



GUJARAT TECHNOLOGICAL UNIVERSITY

BACHELOR OF ENGINEERING SYLLABUS
Minor Degree: DATA SCIENCE (W.E.F 2022-23)

Semester: 4

Subject Code : 114AT01

Subject Name : Introduction to Data Science

Prerequisite:

Rationale:

Course Objective:

1. To provide the knowledge and expertise to become a proficient data scientist;
2. Demonstrate an understanding of statistics and machine learning concepts that are vital for data science;
3. Produce Python code to statistically analyse a dataset;
4. Critically evaluate data visualisations based on their design and use for communicating stories from data;

Course Scheme:

Teaching Scheme			Total Credits	Assessment Pattern and Marks				Total Marks
L	T	PR	C	Theory		Practical		
				ESE (E)	PA(M)	PA (I)	ESE (V)	
3	0	2	4	70	0	0	30	100

Course Content:

Sr. No.	Course Content	No. of Hours	% of Weightage
1	Introduction to Data Science, Different Sectors using Data science, Purpose and Components of Python in Data Science.	07	
2	Data Analytics Process, Knowledge Check, Exploratory Data Analysis (EDA), EDA- Quantitative technique, EDA- Graphical Technique, Data Analytics Conclusion and Predictions.	07	
3	Feature Generation and Feature Selection (Extracting Meaning from Data)- Motivating application: user (customer) retention- Feature Generation (brainstorming, role of domain expertise, and place for imagination)- Feature Selection algorithms.	11	
4	Data Visualization- Basic principles, ideas and tools for data visualization, Examples of inspiring (industry) projects- Exercise: create your own visualization of a complex dataset.	10	
5	Applications of Data Science, Data Science and Ethical Issues- Discussions on privacy, security, ethics- A look back at Data Science- Next-generation data scientists.	07	

Reference Book:

1. Business Analytics: The Science of Data - Driven Decision Making, U Dinesh Kumar, John Wiley & Sons.



GUJARAT TECHNOLOGICAL UNIVERSITY

BACHELOR OF ENGINEERING SYLLABUS

Minor Degree: DATA SCIENCE (W.E.F 2022-23)

Semester: 4

Subject Code : 114AT01

Subject Name : Introduction to Data Science

2. Introducing Data Science: Big Data, Machine Learning, and More, Using Python Tools, Davy Cielen, John Wiley & Sons.
3. Joel Grus, Data Science from Scratch, Shroff Publisher/O'Reilly Publisher Media
4. Annalyn Ng, Kenneth Soo, Numsense! Data Science for the Layman, Shroff Publisher Publisher
5. Cathy O'Neil and Rachel Schutt. Doing Data Science, Straight Talk from The Frontline. O'Reilly Publisher.
6. Jure Leskovek, Anand Rajaraman and Jeffrey Ullman. Mining of Massive Datasets. v2.1, Cambridge University Press.
7. Jake VanderPlas, Python Data Science Handbook, Shroff Publisher/O'Reilly Publisher Media.
8. Philipp Janert, Data Analysis with Open Source Tools, Shroff Publisher/O'Reilly Publisher Media.

Course Outcome:

After Completion of the Course, Student will able to:

No	Course Outcomes	RBT Level*
01	To explain how data is collected, managed and stored for data science.	UN
02	To understand the key concepts in data science, including their real-world applications and the toolkit used by data scientists.	UN
03	To implement data collection and management scripts using MongoDB.	AP

*RM: Remember, UN: Understand, AP: Apply, AN: Analyze, EL: Evaluate, CR: Create

Suggested Course Practical List:

List of Laboratory/Learning Resources Required:

1. Python Environment setup and Essentials.
2. Mathematical computing with Python (NumPy).
3. Scientific Computing with Python (SciPy).
4. Data Manipulation with Pandas.
5. Prediction using Scikit-Learn
6. Data Visualization in python using matplotlib

* * * * *