

A PROPOSED MINING APPROACH BASED ON BUSINESS INTELLIGENCE BY USING DATA SCIENCES TECHNIQUES

AHMED I. EL SEDDAWY

IS Department,

Arab Academy for Science and Technology and
Maritime Transport -Smart Village, Cairo, Egypt.

Ahmed.Bahgat@aast.edu

MOHAMED HEGAZY MOHAMED

IS Department,

The French University of Egypt (UFE), Cairo, Egypt.

Mohamed.Hegazy@ufe.edu.eg

Received 2021- 1-13; Revised 2021-4-5; Accepted 2021-4-5

Abstract

Business Intelligence (BI) is a concept of applying a set of technologies to convert data into meaningful information. BI methods include information retrieval, data mining, statistical analysis as well as data visualization. Large amounts of data originating in different formats and from different sources can be consolidated and converted to key business knowledge. Data mining is used to search for patterns and correlations within a database of information. Business Intelligence is becoming an important factor that can help organizations in managing, developing and communicating their valuable goods such as information and knowledge. The main goal of data mining is the search for relationships and distinct patterns that exist in datasets but they are "hidden" among the vast amount of data. Data mining was effectively applied in many areas including: inventory control, fault diagnosis, credit scoring, marketing, financial services, scheduling, fraud detection and network management. Data mining can play a significant role for the exploitation of business intelligence in e-business. While techniques and tools for e-commerce are being actively developed, there are far fewer for e-business intelligence. This paper contributes to the area of web usage analysis for e-business intelligence by 'marrying' e-business with data mining, considering business intelligence as the best environment. Four metrics, applied innovatively for the first time in the field of e-business, provide capabilities for examining data about e-shopping in online stores and detailed insight into the effectiveness of different efforts by answering related business questions.. This paper attempts to propose prescient investigation to foresee Consumer Behavior conduct on Jumia Website in Egypt by utilizing conduct information, informatics and investigation approach by utilizing Data Sciences Techniques through grouping Classification Technique by choice tree Algorithm in particular Random Tree the results show that this the Random Tree Algorithm is reasonable for Jumia Datasets for anticipating

* Corresponding author: AHMED I. EL SEDDAWY

IS Department, Arab Academy for Science and Technology and Maritime Transport -Smart Village, Cairo, Egypt

E-mail address: Ahmed.Bahgat@aast.edu

consumer conduct with the goal that more profound knowledge into Consumer conduct can be gotten to help prescient examination so as to improve business choice.

Keywords- *Data Sciences; Classification; Consumer Behavior; Data Mining; Business Intelligent.*

1. Introduction

Buyer direct is the assessment of people, parties, or affiliations and the philosophy they use to pick, secure, and discard things, associations, encounters, or contemplations to fulfill needs and the effects that these strategies have on the purchaser and society. Shopping has changed because of the impact of headway, with the nearness of web based shopping more choices have been given to the customers[5]. This model transformed into the overarching purpose of assembly in the previous decade with different retail brutes joining the two distinct approaches to oversee accomplish most imperative piece of slack. This examination means to get a handle on the connection between's online shopping& physical store shopping. A fundamental energy about customer lead in the virtual condition, as in the physical world, can't be practiced if the fragments influencing the buy choice are overlooked or confused. For example, online customers' tendencies about nonattendance of chance to evaluate things going before buy are viewed as the particular factor influencing the purchasing choice. Web is changing the manner by which buyers shop and purchase things and experiences, and has promptly advanced into a general marvel. Different affiliations have begun utilizing the Web to cut propelling costs, as such lessening the cost of their things and associations so as to remain ahead in remarkably certifiable markets. This appraisal is required to improve our comprehension of on the web and physical buyer lead.

2. Related work

In [1] this paper, we presented a two-stage method for creating accurate classifiers for DNA sequences with interesting and comprehensible features. Data mining is the extraction of hidden predictive information from large databases and it is a powerful new technology with great potential to help companies focus on the most important information in their data warehouses. In paper [2], different feature selection techniques for classification: Big Data concerns large-volume, complex, growing data sets with multiple, autonomous sources. With the fast development of networking, data storage, and the data collection capacity, Big Data is now rapidly expanding in all science and engineering domains, including physical, biological and biomedical sciences. In paper [3], this paper presents an application of business intelligence (BI) for electricity management systems in the context of the Smart Grid domain. This distribution is the basis for performing on an upper level all the business processes for managing the energy demand and other customer services in a Smart Grid. We study here the problem of secure mining of association rules in horizontally partitioned databases. In that setting, there are several sites (or players) that hold homogeneous databases, databases that share the same schema but hold information on different entities. In paper [4] this paper proposed a data mining methodology called business intelligence driven data mining. It combines the knowledge driven data mining and method driven data mining and fills the gap between business intelligence knowledge and existence various data mining methods in e-Business. It setup a four layer frame layer. Outlier detection is the data mining task whose goal is to isolate the observations which are considerably dissimilar from the remaining data. This task has practical applications in several domains such as fraud detection, intrusion detection, data cleaning, medical diagnosis, and many others. IN paper [5] this paper presents an application of business intelligence for electricity management systems in the context of the Smart Grid domain. This distribution is the basis for performing on an upper level all the business processes for managing the energy demand and other customer services. IN paper[6] proposed a data mining methodology called business intelligence driven data mining. It combines the knowledge driven data mining and method driven data mining and fills the gap between business intelligence knowledge and existence various data mining methods in e-Business.

The current writing on customer on-line shopping determinations has especially focusing on making sense of the components which affect the eagerness of buyers to have association in Internet shopping. In the territory of customer conduct research, there are well-known styles of buying conduct that portray the framework which buyers use in making a purchase decision. These designs are extremely significant to business visionaries as they have the cap potential to give a clarification to and anticipate buyers' purchase conduct. The traditional client shopping choice-making idea may be characterised as a continuum extending from habitual hassle-fixing behaviours, thru to restrained hassle-fixing behaviours after which toward large hassle-fixing behaviours[8]. The ordinary structure for assessment of the buyer decision framework is a 5-advance adaptation. Given the adaptation, the customer advances leading from a realm of felt hardship (bother acknowledgment), to the search for realities on problem arrangements. The realities gathered manages the reason for the appraisal of options. The improvement and assessment of purchasing appraisal principles realize the genuine decision to purchase. At long last, post-purchase conduct is significant with inside the promoting and showcasing point of view, since it in this manner impacts buyers' conviction of fulfilment/disappointment with the item/administration. This conventional 5 level adaptation fuses the quintessence of customer conduct beneath most extreme settings. By the by, the control of promoting and showcasing issues at each level with inside the computerized environmental factors must be settled by means of method of methods for individual E-business visionaries. Early level in Internet improvement in phrases of constructing the best committed version of client shopping for behaviour. Decision sequences might be inspired via way of means of the start line of the client, the applicable marketplace systems and the traits of the product in question [1].

The examination can be extensively ordered in two variables: Economic: Consumers endeavor to boost their utility subject to restricted measure of cash and time, Psychological: It considers the attitude of shoppers, Economic Function and Psychological Function. Top of Form Bottom of Form Top of Form. Bottom of Form

Top of Form

Bottom of Form

Top of Form

Bottom of Form

It is over the long haul discovered that during terms, averting a purchasing ride considerable shop has far extra focal points than bearing envisioning the transportation of online bought books. This finding can't resist negating others in which they accept that the great situation from saving visit time may be balanced towards the counterbalancing time spent foreseeing conveying [10]. Direct assortments among clients enthused about purchasing on line instead of customary general stores recommendations for office organization and publicizing method" Despite the touch. Com shakeout, on line deals hold to blast and are anticipated to force more weight on regular dispersion channels. Be that as it may, there's a draping nonattendance of posted observational compositions on how customers keen on buying on line carry on comparative with customers buying in a regular shop. This examine utilizes records from each traditional market scanners and a web supermarket to check anticipated varieties in inclination practices of such customers [13].

Results of examinations finished for 2 thing classes suggest that conversely with standard grocery store clients, on line clients are significantly less rate sensitive, pick tremendous sizes to humbler sizes (or as a base have more weak choices for little sizes), have more solid length trustworthiness, do extra screening on the chance of image names in any case impressively less screening on sizes, and feature all the more remarkable tendency set effects.

Customers will save in monetary when there are esteem headways on unequivocal things. In an internet setting, customers will undoubtedly depend upon the worth signs to choose the idea of a thing which are presented in the website since they can't see or contact the genuine thing. their assessment came about that

purchasers with a shopping objective are more responsive towards unique messages, for instance, "set aside cash" and "markdown" while clients without shopping objective are responsive towards constrained time messages, for instance, "save more" and "free gift".viewed that at a certain point, impression of lower nature of the thing will evoked when there is a worth reduction or there will be requests on the retailer's plan to decrease the expense. Cost will be a basic factor when customers search thing in a web and this would show that Consumer social point are related to the worth signs that are on the whole the more immediately accessible from memory[3]. Online shopping has become colossally in view of different points of interest that customers see the technique for shopping to have. Among them is the flexibility of acquirement where the Consumer can find information similarly as make a purchase at whatever point wherever. Another favorable position is the cash sparing bit of leeway and this comes considering the way that online things are believed to be fundamentally more affordable than physical store products[4].

The expense factor is a direct result of the diminished overheads notwithstanding the charges of safeguarding the shop fronts. In spite of reality that on line stock every now and then suffer dispatching charges onto the absolute last expense, the advancing expense of limit of those product is cheap than the ones of substantial stores. The 1/3 increase is that of solace wherein individuals can do buying at whatever point anyplace. The development of at whatever point net empowered contraptions all in all with loads of buying programs for cell phones have upheld people to practices on line buying[14]. The above gifts incorporate various downsides that have certainly kept drop down the expansion of on line buying. The first and transcendent difficulty is that of net security in expressions of ID burglary and financial assessment card misrepresentation. Individuals in city areas are substantially more liable to search for on line because of the reality people in city districts are additional open to new innovation and musings than the ones in faraway locales. In this manner, with on line looking for being a progressive buying channel, the development dissemination theory expresses that city residents are considerably more liable to be on line purchasers. This examinations focuses to uncover the associations among e-buying and in-shop buying the utilization of an example of Internet clients withinside the Minneapolis-St. Paul metropolitan region through some of diagnostic methodologies [5].

Web based shopping is a type of online business which grants customers to straightforwardly buy products or administrations from a dealer by utilizing the Internet. Different names are: e-shop, e-store, Internet shop, web-store, virtual store and online store. An online shop excites the physical closeness of purchasing items just as administrations from web shop and this cycle of shopping is called business-to-shopper web based shopping[40]. Web based shopping is the training where customers choose to purchase the item through web. Shopper conduct as the investigation of people, gatherings, cycle and associations they use to make sure about, select, and organize of involvement, items, administrations, encounters, or thoughts to fulfill the buyer and society[41]. It adjusts segments from humanism, brain research, the executives and financial aspects. It attempts to know the dynamic cycles of purchasers, both only and in groups. It makes accentuation on singular buyers and their qualities, for example, socioeconomics and conduct factors to comprehend clients' need[42]. It likewise attempts to evaluate the impacts on the customer from family, companions, gatherings, and furthermore society by and large. Investigation of shopper conduct online clarified that the most suitable social appearances of online clients and review the manners in which they discover, partner and gauge item data[43].Correlation of study information with the current client execution hypothesis stemmed in location of a measure of issues identified with an express client gathering. The target of this report is to change these outcomes into a lot of execution exercises at vital and specialized level[38]. Usage of these suggestions will end in better transformation of clients.

There are various types of item, which are extra conceivable to be sold online, for example, book, programming, hardware and music. Intention in such things is that when purchasing these sorts of items, one needn't bother with singular assessment, if not everything items, can be attracted the item clarification and portrayals. Most products in the cell phone family fit to this gathering. As indicated by the new examination on client conduct, there are four distinctive client bunches with different purposes and inspirations. They

additionally found that normal productive assortment of music recordings. An incredible degree of specialized confirmation inside this group will in general be a confident element with regards to item prove research on the web. Selling retail merchandise on the Internet investigates that e-showcasing is the decreased adaptation of "electronic retailing" which fundamentally makes business to purchaser business. While the idea of e-following is no longer in its maturing stage; it stays to change, as reformist online business demands go about as a solid advertiser in creating e-promoting[5].

The possibility of e-showcasing varies from an assorted scope of items and administrations as against direct shopping experience. In this way, web based promoting is simply not limited to customers buy[45]. As customers are all around educated, they can take great choices. This incorporates a reasonable value correlation, item research, and demonstrating the believability of the e-store internet shopping. Web has formed into another distributive channels for some items. Utilizing the web to shop online has become an essential motivation to utilize the web, joined with looking of items and discovering data about them.

Accordingly web has built up an exceptionally serious market, where the opposition over buyers is savage. In some person to person communication website like E Bay, e-shopping is being done, where some retail online shops are accessible. The utilization of Internet in India offers a developing possibility for e-shopping. On the off chance that E-clients know the variables that influence online conduct, and the connections between these elements , further they can manufacture their new advertising procedures to change plausible clients into energetic clients, Consumer conduct is treated as an applied control since certain choices altogether influences clients conduct and anticipated activities[37]. The two principle sees that look for requesting of its mindfulness are cultural and miniature.

Web has changed the method of buyer's store, and has quickly formed into a worldwide viewpoint. Numerous organizations began utilizing internet shopping with the point of diminishing showcasing costs, that will prompt lessen the cost of their items so as to remain forward in exceptionally serious business sectors. Organizations likewise utilize the Internet to convey, interface and circulate data and items[36]. Client utilizes the Internet from multiple points of view for purchasing the items, yet additionally to think about item structures, costs, warrantees and conveyance administrations. Numerous authorities are certain about the eventual fate of web based showcasing business[35]. In amassing to the awesome capability of the E-market, the Internet gives an interesting opening to organizations to extra effectively to arrive at existing and potential clients. Albeit greatest pay of web based managing originates from business-to-business exchange, the experts of business-to-purchaser exchange ought not lose their certainty[37]. Scientists and experts of web based business consistently battle to form an improved vision into purchaser conduct. Alongside the improvement of E-retailing, researchers keep on clarifying E-shoppers' conduct from assorted viewpoint[50]. A considerable lot of the investigations have suspicions which depend on traditional models of customer conduct, and afterward study the legitimacy of E-showcasing [51].The resoning for using andom tree is faster algorithn to create tree and high accuracy.

3. Motivation

Data mining techniques are more and more frequently used on numerical or structured data to discover new knowledge and the benefit of such techniques is well proven. In recent years, the amount of information has been growing quickly within organisations. Thus, several organisations have implemented analytics and business intelligence (BI) for Big Data solutions. To effectively foster the acceptance and usage of BI, this study examines how the motivational factors of both expectancy theory and social exchange theory (SET) affect managers' willingness to utilise BI. However, knowledge captured in textual documentation is also a very valuable information source for any organization, but methods and tools to explore and exploit such data are less mature. Keeping these generalities in mind, it's time to take a closer look at how specific businesses have benefited from data mining and business intelligence. Usinesses can't thrive without

reliable data mining and business intelligence options in the modern world, and the importance of data and analytics will only become more important over the next few decades[40].

4. Study Objective

To grasp which is the most preferred technique for shopping. In like manner to appreciate what Consumer direct into while picking the products for shopping. Shopping has changed as a result of the effect of development with a considerable number individuals slanting toward online shopping to the ordinary physical store shopping. This example turned into the predominant point of convergence in the earlier decade with various retail mammoths joining the two different ways to manage achieve most outrageous favorable position. This assessment intends to understand the relationship between's online shopping and physical store shopping, The Main Objectives are: 1)To predict buyer conduct towards web based shopping, 2) To discover the customer significant properties towards web based shopping, 3) To inspect how socio-segment factors influence Consumer demeanor towards web based shopping, 4) To discover how the example of Online purchasing impacts shopper's mentalities towards web based shopping and 5) To research the variables that assumed a job of boundary to the shopper's mentality towards on the web shopping.

5. Proposed Approach

Egypt is the third most populated country in Africa with its 99.3 million tenants and has recorded an immense improvement in web invasion rates starting late. The amount of web customers rose to 49.2 million around the completion of 2020 (Internet World Stats), speaking to half of the masses. While this rate stays over the African ordinary (35.2%), it stays underneath the overall typical (54.4%)[38].In any case, this denotes a 41-percent expansion contrasted with January 2020 (14 million new clients in a year). The researcher will propose an intelligence model to predict consumer purchasing behavior towards online shopping[40].

6. Results Analysis and Discussion

This section describe the dataset collection, building random tree model, results analysis and discussion.

6.1 Dataset Collection

The information utilized for this investigation were gathered by Jumia online database. The Egyptian online business advertise is blasting and set to be the biggest in Africa on account of a huge populace and quickly expanding web infiltration rates. As a country at a convergence between the Arab world and Africa, both holder Arab and skillet African web business locales are notable in Egypt. In that limit, the local arrival of the Pan-Arab general retailer Souq is among the most standard destinations in the country simply like the Pan-African general retailer Jumia.The datasets for Jumia online are 1003000 models and 13 ostensible qualities as showed in Table 1.

Table 1. Jumia Data Set Collection

ID	Attribute Name	Description	Domain
1	Gender	Consumer Gender Type: Binary [M = Male Or F = Female]	[M, F]
2	Age	Age Type: Nominal [A=20-29 / B=30-39 / C=40-49 / D=50-59 / E= Above 60]	[A, B, C, D, E]
3	Education	Consumer Education Level Type: Nominal [[H,M,L]

		H= High / M= Medium / L = Low]	
4	Types of Goods	Types Of Goods Type: Nominal [A= Physical Goods / B= Digital Goods / C= Services]	[A, B, C]
5	REGION	The Different Regions Of Egypt Type: Nominal [D=Delta / E= Est / W= West / N= North / S= South]	[D, E, W, N, S]
6	Mouse click	Frist Mouse click Type: Nominal [A= Goods / B= Services]	[A, B]
7	Online Shopping References	Type: Nominal [A= Website Advertisements / B= Family Or Friend Recommendation C= E-Mail Links	[A, B, C]
8	Staying Time on website	Time Spent On Website Type: Nominal [A=Less Than one H./ B=1:2 H. / C=Above 2H.]	[A, B, C]
9	Frequency Of Online Purchase	Number Of Purchasing Frequency Of Online Purchase Type: Nominal A=Once A Week / B= More Than Twice In A Week / C= Once In A Month / B= More Than Twice In 3 Months / D= Once In 6 Months / E =Once In A Year	[A, B, C, D, E]
10	Payments Types	Consumer Payments Types Credit Card / Payments Type: Nominal [Cc= Credit Card / CH= Cash]	[CC,CH]
11	Invoice Types for goods	Invoice Types For Goods Type: Nominal [B=Basic Needs / E= Entertainment Needs / Be =Basic & Entertainment	[B, E, BE]
12	Total Price For Invoice	Total Price For Invoice Type: Nominal [A=Less Than 1000./ B=1000-3000 / C=Above 3000]	[A, B, C]
13	Purchasing (Decision Attribute)	Purchasing Goods/Services Online By Consumer Type: Nominal [Y=Yes / N= No]	[Y, N]

6.2 Buliding Random Tree Model

For bulding the model we are using the data set throught add dataste with extination CSV file by using retriive data set process and Select Attribute process 11 form 13 attributes because 11 attributes that were used have high priority with information gain by using Select attributes process after that add Set Role process in this process set target (purchasing attribute) that is the target attribute in the dataset after that we split the data set into two phaseses 80% treaing and 20% testing and using the alorthm namely Random Tree to bulid tree then using Apply model for data set finally we use perfomance process to meger the acuracy and error this model this Random Tree Model for prdecting purchasing as showed in Fig1.

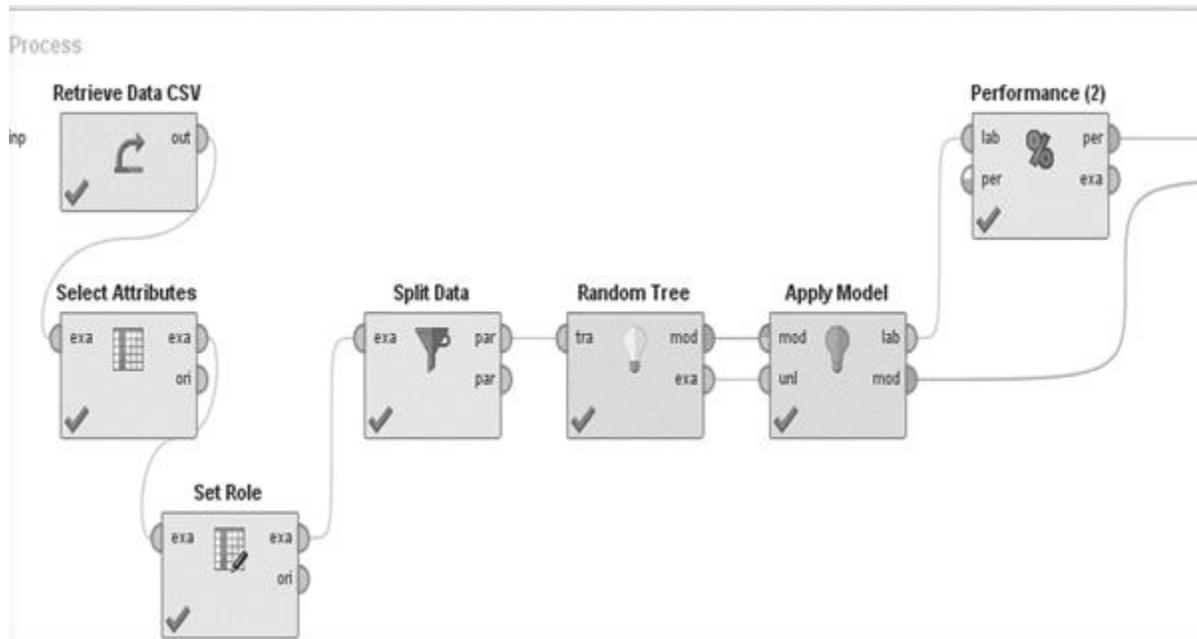


Figure 1: Buliding Random Tree Model

Then add Splite Data process to Splite 1003000 examples 80% for training and 20% for testing after that add Random Tree process to create prdiction tree then add Apply Model process at the end process Performance slecte the acuracy and the errorr this is test model for Random Tree test model for prdecting purchasing as showed in Fig2.

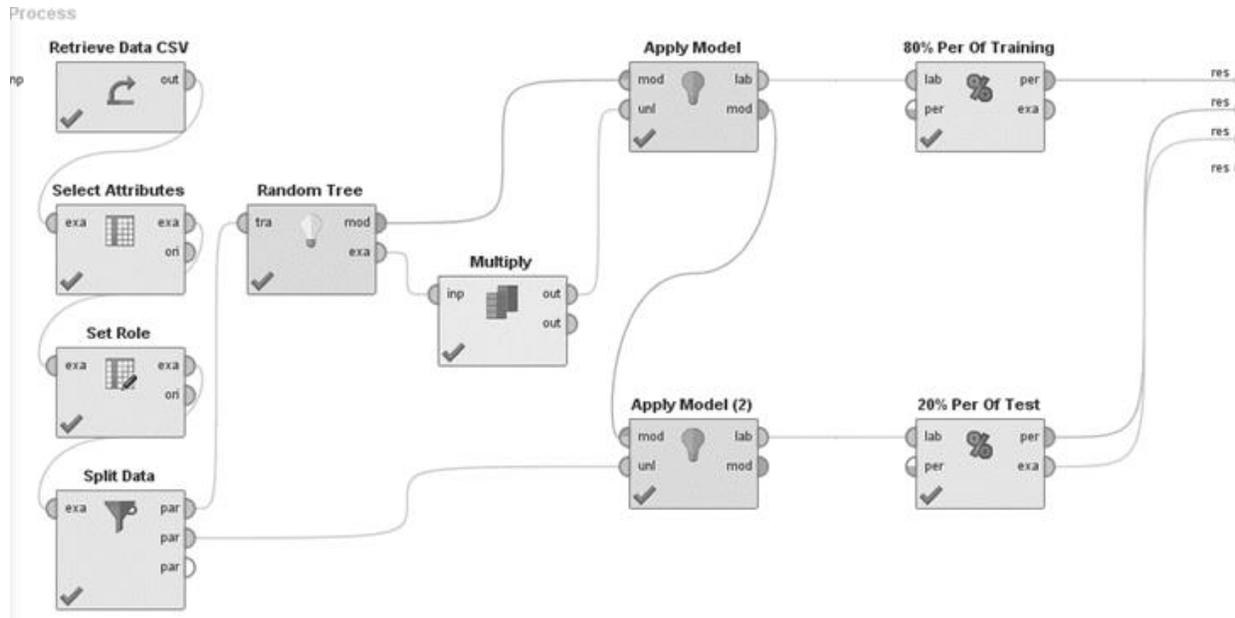


Figure 2: Buliding Tranining and Testing Model

7. Results analysis and discussion

The findings of this study clearly show that Random Tree is suitable algorithm for predicting consumer behaviour for taking decision to purchase or Not. One experiment for predicting consumer behaviour for Jumia online through 1003000 examples and 13 nominal attributes The researcher divided the data to main sections first is training data=80% and second is testing data = 20% of After running the Random Tree Model, the decision tree created as showed in Fig3 in WEKA Tool and another tree created as showed in Fig4.The most root nod in the tree is consumer education level: Three paths were created. In the first path High Level=H, Invoice Types for goods and Types of Goods are the most Attributes in all Attributes.

In the second path is Low Level = L Invoice Types for goods and Types of Goods are the most Attributes in all Attributes. Finally, Medium Level =M the result of purchasing is No purchasing. The accuracy of training model is 96.59 and testing model is 96.56. The time that Random Tree algorithm taken to build the model is 5 Second as showed in Fig7. The tree prove that If education is Medium the decision of Consumer Purchasing is No Purchasing, If education is Low Level and Invoice Types for goods is Entertainment =E and Types of Goods is A= Physical Goods , B= Digital Goods then the decision of Consumer Purchasing is No Purchasing, If education is High Level and Invoice Types for goods is Entertainment =E and Types of Goods, B= Digital Goods then the decision of Consumer Purchasing is No Purchasing, through using data mining WEKA tool for data analysis and obtaining the results as showed in Fig5.

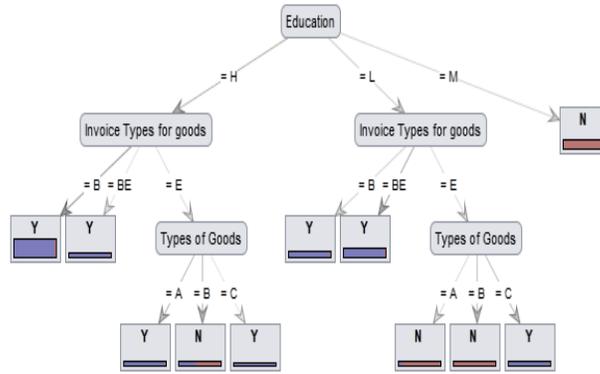


Figure 3: Random Tree

```

Tree
Education = H
| Invoice Types for goods = B: Y (Y=399394, N =11634)
| Invoice Types for goods = BE: Y (Y=11609, N =335)
| Invoice Types for goods = E
| | Types of Goods = A: Y (Y=19235, N =510)
| | Types of Goods = B: N (Y=9007, N =13017)
| | Types of Goods = C: Y (Y=31, N =0)
Education = L
| Invoice Types for goods = B: Y (Y=40465, N =1231)
| Invoice Types for goods = BE: Y (Y=134100, N =3945)
| Invoice Types for goods = E
| | Types of Goods = A: N (Y=302, N =2442)
| | Types of Goods = B: N (Y=43, N =10549)
| | Types of Goods = C: Y (Y=9759, N =268)
Education = M: N (Y=80, N =134444)
    
```

Figure 3: Random Tree by Rapid Miner

RandomTree

=====

```

Frequency of Online Purchase = A
| Staying Time on website = A
| | Gender = M
| | | Education = H: Y (167204/4870)
| | | Education = M: Y (0/0)
| | | Education = L: Y (16678/514)
| | Gender = F: Y (6288/0)
| Staying Time on website = B
| | Gender = M: Y (11462/324)
| | Gender = F
| | | Education = H
| | | | Types of Goods = A: Y (398/0)
| | | | Types of Goods = B: Y (2963/81)
| | | | Types of Goods = C: Y (0/0)
| | | Education = M: N (5932/0)
| | | Education = L
| | | | Types of Goods = A: Y (3/0)
| | | | Types of Goods = B: Y (0/0)
| | | | Types of Goods = C: Y (2963/81)
| Staying Time on website = C
| | Education = H
| | | Types of Goods = A
| | | | Gender = M: Y (3439/81)
| | | | Gender = F: Y (337/61)
    
```

| | | Types of Goods = B: Y (1992/54)
| | | Types of Goods = C: Y (0/0)
| | Education = M: N (20/0)
| | Education = L
| | | Gender = M: Y (68/0)
| | | Gender = F
| | | | Types of Goods = A: Y (37/0)
| | | | Types of Goods = B: Y (0/0)
| | | | Types of Goods = C: Y (2983/81)
Frequency of Online Purchase = B
| Education = H: N (15856/0)
| Education = M
| | Staying Time on website = A: N (187/0)
| | Staying Time on website = B
| | | Types of Goods = A: N (33523/33)
| | | Types of Goods = B
| | | | Gender = M: N (24900/26)
| | | | Gender = F: N (102554/37)
| | | Types of Goods = C: Y (0/0)
| | Staying Time on website = C: N (680/0)
| Education = L
| | Staying Time on website = A: Y (0/0)
| | Staying Time on website = B
| | | Types of Goods = A: N (3045/3)
| | | Types of Goods = B: N (13163/0)
| | | Types of Goods = C: Y (0/0)
| | Staying Time on website = C: N (96/0)
Frequency of Online Purchase = C
| Staying Time on website = A: Y (187/0)
| Staying Time on website = B
| | Types of Goods = A: Y (6/0)
| | Types of Goods = B: Y (0/0)
| | Types of Goods = C: Y (5917/162)
| Staying Time on website = C
| | Types of Goods = A
| | | Education = H: Y (15/0)
| | | Education = M: Y (0/0)
| | | Education = L: Y (36751/1097)
| | Types of Goods = B: Y (0/0)
| | Types of Goods = C
| | | Gender = M: Y (24973/690)
| | | Gender = F: Y (126158/3578)
Frequency of Online Purchase = D
| Education = H
| | Gender = M
| | | Staying Time on website = A: Y (354/0)
| | | Staying Time on website = B: Y (5781/162)
| | | Staying Time on website = C
| | | | Types of Goods = A: Y (165234/4793)
| | | | Types of Goods = B: Y (24/0)
| | | | Types of Goods = C: Y (0/0)
| | Gender = F
| | | Staying Time on website = A
| | | | Types of Goods = A: Y (427/0)
| | | | Types of Goods = B: Y (105/16)
| | | | Types of Goods = C: Y (15/0)

```

| | | Staying Time on website = B
| | | | Types of Goods = A: Y (215/0)
| | | | Types of Goods = B: N (100/36)
| | | | Types of Goods = C: Y (0/0)
| | | Staying Time on website = C
| | | | Types of Goods = A: Y (5851/61)
| | | | Types of Goods = B: Y (328/0)
| | | | Types of Goods = C: Y (60/0)
| Education = M: N (430/0)
| Education = L
| | Types of Goods = A
| | | Gender = M
| | | | Staying Time on website = A: Y (51/0)
| | | | Staying Time on website = B: Y (0/0)
| | | | Staying Time on website = C: Y (16797/514)
| | | Gender = F
| | | | Staying Time on website = A: Y (65/3)
| | | | Staying Time on website = B: N (12/0)
| | | | Staying Time on website = C: Y (634/0)
| | Types of Goods = B
| | | Staying Time on website = A: Y (27/13)
| | | Staying Time on website = B: N (52/0)
| | | Staying Time on website = C: Y (56/0)
| | Types of Goods = C: Y (479/0)
Frequency of Online Purchase = E
| Education = H
| | Staying Time on website = A: Y (185/0)
| | Staying Time on website = B
| | | Types of Goods = A: Y (6/0)
| | | Types of Goods = B: Y (5918/162)
| | | Types of Goods = C: Y (0/0)
| | Staying Time on website = C
| | | Gender = M: Y (24013/663)
| | | Gender = F
| | | | Types of Goods = A: Y (32677/972)
| | | | Types of Goods = B: Y (115003/3246)
| | | | Types of Goods = C: Y (0/0)
| Education = M: Y (0/0)
| Education = L
| | Types of Goods = A: Y (3063/93)
| | Types of Goods = B: Y (14260/421)
| | Types of Goods = C: Y (0/0)
    
```

Figure 5: Random Tree by WEKA

8. Performance measures

After running the model the performance prove that the accuracy of the model is (96.59%) as showed in Fig6

accuracy:96.59%

	true Y	true N	class precision
pred. Y	614593	17923	97.17%
pred. N	9432	160452	94.45%
class recall	98.49%	89.95%	

Figure 6: Accuracy of the model

The classification error of the model is 3.41% as showed in Table 2

Table 2. PerformanceVector of tranning

Predicted Attribute	True = Y	True = N	Precision
Predicted Y	614593	17923	97.17%
Predicted N	9432	160452	94.45%
Class Recall	98.49%	89.95%	
accuracy	96.59%	classification error	3.41%

The performane vector of the exactness of preparing model is 96.59 and the characterization mistake is 3.41 and the precision of testing model is 96.56 and the order blunder is 3.44 as appeared in Table 3

Table 3. PerformanceVector of tranning

Predicted Attribute	True = Y	True = N	Precision
Predicted Y	153427	4555	97.12%
Predicted N	2340	40278	94.51%
Class Recall	98.50%	89.84%	
accuracy	96.56%	classification error	3.44%

In WEKA Tool Time taken to build Prediction model is 1.53 seconds, Size of the tree : 123, Time taken to test Prediction model on test split is 0.3 seconds, Correctly Classified Instances = 196001 = 97.7074 % and Incorrectly Classified Instances = 4599 = 2.2926 % as showed in Fig7.

```

=== Summary ===

Correctly Classified Instances      196001          97.7074 %
Incorrectly Classified Instances     4599            2.2926 %
Kappa statistic                     0.9315
Mean absolute error                 0.0444
Root mean squared error             0.1492
Relative absolute error             12.8106 %
Root relative squared error         35.8049 %
Total Number of Instances          200600
    
```

Figure 7: Acuracy of the model by WEKA Tool

On other side The performane of Model by WEKA Tool prove that Numbers of pridicting prechasing instance are 155723 and Numbers of pridicting No prechasing instance are 25 and Numbers of pridicting prechasing instance are 4574and Numbers of pridicting No prechasing instance are 40278as showed in Table 4.

Table 4. Confusion Matrix ByWEKA Tool

Pred. Attribute	True A =Y	True B =N
Pred. A =Y	155723	25
Pred. B =N	4574	40278
Class Recall	0.973	0.936

9. Discussion

Prescient examination is additionally characterized as the utilization of measurable or AI techniques to make expectations about future or obscure outcomes. Prescient investigation manages the examination of authentic information to give forecasts about future occasions. For forecasts, prescient investigation applies a few information mining strategies and measurements, for example, grouping, arrangement, relapse, and so on. By applying this method, prescient examination assembles a prescient model that speaks to specific conditions between accessible highlights or indicators related with the occasion. Prescient examination endeavors more on understood investigation that centers on prescient displaying which alludes to the kind of information. Examination that makes expectations. About obscure future occasions and uncovers the purposes for them, typically with further investigation. Prescient examination can carry lucidity and consistency to circumstances about conceivable future conduct or dubious conditions. The outcomes demonstrated that arbitrary tree is achievement Algorithm for building choice tree for foreseeing Consumer Behavior and towards online shopping by utilizing WEKA Tool than Rapid Miner as appeared in Figure 8.

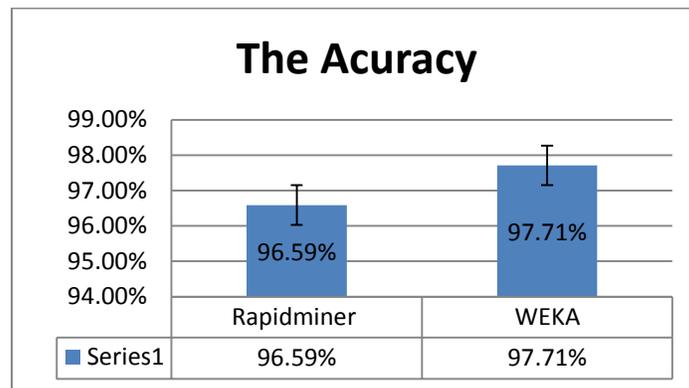


Figure 8: Accuracy of the model by WEKA and Rapidminer

10. Conclusion

Prescient investigation is utilized to foresee patterns, occasions and conduct later on dependent on information which is expected to help improvement information driven choice making. Prescient investigation is utilized more for showcasing purposes, particularly to anticipate consumer conduct where organizations begin recording consumer reactions and exchanges, at that point use information to focus on their advertising efforts. Buyer activity can change rapidly due to new needs and changes in life. Along these lines, social information can now and then be more prescient than segment information. With the conduct informatics , investigation approach, in which there is social information development that maps exchange information into social information and conduct examination that recognizes personal conduct standards, it is relied upon to be increased a more profound comprehension of consumer conduct to help more exact forecast examination for expanding business choice making. Future investigation, is a contextual

investigation utilizing a clustering task technique to anticipate. Consumers conduct with a comparable social examination gatherings. The performane vector of the exactness of preparing model is 96.59 and the characterization mistake is 3.41 and the precision of testing model is 96.56 and the order blunder is 3.44 as appeared in Table 3 In WEKA Tool Time taken to build Prediction model is 1.53 seconds, Size of the tree : 123, Time taken to test Prediction model on test split is 0.3 seconds, Correctly Classified Instances = 196001 = 97.7074 % and Incorrectly Classified Instances = 4599 = 2.2926 % as showed in Fig7. On other side The performane of Model by WEKA Tool prove that Numbers of pridicting prechasing instance are 155723 and Numbers of pridicting No prechasing instance are 25 and Numbers of pridicting prechasing instance are 4574 and Numbers of pridicting No prechasing instance are 40278 as showed in table 4. Finally we aprove that weka tool is sutiable for dataset for predicting behavior and towards online shopping with accuracy 97.71%.

10. REFERENCES

- [1] Linar Zhou, Liwei Dai and Dongsong Zhang, 2019, "Internet Shopping Acceptance Model-a Critical Survey of Consumer Factors In Online Shopping", *Journal Of Electronic Research*, VOL 8, No.1
- [2] M. Ouahilal, M. El Mohgajir, M. Chahhou, and B. E. El Mohajir, 2019, "A relative investigation of prescient calculations for business analytics and choice emotionally supportive networks: Finance as a contextual analysis," in 2019 International Conference on Information Technology for Organizations Development, 2019, pp. 1–6
- [3] M. Krueger, M. Badamchizadeh, and M. R. Feizi-Derakhshi, 2016. "Insect Inspired Fuzzily Deceptive Robots," *IEEE Trans. Fluffy Syst.*, vol. 24, no. 2, pp. 374–387
- [4] Shanthi, R., and Desti, K, 2018 P,urchasers' recognition on web based shopping. *Diary of Marketing and Purchaser Research*, 13, 14-21.
- [5] Sonia Sharma and Munishwar Rai, 2017, "Customer Behaviour Analysis using Web Usage Mining", *IJSRCSE*.
- [6] Toshinori Hayashi, Yuanyuan Wang Yukiko Kawai and Kazutoshi Sumiya M, , 2018 "An E-Commerce Recommender System using Complaint Data and Review Data", *WII'18*, March 11.
- [7] A. Abbeasi, R. Y. K. Laeu, and D. E. Earthy colored, 2015, "Anticipating conduct," *IEEE Intell. Syst.*, vol. 30, no. 3, pp. 35–43
- [8] L. Crao, 2019, "Information Science: A Comprehensive Overview LONGBING," *ACM Comput. Surv.*, vol. 50, no. 3,
- [9] A. H. Yousef, A. F. Faehmy, and H. K. Mohamed, 2018, "On the utilization of prescient investigation procedures for arrange components disappointment expectation in telecom administrators," in *ICENCO 2019 - thirteenth International Computer Engineering Conference: Boundless Smart Societies* vol. 2018–Janua, pp. 250–255.
- [10] Pornpirmon Kachamas, Suphamongkol Akkaradamrongrat, Sukree Sinthupinyo, and Achara Chandrachai, 2019, "Application of Artificial Intelligent in the Prediction of Consumer Behavior from Facebook Posts Analysis," *International Journal of Machine Learning and Computing* vol. 9
- [11] K. Z. K. Zehang, H. Xu, S. Zhao, and Y. Yu, 2018, "Online reviews and impulse buying behavior: The role of browsing and impulsiveness," *Internet Res.*, vol. 28, no. 3, pp. 522–543, Jun.
- [12] L.-Y. Leeong, N. I. Jaafar, and A. Sulaiman, 2018, "What triggers impulse purchase in Facebook commerce," *Int. J. Mobile Commun.*, vol. 16, no. 4, pp. 459–492
- [13] L.-Y. Leoeng, N. I. Jaafar, and A. Sulaiman, 2020, "Understanding impulse purchase in Facebook commerce: Does big five matter?" *Internet Res.*, vol. 27, no. 4, pp. 786–818

- [14] A. H. Moechsin, N. S. Jalood, M. J. Baqer, A. H. Alamoodi, E. M. Almahdi, A. S. Albahri, M. A. Alsalem, K. I. Mohammed, H. A. Ameen, S. Garfan, A. A. Zaidan, B. B. Zaidan, O. S. Albahri, S. A. B. Ariffin, A. Alemran, O. Enaizan, A. H. Shareef, and A. N. Jasim, 2020, "Finger vein biometrics: Taxonomy analysis, open challenges, future directions, and recommended solution for decentralised network architectures," *IEEE Access*, vol. 8, pp. 9821–9845
- [15] Germeawnn F, Lilien G L, Fiedler L, et al, 2020, Do retailers benefit from deploying customer analytics? *Journal of Retailing*, 90(4): 587- 593. 2020.
- [16] Atleauri G, Karpatne A, Kumar V , 2019, Spatio-temporal data mining: A survey of problems and methods. *ACM Computing Surveys (CSUR)*, 51(4): 1-41.
- [17] Urkup C, Bozkaya B, Salman F, 2018, Customer mobility signatures and financial indicators as predictors in product recommendation. *PloS one*,13(7).
- [18] Kaysa E, Dong X, Suhara Y, Balcisoy S, Bozkaya B , 2018, Behavioral attributes and financial churn prediction. *EPJ Data Science*,7(1): 41.
- [19] Birant D, Kut A. ST-DBSCAN, 2018, An algorithm for clustering spatialtemporal data. *Data & Knowledge Engineering*,60: 208–221.
- [20] Han J, Pei J, Mortazavi-Asl B, et al, 2017, Prefixspan: Mining sequential patterns efficiently by prefix-projected pattern growth, *Proceedings of the 17th international conference on data engineering*, IEEE, Washington, DC, USA,215-224.
- [21] Wang S, Cao J, Yru P, 2019, Deep learning for spatio-temporal data mining: A survey, arXiv:1906.04928.
- [22] Nanbo A, Weatanabe S, Halfmann P, et al, 2018, The spatio-temporal distribution dynamics of Ebola virus proteins and RNA in infected cells. *Scientific Reports*,3: 1206.
- [23] Wang S, Luo K, Lieu Y. Spatio-temporal distribution of human lifespan in China. *Scientific Reports*, 2015, 5: 13844. [24] Gaudart J, Poudiougou B, Dicko A, et al, 2018, Space-time clustering of childhood malaria at the household level: a dynamic cohort in a Mali village. *BMC Public Health*,6: 1–13.
- [24] Kulldoerff M, Heffernan R, Hartman J, et al, 2018, A space-time permutation scan statistics for disease outbreak detection. *PLoS Medicine*,2: 216–224.
- [25] Waeng M, Wang A, Li A, 2019, Mining spatial-temporal clusters from Geodatabase. *Proceedings of International Conference on Advanced Data Mining and Applications*,4093: 263–270.
- [26] Pei T, Zheou C, Zhu A, et al, 2020, Windowed nearest neighbour method for mining spatio-temporal clusters in the presence of noise. *International Journal of Geographical Information Science*,24: 925–948
- [27] Zaliaepin I, Geabrielov A, Keilis-borok V, et al. , 2018, Clustering analysis of seismicity and aftershock identification. *Physical Review Letters*,018501: 1–4.
- [28] Kisilervich S, Maensmann F, Nanni M, et al, 2018 Spatio-temporal clustering. In: *Data Mining and Knowledge Discovery Handbook*, New York: Springer Press, 855–874.
- [29] Cao H, Maemoulis N, Cheung D W, 2018, Mining frequent spatio-temporal sequential patterns. *Proceedings of the Fifth IEEE International Conference on Data Mining (ICDM'05)*, IEEE,82–89.
- [30] Zheneg Y, Zhang L, Ma Z, Xie X, Ma W, 2019, Recommending friends and locations based on individual location history. *ACM Transactions on the Web (TWEB)*, 5(1): 5-44.
- [31] Ye Y, Zheng Y, Chen, Feng J, Xie X, 2020, Mining individual life pattern based on location history. *Proceedings of the 10th IEEE International Conference on Mobile Data Management IEEE*, 1-10.
- [32] A. Monrreale, F. Pinelli, R. Trasarti, F. Giannotti, 2019, WhereNext: A location predictor on trajectory pattern mining. *Proceedings of the 15th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, ACM, 637-646.

- [33] Di Clemeente R, Luengo-Oroz M, Travizano M, et al. , 2019, Sequences of purchases in credit card data reveal lifestyles in urban populations. *Nature Communications*,9(1): 1-8.
- [34] N. Lararnjeiro, R. Oliveira, and M. Vieira, 2018, “Applying text classification algorithms in Web services robustness testing,” in *Proc. 29th IEEE Symp. Rel. Distrib. Syst.*, Oct./Nov. ,pp. 255–264.
- [35] F. Beelli, A. Hollmann, and W. E. Wong, 2019, “Towards scalable robustness testing,” in *Proc. 4th Int. Conf. Secure Softw. Integr. Rel. Improvement (SSIRI)*, Jun. ,pp. 208–216.
- [36] P. Runweson and M. Höst, 2020, “Guidelines for conducting and reporting case study research in software engineering,” *Empirical Softw. Eng.*, vol. 14, no. 2, pp. 131–164.
- [37] K. Peffers, T. Tuunanen, M. Rothenberger, and S. Chatterjee, 2019, “A design science research methodology for information systems research,” *J. Manage. Inf. Syst.*, vol. 24, no. 3, pp. 45–77.
- [38] C. Wohwlin, P. Runeson, M. Höst, M. C. Ohlsson, B. C. Regnell, and A , 2020, *Wesslén, Experimentation in Software Engineering: An Introduction*, vol. 6. Norwell, MA, USA: Kluwer.
- [39] Farder, Peeter and Lefigh, McdAlisterm, 2018, “A Promotion-Induced Choice Restriction Model of Consumer Choice,” unpublished working paper, University of Texas at Austin.
- [40] Preeti Sharma and Sanjay, 2019 "An Approach for Customer Behaviour Analysis Using Web Mining", *International Journal of Internet Computing (IJIC)*, vol. 1, no. 2, ISSN 2231-6965.