Matrik: Jurnal Manajemen, Teknik Informatika, dan Rekayasa Komputer
Vol. 25, No. 3, July 2026, pp. 461∼472
ISSN: 2476-9843, accredited by Kemenristekdikti, Decree No: 10/C/C3/DT.05.00/2025
DOI: 10.30812/matrik.v25i3.5866

A Hyperparameter-Adaptive Multilayer Perceptron Framework for Revenue Prediction Based on E-Commerce User Behavior Data

Safrizal[1], Lili Tanti[2], Yan Yang Thanri[2]
[1]Universitas Muhammadiyah Asahan, Kisaran, Indonesia
[2]Universitas Potensi Utama, Medan, Indonesia

Article Info

Article history:

Received October 27, 2025
Revised April 07, 2026
Accepted May 20, 2026

Keywords:

E-commerce Revenue Prediction;
Hyperparameter-Adaptive;
Optimization Multilayer;
User Behavior Data.

ABSTRACT

Accurate revenue prediction remains a critical challenge for e-commerce platforms due to the highly nonlinear and dynamic nature of user behavior. At the same time, many existing machine learning approaches rely on static model configurations that limit predictive robustness. Although various techniques have been proposed for e-commerce revenue prediction, a systematic, performance-driven approach to adapting Multilayer Perceptron hyperparameters remains underexplored. This study proposes a hyperparameter-adaptive Multilayer Perceptron framework for predicting e-commerce revenue based on user behavior data. Revenue prediction is formulated as a binary classification problem, where outcomes are categorized into conversion and non-conversion events. The dataset comprises 12,330 e-commerce user sessions with behavioral and contextual features, including page interactions, session duration, bounce rate, and visitor characteristics. The proposed framework employs iterative hyperparameter adaptation by evaluating multiple MLP configurations with variations in network depth, activation functions, optimization algorithms, and regularization levels. Model performance is assessed using accuracy, precision, recall, F1-score, and Area Under the Curve. Experimental results indicate that the configuration with the Adam optimizer, ReLU activation, and moderate regularization achieves the best performance, yielding 88.93% accuracy and an AUC of 0.91. These findings confirm that hyperparameter-adaptive selection significantly enhances prediction performance compared to static model settings. The proposed framework provides a systematic approach to improving revenue prediction accuracy and offers valuable insights for data-driven decision-making and strategic planning in e-commerce environments.

Corresponding Author:

Safrizal, +62 823-6850-6066,
Faculty of Economics, Science, and Technology,
Universitas Muhammadiyah Asahan, Kisaran, Indonesia,
Email: safrizal75@ummas.ac.id.

How to Cite:
Safrizal, lili Tanti, and Y. Y. Thanri, "A Hyperparameter-Adaptive Multilayer Perceptron Framework for Revenue Prediction Based on E-Commerce User Behavior Data", *MATRIK: Jurnal Manajemen, Teknik Informatika, dan Rekayasa Komputer*, Vol. 25, No. 3, pp. 461-472, July, 2026.

1. INTRODUCTION

The rapid growth of e-commerce platforms has intensified the demand for accurate revenue prediction to support strategic decision-making, marketing optimization, and efficient resource allocation. In e-commerce environments, revenue generation is strongly influenced by user interaction patterns, where customer behavior plays a central role in determining conversion rates and platform profitability [1, 2]. However, user behavior data are inherently nonlinear and dynamic, making revenue prediction a challenging task for machine learning models and increasing the risk of inaccurate forecasts that may lead to ineffective pricing strategies and suboptimal marketing decisions [3, 4].

To address this challenge, numerous studies have explored machine learning techniques for predicting e-commerce revenue. Existing approaches include ensemble learning methods that combine multiple predictors to improve accuracy [1, 5], survey-based machine learning and deep learning techniques for e-commerce analytics [5], and graph-based neural networks such as Hierarchical Bipartite Graph Neural Networks (HiGNN) designed for large-scale recommendation systems [6]. Other studies have proposed deep learning architectures, including Multimodal Interactive Neural Networks and Directed Acyclic Graph Neural Networks (DAGNN), to capture complex customer behavior and sales patterns [7, 8]. In addition, data imbalance handling and explainability techniques, such as SMOTE-enhanced models with SHAP analysis, have been applied to improve prediction performance and interpretability [9]. Despite these advances, many existing approaches emphasize architectural complexity or specific modeling techniques without providing a systematic strategy for adapting model parameters, which limits robustness when applied to diverse and evolving user behavior data [10, 11].

Among various machine learning models, the Multilayer Perceptron (MLP) remains a widely adopted approach due to its flexibility in modeling nonlinear relationships and its relatively low computational complexity compared to recurrent or graph-based architectures [12, 13]. Previous studies have demonstrated that MLP-based models can effectively predict purchasing intention and customer revenue when properly configured [3, 12]. To improve MLP performance, several studies have incorporated metaheuristic optimization techniques, such as genetic algorithms, bat algorithms, and grasshopper optimization, to tune network parameters [14–16]. However, these approaches often rely on external optimization mechanisms or heuristic settings and do not provide a unified, performance-driven framework that systematically adapts MLP hyperparameters within the revenue prediction process. Consequently, MLP models frequently employ fixed or manually selected hyperparameters, which restrict their ability to generalize across heterogeneous e-commerce user behavior patterns [5, 10].

Based on the above limitations, a clear research gap can be identified: although machine learning and deep learning models for e-commerce revenue prediction have been extensively studied, there remains a lack of a comprehensive, systematic, hyperparameter-adaptive framework for Multilayer Perceptron models that explicitly evaluates and selects optimal configurations based on empirical performance. This gap is critical, as hyperparameter selection directly affects model accuracy, robustness, and practical applicability in real-world e-commerce environments [17].

To address this gap, this study proposes a hyperparameter-adaptive Multilayer Perceptron (MLP) framework for predicting e-commerce revenue based on user behavior data. The main contributions of this study are threefold: (1) developing a systematic and performance-driven hyperparameter adaptation framework for MLP models inspired by recent advances in automated hyperparameter tuning [17]; (2) formulating e-commerce revenue prediction as a binary classification problem based on conversion behavior, incorporating behavioral variables such as product search patterns, session duration, visit-to-transaction conversion rates, and responses to promotions [13, 4]; and (3) validating the effectiveness of the proposed framework through comprehensive experimental evaluation using real-world e-commerce user behavior data. Through this approach, the study aims to enhance revenue prediction accuracy while providing strategic insights for data-driven decision-making in e-commerce businesses.

2. RESEARCH METHOD

This study uses a systematic approach with key stages: data collection, preprocessing, model training, and performance evaluation. Each stage is designed to ensure valid and reliable results. Figure 1 illustrates the proposed hyperparameter-adaptive Multilayer Perceptron (MLP) framework for e-commerce revenue prediction, trained on a dataset of 12,330 records with 18 user behavior and contextual features. The framework consists of data collection, preprocessing (normalization, exploratory data analysis, and distribution analysis), followed by an iterative hyperparameter adaptation process involving tuning and regularization. The optimized MLP model is then trained and evaluated using a test-and-score approach to ensure robust and accurate revenue prediction. This adaptive framework supports data-driven business decision-making in e-commerce environments.

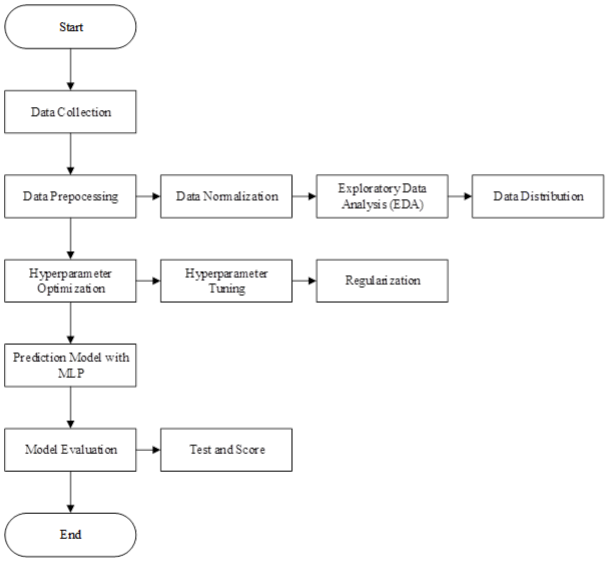

Figure 1. Research Stages

2.1. Research Dataset

This study uses an e-commerce visitor behavior dataset with 12,330 entries and 18 key features. These include quantitative data on page visits and visit duration across administrative, informational, and product categories, as well as metrics like BounceRates and ExitRates. Contextual variables such as SpecialDay, Month, OperatingSystems, Browser, Region, TrafficType, VisitorType, and Weekend are also included. The target variable is Revenue, a binary indicator of whether the visit generated revenue (1) or not (0).

2.2. Multilayer Perceptron (MLP)

This study uses a Multilayer Perceptron (MLP) artificial neural network to capture nonlinear relationships between inputs and outputs, enabling the recognition of complex patterns in the data [18]. The MLP consists of input, hidden, and output layers, trained using optimization algorithms such as Adam or RMSProp and the backpropagation method [19]. ReLU activation functions are applied in the hidden layers, while sigmoid is used in the output layer. The strength of the MLP lies in its ability to learn complex decision boundaries through hierarchical feature representation [20]. Adjusting the number of hidden layers and neurons is crucial to avoid overfitting or underfitting [21]. Dropout and batch normalization techniques are applied to improve the model's generalization and stability. Figure 2 shows a simple MLP architecture with an input layer, four hidden layers, and an output layer. Each neuron in the hidden layers processes inputs from the previous layer using weights, biases, and nonlinear activation functions. The final output is produced by two neurons in the output layer. This structure enables the MLP to capture complex patterns in the data.

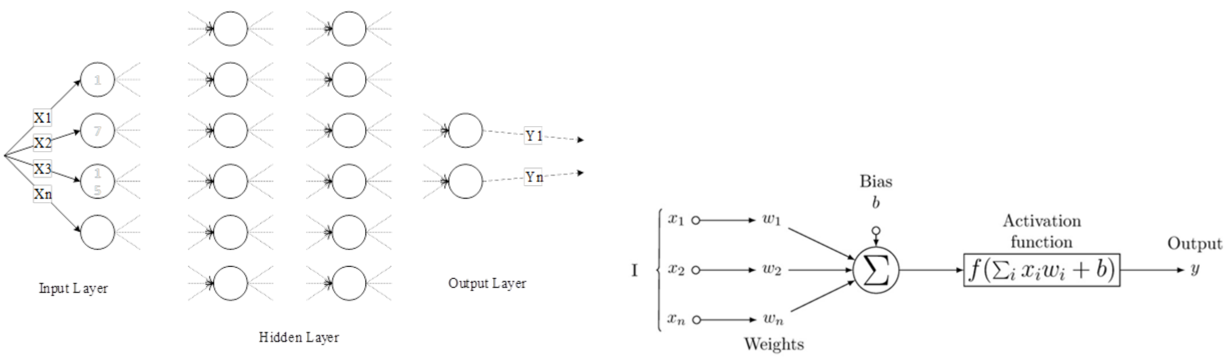

Figure 2. Multi-layer perceptron (MLP) diagram [17]

2.3. Hyperparameter Tuning

Hyperparameter tuning is a crucial step in developing a Multilayer Perceptron (MLP) model to improve e-commerce revenue prediction [22]. Hyperparameters such as the number of hidden layers (2–5 layers), neurons per layer (32–256), activation functions (ReLU in hidden layers and sigmoid in the output layer), learning rate (0.001–0.01 with Adam optimizer), dropout rate (0.2–0.5), and batch size (32, 64, 128) are adjusted using grid search and random search methods [23]. The goal is to balance model complexity to avoid overfitting or underfitting while enhancing training efficiency. Each combination is tested with cross-validation to ensure the model generalizes well to new data. Through this process, the MLP model is expected to achieve optimal performance in accurately and reliably predicting e-commerce revenue based on user behavior patterns. Through this iterative evaluation process, the proposed MLP framework is considered hyperparameter-adaptive, as the final model configuration is determined by comparative performance rather than a predefined static setting.

2.4. Regularization

Regularization is an important step in developing a Multilayer Perceptron (MLP) model to prevent overfitting, a condition where the model becomes too closely fitted to the training data and performs poorly on new data [24]. By applying techniques such as dropout and L2 regularization, the model becomes better able to generalize truly relevant patterns from e-commerce customer behavior data [25]. Dropout randomly disables neurons during training, reducing excessive dependence on certain features, while L2 regularization penalizes large weights to maintain model stability [26]. Implementing regularization directly improves revenue prediction accuracy, making the model not only effective in recognizing complex patterns but also reliable for data-driven business decision-making.

2.5. Exploratory Data Analysis (EDA)

Exploratory Data Analysis (EDA) was conducted thoroughly to identify patterns and characteristics within each customer behavior feature in the dataset [17]. The granular analysis focused on the distribution and variation of features such as the number of pages visited in the Administrative, Informational, and ProductRelated categories, as well as metrics like BounceRates and ExitRates, which reflect visitor engagement and the tendency to leave the site [27]. Additionally, relationships between features and their impact on the target variable, Revenue, were analyzed using visualizations and correlations to uncover specific behavior patterns that contribute to revenue conversion opportunities [2]. This approach helps to gain a detailed understanding of user interactions with the site, which is essential for building an accurate and effective predictive model.

3. RESULT AND ANALYSIS

3.1. Data Collection

The dataset consists of behavioral and contextual features describing e-commerce visitor activity, including the number and duration of visits to administrative, informational, and product-related pages, as well as engagement indicators such as BounceRates, ExitRates, and PageValues. In addition, technical attributes (Operating Systems, Browser, Region, TrafficType) and visitor-related information (VisitorType and Weekend) are included, with Revenue defined as the binary target variable indicating conversion or non-conversion events. The descriptive statistics of the dataset are presented in Table 1.

An initial inspection of Table 1 reveals missing values and extreme outliers, particularly in duration-based features such as ProductRelated_Duration and PageValues. In addition, negative values are observed in several duration attributes (Administrative_Duration, Informational_Duration, and ProductRelated_Duration). These negative values originate from the original dataset recording mechanism, which encoded invalid or unrecorded session durations with negative placeholders. Such values lacked meaningful temporal information and were therefore treated as invalid entries and removed during preprocessing.

Furthermore, the target variable exhibits class imbalance, with revenue-generating sessions (Revenue = 1) accounting for a relatively small proportion of the total observations compared to non-revenue sessions (Revenue = 0). This imbalance is a critical consideration, as it may bias accuracy-oriented evaluation metrics and necessitate the use of complementary metrics, such as F1-score and AUC, for a more reliable performance assessment.

Outliers in duration-related features, especially ProductRelated_Duration, indicate that a small number of sessions involve exceptionally long browsing times. If left untreated, these extreme values may dominate gradient updates during MLP training, thereby negatively affecting convergence stability. Consequently, appropriate preprocessing and normalization steps are required to mitigate their influence on model learning.

Table 1. Summary Statistics

Column	count	mean	std	min	25%	50%	75%	max
Administrative	12316	2.32	3.32	0.00	0.00	1.00	4.00	27.00
Administrative Duration	12316	80.91	176.86	-1.00	0.00	8.00	93.50	3398.75
Informational	12316	0.50	1.27	0.00	0.00	0.00	0.00	24.00
Informational Duration	12316	34.51	140.83	-1.00	0.00	0.00	0.00	2549.38
ProductRelated	12316	31.76	44.49	0.00	700	18.00	38.00	705.00
ProductRelated Duration	12316	1196.04	1914.37	-1.00	185.00	599.77	1466.48	63973.52
BounceRates	12316	0.02	0.05	0.00	0.00	0.00	0.02	0.20
ExitRates	12316	0.04	0.05	0.00	0.01	0.03	0.05	0.20
PageValues	12330	5.89	18.57	0.00	0.00	0.00	0.00	361.76
SpecialDay	12330	0.06	0.20	0.00	0.00	0.00	0.00	1.00
Operating Systems	12330	2.12	0.91	1.00	2.00	2.00	3.00	8.00
Browser	12330	2.36	1.72	1.00	2.00	2.00	2.00	13.00
Region	12330	3.15	2.40	1.00	1.00	3.00	4.00	9.00
TrafficType	12330	4.07	4.03	1.00	2.00	2.00	4.00	20.00

3.2. Data Preprocessing

3.2.1. Missing Value

Handling missing data is an essential step to ensure data quality and model reliability. In this study, rows containing missing values across eight numerical features were removed. These missing entries were sparsely distributed across observations, and the proportion of removed samples was below 0.15% of the total dataset, thereby minimizing the risk of sampling bias or information loss. Figure 3 illustrates the distribution of missing values before and after data cleaning, showing that, prior to preprocessing, each affected feature contained approximately 14 missing entries, whereas after cleaning, all missing values were eliminated. This result confirms that the dataset was successfully cleaned and rendered suitable for subsequent analysis and model training.

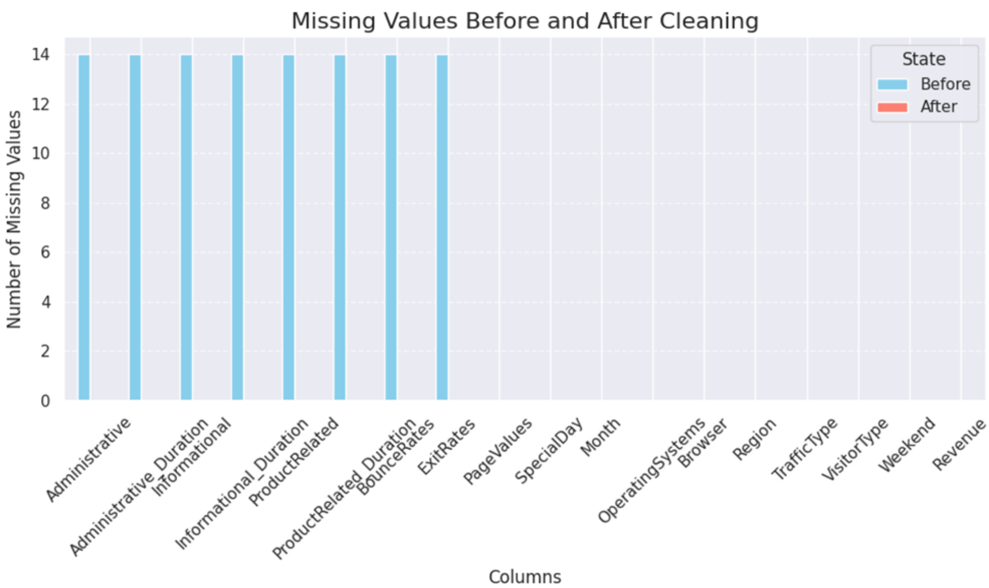

Figure 3. Distribution of missing values before and after data cleaning

Figure 3 compares the number of missing values in each dataset feature before and after the data cleaning process. Prior to cleaning, approximately 14 missing values were observed in each affected numerical feature, as indicated by the blue bars. After preprocessing, all missing values were completely eliminated, as shown by the absence of red bars across all features. This result confirms that the applied data-cleaning strategy successfully removed all incomplete observations, producing a fully cleaned dataset suitable for subsequent exploratory analysis and predictive modeling.

3.2.2. Data Normalization

Despite data cleaning, the dataset still exhibits scale disparities and extreme values, particularly in duration-based features, which may dominate gradient updates during neural network training. To address this issue, interval normalization to the range [-1,1] was applied. Although ReLU activation functions are typically associated with non-negative inputs, normalization to [-1,1] was retained to preserve relative feature distributions and to ensure numerical stability across heterogeneous feature scales. Empirically, this normalization strategy did not degrade model performance and facilitated stable convergence during training, enabling the MLP model to learn robust decision boundaries. The resulting normalized dataset is illustrated in Figure 4.

	Revenue	Administrative	ministrative_Durati	Informational	formational_Duratic	ProductRelated	ductRelated_Durat	BounceRates	ExitRates	PageValues	SpecialDay	
1	FALSE	-0.698	-0.457476	-0.397	-0.245039	-0.6915	-0.624793	3.67263	3.23537	-0.317376	-0.309	Feb
2	FALSE	-0.698	-0.457476	-0.397	-0.245039	-0.6690	-0.59136	-0.457458	1.17459	-0.317376	-0.309	Feb
3	FALSE	-0.698	-0.463131	-0.397	-0.252141	-0.6915	-0.625315	3.67263	3.23537	-0.317376	-0.309	Feb
4	FALSE	-0.698	-0.457476	-0.397	-0.245039	-0.6690	-0.623399	0.575063	1.9989	-0.317376	-0.309	Feb
5	FALSE	-0.698	-0.457476	-0.397	-0.245039	-0.4892	-0.296996	-0.0444492	0.144202	-0.317376	-0.309	Feb
6	FALSE	-0.698	-0.457476	-0.397	-0.245039	-0.2869	-0.544232	-0.131398	-0.380031	-0.317376	-0.309	Feb
7	FALSE	-0.698	-0.463131	-0.397	-0.252141	-0.6915	-0.625315	3.67263	3.23537	-0.317376	1.701	Feb
8	FALSE	-0.397	-0.463131	-0.397	-0.252141	-0.6915	-0.625315	3.67263	3.23537	-0.317376	-0.309	Feb
9	FALSE	-0.698	-0.457476	-0.397	-0.245039	-0.6690	-0.605464	-0.457458	1.17459	-0.317376	3.711	Feb
10	FALSE	-0.698	-0.457476	-0.397	-0.245039	-0.6465	-0.239272	-0.457458	-0.428236	-0.317376	1.701	Feb

Figure 4. Dataset after normalization

Figure 4 illustrates the normalized e-commerce user behavior dataset after applying interval normalization to the range [-1,1]. All numerical features, including page interaction metrics, duration-based variables, and behavioral indicators, were scaled to a common range, while the target variable, Revenue, remained binary. This normalization reduces scale-induced bias during training and supports stable learning in the Multilayer Perceptron (MLP) model.

3.3. Exploratory Data Analysis (EDA)

Exploratory Data Analysis (EDA) was conducted to examine the statistical characteristics and inter-feature relationships of the normalized dataset prior to model development. Pair plots and correlation heatmaps were used to identify distribution patterns, potential feature dependencies, and structural properties that may influence the design of the Multilayer Perceptron (MLP) model.

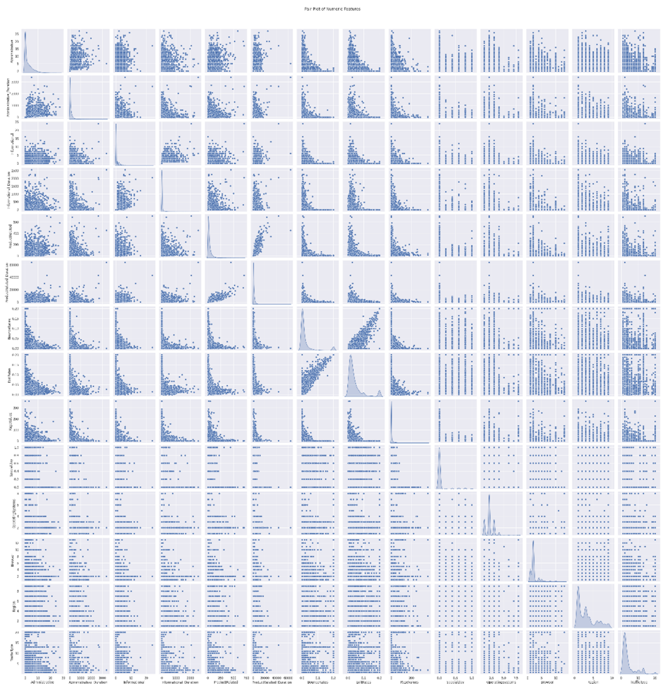

Figure 5. Pair Plot of Numerical Features in the E-Commerce Data

The pair plot in Figure 5 reveals clear structural relationships among numerical features in the e-commerce user behavior dataset. Strong positive associations are observed between page count variables and their corresponding duration features, such as ProductRelated and ProductRelated_Duration, as well as Informational and Informational_Duration. This pattern indicates that increased page visits within a category are consistently associated with longer session durations in that category. In addition, BounceRates and ExitRates exhibit a strong positive correlation, suggesting that users who leave the site early are also more likely to exit without further interaction.

Most numerical features exhibit right-skewed, non-normal distributions, particularly PageValues and ProductRelated_Duration, in which a small number of sessions exhibit extreme values. These distributional characteristics indicate the presence of outliers that may disproportionately influence gradient-based learning if not properly controlled. The observed deviations from normality further justify the application of normalization and regularization strategies in the subsequent modeling stage.

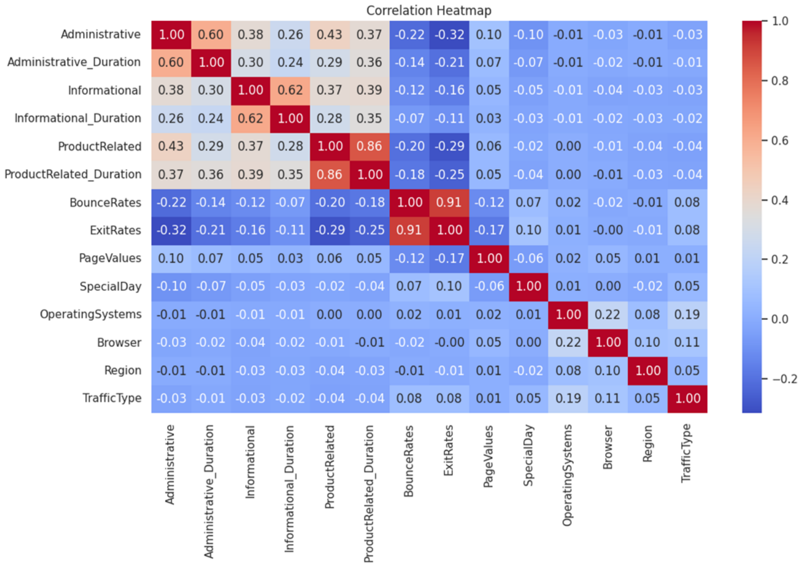

Figure 6. Correlation Heatmap of User Behavior Features in E-Commerce Revenue Prediction

The correlation heatmap in Figure 6 provides a quantitative overview of inter-feature dependencies. High correlation coefficients are observed between page count features and their duration counterparts, including Administrative-Administrative_Duration (0.60), Informational-Informational_Duration (0.62), and ProductRelated–ProductRelated_Duration (0.71). These strong correlations suggest potential feature redundancy, which may increase the risk of overfitting, particularly in deeper MLP architectures. Conversely, features such as OperatingSystems, Browser, Region, and TrafficType exhibit weak correlations with most behavioral variables, indicating that their influence on revenue prediction is likely indirect or context-dependent. Furthermore, the negative correlations between BounceRates and PageValues (-0.39) and between ExitRates and PageValues (-0.20) imply that early exits are associated with lower revenue potential, reinforcing the behavioral relevance of these features for conversion prediction. Overall, insights derived from EDA directly informed the design of the predictive model. The presence of correlated behavioral features motivated the incorporation of regularization mechanisms in the MLP to mitigate redundancy-induced overfitting, while the observed distribution patterns supported the use of normalization to stabilize training. These findings were subsequently translated into architectural and hyperparameter choices in the MLP modeling stage.

3.4. Predictive Model with MLP

Based on insights from the Exploratory Data Analysis (EDA), the Multilayer Perceptron (MLP) model was configured to balance model complexity and generalization. EDA revealed strong correlations between several behavioral features, particularly between page visit counts and their corresponding duration variables, indicating potential feature redundancy. Such redundancy may increase the risk of overfitting, especially when deeper neural network architectures are applied. Therefore, the MLP configurations were intentionally designed with controlled network depth and explicit regularization mechanisms. In this stage, systematic hyperparameter tuning was conducted to evaluate the influence of architectural and optimization choices on revenue prediction performance. The evaluated hyperparameters included the number of hidden layers and neurons, activation functions, optimization algorithms (solvers), regularization strength (alpha), and the maximum number of training iterations. Rather than relying on a single fixed configuration, multiple MLP architectures were tested to enable a performance-driven comparison across different model settings. The complete set of tested configurations is summarized in Table 2. To ensure robustness and reduce the risk of performance bias caused by a single train–test split, each configuration was evaluated using k-fold cross-validation (k = 5). Model performance was reported as the mean value across folds, allowing assessment of both predictive accuracy and stability. Although the absolute performance differences among schemes appear relatively small (approximately 0.6–1%), they can be practically meaningful in large-scale e-commerce environments, where minor improvements can translate into significant revenue impact.

Table 2. Model Configurations

Scheme	Hidden Layer	Activation	Solver	Alpha	Max Iterations	Description
A	50	ReLU	Adam	0.0001	200	Baseline
B	50, 30	ReLU	Adam	0.0001	200	Added a hidden layer
C	50, 40, 30	ReLU	Adam	0.0001	200	More hidden layers
D	50, 40, 30	Tanh	Adam	0.0001	200	Changed activation function
E	50, 40, 30	ReLU	SGD	0.0001	200	Changed solver
F	50, 40, 30	ReLU	Adam	0.1	200	High regularization level
G	50, 40, 30	ReLU	Adam	0.001	200	Moderate regularization level

Scheme A served as the baseline configuration, employing a single hidden layer with 50 neurons, ReLU activation, Adam optimizer, and minimal regularization. Schemes B and C progressively increased network depth to examine whether additional representational capacity improves performance. Scheme D evaluated the effect of replacing ReLU with the Tanh activation function, while Scheme E assessed the impact of using Stochastic Gradient Descent (SGD) instead of Adam. Schemes F and G explored the influence of higher and moderate regularization strengths to mitigate overfitting arising from correlated input features. Although the observed performance differences among schemes were numerically moderate, Scheme E consistently outperformed other configurations across all cross-validation folds, indicating more stable generalization. The superior performance of SGD suggests that controlled, incremental gradient updates help avoid suboptimal local minima in this dataset, particularly under normalized, correlated feature conditions. This result highlights that selecting an optimization strategy can have a greater impact on model performance than architectural depth alone. Overall, this configuration strategy enables the proposed framework to operate in a hyperparameter-adaptive manner, where architectural and optimization decisions are guided by empirical performance rather than predefined assumptions. By explicitly incorporating insights from EDA and validating configurations via cross-validation, the MLP model achieves robust, reliable revenue prediction.

3.5. Model Evaluation

The model evaluation stage assesses the predictive performance and generalization stability of the proposed Multilayer Perceptron (MLP) configurations. Each scheme (A–G) was evaluated using five classification metrics: accuracy, precision, recall, F1-score, and Area Under the Curve (AUC). These metrics were selected to provide a comprehensive assessment of model performance, particularly under potential class imbalance conditions commonly observed in e-commerce conversion data. The evaluation results are summarized in Table 3. Figure 2 shows a simple MLP architecture with an input layer, four hidden layers, and an output layer. Each neuron in the hidden layers processes inputs from the previous layer using weights, biases, and nonlinear activation functions. The final output is produced by two neurons in the output layer. This structure enables the MLP to capture complex patterns in the data. Although the baseline configuration (Scheme A) achieves competitive results and even yields the highest AUC value, its overall classification balance, as reflected in precision and F1-score, is slightly inferior to that of Scheme E. This indicates that while simpler architectures may perform adequately, the optimization strategy and regularization are more decisive in achieving stable, balanced predictions. Furthermore, Schemes B and C, which increase architectural depth, show a gradual decline in performance, confirming that additional layers do not necessarily improve predictive accuracy and may introduce overfitting in the presence of correlated features identified during EDA. The observed performance differences among schemes are numerically moderate (approximately 0.6–1%). However, Scheme E demonstrates consistent superiority across cross-validation folds, indicating stable generalization rather than isolated performance gains. The superior performance of SGD compared to Adam suggests that controlled gradient updates help avoid suboptimal local minima in this dataset, particularly under normalized and correlated feature conditions. This finding highlights that solver selection can have a greater impact on model performance than increasing architectural complexity.

Table 3. Model Evaluation Metrics

Scheme	Accuracy	Precision	Recall	F1-Score	AUC
A	0.883212	0.876529	0.883212	0.878754	0.91079
B	0.876318	0.867982	0.876318	0.870596	0.883128
C	0.86618	0.867233	0.86618	0.866694	0.886349
D	0.857664	0.858641	0.857664	0.858144	0.872312
E	0.889294	0.884048	0.889294	0.885927	0.909978
F	0.864558	0.866165	0.864558	0.865334	0.884775
G	0.859692	0.859966	0.859692	0.859828	0.883058

Figure 7 visualizes the comparative performance of all seven MLP schemes across the five evaluation metrics. Scheme E consistently outperforms other configurations, particularly in accuracy and F1-score, indicating its effectiveness in balancing false positives and false negatives. From a practical e-commerce perspective, this balance is crucial, as misclassification directly affects marketing cost efficiency and conversion targeting strategies. Moreover, the relatively high AUC achieved by Scheme E demonstrates robust discriminative capability across different decision thresholds, which is essential for real-world deployment where operating conditions may vary.

Overall, these results confirm the hyperparameter-adaptive nature of the proposed framework. By systematically evaluating multiple configurations and selecting hyperparameters based on empirical performance rather than predefined assumptions, the framework achieves robust and reliable revenue prediction. The integration of EDA-driven insights, cross-validation evaluation, and multi-metric assessment ensures that the selected model configuration is both statistically stable and practically meaningful for e-commerce applications.

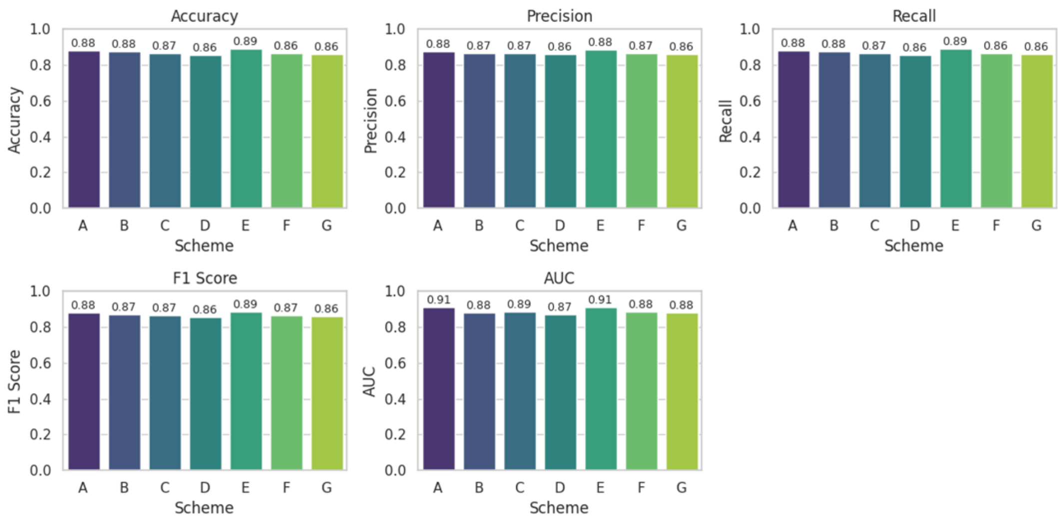

Figure 7. Classifiers overall performance assessment

Figure 7 visualizes the comparative performance of all seven MLP schemes across the five evaluation metrics. Scheme E consistently outperforms other configurations, particularly in accuracy and F1-score, indicating its effectiveness in balancing false positives and false negatives. From a practical e-commerce perspective, this balance is crucial, as misclassification directly affects marketing cost efficiency and conversion targeting strategies. Moreover, the relatively high AUC achieved by Scheme E demonstrates robust discriminative capability across different decision thresholds, which is essential for real-world deployment where operating conditions may vary.

Overall, these results confirm the hyperparameter-adaptive nature of the proposed framework. By systematically evaluating multiple configurations and selecting hyperparameters based on empirical performance rather than predefined assumptions, the framework achieves robust and reliable revenue prediction. The integration of EDA-driven insights, cross-validation evaluation, and multi-metric assessment ensures that the selected model configuration is both statistically stable and practically meaningful for e-commerce applications.

4. CONCLUSION

This study successfully applied the Multi-Layer Perceptron (MLP) model to predict e-commerce revenue from user behavior, demonstrating that hyperparameter optimization significantly improves predictive performance. Through systematic experiments on network architectures, activation functions, solvers, and different levels of regularization, the study revealed that deeper architectures, combined with appropriate regularization and solvers, achieved the best performance across evaluation metrics such as accuracy, precision, recall, F1-score, and AUC. The best model configuration used ReLU activation, the Adam optimizer, and moderate regularization, effectively balancing generalization and accuracy.

These findings emphasize the importance of carefully tuning neural network parameters to capture user behavior patterns that influence purchase decisions. By leveraging behavioral features such as page interactions, session duration, and bounce rate, the MLP proved a powerful tool for predicting revenue on e-commerce platforms. This insight is highly relevant for developing digital marketing strategies, customer segmentation, and personalized recommendations.

Future studies are recommended to explore additional feature engineering techniques, integrate temporal or sequential patterns through recurrent architectures, and apply explainable AI approaches to better understand model decisions transparently. Moreover, validating the model across different e-commerce domains and user demographics will enhance its reliability and broader applicability.

Furthermore, the experimental results confirm that the proposed MLP framework operates in a hyperparameter-adaptive manner. Instead of relying on a fixed network configuration, the framework adaptively determines the most effective hyperparameter setting through iterative performance comparison. This adaptive selection process enhances model robustness and ensures that the chosen configuration is well-aligned with the underlying patterns of e-commerce user behavior influencing revenue generation.

5. ACKNOWLEDGEMENTS

With deepest respect, the authors express their sincere gratitude to Universitas Muhammadiyah Asahan, the primary institution that supported and facilitated this research. The authors also acknowledge the valuable collaboration, academic support, and technical contributions provided by the research team from Universitas Potensi Utama, which significantly contributed to the successful completion of this study.

6. DECLARATIONS

AUTHOR CONTIBUTION

All authors contributed substantially to this research. Safrizal led the conceptualization of the study, research design, and overall framework development. Lili Tanti contributed to data preprocessing, experimental design, and model evaluation. Yan Yang Thanri was responsible for methodology refinement, hyperparameter optimization strategy, and result analysis. All authors contributed to manuscript preparation, reviewed the final version, and approved the manuscript for publication.

FUNDING STATEMENT

This research was supported by Universitas Muhammadiyah Asahan, Kisaran, Indonesia, with collaborative support from Universitas Potensi Utama. The authors gratefully acknowledge the institutional support that enabled the successful completion of this study.

COMPETING INTEREST

The authors declare that they have no competing interests related to this research, its results, or its publication.

REFERENCES

[1] C. Zhan, J. Li, W. Jiang, W. Sha, and Y. Guo, "E-commerce Sales Forecast Based on Ensemble Learning," in *2020 IEEE International Symposium on Product Compliance Engineering-Asia (ISPCE-CN)*, 2020, pp. 1–5, https://doi.org/10.1109/ISPCE-CN51288.2020.9321858.

[2] S. C. Necula and V. D. Păvăloaia, "AI-Driven Recommendations: A Systematic Review of the State of the Art in E-Commerce," vol. 13, no. 9, pp. 1–22, 2023, https://doi.org/10.3390/app13095531.

[3] B. S. Abunasser and S. S. Abu-Naser, "Predicting Customer Revenue in E-commerce Using Machine Learning a Case Study of

the Google Merchandise Store BT - Advances in Intelligent Computing Techniques and Applications," F. Saeed, F. Mohammed, and Y. Fazea, Eds. Cham: Springer Nature Switzerland, June, 2024, pp. 27–38, https://doi.org/10.1007/978-3-031-59711-4_3.

[4] M. Sarkar, E. H. Ayon, M. T. Mia, R. K. Ray, M. S. Chowdhury, B. P. Ghosh, M. Al-Imran, M. T. Islam, and M. Tayaba, "Optimizing E-Commerce Profits: A Comprehensive Machine Learning Framework for Dynamic Pricing and Predicting Online Purchases," *Journal of Computer Science and Technology Studies*, vol. 5, no. 4, pp. 186–193, 2023, https://doi.org/10.32996/jcsts.2023.5.4.19.

[5] X. Zhang, F. Guo, T. Chen, L. Pan, G. Beliakov, and J. Wu, "A Brief Survey of Machine Learning and Deep Learning Techniques for E-Commerce Research," vol. 18, no. 4, pp. 2188–2216, 2023, https://doi.org/10.3390/jtaer18040110.

[6] Z. Li, X. Shen, Y. Jiao, X. Pan, P. Zou, X. Meng, C. Yao, and J. Bu, "Hierarchical bipartite graph neural networks: Towards large-scale E-commerce applications," in *Proceedings - International Conference on Data Engineering*, vol. 2020-April, 2020, pp. 1677–1688, https://doi.org/10.1109/ICDE48307.2020.00149.

[7] P. Ganguly and I. Mukherjee, "Enhancing Retail Sales Forecasting with Optimized Machine Learning Models," in *2024 4th International Conference on Sustainable Expert Systems (ICSES)*, October, 2024, pp. 884–889, https://doi.org/10.1109/ICSES63445.2024.10762950.

[8] D. M. Petroșanu, G. Pîrjan, Alexandru, A. Tăbușcă, D. L. Zirra, and A. Perju-Mitran, "E-Commerce Sales Revenues Forecasting by Means of Dynamically Designing, Developing and Validating a Directed Acyclic Graph (DAG) Network for Deep Learning," *Electronics (Switzerland)*, vol. 11, no. 18, pp. 1–35, 2022, https://doi.org/10.3390/electronics11182940.

[9] S. Jayanthi, T. Suvarna Kumari, S. Inturi, B. Nathan, M. A. J. Sathya, and K. Karmakonda, "Predicting E-Commerce Revenue with SHAP Insights: A Comparative Study of SMOTE-Enhanced Machine Learning Models," vol. 35, no. 4, pp. 115 – 135, 2025, https://doi.org/10.52783/pmj.v35.i4s.4631.

[10] Mohammed Aljbour & İsa Avcı, "Sales Prediction in E-Commerce Platforms Using Machine Learning," in *In International Conference on Forthcoming Networks and Sustainability in the AIoT Er.* Switzerland: Springer, Cham, jun 2024, pp. 207–216, https://doi.org/10.1007/978-3-031-62881-8_17.

[11] A. Valencia-Arias, H. Uribe-Bedoya, J. D. González-Ruiz, G. S. Santos, E. C. Ramírez, and E. M. Rojas, "Artificial intelligence and recommender systems in e-commerce. Trends and research agenda," *Intelligent Systems with Applications*, vol. 24, p. 200435, December, 2024, https://doi.org/10.1016/j.iswa.2024.200435.

[12] C. O. Sakar, S. O. Polat, M. Katircioglu, and Y. Kastro, "Real-time prediction of online shoppers' purchasing intention using multilayer perceptron and LSTM recurrent neural networks," *Neural Computing and Applications*, vol. 31, no. 10, pp. 6893–6908, 2019, https://doi.org/10.1007/s00521-018-3523-0.

[13] S. Chatterjee, R. S. Mishra, S. Raichandani, and P. Joshi, "Response Prediction and Ranking Models for Large-Scale Ecommerce Search," in *Springer Proceedings in Business and Economics*, June, 2021, pp. 199–218, https://doi.org/10.1007/978-981-33-6656-5_17.

[14] S. A. Rather, P. S. Bala, and P. L. Ashokan, "Training Multi-layer Perceptron Using Hybridization of Chaotic Gravitational Search Algorithm and Particle Swarm Optimization," in *International Series in Operations Research and Management Science*, May, 2021, vol. 306, pp. 233–262, https://doi.org/10.1007/978-3-030-70281-6_13.

[15] L. Gajic, D. Cvetnic, M. Zivkovic, T. Bezdan, N. Bacanin, and S. Milosevic, "Multi-layer Perceptron Training Using Hybridized Bat Algorithm," pp. 689–705, June, 2021, https://doi.org/10.1007/978-981-33-6862-0_54.

[16] A. A. Heidari, H. Faris, I. Aljarah, and S. Mirjalili, "An efficient hybrid multilayer perceptron neural network with grasshopper optimization," *Soft Computing*, vol. 23, no. 17, 2019, https://doi.org/10.1007/s00500-018-3424-2.

[17] M. Garouani and M. Bouneffa, "Automated machine learning hyperparameters tuning through meta-guided Bayesian optimization," *Progress in Artificial Intelligence*, pp. 1–10, 2024, https://doi.org/10.1007/s13748-023-00311-y.

[18] O. R. Vincent, A. S. Makinde, and A. T. Akinwale, "A cognitive buying decision-making process in B2B e-commerce using Analytic-MLP," *Electronic Commerce Research and Applications*, vol. 25, pp. 59–69, 2017, https://doi.org/10.1016/j.elerap.2017.08.002.

[19] K. Kelly Isyanta, "Application of Multilayer Perceptron for Digital Society Sentiment Analysis towards Indonesia Biggest E-commerce Platforms," *International Journal of Advanced Trends in Computer Science and Engineering*, vol. 9, no. 4, 2020, https://doi.org/10.30534/ijatcse/2020/171942020.

[20] F. M. M. Mokbal, W. Dan, A. Imran, L. Jiuchuan, F. Akhtar, and W. Xiaoxi, "MLPXSS: An Integrated XSS-Based Attack Detection Scheme in Web Applications Using Multilayer Perceptron Technique," *IEEE Access*, vol. 7, 2019, https://doi.org/10.1109/ACCESS.2019.2927417.

[21] A. Ahmed, K. Saleem, O. Khalid, and U. Rashid, "On deep neural network for trust aware cross domain recommendations in E-commerce," *Expert Systems with Applications*, vol. 174, July, 2021, https://doi.org/10.1016/j.eswa.2021.114757.

[22] K. Madhusudhanan, S. Jawed, and L. Schmidt-Thieme, "Hyperparameter Tuning MLP's for Probabilistic Time Series Forecasting BT - Advances in Knowledge Discovery and Data Mining," D.-N. Yang, X. Xie, V. S. Tseng, J. Pei, J.-W. Huang, and J. C.-W. Lin, Eds. Singapore: Springer Nature Singapore, April, 2024, pp. 264–275, https://doi.org/10.1007/978-981-97-2266-2_21.

[23] B. Bischl, M. Binder, M. Lang, T. Pielok, J. Richter, S. Coors, J. Thomas, T. Ullmann, M. Becker, A. L. Boulesteix, D. Deng, and M. Lindauer, "Hyperparameter optimization: Foundations, algorithms, best practices, and open challenges," vol. 13, no. 2, 2023, https://doi.org/10.1002/widm.1484.

[24] E. Phaisangittisagul, "An Analysis of the Regularization Between L2 and Dropout in Single Hidden Layer Neural Network," in *Proceedings - International Conference on Intelligent Systems, Modelling and Simulation, ISMS*, 2016, pp. 174–179, https://doi.org/10.1109/ISMS.2016.14.

[25] X. Xie, M. Xie, A. J. Moshayedi, and M. H. Noori Skandari, "A Hybrid Improved Neural Networks Algorithm Based on L2 and Dropout Regularization," *Mathematical Problems in Engineering*, vol. 2022, no. 1, pp. 1–19, 2022, https://doi.org/10.1155/2022/8220453.

[26] Z. Farhadi, H. Bevrani, and M. R. Feizi-Derakhshi, "Combining Regularization and Dropout Techniques for Deep Convolutional Neural Network," in *IEEE Global Energy Conference, GEC 2022*, December, 2022, pp. 335–339, https://doi.org/10.1109/GEC55014.2022.9986657.

[27] D. C. Gkikas and P. K. Theodoridis, "Predicting Online Shopping Behavior: Using Machine Learning and Google Analytics to Classify User Engagement," *Applied Sciences (Switzerland)*, vol. 14, no. 23, pp. 1–31, dec 2024, https://doi.org/10.3390/app142311403.