

Which method is commonly used to fill missing numerical values?

- A) Mode
- B) Median
- C) Label Encoding
- D) One-Hot Encoding

উত্তর: B

ব্যাখ্যা: Numerical missing values সাধারণত **median** দিয়ে পূরণ করা হয় কারণ এটি outlier-এর প্রভাব থেকে নিরাপদ।

Which technique converts categorical labels into numbers like 0, 1, 2?

- A) Standardization
- B) Label Encoding
- C) Normalization
- D) Binning

উত্তর: B

ব্যাখ্যা: **Label Encoding** ক্যাটেগরিক্যাল ডেটাকে সংখ্যায় (0, 1, 2) রূপান্তর করে।

Scaling কোন কাজের জন্য করা হয়?

- A) Missing value fill করা
- B) Outlier detect করা
- C) Features একই scale এ আনা
- D) Column drop করা

উত্তর: C

ব্যাখ্যা: Scaling ডেটার সব numerical feature-কে একই স্কেলে আনার জন্য ব্যবহৃত হয়।

Which method is used to handle outliers effectively?

- A) Min-Max Scaling
- B) Mean Imputation
- C) IQR Method
- D) PCA

উত্তর: C

ব্যাখ্যা: Outlier detect ও handle করতে **IQR (Interquartile Range)** পদ্ধতি সাধারণত সবচেয়ে ব্যবহৃত হয়।

Polynomial Features কেন ব্যবহার করা হয়?

- A) Missing value তৈরি করতে
- B) Model complexity বাড়াতে
- C) Categorical encode করতে

D) Outlier remove করতে

উত্তর: B

ব্যাখ্যা: Polynomial features নতুন ফিচার তৈরি করে **model**-এর **non-linear** সম্পর্ক ধরতে সাহায্য করে।

One-Hot Encoding কোন ধরনের ডেটার জন্য ব্যবহৃত হয়?

A) Numerical

B) Continuous

C) Categorical (Non-ordinal)

D) Time-series

উত্তর: C

ব্যাখ্যা: **One-Hot Encoding** Non-ordinal categorical ভেরিয়েবলকে binary columns-এ রূপান্তর করে।

Normalization সাধারণত কোন ক্ষেত্রে ব্যবহৃত হয়?

A) যখন ডেটার range খুব বড়

B) যখন missing value বেশি

C) যখন category বেশি

D) যখন regression করা হয়

উত্তর: A

ব্যাখ্যা: **Normalization** ডেটার মানকে 0 থেকে 1 এর মধ্যে স্কেল করে, যখন ডেটার range বড় হয়।