Feature Scaling Methods

Feature scaling is an important preprocessing step in machine learning, especially when using gradient-based algorithms. Here are the three common methods for feature scaling:

1. Min-Max Scaling

Min-max scaling rescales the feature values to a fixed range, typically between 0 and 1. For a feature x_1, the min-max scaled value $x_1^{(scaled)}$ is given by:

$$x_1^{(scaled)} = \frac{x_1}{\max(x_1)}$$

Where:

- x_1 is the original feature value,
- $\max(x_1)$ is the maximum value of x_1 across all examples in the dataset.

This ensures that the resulting value will lie between 0 and 1.

2. Mean Normalization

Mean normalization centers the feature values around 0. The mean normalized value $x_1^{(mean\ norm)}$ is calculated as:

$$x_1^{(mean\ norm)} = \frac{x_1 - \mu_1}{\max(x_1) - \min(x_1)}$$

Where:

- x_1 is the original feature value,
- μ_1 is the mean value of the feature x_1 across all examples,
- $\max(x_1)$ and $\min(x_1)$ are the maximum and minimum values of x_1 in the dataset.

In this case, the scaled values are centered around zero and can be both positive or negative.

3. Z-Score Normalization

Z-score normalization, also known as standardization, transforms the feature values based on their mean and standard deviation. The z-score normalized value $x_1^{(z)}$ is given by:

$$x_1^{(z)} = \frac{x_1 - \mu_1}{\sigma_1}$$

Where:

- x_1 is the original feature value,
- μ_1 is the mean of the feature x_1,
- σ_1 is the standard deviation of the feature x_1.

Z-score normalization produces values that are centered around zero with a standard deviation of 1.