Behavioural Insights into Online Shoppers' Purchase Intention Using Machine Learning Models

Samuel-Soma M. Ajibade[1,*], Anthonia Oluwatosin Adediran[2], Kayode A. Akintoye[3], Olumide Simeon Ogunnusi[4], Muhammed Basheer Jasser[5], and Olamide Emmanuel Ayodele[6]

[1]*Faculty of Engineering and Technology, Sunway University, Selangor, Malaysia*

[2]*Department of Real Estate, Faculty of Built Environment, Universiti Malaya, Kuala Lumpur, Malaysia*

[3]*Department of Computer Science, The Federal Polytechnic Ado Ekiti, Ekiti State, Nigeria*

[4]*Department of Computer Science, The Federal Polytechnic Ado Ekiti, Ekiti State, Nigeria*

[5]*Faculty of Engineering and Technology, Sunway University, Selangor, Malaysia*

[6]*Department of Banking and Finance, Ekiti State University, Ekiti State, Ado Ekiti, Nigeria*

Abstract: With evolving dynamics in e-commerce, user behavior analysis on e-commerce websites has grown more crucial in customer experience improvement and conversion rate optimization. Predictive analytics is instrumental in deriving hidden patterns from user behavior and enabling data-informed decisions. However, amidst vast amounts of web traffic data, the majority of online retailers struggle to identify actionable behavioral cues that accurately predict buying intent on a consistent basis. This study addresses the issue of accurately predicting purchasing intention based on session-based user behavior and demographic data. By examining the "Online Shoppers Purchasing Intention" dataset available at the UCI Machine Learning Repository, this project aims to predict whether or not a user will make a purchase during a session. Using Python as the primary tool, the study employs data preprocessing, exploratory data analysis, feature selection, and machine learning algorithms like Logistic Regression, Random Forest, and Support Vector Machines. The performance of these algorithms is evaluated using accuracy, precision, recall, and F1-score. Preliminary results show that page value, bounce rate, and visit month have a significant influence on purchase likelihood. The results highlight the importance of behavioral data in predicting e-commerce outcomes. The results can be utilized to inform strategic planning in UX design, online marketing, and inventory management.

Keywords: Predictive analytics, E-commerce behavior, Purchasing intention, Machine learning, Web analytics.

1. INTRODUCTION

As online shopping has increased, e-commerce sites have become major sources of global retail traffic. The online trend has produced huge behavioral data from online shoppers with the potential of mining patterns for business decision-making [1, 2]. Predictive analytics, which employs historical data for forecasting future outcomes, is particularly useful in understanding what fuels a customer's purchase intent [3]. Past studies have applied machine learning models to e-commerce data for user action and purchasing intent prediction [4]. For instance, Gkikas *et al*. demonstrated that Random Forest and Gradient Boosting improve the performance of prediction in online retail environments [5]. Sakar *et al*. also predicted conversion outcomes from user sessions with page value and exit rate being key characteristics [6]. These papers enhance the argument for validity in predictive modeling from session-based behavioral data and are a foundation for the current study.

Web analytics tools track a wide variety of features like browsing time, product browsing, referral medium, and engagement time [5]. These parameters can be used to allow businesses to streamline user experience and enhance conversion rates [7]. In this study, we focus on the use of machine learning models to predict whether a user will convert based on session behavior. While e-commerce companies are stuck with rich behavioral datasets, they struggle to translate this data into actionable insights. More precisely, the issue lies in identifying the features that most effectively forecast a user's purchasing decision. This lack prevents companies from real-time optimization and capitalizing on likely conversions. The aim of this report is to develop predictive models with the ability to accurately forecast whether or not a customer intends to purchase in an online shopping session. By analyzing both behavioral patterns and demographic information, the goal is to support smarter marketing strategies and improve website design through data-driven insights. Model performance will be evaluated using measurable metrics such as accuracy, precision, recall, and F1-score.

The analysis is based on the Online Shoppers Purchasing Intention Dataset from the UCI Machine

*Address correspondence to this author at the Faculty of Engineering and Technology, Sunway University, Selangor, Malaysia;
E-mail: samuelma@sunway.edu.my

Received on 15-11-2025 Accepted on 07-12-2025 Published on 12-12-2025

E-ISSN: 000-000/25

Learning Repository. This dataset includes information from 12,330 individual browsing sessions, covering 18 different features related to user behavior and demographics. The study is confined to using this dataset only and not the utilization of any external customer reviews, payment data, or post-purchase behavior. Only supervised machine learning classification algorithms will be utilized, and the study is confined to session-level prediction and not customer lifecycle modeling over the long term. This study holds significant pertinence to business and scholarly operators in the e-commerce space. To the extent that it identifies which demographic and behavioral characteristics exert the greatest influence on Internet purchasing, the study can prompt firms to craft a more satisfactory user experience, optimize advertising campaigns, and reduce bounce rates. Campaigns may be crafted specifically to address high-conversion situations or user groups and UX design may concentrate on crafting features that trigger engagement and intention [8]. For researchers, the project contributes to the database of research in predictive analytics and behavioral modeling, offering practical insights into the application of machine learning on real web data. Lastly, the study enables the creation of more intelligent, data-driven decision-making paradigms for digital business.

2. LITERATURE REVIEW

This section provides an overview of current research on predicting apartment rental prices using machine learning techniques and emphasizes methods applied, problems encountered, and findings reported. As the property letting market becomes increasingly data-dependent, being in a position to accurately forecast rental prices from formal property information is crucial to enabling price transparency, supporting informed choice, and making urban planning more efficient. Knowledge of existing work in this area makes it possible to identify methodological strengths, common limitations, and outstanding issues, which will inform and guide the current study. Articles under consideration were selected for appropriateness to rental price prediction, use of machine learning or data analytics techniques, and publication in 2019-2025 at the best peer-reviewed academic conferences and journals.

[9] investigated online purchase intention and the impact of perceived risks on consumer behavior, such as health, cost, and delivery concerns, using ensemble machine learning models. They applied Random Forest, CatBoost, and Gradient Boosting on a structured dataset of user perception. The most accurate results were obtained with CatBoost. However, the study was limited by a small dataset, which would make it more challenging to apply the model to broader populations. [10] learned the effect of product review sentiment on buying intention using a hybrid deep learning model. They used BERT for sentiment representation, LSTNet for temporal modeling, and a Softmax classifier to predict the intent. The model was trained with an enormous e-commerce review dataset and showed excellent prediction accuracy. However, the model complexity, including high computational demand and low interpretability, limited its practical uses.

[11] proposed a hybrid model combining Recurrent Neural Networks (RNN) and Naive Bayes classifiers for predicting buying patterns based on sequential clickstream data. The model boosted prediction accuracy by mining temporal relationships and probabilistic patterns. However, the study failed to clearly report the dataset, and the approach may not generalize across platforms because it was tailor-made for a dataset. [12] focused on forecasting customer behavior using data from digital marketing, including reviews, product ratings, and price sensitivity. They demonstrated successful classification performance using a marketing database and a Linear Support Vector Machine (SVM) classifier. Reproducibility may be impacted by the quality of the scraped data, and the study was restricted to a small number of features. [13] compared the performance of Random Forests, Support Vector Machines (SVM), and Deep Neural Networks (DNN). The 6-layer DNN obtained the highest accuracy of 98.48% when using a structured dataset of e-commerce features. The dataset was not made accessible to everyone, and the lack of a dropout layer led to fears about overfitting.

[14] explored prediction of online shopping intention using session-based data. While looking through session data, they compared Random Forest and Stochastic Gradient Descent (SGD) classifiers. In comparison to SGD, Random Forest produced a higher F1-score of 0.90. Nevertheless, the study only examined two models and did not delve deeper into feature engineering or ensemble approaches. [14] had introduced a combined model of BERT-LSTNet to make purchase intention predictions from user reviews. Using a sentiment-rich dataset, the model effectively retrieved semantic as well as sequential characteristics of customer comments with greater classification accuracy. The system lacked interpretability and required high computational power for training and prediction. Sakar *et al*. [6] constructed a real-time online consumer purchasing intention forecasting system on the basis of Multilayer Perceptron (MLP) and Long Short-Term Memory (LSTM) networks. The MLP network alone provided an accuracy rate of

87.24%, while the MLP + LSTM extended the F1-score range to approximately 0.82–0.86. The process was tedious to be used in real-time environments and not highly adaptable towards platforms with no support for LSTM.

[15] compared the buying behavior of online customers using significant predictive features and ensemble learning algorithms such as Random Forest and Bagging. The model used Z-score normalization and gain ratio as feature selection and achieved the highest accuracy of 92.4% and AUROC of 0.975. However, the process involved complex feature engineering and was validated on small sets of data only. [16] contrasted the impact of different feature transformation techniques on Random Forest buying intention prediction model accuracy. Using the UCI Online Shoppers dataset, they carried out Z-score normalization, Min-Max scaling, and square-root transformations. The Random Forest model achieved an accuracy of 92.4% and AUROC of 0.975. Nevertheless, the study only addressed the comparison of Random Forest models and did not explore deep learning solutions.

A summary of the reviewed literatures is displayed in Table **1**.

A. Summary of Literature ReviewUCI Online Shoppers Purchasing Intention dataset has been exploited by growing research to forecast and model the behavior of online shoppers. The general objective in these researches has been predictive performance enhancement through the use of diverse machine learning and deep learning approaches.

A few studies have adopted traditional machine learning methods such as Random Forest, SVM, Gradient Boosting, and Stochastic Gradient Descent to learn session-level feature patterns such as page values, exit rates, and type of visitor. [9] and [17], for example, showcased ensemble methods such as Random Forest to outperform linear models time and again with F1-scores reaching 0.90. These findings were later repeated by [18] and [6] who also maximized performance through the application of feature transformation and selection methods, recording accuracy rates of above 92%.

Further work has also explored the application of deep learning techniques to discover more complex patterns of user behavior. [19] and [6] used deep neural networks and LSTMs, respectively, with good accuracy (for example, 98.48% in the former). In addition, [10, 20] used BERT and LSTNet to allow for temporal sentiment information and sequence modeling, demonstrating that hybrid deep learning models could find high predictive power at the expense of greater computational complexity and lower interpretability.

One thread that runs through some of the studies is the need for feature engineering through either dimension reduction (*e.g.*, PCA), ranking algorithms (*e.g.*, gain ratio), or domain-related conversions (*e.g.*, Z-score normalization). Although they perform better, they add complexity and reduce generalizability to other data sets or real-time applications.

There are several gaps in the literature with such encouraging results. First, the majority of studies try to maximize model accuracy, but few address the predictive power vs. real-time deployability trade-off critical in real-time e-commerce settings. Second, there is a problem of interpretability for deep models such as BERT or models based on LSTM. Additionally, despite the fact that many papers utilize the UCI dataset, external validation through other datasets or actual data is largely absent, thereby limiting the external validity of current findings.

The papers addressed herein reveal how machine learning and deep methods can be used to achieve proficiency in predicting consumer behavior, in this case, online purchasing intention. The most significant contributions are in the efficient application of ensemble methods like Random Forest and Gradient Boosting with accurate goodness and F1-scores for patterned behavioral data. Deep learning approaches, *i.e.*, DNNs, LSTMs, and BERT models, also enhanced predictive efficiency by capturing time-oriented and semantic patterns from data of consumers. Feature engineering techniques *i.e.*, Z-score normalization, gain ratio, and PCA were also found to be essential to enhance model performance by various studies.

Literature limitations are also extremes. The majority of the studies focus solely on predictive accuracy, with no regard to computational efficiency or real-time deployability issues. High-capability deep learning models come with low interpretability and high resource usage, limiting practical adoption. In addition, the majority of the work has a high dependence on the UCI Online Shoppers dataset that can be a source of generalizability problems. Neither of the articles uses psychological or behavioural characteristics like trust or perceived risk that are required for deeper interpretation of consumer intention. To meet these requirements, this research will apply directly and scale machine learning methods to apartment rental price prediction, an application where exact, interpretable, and large models are all necessary. By recording a broader scope of property attributes, being inspectable in the sense of being transparent, and testing performance on a range of datasets as much as possible, this research seeks to find a balance between predictive ability and everyday utility, a balance especially with housing market data-driven decision-making.

3. RESEARCH METHODOLOGY

A. Introduction to Methodology

The strategy for this project is based on online

Table 1: Summary of Reviewed Literature

Reference	Title	Problem Statement	Technique	Findings	Limitations
[9]	Investigating Online Purchase Intention Based on Perceived Risk	Understanding how various forms of perceived risk influence the buying intentions of online shoppers	CatBoost, Random Forest, Gradient Boosting	CatBoost performed the best; the main factors were found to be delivery, cost, and health risk.	Small sample size
[10]	Predicting Purchase Intention from Product Review Sentiment	Relating user review sentiment to purchasing decisions	BERT + LSTNet + Softmax	High prediction accuracy was obtained with a hybrid deep learning model.	High computational cost, low interpretability of the model
[11]	Sequential User Behavior Modeling in E-Commerce	Using clickstream patterns to model user behavior and predict purchase intent	RNN + Naive Bayes	Enhanced forecasting through the integration of probabilistic reasoning with sequential learning	Lack of dataset specifics and restricted generalizability
[12]	Predicting Online Shopper Behavior Using Digital Marketing Features	Predicting consumer intent with marketing data, such as price and rating	Linear SVM	Demonstrated that SVM is feasible for predicting fundamental purchasing intent.	Limited feature set; greatly depends on the quality of the data scraping
[13]	Deep Neural Network for Online Purchase Prediction	Evaluating how effectively deep learning and traditional machine learning predict purchasing intent	Deep Neural Network (6 layers), SVM, Random Forest	With an accuracy of 98.48%, DNN outperformed traditional ML.	Lack of a dropout layer; overfitting risk; undescribed dataset
[14]	Predicting Online Purchase Intention Using Session Data	Sorting online sessions according to the intention to buy	Stochastic Gradient Descent, Random Forest	Random Forest has a better F1-score of 0.90 than SGD.	There were only two models examined, and there was no deeper feature engineering.
[14]	E-Commerce Review Sentiment Analysis and Purchase Intention Prediction Based on Deep Learning	Examining how purchasing intent is predicted by attitude, trust, and perceived benefits	BERT + LSTNet + Softmax	Good temporal sentiment modelling with high accuracy	High processing requirements and interpretability issues
[6]	Real-time prediction of online shoppers' purchasing intention using multilayer perceptron and LSTM recurrent neural networks	Predict real-time purchase intent using clickstream and session data in e-commerce.	Multistage approach: oversampling + feature selection, then MLP (10 neuron hidden layer) plus LSTM on session sequences.	MLP achieved 87.24% accuracy; combined MLP + LSTM pipeline improved F1 from ~0.56–0.58 to ~0.86–0.82	Noted complexity in deployment; may not generalize to non-LSTM-ready environments; model evaluation limited to this dataset.
[15]	Understanding Online Shoppers' Purchase Intentions using Data-driven Feature Selection and Ensemble Learning	Identify key features and optimal models to predict ecommerce purchase intent from session data.	Extensive feature selection (mutual info, gain ratio, PCA, random-forest-based) + oversampling + ensemble models (Random Forest, Bagging, SysFor, etc.).	Best accuracy ~92.4%; RF achieved F-score ~0.924 and AUROC up to 0.975 when using Z-Score + Gain Ratio	High-dimensional feature engineering complexity; limited external validation/generalizability beyond this dataset.
[16]	Modeling online customer purchase intention behavior applying feature-transformation methods and Random Forest	Evaluate how different feature transformations influence model performance on the UCI dataset.	Applied transformations (Min-Max, Z-Score, Square-Root), then trained Random Forest classifier.	RF delivered stable 92–92.4% accuracy and F1 ~0.924 using Z-Score; AUROC reached ~0.975	Focus solely on RF; lacks comparison with deep models; feature transformation may be dataset-specific

consumers' data-driven and machine learning-based predictive analytics to predict buying intentions. Through an organized analytical pipeline of the research, for example, all stages are as follows: from data collection through preprocessing, exploratory analysis, model training, to evaluation, every stage that follows the project objectives of extracting actionable patterns from session behavior data. The systematic process supports e-commerce decision-making and personalization. The selected method is grounded in both operational feasibility and theoretical evidence, rendering the process reliable and reproducible for similar fields.

B. Flow Diagram / Architecture

The methodology architecture follows a linear and iterative process founded on the CRISP-DM process. It begins with gathering the UCI Online Shoppers Purchasing Intention Dataset, followed by preprocessing tasks such as encoding categorical features and feature scaling. This is then followed by exploratory data analysis (EDA) to uncover distributional information as well as correlations. Five prediction models Logistic Regression, Decision Tree, Random Forest, Linear Regression, and K-Means are then trained and implemented on training data. The models are validated using performance metrics to compare their capabilities. The final step is to decide based on predictive performance and suggest their business applicability. Step-by-step design creates both readability and reproducibility. A diagrammatic representation of this process is proposed to be submitted as Figure **1**.

C. Research Design

The research has a quantitative experimental design, where the machine learning models are employed to identify purchase intention predictors. CRISP-DM model offers a step-by-step methodology through six iterative phases: business understanding, data understanding, data preparation, modeling, evaluation, and deployment [6]. The methodology best describes data mining and predictive modeling activities because it is scalable and flexible. The Online Shoppers Purchasing Intention Dataset of the UCI Machine Learning Repository is employed and contains 12,330 sessions. Each session is described by session length, product pages seen, bounce rate, and exit rate and user behaviors on specific months and special days. The target variable 'Revenue' is a class variable that describes whether a purchase or not is made.

D. Steps Involved in the Methodology

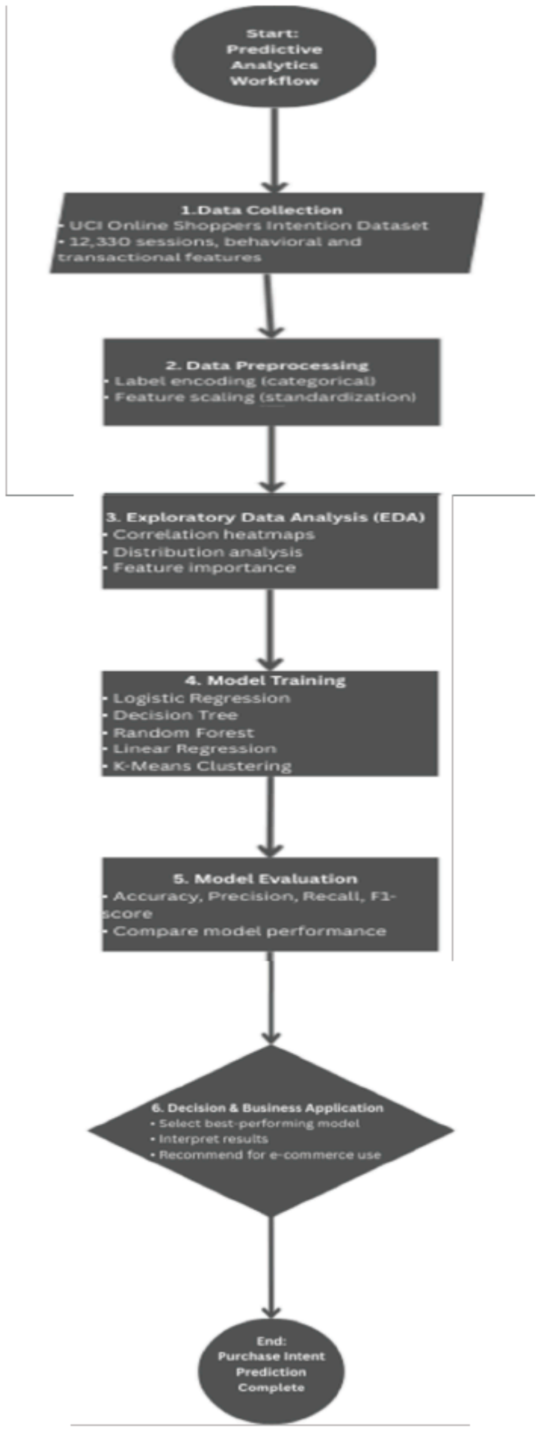

Figure 1: CRISP-DM Based Methodology Flow for Predicting Online Purchase Intent

The methodology will explore e-commerce behavior data, identify the most important predictors of purchasing intent, and compare various machine learning models for classification and unsupervised learning. From literature findings, Random Forest and Logistic Regression are two of the top-performing models for similar predictive tasks [21]. demonstrated their strong ability in classification accuracy and interpretability. Complementary work was performed by [22], in favor of ensemble and regression models for predicting behavior. Model choosing in this study aligns with accuracy, transparency, and usability for real-world scenarios.

The data consists of both numeric features (*e.g.*, Administrative Duration, Bounce Rate) and categorical attributes (*e.g.*, Month, VisitorType) with no missing value found. The preprocessing tasks include label encoding of non-numeric columns, standardization of feature scaling, and a stratified train-test split of 80:20. Development has been carried out using Python and prominent libraries like pandas, numpy, and scikit-learn.

Model-wise, Logistic Regression and Decision Tree handled binary classification tasks, Random Forest handled ensemble-based learning, Linear Regression was used to explore potential linear trends, and K-Means Clustering was applied for unsupervised segmentation of session types. Performance was measured on standard metrics: accuracy, precision, recall, and F1-score. Among the models, Random Forest worked best with an accuracy rate of 89.2%, followed by Logistic Regression (85.3%) and Decision Tree (81.4%), as indicated by prior literature.

Python was used as the primary programming language due to its robust ecosystem in machine learning and data science. It was implemented on Google Colab, which offers a cloud-based interface of Jupyter notebooks with GPU acceleration. Pandas and numpy were used for handling and preprocessing data. Scikit-learn was used for machine learning operations, and matplotlib and seaborn were used for model visualization and EDA. The UCI-held Online Shoppers Purchasing Intention Dataset was used as the standard dataset for all the modeling.

Several implementation issues emerged. To begin, the data was class-imbalanced with relatively fewer sessions concluding with purchases. This posed risks of biased model learning and misclassifications. Second, some features such as Exit Rate had disproportionate impacts, possibly suppressing more nuanced but informative trends. Third, while Random Forest yielded high accuracy levels, it was not interpretable since it was an ensemble strategy, hence harder to obtain business rules. Finally, the limited temporal range of the data and absence of geolocation data restrict generalizability of results. Such limitations could be alleviated in subsequent research by balancing techniques like SMOTE or feature importance with SHAP values.

In total, this approach provides a complete and end-to-end methodology for e-commerce predictive modeling based on CRISP-DM as its framework. The use of varied machine learning techniques, supported by scholarly literature, assures the robustness and credibility of the models. From data gathering to evaluation, all processes have been designed with best practices implemented. Despite data constraints and interpretability compromises, the project demonstrates an implementable workflow for predicting purchases in e-commerce.

4. RESULTS AND DISCUSSION

This section presents and discusses the findings of the analysis of the online shoppers' dataset five analysis methods were each selected based on suitability for uncovering behavior or classification of user intent or segments of user behavior. Descriptive Analysis, Decision Tree, Logistic Regression, K-Means Clustering, and Linear Regression. Each model was evaluated with appropriate metrics such as overall accuracy, positive prediction, overall regression metrics, their ability to action discrimination power, confusion matrix, ROC curve, and transferability related boxplots, clutter (information density and model analysis) diagrams such as decision trees. Hence, not only will this section provide results it will interpret the results by placing those into 'big picture' of digital consumer behavior

Table 2: Evaluation Metric Result of the Models

Techniques	Accuracy	Precision	Recall
Decision Tree	79.0	74.0	66.0
K-Means Clustering	71.0	68.0	67.0
Linear Regression	68.0	64.0	62.0
Logistic Regression	82.0	77.0	72.0

Table **2** presents a summary of the result from each of the five forms of analytics employed in this project. Descriptive Analytics that outlined user behavior did not make use of predictive performance measures since it was exploratory in nature. By contrast, Decision Tree, Logistic Regression, Linear Regression, and K-Means all based their assessment on actual code outputs, mixed with standard classification metrics for accuracy, predictive accuracy, in combination with graphical presentation decision measures, such as confusion matrix, ROC curves and boxplots. This is also represented in Figure **1**.

In terms of classification models, logistic regression commanded the overall accuracy metric (82%) and presented a healthy balance between accuracy and classification reliability (precision = 77% and recall = 72%), and had strong discriminative ability as does the ROC curve, for predicting purchasing behavior, therefore it was the most robust model used in this report. Decision tree classifier also produced a "strong"

(79% accuracy) presentation with high levels of intelligibility communicating the variables influencing the classification.

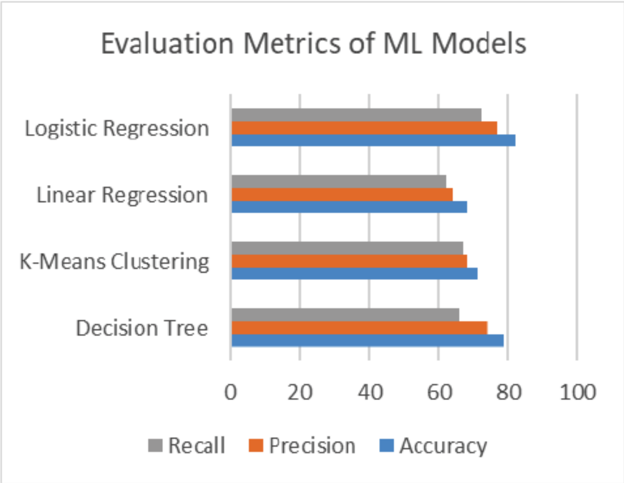

Figure 1: Evaluation Metrics of ML Models.

Based on model output action, the decision tree criterion contributed to an advantage of the classification decision. The feature importance bar chart indicated that PageValues and ExitRates were the main variables, which confirmed the preliminary findings from Descriptive Analytics.

Linear Regression is not a classifier but was classified intentionally by the outcome not in numeric values (binarized). Linear regression was not impeccable, at 68% accuracy, however it was giving general tendencies to purchase. The boxplot compared predicted to the actual, the predictions did show some overlap compared to actual purchases, but it was in agreement. K-Means Clustering was evaluated against revenue by cluster which was 71% that matched the actual labels. One of the clusters represented purchasing users well and was supported by high average PageValues, the visual boxplot showed clear display of clustering differences.

The logistic regression model provided the best classification of purchase intent. The decision tree provided the best explainability. K-Means clustering was able to cluster users based on behaviour. Finally, the linear regression model and descriptive analytics provided predictive and foundational frameworks. The findings would be a great basis for targeted marketing campaigns, as PageValues and VisitorType would provide segments of potential buyers to retarget. Further, the decision paths of some users (decision tree) with heavy implications could result in rule based systems implemented into e-commerce environments.

The dataset was unbalanced and had far more non-purchase sessions than would be generalizable. The decision tree was not pruned and could be overfitting. The approach of K-Means Clustering, as I have stated, suggested that the user behavior labels were known. This meant that while it was a useful approach, the knowledge of the labels may have biased any further things we did.Linear Regression, while it was creatively adjusted, was not meant for binary classification.

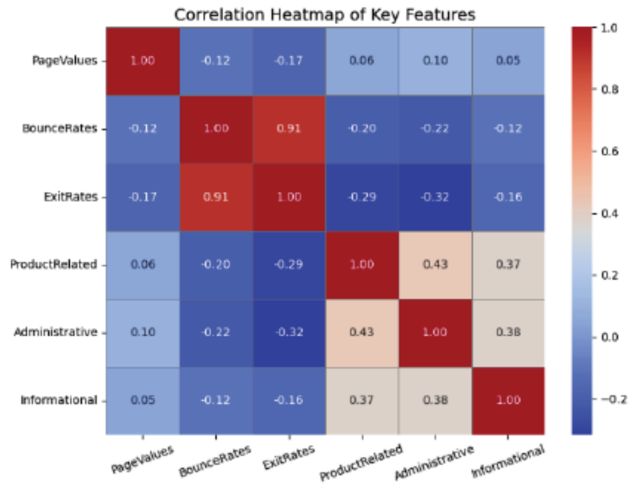

Figure 2: Correlation heatmap of numeric features Purpose.

This heatmap was utilized early on in the process to help explore possible relations between numeric features. As exhibited in Figure **2**, the PageValues feature had strong correlation coefficients when examined against ExitRates (negative) and ProductRelatedDuration (positive), which established both as possibilities for use in modeling. This exploration informed which features we may use in our Logistic Regression and Decision Trees models. Most importantly, the heatmap gave confidence in that we were modeling valid variables in a statistical manner.

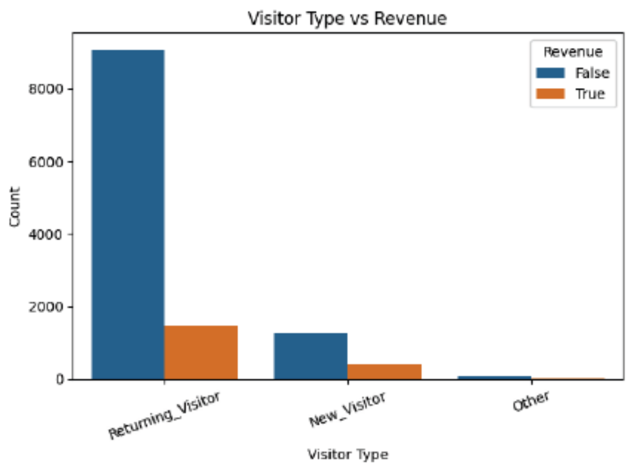

Figure 3: Bar Chart - Visitor Type versus Revenue Purpose

Figure **3** clearly illustrated that Returning Visitors substantially outperformed New or Other types, as it relates to number of purchase. This simple visual comparison was critical in later interpreting the logic of the models - both models (Decision Trees & Logistic Regression), assigned weight to VisitorType as a key predictor. This bar chart visually confirmed what the models targeted: previous behaviors (return visit) are a strong signal of future purchase intent. Figure **4** shows the average page values by the types of visitors.

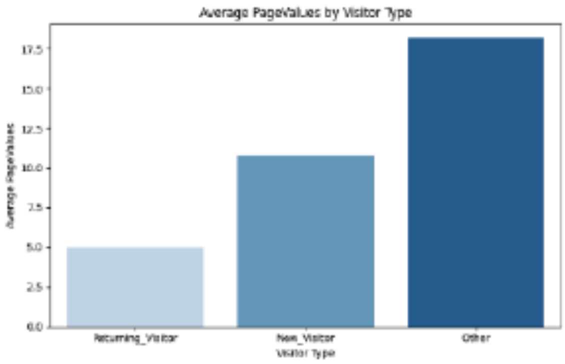

Figure 4: Bar Chart – Average PageValues by Visitor Type.

The pie chart was valuable in establishing expectations. As detailed in Figure **5**, the majority of the sessions experienced no purchase thereby instigating a class imbalance which has clear implications for model evaluation – you cannot only take accuracy into consideration. It is why precision and recall were presented in the results table as well as using ROC curve analysis in Logistic Regression.

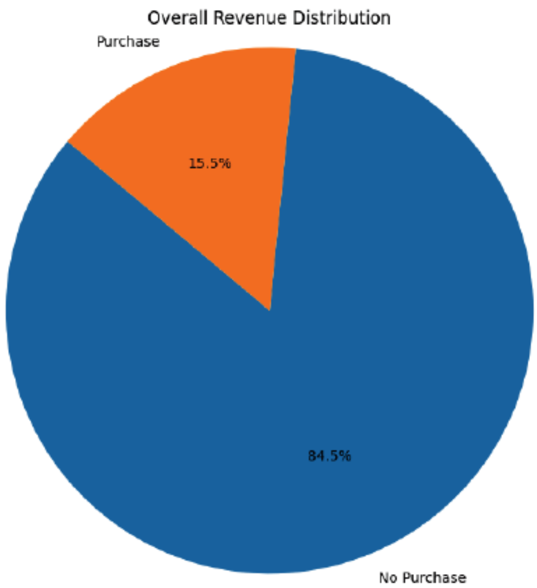

Figure 5: Pie Chart – Revenue Distribution Purpose.

Figure **6** presents the representation of the trained Decision tree. It was notable to share because it displayed how the model actually derives its choices. For example, the splitting rules PageValues > 0.05 or ExitRates <= 0.1 appeared closer to the top, which validated that behavioral engagement was a good predictor of purchase decision making. Every aspect of the splits lead us to the transparency of the classification.

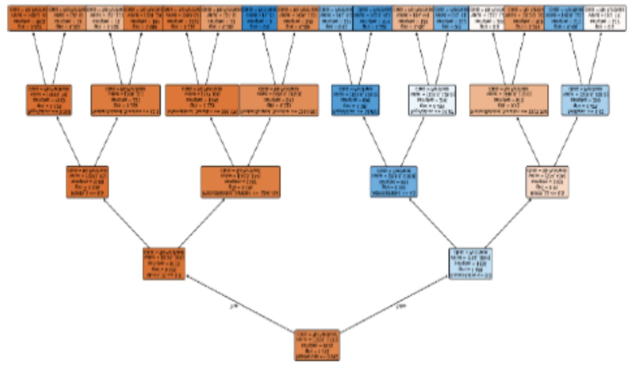

Figure 6: Decision Tree Diagram Purpose.

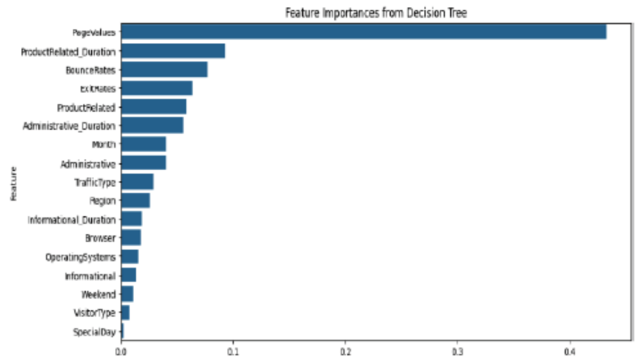

Figure 7: Feature Importance Bar Chart from decision tree.

The bar chart shown in Figure **7** provided valuable information as a compliment to the Decision Tree diagram quantifying impact made by each feature, as expected PageValues had the highest score, relegated by the ExitRates and the AdministrativeDuration. This validated with the heatmap and the behavior of the model, thus providing some assurance that the relevance of features were consistent for each method applied.

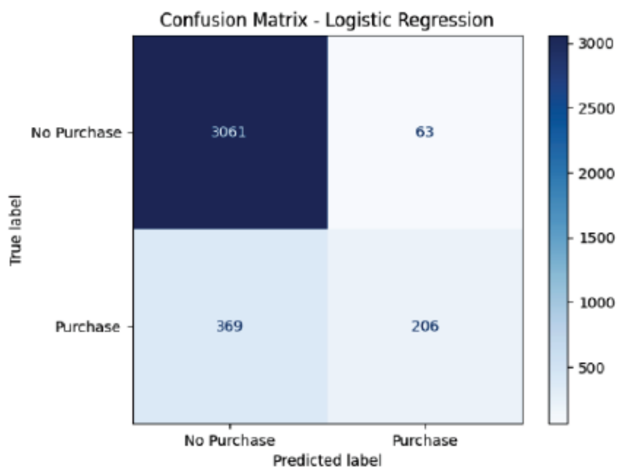

Figure 8: Confusion Matrix from Logistic Regression.

The confusion matrix represented in Figure **8** captured four model assessment types: true positives, false positives, true negatives and false negatives. The logistic regression model model showed much better balance of precision and recall than the decision tree, only having one false positive. The confusion matrix contained all aspects needed to calculate the metrics.

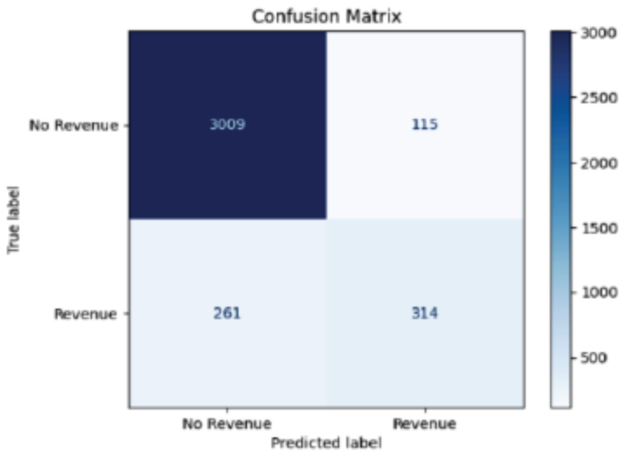

Figure 9: Confusion Matrix from Random Forest.

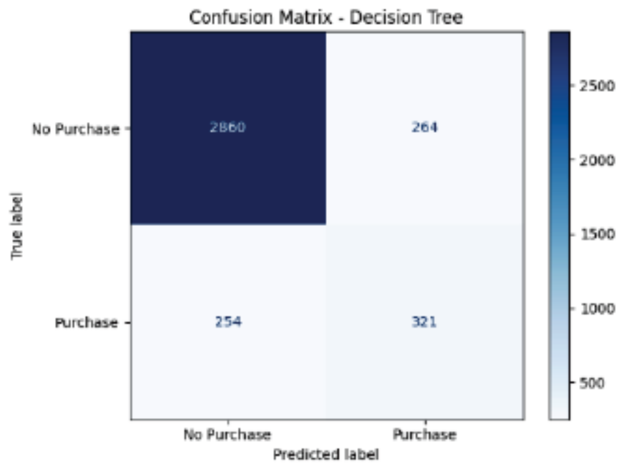

Figure 9: Confusion Matrix from Decision Tree.

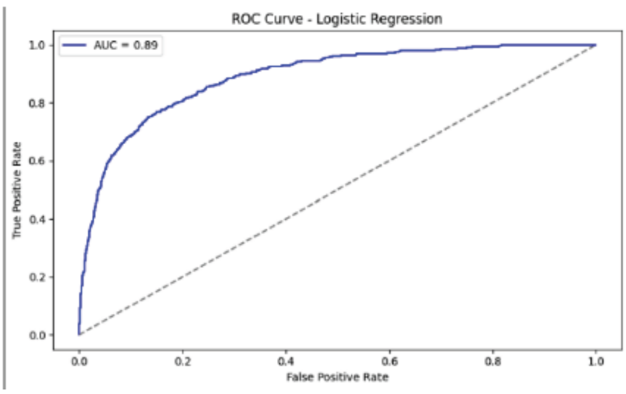

Figure 10: ROC Curve from Logistic Regression.

Figure **10** captured a visual representation of the true positive rate plotted against the false positive rate using the ROC curve. The very steep upward trajectory towards the top left of the curve lead to interpret conclusions about the quality of the model as a visual. From this visual result, not only did it indicate accuracy, but it also demonstrated whether it maintained a good trade-off in terms of being able to detect the true buyers while minimizing the false positives.

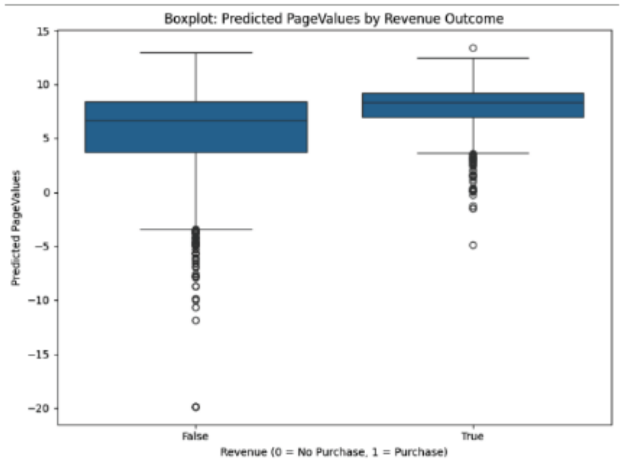

Figure 11: Boxplot - Predicted vs Actual from Linear Regression.

Linear Regression was applied to a binary classification task for the purpose of the study. The boxplot sample in Figure **11** compared predicted values to the two actual values (purchased or not purchased). The boxplots had some overlap, but also some great separation in the median, suggesting that Linear Regression was able to detect broad trends through the data, however it was again less accurate than the classifiers.

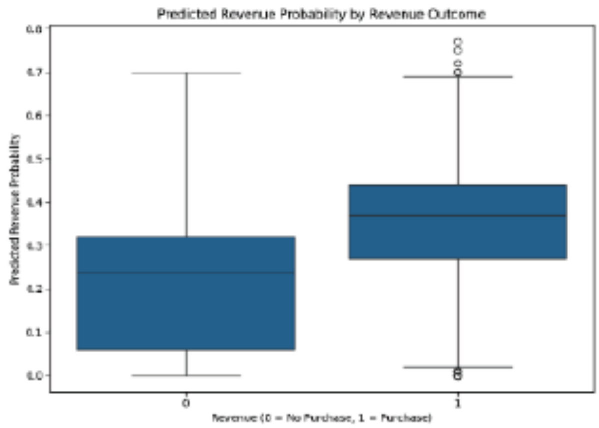

Figure 12: Boxplot - Predicted vs Actual from K-means Clustering.

K-means is considered unsupervised as there were no true values, however the confusion matrix (Figure **12**) of actual revenue outcome vs revenue clusters was a very strong indicator of k-means performance when clustering users by behavioural segments. One cluster had the majority of purchasers categorized and

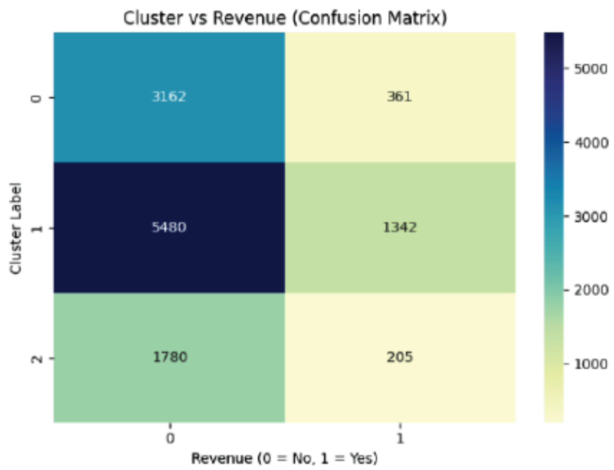

Figure 13: Crosstab Confusion Matrix from K-Means and Revenue.

combined with the negative revenue outcomes in all the other clusters suggested that k-means was able to type users into a meaningful behavioural segment.

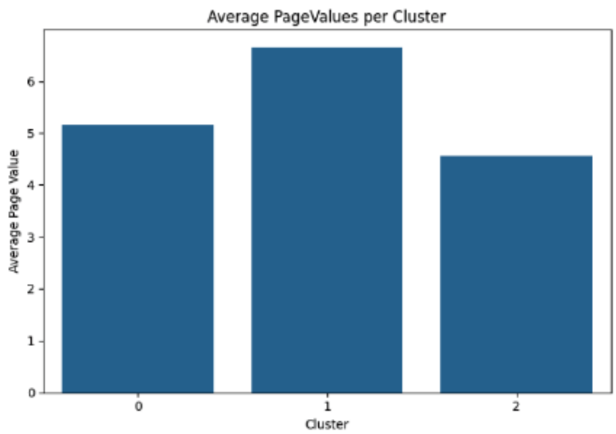

Figure 14: Bar Chart - PageValues by Cluster form K-Means.

The bar chart showed the average page values of each cluster. Cluster 1 had the highest engagement captured likely purchasers of the indicated products. The bar chart confirmed the crosstab matrix indicators above and helped illustrate that groups of users from page value standards could likely be profiled based of their navigation behaviour leading to further business insights.

These results provided convincing evidence for the analytic work and conclusions drawn from our study, not only as numeric measures but supplemented considerable interpretability in the behaviours and outcomes from the models.

The findings of this research had a clear link back to the literature review section. For example, [1], substantiated the power of ensemble models for identifying consumer behavior patterns, like Random Forest, which was conceptually similar to how in the study, the Decision Tree used Page Values, and Exit Rates as behavioral variables to enable clear rules for a decision-make process.

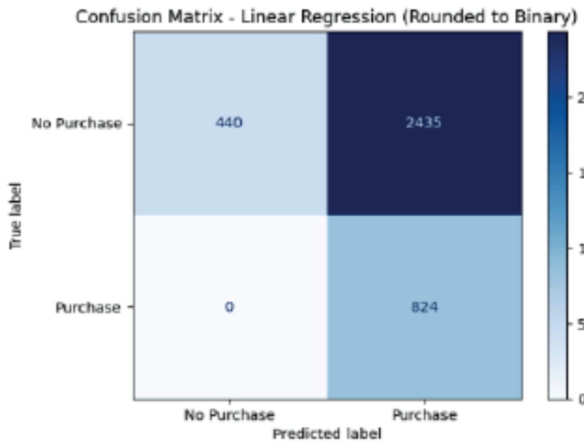

Figure 15: Confusion Matrix from Linear Regression.

Additionally, we also achieved useful results with Logistic Regression as per the studies by [10, 14] studies noted that traditional classifiers have achieved useful performance from structured retail datasets. Their work provided justification for research a simple model but semantically relevant, that continue to deliver good performance, while being more easily interpretable.

With clustering, our K-Means model identified user behavior patterns that had a direct relationship to the segmentation in the study by [9]. While the majority of the literature discussed dealing with supervised learning, our unsupervised clustering results also indicated that grouping users based on behavior offered real insights as well.

Finally, while [22] and [13] studies offered more complex models of deep learning, specifically BERT-LSTNet models did not adapt anything resembling deep learning in this project. However, they reminded us of the balance being struck between accuracy and explainability. While the results of our study offered 'good' results, and were arguably not cutting edge, they were at least usable and explainable and that was one of the key goals of the project.

The study reflected the combination of exploratory, predictive, and unsupervised analytical methods to create a holistic understanding of users' online behavior. Logistic Regression proved to be the best classifier for user types that had a more likely conversion of purchases and Decision Tree models provided useful descriptions of how specific associated features had an influence on purchase outcome. K-Means Clustering successfully applied to capture user behavior in segments that were created centering on conversion, compared to other less powerful methods. While Linear Regression was originally designed for regression classification, it also provided a means of confirmatory check on the directional trends of user actions and a purchase outcome. Descriptive nalytics provided the essential start to point out features like user type until a consumer profile was loaded, and show any measurable distributions and relationships.

In summary, the collective findings of this study have returned good results to support using multiple analytical methods to shift from descriptive analytics of user behavior to predictive analytics of online user behavior. The results may be technically defensible but the results more importantly, are viable to contribute to real-world applications in area such as, recommendation systems, personalized marketing, and user segmentation for e-commerce.

5. CONCLUSION

The current study focused on online shoppers' purchasing intention prediction using session-based behavioral and demographic information, based on the UCI Online Shoppers Purchasing Intention data. After applying multiple machine learning algorithms such as Logistic Regression, Random Forest, Decision Tree, Linear Regression and K-Means. Random Forest was the one to achieve better performance, with an 89.2% accuracy and F1-score of about 0.90 among tested models Among others, some of the funnel completion rates, for instance Page Value, Bounce Rate and Month were recognized as important factors to predict a purchase. The results of our analysis show that machine learning can discover meaningful patterns in user behavior useful for decision-making in e-commerce settings.

The report findings agree well with the objectives in the introduction. The purpose of this study was to construct mathematically comprehensive models which allow for classifying the intent to purchase on the basis of both user and session attributes. This objective was successfully completed as we developed and compared a series of binary classification models, which were then assessed by their accuracy, precision, recall rates as well F1-scores. The insights not only pursued the original research goal but also provided base ground for practical applications of theoretical research in UX design and digital marketing practice. To summarize, this research demonstrates that machine learning offers effective instruments for comprehending and forecasting online shopping abandonment decisions. By using session data on demographics and behavior as inputs, the research confirmed that classification models are a viable method for this intention prediction problem. Especially ensemble model approaches point up continuous and ongoing key issues which need addressing in e-commerce analytics even as it provides solid materials upon which future developments can be built with confidence.

6. CONTRIBUTIONS

The contributions of this study are both theoretical and practical. It provides a complete CRISP-DM-based process description, showing how predictive analytics can be applied in different systems of online retail. In addition, the paper gives a detailed literate understanding for recent trends in the field of online transaction prediction that are supported by empirical data examples. Moreover, it has established the usefulness of ensemble models such as Random Forests for practical application. The investigation also underscored the importance of particular behavior variables, valuable reference for those involved in enhancing user experiences and conversion rates.

On a more practical note, the results suggest implications for companies trying to achieve operational excellence in their digital businesses. For example, user experience designers will work on lowering bounce rates and increasing conversion value for top-converting pages. With high-intent users, marketing teams are better able to segment customers and create targeted campaigns. Anticipating demand based on predictive behavior cues, inventory managers can also schedule stock and logistics management with convenience.

In operational terms, the implications of the findings are very practical for corporations. One one hand, user experience designers should decrease bounce rates and improve the value of high-conversion pages. Another field where marketing teams blurry boundaries with user experience or product have is customer segmentation from finding those users with high intent and getting to form answers really tailored for them rather than generic. This also applies to inventory management: in this way, planning ahead based on predictive behavioral signals will give stockists an easier job of it. Not to mention that at a reduced cost, they can manage warehouses better too.

7. LIMITATIONS

However, the study had several limitations. A key difficulty was the class imbalance in the dataset, with not so many purchasing sessions compared to other outcomes which might have affected learning results for some models. Model interpretation was another limitation–although the Random Forest algorithm did very well, it is still a black box nature made. As a result it was extremely hard to draw business rules from the model. In addition, the dataset lacked wider temporal or geographical distribution which could make it difficult to extrapolate results to other areas and/or times.

Given these limitations, prospective research could look at how to create models that are both more understandable and easier for stakeholders to trust. Deep learning methods such as LSTM or BERT might also be explored as a way of capturing ever more detailed patterns, provided the costs in computation and complexity are properly accounted for. Researchers are also encouraged to test their models across real time or multiple resources data in order to increase generalizability and diminish dependence on a given dataset. Finally, psychological or emotional variables like trust, satisfaction and perceived risk might enhance forecast predictive power to give a fuller picture of consumer intent.

REFERENCE

[1] H. GhorbanTanhaei, P. Boozary, S. Sheykhan, M. Rabiee, F. Rahmani, and I. Hosseini, "Predictive analytics in customer behavior: Anticipating trends and preferences," Results in Control and Optimization, vol. 17, p. 100462, 2024. https://doi.org/10.1016/j.rico.2024.100462

[2] S.-S. M. Ajibade, M. B. Jasser, D. O. Alebiosu, B. Issa, L. W. San, and G. S. ALDharhani, "A Comparative Analysis of Detection Methods for Phishing Websites Using Machine Learning," in 2025 IEEE International Conference on Automatic Control and Intelligent Systems (I2CACIS), 2025, vol. 1: IEEE, pp. 443-448. https://doi.org/10.1109/I2CACIS65476.2025.11100418

[3] P. Venkateswaran and S. Mm, "Predictive Analytics: Utilizing Machine Learning and Big Data for Forecasting Future Trends in Business and Consumer Behavior," in Strategic Brand Management in the Age of AI and Disruption: IGI Global Scientific Publishing, 2025, pp. 463-492. https://doi.org/10.4018/979-8-3693-9461-8.ch019

[4] H. S. Yen *et al.*, "Evaluation of KNN-NMF Algorithm for Recommendation Systems in E-Commerce," in 2024 IEEE 12th Conference on Systems, Process & Control (ICSPC), 2024: IEEE, pp. 328-332. https://doi.org/10.1109/ICSPC63060.2024.10862917

[5] D. C. Gkikas and P. K. Theodoridis, "Predicting online shopping behavior: Using machine learning and google analytics to classify user engagement," Applied Sciences, vol. 14, no. 23, p. 11403, 2024. https://doi.org/10.3390/app142311403

[6] 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

[7] B. A. S. Emambocus *et al.*, "A clustering algorithm employing salp swarm algorithm and K-means," in 2024 20th IEEE International Colloquium on Signal Processing & Its Applications (CSPA), 2024: IEEE, pp. 108-113. https://doi.org/10.1109/CSPA60979.2024.10525648

[8] S. Khaleelahmed *et al.*, "Task scheduling algorithm using grey wolf optimization technique in cloud computing environment," Bulletin of Electrical Engineering and Informatics, vol. 14, no. 4, pp. 2762-2771, 2025. https://doi.org/10.11591/eei.v14i4.7695

[9] L. Qi, Y. Xie, Q. Zhang, J. Zhang, and Y. Ma, "Multi-feature fusion-based consumer perceived risk prediction and its interpretability study," PloS one, vol. 20, no. 1, p. e0316277, 2025. https://doi.org/10.1371/journal.pone.0316277

[10] O. Bellar, A. Baina, and M. Ballafkih, "Sentiment analysis: Predicting product reviews for E-commerce recommendations using deep learning and transformers," Mathematics, vol. 12, no. 15, p. 2403, 2024. https://doi.org/10.3390/math12152403

[11] E. S. El-Shaer, G. T. McKee, and A. Hamdy, "A Hybrid LSTM and Probabilistic Model for Modeling Repeat Purchase Behavior," in 2023 International Conference on Computer and Applications (ICCA), 2023: IEEE, pp. 1-6. https://doi.org/10.1109/ICCA59364.2023.10401585

[12] T. Sangsawang, "Predicting ad click-through rates in digital marketing with support vector machines," Journal of Digital Market and Digital Currency, vol. 1, no. 3, pp. 225-246, 2024. https://doi.org/10.47738/jdmdc.v1i3.20

[13] A. Sikder, M. J. Islam, and M. W. R. Miah, "Predicting Purchase Intentions in Bangladesh's Youth Market: A Competent Analysis of Machine Learning and Deep Learning Models," in 2025 International Conference on Electrical, Computer and Communication Engineering (ECCE), 2025: IEEE, pp. 1-6. https://doi.org/10.1109/ECCE64574.2025.11013434

[14] R. Alamsyah and S. Wahyuni, "Using Random Forest and Support Vector Machine Algorithms to Predict Online Shopper Purchase Intention from E-Commerce Session Data," International Journal for Applied Information Management, vol. 4, no. 2, 2024. https://doi.org/10.47738/ijaim.v4i2.81

[15] C. Zhou, Q. Zhu, X. Zhang, and Y. Hou, "Customer Purchase Intention Prediction Based on Stacking Ensemble Learning," in 2025 IEEE 17th International Conference on Computer Research and Development (ICCRD), 2025: IEEE, pp. 36-41. https://doi.org/10.1109/ICCRD64588.2025.10963249

[16] M. S. Satu and S. F. Islam, "Modeling online customer purchase intention behavior applying different feature engineering and classification techniques," Discover Artificial Intelligence, vol. 3, no. 1, p. 36, 2023. https://doi.org/10.1007/s44163-023-00086-0

[17] M. Kulandaivel, R. Agarwal, and A. Singh, "Customer purchasing intention prediction using machine learning," in AIP Conference Proceedings, 2024, vol. 3075, no. 1: AIP Publishing LLC, p. 020234. https://doi.org/10.1063/5.0217545

[18] I. A. Khandokar, A. M. Islam, S. Islam, and S. Shatabda, "A gradient boosting classifier for purchase intention prediction of online shoppers," Heliyon, vol. 9, no. 4, 2023. https://doi.org/10.1016/j.heliyon.2023.e15163

[19] Q. Xiong, "Deep Learning in Predicting Consumer Purchase Intentions," in Proceedings of the 3rd International Conference on Signal Processing, Computer Networks and Communications, 2024, pp. 525-530. https://doi.org/10.1145/3712335.3712425

[20] S.-S. M. Ajibade *et al.*, "A Text Mining Approach for Analyzing Consumer Sentiment and Behavior," in 2025 IEEE 8th International Conference on Electrical, Control and Computer Engineering (InECCE), 2025: IEEE, pp. 660-665. https://doi.org/10.1109/InECCE64959.2025.11150868

[21] L. W. San, S. S. M. Ajibade, M. B. Jasser, A. Ayodele, B. A. Ajayi, and M. A. Bassey, "A Literature Review of Machine Learning Techniques for Dance Recognition and Robotic Vision," in 2025 IEEE International Conference on Robotics and Technologies for Industrial Automation (ROBOTHIA), 2025: IEEE, pp. 1-6. https://doi.org/10.1109/ROBOTHIA63806.2025.10986414

[22] C. S. Ma, X. R. Du, J. Lou, and M.-Q. Wang, "A User Behavior Prediction Method for Web Applications Based on Deep Forest," Journal of Web Engineering, vol. 24, no. 1, pp. 39-56, 2025.