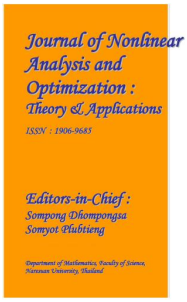
Journal of **N**onlinear **A**nalysis and **O**ptimization
Vol. 15, Issue. 1, No.15 : 2024
ISSN : **1906-9685**



ONLINE USER PERFORMANCE THROUGH GRAPHICAL PATTERN

Sandhi Manjusri (20641A6759) B. Tech Student, CSE (Data Science), Vaagdevi College Of Engineering

Banothu Mounika (20641A6704) B. Tech Student, CSE (Data Science), Vaagdevi College Of Engineering

Durnala Pranay Kumar (20641A6723) B. Tech Student, CSE (Data Science), Vaagdevi College Of Engineering

Yakara Siddartha (20641A6766) B. Tech Student, CSE (Data Science), Vaagdevi College Of Engineering

Mr. Sayyed Hasanoddin Assistant Professor, CSE (Data Science), Vaagdevi College Of Engineering

ABSTRACT

Online shopping has experienced significant growth, with people opting to purchase products via the internet, enjoying the convenience of doorstep delivery. This trend has greatly simplified and expedited people's lives. As online shopping continues to surge, a vast amount of data on users' online activities becomes available on the web. Leveraging such data presents numerous application possibilities, including extracting user behavior and online customers classification from web data. In this study, we propose a system designed to extract and analyze users' online shopping behavior, presenting it in a graphical format. This graphical representation aids administrators in the decision-making process. The system utilizes a graphical hidden state model based on statistical features and integrates various information sources to simulate the

decision-making process effectively. Our proposed system demonstrates significant improvements, achieving nearly a 30% enhancement on million-click datasets. The system is implemented as an online web application, displaying numerous products on web pages for users to view and purchase. Sequential user behavior patterns are carefully tracked by the system and presented in graphical format, providing valuable insights during the decision-making process. The graphical representation allows the admin to identify frequently purchased products and understand current market demands, empowering them to make data-driven decisions. The ease of viewing data in graphical format enables faster decision-making and facilitates prompt and effective solutions.

1. INTRODUCTION

Generally, data mining (sometimes called data or knowledge discovery) is the process of analyzing data from different perspectives and summarizing it into useful information -information that can be used to increase revenue, cuts costs, or both. Data mining software is one of a number of analytical tools for analyzing data. It allows users to analyze data from many different dimensions or angles, categorize it, and summarize there patronships identified. Technically, data mining is the process of finding correlations or patterns among dozens of fields in large relational databases.

How Data Mining Works?

While large-scale information technology has been evolving separate transaction and analytical systems, data mining provides the link between the two. Data mining software analyzes relationships and patterns in stored transaction data based on open-ended user queries. Several types of analytical software are available: statistical, machine learning, and neural networks. **Generally, any of four types of relationships are sought:**

- **Classes**: Stored data is used to located at impede terminal groups. For example, are staurantchaincouldminecustomerpurchasedatatodeterminewhencustomers visit and what they typically order. This information could be used to increase traffic by having daily specials.

- **Clusters**: Data items are grouped according to logical relationships or consumer preferences. For example, data can be mined to identify market segments or consumer

affinities.

- **Associations**: Data can be mined to identify associations. The beer-diaper example is an example of associative mining.

- **Sequential patterns**: Data is mined to anticipate behavior patterns and trends. For example, an outdoor equipment retailer could predict the likelihood of a backpack being purchased based on a consumer's purchase of sleeping bags and hiking shoes.

Datamining consists of five major elements:

1) Extract, transform, and load transaction data onto the data warehouse system.
2) Store and manage the data in a multidimensional data base system.
3) Providedataaccesstobusinessanalystsandinformationtechnologyprofessionals.
4) Analyze the data by application software.
5) Present the data in a use full format, such as a graph or table.

Different levels of analysis are available:

- **Artificial neural networks**: Non-linear predictive models that learn through train in grand resemble biological neural networks in structure.

- **Genetic algorithms**: Optimization techniques that use process such as genetic combination, mutation, and natural selection in a design based on the concepts of natural evolution.

- **Decision trees**: Tree-shaped structures that represent sets of decisions. These decisions generate rules for the classification of a dataset. Specific decision tree methods include Classification and Regression Trees (CART) and Chi Square Automatic Interaction Detection (CHAID). CART and CHAID are decision tree techniques used for classification of a dataset. They provide a set of rules that you can apply to a new (unclassified) dataset to predict which records will have a given outcome. CARTsegmentsadatasetbycreating2-waysplitswhileCHAID segments using chi-square tests to create multi-way splits. CART typically requires less data preparation than CHAID.

- **Nearest neighbor method**: A technique that classifies each record in a data set based on a combination of the classes of the *k* record(s) most similar to it in a historical dataset (where *k*=1). Sometimes called the *k*-nearest neighbor technique.

- **Rule induction** : The extraction of use fulfil- the rules from data based on statistical significance .
- **Data visualization**: The visual interpretation of complex relationships in multidimensional data. Graphics tools are used to illustrated a relationship.

Characteristics of Data Mining:

- **Large quantities of data**: The volume of data so great it has to be analyzed by automated techniques e.g. satellite information, credit card transactions etc.
    - Noisy, in complete data: Imprecise data is the characteristic of all data collection.
    - Complex data structure: conventional statistical analysis not possible

Heterogeneous data stored in legacy systems.

Benefits of Data Mining:

It's one of the most effective services that are available today. With the help of data mining, one can discover precious information about the customers and their behavior for a specific set of products and evaluate and analyze, store, mine and load data related to them

1. An analytical CRM model and strategic business-related decisions can be made with the help of data mining as it helps in providing a complete synopsis of customers
2. Anendlessnumberoforganizationshaveinstalleddataminingprojectsandithashelped them see their own companies make an unprecedented improvement in their marketing strategies (Campaigns)
3. Data mining is generally used by organizations with a solid customer focus. For its flexible nature as far as applicability is concerned is being used vehemently in applications to foresee crucial data including industry analysis and consumer buying behaviors
4. Fastpacedandpromptaccesstodataalongwitheconomicprocessingtechniqueshavem adedata mining one of the most suitable services that a company seek

Advantages of Data Mining:

1.Marketing/Retail:

Data mining helps marketing companies build models based on historical data to predict who will respond to the new marketing campaigns such as direct mail, online marketing campaign…etc. Through there sults, marketer will have appropriate approach to sell profitable products to targeted customers.

Datamining brings a lot to benefit store tail companies in the same way as marketing. Through market basket analysis, as to recap have an appropriate production arrange a way that customers can buy frequent buying products together with pleasant. In addition, it also helps the retail companies offer certain discounts for particular products that will attract more customers.

2.Finance/Banking

Dataminin ggivesfinancialinstitutionsinformationaboutloaninformationandcreditreport ing. By building a model from historical customer's data, the bank and financial institution can determine good and bad loans. In addition, data mining helps banks detect fraudulent credit card transactions to protect credit card's owner.

3.Manufacturing

By applying data mining in operational engineering data, manufacturers can detectfaultyequipmentsanddetermineoptimalcontrolparameters.Forexamplesemi-conductor manufacturers has a challenge that even the conditions of manufacturingenvironmentsatdifferentwaferproductionplantsaresimilar,thequalityofwaferarelo tthesameandsomeforunknownreasonsevenhasdefects.Datamininghasbeenapplyingtodeterminet herangesofcontrolparametersthatleadtotheproductionofgoldenwafer.Then those optimal control parameters are used to manufacture wafers with desired quality.

2. LITERATURE SURVEY

A methodology to detect temporal regularities in user Behavior for Anomaly Detection.

Network security, and intrusion detection in particular, represents an area of increased interest insecurity community over last several years. However, the majority of work in this area has been concentrated up on implementation no misuse detection systems for intrusion patterns monitoring among network traffic. In anomaly detection the classification was mainly based on statistical or sequential analysis of data often neglecting temporal events 'information as well as existing relations between them. In this paper we consider an anomaly detection problem as one of classification of user behavior in terms of incoming multiple discrete sequences. We present an approach that allows creating and maintaining user behavior profiles relying not only on sequential information but taking into account temporal features, such as events' lengths and possible relations between them. We define a user profile as a number of predefined classes of action switch accumulated temporal statistics for every class, and matrix of possible relations between classes

Target Vue : Visual analysis of an omalous user behaviors in Online Communication Systems

Users with anomalous behaviors in online communication systems (e.g. email and social media platforms) are potential threats to society. Automated anomaly detection based on advanced machine learning techniques has been developed to combat this issue; challenges remain, though, due to the difficulty of obtaining proper ground truth form odal training and evaluation. Therefore, substantialhumanjudgmentontheautomatedanalysisresultsisoftenrequiredtobetteradjustthe performance of anomaly detection. Unfortunately, techniques that allow users to understand the analysis results more efficiently, to make a confident judgmental bout anomaly, and to explore data in their context, are still asking. In this paper, we propose a novel visual analysis system, Target Vue, which detects anomaloususersviaanunsupervisedlearningmodelandvisualizesthebehaviorsofsuspicioususers in behavior-rich context through novel visualization designs and multiple coordinated contextual views. Particularly, Target Vue incorporates three new ego-centric glyphs to visually summarize a user's behaviors which effectively present the user's communication activities, features, and social interactions. An efficient layout method is proposed to place these glyphs on a triangle grid, which captures similarities among users and facilitates comparisons of

behaviors of different users. We demonstrate the power of Target Vue through its application in a social bot detection challenge using Twitter data, a case study based one mail records, and anointer view with expert users. Our evaluation shows that Target Vue is beneficial to the detection of users with anomalous communication behavior.

Mining Social Networks for Anomalies: Methods and Challenges

Online social networks have received a dramatic increase of interest in the lastdecadeduetothegrowthofInternetandWeb2.0.Theyareamongthemostpopularsites on the Internet that are being used in almost all areas of life includingeducation,medical,entertainment,business,andtelemarketing.Unfortunately,they have become primary targets for malicious users who attempt to performillegalactivitiesandcauseharmtootherusers.Theunusualbehaviorofsuchuserscan be identified by using anomaly detection techniques. Anomaly detection insocialnetworksreferstotheproblemofidentifyingthestrangeandunexpected behavior of users by exploring the patterns hidden in the networks, as the patterns of interaction of such users deviate significantly from the normal users of the networks. Even though multitude of anomaly detection methods have been developed for different problem settings, this field is still relatively young and rapidly growing. Hence, there is a growing need for an organized study of the work done in the area of anomaly detect if on in social networks. In this paper, we provide a comprehensive review of a large set of methods for mining social networks for anomalies by providing a multi-level taxonomy to categorize the existing techniques based on the nature of input network, the type of an omalies they detect, and the underlying anomaly detection approach. In addition, this paper highlights the various application scenarios where the methods have been used, and explores the research challenges and open issues in this field

3.PROBLEM STATEMENT

Coefficient of reliability is calculated at each step of classification. During classification a coefficient of reliability is changed. Based on this, normal or anomalous user behavior is identified. While classifying the user behavior the system monitors deviations between expected user behavior and current one. Coefficient of reliability [5] is used to estimate the value of deviation in user behavior. If this value crosses a certain threshold, then it is

considered as a case of abnormal behavior. That means the parameters in actions of user are not in admissible intervals. Coefficient of reliability is measuring the same individual twice and it correlates the 2 sets of measures. Every user action class was characterized by statistic parameters of time distribution –mean and standard deviation. Deviations from current values of sequential and temporal parameters are considered as consequence of abnormal behavior [7]. The tools which are used to classify user behavior are N action classes and a relational matrix. These tools describe the model of user behavior.

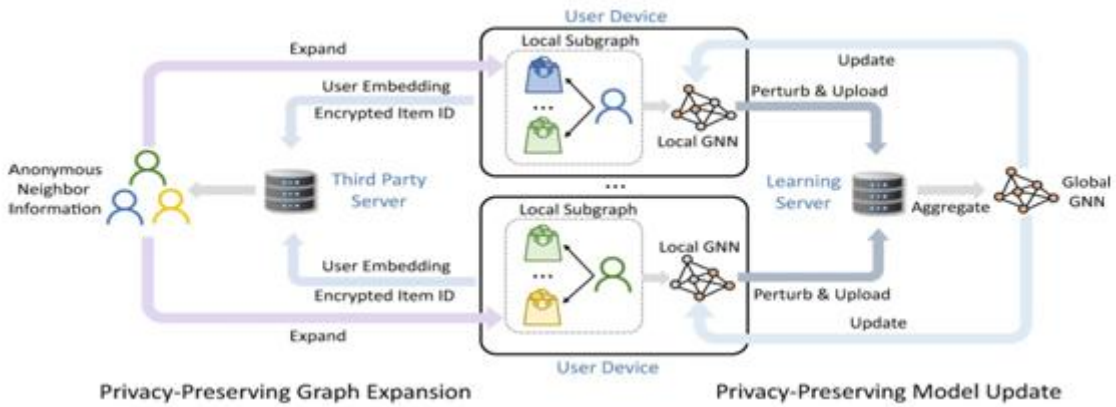
3.1 LIMITATIONS OF SYSTEM

Users who don't have internet connection can't access the system. Systemdoesnotdisplayproductstotheuserwhicharepurchasedfrequently

4.PROPOSED SYSTEM

In this work, we propose to overcome the limitations of prior works in user preference modeling by exploring local and global user behavior patterns on a user successive behavior graph (SBG), which is constructed by utilizing short-term actions of all users. We then exploit high-order relations in the SBG [6] to capture implicit collaborative patterns and preference signals with an efficient jumping graph convolution and learn enriched product representations for user preference modeling. Our approach addresses the a for mentioned problems in the following two aspects.

5.SYSTEM ARCHITECTURE



6.IMPLEMENTATION

6.1AdminLogin: Admin can access the authorized modules by log in to the system using his credentials.

- Add Products: Admin can add products by entering product detail slike product image, cost, and description.
- View User: Admin can view registered user details.
- View Products: Admin can view products which added into the database.
- View User Behavior in graphical format: User behavior pattern isfetchedbyapplyingeffectivealgorithmandisdisplayedingraphicalformat.
- View Feedback: Admin can view feedback of the user.

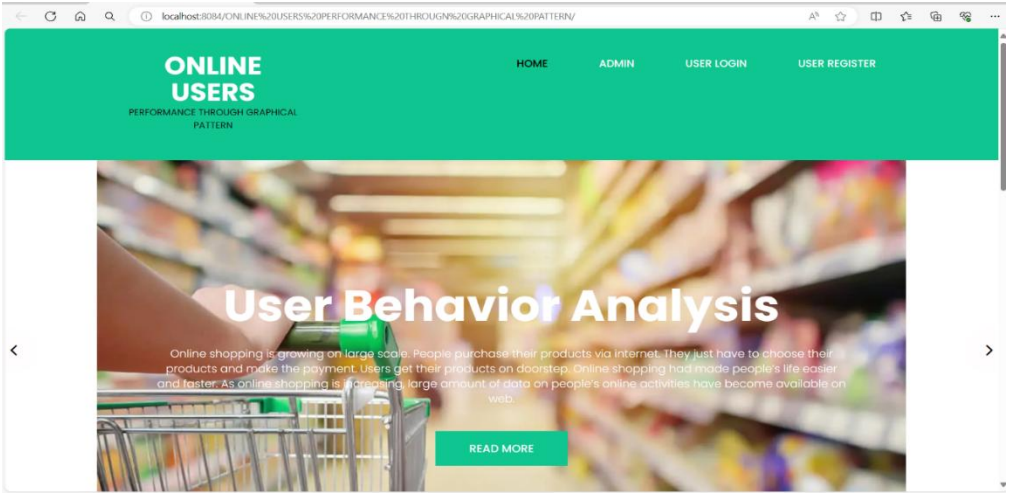
6.2 User Login: User must register with his details and system will provide him with id and password. He must use this user id and password to login to the system.

- View Products: User can view the products and their cost.
- Products Details: User must select the product of his choice and view further details of that product.
- Add To Cart: User can add products into cart, if he wants to purchase the product.

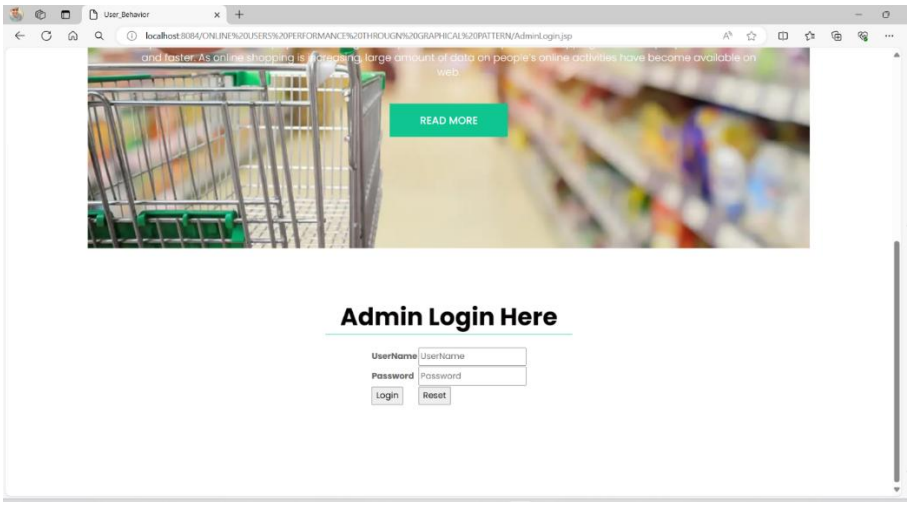
- Make payment: System will display total cost. User can make payment by selecting the model of payment.
- Add Feedback: Customer can add feedback about any product

7.OUTPUT EXPERIMENTS

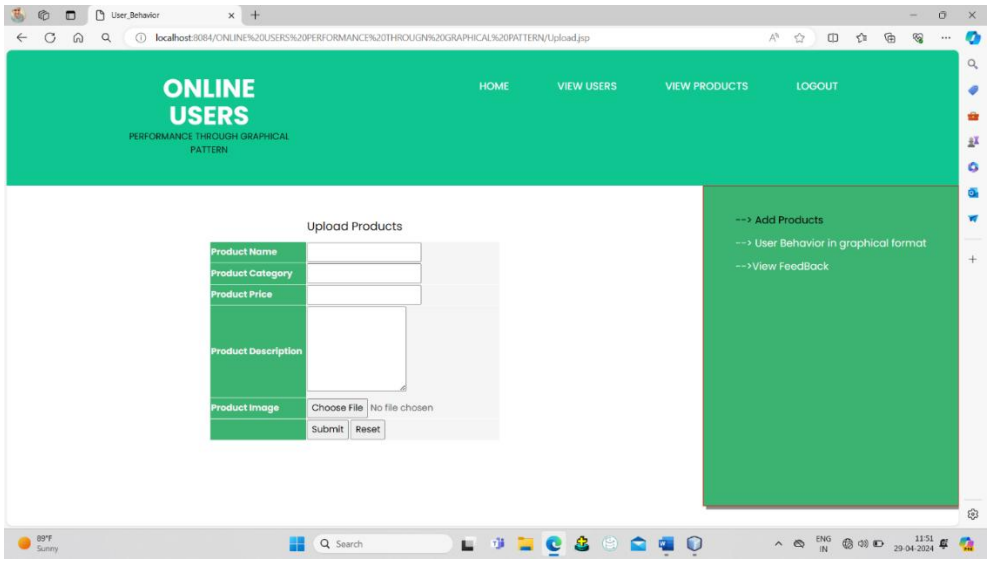
Home Screen



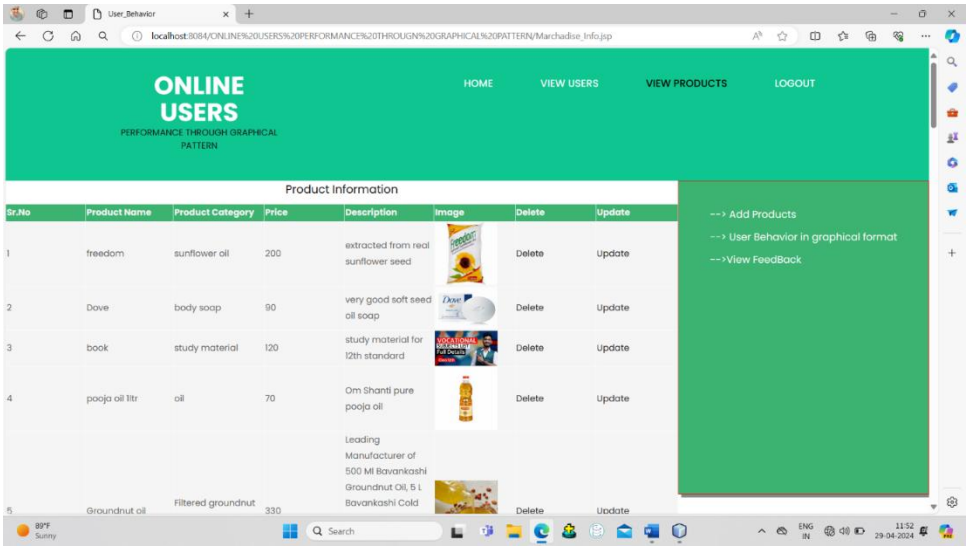
Admin Login



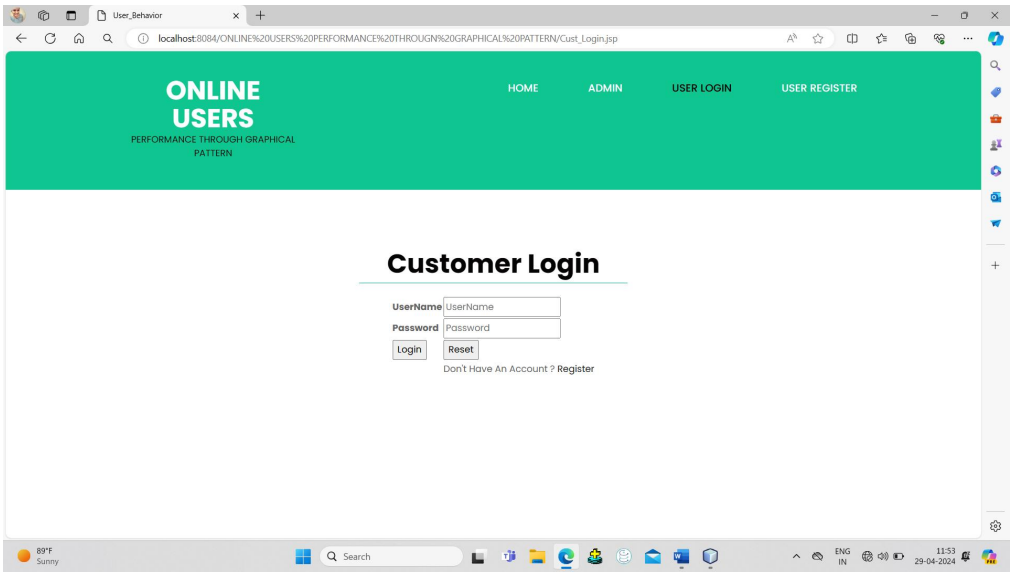
Add Products



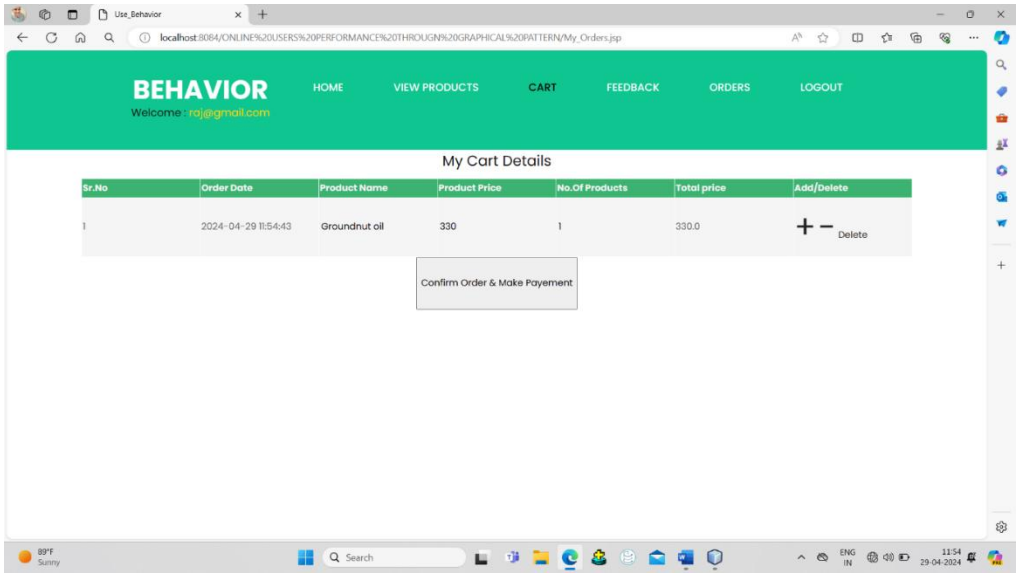
View products



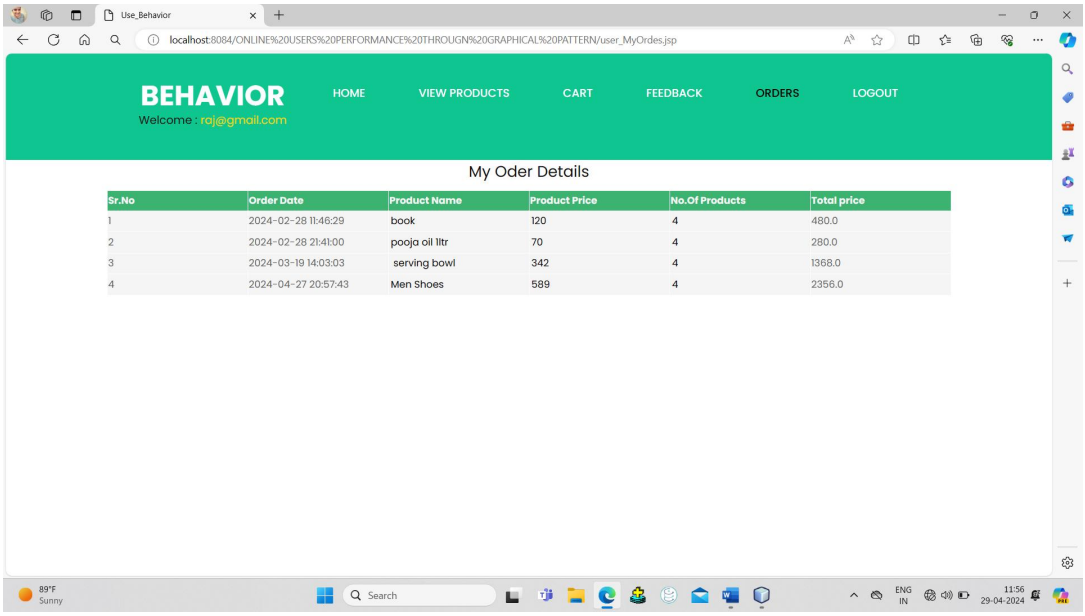
Customer Login


ONLINE USERS
PERFORMANCE THROUGH GRAPHICAL PATTERN
HOME
ADMIN
USER LOGIN
USER REGISTER
Customer Login
UserName
Password
Login
Reset
Don't Have An Account ? Register

View cart


BEHAVIOR
Welcome : raj@gmail.com
HOME
VIEW PRODUCTS
CART
FEEDBACK
ORDERS
LOGOUT
My Cart Details
Sr.No
Order Date
Product Name
Product Price
No.Of Products
Total price
Add/Delete
2024-04-29 11:54:43
Groundnut oil
330
1
330.0
Delete
Confirm Order & Make Payement

Orders



8. CONCLUSION

This analytical study discusses different characteristics of online user behavior models by using various methods and algorithms. Various domains are considered while detecting normal and anomalous online user behavior. The characteristics of Intrusion detection system (IDS)is studied which was used to classify user behavior. Target Vue is the Visualization system studied with different communication features of user behavior. The feature extraction technique like PCA is used to detect anomaly in user behavior. REP Tree is found as the best classification algorithm which is studied with different parameters. Different neural network algorithms studied comparatively which are used for accurate sales prediction task. Different Metrics are studied which were used to analyze users' behavior on 2 OSN platforms. How search engine's service can possible to improve is studied by analyzing users' queries.

9. REFERENCES

[1] Alexandr Seleznyov, Finland,“ A methodology to detect temporal regularities in user behavior for anomaly detection”, Network Security and Intrusion Detection, part 9, pp. 339-352, 2016.

[2] Nan Cao, Cong lei Shi, Sabrina Lin, Jie Lu, Yu-Ru Lin, Ching-Yung Lin, “Target Vue: Visual Analysis of Anomalous User Behaviors in Online Communication Systems”, IEEE Transactions on Visualization and Computer Graphics, vol. 22, pp. 280-289,1 January 2016.

[3] P V Bindu, P Santhi Thilanga, India, “Mining Social Networks for Anomalies: Methods and Challenges”, Journal of Network and Computer Applications, Elsevier, pp. 1-22, 25 Feb2016.

[4] Meng Bi, Jian Xu, Mo Wang, Fuca Zhou, “Anomaly detection model of user behavior based on principal component analysis”, Springer, 21January 2016.

[5] Guirong Chen, Ning Wang, Fenian Zhang, Hua Jiang, China, “Understanding the Time Characteristic of User Behavior on Online Forums”, IEEE International Conference on Big Data(Big Data), 978-1-4799-9926-2/15,pp. 2300-2306, 2015.

[6] Francesco Buccafurri, Gianluca Lax, Serena Nicolazzo, Antonino Nocera, “Comparing Twitter and Facebook user behavior: Privacy and other aspects”, Computers in Human Behavior, Elsevier, 0747-5632, pp. 87-95, 10 June 2015.

[7] Mona Gupta , Happy Mittal , Parag Singla , Amitabha Bagchi, “Characterizing comparison shopping behavior: A case study”, ICDE Workshops, IEEE, 978-1-4799-3481-2/14, pp. 115-122, 2014.

[8] Ke XIE, Huijia YU, Rongwei CEN, “Using log mining to analyze user behavior on search engine”, Front. Electr. Electron. Eng., Higher Education Press and Springer-Verlag Berlin Heidelberg 2011, pp. 254-260, 2011.