



Gujarat Technological University

Graduate School of Engineering and Technology

Course Title:	Implementation of Data Science Life Cycle on Real-time Dataset
Course Offered Under:	Internship for UG students entering in final year
Duration (Contact hours):	02 weeks (05 Days/week, Max. 30 hours)
Mode:	Face to Face (10:30 am to 01:30 pm)
Fees:	The fee per student for the two-week internship as Rs. 2000 + 18% GST = 2360/-.
Max batch size:	20 interns
Course Coordinators:	Prof. S. K. Hadia, Prof. Mahesh Panchal, Ms. Anal Raval

Rationale

The course bridges the gap between theoretical concepts of Data Science and its usage in developing the projects.

Course Content

Module No.	Module Content	No. of Hours	Resource Person
01.	Data Collection, Validation and Labelling Overview, Importance of Data, Data analyst pipeline, Responsible Data, Detection of Data Issue	03	Prof. S K Hadia Prof. Mahesh Panchal
02.	Data Exploration & Preprocessing Introduction to preprocessing, Data Manipulation, Data Preparation and Feature Engineering, Feature Selection & Transformation	03	Ms. Arpita Maheriya Ms. Aanal Raval
03.	Data Exploration & Preprocessing Importing Dataset, Exploring and analyzing datasets	03	Ms. Arpita Maheriya Ms. Aanal Raval
04.	Model Building and Evaluation (Part-1) Supervised, Unsupervised Methods, Matching data points and explaining the concept with a real-life example	03	Prof. Mahesh Panchal
05.	Model Building and Evaluation (Part-2) Semi supervised methods, model monitoring- data drift and model drift analysis	03	Prof. Mahesh Panchal
06.	Model Visualization Using Matplotlib and pandas for visualization, Advanced techniques- seaborn, bokeh	03	Prof. Soniya Jain



Gujarat Technological University

Graduate School of Engineering and Technology

07.	Advanced Statistics for Data Understanding (Part-1) data sampling, probability distribution, descriptive statistics, inferential statistics, Central Limit Theorem	03	Prof. Soniya Jain Prof Komal Prajapati
08.	Advanced Statistics for Data Understanding (Part-2) Confidence Intervals, Hypothesis Testing, Z-test, T-test, Chi-Square Test, F-Test and ANOVA	03	Prof. Soniya Jain Prof Komal Prajapati
09, 10	Development of Mini Project	06	Prof. S. K. Hadia Prof. Mahesh Panchal Prof. Soniya Jain Ms. Anal Raval Ms. Arpita Maheriya

Learning Outcomes:

After completion of the Internship, the learners will be able to:

No	Course Outcome	RBT Level*
01	Understand the working mechanisms of data pre-processing, visualization, statistical and machine learning methods.	UN
02	Apply data pre-processing and data visualization methods on real world data.	AP
03	Apply machine learning methods to develop predictive model from clean data.	AP
04	Analyze the working methodology of machine learning methods.	AN
05	Evaluate the performance of statistical and machine learning methods using suitable performance metrics.	EV

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