Scaling is one of the important steps while data pre-processing. It ensures that all the features are treated equally by the algorithm. Sometimes, due to varying range of different features some of the algorithms tend to add weight as per the values in those features.

Scaling is required in gradient based algorithms like Linear/Logistic Regression and Artificial Neural Networks. Scaling is irrelevant in tree based or distance-based algorithms like Support Vector Machine, K-Nearest Neighbours, Decision Trees, Random Forest, XG Boost.

Normalization and Standardization are both different types of scaling techniques.

**Normalization**

It converts all the feature values in the range of 0 to 1.

X’ = (x – min(x)) / (max(x) – min(x))

**Standardization**

Standardization, also known as z-score normalization, is conducted to transform data to have mean of zero and standard deviation of 1.

Z = (x – mean(x)) / std

**When to use what?**

1. In case of neural networks, Normalization is preferred since we do not assume any data distribution.
2. Standardization is preferred when data follows gaussian distribution
3. In case the data has lot of outliers, standardization is preferred.

Reference: <https://www.youtube.com/watch?v=bqhQ2LWBheQ>