**Data Preprocessing**

**The Steps of Data Preprocessing:**

The machine learning process involves 3 steps:

1. **Data Pre-Processing:**
   1. Import the data
   2. Clean the data
   3. Split into training and test sets
2. **Modelling:**
   1. Build the model
   2. Train the model
   3. Make predictions
3. **Evaluation:**
   1. Calculate performance metrics
   2. Make a verdict

**Feature Scaling:** It is the process of converting character values into relational quantities, basically they become values in comparison to each other.

Feature scaling is always applied to columns, never rows.

Two main methods of feature scaling: Normalisation and Standardisation.

Normalisation:

It is the process of taking the minimum inside a column, subtracting that from every value in the column, and dividing by the difference of the maximum and minimum.

Range: [0:1]

Standardisation:

­­Similar to Normalisation, we just subtract the average and divide by the standard deviation.

Range: [-3:+3]