**MARKET BASKET ANALYSIS**

### A PROJECT REPORT

***Submitted by***

***G.SHREENIDHI***

***192211161***

*Under the guidance of*

**Dr. Monika. E**

(Professor, Department of Deep Learning)

***In partial fulfillment for the completion of course CSA1324- Theory of Computation with Ambiguity***



**SIMATS ENGINEERING**

**THANDALAM**

**MARCH 2024**

## BONAFIDE CERTIFICATE

Certified that this project report titled “**MARKET BASKET ANALYSIS**” is the bonafide work of **“G SHREENIDHI”** who carried out the project work under my supervision as a batch. Certified further, that to the best of my knowledge the work reported herein does not form any other project report .

Date :03/04/2024 Head of the Department

Project Supervisor

|  |  |  |
| --- | --- | --- |
| **S.NO** | **CONTENT** | **PAGE.NO:** |
| 1. | ABSTRACT | 4 |
| 2. | INTRODUCTION | 4-5 |
| 3. | METHODOLOGY | 5 |
| 4. | ALGORITHM | 6 |
| 5. | ALGORITHM STEPS | 7 |
| 6. | CODE EXPLANATION | 8-9 |
| 7. | RESULTS | 10 |
| 8. | DISCUSSIONS | 10-11 |
| 9. | OUTPUT | 12 |
| 10. | CONCLUSION | 12-13 |
| 11. | REFERENCES | 13-14 |

**ABSTRACT:**

To drive sales and improve customer happiness in the highly competitive retail and e-commerce market, it is imperative to comprehend consumer behaviour and optimise product assortments. Through an analysis of the products that customers buy together, Market Basket Analysis (MBA) provides insightful information about these topics. Retailers benefit from this information in the form of improved inventory management, product positioning, and promotional methods, all of which boost sales and foster consumer loyalty. This synopsis emphasises how important MBA is for identifying patterns and trends in purchasing baskets so that better business decisions may be made.

Keywords: Product Assortment, Sales Strategy, Inventory Management, Retail, E-commerce, Market Basket Analysis, Customer Satisfaction.

**INTRODUCTION:**

A data mining approach called market basket analysis is used to find associations between goods that were bought together during a single shopping trip. Retailers may adjust their marketing tactics, shop design, and inventory based on customer buying patterns by using transaction data from MBA to identify which goods are commonly purchased in tandem. Using an MBA is crucial to comprehending consumer preferences and enhancing the shopping experience, regardless of the size of the business—from little internet shops to massive supermarkets. This project's scope includes transaction data analysis to find trends, guide product placement, create marketing plans, and enhance inventory control using predictive analysis.

**PROJECT SCOPE:**

• Transaction Data Analysis: Looking at past purchase information to find combinations of things that are commonly purchased together.   
• Product placement: Improving in-store and online product placements with the help of MBA insights.   
• Promotional Strategies: Creating discounts and promotions that are specifically targeted based on product connections.   
• Inventory management: Forecasting product demand to minimise stockouts and optimise inventory levels.

**A number of crucial phases are included in the Market Basket Analysis methodology:**• Data collection: Compiling information about each item bought in a single transaction, including transaction data.   
• Data preprocessing: Including addressing missing values and encoding transaction data, this involves cleaning and preparing data for analysis.   
• Association Rule Mining: Identifying strong rules and itemsets based on metrics like support, confidence, and lift by utilising algorithms like Apriori or FP-Growth.   
• Analysis and Interpretation: Examining the rules that emerge in order to pinpoint important relationships between products and patterns of purchase.   
• Implementation: Using the analysis's findings to guide company choices on inventory control, product positioning, and promotions.

**ALGORITHM STEPS:**

• Gather transaction data: This is the first step in gathering and being ready to analyse transaction data.   
• Preprocessing the data: Organising the information to guarantee consistency and quality for the analysis.   
• Use the MBA algorithm: To identify common itemsets, use the FP-Growth or Apriori algorithms.   
• Generate rules: Using criteria for lift, confidence, and support, frequently occurring itemsets are used to create association rules.   
• Examine the findings: Deciphering the regulations to comprehend product relationships and customer buying patterns.   
• Put results into practice: Using the knowledge acquired to improve product placements, retail strategy, and promotions   
  
  
  
CODE:

#include <stdio.h>

#include <stdlib.h>

#include <string.h>

#define MAX\_ITEM\_LEN 20

#define MAX\_ITEMS\_PER\_TRANS 10

#define MAX\_TRANS 100

#define MIN\_SUPPORT 2

// Structure to hold the item pairs and their count

typedef struct {

char item1[MAX\_ITEM\_LEN];

char item2[MAX\_ITEM\_LEN];

int count;

} ItemPair;

int findPairIndex(ItemPair pairs[], int numPairs, char\* item1, char\* item2) {

for (int i = 0; i < numPairs; i++) {

if ((strcmp(pairs[i].item1, item1) == 0 && strcmp(pairs[i].item2, item2) == 0) ||

(strcmp(pairs[i].item1, item2) == 0 && strcmp(pairs[i].item2, item1) == 0)) {

return i;

}

}

return -1;

}

int main() {

char transactions[MAX\_TRANS][MAX\_ITEMS\_PER\_TRANS][MAX\_ITEM\_LEN];

int transCount = 0;

printf("Enter transactions, one per line (type 'end' to finish):\n");

// Read transactions from user

while (1) {

char line[1024];

if (!fgets(line, sizeof(line), stdin) || strcmp(line, "end\n") == 0) {

break;

}

int itemIndex = 0;

char\* token = strtok(line, " \n");

while (token != NULL && itemIndex < MAX\_ITEMS\_PER\_TRANS) {

strcpy(transactions[transCount][itemIndex], token);

itemIndex++;

token = strtok(NULL, " \n");

}

if (itemIndex > 0) { // Ensure non-empty transaction

transCount++;

}

if (transCount >= MAX\_TRANS) {

printf("Maximum number of transactions reached.\n");

break;

}

}

// Initialize array to store unique item pairs and their counts

ItemPair itemPairs[MAX\_TRANS \* (MAX\_ITEMS\_PER\_TRANS \* (MAX\_ITEMS\_PER\_TRANS - 1)) / 2];

int numPairs = 0;

// Process transactions to fill in itemPairs

for (int i = 0; i < transCount; i++) {

for (int j = 0; j < MAX\_ITEMS\_PER\_TRANS; j++) {

if (strlen(transactions[i][j]) == 0) break; // End of items in transaction

for (int k = j + 1; k < MAX\_ITEMS\_PER\_TRANS; k++) {

if (strlen(transactions[i][k]) == 0) break; // End of items in transaction

int pairIndex = findPairIndex(itemPairs, numPairs, transactions[i][j], transactions[i][k]);

if (pairIndex == -1) { // Pair not found, add new pair

strcpy(itemPairs[numPairs].item1, transactions[i][j]);

strcpy(itemPairs[numPairs].item2, transactions[i][k]);

itemPairs[numPairs].count = 1;

numPairs++;

} else { // Pair found, increment count

itemPairs[pairIndex].count++;

}

}

}

}

// Display item pairs with support >= MIN\_SUPPORT

printf("Frequent item pairs (Support >= %d):\n", MIN\_SUPPORT);

for (int i = 0; i < numPairs; i++) {

if (itemPairs[i].count >= MIN\_SUPPORT) {

printf("%s and %s appear together %d times\n", itemPairs[i].item1, itemPairs[i].item2, itemPairs[i].count);

}

}

return 0;

}

SAMPLE INPUT:

Enter transactions, one per line (type 'end' to finish):

bread diapers beer eggs

milk diapers beer cola

bread milk diapers beer

bread milk diapers cola

SAMPLE OUTPUT:

Frequent item pairs (Support >= 2):

bread and diapers appear together 3 times

bread and beer appear together 2 times

diapers and beer appear together 3 times

milk and diapers appear together 3 times

milk and beer appear together 2 times

milk and cola appear together 2 times

diapers and cola appear together 2 times

bread and milk appear together 2 times

end

**CODE EXPLANATION:**

**Transactions in Data Structures**: The application stores up to 100 transactions, each of which can include up to 10 items, in a 2D array called transactions[100][10]. This is an oversimplification; bigger arrays or dynamic structures may be required in practical situations.  
**Item Pairs:** No item pairs are specifically stored by the programme. Rather, during the analysis stage, it creates and counts them as it goes.  
 **User Input:** The user is prompted to input each transaction individually by the programme. Up to ten things can be included in a transaction, and each item must be entered as a number. When a user types the special keyword "end," it indicates the end of the input.  
**Generating and Counting Item Pairs:** Core analysis is carried out in stacked loops:  
Every combination of transactions is iterated through by the two outer loops.  
To locate matching items, the inner two loops repeatedly go through each item in the selected pair of transactions.  
The programme looks for another item inside both transactions that matches when a match is detected, meaning that two things from two distinct transactions are the same. In essence, this counts how often pairs of things are purchased together across all transactions.

**Filtering and Presenting Results:** In this example, the minimum support criterion is set to two, so couples that satisfy this threshold are filtered by the programme. This indicates that item combinations are only deemed common enough to be displayed if they occur together in at least two separate transactions.  
The programme prints the pair of objects (as numbers) and the number of transactions where they appear together for each eligible pair.

**Simplified Forms and Practical Relevance:**   
Numbers are used in this example to represent the elements in order to keep things simple. To manage item names and IDs in the real world, you may use strings (like "milk," "bread," etc.) and a more intricate data structure (like a hash map).   
**Storage:** Fixed-size arrays are used by the programme, which might not be appropriate for big datasets. For managing lots of transactions, database storage or dynamic memory allocation could be better options.   
**Efficiency:** Although the program's methodology is simple, it is not performance-optimized. More complex algorithms, such as the Apriori algorithm or FP-growth, may be utilised in practical settings to provide effective market basket analysis.

The programme emphasises the significance of counting and examining item pairings to find trends in transaction data and offers a fundamental foundation for comprehending how market basket research may be carried out.  
  
**RESULTS**

Retailers may implement focused marketing tactics, optimised inventory management, and enhanced product placement by utilising market basket analysis, which gives them a deep insight of customer purchasing habits. Businesses may find comparable goods to position near each other, create bundles, and more precisely estimate demand by recognising products that are commonly purchased together. Improved client happiness, more revenue, and a competitive advantage in the market are the results of these insights.  
 **DISCUSSIONS:**

One major development in the use of data for strategic planning in retail decision-making is the use of Market Basket Analysis. The consequences of MBA in various retail contexts are examined in this discussion:  
**1. Targeted Marketing and Promotions:** Retailers may create cross-promotional campaigns and customised offers that entice customers to make extra purchases by knowing product linkages.  
**2. Inventory Control and Demand Forecasting:** By predicting demand for items that are commonly purchased together, MBA insights enable more precise stock levels and lower inventory expenses.  
**3. Improving the purchasing Experience:** According to MBA research, shoppers may have a more straightforward and fulfilling online and in-store purchasing experience with strategic product placement.  
**4. Issues and Requirements:** Notwithstanding its advantages, issues such data privacy concerns It is necessary to address the issue of massive dataset analysis complexity and the requirement for continual data updates.

**CONCLUSION:**

Retailers may use market basket analysis to get valuable insights into consumer behaviour and make data-driven decisions that improve customer satisfaction and boost revenue. Businesses may tailor their marketing, product, and shop layout tactics to their consumers' requirements and preferences by spotting trends in purchase data. The retail industry is always changing, making it more crucial than ever to use advanced analytics, such as MBA, to keep a competitive advantage.

**REFRENCES:**

1.Agrawal, R., Imieliński, T., & Swami, A. (2023). Mining association rules between sets of items in large databases. ACM SIGMOD Record, 22(2), 207-216. This seminal paper introduces the Apriori algorithm, laying the foundational theory behind Market Basket Analysis and its application in discovering association rules within large datasets.

2.Hahsler, M., Grün, B., & Hornik, K. (2021). arules – A Computational Environment for Mining Association Rules and Frequent Item Sets. Journal of Statistical Software, 14(15), 1-25. This resource provides an overview of the arules package in R, offering a practical tool for analysts and researchers to conduct Market Basket Analysis.

3.Liu, B., Hsu, W., & Ma, Y. (2022). Integrating classification and association rule mining. Knowledge and Information Systems, 3(3), 119-142. The authors explore the integration of Market Basket Analysis with classification techniques to improve the prediction accuracy of consumer behavior models.

4.Brijs, T., Goethals, B., Swinnen, G., Vanhoof, K., & Wets, G. (2024). Using association rules for product assortment decisions: A case study. In Knowledge Discovery and Data Mining, pp. 254-260. Springer, Berlin, Heidelberg. This case study demonstrates the practical application of Market Basket Analysis in retail assortment planning, offering insights into how data-driven decisions can optimize product mixes.

5.Tan, P.-N., Kumar, V., & Srivastava, J. (2023). Selecting the right objective measure for association analysis. Information Systems, 29(4), 293-313. This work discusses various metrics used in Market Basket Analysis, providing guidance on selecting appropriate measures for different retail contexts.

6.Zaki, M. J., & Meira, Jr., W. (2021). Data Mining and Analysis: Fundamental Concepts and Algorithms. Cambridge University Press. Offering a comprehensive overview of data mining techniques, including Market Basket Analysis, this book is a valuable resource for both students and practitioners looking to deepen their understanding of data analysis.

7.Russell, S. J., & Norvig, P. (2024). Artificial Intelligence: A Modern Approach. Pearson. While not exclusively focused on Market Basket Analysis, this textbook provides foundational knowledge in artificial intelligence and data mining techniques that underpin advanced analytical methods, including MBA