RSET
RAJAGIRI SCHOOL OF
ENGINEERING & TECHNOLOGY
(AUTONOMOUS)

B. TECH.
IN
COMPUTER SCIENCE & BUSINESS STUDIES

SEMESTER V
(2021 ADMISSIONS)

SYLLABUS

RSET
RAJAGIRI SCHOOL OF ENGINEERING & TECHNOLOGY (Autonomous)

Rajagiri Valley, Kakkanad,
Kochi 682 039, Kerala, INDIA
www.rajagiritech.ac.in

COURSE CODE	COURSE NAME	L	T	P	CREDIT	YEAR OF INTRODUCTION
101009/IT522S	Machine Learning Lab	0	0	2	1	2021

1. Preamble

Database Management Systems Lab course is intended to provide students a hands-on experience in Machine Learning concepts. It also provides a strong formal foundation in various machine learning techniques and practice to the students. It gives an exposure to design and develop applications.

2. Prerequisite

Nil

3. Syllabus

Lab Sessions:

1. Introduction to WEKA
2. Introduction to R
3. Classification of some public domain datasets in UCI ML repository

Mini projects in the Lab:

1. Implementation of one clustering algorithm
2. Implementation of one association rule mining algorithm
3. Implementation of one anomaly detection algorithms
4. Implementation of EM algorithm for some specific problem

List of Experiments

Experiment No.1: Explore WEKA Machine Learning Toolkit

- Downloading and/or installation of WEKA data mining toolkit,
- Understand the features of WEKA toolkit such as Explorer, Knowledge Flow interface, Experimenter, command-line interface.

• Navigate the options available in the WEKA (ex. Select attributes panel, Preprocess panel, Classify panel, Cluster panel, Associate panel and Visualize panel)

Experiment No.2:

• Study the arff file format
• Explore the available data sets in WEKA
• Load a data set (ex. Weather dataset, Iris dataset, etc)
• Load each dataset and observe the following:

a. List the attribute names and their types
b. Number of records in each dataset
c. Identify the class attribute (if any)
d. Plot Histogram
e. Determine the number of records for each class
f. Visualize the data in various dimensions

Experiment No.3: Implementation of various plots R

Experiment No.4: Data manipulation with R, Pre-processing and preparing a dataset for Classification.

Experiment No.5: Demonstrate the following Classification algorithms using some public domain datasets in UCI ML repository and compute the accuracy of the classifier, considering few test datasets.

i. Naive Bayes classification
ii. Decision Tree
iii. Random Forests
iv. Ensembles of classifiers including bagging and boosting

Experiment 6: Demonstrate the following Classification algorithms using some public domain datasets in UCI ML repository and compute the accuracy of the classifier, considering few test datasets.

i. k-Nearest neighbor classification
ii. Support Vector Machines
iii. Artificial neural networks (back propagation)

Mini projects in the Lab:

1. Implementation of clustering algorithms
2. Implementation of association rule mining algorithm (Apriori)
3. Implementation of anomaly detection algorithms
4. Implementation of EM algorithm for some specific problem

4. Text Books

1. Christopher M. Bishop, *Pattern Recognition and Machine Learning*, Springer, 2006.
2. Ethem Alpaydın, Introduction to Machine Learning (Adaptive Computation and Machine Learning), MIT Press, 2004.

5. Reference Books

1. R.O. Duda, P.E. Hart, D.G. Stork, *Pattern Classification*, 2/e, Wiley, 2001.
2. C. Bishop, Pattern Recognition and Machine Learning, Springer, 2007.
3. E. Alpaydin, *Introduction to Machine Learning*, 3/e, Prentice-Hall, 2014.
4. Rostamizadeh, A. Talwalkar, M. Mohri, *Foundations of Machine Learning*, MIT Press.
5. Andrew R Webb, Keith D Kopsey, Statistical Pattern Recognition, 3/e, Wiley, 2011.

6. Course Outcomes

After the completion of the course the student will be able to

CO1: Familiarize the usage of Machine Learning tools such as WEKA and R.

CO2: Implement and apply machine learning algorithms to analyze the complex problems.

CO3: Select and apply appropriate algorithms for solving a of real-world problems.

CO4: Understand the associations in data and infer knowledge for future predictions.

.

6. Mapping of Course Outcomes with Program Outcomes

	PO1	PO2	PO3	PO4	PO5	PO6	PO7	PO8	PO9	PO10	PO11	PO12
CO1	1	2	2	2	3	1	-	-	-	-	-	1
CO2	2	2	2	3	3	1	-	-	-	-	-	1
CO3	3	2	2	3	3	1	-	-	-	-	-	1
CO4	2	3	3	2	3	1	-	-	-	-	-	1

7. Mark Distribution

Total Marks	Continuous Internal Evaluation (CIE)	End Semester Examination (ESE)	ESE duration
150	75	75	2.5 hours

8. Continuous Internal Evaluation Pattern

Attendance :15 marks
Continuous Assessment :30 marks
Internal Test (Immediately before the second series test) :30 marks

9. End Semester Examination Pattern

The following guidelines should be followed regarding award of marks

1. a. Algorithm : 10 Marks
   b. Implementation of Project : 15 Marks
2. Performance, result and inference
   (usage of application tool and troubleshooting) : 25 Marks
3. Viva voce : 20 Marks
4. Record : 5 Marks

General instructions: Practical examination to be conducted immediately after the second series test covering the entire syllabus given below. Evaluation is a serious process that is to be conducted under the equal responsibility of both the internal and external examiners. The number of candidates evaluated per day should not exceed 20. Students shall be allowed for the University examination only on submitting the duly certified record. The external examiner shall endorse the record.

****