International Journal of Scientific Research in Engineering and Management (IJSREM)
Volume: 08 Issue: 12 | Dec - 2024 SJIF Rating: 8.448 ISSN: 2582-3930

The Use of Data Mining Procedures in Understanding Customer Inclinations Within the Social and Entertainment Industry

Niranjana S R [1]

[1]Computer Science Department, Sree Chitra Thirunal College of Engineering

---***---

Abstract - Different techniques and methods have been established that have analysed consumers in shopping patterns and their usability and the real-world solutions required by them. Such information-mining techniques and analyses provide useful insights to organizations that can be really profitable while mining databases; though different techniques have some merit but also some disadvantages when their limitations are considered.

Television dominates the entertainment landscape, with service providers always offering diverse programs to suit the tastes of different customers. The choices of viewers vary according to locality and season, and service providers usually bundle channels into pre-defined packages. Data mining techniques can be used to analyse customer preferences, thereby creating new, customized packages or groups of channels that suit individual needs.

This paper explores consumer behaviour, focusing on the psychological factors that influence purchasing decisions and demonstrates how data mining methods can improve traditional approaches. Using an experiment based on association rule mining, the study derives rules for identifying trusted customers from sales data in the supermarket industry.

***Key Words*:** Consumer behaviour. Data mining, Association rules, Television, Supermarket

1.INTRODUCTION

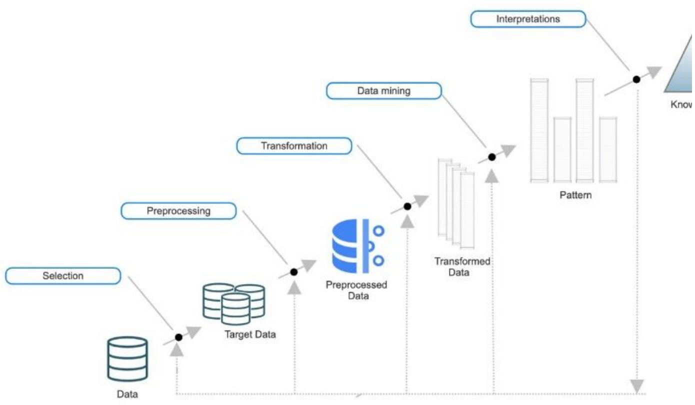
Social media excitement stages have gotten to be a key portion of our day by day lives, facilitating assorted substance extending from moving and comedy to singing and reality appears. The notoriety of social media substance has been on the rise within the past few a long time, taking the consideration of clients who are energetically devouring recordings, reels, and scenes shared over stages. The victory and viewership of such substance are measured through different apparatuses and measurements, counting engagement rates, likes, offers, and sees. These measurements are fundamental in deciding which programs and makers are most well-known and what gatherings of people incline toward and how far-reaching particular channels or pages are.

Analytics instruments implanted inside stages collect information on the time went through, the sort of substance seen, and client interaction. Examining this information over a period of time makes a difference to distinguish patterns and designs, which empowers comparison and contrasts between distinctive makers, pages, or substance sorts. Calculations such as K-means and incremental K-means advance improve the examination with information clustering that makes a difference the distinguishing proof of patterns in client engagement.

Similarly, understanding shopper conduct has gotten to be fundamental in numerous businesses, counting social media, to plan suitable promoting techniques. Customer conduct considers are centred on the mental and natural variables that impact clients to lock in with particular substance or make buys through a social media stage. It makes a difference get it how individuals think, feel, and make choices between options and how their environment and inspiration shape their choices. This understanding permits organizations to fine-tune promoting campaigns and procedures based on client inclinations, driving to more prominent engagement and fulfilment. Combining client engagement analytics with customer conduct bits of knowledge, stages and businesses can advance optimize their procedures to meet group of onlookers interface and maximize substance success.

Data mining: It is the method of getting important experiences and designs from huge datasets. It is sorting through parcels of information to set up patterns, make connections, and unravel issues with information examination. Typically basic for organizations to anticipate future patterns and make choices based on data. **Fig-1** : Data Mining

International Journal of Scientific Research in Engineering and Management (IJSREM)
Volume: 08 Issue: 12 | Dec - 2024 SJIF Rating: 8.448 ISSN: 2582-3930



Data mining is, at its centre, the disclosure of information from information, not the extraction of crude information itself. It is an intrigue field that combines methods from machine learning, insights, and database systems. The ultimate objective is to convert crude information into significant information that can be effortlessly caught on and utilized effectively.

We have to be extricate data isn't the as it were handle; DM too includes other forms like Information Cleaning, Information Integration, Information Change, Information Mining, Design Assessment, and Information Introduction. After completing all of these forms, we seem utilize this data in a few applications such as Extortion Location, Advertise Investigation, Generation Control, Science Investigation, etc.

2. Data Mining and Consumer Behaviour

2.1 Applications of Data Mining

2.1.1 Data Mining in Finance

Data mining is broadly utilized within the money related division to upgrade client fulfilment and dependability. By examining client conduct, money related teach can foresee patterns and customize their offerings, such as personalized credits, protections bundles, and venture plans. It reveals covered up designs in budgetary exchanges, making a difference recognize relationships that show dangers or openings. Besides, information mining makes a difference to identify extortion through chronicled information investigation of inconsistencies such as designs of odd exchanges or unauthorized exercises in accounts in guaranteeing money related security and hazard management.

2.1.2 Market Basket Analysis

Market basket analysis (MBA) may be a common method connected in retail and e-commerce businesses. MBA makes a difference within the understanding of item affiliations and the buy design of businesses. By dissecting which products are regularly bought together, businesses can create combo offers, bundle bargains, and focused on advancements to extend deals. For occasion, by offering coffee and mugs together or advertising extras at a rebate with electronic products, the fulfilment of the client increments and income develops. MBA too makes a difference optimize rack courses of action and store formats to create clients purchase more. These bits of knowledge offer assistance businesses deliberately choose on overseeing their inventories and supply chains.

2.1.3 Marketing and Retail

In showcasing, information mining is utilized to create personalized strategies by considering client socioeconomics, obtaining conduct, and inclinations. Businesses can in this way shape their showcasing campaigns, maximizing reach and viability. They can spot the leading clients for them and draw in modern ones by having focused on advancements, dependability programs, and proposals. Information mining too estimates patterns, in this way permitting retailers to stock the proper items at the correct time and subsequently diminishing overstock and stockouts. The company makes strides the item arrangement by knowing client inclinations and plans way better advertisements.

2.1.4 Customer Segmentation

Data mining helps client division. Clients are categorized into diverse bunches depending on their characteristics such as buying propensities, inclinations, or investing propensities. By clustering, organizations might target specific fragments such as the client with tall values, moo clients, or potential modern clients. Take for case the case where telecom companies may bunch the same paid ahead of time clients based on sum of revive or indeed minutes of calls and come to offer extraordinary bargains or offer them extraordinary offers, all this makes a difference move forward the hone of CRM and eventually contributes more to profitability.

2.1.5 Telecommunication

Telecommunication firms have a solid reliance on information mining to bargain with gigantic client information and optimize their administrations. Information mining makes a difference firms identify false action, like unauthorized get to or abuse of administrations, thus keeping up believe with clients and their security. It makes a difference examine client churn rates, distinguishing why clients are taking off and coming up with procedures to keep them. Moreover, information mining permits telecom suppliers to get it client inclinations, for case, crest calling times or information utilization designs, and tailor their offerings to those inclinations. The experiences offer assistance suppliers recognize items and administrations that are most productive, guaranteeing superior asset assignment and progressed commerce results.

2.2 Consumer Behaviour

It alludes to the ponder of how people, bunches, or organizations select, buy, utilize, and arrange of items, administrations, or thoughts to fulfill their needs. This consider too looks at the impact of these forms on both shoppers and society.The perspective of buyer conduct is in understanding impacts at the smaller scale, bunch, or the level of organizations that impact their decisions.Peer impact decides clothing choice. The collective firm choice in choosing which item to present may well be seen as receiving certain items in it.By deciding customer conduct, a advertiser can arrange his/her item with regard to understanding its applications in encouraging utilization.

2.3 Applications of Consumer Behaviour Public Policy

2.3.1 Marketing strategy

Shopper conduct experiences frame a spine for compelling marketing.

For occurrence, knowing that kids are more touchy to promotions almost nourishment items amid cartoons that discuss after school gives marketers an opportunity to discuss nibble commercials amid late afternoon.

2.3.2 Public Policy:

For case, when the government realized there are wellbeing suggestions of cigarette utilization, at that point the FDA will demand on including caution stickers on cigarette bundles to act as a disheartening operator against its usage.

2.3.3 Social Marketing:

This is additionally an approach whereby thoughts spread, not a item. It majorly applies to open wellbeing publicizing and natural assurance campaigns among others.

2.3.4 Consumer Awareness and Cost Sensitivity

Understanding the conduct of clients guarantees that they make sensible choices. For case, mindfulness of cost premiums for bigger bundle sizes, such as a one-litre bottle of ketchup that costs much more than two half-litre bottles, makes buyers pay consideration to unit cost labels to urge the finest esteem for their cash.

2.4 Applications of Data Mining in Consumer Behaviour Analysis

2.4.1. Shopping Habits and Seasonal Patterns

Retailers use data mining to analyze customer shopping patterns during festive seasons or other special events. For example:

- Identifying which products sell the most during holidays.
- Finding correlations between products, such as customers who buy TVs also being likely to purchase home theater systems.
- Analyzing customer behavior and then organizing store layouts accordingly.
- For example, during festive seasons, stores may place high-demand electronic items like TVs and LCDs next to mobile phones, offering attractive bundles in prominent displays. High-demand items are placed together to encourage bundled sales.

2.4.2. Demand Forecasting and Inventory Management

Using data mining, weekly shopping patterns and product preferences can be analyzed to optimize inventory. Retailers can:

- Identify products that need to be stocked regularly.
- Monitor product demand at specific times (e.g., the start vs. end of the month).
- Adjust stock to align with patterns, such as increased spending at the start of the month when customers have more disposable income and reduced spending at the end of the month.
- Similarly, during school holidays or back-to-school seasons, demand for certain products increases.
- Retailers can analyze past trends to maintain optimal stock levels and capitalize on demand during these periods.

2.4.3. Strategic Pricing and Promotions

Retailers can analyze patterns to decide when to offer promotions or discounts. For example:

- Reducing prices during slow days to increase foot traffic.
- Avoiding discounts on high-demand days to maximize profits

2.4.4. Enhanced Customer Satisfaction

By leveraging data insights, stores can personalize the shopping experience for customers. This improves customer satisfaction through strategic product placement and predictive stocking, which reduces the likelihood of out-of-stock items.

3. Literature Review

In the paper *"Market Basket Analysis in Different Store Environments"*, authors Yen-ling Chen, Kwei Tang, Ren-Jie Shen, and Ya-Han Hu identify two major challenges in applying existing methods within multi-store scenarios. The primary challenge is that purchasing patterns are time-dependent, such as regular products. The second challenge involves finding common association patterns across subsets of stores. To address these issues, the authors present an Apriori-like algorithm that automatically extracts association rules in a multi-store environment. [3]

Many studies have been conducted on consumer purchasing behavior and its application to real-world problems. Data mining techniques are considered effective tools for analyzing customer behavior. However, these techniques have both advantages and limitations, requiring careful selection of appropriate tools for database mining. In the paper *"A Data Mining Approach to Consumer Behavior"*, Junzo Watada and Kozo Yamashiro suggest improvements to data mining analysis by utilizing several techniques such as fuzzy clustering, principal component analysis, and discriminant analysis. Their work addresses and advances many of the limitations of traditional methods. [4]

S. Vijaylaxmi, V. Mohan, and S. Suresh Raju explain in their study *"Mining of Users' Access Behavior for Frequent Sequential Pattern from Web Logs"* that sequential pattern mining is a subset of association rule mining. For a given transaction database T, an association rule X→Y holds with confidence τ% of transaction set T if σ% of T supports X∪Y. Association rule mining involves two steps: (1) identifying frequent patterns based on a minimum support threshold, and (2) generating association rules using a minimum confidence threshold. [5]

In the paper *"Mining Utility-Oriented Association Rules"*, Parvinder S. Sandhu, Dalvinder S. Dhaliwal, and S. N. Panda propose an effective approach that takes into account profit and quantity for mining important association rules. The approach starts with the classic Apriori algorithm to generate a set of association rules using its anti-monotone property, which states that for any k-itemset to be frequent, all (k-1) subsets of this itemset must also be frequent. The generated rules are then filtered using weightage (W-gain) and utility (U-gain) constraints. A combined utility-weighted score (UW-Score) is calculated for each rule to identify high-utility rules that are useful for business development. Experimental results validate the approach's effectiveness in generating high-utility association rules. [6]

B. Yıldız and B. Ergenç, in the paper *"Comparison of Two Association Rule Mining Algorithms Without Candidate Generation"*, highlight the role of association rule mining in discovering relationships among itemsets in databases. The Apriori algorithm has significantly advanced data mining research, but suffers from the drawback of candidate generation, requiring multiple passes over the data. The FP-Growth and Lattice Apriori algorithms circumvent this bottleneck by using compact data structures to store frequent itemsets, thus avoiding candidate generation. Their study compares these two algorithms step by step using synthetic datasets. Key findings include: (1) algorithm performance depends on dataset characteristics and threshold values, (2) Lattice Apriori outperforms FP-Growth at threshold values below 10%, and (3) while building the lattice data structure incurs higher costs, finding itemsets is faster. [7]

In the paper *"Extraction of Interesting Association Rules Using Genetic Algorithms"*, Peter P. Wakabi-Waiswa and Venansius Baryamureeba describe the extraction of interesting and unexpected rules from large datasets using association rule mining. Traditional metrics like support and confidence fail to fully capture the interestingness of rules. The study frames association rule mining as a multi-objective problem involving metrics such as predictive accuracy, interpretability, and interestingness. The authors propose a Pareto-based multi-objective evolutionary algorithm, using specific crossover and mutation operators along with elitism to extract valuable rules. Experimental results show high predictive accuracy for the generated rules. [8]

Jyothi Pillai, in *"User-Centric Approach to Itemset Utility Mining in Market Basket Analysis"*, examines how business insights derived from data mining help predict company performance and identify profitable trends. Traditional association rule mining assumes equal significance and frequency for all items, which is not realistic in the real world. Utility mining addresses the limitations of traditional methods by focusing on rare itemsets with high utility, based on factors like cost, profit, and revenue. This paper emphasizes the importance of finding rare, high-utility itemsets and highlights their significance in decision-making areas such as retail, healthcare, and fraud detection. The study also integrates temporal aspects into utility mining, recognizing their importance in complex real-world problems. [9]

Lastly, in *"Efficient Association Rule Mining for Market Basket Analysis"*, Shrivastava A. and Sahu R. explore the value of data mining in making informed business decisions. Association rule mining, particularly through market basket analysis, helps retailers understand customer purchasing tendencies and develop strategies like store layout optimization and promotional planning. Traditional algorithms, such as Apriori and FP-Tree, are widely used but have issues with candidate set generation, especially when dealing with large or long patterns. This paper enhances these algorithms to identify frequent, profitable patterns for market analysts to make better-informed decisions for business growth. [10]

4. Methodology for Analysing Consumer Behaviour

Over the past few decades, data mining (DM) has emerged as a powerful tool for analyzing consumer purchasing behavior through various techniques. It is a field that has evolved from nothing and now constitutes a $100 billion industry. For every action taken by a consumer in a store, a data point is generated. These actions include what they spend, when they shop (specific days and times), what they frequently buy, how much they purchase, and even shopping patterns specific to their location.

All of this data is subtly mined in the background, often without the consumer's awareness. The industry around processing, analyzing, and selling this data to convert it into actionable insights is massive, and the value attached to it is high.

Data mining is the process of analyzing data from different perspectives to extract valuable information that drives business decisions, such as increasing revenue, reducing costs, or achieving both. It allows companies to view data across various dimensions, categorize it, and uncover relationships or patterns within large datasets. Businesses use data mining to understand internal factors such as pricing strategies, product placement, and employee skills, as well as external factors like customer demographics, economic trends, and competition.

Retailers often use data mining techniques to track shoppers' purchasing habits or trends in consumer behavior during specific seasons and events. For example, they aim to identify high-demand items to promote during holidays, assess associations between frequently purchased items, and recognize behavioral patterns that help optimize store layouts. One retailer, for instance, may display electronic devices like TVs and mobile phones during a festive period with attractive offers, while placing complementary products nearby to encourage bundled sales.

Data mining also aids in inventory management and demand forecasting by analyzing weekly shopping trends and product preferences. Retailers can predict which products need to be stocked in higher quantities and anticipate demand fluctuations based on the time of the month. Spending generally increases at the start of the month when consumers have more disposable income, and decreases toward the month's end. Moreover, events like school vacations or back-to-school periods create a surge in demand for specific products, and retailers can plan accordingly.

Strategic pricing and promotions are another application of data mining. Insights derived from data can identify the optimal times for discounts or promotions. Retailers may choose to avoid offering discounts on high-demand days to maximize profits or use promotions strategically to boost foot traffic on slower days.

The insights gained from data mining not only improve operational efficiency but also contribute to a more personalized shopping experience. Effective product placement, inventory management, and timely promotions ensure that customers find what they need, thereby improving customer satisfaction and loyalty.

The data used in these analyses is collected from a variety of sources. Some of the operational or transactional data sources include sales, prices, inventory, finance, and accounting. Non-operational data, such as industry-wide sales figures, forecasts, and macroeconomic trends, are also utilized. Metadata, which refers to information about the structure of the data itself, such as database schemas or data dictionary definitions, is also a critical part of the process.

By employing data mining techniques, businesses can gain deeper insights into consumer behavior, optimize resources, increase profits, and deliver a better customer experience.

5. Classification of Data Mining Systems

Data mining systems can be classified based on several criteria to better understand their functionality, methods, and applications.

5.1 Classification Based on the Type of Knowledge Mined

Data mining systems can be categorized by the type of data they extract. This is often aligned with data mining functions such as classification, regression, association, clustering, exception detection, and evaluation. Comprehensive data mining systems typically offer multiple integrated functionalities to cover a broad range of analytical needs.

The reflection levels of the data in these systems may vary. Some systems operate at a generalized level, handling raw data, while others work at different levels of abstraction, using varying degrees of deliberation. This means that more advanced systems can perform multiple levels of data discovery, making them more functional and informative.

Another classification is based on the type of data mined: regularities (or patterns) versus anomalies (or exceptions). Techniques for concept description, association analysis, classification, prediction, and clustering often focus on discovering regularities, treating exceptions as noise. However, these techniques can also be adapted to discover exceptions, offering a dual-purpose approach.

5.2 Classification on the Basis of Techniques Used

Data mining systems can also be categorized by the techniques they apply. These methods differ in the level of user interaction, such as independent systems, interactively exploratory systems, or query-driven systems.

The methods employed in data analysis help define these systems. Some examples include:

- Database-oriented or data warehouse-oriented techniques
- Machine learning approaches
- Statistical methods
- Visualization tools
- Pattern recognition techniques
- Neural networks

Advanced data mining systems typically incorporate multiple techniques or integrate them into a cohesive approach, maximizing the strengths of individual methods.

5.3 Classification Based on Application Domain

Data mining systems can also be classified by the domain they are applied to. There are data mining systems specifically designed for applications like telecommunications, DNA analysis, stock market prediction, web analytics, email filtering, and more.

Different applications require the integration of domain-specific techniques, as generic, all-purpose data mining systems may not meet the specialized needs of a given task. Customization ensures that the system effectively addresses the unique requirements of a particular domain.

5.4 Classification Based on Client Interaction

Data mining frameworks can moreover be classified by the level of interaction with the client. These categories include:

Autonomous Frameworks: These frameworks work without much client intercession, consequently selecting information sources, preparing them, and creating results.

Interactive Frameworks: Give clients with the capacity to inquiry information and alter parameters to fine-tune the examination. These frameworks are adaptable and permit for more profound investigation of data.

Query-Driven Frameworks: Clients yield particular questions, and the framework recovers and forms information based on the given ask, guaranteeing tall significance to the user's needs.

5.5 Classification Based on Learning Techniques

Data mining frameworks can be recognized by the learning methods they utilize. This classification includes:

Supervised Learning: The framework employments labeled information to memorize a mapping from inputs to yields. It can at that point apply this learning to anticipate the yield for modern, concealed data.

Unsupervised Learning: In this case, the framework recognizes covered up designs in information without predefined names. Common procedures incorporate clustering and affiliation run the show mining.

Reinforcement Learning: The framework learns by connection with its environment, accepting input, and optimizing its activities over time.

6. Research Methodology

To approve and test unused strategies in information mining, the analyst conducted an in-depth think about on buyer conduct by taking an organization, such as a shopping center or general store, as a test. For the reason of the think about, the analyst collected live value-based information every day and month to month premise, in this manner capturing the obtaining exercises of person customers.

After collecting the information, the analyst explored different strategies and procedures for dissecting the information. Among the choices accessible, the foremost appropriate methods, equations, calculations, and strategies were chosen for dissecting the client database. For this inquire about ponder, Margin Free Bazaar was chosen as the research site because it may be a grocery store found in Trivandrum city. Information with respect to the obtaining exchanges of clients was collected and appropriately arranged.

To dissect the information, the analyst chose affiliation run the show mining, utilizing Showcase Wicker container Investigation, to recognize affiliation between products acquired by a client. This strategy is utilized to find designs or connections within the information gotten.

7. Use of Data Mining Techniques

7.1 Association Rule Mining

The calculation of affiliation run the show mining includes significant affiliations or relationships inside enormous collections of information. It is utilized more viably in dissecting information around client exchanges. All the experiences that can be picked up from such examination offer assistance in vital commerce choices; for case, catalogue plan, cross-marketing methodology, and loss-leader examination. Mining affiliation rules aiming to discover covered up connections inside exchange information of the grocery store are point by point here.

7.2 Customer Transactions with the Study Unit

Market Bushel Examination is one of the most applications of association run the show mining. Typically based on the relationship between things put into a shopping wicker container to dissect the obtaining design of clients. Such connections are accommodating in planning particular showcasing methodologies and rack arrangements.

For illustration, on the off chance that information appears that clients who purchase drain are moreover likely to buy bread (and certain sorts of bread), this information can offer assistance retailers position those items in ways that make them more likely to be obtained together. For case, setting drain and bread following to each other within the store can increment drive buys and drive sales.

The analyst chose to utilize Advertise Wicker container Investigation for this think about since it could be a device that's exceptionally effective to uncover co-occurrence designs among categorical information. By examining value-based information, Showcase Bushel Examination provides association rules that donate bits of knowledge straightforwardly from the dataset, making it actionable..

Table -1: Samples Collected

Transaction ID	Items
1	Bread, Butter, Milk
2	Bread, Butter
3	Bread, Milk, Eggs
4	Butter, Eggs
5	Bread, Butter, Eggs
6	Milk, Eggs
7	Bread, Butter, Milk, Eggs

The value-based information applied in this inquire about were sourced from one operational day within the chosen company, and the number of exchanges on that specific day is shown within the taking after dataset

Based on the information over, Analyst determine the taking after yield of affiliation run the show by utilizing advertise Bushel investigation.

Table -2: Calculated Associated Rule Output

People who bought this item	Also bought the following items	Support	Confidence
Bread, Butter	Milk	3/7≈0.43	3/4=0.75
Bread, Butter	Eggs	2/7≈0.29	2/4=0.50
Bread, Milk	Butter	3/7≈0.43	3/3=1.00
Bread, Milk	Eggs	2/7≈0.29	2/3=0.67
Butter, Milk	Bread	3/7≈0.43	3/3=1.00
Butter, Milk	Eggs	2/7≈0.29	2/3=0.67
Milk, Eggs	Bread	2/7≈0.29	2/3=0.67

The association rule has the following form: X→Y, which means that individuals who bought items from set X are also likely to buy items from set Y.

Support and confidence are two key measures for association rules. **Support** refers to the frequency of transactions that include all items in both sets X and Y. For example, a support of 5% indicates that 5% of all transactions (considered by the researcher for the analysis) include both sets X and Y. In equation form, support can be calculated as the probability of the union of sets X and Y.

7.2.1 Association Rule Analysis:

Support:
The recurrence of exchanges that incorporate all the things in X ∪ Y, separated by the overall number of exchanges:

$$\text{Support}(X \rightarrow Y) = \frac{n(X \cup Y)}{N}$$

where:

- n(X∪Y): Number of transactions containing both X and Y.
- N: Total number of transactions

Confidence:
The conditional likelihood of acquiring Y, given X was obtained:

$$\text{Confidence} = \frac{n(X \cup Y)}{n(X)}$$

where:

- n(X): Number of transactions containing X
- n(X∪Y): Number of transactions containing both X and Y

7.2.2 Key Insights:

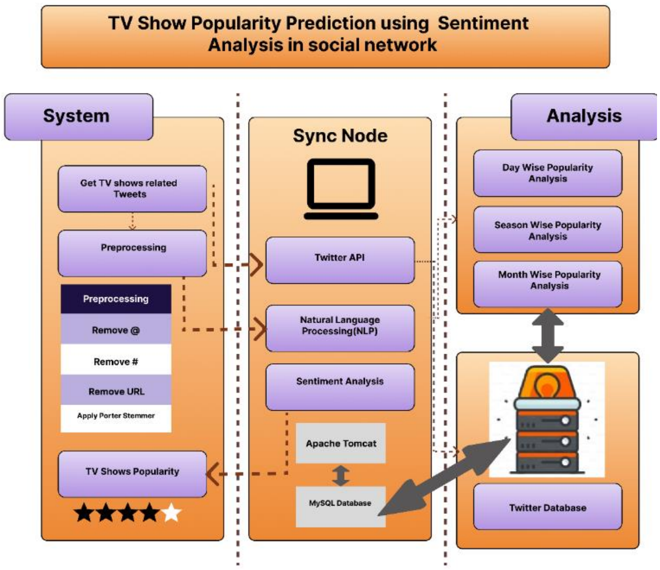
- Bread, Butter → Milk: This rule has a support of 0.43 and a confidence of 0.75, meaning 43% of all transactions include Bread, Butter, and Milk, and 75% of customers who purchase Bread and Butter also buy Milk.
- Milk, Eggs → Bread: This rule has a support of 0.29 and a confidence of 0.67, meaning 29% of all transactions include Milk, Eggs, and Bread, and 67% of customers who purchase Milk and Eggs also buy Bread.

International Journal of Scientific Research in Engineering and Management (IJSREM)
Volume: 08 Issue: 12 | Dec - 2024 SJIF Rating: 8.448 ISSN: 2582-3930

8. Advanced Metrics in Consumer-Centric Marketing

8.1 System used to determine TRP

Fig-2 : Design of a TV Show Popularity Prediction System



The workflow of the system is divided into different steps to predict the popularity of TV shows. In the first step, the system gathers data about the TV shows, including the content reviews related to the specific shows, actors, and directors.

A. Hash tagged Dataset

The system begins by removing duplicate tweets, non-English tweets, and those without relevant hashtags to create a hash-labeled dataset. The remaining tweets are further analyzed for the appearance of frequently occurring hashtags showing positive, negative, or neutral sentiment. These are some of the hashtags on which training and development will be conducted. Here, content mining is applied, along with NLP and IR techniques, which involve essential preprocessing steps like tokenization and normalization. Removal of elements such as @, #, and URLs is included within the step. Preprocessing extracts valuable insights from unstructured text data. Information retrieval ensures relevant reports from a collection are retrieved to meet user requirements effectively.

B. Tokenization

This process involves breaking text into smaller components such as words, symbols, and phrases, known as tokens. It is the primary input for further processes like parsing or data mining. Tokenization splits sentences into individual words, enabling further analysis and creating a foundation from where relevant data can be retrieved in a dataset.

C. Sentiment Analysis

Sentiment analysis is one of the key applications of NLP. It allows data scientists to analyze social media comments and evaluate brand performance concerning customer sentiment. For example, businesses can track customer reviews to identify potential areas for improvement and enhance customer satisfaction. Sentiment analysis reveals valuable insights about audience opinions.

D. Popularity Analysis

Day-wise Popularity Analysis: Popularity is calculated on a daily basis. This can be very useful for live shows or single-day events like cricket matches or award shows. Users can analyze the popularity for specific dates.
Season-wise Popularity Analysis: This analysis calculates the popularity of TV shows across entire seasons. It identifies the types of shows that performed well during specific periods, such as reality shows.
Month-wise Popularity Analysis: This type of analysis evaluates the performance of shows over a month. It helps users track the most-watched programs, such as daily soap operas.

Table -3: Output Table showing TV Shows with Scores

Sr.No	TV Shows	Channel Name	Score
1	Friends	NBC	98542
2	Game of Thrones	HBO	91238
3	The Big Bang Theory	CBS	86745
4	Stranger Things	Netflix	83254
5	Breaking Bad	AMC	78963
6	The Office	NBC	75629
7	Sherlock	BBC One	70418
8	The Crown	Netflix	65842
9	Grey's Anatomy	ABC	60518
10	Suits	USA Network	57849

9. CONCLUSIONS

Data mining systems prove to be effective tools for monitoring customer buying behavior in retail and departmental stores. This study concludes that customers exhibit distinct purchasing patterns. By utilizing these insights, management can streamline their systems and services by offering more of what customers desire, ensuring customer satisfaction and fostering long-term loyalty to the business.

Data mining systems also assist businesses in identifying associations between customers and various products. By analyzing these patterns, businesses can understand how customers switch between brands to meet their needs. This information allows businesses to adjust their offerings based on customer preferences and past purchasing habits.

ACKNOWLEDGEMENT

I would like to express my heartfelt gratitude to all who supported me in the process of completion of this research paper.

First and foremost, I extend my sincere thanks to my mentors and faculty members for their invaluable guidance, encouragement, and constructive feedback in the course of this study. Their expertise and insights have been instrumental in shaping this work.

I am also deeply thankful to the organizations and people who provided the data and resources necessary for conducting this research. Their cooperation allowed me to perform in-depth analysis and derive meaningful conclusions.

Lastly, I wish to acknowledge the unwavering support of my family, friends, and peers whose encouragement and understanding motivated me to complete this work.

This research is an accumulation of collective efforts, for which I am sincerely thankyou to everyone concerned.

REFERENCES

1. Consumer behavior The psychology of marketing Lars parner
2. Consumer behavior – Himanshu S. M. Dec 24 2008
3. Market basket Analysis in Multiple store environment Yen- liang Chen, Kwei Tung, Ren-Jie Shen, Ya- Han Hu. Decision support system (2005) Elsevier
4. A Data Mining Approach to consumer behavior- Junzo Watada, Kozo Yamashiro. Procedings of the first International Conference on Innovative computing Information (2006)
5. Mining of users access behavior for frequent sequential pattern from web logs. S. Vijaylakshmi, V. Mohan, S. Suresh Raja. International Journal of Database Management System (IJDM) Vol 2, August 2010.
6. Mining utility-oriented association rules: An efficient approach based on profit and quantity Parvinder S. Sandhu, Dalvinder S. Dhaliwal and S. N. Panda International Journal of the Physical Sciences Vol. 6(2), pp. 301-307, 18 January, 2011
7. B. Yıldız and B. Ergenç (Turkey) in "Comparison of Two Association Rule Mining Algorithms without Candidate Generation" International Journal of Computing and ICT Research 2010. International Journal of Computing and ICT Research, ISSN
8. "Extraction Of Interesting Association Rules Using Genetic Algorithms" Peter P. Wakabi-Waiswa* Venansius Baryamureeba, International Journal of Computing and ICT Research, Vol. 2 No. 1, June 2008
9. User centric approach to itemset utility mining in Market Basket Analysis Jyothi Pillai / International Journal on Computer Science and Engineering (IJCSE 1 Jan 2011)
10. Efficient Association Rule Mining for Market Basket Analysis Shrivastava A., Sahu R Global Journal of e-Business & Knowledge Management Year:2007,Volume:3,Issue:1
11. Mining association with the collective strength approach- Aggarwal C.C., Yu, P.S. 2001
12. Fast Algorithms for Mining Association Rules- Agrawal R.and Srikat l. Sept 1994 15
13. R.Agrawal, T. Imilienski and A. Swami. Mining Associations Rules between Sets of Items in large databases.Proc. Of the ALM SIGNOD. Int'l conf. on management of Data, pages 207-216,May1993
14. Fast Algorithms for Mining Association Rules R. Agarwal and R. Srikant algorithm for mining association rules. In proceedings of the 20th VLDB conference Santiago ,Chile, 1994
20. Sergey Brin , Rajeev Motwani and Craig silvertrin Beyound Market Baskets : Generalizing association rules for correlation.SIGMOD Record (ACM Special Interest Group on Management of Data).26 (2):265,1997
21. Sergey Brin ,Rajeev Motwani, Jetlrey D.Ullman, and Shulem Tsur. Dynamic interest counting and implication rules for market basket data..SIGMOD Record (ACM Special Interest Group on Management of data). 26(2):255,1997
22. Giadici Paulo. Applied Data mining :Statistical Methods for business and industry-ISBN 9812-53-178-5 23. Data mining Concepts and Techniques Jiauei Han , Michele Kamber,Simon Fraser University ISBN 1-55860-489-8-2001
23. Raorane, Abhijit & Kulkarni, Raja. (2011). Data Mining Techniques: A Source for Consumer Behavior Analysis. International Journal of Database Management Systems. 03. 10.5121/ijdms.2011.3304.
24. Apeksha, Ms & Ajmire, Prafulla. (2019). An Analytical Study of Use of Data mining for Entertainment.