NATIONAL INSTITUTE OF TECHNOLOGY, TIRUCHIRAPPALLI

DEPARTMENT OF COMPUTER APPLICATIONS

COURSE PLAN – PART I

Name of the programme and specialization	MASTER OF COMPUTER APPLICATIONS		
Course Title	DATA MINING LAB		
Course Code	CA707	No. of Credits	2
Course Code of Pre-requisite subject(s)	CA721		
Session	July 2020	Section (if, applicable)	B
Name of Faculty	Dr.P.CHITRA	Department	CA
Official Email	chitrap@nitt.edu	Telephone No.	9715258867
Name of Course Coordinator(s) (if, applicable)	Dr.B.JANET		
Official E-mail	janet@nitt.edu	Telephone No.	0431 2503741
Course Type (please tick appropriately)	✓ Core course	☐ Elective course	

Syllabus (approved in BoS)
- Data sets and data preprocessing using ETL Tool, association rule mining, classification, clustering and regression

COURSE OBJECTIVES
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

MAPPING OF COs with POs

Course Outcomes	Programme Outcomes (PO) (Enter Numbers only)
• Work with ETL tools	PO1
• Demonstrate the classification, clustering and etc. in large data sets	PO2
• Ability to add mining algorithms as a component to the exiting tools.	PO3
• Ability to apply mining techniques for realistic data.	PO4

NATIONAL INSTITUTE OF TECHNOLOGY, TIRUCHIRAPPALLI

COURSE OVERVIEW

- This course is designed to expand students' knowledge and skills gained in database management courses and look in depth at data warehousing and data mining methods.
- The course examines the database architecture and technologies required for solving complex problems of data and information management, information retrieval, and knowledge discovery facing modern organizations. Case studies of organizations using these technologies to support business intelligence gathering and decision making are examined.
- This course also provides hands-on experience with state-of-the-art data warehousing and data mining methods and tools.

COURSE TEACHING AND LEARNING ACTIVITIES (Add more rows)

S.No.	Week/Contact Hours	Topic	Mode of Delivery
1	Week 1	Introduction about data mining and WEKA tool (Demo Class)	PPT through video conferencing in MS Teams
2	Week 2	Basic Exercises on Data Mining	Google Classroom
3	Week 3	Data using ETL tools	Google Classroom
4	Week 4	Data using Preprocessing	Google Classroom
5	Week 5	Implementation of measures of proximity	Google Classroom
6	Week 6	Data using the trees generation	Google Classroom
7	Week 7	Demo on Classification tools	Google Classroom
8	Week 8	Learning and implementing k-means clustering	Google Classroom

9	Week 9	Learning Naïve and Decision Tress classifier in WEKA	Google Classroom
10	Week 10	Implementation of outlier detection algorithms (nearest neighbor and Mahalanobis)	Google Classroom
11	Week 11	Feature Selection, Cleaning, and Preprocessing to Construct an Input from Data	Google Classroom
12	Week 12	Calculating Proximity of Two Binary Object Vectors With Simple Matching, Jaccard Similarity, Cosine Similarity	Google Classroom
13	Week 13	Correlation Analysis for Two Features	Google Classroom

COURSE ASSESSMENT METHODS (shall range from 4 to 6)

S.No.	Mode of Assessment	Week/Date	Duration	% Weightage
1	Lab Assessment 1	—	2 hours	10
2	Lab Assessment 2	—	2 hours	10
3	Lab Assessment 3	—	2 hours	10
4	Lab Assessment 4	—	2 hours	10
CPA	Compensation Assessment*	—	1 hour	10
5	Project	—	-	20
6	Final Assessment *	—	2 hours	30

*mandatory; refer to guidelines on page 4

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

- Understand data mining principles and techniques: Introduce DM as a cutting edge business intelligence method and acquaint the students with the DM techniques for building competitive advantage through proactive analysis, predictive modelling, and identifying new trends and behaviours.

Learning objectives include:
- Building basic terminology.
- Learning how to gather and analyze large sets of data to gain useful business understanding. 4. Learning how to produce a quantitative analysis report/memo with

the necessary information to make decisions.
- Describing and demonstrating basic data mining algorithms, methods, and tools
- Identifying business applications of data mining
  Overview of the developing areas - web mining, text mining, and ethical aspects of data mining.

COURSE POLICY (including compensation assessment to be specified)

- Practical exposure on implementation of well-known data mining tasks.
- Exposure to real life data sets for analysis and prediction.
- Learning performance evaluation of data mining algorithms in a supervised and an unsupervised setting.
- Handling a small data mining project for a given practical domain.

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

FOR APPROVAL

Course Faculty _____ CC- Chairperson _____ HOD _____

Dr.P.J.A. Alphonse

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.