Ludwig-Maximilians-Universität München
Institut für Informatik
Lehrstuhl für Mobile und Verteilte Systeme
Dr. Juan Bernabé-Moreno

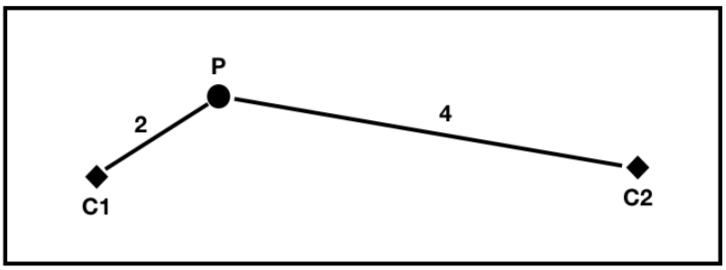
mobile and distributed systems group

Exercise Sheet 3
Intelligent Systems - WS 23/24

Exercise 1: Fuzzy Clustering

(Pkt.)

Given is a point cloud with the two cluster centers $C1$ and $C2$. For a clearer representation, only a single data point P is mapped.



Furthermore, the two distances $d(P, C1) = 2$ and $d(P, C2) = 4$ as well as the to be minimized objective function are given:

$$J(X, B, U) = \sum_{i=1}^{c} \sum_{j=1}^{n} u_{ij}^{w} d^2(\overrightarrow{\beta_i}, \overrightarrow{x_j})$$

X is the set of data points, B is the set of cluster prototypes, and U is a fuzzy partition matrix. As fuzzifier $w = 2$ was chosen.

a. In each of the following cases, calculate the resulting value of the objective function J when, for the given data point P, the following degrees of membership have been calculated:

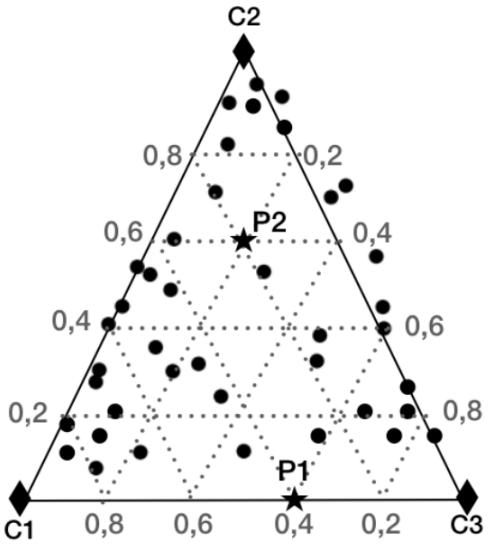
   (i) $\overrightarrow{U_1} = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$

   (ii) $\overrightarrow{U_2} = \begin{pmatrix} 0.9 \\ 0.1 \end{pmatrix}$

   (iii) $\overrightarrow{U_3} = \begin{pmatrix} 0.8 \\ 0.2 \end{pmatrix}$

b. Which of the given degrees of membership from a) is to be preferred for the minimization of the objective function?

c. Now let us assume that we have a data set represented in coefficient space with auxiliary lines drawn in for membership weighting as well as the three cluster centers C1, C2, C3 and the two marked points P1 and P2.



Fill in the following table by determining the fuzzy affiliations of the two points P1 and P2 to the cluster centers C1, C2, C3.

	C1	C2	C3
P1			
P2			

Exercise 2: Finetuning of Foundation Models

(Pkt.)

Finetuning foundation models is an important approach for training highly specialized ML models. In this exercise, we are finetuning the *distilBERT* model from Hugging Face: `https://huggingface.co/distilbert/distilbert-base-uncased`. Install the `requirements.txt` since we are using *torch*, *transformers* and *datasets*. Complete the code in file `finetuning.py` to finetune the *distilBERT* model on the *IMDb* film review text corpus.