Basics of Machine Learning - Notes

What is Machine Learning (ML)?

Machine Learning is a subset of Artificial Intelligence. It enables systems to learn patterns from data and improve without being explicitly programmed.

Types of Machine Learning

- Supervised Learning Learn from labeled data (e.g., Linear Regression, Decision Trees)
- Unsupervised Learning Learn from unlabeled data (e.g., K-Means Clustering)
- Reinforcement Learning Learn via rewards/punishments (e.g., Q-Learning)

Basic Terminologies

- Model: The algorithm or mathematical representation used to predict or classify.
- Features: Independent variables or input data.
- Labels: Output or result we want to predict.
- Training: Process of learning patterns from data.
- Testing: Evaluating model on new, unseen data.

Steps in a Machine Learning Project

1. Problem Definition
2. Data Collection
3. Data Cleaning and Preprocessing
4. Feature Selection/Engineering
5. Model Selection
6. Training the Model
7. Model Evaluation
8. Deployment

Popular Algorithms

Linear Regression Supervised Predict continuous value

Logistic Regression Supervised Classification problems

Decision Tree Supervised Classification/Regression

K-Means Clustering Unsupervised Grouping similar data

Naive Bayes Supervised Text classification

SVM Supervised High-dimensional data

Evaluation Metrics

- Accuracy
- Precision & Recall
- F1 Score
- Confusion Matrix
- MAE
- RMSE

Common Libraries (Python)

- scikit-learn
- pandas
- numpy
- matplotlib
- tensorflow, keras
- pytorch

Basic Code Example (Python)

```
from sklearn.datasets import load_iris
from sklearn.model_selection import train_test_split
```

```
from sklearn.tree import DecisionTreeClassifier


data = load_iris()

X_train, X_test, y_train, y_test = train_test_split(data.data, data.target, test_size=0.2)


model = DecisionTreeClassifier()

model.fit(X_train, y_train)


accuracy = model.score(X_test, y_test)

print(f"Accuracy: {accuracy}")
```