

Day-10

Outliers

Outliers are Sharma ji ka beta !!!

Outlier removal technique:

1. Z-Score Method:

- Description: Identifies outliers by computing the Z-score of each data point, which measures how many standard deviations away the point is from the mean. Points with Z-scores above a threshold are considered outliers.
- Generally outside $(-3\sigma \text{ and } 3\sigma)$ is considered outliers.

2.IQR (Interquartile Range) Method:

- Description: Identifies outliers by calculating the range between the first quartile (Q1) and third quartile (Q3).
- Outliers are those outside the range of $[Q1 - 1.5*QR]$ and $[Q3 + 1.5*QR]$.

3.Percentile

- Percentile-Based involves using percentiles (e.g., 1st and 99th) to define the range for capping. Values below the lower percentile or above the upper percentile are capped to these percentile values.

4.Capping

- Capping involves setting a maximum (upper cap) and minimum (lower cap) value for a feature. Data points beyond these caps are adjusted to the cap values.