SvedbergOpen
DISSEMINATION OF KNOWLEDGE
International Journal of Artificial Intelligence and Machine Learning
Publisher's Home Page: https://www.svedbergopen.com/

Research Paper Open Access

A Scalable Artificial Intelligence Framework for Intelligent Data Processing and Decision Support in Modern Computing Systems

Dr. Reghunath K[1*], Dr.G. Indumathi[2], Dr. R.Z Inamul Hussain[3], Chandrakant B Kadu[4], Ranjanbanerjee[5]

[1*]Associate Professor, Specialization in Computer Science & Applications,Department of Computer Applications, Marian Academy of Management Studies, Mahatma Gandhi University, Pin code: 686673, City: Kothamangalam, Country: India, Orcid Id: 0009-0001-7136-4851, Email Id: sreereghunath@gmail.com
[2]Assistant professor Specialization in MANETS, Cognitive radio adhoc networks, Machine learning, Department of Computer science and engineering, University SRM Institute of Science and technology ramapuram, Pin code: 600089, City: Bharathi Salai, Ramapuram, Chennai, Tamil Nadu, Country: India, Orcid Id: 0000-0001-5109-4826, Email Id: indumatg2@srmist.edu.in
[3]Associate professor, Department of computer science and engineering, C Abdul Hakeem college of engineering and technology, Melvisharam Tamil Nadu 632509, Email Id: inamulhasan.rz@gmail.com Orcid Id: 0009-0000-3574-8818
[4]Associate Professor, Instrumentation and Control Engineering, Pravara Rural Engineering College Loni, Email Id: chandrakant.kadu@pravara.in, Orcid Id: 0000-0001-9021-1108
[5]Assistant Professor, Department of Computer Science & Engineering, Brainware University, Barasat, Kolkata – 125, West Bengal, India, City / Pincode:700125, Orcid Id: 0009-0003-1950-7530, Email Id: rbkpccst@gmail.com

ABSTRACT
Today's computing systems are increasingly required to process heterogeneous data and transform it into trustworthy operational knowledge, and artificial intelligence (AI) based decision support is crucial for such systems. For e-commerce applications, user-session data are rich with important behavioral clues, and finding meaningful purchase-intention patterns is difficult due to the mixed data types, class imbalance, and the requirement for interpretable predictive outputs. In this study, a scalable artificial intelligence framework for intelligent data processing and decision support is proposed by using the Online Shoppers Purchasing Intention Dataset. The initial data set consisted of 12,330 records with 18 columns, 17 of which were input features and the last one was the target variable Revenue. Following duplicate removal, 12,205 records were left for analysis. The suggested model included data cleaning, exploratory analysis, preprocessing, model training, model evaluation, and feature-importance interpretation. The data were standardized for continuous variables, categorical variables were encoded, and an 80:20 train-test split was performed on the data in a stratified manner. The following classification models were tested: Logistic Regression, Decision Tree, Random Forest and Gradient Boosting. The accuracy, precision, recall, F1-score, ROC-AUC, Precision-Recall AUC, confusion matrix analysis, training time, and prediction time were used as measures of performance. The results revealed that Random Forest had the highest F1 score which shows the best balance between precision and recall, whereas Gradient Boosting had the highest ROC-AUC and accuracy. The most important predictor was PageValues, as determined by feature-importance analysis. The proposed framework facilitates e-commerce decision making and prediction of purchase intention, customer targeting, campaign planning and conversion optimization.

Keywords: Artificial Intelligence, Intelligent Data Processing, Decision Support System, Machine Learning, E-Commerce Analytics, Purchase Intention Prediction, Scalable Computing.

INTRODUCTION

Today's computing systems produce massive amounts of diverse data via digital platforms, enterprise applications, cloud services, sensors and web-based user interactions. These data environments are becoming more complex and there is a great need for intelligent systems that can process data automatically, predictively analyze and support decision-making. Machine learning has emerged as a key ingredient of such intelligent systems, as it can be used to learn patterns from data and provide adaptive decision making in application areas [1]. At the same time, artificial intelligence has gained more and more

popularity in decision-support systems, to enhance analytical efficiency, minimize manual efforts, and facilitate evidence-based decision making [2]. In the field of information systems research, predictive analytics is regarded as one of the significant methods of converting the raw data of operations into knowledge that can be used to make decisions, especially when a business needs to predict future events, user behavior, or business outcomes [3].

Intelligent data processing is particularly crucial in e-commerce settings, where online platforms constantly track user-session behavior, navigation patterns, product interactions, traffic sources, user characteristics, and transaction results. These data can be utilized to foretell purchase intent and inform choices on customer targeting, marketing campaigns, website optimization, and conversion improvement. Sakar et al. presented the real time prediction of purchasing intention of online shoppers by using multilayer perceptron and long short-term memory recurrent neural network, showing the importance of behavioural session data for the online purchase prediction [4]. This Online Shoppers Purchasing Intention Dataset has since been used as a benchmark dataset to test machine-learning models in predicting purchasing intentions in e-commerce [5]. Recently, online customer purchase intention has been further explored with the use of feature engineering and classification techniques [6], machine-learning based analysis of consumer purchase triggers [7], classification of customer purchasing behavior [8], anonymous clickstream-based purchase-intention modelling [9] and prediction of user engagement using Google Analytics [10]. From these studies, it is found that predicting purchase-intention in e-commerce is an ongoing research issue and machine learning models can be beneficial in understanding the user behaviour and business decision-making.

However, there are still some challenges to overcome to create robust AI systems for intelligent data processing and decision support. First, e-commerce data sets typically include a variety of data types, such as continuous numerical data, categorical data, encoded data and Booleans indicators. Second, online purchase-intention data sets are often class imbalanced as only a small fraction of browsing sessions end up in a purchase. Third, the accuracy of a predictive model is not enough to assess decision-support models, particularly if the minority class is the business-critical outcome. Fourth, many studies only consider the performance of model prediction, and less attention is paid to the entire method pipeline, such as data cleaning, data preprocessing, model comparison, robust evaluation, efficiency of calculation and analysis of explainable features.

In response to these challenges, this study introduces a scalable Artificial Intelligence (AI) system for intelligent data processing and decision support in modern computing systems. The Online Shoppers Purchasing Intention Dataset is used to evaluate the framework and the framework is designed to process the heterogeneous e-commerce session data through a structured pipeline of data collection, data cleaning, exploratory analysis, data preprocessing, model training, model evaluation and decision-support interpretation. A reproducible machine-learning workflow was realized with the use of Scikit-learn Python library that offers common tools for data preprocessing, model construction, and assessment [11]. Logistic Regression, Decision Tree, Random Forest and Gradient Boosting models are compared to see which model is suitable for the purchase-intention prediction. Random Forest is included due to its capability of enhancing the generalization using ensemble learning [12] and Gradient Boosting is included as it has strong predictive performance by sequential error correction [13]. As a baseline model, Decision Tree is used as an interpretable model, which is based on the classical classification and regression tree framework [14].

The primary goal of this study is to create and test a scalable AI-based system able to transform the raw ecommerce session data into predictive and interpretable decision-support outputs. The specific aim is not only to classify whether a user session will result in a purchase or not, but also to assess the behavior of the model when faced with class imbalance and determine which features are most influential for the purchase predictions. The minority class (purchase generating class) is underrepresented and therefore evaluation takes into account minority class learning challenges and includes metrics like precision, recall, F1-score, ROC-AUC, Precision-Recall AUC and confusion matrix analysis. The problem of class imbalance is of great methodological importance in binary classification since the performance of the minority class is not evaluated correctly, standard classifiers may become biased towards the majority class [15].

This study is a contribution in three aspects. Firstly, it offers a comprehensive and replicable AI architecture for intelligent data processing and decision-making based on a real e-commerce purchase-intention dataset. Secondly, it offers a comparative assessment of the baseline and ensemble machine-learning models based on performance metrics appropriate for imbalanced classification. Third, it includes feature-importance analysis for interpretability and use in practice. The proposed framework is thus applicable to the current computing systems where scalable data processing, predictive modelling and explainable decision support is needed to achieve effective business intelligence.

2. MATERIALS AND METHODS

The dataset, proposed AI-based framework, data pre-processing, machine-learning models and evaluation

metrics for developing an intelligent decision-support system for online purchase-intention prediction are described in this section. The methodological design is structured in a pipeline from data collection to decision-support interpretation, making sure that the proposed framework is still applicable for a scalable and efficient implementation in modern computing environments.

2.1 DATASET DESCRIPTION

The experimental analysis was based on Online Shoppers Purchasing Intention Dataset [5] that consists of user-session level behavioral data from an ecommerce setting. Administrative page visits, informational page visits, product related page visits, duration of page visited, bounce rate, exit rate, page value, traffic related characteristics, visitor type, weekend activity and purchase outcome are all part of the dataset. The target variable is Revenue, which represents if a user session had a purchase or not. Hence, the prediction task was modeled as a binary classification problem, in which the model classifies the class of the session to be either non-purchase class or purchase generating class. The original data set had 12330 records, 18 columns and 17 input features. In the data cleaning process, 125 duplicate data were deleted in the final data cleaning process, resulting in 12,205 data. The characteristics of the data set and the data preprocessing configuration used in the study are summarized in Table 1.

Table 1. Dataset Description and Preprocessing Summary

Aspect	Description
Dataset name	Online Shoppers Purchasing Intention Dataset
Application domain	E-commerce purchase intention prediction
Original records	12,330
Cleaned records	12,205
Input features	17
Target variable	Revenue
Problem type	Binary classification
Missing values	0
Duplicate records removed	125
Numerical preprocessing	StandardScaler
Categorical preprocessing	OneHotEncoder
Train-test split	80:20 stratified split
Evaluation metrics	Accuracy, Precision, Recall, F1-score, ROC-AUC, PR-AUC, Confusion Matrix

Table 1 indicates that the dataset is suitable for evaluating an artificial intelligence framework because it contains heterogeneous numerical, categorical, and Boolean variables. The presence of both behavioral and technical attributes makes it appropriate for intelligent data processing and decision-support modeling.

The class distribution of the target variable was also examined before model development. Since purchase-generating sessions are less frequent than non-purchase sessions, the dataset exhibits class imbalance. Table 2 presents the distribution of the target variable after duplicate removal.

Table 2. Target Class Distribution

Revenue Class	Count	Percentage
False	10,297	84.3671
True	1,908	15.6329

Table 2 shows that the positive class, represented by Revenue=True, accounts for only 15.6329% of the cleaned dataset. This imbalance influenced the selection of evaluation metrics, particularly the use of F1-score and Precision-Recall AUC in addition to accuracy.

2.2 PROPOSED SCALABLE AI FRAMEWORK

The proposed framework aims to facilitate intelligent data processing and decision making in a modern computing environment. It is based on a modular pipeline that involves seven main steps: dataset collection, data cleaning, exploratory analysis, data preprocessing, model training, model evaluation and decision support. These stages all help to turn raw e-commerce session data into actionable predictive

insights.
The framework starts with the collection of datasets, which consists of loading raw data from the session level into the framework for analysis. Data cleaning can then be done to find missing values, duplicate data, and inconsistencies in the data structure. Target imbalance, categorical distributions, numerical patterns, correlations, outliers and skewness are explored using exploratory analysis. This is followed by preprocessing the heterogeneous variables to be understood by machine learning. These data are then processed and used for training several machine learning models. Lastly, the results of model outputs, evaluation scores, and feature-importance results are interpreted for decision making. The overall process of the proposed AI framework is shown in Figure 1. This modularity can be of great benefit for scalability, as each module can be altered, extended or optimized separately, based on the computing environment and the decision-support requirement.

Figure 1. Proposed Scalable AI Framework for Intelligent Data Processing and Decision Support

The proposed framework is particularly suitable for e-commerce analytics because it connects predictive modeling with practical decision support. Instead of only classifying user sessions, the framework also identifies important behavioral features that can guide conversion optimization, targeted marketing, and visitor segmentation.

2.3 DATA CLEANING AND EXPLORATORY ANALYSIS

The dataset was analysed for missing values, duplicate records, outliers and class imbalance prior to model training. There were no missing values in the dataset and it was not necessary to impute missing values. But 125 duplicate records were detected and excluded from model learning to avoid biased learning. Following cleaning, there were 12,205 records.
Outlier analysis was done using interquartile range method. There were outlying values for several variables: PageValues, ProductRelated_Duration, Administrative_Duration, and Informational_Duration. These outliers were kept as they might indicate real customer behavior and not data-entry errors: High page duration, high number of page visits and high page value. The removal of such observations may make the model less useful for decision support. An exploratory analysis was also conducted which showed that the target variable was imbalanced. Revenue=False was the majority class while Revenue=True was the minority class. Figure 2 shows the result of duplicate removing and the final target class distribution.

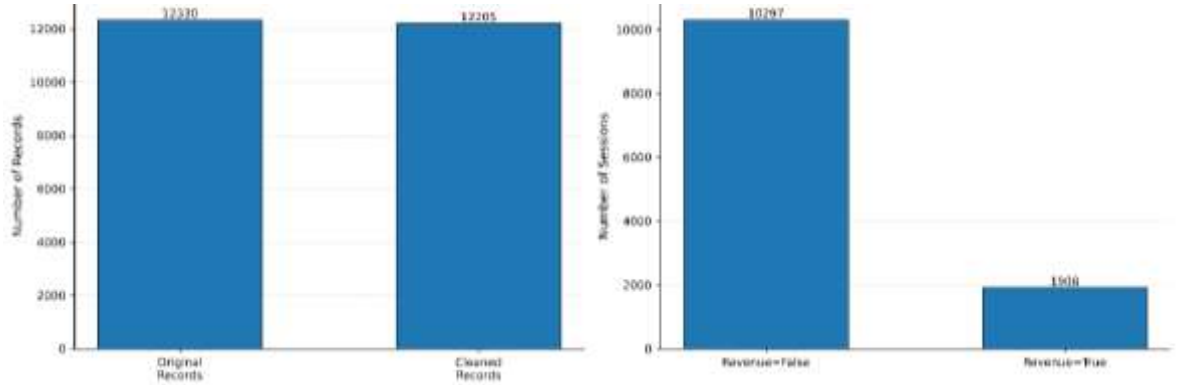

Figure 2. Dataset Cleaning and Target Class Distribution Summary

The effect of duplicate removal on the size of the dataset was minimal but important as seen in Figure 2. It also shows that there is an imbalance between the purchase and non-purchase sessions and so evaluation measures that take into account the quality of prediction of minority class are required.
The results from the exploratory analysis informed the preprocessing and modeling decisions. The data included both continuous numerical and categorical variables, so a pipeline-based approach to preprocessing was used to ensure consistency between all machine-learning models.

2.4 DATA PREPROCESSING

The target variable Revenue was recoded into numeric format (False=0, True=1). This change enabled the classification models to treat the target variable as a binary variable. The input features were split into continuous numerical features and categorical features. The numerical features were continuous and scaled using StandardScaler which removes the mean and scales to unit variance. This is especially helpful for models like Logistic Regression that can be affected by the scale of features when the model is

converging or when estimating coefficients. OneHotEncoder was used to convert the categorical variables. This approach results in each category being transformed into a binary indicator variable, and thus categorical data like Weekend, OperatingSystems, Browser, Region, TrafficType and Month can be utilized by machine-learning algorithms. Some of the categorical variables were originally coded numerically, but were coded as categorical attributes because the numeric values were labels for categories. Stratified 80:20 train-test split was used to ensure that the purchase and non-purchase sessions were maintained in the train and test sets, respectively. The data set was imbalanced and required stratification. If no stratification, minority class may be underrepresented in training or testing set which results in unreliable model evaluation. The original feature space was expanded to a machine-learning ready space through preprocessing and encoding. With this transformation, the models were able to handle all of the different types of data in a single pipeline.

2.5 MACHINE LEARNING MODELS

Logistic Regression, Decision Tree, Random Forest and Gradient Boosting were chosen to assess the proposed framework. These models were selected due to their relatively good interpretability, computational efficiency, and prediction power.

The linear classifier used as a baseline was Logistic Regression. It is simple, efficient, and can be used as a reference model for binary classification, for comparison of the performance of the more complex algorithms.

Decision Tree was added as a baseline interpretable model. Decision Trees are easy to understand because they split observations according to rules, and classify them. They can, however, be sensitive to data variation and can over fit if not controlled appropriately.

Random Forest was chosen as a good ensemble model. Constructs several decision trees and aggregates their predictions to minimize overfitting and to enhance generalization. Random Forest can be used with heterogenous tabular data and it can manage nonlinear relationship among the user's behavior and purchase intention.

The Sequential Ensemble learning method was chosen as Gradient Boosting. Constructs models incrementally, with each successive learner trying to fix previous learners' mistakes. This renders Gradient Boosting well suited for high predictive performance especially in structured classification problems.

All the models were realized in a pipeline fashion where the preprocessing and classification steps were cascaded into a single pipeline. This guaranteed that the preprocessing steps were applied in the same way when training and testing.

2.6 EVALUATION METRICS

The evaluation of the models was done with various performance metrics: Accuracy, Precision, Recall, F1-score, ROC-AUC, Precision-Recall AUC, Confusion Matrix, training time and prediction time. Multiple metrics were needed as the target variable was imbalanced.

The overall accuracy is the percentage of correct predictions. In an imbalanced dataset, however, accuracy is not a good indicator as the model could be accurate by predicting the majority class most of the time. Thus, other metrics were employed to evaluate the quality of prediction of minority class. Precision is the percentage of purchase sessions predicted that actually were purchase sessions. Recall is a measure of how many purchase sessions the model was able to detect correctly. F1-score takes into account both precision and recall, and is a balanced score for this reason, especially for imbalanced classification. ROC-AUC is a measure of how well the model discriminates between the positive and negative classes at various classification thresholds. Precision-Recall AUC was also used as it is more informative in the case of rare positive classes. PR-AUC was particularly significant in this study as the ratio of Revenue=True was only 15.6329% of the cleaned data. True negatives, false positives, false negatives and true positives were studied using confusion matrix. This enabled interpretation of the model errors, in particular the balance between purchase session identification and false identification of a non-purchase session as a purchase session.

The time taken for training and prediction was measured to check the computational efficiency. These metrics help to demonstrate the scalability part of the proposed framework, by indicating if the models can process the dataset in a timely manner and still provide a useful prediction.

3. Results

This section shows the empirical results which has been derived from the proposed AI framework. The dataset description and summary of data preprocessing, the design of the framework and the distribution of classes have already been reported in Materials and Methods section, thus, the present section

addresses the predictive performance, classification behavior, threshold-based evaluation and interpretation of feature importance. The results show the effectiveness of the proposed framework to

achieve intelligent data processing and purchase-intention decision support.

3.1 DATASET AND TARGET DISTRIBUTION RESULTS

The cleaned data set, as described above in Table 2 and Figure 2, contained 12,205 sessions, of which 10,297 weren't purchases and 1,908 did result in purchases. The number of positive class (Revenue=True) was 15.6329% and the number of negative class (Revenue=False) was 84.3671% of the cleaned data.

This distribution verifies class imbalance situation in the prediction task. Such an environment makes only accuracy as a whole unreliable in order to measure the effectiveness of a model as a classifier might get a high accuracy score by just predicting the class with higher frequency. Hence, precision, recall, F1-score, ROC-AUC, Precision-Recall AUC and confusion matrix were chosen as the main metrics for evaluating the experimental results. These metrics are more accurate in determining the model's effectiveness in recognizing sessions that lead to purchases.

The imbalance observed also has practical implications with regards to the decision-support systems. It is commonly agreed that only a fraction of people who look at products end up buying them in the world of e-commerce. Thus, an effective intelligent decision support system should not only have a high accuracy rate in the classification of the majority class (non-purchase) but should also have a reasonably high accuracy rate on the minority class (purchase).

3.2 MODEL PERFORMANCE RESULTS

The performance of four machine-learning models was evaluated: Logistic Regression, Decision Tree, Random Forest, and Gradient Boosting. Table 3 presents the comparative performance of these models in terms of accuracy, precision, recall, F1-score, ROC-AUC, training time, and prediction time.

Table 3. Model Performance Comparison

Model	Accuracy	Precision	Recall	F1-score	ROC-AUC	Training Time (s)	Prediction Time (s)
Random Forest	0.8955	0.6480	0.7277	0.6856	0.9229	0.4077	0.0506
Gradient Boosting	0.9037	0.7168	0.6361	0.6741	0.9367	2.6362	0.0165
Logistic Regression	0.8460	0.5050	0.7958	0.6179	0.9097	0.1438	0.0062
Decision Tree	0.8554	0.5404	0.5079	0.5236	0.7139	0.1089	0.0111

The results show that Random Forest achieved the highest F1-score of 0.6856, indicating the best balance between precision and recall. This is particularly important in the present study because the purchase class is underrepresented. A high F1-score suggests that the model is able to identify purchase sessions while controlling false positive predictions.

Gradient Boosting had the best accuracy (0.9037) and best ROC-AUC (0.9367) showing good overall classification and probability-ranking performance. While its F1-score was slightly less than Random Forest, it showed higher precision and ROC-AUC, indicating that it was more conservative and discriminative in assigning purchase probabilities.

Logistic Regression had the highest recall (0.7958), which meant that it identified more purchase sessions. Its lower precision value of 0.5050, however, suggests that this increase in recall came at the cost of the higher number of false positive predictions. Decision Tree had the lowest F1 score and ROC-AUC values among the tested models, indicating that this single tree-based model was not as effective as the ensemble-based models for this dataset. A graph showing the key performance indicators for the four models is shown in figure 3.

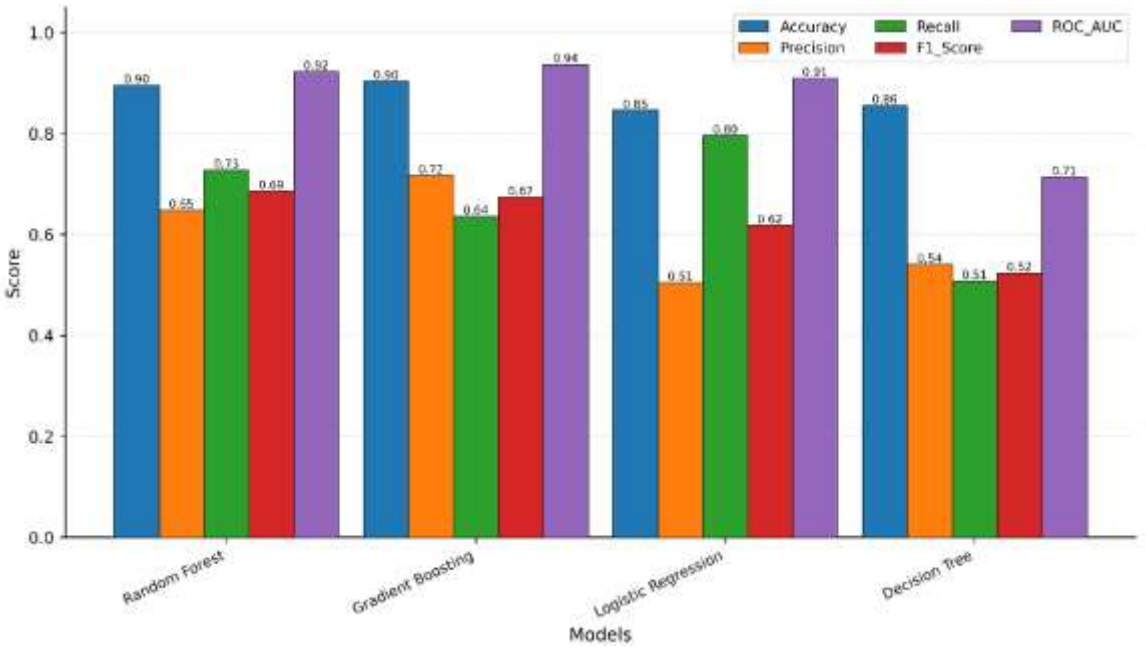

Figure 3. Machine Learning Model Performance Comparison

Figure 3 confirms the advantage of ensemble-based learning models. Random Forest yielded the best balanced classification result, while Gradient Boosting had the highest overall accuracy and ROC-AUC. This shows that ensemble techniques are more appropriate for the proposed Decision-support framework than a single Decision Tree model.

3.3 CONFUSION MATRIX RESULTS

To further evaluate classification behavior, confusion matrix analysis was performed for all models. This analysis provides a practical interpretation of model predictions by showing true negatives, false positives, false negatives, and true positives. Table 4 presents the confusion matrix summary.

Table 4. Confusion Matrix Summary

Model	True Negative	False Positive	False Negative	True Positive	Correct Predictions	Incorrect Predictions
Random Forest	1,908	151	104	278	2,186	255
Gradient Boosting	1,963	96	139	243	2,206	235
Logistic Regression	1,761	298	78	304	2,065	376
Decision Tree	1,894	165	188	194	2,088	353

The confusion matrix results reveal that Gradient Boosting had the largest number of correct predictions (2,206 sessions) and the lowest false positive predictions (96 sessions). This means that Gradient Boosting was successful in minimizing the number of times it predicted a purchase that was incorrect, and it was more selective in assigning the positive class.

Random Forest demonstrated high balance between purchase session detection and control of false positives. It correctly captured 278 purchase sessions, with 151 false positives and 104 false negatives. This balance enables it to achieve the highest F1 score of the tested models.

Logistic Regression was able to correctly identify a purchase session in 304 sessions, the highest number of true positives among all models. It also, however, produced a significant number of false positives (298) showing that its higher recall was at the cost of less precision. The results of Decision Tree were not as good as other classifiers in terms of detection of minority class with 194 true positives and 188 false negatives, which is in line with the results of its limited suitability to imbalanced classification setting.

From the overall results, the confusion matrix analysis shows that the Random Forest and Gradient Boosting classifiers exhibited better and stable classification behavior as compared to Logistic Regression and Decision Tree. Random Forest was more balanced in treating the minority class and Gradient Boosting was more conservative with fewer false positives.

3.4 ROC AND PRECISION-RECALL RESULTS

ROC and Precision-Recall analyses were conducted to examine model performance across classification thresholds. ROC-AUC measures the ability of a model to distinguish between purchase and non-purchase sessions, while Precision-Recall analysis focuses more directly on the positive class. Since the positive class represented only 15.6329% of the cleaned dataset, Precision-Recall analysis was particularly important for evaluating practical purchase-session detection. Figure 4 presents the combined ROC and Precision-Recall curves for all evaluated models.

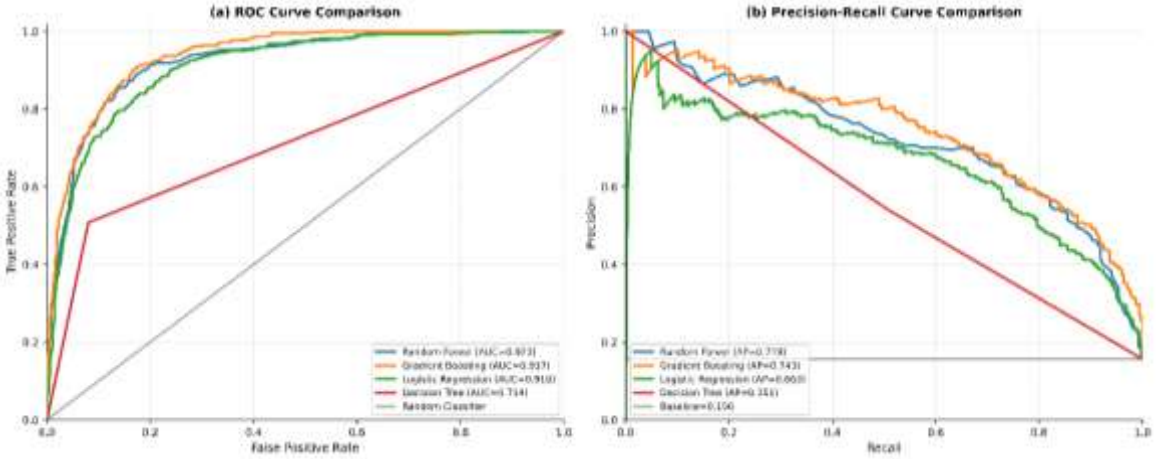

Figure 4. ROC and Precision-Recall Analysis of Baseline Models

The results of ROC curve indicate that the Gradient Boosting has the highest ROC-AUC of 0.9367 followed by the Random Forest which is 0.9229. Logistic Regression also yielded competitive threshold-independent results with a ROC-AUC of 0.9097 while Decision Tree yielded significantly poor results with a ROC-AUC of 0.7139.

The difference between the two algorithms (Precision-Recall curve) also shows that Gradient Boosting and Random Forest outperform each other for purchase generating sessions. The model Gradient Boosting had the best mean precision, followed by Random Forest. Decision Tree exhibited much lower Precision-Recall and therefore was not as reliable in predicting the minority class.

These findings validate that ensemble-based models are more effective for Purchase Intention prediction in e-commerce domain where classes are imbalanced. This ROC/PR evaluation also points to the need for multiple metrics and not just accuracy.

3.5 FEATURE IMPORTANCE RESULTS

Feature-importance analysis was conducted for Random Forest and Gradient Boosting to identify the most influential variables contributing to purchase-intention prediction. This analysis supports the interpretability component of the proposed decision-support framework by linking predictive outcomes to meaningful behavioral and session-level indicators. Table 5 presents the top-ranked important features for the two ensemble models.

Table 5. Top Feature Importance Summary

Model	Rank	Feature	Importance
Random Forest	1	PageValues	0.330906
Random Forest	2	ExitRates	0.095878
Random Forest	3	ProductRelated_Duration	0.077646
Random Forest	4	ProductRelated	0.065109
Random Forest	5	BounceRates	0.052791
Random Forest	6	Administrative_Duration	0.047801
Random Forest	7	Administrative	0.040860
Random Forest	8	Month_Nov	0.026453
Random Forest	9	Informational_Duration	0.021078

Random Forest	10	Informational	0.015344
Gradient Boosting	1	PageValues	0.762279
Gradient Boosting	2	Month_Nov	0.046340
Gradient Boosting	3	BounceRates	0.042483
Gradient Boosting	4	ProductRelated_Duration	0.029540
Gradient Boosting	5	Administrative	0.029199
Gradient Boosting	6	ExitRates	0.024815
Gradient Boosting	7	ProductRelated	0.012866
Gradient Boosting	8	Administrative_Duration	0.012124
Gradient Boosting	9	Month_May	0.007962
Gradient Boosting	10	VisitorType_Returning_Visitor	0.006621

The feature-importance results show that PageValues was the most influential predictor in both Random Forest and Gradient Boosting. In Gradient Boosting, PageValues' importance score was 0.762279, which was the highest among the various importance scores. This means that sessions on high-value pages will be more likely to lead to revenue.

Other significant factors were ExitRates, BounceRates, ProductRelated, ProductRelated_Duration and Month_Nov. These variables have a meaning in the context of e-commerce. Product-related activity and product-related duration are measures of user engagement with product, whereas bounce and exit rates are measures of disengagement or end of session. The significance of Month_Nov indicates that there is a seasonal purchasing behavior. Figure 5 shows graphically the top feature-importance rankings for Random Forest and Gradient Boosting.

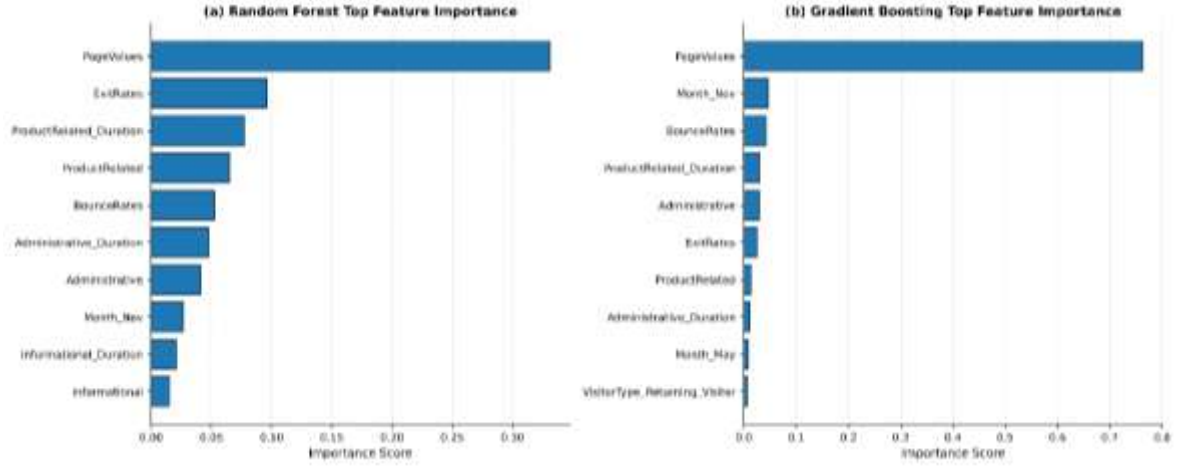

Figure 5. Feature Importance Analysis for Decision Support

Figure 5 shows that the proposed framework provides interpretable outputs in addition to predictive performance. The identification of PageValues, engagement-related variables, and seasonal indicators supports practical decision-making in e-commerce systems. These findings can assist in targeted marketing, campaign timing, visitor segmentation, and conversion optimization. In summary, the results demonstrate that the proposed scalable AI framework can process heterogeneous e-commerce session data, evaluate multiple predictive models, handle imbalanced classification, and generate interpretable decision-support insights. Random Forest provided the strongest balanced classification performance, whereas Gradient Boosting achieved the highest accuracy, ROC-AUC, and probability-ranking capability.

4. DISCUSSION

The results show that the e-commerce session data, which are heterogeneous, can be effectively processed using the proposed scalable artificial intelligence framework, and that purchase-intention prediction can be achieved. It is a methodological pipeline that combines data cleaning, exploratory analysis, data-preprocessing, model training, model evaluation and feature-importance interpretation. It is ideal for

modern computing systems as it can take raw behavioral and technical session attributes and turn it into a predictive output that can help make decisions. The purified data set included different categories of features such as numerical, categorical, encoded, Boolean, etc. and the proposed framework could scale, encode, split it using the stratification method and classify using the model-based classification. The results show that this AI-assisted framework can offer a predictive performance and analytical support in an interpretable way for e-commerce decision environments.

Comparative results with the model demonstrated that the Random Forest model obtained the maximum F1 score and the Gradient Boosting model obtained maximum ROC-AUC and Precision-Recall. Random Forest worked well because it stabilizes individual DTs and helps to ensure some kind of generalization over the different tabular features. The highest F1 score confirms it had the best balance between precision and recall, which is crucial when the goal is to identify quality sessions that can drive purchases, while minimizing the number of false positive sessions. Gradient Boosting, however, demonstrated more effective threshold-independent discrimination, which is because while training learners sequentially, it corrects the errors made by the preceding learner. This justified its excellent ROC-AUC and Precision-Recall score as it demonstrated an ability to better classify purchase and non-purchase sessions across all classification thresholds. In this context, the use of ROC analysis is valid as it is a method which evaluates the discriminative capability of classifiers not at a fixed threshold but at any level [16].

Models' performances were also interpreted with respect to the class distribution of the data set. Because the number of samples in the Revenue=True class was only 15.6329% of the cleaned data, using accuracy as the only metric to assess the effectiveness of the model was not adequate. A model can become quite accurate by simply being biased towards the majority class (non-purchasing sessions) and not identifying purchase generating sessions. Therefore, precision, recall, F1-score, ROC-AUC and Precision-Recall AUC were more meaningful to assess model behaviour. Precision and recall were used to explain the trade-off between correctly identifying purchase sessions and avoiding incorrect purchase prediction and F1-score summarized this trade-off by using a balanced measure. Precision-Recall was particularly significant as it deals with the performance of classifiers on positive class and is more informative than ROC analysis for imbalanced datasets [17].

For both Random Forest and Gradient Boosting, the feature importance plot showed that the most important feature is PageValues. This means that the pages visited were very strongly correlated with purchase. There are other interesting factors like ExitRates, BounceRates, ProductRelated, ProductRelated_Duration, Month_Nov for additional behavioral data. A high product-related activity and product-related duration might indicate that the user is highly interested in making a purchase, while a high bounce rate and exit rate could mean that the user has been on the site for a short period and may have had little interest in the products. The result of Month_Nov shows that the seasonality shopping attributes are significant leading to the prediction of purchase-intention. These can be used to segment visitors, target marketing, optimize conversions, and time campaigns for a decision support standpoint. Explainable machine learning predictions, for instance, in the context of decision support, lead users to understand the factors that affect predictions made by automated systems [19] and this is important for the interpretable model outputs as well.

Today's computing systems are also being affected by the proposed framework. It offers a structured approach at converting the e-commerce data of the operations into useful data for decision support. The framework can serve organizations to distinguish purchase-oriented customers, customize marketing actions and evaluate them robustly, and interpret the results. The greater success of the ensemble methods also suggests that scalable tree-based learning methods should be used in structured data settings. Previous research on scalable tree boosting has demonstrated the ability of ensemble-based approaches to deliver efficient and performant predictive solutions for large-scale data mining problems [18]. These kinds of predictive models are useful in business analytics not only for predicting but also for data driven reasoning and decision making in the operations [20].

However, it has to be noted that there are some drawbacks of this study. One drawback was that the analysis was based on just one benchmark data set, which may restrict the applicability of the findings to other ecommerce sites or industrial applications. Second, the data is historical data from sessions and isn't real-time streaming behavior or continually updated user profiles. Thirdly, the proposed framework underwent experimental testing, but was never implemented as a real-time decision-support system. Fourth, the study excluded deep learning models which might be applicable for sequential modelling of clickstream data or in case of massive behavioural data. Fifth, the dataset is imbalanced and while the proper evaluation metrics were employed, more imbalance handling techniques can be employed to further enhance minority class detection. Lastly, there are some encoded categorical variables, like operating system, browser, region, traffic type, which do not have detailed semantic labels, so it's difficult to make deeper business interpretation. Future research can aim to overcome these limitations by testing the framework with larger and more varied datasets, real-time deployment, comparison with other deep learning models, and by considering more sophisticated explainability methods.

5. CONCLUSIONS

This study proposed a scalable AI model that can be employed in the modern computing system for intelligent data processing and decision support using the Online Shoppers Purchasing Intention Dataset. The components of this framework were: Data cleaning, Exploratory Data Analysis, Data Preprocessing, Classification using machine learning, Model Evaluation, and Interpretation of the results (Feature importance). After removing the duplicate records, the final data set was 12,205 sessions, but only 15.6329% of these sessions were purchase generating sessions, which is clear indication of an imbalance classification problem. The following models were tested: Logistic Regression, Decision Tree, Random Forest and Gradient Boosting. The findings revealed that Random Forest outperformed with the best score of 0.6856 for F1-score, which reflects a good balance of precision and recall, while Gradient Boosting showed the highest accuracy (0.9037) and ROC-AUC (0.9367) scores, which are indicative of the best probability-ranking performance. The feature-importance analysis revealed PageValues as the most important feature followed by some engagement related features (ExitRates, BounceRates, ProductRelated, ProductRelated_Duration and Month_Nov). Based on these findings, it is likely that this proposed framework will be able to facilitate purchase-intention prediction, customer targeting, campaign planning and conversion optimization. Future studies could include expanding the framework with larger multi-domain datasets, deploying it in real-time, developing better explainability techniques, implementing deep learning models, and hyperparameter optimization.

REFERENCES

1. M.N. Injadat, A. Moubayed, A. Bou Nassif, A. Shami. “Machine Learning Towards Intelligent Systems: Applications, Challenges, and Opportunities”, Artificial Intelligence Review, LIV, pp. 3299–3348, 2021.
2. S. Gupta, S. Modgil, S. Bhattacharyya, I. Bose. “Artificial Intelligence for Decision Support Systems in the Field of Operations Research: Review and Future Scope of Research”, Annals of Operations Research, CCCVIII, pp. 215–274, 2022.
3. G. Shmueli, O.R. Koppius. “Predictive Analytics in Information Systems Research”, MIS Quarterly, XXXV (3), pp. 553–572, 2011.
4. C.O. Sakar, S.O. Polat, M. Katircioglu, Y. Kastro. “Real-Time Prediction of Online Shoppers’ Purchasing Intention Using Multilayer Perceptron and LSTM Recurrent Neural Networks”, Neural Computing and Applications, XXXI (10), pp. 6893–6908, 2019.
5. UCI Machine Learning Repository, “Online Shoppers Purchasing Intention Dataset”, Available at: https://archive.ics.uci.edu/dataset/468/online+shoppers+purchasing+intention+dataset .
6. M.S. Satu, S.F. Islam. “Modeling Online Customer Purchase Intention Behavior Applying Different Feature Engineering and Classification Techniques”, Discover Artificial Intelligence, III, Article no. 36, 2023.
7. S.K. Trivedi, P. Patra, P.R. Srivastava, J.Z. Zhang, L.J. Zheng. “What Prompts Consumers to Purchase Online? A Machine Learning Approach”, Electronic Commerce Research, 2022.
8. G. Chaubey, P.R. Gavhane, D. Bisen, S.K. Arjaria. “Customer Purchasing Behavior Prediction Using Machine Learning Classification Techniques”, Journal of Ambient Intelligence and Humanized Computing, 2022.
9. Z. Wen, W. Lin, H. Liu. “Machine-Learning-Based Approach for Anonymous Online Customer Purchase Intentions Using Clickstream Data”, Systems, XI (5), Article no. 255, 2023.
10. D.C. Gkikas, P.K. Theodoridis. “Predicting Online Shopping Behavior: Using Machine Learning and Google Analytics to Classify User Engagement”, Applied Sciences, XIV (23), Article no. 11403, 2024.
11. F. Pedregosa, G. Varoquaux, A. Gramfort, V. Michel, B. Thirion, O. Grisel, M. Blondel, P. Prettenhofer, R. Weiss, V. Dubourg, J. VanderPlas, A. Passos, D. Cournapeau, M. Brucher, M. Perrot, E. Duchesnay. “Scikit-Learn: Machine Learning in Python”, Journal of Machine Learning Research, XII, pp. 2825–2830, 2011.
12. L. Breiman. “Random Forests”, Machine Learning, XLV, pp. 5–32, 2001.
13. J.H. Friedman. “Greedy Function Approximation: A Gradient Boosting Machine”, The Annals of Statistics, XXIX (5), pp. 1189–1232, 2001.
14. L. Breiman, J.H. Friedman, R.A. Olshen, C.J. Stone. Classification and Regression Trees, Wadsworth International Group, Belmont, California, 1984.
15. N.V. Chawla, K.W. Bowyer, L.O. Hall, W.P. Kegelmeyer. “SMOTE: Synthetic Minority Over-sampling Technique”, Journal of Artificial Intelligence Research, XVI, pp. 321–357, 2002.
16. T. Fawcett. “An Introduction to ROC Analysis”, Pattern Recognition Letters, XXVII (8), pp. 861–874, 2006.
17. T. Saito, M. Rehmsmeier. “The Precision-Recall Plot Is More Informative than the ROC Plot When Evaluating Binary Classifiers on Imbalanced Datasets”, PLOS ONE, X (3), Article no. e0118432, 2015.

18. T. Chen, C. Guestrin. "XGBoost: A Scalable Tree Boosting System". In Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining(KDD), pp. 785–794, 2016.
19. S.M. Lundberg, S.I. Lee. "A Unified Approach to Interpreting Model Predictions". In Proceedings of Advances in Neural Information Processing Systems, pp. 4765–4774, 2017.
20. F. Provost, T. Fawcett. Data Science for Business: What You Need to Know About Data Mining and Data-Analytic Thinking, O'Reilly Media, Sebastopol, California, 2013.