SAMPLE EXAM

Deep Learning – SoSe 2022 – Prof. Dr. Constantin Pape

1 NOTES

1. Questions that require a written answer can be answered in a few sentences. No need to write long answers!
2. Multiple choice questions can have multiple (or no) correct answer unless it's specified to pick one.
3. The sample exam consists of 7 exercises. Each exercise consists of one or multiple questions / task and gives a maximum of 10 points. The full exam will consist of 10 such exercises.
4. I have tried to give a good idea of the kinds of exercises that will be part of the exam, but did not have enough time to balance the questions. Please be prepared that the easier exercises may contain more questions in the exam.

2 EXERCISES

2.1 Perceptrons

2.1.1 Under what conditions can a Perceptron learn to classify a training set of points perfectly?

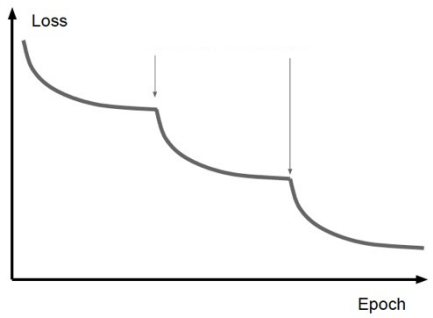
[ ] The two classes are separable by an arbitrary decision boundary.
[ ] The two classes are linearly separable.
[ ] Perfect classification is not possible.

2.1.2 Draw an example of a training set of 10 points that cannot be classified perfectly by any Perceptron.

2.2 Loss functions & optimization

2.2.1 Does optimization by gradient descent always end up in the global mininum? Why/why not?

2.2.2 What happend at the points of the arrows?



2.2.3 What diagnostic function does overfitting the neural network on a single minibatch of training data serve?

2.3 **CNNs**

Suppose with have a CNN with the following layers:

0. Input images: RGB with 12×12 pixels
1. 4 kernels of size 4×4, stride 2, no zero padding (= “valid” padding)
2. 8 kernels of size 3×3, stride 1, no zero padding (= “valid” padding)
3. 16 kernels of size 3×3, stride 1, no zero padding (= “valid” padding)

2.3.1 What is the receptive field size in pixels of a unit in the second layer?

2.3.2 What is the dimensionality of the output?

2.3.3 How many parameters does the network have? Let’s assume there is no batch normalization and we don’t use biases in the convolutional layers. You can write down the formula without calculating the final result.

2.3.4 In a convolutional neural network, the number of parameters is independent of the image size.

[ ] Correct

[ ] Wrong

2.4 Regularization & Normalization

2.4.1 Name at least two commonly used mechanisms for regularizing neural network training.

2.4.2 Which of the following statements about Dropout (p=0.5) are **correct** (multiple answers possible)?

[ ] During training 50% of the weights are set to zero in each iteration.

[ ] 50% of the units are removed from the network at random.

[ ] During inference, activations have to be scaled by a factor of 2.

[ ] Dropout can be thought of as training an ensemble of models.

[ ] The randomness introduced by Dropout during training hurts model performance.

2.4.3 For which of the following normalization schemes is the following statement **correct**?

(multiple answers possible)

“_____ behaves differently during training and inference time.”

[ ] Batch normalization

[ ] Layer normalization

[ ] Group normalization

[ ] Instance normalization

2.5 Learning with few labels

2.5.1 A friend asks you for advice on a project that involves recognizing different car models. She has collected a labeled dataset containing 10 images for each of 50 different car models. Which of the following approaches would you recommend if classification accuracy is the primary objective? (pick one)
[ ] Train a ResNet-18 from scratch
[ ] Use a ResNet-18 pre-trained on ImageNet, replace and train only classification layer
[ ] Use logistic regression
[ ] Search for a small CNN architecture that performs well
[ ] Use a ResNet-152 pre-trained on ImageNet replace and train only classification layer

2.5.2 What is the difference between self-supervised and semi-supervised learning?

2.5.3 List 3 types of weak labels that can be used for weakly supervised object detection.

2.6 Generative Models

2.6.1 Which of the following generative models allows us to compute the likelihood of an image under the model?
[ ] Variational Autoencoder
[ ] GAN
[ ] PixelCNN

2.6.2 What is the training objective of the discriminator in a GAN?

2.7 Advanced

Proof that end-to-end learning of instance segmentation is not possible using a “simple” loss function like cross entropy. Hint: laying out a sketch of the proof is sufficient, it does not need to be rigorous, but should convey the idea of the proof.