UNIVERSITY
OF SKÖVDE

School of Information Technology

WRITTEN RE-EXAMINATION

Course: Big Data Programming

Examination: Final written re-exam

Course code: IT739A

Credits for written examination: 5.5hp

Date: May 13, 2026

Examination time: 8:15 – 12:30

Examination responsible: Richard Senington

Teachers concerned: Juhee Bae, Richard Senington, Gunnar Mathiason

Aid at the exam/appendices

Other

Instructions

- ☐ Take a new sheet of paper for each teacher.
- ☒ Take a new sheet of paper when starting a new question.
- ☒ Write only on one side of the paper.
- ☒ Write your name and personal ID No. on all pages you hand in.
- ☒ Use page numbering.
- ☒ Don´t use a red pen.
- ☒ Mark answered questions with a cross on the cover sheet.

Grade points

Examination results should be made public within 18 working days

Good luck!

UNIVERSITY
OF SKÖVDE

IT739A, Re-Exam, May 13th, 2026, at 8:15-12:30

Instructions that must be followed:

- The max number of points listed with each question indicates how thoroughly you are expected to answer the question.
- Number your answer clearly with the same numbering as the question and its sub-questions.
- Be clear in your writing. Sometimes fewer but more concise formulations are better than writing a lot of text.
- More points assigned indicate that a more detailed answer is expected.

The grading of your answers considers:

- The **correctness** of your answer/explanation
- The **clarity and logical cohesion** of your answer/explanation
- The **conciseness** of your answer/explanation

Grading levels:

- There are 5 questions (with sub-questions), which give a maximum 10 points each.

1) You need to pass 5 points per each of the 5 questions to pass the exam (25 points).
2) Accumulated points above that threshold linearly determines the exam grading (the exam maximum is 50 points).

- Exam grading range is A to F, which also determines the final course grading:
  A: 45-50 points, B: 40-44 points, C: 35-39 points, D: 30-34 points, E: 25-29 points.

Examination goals:

This exam assesses the following Course Curriculum goals

- *critically evaluate the relevance and application of advanced AI methods,*
- *in an in-depth manner account for the development of advanced AI methods,*
- *apply and demonstrate appropriate analysis strategies using advanced AI tools and techniques,*
- *in an in-depth manner account for ethical use of AI to solve complex problems*

The exam questions begin on the next page.

UNIVERSITY
OF SKÖVDE

Re-Exam questions:

1. Deep Learning, Keras and Tensorflow (10 points):

1. Explain how the Loss function and the Optimizer interact while training a neural network in a typical supervised learning setup. What are their differences? (2p)
2. Why do you need to normalize or standardize your input data when training neural networks? (1p)
3. What is the effect of using "dropout" when training a neural network? Briefly describe the training process for how to achieve this? (2p)
4. Describe a way to detect overfitting of a neural network during training, using plotting of the accuracy and loss parameters. (2p)
5. Explain how the weights in a Convolutional layer are stored and updated using kernels/filters. When are the weights updated in these kernels in a network with stacked Convolutional layers? (3p)

2. Anomaly Detection (10 points):

1. How do variational auto-encoders vary the basic design of an autoencoder, and why might it help anomaly detection? (3p)
2. Describe how a loss function for an Auto-Encoder works, and hence how it is used to train an Auto-Encoder. (2p)
3. YOLO can effectively identify different kinds of objects in an image, and where they are. This is a useful property for anomaly detection, but what is the problem with using YOLO for this? (1p)
4. Describe how the PADIM algorithm works. (4p)

3. Recurrent Neural Networks (10 points):

1. Describe what a recurrent network is, how it differs from a standard neural network and why it is effective for working with data series. (3p)
2. Give one advantage (1) and one disadvantage (1) of recurrent networks vs Transformer Networks for time series analysis. (2p)
3. What is the vanishing gradient problem in the context of recurrent networks? (1p)
4. Describe GRU networks and explain what the difference is between LSTM and GRU type recurrent networks. (4p)

4. Concepts and ethics in Big Data (10 points):

1. One of the 5Vs of Big Data is Volume. Please describe the Volume aspect of big data and how it relates to Deep Learning. (3p)

UNIVERSITY
OF SKÖVDE

2. Explain how graph databases and a Hadoop distributed filesystem differ for big data storage. (2p)
3. Discuss the ethical implications of sustainability in relation to Big Data and Deep Learning AI techniques. (2p)
4. Discuss the ethical implications of image generation and manipulation using deep learning techniques. (3p)

5. Large Language Models (LLM) (10 points):

1. Briefly describe the role of the reward model in RLHF. (1p)
2. Provide an example of bias or undesired behavior in RLHF. (1p)
3. A language model generates responses based on user input. Briefly describe the main steps in the LLM text-generation pipeline, from input text to generated output (mention the key components involved) and explain why the model can predict the next token effectively. (3p)
4. A company deploys a chatbot that produces fluent but sometimes incorrect answers. What evaluation method would you use to detect this issue, what would you measure, and why? (2p)
5. Design a system for a customer-support assistant that must:
   - Follow strict company tone.
   - Use internal documents.
   - Be cost-efficient.

   Explain how you would combine at least two of the following: prompting, RAG, LoRA, RLHF. Justify your design. (3p)