Course Name	**Machine Learning Specialty**
Course Duration	5 Day (40 hours)
Time Division	**Break:** 15 + 45 + 15 mins
Course Outcomes	Understand the fundamentals of Data Science and Machine Learning.
	Analyze and preprocess data proficiently using Python.
	Apply Supervised Machine Learning techniques for regression and classification.
	Apply Unsupervised Machine Learning for clustering and natural language
	Introduction to Deep learning concepts

Important Note:

- Courseware – Reference material/ppt along with lab files/exercises will be provided

Module	Content
Module 01	**Introduction to Data Science & Machine Learning**
1.1	Need for Data Science and Machine Learning
1.2	Types of Analytics
1.3	Lifecycle of a Data Science project
1.4	Skills for a Data Scientist role
1.5	Types of Machine Learning
Module 02	**Python for Data Analysis & Pre-processing**
	Introduction to Python
2.1	Python Libraries – NumPy, Pandas, matplotlib, Seaborn scikit-learn, TensorFlow, Keras, Pytorch
2.2	Exploratory Data Analysis (EDA)
2.3	Data Cleaning Techniques, Handling Missing Data, Handling Categorical Data
2.4	Introduction to EDA, 2D Scatter-plot, 3D Scatter-plot, Pair plots
2.5	Univariate, Bivariate, and Multivariate Analysis, Box-plot
	Data Pre-Processing
2.6	Need for Data Pre-Processing
2.7	Handling Missing Values
2.8	Label-Encoding for Categorical Data
2.9	Hot-Encoding for Categorical Data Explained
	Data Transformation
2.10	Need for Data Transformation
2.11	Concept of Data Normalization
2.12	Data Normalization Techniques - Standard Scalar & Minmax
2.13	Train, Test & Validation of Data
Module 03	**Supervized Machine Learning – Regression**
3.1	**Simple Linear Regression**
3.2	Concept of Linear Regression
3.3	Ordinary Least Square and Regression Errors
3.4	Data Processing & Train and Test of Model
3.5	Model Evaluation Parameters like R-squared, Score, RMSE and their Interpretations
3.6	Prediction Plot & its Interpretation
3.7	Hands-on Problem
3.8	**Multiple Linear Regression**
3.9	Concept of Multiple Linear Regression
3.10	Degrees of Freedom
3.11	Adjusted R-Squared
3.12	Assumptions of Multiple Linear Regression - Linearity, Multicollinearity, Autocorrelation, Endogeneity, Normality of Residuals, Homoscedasticity, etc.

3.13	Concept of time-lag data in Autocorrelation
3.14	Concept of Dummy variable trap
3.15	Hands-on Problem
Module 04	**Supervised Machine Learning - Classification**
	Logistic Regression
4.1	Concept of Logistic Regression
4.2	Concept of Stratification
4.3	Concept of Confusion Matrix
4.4	Hands-on Problem
	Support Vector Machine (SVM)
4.5	Common Sensical Intuition of SVM
4.6	Mathematical Intuition of SVM
4.7	Different types of SVM Kernel Functions
4.8	Hands-on Problem (Preferred: IRIS Classification Problem)
	Decision Tree Classifier
4.9	Intuition Behind Decision Trees
4.10	Optimal Model Selection Criterion in Decision Tree
4.11	Hands-on Problem
	Random Forest Classifier
4.12	Ensemble Learning and Random Forests
4.13	Bagging and Boosting
4.14	Hands-on Problem
	Evaluation Metrics for Classification Models
4.15	Need for Evaluation and Accuracy Paradox
4.16	Different Measures for Classification Models - Accuracy, Precision, Recall, F1 Score, etc.
4.17	Threshold and Adjusting Thresholds
4.18	AUC ROC Curve
4.19	Hands-on Problem
Module 05	**Feature Selection and Dimensionality Reduction**
	Univariate Feature Selection
5.1	Feature Selection Importance
5.2	Concept of Univariate Feature Selection
5.3	F-Test for Regression and Classification
5.4	Hands on F-test (p value analysis)
5.5	Chi-Squared for Classification
5.5	Feature Selection Techniques - SelectKBest, SelectPerecntile & Generic Univariate Select
5.6	Hands-on Chi-squared (p value analysis)
	Recursive Feature Elimination (RFE)
5.7	Concept of Recursive Feature Elimination (RFE)
5.8	Feature Importance Score/Feature Ranking
5.9	Hands-on RFE
	Principle Component Analysis (PCA)
5.10	Need to reduce dimensions and Importance of PCA
5.11	Mathematical Intuition of PCA & Steps to calculate PCA
5.12	Hands-on PCA (Model Comparisons with PCA & without PCA recommended)
Module 06	**Cross validation & Hyperparameter Tuning**
	Cross Validation
6.1	Importance of Cross Validation
6.2	Parameter & Implementation of Cross Validation

6.3	Hands-on Problem (Drawing inference from results)
	Hyperparameter Tuning
6.4	Concept of Hyperparameter Tuning
6.5	Grid Search & Randomized Search
6.6	Hands-on GridSearchCV (analyse results)
Module 07	**Supervized Machine Learning – Natural Language Processing**
7.1	Introduction to NLP
7.2	Basic Concepts of NLP: Tokenization, stop words, Stemming, Lemmatization, etc.
7.3	Tfidf Vector and its mathematical intuition
7.4	Recommendation system example
Module 08	**Supervized Machine Learning – Clustering**
8.1	Introduction to Clustering
8.2	Mathematical intuition behind cluster formation
8.3	Elbow method & its mathematical intuition
8.4	K-means Clustering Implementation (numerical)
8.5	K-means Clustering Implementation (natural language processing)
8.6	Introduction to Clustering
Module 09	**Introduction to Deep Learning**
9.1	Need & Applications of Deep Learning
9.2	Working of Artificial Neural Network
9.3	Backend (TensorFlow) & Frontend (Keras)
9.4	Concept of Tensor
9.5	Keras Model Building Overview - Construct, Compile & Evaluate
9.6	Activation Function
9.7	Loss Functions
9.8	Optimization Techniques
9.9	Evaluation metrics for Deep Learning