

DATA MINING LAB

Course Code: 22CD1104

L T P C
0 0 3 1.5

COURSE OUTCOMES: At the end of the Course the student shall be able to

- CO1:** Use different features of WEKA tool (L3)
- CO2:** preprocess the data for mining (L3)
- CO3:** determine association rules (L3)
- CO4:** model various classifiers. (L3)
- CO5:** examine clusters from the available data (L3)

-

LIST OF PROGRAMS:

(Any 12 programs from the following to be performed)

1. Basics of WEKA tool a. Investigate the Application interfaces. b. Explore the default datasets.
2. Pre-process a given dataset based on the following: a. Attribute Selection b. Handling Missing Values
3. Pre-process a given dataset based on the following: a. Discretization b. Eliminating Outliers
4. Create a dataset in ARFF (Attribute-Relation File Format) for any given dataset and perform Market-Basket Analysis.
5. Generate Association Rules using the Apriori algorithm.
6. Generate Association Rules using the FP-Growth algorithm.
7. Build a Decision Tree using ID3 algorithm.
8. Demonstrate classification process on a given dataset using Naïve Bayesian Classifier.
9. Demonstrate classification process on a given dataset using Rule based Classifier.
10. Demonstrate classification process on a given dataset using Nearest neighbor Classifier.
11. Build a distance matrix of the given data using various distance measures.
12. Cluster the given dataset by using the k-Means algorithm and visualize the cluster mean values and standard deviation of dataset attributes.
13. Cluster the given dataset using a hierarchical clustering algorithm.
14. Cluster the given dataset using the DBSCAN algorithm.
15. Detect anomalies using any clustering algorithm.
16. Implement density based outlier detection.

REFERENCE BOOKS:

1. Pang-Ning Tan, Michael Steinbach, Vipin Kumar, *Introduction to Data Mining*, 2nd edition, Pearson education, 2018.
2. Jiawei Han & Micheline Kamber, *Data Mining, "Concepts and Techniques"*, 3rd edition, Morgan Kaufmann Publishers, 2012.
3. Margaret H Dunham, *Data Mining Introductory and advanced topics*, 6th edition, Pearson Education, 2009
4. Arun K Pujari, *Data Mining Techniques*, 1st edition, University Press, 2005.
5. GK Gupta, *Introduction to Data Mining with Case Studies*, 3rd edition, Prentice Hall, 2014.

6. [K. P. Soman](#), [ShyamDiwakar](#), [V. Ajay](#), *DataMiningTheoryandPractice*, 1st edition, PHI, 2006.

WEB REFERENCES:

1. www.cs.waikato.ac.nz/ml/weka/downloading.html