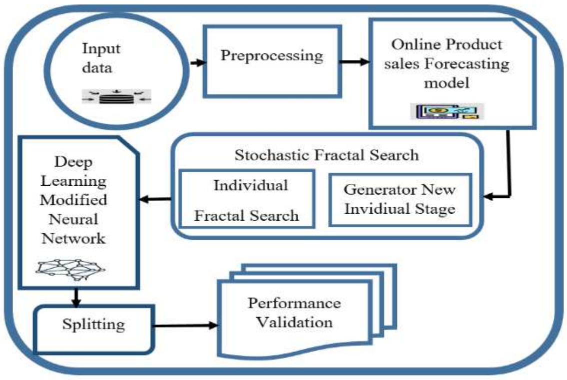

Fig. 1. Proposed SFS –DLMNN method.

A. Data Collection

This study's dataset will be made up of transactional data. This is because we intend to produce a transactions forecast, which requires utilizing all prior transactional data to estimate future sales. The open-source and unrestricted-use transactional dataset will come from an online store. The dataset was obtained from the Brazilian E-Commerce Public Dataset by List Store, [13,14]available on Kaggle.com.

B. Pre-processing

The participation data is standardized at the pre-processing level after choosing the proper characteristics to create a suitable recognition for deep-learning estimators. When the designated article does not represent pretty standard, generally dispersed information, the methods used by deep learning estimators can occasionally behave incorrectly[15-16]. The extracted features in this study are updated using the Standard Scaler method, which removes the mean and scaled to unit variance [15].

C. Forecasting Model for Online Product Sales

The fully connected layer is by far the most basic structure in deep learning, to begin with. The layers of the network are thick, and the links between neurons are perfect. A deep learning structure known as a convolutional (CNN), which excels in the field of vision, was inspired by biological vision systems. Convolution neural networks consist of four main types of layers: input, result, convolution, and lakeThe overall architecture is composed of several layers that resemble a completely connected layer[17-19]. The hidden layer's convolution and pooling layers can be ordered and combined to enhance the model's functionality. The parameter sharing of the convolutional kernel in the convolution layer, which leads to a significant decrease in the overall number of model parameters, is a substantial benefit. The pooling layer can reduce model parameters and extract more valuable features indefinitely [16].

D. Stochastic Fractal Search (SFS)

The SFS algorithm is inspired by fractal and biological growth in nature. Individual update and individual fractal are the two phases of the algorithm[20]. To simultaneously eliminate people who perform poorly and keep people who serve well, a decision should be made after each update based on the scope of the suitability price.

E. Individual Fractal Stage

The upper and lower boundaries of the planning space are denoted by the symbols U and L in the formula, respectively. In particular, U stands for the upper border [17].

$$P_i = L + \varepsilon * (U - L) \quad (1)$$

After initializing individual populations following ((2) or (3)), the algorithm generates new individuals using the Gaussian walking distribution. Although the random fractal principle has a wide range of applications, the distinct fractal period uses the Gaussian stride distribution method because of its superior performance in the unexpected direction.

$$GW_1 = Gaussian(\mu_{MP,}\ \sigma) + (\varepsilon * BP - \varepsilon^1\ * pi) \quad (2)$$

$$GW_2 = Gaussian(\mu_p, \sigma) \quad (3)$$

In the formulation, where BP stands for the newfangled separate situation, μ p stands for the average value of Pi, and Gaussian stands for the Gaussian .

In the formulation, where BP stands for the newfangled separate situation, μ p stands for the average value of Pi, and Gaussian stands for the Gaussian distribution. P ( ) I N = 1,2,3,... . N denotes the population's size, and ′ denotes a chance value between [0,1] and variance.

$$\sigma = \left| \frac{\log(k)}{k} * \left(p_i - BP\right) \right| \quad (4)$$

The formula's iterations are indicated by the letter K.

F. Creates New Personas Stage

The SFS algorithm suggests a customized updating strategy to recover the algorithm's growth and examination competencies. The first updating phase and the second updating phase are the two phases that make up the strategy.

People in the population P′ must use the following formula to determine the independent selection chance after each fractal produces new individuals [18,19]:

$$pa_i = \frac{rank\left(p^1{}_i\right)}{N} \quad (5)$$

Pi ′ denotes an individual's capacity for adaptation within a population of P′, while rank P(i)′ denotes the order of flexibility within a population of P′. The rank function in the formula sorts items from small to large.

$$P''_{i.k} = P''_{r,k} - \varepsilon * \left(P''_{t,k} - P''_{i,k}\right) \quad (6)$$

A person has a lower chance of being selected for updating if they perform better, according to the formula, rt N, 1,2,3,..., ∈[ ] and r t ≠ . r t ≠.from Calculation (6). Conversely, the likelihood of being chosen for an update increases with the lower performance of the individual. When the newly updated individual's fitness level surpasses the parent's, the parent is replaced.

$$\begin{cases} P_i^{''''} = P_i^n - \varepsilon^1 * (P_t'' - BP)\varepsilon^1 \le 0.5 \\ P^{''''}{}_i = P_i^{''''} + \varepsilon^1 * \left(p_t'' - p_r''\right)\varepsilon^1 \ge 0.5 \end{cases} \quad (7)$$

Individuals Pt ″ and Pr ″ are two people chosen at random from the population P ″ in the formula, and ε ′ are Gaussian random numbers with uniform distribution. When the fitness level of the new person exceeds that of the parent, the parent must also be replaced. This update phase differs from the first in that it updates the people who perform poorly to prevent the algorithm from reaching a local optimum.

G. Deep learning-modified neural networks

Big data is characterized as vast quantities that are difficult to mine using conventional techniques because they contain a lot of information in various forms and noise. This essay suggests a DLMN. Cleaving is the first step in pre-processing input data. The pre-processed data is then used to derive the frequency of the entropy components as depicted in fig 2.

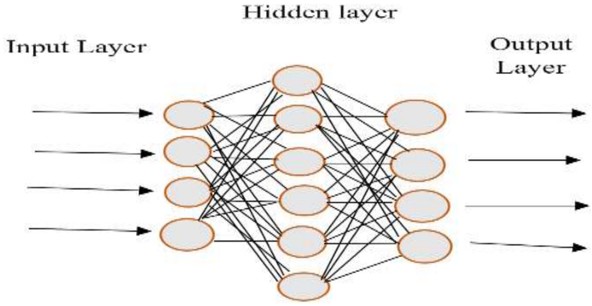

Fig. 2. Proposed DLMNN model.

$$\Delta D_{s^1} = \left\{ D_{s^1}, D_{s^2}, D_{s^3}, ...D_{n^j} \right\} \quad (8)$$

H. Splitting

This is how the information has been divided into manageable, discrete chunks. Large amounts of data are divided into manageable chunks during the preprocessing stage, enabling the identification of information highlights. The next step is tokenization[20].

The dataset Ds ¼ Ds1, Ds2, Ds3, and...... The input is assumed to be Dn. Every document in this process is given an equal number of Cr1s, Cr2s, Cr3s, etc. The split work, for instance, on document Ds1 is designated as Eq.

$$B_1 = D_{s_1} \left\{ C_{r^1}, C_{r^2}, C_{r^3}, ...C_{r^n} \right\} \quad (9)$$

B1 denotes the split document, and r n indicates the word count per document. Similarly, each record in the datasets Ds1, Ds2, and Ds3. According to the fragmented joint group B[i], Dn splits.

$$B_{[1]} = B_1 + B_2 + B_3 ....B_n \quad (10)$$

IV. Result And Discussion

A. Experimental Validation

This work conducts an observational study on 13000 vacuuming samples based on a previously published model training technique and incorporating several influencing parameters for online product sales prediction. To demonstrate the model's predictive power and flexibility in predicting product sales across industries, this paper examines model training results. Data pre-processing can quickly lead to instability in the model training process because of the varying dimensions of the selected indicators. As a result, all indicator data in this paper have been standardized.

B. Performance Metrics

The mean absolute error (MAE), mean bias error, and root mean square error (RMSE) will be used as performance indicators for the experiment. The RMSE metric can be calculated using the formula below, and prediction performance can be evaluated using it as well:

$$RMSE = \sqrt{\frac{\Sigma^{n}_{i=1} (H_{p,i} - H_i)^2}{n}} \tag{11}$$

Where Hi is the accurate restrained price and Hp, i is the expected price. There are n total values in the set. The MAE determines the location of predictions' average number of errors. The calculation is as follows:

$$MAE = \frac{\Sigma^{n}_{i=1} \left| H_{p,i} - H_i \right|}{n} \tag{12}$$

The tested model's under- or overprediction can be found using the MBE. By averaging the dimensional disparities between the calculated and observed values, it calculates the average bias of prediction.

$$MBE = \frac{\Sigma^{n}_{i=1} \left| H_{p,i} - H_i \right|}{n} \tag{13}$$

1) RMSE Analysis

TABLE. I. RMSE ANALYSIS FOR SFS-DLMNN TECHNIQUE WITH EXISTING SYSTEM

No of node	PSO	WOA	BRNN	GA	SFS-DLMNN
100	52.87	47.32	42.87	37.12	32.98
200	53.23	48.78	43.14	38.88	33.18
300	54.98	49.12	44.87	39.13	34.55
400	55.18	50.33	45.66	40.54	35.98
500	56.44	51.67	46.12	41.22	36.12

Figure 3 and Table 1 display an RMSE comparison of the SFS-DLMNN methodology with other popular methods. The graph shows that machine learning led to better outcomes with a lower RMSE value. The RMSE, for instance, is 32.98% for the SFS-DLMNN model with node 100, while it is only marginally better for the PSO, WOA, BRNN, and GA models, at 52.87%, 47.32%, 42.87%, and 37.12%, respectively. On the other hand, the SFS-DLMNN model has been shown to perform at its peak for various nodes with low RMSE values. Similar to this, the PSO, WOA, BRNN, and GA models have RMSE values of 56.44%, 51.67%, 46.12%, and 41.22%, respectively, whereas the SFS-DLMNN model's RMSE value under 500 nodes is 36.12%.

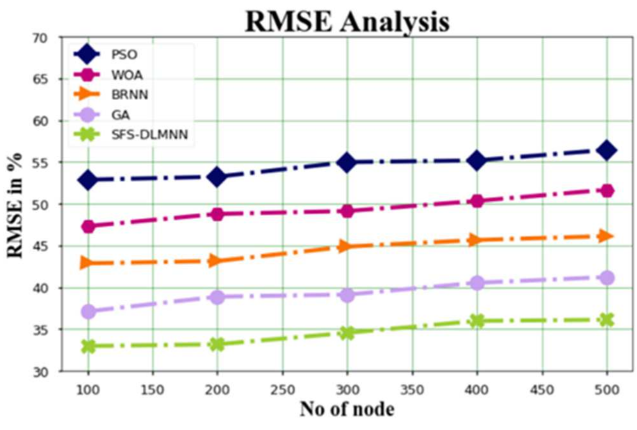

Fig. 3. RMSE Analysis for SFS-DLMNN technique with existing system.

2) MAE Analysis

TABLE. II. MAE ANALYSIS FOR SFS-DLMNN TECHNIQUE WITH EXISTING SYSTEM

No of node	PSO	WOA	BRNN	GA	SFS-DLMNN
100	64.55	59.35	54.18	49.32	45.78
200	66.18	60.12	55.98	49.12	44.19
300	66.19	61.77	56.34	51.66	46.78
400	67.11	62.87	57.12	52.89	47.23
500	66.22	63.19	58.33	53.55	48.55

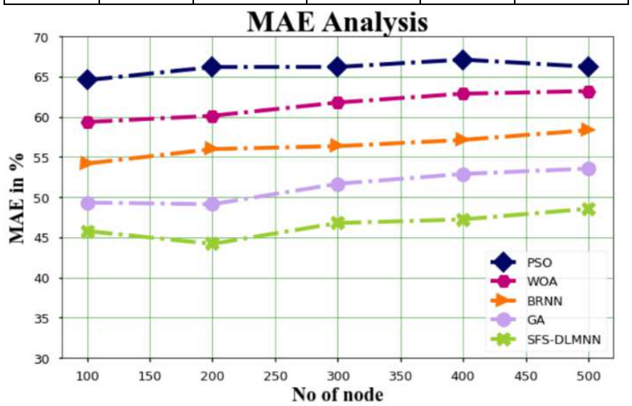

Fig. 4. MAE Analysis for SFS-DLMNN method with existing system.

An MAE comparison of the SFS-DLMNN approach with other current methods is shown in Figure 4 and Table 2. The graph demonstrates that the machine learning method yielded superior outcomes with a lower MAE value. For instance, the SFS-DLMNN model's MAE value for node100 is 45.78%, whereas the MAE values for the PSO, WOA, BRNN, and GA models are slightly higher at 64.55%, 59.35%, 54.18%, and 49.32%, respectively. On the other hand, the SFS-DLMNN model has shown to perform at its peak for a number of nodes with low MBE values. Similar to this, the MAE values for the PSO, WOA, BRNN, and GA models are 66.22%, 63.19%, 58.33%, and 53.55%, respectively, while the MBE value for the SFS-DLMNN is 48.55% under 500 nodes.

3) MBE Analysis

TABLE. III. MBE ANALYSIS FOR SFS-DLMNN METHOD WITH EXISTING SYSTEM

No of node	**PSO**	**WOA**	**BRNN**	**GA**	**SFS-DLMNN**
100	28.17	43.98	33.67	38.17	21.98
200	27.55	42.19	32.16	37.55	20.19
300	30.16	45.12	34.13	38.11	22.98
400	31.22	46.22	36.22	41.26	26.77
500	32.66	47.76	37.98	42.45	27.12

An MBE comparison of the SFS-DLMNN methodology with other current approaches is shown in Figure 5 and Table 3. The graph demonstrates that using machine learning led to better outcomes with a lower MBE value. For instance, the SFS-DLMNN MBE value for node100 is 21.98%, whereas the MBE values for the PSO, WOA, BRNN, and GA models are slightly better at 28.17%, 43.98%, 33.67%, and 38.17%, respectively. On the other hand, the SFS-DLMNN model has shown to perform at its peak for a number of nodes with low MBE values. Similar to this, the MBE value of the SFS-DLMNN is 27.12% under 500 nodes, whereas the MBE values of the PSO, WOA, BRNN, and GA models are 32.66%, 47.76%, 37.98%, and 42.45%, respectively.

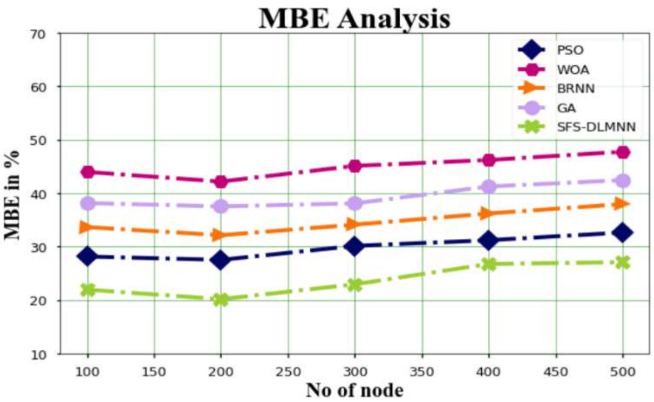

Fig. 5. MBE Analysis for SFS-DLMNN method with existing system.

V. Conclusion

While working on this project, it became evident that there are numerous approaches to using deep learning to predict e-commerce transactions. The investigator, however, could only emphasize four procedures frequently applied to sales forecasting. The researcher could shape and trial each of the designated deep knowledge reproductions. The most efficient procedure is determined by the range of the model's predictions where the predicted and actual values roughly coincide. The researcher will simultaneously implement this ideal algorithm into a web application they create. To better understand the pattern and outcomes, the researcher advises using a different time series model in the future. To identify pattern differences and how they affect sales, gather data from other e-commerce sites. Finally, the dataset size should be increased since 20 months of data must be raised for accurate forecasting and trend comparison.

References

[1] A. Intezari, and B. McKenna, "Creating organizational culture Leadership Regional," Global Perspectives, vol. 160, pp. 160–189,2018.

[2] K. Weerts, W. Vermeulen, and S. Witjes, "On corporate sustainability integration research Analyzing corporate leaders experiences and academic learnings from an organizational culture perspective," Journal of Cleaner Production, vol. 203, no.4, pp.1201–1215, 2018.

[3] K. Zhao, and C. Wang, "Sales Forecast in E-commerce using Convolutional Neural Network," pp. 1708.07946, 2017.

[4] N. Ganatra, and A. Patel, "A comprehensive study of deep learning architectures applications and tools," International Journal of Computer Sciences and Engineering, vol.12, no. 6, pp.701–705 2018.

[5] E. M. Hassib, A. I. El-Desouky, E. M. El-Kenawy, and S. M. El-Ghamrawy "An imbalanced big data mining framework for improving optimization algorithms performance," IEEE Access, vol.7 1,pp.170774–170795, 2019.

[6] C. P. Ranjith, B. M. Hardas, M. S. K. Mohideen, N. N. Raj, and N. R. Robert, "Robust Deep Learning Empowered Real Time Object Detection for Unmanned Aerial Vehicles based Surveillance Applications," Journal of Mobile Multimedia, pp.451-476,2023.

[7] Q. Qu, C. Liu, and X. Bao, "E-commerce enterprise supply chain financing risk assessment based on linked data mining and edge computing Mobile Information Systems," pp.9938325, 2021.

[8] J. Liu, S. He, and Q. Zhu, "Research on online agricultural product sales forecasting model based on deep learning," Computer Application Research, vol.34, no 8, pp.2291–2293,2017.

[9] N. S. Arunraj, and D. Ahrens "A hybrid seasonal autoregressive integrated moving average and quantile regression for daily food sales forecasting,"Int J Prod Econ, vol.170, pp.321–335, 2015.

[10] H. Haviluddin, R. Alfred and J. H. Obit"A performance comparison of statistical and machine learning techniques in learning time series data," Adv Sci Lett, vol. 21, no.10,pp.3037–3041,2015.

[11] M. M. Fouad, A. I. El-Desouky, R. Al-Hajj, and E. S. M. El-Kenawy "Dynamic group-based cooperative optimization algorithm," IEEE Access, vol. 8, pp. 148378–148403, 2020.

[12] M. Prakash, and C. J. Kavitha Priya, "An Analysis of Types of Protocol Implemented in Internet of Things Based on Packet Loss Ratio," In Proceedings of the Second International Conference on Information and Communication Technology for Competitive Strategies, pp. 1-6, 2016.

[13] M. Arul Vinayakam Rajasimman, R. K. Manoharan, N. Subramani, and M. Aridoss, M. G. Galety (2023) "Robust Facial Expression Recognition Using An Evolutionary Algorithm with a Deep Learning Model," Applied Sciences ,vol. 13, no. 1, pp. 468, 2023. https://doi.org/10.3390/app1301048

[14] J. Rene Beulah, C. Kranthi Kumar Singamaneni, C. Pretty Diana Cyril, S. Neelakandan and S. Velmurugan "Automated speech based evaluation of mild cognitive impairment and Alzheimer's disease detection using with deep belief network model," International Journal of Healthcare Management, 2022. DOI: 10.1080/20479700.2022.2097764.

[15] D. Anuradha, Javid Iqbal, Mohammad Gouse Galety, Ruby Singh and S. Neelakandan "A novel convolutional neural network with gated recurrent unit for automated speech emotion recognition and classification.," Journal of Control and Decision, 2022. DOI: 10.1080/23307706.2022.2085198.

[16] G. L. N. Murthy, N. Arulkumar and Nilesh Shelke, "An efficient low complexity compression based optimal homomorphic encryption for secure fiber optic communication," Optik, pp. 252:168545, 2021. https://doi.org/10.1016/j.ijleo.2021.168545.

[17] P. Mohan, N. Subramani, and M. Subramanian, S. Meckanzi "Handcrafted Deep-Feature-Based Brain Tumor Detection and Classification Using MRI Images," Electronics, vol. 11pp. 4178,2022. https://doi.org/10.3390/electronics11244178.

[18] K. Sathishkumar, V.E, "Deep learning based sentiment analysis and offensive language identification on multilingual code-mixed data," Sci,vol. 12, pp. 21557 https://doi.org/10.1038/s41598-022-26092-3.

[19] Ahmed Mohammed Metwally, and Madhappan Santhamoorthy, M. Satyanarayana Gupta"Metaheuristics with Deep Transfer Learning Enabled Detection and classification model for industrial waste management,"Chemosphere, pp. 136046, 2022. https://doi.org/10.1016/j.chemosphere.2022.136046.

[20] Deepak Kumar Jain, Xue Liu, and Mohan Prakash "Modeling of human action recognition using hyperparameter tuned deep learning model," Journal of Electronic Imaging, vol. 32,no. 1, pp. 011211, 2022. https://doi.org/10.1117/1.JEI.32.1.011211.