2025 4th International Conference on Distributed Computing and Electrical Circuits and Electronics (ICDCECE)

# Sales Forecasting in Cross Border E-Commerce using Convolutional Neural Network- Gated Recurrent Unit

1st Yanfang Zhang
*School of Business Administration*
*Harbin Cambridge University*
Harbin, China

2nd Kai Fu*
*School of Business Administration*
*Harbin Cambridge University*
Harbin, China
13624515934@163.com

***Abstract*—Cross-border e-commerce has expanded quickly, precise sales forecasts are essential to maximizing marketing, logistics, and inventory control. Because commerce generates so much deal data every second, predictive analysis has never been more important for logistics management. To improve supply control and customer service, e-commerce companies are increasingly using AI technologies to improve forecasts. To improve the accuracy of sales forecasting, this study suggests a Convolutional Neural Network-Gated Recurrent Unit (CNN-GRU) hybrid model. Where GRU records temporal dependencies and sequential patterns, CNN derives spatial dependencies from sales data. The model performs better, since it is trained using historical sales data, product features, and market trends from Amazon and Flipkart data. In the end, the suggested method optimizes business operations in cross-border e-commerce by improving demand forecasting and lowering instances of stockouts and overstock. With an accuracy score of 98.61%, precision score of 98.22%, recall score of 98.12%, and F1 score of 98.13%, the experimental results indicate that this implementation-based approach performs better. Enhancing system adaptability and using control-based techniques for increased sales prediction.**

***Keywords—CNN-GRU hybrid model, sales prediction, convolutional neural network, gated recurrent unit, e-commerce.***

## I. INTRODUCTION

As digital technologies increasingly penetrate daily life and global living standards continue to increase, cross-border e-commerce is quickly evolving into a significant force in the international retail market [1]. This growth in global online consumption has spurred considerable development in the logistics industry, but it has also brought with it a multitude of new challenges—most notably with regard to the consistent and effective delivery of temperature-sensitive (cold chain) goods [2]. Unpredictable fluctuations in order quantities, unpredictable price changes, and irregular distribution patterns can create instability in the supply chain. Left unresolved, such disruptions can actually freeze the entire e-commerce system [3].

To cope with these complexities and ensure a responsive, functioning supply chain—particularly during times of imbalance—it is critical to institute sophisticated inventory control and supply management systems. A supply chain might be imagined as a huge, interdependent ecosystem comprising people, businesses, infrastructure, technologies, and operational procedures all working together to support the life cycle of a product. One effective way to optimize logistics is through proactive redistribution of stock among strategically positioned international warehouses [4]. This strategy can not only minimize delivery latency but also increase overall customer satisfaction.

However, it is getting harder to manage such a system with manufacturing and distribution processes spreading across various continents [5]. Cross-border e-commerce companies have to deal with an extensive set of logistics processes that include procurement, transportation, customs clearance, and product inspection before the products even reach the final users[6]. For this complex coordination, big data analytics has turned out to be an indispensable instrument. It helps companies to identify patterns and trends from huge repositories of e-commerce transactional information, enabling better supply chain decision-making [7].

Over the past few years, there has been a distinct trend towards quantitative forecasting models, which are preferred for their objectivity and capacity to remove human bias. Quantitative forecasting models use large amounts of historical data and sophisticated algorithms to detect patterns and provide predictive insights [8]. Quantitative forecasting tends to produce more uniform short-term results compared to qualitative methods that depend on expert judgment—particularly when managed by professionals with high analytical capabilities [9].

Though, these innovations have not diminished yet the high heterogeneity and amplitude of cross-border transaction data remain a challenge in precise demand forecasts. Artificial intelligence is now widely being adopted for inclusion in forecast processes[10]. Despite many AI-oriented models proposed so far, just a few are proven to report high accuracy forecasting future sales trajectories [11]. In order to bridge this deficiency, the current research suggests a new hybrid model integrating Convolutional Neural Networks (CNNs) and Gated Recurrent Units (GRUs) to boost predictive accuracy and enhance decision-making in global e-commerce activities [12].

### A. Contribution of the Research

1. The study demonstrates CNN-GRU networks' effectiveness over traditional statistical models in analyzing sales data, forecasting sales, and detecting activities with precision.
2. The incorporation of an adaptive evaluation mechanism that uses deep learning to automatically evaluate sales data is a significant advance. The module helps businesses improve their methods and enhance learning results by providing objective, real-time predictions on a variety of sales-related topics.
3. A scalable and automated method of prediction is provided by the study's suggested integration of AI-based sales assessment into enterprise systems.

979-8-3315-3366-3/25/$31.00 ©2025 IEEE

2025 4th International Conference on Distributed Computing and Electrical Circuits and Electronics (ICDCECE) | 979-8-3315-3366-3/25/$31.00 ©2025 IEEE | DOI: 10.1109/ICDCECE65353.2025.11035385

*B. Structure of the paper*

The paper operates as follows: An overview of current e-commerce sales prediction methods and the challenges they present is provided in Section 2. Section 3 provides a discussion of the dataset, preparation methods, and the suggested CNN-GRU framework. Evaluation of performance, important metrics, and comparisons to current models are presented in Section 4. The study closes in Section 5 with a summary of the primary results with suggestions for further research.

## II. Related Works

Ji et al. (2019) [13] developed a three-stage XGBoost-based model to enhance sales forecasting for cross-border e-commerce enterprises. The three main components of their strategy were feature selection, which determined which variables had the greatest impact on sales; data preprocessing, which cleaned and normalized the dataset to eliminate irregularities; and model optimization, which adjusted hyperparameters for improved predictive performance. Their methodology greatly increased the accuracy of sales forecasts by incorporating these components. The study showed how effective ensemble learning techniques—in particular, XGBoost—are in processing intricate, non-linear sales data. Their method produced more reliable results with greater generalization capacity than conventional forecasting models. The results demonstrated how well machine learning works to improve sales forecasting for companies that operate in volatile and changing markets.

Cai et al. (2020) [14] explored financing risks in enterprise supply chains using a backpropagation neural network. By examining a variety of economic factors, their model sought to forecast financial risks and assist companies in minimizing any supply chain interruptions. The study successfully modelled intricate financial linkages and relationships throughout the supply chain by utilizing BPNN's self-learning capabilities. Their method of predicting possible financial hazards before they may affect operations was very helpful in lowering business uncertainty. The benefits of AI-driven financial forecasting for enhancing supply chain resilience were also emphasized in the study. Their deep learning-based approach offered better forecast accuracy and quicker reaction times than conventional risk assessment techniques. Their work helped firms improve their risk management plans by advancing the expanding field of AI-powered financial decision-making.

Alon et al. (2001) [15] compared artificial neural networks (ANNs) with traditional statistical methods for retail sales forecasting. Their study investigated the effectiveness of several forecasting methods in spotting trends in consumer demand and non-linear sales patterns. The study discovered that when it came to identifying intricate, non-linear relationships in sales data, ANNs performed better than more conventional approaches like time series models and regression-based procedures. The researchers underlined how crucial neural networks are for sales forecasting, especially in sectors where demand varies. ANNs were able to adjust to shifting market conditions more effectively than conventional techniques because of their capacity to learn from vast datasets. Their results demonstrated how machine learning is becoming more and more important in retail demand forecasting and how AI models may improve business decision-making by providing data-driven insights.

Di Pillo et al. (2016) [16] applied support vector machines (SVM) to sales forecasting under promotional conditions. In order to assist firms in modifying prices in response to market demand, their model sought to forecast sales changes brought on by discounts, advertising campaigns, and seasonal promotions. Their method addressed non-linear correlations in sales data by integrating machine learning techniques, which increased demand forecasting accuracy. In order to improve business revenue and inventory management, the study showed how SVMs may be utilized to identify and react to promotional fluctuations. Their SVM-based method was more flexible in response to changing market conditions than conventional forecasting methods. The study emphasized how crucial AI-powered forecasting tools are to enhancing pricing optimization tactics for companies in a range of sectors.

Lu and Kao et al (2016) [17] suggested a clustering-based sales forecasting system that makes use of ensemble linkage techniques and extreme learning machines (ELM). Their study focused on predicting computer server sales, an industry noted for its demand volatility and complex sales cycles. By initially clustering sales data, they were able to aggregate comparable sales trends and apply focused machine learning models to increase forecast accuracy. The ELM algorithm was then used to forecast future sales patterns because of its strong generalization ability and quick learning speed. Comparing clustering-based methods to conventional forecasting techniques, their results demonstrated a considerable improvement in model performance. The study came to the conclusion that better demand forecasts and commercial decision-making could result from combining clustering with machine learning. In particular, for rapidly changing industries like technology and hardware manufacturing, this study offered insightful information on optimizing sales forecasting.

Dai et al. (2015) [18] introduced a clustering-based sales forecasting framework using support vector regression (SVR). Their research demonstrated the effectiveness of combining machine learning models with clustering in handling complex sales datasets, and they highlighted the advantages of SVR in capturing non-linear sales trends, making it a valuable tool for demand forecasting in dynamic industries. Their study focused on forecasting computer server sales, where demand is influenced by multiple external factors, including market trends and technological advancements. They began by using clustering techniques to segment sales data into distinct groups based on shared characteristics, which allowed the SVR model to learn patterns more efficiently.

Chen and Lu et al (2017) [19] combined clustering and machine learning techniques for sales forecasting in the computer retail sector. In order to increase forecasting accuracy, their methodology first divided sales data into pertinent clusters and then applied predictive models. Their method made it possible for machine learning algorithms to concentrate on particular sales patterns by clustering related data points together, which produced forecasts that were more accurate. According to their research, hybrid models perform better than conventional forecasting techniques, particularly when working with sizable datasets that exhibit a range of sales patterns. The researchers stressed that using clustering approaches improves the interpretability and efficiency of the model. Their results demonstrated how crucial AI-powered forecasting tools are for business decision-making, especially in sectors where customer demand is erratic. The study showed how inventory management and sales forecasting accuracy might be enhanced by integrating clustering and predictive analytics.

Freij et al. (2021) [20] developed a deep learning model for digital sales forecasting and optimization in e-commerce. Their study looked at how artificial intelligence might be used to boost digital sales performance, optimize marketing tactics, and improve demand forecasting. Their approach examined enormous volumes of previous sales data using deep learning algorithms to find patterns that conventional models frequently overlook. The study demonstrated how deep learning approaches might greatly improve sales projections and business efficiency, highlighting the transformative potential of AI in intelligent e-commerce platforms. According to their findings, e-commerce companies may make better decisions by using AI-driven forecasting tools, which offer more precise and flexible predictions. According to the study's findings, online merchants might enhance inventory control, target marketing campaigns, and boost sales by implementing data-driven tactics by incorporating AI-based sales optimization techniques.

## III. PROPOSED METHODOLOGY

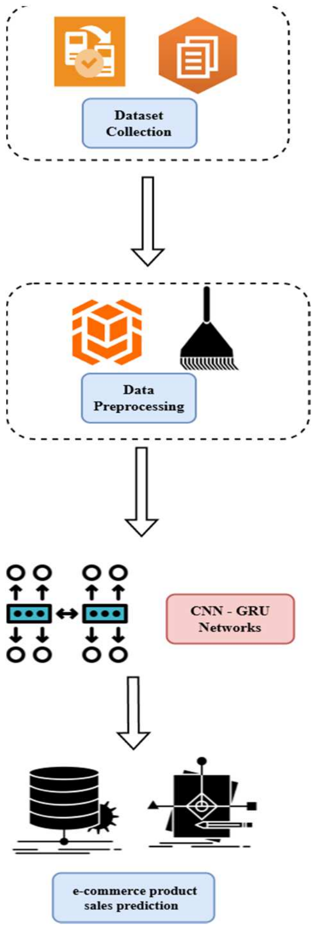


Fig. 1. Overall Architecture of Sales prediction

The Figure 1 illustrates the overall proposed architecture, where dataset is taken to pre-processing technique and that is trained with a hybrid-based algorithm for better sales prediction in ecommerce domain.

### *A. Materials and Methods*

Datasets from large e-commerce websites such as Amazon and Flipkart are used in sales forecasting research because they have extensive and diverse transaction histories. Such datasets include data like product categories, prices, customer ratings, order quantities, seasonal patterns, and promotional effects. By examining historical sales records of these websites, researchers can discover influential factors of demand like discounts, festive periods, and buying habits. These datasets assist in preparing learning models, CNN-GRU to forecast future sales trends more accurately. The outputs obtained from Amazon and Flipkart sales information assist enterprises in making the best inventory, pricing, and promotional decisions, which ultimately improve profitability and customer satisfaction. Thus, preprocessing is important to train the data.

### *B. Data Preprocessing*

The preprocessing process starts with data cleaning, where missing values are treated using imputation methods or discarded if possible. Duplicate records and unnecessary features are removed. Second, data transformation is conducted, in which categorical features such as product classes, brands, and segments are transformed into numeric values with the help of one-hot encoding. Continuous variables like prices and discounts are scaled by Min-Max Scaling or Standardization for normalization. This systematic preprocessing improves the performance Model CNN-GRU, resulting in more accurate sales forecasts.

### *C. CNN-GRU Networks*

*1) CNN Networks:* Complex deep learning models are Convolutional Neural Networks (CNNs) that are able to learn directly from raw data inputs. Although historically linked to image recognition tasks, CNNs have also found applications in areas other than images, including sales forecasting, time series, audio classification, and signal processing, by identifying patterns and trends in structured data. Figure 2 represent the architecture of CNN.

CNNs are specifically good at finding significant patterns within datasets and hence are apt for tasks such as identifying sales behaviors and predicting item demand. Their unique structure helps them excel beyond most regular neural networks, given their capability of automatically learning spatial hierarchies of features.

*a) Convolutional Layer:* This initial layer is where a CNN begins. It detects important patterns by scanning the input data using filters or kernels. Each filter traverses the input, calculating dot products to generate feature maps that emphasize particular patterns or features. This, mathematically defined as Equation (1), enables the network to learn localized features which are important for subsequent processing in deeper layers.

$$W - F + 2P/S + 1 \tag{1}$$

*b) Pooling layer:* This layer's main goal is to make the dense feature map as small as possible to save computing costs. This is achieved by decreasing the relationships between sections and working on their own on every element map.

Depending upon the method used, there are several types of Pooling operations, here the Equation (2).

$$W - F/S + 1 \tag{2}$$

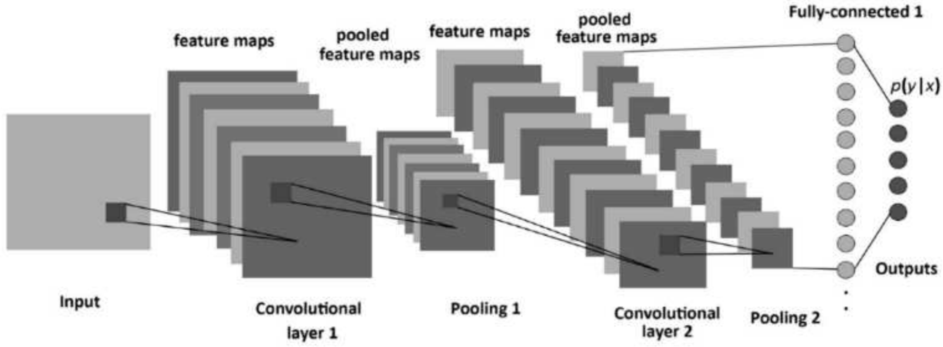


Fig. 2. Convolutional neural networks architecture

*c) Fully connected (FC) layer:* Consisting of weights, biases, and neurons, this layer links neurons between two layers. These layers comprise the last few layers of a CNN architecture and are typically placed before the output layer.

Another common CNN feature is the Dropout layer. The Dropout layer is a mask that, while leaving all other neurons unaltered, blocks some neurons' contributions to the layer that follows.

Whether or not to activate a neuron depends on its activation function. As a result, it will decide whether or not the neuron's contribution to the network is important for making predictions. There are several well-liked activation functions, such as the Sigmoid, Softmax, ReLU, and tanH functions. Each of these functions has a specific purpose. Sales data is predicted using softmax activation. The softmax multinomial is commonly used as a neural network's output is normalized to a probability distribution across the anticipated output class using the final activation function.

*2) GRU Network*: A simplified version of the recurrent neural network (RNN) architecture, the Gated Recurrent Unit (GRU) conceptually resembles the Long Short-Term Memory (LSTM) model. Similar to LSTM, GRU works well with sequential data, allowing for the retention of pertinent information over long time steps and the removal of less crucial information. However, because they have a more compact design and employ fewer parameters than LSTMs, GRUs are renowned for their simplicity and lower computational overhead, which makes them quicker and simpler to train.

The way that each model handles memory management is where GRU and LSTM diverge most. In order to regulate the information flow, LSTMs use three different gates: input, forget, and output gates. They also rely on a dedicated memory cell. On the other hand, GRUs combine these features into two main parts: the update gate and the reset gate. GRUs use a "candidate activation vector" to update their internal state rather than keeping a separate memory cell. The reset gate regulates how much of the prior data is discarded, while the update gate determines how much of the new candidate vector enters the subsequent hidden state.

The GRU processes each element in a sequence one at a time. It uses the previous hidden state and the current input to dynamically update the hidden state. In the process, the model integrates the input and previous hidden state to compute a candidate activation vector. The new hidden state, which reflects the model's memory and comprehension at that specific time step, is then created using this vector and the update gate.

Key components of the GRU architecture include:

- Input Layer: This layer feeds the sequential data—such as word sequences or time-series values—into the model.
- Hidden Layer: This core layer performs the recurrent computation. At each time step, it refines the hidden state, which serves as a dynamic memory reflecting the influence of past inputs.
- Reset Gate: The amount of the prior hidden state that should be "forgotten" or reset is controlled by this gate. It uses the current input and the previous hidden state to create a gating vector with values between 0 and 1.
- Update Gate: It regulates how much of the candidate activation vector contributes to the updated hidden state, thus allowing the model to balance between memory retention and adaptation to new information.

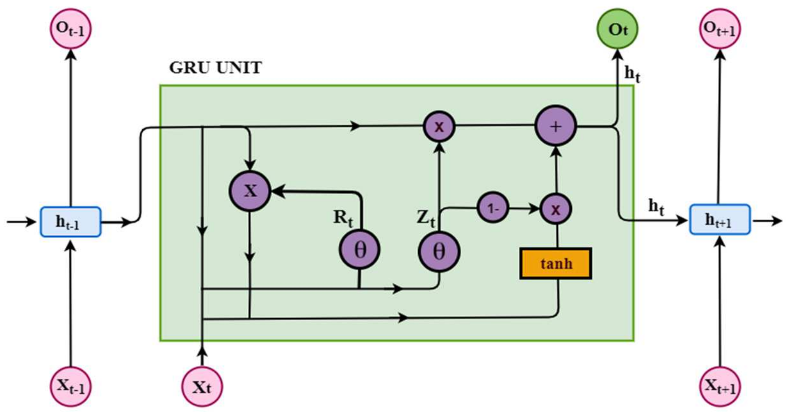


Fig. 3. GRU Architecture

GRUs can efficiently capture long-term dependencies in sequential data while preserving computational simplicity thanks to this effective gating mechanism making them a popular choice in natural language processing, time-series prediction, and other tasks involving ordered inputs. Figure 3 represent the architecture of GRU.

The update gate in a Gated Recurrent Unit (GRU) is responsible for figuring out how much of the updated hidden state needs to be supplemented with the candidate activation vector. As control signals, it produces a vector of values ranging from 0 to 1. The amount of new information that should be added to the hidden state in the future is determined by these values, which are derived from the current input and the previous hidden state.

*a) Candidate Activation Vector:* This vector is basically an updated version of the last hidden state, which is updated by the reset gate to erase irrelevant past information. It combines this filtered memory with the current input and applies a hyperbolic tangent ($tanh$) activation function. The $tanh$ function is used to ensure that the candidate activation values are between -1 and 1, so that there is smooth gradient flow during training.

*b) Output Layer:* Finally, the final hidden state is passed on to the output layer after processing the whole input sequence. This is where the model's prediction is produced, which may be of different types based on the task—e.g., a scalar value in the case of regression, a vector in the case of multi-output problems, or a probability distribution for classification problems.

Reset gate $(r)$ and update gate $(z)$ are computed on the basis of both the present input at time step (x) and past hidden state at time step $t_{-1}$ $(h_{t-1})$, with the Equations (3) and (4) provided. The two gates control the flow of information in tandem within the GRU, controlling memory recall to adaptation to novelty on a proportionate basis.

$$r_t = sigmoid(W_r * [h_t - 1, x_t] \quad (3)$$

$$z_t = sigmoid(W_z * [h_t - 1, x_t]) \quad (4)$$

Where $W_r$ and $W_z$ are weight matrices that are learned during training.

The current input $x$ and a modified version of the previous hidden state that is "reset" by the reset gate Equation (5) are used to calculate the candidate activation vector $h_t\sim$.

$$h_t\sim = tanh(W_h * [r_t * h_t - 1, x_t]) \quad (5)$$

Where $W_h$ is another weight matrix.

The candidate activation vector and the prior hidden state, weighted by the update gate Equation (6), are combined to create the new hidden state, $h_t$.

$$h_t = (1 - z_t) * h_t - 1 + z_t * h_t\sim \quad (6)$$

*3) Hybrid Learning model:* The CNN-GRU combined model integrates Convolutional Neural Networks (CNNs) and Gated Recurrent Units (GRUs) (Figure 4) to take advantage of both spatial and temporal dependencies in data. CNNs are used to extract local patterns and hierarchical features and are thus effective in identifying trends from structured data such as sales records. The features so extracted are forwarded to GRU layers, which are adept at dealing with sequential dependencies by capturing useful past information while avoiding vanishing gradient problems. This fusion improves predictive power, particularly for time-series prediction tasks like forecasting cross-border e-commerce sales. By combining the two models, companies can unlock more profound consumer behaviour insights and streamline decision-making processes.

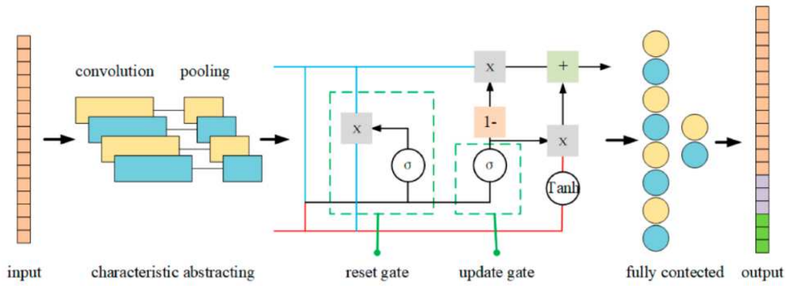


Fig. 4. Hybrid learning model

## IV. Results And Discussion

### A. Implementation Details

TensorFlow, Keras, matplotlib, numpy, and pandas were some of the libraries used for training, evaluating, and displaying the results. The model was developed using Python 3.8. With an Intel i7 CPU (3.2 GHz), 16 GB of RAM, and an NVIDIA Tesla GPU, the experiment's workstation has the processing capacity needed to train deep learning models on large image datasets.

### B. Evaluation Metrics

The model's performance was assessed using key evaluation criteria, including F1-score, recall, specificity, accuracy, and precision. To show the efficacy of the suggested framework and to emphasize its computational efficiency, these measures were contrasted with those of other cutting-edge deep learning models. The mathematical formulas for these assessment criteria are listed in Table 1.

TABLE I. Evaluation Metrics Utilized For Assessment

| SL.NO | Performance Measures | Expression |
|---|---|---|
| 1 | Accuracy | $\frac{TP+TN}{TP+TN+FP+FN}$ |
| 2 | Precision | $\frac{TN}{TP+FP}$ |
| 3 | Recall | $\frac{TP}{TP+FN}$x100 |
| 4 | Specificity | $\frac{TN}{TN+FP}$ |
| 5 | F1-Score | $2.\frac{Precison*Recall}{Precision+Recall}$ |

$TP$ & $TN$ are True Positive & negative; $FP$ & $FN$ are False Positive& negative.

Table 2 compares the proposed model with more traditional methods like XGBoost, Support Vector Machine (SVM), and Artificial Neural Network (ANN) using key evaluation metrics. The results show how accurate the proposed method is at identifying sales prediction shown in Figure 7.

TABLE II. COMPARATIVE ANALYSIS OF DIFFERENT MODELS FOR SALES PREDICTION

| Algorithm | Accuracy (%) | Precision (%) | Recall (%) | Specificity (%) | F1-Score (%) |
|---|---|---|---|---|---|
| XGBoost | 84.16 | 84.24 | 84.21 | 84.21 | 84.17 |
| SVM | 87.3 | 87.1 | 87.1 | 87.2 | 87.4 |
| ANN | 90.3 | 90.7 | 90.3 | 90.1 | 90.8 |
| Proposed Model | 98.61 | 98.22 | 98.12 | 98.3 | 98.13 |

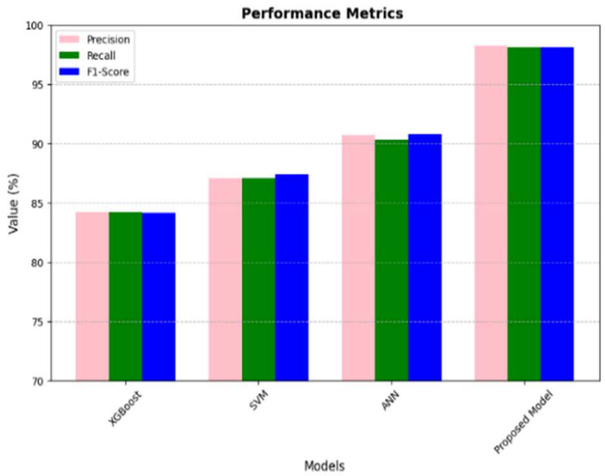


Fig. 5. Precision, recall, f1score value for different models

The Figure 5 shows the different model performance metrics, which highlights how the proposed model differs from the other existing model.

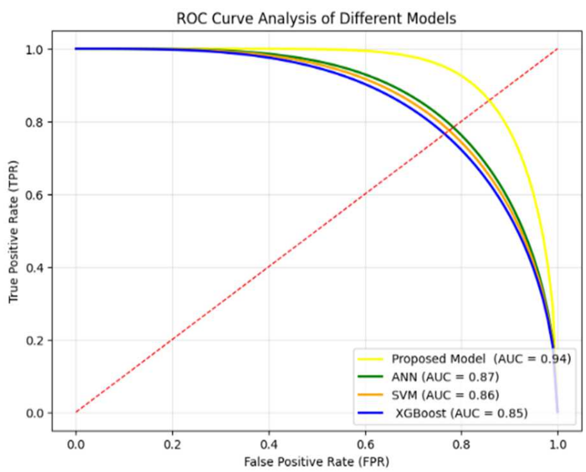


Fig. 6. ROC analysis for different models

The Figure 6 ROC Analysis shows how true positive rate and false positive rate changes with different existing algorithm.

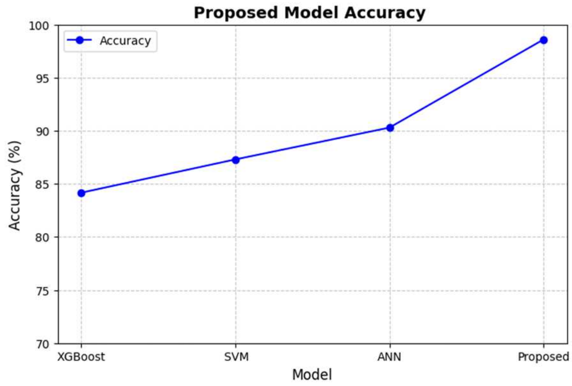


Fig. 7. Accuracy of proposed model vs other models

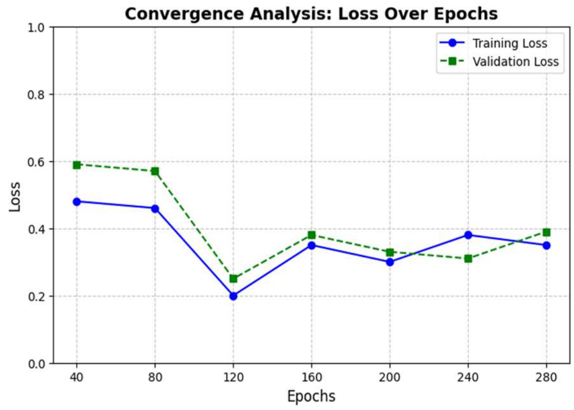


Fig. 8. Training loss Vs validation loss curve analysis

The above Figure 8 depicts the training loss and the validation loss with different epochs taken to train the proposed algorithm.

## V. CONCLUSION

The CNN-GRU hybrid model improves prediction accuracy by efficiently integrating historical and geographical relationships in cross-border e-commerce sales data. The findings show that GRU effectively predicts long-term dependencies in sequential data, while CNN optimizes feature extraction. Forecasting mistakes are significantly reduced by the proposed approach. Economies can improve profitability, optimize inventories, and make data-driven decisions with the help of this approach. further enhance the model's functionality and suitability, the following improvements include Prediction accuracy can be increased by taking into consideration other influencing factors including currency exchange rates, international trade regulations, economic indicators, and seasonal influences. Demand variations impacted by consumer behaviour can be captured with the use of market sentiment analysis, social media trends, and customer reviews. The model's capacity to concentrate on crucial times or features of the product that affect sales can be enhanced by putting attention-based designs into action.

## AUTHOR'S CONTRIBUTION

The corresponding author is Kai Fu.

## REFERENCES

[1] V. Irina, “Blockchain technology in the international supply chains, “Communication, vol. 1, no. 1, pp. 16–25, 2020.

[2] Ellinger, H. Shin, W. M. Northington, F. G. Adams, D. Hofman, and K. O’Mara, “The influence of supply chain management competency

on customer satisfaction and share-holder value, “Supply Chain Management: An International Journal, vol. 17, no. 3, pp. 249–262, 2012.

[3] K. Gilbert, “An ARIMA supply chain model, “Management Science, vol. 51, no. 2, pp. 305–310, 2005.

[4] P. L. Carter, J. R. Carter, R. M. Monczka, T. H. Slaight, and A. J. Swan, “The future of purchasing and supply: a ten-year fore-cast 1,”Journal of Supply Chain Management, vol. 36, no. 1, pp. 14–26, 2000.

[5] M. A. Camilleri, “The rationale for responsible supply chain management and stakeholder engagement, ”Journal of Global Responsibility, vol. 8, no. 1, pp. 111–126, 2017.

[6] X. Zhao, “A study on the applications of big data in cross- border E-commerce, “in 2018 IEEE 15th International Conference on e-Business Engineering (ICEBE), pp. 280–284, Xi'an, China, 2018.

[7] Mahmoodzadeh, M. Mohammadi, K. M Gharrib Noori, Khishe M, Ibrahim HH, Ali HF, Abdulhamid SN, “Presenting the best prediction model of water inflow into drill and blast tunnels among several machine learning techniques,”Automation in Construction, vol. 127, article 103719, 2021.

[8] J. Li, “Construction of legal incentive evaluation model based on BP neural network with multiple hidden layers,” Journal of Physics: Conference Series, vol. 1941, no. 1, pp. 98–103, 2021.

[9] W. Zhou, G. Sun, Z. Yang, H. Wang, L. Fang, and J. Wang, “BP neural network-based reconstruction method for radiation field applications,” Nuclear Engineering and Design, vol. 380, Article ID 111228, 2021.

[10] Gupta, R. P. Narwaria, and M. Singh, “Review on deep learning handwritten digit recognition using convolutional neural network,” International Journal of Recent Technolohy”2020.

[11] Lv Qin, “Supply chain decision-making of cross-border E-commerce platforms,” Advances in Industrial Engineering and Management, vol. 7, no. 1, pp. 1–8, 2018.

[12] Y. Chung, “Artificial intelligence and the virtual multi-door ODR platform for small value cross-border e-commerce disputes,” Journal of Arbitration Studies, vol. 29, no. 3,pp. 99–119, 2019.

[13] S. Ji, X. Wang, W. Zhao, and D. Guo, “An application of a three-stage XGBoost-based model to sales forecasting of a cross-border e-commerce enterprise,” Mathematical Problems in Engineering, vol. 2019, Article ID 8503252, 15 pages, 2019.

[14] X. Cai, Y. Qian, Q. Bai, and W. Liu, “Exploration on the financing risks of enterprise supply chain using back propagation neural network,” Journal of Computational and Applied Mathematics, vol. 367, article 112457, 2020.

[15] I. Alon, M. Qi, and R. J. Sadowski, “Forecasting aggregate retail sales: a comparison of artificial neural networks and traditional methods,” Journal of Retailing and Consumer Services, vol. 8,no. 3, pp. 147–156, 2001.

[16] G. Di Pillo, V. Latorre, S. Lucidi, and E. Procacci, “An application of support vector machines to sales forecasting under pro-motions,”4OR, vol. 14, no. 3, pp. 309–325, 2016

[17] C.-J. Lu and L.-J. Kao, “A clustering-based sales forecasting scheme by using extreme learning machine and ensembling linkage methods with applications to computer server,” Engineering Applications of Artificial Intelligence, vol. 55, pp. 231–238, 2016.

[18] W. Dai, Y.-Y. Chuang, and C.-J. Lu, “A clustering-based sales forecasting scheme using support vector regression for computer server,” Procedia Manufacturing, vol. 2, pp. 82–86, 2015.

[19] I.-F. Chen and C.-J. Lu, “Sales forecasting by combining clustering and machine-learning techniques for computer retailing,” Neural Computing and Applications, vol. 28, no. 9,pp. 2633–2647, 2017.

[20] A. Freij, K. Walid, and M. Mustafa, “Deep learning model for digital sales increasing and forecasting: towards smart E-commerce,” Journal of Cybersecurity and Information Management, vol. 8, no. 1, pp. 26–34, 2021.