Machine Learning with Python (3-4 Months)

- 1. Data Splitting: Training Data & Test Data
 - What is Training data?
 - What is Testing data?
 - Why do we need to separate data for training and testing?
 - How to randomly separate training and testing data
 - train_test_split function
- 2. Supervised Learning Linear Regression
 - Understanding Regression and all its form
 - Assumptions of Regression
 - Regression formula and calculation
 - Simple Linear Regression
 - Multiple Linear Regression
 - Performance Metrics: Accuracy scores, R2 score, adjusted R2 score
 - MSE and RMSE
 - Advantages and disadvantages
- 3. Supervised Learning Logistic Regression
 - Concept of Logistic Regression
 - Logistic regression formula and calculation
 - Univariate and Multivariate Logistic Regression
 - Dealing with categorical independent variable
 - Advantages and disadvantages
- 4. Feature Selection
 - What is feature selection?
 - Understanding the need of feature selection
 - Univariate selection
 - Feature Importance
 - Recursive Feature Elimination (RFE)
- 5. Data Transformation and Normalization
 - Encoding:
 - Label Encoding

- One-Hot Encoding, Dummy Variables
- Scaling:
 - o Standard Scaler
 - o Min-Max Scaler
- 6. Linear and Logistic regression with Data Transformation
 - Linear Regression and its performance evaluation
 - Logistic Regression and its performance evaluation
 - Model optimization
- 7. Decision Trees
 - Concept of decision trees
 - Importance and Usage of Ginny index and Entropy
 - Visualizing decision trees nodes and splits
 - Working of the decision tree algorithm
 - Evaluating decision tree models:
 - o Accuracy
 - \circ Precision
 - o Recall
 - o Confusion Matrix
 - o ROC Curve
 - Advantages and disadvantages
- 8. K-Nearest Neighbor
 - Concept of KNN algorithm
 - What is K in KNN?
 - What should be the value of K?
 - Elbow Curve Method
 - Silhouette Score
 - KNN model evaluation
 - Advantages and disadvantages
- 9. Gaussian Naive Bayes Classifier
 - Understanding the mathematical intuition behind Naive Bayes
 - Assumptions of Naive Bayes
 - Advantages and disadvantages
 - Case studies and use cases

- 10. Support Vector Machines(SVM)
 - What is Support Vector Machines
 - How does SVM work?
 - Different types of SVM
 - SVM Kernels
 - SVM Use Cases
 - Implementation of SVM?
- 11. Unsupervised Learning: Clustering
 - Introduction to Clustering
 - K-Means Clustering: Linkage, Use of Elbow Curve & Silhoutte Score
 - Hierarchical Clustering Agglomerative & Divisive
 - Distance Matrix & Dendrogram
 - Implementation of Clustering on unlabeled data
- 12. Principal Component Analysis(PCA)
 - Understanding noise in the data
 - Dimensionality Reduction
 - Capturing Variance The concept of Principal Component
 - Assumptions of PCA
 - The working of the PCA algorithm
 - Practical implementation of PCA on a dataset
 - Advantages of Using PCA
- 13. Ensemble Modelling:
 - What are ensemble models?
 - Popular Ensembles
 - Bagging
 - o Random Forest
 - Concept of Random Forest
 - Comparison with decision trees
 - Feature Importance in Random Forest
 - OOB Error in Random Forest
 - Hyperparameter Tuning in Random Forest
 - Boosting Algorithms
 - o AdaBoost
 - Doubts Clearing/ Assessments

14. Model Validation & Tuning

- K-fold Cross validation
- Hyperparameter Tuning
 - GridSearchCV
 - RandomizeSearchCV

Artificial Intelligence with Python (3 Months):

Artificial Intelligence:

- Definition and History of Al
- Key milestones in AI history
- Narrow AI vs. General AI vs. Superintelligence
- AI in various fields: healthcare, finance, robotics, etc.
- Ethical Considerations
- Impact of AI on Society
- Knowledge representation in AI

Deep Learning:

- Introduction to Neural Networks
- Backpropagation and Optimization
- Convolutional Neural Networks (CNNs) for image data
- Recurrent Neural Networks (RNNs) and LSTMs for sequential data
- Modern architectures (e.g., Transformers)

Computer Vision:

- Image Preprocessing and Augmentation
- Convolutional Neural Networks (CNNs)
- Object Detection (e.g., YOLO, SSD)

- Image Segmentation (e.g., U-Net)
- Advanced topics: GANs for image generation

Natural Language Processing (NLP):

- Text Preprocessing and Tokenization
- Word Embeddings (e.g., Word2Vec, GloVe)
- Recurrent Neural Networks (RNNs) and LSTMs
- Transformers and Attention Mechanisms (e.g., BERT, GPT)
- NLP Tasks: Text Classification, Named Entity Recognition, Machine Translation