

# Mining massive Datasets WS 2017/18

## Problem Set 8

Rudolf Chrispens, Marvin Klaus, Daniela Schacherer

December 13, 2017

### Exercise 01

- a) The probability that C1 and C2 are matched?

QuestionnairesCount =  $b ( Q_1, \dots, Q_b )$

QuestionCount =  $r$

QuestionCountSum =  $b * r$

MatchCondition =  $\text{if}(mQ_x(r) == fQ_x(r)) // m = \text{person1}, f = \text{person2}, x = \text{element of } b$

*Probability – of – an – event – happening* =  $\frac{\text{Number-of-ways-it-can-happen}}{\text{Total-number-of-outcomes}}$

Propability that 1 Questionair matsches:  $(\frac{1}{r})^r * b$

HIER BRAINSTORME ICH EINFACH NICHT BEACHTEN :D

- b) The probability that exactly two (no matter which) questionnaires match, i.e. have the same answers for both C1 and C2.

### Exercise 02

MISSING

### Exercise 03

MISSING

### Exercise 04

MISSING

### Exercise 05

MISSING

## **Exercise 06**

MISSING

## **Exercise 07**

MISSING