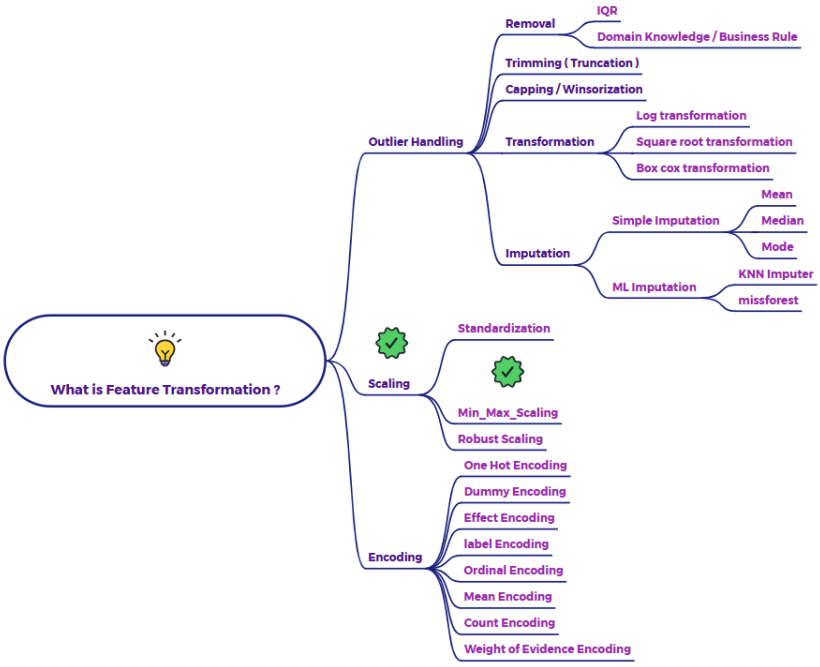
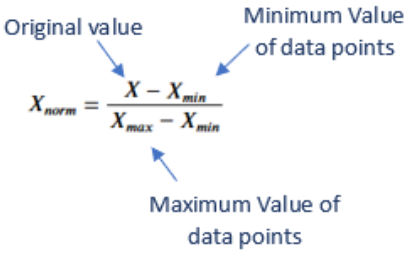
Explain Min-max scaling



Min-Max Scaling (Normalization)

1. **Explanation of Min-Max Scaling**
    - Min-Max scaling is a technique that scales numerical data to fit within a specific range, typically between 0 and 1. It transforms values linearly, based on the minimum and maximum values of the original data.
2. **How to Calculate Min-Max Scaling**



The formula for Min-Max scaling is:

x'=(x−xmin)/(xmax−xmin)

Where:

- (x): Original value
- (x_{min}): Minimum value of the variable
- (x_{max}): Maximum value of the variable

Example:

Let's say we have the following data for a variable "Income" (in thousands of dollars): 30, 50, 20, 60, 40

Calculate the minimum value ((x_{min})): 20

Calculate the maximum value ((x_{max})): 60

Scale each value:

- For 30: (30 - 20) / (60 - 20) = 0.25
- For 50: (50 - 20) / (60 - 20) = 0.75
- For 20: (20 - 20) / (60 - 20) = 0
- For 60: (60 - 20) / (60 - 20) = 1
- For 40: (40 - 20) / (60 - 20) = 0.5

So, the Min-Max scaled "Income" values are: 0.25, 0.75, 0, 1, 0.5

3. **When to Use Min-Max Scaling**
    - When you need your data to be within a specific range, such as 0 to 1.
    - When you are using algorithms that are sensitive to the scale of the input data, such as:
        - Neural Networks
        - K-Nearest Neighbors (KNN)
    - When you do not assume any specific distribution for your data.
4. **Strengths and Weaknesses of Min-Max Scaling**
    - **Strengths:**
        - Simple to implement and easy to understand.
        - Preserves the original distribution of the data.
        - Useful when you need data in a bounded range.
    - **Weaknesses:**
        - Sensitive to outliers. Outliers can significantly affect the (x_{min}) and (x_{max}) values, leading to poor scaling for the rest of the data.
        - Not robust to new data points that fall outside the original range used to calculate (x_{min}) and (x_{max}).