

Digital Transformation – AI & ML

Course curriculum

M1. Regression

Regression is a widely used statistical learning method, and this course will enable participants to have a deeper understanding of regression models both from theoretical and implementation perspective.

The course covers concepts such as lasso regression, ridge regression and the interpretability of the predicted models.

- Overview of certificate programme in ML & AI
- Introduction to Regression
- Mathematics Foundations
- Model Building using Least squares
- Model Accuracy & Selection
- Overfitting
- Interpretability of regression models

M2. FEATURE ENGINEERING (4 WEEKS)

Feature Engineering is an important step to develop and improve performance of Machine Learning models. In this course, students will learn different data wrangling techniques that help transforming the raw data to an appropriate form on which learning algorithms can be applied.

This course enables students to identify and implement appropriate feature extraction and pre-processing techniques. The Visualization techniques will also be taught in this course.

- Overview of Feature Engineering
- Data Pre-processing
- Dimensionality Reduction
- Visualization

M3. CLASSIFICATION (8 WEEKS)

The course on Classification lays down a strong foundation on the algorithmic perspective of popular classification algorithms – k-NN, Naïve Bayes, Decision Tree,

Logistic Regression and SVM. The implementation details of these models, along with tuning of parameters will be illustrated. The course also covers concepts such as ensemble methods like bagging, boosting, Random Forest, and interpretability of the predicted models.

- Overview of the Classification Module
- Nearest-neighbour Methods
- Naïve Bayes Classifier
- Logistic Regression
- Decision Tree
- Optimization Foundations for Support Vector Machines
- Support Vector Machines
- Support Vector Machines in overlapping class distributions & Kernels
- Ensemble Methods

M4. UNSUPERVISED LEARNING & ASSOCIATION RULE MINING (7 WEEKS)

The course on Unsupervised Learning & Association Rule Mining focuses on finding natural groups or clusters that are present in the data. The course will cover clustering algorithms like K-means, Hierarchical & DBSCAN algorithms, Hidden Markov Models for time series prediction, and market basket analysis to generate the interesting rules from a transactional database.

- Introduction to Unsupervised Learning, Clustering
- K-Means Algorithm, K-Means – Variations, Detecting Outliers
- Math Fundamentals for EM Algorithm, EM Algorithm, Clustering for Customer Segmentation
- Hierarchical Clustering

- Density Based Clustering, Clustering for Anomaly Detection
- Assessing Quality of Clustering, Significance of Clustering – Interpreting/ summarizing Clusters by businesses
- Association Rule Mining, A priority Algorithm
- Time series Prediction and Markov Process, Hidden Markov Model

M5. Text Mining (5 weeks)

Text mining is the process of deriving high-quality information from text and this is the fifth course of the program. This course aims to equip students with adequate knowledge in extracting the relevant text data and skills to identify patterns therein. This course covers topics like converting documents to vectors,

parts of speech tagging, topic modelling, and sentiment analysis and recommender systems.

- Document vectorization and Parts of Speech Tagging
- Introduction to Part of speech tagging, Part of speech tagging using HMM-1, Implementing POS Tagging in Python
- Topic modelling using LDA
- Introduction to Sentiment Analysis
- Recommender Systems

M6. Deep Learning and ANN (6 weeks)

Deep learning is an evolving subfield of Machine Learning, and this course starts with traditional Neural Networks followed by sequential networks, Convolution Networks, Auto encoders and Generative deep learning models. The implementation details of these deep learning models along with tuning of the parameters will be illustrated in this course.

- Artificial Neural Network
- Sequence Modelling in Neural Network
- Deep Learning
- Convolution Networks with Deep Learning
- Auto encoders with Deep Learning
- Generative deep learning models

CAPSTONE PROJECT (8 WEEKS)

During the 8-week Capstone Project, participants will work in teams to design and solve a real-world business problem encompassing data science pipeline using AI&ML techniques under the mentorship of faculty members and senior Industry practitioners.