ABSTRACT

Data mining is interesting patterns and trend finding process in large database. Outlier defined as a data point in database where is different than data point from common database with fixed size. Even outlier have an abnormal behaviour, often contain important information. Outlier detection have important role in fraud detection, intrusion detection, and network monitoring application.

Finding an outlier usually using proximity based on existing remain data. In high dimensional data, data become spare, finally proximity notion data become failed.

In this final assignment, will doing methods comparison finding outlier in high dimensional data. Existing methods which will be use is Clustering-based, Density-based, and Distance-based. Where each methods support on high dimensional data.

Keywords : data mining,outlier, outlier detection, outlier detection method.