Wolfgang K. Seiler
Tel. 2515

B6, B4.18

November 9, 2018

Mathematics and Information, Exercise sheet 9

Problem 1: (*7 points*)

a) Some collection contains six documents, in which the following words occur:

D_1: *Shannon, Entropy, Information*
D_2: *Boltzmann, Entropy, Clausius, Heat*
D_3: *Information, Shannon, Code, Cryptanalysis, Key*
D_4: *Kelly. Shannon, Bet, Portfolio, Information*
D_5: *Las Vegas, Shannon, Kelly*
D_6: *Shannon, Juggling, Roboter, Unicycle*

Construct a term-document-matrix for this collection in which all column vectors have length one!

b) Code the query *Information Shannon* by a unit vector and compute the cosine of the angle between this vector and each of the six document vectors!

Problem 2: (*6 points*)

a) Let (t_i, x_i), $i = 1, \ldots, 100$ be data points for which a relation of the form $x_i = a \sin t_i + b \sin 2t_i + c \sin 3t_i + d \sin 4t_i$ is supposed to hold. Which system of linear equations gives the best values for the coefficients a, b, c, d in the sense of least squares?

b) How can you proceed if a connection of the form $x_i = a \cos(t_i + c)$ with unknown parameters a, c is suspected?

Problem 3: (*7 points*)

Determine parameters $a, b, c \in \mathbb{R}$ such that the relation $z = a + bx + cy$ holds approximately for the following points $P_i = (x, y, z) \in \mathbb{R}^3$:

$$P_1 = (1,1,1), \quad P_2 = (1,2,3), \quad P_3 = (1,3,2), \quad P_4 = (2,3,4), \quad P_5 = (0,4,5), \quad P_6 = (1,-1,3)$$

Can be submittel till Thursday, November 15, 2018 at 11.55h