SSRG International Journal of Electronics and Communication Engineering
ISSN: 2348-8549/ https://doi.org/10.14445/23488549/IJECE-V12I5P110
Volume 12 Issue 5, 118-125, May 2025

Original Article

Leveraging Deep Learning to Forecast E-Commerce Product Fulfillment

Revelle Akshara[1], S Kranthi Reddy[2], Ajay Jain[3]

[1,2]Department of Computer Science and Engineering, Vignan Institute of Technology and Science, Telangana.
[3]Faculty of Engineering and Technology, Mansarovar Global University, Sehore, Madhya Pradesh.

[1]Corresponding Author : akshara.revelle@gmail.com

Received: 06 March 2025 Revised: 09 April 2025 Accepted: 10 May 2025 Published: 27 May 2025

Abstract *- In today's rapidly growing e-commerce market, many businesses face a major challenge: understanding a customer's behaviour and predicting its fulfilment based on various attributes. To address this challenge, we propose an approach using deep learning techniques. By leveraging advanced neural network architectures such as Artificial Neural Networks (ANN), Multi Layer Perceptrons (MLP), and Deep Neural Networks (DNN), we aim to uncover complex patterns and relationships within the dataset that traditional methods might overlook. This approach involves data collection, preprocessing data, correlation analysis, implementing and fine-tuning deep learning models, and assessing their performance using evaluating metrics. Among the implemented models, the DNN outperformed others, achieving 72.2% accuracy, 78.4% precision, 77.3% recall, and a 77.8% F1 score. The results demonstrate the DNN efficiency and accuracy of enhancing customer satisfaction in the e-commerce sector, ultimately fulfilling the products and driving sales growth.*

Keywords *- Correlation analysis, Deep learning algorithms, E-Commerce, Performance metrics, Preprocessing.*

1. Introduction

Rapidly growing E-commerce has transformed how businesses operate, and customers purchase the requirements. Online shopping platforms now play a vital role in the global economy, offering various products to different groups of customers [1], making the e-commerce market highly competitive and requiring them to adapt quickly to changing customer preferences and market conditions [2].

In the highly competitive market, one of the critical challenges that many e-commerce businesses face is the fulfilment of the product, which indicates customer satisfaction. Pricing strategies and promotional offers are essential factors in e-commerce as they significantly influence customer purchase decisions and overall sales performance. Effective pricing strategies and targeted offers can drive customer engagement and fulfilment, gradually increasing sales [3].

Although a huge amount of e-commerce data is available, it is often not used to predict product fulfilment efficiently. Many existing approaches focus mostly on the analysis of general sales or customer behavior, but they do not fully explore the factors that influence product fulfilment. Predicting product fulfilment requires a sophisticated analysis to interpret customer behavior based on pricing strategies and promotional offers [4].

Traditional techniques like basic statistics or rule-based models might miss complex patterns in the data, leading to less accurate predictions. This highlights the need for an advanced, data-driven approach like deep learning to predict the fulfilment of the product accurately. This study focuses on using data preprocessing and correlation analysis to understand the factors affecting product fulfillment. We apply deep learning techniques to predict product fulfillment based on these insights.

2. Related Work

Many researchers have explored ways to analyze customer data, especially to predict customer behavior in purchasing products and improve customer satisfaction in e-commerce. However, there is still a need to study how e-commerce big data and artificial intelligence can be combined to gain customer insights and improve decision-making, ultimately leading to product fulfilment [5].

Gandomi et al. (2015) emphasized how big data analytics enables businesses to extract meaningful insights from vast datasets and streamline operations to improve customer satisfaction [6].

Kietzmann et al. (2018) explored the transformative role of Artificial Intelligence (AI) in advertising by understanding customer insights through analysis and machine learning [7].

Wong and Marikannan (2020) compared machine learning models to predict customer satisfaction in e-commerce and identified delivery time as a key influencing factor; they also found Random Forest as an efficient model, achieving up to 87.5% accuracy [8].

Zhuang et al. (2021) highlighted how advancements in analytics have significantly influenced customer behavior and overall e-commerce operations [9].

Ho, George To Sum, et al. (2022) proposed a model using Gaussian regression to assess how errors in demand forecasting influence order-picking performance in e-commerce fulfilment centers. Their findings highlight the importance of accurate demand prediction for improving operational efficiency [10].

Alsmadi et al. (2023) reviewed past studies to understand how analytics supports the growth of e-commerce, especially during challenging times like the COVID-19 pandemic. They emphasized the importance of integrating knowledge from different fields and encouraged further research to fully use the potential of big data analytics in businesses, highlighting how it can transform different industries [11].

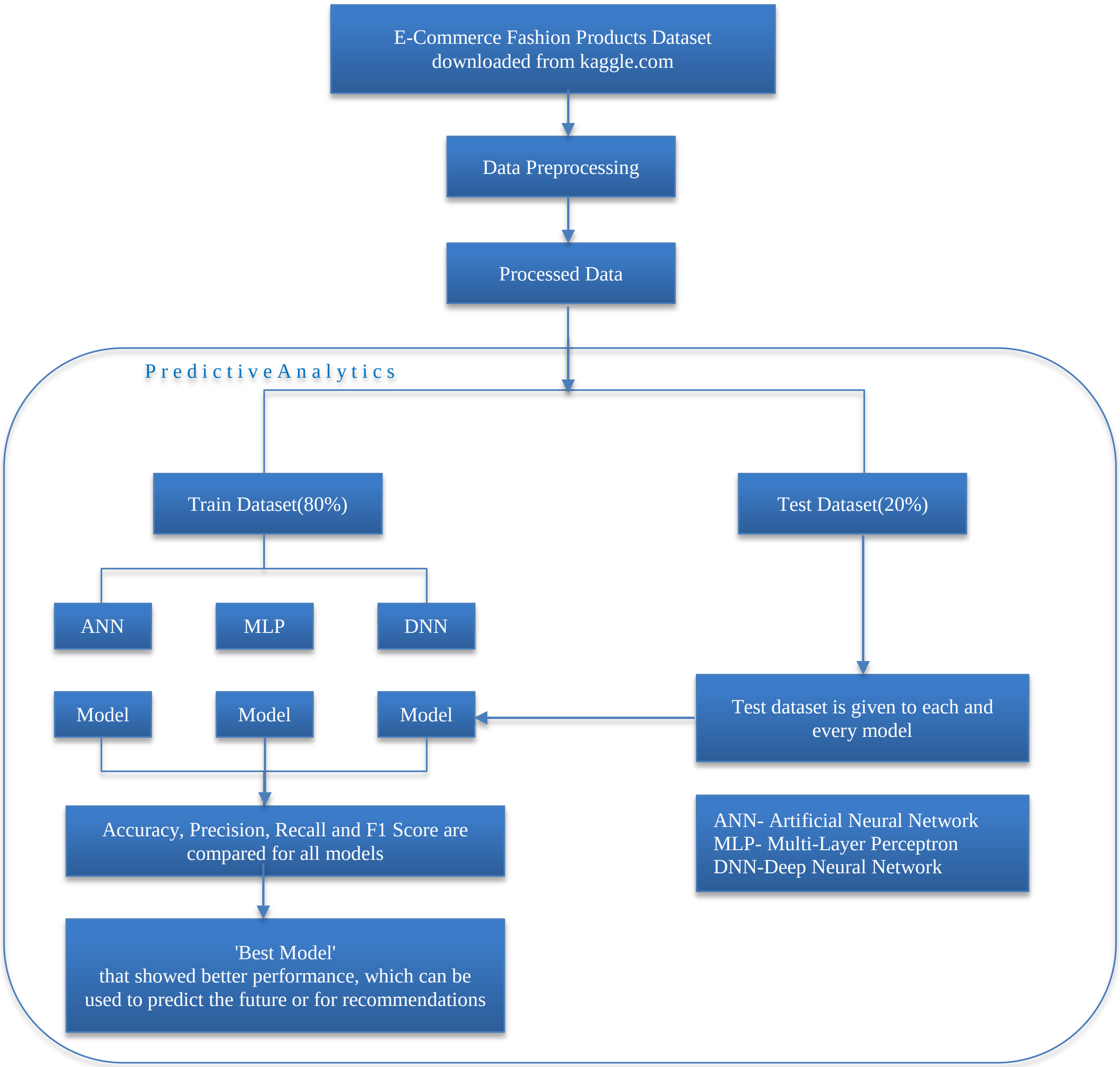

Fig. 1 Architecture of the proposed work

Nijjer et al. (2023) focus on the growing use of AI and ML in customer analytics, highlighting their impact on personalized marketing and customer experience [12].

Ramkumar et al. (2023) demonstrate how real-time data enhances marketing and operations by enabling more personalized and strategic decision-making. Their findings indicate that predictive analytics and data mining play a key role in understanding customer behavior and improving business strategies to stay competitive in e-commerce [13].

All these studies have demonstrated the importance of using analytics and machine learning to improve customer satisfaction and delivery performance in e-commerce. However, there is still a need to accurately predict product fulfilment based on various attributes related to the product.

To address this, a framework using deep learning on product-related datasets to predict product fulfilment to enhance overall customer satisfaction which was less explored in previous studies.

3. Proposed Work

The proposed work is depicted in Figure 1, which entails a comprehensive analysis of an e-commerce product-related dataset using deep learning that aims to predict product fulfilment, i.e., customer satisfaction, to enhance the performance of e-commerce businesses.

3.1. Key Stages

3.1.1. Data Collection and Preprocessing

Initially, a product-related e-commerce dataset is collected and undergoes rigorous preprocessing to ensure it is ready for analysis. This involves fixing missing data, removing duplicates, and standardizing data format to make it easier for further analysis.

3.1.2. Correlation Analysis

With the pre-processed data, correlation analysis is performed to extract the essential features that affect product fulfilment.

3.1.3. Predictive Analytics

Building upon the essential features extracted from correlation analysis, predictive analytics methods are employed to develop robust deep learning models. These models are trained on historical data to forecast future trends and outcomes, enabling businesses to anticipate changes and proactively adapt their strategies.

Through this proposed work, e-commerce businesses can make decisions that optimize their strategies and ultimately make them thrive in the dynamic e-commerce landscape.

3.2. Data Collection and Processing

The proposed work is demonstrated using an e-commerce dataset downloaded from Kaggle.com, which contains some essential attributes related to different products as depicted in Figure 2.

In Figure 2, the 'maincateg' attribute corresponds to the products category reflecting the gender of the customer (men or women) and the 'fulfilled1' attribute is the class label that divides products into two groups, either fulfilled (1) or not fulfilled (0).

Missing values are identified and handled using fusion imputation, showing the lowest Mean Square Error of 0.1808 than that of KNN and Regression, as depicted in Figure 3. After handling missing values, duplicates were removed, thereby reducing the dataset to 5085 unique entries. Then, label encoding was applied to convert the categorical values into numerical, enhancing data compatibility with deep learning algorithms.

```
print(df.shape)
df.head()
```

```
(15730, 15)
```

	id	title	Rating	maincateg	price1	actprice1	Offer %	norating1	noreviews1	star_5f	star_4f	star_3f	star_2f	star_1f	fulfilled1
0	16695	Fashionable & Comfortable Bellies For Women (...	3.9	Women	698	999	30.13%	38.0	7.0	17.0	9.0	6.0	3	3	0
1	5120	Combo Pack of 4 Casual Shoes Sneakers For Men ...	3.8	Men	999	1999	50.03%	531.0	69.0	264.0	92.0	73.0	29	73	1
2	18391	Cilia Mode Leo Sneakers For Women (White)	4.4	Women	2749	4999	45.01%	17.0	4.0	11.0	3.0	2.0	1	0	1
3	495	Men Black Sports Sandal	4.2	Men	518	724	15.85%	46413.0	6229.0	1045.0	12416.0	5352.0	701	4595	1
4	16408	Men Green Sports Sandal	3.9	Men	1379	2299	40.02%	77.0	3.0	35.0	21.0	7.0	7	7	1

Fig. 2 Dataset overview: dimensions and sample entries

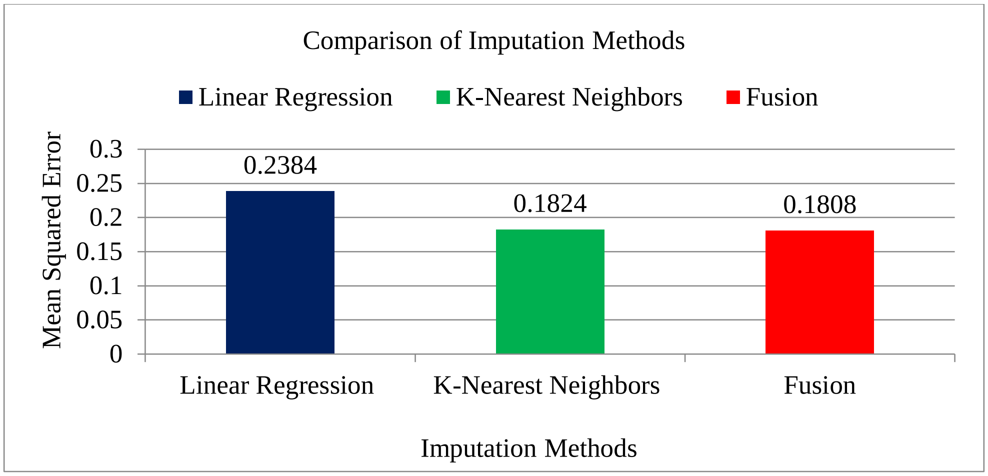

Fig. 3 A Comparative analysis of imputation techniques

The dataset after preprocessing is depicted in Figure 4.

3.3. Correlation Analysis

Correlation analysis helps in determining essential features. It is performed and depicted in Figure 5. From the correlation analysis, Price, Offer%, and Rating are identified as the three essential features and the outcome attribute Fulfilled is highly correlated to Price.

3.4. Predictive Analytics

Predictive analytics techniques involved in building deep learning models to predict product fulfilment.

Deep learning algorithms, namely, Artificial Neural Network (ANN), Multi-Layer Perceptron (MLP), and Deep Neural Network (DNN), are applied to predict product fulfilment based on historical data.

(5085, 15)

	id	title	Rating	maincateg	price1	actprice1	Offer %	norating1	noreviews1	star_5f	star_4f	star_3f	star_2f	star_1f	fulfilled1
0	16695	1257	3.9	1	698	999	30.13	38.0	7.0	17.0	9.0	6.0	3	3	0
1	5120	802	3.8	0	999	1999	50.03	531.0	69.0	264.0	92.0	73.0	29	73	1
2	18391	686	4.4	1	2749	4999	45.01	17.0	4.0	11.0	3.0	2.0	1	0	1
3	495	2039	4.2	0	518	724	15.85	46413.0	6229.0	1045.0	12416.0	5352.0	701	4595	1
4	16408	2109	3.9	0	1379	2299	40.02	77.0	3.0	35.0	21.0	7.0	7	7	1

Fig. 4 Processed data

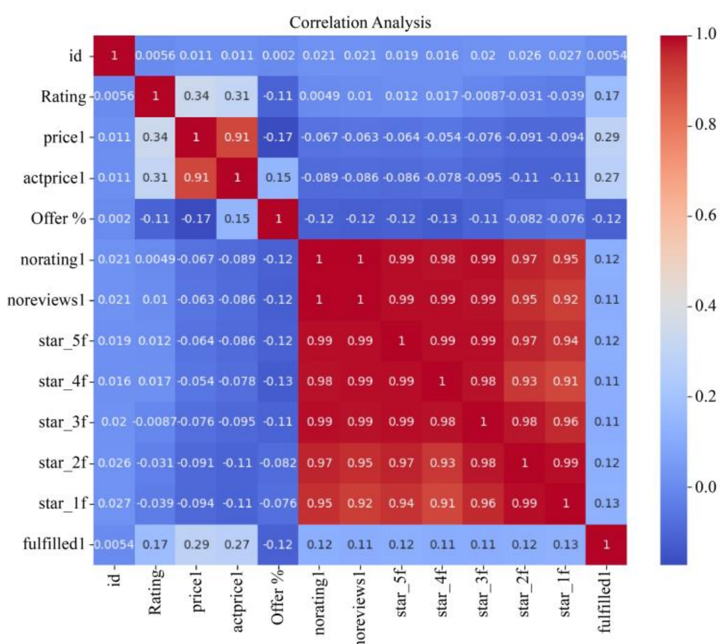

Fig. 5 Correlation analysis

Models' performance is assessed using accuracy, precision, recall, and F1 score metrics.

- Accuracy is the proportion of total correct predictions [14].

$$Accuracy = \frac{TP+TN}{TP+TN+FP+FN} \quad (1)$$

- Precision is the proportion of total true positive predictions among the total positive predictions [15].

$$Precision = \frac{TP}{TP+FP} \quad (2)$$

- Recall (Sensitivity) is the proportion of total true positive predictions among the total actual positive cases [16].

$$Recall = \frac{TP}{TP+FN} \quad (3)$$

- F1 Score is the harmonic mean of precision and recall that provides a balance between them [17].

$$F1 = 2 * \frac{Precision*Recall}{Precision+Recall} \quad (4)$$

Following are the formulae used by deep learning algorithms applied:

3.4.1. Activation Functions

- Rectified Linear Unit (ReLU) to add non-linearity to the model, which makes the model learn complex patterns.

$$\mathrm{ReLU}(\mathrm{x}) = \max(0, \mathrm{x}) \quad (5)$$

- Sigmoid Function to convert the output into a probability for binary classification.

$$\sigma(x) = \frac{1}{1+e^{-x}} \quad (6)$$

3.4.2. Binary Cross-Entropy Loss Function

It measures how well the predicted probabilities y^ match the actual labels y.

$$Loss = -[ylog(\hat{y}) + (1-y)\log(1-\hat{y})] \quad (7)$$

3.4.3. Adam Optimizer

It merges the advantages of previous versions of stochastic gradient descent: Adaptive Gradient Algorithm (AdaGrad) and Root Mean Square Propagation (RMSProp), making it a powerful tool for training models with huge data and complex parameter sets.

$$m_t = \beta_1 m_{t-1} + (1-\beta_1) g_t \quad (8)$$

Where m_t is the first moment estimate at time step t, g_t is the gradient at time step t, and β_1 and β_2 are the decay rates.

$$\upsilon_t = \beta_2 \upsilon_{t-1} + (1-\beta_2) {g_t}^2 \quad (9)$$

Where υ_t is the second-moment estimate at time step t.

$$\widehat{m_t} = \frac{m_t}{1-{\beta_1}^t} \quad (10)$$

$$\widehat{\upsilon_t} = \frac{\upsilon_t}{1-{\beta_2}^t} \quad (11)$$

$$\theta_t = \theta_{t-1} - \eta \frac{\widehat{m_t}}{\sqrt{\widehat{\upsilon_t}}+\epsilon} \quad (12)$$

Where θ_t represents the model parameters at time step t and η is the learning rate.

3.5. Artificial Neural Networks (ANNs)

Machine Learning (ML) models are designed to mimic how the human brain operates. They integrate with many layers of connected nodes, or neurons, that work together to recognize patterns and make predictions [18].

An ANN model is developed for binary classification with numerical data with hidden layers that use ReLU activation and an output layer that uses a sigmoid function. It is trained using an Adam optimizer and a binary cross-entropy loss function. For evaluating the performance, a confusion matrix of the ANN model was constructed, as shown in Figure 6(a), and its training and validation loss, as well as accuracy plots, are depicted in Figures 6(b) and (c), respectively.

3.5.1. Multi-Layer Perceptron (MLP)

MLPs are commonly used in predictive modeling with numerical data as they effectively uncover and model detailed insights [19].

The same functions used in the above ANN model are also used in MLP.

For evaluating the performance, a confusion matrix of the MLP model was constructed, as shown in Figure 7(a), and its training and validation loss, as well as accuracy plots, are depicted in Figures 7(b) and (c), respectively.

3.5.2. Deep Neural Networks (DNNs)

These are a type of ANN with multiple hidden layers between the input and output layers, which help the network to learn detailed insights from the data by building up features in stages [20].

The same functions used in the above models are also used in DNN.

For evaluating the performance, a confusion matrix of the DNN model was constructed as shown in Figure 8(a), and its training and validation loss, as well as accuracy plots, are depicted in Figures 8(b) and (c), respectively.

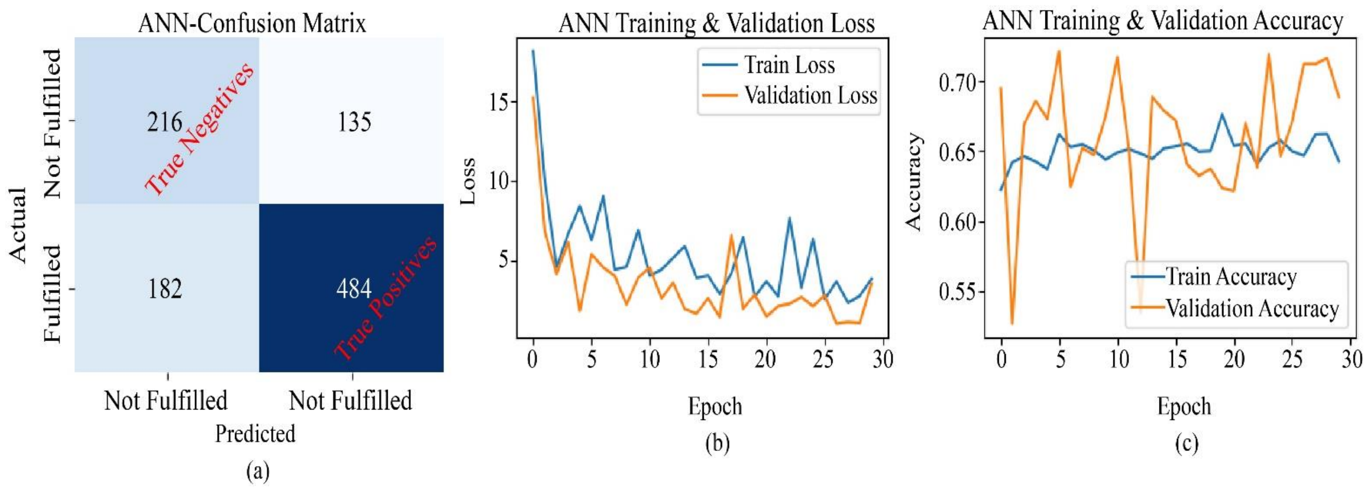

Fig. 6 (a) Confusion matrix of ANN model, (b) Training & validation loss of ANN model, and (c) Training & validation accuracy of ANN model.

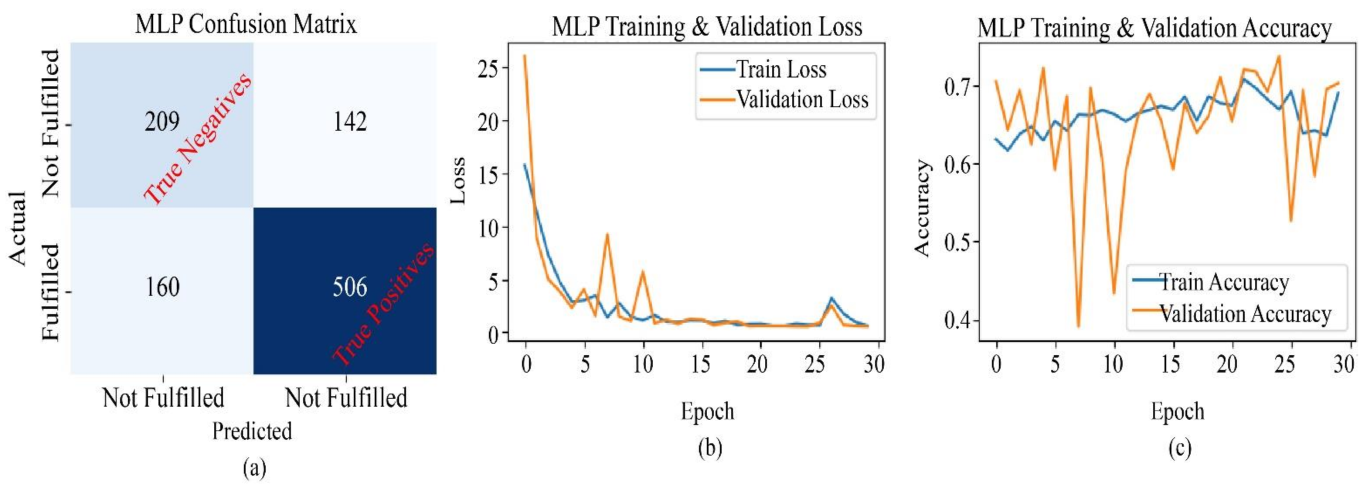

Fig. 7 (a) Confusion matrix of MLP model, (b) Training & validation loss of MLP model, and (c) Training & validation accuracy of MLP model.

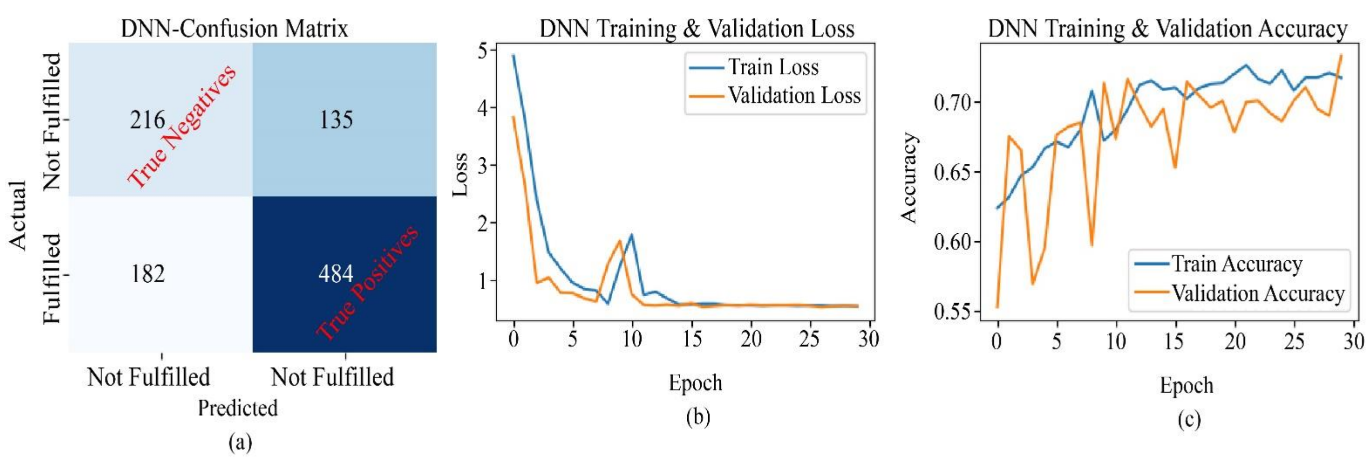

Fig. 8 (a) Confusion matrix of DNN model, (b) Training & validation loss of DNN model, and (c) Training & validation accuracy of DNN model.

Figure 9 depicts the comparative analysis of evaluation metrics across deep learning algorithms applied.

Based on the analysis, the Deep Neural Networks (DNNs) model emerged as the best-performing model among the deep learning algorithms applied. It achieved the highest accuracy of 72.2%, precision of 78.4%, recall of 77.3%, and F1-score of 77.8%. This performance demonstrates that the DNN model effectively learns complex patterns in the data, offering reliable predictions with balanced precision and recall.

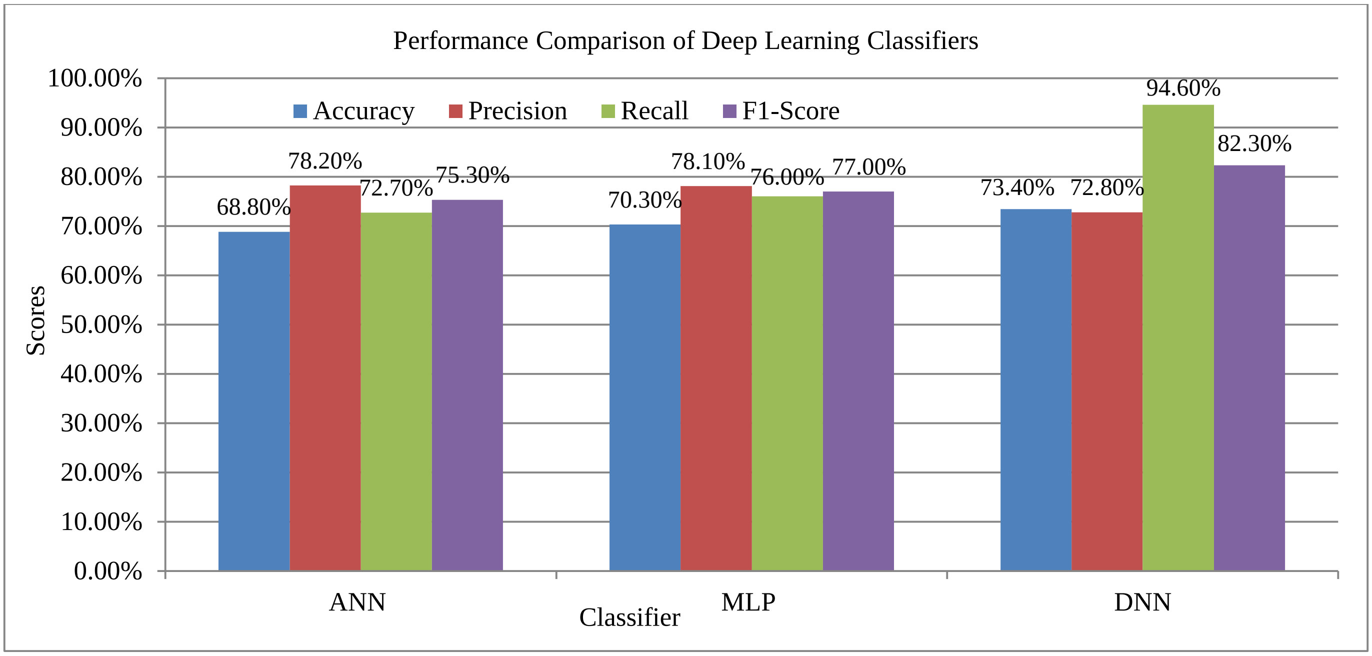

Fig. 9 Performance comparison of deep learning algorithms

4. Conclusion and Future Scope

In today's rapidly growing and competitive world of e-commerce, predicting product fulfilment based on product-related data plays a vital role in enhancing customer satisfaction and improving business efficiency. This study highlights how deep learning models predict product fulfilment accurately by analyzing hidden insights from the data. Among ANN, MLP, and DNN deep learning models, the DNN model has shown superior outcomes with 72.2% accuracy, 78.4% precision, 77.3% recall, and a 77.8% F1-score, which indicates the model's capability in enhancing customer satisfaction and increasing business efficiency.

4.1. Future Scope

However, the model's performance depends on the quality and type of data used. It may also face challenges such as overfitting due to its complexity and might need adjustments to work well across different e-commerce platforms.

Future research can enhance accuracy through hyperparameter tuning, deeper architectures, and advanced models like RNNs, further refining fulfilment predictions. Exploring cross-platform validation and incorporating diverse datasets could also improve the robustness and generalizability of the proposed approach.

References

[1] Vasco Santos et al., "E-Commerce: Issues, Opportunities, Challenges, and Trends," *Promoting Organizational Performance through 5G and Agile Marketing*, pp. 224-244, 2023. [CrossRef] [Google Scholar] [Publisher Link]

[2] Celia Pricilla Mesatania, "Factors Influencing Online Buying Behavior: A Case of Shopee Customers," *Management Science and Business Decisions*, vol. 2, no. 1, pp. 18-30, 2022. [CrossRef] [Google Scholar] [Publisher Link]

[3] Muhammad Masyhuri, "Pricing Strategies Application amongst the Top E-Commerce Southeast Asian Countries," *Asian Journal of Management Analytics*, vol. 2, no. 4, pp. 379-390, 2023. [CrossRef] [Publisher Link]

[4] Muneeb Iqbal et al., "Enhancing Customer Satisfaction in E-Commerce: The Role of Service Quality and Brand Trust," *Forum for Economic and Financial Studies*, vol. 1, no. 1, pp. 1-14, 2023. [CrossRef] [Google Scholar] [Publisher Link]

[5] Mohammed M. Mohammed et al., "Current Directions and Future Research Priorities of Customer Data Analysis," *Journal of Information Systems and Informatics*, vol. 2, no. 2, pp. 300-311, 2020. [CrossRef] [Google Scholar] [Publisher Link]

[6] Amir Gandomi, and Murtaza Haider, "Beyond the Hype: Big Data Concepts, Methods, and Analytics," *International Journal of Information Management*, vol. 35, no. 2, pp. 137-144, 2015. [CrossRef] [Google Scholar] [Publisher Link]

[7] Jan Kietzmann, Jeannette Paschen, and Emily Treen, "Artificial Intelligence in Advertising: How Marketers Can Leverage Artificial Intelligence along the Consumer Journey," *Journal of Advertising Research*, vol. 58, no. 3, pp. 263-267, 2018. [CrossRef] [Google Scholar] [Publisher Link]

[8] Ann-Nee Wong, and Booma Poolan Marikannan, "Optimising E-Commerce Customer Satisfaction with Machine Learning," *Journal of Physics: Conference Series: International Conference on Computational Physics in Emerging Technologies*, Mangalore, India, vol. 1712, no. 1, pp. 1-9, 2020. [CrossRef] [Google Scholar] [Publisher Link]

[9] Weiqing Zhuang et al., "Big Data Analytics in E-Commerce for the U.S. and China through Literature Reviewing," *Journal of Systems Science and Information*, vol. 9, no. 1, pp. 16-44, 2021. [CrossRef] [Google Scholar] [Publisher Link]

[10] G.T.S. Ho et al., "A Forecasting Analytics Model for Assessing Forecast error in E-Fulfilment Performance," *Industrial Management & Data Systems*, vol. 122, no. 11, pp. 2583-2608, 2022. [CrossRef] [Google Scholar] [Publisher Link]

[11] Ayman Abdalmajeed Alsmadi et al., "Big Data Analytics and Innovation in E-Commerce: Current Insights and Future Directions," *Journal of Financial Services Marketing*, vol. 29, pp. 1635-1652, 2024. [CrossRef] [Google Scholar] [Publisher Link]

[12] Shivinder Nijjer et al., "Customer Analytics: Deep Dive into Customer Data," *Encyclopedia of Data Science and Machine Learning*, pp. 1092-1107, 2023. [CrossRef] [Google Scholar] [Publisher Link]

[13] Anagha Ramkumar et al., "Big Data Analytics and its Application in E-Commerce," *AIP Conference Proceedings*, vol. 2736, no. 1, pp. 1-7, 2023. [CrossRef] [Google Scholar] [Publisher Link]

[14] Marina Sokolova, and Guy Lapalme, "A Systematic Analysis of Performance Measures for Classification Tasks," *Information Processing & Management*, vol. 45, no. 4, pp. 427-437, 2009. [CrossRef] [Google Scholar] [Publisher Link]

[15] David M.W. Powers, "Evaluation: from Precision, Recall and F-Measure to ROC, Informedness, Markedness and Correlation," *Arxiv*, pp. 1-27, 2020. [CrossRef] [Google Scholar] [Publisher Link]

[16] Jesse Jon Davis, and Mark Harlan Goadrich, "The Relationship between Precision-Recall and ROC Curves," *Proceedings of the 23rd International Conference on Machine Learning*, Pittsburgh Pennsylvania USA, pp. 233-240, 2006. [CrossRef] [Google Scholar] [Publisher Link]

[17] C. J. Van Rijsbergen, *Information Retrieval*, 2nd ed., Buttersworth, 1-209 1979. [Google Scholar] [Publisher Link]

[18] Zoran Kalinić et al., "Neural Network Modeling of Consumer Satisfaction in Mobile Commerce: An Empirical Analysis," *Expert Systems with Applications*, vol. 175, 2021. [CrossRef] [Google Scholar] [Publisher Link]

[19] Maha Zaghloul, Sherif Barakat, and Amira Rezk, "Predicting E-Commerce Customer Satisfaction: Traditional Machine Learning vs. Deep Learning Approaches," *Journal of Retailing and Consumer Services*, vol. 79, 2024. [CrossRef] [Google Scholar] [Publisher Link]

[20] Neha Chaudhuri et al., "On the Platform but Will They Buy? Predicting Customers' Purchase Behavior Using Deep Learning," *Decision Support Systems*, vol. 149, 2021. [CrossRef] [Google Scholar] [Publisher Link]