



Personal Recommender Model and Predicting Consumer Behavior in Digital Marketing Based on Deep Learning

A. Mirzaee^{1,*}, M. Zeynali², A. Ghorbanzadeh³, P. Ghorbanzadeh⁴

¹ Department of computer engineering, Afagh Institute of Higher Education, Urmia, Iran

² Department of computer engineering, Islamic Azad University, Urmia Branch, Urmia, Iran

³ Department of computer engineering, Kamal Institute of Higher Education, Urmia, Iran

³ Department of computer engineering, Urmia University of Technology, Urmia, Iran

ARTICLE INFO	ABSTRACT
<p>Article History: Received 30 March 2024 Received in revised form 9 June 2024 Accepted 4 September 2024 Available online 13 September 2024</p>	<p>This study presents the design and evaluation of a personalized recommender system aimed at predicting consumer behavior within digital marketing environments. The proposed model integrates the strengths of Long Short-Term Memory (LSTM) networks and Recurrent Neural Networks (RNNs) to effectively process and learn from sequential data. LSTM units are employed to capture long-range temporal dependencies in user interactions, while RNN layers provide a framework for processing dynamic sequences of consumer activity over time. The model is trained using a dataset derived from real-world digital marketing platforms, which includes detailed logs of user preferences, purchase history, and browsing patterns. By learning from these behavioral indicators, the hybrid LSTM-RNN model is able to generate highly personalized recommendations tailored to individual consumers. Experimental results indicate that the proposed architecture achieves a high level of predictive accuracy, outperforming traditional recommendation methods in several key performance metrics such as precision, recall, and F1-score. These findings underscore the effectiveness of deep learning approaches in modeling complex consumer behaviors and highlight the potential of neural network-based recommender systems in optimizing marketing campaigns and enhancing user engagement. Ultimately, this research contributes to the advancement of intelligent recommendation technologies in the digital marketing domain, offering practical implications for businesses aiming to deliver more targeted and responsive customer experiences.</p>
<p>Keywords: Personal Recommender Model, Predicting Consumer Behavior, Digital Marketing, LSTM, RNN, Deep Learning</p>	

1. INTRODUCTION

In the realm of digital marketing, there is a growing need for personalization and predicting consumer behavior to enhance user experiences and drive marketing strategies effectively. Presently, many businesses struggle to tailor their recommendations to individual consumers due to limited understanding of their preferences and behavior patterns. This limitation results in inefficient customer targeting, reduced customer engagement, and missed revenue

* Corresponding Author: ashraf.mirzaee@gmail.com

Department of computer engineering, Afagh Institute of Higher Education, Urmia, Iran



opportunities. To address this challenge, a personal recommender model along with accurate prediction of consumer behavior is required. Leveraging the advancements in deep learning and neural networks, the combination of Long Short-Term Memory (LSTM)[1] and Recurrent Neural Network (RNN)[2] techniques is expected to offer valuable insights into consumer preferences and behavior, thus enabling businesses to deliver personalized recommendations.

However, developing an efficient personal recommender model and accurately predicting consumer behavior using LSTM and RNN pose several challenges. First, acquiring and preprocessing large-scale, heterogeneous consumer behavior data from various digital marketing channels (e.g., website browsing history, purchase records, social media interactions) is a complex task. This involves extracting relevant features and cleaning the data to ensure its quality and integrity. Additionally, the integration and optimization of the LSTM and RNN models to effectively handle the dynamic and sequential nature of consumer behavior patterns is a non-trivial task. Balancing model complexity and accuracy while avoiding overfitting or underfitting is crucial. Furthermore, training the model on extensive datasets with nuanced consumer behaviors, such as time-dependent preferences and evolving trends, requires significant computational resources and efficient algorithms. Moreover, interpretability and explainability of the recommender system's predictions are essential. The black-box nature of deep learning models often hinders the understanding of how and why specific recommendations are generated. Addressing this challenge involves developing methodologies to incorporate transparency and interpretability into the model, thus enabling businesses to build trust with consumers and justify the recommendations made. Lastly, the real-time implementation and scalability of the personal recommender model pose significant challenges. Deploying the model on large-scale digital platforms demands efficient algorithms and infrastructure to handle a high volume of concurrent user interactions while ensuring minimal latency.

In summary, the problem to be addressed is the development of a personal recommender model utilizing LSTM and RNN techniques that accurately predicts consumer behavior in digital marketing. This involves overcoming challenges related to data acquisition and preprocessing, model integration and optimization, interpretability and explainability, and real-time deployment and scalability. Solving these challenges would empower businesses to deliver personalized recommendations, resulting in enhanced user experiences, increased customer engagement, and improved marketing strategies. The objectives of the research are:

- To develop a personal recommender model based on LSTM-RNN for digital marketing purposes.
- To investigate the effectiveness and accuracy of the LSTM-RNN model in predicting consumer behavior in digital marketing.
- To analyze the factors influencing consumer behavior in digital marketing and identify patterns that can be utilized in the LSTM-RNN model.
- To explore the impact of personalized recommendations on consumer engagement and purchasing behavior in digital marketing.
- To evaluate the performance and compare the LSTM-RNN model with other existing models in terms of accuracy and prediction capability for consumer behavior in digital marketing.
- To identify the limitations and challenges of using LSTM-RNN for predicting consumer behavior in digital marketing and propose potential solutions or improvements.
- To provide practical recommendations for businesses and marketers on utilizing the LSTM-RNN model for personalized recommendations and improving overall consumer experiences in digital marketing.

The structure of the article is as follows: in the second section, we will review related works. In section 3, we explain the concepts of LSTM and RNN. In section 4, the proposed method is described. In section 5, the results are evaluated, and then in section 6, conclusions and future suggestions are expressed.

2. LITERATURE REVIEW

In [3], emphasizes that many of the manual and repetitive tasks of a marketer's life can be replaced by artificial intelligence, and that using machines to work with humans is the key to better marketing results. The challenges and

ethical aspects leading to slow or no adoption of AI have been addressed, and one of the main obstacles is that humans still do not trust the technology and are not yet ready for this cultural change. Based on these findings, decision makers and business managers should prepare their companies and employees to implement artificial intelligence in marketing.

In [4], by analyzing recent articles related to machine learning and artificial intelligence in digital marketing, this paper provides a systematic review of recent progress in the work at hand. Apart from examining the legal and ethical context in which these techniques are being developed, they also explored the challenges that artificial intelligence (AI) will present in the future of digital marketing.

In [5], aims to predict customer behavior in digital marketing. A well-performing optimization algorithm called particle swarm optimization (PSO) is used to train the Neural Network (NN) and compared with several other machine learning algorithms to validate the effectiveness of the proposed model. PSO is a computational method that optimizes a problem by iteratively trying to improve a candidate solution according to a given quality criterion. In this paper, the accuracy of the developed PSO-NN achieves the best results compared to the conventional NN in any number of hidden neurons. Considering the number of hidden neurons as 25, the accuracy of the introduced PSO-NN was 4.2% of the advanced NN. Therefore, it is confirmed that the introduced PSO-NN performs well in predicting customer behavior under digital marketing.

In [6], the aim is to investigate consumer behavior in digital platforms and how machine learning affects customer loyalty. The researcher adopted a qualitative method to understand consumer behavior and how machine learning (ML) contributes to customer loyalty. A correlational research design was used to conduct this research. In addition, secondary data is used in the study to obtain reliable results. The researcher analyzes all journals in the general marketing literature. In addition, this research creates a less complicated process for predicting consumer purchasing behavior on digital platforms. It has been found that customer loyalty can be tracked using machine learning and the ability to collect customer preference data on digital platforms. In addition, machine learning is found to be better at identifying customer loyalty.

In [7], suggests that this strategy is one of the main areas of success for this business model, as customers who join following recommendations are more active and therefore more profitable and more loyal to the brand. However, new users referred to these sites vary greatly in terms of the number of transactions they complete on the site. This study advances research on designing recommendation-based digital marketing strategies by providing companies with a predictive model. The model uses data science, including machine learning and big data methods, to personalize financial incentives for users based on the quality of new customers who visit the Cashback website. Therefore, companies can optimize and maximize the return on their marketing investment.

In [8], investigates the analysis of social media data using machine learning tools. This new approach to social media marketing strategy development uses the Waikato Environment for Knowledge Analysis (WEKA). WEKA is compared with other algorithms of interest and found to outperform its peers, especially with respect to parameters such as precision, recall, and F-measure, indicating that WEKA outperforms other methods.

In [9], investigated the impact of artificial intelligence on brand preference for retail banks in Hong Kong. Structural equation modeling was used to analyze 300 responses collected from a questionnaire survey of Generation Z individuals. Findings suggest that AI marketing efforts influence brand experience, brand preference, and repurchase intention. Among AI marketing efforts, information, accessibility, and customization influenced brand experience, while engagement did not significantly influence it. Brand experience also mediates the relationship between AI marketing efforts and brand preference. This study will help micro banks to design AI marketing activities and formulate better marketing and branding strategies for customer attraction and retention.

In [10], aims to develop a comprehensible data mining forecasting approach, in order to achieve two goals: (1) provide a highly accurate and robust demand forecasting model for refurbished products. and (2) to clarify the nonlinear effect of online market factors as predictors of customer demand. Based on Amazon's real data set, the results show that forecasting the demand of refurbished product is a complex and nonlinear problem, and by using advanced machine learning techniques, our proposed approach can predict product demand with high accuracy. In terms of practical implications, the importance of market factors is ranked according to their demand forecasting power, while their effects on demand are analyzed through their partial dependence diagrams. Several insights for

management are revealed by a thorough comparison of the sales impact of these market factors on remanufactured and new products.

The integration of a personalized recommender model based on Long Short-Term Memory (LSTM) and Recurrent Neural Networks (RNN) offers a promising solution to several persistent challenges in the digital marketing landscape. One of the foremost issues addressed by this approach is information overload. In the era of big data, consumers are frequently overwhelmed by the volume of online content, making it difficult to locate information that aligns with their individual interests. The proposed model effectively mitigates this problem by delivering tailored recommendations derived from users' historical behaviors and preferences.

Another critical challenge is user engagement. Sustaining consumer interest in marketing content is vital for brand visibility and campaign success. By accurately forecasting user behavior, the LSTM-RNN model facilitates the creation of dynamic, personalized campaigns that captivate users and encourage continued interaction. This, in turn, directly influences conversion rates. By identifying patterns in consumer preferences, marketers can recommend products or services that better align with users' needs, significantly enhancing the likelihood of converting leads into customers.

Additionally, the model contributes to customer retention and brand loyalty by enabling the delivery of consistently relevant content and personalized experiences. Such interactions foster long-term relationships between consumers and brands. In terms of advertising, the recommender system refines ad targeting by ensuring that users receive content tailored to their specific interests, which not only improves campaign efficiency but also reduces ad fatigue.

Finally, improved customer satisfaction is achieved by aligning digital experiences with user expectations and preferences. Collectively, the proposed LSTM-RNN framework addresses several key limitations in current digital marketing strategies, offering a comprehensive, intelligent solution that enhances personalization, engagement, and overall effectiveness.

3. CONCEPTS

3.1. Long Short-Term Memory (LSTM): Architecture and Applications

Long Short-Term Memory (LSTM) networks are a specialized form of Recurrent Neural Networks (RNNs) specifically designed to address the limitations of traditional RNNs in learning long-term dependencies. One of the major drawbacks of conventional RNNs is the vanishing and exploding gradient problem, which hinders their ability to retain information across lengthy input sequences. LSTMs effectively overcome this issue through the use of memory cells and gating mechanisms, enabling them to preserve and manipulate information over extended time intervals. This makes them highly suitable for sequential and time-series data processing.

An LSTM unit is composed of the following key components:

- **Cell State:** This is the central memory element that runs through the entire sequence. It serves as a conduit for information, regulated by three gates: the forget gate, input gate, and output gate.
- **Forget Gate:** This gate decides what information should be discarded from the cell state. It takes the previous hidden state and the current input, applies a sigmoid activation function, and generates a **forget vector**. This vector, when multiplied element-wise with the previous cell state, removes irrelevant information.
- **Input Gate:** This gate determines which new information should be added to the cell state. It uses a sigmoid function to produce an input vector, and a tanh activation to create a candidate vector. The combination of these vectors allows the cell to selectively incorporate relevant data.
- **Update Mechanism:** The cell state is updated by combining the outputs of the forget and input gates. Specifically, the old cell state is multiplied by the forget vector, and the product of the input vector and the candidate vector is added, producing the new cell state.

- **Output Gate:** This gate controls the information to be output as the new hidden state. It uses the current input and previous hidden state to generate an output vector via a sigmoid function. The updated cell state is then passed through a tanh function and element-wise multiplied by the output vector to produce the hidden state.

LSTM networks have proven effective across a range of sequential learning tasks, such as speech recognition, language modeling, machine translation, and sentiment analysis. Their ability to maintain long-term contextual information enables superior performance in domains that require understanding of long-range dependencies.

3.2. Recurrent Neural Networks (RNN)

Recurrent Neural Networks (RNNs) represent a class of artificial neural networks designed to process sequential or time-series data, making them particularly effective for tasks where temporal or contextual relationships are critical—such as speech recognition, language translation, handwriting recognition, and sentiment analysis.

The hallmark feature of RNNs lies in their recurrent connections, which enable the network to retain and utilize information from previous time steps. This memory-like mechanism allows RNNs to capture temporal dependencies by passing the hidden state from one time step to the next, thereby allowing past information to influence current and future computations.

The fundamental computational unit in an RNN is the recurrent cell. Among the most widely used variants are:

- Long Short-Term Memory (LSTM) cells, which are capable of learning both short-term and long-term dependencies, and
- Gated Recurrent Units (GRUs), which offer a simplified alternative to LSTMs while maintaining comparable performance.

RNNs process sequences in a step-by-step manner: at each time step, the current input is combined with the hidden state from the previous step. This combination is passed through a non-linear activation function, allowing the model to learn complex temporal patterns. The resulting hidden state is then propagated forward to the next time step, effectively summarizing the sequence information seen so far.

Training RNNs involves minimizing a loss function using optimization algorithms such as Backpropagation Through Time (BPTT), which extends standard backpropagation to sequential data by unfolding the network over time. However, RNNs are prone to the vanishing or exploding gradient problem, especially with long sequences. This issue can hinder convergence and learning efficiency.

To address these challenges, several techniques are commonly employed, including:

- Gradient clipping to prevent gradient explosion,
- Careful weight initialization, and
- The adoption of LSTM or GRU architectures, which are more resilient to these training difficulties.

In summary, RNNs are powerful tools for modeling sequences due to their ability to capture and propagate contextual information over time. Their extensive use across natural language processing, time-series forecasting, and sequential decision-making highlights their foundational role in modern machine learning.

4. PROPOSED METHOD

4.1. Methodology

The proposed approach for consumer behavior prediction in digital marketing using an LSTM-RNN model comprises the following steps:

a. Data Collection and Preparation

The process begins with assembling a comprehensive dataset that encapsulates various aspects of consumer behavior derived from digital marketing campaigns. Key variables may include customer demographics, purchase history, browsing patterns, and responses to promotional activities. Prior to modeling, the dataset undergoes preprocessing to ensure cleanliness, consistency, and appropriate formatting.

b. Defining the Target Variable

We identify the specific consumer behavior to be predicted—such as purchase likelihood or click-through rate—which serves as the target variable for training the model.

c. Feature Engineering

Relevant features are extracted to enhance the model's predictive capacity. This includes generating temporal features such as time of day, day of the week, or seasonal indicators that may influence consumer activity. For example, user engagement may vary between weekdays and weekends or across different times of day.

d. Data Splitting

The dataset is partitioned into training and testing subsets. The training set is used to develop the LSTM-RNN model, while the testing set serves to evaluate its predictive performance.

e. Data Preprocessing

To facilitate efficient training, feature values are normalized or scaled, ensuring that all input variables fall within a similar range. This step helps prevent dominance of features with larger magnitudes and supports stable convergence during training.

4.2. LSTM-RNN Model Architecture Design

The architecture of the LSTM-RNN model is constructed with careful consideration of the following components:

- **Number of Hidden Layers:** Initially, one or two hidden layers are implemented. The depth can be increased based on task complexity to capture intricate patterns in the data.
- **LSTM Units per Layer:** The number of LSTM units determines the memory and representational power of each layer. While more units may capture finer-grained patterns, they also increase computational cost. Hyperparameter tuning is employed to balance performance and efficiency.
- **Dropout Rate:** To mitigate overfitting, dropout regularization is introduced, randomly deactivating a fraction of LSTM units during training. Typical dropout rates range from 0.2 to 0.5, and optimal values are determined empirically.
- **Bidirectional LSTM:** To improve contextual understanding, bidirectional LSTM layers are used to process input sequences in both forward and backward directions. This bidirectional structure enables the model to leverage information from both past and future time steps, enhancing prediction accuracy.
- **Additional Layers and Techniques:** Fully connected (dense) layers may be added before or after the LSTM layers to refine feature representations. Techniques such as batch normalization and early stopping are also applied to stabilize training and prevent overfitting.

4.3. Model Training

The training process involves feeding the preprocessed data into the LSTM-RNN model and optimizing its parameters using an appropriate algorithm, such as Stochastic Gradient Descent (SGD) or its variants. The training continues iteratively until the model reaches a satisfactory level of performance on the training set.

4.4. Model Evaluation

The trained model is validated using the test set to assess its predictive effectiveness. Performance metrics such as accuracy, precision, recall, and F1-score are computed to provide a comprehensive evaluation of the model's ability to forecast consumer behavior in a digital marketing context.

5. EVALUATION RESULTS

We use Online Retail dataset. This dataset contains transactional data from an online retail store. It includes customer information, product details, and purchase history, which can be used to analyze consumer behavior in terms of browsing, purchase patterns, and segmentation. We can download it here:

<https://archive.ics.uci.edu/ml/datasets/online+retail>.

This dataset has 541909 records with 8 features. In this research, we chose the KFOLD method with 4 repetitions as the training and test data allocation for the data set. The method of data selection is also random. The k-fold cross-validation method is commonly used in machine learning to evaluate the performance of predictive models, including recommender systems. The k-fold method ensures that the available data is fully used for training and testing. Each sample in the dataset gets a chance to appear in the training and test sets in different iterations. This helps capture a wider range of patterns and variations in the data, resulting in more robust and accurate assessments. The software used in this research is MATLAB version 2020. The settings of LSTM-RNN parameters are shown in Table 1. To evaluate the proposed method (LSTM- RNN), we compare it with RNN, LSTM methods and the article (Reddy et al., 2020)[5]. For each of k, two confusion and ROC charts have been considered and evaluation criteria have been calculated. For the correctness of the model, it is necessary to divide the available data into three parts: training, testing and validation. They generate the data of the training section and the data of the testing section tests the data with the help of a number of records. The data of the validation section also check the accuracy of the generated model. There are different indicators to evaluate the accuracy, which can be called sensitivity, transparency, accuracy, and accuracy. The accuracy rate of a method on the training data set is the percentage of observations from the training set that are correctly classified by the method used. Test data are used to calculate this index. It is also possible to calculate the error rate or incorrect classification based on the accuracy index. To calculate the accuracy of the model, the confusion matrix can be used (Figure 1). This matrix is a useful tool for analyzing how the classification method works in detecting data or observations of different categories.

Where TP is the number of fault modules, which are correctly predicted to be fault-prone; FP is the number of error-free modules, which are incorrectly predicted as error-prone; TN is the number of error-free modules, which are correctly predicted as error-prone, and finally, FN is the number of error-free modules, which are incorrectly predicted as error-prone. They have been seen.

		Actual Class	
		Positive (P)	Negative (N)
Predicted Class	Positive (P)	True Positive (TP)	False Positive (FP)
	Negative (N)	False Negative (FN)	True Negative (TN)

Fig. 1. The confusion matrix

Table 1. LSTM-RNN parameters

Parameters
Number of Layers: 2-3
Number of LSTM Units: 50-100
Input Size: 8
Sequence Length: 10-50
Dropout Rate: 0.2-0.5
Batch Size: 32-128
Number of hidden layers: 1-2

6. EVALUATION METRICS

6.1. Misclassification rate (MR):

The most common scale for evaluating the performance of error prediction models is MR, which is the ratio of the number of incorrectly classified modules to the total number of modules. From the confusion matrix, MR can be obtained according to the following relation:

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

6.2. The cost of misclassification (ECM):

ECM is a standard for comparing the efficiency of different classification models. A function of the error cost in classification based on the classification error of type | (Err|) (an npf module is classified as fp) and type classification error || (Err||) (an fp module is classified as npf), which the ECM uses to calculate the ratio of these different costs. The evaluation of classification models is very important in presenting the difference in the ratio of costs, because the usefulness of a model depends on the cost of its wrong classification. from the confusion matrix of errors Err| and Err|| can be obtained:

$$EE_{||} = \frac{FN}{TP+FN} \quad EE_{|} = \frac{FP}{TN+FP} \quad (2)$$

Because the cost of these errors is different and we need a single measure of cost, we use ECM as a result. which is calculated according to the following relationship:

$$ECM = C_{|}Err_{|}P_{ndf} + C_{||}Err_{||}P_{df} \quad (3)$$

C and C|| Costs related to errors $Err_{|}$ and $Err_{||}$ and P_{df} and P_{ndf} are the probabilities of error-prone and error-prone modules, respectively.

Normalized cost of misclassification (NECM):

In many cases, it is not possible to obtain the cost of each misclassification separately, so the normalized (NEMC)EMC is used:

$$NECM = Err_{|}P_{ndf} + \frac{C_{|}}{C_{||}}Err_{||}P_{df} \quad (4)$$

6.3. Sensitivity (S):

This measure shows the accuracy of the prediction model and is defined as the percentage of classes that are correctly predicted prone to error.

$$sensitivity = \frac{\text{Number of moduls correctly predicted as fault prone}}{\text{total number of actual faulty module}} * 100 \quad (5)$$

6.4. Specificity

This criterion, like sensitivity, shows the accuracy of the prediction model, which is defined as the percentage of classes that are correctly predicted and not prone to error:

$$Specificity = \frac{\text{Number of moduls correctly predicted as non fault prone}}{\text{total number of actual non faulty module}} * 100 \quad (6)$$

6.5. Accuracy (Ace)

Accuracy is defined as the number of correctly predicted classes (including errors and errors) to the total number of classes:

$$Acc = \frac{TP+TN}{TP+TN+FP+FN} \quad (7)$$

6.6. Precision (P)

Accuracy shows the number of error-prone classes that are correctly predicted as error-prone by the model. Its best value is 1. More accuracy, i.e. less FP (error-free elements misclassified as error-prone):

$$Precision = \frac{TP}{TP+FP} \quad (8)$$

6.7. Recall

Recall indicates the number of error-prone classes predicted by the model as error-prone. Its best value is 1. High recall means less FN:

$$Recall = \frac{TP}{TP+FN} \quad (9)$$

6.8. F-measure

The F-measure (or F-score) is a metric that combines precision and recall into a single performance score, providing a balanced evaluation of a classification model. It is particularly useful when there is an uneven class distribution or when both false positives and false negatives carry significant consequences. The F-measure is calculated as the harmonic mean of precision and recall:

$$F - measure = \frac{2 \times Precision \times Recall}{Precision + Recall} \quad (10)$$

This metric assumes equal importance (weight) for precision and recall when the weighting factor $\alpha=1$. The *F-measure* ranges from 0 to 1, with values closer to 1 indicating better classification performance.

6.9. Consistency:

Consistency is a metric that reflects the stability and reliability of a model's predictions. A higher consistency value indicates a more accurate and dependable model. When the number of true positives (TP) equals the sum of true positives and false negatives (TP + FN), consistency reaches its maximum value of 1.

The consistency metric is computed using the following formula:

$$Consistency = \frac{dn - k^2}{k(n - k)} \quad (11)$$

where:

- d is the number of classes predicted as error-prone by the model in each dataset,
- k (i.e., TP) represents the actual number of error-prone classes in the dataset, and
- n is the total number of instances (i.e., $k=TP+FN$).

This formulation allows researchers to assess the alignment between predicted and actual distributions of error-prone classes, offering a complementary perspective to conventional performance metrics.

6.10. Analysis of the receiver operating characteristic (ROC)

ROC analysis is an effective method in evaluating the performance of the prediction model. The ROC curve is defined as a plot of sensitivity on the Y coordinate and against it 1 feature on the X coordinate. While the ROC curve is being constructed, we select a number of cut-off points between 0 and 1 and calculate the sensitivity and specificity at each cut-off point. The ROC curve is used to obtain the desired cut-off point that maximizes both sensitivity and specificity.

The ROC curve depicts the benefits of using the model versus the costs of using the model at different threshold values. In fact, the ROC curve allows us to evaluate the performance of a prediction model in general without considering any specific cut-off value.

6.11. Area under the ROC curve (AUC)

The area under the ROC curve can be inferred as a statistical descriptor to estimate whether a prediction model is more likely to detect a sample with error than a sample without error. An AUC of less than 0.5 means that the TP rate is very low. Therefore, we also use AUC to evaluate the performance of the predictors.

6.12. Balance (B)

Balance is a standard for measuring accuracy by calculating the Euclidean distance of the correct classification rate from the best value, which is 1, and is often used by software engineers in practice:

$$B= 1 - \sqrt{\frac{1}{2} \left(\frac{FN}{TP+FN}\right)^2 + \left(\frac{FP}{TP+FP}\right)^2} \tag{12}$$

7. EXPERIMENTS RESULTS

The confusion matrix for the proposed method with RNN, LSTM methods and the article (Reddy et al., 2020) for class 1: purchase and class 2: non-purchase is shown in figures 2 to 5 [5].

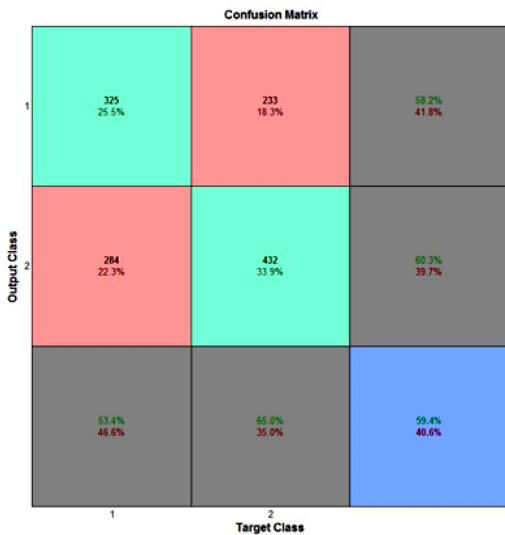


Fig. 3. LSTM

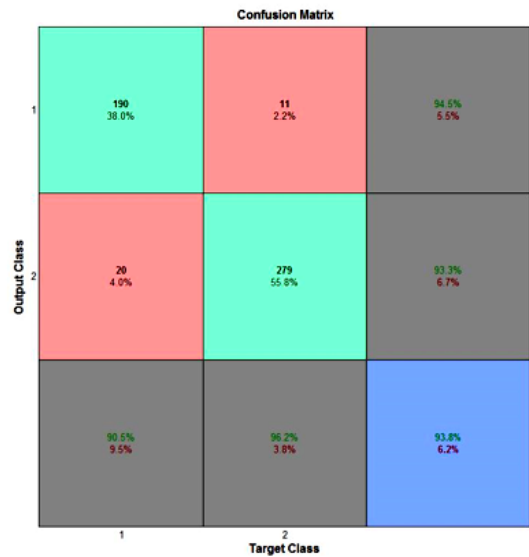


Fig. 2. The proposed method

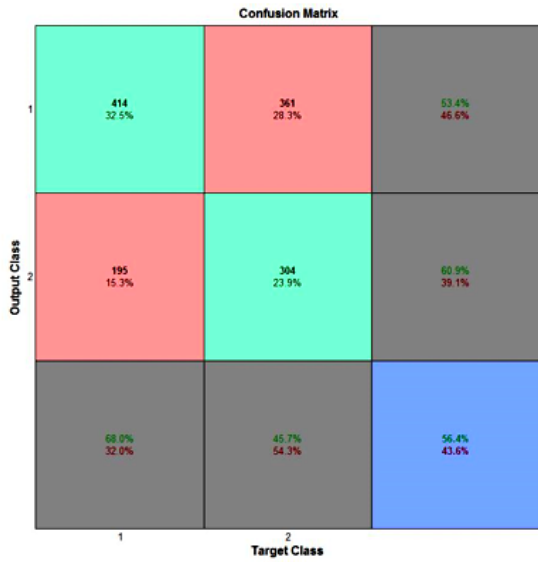


Fig. 5. (Reddy et al., 2020)

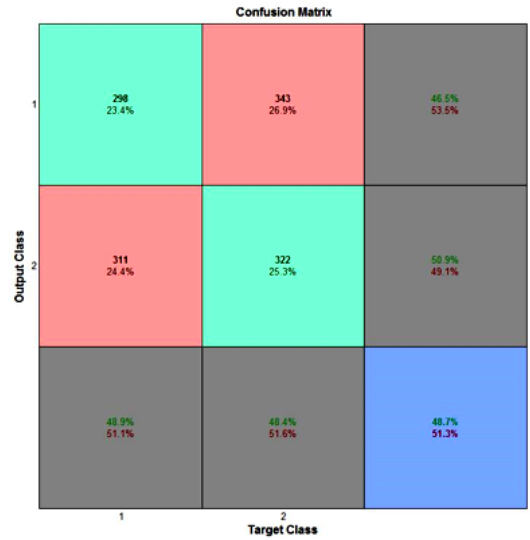


Fig. 4. RNN

Comparison of the confusion matrix of the proposed method compared to other methods shows that the proposed method has performed the classification results better compared to other methods. The ROC diagram for the proposed method with RNN, LSTM methods and the article (Reddy et al., 2020) is shown in Figures 6 to 9 [5].

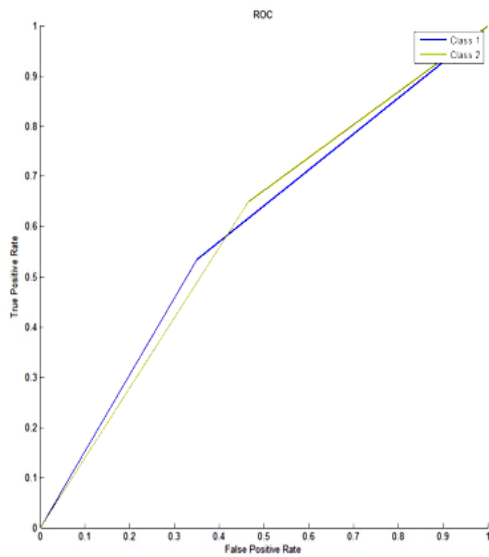


Fig. 7. LSTM

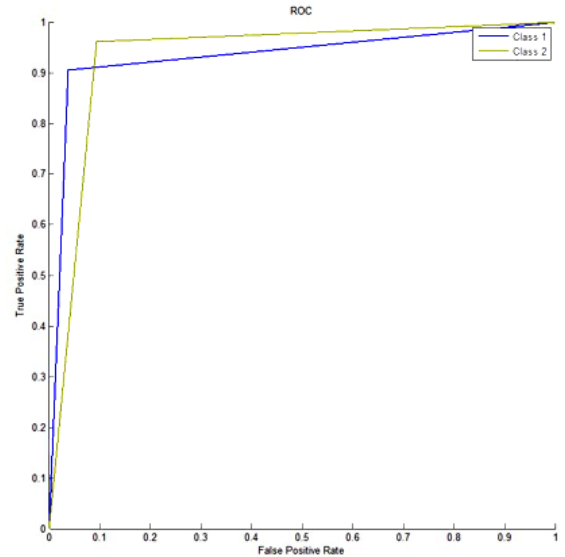


Fig. 6. The proposed method

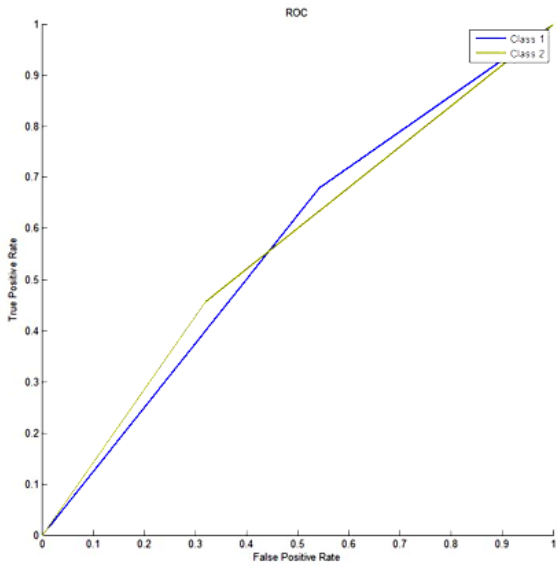


Fig. 9. (Reddy et al., 2020)

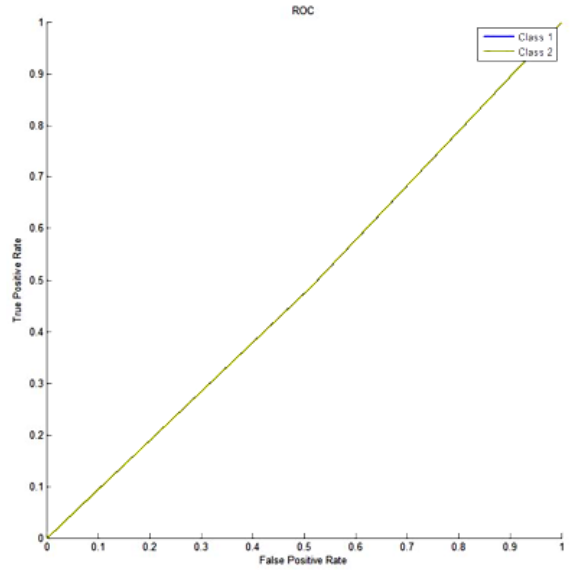


Fig. 8. RNN

Comparison of the ROC diagram of the proposed method compared to other methods shows that the proposed method has performed a better classification compared to other methods, because the graphs of two classes are closer to the ideal state.

The results obtained from the general comparisons for the proposed method with LSTM, RNN methods and the basic article (Reddy et al., 2020) are given in Tables 2 to 5 [5].

Table 2. Results of the Proposed Method

Metric	Mean	1k	2k	3k	4k
MR	0.0745	0.1380	0.0440	0.0540	0.0620
ECM	100622.0	16040157.0	675802.0	80255.0	94253.0
NECM	0.050311	80.20078	33.7901	40.127	47.126
Sensitivity	0.900233	0.74789916	0.9651741	0.942584	0.945274
Specificity	0.949261	0.96564885	0.9498328	0.948454	0.93311
Accuracy (ACC)	0.9255	0.862	0.956	0.946	0.938
Precision (P)	0.928527	0.95187166	0.9282297	0.929245	0.904762
Recall	0.900233	0.74789916	0.9651741	0.942584	0.945274
F-measure (F1)	0.911107	0.83764706	0.9463415	0.935867	0.924574
Consistency	0.817621	0.51889153	0.9417628	0.901347	0.908484
AUC	0.924747	0.85677401	0.9575035	0.945519	0.939192
Balance	0.915309	0.82009053	0.9568167	0.945440	0.938889
Execution Time (s)	2513.33	6204.7034	7503.1432	3920.61	2904.87

Table 3. Results of the LSTM

Metric	Mean	1k	2k	3k	4k
MR	0.432544	0.4342105	0.394737	0.407895	0.493333
ECM	0.595086	0.6213115	0.646853	0.458333	0.653846
NECM	0.297543	0.3106557	0.323427	0.229167	0.326923
Sensitivity	0.562101	0.5573771	0.615385	0.583333	0.945274
Specificity	0.623864	0.6000000	0.545455	0.750000	0.600000
ACC	0.567456	0.5657895	0.605263	0.592105	0.506667
Precision (P)	0.090113	0.8500000	0.888889	0.976744	0.888889
Recall	0.562101	0.5573771	0.615385	0.583333	0.492308
F-measure (F)	0.069116	0.6732673	0.727273	0.730435	0.633663
Consistency	-0.156083	-0.242623	-0.657343	-0.091667	-0.807692
AUC	0.592982	0.5786885	0.580420	0.666667	0.546154
Balance	0.589123	0.5781500	0.578965	0.656408	0.542971
Exec. Time (s)	22198.22	85280.68	63043.92	5727.24	3073.91

Table 4. Results of the RNN

Metric	Mean	1k	2k	3k	4k
MR	0.254167	0.2236842	0.1184211	0.407895	0.266667
ECM	0.314738	0.3284003	0.2121849	0.458333	0.260033
NECM	0.157369	0.1642001	0.1060924	0.229167	0.130017
Sensitivity	0.733298	0.7674419	0.928571	0.583333	0.653846
Specificity	0.818613	0.7878788	0.823529	0.750000	0.913043
ACC	0.745833	0.7763158	0.881579	0.592105	0.733333
Precision (P)	0.903214	0.8250000	0.866667	0.976744	0.944444
Recall	0.733298	0.7674419	0.928571	0.583333	0.653846
F-measure (F)	0.798724	0.7951807	0.896552	0.730435	0.772727
Consistency	-0.043517	0.4644116	0.840336	-0.091667	-0.128760
AUC	0.775956	0.7776603	0.876050	0.666667	0.783445
Balance	0.761711	0.7774256	0.865382	0.656408	0.747627
Exec. Time (s)	7736.969	44637.85	41357.96	3737.26	1974.564

Table 5. Results from Reddy et al. (2020)

Metric	Mean	1k	2k	3k	4k
MR	0.204693	0.025000	0.131579	0.210526	0.226667
ECM	0.287120	0.3591398	0.216846	0.342509	0.229984
NECM	0.143560	0.1795699	0.108423	0.171254	0.114992
Sensitivity	0.786342	0.7333333	0.888889	0.829268	0.693878
Specificity	0.819709	0.7741936	0.838710	0.742857	0.923077
ACC	0.709125	0.8620000	0.690000	0.720000	0.716500
Precision (P)	0.862258	0.8250000	0.888889	0.790698	0.944444
Recall	0.786342	0.7333333	0.888889	0.829268	0.693878
F-measure (F)	0.818721	0.7764706	0.888889	0.809524	0.800000
Consistency	0.455014	0.3462366	0.7275986	0.629268	0.116954
AUC	0.803026	0.7537634	0.863799	0.786063	0.808477
Balance	0.793244	0.7529174	0.861508	0.781744	0.776809
Exec. Time (s)	80331.4	20269786.4	3184428.4	591441.3	208742.4

8. CONCLUSION

The integration of Personal Recommender Models and LSTM-RNN algorithms in predicting consumer behavior in digital marketing has yielded promising results. LSTM-RNN's strength lies in its capacity to process and learn from sequential data, making it highly effective in identifying consumer preferences and behavioral patterns. These

predictive insights enable businesses to make data-driven decisions regarding product development, pricing strategies, and customer retention efforts.

To enhance the effectiveness of these models, future work could focus on improving scalability to handle larger datasets and enable real-time adaptability. Incorporating diverse data sources such as social media activity and user demographics could further refine prediction accuracy and offer a more granular understanding of individual consumers. This would facilitate even more targeted and personalized marketing strategies.

Nevertheless, ethical considerations surrounding the use of personal data must remain a priority. As predictive models become more sophisticated, ensuring data privacy, user consent, and transparency is essential to maintaining consumer trust.

Overall, the application of LSTM-RNN and Personal Recommender Models offers significant potential to transform digital marketing. With continued development, these technologies can usher in a new era of personalized, efficient, and ethically responsible marketing practices.

Declaration

We acknowledge that we used ChatGPT to enhance the academic writing of our manuscript while ensuring the originality and integrity of our work.

Transparency Statement

The data supporting this study are available upon reasonable request to the corresponding author, subject to ethical and confidentiality considerations.

Acknowledgments

We would like to express our gratitude to all individuals who contributed to this project.

Declaration of Interest

The authors declare that they have no competing interests.

Funding

This research received no specific grant from any funding agency, commercial, or not-for-profit sectors.

REFERENCES

- [1] Timotheou, S. (2010). The random neural network: A survey. *The Computer Journal*, 53(3), 251–267. <https://doi.org/10.1093/comjnl/bxp032>
- [2] Smagulova, K., & James, A. P. (2019). A survey on LSTM memristive neural network architectures and applications. *The European Physical Journal Special Topics*, 228(10), 2313–2324. <https://doi.org/10.1140/epjst/e2019-900046-x>
- [3] Ribeiro, T., & Reis, J. L. (2020). Artificial intelligence applied to digital marketing. In *Trends and Innovations in Information Systems and Technologies: Volume 2 8* (pp. 158–169). Springer. https://doi.org/10.1007/978-3-030-45691-7_15
- [4] Kaponis, A., & Maragoudakis, M. (2022, September). Data analysis in digital marketing using machine learning and artificial intelligence techniques, ethical and legal dimensions, state of the art. In *Proceedings of the 12th Hellenic Conference on Artificial Intelligence* (pp. 1–9). <https://doi.org/10.1145/3549737.3549756>
- [5] Reddy, N. S. (2020, July). Particle swarm optimized neural network for predicting customer behaviour in

digital marketing. In 2020 Second International Conference on Inventive Research in Computing Applications (ICIRCA) (pp. 70–77). IEEE. <https://doi.org/10.1109/ICIRCA48905.2020.9182801>

- [6] Sriram, V. P., Shaikh, A. A., Sumana, B. K., Kumar, A., Dhiman, V., & Naved, M. (2022, June). Consumer behaviour on digital marketing platforms—Specifically in terms of consumer loyalty using machine learning. In *Proceedings of Second International Conference in Mechanical and Energy Technology: ICMET 2021, India* (pp. 377–386). Springer. https://doi.org/10.1007/978-981-19-0108-9_40
- [7] Ballestar, M. T., Grau-Carles, P., & Sainz, J. (2019). Predicting customer quality in e-commerce social networks: A machine learning approach. *Review of Managerial Science*, 13, 589–603. <https://doi.org/10.1007/s11846-018-0316-x>
- [8] Senyapar, H. N. D. (2024). Artificial intelligence in marketing communication: A comprehensive exploration of the integration and impact of AI. *Technium Social Sciences Journal*, 55, 64–81. <https://doi.org/10.47577/tssj.v55i1.10651>
- [9] Ho, S. P. S., & Chow, M. Y. C. (2024). The role of artificial intelligence in consumers' brand preference for retail banks in Hong Kong. *Journal of Financial Services Marketing*, 29(4), 292–305. <https://doi.org/10.1057/s41264-022-00207-3>
- [10] Van Nguyen, T., Zhou, L., Chong, A. Y. L., Li, B., & Pu, X. (2020). Predicting customer demand for remanufactured products: A data-mining approach. *European Journal of Operational Research*, 281(3), 543–558. <https://doi.org/10.1016/j.ejor.2019.08.015>