**Week 1 - Machine Learning Basics**

**1. What is Machine Learning?**

**Machine Learning (ML)** is a subset of Artificial Intelligence (AI) that enables computer systems to learn from data and improve their performance on tasks without being explicitly programmed. Instead of following hard-coded instructions, ML algorithms identify patterns within data and make predictions or decisions based on that information.

**Key Characteristics:**

* **Data-Driven Learning**: ML models improve their performance as they are exposed to more data.
* **Adaptability**: These models can adjust to new inputs and make accurate predictions.
* **Automation**: ML automates analytical model building, allowing systems to make decisions with minimal human intervention.

**Real-World Applications:**

* **Healthcare**: Predicting disease outbreaks and patient diagnoses.
* **Finance**: Detecting fraudulent transactions and credit scoring.
* **Retail**: Personalizing shopping experiences and inventory management.
* **Transportation**: Optimizing routes and enabling autonomous vehicles.

**2. What is Supervised Machine Learning?**

**Supervised Machine Learning** is a type of ML where the model is trained on a labeled dataset. This means that each training example is paired with an output label. The model learns to map inputs to the correct output, enabling it to make predictions on new, unseen data.

**How It Works:**

* **Training Phase**: The model is provided with input-output pairs and learns the relationship between them.
* **Prediction Phase**: Once trained, the model can predict outputs for new inputs based on the learned relationships.

**Examples:**

* **Spam Detection**: Classifying emails as 'spam' or 'not spam'.
* **Image Recognition**: Identifying objects within images.
* **Credit Scoring**: Predicting the likelihood of a borrower defaulting on a loan.

**3. What is Regression in Machine Learning?**

**Regression** is a supervised learning technique used when the output variable is a continuous value. The goal is to model the relationship between input variables and the continuous output, allowing for predictions of numerical values.

**Key Points:**

* **Continuous Output**: Unlike classification, regression predicts quantities.
* **Modeling Relationships**: It identifies how changes in input variables affect the output.

**Common Algorithms:**

* **Linear Regression**: Assumes a linear relationship between inputs and output.
* **Polynomial Regression**: Models nonlinear relationships.
* **Support Vector Regression (SVR)**: Uses support vector machines for regression tasks.

**Applications:**

* **Real Estate**: Predicting house prices based on features like size and location.
* **Economics**: Forecasting economic indicators like GDP growth.
* **Environmental Science**: Estimating pollution levels based on various factors.

**4. What is Classification in Machine Learning?**

**Classification** is another form of supervised learning where the output variable is categorical. The model learns to assign inputs into predefined categories or classes.

**Key Points:**

* **Categorical Output**: Outputs are discrete labels, such as 'yes' or 'no'.
* **Decision Boundaries**: The model learns boundaries that separate different classes.

**Common Algorithms:**

* **Logistic Regression**: Used for binary classification problems.
* **Decision Trees**: Splits data into branches to make predictions.
* **k-Nearest Neighbors (k-NN)**: Classifies based on the majority class among the nearest neighbors.

**Applications:**

* **Medical Diagnosis**: Classifying whether a tumor is benign or malignant.
* **Customer Segmentation**: Grouping customers based on purchasing behavior.
* **Language Processing**: Identifying the sentiment of a text as positive, negative, or neutral.