10. DATA MINING CONCEPTS AND TECHNIQUES

Discipline	Computer Science				
Course Code	UK3DSCCSC209				
Course Title	**DATA MINING CONCEPTS AND TECHNIQUES**				
Type of Course	DSC				
Semester	III				
Academic Level	2				
Course Details	Credit	Lecture per week	Tutorial per week	Practical per week	Total Hours/Week
	4	3 hours	-	2 hours	5 hours
Pre-requisites	Nil				
Course Summary	This course, Data Mining Concepts and Techniques, introduces the student to the world of data and the various methods and models used in transforming, Classifying and analysing data.				

Detailed Syllabus:

Module	Unit	Content	Hrs (T+P)
I	**Basics of Data Mining**		**15**
	1	Definition of data, Information and Data analysis	
	2	Fundamentals of Data Mining , Data mining stages.	
	3	Applications of Data mining, Data Pre –processing.	
	4	Need for Pre-processing the Data, Data Cleaning.	
II	**Data Integration and Transformation**		**15**
	5	Data Reduction	
	6	Introduction to data warehouse;	

	7	Business Intelligence.	
III	**Classification Models**		**15**
	9	Classification and Prediction	
	10	Issues Regarding Classification and Prediction.	
	11	Classification by Decision Tree Induction	
	12	KNN, Bayesian Classification	
	13	Neural networks	
	14	Support VectorMachines.	
IV	**Association Rules Mining**		**15**
	15	Mining Frequent Patterns	
	16	Associations and Correlations	
	17	Efficient and Scalable Frequent Itemset Mining Methods	
	18	Mining various kinds of Association Rules	
	19	From Association Mining to Correlation Analysis.	
V		**Flexi module(Not included for End Semester Examination)**	**15**
	20	Understanding the Hadoop	
	21	Distributed File System (HDFS) Getting Data into Hadoop	
	22	Understanding Data Processing in Hadoop	

TEXT BOOK

1. Han, J., Pei, J., & Kamber, M. (2011). *Data mining: concepts and techniques*. Elsevier.

REFERENCES

2. Hall M, Frank E, Holmes G, P.fahringer B. Reutemann P &Witten, I.H(2009), The WEKA data mining software: an update. *ACM SIGKDD explorations newsletter*, *11*(1), 10-18.
3. Gupta, G.K (2014) Introduction to Data Mining with Case Studies, 2014, Prentice Hall India.

Lab Exercises

Practical using Python/WEKA Tool

1. List all the categorical (or nominal) attributes and the real-valued attributes separately

2. Calculate: mean, median, mode
3. Demonstration of data preprocessing on dataset
4. Demonstration of data preprocessing on dataset based on missing values
5. Demonstration of Association rule process on dataset using Apriori Algorithm
6. Demonstration of classification rule process on dataset using decision tree induction
7. Demonstration of classification rule process on dataset using naive bayes algorithm
8. Demonstration of clustering rule process on dataset using various clustering methods
9. Practising outlier detection in clustering on dataset

Course Outcomes

No.	Upon completion of the course the graduate will be able to	Cognitive Level	PSO addressed
CO1	Cite the fundamentals of data mining	U	PSO- 1,
CO2	Summarise about pre-processing techniques	U	PSO- 1,2
CO3	Illustrate the data integration, transformation and reduction techniques	Ap	PSO- 1,2,3
CO4	Experiment with classification and prediction models.	Ap	PSO- 1,2,3

R-Remember, U-Understand, Ap-Apply, An-Analyse, E-Evaluate, C-Create

Note: 1 or 2 COs/module

Name of the Course: DATA MINING CONCEPTS AND TECHNIQUES

Credits: 3:0:1 (Lecture: Tutorial: Practical)

CO No.	CO	PO/PSO	Cognitive Level	Knowledge Category	Lecture (L)/ Tutorial (T)	Practical (P)
1	Cite the fundamentals of data mining	PO-3,6,7 PSO-1,	U	F, C. P	L	P
2	Summarize about pre-processing techniques	PO-3,6,7	U	F, C, P	L	P

		PSO-1,2				
3	Illustrate the data integration, transformation and reduction techniques	PO-3,5,6,7 PSO-1,2,3	Ap	F, C, P	L	P
4	Experiment with classification and prediction models.	PO-3,5,6,7 PSO-1,2,3	Ap	F, C, P	L	P

F-Factual, C- Conceptual, P-Procedural, M-Metacognitive

Mapping of COs with PSOs and POs:

	PO1	**PO2**	**PO3**	**PO4**	**PO5**	**PO6**	**PO7**	**PO8**	**PSO1**	**PSO2**	**PSO3**	**PSO4**
CO 1	-	-	3	-	-	3	3	-	2	1	-	-
CO 2	-	-	3	-	-	3	3	-	2	1	-	-
CO 3	-	-	3	-	1	3	3	-	2	1	2	-
CO 4	-	-	3	-	1	3	3	-	2	1	2	-

Correlation Levels:

Level	**Correlation**
-	Nil
1	Slightly / Low
2	Moderate / Medium