

Anomaly detection

The goal of this homework is to implement a simple anomaly detector and observe the effect of its hyperparameters on the accuracy of the method.

Instructions

1. In provided data, you have normal samples (`normal.txt`) and anomalous samples (`easy.txt`).
2. Load the data to R and split normal data to training and testing part. All anomalous data will be used for testing.
3. Implement anomaly k-nearest neighbor anomaly detector as was discussed on lectures.
4. Plot an ROC curve for a given detector for various k .
5. Visualize, to which part of the space the detector assigns high / low anomaly score using heatmap.
6. Estimate area under ROC curve for each provided problem from ten repetitions.