Online Exam Machine Learning (AI5031)

Prof. Dr. Alexander Gepperth
Datum: February 23^{rd}, 2021

Start: 8.15
End: 9.45
Submission period: 9.45-10.15
Net exam time: 90 minutes
Max points: 90

Exercise 1: Derivatives (20P)

1.1 Simple derivatives (11P)

Compute the derivative $h'(x)$. Always give rule and decomposition!!
a) $h(x) = \sin(\sin(x))$
b) $h(x) = 3\sin(x) + 5\cos(x)$
d) $h(x) = \ln(\sin(\cos(5.2)))$
e) $h(x) = \exp(x)\cos(x)$

1.2 Higher-dimensional derivatives (3P)

Compute the following derivatives, no derivation required just a short justification
a) $\frac{\partial}{\partial x_2}\sum_{i=1}^{5} x_i$
b) $\frac{\partial}{\partial x_1 6}\sum_{i=1}^{5} x_i$
f) $\frac{\partial}{\partial x_3}\cos(x_3)x_4$

1.3 Gradient descent (6P)

Consider the function $f(\vec{x}) = 0.5x_1^2 + 5\pi$. Assuming a starting point of $\vec{x}(0) = [2, 3]$ and a step size of $\epsilon = 0.1$: perform 2 steps of gradient ascent, that is, compute the values of $\vec{x}(1)$, $\vec{x}(2)$ by iteration. Verify that the value of f is decreasing!

Exercise 2: numpy and matplotlib (25P)

In this whole exercise, assume that the 3D array 'traind' contains 2000 images of dimension 20x20, no need to generate it!

2.1 Basics, slicing (6P)

Give code snippets for the following operations:
a) Slice out the 2nd sample into an array x
b) Set the 2 lowermost rows of sample at index 1000 to -1
c) Slice out the sample with index 10, keeping every 3rd row!

2.2 Reduction and broadcasting (6P)

a) Compute the row-wise mean for all samples
b) Compute the column-wise max for all samples
c) Compute the mean image, and add it to all samples via broadcasting

2.3 Fancy indexing and mask indexing (5P)

a) copy out images with indices 3,4 and 17 using a single assignment operation!
b) copy out all images whose sum is larger than 200

2.4 Matplotlib (8P)

a) plot the function x^2 between 1 and 4 using 100 support points!
b) display the image with index 9
c) generate 100 numbers distributed according to a uniform distribution between 0 and 1, and display their histogram!

Exercise 3: TensorFlow (15P)

3.1 Softmax (5P)

Write a python function S(x) which takes an 1D TF tensor and returns the softmax, also as a TF tensor!

3.2 Affine layer (5P)

Given a 2D TF tensor ‘X’ of shape (100,20): give a snippet that performs an affine layer transformation on X. Include the creation of all required TF variables!

3.3 Computing gradients (5P)

Assuming a function f(X, W) that performs some computation on ‘X’ using the TF tensor ‘W’: give a code snippet that computes the gradient of f w.r.t. W!

Exercise 4: Back-propagation (10P)

4.1 Back-prop through affine layers (5P)

Assuming that, in affine layer X, we know $\frac{\partial \mathcal{L}}{\partial A^{(X)}}$.
Derive a symbolic expression for $\frac{\partial \mathcal{L}}{\partial A^{(X-1)}}$.

4.2 Back-prop through transfer function layers (5P)

Assuming that, in a transfer function layer X, we know $\frac{\partial \mathcal{L}}{\partial A^{(X)}}$. The transfer function layer computes $A_{ij}^{(X)} = -2A_{ij}^{(X-1)}$.
Derive a symbolic expression for $\frac{\partial \mathcal{L}}{\partial A^{(X-1)}}$.

Exercise 5: Discrete probability theory (20P)

5.1 Probabilities (5P)

Given are two sets of experimental results:

$$\begin{array}{lcccccccccccc} \textbf{Exp 1 (X):} & 1 & 2 & 2 & 1 & 3 & 1 & 1 & 2 & 2 & 2 & 1 & 1 \\ \textbf{Exp 2 (Y):} & 2 & 1 & 1 & 2 & 1 & 2 & 3 & 1 & 3 & 1 & 1 & 3 \end{array} \tag{1}$$

Compute (no justification):
a) $\#(Y=1)$, $\#(Y=2)$, $\#(Y=3)$
b) $\#(X=1, Y=1)$, $\#(X=1, Y=2)$, $\#(X=1, Y=3)$
c) $p(X=1, Y=1)$, $p(X=1, Y=2)$, $p(X=1, Y=3)$
d) $p(X=1|Y=1)$, $p(X=1|Y=2)$, $p(X=1|Y=3)$
e) $p(Y=1|X=1)$
f) $p(X=1|Y\neq 1)$

5.2 Confusion matrix (5P)

Evaluating a classifier yields the following confusion matrix:

$$\begin{pmatrix} 10 & 60 & 10 \\ 70 & 10 & 0 \\ 0 & 20 & 60 \end{pmatrix} \tag{2}$$

Compute:
a) The number of samples in the test set
b) $p(\hat{y}=1|\hat{t}=2)$
c) $p(\hat{t}=1)$, $p(\hat{t}=2)$, $p(\hat{y}=2)$
d) The probability of an incorrect classification under the restriction that the classifier output is 2
e) The probability of an incorrect classification!

5.3 Binary Classification 1 (5P)

A classifier which always chooses the negative class is evaluated on a test set in which the proportion of positive to negative samples is 1:3. Compute with justification:
a) tpr, fnr
b) fpr, tnr
c) the classification error

5.4 Binary classification 2 (5P)

A binary classifier with a sensitivity of 0.9 and a specificity of 0.5 is evaluated on a balanced dataset. Compute with justification:

a) tpr, tnr
b) fpr, fnr
c) the classification accuracy
d) precision using Bayes' theorem!