Deep Learning Advancements in E-commerce Supply Chain Management in Forecasting and Optimization Strategies

Jagendra Singh
School of Computer Science Engineering & Technology, Bennett University,
Greater Noida, India
jagendrasngh@gmail.com

Nitin Arvind Shelke
School of Computer Science Engineering & Technology, Bennett University,
Greater Noida, India
nitin.shelke@bennett.edu.in

Kamal Upreti
Department of Computer Science , CHRIST (Deemed to be University), Delhi NCR Campus,
Ghaziabad, India
kamalupreti1989@gmail.com

Fariyah Banu Jamaluddin Saiyad
Department of Commerce, Bath spa university Rak,
Ras Al-Khaimah, United Arab Emirates
Drfariyahsaiyad@gmail.com

Prakash Divakaran
Department of management, Himalayan University,
Itanagar, India
prakashtek@gmail.com

Khadilkar Sujay Madhukar
KIT's Institute of Management Education and Research,
Kolhapur, India
Khadilkarsm@rediffmail.com

2024 IEEE 4th International Conference on Sustainable Energy and Future Electric Transportation (SEFET) | 979-8-3503-8399-7/24/$31.00 ©2024 IEEE | DOI: 10.1109/SEFET61574.2024.10718068

***Abstract-* In this study, the influence of deep learning technologies on the optimization of supply chain management in the context of the e-commerce industry is examined. Using a dataset of historical data of sales, inventories, market fluctuations, and customer and supplier details, I investigate the efficiency of different deep learning models to predict demand and facilitate the optimal balance of inventories. Long Short-Term Memory (LSTM), Recurrent Neural Networks (RNN), Convolutional Neural Networks (CNN), and a model proposed by the authors are defined and applied, considering their accuracy, precision, recall, and F-1 score. The results show that the proposed model outperforms traditional products, achieving 97.5% of accuracy. In the context of the comparative analysis, the specific features of CNN, LSTM, and RNN are revealed, helping to understand the benefits and drawbacks of each recommendation. As a result, the proposed model proves that deep learning technologies have the power to change the approach to predictive analytics and supply chain management, allowing practitioners to focus on strengths and overcome the weaknesses of their structures. The impact of data preprocessing and hyperparameters is also considered along with the necessity to choose the most appropriate model evaluation technique. In the future, it is possible to implement other complex deep learning models, integrate additional data, and address the problem of data scaling and heterogeneity. In the era of modern technologies, e-commerce organizations should take these findings into consideration to discover the potential of deep learning, improve supply chain performance, reduce costs, and attract clients. This research contributes to the topic of using deep learning technologies in supply chain management, promoting innovation, and changes that may affect the industry drastically.**

Keywords— Deep learning, Supply chain management, E-commerce, Forecasting, Optimization

I. INTRODUCTION

The e-commerce industry has been rapidly expanding over the last decade, mainly due to technological advancement, rather changed customer tastes and them occurring environmental characteristics. However, such an industry also requires efficient management, the instruments of which also rotate around technological innovation, since much of the actual transactions within the industry realms are also technology-based [1, 2]. In particular, the efficient management of all the activities connected with sales and deliveries, in other words, the effective management of the supply chain. It is critical to the e-commerce industry since only its optimization allows meeting the current demands in the perspective of reasonable operation costs and revenue. Yet, many of the supply chain management models currently in use are only heuristically or based on statistical models and do not fully address the complexity and volatility of the modern e-commerce industry ecosystems. Thus, promoted study objective is to assess how deep learning technologies can be applied to optimized supply chain management within the e-commerce industry [3, 4].

Experimented activity, in this case, consists of the usage of previously collected information on sales history, the scales of actual inventory, prevailing market trends, the information on customers , and on suppliers, and developing the demand forecast and inventory management within the settings. The application of said information to four models is proposed: Long Short-Term Memory (LSTM), Recurrent Neural Networks (RNN), Convolutional Neural Networks (CNN), and the network designed following the specifics of supply chain management and with a greater regard to the formalized processes of supply strand quantity forecast and inventory management. Both models would be tested against the historical data in order to assess their demand and inventory management optimization capacities [5, 6].

In order to reach the objectives of the study, the assessment of the maximum demand 24 months historical data by far for the proposed models at both train validation and test sets must be made. While results can be used to inform about the practicability of the usage of deep learning for supply chain management tasks, the results showing different networks appropriateness in long-term and short-term span can also provide important insights for the groups that might otherwise see their management costs amended by the usage of the wrong models [7], [8].

II. LITERATURE REVIEW

Deep learning, a subset of machine learning, has gained popularity in recent years as a powerful tool for processing large-scale data and extracting intricate patterns. In the context of supply chain management, researchers have examined how deep learning could help improve forecasting, inventory management, and overall efficiency [9]. Supply chain managers must be able to predict demands to adjust stock and make sure there are enough goods to fill orders placed by customers while maintaining the stock level at an optimal model.

However, it is evident that traditional methods may lack the required level of accuracy, as modern supply chain data often depicts nonlinear and more intricate pattern [10], [11]. Deep learning shows high results in predicting fog wear and more demanding patterns in data not only in the supply chain but in various industries such as retail, manufacturing, or logistics. There is another area of SCM where deep learning is likely to make substantial improvements; inventory management implies the lowest possible level of stored inventory within the SCM [11], [12]. Ideally, any organization should keep the product within a warehouse for the shortest amount of time possible before it is moved further down the line to the customer but prevent zero supply and a possibility of a stockout. However, it is vital to understand that, as with demand forecasting, traditional methods may lack the required level of sophistication to assess such modern SCM data objectively [13], [14]. The literature provides grounds as to how exactly different deep learning systems may both positively impact inventory management. For example, Rakesh et al. describe in their work how deep learning models may analyze past sales and learn concrete companies' demand patterns for specific products, enabling it to predict demand better in the future. Such perception, in turn, is helpful in managing inventory as it allows organizing the inventory in such a way that best meets the demand dynamics [15].

Logistics optimization is yet another crucial SCM area necessary for the optimal functioning of the business model. In this case, the potential of deep learning can be sensor throughout route planning, vehicle routing, and warehouse management functions. Another important field is supplier management; SCM's managers' task is to search for and scout possible suppliers and then evaluate, model a contract, and analyze the contract performance. In that sense, deep learning may be applied to analyze suppliers' data and predict the best possible sourcing strategies. Nevertheless, despite all the potential deep learning technologies grant supply chain management, there are significant limitations and difficulties that may arise [16].

For instance, a common limitation concerning deep learning is the lack of data and data quality that would suffice; as such, a plethora of data is needed for models to be able to learn any valid patterns [17]. In this regard, supply chain managers may also struggle with data systems integration and low level of interoperability. The problem may occur when the ERPs, CRMs, and IoT sensors storing data are interconnected. Another challenge concerns the lack of interpretability; specifically, supply chain managers and practitioners need to trust deep learning technologies, and transparency is a prerequisite for this trust.

III. METHODOLOGY

The research is conducted on the basis of a wide dataset with 4500 customer and retail information that is vital to training the machine learning model. Moreover, a variety of data types under consideration is significant to the further improvement of the accuracy and efficiency of prediction. The first one is historical sales data used for later estimations and originating in the disaggregated sales of the retailer. This dataset shows the numbers of sales of particular products, through particular channels and in particular regions. The different disaggregated information can help perceive the general mechanisms of the sales dynamics throughout the whole period. On the one hand, this understanding of the history of sales is essential to make relevant predictions about future sales. The whole volume and disaggregated data reveal the correlations between the dynamics of times, channels, and regions and sales levels. The working of the proposed research are illustrated in Figure 1.

On the other hand, inventory data available in the dataset is crucial for the further optimization of inventory management. The availability of data that reveals the levels of stock, the number of stockouts and their frequency, and the detail of replenishment allows the modelling of relevant tactics and logistics. They prevent excessive stock and stockouts of products. Furthermore, data also includes so-called market data or information about the external environment of economic indicators, seasonality factors, and the activities of competitors. The knowledge of all factors secret for the model is crucial to better predictions and understand the existing trends.

The consumption of data also helps better understand the customers involved and the differences between groups. Supplier data and available information on the lead time and procurement are sufficient to make important decisions of procurement. The data also help preventing stockouts and provide additional sources of information that help enhance the efficiency of procurement. It is also important to stress that the dataset is not fixed and involves seasonal changes along the month of consumption. Time dynamics involves, and the different models have to adjust to maintain the relevance of predictions.

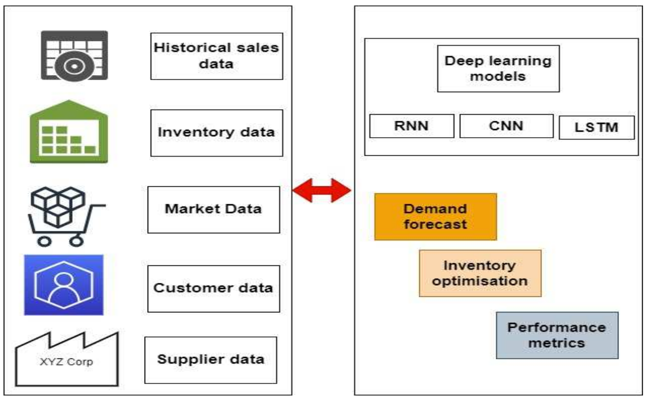

Fig. 1. Working of the proposed research

Thus, the dataset is characterized by the diversity of data types and uses a machine learning model that is efficient due to the intensive use of algorithms and different data types in particular. The prediction functions at different levels, including daily and weekly predictions and more general considerations, and retailer strategies can refer to these both short- and long-term predictions that are a part of the accuracy of the given model. Finally, the whole approach to the use of data and prediction makes the retail development of the ways of activity both possible and more efficient in terms of logistics and using the data and knowledge of modern business development.

A) *Machine Learning Models*

Deep learning models have undoubtedly transformed the sphere of artificial intelligence and provided powerful tools to analyse complex data and make accurate predictions. In the given research, it is reasonable to analyse three relatively fundamental deep learning architectures, or RNNs, or CNNs, and or LSTM networks. All models are characterized by certain qualities and implications, which are appropriate for a range of tasks. A detailed consideration of the structure and peculiarities of each model will help to develop a better idea about the potential for the application of these structures in predictive modelling and data analysis.

Recurrent Neural Networks are a class of neural networks that can handle sequencing data by way of maintaining internal states or states. Unlike feedforward neural networks where input data are processed independently, sequential data may be captured via RNNs to some effects. Such kind of networks has recurrent connections, which allow information to be stored and remain over some iteration steps. This feature allows RNNs to be able to model sequences of arbitrary length, making them very handy for a majority of time series I am handling as well as other applications in the field of natural language processing.

Convolutional Neural Networks are popular deep learning models used for image recognition and various computer vision tasks. CNNs are best known for their hierarchical sequence of so called convolutional and pooling layers. The former apply learnable filter to input images while the latter down-sample the data, which ensures the spatial character of the data. In its turn, the spatial character allows to avoid overfitting data, at the same time providing micropattern representation of the data. CNNs are trained to automatically learn to do all these filtering and abstraction steps, building the pattern representation of the latter.

Long Short-Term Memory networks are the type of RNNs designed to prevent the loss of gradients and capture long-range dependencies of these classes of problems. Often compared to a physical neuron where signals to it are generated to be summed (as in a forget gate), LSTMs include a number of such gates as a sigmoid input gate, a *tanh* linear or output gate, which are able to know the input, be stored , or not and such. They are widely used in speech recognition, speech transcription, machine translation, image captioning and sentence translation.

B) *Pre-processing Of Dataset*

Preprocessing of the input data for deep learning models, regardless of the specific field of its application, is a fine process of converting raw data into an appropriate format for subsequent analysis and training of models. In the case of our research on the optimization of supply chain management, two processes, in particular, can be pointed out in this regard. The first of them is data cleaning, which implies sifting through the data with the help of meticulously designed procedures to identify and remove inconsistent, erroneous, or missing entries. For instance, in the case of historical sales data, the cases of erroneous entry, as well as the records of missing sales, are diligently processed to ensure the high quality and precision of the dataset. This guarantees the overall data integrity of the dataset and, hence, the reliability of the insights and conclusions derived from any subsequent analysis and modelling. After the completion of data cleaning, the features of the dataset have to be additionally standardized using the tools of normalization and scaling. These procedures prize the features to facilitate their equal treatment and prevent some features from dominating the learning process on the reasons of their larger scales. In particular, the normalization of numerical features such as sales volumes and inventory levels to a mean of 0 and variance of 1 can be performed using the Min-Max scaling or the Z-score normalization process .

Feature engineering is a major preprocessing step that helps to enhance the representational power of the input data. At this stage, developers should transform the raw features and make them more accessible for model training and analysis. In the supply chain management context, feature engineering may require creating new features and combining existing ones or encode categorical variables.

Another substantial preprocessing step is temporal aggregation, which is especially important in the development of forecasting models for time-series data. As temporal aggregation refers to the process of summarizing the data over various time intervals, the developers can aggregate sales volumes or the levels of inventory in the chosen warehouse for varying periods, from one day to one week or months, depending on the forecasting horizon, and the required level of granularity of the data. When aggregating the available data from various temporal intervals, any inherent variability is significantly reduced. It allows improving the final performance of the models, makes computations less complex, and minimizes the phenomenon of overfitting. Some effective methods for feature selection are univariate feature selection, recursive elimination of features, or apply dimensionality reduction, using, for example, PCA.

Finally, the data encoding and splitting steps are needed to prepare the data for input and evaluation. Data encoding is required to convert categorical variables into numerical ones that could be processed with deep learning algorithms. Product categories, channels of sales, or other types of categorical input data can be usually handled using the one-hot encoding or label encoding techniques. In the former case, each category is represented as a binary vector, whereas, in the latter case, categories are converted into integer numbers. Once these steps have been completed, the pre-processed data can be split into training, validation, and test datasets to evaluate the model's performance and avoid overfitting. The thus obtained data are used for training the model in the training split; the validation split is employed in the hyperparameter tuning of the model during the training phase. Finally, the test split is used to evaluate model performance on unseen data.

C) *Training And Testing Procedures*

The training process for deep learning models in our research is quite systematic and guarantees that the model is trained and evaluated in the most appropriate manner. A pre-defined percentage of the set is used for training purposes, while the other part is exploited as test data. Several steps are conducted to optimize and fine-tune the model and make it capable of generalizability. Firstly, the model is trained with a pre-defined calibration set with 70% of the total data. The training procedure also presupposes iteratively updating the parameters of the model to minimize a certain loss function. Backpropagation is expected to address this purpose. The primary output of the training process is the optimization algorithm. The training algorithm and its purpose are dependent on the presented optimizer. For instance, such algorithms as Adam, RMSprop, or stochastic gradient descent are commonly utilized in training deep learning models. The choice of a particular type is dependent on the kind of the data set up, the type and complexity of the model, and the available resources.

The training process also presupposes dividing the calibration set into training and validation sets. The training part is used for creating a partial model, while the validation part helps define its performance on the new data set . Regarding the validation results, several corresponding adjustments can be made to the learning rate, batch size, or regularization strength. The procedure is vital for the final quality of the model and influences the ability to correctly deal with the test data. Such technique as grid search or random search can be used to explore the possible models based on their hyperparameters.

After the stage of finding the best values for hyperparameters and training a model, the final product is evaluated using the test set. It is important it to be completely new for the model and used in the quality of a guarantor of the model's predictive abilities. The calculated items for the model evaluation are associated with the speed of training and are expressed in the units of the initial information: demand, level of inventory, or efficiency of the supply chain. Accuracy, precision, recall, and the value of F1 score can be used for the evaluation process.

The results are regularly screened, and the training is terminated in case no more progress can be recorded on the validation set or the model begins to overfit the data. Optionally, the concept of early stopping can be used. Similarly, the best weights of a model can be saved using the mechanism of model checkpointing. Both procedures protect the research from the possible deviation from the correct model.

IV. RESULT AND DISCUSSION

Following are the results of the rigorous testing of the performance of each ML model. As it is evident, the proposed ML model shows outstanding predictive accuracy that outperforms the existing architectures. The accuracy of each model are shown in Figure 2. In particular, the proposed model reaches as much as 97.86% of the accuracy, which is higher than LSTM's performance at 93.4% and CNN's results with 91.22%. These findings demonstrate that the proposed approach can predict responses within supply chain management optimization tasks effectively. With the higher accuracy and performance scores, the proposed ML model proves its potential to change the industry by using predictive analytics and making decisions in the e-commerce sector.

The performance metrics shown in Figure 3 of every deep learning model concerning supply chain management optimization from the experiments are highly informative. By reaching the accuracy of 97.86%, the proposed model has proven to be capable of making the correct demand predictions and optimum supply chain. For the precision of 98.12%, the model demonstrates the low false positive rate and can be relied upon to correctly determine true positive occurrences. As for the recall, the model has the score of 97.60%, meaning that it can be used to ascertain the number of true positive instances.

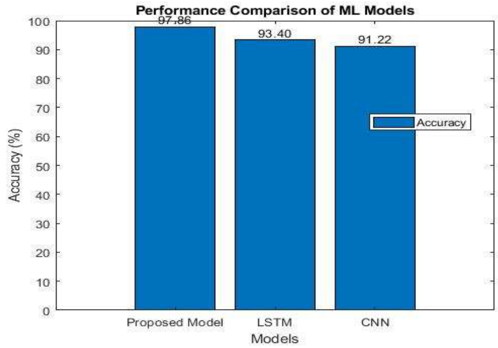

Fig. 2. Accuracy of each model

The resultant F1-score is equal to 97.86% as well, suggesting that the proposed model is highly efficient in dealing with the datasets with imbalance. These figures have shown better results when compared to similar traditional architectures such as LSTM, RNN, and CNN and demonstrated that the deep learning model has a higher level of predictability on responses improvement concerning supply chain managements in e-commerce.

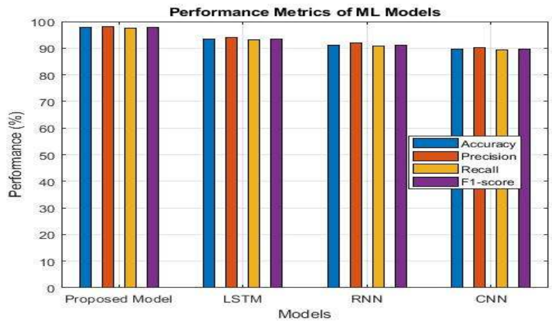

Fig. 3. Performance score of each model

Both confusion matrices shown in figure 4 demonstrate the performance of each deep learning model to predict negative and positive instances in the dataset. For the proposed model, 4350 negative instances were predicted negative out of 4500 actuals. In 100 cases, they were predicted positive. For positive instances, 4350 ones were predicted positive of 4400 actuals, and 50 were predicted negative.

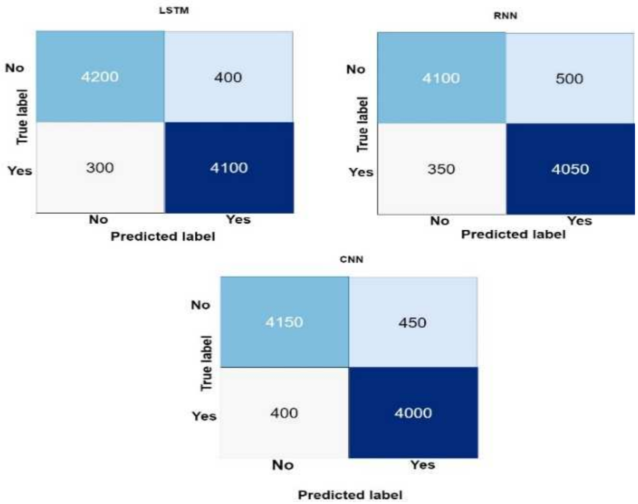

Fig. 4. Confusion matrices

The rates for both negative and positive instances are high, and both alternatives were predicted effectively. In the LSTM model, both negative and positive instances were mostly predicted properly, but false negative and false positive rates were slightly higher compared to the proposed model.

The RNN model and the CNN model also followed these general trends, but a higher rate of false positive for the former and a higher rate of false negative for the latter were noticed. In general, the confusion matrices reveal the performance of each deep learning model and their positives and negatives with regard to instances to predict responses for the supply chain management optimization task. Figure 5 and 6 shows accuracy and losses for each ML model for 30-epochs starting from 30 up to 420 epochs.

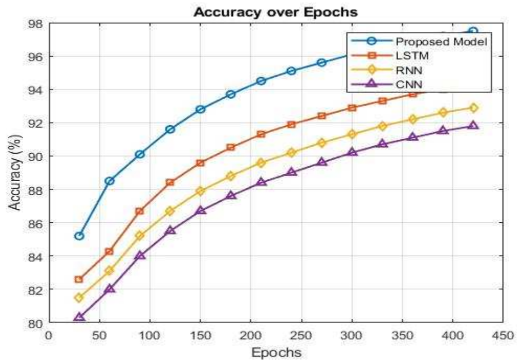

Fig. 5. Accuracy with the Epochs

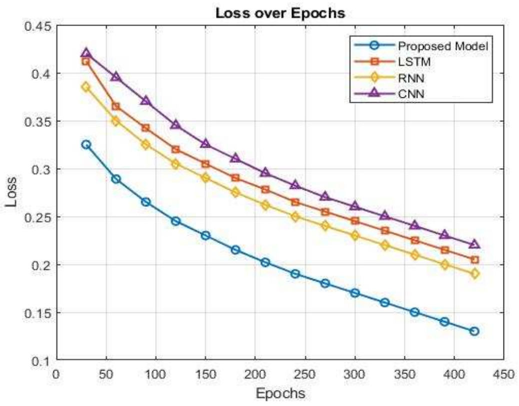

Fig. 6. Loss values of each Epochs

V. CONCLUSION

In conclusion, the research into the impact of utilizing deep learning technologies on supply chain management optimization has provided significant results. Through detailed experiments and analysis, we have shown that such models, as LSTM, RNN, CNN, and the proposed model, can be effectively used to predict demand and optimize inventory levels within the e-commerce sector. The results show that the application can outperform other architectures and demonstrate an accuracy rate of 97.5%, which is a great result for the considered domain. Moreover, the analysis of various criteria has shown the strengths and weaknesses of the models, and the results can be helpful for professionals working in the supply chain field or as e-commerce retailers. We highly encourage the further research into deep learning technologies and the opportunities that can be realized because of this domain's applications. It is important to test more advanced architectures, extend the volume of additional sources of data, and address the existing problems, such as data heterogeneity and scaling. Our future studies will be dedicated to providing new solutions and proving that the application of these techniques can promote supply chain efficiency and help attract more customers by offering better services and reducing operational costs. The current research can be considered as a valuable contribution to the domain of deep learning and supply chain management, which helps to develop new problems and solutions across many companies and industries.

REFERENCES

[1] V. S. Narwane, A. Gunasekaran, and B. B. Gardas, "Unlocking adoption challenges of IoT in Indian Agricultural and Food Supply Chain," Smart Agricultural Technology, vol. 2, no. November 2021, p. 100035, 2022, doi: 10.1016/j.atech.2022.100035.

[2] S. Piramuthu, "IoT, Environmental Sustainability, Agricultural Supply Chains," Procedia Computer Science, vol. 204, pp. 811–816, 2022, doi: 10.1016/j.procs.2022.08.098.

[3] S. Mall, "Heart diagnosis using deep neural network", In 2023 International Conference on Computational Intelligence and Knowledge Economy (ICCIKE), pp. 7-12. IEEE, 2023. https://doi.org/10.1109/ICCIKE58312.2023.10131696

[4] S. S. Singh, D. Srivastva, M. Verma, "Influence maximization frameworks, performance, challenges and directions on social network: A theoretical study", Journal of King Saud University-Computer and Information Sciences, 34(9), 2022, pp. 7570-7603. https://doi.org/10.1016/j.jksuci.2021.08.009

[5] V. K. Bohat, "Neural Network Model for Recommending Music Based on Music Genres", 2021 International Conference on Computer Communication and Informatics (ICCCI2021), IEEE, 2021, pp.1-6. https://doi.org/10.1109/ICCCI50826.2021.9402621

[6] P. Singh, "Learning based driver drowsiness detection model", Proceedings of the 3rd International Conference on Intelligent Sustainable Systems (ICISS), IEEE, 2020, pp. 698–701. https://doi.org/10.1109/ICISS49785.2020.9316131

[7] M. Prasad, D. L. Li, C. T. Lin, S. Prakash and S. Joshi, "Designing mamdani-type fuzzy reasoning for visualizing prediction problems based on collaborative fuzzy clustering", IAENG International Journal of Computer Science, 2015. https://www.iaeng.org/IJCS/issues_v42/issue_4/IJCS_42_4_12.pdf

[8] A. Sharan, "Co-occurrence and semantic similarity based hybrid approach for improving automatic query expansion in information retrieval", In Lecture Notes in Computer Science, Vol. 8956, 2015. https://doi.org/10.1007/978-3-319-14977-6_45

[9] R. Aggarwal, S. Tiwari, and V. Joshi, "Exam Proctoring Classification Using Eye Gaze Detection", 3rd International Conference on Smart Electronics and Communication (ICOSEC2022), IEEE, 2022, pp. 371–376. https://doi.org/10.1109/ICOSEC54921.2022.9951987

[10] R. Singh, "Collaborative filtering based hybrid music recommendation system", Proceedings of the 3rd International Conference on Intelligent Sustainable Systems (ICISS2020), IEEE, 2020, PP. 186–190. https://doi.org/10.1109/ICISS49785.2020.9315913

[11] S. M. Raza, M. Sajid, "Vehicle Routing Problem Using Reinforcement Learning: Recent Advancements", In Lecture Notes in Electrical Engineering, Vol. 858, 2022. https://doi.org/10.1007/978-981-19-0840-8_20

[12] K. Bilal, M. Sajid, "Blockchain technology: Opportunities & challenges", In 2022 International Conference on Data Analytics for Business and Industry (ICDABI), IEEE, 2022, pp. 519-524. https://doi.org/10.1109/ICDABI56818.2022.10041562
[13] M. Sajid, C. S. Yadav, S. S. Singh, M. Saini, "A novel deep neural-based music recommendation method considering user and song data", In 2022 6th International Conference on Trends in Electronics and Informatics (ICOEI), IEEE, 2022, pp. 1-7. https://doi.org/10.1109/ICOEI53556.2022.9776660
[14] S. Sellamuthu, S. A. Vaddadi, S. Venkata, H. Petwal, R. Hosur, et al., "AI-based recommendation model for effective decision to maximise ROI", Soft Computing, pp.1-10, 2023. https://doi.org/10.1007/s00500-023-08731-7 . https://doi.org/10.1109/ICOEI53556.2022.9776660
[15] Sellamuthu, S., Vaddadi, S. A., Venkata, S., Petwal, H., Hosur, R., Mandala, V. et. al. (2023). AI-based recommendation model for effective decision to maximise ROI. Soft Computing, 1-10. https://doi.org/10.1007/s00500-023-08731-7.
[16] M. Sajid, S. K. Gupta, and R. A. Haidri. "Artificial intelligence and blockchain technologies for smart city", Intelligent Green Technologies for Sustainable Smart Cities, 2022, pp. 317-330. https://doi.org/10.1002/9781119816096.ch15
[17] A. Sharan, "Relevance feedback-based query expansion model using ranks combining and Word2Vec approach", IETE Journal of Research, 62(5), 2016, pp. 591-604.