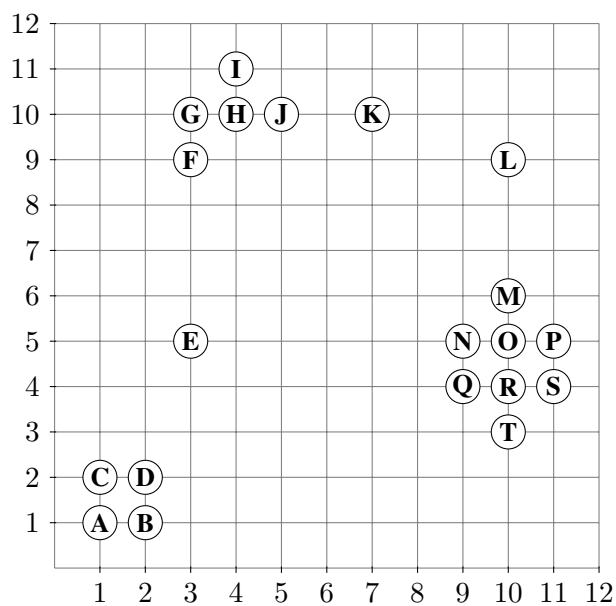


Knowledge Discovery in Databases
 SS 2012

Übungsblatt 9: Outlier Detection

Aufgabe 9-1 Outlier Scores



As distance function, use Manhattan distance $L_1(a, b) := |a_1 - b_1| + |a_2 - b_2|$.

Compute the following (without including the query point when determining the k NN):

- LOF using $k = 2$ for the points E , K and O .
- LOF using $k = 4$ for the points E , K and O .
- k NN distance using $k = 2$ for all points.
- k NN distance using $k = 4$ for all points.
- aggregated k NN distances for $k = 2$ and $k = 4$ for all points
 (aggregated k NN distance = sum of the distances to all the k NN!)

Discuss the choice of k for this data set.

Compare your results to the ELKI results using algorithms `outlier.LOF`, `outlier.KNNOutlier` and `outlier.KNNWeightOutlier`!