# NATIONAL INSTITUTE OF TECHNOLOGY, TIRUCHIRAPPALLI

## DEPARTMENT OF COMPUTER APPLICATIONS

| COURSE PLAN – PART I | | | |
|---|---|---|---|
| **Name of the programme and specialization** | **MCA** | | |
| **Course Title** | **DATA MINING LAB** | | |
| **Course Code** | **CA 707** | **No. of Credits** | **2** |
| **Course Code of Pre-requisite subject(s)** | **CA 721** | | |
| **Session** | **July 2021** | **Section (if, applicable)** | **A & B** |
| **Name of Faculty** | **Dr. Balaji Ganesh R** | **Department** | **Computer Applications** |
| **Official Email** | **rbalaji@nitt.edu** | **Telephone No.** | **8220037222** |
| **Name of PAC Chairman** | **Dr. (Mrs.) B. Janet, Assistant Professor** | | |
| **Official E-mail** | **janet@nitt.edu** | **Telephone No.** | **0431-2503742** |
| **Course Type (**please tick appropriately**)** | ☑ **Core course** | ☐ **Elective course** | |

**Syllabus (approved in BoS)**

Exercises to

- Understand the data sets and data preprocessing using ETL tools
- Demonstrate the working of algorithms for data mining tasks such association rule mining, classification, clustering and regression

**COURSE OBJECTIVES**

To make the students

- Familiarize with the ETL (Extract, Transform, Load) tools like Weka for Data exploration and processing
- Demonstrate the various Data mining tasks with Weka Explorer and Sample Datasets

**MAPPING OF COs with POs**

| **Course Outcomes** | **Programme Outcomes (PO)** (Enter Numbers only) |
|---|---|
| 1. Work with ETL tools | **1, 2, 5** |
| 2. Demonstrate the classification, clustering and etc. in large data sets. | **1, 2, 3** |
| 3. Ability to add mining algorithms as a component to the exiting tools | **1, 2, 3, 4** |
| 4. Ability to apply mining techniques for realistic data. | **1, 2, 8, 9** |

## COURSE PLAN – PART II

### COURSE OVERVIEW

This course makes students to install, configure networks and build the networks according to the requirement and implement the network principles using Simulation software Packet Tracer

### COURSE TEACHING AND LEARNING ACTIVITIES

| S. No. | Week/Contact Hours | Topic | Mode of Delivery |
|---|---|---|---|
| 1 | Week 1 | Introduction to Weka ETL tool, Data Loading, Visualization | PPT, Demo in MS Teams using Weka |
| 2 | Week 2 | Data Preprocessing, Cleaning, Discretization | PPT, Demo in MS Teams using Weka |
| 3 | Week 3 | Association Rule Mining using Apriori Algorithm | PPT, Demo in MS Teams using Weka |
| 4 | Week 4 | Association Rule Mining using predictiveApriori Algorithm | PPT, Demo in MS Teams using Weka |
| 5 | Week 5 | Classification Rule process on sample dataset (ZeroR classifier) | PPT, Demo in MS Teams using Weka |
| 6 | Week 6 | Classification Rule process on sample dataset (Decision Tree- j48 algorithm) | PPT, Demo in MS Teams using Weka |
| 7 | Week 7 | Classification Rule process on sample dataset (Decision Tree-id3 algorithm) | PPT, Demo in MS Teams using Weka |
| 8 | Week 8 | Demonstration of Naïve Bayes algorithm | PPT, Demo in MS Teams using Weka |
| 9 | Week 9 | Prediction using Regression Model (Linear Regression, KNN) | PPT, Demo in MS Teams using Weka |
| 10 | Week 10 | Clustering using simple K-means, EM, Hierarchical Clustering | PPT, Demo in MS Teams using Weka |

**COURSE ASSESSMENT METHODS (**shall range from 4 to 6**)**

| Sl. No. | Mode of Assessment | Week/Date | Duration | % Weightage | Mode of Conduct |
|---|---|---|---|---|---|
| 1 | Exercises Demonstration – I | 6th Week | - | 15 | MS Teams / Cisco Packet Tracer |
| 2 | Exercises Demonstration – II | 10th Week | - | 15 | MS Teams / Cisco Packet Tracer |
| 3 | Laboratory Report | Every week | - | 10 | Google Forms |
| 4 | Online Assessment – Objective type | 10th Week | 1 hour | 15 | Instructure Canvas |
| 5 | Project | 11th Week | - | 15 | MS Teams |
| CPA | Compensation Assessment* | 12th Week | 1 hour | 15 | Instructure Canvas |
| 6 | Final Assessment & Viva Voce Examination | 12th Week | 3 hours | 30 | MS Teams |

***mandatory; refer to guidelines on page 4**

**COURSE EXIT SURVEY (**mention the ways in which the feedback about the course shall be assessed)

- The students through the class representative may give their feedback at any time to the faculty which will be duly addressed.
- The students may give their feedback during class committee meetings.

**COURSE POLICY (**including compensation assessment to be specified**)**

**Compensation Assessment**

One compensation assessment for absentees in assessment (other than the final assessment) is mandatory. Only genuine cases of absence shall be considered.

**ATTENDANCE POLICY** (A uniform attendance policy as specified below shall be followed)

- At least 75% attendance in each course is mandatory.
- A maximum of 10% shall be allowed under On Duty (OD) category.
- Students with less than 65% of attendance shall be prevented from writing the final assessment and shall be awarded 'V' grade.

**ACADEMIC DISHONESTY & PLAGIARISM**

- Possessing a mobile phone, carrying bits of paper, talking to other students, copying from others during an assessment will be treated as punishable dishonesty.
- Zero mark to be awarded for the offenders. For copying from another student, both students get the same penalty of zero mark.
- The departmental disciplinary committee including the course faculty member, PAC chairperson and the HoD, as members shall verify the facts of the malpractice and award the punishment if the student is found guilty. The report shall be submitted to the Academic office.
- The above policy against academic dishonesty shall be applicable for all the programmes.

**ADDITIONAL INFORMATION, IF ANY**

- Weka Tool can be downloaded from https://waikato.github.io/weka-wiki/downloading_weka/
- **Reference Book:** "The Weka Workbench- Data Mining: Practical Machine Learning Tools and Techniques", Eibe Frank, Mark A.Hall, Morgan Kaufmann, Fourth Edition, 2016
- **NPTEL resources:** Learning Analytics Tools, By Prof. Ramkumar Rajendran, IIT Bombay available at https://nptel.ac.in/courses/127/101/106101224

**FOR APPROVAL**

| | | |
|---|---|---|
| **Course Faculty** | **CC- Chairperson** | **HOD** |
| **Dr.R.Balaji Ganesh** | **Dr.(Mrs.) B.Janet,** | **Prof. Dr.P.J.A. Alphonse** |
| **PDF** | **Assistant Professor** | **Professor and Head** |

## Guidelines

a) The number of assessments for any theory course shall range from 4 to 6.

b) Every theory course shall have a final assessment on the entire syllabus with at least 30% weightage.

c) One compensation assessment for absentees in assessments (other than final assessment) is mandatory. Only genuine cases of absence shall be considered.

d) The passing minimum shall be as per the regulations.

| B.Tech. Admitted in | | | | P.G. |
|---|---|---|---|---|
| 2018 | 2017 | 2016 | 2015 | |
| 35% or (Class average/2) whichever is greater. | | (Peak/3) or (Class Average/2) whichever is lower | | 40% |

e) Attendance policy and the policy on academic dishonesty & plagiarism by students are uniform for all the courses.

f) Absolute grading policy shall be incorporated if the number of students per course is less than 10.

g) Necessary care shall be taken to ensure that the course plan is reasonable and is objective.