

Course Code: ELE-622

Course Title: Data Science and Machine Learning

Number of Credits: 04

Total Hours: 60

Total Marks: 100

Effective from AY: 2022-23

Prerequisites for the course		
Should have the basic knowledge of linear algebra and Python programming language.		
Objectives of Course		
This course is intended to:		
• Introduce the mathematical foundations required for Data Science and Machine learning. • Introduce the statistical and probabilistic concepts. • Learn the data analysis and processing. • To study different machine learning and deep learning techniques.		
Course Content		
Unit I	Introduction	05 Hours
Introduction to Data Science, Data Science Process, Fundamentals of linear algebra		
Unit II	Foundations of Data Science	08 Hours
Python for Data Science- Programming basics, libraries – Numpy, SciPy, Scikit-Learn, Tensor flow, Keras, Py-torch, Pandas, Matplotlib, Seaborn. Statistical concepts – Descriptive and Inferential Statistics, Probability		
Unit III	Data Visualization, Exploration and Manipulation	12 Hours
Types of Data, Visualization, Exploratory Data Analysis, Data Cleansing, Data Manipulation, Feature Extraction and Feature Selection		
Unit IV	Regression and Classification	10 Hours
Linear Regression, Logistic Regression, Multivariate Regression, Support Vector Machine, Decision Tree, Random Forest, Naive Bayes, Regularization, Goodness of fit		
Unit V	Machine Learning	15 Hours
Machine Learning Process Flow, Overfitting and under fitting, Bias-Variance, Types of machine learning, Dimensionality Reduction-Principal Component Analysis, K-Nearest Neighbour, Clustering- K-means		
Unit VI	Deep Learning	10 Hours
Neural Networks, Convolutional Neural Networks, Transfer Learning, Reinforcement learning, Recurrent Neural Networks		
case studies		
1. Implement data cleansing and manipulation operations 2. Implement Support Vector Machine algorithm for multiclass classification 3. Implement clustering algorithm 4. Demonstrate the overfitting and under fitting conditions 5. Face recognition using deep convolutional neural network and using transfer learning.		
Pedagogy		
lectures/ Experiential Learning		
Course Outcome		
The students will:		
1. Understand the fundamental concepts of data science and machine learning. 2. Perform data processing technique using python.		

3. Explain and implement the machine learning methods.

References/Readings

1. Trevor Hastie, Robert Tibshirani, Jerome Friedman, 'The Elements of Statistical Learning - Data Mining, Inference, and Prediction', Springer New York, 2013
2. Joel Grus, 'Data Science from Scratch - First Principles with Python', 2nd Edition, O'Reilly Media, 2019.
3. Cathy O'Neil and Rachel Schutt, 'Doing Data Science - Straight Talk From The Frontline', O'Reilly. 2013.
4. Laura Igual, Santi Seguí, Eloi Puertas, Petia Radeva, Oriol Pujol, Sergio Escalera, Francesc Dantí, Lluís Garrido, 'Introduction to Data Science - A Python Approach to Concepts, Techniques and Applications', Springer International Publishing, 2017
5. Samir Madhavan, 'Mastering Python for Data Science', Packt Publishing, 2015
6. Ian Goodfellow, Yoshua Bengio, Aaron Courville, 'Deep Learning', MIT Press, 2016
7. Zhi-Hua Zhou, 'Machine Learning', Springer Nature Singapore, 2021