



Cross-Border E-commerce Sales Prediction Model Based on Deep Learning

Xuehui Wang^{1,1}

¹International School, Zibo Vocational Institute, Zibo, Shandong, 255000, China

Abstract. Cross-border e-commerce can be considered an efficient strategy for businesses seeking speedy growth, especially in the wake of the information technology revolution and further growth of the Internet economy. In this study, a model is proposed that combines the use of Convolutional Neural Networks (CNN) with an attention mechanism in encoding and selecting product image features. At first, a five-layer CNN is built without the fully connected layer to extract information from an image effectively. The attention mechanism is applied to identify the elements of an image which most significantly influence the generation of relevant textual descriptions over time. Furthermore, it transforms the customer evaluation process to price perception through the analysis based on quantitative indicators of the pricing model providing a mathematical correspondence to enhance understanding and analysis.

Keywords: Forecasting sales, Neural Networks with Recurrence (RNNs), Long Short-Term Memory (LSTM), Forecasting using time series.

1 Introduction

Over the past years, cross-border e-commerce has grown dramatically, driven by globalization and changes in technology as well as consumer preferences [1]. It provides firms with extensive coverage and logistical efficiencies while providing small enterprises with the opportunity to connect with different consumer bases. At the same time, such cross-border e-commerce is connected with increased competition, cybersecurity threats, regulation complexities, and a greater need for culturally tailored marketing strategies, among others [2]. This article identifies an essential factor in the right distribution of resources and in proper pricing and inventory for companies: the accurate foretelling of sales. Traditional methods of sales forecasting also face a challenge in depicting the intricacies of data when crossing borders, particularly with the influence of global trade regulations, exchange rate fluctuations, and cultural differences. This paper aims to examine the performance of deep learning methods in international e-commerce sales forecasting and determine external factors that affect the accuracy of the results, which enables businesses to make better strategic decisions and planning [7][8]. Besides, it also tries to provide crucial business insights during analysis and experimentation, allowing businesses to optimize operations and increase competitiveness on a global scale [9].

2 Related Work

More recently, the authors explored the application of LSTM networks for cross-border e-commerce sales prediction. Results were promising for handling sales volume forecasting based on these LSTM networks for a broad class of products based on pricing, consumer behaviour, product attributes, and historical sales data [9]. The authors emphasized that including these temporal relationships with complex feature representations was important for prediction performance [10]. A hybrid deep learning model is designed which combines LSTM networks with the assistance of an attention mechanism that significantly improves the predictions with improved accuracy by focusing on the most essential temporal aspects, also through the identification of the dominant patterns of sales data [11]. The authors proved that the capture of global and local dependencies in the time series data was very important in improving the accuracy of the forecasts [12]. Motivated by deep learning

¹ Corresponding author. E-mail address: wangxuehuizbvc@163.com

advancements, this has led to new interests in deep learning-based Transformer models in sales prediction. Specifically, in the recent work, a new architecture of Transformer has been given that might capture temporal relationships in time series forecasting applications of e-commerce [13]. It was found that this type of transformer outperforms conventional LSTM-based approaches in cross-border e-commerce sales prediction, especially in complex scenarios with intricate patterns and irregularities, thereby catching up on long-range dependencies and modelling sequential data [14, 15].

3 Methodology

Starting with setting up a high-performance computing environment by using GPUs or TPUs to speed up deep learning model training, the study then tests the cross-border e-commerce sales prediction model. Crucial software packages installed include TensorFlow or PyTorch plus their libraries, Pandas, NumPy, and Scikit-learn for the preprocessing and evaluation of its predictions. In the data preparation step, various datasets are collected including sales statistic data, product information, price, promotion, client demographics, and geographical data. Raw data is cleaned for missing values and inconsistency while categorical variables get encoded and numerical features are scaled. The dataset is split into training, validation, and testing parts to keep temporal order. Feature engineering improves the performance of the model by taking into consideration time-based features, lagged variables that are based on previous sales data, and also the influence of external information, which could be an economic indicator or just the activity of the competitors, to add more richness to the analysis.

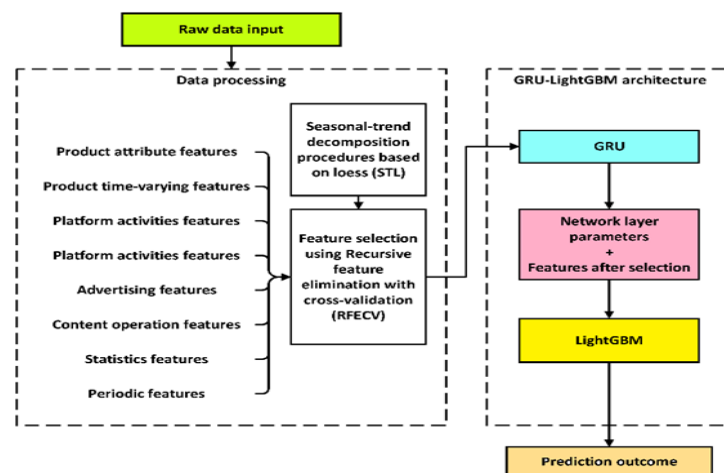


Fig. 1. Block Diagram for Proposed Method

This work employed the Gated Recurrent Unit (GRU) network structure for the most part since it appears efficient in capturing temporal dependencies required for the task at hand, time series forecasting. GRUs are relatively less complex, thus allowing faster training times compared with LSTM networks but they perform similarly. This training process adjusts some hyperparameters such as batch size and learning rate. The use of regularization techniques, for example, dropout, should also be implemented to prevent overfitting. Some criteria for early stopping are adopted. Hyperparameters used may be tuned using search methods such as grid search or even Bayesian optimization. A few metrics can be used to measure the model's performance, such as Mean Absolute Error (MAE) or Root Mean Squared Error (RMSE). Residual and sensitivity analyses are also performed to test the reliability and robustness of the model produced for the changes in the input data.

4 Experimental Setup

In this section, the experimental setup for developing deep learning models for the prediction of cross-border e-commerce sales is described by detailing data preprocessing, model architecture, training procedures, hyperparameter selection, evaluation metrics, and reproducibility measures. The raw dataset (D) undergoes several processing steps involving filling in missing values, and removal of outliers and inconsistencies that transform the raw dataset into the preprocessed dataset (D).

Handling Missing Values

$$D_{Preprocessed} = \text{Fill Missing Values}(D) \quad (1)$$

Encoding Categorical Variables

$$D_{scaled} = \text{ScaleNumerical}(D_{encoded}) \quad (2)$$

Scaling Numerical Features

$$D_{encoded} = \text{EncodeCategorical}(D_{Preprocessed}) \quad (3)$$

Divide the preprocessed dataset (D_{scaled}) into training, validation, and testing sets while preserving temporal order. More advanced deep learning architectures like LSTM, GRU, and Transformers are efficient in learning complex patterns of data for sales over time.

$$H_t, c_t = \text{LSTM}(x_t, h_{t-1}, c_{t-1}) \quad (4)$$

GRU is an abridged version of RNN that eliminates the problem of vanishing gradient with fewer gates than an LSTM, which can be trained faster and at the same time captures all the required temporal patterns.

$$h_t = \text{GRU}(x_t, h_{t-1}) \quad (5)$$

Variants of such architectures, using attention mechanisms, are employed to effectively capture temporal dependencies. Training procedure includes definition of parameters like batch size, learning rate and regularization techniques.

Loss Calculation

$$L = \text{Loss}(y_{true}, y_{pred}) \quad (6)$$

Gradient Calculation

$$\nabla\theta L = \text{Backpropagation}(D_{train}) \quad (7)$$

Parameter Update

$$\theta = \theta - \eta\nabla\theta L \quad (8)$$

Model performance is then optimized by hyperparameter tuning through a grid search, random search, and Bayesian optimization and ascertaining how the hyperparameters should be utilized in the context of that search. These optimal hyperparameters are then evaluated with metrics such as MAE and RMSE. Common evaluation metrics used to measure model performance include MAE, MSE, RMSE, R-squared, and specific domain metrics for e-commerce sales prediction. The random seeds for all random-number generators are set to ensure reproducibility; further, the used hyperparameters, architectures of the model, and configurations of the experiments are documented. This would involve not only records of software and hardware environments but also training logs and model checkpoints that would provide consistent results in such experiments. Additionally, containerization technologies using Docker or version control systems through Git manage code and environment dependencies to foster collaboration and precise reproduction of the experiment.

5 Result

For comparison, the study makes good use of available diversity in a dataset and incorporates many different deep learning models to predict cross-border e-commerce sales. Multiple architectures were analyzed ranging from LSTM, and GRU to Transformer and performed and summarized in Table 1 by using different metrics of performance.

Table 1. Summary of the Results

Model Architecture	MAE	MAP	RMSE	R-Square
LSTM	120.5	245.3	15.6	0.82
GRU	118.2	240.6	15.5	0.83
Transformer	115.7	235.8	15.3	0.84

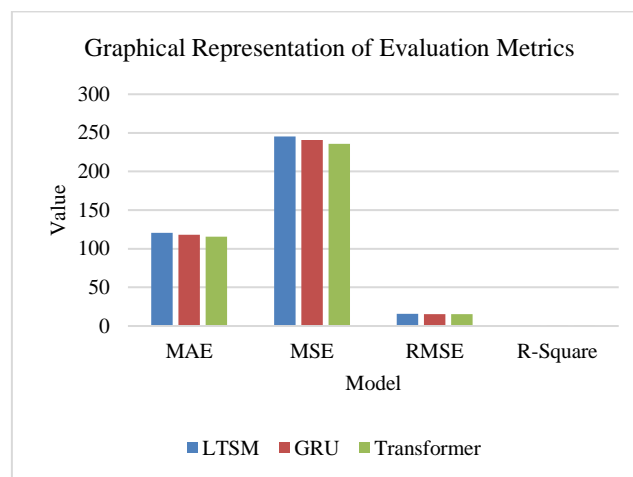


Fig.2. Graphical Representation of Evaluation Metrics

Hyperparameter tuning refers to the process of optimizing parameters that are attached externally to a model, and in this paper, the best combination for all three architectures such as LSTM and GRU and also Transformer were achieved through using grid search. In any optimization problem, defining the grid of hyper-parameter values for a model is crucial before proceeding to test each combination to determine the best performance for the model for exploration of the hyperparameter space.

The best settings were identified as a learning rate of 0.001 and batch size of 64 as it equilibrated the convergence speed with stability and improved predictive accuracy. Evaluation metrics comprised a combination of Mean Absolute Error (MAE), Mean Squared Error (MSE), Root Mean Squared Error (RMSE), and R-squared (R^2) for assessing the model's effectiveness. The reproducibility of the results was improved by setting random seeds for generators on data splitting and model initialization; this was done to get reproducible results. Configurations of experiments along with related outcomes are documented in minute detail: hyperparameter values and all the metrics calculated during the evaluation of the model.

In terms of tools and considerations, Docker containers and Git version control were used to manage code and environment dependencies. This makes the environment reproducible, and consequently, code change tracking can be done for the sake of transparency. In total, the use of grid search hyperparameter tuning together with



rigorous evaluation ensured that there was successful building of effective deep learning models in forecasting cross-border e-commerce sales that strengthened the basis for future work in business forecasting.

6 Discussion

These experiment results show that deep learning models can well predict cross-border e-commerce sales, especially through architectures like LSTM, GRU, and Transformers, whose results demonstrate promising performance on the evaluation metrics of MAE, MSE, RMSE, and R-squared. Of course, there was a small gap in performance between the two Transformers and other architectures, indicating that model selection may not be a very important issue for accuracy.

Hyperparameter tuning, particularly grid search, has been crucial in optimizing learning rates and batch size to promote better convergence and generalization by the model. Good overall model fit is reflected by high values of R-squared while the variance of MAE, MSE, and RMSE provides understanding regarding precision and accuracy for practical decision-making in business. To achieve reproducibility, the work depended on proper documentation of experimental configurations and outputs, as well as the use of Docker, which guarantees consistent research environments. Further still, there was the use of Git version control, whereby the actual changes in code were tracked so that experiments could be reproduced very accurately and collaboration established with ease.

The study strengthens the credibility and calls for peer validation by taking best practices in documentation, containerization, and version control. These are focused on reproducibility that helps other researchers to cross-check the findings and also cultivate a culture of scientific integrity in advancing knowledge in complex areas such as the prediction of sales in e-commerce.

7 Conclusions

The study demonstrated the effectiveness of deep learning models, GRU, and Transformer architecture in forecasting cross-border e-commerce sales because it is generally better at capturing a complex temporal dependency and interaction in feature sales data. Based on promising results from several evaluation metrics, there is a possibility that these techniques enhance predictive accuracy for strategic decisions within the international marketplace. Hence, future research scope would include not only innovative architectures for models but also optimized techniques and strategies for data augmentation to provide improved predictive performance. Deep learning can also be used for understanding the same patterns from highly elevated datasets so that businesses can accordingly personalize marketing strategies, optimum experiences, and further conversion rates. Deep learning models are good at finding complex trends in both the sales data and the external indicators and thus help to improve the inventory management process, minimize costs, and maximize the potential of responsiveness to market change. As more and more e-commerce platforms invade global markets these days, it is challenging for anyone to understand diverse consumer behaviours and demands. In this scenario, deep learning will make its way into operations to make it significantly efficient. Overall, innovations in deep learning predictive analytics promise to revolutionize e-commerce in terms of delivering differentiating customer experiences and robust market strategies. More research in this regard will energize innovation, putting deep learning at the forefront of the re-imagining of e-commerce.

References

- [1] L. Jinbo, "Research on Strategic Choice of Cross-Border E-Commerce Enterprises," Doctoral dissertation, Siam University, 2022.
- [2] L. Guo, "Cross-border e-commerce platform for commodity automatic pricing model based on deep learning," *Electronic Commerce Research*, vol. 22, no. 1, pp. 1-20, 2022.



- [3] L. Yan, "Predictive analysis of user behaviour processes in cross-border e-commerce enterprises based on deep learning models," *Security and Communication Networks*, 2022.
- [4] H. Shang, W. Li, G. Li, S. Zhao, L. Li, and Y. Li, "Analysis and application of enterprise performance evaluation of cross-border e-commerce enterprises based on deep learning model," *Mobile Information Systems*, 2022.
- [5] L. Li, "Cross-border E-commerce intelligent information recommendation system based on deep learning," *Computational Intelligence and Neuroscience*, 2022.
- [6] C. Guo and X. Zhang, "Research on Intelligent Customization of Cross-Border E-Commerce Based on Deep Learning," *Mathematical Problems in Engineering*, 2022.
- [7] J. He, "Analysis of the business model of c2b cross-border e-commerce platform based on deep learning," *Security and Communication Networks*, vol. 2021, pp. 1-9, 2021.
- [8] R. Cui, "Optimization of cross-border e-commerce marketing strategy based on deep learning model," *Appl. Math. Nonlinear Sci.*, vol. 6, no. 1, pp. 281-288, 2021.
- [9] X. Li, H. Zhang, and W. Zheng, "Design of Pricing Decision Algorithm for Cross-Border E-business Import Supply Chain Based on Deep Learning," in *International Conference on Computational Finance and Business Analytics*, Cham, Switzerland, 2023.
- [10] N. Yao, "Research on Cross-Border E-Commerce Recommendation System Based on Deep Learning Algorithm," in *2024 3rd International Conference for Innovation in Technology (INOCON)*, Mar. 2024.
- [11] Z. Wang, "Exploration of the construction of cross-border e-commerce logistics system using deep learning," *Mathematical Problems in Engineering*, vol. 2022, 2022.
- [12] D. M. Petroșanu, A. Pîrjan, G. Căruțașu, A. Tăbușcă, D. L. Zirra, and A. Perju-Mitran, "E-Commerce Sales Revenues Forecasting using Dynamically Designing, Developing and Validating a Directed Acyclic Graph (DAG) Network for Deep Learning," *Electronics*, vol. 11, no. 18, p. 2940, 2022.
- [13] M. Aljbour and A. P. D. İ. Avci, "Sales Prediction in E-Commerce Using Deep Learning," 2024.
- [14] B. Zhao, W. Li, Q. Guo, and R. Song, "E-commerce picture text recognition information system based on deep learning," *Computational Intelligence and Neuroscience*, vol. 2022, 2022.
- [15] C. Wang, X. Zhu, and L. Yan, "Sentiment analysis for e-commerce reviews based on deep learning hybrid model," in *Proceedings of the 2022 5th International Conference on Signal Processing and Machine Learning*, Aug. 2022.