RASHTRASANT TUKADOJI MAHARAJ NAGPUR UNIVERSITY, NAGPUR

FOUR YEAR BACHELOR OF TECHNOLOGY (B. Tech..) DEGREE COURSE

SEMESTER: VI (C.B.C.S.)

BRANCH: COMPUTER SCIENCE AND ENGINEERING

Examination Scheme and Syllabus

Sixth Semester:-

S. N.	Subject	Teaching Scheme			Evaluation Scheme			Credits	Category
		L	T	P	CA	UE	Total		
1	Compiler Design	4	-	-	30	70	100	4	PCC-CS
2	Compiler Design -Lab	-	-	2	25	25	50	1	PCC-CS
3	Elective-II	3	-	-	30	70	100	3	PEC-CS
4	Elective-III	3	-	-	30	70	100	3	PEC-CS
5	Open Elective-I	3	-	-	30	70	100	3	OEC
6	Professional Skills Lab II	-	-	2	25	25	50	1	PCC-CS
7	Hardware Lab	-	-	2	25	25	50	1	ESC
8	Mini Project	-	-	6	50	50	100	3	PROJ-CS
9	Economics of IT Industry	2	-	-	15	35	50	2	HSMC
10	Intellectual Property Rights (AuditCourse)	2	-	-	50	-	-	Audit	PCC
	Total	**17**	**-**	**12**			**700**	**21**	

Elective-II: - 1. Machine Learning 2. Internet of Things 3. Cluster and Cloud Computing

Elective-III: - 1. Data Science 2. Distributed Operating Systems 3. Human Computer Interaction

Open Elective 1:- 1. Linux Fundamentals 2. Android Application Development 3. Blockchain Technologies

RASHTRASANT TUKADOJI MAHARAJ NAGPUR UNIVERSITY, NAGPUR

FOUR YEAR BACHELOR OF TECHNOLOGY (B. Tech..) DEGREE COURSE

SEMESTER: VI (C.B.C.S.)

BRANCH: COMPUTER SCIENCE & ENGINEERING

Subject: Elective 3: **Data Science** Subject Code : **BTECH_CSE-603.1T**

Load	Credits	College Assessment Marks	University Evaluation	Total Marks
36 Hrs.	**03**	**30**	**70**	**100**

Aim: To apply data science concepts and methods to solve problems in real-world contexts and to communicate these solutions effectively.

Prerequisite(s): Preliminary Linear Algebra

Course Objectives:

1	To understand the basic concepts of Data science.
2	Demonstrate an understanding of statistics and classification concepts that are vital for data science.
3	Demonstrate the implementation of Data Science experiments through Python or R Language.

Course Outcomes:

At the end of this course Student will be able to:

1	Understanding the significance of exploratory data analysis in Data Science.
2	Demonstrate the usage of Random Sampling and bias in a given dataset.
3	Analysis of various Statistical Experiments through various types popular Testing methods.
4	Design and analysis of regression techniques to estimate outcomes and detect anomalies.
5	Ability to implement classification Techniques.

SYLLABUS:

UNIT I

Exploratory Data Analysis

Elements of Structured Data, Rectangular Data, Estimates of Location, Estimates of Variability, Exploring the Data Distribution, Exploring Binary and Categorical Data, Correlation, Exploring Two or More Variables

UNIT 2

Data and Sampling Distributions

Random Sampling and Sample Bias, Selection Bias, Sampling Distribution of a Statistic, The Bootstrap, Confidence Intervals, Normal Distribution, Long-Tailed Distribution, Student's t-Distribution. Binomial Distribution, Chi-Square Distribution, F-Distribution

UNIT 3

Statistical Experiments and Significance Testing

A/B Testing, Hypothesis Tests, Resampling, Statistical Significance and p-Values, Multiple Testing, Degrees of Freedom, ANOVA, Chi-Square Test, Multi-Arm Bandit Algorithm. Power and Sample Size

UNIT 4:

Regression and Prediction

Simple Linear Regression, Multiple Linear Regression, Prediction Using Regression, Factor Variables in Regression, Interpreting the Regression Equation, Regression Diagnostics, Polynomial and Spline Regression

UNIT 5:

Classification

Naive Bayes, Discriminant Analysis, Logistic Regression, Evaluating Classification Models, Strategies for Imbalanced Data

Text books:

1. Peter Bruce, Andrew Bruce and Peter Gedeck, Practical Statistics for Data Scientists, 2nd Edition, Oreilly.
2. R Programming for Data Science – Roger D.Peng, Learn Pub Book, Learn Publishing.
3. Sanjivranjan Das, Data Science: Theories, Models, Algorithms and Analytics.
4. Cathy O'Neil and Rachel Schutt, Doing Data Science, Straight Talk.

Reference books:

1. Allen B. Downey, Think Python: How to Think Like a Computer Scientist, (2nd Edition), O'Reilly, 2015. ISBN-978-1-491-93936-9.
2. R for dummies – Andrie de vries and Joris Meys, A John Wiley sons, Ltd. Publication.