

1.1 OVERVIEW

- Customer churn refers to when a customer (player, subscriber, user, etc.) ceases his or her relationship with a company.
- Online businesses typically treat a customer as churned once a particular amount of time has elapsed since the customer's last interaction with the site or service.
- The full cost of customer churn includes both lost revenue and the marketing costs involved with replacing those customers with new ones.
- Reducing customer churn is a key business goal of every online business.

1.2 MOTIVATION

- The prediction of customer churn is an important research direction of customer churn management.
- Cost of acquiring new customers is typically high. This makes predictive models of customer churn appealing as they enable companies to maintain their existing customers at a higher rate.
- Given the importance of customers as the most valuable assets of organizations, customer retention seems to be an essential, basic requirement for any organization.
- It is generally too late to take retention actions after a consumer churns out.
- The ability to predict that a particular customer is at a high risk of churning, while there is still time to do something about it, represents a huge additional potential revenue source for every online business.
- Identifying unhappy customers early on gives you a chance to offer them incentives to stay.

- If a business can predict when a customer is at high risk of churning when there is still time for the company to do something about it, would be a huge additional potential revenue source for every online business.

1.3 PROBLEM DEFINITION

Implementation of application for Customer Churn prediction on e-commerce website using

- Session parameters
- Classification technique

and Calculating Churn Probability for classifying about to churn customers on the basis of risk associated with them.

1.4 PROJECT SCOPE

- This project involves first recording behavior and flow of customer's actions on a website to generate behavioral patterns.
- Classifying customers into groups of loyal and churn customers using a classification technique.
- Classifying about to churn customers according to risk associated with them.
- Generate reports regarding various ratios.
- User logging in through admin side will be able to view a detailed analysis of customer's data.

1.5 LIMITATION

- Reducing customer churn requires consistent effort and focus over a period of time. The results may not be immediate but they are decidedly sweet.
- Data set is not fixed.
- Continuous Internet connection is required.
- It uses lots of memory on the server to store data.

1.6 METHODOLOGIES FOR PROBLEM SOLVING

- K-means
- Artificial Neural Network
 - i. It identifies those customers who have a clear intention of terminating the use of the website, and further divides that set into sub-sets based on churn probability.
 - ii. This gives a clear picture of which customer relationships are worth investing in further and which are just costing the organization more money than they bring.

1. Predicting Customer Behaviour in Online Shopping Using SVM Classifier

Each and every activity of a customer is stored as a byte of data in a database to collect information such as how the customer spends their valuable time, day in buying decision. Most frequents items bought and quantity of buy is also considered. The dataset is used to analyze and categorize the customer based on their purchase behavior. The classification is performed by SVM algorithm. The inventory data set and sales data set which is available in the internet is used in this work and the performance is evaluated by using the algorithms.

2. Research on E-commerce User Churn Prediction Based on Logistic Regression

Based on logistic regression model, this paper established an e-commerce user churn prediction model through preliminary research on e-commerce customer churn behaviour. By using the factor analysis method, the user's online duration, number of logins, attentions, and other user behavior factors were analyzed which concludes the factor affecting the loss of users. Finally, the empirical study proved that the proposed model can predict user churn behavior in a high confidence level. It helps e-commerce platform to formulate operational strategy more precisely, provide users with personalized recommendations, increase user activity, retain users and improve the economic effects of e-commerce platform.

3. Customer Churn Analysis - A Case Study

Churning customers are analyzed in banking context and forecast of churning customers based on a logistic regression model. In the normal regression model with n inputs the output is obtained by using the formula:

$$P(X) = \alpha + b_1x_1 + b_2x_2 + \dots + b_nx_n$$

in the logistic regression model the corresponding formula is:

$$Q(X) = 1/(1 + e^{-(\alpha + b_1x_1 + b_2x_2 + \dots + b_nx_n)})$$

The results of the continuous probabilities that are produced by the logistic regression model will be discriminated into two groups by using a threshold value.

Lift Curve was used for analysis of the estimated results of the logistic regression.

4. Churn Analysis - Predicting Churners

Churn Analysis is the calculation of the rate of attrition in the customer base of any company. It involves identifying those consumers who are most likely to discontinue using a service or product. Churn analysis is extremely helpful in developing a sustainable and robust strategy for customer retention in a company. When a company is aware of the percentage of customers who end their relationship with them in a given time period they can easily come up with a detailed analysis of the causes for the churn rate using churn analysis. This helps in developing effective customer retention programs for the company. The analysis focused on churn prediction is based on rule based classification. However the incompleteness of the dataset made it hard to create a predictive model. The findings indicate that we have to use a complete and big dataset if we want to achieve any kind of accuracy.

3.1 ASSUMPTION AND DEPENDENCIES

Customer must have an account in order to proceed with any activity. As motive of the project is recording behavior of customer, customer should log in and then can proceed with viewing products and further.

3.2 FUNCTIONAL DEPENDENCIES

- Customer level
 1. User Registration (Sign Up/ Log In)
 - User should create an account by providing required details and authorized credentials.
 - User should have an unique account to proceed with log in and further activities. (View products, buy product etc.)
 2. View Products
 - Customer can view different products and either can buy a product or add it to cart.
 - Details of product viewed by each customer are recorded for average product view rate per session.
 3. Cart Details
 - Details of products added to cart are maintained to calculate Cart abandoned rate, cart change percentage with respect to individual customer.
 4. Product buy details
 - Details of products bought by each customer like transaction amount, transaction number are saved.
 5. Session Management
 - For each particular session session_id, session_time, session_day are recorded.

- Admin Level
 1. Classification
 - Training dataset containing all attributes is used for modelling the classifier.
 - Customers are classified into loyal and about to churn categories.
 2. Churn Probability
 - Using different attributes like cart change percentage, customer lifetime value churn probability for each customer is calculated.
 - Customers are then divided into three categories as High, Medium and Low according to churn probability

3.3 EXTERNAL INTERFACE REQUIREMENTS

3.3.1 User Interfaces

The project will be basically a website designed using html, css and php. The two main modules of User interface will be Customer (Client side) and Admin (Server side).

The flow of user interface will be: -

- Customer
 1. User Registration (Sign up/ Log in)
 - User can create account by providing required details like name, email in the sign up form.
 - User must log in by providing unique credentials (username, password)
 2. View Products

A view containing different categories of products and corresponding product details.

 - Navigation bar containing different categories.
 - Different sub- categories are displayed using dropdown list.
 3. Give Order/Add to cart

- Each product will have two options Buy Now and Add to cart (Buttons)
- Admin
 - Admin side has three different functionalities has shown using buttons-
 - 1. Classification
 - Graph displaying classification loyal and about to churn customers.
 - Number of customers belonging to each category is displayed.
 - 2. Churn Probability Calculation
 - i. Display churn probability
 - Customer ID is taken from user using input type text and churn probability is returned.
 - ii. Display graph
 - Graph is displayed showing different categories of churn probability like high, medium and low.
 - 3. View Customer Session Parameters
 - Customer ID is taken and different session parameters for the particular customer are returned.
 - Drop down list containing different attributes and graph for particular attribute is displayed.

3.3.2 Software Interfaces

1. Different Python libraries used are-
 - i. scikit-learn -To build model for prediction.
 - ii. matplotlib - To plot different graphs
 - iii. datetime – To fetch date and time for session parameters
 - iv. gspread- For connection between spyder and google sheet
 - v. xlrd- To fetch data from google sheet
 - vi. auth2client- For authentication of google account

2. Flask

This platform is used represent the predictions from python on web.
It has its own port number i.e localhost/5000.

3.3.3 Communication Interfaces

1. Google Sheets

- All the session parameters are maintained in Google sheet from customer side
- Same Google sheet is used for classification at admin side.

2. WAMP (Apache server)

- It is an open source component.
- By running apache server we can test web pages in website without publishing them live on internet.
- Also used to implement PHP and MySQL connectivity.

3. The browser (Internet) will act as an interface between local website and online Product API.

3.4 NON-FUNCTIONAL REQUIREMENTS

3.4.1 Performance Requirements

- Response time will be increased if system has a good internet connection.
- Increase in utilization of all resources tends to have good performance of a system at customer/admin/company level.
- To have good throughput of a system we should have large training dataset which will then help to build a scalable prediction model.

3.4.2 Security Requirements

- Each customer should have different account for security purpose.
- Any customer should not be allowed to make changes at admin side.

3.4.3 Software Quality Requirements

1. Scalability – The system can work properly even in case of large increment in session dataset without any impact on performance. This is basically ability of system to be enlarged.
2. Availability – The proposed churn prediction system is functional and working at required time.
3. Reusability – The proposal system can work for various other websites like any entertainment website. It will be beneficial in sectors where churn prediction becomes important.
4. Usability – This system can be easily understood by users i.e. stakeholders in company. The proposed system allows them to get desired results about churning of customers efficiently. This system will be efficient for both new and experienced users.
5. Maintainability – This system is maintainable and can be easily changed according to changed data by development team.
6. Manageability – The quality of system can be administrated by team in cases of addition of more session data, changing configurations, doing backups, error correction.

7. Performance – Good performance can be achieved by using more training data with variations.

3.5 SYSTEM REQUIREMENTS

3.5.1 Database Requirements

1. A structured database MySQL is used to store Customers details, Product Details and also Session Parameters.
2. A variable training dataset will be used on the basis of which system will build a model to predict customer behavior.
3. A test dataset will be used on which actual analytical algorithm will be performed.

3.5.2 Software Requirements

1. OS- Windows
2. Platforms
 - i. Visual Studio Code (To build a website using HTML, CSS, PHP)
 - ii. Google sheet (Communication between customer and admin)
 - iii. Anaconda (Spyder - To perform analytics)
 - iv. Flask
3. Scripting Languages – HTML,CSS,PHP
4. Programming Languages – Python

3.5.3 Hardware Requirements

1. Intel(R) core i7 processor
2. 4GB RAM
3. Windows 8.0 or higher
4. A network support

3.6 SDLC MODEL TO BE APPLIED

Agile software development life cycle

- Basic outline of Agile SDLC model –

1. Plan
2. Develop
3. Test/QA
4. Deliver
5. Assess

- Different phases in agile SDLC model –

1. Concept

After finalizing the domain of project concept is finalized based on current real-time business opportunities and technical feasibility. Data Analytics is the finalized domain for the project and based on that Churn prediction concept is finalized as it is basic requirement for every company.

2. Inception

In this phase project is initiated by finalizing diagram requirements. Once we identified the project, worked to determine requirements like dataset. We defined user flow diagrams or high-level UML diagrams to demonstrate how the new feature should function. Team member responsibilities are decided.

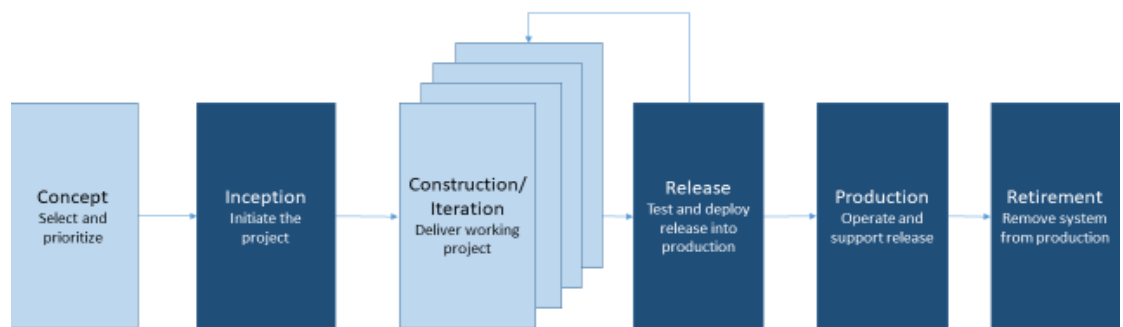


Fig.3.6: SDLC model

3. Construction/Iterations

Once we defined requirements first iteration of the project is implemented. Iteration focused on achieving the basic goal for the system and then it is revised according to change in requirements and environment. Additional sprints are added to expand upon the overall product.

4. Transition

Finishing up this software iteration with the following steps:

- Test the system -System is tested to check if functionalities are satisfying the main requirements of system.
- Address any defects
- Finalize system and user documentation –Code is visualized through UML diagrams or demonstrates user flows so everyone understands how the system functions and how they can build upon it further.
- Release the iteration into production.

5. Production

System is kept running and output is verified at each time from the perspective of each user.

6. Retirement

If the system becomes redundant for the business model we can replace it with new release of the system with some new required additions and old release is removed from production.

4.1 System Architecture

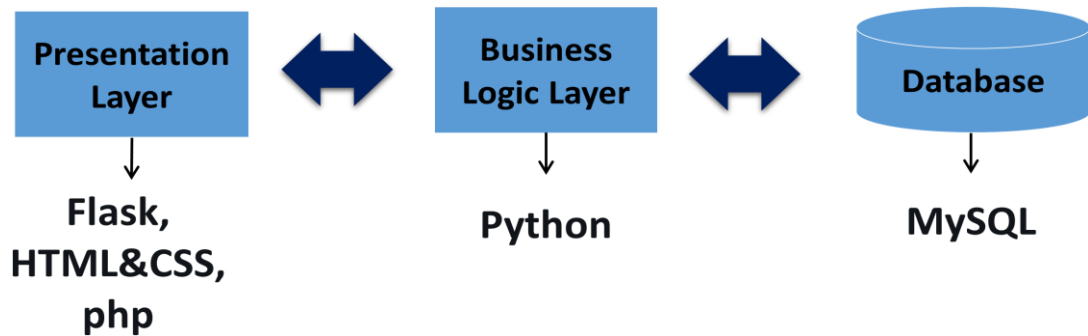


Fig. 4.1: System Architecture

1. Presentation Layer

- Flask is used to present the output obtained from python platform which mainly represents the **admin side**.
- HTML& CSS is used for building and styling for website
- Php is used for running webpage on browser without hosting them and for session management.

2. Business Logic layer

- Python is used for classification purposes as well as calculation of some important parameters used for churn probability is done.
- pygal from python is used for data visualization.

3. Database

- MySQL is used for storing the data regarding customer's information, product lists and session parameters.

4.2 Entity Relationship Diagram

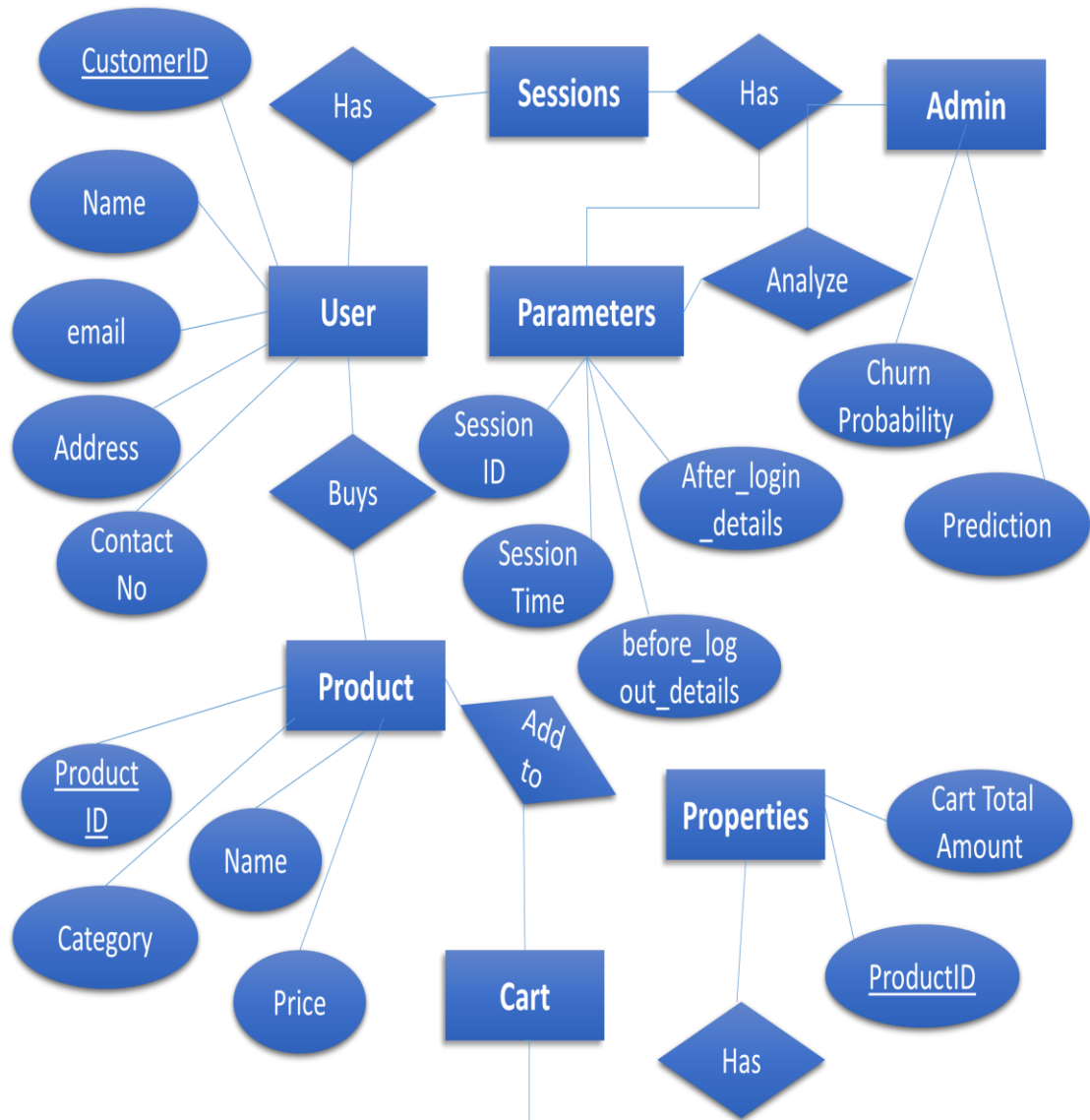


Fig. 4.2 : ER Diagram

4.3 UML Diagram

4.3.1 Use Case

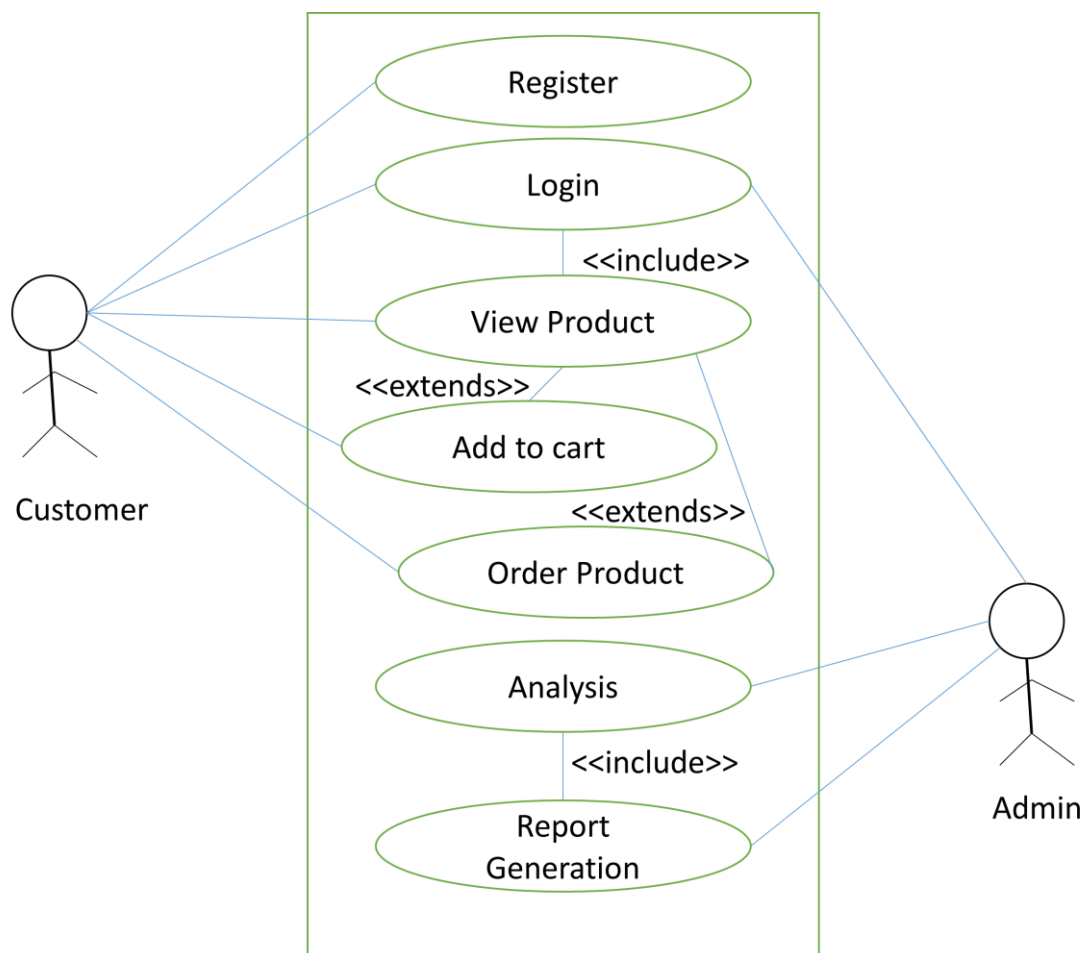


Fig. 4.3.1 :Use Case Diagram

4.3.2 Class Diagram

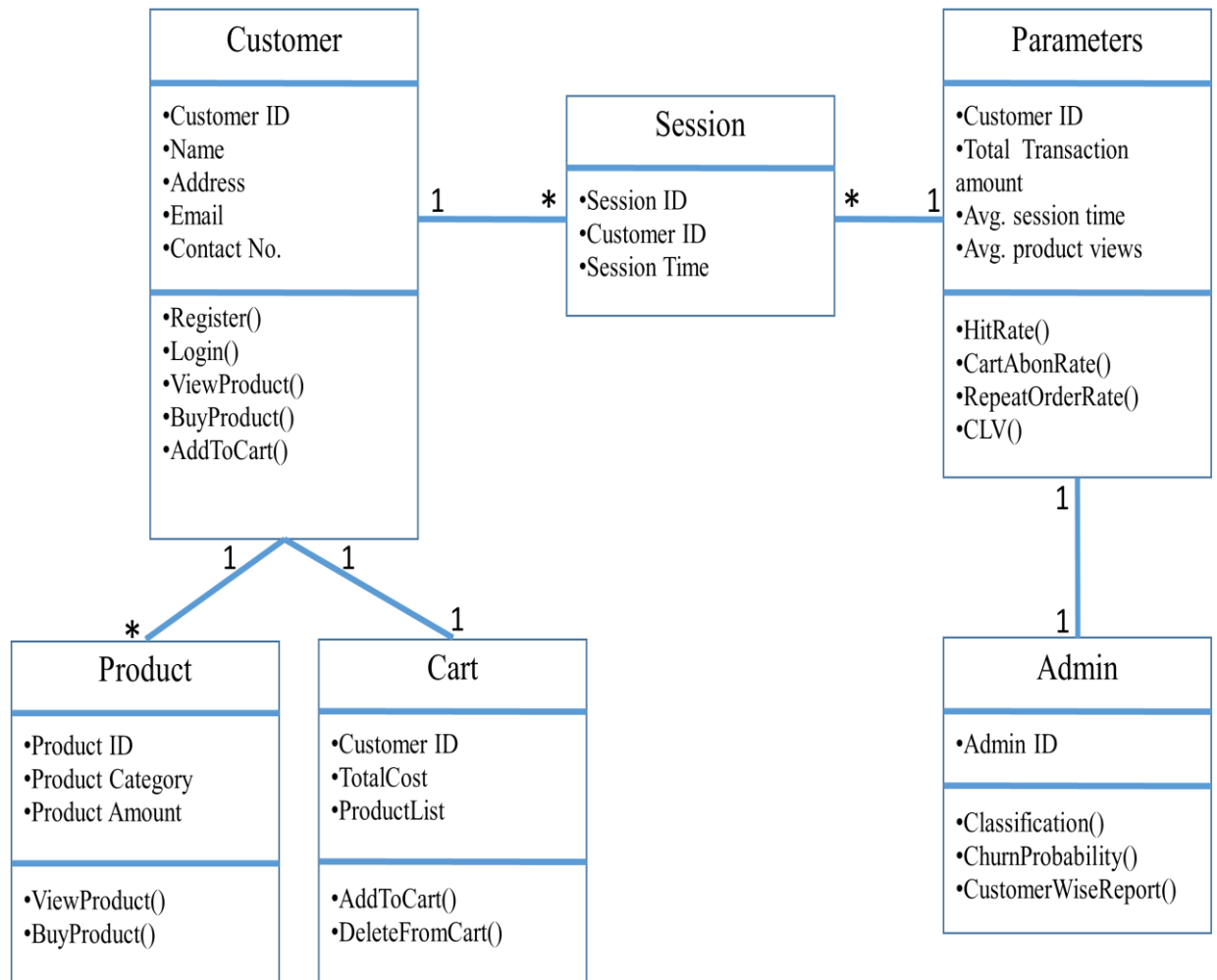


Fig. 4.3.2 : Class Diagram

4.3.3 State Chart

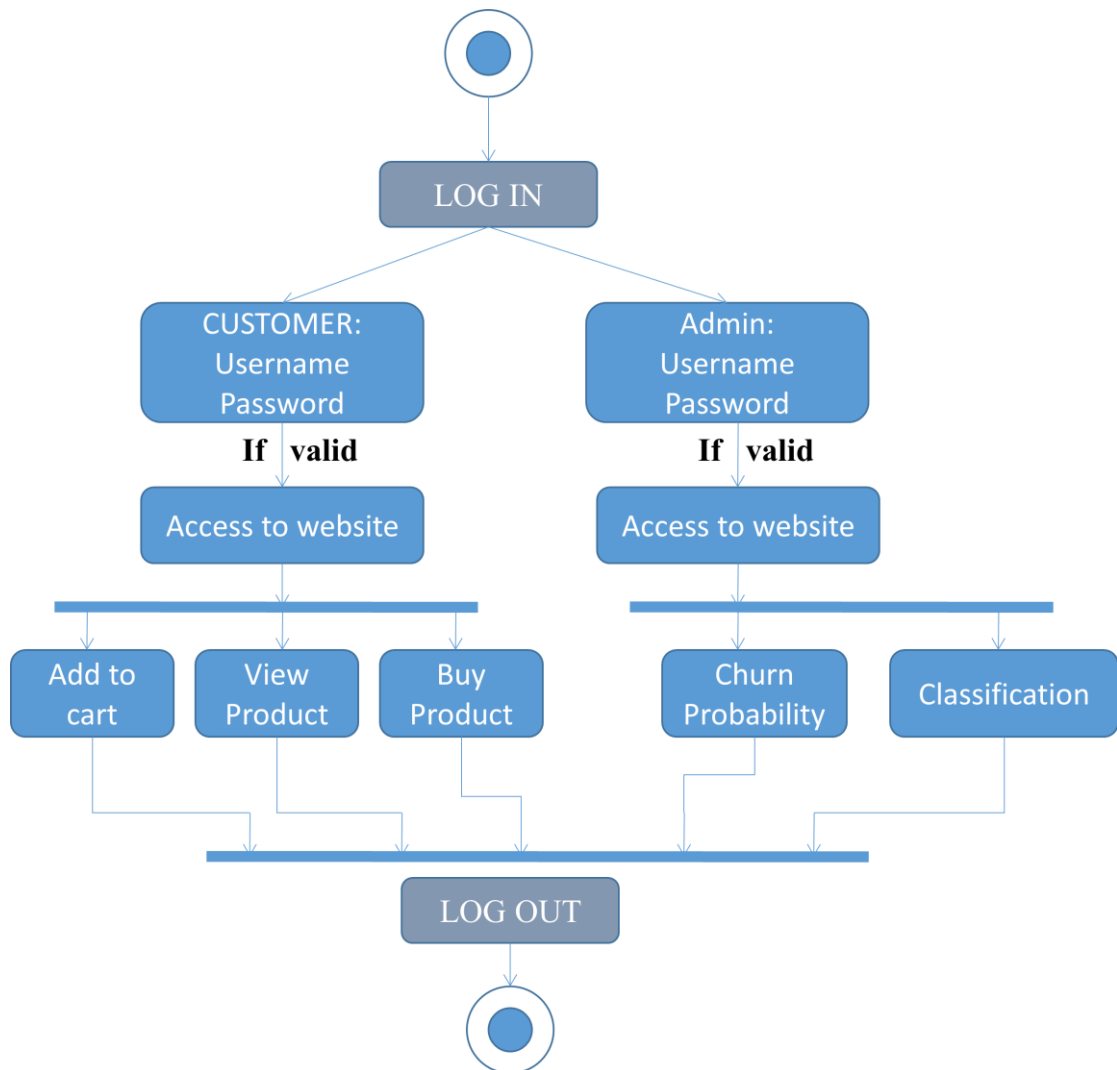


Fig. 4.3.3 : State Chart Diagram

4.3.4 Sequence Diagram

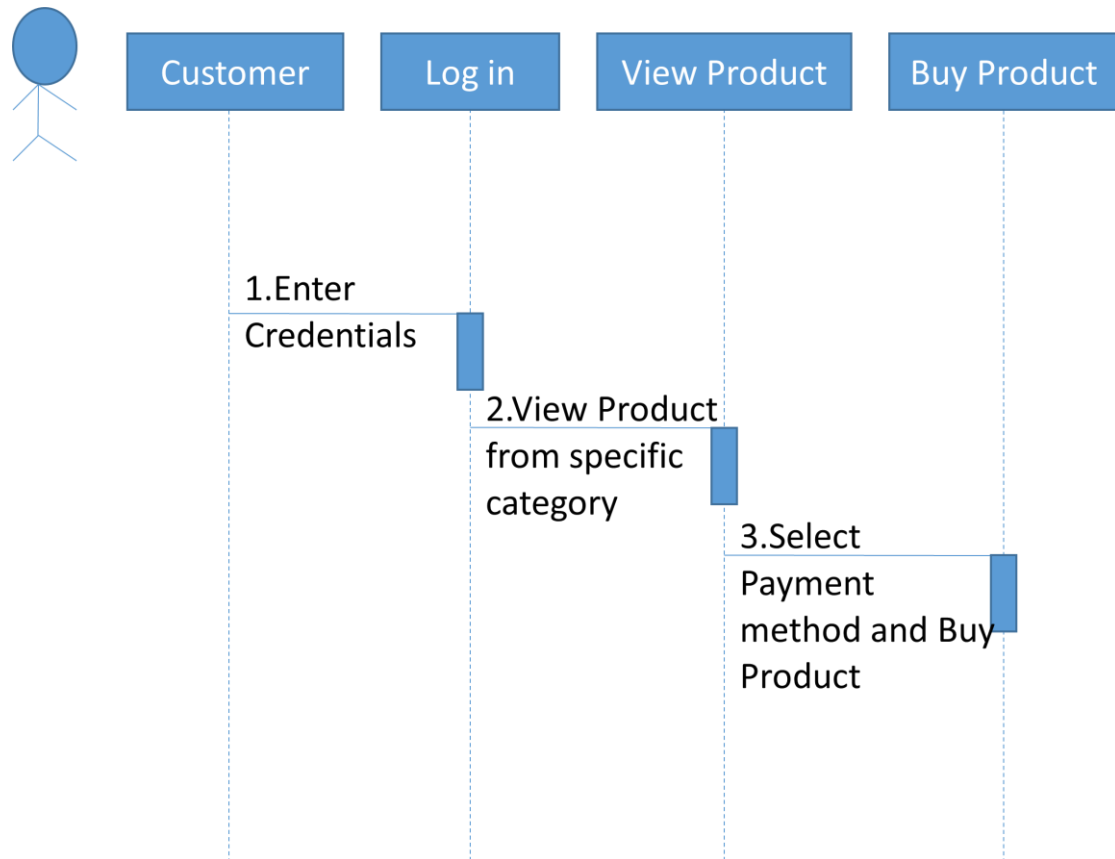


Fig. 4.3.4 : Sequence Diagram

5.1 PROJECT ESTIMATE

5.1.1 Project Resources

- Operating System : Windows
- Programming Languages : Python, PHP
- Platform: Spyder, Flask, Visual Studio Code

5.2 PROJECT SCHEDULE

5.2.1 Project Task Set

Sr.No.	Task	Period	Expected Deliverables	Target Achieved
1	Thoroughly study of the domain	Jun-Jul 18	To have knowledge of the domain and understanding the critical issues	Clear understanding of the domain
2	Survey of existing systems	Aug 18	To understand the proposed systems and their drawbacks	Selected the limitations that can be addressed in our system
3	System framework	Sept-Oct 18	To come up with the framework of model	Framework of the system

Table 5.2.1 : Project Task Set

5.2.2 TimeLine Chart

	Task Name	Q3			Q4			Q1			Q2		
		Jul	Aug	Sep	Oct	Nov	Dec	Jan	Feb	Mar	Apr	May	Jun
1	Decide concept		Decide concept										
2	Requirements Gathering		Requirements Gathering										
3	Plan			Plan									
4	Develop						Develop				Develop		
5	Test/QA										Test/QA		
6	Deliver										Deliver		
7	Assess										Assess		

Fig 5.2.2 : Timeline Chart

5.3 TEAM ORGANIZATION

5.3.1 Team Structure

Sr. No.	Team Member Name	Individual Contribution
1.	Rachana M. Fulsundar	Literature Survey, Churn Probability, Dataset Creation
2.	Pallavi R. Kale	Literature Survey, Classification implementation, Churn Probability
3.	Kimaya K. Khilare	Literature Survey, MySQL database, Front-end
4.	Dhanshree B. Lonkar	Literature Survey, Database, Session Management

Table 5.3.1 : Team Structure

5.3.2 Management reporting and communication

Project meeting with respective project guide took place regularly. In every meeting, progress of the project, areas of improvement, how to plan out remaining project and how to deal with the problems and deviations being faced during implementation of project were highlighted, discussed and solved. At the end of each project a target was set upon which the project group used to work and come up for the next meeting.

6.1 OVERVIEW OF PROJECT MODULES

The two main modules of the system will be Customer (Client side) and Admin (Server side).

- Customer
 1. User Registration (Sign up/ Log in)
 - User can create account by providing required details like name, email in the sign up form.
 - User must log in by providing unique credentials (username , password)
 2. View Products
 - A view containing different categories of products and corresponding product details.
 3. Give Order/Add to cart
 - Customer can either add product to cart or can by product or can just view and leave.
 4. Logout
 - Customer can leave the website anytime by logging out.
- Admin

Admin side has three different functionalities has shown using buttons-

 1. Clustering
 - Using K-means clustering algorithm we divide data into two categories loyal and about to churn customers.
 2. Churn Probability Calculation

Using Artificial Neural Network based on session parameters churn probability for each customer is calculated.

 - i. Display churn probability- Customer ID is taken from user and respective churn probability is displayed.
 - ii. Display graph - Graph is displayed showing different categories of churn probability like high, medium and low.

3. View Customer Session Parameters

- Customer ID is taken and different session parameters like Customer lifespan value, Hit rate, Repeat Order rate etc. for the particular customer are returned.

6.2 TOOLS AND TECHNOLOGIES USED

1. HTML, CSS, Bootstrap

- Dummy E-Commerce website is designed using HTML, CSS.
- Bootstrap is used more inbuilt classes for different widgets in website. It will help to make User Interface more user friendly and beautiful.

2. PHP

- PHP is used for freedom to use operating system and web server.
- MySQL database connectivity with website designed is implemented using PHP where details of products, customers and session parameters is stored.

3. Anaconda – Spyder

- Spyder is platform for Python programming. Backend programming i.e Classification and Artificial Neural Network algorithms are implemented in Spyder.

4. Google Sheets

- Google sheet is used to store live session parameter data from website.
- Backend algorithms are implemented by fetching the data from Google sheet.

5. Flask

- Flask is used to represent the output of prediction algorithm on web browser.

6.3 ALGORITHM DETAILS

1. Cart Abandoned = 0 if before_logout_cart=0
 1 if before_logout_cart!= 0
2. Average Session Time =
$$\frac{\text{Sum of top N session times}}{N}$$
3. Average number of Product Viewed =
$$\frac{\text{Sum of all product views}}{\text{Number of sessions}}$$
4. Average session Amount =
$$\frac{\text{Total of all non-zero transactions}}{\text{Number of non-zero transactions}}$$
5. Cart Abandoned Rate =
$$\frac{\text{Number of Time cart abandoned}=1}{\text{Number of Sessions}}$$
6. Hit Rate =
$$\frac{\text{Number of product bought}}{\text{Average number of product bought}}$$
7. Repeat Purchase Rate =
$$\frac{\text{Number of Sessions which Included transaction}}{\text{Total number of sessions}}$$
8. Customer lifespan = Last transaction month - First transaction month
9. Customer Lifetime Value = Average Transaction Amount *
 Repeat Purchase Rate * Customer lifespan

.1.1 Churn Prediction Module

1. Logistic Regression:

Logistic regression is basically a supervised classification algorithm. In the classification problem, the target variable (or output), y can take only discrete values for given set of features (or inputs) x . Logistic regression predicts the probability of an outcome that can only have two values (i.e. a dichotomy). The prediction is based on the use of one or several predictors (numerical and categorical). As the data we have is in binary form so we are using logistic regression achieve the higher accuracy.

2. Algorithm

1. Logistic regression model is built on our session wise dataset.
From `sklearn.linear_model LogisticRegression` library is used.
`Classifier_name = LogisticRegression()`
`LogisticRegression.fit(x,y)`
`y_prediction= Classifier_name.predict(x)`
2. Outcome values as loyal (0) and about to churn (1) is predicted on test dataset on previously build module.

.1.2 Churn Probability Module:

This module will give us the severity of churn of specific customer. The churn probability of any customer can be calculated only if the classification module has predicted YES for churn. For that class only, we can tell the severity like High, Medium, Low and approximate value of churn probability also. All operation described above can be simply done by using **Artificial neural network**.

1. Artificial Neural Network (ANN):

- An artificial neuron network (ANN) is a computational model based on the structure and functions of biological neural networks. Information that flows through the network affects the structure of the ANN because a neural network changes - or learns, in a sense - based on that input and output

- ANNs have three layers that are interconnected. The first layer consists of input neurons. Those neurons send data on to the second layer, which in turn sends the output neurons to the third layer.

2. Algorithm:

1. Fetch the data from file where all parameter of individual customers are stored.
2. Get the values of hit rate, cart abandonment rate, repeat order rate, customer life value, customer lifespan , average session time and average product views of every customer.(This values will be considered as input parameters for input layer of artificial neural network)
3. Synaptic weights for all parameters are assigned randomly at initial stage
4. Weights are multiplied to the parameter values of each tuple and the matrix of outputs is formed.
5. Error is calculated: $\text{Error} = \text{Expected Output} - \text{Calculated Output}$
6. Adjustment is calculated : $\text{adjustment} = \text{Error} * \text{sigmoid_derivative}(\text{outputs})$
7. Again the synaptic weights are updated according to the error and adjustment :
8. Synaptic
 $\text{weights} += \text{matrix_multiplication}(\text{input_layer.T}, \text{adjustment})$
9. Counter values for specific output is calculated
10. Repeat steps 4 to 8 for decided threshold value
11. Counter having maximum value is considered as class or range of churn probability of that output.



Fig no. 6.3

7.1 TYPE OF TESTING

7.1.1 Unit Testing

Unit testing is a software development process in which the smallest testable parts of an application, called units, are individually and independently scrutinized for proper operation. Unit testing can be done manually but is often automated. In the scenario where requirements change rapidly, manual testing is preferred over automation testing.

After the design of proposed model (in the first half of project development phase). Black box testing was carried out. In this testing knowledge of programming is not necessarily essential. Moreover, knowledge of implementation is not required for this testing. This testing mainly focuses on the functionality of the system.

The tests cases were executed for comparison of the actual outputs with the expected outputs. Defects if any fixed and re-tested.

As soon as the coding for trust management for trust management model is complete (second half of project development phase). White box testing was carried out. In this form of testing knowledge of programming is must. It mainly focuses on the structure means program/code of the system. Little tests for each process or series of processes were developed. It involved testing a series of predefined inputs against expected or desired outputs so that when a specific input does not result in the expected output, you have encountered a bug. This let's know about the internal security holes and broken or poorly structured paths in the coding processes.

7.1.2 Integration Testing

Integration testing is a software testing methodology used to test individual software components or units of code to verify interaction between various software components and detect interface defects. Components are tested as a single group or organized in an iterative manner. After the integration testing has been performed on the components, they are readily available for system testing.

Integration Testing is defined as a type of testing where software modules are integrated logically and tested as a group. A typical software project consists of multiple software modules, coded by different programmers. Integration Testing focuses on checking data communication amongst these modules. The main function or goal of this testing is to test the interfaces between the units/modules. The customer side and the admin side is integrated and then tested together.

Big bang approach integrates all the modules in one go i.e. it does not go for integrating the modules one by one. In big bang approach of integration testing, all components are integrated together at once and then tested. It is also convenient for small systems. It verifies if the system works as expected or not once integrated. If any issue is detected in the completely integrated module, then it becomes difficult to find out which module has caused the issue.

7.2 TEST CASES AND TEST RESULTS

Precondition /Assumption	Test Case Id	Test Case Name	Test Case Description	Test steps	Expected Result	Actual Result
Open web browser and enter the given URL in the address bar.	TC1	Validate Name	To verify that name should contain only characters	Enter user name	Error message, "Please enter valid name"	Error Message displayed
Open web browser and enter the given URL in the address bar.	TC2	Validate email	To verify that entered text should be in email format	Enter email id	Error message, "Please enter valid email address"	Error Message displayed

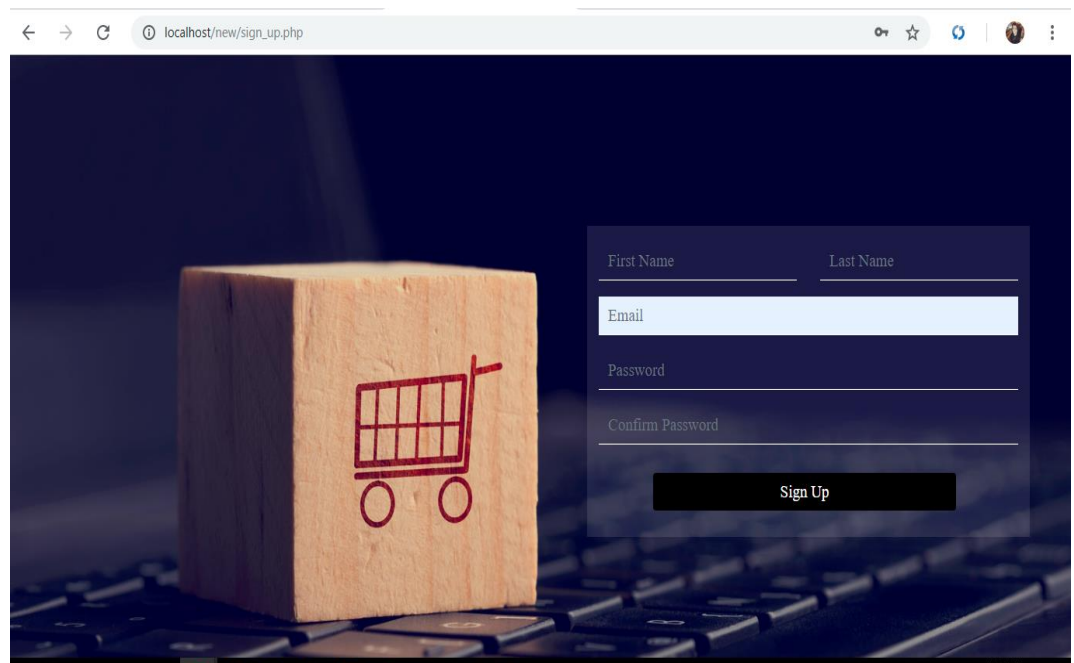
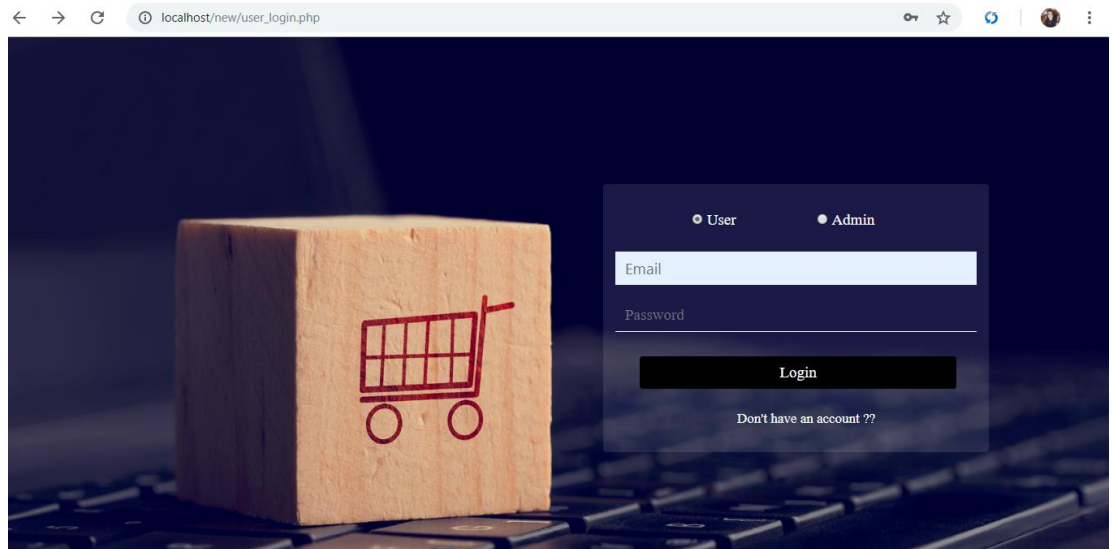
Open web browser and enter the given URL in the address bar.	TC3	Validate Password	Verify that password should contain more than the 8 characters	Enter password of length seven containing at least one number and special character	Error message according to password incorrectness i.e. "Password should contain at least one special character"	Error message displayed
Open web browser and enter the given URL in the address bar.	TC4	Validate navigation	Verify that user is navigated to correct position.	Click on the link.	User navigated to correct position	User navigated to correct position
Open web browser and enter the given URL in the address bar.	TC5	Validate Button Click	Verify that button is getting clicked.	Click on button	Respective action should get performed	Respective action performed
Database should exist and database connection program should be written.	TC6	Validate database connection	Verify that database is connected.			
Program for insert query should be written	TC7	Validate database insert query	Verify that data is getting inserted.			

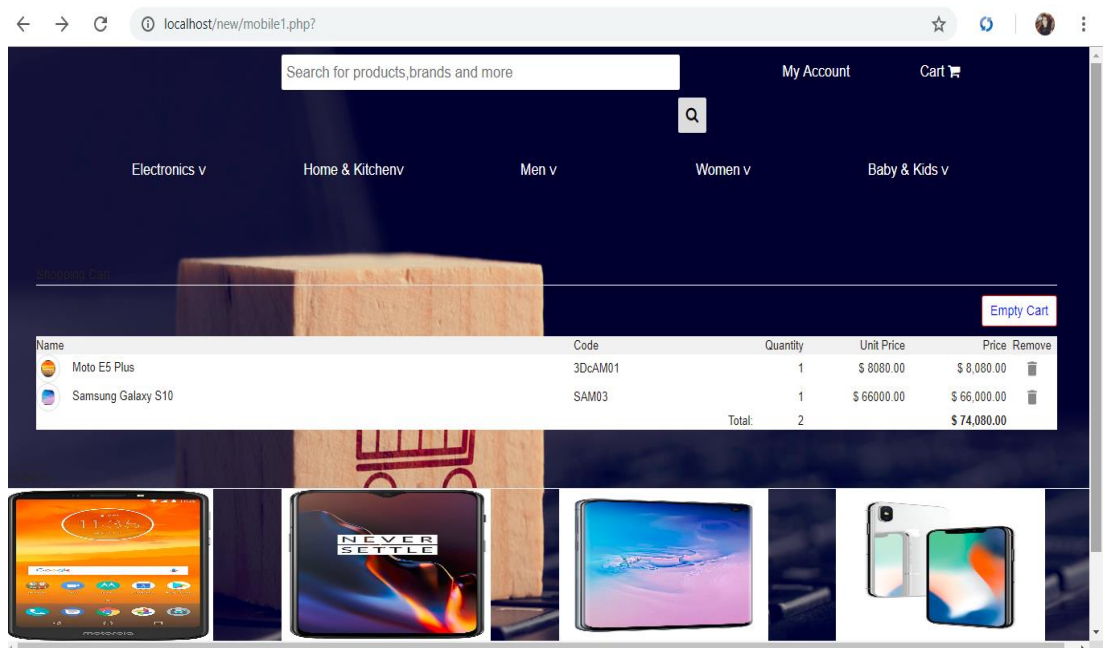
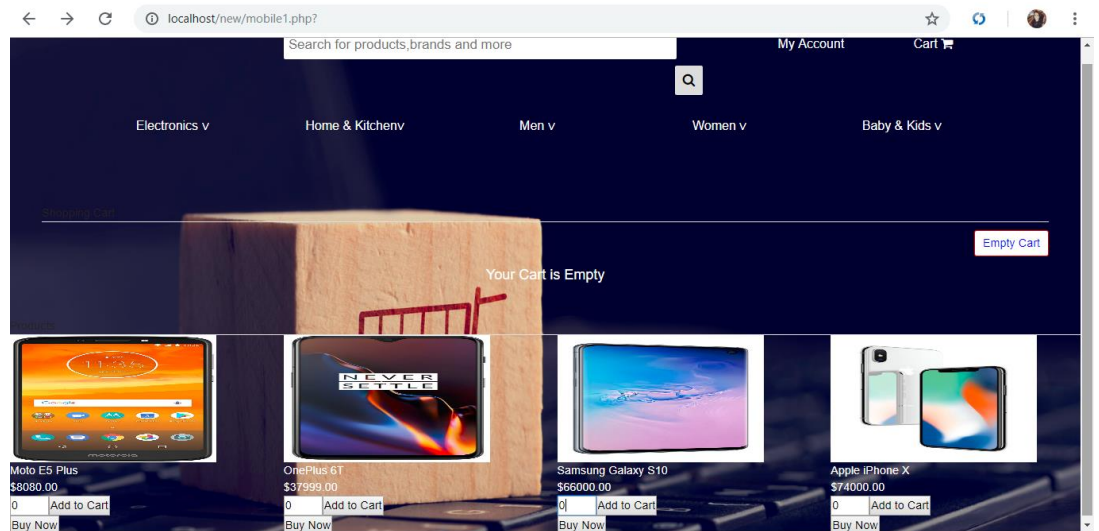
Table no 7.2

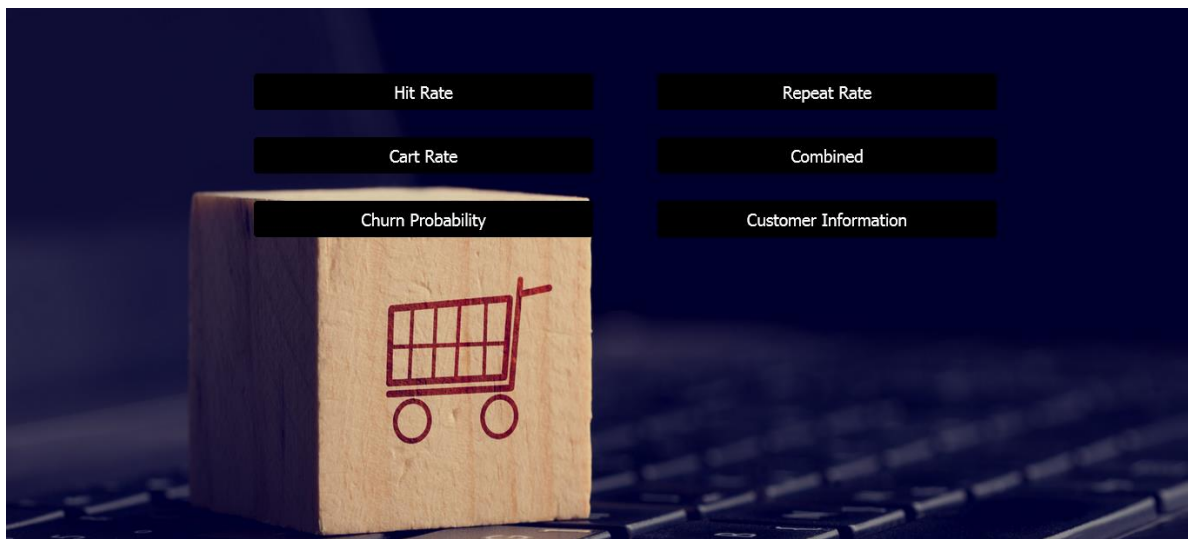
8.1 OUTCOMES

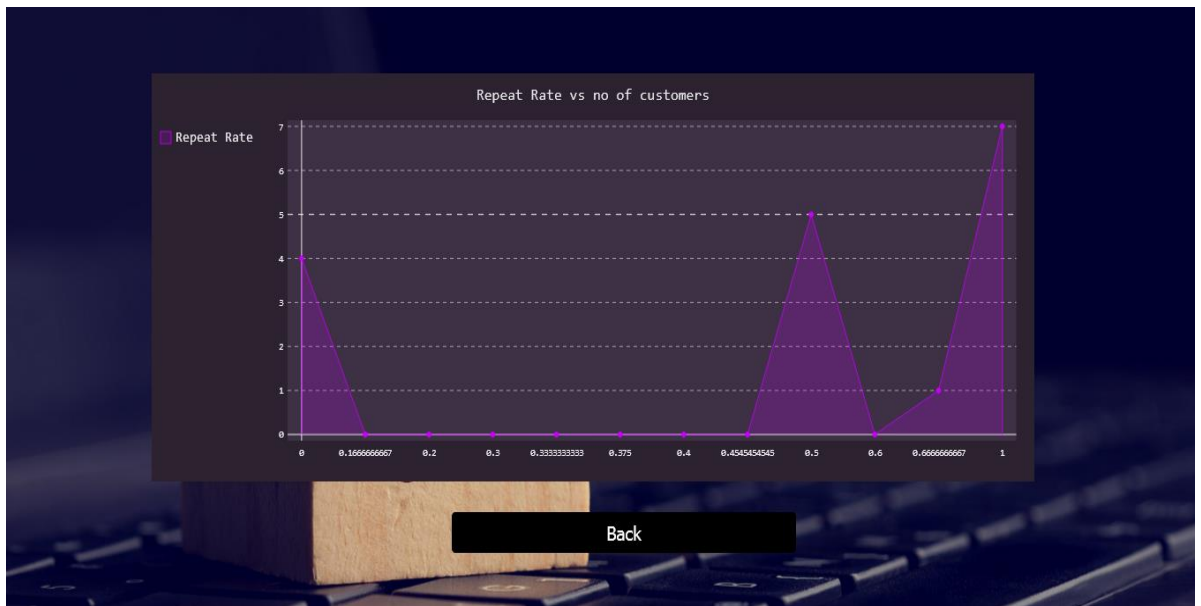
- Customer
 - An dummy E-commerce website having Log in and Signup facilities.
 - New customer data is stored in MySQL database. While logging in only authenticated customers are allowed to proceed.
 - For particular customer cart and transaction details for after_login, before_logout are stored as session parameters as input to admin.
- Admin
 - Graphs for individual customer rates like hit rate, repeat rate etc.
 - Combined graph of different rates for comparison.
 - Pie chart displaying number of churning customers belonging to each category i.e high/medium/low.
 - Details of rates are displayed customer wise.
 - Regression graph representing customers belonging to loyal or about to churn class.

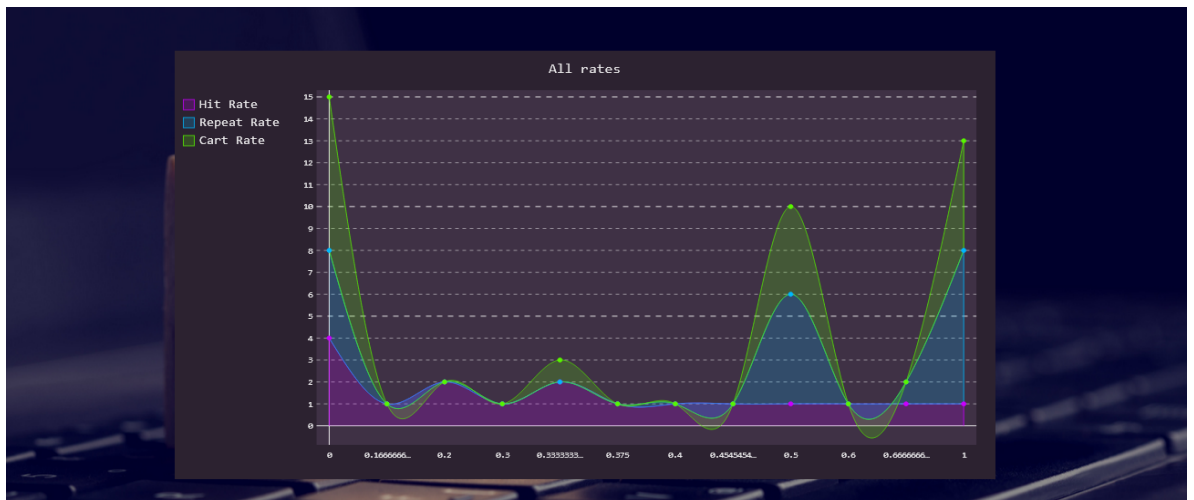
8.2 SCREENSHOTS





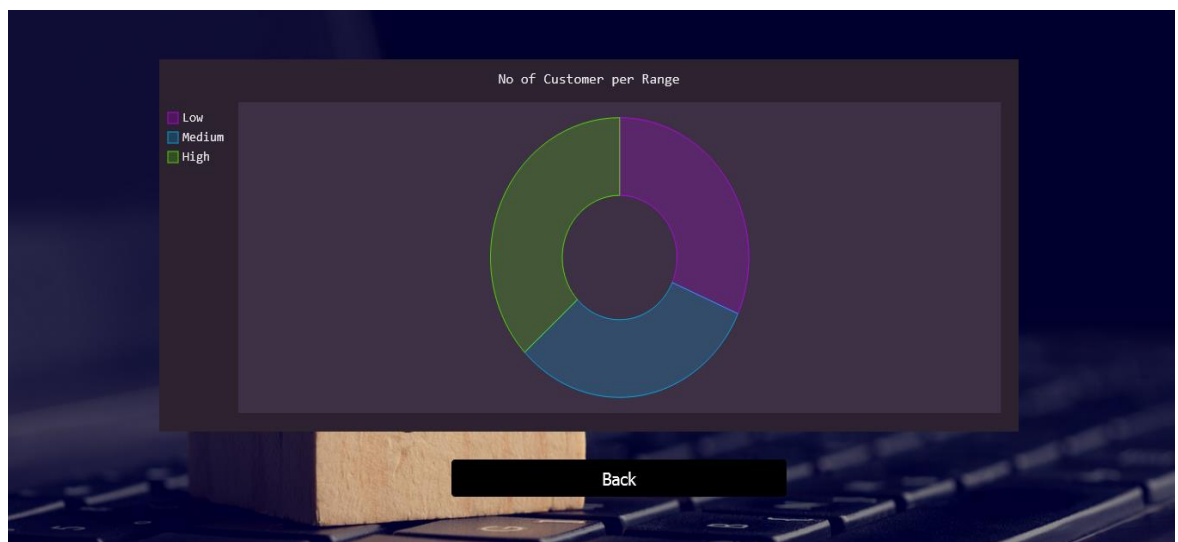






Customer ID	1512
Average Session Time	1349.6666666666667
Average Product view	3.0
Average Transaction Amount	5440.0
Cart Abond rate	0.3333333333333333
Hit rate	0.4545454545454545
Repeat Rate	0.6666666666666666
Customer Life Span	99.0
Customer Life Value	359040.0
Prediction	1
Churn Probability	low

Back



9.1 CONCLUSION

Report generated by system will be transferred to companies which will help them to analyze the behavior of customers. The system tells whether the customer is loyal or is about to churn. Companies will have exact statistics about activities occurring on any product by any customer which will help them to take correct and strategic decisions. It will give the churn probability of the happy and unhappy customers. Speed of decision making process can be increased which will result in better management.

9.2 APPLICATIONS

- **Banks**

In banking and retail sector, customer behaviour predictions can be used to identify the customers which are most likely to defect before they end their relationship. It helps to keep right customer longer and predicts actions which will help to earn their loyalty. Applying this technology would also help discover the churn patterns and develop profiles of users who have left, to get an insight of why they left and discover strategies to keep them satisfied. Predictive analytics allows banks and financial firms to keep up their relationship with the customers by giving them right services and products for their need and matching individual preferences in the most sorted way.

- **For companies of Furniture or House Buyers**

In companies of Furniture or House Buyers like Furlenco and magicbricks, the system can be used to find out the profitable and non-profitable products on website as well as analysis on happy customers and unhappy customers is also possible which will help companies to find out reasons behind non-profitable products.

- **Entertainment**

The system will help in series like Netflix, Amazon Prime to analyze the activities of customer so they can recommend them other sources of entertainment according to their previous choices.

- **Telephone Services Providers**

Telephone Services providers use a variety of different metrics to determine when customers are about to churn, or leave. It is profitable for companies to explore the reasons why customers are leaving, and then target at-risk customers with enticing offers. Providing clients with an easy way to get questions answered and issues handled is the key to maintaining cellular clients.

9.3 FUTURE SCOPE

- Using Neuro-fuzzy model to calculate churn probability more accurately .
- Tracking cursor movements.
- Recommending products and services to customers according to Churn probability.
- Recommending products and services to customers by analyzing their web history.

APPENDIX A: GLOSSARY

Name	Description
Customer Churn	Customer ceases his or her relationship with a company
Churn prediction	detecting customers who are likely to cancel a subscription to a service
Churn probability	Probability of customer to churn
Hit rate	Ratio of number of products bought to the total number of product views
Cart abandonment	1/0 1 - if customer leaves his cart non-empty 0 -if customer buys all items in his cart
Session time	Time elapsed between login and logout from browser
Repeat Purchase Rate	Ratio of number of sessions included transaction to the total number of sessions
Customer Life Span	Time for which customer was active using the service
Customer Lifetime Value (CLV)	prediction of the net profit attributed to the entire future relationship with a customer
Logistic Regression	Classification technique
Synaptic Weights	the “strength” of the input in determining the output

APPENDIX B: PLAGIARISM REPORT

RESULTS

100% Completed: 100% Checked 0% Plagiarism 100% Unique



Sentence Wise Result



Matched Sources



Document View

Unique	The project will be basically a website designed using html, c...
Unique	The two main modules of User interface will be Customer (Cli...
Unique	- User can create account by providing required details like n...
Unique	- User must log in by providing unique credentials (username,...
Unique	A view containing different categories of products and corres...
Unique	- Different sub- categories are displayed using dropdown list.

100% Completed: 100% Checked 0% Plagiarism 100% Unique



Sentence Wise Result

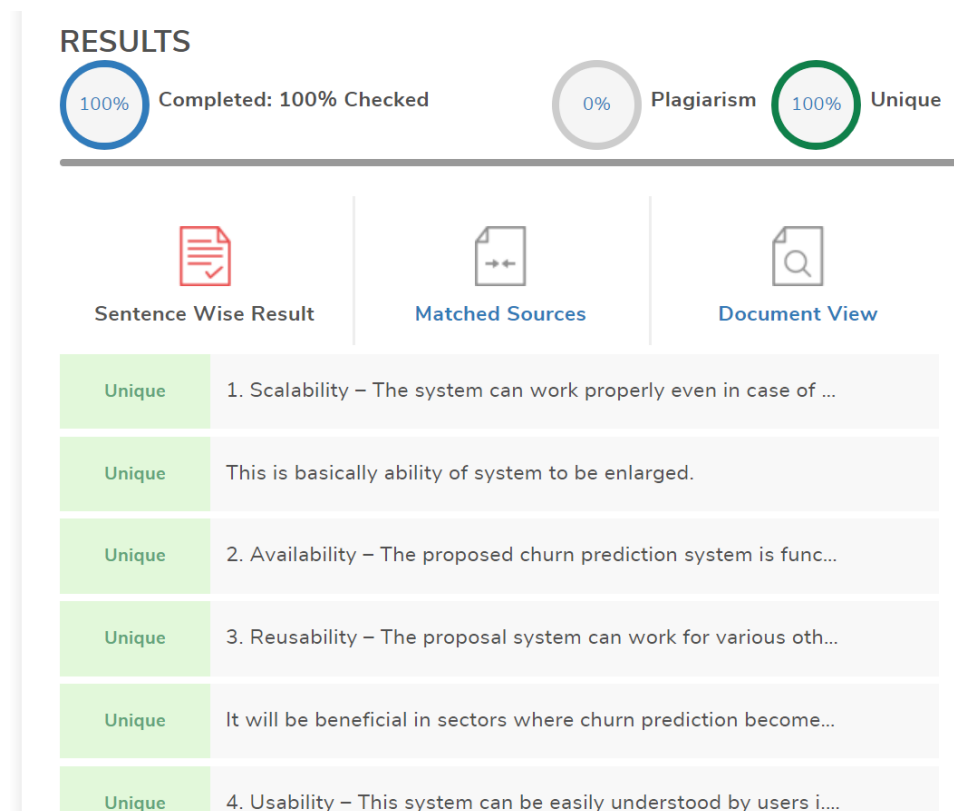
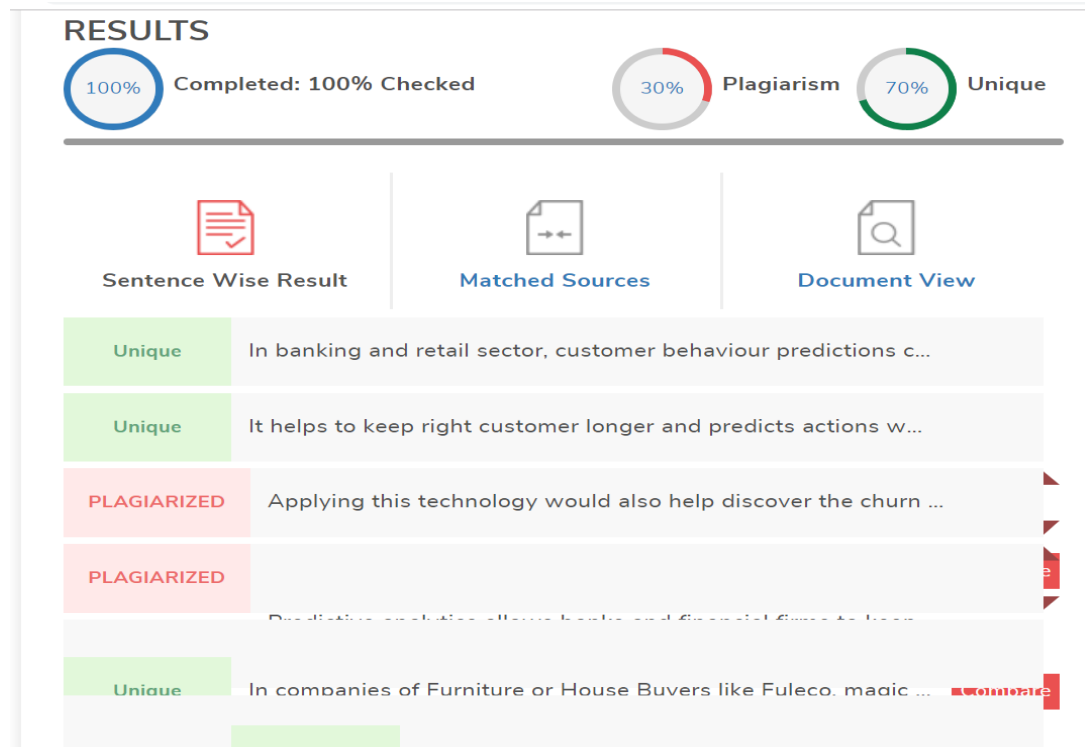


Matched Sources



Document View

Unique	- User should create an account by providing required details ...
Unique	- User should have an unique account to proceed with log in ...
Unique	- Customer can view different products and either can buy a ...
Unique	- Details of product viewed by each customer are recorded fo...
Unique	- Details of products added to cart are maintained to calculat...
Unique	- Details of products bought by each customer like transactio...



REFERENCES

- Churn Analysis: Predicting Churners Navid Forhad, Md. Shahriar Hussain, Rashedur M Rahman Department of Electrical and Computer Engineering, North South University Plot-15, Block-B, Bashundhara, Dhaka 1229,
- T. Mutanen, Customer churn analysis – a case study, Technical Report, Retrieved from, http://www.vtt.fi/inf/julkaisut/muut/2006/customer_churn_case_study.pdf, April 12, 2014.
- Predicting Customer Behavior in Online Shopping Using SVM Classifier Dr.K.Maheswari, P.Packia Amutha Priya ,Department of Computer Applications Kalasalingam University, Anand Nagar, Krishnankoil. Tamil Nadu, India.
- Research on E-commerce User Churn Prediction Based on Logistic Regression Qiu Yanfang, Li Chen ,School of Information Management Beijing Information Science and Technology University Beijing, China
- <https://ieeexplore.ieee.org>
- <https://www.pointillist.com>

