International Journal of
INTELLIGENT SYSTEMS AND APPLICATIONS IN ENGINEERING
ISSN:2147-6799 www.ijisae.org Original Research Paper

A Novel Approach to Predict Consumers Behaviour using Implicit Product Properties in E-Commerce using Deep Learning Techniques

Vidya Rajasekaran[1], Dr. Latha Tamilselvan[2]*

Submitted: 23/08/2023 **Revised**: 08/10/2023 **Accepted**: 21/10/2023

Abstract: The proposed model presents a novel approach for predicting the behavior of consumers using implicit product properties through deep learning techniques. The existing works only focus on the purchase intention of consumers in particular sessions. They do not focus on the implicit properties of the products and the acts performed by the consumers. This model extracts valuable insights into a product through the consumer's behavior in their journey starting from viewing the product, adding the product to their cart, and finally purchasing the product. The key variables considered for this study are based on the perspective of the consumers and the products. The implicit product properties like customers' preference for a product, their perception of the quality of the product, and their action in purchasing the product generate various data for analysis. These inputs can accurately predict the consumer's behavior in purchasing a product. A novel CPCPA approach is proposed to calculate the predictive score of the developed model. Then a comparative analysis of deep learning along with machine learning outcomes for the same dataset is carried out and the resulting metrics prove the developed deep learning model outperforms in terms of performance. A very clear deep analysis and understanding of the consumer's behavior will support firms to build solutions resulting in enhanced business outcomes.

Keywords: *Consumer Behavior, Machine Learning, Deep Learning, Implicit Product Analytics, E-commerce.*

1. Introduction

The E-Commerce platform has revolutionized the way of shopping and where the consumers check on several insights on a product and finalize their purchase. The resultant traces of the consumers in the e-commerce firm are stored as data. Every act they perform is monitored and recorded. This record remains an invaluable asset for the analysts and the data scientist to extract meaningful information from those data and in turn, find ways to gain profit from it. This research work focuses on predicting the behavior of consumers through the implicit product properties of the products and the consumer's act of viewing, adding to the cart, and finally purchasing the product. Consumer behavior is an important factor that needs to be understood by the firm to enhance its marketing strategies and offer tailored services based on the individual consumer's expectations and needs.

Traditional practices of understanding the behavior of consumers rely on the explicit features of the consumers like the reviews of the consumers, their ratings, and so on. The gap to understand the implicit product properties which help to identify the factor behind consumer decision-making is concentrated in this paper.

The implicit properties considered in our research focus on the characteristics of the product based on the consumer's browsing and purchase activities. Working on these implicit properties helps to make more accurate predictions about the behavior of the consumers. Once a clear understanding is made of the behaviour of consumers the firm can concentrate on applying targeted marketing strategies. The process of extracting complex relationships in a dataset can be achieved using deep learning techniques. Deep learning algorithms can be used to capture intricate features and data dependencies to understand consumer behavior.

Our research tries to propose a novel approach to predict the behavior of consumers using implicit product properties through deep learning techniques for enhanced decision-making.

These prediction techniques can be used to accurately predict consumer behavior in a real-time e-commerce environment. Our study and proposed model contribute to the next level of advancement in the field of understanding consumer behavior.

2. Literature Survey

In [1], the researchers aimed to evaluate both the implicit and explicit qualities of the protein products that attracted the consumers. The resulting outcome found differences in consumer preferences which can be used by the manufacturers to focus on obtained insights to meet the consumer's preferences.

In paper [2], the research focuses on collecting the

[1] *B. S. Abdur Rahman Crescent Institute of Science and Technology, Chennai, India*
ORCID ID: 0000-3343-7165-777X
[2] *B. S. Abdur Rahman Crescent Institute of Science and Technology, Chennai, India*
ORCID ID: 0000-3343-7165-777X
* *Corresponding Author Email: latha.tamil@crescent.education*

consumer's implicit and explicit data to understand their perceptions of products for nutrition-based information.

The research work in [3] proposes a short-term context for enhancing the performance of the consumer's search for the products. The study proposes the effectiveness of using implicit feedback of user clicks comprising both long and short-term contexts. The author lacks to explore an understanding of the short-term context for personalized product search.

The research article [4] explores the different factors influencing consumers' decision-making process. The study considers the implicit attributes and their impact on consumers' decision-making.

The article [5] focuses on implicit attitudes and their applications in consumer behavior-related research. The study delves into the theoretical concepts of implicit attitudes and their significance in understanding the logic behind consumer decision-making.

The paper [6] addresses the usage of implicit attitudes to enhance the user's experience to increase the consumer's satisfaction with a food delivery application.

There are several research works discussed regarding the process of predicting consumer behavior [7] [8] [9]. They focus on predicting the consumer's behavior for different applications. The research papers [10] to [17] deal with the process of predicting consumer behavior using deep learning techniques. Different deep learning algorithms and evaluation criteria are applied to find the insights of different applications to observe consumer behavior.

The authors in [18] [19] [20] state the process of predicting the behaviour of consumers using sentiment analysis and analytical approaches. They generate the categorical input into numerical to predict the consumer score and in turn predict the consumer behavior. The several ideas and techniques gained from the above research works are used to develop a novel system to predict the behavior of consumers.

3. Proposed Methodology

The main goal of this approach is to identify the consumer's behavior based on product preference using the data on the consumer's engagement with a product based on the views, adding the product to the cart, and purchasing the product. Then a novel approach is proposed to predict the behavior of the consumers based on the above-obtained results. The developed methodology uses deep learning techniques to recognize complex patterns in the dataset as shown in Fig 1.

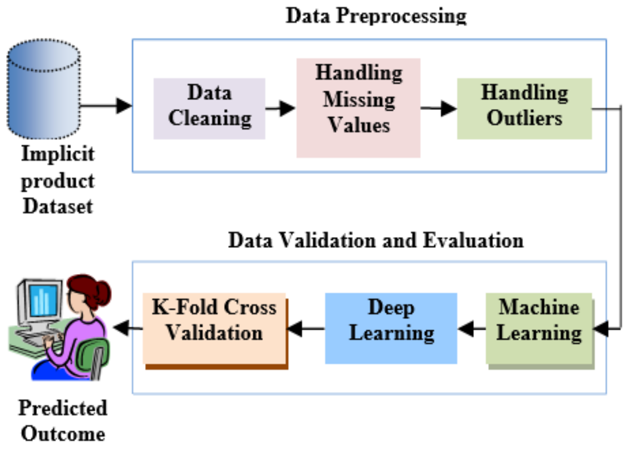

Fig - 1 Process Flow Block Diagram

3.1. Data pre-processing

In this work, a three-stage data pre-processing approach is applied to enhance the accuracy of the model. We start with data cleaning and identifying the missing variables. The dataset used for this research is obtained from the link [11]. The attributes like timestamp, visitor_id, event, item_id, and transaction_id were considered for pre-processing. The missing values are replaced with relevant values as shown in Table 1. Then the duplicates and outliers were found and eliminated.

Finally, the categorical variables are transformed into numerical ones for further analysis and training using machine learning and deep learning techniques as represented in Table 2.

Table 1. Missing Value Identification

Missing values in each variable before pre-processing		Missing values in each variable after pre-processing	
timestamp	0	timestamp	0
visitor_id	0	visitor_id	0
event	0	event	0
item_id	0	item_id	0
transaction_id	1039921	transaction_id	0

Table 2. Categorical Conversion

timestamp	visitor_id	Event	item_id	transaction_id
1.433220e+12	257597	2	355908	0

1.433220 e+12	992329	2	248676	0
1.433220 e+12	111016	2	318965	0
1.433220 e+12	483717	2	253185	0
1.433220 e+12	951259	2	367447	0

Table 3 contains a description of the descriptive variables present in our study. There are totally five variables in the dataset that can be classified into two sets: the first set is based on the perspective of the product and the second based on the consumer's perspective.

Table 3. Descriptive variables and their statistics

	Mean	**Std. Dev**	**Min**	**Max**
timestamp	1.436536 e+12	2.585727 e+09	1.433140 e+12	1.44057 0e+12
visitor_id	7.024382 e+05	4.051996 e+05	1.433140 e+12	1.44057 0e+12
event	1.940714 e+00	4.051996 e+05	1.433140 e+12	1.44057 0e+12
item_id	2.346057 e+05	4.051996 e+05	1.433140 e+12	1.44057 0e+12
transaction_id	7.270382 e+01	4.051996 e+05	1.433140 e+12	1.44057 0e+12

3.2. Data analysis

Data analysis is an important step in the process of data processing. The characteristics of the dataset are explored and insights are extracted to draw trends and patterns of the dataset. The insights are extracted based on two different perspectives: the consumer perspective (P1) and the products perspective (P2). The results are computed using the formulas depicted in Table 4 and Table 5. The results of the analysis are depicted using graphs in the below Figures 2 – 9. These results are used as the foundation for further analysis in our research. Figure - 2 represents the resulting analysis of each product along with their number of views. The product with the count on the number of views is calculated and rated as per the views received. This data insight helps to understand the interest of consumers to view the product and the popularity of the product.

Table 4. Table based on Consumers Perspective (P1)

Description	Formula
Percentage of the consumers individually involved in a particular session	(No. of individual consumers involvement in a particular session / Total number of consumers involvement in a particular session) * 100
Percentage of the views made by the individual consumer	(No. of views of individual consumer in a particular session / Total number of views made by all consumers in a particular session) * 100
Percentage of items added to the cart by the consumers	(No. of items added to cart by individual consumers in a particular session / Total number of items added to the cart for all consumers in a particular session) * 100
Percentage of items purchased by individual consumers	(No. of items purchased by the individual consumer in a particular session / Total number of items purchased by all consumers in a particular session) * 100

Table 5. Table based on Products Perspective (P2)

Description	Formula
Percentage of the individual product involved in a particular session	(No. of individual product involved in a particular session / Total number of products involved in a particular session) * 100
Percentage of the individual products viewed	(No. of views of individual product in a particular session / Total number of products viewed in a particular session) * 100
Percentage of individual products added to cart	(No. of individual products added to cart in a particular session / Total number of items added to the cart in a particular session) * 100
Percentage of individual products sold	(No. of items purchased by the individual consumer in a particular session / Total number of items sold in a particular session) * 100

In Figure - 3 the representation of the count of the products added to the cart by the consumers is explored. The representation depicts the item which is added to the cart and this analysis helps us to understand the items mostly wanted by the consumers and the ones where the focus needs to be made and convert to sales.

Figure - 4 illustrates the final count on the sales of the product. The analyst can gain insights into the maximum sales for each product. This insight can be used to understand the consumer's preference for the products and the products where the firm can focus to enhance sales.

Figure - 5 depicts the count on the product views from a consumer point of view. The consumers making views on the individual products are calculated and this data can be used to understand the trending products and the products in demand where the focus is necessary.

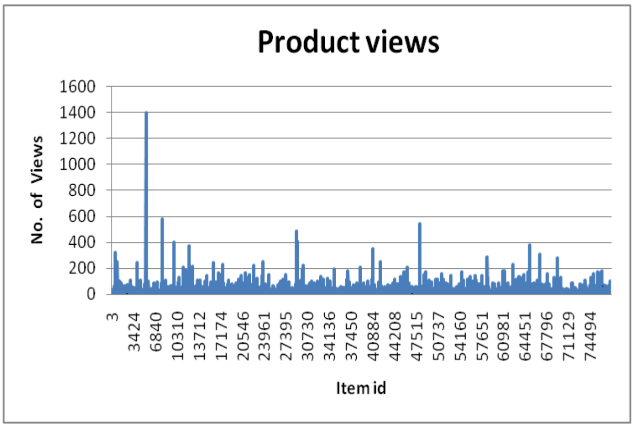

Fig - 2 Product views representation

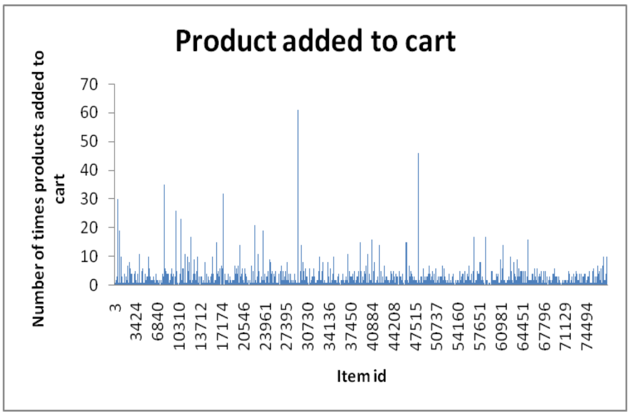

Fig - 3 Product added to cart representation

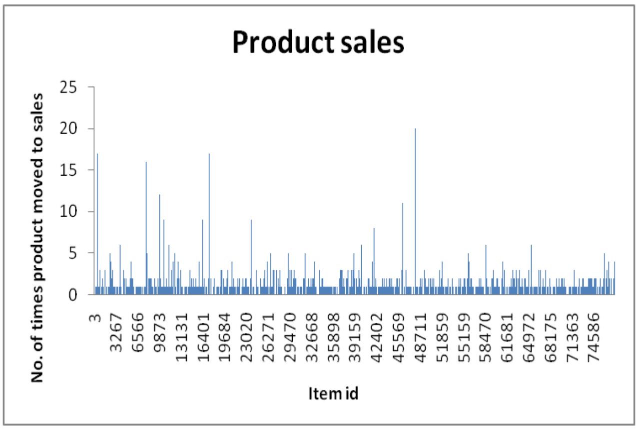

Fig - 4 Product sales representation

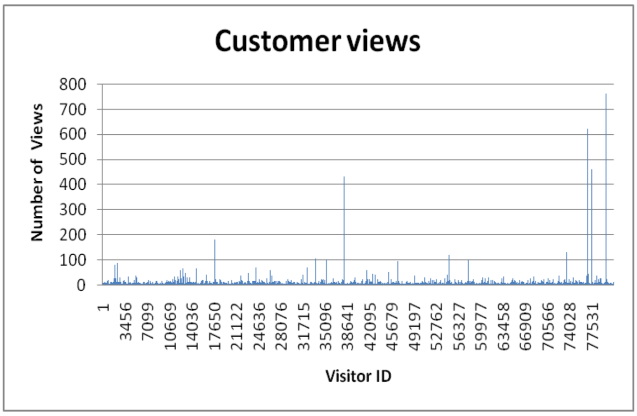

Fig - 5 Consumer views representation

Figure - 6 denotes the count of the products added to the cart from the consumer's point of view. This data helps us to understand the behavior and preference of individual consumers.

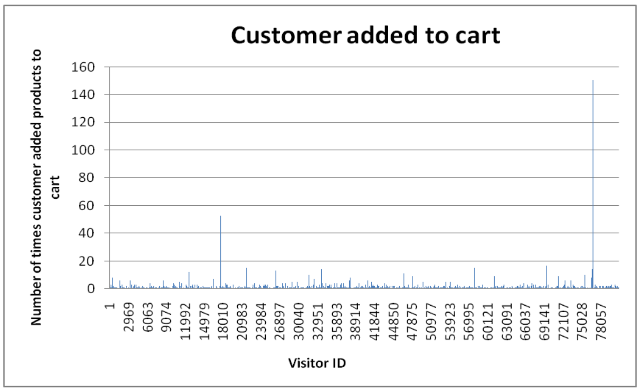

Fig - 6 Consumer add to cart representation

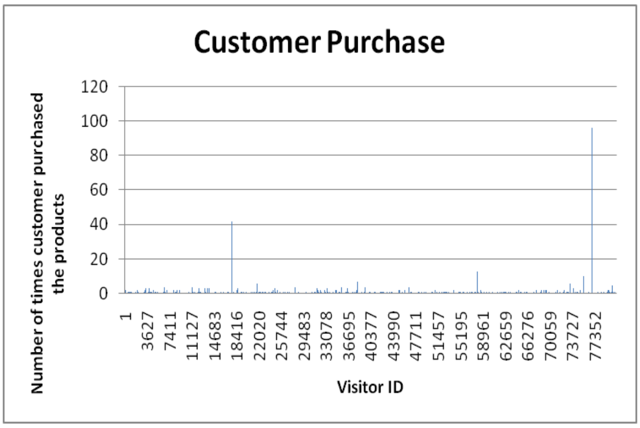

Fig - 7 Consumers purchase representation

Figure - 7 denotes the final purchase made by the consumers. The products with maximum and minimum purchases are helpful to understand the consumers buying and their desire for products.

After preprocessing and analyzing the dataset, we develop a novel CPCPA Algorithm to predict the behavior of consumers using implicit product properties.

3.3. CPCPA Model

A novel model is developed to predict the consumers buying behavior using the predictor variables represented in Table 4 and Table 5. The Consumer and Product Centric Prediction Algorithm (CPCPA), is developed to predict the buying behaviour of consumers.

3.3.1. CPCPA

1. Initialize
2. Declare the Input Variables c1,c2,c3,c4,p1,p2,p3,p4;
3. Compute P1 using $P1= [(w1_{c1} * c1) + (w2_{c2} * c2)+ (w3_{c3} * c3)+ (w4_{c4} * c4)]$
4. Compute P2 using $P2= [(w1_{p1} * p1) + (w2_{p2} * p2)+ (w3_{p3} * p3)+ (w4_{p4} * p4)]$
5. Calculate $P_{CPCPA} = [(0.7 * P1) + (0.3 * P2)]$
6. Store the results
7. Prediction of Results

    a. If P_{CPCPA} > 0.75 and < 1 then

    Print (“Consumers intention to buy the product x is high”)

    b. Else If P_{CPCPA} > 0. 5 and < 0.75 then

    Print (“Consumers intention to buy the product x is moderate”)

    c. Else If P_{CPCPA} > 0. 25 and < 0.5 then

    Print (“Consumers intention to buy the product x is less”)

    d. Else If P_{CPCPA} > 0. 1 and < 0.25 then

    Print (“Consumers intention to buy the product x is very low”)

    e. Else If P_{CPCPA} = = 0 then

    Print (“Consumers intention to buy the product is nil”)

    f. End If

8. Stop

The above algorithm is used to compute the P1 and P2 scores based on the parameters represented in Table 4 and Table 5. Separate weights are multiplied with each parameter based on the importance of the parameter. After calculating P1 and P2 scores the final score is computed. A weight of 70% and 30% is multiplied with P1 and P2 scores to calculate the final value of PCPCPA. Thus based on the resulting values the consumer’s intention and their behaviour can be predicted. We propose a decision theory based on the outcomes. If the score obtained is above 75% the chances to buy a product are high. If it’s above 50% and below 75% the chances are moderate. If it’s less than 50% and above 25% the chances are less. If it’s less than 25% and above 1% there the chances are very low. Finally, if the resulting score is 0, it denotes that the chances are nil.

Figure - 8 and Figure - 9 represent the P1 and P2 scores calculated. P1 score denotes the sales based on individual consumers and P2 denotes the sales based on the individual products.

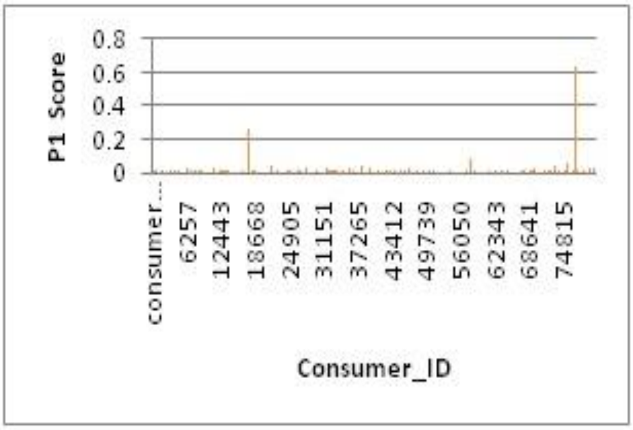

Fig - 8 Graph Indicating P1 Score of Individual Consumers

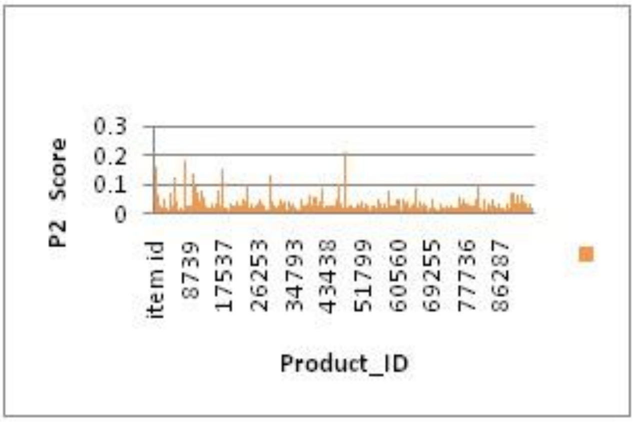

Fig - 9 Graph Indicating P2 Score of Individual Products

The final computed values for P1 and P2 with the weight factors are represented in Figure - 10 and Figure - 11.

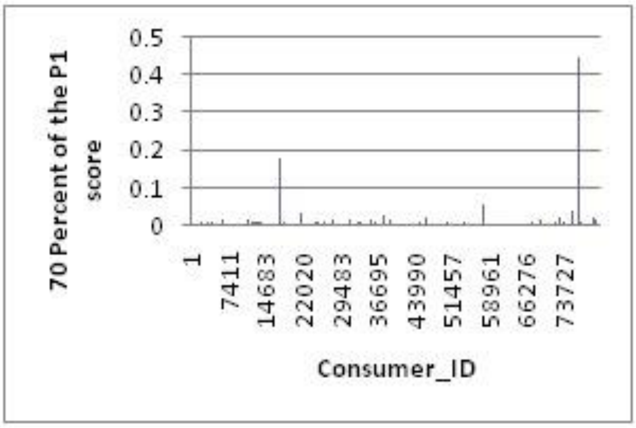

Fig - 10 P1 Score of Individual Consumers with Weight Factor

This final score of P1 and P2 is used to understand consumer's preference for a product and the product's preference in the market.

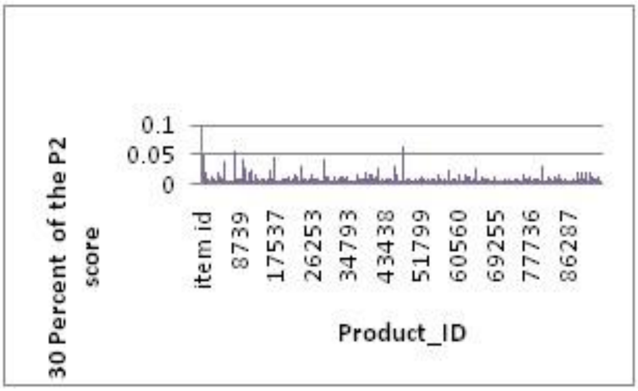

Fig - 11 P2 Score of Individual Consumers with Weight Factor

The computed P_{CPCPA} score as represented in (1) can be used to understand and predict the consumers purchase behavior. These insights can also be further used to predict the seasonal purchase of consumers and product sales.

$$P_{CPCPA} = [(0.7 * P1) + (0.3 * P2)] \quad (1)$$

The values 0.7 and 0.3 represent the weight factor multiplied to the derived outcomes P1 and P2.

3.4. Prediction results based on Consumers Perspective (P1)

The evaluation metrics are calculated using the below formulas. Accuracy of the model is calculated using the formula as shown in (2), Recall as shown in (3) and F1 Score as shown in (4).

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

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

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

Table 6. Prediction Outcomes (P1)

Metrics	Machine Learning				Deep Learning
	DT	RF	SVM	ANN	RNN
Accuracy	0.71	0.73	0.71	0.74	0.79
Recall	0.83	0.82	0.80	0.84	0.86
F1 score	0.87	0.86	0.81	0.88	0.89
ROC-AUC Curve	0.72	0.74	0.74	0.72	0.79

Table 6 and Table 7 represent the prediction results of the model based on consumer and product perspective Prediction results based on Products Perspective (P2).

Table 7. Prediction Outcomes (P2)

Metrics	Machine Learning				Deep Learning
	DT	RF	SVM	ANN	RNN
Accuracy	0.62	0.65	0.62	0.69	0.71
Recall	0.75	0.72	0.71	0.76	0.79
F1 score	0.79	0.73	0.77	0.80	0.81
ROC-AUC Curve	0.63	0.65	0.65	0.63	0.72

Figure 12 and 13 represent the resulting evaluation metrics of P1 and P2 scores.

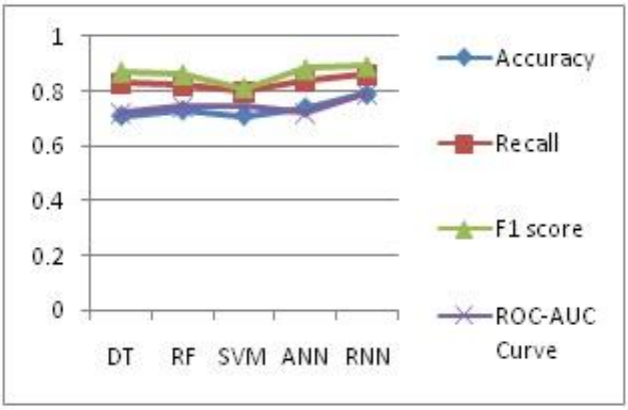

Fig - 12 P1 Evaluation Metrics

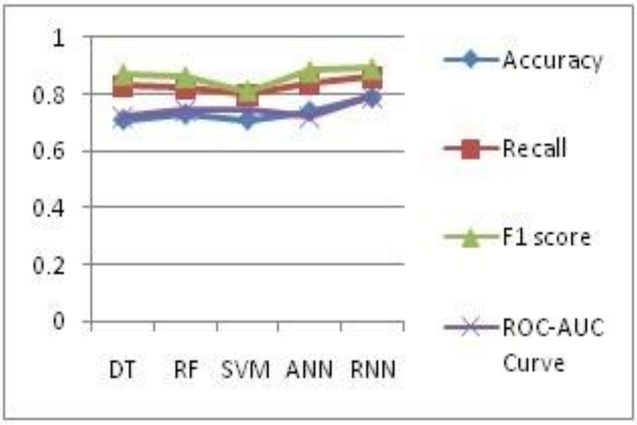

Fig - 13 P2 Evaluation Metrics

4. Results and Discussion

The developed model is evaluated using both machine learning and deep learning algorithms. Even though the prediction results of machine learning algorithms like Decision Trees (DT), Random Forest (RF), Support Vector Machine (SVM), Artificial Neural Networks (ANN) are producing results with good accuracy, we tried to test the model using deep learning algorithm Recurrent Neural Network (RNN). The resulting outcomes of accuracy score, Recall, F1 score, ROC-AUC curve denotes that the deep learning models outstands other machine learning algorithms with enhanced results.

5. Conclusions

The proposed novel approach acts as an enhanced solution for predicting consumer behavior in e-commerce applications. The proposed CPCPA algorithm can be used to understand the consumer's preferences which can be used by business firms to make informed decisions and apply enhanced marketing strategies. Further research can be done by adding the explicit product properties with the implicit ones to make more accurate predictions.

Conflicts of Interest

The authors declare no conflicts of interest.

References:

[1] Harwood, W. S., & Drake, M. (2019). Understanding implicit and explicit consumer desires for protein bars, powders, and beverages. Journal of Sensory Studies, 34(3), e12493. https://doi.org/10.1111/joss.12493

[2] Franco-Arellano, B., Vanderlee, L., Ahmed, M., Oh, A., & L'Abbé, M. R. (2020). Consumers' implicit and explicit recall, understanding and perceptions of products with nutrition-related messages: An online survey. International Journal of Environmental Research and Public Health, 17(21), 8213. https://doi.org/10.3390/ijerph17218213

[3] Teichert, T., Graf, A., Rezaei, S., Wörfel, P., & Duh, H. (2019). Measures of implicit cognition for marketing research. Marketing ZFP, 41(3), 48-76. https://doi.org/10.15358/0344-1369-2019-3-48

[4] Roy, R., & Naidoo, V. (2021). The role of implicit lay belief, SEC attributes and temporal orientation in consumer decision making. Journal of Business Research, 122, 411-422. https://doi.org/10.1016/j.jbusres.2020.09.024

[5] Ackermann, C.-L., & Mathieu, J.-P. (2015). Implicit attitudes and their measurement: Theoretical foundations and use in consumer behavior research. Recherche et Applications En Marketing (English Edition), 30(2), 55–77. https://doi.org/10.1177/2051570715579144

[6] Gardan DA, Epuran G, Paștiu CA, Gardan IP, Jiroveanu DC, Tecau AS, Prihoanca DM. Enhancing Consumer Experience through Development of Implicit Attitudes Using Food Delivery Applications. Journal of Theoretical and Applied Electronic Commerce Research. 2021; 16(7):2858-2882. https://doi.org/10.3390/jtaer16070157

[7] Al Mamun, A., Nawi, N. C., Hayat, N., & Zainol, N. R. (2020). Predicting the purchase intention and behaviour towards green skincare products among Malaysian consumers. Sustainability, 12(24), 10663. https://doi.org/10.3390/su122410663

[8] Kavitha, S., Mohanavalli, S., & Bharathi, B. (2018). Predicting learning behaviour of online course learners' using hybrid deep learning model. 2018 IEEE 6th International Conference on MOOCs, Innovation and Technology in Education (MITE). https://doi.org/10.1109/mite.2018.8747136

[9] Koay, K. Y., Tjiptono, F., & Sandhu, M. S. (2021). Predicting consumers' digital piracy behaviour: Does past experience matter? International Journal of

Emerging Markets, 17(9), 2397-2419. https://doi.org/10.1108/ijoem-09-2020-1067

[10] Koehn, D., Lessmann, S., & Schaal, M. (2020). Predicting online shopping behaviour from clickstream data using deep learning. Expert Systems with Applications, 150, 113342. https://doi.org/ 10.1016/j.eswa.2020.113342

[11] Kushwaha, A. K., Kar, A. K., & Ilavarasan, P. V. (2021). Predicting retweet class using deep learning. Trends in Deep Learning Methodologies, 89-112. https://doi.org/10.1016/b978-0-12-822226-3.00004-0

[12] Naseer, F., Khan, M. N., Rasool, A., & Ayub, N. (2023). A novel approach to compensate delay in communication by predicting teleoperator behaviour using deep learning and reinforcement learning to control telepresence robot. https://doi.org/10.22541/ au.167262775.55104384/v1

[13] Noor, T. H., Almars, A. M., Atlam, E., & Noor, A. (2022). Deep learning model for predicting consumers' interests of IoT recommendation system. International Journal of Advanced Computer Science and Applications, 13(10). https://doi.org/10. 14569/ijacsa.2022.0131022

[14] SHAH DHARIA, R. (2020). Predicting glaucoma interventions with deep learning networks. https://doi.org/10.26226/ morressier.5e61463c250f7ee26aeb4217

[15] Baratzadeh, F., & Hasheminejad, S. M. (2022). Customer behaviour analysis to improve detection of fraudulent transactions using deep learning. https://doi.org/10.21203/rs.3.rs-1164359/v1

[16] Das, P., & Singh, V. (2022). Customer insights analysis using deep learning. 2022 Seventh International Conference on Parallel, Distributed and Grid Computing (PDGC). https://doi.org/10.1109/ pdgc56933.2022.10053272

[17] Mahalekshmi, A., & Chellam, G. H. (2022). Analysis of customer churn prediction using machine learning and deep learning algorithms. International journal of health sciences. https://doi.org/10.53730/ijhs.v6ns1.7861

[18] An analysis of customer perception about Bancassurance an empirical study. (2021). Recent trends in Management and Commerce, 2(2). https://doi.org/10.46632/rmc/2/2/23

[19] Rajasekaran, V., & Priyadarshini, R. (2021). An e-Commerce prototype for predicting the product return phenomenon using optimization and regression techniques. Communications in Computer and Information Science, 230-240. https://doi.org/10. 1007/978-3-030-88244-0_22

[20] Purohit, A. (2021). Sentiment analysis of customer product reviews using deep learning and compare with other machine learning techniques. International Journal for Research in Applied Science and Engineering Technology, 9(VII), 233-239. https://doi.org/ 10.22214/ijraset.2021.36202

[21] M, S. ., Ch, S. C. ., Indira, D. ., & Kumar, G. D. . (2023). Predicting Power Consumption of Individual Household using Machine Learning Algorithms. International Journal on Recent and Innovation Trends in Computing and Communication, 11(3s), 247–252. https://doi.org/10.17762/ijritcc.v11i3s.6192

[22] Kshirsagar, D. P. R. ., Patil, D. N. N. ., & Makarand L., M. . (2022). User Profile Based on Spreading Activation Ontology Recommendation. Research Journal of Computer Systems and Engineering, 3(1), 73–77. Retrieved from https://technicaljournals.org/RJCSE/index.php/journal/article/view/45