

ASSIGNMENT 6

25

Aim :

to understand the Association in machine learning using Market - Basket Analysis

OBJECTIVE :

create Association Rules for the market - Basket Analysis for the given Threshold.

THEORY :

ASSOCIATION RULE :

Association rule mining finds interesting associations and relationships among large sets of data items. This rule shows how frequently a itemset occurs in a transaction. A typical example is market Based Analysis.

Given a set of transactions, we can find rules that will predict the occurrence an item based on the occurrences of other items in the transaction.

TID Items

1. Bread, milk
2. Bread, Diaper, Beer, Eggs
3. milk, Diaper, Beer, Coke
4. Bread, milk, Diaper, Beer
5. Bread, Milk, Diaper, coke

Before we start defining the rule, let us first see the basic definitions.

support count() - frequency of occurrence of a itemset.

Here. $(\{ \text{milk, bread, Diaper} \}) = 2$

ASSIGNMENT 6

25

Aim :

to understand the Association in machine Learning using Market - Basket Analysis

OBJECTIVE :

create Association Rules for the market - Basket Analysis for the given Threshold.

THEORY :

ASSOCIATION RULE :

Association rule mining finds interesting associations and relationships among large sets of data items. This rule shows how frequently a itemset occurs in a transaction. A typical example is market Based Analysis.

Given a set of transactions, we can find rules that will predict the occurrence an item based on the occurrences of other items in the transaction.

TID Items

1. Bread, milk
2. Bread, Diaper, Beer, Eggs
3. milk, Diaper, Beer, Coke
4. Bread, milk, Diaper, Beer
5. Bread, Milk, Diaper, coke

Before we start defining the rule, let us first see the basic definitions.

support count() - frequency of occurrence of a itemset.

Here $(\{milk, Bread, Diaper\}) = 2$

EXAMPLE:

From the above table. $(\text{milk}, \text{Dialer}) = (\text{Beer})$

$S = (\text{milk}, \text{Dialer}, \text{Beer})$ [T]

≈ 15

≈ 0.4

$C = (\text{milk}, \text{Dialer}, \text{Beer})$ $(\text{milk}, \text{Dialer})$

≈ 13

≈ 0.67

The Association rule is very useful in analyzing datasets. The data is collected using bar-code scanners in supermarkets. Such database consists of large number of transaction records which list all items bought by a customer on a single purchase.

APRIORI ALGORITHM:

Apriori Algorithm is a machine learning algorithm which is used to gain insight into the structured relationships between different items involved. The most prominent practical application of the algorithm is to recommend products based on the products already present in user's cart. Walmart especially has made great use of the algorithm in suggesting products to users.

IMPLEMENTATION OF ALGORITHM IN PYTHON

Step 1. Importing the required libraries

- step 2. Loading and exploring the data.
- step 3. cleaning the data.
- step 4. splitting the data according to the region of transaction.
- step 5. Hot encoding the data.
- step 6. Building the models and analyzing the results.

APPLICATIONS:

so what kind of items are we talking about?
 There are many applications of association

- Product recommendation - like Amazon's customers who bought this, also bought this.
- music recommendation - like FM's artist recommendations.
- Medical diagnosis - like with diabetes really cool stuff
- content optimisation - like in magazine websites or blogs

CONCLUSION:

Thus we have studied and implemented the Association rule in machine Learning market-basket Analysis