Machine Learning Specialty

Course Duration	5 Day (40 hours)
Target Audience	Machine Learning Enthusiasts, Data Analyst, Data Scientist
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.
	Introduce Unsupervised Machine Learning and Deep Learning concepts.

Module	Content
Module 01: Introduction to Data Science & Machine Learning	Need for Data Science and Machine Learning
	Types of Analytics
	Lifecycle of a Data Science project
	Skills for a Data Scientist role
	Types of Machine Learning
Module 02: Python for Data Analysis & Pre-Processing	**Introduction to Python**
	Python Libraries – NumPy, Pandas, matplotlib, Seaborn scikit-learn, TensorFlow, Keras, Pytorch
	Exploratory Data Analysis (EDA)
	Data Cleaning Techniques, Handling Missing Data, Handling Categorical Data
	Introduction to EDA, 2D Scatter-plot, 3D Scatter-plot, Pair plots
	Univariate, Bivariate, and Multivariate Analysis, Box-plot
	Data Pre-Processing
	Need for Data Pre-Processing
	Handling Missing Values
	Label-Encoding for Categorical Data
	Hot-Encoding for Categorical Data Explained
	Data Transformation
	Need for Data Transformation
	Concept of Data Normalization
	Data Normalization Techniques - Standard Scalar & Minmax
	Train, Test & Validation of Data
	Simple Linear Regression
	Concept of Linear Regression

Module 03: Supervised Machine Learning - Regression	Ordinary Least Square and Regression Errors
	Data Processing & Train and Test of Model
	Model Evaluation Parameters like R-squared, Score, RMSE and their Interpretations
	Prediction Plot & its Interpretation
	Hands-on Problem
	Multiple Linear Regression
	Concept of Multiple Linear Regression
	Degrees of Freedom
	Adjusted R-Squared
	Assumptions of Multiple Linear Regression - Linearity, Multicollinearity, Autocorrelation, Endogeneity, Normality of Residuals, Homoscedasticity, etc.
	Concept of time-lag data in Autocorrelation
	Concept of Dummy variable trap
	Hands-on Problem
Module 04: Supervised Machine Learning - Classification	Logistic Regression
	Concept of Logistic Regression
	Concept of Stratification
	Concept of Confusion Matrix
	Hands-on Problem
	Support Vector Machine (SVM)
	Common Sensical Intuition of SVM
	Mathematical Intuition of SVM
	Different types of SVM Kernel Functions
	Hands-on Problem (Preferred: IRIS Classification Problem)
	Decision Tree Classifier
	Intuition Behind Decision Trees
	Optimal Model Selection Criterion in Decision Tree
	Hands-on Problem
	Random Forest Classifier
	Ensemble Learning and Random Forests
	Bagging and Boosting
	Hands-on Problem
	Evaluation Metrics for Classification Models
	Need for Evaluation and Accuracy Paradox
	Different Measures for Classification Models - Accuracy, Precision, Recall, F1 Score, etc.
	Threshold and Adjusting Thresholds
	AUC ROC Curve
	Hands-on Problem
	Univariate Feature Selection

Module 05: Feature Selection and Dimensionality Reduction	Feature Selection Importance
	Concept of Univariate Feature Selection
	F-Test for Regression and Classification
	Hands on F-test (p value analysis)
	Chi-Squared for Classification
	Feature Selection Techniques - SelectKBest, SelectPerecntile & Generic Univariate Select
	Hands-on Chi-squared (p value analysis)
	Recursive Feature Elimination (RFE)
	Concept of Recursive Feature Elimination (RFE)
	Feature Importance Score/Feature Ranking
	Hands-on RFE
	Principle Component Analysis (PCA)
	Need to reduce dimensions and Importance of PCA
	Mathematical Intuition of PCA & Steps to calculate PCA
	Hands-on PCA (Model Comparisons with PCA & without PCA recommended)
Module 06: Cross validation & Hyperparameter Tuning	**Cross Validation**
	Importance of Cross Validation
	Parameter & Implementation of Cross Validation
	Hands-on Problem (Drawing inference from results)
	Hyperparameter Tuning
	Concept of Hyperparameter Tuning
	Grid Search & Randomized Search
	Hands-on GridSearchCV (analyse results)
Module 07: Unsupervized Machine Learning – Natural Language Processing	Introduction to NLP
	Basic Concepts of NLP: Tokenization, stop words, Stemming, Lemmatization, etc.
	Tfidf Vector and its mathematical intuition
	Recommendation system example
Module 08: Unsupervized Machine Learning – Clustering	Introduction to Clustering
	Mathematical intuition behind cluster formation
	Elbow method & its mathematical intuition
	K-means Clustering Implementation (numerical)
	K-means Clustering Implementation (natural language processing)
Module 09:	Need & Applications of Deep Learning

Deep Learning	
	Working of Artificial Neural Network
	Backend (TensorFlow) & Frontend (Keras)
	Concept of Tensor
	Keras Model Building Overview - Construct, Compile & Evaluate
	Activation Function
	Loss Functions
	Optimization Techniques
	Evaluation metrics for Deep Learning