

ARTIFICIAL INTELLIGENCE

PROJECT NAME:-MARKET BASKET ANALYSIS(INSIGHTS)

PHASE 4 SUBMISSION BY:- JAGAN VIGNESH.B

NM ID:- DB66480CB493E7DA570430FCB88B1F03

YEAR:- 3RD

DEPT:- COMPUTER SCIENCE AND ENGINEERING (C.S.E)

##TABLE OF CONTENT:--

Data Collection:

2. DATA PREPROCESSING:

3. DATA EXPLORATION:

3. DATA EXPLORATION:

4. MARKET BASKET ANALYSIS:

a. Convert Data to a One-Hot Encoded Format:

b. Finding Frequent Itemsets:

c. Discovering Association Rules:

d. Review and Interpret the Rules:

5. VISUALIZATION:

6. INTERPRETATION AND INSIGHTS:

7. BUSINESS APPLICATION:

8. CONTINUOUS MONITORING AND REFINEMENT:

MARKET BASKET ANALYSIS is a data mining technique used to discover patterns and relationships between items that tend to be purchased together. It's widely used in retail and e-commerce for optimizing product placement, recommendation systems, and more. In this example, I'll guide you through the process of performing Market Basket Analysis using Python with the popular library, **mlxtend**.

```
# Import necessary libraries

import pandas as pd

from mlxtend.frequent_patterns import apriori
from mlxtend.frequent_patterns import association_rules


# Create a sample dataset (You can load your own dataset)

data = {'Transaction': [1, 1, 2, 2, 3, 3, 4, 4, 5],
        'Item': ['A', 'B', 'A', 'B', 'A', 'C', 'A', 'B', 'C']}

df = pd.DataFrame(data)


# Convert the dataset into a one-hot encoded format

basket = pd.get_dummies(df, columns=['Item'], prefix='', prefix_sep='')


# Group by transaction and sum the items

basket = basket.groupby('Transaction').sum()
```

```
# Convert item counts to binary values (1 if an item appears in the transaction, 0 otherwise)
basket[basket >= 1] = 1

# Use Apriori algorithm to find frequent itemsets
frequent_itemsets = apriori(basket, min_support=0.5, use_colnames=True)

# Find association rules from the frequent itemsets
rules = association_rules(frequent_itemsets, metric='lift', min_threshold=1.0)

# Display the frequent itemsets and association rules
print("Frequent Itemsets:")
print(frequent_itemsets)

print("\nAssociation Rules:")
print(rules)
```

This code creates a sample dataset, converts it into a one-hot encoded format, finds frequent itemsets using the Apriori algorithm, and then generates association rules based on the frequent itemsets.

You can adjust the **min_support** and **min_threshold** parameters to control the minimum support and lift values for frequent itemsets and association rules, respectively, to suit your specific use case. Additionally, replace the sample dataset with your own transaction data.

Make sure to install the required libraries and adapt the code as needed for your specific dataset and analysis.

Creating a Market Basket Analysis (MBA) project involves multiple steps. Let's break it down topic by topic:

1. Data Collection:

To start an MBA project, you need transaction data. You can collect this data from your point of sale (POS) systems or use publicly available datasets. For example, you can use data from a grocery store, an online retail platform, or any other business where customers make multiple purchases in a single transaction.

2. Data Preprocessing:

Data preprocessing is essential to clean and prepare the data for analysis. This step involves handling missing values, data encoding, and structuring the data in a way that's suitable for MBA. The data should be organized as a table with transactions and items.

3. Data Exploration:

Before diving into MBA, it's helpful to explore the data to get a better understanding of customer behavior. You can create visualizations, calculate basic statistics, and identify key trends in the data.

4. Market Basket Analysis:

Here's how you can perform Market Basket Analysis using Python:

a. Convert Data to a One-Hot Encoded Format:

```
import pandas as pd

from mlxtend.frequent_patterns import apriori
from mlxtend.frequent_patterns import association_rules

# Load your data into a DataFrame

# Convert data to one-hot encoded format

basket = pd.get_dummies(your_data, columns=['Item'], prefix="",
prefix_sep='')

# Group by transaction and sum the items

basket = basket.groupby('Transaction').sum()

# Convert item counts to binary values (1 if an item appears in the
transaction, 0 otherwise)

basket[basket >= 1] = 1
```

b. Finding Frequent Itemsets:

```
# Use Apriori algorithm to find frequent itemsets

frequent_itemsets = apriori(basket, min_support=0.1,
use_colnames=True)
```

c. Discovering Association Rules:

```
# Find association rules from the frequent itemsets  
  
rules = association_rules(frequent_itemsets,  
metric='lift', min_threshold=1.0)
```

d. Review and Interpret the Rules:

```
# Display the association rules  
  
print(rules)
```

5. Visualization:

Visualize the results of your Market Basket Analysis. You can use libraries like Matplotlib or Seaborn to create bar charts or other visualizations to make the findings more accessible.

6. Interpretation and Insights:

The generated association rules will provide insights into item relationships. Interpret the rules to understand which items are frequently purchased together. Look for high confidence and lift values to identify strong associations.

7. Business Application:

Use the insights gained from the analysis to inform business decisions. For example, you can optimize product placement in stores, create product bundles, improve recommendations in e-commerce, or design targeted marketing campaigns.

8. Continuous Monitoring and Refinement:

Market Basket Analysis isn't a one-time task. It's an ongoing process. Continuously collect data, re-run the analysis, and adjust strategies based on changing customer behavior.

Remember that the success of your MBA project depends not only on the technical aspects but also on your ability to interpret and apply the results to your specific business goals.