BHARATHIDASAN ENGINEERING COLLEGE

TEAM LEADER

REGISTER NO : 510521205014

NAME : A.GURUDHEEP

PROJECT SUBMISSION PHASE-2

MARKET BASKET INSIGHTS

INTRODUCTION:

Market basket insights involve the analysis of transaction data to discover which items or products are commonly purchased together by customers during a single shopping trip or transactions

Implementation of market basket analysis requires a background in statistics and data science, as well as some algorithmic computer programming skills. For those without the needed technical skills, commercial, off-the-shelf tools exist.

This is a technique that gives the careful study of purchases done by a customer in a supermarket. This concept identifies the pattern of frequent purchase items by customers. This analysis can help to promote deals, offers, sale by the companies, and data mining techniques helps to achieve this analysis task.

ADVANCED ASSOCIATION ANALYSIS TECHNIQUES FOR MARKET BASKET INSIGHTS:

- Sequential Pattern Mining
- Temporal Analysis
- Frequent Pattern Growth Algorithms
- Multi-Dimensional Association Analysis
- Constrained Association Mining
- Hybrid Recommendation Systems
- Market Basket Analysis with Text Data
- Dynamic Pricing and Promotion
- Graph-Based Association Analysis
- Machine Learning Integration

SEQUENTIAL PATTERN MINING:

This technique focuses on identifying patterns of item sequences within transactions. It helps reveal the order in which items are purchased, allowing businesses to optimize product placement and recommendation strategies based on sequential behavior.

TEMPORAL ANALYSIS:

Temporal analysis considers the time dimension in market basket data. It helps in understanding how item associations change over time, such as seasonality or trends in purchasing behavior, enabling businesses to adapt their strategies accordingly.

FREQUENT PATTERN GROWTH ALGORITHMS:

These algorithms, like FP-Growth, efficiently discover frequent itemsets even in large datasets. They use a compact data structure called a frequent pattern tree to speed up the mining process

MULTI-DIMENSIONAL ASSOCIATION ANALYSIS:

This approach extends association analysis to consider multiple dimensions or attributes, such as customer demographics or location. It helps identify associations that are specific to certain customer segments or store locations.

CONSTRAINED ASSOCIATION MINING:

Constrained mining allows users to incorporate additional constraints into association rules, such as minimum support thresholds or item exclusions. This ensures that the discovered rules are more relevant and actionable.

HYBRID RECOMMENDATION SYSTEMS:

Combining association analysis with other recommendation techniques, like collaborative filtering or content-based filtering, can lead to more accurate and personalized product recommendations.

MARKET BASKET ANALYSIS WITH TEXT DATA:

Integrating natural language processing (NLP) techniques can help analyze textual information associated with transactions, such as product reviews or customer feedback, to gain deeper insights into customer preferences and sentiment.

DYNAMIC PRICING AND PROMOTION:

By analyzing market basket data in real-time, businesses can dynamically adjust pricing and promotional strategies to maximize revenue and customer satisfaction.

GRAPH-BASED ASSOCIATION ANALYSIS:

Representing transactional data as a graph, where items are nodes and associations are edges, can reveal complex relationships and communities within the data, aiding in targeted marketing and product bundling strategies.

MACHINE LEARNING INTEGRATION:

Advanced machine learning algorithms, such as deep learning or reinforcement learning, can enhance association analysis by modeling complex patterns and making real-time recommendations.

VISUALIZATION TOOLS FOR ENHANCED INSIGHTS PRESENTATION:

- Tableau
- Power BI
- QlikView/Qlik
- Looker
- D3.j
- Google Data Studio:
- Plotly
- Matplotlib
- Infogram
- Sisense
- Periscope Data

- IBM Cognos Analytics
- Highchart

Tableau:

Tableau is a widely used data visualization tool that allows users to create interactive and shareable dashboards. It supports a wide range of data sources and offers a drag-and-drop interface for building visualizations.

Power BI:

Microsoft Power BI is another popular choice for data visualization and business intelligence. It offers robust data connectivity, powerful data modeling capabilities, and the ability to create interactive reports and dashboards.

QlikView/Qlik sense:

QlikView and Qlik Sense are data visualization and business intelligence platforms that enable users to explore and visualize data for better insights. Qlik Sense is particularly known for its self-service capabilities.

Looker:

Looker is a data exploration and business intelligence platform that allows users to create and share interactive data visualizations and dashboards. It integrates well with various data sources.

D3.js:

D3.js is a JavaScript library for creating custom and interactive data visualizations. It offers full control over the design and interactivity of visualizations but requires coding expertise.

Google Data Studio:

This free tool by Google allows users to create interactive and shareable reports and dashboards using data from various sources, including Google Analytics, Google Sheets, and more.

Plotly:

Plotly is a versatile Python library for creating interactive and publication-quality graphs. It supports a wide range of chart types and can be used in Jupyter notebooks, web applications, and more.

Matplotlib:

Matplotlib is a popular Python library for creating static, animated, and interactive visualizations. It is highly customizable and commonly used for data exploration and research.

Infogram:

Infogram is an easy-to-use infographic and chart maker tool that simplifies the process of creating visually appealing charts, infographics, and reports.

Sisense:

Sisense is a business intelligence platform that offers data visualization capabilities, allowing users to create interactive dashboards and reports for data analysis.

Periscope Data:

Now part of Sisense, Periscope Data provides a platform for data visualization and analysis, enabling teams to collaborate on data projects and create impactful visualizations.

IBM Cognos Analytics:

IBM's Cognos Analytics is a business intelligence and reporting tool that supports data visualization, reporting, and dashboard creation, with AI-driven insights.

Highcharts:

Highcharts is a JavaScript charting library that enables the creation of interactive and responsive charts for web applications and websites.

Choosing the right visualization tool depends on your specific needs, data sources, technical skills, and budget. It's important to select a tool that aligns with your objectives and allows you to present insights in a compelling and informative manner.

BENEFITS OF VISUALIZATION TOOLS IN MARKET BASKET ANALYSIS:

- **CLARITY AND UNDERSTANDING:**

Visualization tools transform large and complex datasets into intuitive charts, graphs, and dashboards. This simplifies the data and allows users to quickly grasp patterns, trends, and relationships in market basket data.

- **DISCOVER PATTERNS AND TRENDS:**

Visualizations make it easier to identify patterns and trends in customer purchasing behavior, helping businesses uncover valuable insights that might not be apparent in raw data.

- **REAL-TIME INSIGHTS:**

Interactive visualization tools enable businesses to monitor market basket data in real-time, making it possible to respond quickly to changing customer behavior and adapt marketing and sales strategies.

- **ENHANCED COMMUNICATION:**

Visualizations are a powerful communication tool. They allow teams and decision-makers to share insights with colleagues, partners, and stakeholders effectively, ensuring a common understanding of the data.

- **CUSTOMIZATION:**

Visualization tools offer flexibility in designing charts and graphs to suit specific needs. Users can customize colors, labels, and other elements to create visuals that convey the intended message.

- **DRILL-DOWN CAPABILITY: :**

_____Users can drill down into visualizations to explore data at different levels of granularity. This interactive feature helps in understanding the details behind the summarized insights.

- **CROSS-FUNCTIONAL INSIGHTS:**
  Visualizations can be understood by a wide range of stakeholders, including those without a deep understanding of data analysis. This enables cross-functional teams to collaborate effectively.
- **DATA VALIDATION**:
  _____Visualization tools allow users to visually validate the results of market basket analysis. If the relationships and patterns identified make sense when presented visually, it adds credibility to the insights.
- **BETTER DECISION-MAKING:**
  _____Visualizations facilitate data-driven decision-making by making insights easily accessible and understandable. This, in turn, leads to more informed and effective decision-making.
- **SEGMENTATION AND TARGETING:**
  _____Visualizations can help in segmenting customers based on their purchasing behavior and preferences, enabling businesses to tailor marketing and product recommendation strategies for different segments.
- **INVENTORY MANAGEMENT:**
  _____By visualizing sales patterns and product associations, businesses can optimize inventory management, ensuring the right products are in stock and minimizing overstock or understock situations.
- **EFFECTIVE REPORTING:**
  Visualization tools are valuable for creating visually appealing and insightful reports for stakeholders, investors, and management. Well-designed reports can convey key findings more persuasively.
- **OPTIMIZED MARKETING:**
  Visualizations help in identifying cross-selling and upselling opportunities, enabling more targeted and effective marketing campaigns. This, in turn, can increase sales and revenue.
- **CUSTOMER RETENTION:**
  _____Visualizations can help in identifying customer preferences, allowing businesses to offer personalized recommendations and discounts, which can enhance customer retention.
- **BENCHMARKING:**
  _____Visualizations facilitate benchmarking by comparing the performance of different product categories, stores, or customer segments. This helps in identifying areas for improvement.

In summary, visualization tools enhance market

APPLICATION OF MARKET BASKET INSIGHTS:

BASKET ANALYSIS VISUALIZATION:

Visualizations can illustrate the most frequently co-occurring items in a customer's basket, enabling businesses to identify popular product combinations and patterns. Bar charts, heatmaps, and word clouds can help convey these insights effectively.

Support, Confidence, and Lift Visualizations: Visualization tools can display support, confidence, and lift values associated with different itemsets or association rules. This makes it easier to evaluate the significance of these metrics and choose the most relevant rules for further action.

INTERACTIVE DASHBOARDS: Dashboards provide a dynamic and interactive way to explore market basket insights. Users can filter, drill down, and interact with the data, allowing for real-time exploration and decision-making. This is particularly useful for retailers who want to analyze their sales data.

SEASONAL AND TEMPORAL PATTERNS:Time series visualizations can help identify seasonal and temporal patterns in market basket data. Line charts and time series plots can show how item associations change over time, allowing for better preparation for seasonal demand fluctuations.

CUSTOMER SEGMENTATION: Visualization tools can be used to create customer segments based on their shopping behavior. Businesses can then use these segments to tailor marketing campaigns or adjust product offerings for different groups of customers.

PRODUCT PLACEMENT OPTIMIZATION: Heatmaps or store layout visualizations can show how items are typically arranged in stores and their relative popularity. This information can help optimize product placement for increased sales.

DYNAMIC PRICING AND PROMOTIONS: Visualization can assist in visualizing the impact of different pricing and promotional strategies on sales and customer behavior. This helps businesses make informed decisions on discounts, coupons, and promotions.

MARKET BASKET INSIGHTS REPORTS: Visualizations are an integral part of reports that convey market basket insights to stakeholders. Combining charts, graphs, and tables can present a comprehensive view of findings and recommendations.

MARKET BASKET ANALYSIS FOR E-COMMERCE: Visualizations can show frequently purchased items and their relationships, making it easier to recommend related products to online shoppers.

CROSS-SELLING OPPORTUNITIES: Visualizations can highlight cross-selling opportunities by illustrating complementary or frequently associated products. This information can guide recommendations and upselling strategies.

TEXTUAL DATA INTEGRATION: Visualization tools can combine market basket analysis results with textual data from customer reviews or feedback, providing a more comprehensive understanding of customer preferences.

INVENTORY MANAGEMENT: Visualizations can be used to monitor inventory levels and predict demand based on market basket insights. Heatmaps and bar charts can show the popularity of items and help businesses optimize stock levels.

GEOSPATIAL INSIGHTS: For businesses with multiple physical locations, geospatial visualizations can help identify regional variations in market basket behavior. These insights can be used for targeted marketing and inventory planning.

Incorporating visualization tools into market basket analysis allows businesses to communicate their findings more effectively, make data-driven decisions, optimize marketing strategies, and enhance the customer experience. These tools provide a bridge between raw data and actionable insights, ultimately contributing to improved profitability and customer satisfaction.

EXAMPLE FOR MARKET BASKET INSIGHTS:

```
from mlxtend.frequent_patterns import apriori

from mlxtend.frequent_patterns import association_rules

import pandas as pd

# Sample dataset (replace with your own data)

data = {

   'TransactionID': [1, 1, 2, 2, 3, 3, 4, 4, 5, 5],

   'Item': ['Milk', 'Bread', 'Milk', 'Diapers', 'Bread', 'Diapers', 'Milk', 'Bread', 'Diapers', 'Beer']

}

df = pd.DataFrame(data)

# One-hot encode the dataset

basket = pd.crosstab(df['TransactionID'], df['Item']).applymap(lambda x: 1 if x > 0 else 0)

# Find frequent itemsets

frequent_itemsets = apriori(basket, min_support=0.2, use_colnames=True)

# Generate association rules

association_rules = association_rules(frequent_itemsets, metric='lift', min_threshold=1.0)

# Display the frequent itemsets

print("Frequent Itemsets:")
```

```
print(frequent_itemsets)
# Display the association rules
print("\nAssociation Rules:")
print(association_rules)
```

SAMPLE OUTPUT:

```
Frequent Itemsets:
   support      itemsets
0      0.4       (Beer)
1      0.4      (Bread)
2      0.4     (Diapers)
3      0.4       (Milk)
4      0.2  (Beer, Milk)
5      0.2  (Milk, Bread)
6      0.2  (Bread, Diapers)
7      0.2  (Milk, Diapers)


Association Rules:
  antecedents consequents  antecedent support  ...      lift  leverage  conviction
0     (Beer)      (Milk)              0.4  ...  1.666667      0.08         1.2
1     (Milk)      (Beer)              0.4  ...  1.666667      0.08         1.2
2    (Bread)      (Milk)              0.4  ...  1.666667      0.08         1.2
3     (Milk)     (Bread)              0.4  ...  1.666667      0.08         1.2
4    (Diapers)    (Bread)              0.4  ...  1.666667      0.08         1.2
5    (Bread)    (Diapers)              0.4  ...  1.666667      0.08         1.2
6     (Milk)    (Diapers)              0.4  ...  1.666667      0.08         1.2
7    (Diapers)     (Milk)              0.4  ...  1.666667      0.08         1.2
```

CONCLUSION:

Cross-selling and upselling is the secret mantra of the retail industry that pushed the consumer to buy more. It has become a thriving factor for such industries that harness patterns with market basket analysis in data mining and derive customer insights to upscale their brand's performance.

An urban legend states that a grocery store increased its sales after they placed beer and diapers together because of the market basket analysis that stated that beer and diapers were both purchased frequently by men.

Organizations are using this technique wisely and making billions by playing with the mind of the customer. It is an effective way of improving your sales without having to put extra effort into marketing that won’t give you results as incredible as with this technique. So go ahead and try it on all the data you have in your repository to recognize patterns that may surprise you to the roots.