

Branch Name:	Bachelor of Technology (INFORMATION TECHNOLOGY)
Program Code:	ET203
Course Title:	Mathematical Foundation of Data Science (Theory)
Course Code:	2ET1030505T
Pre-requisite Course:	Basic knowledge of mathematics

Course Objectives:

1. Ability to use mathematical concepts in the field of data science.
2. Employ the techniques and methods related to the area of data science in a variety of applications.
3. Apply logical thinking to understand and solve the problem in context.

Teaching and Examination Scheme:

Credit	Teaching Scheme Hours per week				Evaluation Scheme (Marks)								
					Theory (Marks)				Practical (Marks)				Total (Marks)
	L	T	P	Total	SEE		CCA		SEE		CCA		
					Min.	Max.	Min.	Max.	Min.	Max.	Min.	Max.	
3	3	-	-	3	24	60	16	40	-	-	-	-	100

Course Contents:

Unit No	Topics	Total Hours	Weightage (%)
1	Introduction: Basics of Data Science: Introduction; Typology of problems; Importance of linear algebra, statistics and optimization from a data science perspective; Structured thinking for solving data science problems.	8	20
2	Linear Algebra: Matrices and their properties (determinants, traces, rank, nullity, etc.); Eigenvalues and eigenvectors; Matrix factorizations; Inner products; Distance measures; Projections; Notion of hyperplanes; half-planes.	8	20
3	Probability, Statistics and Random Processes: Probability theory and axioms; Random variables; Probability distributions and density functions (univariate and multivariate); Expectations and moments; Covariance and correlation; Statistics and sampling distributions; Hypothesis testing of means, proportions, variances and correlations; Confidence (statistical) intervals; Correlation functions; White-noise process.	10	25
4	Optimization: Unconstrained optimization; Necessary and sufficiency conditions for optima; Gradient descent methods; Introduction to non-gradient techniques; Introduction to least squares optimization; Optimization view of machine learning. Introduction to Data Science Methods: Linear regression as an exemplar function approximation problem; Linear classification problems.	10	25
5	Basic of python Required for data science : Numpy, pandas, matplotlib.lib libraries	8	10

Text Books:

1. G. Strang . Introduction to Linear Algebra, Wellesley-Cambridge Press, Fifth edition, USA, 2016.
2. Bendat, J. S. and A. G. Piersol. Random Data: Analysis and Measurement Procedures. 4th Edition. John Wiley & Sons, Inc., NY, USA, 2010
3. Montgomery, D. C. and G. C. Runger. Applied Statistics and Probability for Engineers. 5th Edition. John Wiley & Sons, Inc., NY, USA, 2011.
4. David G. Luenberger . Optimization by Vector Space Methods, John Wiley & Sons (NY), 1969.
5. Cathy O’Neil and Rachel Schutt . Doing Data Science, O’Reilly Media, 2013.

List of Open Source Software/learning website:

<http://www.simplelearn.com>

<http://www.geeksforgeeks.com>

Course Learning Outcomes (CLO):

Course Outcomes	Description	Bloom’s Taxonomy Level
CLO1	Demonstrate understanding of basic mathematical concepts in data science, relating to linear algebra, probability, and calculus.	2 Understanding
CLO2	Employ methods related to these concepts in a variety of data science applications.	2 Understanding
CLO3	Apply logical thinking to problem-solving in context.	2 Understanding, 3 Applying
CLO4	Demonstrate skills in writing mathematics.	3 Applying, 2 Understanding

Mapping of CLO with PO and PSO:

Course Learning Outcomes	Program Outcomes (POs)												Program Specific Outcomes(PSOs)			
	PO1	PO2	PO3	PO4	PO5	PO6	PO7	PO8	PO9	PO10	PO11	PO12	PSO1	PSO2	PSO3	PSO4
CLO1	M	M							L				M			
CLO2	M				L				M					L		
CLO3	L		L			L								M		
CLO4						L									M	

H: High, M: Medium, L: Low

Reference University:

1. University of Adelaide
2. IIT Kanpur
3. Gujarat Technological University