GANPAT UNIVERSITY

FACULTY OF COMPUTER APPLICATIONS

Programme	Master of Computer Applications	**Branch/Spec.**	Computer Application
Semester	III	**Version**	1.0.0.0
Effective from Academic Year	2023-24	**Effective for the batch Admitted in**	June 2022
Subject Code	P13A5DS2	**Subject Name**	Data Science-II

Teaching scheme						**Examination scheme (Marks)**			
(Per week)	**Lecture (DT)**		**Practical (Lab.)**		**Total**		**CE**	**SEE**	**Total**
	L	TU	P	TW					
Credit	2	0	2	0	4	Theory	40	60	100
Hours	2	0	4	0	6	Practical	20	30	50

Objective:

- To Learn advance Data Science concepts.

Pre-requisites:

- Students must have basic computer and mathematics knowledge, studied fundamentals of data science in the previous semester, and statistics.

Course Outcomes :

- **1 = Slight (Low); 2 = Moderate (Medium); 3 = Substantial (High); "-" = No Correlation**

Name of CO	**Description**
CO1	Explain data science concepts, data types, processes, preprocessing techniques, applications, and challenges.
CO2	Apply statistical methods and hypothesis testing techniques for data analysis.
CO3	Analyze and apply machine learning algorithms for supervised and unsupervised learning problems.
CO4	Evaluate artificial neural network models and deep learning techniques for complex data problems.
	Mapping of CO and PO

CO \ PO	**PO1**	**PO2**	**PO3**	**PO4**	**PO5**	**PO6**	**PO7**	**PO8**
CO1	3	1	–	2	–	–	–	1
CO2	3	3	–	2	–	–	–	1
CO3	2	3	3	3	–	–	–	2
CO4	2	2	3	3	–	–	–	2

Content:

Unit	**SECTION-I**	**Hrs**

1.	Introduction Introduction to Data Science Technology and techniques, Types of data, Components of Data Science, Data Science Process, Data Pre- processing overview, Data cleaning, Data Transformation, Data Reduction, Applications, and Challenges of Data Science.	9
2.	Statistical Analysis of Data. Sampling methods, Sampling distribution, Central limit theorem, Hypothesis testing: Null and alternative Hypothesis, Type I & Type II error, one and two tailed test, rejection using p-value and critical value approach, Analysis of variance (1-way, two-way), Chi-square test for goodness of fit and independence.	9
	Section -II	
3.	Advanced Machine Learning algorithms: Machine Learning Overview, Supervised, Unsupervised, and Reinforcement learning, Clustering, KNN, SVM (Support Vector Machine), Naive Bayes.	9
4.	Artificial Neural Networks. Overview of artificial deep networks, Bias and Variance, the curse of dimensionality, Vanilla, MLP, Flow graphs and Back propagation, Universal Approximation theorem, Feature Representation, CNN (Convolution neural networks).	9

Practical Content:	
List of programs specified by the subject teacher based on above mentioned topics.	
Text Books:	
1	"Practical Statistics for Data Scientists" by Peter Bruce and Andrew Bruce,Publisher : O'Reilly; 1st edition)6 June2017.(
2	"Introduction to Machine Learning with Python: A Guide for data" by Andreas C. Mueller and Sarah Guido Publisher : O'Reilly; 1st edition)7 October2016(.
MOOC/Certification Courses:	
1	https://www.coursera.org/specializations/machine-learning-introduction
2	https://www.coursera.org/specializations/applied-data-science
3	https://nptel.ac.in/courses/106106179

Question Paper Scheme:

University Examination Duration: 3 Hours

Note for Examiner: -

(I) Questions 1 and 4 are compulsory with no options.

(II) Internal options should be given in questions 2, 3, 5 and 6.

SECTION – I

Q.1 –8 Marks

Q.2 –11 Marks

Q.3 –11 Marks

SECTION - II

Q.4 –8 Marks

Q.5 –11 Marks

Q.6 –11 Marks