MARKET BASKET INSIGHTS

TEAM MEMBER

810621104004: ANITHA.S

PROJECT: MARKET BASKET INSIGHTS

DEFINITION :

Market Basket insights is a mathematical modeling technique based upon the theory that if you buy a certain group of items, you are likely to buy another group of items.

It is used to analyze the customer purchasing behavior and helps in increasing the sales and maintaining ventory by focusing on the point of sale transaction data.

Given a dataset, the Apriori Algorithm trains and identifies product baskets and product association rules

TERMINOLOGY :

Transaction is a set of items (Item set).

Confidence : It is the measure of uncertainty or trust worthiness associated with each discovered pattern.

Support : It is the measure of how often the collection of items in an association occur together as percentage of all transactions

Frequent item set : If an item set satisfies minimum support, then it is a frequent item set.

Strong Association rules: Rules that satisfy both a minimum support threshold and a minimum confidence threshold

In Association rule mining, we first find all frequent item sets and then generate strong association rules from the frequent item sets.

Apriori algorithm is the most established algorithm for finding frequent item sets meaning. The basic principle of Apriori is "Any subset of a frequent item set must be frequent".

We use these frequent item sets to generate association rules.

APRIORI ALGORITHM :

Ck: Candidate itemset of size k.

Lk: Frequent item set of size kL1={frequent items};

For (k=1; Lk!=0; k++) do beginCk+1= Candidates generated from Lk;

For each transaction t in the database do Increment the count of all candidates in Ck+1 that are contained in t.

Lk+1=candidates in Ck+1 with min_support End.

Return UkLk;

DEMO 1

Installations

Oracle 10g enterprise edition

SQL Plus

Oracle Data Miner Client

Demo 1 – Data Preparation

- **Download the sample data, which is in excel sheet.**
- **write macro to convert data in excel sheet to insert queries**
- **Create a table and execute the seinsert queries in SQL plus**
- **As we are connected to Oracle server, this table is then found in Oracle database**

Demo-1 Connections

Connect Oracle Data Miner Client to Oracle Database

- **Make sure the oracle listener is listening**
- **Database instance „ora478" is started.**
- **The port used is 1521**
- **Give the hostname asoracle.itk.ilstu.edu**

Demo-1

- **Perform the activity, after installations and connections are made.**

DEMO 2

- **Download Oracle 10g on your system and install it**
- **Select the sample schema option during the custom installation**
- **Launch Oracle Data Miner Client**

•In order to use this sample scheme for our activity, we should have the system administrator privileges.

•The username is SH and password is password

Demo -2

•Administrator should perform some grants in sql plus to build this activity. They are alter user sh account unlock;

alter user sh identified by password;

grant create table to sh;

grant create sequence to sh;

grant create session to sh;

grant create view to sh;

grant create procedure to sh;

grant create job to sh;

grant create type to sh;

grant create synonym to sh;

grant execute on ctxsys.ctx_ddl to sh;

Demo-2

The points to be noted before starting the activity are:

•Make sure the oracle listener is started

•Database instance „ORCL‟ isstarted.

- **The port used is 1521**
- **Give the hostname as 127.0.0.1,which is a general hostname.**

Demo-2

- **Finally, the results from the model are published to a table, and this table forms the raw source for the new OLAP product dimension.**
- **At this point there is no information relating to revenue, costs or quantity. So, we need to extend the activity beyond association analys is to OLAP.**

OLAP

- **We have to correctly format the results obtained from Association analysis for dimension mapping in OLAP. This can be done using OLAP DML or PL/SQL.**
- **In our activity we create a separate dimension that can hold the results from algorithm. For each dimension we can create Levels, hierarchy OLAP- Analytic workspace**
- **Launch Analytic workspace and give the login details as Username-sh Connection information-127.0.0.1:1521:orclThis connects to Oracle samples chema SH on 1521 port and local host 127.0.0.1 and orcl database instance. Attributes and mappings.**

Demo 3- OLAP Analytic Workspace

- **Perform the activity and show the mappings**

We have shown how Market basket analysis using association rules works in determining the customer buying patterns. This can be further extended using OLAP Analytic workspace as shown in demo-3, to add dimensions and cube to identify other measures like costs, revenue and quantity.