ARTIFICIAL INTELLIGENCE
PROJECT:- MARKET BASKET ANALYSIS(MBA)

YEAR:- 3RD
DEPT:- COMPUTER SCIENCE & ENGINEERING(C.S.E)

TABLE OF CONTENTS:-

****START****

****INTRODUCTION****

Market Basket Analysis (MBA) is a data mining technique used to discover associations between items purchased by customers. In this Python tutorial, we'll walk through the process of performing MBA step by step, topic by topic.

1.DATA PREPARATION:

Before you can perform market basket analysis, you need transaction data that records the items purchased by customers. Start by loading your data into Python using a library like pandas. Ensure that your data is structured with one row per transaction and the items listed in columns.

```
import pandas as pd

# Load your transaction data data =
pd.read_csv('transaction_data.csv')
```

2. DATA EXPLORATION:

Explore your data to get a better understanding of its structure. You can check the first few rows using **head()**, find unique items, and check for any missing values.

3.DATA PREPROCESSING:

Preprocess the data as needed. Common preprocessing steps include removing duplicates and converting the data into a suitable format for MBA.

```
# Remove duplicates
data = data.drop_duplicates()

# Convert data to a one-hot encoded
format basket = pd.get_dummies(data,
columns=['item'])
```

4.APRIORI ALGORITHM:

Use the Apriori algorithm, provided by the **mlxtend** library, to find frequent itemsets. The frequent itemsets are item combinations that occur above a specified minimum support threshold.

```
from mlxtend.frequent_patterns import apriori

min_support = 0.1 # Adjust the support threshold as
needed frequent_itemsets = apriori(basket,
min_support=min_support, use_colnames=True)
```

5.ASSOCIATION RULES:

Generate association rules from the frequent itemsets. Association rules consist of antecedents and consequents and are evaluated based on metrics like confidence and lift.

```
from mlxtend.frequent_patterns import
association_rules

min_confidence = 0.5 # Adjust the confidence
threshold as needed
rules = association_rules(frequent_itemsets,
metric='confidence', min_threshold=min_confidence)
```

6.VISUALIZE RESULTS:

You can visualize the frequent itemsets and association rules using various plots or tables. For example, you can create a table to display the results:

```
# Display frequent itemsets
print("Frequent Itemsets:")
print(frequent_itemsets)

# Display     association
rules  print("Association
Rules:") print(rules)
```

7. Interpret Results:

Interpret the results to gain insights. For instance, you can identify which items are commonly bought together and the strength of those associations based on the confidence and lift values in the association rules.

8. Fine-Tuning:

Experiment with different support and confidence thresholds to refine your analysis. Additionally, consider filtering and post-processing the results based on your specific business objectives.

9. Real-World Application:

Apply the insights gained from the analysis to improve business strategies, such as product placement, marketing, and customer recommendations.

This step-by-step guide provides a solid foundation for performing market basket analysis in Python. Depending on your specific use case and dataset, you may need to adapt and extend these steps as necessary.

****END****

-:THANK YOU:-