**PROJECT TITLE:**

**costumer segmentation using data science**

**PROBLEM DEFINITION:**

* The problem is to implement data science techniques to segment customers based on their behavior, preferences, and demographic attributes. The goal is to enable businesses to personalize marketing strategies and enhance customer satisfaction. This project involves data collection, data preprocessing, feature engineering, clustering algorithms, visualization, and interpretation of results.

**PRE-PROCESSING:**

**STEPS:**

1. **DATA CLEANING**
2. **HANDLE MISSING VALUES**
3. **CATEGORICAL TO NUMERICAL REPRESENTATIONS.**

**DATA CLEANING:**

Data cleaning is the process of fixing or removing incorrect, corrupted, incorrectly formatted, duplicate, or incomplete data within a dataset. When combining multiple data sources, there are many opportunities for data to be duplicated or mislabeled.

**HANDLE MISSING VALUES:**

1. Deleting Rows with missing values
2. Impute missing values for continuous variable
3. Impute missing values for categorical variable
4. Other Imputation Methods
5. Using Algorithms that support missing values
6. Prediction of missing values
7. Imputation using Deep Learning Library — Datawig.

**CATEGORICAL TO NUMERICAL REPRESENTATIONS:**

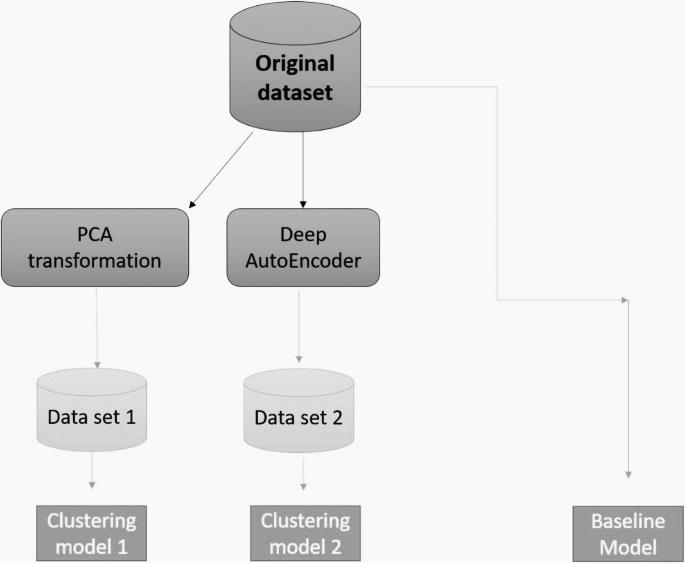
1. cat.codes Attribute
2. replace
3. Label Encoder

**ALGORITHM:**

1. DEEP LEARNING TECHNIQUE : PCA
2. ATTENTION MECHANISMS

**PCA :**

Using unsupervised learning techniques of Principal Component Analysis (PCA) and KMeans to identify customer segments of the German population that were popular or less popular with a mail-order sales in Germany.

****

**ATTENTION MECHANISM:**

An attention mechanism is an Encoder-Decoder kind of neural network architecture that allows the model to focus on specific sections of the input while executing a task. It dynamically assigns weights to different elements in the input, indicating their relative importance or relevance.

**PROJECT WORKFLOW:**

