# DI504 Foundations of Deep Learning

## 01/12/2021

online (20:00-23:59)

- This is an online midterm exam with five questions. You are to upload the exam to ODTUClass tonight until 23:59. Please do not email your answers.
- Passing the work of others off as your own is a breach of academic ethics and also of the University's disciplinary rules.
- A simple photo of your sheet or word editing program outputs are both acceptable for your hand-written answers.
- Good luck. Please contact me by email, if you have any questions. I will answer them as fast as I can.

**Question 1 (20 Points):** Consider the function below:

$$f(x) = max\left[\begin{bmatrix}1\\0\end{bmatrix}, e^{\left(\mathbf{W}\cdot\mathbf{x}+\begin{bmatrix}0\\1\end{bmatrix}\right)}\right]$$

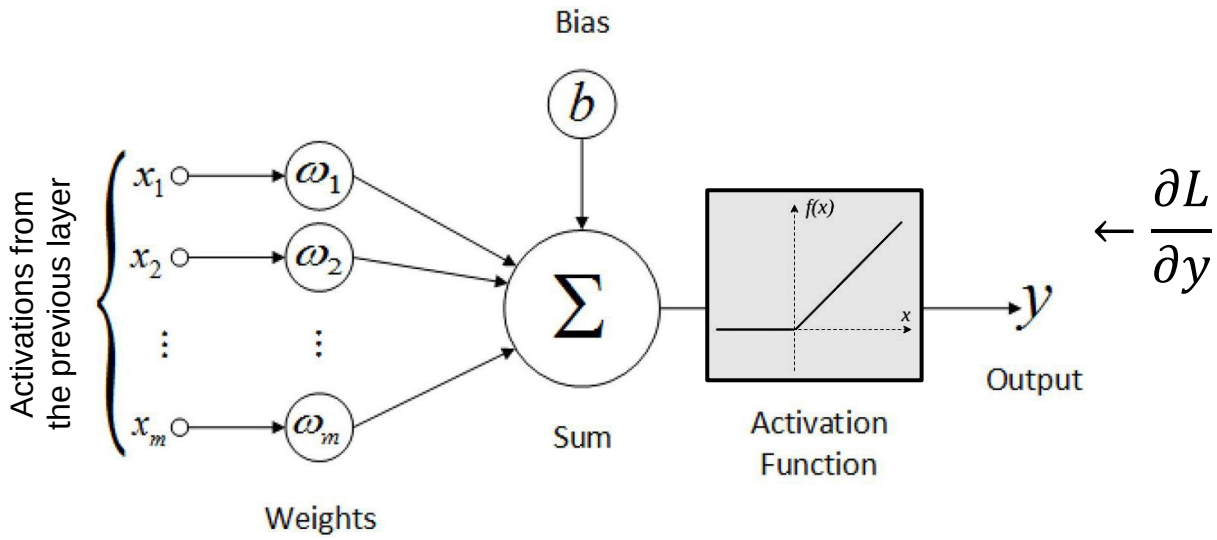
**a) (5 Points)** Draw the computational graph of $f(x)$, where each node is an atomic operation (i.e. product, summation, exponential, element-wise maximum, etc.)

**b) (5 Points)** On the computational graph, calculate and show each node output (activations) given that:

- $\mathbf{W} = \begin{bmatrix}-0.1 & -0.5\\ +0.3 & +0..2\end{bmatrix}$ and $\mathbf{x} = \begin{bmatrix}+0.3\\ -0.1\end{bmatrix}$

**c) (10 Points)** For the given values in part b, back-propagate and show the gradient values at each node, at **W**, and at **x**, given that the gradient at the output of the computational graph is $\begin{bmatrix}+1.0\\ +1.0\end{bmatrix}$.

**Question 2 (20 Points):** Consider the artificial neuron with a Rectified Linear Unit (ReLU) activation function.




Given the ReLU layer as $y = \mathbf{A}(\mathbf{W}^{\mathrm{T}} \cdot \mathbf{x} + \mathbf{b})$, where A is $max(0, \mathbf{x})$ function and provided that $\mathbf{x}, \mathbf{W}, \mathbf{b}, \frac{\partial \mathbf{L}}{\partial \mathbf{y}}$ has the values given below:

$\mathbf{W} = [-0.1 \quad -0.2 \quad 0 \quad -0.1 \quad -0.3]^T$
$\mathbf{x} = [0.7 \quad 0.9 \quad 0.9 \quad 0.2 \quad 0.9]^T$
$\mathbf{b} = [0.5]$
$\frac{\partial \boldsymbol{L}}{\partial \boldsymbol{y}} = [-0.124]$

There is a problem with this neuron. It cannot learn.

**a) (10 Points) Mathematically** explain and show the problem with this neuron.

**b) (10 Points)** Is there a way/possibility that this neuron may start to learn again in the succeeding iterations? Please explain how (or how not).

**Question 3 (20 Points):** We have a linear score function with L1 regularization, that accept Dx1 (D=5) vectors and creates scores for N=4 number of classes.

$$s_{Nx1} = f(\mathbf{x}; \mathbf{W}; \mathbf{B}) = (\mathbf{W}_{NxD} \cdot \mathbf{x}_{Dx1} + \mathbf{B}_{Nx1}) + \sum_{i=1}^{D} |\mathbf{W}_{N,i}|$$

Feeding these scores to a Multinomial (i.e. multiple class) Logistic Regression (i.e. Soft-Max) function we can obtain class probabilities for each class, for a given vector **x.**

$$P(y = k \mid \mathbf{x} = x_i) = \frac{e_k^s}{\sum_j (e_j^s)}$$

If $\mathbf{W} = \begin{bmatrix} -0.8 & -0.5 & +0.5 & +0.2 & -0.2 \\ +0.8 & -0.3 & 0 & -0.1 & -0.3 \\ 0 & +0.6 & +0.6 & +0.2 & -0.6 \\ -0.2 & -0.4 & +0.1 & -0.1 & +0.5 \end{bmatrix}$ and $\mathbf{B} = \begin{bmatrix} -0.3 \\ +0.1 \\ -0.7 \\ +0.9 \end{bmatrix}$

Find (and draw) the (discrete) class probability distribution functions (PDFs, i.e. scores) for the input vectors:

a) $\mathbf{x_A} = [\ 0.0 \quad 0.0 \quad 0.0 \quad 0.0 \quad 0.0]^T$

b) $\mathbf{x_B} = [+1.2 \quad +4.3 \quad +5.6 \quad +3.4 \quad +7.8]^T$

c) $\mathbf{x_B} = [-2.1 \quad +3.4 \quad +4.5 \quad -2.8 \quad +7.1]^T$

d) $\mathbf{x_D} = [-2.1 \quad +3.4 \quad +4.5 \quad -2.8 \quad +7.1]^T$ without L1 regularization.

**Question 4 (25 Points):** Consider the CNN architecture given in the table below. You are asked to fill the blank parameters in the table with proper values. Please note your calculations for each layer under the table. The table continues in the next page. **Please read the parameters very carefully!**

| *no.* | *name* | *type* | *parameters* | *activations (layer output)* | *learnables (weights/biases)* |
|---|---|---|---|---|---|
| **L0** | **data** | **Input** | - | **3001 x 33 x 1** | - |
| | **conv1_1** | **Convolution** | **filters: 64**<br>**size: 200 x 1 x 1**<br>**padding: [3 3 ; 0 0]**<br>**stride: [1 1]** | ? | Weights: 200 x 1 x 1 x 64<br>Bias: 1 x 1 x 64 |
| **L1** | **ReLU1_1** | **Rectified Linear Unit** | - | | - |
| | **conv1_2** | **Convolution** | **filters: 64**<br>**size: 3 x 1 x 64**<br>**padding: [1 1 ; 0 0]**<br>**stride: [1 1]** | ? | Weights: ?<br>Bias: |
| | **ReLU1_2** | **Rectified Linear Unit** | - | | - |
| **L2** | **pool1** | **Max-Pooling** | **size: 6 x 1**<br>**padding: [0 0 ; 0 0]**<br>**stride: [6 1]** | ? | Weights: ?<br>Bias: |
| | **conv2_1** | **Convolution** | **filters: 128**<br>**size: 3 x 1 x 64**<br>**padding: [1 1 ;0 0]**<br>**stride: [1 1]** | ? | Weights: ?<br>Bias: |
| **L3** | **ReLU2_1** | **Rectified Linear Unit** | - | | - |
| | **conv2_2** | **Convolution** | **filters: 128**<br>**size: 3 x 1 x 128**<br>**padding: [1 1 ; 0 0]**<br>**stride: [1 1]** | ? | Weights: ?<br>Bias: |
| | **ReLU2_2** | **Rectified Linear Unit** | - | | - |
| **L4** | **pool2** | **Max-Pooling** | **size: 6 x 1**<br>**padding: [0 0 0 0]**<br>**stride: [6 1]** | ? | Weights: ?<br>Bias: |
| | **conv3_1** | **Convolution** | **filters: 128**<br>**size: 3 x 1 x 128**<br>**padding: [1 1 ; 0 0]**<br>**stride: [1 1]** | ? | Weights: ?<br>Bias: |
| **L5** | **ReLU3_1** | **Rectified Linear Unit** | - | | - |
| **L6** | **conv3_2** | **Convolution** | **filters: 128**<br>**size: 3 x 2 x 128**<br>**padding: [1 1 ; 1 1]**<br>**stride: [1 3]** | | Weights: ?<br>Bias: |

| | | | | | |
|---|---|---|---|---|---|
| | **ReLU3_1** | **Rectified Linear Unit** | - | ? | - |
| | **pool3** | **Max-Pooling** | **size: 6 x 1**<br>**padding: [0 0 0 0]**<br>**stride: [6 1]** | ? | Weights:<br>?<br>Bias: |
| | **fc7** | **Fully Connected** | **nodes: 256** | ? | Weights:<br>?<br>Bias: |
| **L7** | **ReLU6** | **Rectified Linear Unit** | **-** | | - |
| | **fc8** | **Fully Connected** | **nodes: 256** | ? | Weights:<br>?<br>Bias: |
| **L8** | **ReLU7** | **Rectified Linear Unit** | **-** | | - |
| **L9** | **fc9** | **Fully Connected** | **nodes: 10** | ? | Weights:<br>?<br>Bias: |
| | **prob** | **Soft-max** | **nodes: 10** | ? | - |
| **L10** | **output** | **Classification Output** | | **10 categories** | - |

**Question 5 (15 Points):** Below you see the Log file of a semantic segmentation experiment (logs continue for 3 pages).

**a) (5 Points)** If the minibatch size is 32 samples, what is the size of the training set (roughly)?
**b) (10 Points)** Is the model learning well? Is it overfitting? What parameter would you change in the next experiment? Why?
(You can make assumptions, but please state them clearly).

| Epoch | Iteration | Time Elapsed (hh:mm:ss) | Mini-batch RMSE | Validation RMSE | Mini-batch Loss | Validation Loss | Base Learning Rate |
|---|---|---|---|---|---|---|---|
| 1 | 1 | 00:00:41 | 47.05 | 49.57 | 1106.7023 | 1228.5365 | 3.0000e-07 |
| 1 | 50 | 00:06:34 | 52.77 | | 1392.1007 | | 3.0000e-07 |
| 1 | 100 | 00:12:33 | 50.32 | | 1266.2394 | | 3.0000e-07 |
| 1 | 150 | 00:18:32 | 52.63 | | 1385.0675 | | 3.0000e-07 |
| 1 | 155 | 00:19:42 | 46.88 | 49.62 | 1098.9186 | 1231.2419 | 3.0000e-07 |
| 2 | 200 | 00:25:07 | 43.04 | | 926.3970 | | 3.0000e-07 |
| 2 | 250 | 00:31:08 | 43.95 | | 965.8181 | | 3.0000e-07 |
| 2 | 300 | 00:37:08 | 46.25 | | 1069.6370 | | 3.0000e-07 |
| 2 | 310 | 00:38:54 | 46.70 | 49.40 | 1090.5907 | 1220.2974 | 3.0000e-07 |
| 3 | 350 | 00:43:45 | 44.60 | | 994.6558 | | 3.0000e-07 |
| 3 | 400 | 00:49:45 | 51.97 | | 1350.4268 | | 3.0000e-07 |
| 3 | 450 | 00:55:44 | 61.09 | | 1866.0005 | | 3.0000e-07 |
| 3 | 465 | 00:58:06 | 46.56 | 49.19 | 1084.0491 | 1209.9977 | 3.0000e-07 |
| 4 | 500 | 01:02:19 | 54.38 | | 1478.5344 | | 3.0000e-07 |
| 4 | 550 | 01:08:18 | 48.22 | | 1162.5237 | | 3.0000e-07 |
| 4 | 600 | 01:14:17 | 51.07 | | 1304.2673 | | 3.0000e-07 |
| 4 | 620 | 01:17:16 | 46.49 | 49.01 | 1080.6145 | 1201.0879 | 3.0000e-07 |
| 5 | 650 | 01:21:02 | 48.38 | | 1170.4280 | | 3.0000e-07 |
| 5 | 700 | 01:27:08 | 40.79 | | 832.0249 | | 3.0000e-07 |
| 5 | 750 | 01:33:14 | 50.24 | | 1261.8347 | | 3.0000e-07 |
| 5 | 775 | 01:36:49 | 46.41 | 48.85 | 1076.9188 | 1192.9679 | 3.0000e-07 |
| 6 | 800 | 01:39:55 | 48.57 | | 1179.6279 | | 3.0000e-07 |
| 6 | 850 | 01:46:02 | 48.39 | | 1170.7449 | | 3.0000e-07 |
| 6 | 900 | 01:52:03 | 56.25 | | 1581.9155 | | 3.0000e-07 |
| 6 | 930 | 01:56:15 | 46.32 | 48.68 | 1072.9183 | 1185.1089 | 3.0000e-07 |
| 7 | 950 | 01:58:40 | 48.96 | | 1198.6907 | | 3.0000e-07 |
| 7 | 1000 | 02:04:42 | 47.13 | | 1110.6980 | | 3.0000e-07 |

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| 7 | 1050 | 02:10:45 | 47.93 | | 1148.4097 | | 3.0000e-07 |
| 7 | 1085 | 02:15:35 | 46.26 | 48.56 | 1069.8335 | 1179.0105 | 3.0000e-07 |
| 8 | 1100 | 02:17:25 | 44.67 | | 997.8920 | | 3.0000e-07 |
| 8 | 1150 | 02:23:26 | 42.23 | | 891.6169 | | 3.0000e-07 |
| 8 | 1200 | 02:29:28 | 58.46 | | 1708.7371 | | 3.0000e-07 |
| 8 | 1240 | 02:34:50 | 46.20 | 48.45 | 1067.4299 | 1173.8594 | 3.0000e-07 |
| 9 | 1250 | 02:36:04 | 41.59 | | 864.7968 | | 3.0000e-07 |
| 9 | 1300 | 02:42:04 | 47.19 | | 1113.6592 | | 3.0000e-07 |
| 9 | 1350 | 02:48:06 | 50.12 | | 1256.0687 | | 3.0000e-07 |
| 9 | 1395 | 02:54:05 | 46.11 | 48.35 | 1063.0153 | 1168.9586 | 3.0000e-07 |
| 10 | 1400 | 02:54:43 | 46.11 | | 1063.2104 | | 3.0000e-07 |
| 10 | 1450 | 03:00:44 | 44.56 | | 992.9717 | | 3.0000e-07 |
| 10 | 1500 | 03:06:48 | 63.51 | | 2016.8215 | | 3.0000e-07 |
| 10 | 1550 | 03:13:24 | 46.01 | 48.23 | 1058.6130 | 1163.0098 | 3.0000e-07 |
| 11 | 1600 | 03:19:29 | 49.71 | | 1235.7737 | | 3.0000e-07 |
| 11 | 1650 | 03:25:32 | 48.45 | | 1173.6980 | | 3.0000e-07 |
| 11 | 1700 | 03:31:33 | 50.29 | | 1264.4790 | | 3.0000e-07 |
| 11 | 1705 | 03:32:43 | 45.92 | 48.15 | 1054.4017 | 1159.3719 | 3.0000e-07 |
| 12 | 1750 | 03:38:10 | 41.64 | | 867.0588 | | 3.0000e-07 |
| 12 | 1800 | 03:44:13 | 43.24 | | 935.0587 | | 3.0000e-07 |
| 12 | 1850 | 03:50:16 | 44.82 | | 1004.4449 | | 3.0000e-07 |
| 12 | 1860 | 03:52:02 | 45.82 | 48.07 | 1049.5654 | 1155.1460 | 3.0000e-07 |
| 13 | 1900 | 03:56:53 | 42.69 | | 911.4063 | | 3.0000e-07 |
| 13 | 1950 | 04:02:55 | 49.49 | | 1224.7085 | | 3.0000e-07 |
| 13 | 2000 | 04:08:58 | 58.71 | | 1723.1716 | | 3.0000e-07 |
| 13 | 2015 | 04:11:21 | 45.70 | 47.95 | 1044.1409 | 1149.7344 | 3.0000e-07 |
| 14 | 2050 | 04:15:36 | 53.01 | | 1404.9589 | | 3.0000e-07 |
| 14 | 2100 | 04:21:40 | 45.47 | | 1033.6006 | | 3.0000e-07 |
| 14 | 2150 | 04:27:43 | 48.41 | | 1171.8534 | | 3.0000e-07 |
| 14 | 2170 | 04:30:43 | 45.63 | 47.88 | 1041.1510 | 1146.1754 | 3.0000e-07 |
| 15 | 2200 | 04:34:21 | 46.98 | | 1103.5192 | | 3.0000e-07 |
| 15 | 2250 | 04:40:22 | 39.54 | | 781.7034 | | 3.0000e-07 |
| 15 | 2300 | 04:46:27 | 48.65 | | 1183.5387 | | 3.0000e-07 |
| 15 | 2325 | 04:50:01 | 45.54 | 47.80 | 1036.8938 | 1142.4320 | 3.0000e-07 |
| 16 | 2350 | 04:53:04 | 46.73 | | 1091.9877 | | 3.0000e-07 |
| 16 | 2400 | 04:59:07 | 47.22 | | 1114.6563 | | 3.0000e-07 |
| 16 | 2450 | 05:05:09 | 52.27 | | 1365.8966 | | 3.0000e-07 |
| 16 | 2480 | 05:09:21 | 45.40 | 47.73 | 1030.6560 | 1138.9236 | 3.0000e-07 |

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| 17 | 2500 | 05:11:48 | 47.76 | | 1140.4403 | | 3.0000e-07 |
| 17 | 2550 | 05:17:50 | 45.37 | | 1029.3501 | | 3.0000e-07 |
| 17 | 2600 | 05:23:52 | 46.66 | | 1088.3760 | | 3.0000e-07 |
| 17 | 2635 | 05:28:41 | 45.27 | 47.65 | 1024.8290 | 1135.4761 | 3.0000e-07 |
| 18 | 2650 | 05:30:31 | 42.42 | | 899.6440 | | 3.0000e-07 |
| 18 | 2700 | 05:36:34 | 41.10 | | 844.4760 | | 3.0000e-07 |
| 18 | 2750 | 05:42:36 | 55.18 | | 1522.3647 | | 3.0000e-07 |
| 18 | 2790 | 05:48:00 | 45.18 | 47.55 | 1020.6840 | 1130.5330 | 3.0000e-07 |
| 19 | 2800 | 05:49:13 | 39.50 | | 780.0601 | | 3.0000e-07 |
| 19 | 2850 | 05:55:17 | 45.95 | | 1055.8196 | | 3.0000e-07 |
| 19 | 2900 | 06:01:18 | 48.27 | | 1165.1357 | | 3.0000e-07 |
| 19 | 2945 | 06:07:18 | 45.01 | 47.46 | 1013.1122 | 1126.1296 | 3.0000e-07 |
| 20 | 2950 | 06:07:56 | 44.21 | | 977.4769 | | 3.0000e-07 |
| 20 | 3000 | 06:14:00 | 43.51 | | 946.6315 | | 3.0000e-07 |
| 20 | 3050 | 06:20:03 | 61.29 | | 1878.5277 | | 3.0000e-07 |
| 20 | 3100 | 06:26:40 | 44.89 | 47.40 | 1007.7694 | 1123.5553 | 3.0000e-07 |
| 21 | 3150 | 06:32:43 | 48.00 | | 1151.9200 | | 3.0000e-07 |
| 21 | 3200 | 06:38:46 | 47.08 | | 1108.3530 | | 3.0000e-07 |
| 21 | 3250 | 06:44:48 | 49.03 | | 1201.8611 | | 3.0000e-07 |
| 21 | 3255 | 06:45:58 | 44.78 | 47.30 | 1002.7260 | 1118.5311 | 3.0000e-07 |
| 22 | 3300 | 06:51:26 | 40.95 | | 838.4462 | | 3.0000e-07 |
| 22 | 3350 | 06:57:30 | 42.18 | | 889.6908 | | 3.0000e-07 |
| 22 | 3400 | 07:03:33 | 43.73 | | 956.3016 | | 3.0000e-07 |
| 22 | 3410 | 07:05:19 | 44.74 | 47.26 | 1000.9611 | 1116.7350 | 3.0000e-07 |
| 23 | 3450 | 07:10:11 | 41.66 | | 867.9470 | | 3.0000e-07 |
| 23 | 3500 | 07:16:17 | 47.19 | | 1113.4923 | | 3.0000e-07 |
| 23 | 3550 | 07:22:20 | 57.33 | | 1643.5494 | | 3.0000e-07 |
| 23 | 3565 | 07:24:42 | 44.64 | 47.21 | 996.2479 | 1114.2549 | 3.0000e-07 |
| 24 | 3600 | 07:28:57 | 52.36 | | 1371.0277 | | 3.0000e-07 |
| 24 | 3650 | 07:35:03 | 43.71 | | 955.2274 | | 3.0000e-07 |
| 24 | 3700 | 07:41:07 | 46.70 | | 1090.3457 | | 3.0000e-07 |
| 24 | 3720 | 07:44:05 | 44.45 | 47.06 | 987.8143 | 1107.2830 | 3.0000e-07 |
| 25 | 3750 | 07:47:45 | 45.64 | | 1041.5979 | | 3.0000e-07 |
| 25 | 3800 | 07:53:50 | 39.39 | | 775.6584 | | 3.0000e-07 |
| 25 | 3850 | 07:59:51 | 47.32 | | 1119.5339 | | 3.0000e-07 |
| 25 | 3875 | 08:03:26 | 44.34 | 47.00 | 982.9999 | 1104.6479 | 3.0000e-07 |