

Course Design Document

Course Code	
Course Name	Machine learning with Python

Duration (in days)	10	Proficiency Level	Fundamental
Pre-requisites	Python	Target Audience	Campus Hires

Learning Outcome

- Introduction to ML
- Regression Algorithms
- Classification - Logistic Regression
- KNN-Classification Algorithm
- Naive Bayes Algorithm
- Decision tree Algorithm
- Support Vector Machines (SVM)
- Dimensionality Reduction for Machine Learning
- Ensemble Techniques
- Feature Engineering Techniques
- Un-Supervised Algorithms
- Reinforcement Learning

Day-wise Session Plan

Day	Module Name	Topics	Hours
1	Introduction to ML	<ul style="list-style-type: none"> • Introduction to Machine Learning • Machine Learning Modelling Flow • Parametric and Non-parametric Algorithms • Types of Machine Learning • Performance Measures • Bias-Variance Tradeoff 	2
	Regression Algorithms	<ul style="list-style-type: none"> • Introduction of Linear Regression • Types of Linear Regression • Assumptions of Linear Regression- Multicollinearity, Homoscedasticity • Simple regression implementation in python • Multiple regression implementation in python • Linear Regression with OLS Model • Performance Metrics to Evaluate Model • Feature Scaling • Feature Selection • Regularization - Lasso & ridge regression model • Real World Problem: Energy Consumption Forecasting 	
2	Classification - Logistic Regression	<ul style="list-style-type: none"> • Logistic Regression Overview • Comparison to linear regression • Binary logistic regression • Sigmoid function • Maximum Likelihood Estimation • Decision boundary • Logistic regression implementation in python 	8

		<ul style="list-style-type: none"> • Training & Testing • Making predictions • Performance Metrics-Confusion Matrix, Precision and Recall • Medical Diagnosis (Diabetes / Heart Disease prediction) 	
3	KNN-Classification Algorithm	<ul style="list-style-type: none"> • Understanding classification and its uses • When do we use KNN Algorithm? • How does the KNN works? • Euclidean Distance - Finding Nearest Neighbour • Implement KNN algorithm with sklearn • Training & Testing • Evaluation of KNN model • Summarize steps & application of KNN • Customer Churn Prediction (Telecom / SaaS) 	8
4	Naive Bayes Algorithm	<ul style="list-style-type: none"> • What is Naive Bayes algorithm? • How Naive Bayes Algorithms Work? • Types of Naive Bayes algorithm: <ol style="list-style-type: none"> 1. GaussianNB() 2. BernoulliNB() 3. MultinomialNB() • Pros and Cons in Naive Bayes • Steps to build Naive Bayes Model in Python • Improve the power of Naive Bayes Model • Credit Card Fraud Detection (Banking/FinTech) 	8
5	Decision tree Algorithm	<ul style="list-style-type: none"> • Decision tree overview • Working Principle • Entropy based Decision Tree • Gini index based Decision tree • Building Decision tree model in python • Training and visualizing Decision Trees • Measuring the Performance-Accuracy,Classification report • Evaluation, overfitting, and regularization • Hyperparameter tuning in Decision Trees • Optimizing Decision Tree Performance • Medical Diagnosis – Diabetes / Heart Disease Prediction 	8
6	Support Vector Machines (SVM)	<ul style="list-style-type: none"> • Understanding Vectors & Hyperplane • Support Vector • Large Margin Intuition • Kernels and Types of Kernels • Linear SVM-linear • NonLinear SVM-rbf • SVM implementation in python • Cost Function and Gradient Updates • Challenges & Advantages of SVM • Employee Attrition Prediction (HR Analytics) 	8
7	Dimensionality Reduction for Machine Learning	<ul style="list-style-type: none"> • Dimensionality reduction • Curse of dimensionality • Tools and libraries used for dimensionality reduction • Algorithms used for dimensionality reduction • Decomposition algorithms: • 1.Principal Component Analysis(PCA) 	8

		<ul style="list-style-type: none"> • 2.Singular Value Decomposition (SVD) • Discriminant Analysis-Linear Discriminant Analysis(LDA) • Applications 	
8	Ensemble Techniques	<ul style="list-style-type: none"> • What is an ensemble? • Why is it needed? • The Logic behind ensemble techniques. • Boosting Techniques- AdaBoost,GradientBoosting • Bagging Techniques- Bagging,RandomForest,ExtraTrees • Voting Techniques 	4
	Feature Engineering Techniques	<ul style="list-style-type: none"> • Introduction to Feature Engineering • Imputation • Treating Outliers • Log Transform • Encoding Techniques • Feature Scaling • Feature Extraction • Handling Overfitting & Underfitting • Cross Validation • K Fold Cross Validation • Implementation of K-Fold Cross Validation Technique • Leave one out Cross Validation(LOOCV) Concept • Implementation of Leave one out Cross Validation(LOOCV) • Up sampling and Downsampling • Gridsearch and Gridsearchcv • RandomizedSearchcv 	4
9	Un-Supervised Algorithms	<ul style="list-style-type: none"> • Common unsupervised learning approaches • Types of cluster techniques • K-means clustering • Working Principle of K-Means • Implementation of KMeans with python • Hierachal clustering implementation • Performance metrics • Retail Customer Segmentation 	8
10	Reinforcement Learning	<ul style="list-style-type: none"> • Reinforcement Learning Overview • Bandit algorithms –Median Elimination, Policy Gradient • Full RL & MDPs • Dynamic Programming • Function Approximation • Hierarchical RL • Markov Decision Process • Shortest Path Problem • Q- Learning Algorithm • Deep Reinforcement Learning • Reinforcement Learning on Machine Learning 	8