LAKIREDDY BALI REDDY COLLEGE OF ENGINEERING
(AUTONOMOUS)
Accredited by NAAC & NBA (CSE, IT, ECE, EEE & ME)
Approved by AICTE, New Delhi and Affiliated to JNTUK, Kakinada
L.B.Reddy Nagar, Mylavaram-521230, Krishna Dist, Andhra Pradesh, India
DEPARTMENT OF INFORMATION TECHNOLOGY

COURSE HANDOUT

PART-A

Name of Course Instructor : Mr.Michael Sadgun Rao Kona
Course Name & Code : Data Mining using Python Lab & 20CS58
L-T-P Structure : 0-0-3 Credits : 3
Program/Sem/Sec : B.Tech., IT., IV-Sem., B section A.Y : 2021-22

PRE-REQUISITE : Python Programming

COURSE EDUCATIONAL OBJECTIVES (CEOs): The objective of this lab is to Practical exposure on implementation of well-known data mining algorithms and Learning performance evaluation of data mining algorithms in a supervised and an unsupervised setting.

COURSE OUTCOMES (COs): At the end of the course, students are able to

CO 1	Apply preprocessing techniques on real world datasets.(**Apply-L3**)
CO 2	Apply apriori algorithm to generate frequent itemsets.**(Apply L3)**
CO 3	Apply Classification and clustering algorithms on different datasets.(**Apply L3**)
CO 4	Improve individual / teamwork skills, communication & report writing skills with ethical values.

COURSE ARTICULATION MATRIX (Correlation between COs, POs & PSOs):

COs	PO 1	PO2	PO 3	PO 4	PO 5	PO 6	PO 7	PO 8	PO 9	PO 10	PO 11	PO 12	PSO1	PSO2	PSO3
CO1	-	-	2	-	1	-	-	-	-	-	-	-	3	1	
CO2	-	-	-	2	1	-	-	-	-	-	-	-	-	3	1
CO3	-	-	-	2	1	-	-	-	-	-	-	-	-	3	2
CO4	-	-	-	-	-	-	-	2	2	2	-	-	-	-	-

Note: Enter Correlation Levels **1** or **2** or **3.** If there is no correlation, **put '-'**
1- Slight (Low), **2 –** Moderate (Medium), **3 -** Substantial (High).

Web Resources:

1. https://analyticsindiamag.com/data-pre-processing-in-python/
2. https://towardsdatascience.com/decision-tree-in-python-b433ae57fb93
3. https://towardsdatascience.com/calculate-similarity-the-most-relevant-metrics-in-a-nutshell-9a43564f533e
4. https://www.springboard.com/blog/data-mining-python-tutorial/
5. https://medium.com/analytics-vidhya/association-analysis-in-python-2b955d0180c
6. https://www.datacamp.com/community/tutorials/naive-bayes-scikit-learn
7. https://www.analyticsvidhya.com/blog/2019/05/beginners-guide-hierarchical-clustering/
8. https://towardsdatascience.com/dbscan-algorithm-complete-guide-and-application-with-python-scikit-learnd690cbae4c5d

PART-B

COURSE DELIVERY PLAN (LESSON PLAN):

S NO	Date (Tentative)	Actual Date	Topics to be covered	Teaching Learning Methods	HOD Signature
1	08.03.2022, 15.03.2022, 22.03.2022		1. Demonstrate the following data preprocessing tasks using python libraries. a) Loading the dataset b) Identifying the dependent and independent variables. c) Dealing with missing data	**TLM4**	
2	29.03.2022, 05.04.2022		Demonstrate the following data preprocessing tasks using python libraries. a) Dealing with categorical data. b) Scaling the features. c) Splitting dataset into Training and Testing Sets	**TLM4**	
3	12.04.2022		Demonstrate the following Similarity and Dissimilarity Measures using python a) Pearson's Correlation b) Cosine Similarity c) Jaccard Similarity d) Euclidean Distance e) Manhattan Distance	**TLM4**	
4	19.04.2022		Build a model using a linear regression algorithm on any dataset.	**TLM4**	
Mid - I Examinations 25.04.2022 to 30.04.2022					
5	03.05.2022		Build a classification model using Decision Tree algorithm on iris dataset	**TLM4**	
6	10.05.2022		Apply Naïve Bayes Classification algorithm on any dataset	**TLM4**	
7	17.05.2022		Generate frequent item sets using Apriori Algorithm in python and also generate association rules for any market basket data.	**TLM4**	
8	24.05.2022		Apply the K- Means clustering algorithm on any dataset.	**TLM4**	
9	31.05.2022		Apply Hierarchical Clustering algorithms on any dataset.	**TLM4**	
10	07.06.2022		Apply the DBSCAN clustering algorithm on any dataset.	**TLM4**	
11	14.06.2022		Internal Exam	**TLM4**	

Teaching Learning Methods			
TLM1	Chalk and Talk	**TLM4**	Demonstration (Lab/Field Visit)
TLM2	PPT	**TLM5**	ICT (NPTEL/Swayam Prabha/MOOCS)
TLM3	Tutorial	**TLM6**	Group Discussion/Project

PART-C

EVALUATION PROCESS (R17 Regulations):

Evaluation Task	Marks
Day-to-day work	D1=05
Record	R1=05
Internal Test	IT1=5
Continuous Internal Evaluation (CIE)=D1+R1+IT1	15
Procedure/Algorithm	P1=5
Experimentation/Program execution	E1=10
Observations/Calculations/Validation	O1=10
Result/Inference	R1=5
Viva voce	V1=5
Semester End Examination (SEE)= P1+ E1+ O1+ V1	30
Total Marks = CIE+SEE	50

PART-D

PROGRAMME OUTCOMES (POs):

PO 1	**Engineering knowledge**: Apply the knowledge of mathematics, science, engineering fundamentals, and an engineering specialization to the solution of complex engineering problems.
PO 2	**Problem analysis**: Identify, formulate, review research literature, and analyze complex engineering problems reaching substantiated conclusions using first principles of mathematics, natural sciences, and engineering sciences.
PO 3	**Design/development of solutions**: Design solutions for complex engineering problems and design system components or processes that meet the specified needs with appropriate consideration for the public health and safety, and the cultural, societal, and environmental considerations.
PO 4	**Conduct investigations of complex problems**: Use research-based knowledge and research methods including design of experiments, analysis and interpretation of data, and synthesis of the information to provide valid conclusions.
PO 5	**Modern tool usage**: Create, select, and apply appropriate techniques, resources, and modern engineering and IT tools including prediction and modelling to complex engineering activities with an understanding of the limitations
PO 6	**The engineer and society**: Apply reasoning informed by the contextual knowledge to assess societal, health, safety, legal and cultural issues and the consequent responsibilities relevant to the professional engineering practice
PO 7	**Environment and sustainability**: Understand the impact of the professional engineering solutions in societal and environmental contexts, and demonstrate the knowledge of, and need for sustainable development.
PO 8	**Ethic**s: Apply ethical principles and commit to professional ethics and responsibilities and norms of the engineering practice.
PO 9	**Individual and team work**: Function effectively as an individual, and as a member or leader in diverse teams, and in multidisciplinary settings.
PO 10	**Communication**: Communicate effectively on complex engineering activities with the engineering community and with society at large, such as, being able to comprehend and write effective reports and design documentation, make effective presentations, and give and receive clear instructions.
PO 11	**Project management and finance**: Demonstrate knowledge and understanding of the engineering and management principles and apply these to one's own work, as a member and leader in a team, to manage projects and in multidisciplinary environments.
PO 12	**Life-long learning**: Recognize the need for, and have the preparation and ability to engage in independent and life-long learning in the broadest context of technological change.

PROGRAMME SPECIFIC OUTCOMES (PSOs):

PSO 1	Organize, Analyze and interpret the data to extract meaningful conclusions.
PSO 2	Design, Implement and Evaluate a computer-based system to meet desired needs
PSO 3	Develop IT application services with the help of different current engineering tools.

Course Instructor	Course Coordinator	Module Coordinator	HOD
Mr.Michael Sadgun Rao Kona	Mr.Michael Sadgun Rao Kona	Dr.K.Lavanya	Dr B.Srinivasa Rao