

DATA.ML.200 Deep Learning

Example exam X.X.2024

Exam prepared by Tuomas Virtanen

Use of a non-programmable calculator is allowed (but not necessary for the completion). Indicate if you have completed the project work or exercises of the course some previous year.

THE EXAMPLE EXAM HAS A BIT MORE QUESTIONS THAT AN ACTUAL EXAM WILL HAVE.

- Describe shortly (max. 25 words each) the what the following operations do (max. 10 points each)
 - Softmax
 - Batch normalization
 - Pooling
 - Backpropagation
 - Dropout
- Describe shortly (max. 25 words each) for what purpose the following methods are used in deep learning (max. 10 points each)
 - Attention
 - Gradient descent
 - Sigmoid
 - Binary cross-entropy
 - Positional encoding
- Three tasks (a-c) below need to be solved by deep neural network based methods based on three available architectures: 1) convolutional neural network, 2) recurrent neural network, 3) transformer. Choose the most suitable architecture for each task. Justify shortly (maximum 40 words) each answer. Each architecture can only be used once (max. 10 points each).
 - machine translation
 - real-time speech denoising
 - image denoising
- Explain how a deep neural network is trained in supervised manner, if there is a dataset available with pairs consisting of input and target outputs. Use maximum one page for your answer (max 30 points).
- Describe how Transformers work. Use maximum one page for your answer (max 30 points).
- Explain how a typical CNN-based image classification system operates. Use maximum one page for your answer (max. 30 points)
- Explain how generative adversarial network (GAN) can be trained to generate images of dogs, given a training dataset of dog images (max. 30 points)
- The formula below represents the processing in one self-attention layer (excluding scaling by \sqrt{D}). Explain what is the purpose of different parts of the formula (max 20 points).

$$\mathbf{Y} = \text{softmax}(\mathbf{QK}^T)\mathbf{V}$$