	Hope Foundation's **Finolex Academy of Management and Technology, Ratnagiri** **Department of Information Technology**		
Subject name	**Business Intelligence Lab**		**Subject Code: ITL602**
Class	**TE IT**	**Semester – VI (CBCGS)**	**Academic year: 2018-19 (FH 2019)**
Name of Student			**QUIZ Score :**
Roll No		**Assignment/Experiment No:**	**07**
Title:	**Implementation of Association Mining/ Apriori algorithm using Java.**		

1. Course objectives applicable: LO4

2. Course outcomes applicable: LO4

3. Learning Objectives:

1. To learn working of Apriori algorithm
2. To learn frequent pattern identification.
3. To learn association rule mining.

4. Practical applications of the assignment/experiment: Association rule identification in Market-Basket analysis.

5. Prerequisites:

1. Apriori property and algorithm theoretically.
2. Concepts of association rules

6. Hardware Requirements:

1. PC with minimum 2 GB RAM

7. Software Requirements:

1. Windows / Linux
2. JDK 8.0 / Python 3.6

8. Viva Questions (if any): (Online Quiz will be taken separately batch-wise)

1. What is support and confidence?
2. What is prior probability and posterior probability?
3. What are association rules?
4. How association rules can be used in data mining?

9. Experiment/Assignment Evaluation:

Sr. No.	Parameters	Marks obtained	Out of
1	Technical Understanding (Assessment may be done based on Q & A **or** any other relevant method.)		6
2	Neatness/presentation		2
3	Punctuality		2
Date of performance (DOP)		**Total marks obtained**	**10**
Date of checking (DOC)		**Signature of teacher**	

10. Theory: **<<handwritten work>>**

Apriori is an algorithm for frequent item set mining and association rule learning over transactional databases. It proceeds by identifying the frequent individual items in the database and extending them to larger and larger item sets as long as those item sets appear sufficiently often in the database. The frequent item sets determined by Apriori can be used to determine association rules which highlight general trends in the database: this has applications in domains such as market basket analysis.

Apriori uses a "bottom up" approach, where frequent subsets are extended one item at a time (a step known as *candidate generation*), and groups of candidates are tested against the data. The algorithm terminates when no further successful extensions are found.

Apriori uses breadth-first search and a Hash tree structure to count candidate item sets efficiently. It generates candidate item sets of length **k** from item sets of length **k-1**. Then it prunes the candidates which have an infrequent sub pattern. According to the downward closure lemma, the candidate set contains all frequent **k-length** item sets. After that, it scans the transaction database to determine frequent item sets among the candidates.

Using apriori algorithm Identify all the frequent itemsets with support_count = 2

Itemset
{1,2,3,4}
{1,2,4}
{1,2}
{2,3,4}
{2,3}
{3,4}
{2,4}

11. Performance Steps:

1. Implement a Java/Python program to find frequent itemsets in the given data.

2. Run the program to display the frequent itemsets with support_count = 2.

12. Results:

<< Add the hard-copy of source code and output screen shots >>

13. Learning Outcomes Achieved

1. Understanding of Apriori property and the algorithm.
2. Understanding of frequent pattern identification in given data set.
3. Understanding of association rules in frequent patterns and mining of rules.

14. Conclusion:

1. **Applications of the studied technique in industry:** To perform analysis in transactional data.
2. **Engineering Relevance:** To perform frequent pattern analysis in data analytics.
3. **Skills Developed:** Understanding of frequent patters and association rules.

15. References:

[1] Han, Kamber, "Data Mining Concepts and Techniques", Morgan Kaufmann 3rd Edition.

[2] P. N. Tan, M. Steinbach, Vipin Kumar, "Introduction to Data Mining", Pearson Education.

[3] Michael Berry and Gordon Linoff, "Data Mining Techniques", 2nd Edition Wiley Publications

[4] https://en.wikipedia.org/wiki/Apriori_algorithm

[5] https://en.wikipedia.org/wiki/Association_rule_learning