Journal of
Data Acquisition and Processing

ISSN: 1004-9037
https://sjcjycl.cn/
DOI: 10.5281/zenodo.778229

ANALYSIS OF CONSUMER BEHAVIOUR USING MACHINE LEARNING TO DETERMINE THE EFFECT OF PRODUCT INFORMATION READING ON E-COMMERCE WEBSITES

Subhash Chandra Jat[1] Neha Malav[2]

[1]Rajasthan College of Engineering for Women, Jaipur
[2]Rajasthan College of Engineering for Women, Jaipur

Abstract. A large number of people now engage in electronic commerce, thanks to the proliferation of contemporary gadgets and their advanced applications, as well as technological breakthroughs that have enabled products and services to be available over the Internet. Customers may feel overwhelmed and find it challenging to locate the ideal product due to the huge selection and diversity of products available on e-commerce websites. Due to the increased competition among international commercial sites brought about by these reasons, it is more important than ever to function profitably and efficiently. The performance of e-commerce platforms is something that can be enhanced by machine learning-based systems, which make it easier for users to locate products that suit their preferences. Many different machine learning algorithms serve this function. A major purpose of this study work is to assist proprietors of E-commerce websites based on machine learning methods to analyse consumer behaviour. An e-commerce site's performance can be improved with the help of a system developed in this study; this system uses an analysis of customers' behaviours to inform decision-making. To tackle these problems and generate highly accurate predictions with grid search for stacked model results, machine learning-based stacking algorithms (XGBoost and CatBoost) are developed. Utilising criteria like F1-Score, ROC-AUC, Accuracy, Precision, and Recall to assess system performance, experimental findings demonstrate the project's contribution. According to the experimental results, applying ML techniques enhances decision-making, which in turn increases the accuracy of suggestion lists that are provided to customers.

Keywords:*E-Commerce Websites, Customer Behavior, Machine Learning, Artificial intelligence, Supervised machine learning algorithms, Stacking, XGBoost and CatBoost, GridSearchCV,*

1 Introduction

Consumer behaviour in online transactions is a rapidly expanding field of critical research, thanks in large part to the growth of e-commerce. Predicting customer behaviour in the context of e-commerce is becoming more and more important as a result of the shift that is now occurring between visiting physical stores and shopping online. Sales volume for e-commerce systems (EC) has increased significantly in recent years, particularly with a rapid advancement of technology and an expansion of online services[1][2][3]. Because of this reality, a lot of big businesses have emerged, and there is now more competition among them to draw in as many customers as possible and generate the maximum revenues[4][5][6]. A larger product selection, more frequent sales and discounts, simplified payment processes, and an emphasis on the customer's needs in product discovery are all necessary for a business to remain competitive[7][8]. One method to simplify shopping for consumers is to implement a recommendation system. This would include creating a list that, based on consumers' past purchases, would suggest products to them. Creating recommendation systems that increase the efficacy of commercial sites has been the subject of several published research[9].

Many aspects of our everyday lives have been dramatically altered since the advent of the internet. In today's internet-based culture, electronic commerce is one of the fastest-growing industries. Customers are excited to shop at online stores such as Flipkart, eBay, and Amazon. Users of internet platforms can also leave reviews for products they have bought. It may be beneficial to make decisions about marketing tactics, how to improve goods and services, and how to treat consumers and sellers[10]. These days, before making any purchases or using any services, customers are eager to read evaluations. This provides an opening for opinion spammers, who may post fake reviews in an effort to raise or reduce the profile of companies, products, and services. However, a number of variables, most notably technology improvements, contribute to a number of hopeful trends in the global e-commerce environment[11].

These days[10], machine learning and data mining are commonplace in consumer behaviour models. Predicting how customers would act is a difficult and uncertain task. As a result, in building consumer behaviour models, the right approach must be taken. Once developed, a prediction model might be difficult to adjust so that a marketer can determine with accuracy what marketing actions to take for each individual consumer or group of customers. Therefore, the appropriate method needs to be used when creating models of consumer behaviour. After a prediction model is developed, it might be difficult to make changes to it so that a marketer is able to accurately identify the best course of action for each customer or group of customers. This study aims to showcase several research projects that have been undertaken on the analysis of consumer behaviour using different data mining and ML techniques[12][13]. A machine learning forecast model can help business people figure out and guess the factors that will most likely change how their customers will act. Because of this, the current study finds better ways to help businesses make decisions by using machine learning and data analytics to better understand how customers behave[14].

1.1 Motivation

Due to the exponential growth of e-commerce, businesses need to study how consumers utilise and navigate various online shopping platforms. Interestingly, a significant but little-studied aspect of online consumer behaviour is the amount of time that customers spend reading product information. The results of this research could help inform more successful website design and marketing strategies by shedding light on the decision-making processes of customers[15]. Furthermore, investigations into user retention mechanisms, specifically within the context of online learning platforms, contribute additional insight into the behaviour of continuance intention[16]. This paper aims to summarise previous efforts to analyse consumer behaviour by means of machine learning and data mining techniques. Our research makes several significant contributions to the study of electronic commerce:

- In this research, we offer a novel ML approach to gaining insight into online shopper behaviour by mining the UCI-ML Repository's online Shoppers' Purchasing Intention Dataset.
- Numerous factors, including the length of time spent on product-related information, the frequency with which users navigate to administrative pages, the frequency with which visitors leave the site, and whether or not they are repeat customers, influence the characteristics of visitors to e-commerce websites.
- We used clustering algorithms, box plots, pair plots, and descriptive statistics for exploratory data analysis. Additionally, as the article discusses, we assessed the validity of parameter performance using particular measures.
- Using the GridSearch method, we determined the best parameters for our stacked models (which included XGBoost and CatBoost classifiers) and performed validation on our work.
- Determining the optimum performance metrics, like confusion matrix, ROC-AUC, f1-score, recall, accuracy, and precision, with which to assess the models.
- We analysed the data and built the models using the help of Python and a number of libraries like scikit-learn, statsmodels, pandas, and numpy.

This article is organised as follows: Section 1. From the published works, Section 2 analyses model performance on comparable tasks. The third section gives an overview of a study's methodology and research strategy. A study's findings and an in-depth examination of those findings are presented in Section 4. At last, Section 5 gives a summary of a study's results and discusses how they should inform practise and future studies.

2 Literature survey

Most of the empirical research on e-commerce found in scientific literature focuses on verifying theoretical ideas and testing hypotheses. ML is an important tool in this field that has been used a lot for a wide range of tasks, such as studying customer behaviour, making recommender systems, and building predictive models. A key part of making good suggestions is analysing how people behave, and researchers from various areas have come up with different ways to solve the issue from the point of view of how people behave [7]. Machine learning models have proven to be effective at anticipating customers' intentions to buy and identifying other products that customers might find interesting.

Sabina-Cristiana Necula et al. (2013) explore how online shoppers' actions are affected by the amount of time they spend reading product descriptions. By studying clickstream data using ML algorithms, we present a method for examining non-linear correlations in datasets and provide unique insights into the fundamental architecture of customer clusters. According to our results, a customer's purchasing decision is heavily impacted by the time they spend reading product-related content in addition to other aspects like customer type, bounce rates, and exit rates. This research fills a gap in the literature of e-commerce studies and suggests ways to improve the layout and promotion of online stores[11].

Ayush Maurya et al. (2023), research analyses the effect of numerous factors that allow E-commerce platforms to obtain previous knowledge about customer purchase tendencies. To advance to predictive analytics, ML algorithm like naive Bayes classification, SVC, logistical regression, DT, KNN classification, and random forest classification are applied to the dataset to anticipate product sales. This proposed research reduces e-waste and promotes sustainable development by anticipating product sales[17].

Wasim Yasin et al. (2023), The study discovers and examines customer purchasing patterns by employing a number of machine learning methods, and it discusses the difficulties and issues inherent to e-commerce applications. The results of the implementation show that the KNN algorithm outperformed the competition when compared to other machine learning methods. The effectiveness of these strategies has been analysed using several different matrices. The model is evaluated by using a data set derived from an online retailer (specifically, the Amazon dataset made available through Kaggle.com). As the study's findings demonstrate, the KNN approach is superior to other ML algorithms like Naive Bayes, Random Forest, Logistic Regression, etc., when it comes to computational and predictive power[18].

Ishrat Jahan et al. (2022), In order to enhance a accuracy of churn prediction and simplify an identification of non-churn consumers, a customer churn forecasting framework was designed employing the best classifier for insight and suggestion. The proposed strategy has been shown to accurately predict customer churn in experiments. Models are rated based on their accuracy and F1-score. The outcomes of an experimental evaluation showed that CatBoost achieved a highest levels of accuracy and F1-score in the Dataset[19].

Salma AbdulazizAlquhtani et al. (2022), seeks to propose a model for analysing user sentiment gleaned from online reviews on the e-commerce platform utilising ML classifiers such as NB, LR, SVM, and NN. Latent semantic analysis was utilised to analyse the most popular terms in user evaluations. Finally, an examination of customers' enthusiasm for online buying has been conducted to categorise feedback left on an e-commerce site. In addition, we compared the outcome of several classifiers' performances on the e-commerce dataset. According to a findings, Logistic regression is the most effective of the four classifiers tested[20].

Palvi Sharma et al. (2021), SVM, LR, and ANN are all examples of classification algorithms used to sort customer feedback into positive and negative categories. This research uses data from the Kaggle dataset, which consists of 23486 customer evaluations, to better understand the e-commerce market for women's clothes. A purpose of this research is to comprehend how consumers see both traditional and internet buying. This study developed an analysis using the ANN, LR, and SVM machine learning algorithms, which were simulated using Python. Calculating the accuracy of ANN, LR, and SVM reveals that ANN is more accurate (i.e., 88%) than SVM (i.e., 80%) and LR (i.e., 75%)[21].

Xi Niu et al. (2017), examines how customers search online at Walmart.com, a sizable e-commerce platform. We use a contemporary machine-learning technique called random forest in conjunction with logistic regression to create two computational models that will help us estimate consumer purchase conversion more precisely based on their search behaviour. In order to provide metrics to measure online users' search activity, we also incorporate literature on information retrieval. With a very high accuracy rate of 76% for forecasting which customers will purchase the item they viewed, a RF model performs better than another models[22].

In contrast to previous research, our strategy takes into account a wider range of factors gleaned from clickstream data rather than just the pages directly influencing a customer's choice to stay on the site or exit it. This all-encompassing strategy enables us to record the nuanced behaviours of internet shoppers that conventional approaches frequently miss. Unlike other approaches that depend on basic metrics like page views or click-through rates, our approach uses ML to examine complex patterns in the dataset of online shoppers' purchasing intentions. This provides a more precise and thorough understanding of customer behaviour and helps develop strategies for retaining and converting customers.

3 Research Methodology

This section offers graphic representations of the data and a detailed discussion of the used dataset. This section also covers every step of the suggested model, including every hyperparameter.

3.1 Problem Statement

The research project addresses a number of interconnected difficulties related to the dataset, preprocessing, balance, feature selection, and classic machine learning models. First, the dataset has a significant class imbalance, making it difficult to effectively model and anticipate minority class occurrences, which are shopping sessions. In addition, data pretreatment requires thorough attention to concerns such as missing data, categorical variable encoding, numerical attribute scaling, and outlier management. To prevent dimensionality difficulties, the difficulty of feature selection emerges, mandating the identification of the most relevant features for predictive modelling. Finally, the use of classical machine learning models offers a trade-off involving model complexity and interpretability, which needs careful consideration in order to fit with the study aims.

The research is to look at how customers' spending time reading product information on e-commerce websites effects their purchase behaviour. The study tries to explore if the length of interaction with product information, such as product descriptions, reviews, and specs, corresponds with the possibility of completing a purchase by utilising machine learning methods and analysing a dataset of user sessions. Ultimately, the objective is to find trends and insights that might assist e-commerce organisations in optimising their online buying experiences and perhaps increasing conversion rates via content and design that is tailored to user behaviour.

3.2 Proposed Methodology

In e-commerce, it is no longer wise to focus on acquiring new customers because keeping current ones is significantly less expensive. The high expenses of recruiting new consumers and the fierce competitiveness in the business-to-consumer (B2C) e-commerce industry have organizations putting a greater emphasis on minimizing their client turnover rate. E-commerce companies usually keep a lot of information about their current customers, like what they look for, how often they buy things, and so on. AI can be utilizied to look at how customers behave and guess when they might stop buying, which lets businesses use focused marketing to keep those customers. An aim of this study is to increase the accuracy of the suggested system by developing a customer behaviour analysis within an e-commerce framework utilising the best machine learning classifiers (XGBoost &CatBoost's Performing Stacked model) and grid search prediction for the Stacked Model Classifier. This research made use of the Online Shoppers Purchasing Intention Dataset, which is housed in the UCI ML repository. The framework consists of seven parts: data collecting, data splitting, data balancing, data preprocessing, feature engineering, exploratory data analysis (EDA), classification, and comparison between several models following model tuning. According to experimental findings, the suggested strategy has excellent recall, accuracy, precision, f1-score, and ROC curve predictive capabilities for customer behaviour.

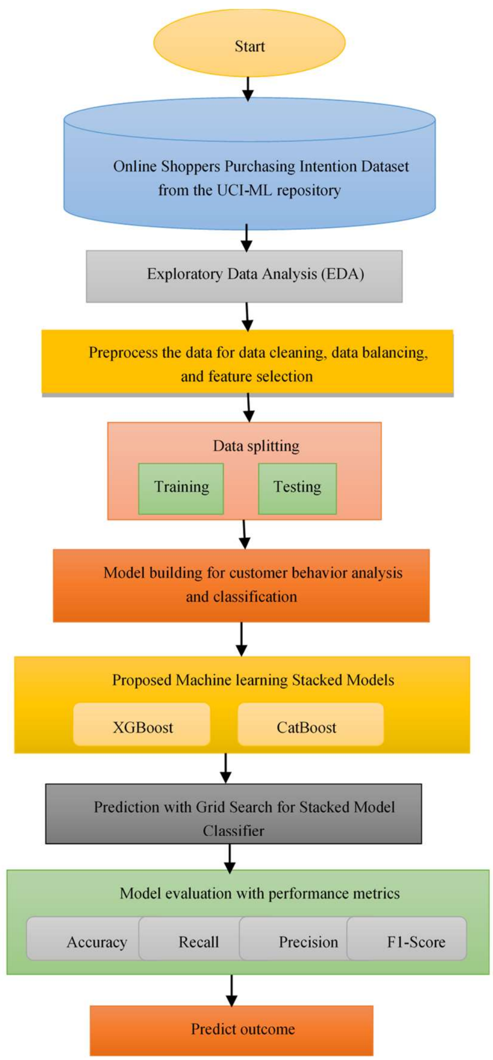

Fig. 1. Flow Chart for Proposed Methodology

The flowchart of the suggested methodology for the machine learning-based customer behaviour analysis of e-commerce websites is displayed in Figure 1 above. The whole methodology and flowchart process described below very briefly:

Data Collection

A crucial component of ML is datasets. One of a most difficult tasks is gathering data, particularly when it comes to e-commerce companies. The UCI ML library's Online Shoppers Purchasing Intention Dataset was utilised in this study. In our dataset, we have 12330 rows and 18 columns in which we have one targeted column called ‘Revenue’. This ‘Revenue’ in this dataset is dependent value, and all other are independent features.

Visualizing the Dataset

Open the dataset files now to see a different feature that contain the data, such as the product name, amount, review, purchase time, visits, add to cart, and so on.

Data preprocessing

Inaccurate specifications, duplicates, missing values, and uneven distribution are just some of the problems that might arise when working with raw feature data generated by extracting features from a dataset. Data preprocessing is a crucial step that underpins ML effectiveness, as ML on such raw feature data may be erroneous. Removing noise, enhancing data quality, and getting rid of unnecessary information all require data cleaning. This procedure involves identifying, fixing, or removing data mistakes. Our dataset underwent the following data-cleaning procedures:

- Checking if a null value is present in the data or not. Using isnull (). Sum ().
- Getting information of our dataset by using info (), names of all the columns, non-null count, and the datatype of each column.
- Getting statistics for the numeric columns in the data frame. Using. Describe ().
- Using Label Encoder to convert objects and Boolean values to integers to deal with the dataset. Using. LabelEncoder().

Feature importance’s (Extra Trees)

The term "feature importance" refers to a group of methods used to rank the significance of individual features used to train a predictive model. It reduces the amount of features required for input[23]. In this work, we extend the decision tree with a tree classifier to account for the significance of individual features. While both RF and Extra Trees construct multiple trees and divide nodes according to random subsets of attributes, Extra Trees uses random divisions to do what RF achieves with sample swapping.Data Balancing with SMOTE

Most practical applications have an asymmetrical class distribution, in which some labels appear far more frequently than others. A good illustration of the issue of class imbalance is when a customer reads product details on e-commerce websites, where there are significantly less class labels than there usually are. Most ML algorithms struggle to perform well in situations when the number of classes is not uniform (i.e., the predictive model tends to incorrectly label the minority case as the majority example).

The majority of the prediction model performs the worst when there is an uneven distribution of classes. Most machine learning methods fail to perform well in situations when the distribution of classes is not uniform (i.e., the predictive model tends to mistake the minority case for the majority example). The SMOTE was employed in this study to examine how consumers behaved when reading product details on e-commerce websites.

SMOTE [24] is a technique for statistical sampling that steadily increases a dataset's sample size. On the basis of edge cases, the module creates new instances. The proportion of just minority cases rises when the complete dataset is used as input. As predicted, SMOTE had no effect on the vast majority of cases.

It's clear that the new cases are distinct from the minority of existing ones. The algorithm instead takes a sample of the feature space for each target class and their close neighbours. Then, using characteristics from the target instance and its neighbours, the algorithm generates new cases. This approach makes more attributes available to every class and results in more inclusive sampling[25].

Data Splitting

Separating the preprocessed dataset into a training and testing set is a last stage in a data preparation process. A total of 80% of a data was utilized for training purposes, while 20% was employed as a test set for each experiment. After a machine has been trained using one dataset, it is put through its paces using another dataset for testing.

Classification Models

This study examines the effects of spending time perusing product details on e-commerce websites in order to analyse consumer behaviour using stacked machine learning models, specifically XGBoost &CatBoost. To particularly address the needs of our study and the characteristics of our data, we altered these approaches rather than applying them as is. To find a best parameters for every ML model, we used GridSearch for parameter tuning. This is a radical departure from the customary method of direct application. We also had to utilise the StandardScaler to normalise the numerical components because our models used gradient solvers, which are somewhat stochastic.

Stacked Models (XGBoost and CatBoost)

Stacking is the method of fitting several categorical projects to identical data and then utilizing the other project for learning to achieve integration of the forecasting in the most suitable way. Stacking is yet another ensemble methodology, which is mostly known as “stacked generalization”. This methodology operates by permitting the training method to ensemble the forecasting made by numerous other identical learning processes. This Stacking ensemble methodology is found to be applied in density computations, regression, classifications, and distance learning. This could also be utilized to compute the rate of errors that take place at the time of bagging. For the customer behavior analysis used Stacking with XGBoost and CatBoost ML method. A following section provides a proposed model overview briefly:

Extreme Gradient Boosting (XGBoost)

In 2001, Friedman proposed the gradient-boosting decision tree (GBDT). It is a version known as extreme gradient boosting (XGBoost)[26]. The XGBoost package includes a simple linear model solver as well as a tree-learning technique. It simplifies a wide variety of tasks using objects, including regression, ranking, & classification. XGBoost makes use of the objective function formula. To make the model's predictions more accurate, this objective function limits the difference between the predictions and the real aim. XGBoost's goal function equation is[27]:

$$L(t) = \sum_{i=1}^{n} l\left(y_i, \hat{y}_i^{(t-1)} + f_t(x_i)\right) + \Omega\,(f_t) \dots \dots (1)$$

where L measures the dissimilarity between a forecast yˆi and the actual target yi, and Ω keeps the model from being too specialised. An increasingly popular machine learning algorithm, XGBoost is developed employing multiple cores [28]and has a number of hyperparameters that can be optimized to enhancea model's performance and speed by reducing the likelihood of overfitting, increasing generalization performance, and shortening the computation time[29].

CatBoost or "Categorical Boosting"

Yandex created the gradient-boosting decision tree (GDBT) model known as CatBoost. In machine learning, CatBoost, short for "Categorical Boosting," is a technique derived from the more well-known gradient boosting. CatBoost is an adaptation of the ordering principle, a basic gradient boosting technique that prevents target leaking, and a new algorithm for processing category information.CatBoost is frequently employed in real-world applications on datasets that contain both category and numeric variables[30]. When compared to classic GBDT, it has two major algorithmic improvements:

- Rather of using the traditional method, it employs an ordered boosting technique driven by permutations.
- It uses an original algorithm for analysing categorical features.

These upgrades were made to combat a particular kind of target leakage that could have resulted in incorrect predictions in earlier GBDT implementations[31][32].

Because CatBoost is a sophisticated ML method, it cannot be expressed using a single formula. This approach incorporates multiple techniques, including decision trees, gradient boosting, and handling of categorical features. Equation (2) below illustrates how the algorithm iteratively constructs tiny trees using gradient boosting techniques to minimise the anticipated loss and increase the accuracy of the model[27]:

$$h^t = agrminE\left(\frac{\delta Ly}{\delta F^{t-1}} - h\right)^2 \approx argmin\frac{1}{n}\left(\frac{\delta Ly}{\delta F^{t-1}} - h\right)^2 h \in H \dots \dots \dots \dots (2)$$

It is also meant to deal with categorical characteristics better than conventional gradient boosting methods by employing improved target-based statistics that aid in lowering the computational cost of processing such features[33]. CatBoost processes features that fit into one or more of several specified categories using binning, target statistics encoding, and one-hot encoding. This opens the way for the algorithm to effectively analyse category features and enhance prediction accuracy. The equation for estimating the ith categorical variable using the k-th element is presented below[27]:

$$\hat{x}_k^i = \frac{\sum_{x_j \in D_k} 1_{\left\{x_j^i = x_k^i\right\}} . y_j + ap}{\sum_{x_j \in D_k} 1_{\left\{x_j^i = x_k^i\right\}} . + a} \dots \dots \dots (3)$$

where a must be greater than zero and p (prior) is often set to the average of the targets in the dataset D used for training.

Prediction with Grid Search for Stacked Model Classifier

To identify the best parameters for each ML model, we used GridSearchCV instead of the more commonplace direct application to obtain the ideal parameters. Because the GridSearchCV approach we used to adjust the parameters requires repeatedly training and testing the models for every set of parameters, this adds even more complexity to the equation. Despite this, considering the scale of our dataset, our solution is still computationally feasible. We didn't use the current machine learning techniques directly. Instead, we used preprocessing and hyperparameter modification to tailor them to our data and the goals of our investigation. This made it possible for us to develop models that, using clickstream data from e-commerce, precisely forecast what customers will decide to buy.

4 Result Analysis and Discussion

The experimental results of proposed machine learning models for customer behaviour analysis in e-commerce websites provided in this section. The study is implemented using Python, Jupyter Notebook and Google Colab software technologies that work on Intel Core i7 CPU and an NVIDIA graphics processing unit (GPU), Windows 10 or 11, 32GB of RAM, & 512GB of external space used to carry out the operation. Also used some Python libraries, namely NumPy, pandas, matplotlib, seaborn, and SK-learn, etc., for the visualization of proposed model implementation results and dataset analysis. In this work used Online Shoppers Purchasing Intention Dataset from the UCI ML repository that split into train and tests for the machine learning implementation for the customer behaviour analysis in e-commerce websites. Evaluated the model performance according to ROC curve, f1-score, recall, accuracy, and precision. Following section gives the dataset description, EDA, performance metrix, model simulation results and comparative analysis and discussion.

4.1 Data Description

The research used UCI's machine learning library to access the Online Shoppers Purchasing Intention Dataset. A dataset consists of 10 numerical and 8 category features. Class labels can be derived from the 'Revenue' field. A total of 12,330 sessions' worth of feature vectors make up the dataset. The goal of creating this dataset was to eliminate any potential bias related to a particular campaign, day, user profile, or time frame. Each session was the property of a distinct user for a period of one year. Out of the 12,330 sessions in the dataset, 10,422 (84.5%) concluded without any shopping, while 1908 (the remaining class sample) finished with some purchasing.

```
df.describe()
```

	Administrative	Administrative_Duration	Informational	Informational_Duration	ProductRelated	ProductRelated_Duration	BounceRates	ExitRates	PageValues
count	12330.000000	12330.000000	12330.000000	12330.000000	12330.000000	12330.000000	12330.000000	12330.000000	12330.000000
mean	2.315166	80.818611	0.503569	34.472398	31.731468	1194.746220	0.022191	0.043073	5.889258
std	3.321784	176.779107	1.270156	140.749294	44.475503	1913.669288	0.048488	0.048597	18.568437
min	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000
25%	0.000000	0.000000	0.000000	0.000000	7.000000	184.137500	0.000000	0.014286	0.000000
50%	1.000000	7.500000	0.000000	0.000000	18.000000	598.936905	0.003112	0.025156	0.000000
75%	4.000000	93.256250	0.000000	0.000000	38.000000	1464.157214	0.016813	0.050000	0.000000
max	27.000000	3398.750000	24.000000	2549.375000	705.000000	63973.522230	0.200000	0.200000	361.763742

Fig. 2. Statistical distribution of Online Shoppers Purchasing Intention Dataset

The statistical distribution of the dataset describing the purchase intentions of online shoppers is shown in Figure 2. 12330 rows and 18 columns make up the dataset. the x-axis represents the features such as Administrative, AdministrativeDuration, Informational, Informational_Duration, Product_Relateded, Product_Relateded_Duration, BounceRates, ExitRates, and PageValues, and the y-axis represents a statistical value, like count, mean, std (standard deviation), min, max value and 25%, 50%, 75% quartile, respectively.

4.2 Exploratory data analysis

One way to explore and analyse data is through the use of exploratory data analysis. Utilising these methods, data scientists can more effectively detect outliers, uncover trends, validate assumptions, and get useful insights[34]. To investigate the connections between the many variables, we first generate a correlation matrix. Attribute elimination should be considered if there is a significant correlation between them. Multiple data visualisation methods are employed to examine a distribution of attributes and a data balance after gain a deeper comprehension of a dataset and its characteristics[25].

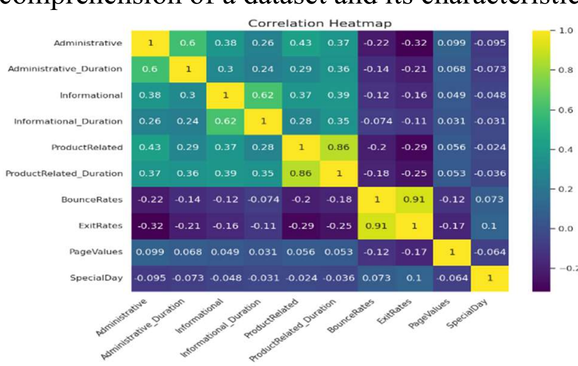

Fig. 3. Correlation Heatmap Matrix for 0to 10 Features

Figure 3 describes the correlation heatmap for the starting 10 columns. The x-axis represents the features, such as Administrative, AdministrativeDuration, Informational, Informational_Duration, Product_Relateded, Product_Relateded_Duration, BounceRates, ExitRates, and PageValues, and SpecialDay, and the y-axis also represents these features.A heatmap depicting these associations is shown in Figure 3. Highly associated variables include Bounce Rate, Exit Rate, ProductRelated, and ProductRelated_Duration.

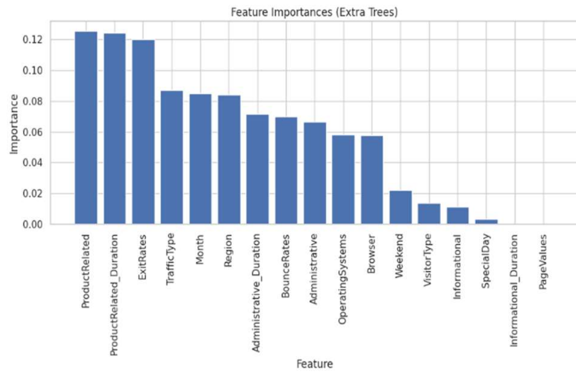

Fig. 4. Feature importance of the extra tree classifier

Figure 4 shows how the extra tree classifier takes feature significance into account. The duration of time utilised reading ProductRelated data and the amount of time spent doing administrative tasks on the website's pages show up to have a substantial influence on the outcome, in addition to PageValues, ExitRates, whether the customer is a return visitor and the month. This is the case despite the fact that PageValues and ExitRates are still important factors.

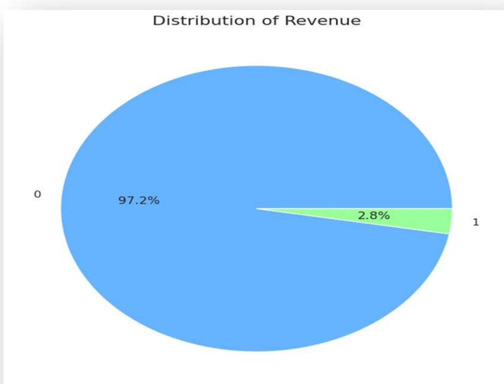

Fig. 5. Pie chart of Data distribution of Revenue of input dataset

A following figure 5 displays a Data distribution of Revenue of input dataset, this dataset contains two classes, 0 and 1, from the online shopping dataset. The 0 class contain 97.2% dataset and 1 class contain only 2.8% dataset area.

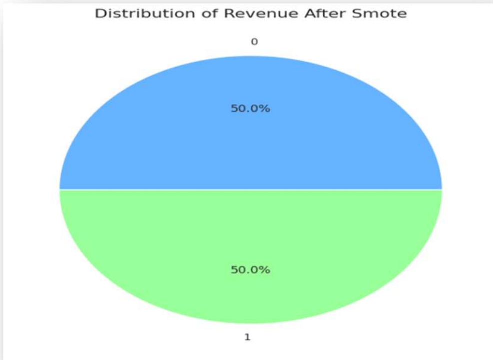

Fig. 6. Pie chart of Data Distribution of Revenue After SMOTE

The following Figure 6 shows the Data distribution of Revenue of input dataset.After SMOTE data balancing technique, this dataset contains two classes, 0 and 1, from the online shopping dataset. The 0 and 1 class data balanced with 50-50% dataset used SMOTE.

4.3 Evaluation Parameters

Evaluation metrics provide a means of explaining a model's performance. One important aspect of evaluation measures is their capacity to distinguish between different model outcomes. Confusion Matrix N is the number of classes being predicted in this N X N matrix[35]. Table 1's confusion matrix will be applied to this article.

Table 1. Representation of cells in confusion matrix

	Predicted:0	Predicted:1
Actual:0	TN	FP
Actual:1	FN	TP

1) **Accuracy Score**

Accuracy is defined as a proportion of accurate forecasts to all predictions. The accuracy was determined using equation (4).

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \dots . (4)$$

2) **Recall Score**

Recall is a percentage of TP cases that are correctly identified. A recall was computed using equation (5).

$$R = \frac{TP}{TP + TN} \dots . (5)$$

3) **Precision Score**

Precision is the percentage of affirmative cases that are correctly detected. The precision score indicates how flawless the forecasting model is. Equation (6) was utilised to calculate precision.

$$Precision = \frac{TP}{TP + FP} \dots . (6)$$

4) **F1 Score**

The F1-Score is the precision and recall values of the harmonic mean of a classification issue. The following equation (7) was used to determine F1.

$$F1 - score = 2 \times \frac{precision * recall}{precision + recall} \dots . (7)$$

5) **ROC curve (receiver operating characteristic curve)**

A ROC curve is a graph that illustrates how well a classification model performs over all of the different levels of categorization. This curve shows two variables. TPR.

4.4 Experimental Results of Stacked Model

Stacked Model (XGBoost and CatBoost) for analysing customer behaviours on e-commerce websites will be evaluated using the recall score, f1 score, accuracy score, precision score, and confusion matrix metrics in this study.

The following Table 2 shows the proposed stacked model train and test performance in terms of evaluation metrix. Training model obtained 100% precision, recall, f1-score, accuracy, and AUC-ROC score,while proposed stacked model achieved 98% and 99% parameter performance in test dataset for the customer behavior analysis in e-commerce website.

Table 2. Proposed stacked model train and test performance in terms of evaluation metrix

Measures		**Train**
Accuracy	98	100
Precision	99	100
Recall	98	100
F1-score	98	100

```
Classification Report of test data:
              precision    recall  f1-score   support

           0       0.98      0.99      0.98      1792
           1       0.99      0.98      0.98      1851

    accuracy                           0.98      3643
   macro avg       0.98      0.98      0.98      3643
weighted avg       0.98      0.98      0.98      3643
```

Fig. 7. Testing classification report of proposed stacked model

Figure 6 displays the outcome of testing a suggested stacking model's classification accuracy on a dataset of online purchases classified as either 0 or 1. The class 0 and 1 model obtain 98% and 98% parameters performance with support 1792 and 1851, overall performance of proposed model 98% with support 3643, respectively.

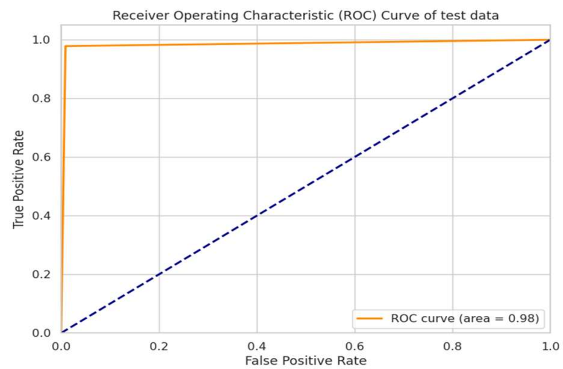

Fig. 8. Testing ROC curve of proposed stacked model

The proposed stacked model implemented ROC curve of testing dataset is shown in Figure 7. Model obtained 98% ROC curve area of Stacked model for CBA in online shopping classification.

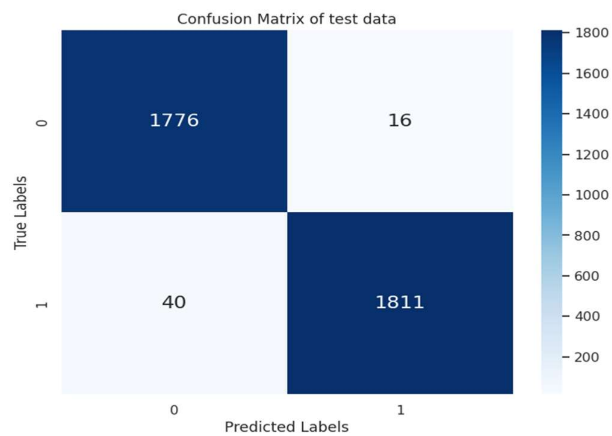

Fig. 9. Testing confusion matrix of proposed stacked model

The proposed stacking model's confusion matrix for the Testing online shoppers' purchase intention dataset, which comprises two classes (0 and 1). The matrix's x-axis displays predicted labels, while a y-axis represents the actual labels from the dataset. Proposed stacked model predicted true predication instances of 3,587 and false prediction instances of 56, respectively.

Classification Report of train data:

	precision	recall	f1-score	support
0	1.00	1.00	1.00	5494
1	1.00	1.00	1.00	5435
accuracy			1.00	10929
macro avg	1.00	1.00	1.00	10929
weighted avg	1.00	1.00	1.00	10929

Fig. 10. Training classification report of proposed stacked model

Figure 10 is report generated by the proposed stacked model during training, with the dataset consisting of items labelled as either 0 or 1. From the class 0 and 1 model, 100% parameters performance with support 5494 and 5435, overall performance of proposed model 10% with support 10929, respectively.

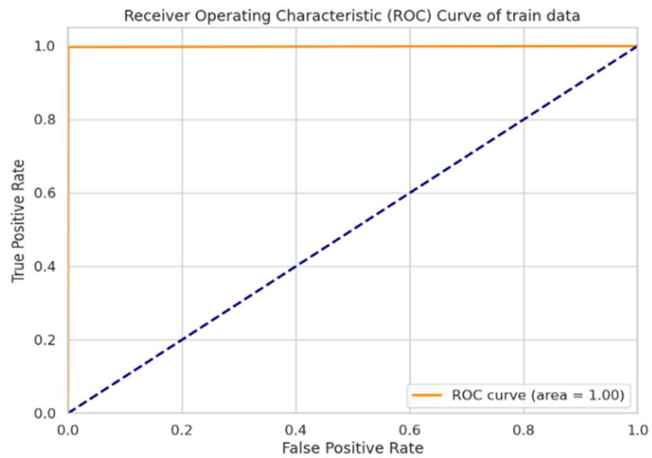

Fig. 11. Training ROC curve of proposed stacked model

The proposed stacked model implemented ROC curve of Training dataset shows in Figure 11. Model obtains 100% ROC curve area of Stacked model for CBA in online shopping classification.

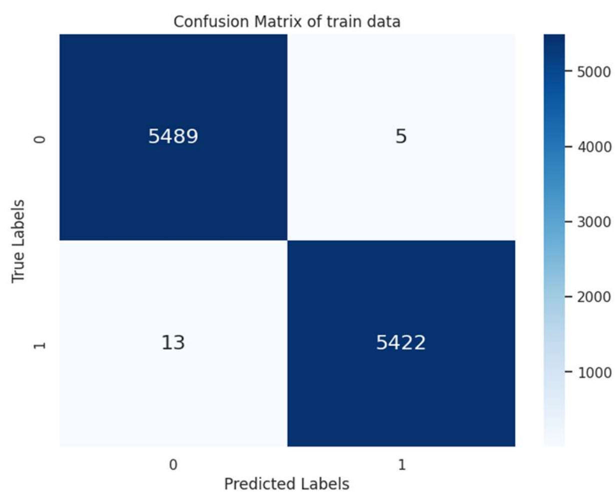

Fig. 12. Training Confusion matrix of proposed stacked model

A confusion matrix of proposed stacked model for the training online shoppers purchasing intention dataset that contains 0 and 1 classes is shown in Figure 12. Proposed stacked model predicted true predication instances of 10,911 and false prediction instances of 10, respectively.

Prediction with Grid Search for Stacked Model Classifier

Evaluated Models for Grid Search: -Accuracies, ROC, Report, Classification report, f1-score, recall, Confusion matrix. Table 3 shows the Proposed stacked model with grid search performance with 98% and 99% in terms of evaluation metrix.

Table 3. Proposed stacked model with grid search performance in terms of evaluation metrix

Measures	
Accuracy	98
Precision	99
Recall	98
F1-score	98

```
Classification Report of GRID SEARCH test data:
              precision    recall  f1-score   support

           0       0.98      0.99      0.98      1792
           1       0.99      0.98      0.98      1851

    accuracy                           0.98      3643
   macro avg       0.98      0.98      0.98      3643
weighted avg       0.98      0.98      0.98      3643
```

Fig. 13. Testing classification report of proposed stacked model with Grid Search

Figure 13 shows the testing classification report of proposed stacked model stacked model with Grid Search on online shopping dataset that has two classes, 0 and 1. The class 0 and 1 model obtain 98% and 98% parameters performance with support 1792 and 1851, overall performance of proposed model 98% with support 3643, respectively.

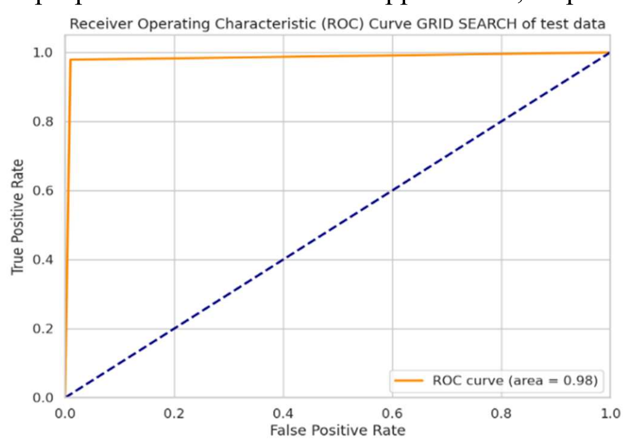

Fig. 14. Testing ROC curve of proposed stacked model with Grid Search

The proposed stacked model with Grid Search implemented ROC curve of testing dataset shows in Figure 14. Model obtained 98% ROC curve area of Stacked model for CBA in online shopping classification.

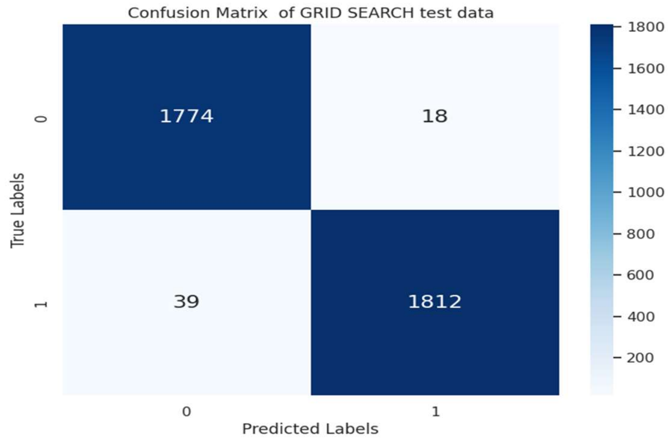

Fig. 15. Testing confusion matrix of proposed stacked model with Grid Search

A confusion matrix of proposed stacked model with Grid Search shows in figure 15. Proposed stacked model predicted true predication instances of 3,586 and false prediction instances of 55, respectively.

Comparison between base and propose Models and discussion

The following table 4 displays a comparison among base (Random Forest and Decision Tree) & proposed model (Stacked model) with grid search for the e-commerce websites-based customer behaviors analysis. Compare the proposed and base models' performance according to precision, recall, f1-score, accuracy, and AUC-ROC measure with Online+Shoppers+Purchasing+Intention+Dataset.

Table 4. Base and proposed models' comparison

Performance Measures	**Proposed Stacked Model**	**Base Random Forest**	**Base Decision Tree**
Accuracy	98	97	92
Precision	99	96	88
Recall	98	99	97
F1-score	98	97	93
AUC-ROC	98	98	92

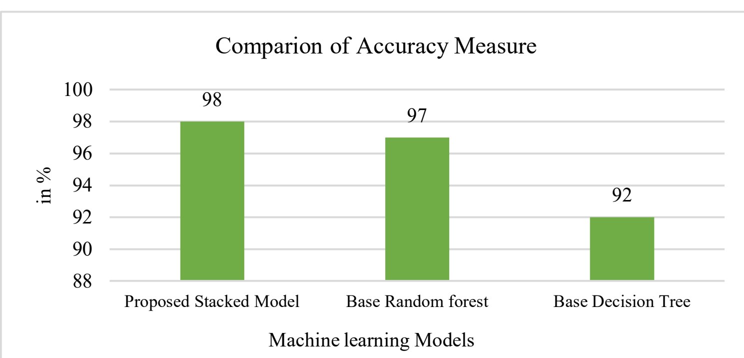

Fig. 16. Accuracy comparison between base and proposed machine learning models with grid search for CBA

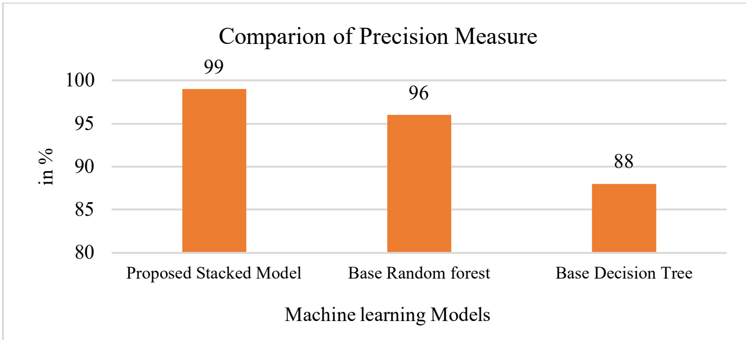

Fig. 17. Precision comparison between base and proposed machine learning models with grid search for CBA

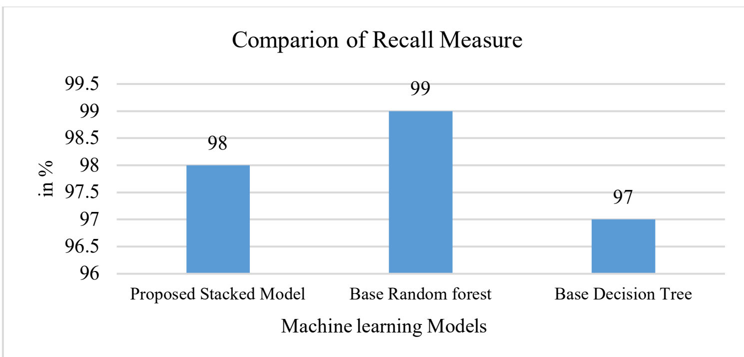

Fig. 18. Recall comparison between base and proposed machine learning models with grid search for CBA

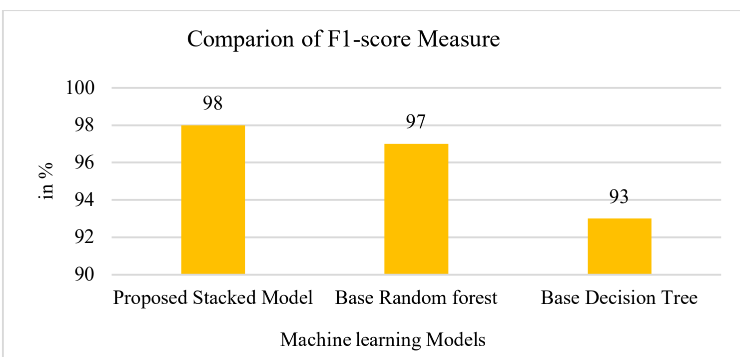

Fig. 19. F1_Score comparison between base and proposed machine learning models with grid search for CBA

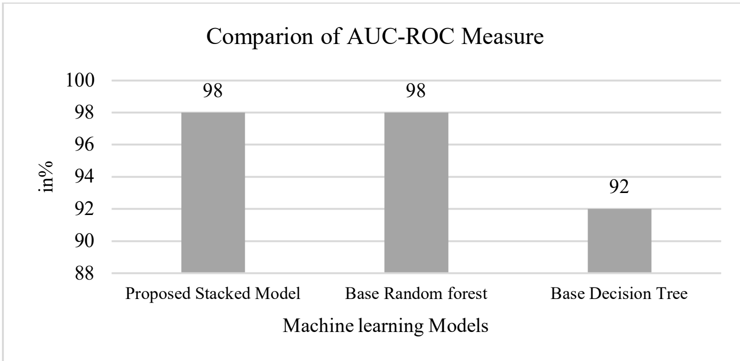

Fig. 20. F1_Score comparison between base and proposed machine learning models with grid search for CBA

The performance measurements of the suggested stacked model show considerable gains when compared to the Existing RF model and an Existing Decision Tree model, with the use of grid search for Consumer Behaviour Analysis (CBA). The stacked model that was presented surpasses both of the basic models with regard to all criteria. It reaches an impressive accuracy of 98%, demonstrating its capacity to accurately forecast user behaviour. In addition to this, its precision is very high (99%), which indicates that it is able to reduce the number of false positives it produces. This is an extremely important quality in the context of e-commerce. In addition, the stacked model performs very well in terms of recall (with a score of 98%) and F1-score (98%), demonstrating its capability of identifying true positives and offering a satisfactory compromise between accuracy and recall. The fact that the stacked model has a high discriminating power is further shown by the fact that its AUC-ROC score is 98%. Based on these findings, it seems that the suggested stacked model, which was optimised using grid search, provides a solid and comprehensive solution for Consumer Behaviour Analysis, giving improved predictive performance in comparison to the separate base models.

5 Conclusion and Future Work

Customer purchasing behaviour research is a key part of coming up with smart ways for online stores and their customers to talk to each other. Variety of consumers purchasing behaviour helps the owners of e-commerce sites to improve business and serving their consumers properly. The main objective of this research work is to assist the owners of E-commerce websites. Consumers purchasing intentions help the owners to have precise knowledge and make proper schemes about their products using consumer's sentiment. Here, consumer's behavior is visualized in various aspects, and different machine learning algorithms are used to identify their shopping demands. In this work, we describe a customer purchasing behaviour analysis system that employs supervised machine learning approaches stacked with XGBoost and CatBoost to support the recent uptick in the popularity of online shopping. We found that Stacked model provides better prediction than any other algorithms, with accuracy 98%. The experimental data demonstrated improved performance over competing systems. Such factors identified by the owners of e-commerce organizations will help them take necessary steps to provide a better service to the consumers and add more lifetime value to their business. In the future, after a purchase has been made, customers may be asked to fill out a survey with a series of specific questions designed to improve the site's functionality and provide valuable data for the recommendation system.

References

[1] H. I. Lee, I. Y. Choi, H. S. Moon, and J. K. Kim, "A multi-period product recommender system in online food market based on recurrent neural networks," *Sustain.*, 2020, doi: 10.3390/su12030969.

[2] A. K and D. S. K.G, "Crop Yield Prediction, Forecasting and Fertilizer Recommendation using Voting Based Ensemble Classifier," *Int. J. Comput. Sci. Eng.*, 2020, doi: 10.14445/23488387/ijcse-v7i5p101.

[3] A. Adadi, "A survey on data-efficient algorithms in big data era," *J. Big Data*, 2021, doi: 10.1186/s40537-021-00419-9.

[4] E. Elahi and A. Chandrashekar, "Learning Representations of Hierarchical Slates in Collaborative Filtering," 2020. doi: 10.1145/3383313.3418484.

[5] A. R. Abbas, "An Adaptive E-Commerce: Applying of Psychological Testing Method to Improve Buying Decision Process," *Eng. Technol. J.*, 2015, doi: 10.30684/etj.33.2b.7.

[6] K. Bortko, P. Bartków, J. Jankowski, D. Kuras, and P. Sulikowski, "Multi-criteria evaluation of recommending interfaces towards habituation reduction and limited negative impact on user experience," 2019. doi: 10.1016/j.procs.2019.09.399.

[7] M. Y. Gheraibia and C. Gouin-Vallerand, "Intelligent mobile-based recommender system framework for smart freight transport," 2019. doi: 10.1145/3342428.3342697.

[8] J. Jankowski, J. Hamari, and J. Wątróbski, "A gradual approach for maximising user conversion without compromising experience with high visual intensity website elements," *Internet Res.*, 2019, doi: 10.1108/IntR-09-2016-0271.

[9] F. T. A. Hussien, A. M. S. Rahma, and H. B. Abdulwahab, "An e-commerce recommendation system based on dynamic analysis of customer behavior," *Sustain.*, 2021, doi: 10.3390/su131910786.

[10] P. Sharma, S. Kumar, and L. Kumar, "Customer Behavior Analysis in E-Commerce using Machine Learning Approach : A Survey," *For. Chem. ...*, no. April, 2022, doi: 10.32628/CSEIT239028.

[11] S. C. Necula, "Exploring the Impact of Time Spent Reading Product Information on E-Commerce Websites: A Machine Learning Approach to Analyze Consumer Behavior," *Behav. Sci. (Basel).*, 2023, doi: 10.3390/bs13060439.

[12] V. Shrirame, J. Sabade, H. Soneta, and M. Vijayalakshmi, "Consumer Behavior Analytics using Machine Learning Algorithms," 2020. doi: 10.1109/CONECCT50063.2020.9198562.

[13] E. Manohar, P. Jenifer, M. S. Nisha, and B. Benita, "A collective data mining approach to predict customer behaviour," 2021. doi: 10.1109/ICICV50876.2021.9388558.

[14] A. AlShamsi, "Understanding Customer Behaviour in Restaurants based on Data Mining Prediction Technique," 2022.

[15] T. Li, Y. Li, M. A. Hoque, T. Xia, S. Tarkoma, and P. Hui, "To What Extent We Repeat

Ourselves? Discovering Daily Activity Patterns Across Mobile App Usage," *IEEE Trans. Mob. Comput.*, 2022, doi: 10.1109/TMC.2020.3021987.

[16] G. Chen, P. Chen, W. Huang, and J. Zhai, "Continuance Intention Mechanism of Middle School Student Users on Online Learning Platform Based on Qualitative Comparative Analysis Method," *Math. Probl. Eng.*, 2022, doi: 10.1155/2022/3215337.

[17] A. Maurya, S. Pratap, P. Pratap, and A. Dwivedi, "Analysis of Behavioural Data of Customer for the E-Commerce Platform by using Machine Learning Approach," 2023. doi: 10.1109/CISES58720.2023.10183475.

[18] W. Yasin, K. Harimoorthy, and D. Vetriveeran, "Identification of Consumer Buying Patterns using KNN in E-Commerce Applications," in *2023 3rd International Conference on Pervasive Computing and Social Networking (ICPCSN)*, 2023, pp. 187–192. doi: 10.1109/ICPCSN58827.2023.00036.

[19] I. Jahan and T. Farah Sanam, "An Improved Machine Learning Based Customer Churn Prediction for Insight and Recommendation in E-commerce," 2022. doi: 10.1109/ICCIT57492.2022.10054771.

[20] S. A. Alquhtani and A. Muniasamy, "Analytics in Support of E-Commerce Systems Using Machine Learning," 2022. doi: 10.1109/ICECET55527.2022.9872592.

[21] P. Sharma, R. Kumar, and M. Gupta, "Impacts of Customer Feedback for Online-Offline Shopping using Machine Learning," 2021. doi: 10.1109/ICOSEC51865.2021.9591939.

[22] X. Niu, C. Li, and X. Yu, "Predictive analytics of E-commerce search behavior for conversion," *AMCIS 2017 - Am. Conf. Inf. Syst. A Tradit. Innov.*, vol. 2017-August, pp. 1–10, 2017.

[23] M. Kanchana, V. Chadda, and H. Jain, "Credit card fraud detection," *Int. J. Adv. Sci. Technol.*, vol. 29, no. 6, pp. 2201–2215, 2020, doi: 10.48175/ijarsct-7776.

[24] N. V. Chawla, K. W. Bowyer, L. O. Hall, and W. P. Kegelmeyer, "SMOTE: Synthetic minority over-sampling technique," *J. Artif. Intell. Res.*, 2002, doi: 10.1613/jair.953.

[25] V. Chang, M. A. Ganatra, K. Hall, L. Golightly, and Q. A. Xu, "An assessment of machine learning models and algorithms for early prediction and diagnosis of diabetes using health indicators," *Healthc. Anal.*, vol. 2, p. 100118, 2022, doi: https://doi.org/10.1016/j.health.2022.100118.

[26] J. H. Friedman, "Greedy function approximation: A gradient boosting machine," *Ann. Stat.*, 2001, doi: 10.1214/aos/1013203451.

[27] G. Ryan, P. Katarina, and D. Suhartono, "MBTI Personality Prediction Using Machine Learning and SMOTE for Balancing Data Based on Statement Sentences," *Inf.*, 2023, doi: 10.3390/info14040217.

[28] R. Santhanam, N. Uzir, S. Raman, and S. Banerjee, "Experimenting XGBoost Algorithm for Prediction and Classification of Different Datasets," *Int. J. Control Theory Appl.*, 2017.

[29] T. Chen and C. Guestrin, "XGBoost: A scalable tree boosting system," 2016. doi: 10.1145/2939672.2939785.

[30] D. N. Gono, H. Napitupulu, and Firdaniza, "Silver Price Forecasting Using Extreme Gradient Boosting (XGBoost) Method," *Mathematics*, 2023, doi: 10.3390/math11183813.

[31] J. T. Hancock and T. M. Khoshgoftaar, "CatBoost for big data: an interdisciplinary review," *J. Big Data*, vol. 7, no. 1, 2020, doi: 10.1186/s40537-020-00369-8.

[32] L. Prokhorenkova, G. Gusev, A. Vorobev, A. V. Dorogush, and A. Gulin, "Catboost: Unbiased

boosting with categorical features," 2018.

[33] S. Hussain *et al.*, "A novel feature engineered-CatBoost-based supervised machine learning framework for electricity theft detection," *Energy Reports*, 2021, doi: 10.1016/j.egyr.2021.07.008.

[34] A. M. Malik, A. K. Sagar, and S. Sahana, "Prediction of Cardiopathy Using Exploratory Data Analysis," 2021. doi: 10.1109/ICCCA52192.2021.9666241.

[35] M. F. AKÇA and O. SEVLİ, "Predicting acceptance of the bank loan offers by using support vector machines," *Int. Adv. Res. Eng. J.*, 2022, doi: 10.35860/iarej.1058724.