**Telecust---KNN-Project**

Imagine a telecommunications provider has segmented its customer base by service usage patterns, categorizing the customers into four groups. If demographic data can be used to predict group membership, the company can customize offers for individual prospective customers. It is a classification problem. That is, given the dataset, with predefined labels, we need to build a model to be used to predict class of a new or unknown case. The example focuses on using demographic data, such as

1. Region,
2. Age, and
3. marital Status to predict usage patterns.

**The target field, called custcat, has four possible values that correspond to the four customer groups, as follows:**

1- Basic Service 2- E-Service 3- Plus Service 4- Total Service

**Aim : Our objective is to build a classifier, to predict the class of unknown cases. We will use a specific type of classification called K nearest neighbour.**

**Libraries USed**

1. pandas
2. numpy
3. matplotlib.pyplot
4. seaborn as sns
5. sklearn import preprocessing
6. sklearn.model\_selection import train\_test\_split
7. sklearn.neighbors import KNeighborsClassifier
8. sklearn import metrics
9. sklearn.metrics import confusion\_matrix
10. sklearn.metrics import classification\_report

**Steps Used**

**1. Extract Features**

a. Features and target should not have any null values b. Features should be numeric (handle categorical values) c. Features should be of the type array/ dataframe d. Features should be having some rows and columns e. Features should be on the same scale

**2. Split the dataset into training and testing datasets**

**3. Train the model on the training set**

a. Import the necessary library b. Instantiate an object c. .fit() on training data 4. Test/ Evaluate the model on the testing set a. .predict() on test data b. Evaluate the performance (Compare the Actual vs. Predicted values) c. If performance is not upto the mark, use performance optimization (Dimensionality Reduction, HPO, Change the ML algorithm)

**Summary :**

Successfully built the KNN model and tested, found that the accuray in not that great and hence, further performance optimization is required like (Dimensionality Reduction, HPO, Change the ML algorithm)

Note : I have solved this case study just to "build and understand the KNN". (Helpful for Beginners to understand the concept)