Documentation of Scaling Techniques

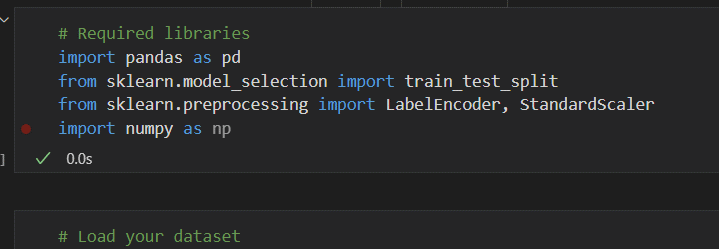
Introduction

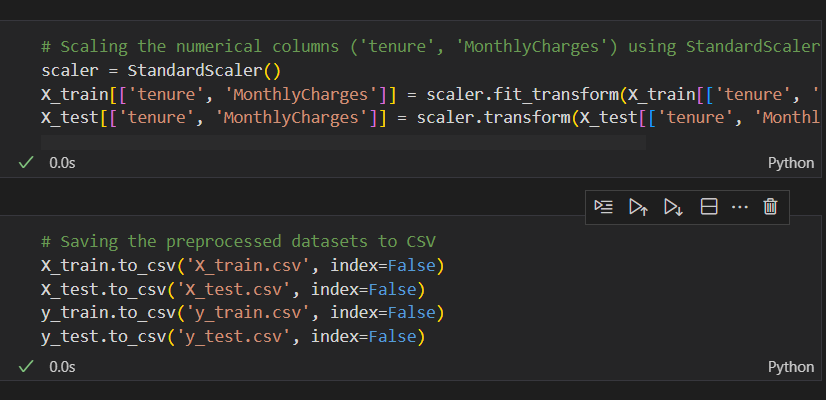
Here the scaling methods used on the dataset of the clustering analysis project that is intended to predict customer churn are described. Scaling is often a prerequisite step in machine learning to while normalizing the features so they all have an equivalent importance.

Applied Scaling Technique

In this project, we used a StandardScaler method to scale the feature set to unit variance. They all standardize the data in a manner that their mean becomes 0 and the standard deviation becomes 1. It is important to standardize, for example, for methods that involve distance such as K-Means clustering algorithm.

Code snippets:





Conclusion

The features in the dataset were normalized using the StandardScaler. This type of preprocessing is critical to the subsequent K-Means clustering model resulting to formation of the right relationship between the customer churn and corresponding predictions.