# Perceptron

## Introduction

The Perceptron is one of the earliest supervised learning algorithms designed for binary classification. It is a simple yet powerful model that learns a linear decision boundary to separate two classes. The Perceptron relies on an iterative learning process where it adjusts its weights based on errors, gradually improving its classification accuracy. While it only works for linearly separable data, it laid the foundation for more advanced machine learning models, including neural networks.

## Mathematical Foundation

### 1. Linear Combination

For a given feature vector $x = (x_1, x_2, \ldots, x_n)$, with corresponding weights $w = (w_1, w_2, \ldots, w_n)$ and bias $b$, the Perceptron computes:

$$\hat{y} = \text{sign}(w^T x + b) \tag{1}$$

where:

- $w^T x$ represents the dot product of the weights and input features.
- $b$ is the bias term, which shifts the decision boundary.
- The sign function outputs 1 if $w^T x + b \geq 0$, and $-1$ otherwise.

### 2. Weight Update Rule

If a data point $(x, y)$ is misclassified, the Perceptron updates its weights and bias using:

$$w \leftarrow w + \eta \cdot y \cdot x \tag{2}$$

$$b \leftarrow b + \eta \cdot y \tag{3}$$

where:

- $\eta$ is the learning rate, controlling the step size of updates.

- $y$ is the actual class label (1 or $-1$).
- $x$ is the input feature vector.

## Algorithm Steps

1. Initialize weights and bias to small random values or zeros.
2. Repeat for a fixed number of iterations or until convergence:
   (a) Compute the prediction $\hat{y}$ using the current weights and bias.
   (b) If $\hat{y} \neq y$, update weights and bias using the update rule.
3. Stop when all points are correctly classified (for linearly separable data) or when the iteration limit is reached.

## Key Characteristics

- Type: Binary classification model.
- Learning: Supervised.
- Decision Boundary: Linear.
- Activation Function: Sign function.

## Strengths

- Simple and computationally efficient.
- Guaranteed to converge if the data is linearly separable.
- Serves as the foundation for more complex models, such as neural networks.

## Weaknesses

- Can only classify linearly separable datasets.
- Produces hard binary decisions with no probabilistic outputs.
- The learning rate impacts convergence speed but does not influence the final solution for separable data.

## Applications

Despite its simplicity, the Perceptron has been applied in various domains:

- Image Recognition: Early applications of the Perceptron included digit and character recognition.
- Spam Filtering: Used in basic email spam classifiers before more advanced methods like Naive Bayes and deep learning.
- Medical Diagnosis: Applied to classify diseases based on medical features.
- Finance: Utilized in risk assessment models to classify borrowers as low or high risk.

## Conclusion

The Perceptron is a fundamental machine learning algorithm that introduced key ideas such as weight updates based on errors and iterative learning. While its ability is limited to linearly separable data, it remains an essential concept in machine learning. Many modern models, including deep neural networks, build upon the principles established by the Perceptron.