



What is Machine Learning (ML)?

Definition:

Machine Learning is a branch of Artificial Intelligence (AI) that focuses on building systems that **learn from data** — without being explicitly programmed for every rule.

Instead of hardcoding rules, we give the machine **examples**, and it **learns patterns** from the data to make predictions or decisions.

Let's dive into **Machine Learning (ML) fundamentals**, focusing on the key concepts of **Supervised Learning**, **Unsupervised Learning**, and **Model Evaluation**.



1. Supervised Learning

Definition:

In supervised learning, the algorithm is trained on a **labeled dataset**, meaning each training example includes the **input data** and the **correct output** (label). The goal is to learn a mapping from inputs to outputs.

Examples:

Use Case	Input	Output (Label)	Type
Spam detection	Email text	Spam or Not Spam	Classification
House price prediction	Features (size, location)	Price (\$)	Regression
Image classification	Image pixels	Object category	Classification

Common Algorithms:

- **Linear Regression** (for regression)
- **Logistic Regression** (for binary classification)
- **Decision Trees**
- **Random Forest**
- **Support Vector Machines (SVM)**

- **K-Nearest Neighbors (KNN)**
 - **Neural Networks**
-

□ **2. Unsupervised Learning**

Definition:

Unsupervised learning works on **unlabeled data**. The goal is to find **hidden patterns** or **structures** in the data.

Examples:

Use Case	Input	Output	Type
Customer segmentation	Purchase history	Customer groups	Clustering
Dimensionality reduction	High-dimensional data	Lower-dim data	Compression
Anomaly detection	Network traffic	Outliers	Detection

Common Algorithms:

- **K-Means Clustering**
 - **Hierarchical Clustering**
 - **DBSCAN**
 - **Principal Component Analysis (PCA)**
 - **Autoencoders**
-

✓ 3. Model Evaluation

Once you've trained a model, you need to measure how well it's performing. Different techniques are used depending on whether you're doing **classification** or **regression**.

□ a. For Classification

🔍 Confusion Matrix:

Predicted Positive Predicted Negative
Actual Positive True Positive (TP) False Negative (FN)
Actual Negative False Positive (FP) True Negative (TN)

📊 Key Metrics:

Metric	Formula	Meaning
Accuracy	$(TP + TN) / \text{Total}$	Overall correctness
Precision	$TP / (TP + FP)$	How many predicted positives were correct
Recall	$TP / (TP + FN)$	How many actual positives were found
F1 Score	$2 \times (\text{Precision} \times \text{Recall}) / (\text{Precision} + \text{Recall})$	Balance between precision and recall
ROC-AUC	Area under the ROC curve	Measures tradeoff between TPR and FPR

📉 b. For Regression

Metric	Formula	Meaning
Mean Absolute Error (MAE)	$\text{avg}(y_{\text{true}} - y_{\text{pred}})$	$y_{\text{true}} - y_{\text{pred}}$
Mean Squared Error (MSE)	$\text{avg}((y_{\text{true}} - y_{\text{pred}})^2)$	Penalizes larger errors more
Root Mean Squared Error (RMSE)	$\sqrt{\text{MSE}}$	Square root of MSE
R ² Score (R-squared)	$1 - (SS_{\text{res}} / SS_{\text{tot}})$	Proportion of variance explained by the model

□ c. Cross-Validation

Instead of training and testing on the same dataset, we use **cross-validation** to ensure the model generalizes well.

🌀 K-Fold Cross-Validation:

- Split data into **K equal parts**
- Train on **K-1** folds, test on the remaining one
- Repeat **K times** and average the results

🌀 Summary Table

Concept	Description	Example Algorithms
Supervised Learning	Learn from labeled data	Linear/Logistic Regression, SVM, Random Forest
Unsupervised Learning	Discover patterns in unlabeled data	K-Means, PCA, Autoencoders
Model Evaluation (Classification)	Accuracy, Precision, Recall, F1, ROC-AUC	Confusion matrix
Model Evaluation (Regression)	MAE, MSE, RMSE, R ² Score	Residual plots, error metrics
Cross-Validation	Robust model testing method	K-Fold, Stratified CV

All above ML Basics concepts guide given step by step with code examples :-

<https://geocities.ws/ommauryasir/ai/ai-model/What%20is%20an%20AI%20Model%20and%20How%20to%20develop%20ai%20model.pdf>

Check Simple Example On ML Concept With Code Example :-

House Price Prediction in India (Simulated)

🔗 We'll Use:

- scikit-learn for model training (LinearRegression)
- pandas for a custom dataset
- Features relevant to Indian housing

<https://geocities.ws/ommauryasir/ai/ai-model/ai-model-for-house-price-prediction.pdf>

Here's your Full Pytorch Code for Indian housing price prediction, to accept user input after training — so users can predict prices for their own custom home features (like size, location, bedrooms, and bathrooms).

<https://geocities.ws/ommauryasir/ai/ai-model/pytorch-housing-ai-model.pdf>

House price prediction model with user inputs using Tensorflow :-

<https://geocities.ws/ommauryasir/ai/deep-learning/House%20price%20prediction-model-with-tensorflow.pdf>