	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:**	**04**
Title:	**To use WEKA to implement the Association Mining Algorithms**		

1. Course objectives applicable: LO3

2. Course outcomes applicable: LO3

3. Learning Objectives:

1. To learn association rules
2. To understand identifying frequent patterns
3. To understand practical implementation Apriori algorithms

4. Practical applications of the assignment/experiment: To find association of data items and frequent patterns

5. Prerequisites:

1. Concepts of Frequent patterns and association rules
2. Apriori property and its use in the algorithm

6. Hardware Requirements:

1. PC with minimum 2 GB RAM

7. Software Requirements:

1. Windows 8.1 or higher
2. WEKA 3.8 or higher

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

1. What is meant by market-basket analysis?
2. What is association rule?
3. How frequent patterns can be distinguished from supervised learning?
4. What is apriori property?

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>>

Association rule mining, at a basic level, involves the use of machine learning models to analyze data for patterns, or co-occurrence, in a database. It identifies frequent if-then associations, which are called *association rules*.

An association rule has two parts: an antecedent (if) and a consequent (then). An antecedent is an item found within the data. A consequent is an item found in combination with the antecedent.

Association rules are created by searching data for frequent if-then patterns and using the criteria *support* and *confidence* to identify the most important relationships. Support is an indication of how frequently the items appear in the data. Confidence indicates the number of times the if-then statements are found true. A third metric, called *lift*, can be used to compare confidence with expected confidence.

With the Apriori algorithm, candidate itemsets are generated using only the large itemsets of the previous pass. The large item-set of the previous pass is joined with itself to generate all itemsets with a size that's larger by one. Each generated item-set with a subset that is not large is then deleted. The remaining itemsets are the candidates. The Apriori algorithm considers any subset of a frequent item-set to also be a frequent item-set. With this approach, the algorithm reduces the number of candidates being considered by only exploring the itemsets whose support count is greater than the minimum support count.

In data mining, association rules are useful for analyzing and predicting customer behavior. They play an important part in customer analytics, market basket analysis, product clustering and catalog design and store layout.

11. Performance Steps:

1. Start WEKA
2. Open the data set (ARFF file)
3. Select the Association rule mining technique (apriori) and run the trial
4. Record the reading
5. Identify the frequent item sets with given conditions.

12. Results:

<< Add the hard-copy of output screen shots >>

13. Learning Outcomes Achieved

1. Understanding of Association Rules.
2. Understanding the techniques for identifying frequent patterns.
3. Understanding of apriori algorithm.

14. Conclusion:

1. **Applications of the studied technique in industry:** To perform market-basket analysis.
2. **Engineering Relevance:** To identify frequent patterns analysis results for decision making.
3. **Skills Developed:** Understanding of association rules mining techniques.

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/Association_rule_learning

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