Training Sessions Outline on Python for Machine Learning and Neural Networks

Session 1: Introduction to Python for Machine Learning

Objectives:

- Set up and use virtual environments.
- Understand basic operations with NumPy.
- Perform data manipulation with Pandas.
- Create visualizations with Matplotlib.

Agenda:

1. **Using Virtual Environments:**
   - Introduction to virtualenv and its benefits.
   - Setting up and using virtualenv in Windows and Ubuntu.
2. **Introduction to Python Libraries:**
   - Basic operations with NumPy (simulated data and provided dataset).
   - Data manipulation with Pandas (simulated data and provided dataset).
   - Data visualization with Matplotlib (simulated data and provided dataset).

Session 2: Introduction to Supervised Learning

Objectives:

- Understand the basics of supervised learning.
- Learn about different datasets (training, validation, testing).
- Apply a simple classifier on simulated data and evaluate using accuracy.

Agenda:

1. **Introduction to Supervised Learning:**
   - What is supervised learning?
   - Importance of different datasets (training, validation, testing).
2. **Implementing a Simple Classifier:**
   - Build and train a simple classifier on simulated data.
   - Evaluate the classifier using accuracy.

Homework:

- Implement a simple classifier on provided datasets.
- Evaluate the classifier’s performance.

Session 3: Data Preprocessing

Objectives:

- Understand the importance of data preprocessing.
- Perform data cleaning, one-hot encoding, normalization, and centering.

Agenda:

1. **Data Cleaning:**
   - Handling missing data.
   - Removing duplicates.
2. **Data Transformation:**
   - One-hot encoding for labels.
   - Normalization and centering.
   - Other relevant preprocessing steps.

Homework:

- Clean and preprocess a provided dataset.

Session 4: Cross-Validation Techniques

Objectives:

- Understand the importance of cross-validation.
- Learn different cross-validation techniques.
- Apply cross-validation to evaluate model performance.

Agenda:

1. **Introduction to Cross-Validation:**
   - Definition and purpose of cross-validation.
   - Types of cross-validation (k-fold, stratified k-fold, etc.).
2. **Implementing Cross-Validation:**
   - Applying cross-validation to different models.
   - Evaluating model performance using cross-validation.

Homework:

- Implement cross-validation on provided datasets and evaluate model performance.

Session 5: Preventing Data Leakage

Objectives:

- Understand data leakage and its impact.
- Learn methods to prevent data leakage.

Agenda:

1. **Introduction to Data Leakage:**
   - Definition and examples.
   - Impact on model performance.
2. **Preventing Data Leakage:**
   - Proper train-test split.

Session 6: Feature Selection Methods

Objectives:

- Understand the importance of feature selection.
- Implement feature selection methods.

Agenda:

1. **Introduction to Feature Selection:**
   - Importance and benefits.
2. **Greedy Forward LDA-Based Feature Selection:**
   - Explanation and implementation.

Homework:

- Implement greedy forward LDA-based feature selection on a provided dataset.

Session 7: Support Vector Machines (SVM)

Objectives:

- Understand the theory behind SVM and RBF SVM.
- Implement and evaluate SVM models.

Agenda:

1. **Introduction to SVM:**
   - Theory and concepts of SVM.
   - Importance and applications of SVM.
2. **Implementing SVM Models:**
   - Build and train SVM and RBF SVM on simulated data.
   - Evaluate the models.

Homework:

- Implement and evaluate SVM models on provided datasets.

Session 8: Decision Trees and Random Forests

Objectives:

- Understand the theory behind decision trees and random forests.
- Implement and evaluate tree-based models.

Agenda:

1. **Introduction to Decision Trees:**
   - Theory and concepts of decision trees.
   - Importance and applications of decision trees.
2. **Implementing Decision Trees and Random Forests:**
   - Build and train decision trees and random forests on simulated data.
   - Evaluate the models.

Homework:

- Implement and evaluate decision trees and random forests on provided datasets.

Session 9: Introduction to Unsupervised Learning

Objectives:

- Understand the basics of unsupervised learning.
- Implement simple unsupervised learning models.

Agenda:

1. **Introduction to Unsupervised Learning:**
   - What is unsupervised learning?
   - Types of unsupervised learning algorithms.
2. **Implementing Models:**
   - K-means clustering.
   - Principal Component Analysis (PCA).

Homework:

- Implement and evaluate simple unsupervised learning models on provided datasets.

Session 10: Neural Networks - MLP with Simulated Data

Objectives:

- Understand the basic structure and functioning of neural networks.
- Implement a simple MLP model on simulated data.

Agenda:

1. **Basics of Neural Networks:**
   - Structure and activation functions.
   - Forward and backward propagation.
2. **Implementing MLP:**
   - Using TensorFlow/Keras to build a simple MLP.
   - Training and evaluating the MLP on simulated data.

Homework:

- Build and train an MLP on provided datasets.

Session 11: Neural Networks - MLP with Real Dataset

Objectives:

- Apply an MLP model to a real dataset.
- Evaluate the performance of the MLP model.

Agenda:

1. **Implementing MLP on Real Data:**
    - Build and train an MLP on the provided dataset.
    - Evaluate the model's performance.

Homework:

- Fine-tune the MLP model and improve its performance on the provided dataset.

Session 12: Regularization Techniques in Neural Networks

Objectives:

- Understand the importance of regularization in neural networks.
- Implement dropout and batch normalization.

Agenda:

1. **Introduction to Regularization:**
    - Overfitting and the need for regularization.
    - Techniques: Dropout and Batch Normalization.
2. **Implementing Regularization:**
    - Adding dropout to neural networks.
    - Applying batch normalization.

Homework:

- Apply dropout and batch normalization to neural networks and evaluate the performance.

Session 13: Convolutional Neural Networks (CNN) with Simulated Data

Objectives:

- Understand the basics of CNNs.
- Implement a simple CNN model on simulated data.

Agenda:

1. **Introduction to CNNs:**
    - Structure and functioning of CNNs.
    - Convolutional layers, pooling layers, and fully connected layers.
2. **Implementing CNN:**
    - Using TensorFlow/Keras to build a simple CNN.
    - Training and evaluating the CNN on simulated data.

Homework:

- Build and train a CNN on provided datasets.

Session 14: Convolutional Neural Networks (CNN) with Real Dataset

Objectives:

- Apply a CNN model to a real dataset.
- Evaluate the performance of the CNN model.

Agenda:

1. **Implementing CNN on Real Data:**
    - Build and train a CNN on the provided dataset.
    - Evaluate the model’s performance.

Homework:

- Fine-tune the CNN model and improve its performance on the provided dataset.