University of Southern Denmark
IMADA
Arthur Zimek
Morteza Chalabi Hajkarim

DM 555: Data Mining and Statistical Learning
Autumn term 2016

Repetition 1

Repetition 1-1

(a) Describe the KDD process model.

(b) Characterize supervised vs. unsupervised learning methods.

(c) What are feature spaces and feature transformations?

(d) What are L_p-norms? How can we attribute different importance to features or to feature combinations in the distance computation?

(e) What is the "Apriori principle"?

(f) What is the frequent pattern mining problem? What are association rules? What are *good* association rules?