**Introduction to Machine Learning**

**Morning Session:**

**Overview of Supervised and Unsupervised Learning Algorithms:**

In the morning session I began with an in-depth overview of the two primary types of machine learning algorithms: supervised and unsupervised learning.

**Supervised Learning:**

* In supervised learning, the algorithm is trained on a labeled dataset, which means that each training example is paired with an output label.
* Common supervised learning algorithms include:
* Linear Regression: Used for predicting continuous values.
* Logistic Regression: Used for binary classification problems.
* Decision Trees: Used for both classification and regression tasks.
* Support Vector Machines (SVM): Used for classification and regression problems.
* K-Nearest Neighbors (KNN): A simple algorithm that stores all available cases and classifies new cases based on a similarity measure.

**Unsupervised Learning:**

* In unsupervised learning, the algorithm is given data without explicit instructions on what to do with it. The goal is to infer the natural structure present within a set of data points.
* Common unsupervised learning algorithms include:
* K-Means Clustering: Used to partition data into K distinct clusters based on distance.
* Hierarchical Clustering: Creates a tree of clusters to understand the data hierarchy.
* Principal Component Analysis (PCA): Used for dimensionality reduction while preserving as much variability as possible.
* Association Rules: Used to find relationships between variables in large databases.

**Hands-on with Scikit-learn: Building and Evaluating Simple Machine Learning Models:**

To deepen my understanding of machine learning, I dedicated significant time to studying online resources and tutorials on Scikit-learn, a powerful and user-friendly Python library for machine learning. Scikit-learn provides simple and efficient tools for data analysis and modeling, making it an excellent choice for both beginners and experienced practitioners.

**Scikit-learn Overview:**

Scikit-learn, built on top of NumPy, SciPy, and Matplotlib, is designed to interoperate with these libraries, providing a wide array of machine learning algorithms for classification, regression, clustering, and dimensionality reduction. The library emphasizes easy-to-use APIs and efficient, reusable code, enabling rapid prototyping and experimentation with various machine learning models.

**Building and Evaluating Machine Learning Models:**

The online study focused on building and evaluating simple machine learning models using Scikit-learn, covering several essential steps:

* **Data Preprocessing:**
  + Importing and cleaning data
  + Handling missing values
  + Feature scaling and normalization
  + Encoding categorical variables
* **Model Selection:**
  + Choosing appropriate algorithms for different tasks (e.g., classification, regression, clustering)
  + Understanding the parameters and hyperparameters of various models
* Model Training:
  + Splitting data into training and testing sets to avoid overfitting
  + Training models using the training data
* Model Evaluation:
  + Making predictions on the test data
  + Evaluating model performance using metrics such as accuracy, mean squared error, and R-squared