

Task - KNN from Scratch

Dataset - Click here to download the dataset: [diamonds.csv](#)

Data Description - [description.txt](#)

Task - Predict the diamond price. 

Write the KNN code from scratch and make it work on the given dataset ?

Step - 1: Load the data

Step - 2: Identify input and output variables

Step - 3: Split the data - Test and Train (recommended 75:25 split)

Step - 4: Data Preprocessing on X_train (You can use sklearn for data preprocessing)

- Categorical Data Encoding
- Numerical Data Rescaling

Step - 5: Data Preprocessing on X_test

Step - 6: Build the model and predict on X_test (**SCRATCH IMPLEMENTATION**)

Implement KNN Algorithm from scratch and do the predictions for test data. You should not use the sklearn KNN algorithm here. Write the complete code implementation from scratch for KNN algorithm.

Step - 7: Evaluate your model

Step - 8: Train a model using sklearn KNN Algorithm and compare the results with your scratch implementation

Dos -

1. You can use sklearn for standardizing and splitting the dataset to train and test
2. Scratch implementation of sklearn's KNN fit function
3. You can use library functions for LabelEncoder or OneHotEncoder
4. Use Jupyter Notebook to implement your work
5. Eat, Code, Sleep, Repeat

Don't -

1. Plagiarism
2. Procrastination