Romika Sairam

October 14, 2021

CSE454

Time Series: Planning

Summary:

The objective of the time series project is to take in a given dataset, read it in and create two data representations of the dataset using PAA and SAX. Here, the goal is to reduce the number of samples needed to represent the given data. The second piece of this project is the classification, which is done through Euclidean Distance and Manhatten Distance. Here, our goal is to make valid conclusions on which one is better and if there are any noticeable trends.

Inputs: The input in this case is the dataset that we are given.

Outputs: The output is the two sets which are new representations of the datasets and the classifications.

Constraints: The only constraints we have in this project are the specified methods to represent and classify the data. Another constraint is to only use the given dataset, to perform these operations.

Assumptions: For this project, we can assume the given dataset is credible and is accurate for us to use it for further calculations and changes.

Ethical Issue: An ethical concern for this project is the dataset we are using. There is a possibility that the readers of this report for this project, could think that is is not a credible dataset since we are retrieving it from another website. Credibility could be questioned.