

Data Preprocessing & Feature Engineering Assignment-2

Q. write short description of encoding & its methods.

Ans: Encoding in pandas refer to the process of transforming categorical data into numerical form so that it can be easily analyzed and used in machine learning models.

Pandas provide several methods for encoding categorical data, including:

- **Label encoding:** This method replaces each unique value in a categorical column with a numerical label. The labels are assigned based on the order in which the unique values appear.
- **One-Hot encoding:** This method creates a new column for each unique value in a categorical column & assigns a binary value (0 or 1) to indicate whether that value is present in the original column.
- **Binary encoding:** This method converts each unique value in a categorical column into a binary code & creates new column for each digit in the code.

- Ordinal Encoding: This method assigns a numerical value to each unique value in a categorical column based on a specified order.