Perceptron and Linear Classifiers

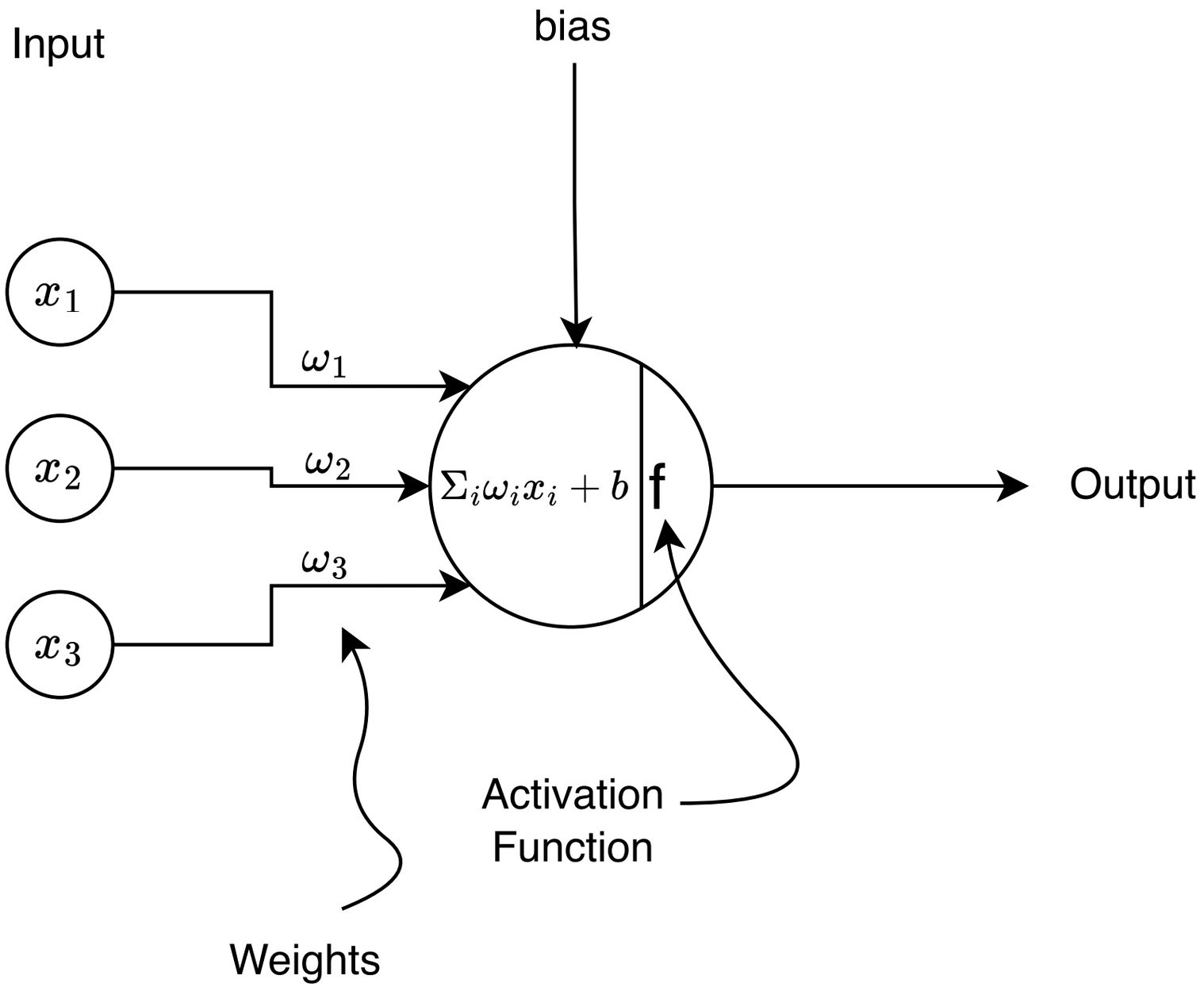
Foundations of Simple Machine Learning Models

Sina Moghimi, Autumn 2024

Outline

- Single Neuron Model
- Mathematical Representation
- Linear Separability Concept
- Decision Boundaries
- Training the Perceptron
- Limitations of Linear Models

Artificial Neuron (Perceptron) Structure



Mathematical Representation

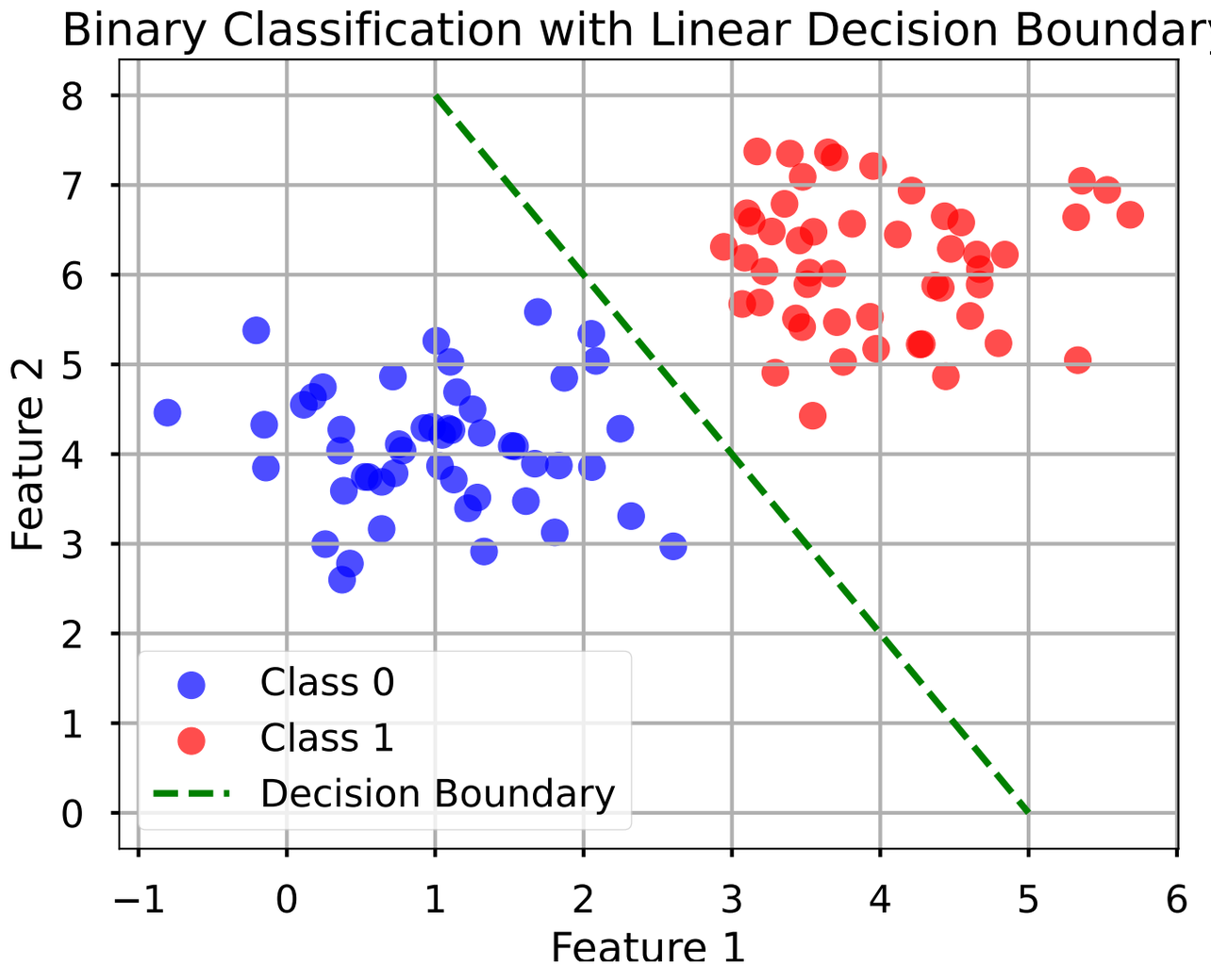
$$Output = f(bias + \sum_{i=1}^{n} \omega_i x_i)$$

Linear Separability

- It defines whether the data can be divided into distinct classes using a straight line (in 2D), a plane (in 3D), or a hyperplane (in higher dimensions)

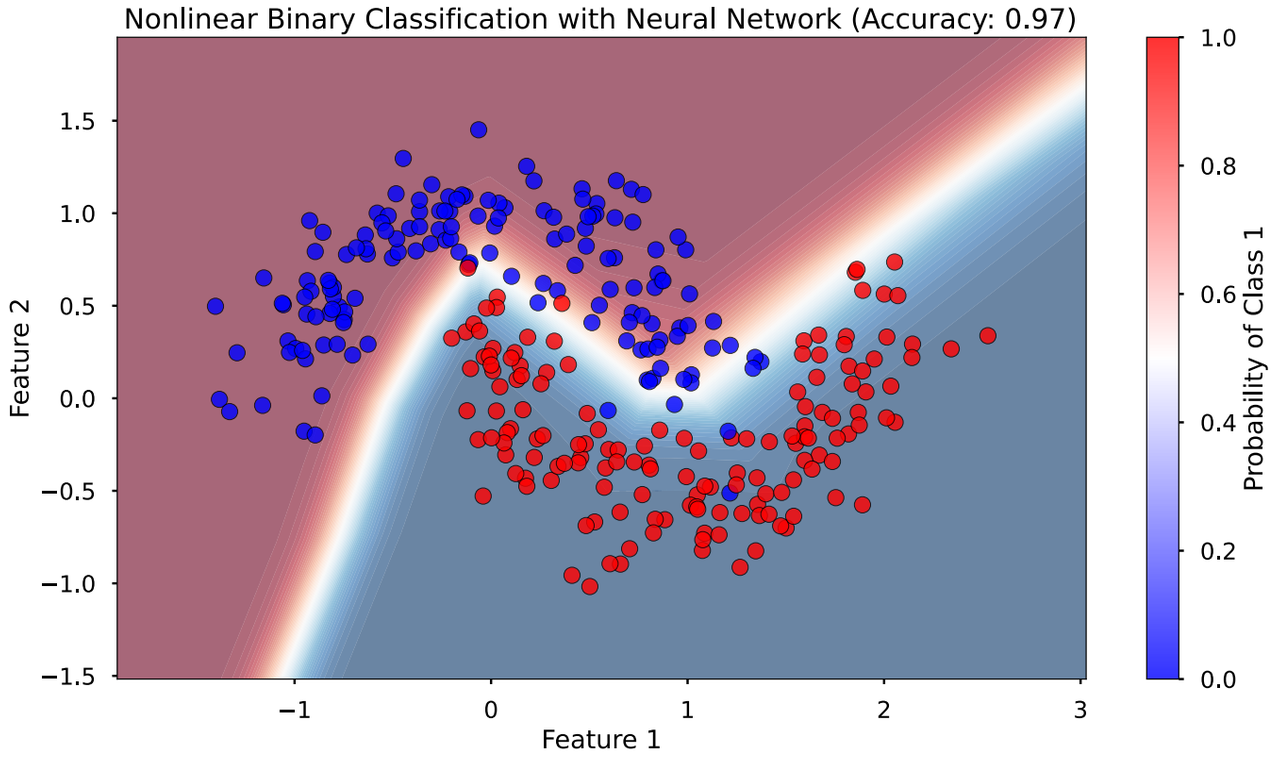
- *Line equation* can be interpreted as a **single neuron** with **1 input** and **linear activation function.** ***w*** *would be the slope, and* ***b*** *would be the intercept.*

- Usage:
  - Classifying emails as normal or spam
  - Classifying images as cats or dogs
  - …

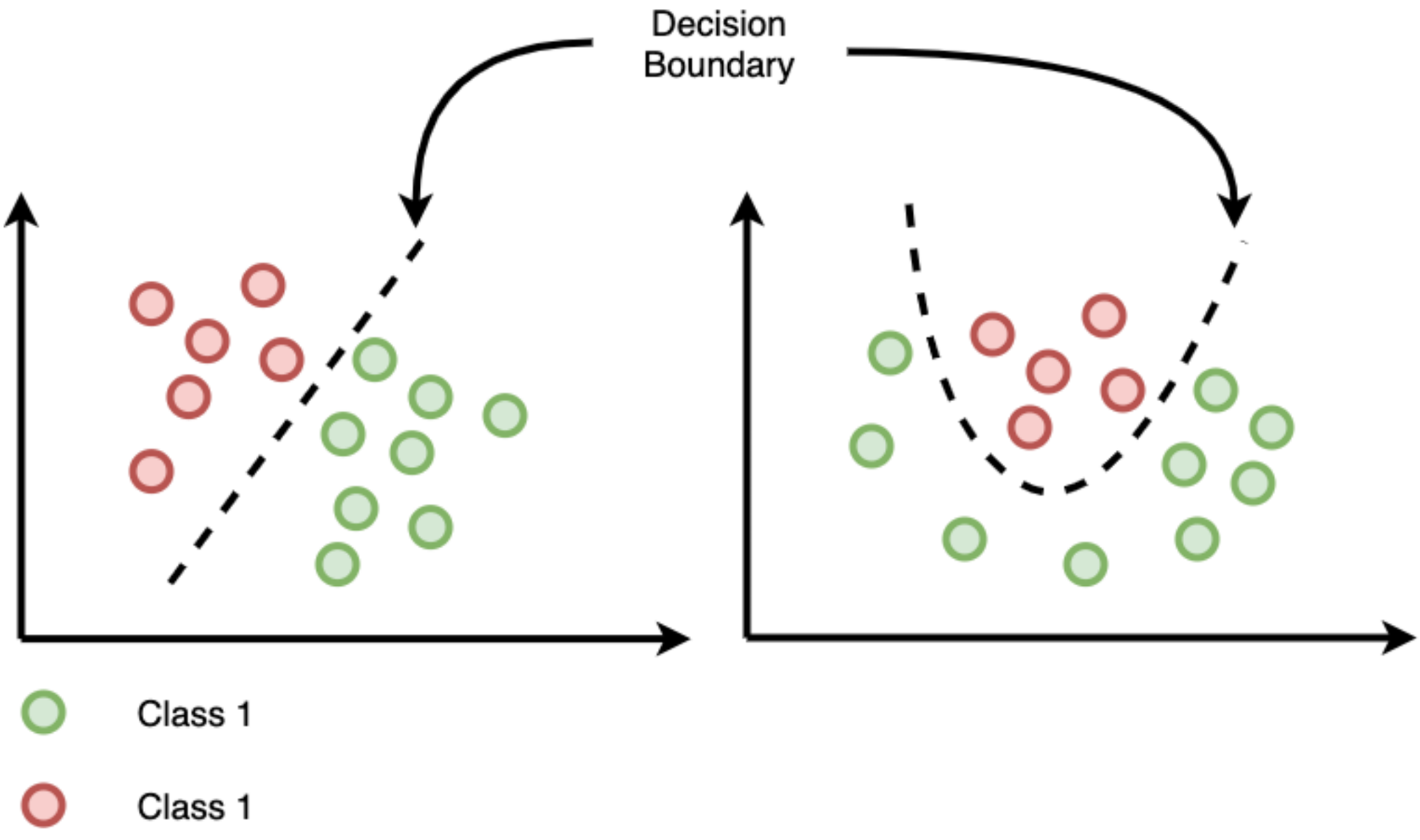


Decision Boundaries

- Decision Boundary, is the boundary in the feature space that separates different classes. It can be a line, plane, hyperplane.
- Types
  - Linear
    - Occurs in simple models where objects belonging to different classes, can be easily separated by a straight line(previous slide) /plane/hyperplane.
  - NonLinear
    - Occurs in complex models(neural networks with hidden layers) and nonlinear activation functions.
    - Enables complex separations such as curved or irregular boundaries



Linearity vs. NonLinearity



https://towardsdatascience.com/logistic-regression-and-decision-boundary-eab6e00c1e8

Training a Perceptron

1. Get a set of inputs and compute the weighted sum

$$z = \omega_1 x_1 + \omega_2 x_2 + \ldots + \omega_n x_n + b$$

2. Apply a step activation function

$$y = \begin{cases} 1, & if \quad z \geq 0 \\ 0, & if \quad z < 0 \end{cases}$$

3. Calculate error

$$error = y_{true} - y_{pred}$$

4. Update Weights

$$\omega_i \leftarrow \omega_i + \eta \cdot error \cdot x_i$$

5. Update Bias

$$b \leftarrow b + \eta \cdot error$$

6. Repeat until error gets near a desirable value like 0.0

МФТИ

Limitations of Linear Models

- XOR problem
- Inability to solve non-linear problems
- Oversimplification of complex datasets