SIFI2300 - Artificial Intelligence and Digitalization

25.03.2026

- Ziyu (ziyuzh@uio.no)
- Bachelor in Biotechnology
- Masters in Bioinformatics
- PhD student in Scientific Computing and Machine Learning

Communication channels

- Most recommended:
  - Mattermost
  - Email ziyuzh@uio.no

From one student to another

- Create an overview of the course
- Do weekly excercises
- Write code for the algorithms you struggle with



The *k*-Means Algorithm

- **Initialisation**
  - choose a value for k
  - choose k random positions in the input space
  - assign the cluster centres $\boldsymbol{\mu}_j$ to those positions
- **Learning**
  - repeat
    - for each datapoint $\mathbf{x}_i$:
      - compute the distance to each cluster centre
      - assign the datapoint to the nearest cluster centre with distance

      $$d_i = \min_j \; d(\mathbf{x}_i, \boldsymbol{\mu}_j). \qquad (14.1)$$
    - for each cluster centre:
      - move the position of the centre to the mean of the points in that cluster (N_j is the number of points in cluster j):

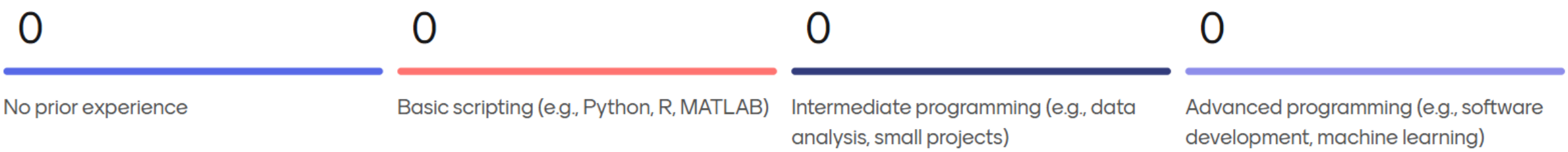
      $$\boldsymbol{\mu}_j = \frac{1}{N_j} \sum_{i=1}^{N_j} \mathbf{x}_i \qquad (14.2)$$
    - until the cluster centres stop moving
- **Usage**
  - for each test point:
    - compute the distance to each cluster centre
    - assign the datapoint to the nearest cluster centre with distance

$$d_i = \min_j \; d(\mathbf{x}_i, \boldsymbol{\mu}_j). \qquad (14.3)$$

Join at menti.com | use code 4481 0344

Mentimeter

What prior experience do you have with programming?



Context today

- Coding in Jupyter Notebook
- Prepare for the exercises later

Jupyter notebook overview

How to start

- Cloud-based (no installation needed)
  - Google Colab: example
- Local installation (on your computer)
  - Anaconda (Navigator + JupyterLab/Notebook): Setup with Anaconda – Universitetet i Oslo

Coding with AI Assistants

- Learn the basics first https://colab.research.google.com/github/jckantor/CBE30338/blob/master/docs/01.02-Python-Basics.ipynb
- Ask questions, then learn
- Solve actual problems
- Don’t blindly trust it
- Refer to the documentation
- Debugging

Useful links

Always check official websites for documentation:

- Numpy https://numpy.org/numpy-tutorials/index.html
- Pandas https://pandas.pydata.org/docs/getting_started/index.html#getting-started
- Scikit-learn https://scikit-learn.org/stable/auto_examples/index.html

- Matplotlib
- https://matplotlib.org/stable/

Break 15 min

Individual help