Types of ML Techniques

Types of Machine Learning Techniques

Summary

- **Supervised Machine Learning**: A technique where the model is trained on a dataset containing both **independent features** (input) and a **dependent feature** (output) to predict future outcomes.
- **Regression**: A sub-type of supervised learning used when the dependent variable is **continuous** (e.g., house prices).
- **Classification**: A sub-type of supervised learning used when the dependent variable is **categorical** (e.g., Pass/Fail), further divided into **binary** and **multi-class** classification.
- **Unsupervised Machine Learning**: A technique used when there is **no output feature**; the goal is to identify patterns or **clusters** within the data (e.g., customer segmentation).
- **Reinforcement Learning**: A technique where an application learns autonomously by performing actions and receiving **rewards**, similar to how a child learns from experience.

Supervised Machine Learning

In supervised machine learning, the goal is to solve a problem statement using a dataset that includes specific features. These features are categorized into two types:

- **Independent Features**: The input data (e.g., size of a house, number of rooms).
- **Dependent Feature** (or Output Feature): The target variable that changes based on the independent features (e.g., price of the house).

The model is trained on this data to predict the dependent feature for new inputs without human intervention.

Types of Problem Statements

Supervised learning generally addresses two types of problems:

1. Regression

A problem is considered **regression** if the dependent (output) feature consists of **continuous** values.

Example: House Price Prediction In this scenario, the price (dependent feature) varies continuously based on the size and number of rooms (independent features).

Size (sq ft)	No. of Rooms	Price ($)
5000	5	450,000
6000	6	500,000

2. Classification

A problem is considered **classification** if the dependent feature is **categorical**.

- **Binary Classification**: The output has exactly two categories (e.g., Pass or Fail).
- **Multi-class Classification**: The output has more than two categories.

Example: Student Performance Predicting if a student passes or fails based on study and play hours.

Study Hours	Play Hours	Result (Output)
7	3	Pass
2	6	Fail

Algorithms

Common algorithms used in supervised learning include:

- **For Regression**:
  - Linear Regression
  - Ridge and Lasso Regression
  - Elastic Net
- **For Classification**:
  - Logistic Regression
- **For Both (Regression & Classification)**:
  - Decision Tree
  - Random Forest
  - AdaBoost
  - XGBoost
  - CatBoost

Unsupervised Machine Learning

Unlike supervised learning, **unsupervised machine learning** deals with data where the **output feature is unknown**. The objective is not to predict a value but to discover structure, such as finding similar **clusters** or groups within the data.

Example: Customer Segmentation

An e-commerce company may want to group customers to send targeted discount coupons. The dataset might include **Salary** and **Spending Score** (1-10).

- **Cluster 1**: High Salary, High Spending Score.
- **Cluster 2**: Medium Salary, High Spending Score.
- **Cluster 3**: Low Salary, Low Spending Score.

By creating these clusters, a company can identify which customers (e.g., those with high salary and spending) should receive specific offers.

Algorithms

Common algorithms used for clustering include:

- K-Means Clustering
- Hierarchical Mean (Hierarchical Clustering)
- DBSCAN Clustering

Reinforcement Learning

Reinforcement Learning is a distinct module where an application learns to perform tasks autonomously by receiving **rewards** for its actions.

Concept & Analogy

The process is often compared to a **baby learning to walk**:

1. **Trial**: The baby attempts to walk.
2. **Failure**: The baby falls and may feel pain (negative feedback).
3. **Learning**: The baby adjusts their behavior to avoid pain and eventually walks successfully.

Similarly, in reinforcement learning, the system performs actions and optimizes its strategy based on the rewards it receives (e.g., positive reinforcement for a correct action, similar to a student getting good marks).