Assignment-4
Basics of Feature Engineering

October 25, 2025

1. Create a dataset of house properties containing numeric columns for length, breadth, and price. Construct a new feature that represents the total floor area and analyze its effect on the target variable.

2. Prepare a dataset containing details such as gender, city, and qualification. Convert all categorical columns into numeric features suitable for machine learning models.

3. Given a dataset containing employee satisfaction levels ("Excellent", "Good", "Average", "Poor"), convert them into numeric values representing their order.

4. Build a dataset of product prices and transform the continuous price column into three categories such as "Low", "Medium", and "High".

5. Using a dataset with multiple numeric features, apply dimensionality reduction to obtain two principal components and visualize the transformed data.

6. Generate a numeric dataset and decompose it into three matrices using SVD. Reconstruct the approximate original dataset from the decomposed matrices.

7. Using a labeled dataset, apply LDA to reduce dimensionality and project the data into a lower-dimensional space while maintaining class separability.

8. Load any dataset and select a subset of columns manually using .iloc. Train a simple model using only the selected subset and compare its performance with the full dataset.

9. Create a dataset with two or more highly correlated features. Detect and remove redundant features using appropriate correlation analysis or other methods.