Machine Learning Using Python

**Introduction:**

Python is a popular platform used for research and development of production systems. It is a vast language with a number of modules, packages, and libraries that provide multiple ways of achieving a task. Python and its libraries like NumPy, SciPy, Scikit-Learn, Matplotlib are used in data science and data analysis. They are also extensively used for creating scalable machine learning algorithms. Python implements popular machine learning techniques such as Classification, Regression, Recommendation, and Clustering. Python offers a ready-made framework for performing data mining tasks on large volumes of data effectively in lesser time. It includes several implementations achieved through algorithms such as linear regression, logistic regression, Naïve Bayes, k-means, K nearest neighbor, and Random Forest.

**Module:1**

**1.Python Basics**

a.Introduction

b.Control Statements

c.List

d.String

e.Tuples

f. Dictionary

g.Introduction to classes and Objects

**2.Python for Data Science**

a.Numpy

b.Pandas

c.Matplotlib

d.seaborn

**Module:2**

1.**Introduction to ML**

1. What is ML,AI,DL,DS
2. Difference between Machine Learning,Artificial Intelligence,Deep Learning,Data Science
3. Types Of Machine Learning

2.**Life cycle of ML Project**

**3. EDA(Exploratory data analysis)**

3.**Feature Engineering**

a. Handling missing values

b.Handling outliers

c.Categorical Encoding

d.Normalization and Standardization

4.**Feature Selection**

1.Correlation and its types

2.Univariant Selection

3. Feature Selection with Decision Trees

**Module:3**

1. **Supervised Machine Learning techniques**

a)**Regression**

i) Linear Regression

1.Simple Linear Regression

2. Multi Linear Regression

3. Polynomial Regression

ii) Lasso and Ridge Regression

iii)Performance Metrics(R2 score,Adj R2 Score)

iv)Bias and Variance TradeOff, MultiCollinearity

v)Underfitting, Overfitting, Best fit

b)**Classification** **PART-1**

i)Logistic Regression

ii)Decision Tree

1.ID3 Algorithm

2.CART Algorithm

iii)KNN

iv)Performance Metrics

1.Confusion Matrix

2.Precision

3.Recall

4.F1\_Score

5.ROC and AUC

v) Cross Fold validation

**Module:4**

1. **Classification PART-2**

i)Random Forest(Ensemble Technique)

ii)SVM

1. **UnSupervised Machine Learning techniques**

a)Clustering Technique

i)K-Means

ii)Hierarchical

**Projects:**

**1.Flight fare prediction project**

**2.Credit card fraud detection**

**3. Phising Classifier**

Hardware Requirements:

* i3 or above Processor is required
* 4 GB or above RAM is recommended
* Good Internet Connectivity
* OS-Windows 10 is Preferable