Hybrid model of Bi-LSTM and CNN for Multivariate Time Series Classification for ecommerce sales Forecasting

Sanjay Saini

Research Proposal

NOVEMBER 2023

Abstract

The research introduces a Hybrid Model amalgamating Bi-LSTM and CNN architectures for the classification of multivariate time series data in the domain of e-commerce sales forecasting. The proposed model addresses the complexities associated with predicting sales patterns influenced by diverse contextual factors. Leveraging the strengths of Bi-LSTM in capturing long-term dependencies and CNN in extracting spatial features, the hybridization aims to provide a robust framework for accurate feature extraction and sequence learning. The model's efficacy will be rigorously evaluated using real-world e-commerce sales datasets, comparing its performance against traditional time series forecasting methods as well as standalone Bi-LSTM and CNN models. Through this research, I seek to advance the precision of e-commerce sales forecasts, contributing to the optimization of business operations and decision-making in the dynamic and competitive landscape of online retail.

Table of Contents

LIST OF FIGURES

LIST OF TABLES

LIST OF ABBREVIATIONS

CNN…………..	Convolutional Neural Network
Bi-LSTM …..	Bidirectional Long-Short Term Memory
LSTM ……...	Long-Short Term Memory
DTW…………	Dynamic Time Warping
RNN …………	Recurrent Neural Network
SAX………...	Symbolic Aggregate approXimation
TSC…………	Time series classification
FCN …………	Recurrent Neural Network
ARIMA……...	Autoregressive Integrated Moving Average
SARIMA…………	Seasonal AutoRegressive Integrated Moving Average

1. Background.

Time series analysis in e-commerce involves the study of data points collected over time to forecast future trends. Traditional statistical methods, such as ARIMA and Exponential Smoothing, have been widely used for this purpose. However, the dynamic and complex nature of e-commerce data, characterized by high volume, velocity, and variety, poses significant challenges to these traditional methods. The need for more advanced techniques to accurately predict sales trends has become increasingly evident, especially with the rapid growth of online retail and the vast amount of data generated by e-commerce platforms.

The limitations of traditional time series analysis methods in handling complex e-commerce data have led to the adoption of machine learning techniques. Machine learning offers the ability to automatically learn and improve from experience without being explicitly programmed. Early machine learning approaches in time series forecasting involved simpler models like decision trees and linear regression. However, these models often fell short in capturing the non-linear relationships inherent in multivariate time series data.

Deep learning, a subset of machine learning, has gained prominence due to its ability to learn complex patterns in large datasets. Two significant architectures in deep learning, LSTM and CNN, have shown remarkable success in time series analysis. LSTM, a type of recurrent neural network, is particularly adept at processing sequential data, capturing long-term dependencies, and retaining information over extended periods. This makes it well-suited for analyzing time series data where past information is crucial for predicting future trends. On the other hand, CNNs, known for their success in image processing, have been adapted for time series classification. Their ability to detect local patterns and features in data makes them suitable for identifying short-term trends and anomalies in time series.

The integration of LSTM and CNN into a hybrid model leverages the strengths of both architectures. While LSTM models temporal dependencies, CNN excels in extracting spatial features from multivariate time series data. This synergy enhances the model's ability to handle the complexities of e-commerce data, which often involves multiple influencing factors such as customer behavior, market trends, and external economic indicators. The hybrid model aims to provide a more accurate and reliable tool for e-commerce sales forecasting, essential for strategic decision-making in the dynamic e-commerce sector.

The application of LSTM-CNN hybrid models in e-commerce sales forecasting involves predicting future sales volumes based on historical data and various external factors. Accurate forecasting is crucial for inventory management, marketing strategy, and overall business planning. The hybrid model processes multivariate time series data, including sales figures, customer traffic, seasonal trends, and promotional activities, to generate more accurate and reliable sales forecasts.

The hybrid model of BiLSTM and CNN represents a significant advancement in the field of e-commerce sales forecasting. By combining the strengths of BiLSTM and CNN, this model offers a robust framework for multivariate time series classification, capable of handling the complexities and variabilities inherent in e-commerce data. As e-commerce continues to evolve, the adoption of such advanced forecasting models will be crucial for businesses seeking to optimize their operations and stay competitive in the digital marketplace.

2. Related Work

- **Time Series Analysis in E-commerce** (Ensafi et al., 2022) The document provides a foundational understanding of time series analysis in the context of e-commerce. It emphasizes the importance of accurate sales forecasting for inventory management, pricing strategies, and overall business planning. The paper also discusses traditional statistical methods for time series forecasting, highlighting their limitations in handling complex, multivariate data common in e-commerce platforms.
- **Evolution of Machine Learning Techniques in Forecasting** (Ensafi et al., 2022; Ko et al., 2021; Sirisha et al., 2022) These Sources delve into the evolution of machine learning techniques in forecasting. They compare traditional methods like ARIMA with more advanced machine learning approaches, including neural networks. These papers underscore the superior performance of machine learning models in capturing non-linear relationships and complex patterns in data, which are crucial for accurate e-commerce sales forecasting.
- **The Rise of Deep Learning: LSTM and CNN** (Jaiswal & Gupta, 2023; Ko et al., 2021; Sirisha et al., 2022; Wei et al., 2022)The documents focus on the rise of deep learning techniques, particularly LSTM and CNN, in time series analysis. LSTM, with its ability to remember long-term dependencies, is shown to be particularly effective in

handling sequential data. Meanwhile, CNNs, known for their prowess in pattern recognition in images, are being increasingly applied to time series data for feature extraction.

- **Hybrid Models: Combining LSTM and CNN** The Sources (Akhtar & Shah, 2021; Karim et al., 2017; Khan et al., 2021; Wei et al., 2022; Zhao et al., 2022) explore the development of hybrid models that combine LSTM and CNN. These models leverage the strengths of both architectures: LSTMs for capturing temporal dependencies and CNNs for extracting spatial features. The papers demonstrate how such hybrid models outperform standalone LSTM or CNN models in various time series classification tasks.
- **Application in E-commerce Sales Forecasting** (Ensafi et al., 2022) In documents the application of these advanced models in the context of e-commerce sales forecasting is specifically addressed. They provide case studies and empirical results showing the effectiveness of LSTM-CNN hybrid models in predicting sales, considering multiple variables like customer behavior, seasonal trends, and promotional activities.
- **Challenges and Future Directions** (Khan et al., 2021; Ko et al., 2021; Sirisha et al., 2022; Wei et al., 2022) Finally, sources discuss the challenges in implementing these advanced models in a real-world e-commerce setting. Issues such as data quality, computational complexity, and model interpretability are highlighted. The papers also suggest future research directions, including the integration of external data sources and the exploration of more sophisticated hybrid models.

Conclusion The reviewed literature collectively underscores the potential of hybrid LSTM-CNN models in enhancing the accuracy and efficiency of multivariate time series classification, particularly in the domain of e-commerce sales forecasting. These models represent a significant advancement over traditional methods, offering a more nuanced understanding of complex, dynamic market trends. Future research in this area will explore further innovations and improvements in predictive analytics for e-commerce.

3. Research Questions

This thesis tries to answer the following questions:

1. How does the integration of BiLSTM and CNN enhance the accuracy and efficiency of multivariate time series classification in e-commerce sales forecasting?

This question investigates the synergistic effect of combining BiLSTM and CNN in a hybrid model. It seeks to understand how the bidirectional nature of LSTM, which captures temporal dependencies, complements the spatial feature extraction capabilities of CNN. The focus is on assessing the model's performance in accurately forecasting sales by analyzing complex, multivariate e-commerce data, as highlighted in sources (Karim et al., 2019; Khan et al., 2021; Ko et al., 2021; Sirisha et al., 2022; Wei et al., 2022; Zhao et al., 2022).

2. What are the challenges and limitations associated with the implementation of a BiLSTM-CNN hybrid model in real-world e-commerce environments, and how can these be addressed?

   Drawing from the discussions in sources (Karim et al., 2019; Khan et al., 2021; Ko et al., 2021; Sirisha et al., 2022; Wei et al., 2022; Zhao et al., 2022), this question delves into the practical aspects of deploying the hybrid model. It explores challenges such as computational complexity, data quality, and model interpretability. The aim is to identify strategies to mitigate these challenges, ensuring the model's applicability and effectiveness in diverse e-commerce settings.

3. Can the BiLSTM-CNN hybrid model be effectively adapted to incorporate external factors (such as economic indicators, social media trends, and seasonal variations) that influence e-commerce sales, and how does this impact its forecasting accuracy?

   This question, inspired by insights from sources (Karim et al., 2019; Khan et al., 2021; Ko et al., 2021; Sirisha et al., 2022; Wei et al., 2022; Zhao et al., 2022), explores the model's adaptability and responsiveness to external factors that significantly impact e-commerce sales. It aims to assess the extent to which incorporating such variables enhances the model's predictive accuracy and provides a more holistic view of market dynamics.

4. Aim and Objectives

4.1. Aim

To develop and evaluate a hybrid model combining BiLSTM and CNN for enhancing the accuracy and efficiency of multivariate time series classification in e-commerce sales forecasting.

4.2. Objectives:

- To Investigate the Integration of BiLSTM and CNN for Time Series Analysis: Objective is to explore how the combination of BiLSTM and CNN can be optimized for time series data, particularly focusing on the unique challenges presented by e-commerce sales data. This involves understanding the strengths of each model in handling temporal and spatial dependencies within the data, as discussed in sources (Ensafi et al., 2022), and (Khan et al., 2021).
- To Evaluate the Performance of the Hybrid Model in Multivariate Forecasting: Aimed at assessing the model's ability to accurately forecast e-commerce sales by analyzing multiple variables. This includes evaluating the model's performance against traditional and single-architecture models, as highlighted in sources(Akhtar & Shah, 2021; Karim et al., 2017; Khan et al., 2021; Ko et al., 2021; Sirisha et al., 2022; Wei et al., 2022).
- To Address Computational and Practical Challenges: This objective focuses on identifying and overcoming the computational challenges, such as model complexity and data requirements, ensuring the model's practical applicability in real-world settings, as discussed in sources (Akhtar & Shah, 2021; Ensafi et al., 2022; Karim et al., 2017; Khan et al., 2021; Ko et al., 2021; Lei & Wu, 2020; Sirisha et al., 2022; Wang et al., 2020; Zhu et al., 2022).
- To Adapt the Model for Dynamic E-commerce Environments: The goal here is to enhance the model's adaptability to rapidly changing e-commerce environments. This includes incorporating external factors like market trends, consumer behavior, and seasonal variations, which are crucial for accurate forecasting, as indicated in sources (Akhtar & Shah, 2021; Ensafi et al., 2022; Karim et al., 2017; Khan et al., 2021; Ko et al., 2021; Lei & Wu, 2020; Sirisha et al., 2022; Wang et al., 2020; Zhu et al., 2022).
- To Contribute to the Field of E-commerce Predictive Analytics: The final objective is to contribute valuable insights and methodologies to the field of predictive analytics in e-commerce. This involves not only advancing the technical understanding of hybrid modeling techniques but also providing practical guidelines for their implementation in e-commerce sales forecasting.

5. Significance of the Study

The significance of the study lies in its potential to revolutionize the accuracy and efficiency of predictive analytics in the e-commerce sector. As e-commerce platforms generate vast amounts of complex, multivariate data, traditional forecasting methods often fall short in capturing the intricate patterns and relationships within this data. The proposed hybrid model leverages the strengths of both BiLSTM and CNN, offering a more sophisticated approach to understanding and predicting consumer behavior and market trends.

This research stands to contribute significantly to the field of e-commerce by providing a robust tool for forecasting sales, which is crucial for inventory management, marketing strategies, and overall business planning. The ability of the hybrid model to incorporate and analyze multiple variables, including external factors like economic indicators and seasonal variations, can lead to more informed and strategic decision-making processes for e-commerce businesses. Furthermore, by addressing the practical challenges of implementing such advanced models, this study tries to bridge the gap between theoretical research and real-world applications, making it a valuable resource for both academics and practitioners in the field of e-commerce and data science.

6. Scope of the Study

The scope of the study encompasses several key areas in the realm of advanced data analytics and e-commerce strategy. Primarily, the study focuses on the development and optimization of a hybrid machine learning model that integrates the strengths of BiLSTM and CNN. This model is specifically tailored to address the complexities of multivariate time series data prevalent in e-commerce platforms, as highlighted in the provided documents.

The research will delve into the comparative analysis of this hybrid model against traditional time series forecasting methods and standalone machine learning models, assessing its efficacy in handling large-scale, dynamic e-commerce data. A significant part of the study is dedicated to evaluating the model's ability to accurately predict sales trends by processing and learning from various data dimensions, including customer behavior patterns, seasonal fluctuations, and market changes.

Additionally, the study aims to explore the practical aspects of implementing this advanced predictive model in real-world e-commerce settings. This includes addressing challenges related to computational resources, data quality, and model scalability. The scope extends to providing insights and recommendations for e-commerce businesses on leveraging this model for strategic decision-making, thereby enhancing their operational efficiency and market responsiveness.

7. Research Methodology

7.1. Introduction

This research aims to develop a hybrid model combining BiLSTM and CNN for forecasting e-commerce sales. The methodology focuses on leveraging the sequential data processing capability of BiLSTM and the feature extraction proficiency of CNN. This approach is designed to handle the complexities and variabilities in e-commerce data, providing a more accurate and efficient forecasting tool. The methodology encompasses dataset description, data preparation, algorithm selection, implementation, and evaluation, ensuring a comprehensive approach to model development and validation.

7.2. Data Preparation

Data preparation will involve cleaning, normalization, and transformation of the e-commerce dataset. This step is crucial to ensure the quality and consistency of the data fed into the model. Techniques like handling missing values, outlier detection, and feature engineering will be employed. The data will be segmented into training, validation, and testing sets to facilitate model training and evaluation. Special attention will be given to the temporal nature of the data, ensuring that time-dependent characteristics are appropriately handled and represented in the model.

7.3. Algorithms & Techniques Description

The core of the research methodology is the integration of BiLSTM and CNN algorithms. BiLSTM will be used for its ability to capture long-term dependencies in time series data, processing information in both forward and backward directions. CNN will be employed for its proficiency in extracting spatial features and patterns from the data. The synergy of these algorithms aims to harness their respective strengths, creating a robust model for time series classification. The research will also explore various architectural configurations and hyperparameter settings to optimize the model's performance.

7.4. Implementation

Implementation will involve coding the hybrid model using Python and ML libraries like TensorFlow or Keras. The model will be trained on the prepared dataset, with iterative adjustments and optimizations based on performance metrics. The implementation phase will also include the integration of the model into a simulated e-commerce environment to test its real-world applicability. This phase is critical for understanding the practical challenges and limitations of the model in a dynamic e-commerce setting.

7.5. Evaluation

The evaluation of the model will be conducted using a range of metrics such as accuracy, precision, recall, and F1 score. The model's forecasting performance will be compared against traditional time series models and standalone deep learning models. Evaluation will also consider the model's computational efficiency and scalability. The robustness of the model will be tested under various scenarios, including changes in market trends and consumer behavior. The evaluation aims to validate the effectiveness of the hybrid model in accurately forecasting e-commerce sales and its potential as a tool for strategic decision-making in the e-commerce.

8. Required Resources

8.1. Hardware requirements

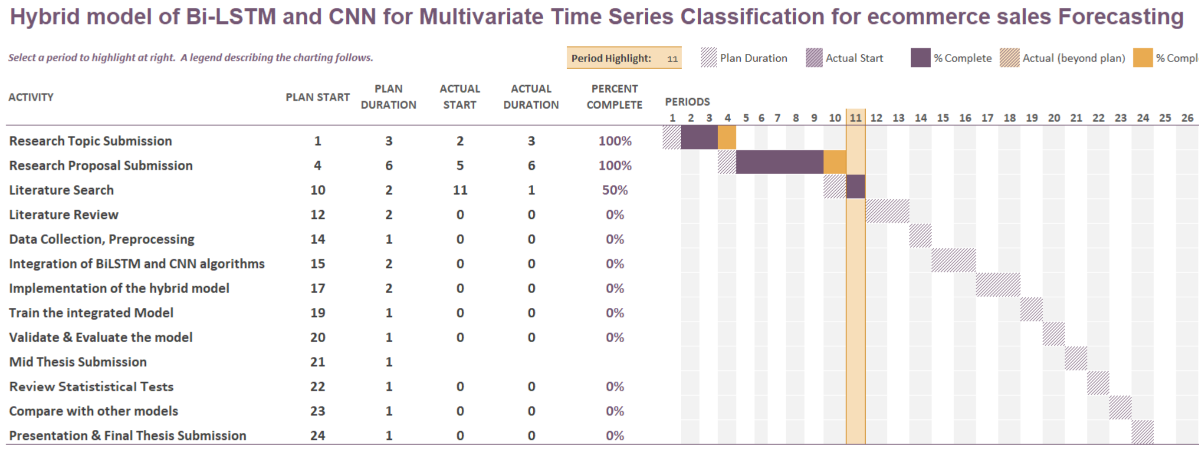
1. High-performance CPUs and GPUs.
2. Large RAM capacity.
3. High-speed SSD storage.
4. Stable internet connection.
5. Backup power solutions.

8.2. Software requirements

1. Python 3.7+.
2. TensorFlow, Keras, or PyTorch.
3. Pandas and NumPy libraries.
4. Matplotlib or Seaborn for visualization.
5. Jupyter Notebook or PyCharm IDE.
6. Git for version control.
7. Database management software.
8. Linux or Windows OS with GPU support.

9. Research Plan

9.1. Gantt Chart



Hybrid model of Bi-LSTM and CNN for Multivariate Time Series Classification for ecommerce sales Forecasting

Select a period to highlight at right. A legend describing the charting follows.

Period Highlight: 11 | Plan Duration | Actual Start | % Complete | Actual (beyond plan) | % Compl

ACTIVITY	PLAN START	PLAN DURATION	ACTUAL START	ACTUAL DURATION	PERCENT COMPLETE
Research Topic Submission	1	3	2	3	100%
Research Proposal Submission	4	6	5	6	100%
Literature Search	10	2	11	1	50%
Literature Review	12	2	0	0	0%
Data Collection, Preprocessing	14	1	0	0	0%
Integration of BiLSTM and CNN algorithms	15	2	0	0	0%
Implementation of the hybrid model	17	2	0	0	0%
Train the integrated Model	19	1	0	0	0%
Validate & Evaluate the model	20	1	0	0	0%
Mid Thesis Submission	21	1			
Review Statististical Tests	22	1	0	0	0%
Compare with other models	23	1	0	0	0%
Presentation & Final Thesis Submission	24	1	0	0	0%

Figure 1: Project Planner

Note: 1 Period = I Calendar Week

9.2. Risk Mitigation and Contingency Plan

The potential risks and their corresponding contingencies are listed below:

Table 1: Risks and Contingency Plan

Risk	Contingency
Unable to perform research tasks due to personal problems or health issues or professional commitments affecting project timeline	▪ Avail extensions if necessary. ▪ Plan for buffer time in project management.
Unavailability of specialized hardware such as GPUs.	▪ Use cloud GPUs. ▪ Check with Upgrad if paid version required

10. References

Akhtar, S., & Shah, A. (2021). *Comparative Analysis of LSTM-FCN on Multiple Datasets*.

Ensafi, Y., Amin, S. H., Zhang, G., & Shah, B. (2022). Time-series forecasting of seasonal items sales using machine learning – A comparative analysis. *International Journal of Information Management Data Insights*, *2*(1). https://doi.org/10.1016/j.jjimei.2022.100058

Jaiswal, S., & Gupta, P. (2023). Diabetes Prediction Using Bi-directional Long Short-Term Memory. *SN Computer Science*, *4*(4). https://doi.org/10.1007/s42979-023-01831-z

Karim, F., Majumdar, S., & Darabi, H. (2019). Insights into lstm fully convolutional networks for time series classification. *IEEE Access*, *7*, 67718–67725. https://doi.org/10.1109/ACCESS.2019.2916828

Karim, F., Majumdar, S., Darabi, H., & Chen, S. (2017). *LSTM Fully Convolutional Networks for Time Series Classification*. https://doi.org/10.1109/ACCESS.2017.2779939

Khan, M., Wang, H., Riaz, A., Elfatyany, A., & Karim, S. (2021). Bidirectional LSTM-RNN-based hybrid deep learning frameworks for univariate time series classification. *Journal of Supercomputing*, *77*(7), 7021–7045. https://doi.org/10.1007/s11227-020-03560-z

Ko, M. S., Lee, K., Kim, J. K., Hong, C. W., Dong, Z. Y., & Hur, K. (2021). Deep Concatenated Residual Network with Bidirectional LSTM for One-Hour-Ahead Wind Power Forecasting. *IEEE Transactions on Sustainable Energy*, *12*(2), 1321–1335. https://doi.org/10.1109/TSTE.2020.3043884

Lei, Y., & Wu, Z. (2020). Time series classification based on statistical features. *Eurasip Journal on Wireless Communications and Networking*, *2020*(1). https://doi.org/10.1186/s13638-020-1661-4

Sirisha, B., Goud, K. K. C., & Rohit, B. T. V. S. (2022). A Deep Stacked Bidirectional LSTM (SBiLSTM) Model for Petroleum Production Forecasting. *Procedia Computer Science*, *218*, 2767–2775. https://doi.org/10.1016/j.procs.2023.01.248

Wang, B., Jiang, T., Zhou, X., Ma, B., Zhao, F., & Wang, Y. (2020). Time-series classification based on fusion features of sequence and visualization. *Applied Sciences (Switzerland)*, *10*(12). https://doi.org/10.3390/APP10124124

Wei, Z., Zhu, Q., Min, C., Chen, Y., & Wang, G. (2022). Bidirectional Hybrid LSTM Based Recurrent Neural Network for Multi-view Stereo. *IEEE Transactions on Visualization and Computer Graphics*. https://doi.org/10.1109/TVCG.2022.3165860

Zhao, L., Mo, C., Ma, J., Chen, Z., & Yao, C. (2022). LSTM-MFCN: A time series classifier based on multi-scale spatial–temporal features. *Computer Communications*, *182*, 52–59. https://doi.org/10.1016/j.comcom.2021.10.036

Zhu, H., Zhao, P., Chan, Y. P., Kang, H., & Lee, D. L. (2022). Breast Cancer Early Detection with Time Series Classification. *International Conference on Information and Knowledge Management, Proceedings*, 3735–3745. https://doi.org/10.1145/3511808.3557107