University of Antwerp

Artificial Neural Networks 2020 – 2021

Theory Exam - May 25th 2021

Name: ……

GUIDELINES

- Start by writing your name on each page and the additional pages
- This is a closed book exam.
- Make sure your mobile phone is switched off and place it at the front of the room.
- Be as complete as possible in your answers. Motivate clearly why you provide this answer. **Usually answers in 1 sentence are not enough!**
- All forms of communication are prohibited.

QUESTION 1: PERCEPTRON LEARNING

A) Parameters of artificial neurons include weights applied to the input and a bias term. 1) Indicate what is the function of the bias term in an artificial neuron, and 2) Indicate what would be the effect of removing it from an artificial neuron trained with it.

QUESTION 2: DEEP NEURAL NETWORKS

A) One of the defining characteristics of models following the "Deep Learning" paradigm is the relatively high number of internal layers that define the "deep" architectures that they posses. Indicate the advantages and disadvantages that having a deep architecture may bring.

B) Describe the training procedure of a deep neural network ***f*** trained for a classification task of ***C*** classes from a dataset $\mathbf{X = [\ x_i, y_i\ ]}$ where each example $\mathbf{x_i}$ is paired with an annotation $\mathbf{y_i}$. Describe relevant factors related to the forward-pass, backward-pass, optimization and stopping criteria involved in the training procedure.

QUESTION 3: LEARNING AND OPTIMIZATION

A) Drop-out is a technique that is commonly used for regularizing the weights of a neural network during training. 1) Indicate how this technique works and 2) indicate what are the advantages of applying it during training.

B) Following the stocastic gradient descent algorithms, parameters of a deep neural network are updated based on the equation $\theta_{t+1} = \theta_t - \alpha_t \nabla_\theta L(\theta_t)$ where α_t is the learning rate set during training. Early works applied a fixed learning rate while follow-up work adopted dynamic learning rates. 1) Indicate how learning rates are applied in a dynamic manner in practice. 2 Indicate what are the advantages, and potential disadvantages, of using a dynamic learning rate instead of a fixed.

QUESTION 4: CONVOLUTIONAL NEURAL NETWORKS

A) Explain how the Squeeze and Excitation (SE) Module propagates information computed by a given filter to different spatial locations in other channels.

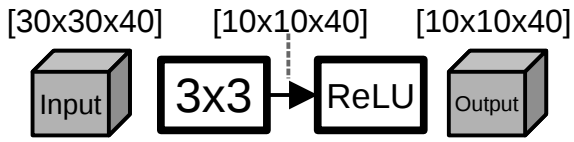
B) A company is developing an image recognition algorithm based on convolutional neural networks. They have produced two different architectures that are capable of

Name: ……

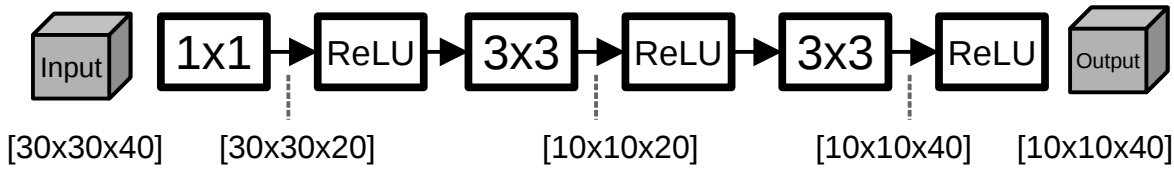
producing the same internal representation from a given input, thus achieving the same classification performance. The diagram below illustrates the two models in question, each block represents a layer. Each convolutional layer indicates its kernel size within it. The size of the intermediate representation after different operations is indicated by the dashed lines.

- Indicate what are the main differences between these two architectures, potential strengths and weaknesses that they may have and scenarios when it would be better using one over the other. Justify your answer appropriately.

Architecture-1



Architecture-2



QUESTION 5: MODELING SEQUENCES WITH NEURAL NETWORKS

A) Indicate the main difference between Recurrent Neural Networks and Long Short - Term Memory Networks.

B) Data related to sequence modelling problems is provided as a set of valid sequences $\{x\}_i$. Describe the training procedure of Recurrent architectures for sequence modelling problems and indicate how they are optimized without the need of annotations.

QUESTION 6: TRANSFER LEARNING

A) Describe the procedure of response-based knowledge distillation.

QUESTION 7: DEEP GENERATIVE NETWORKS

A) Describe the training procedure of denoising autoencoders and indicate what is the added value of this procedure when compared to that of standard autoencoders.

B) Describe two differences between autoencoders and variational autoencoders.

QUESTION 8: INTERPRETATION AND EXPLANATION

A) Describe two differences between input modification methods and deconvolution-based methods for model explanation.