Experiment 04 : Implement Data Mining algorithms for cognitive computing.

Learning Objective : Students should be able to apply Data Mining algorithms for Cognitive Computing and solve problems.

Tool : RapidMiner

Theory :

Data mining algorithms are commonly used in cognitive computing to extract insights from large and complex datasets. Here are some common data mining algorithms that can be used in cognitive computing:

Association Rule Mining: This algorithm is used to find patterns in data that occur together. For example, it can be used to identify which products are commonly purchased together in a store.

Clustering: This algorithm is used to group data points that are similar to each other. For example, it can be used to group customers based on their purchase history.

Decision Trees: This algorithm is used to build a model that predicts an outcome based on a set of conditions. For example, it can be used to predict whether a customer is likely to buy a product based on their demographic information.

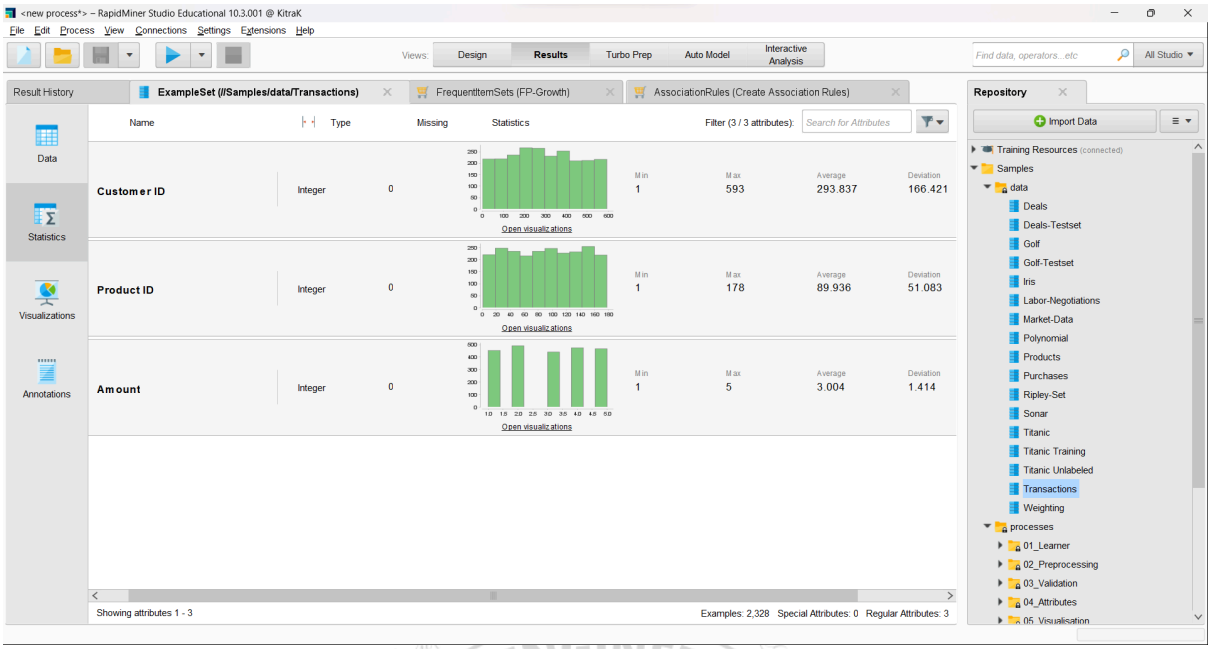
Neural Networks: This algorithm is used to build a model that mimics the behavior of the human brain. For example, it can be used to recognize patterns in images or speech.

Support Vector Machines: This algorithm is used to find a line or a plane that separates two classes of data. For example, it can be used to classify emails as spam or not spam.

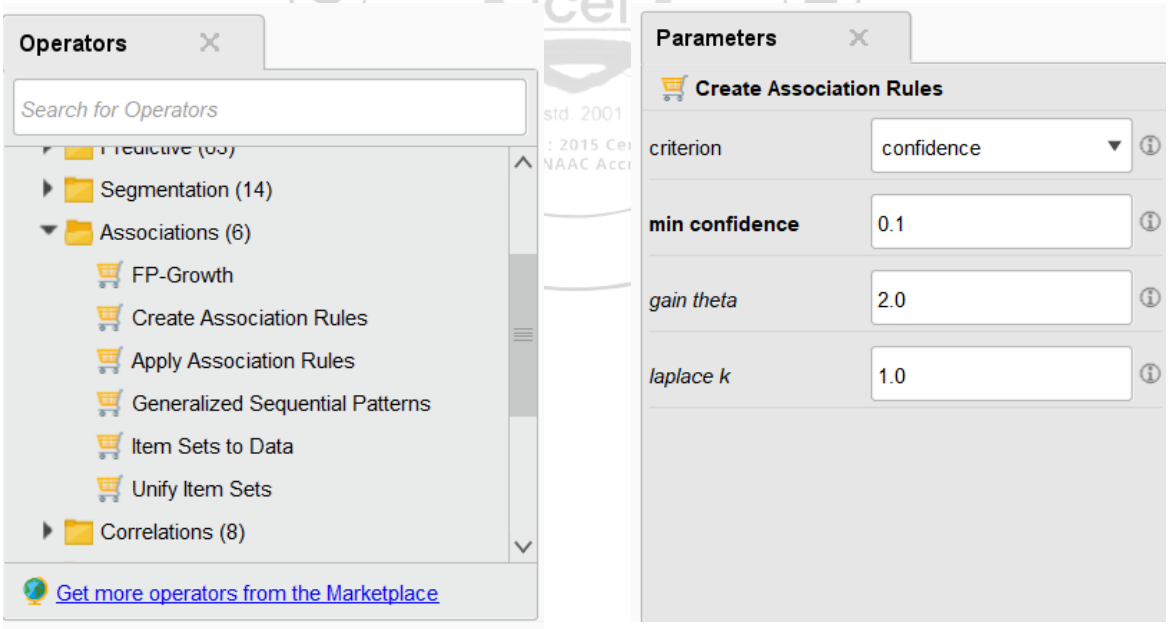
These algorithms can be implemented using various programming languages such as Python, R, and Java. The choice of programming language depends on the specific requirements of the project and the availability of libraries for implementing the algorithms.

Implementation :

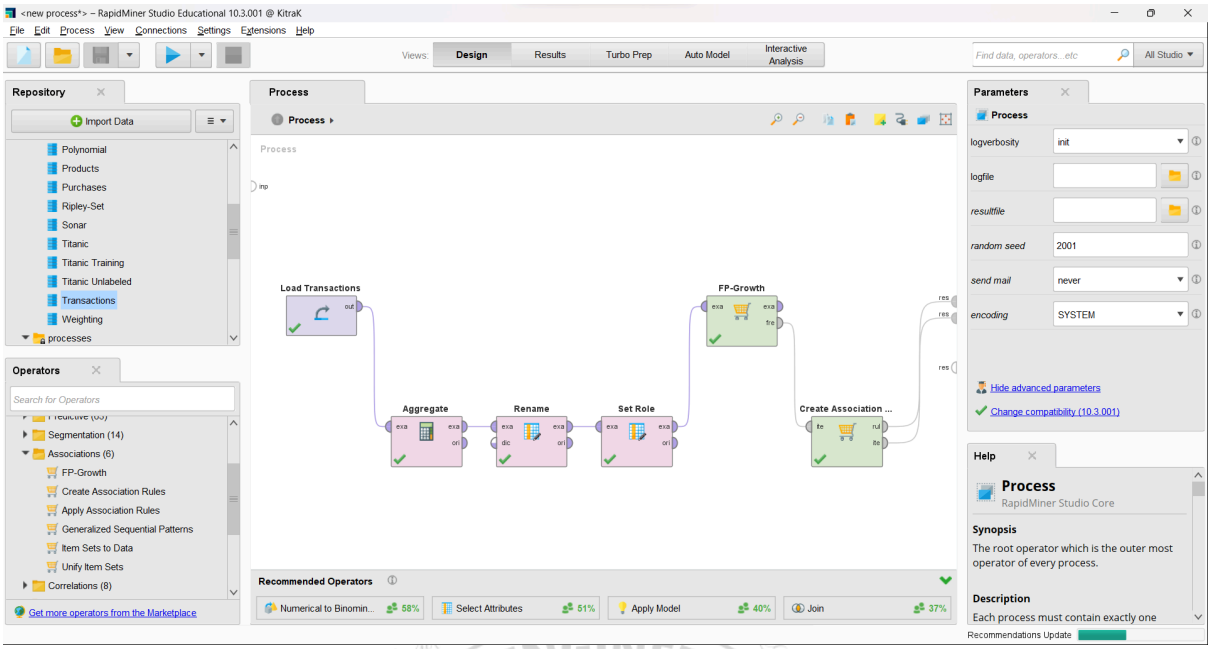
1. **Data Preparation and Exploration :** Before diving into association rule mining, it's important to prepare and explore the transaction dataset. This involves tasks such as data cleaning, handling missing values, and understanding the distribution of the data. Exploratory data analysis can help in understanding the frequency of different products being purchased together and the overall transaction patterns.



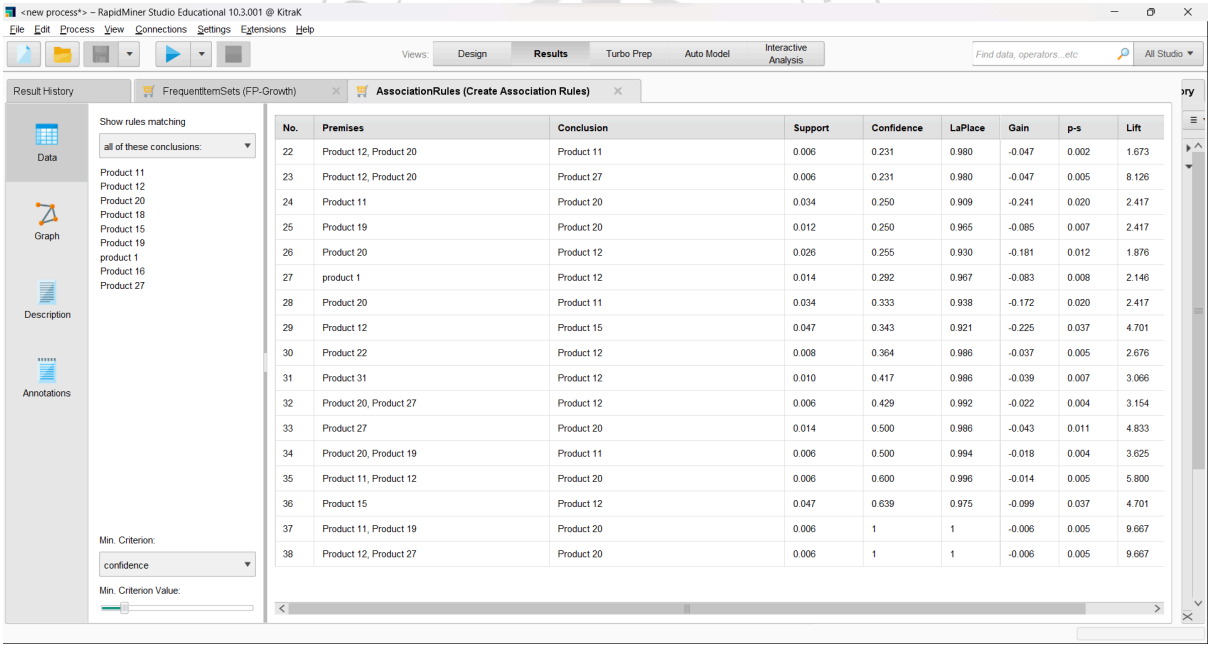
2. **Feature Selection and Model Selection :** In the context of association rule mining, feature selection involves identifying the relevant columns for analysis, such as transaction ID and product ID. Model selection typically refers to choosing the appropriate algorithm for mining association rules, such as Apriori or FP-growth.



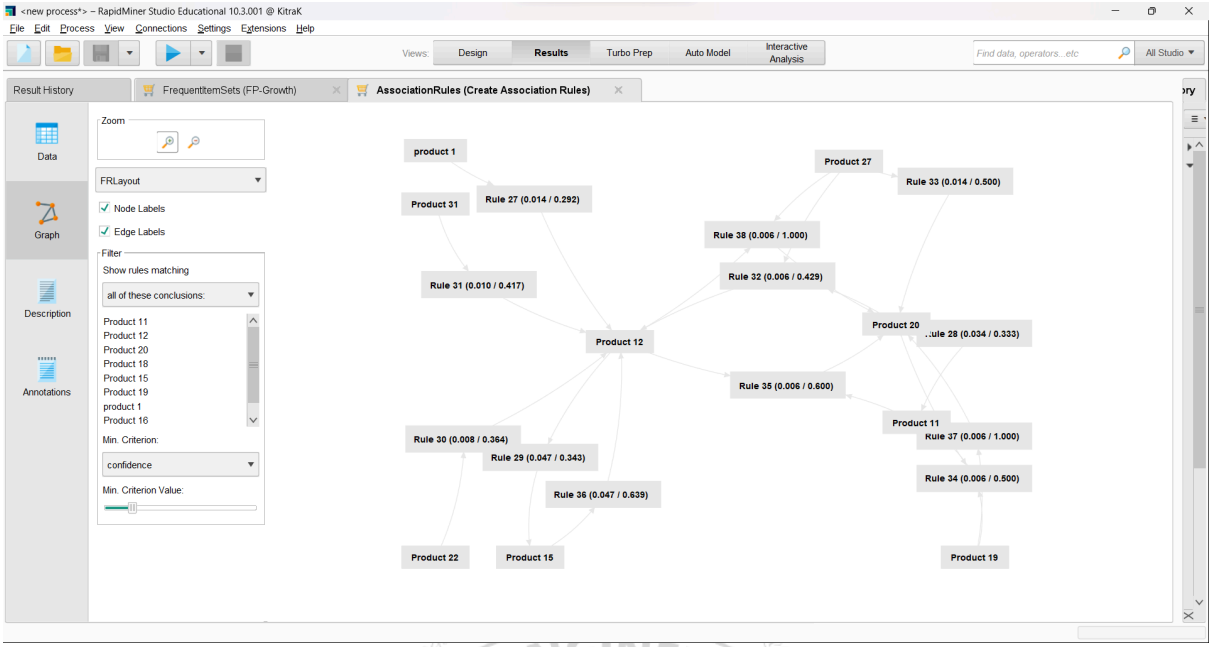
3. **Model Training :** Training an association rule mining model involves applying the selected algorithm to the prepared dataset to identify frequent itemsets and generate association rules based on predefined support and confidence thresholds.



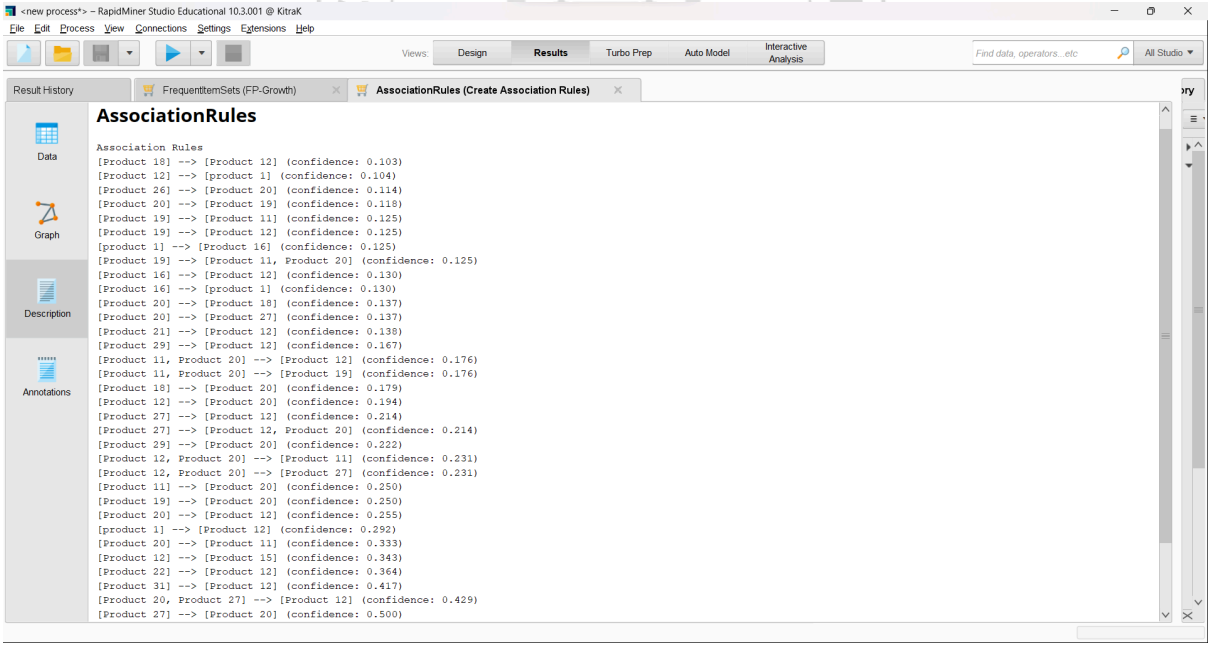
4. **Model Evaluation :** Once the association rules are generated, they can be evaluated based on metrics such as support, confidence, and lift to assess their significance and usefulness in capturing patterns within the transaction data.



No.	Premises	Conclusion	Support	Confidence	LaPlace	Gain	p-s	Lift
22	Product 12, Product 20	Product 11	0.006	0.231	0.980	-0.047	0.002	1.673
23	Product 12, Product 20	Product 27	0.006	0.231	0.980	-0.047	0.005	8.126
24	Product 11	Product 20	0.034	0.250	0.909	-0.241	0.020	2.417
25	Product 19	Product 20	0.012	0.250	0.965	-0.085	0.007	2.417
26	Product 20	Product 12	0.026	0.255	0.930	-0.181	0.012	1.876
27	product 1	Product 12	0.014	0.292	0.967	-0.083	0.008	2.146
28	Product 20	Product 11	0.034	0.333	0.938	-0.172	0.020	2.417
29	Product 12	Product 15	0.047	0.343	0.921	-0.225	0.037	4.701
30	Product 22	Product 12	0.008	0.364	0.986	-0.037	0.005	2.676
31	Product 31	Product 12	0.010	0.417	0.986	-0.039	0.007	3.066
32	Product 20, Product 27	Product 12	0.006	0.429	0.992	-0.022	0.004	3.154
33	Product 27	Product 20	0.014	0.500	0.986	-0.043	0.011	4.833
34	Product 20, Product 19	Product 11	0.006	0.500	0.994	-0.018	0.004	3.625
35	Product 11, Product 12	Product 20	0.006	0.600	0.996	-0.014	0.005	5.800
36	Product 15	Product 12	0.047	0.639	0.975	-0.099	0.037	4.701
37	Product 11, Product 19	Product 20	0.006	1	1	-0.006	0.005	9.667
38	Product 12, Product 27	Product 20	0.006	1	1	-0.006	0.005	9.667



Result and Discussion :



Learning Outcomes : Students should have the ability to

LO 1: Formulate the problem using AI and CC Approach.
LO 2: Solve the problem using Data Mining Algorithm.

Course Outcomes :

CO : Understand and Apply future directions of Cognitive Computing.

Conclusion :

_____

_____

Viva Questions :

Q1. How do you evaluate the performance of a data mining model, and what metrics do you use?

Q2. What is data mining, and how does it relate to cognitive computing?

Q3. How do you prepare data for data mining, and what techniques can you use to select important features?

For Faculty Use

Correction Parameters	Formative Assessment [40%]	Timely completion of Practical [40%]	Attendance / Learning Attitude [20%]	Total
Marks Obtained				