Assignment 3

CS5370: Deep Learning for Vision/AI5100: Deep Learning/AI2100: Deep Learning
IIT-Hyderabad
Jan-Apr 2021

Max Marks: 50
Due: 11th Mar 2021 11:59 pm

Instructions

- Please use Google Classroom to upload your submission by the deadline mentioned above. Your submission should comprise of a single ZIP file, named `<Your_Roll_No>_Assign3`, with all your solutions, including code.
- For late submissions, 10% is deducted for each day (including weekend) late after an assignment is due. Note that each student begins the course with 10 grace days for late submission of assignments, of which upto 4 grace days can be used for a single assignment. Late submissions will automatically use your grace days balance, if you have any left. You can see your balance on the Marks and Grace Days document, soon to be shared under the course Google drive.
- You have to use PYTHON for the programming questions.
- Please read the department plagiarism policy. Do not engage in any form of cheating - strict penalties will be imposed for both givers and takers. Please talk to instructor or TA if you have concerns.

1 Theory (15 marks)

You can submit your response as a PDF document, which can be typed out in LaTeX/Word, or handwritten and scanned. If handwritten, please ensure legibility of answers.

1. *(2 marks)* Briefly give the math for feedforward and backprop steps through a single residual block in ResNets (no need for the full network; you can make necessary assumptions accordingly).
2. *(2 marks)* A convolutional neural network has 4 consecutive 3×3 convolutional layers with stride 1 and no pooling. How large is the support of (the set of image pixels which activate) a neuron in the 4th non-image layer of this network?
3. *(2 marks)* Consider a fully-connected deep network with one hidden layer. If the number of hidden units is increased, what effect would this have on bias and variance of the learned model? Explain briefly.
4. *(3 marks)* Consider a two-layer neural network as given below:

$$y_k(\mathbf{x}, \mathbf{w}) = \sigma\left(\sum_{j=1}^{M} w_{kj}^{(2)} h\left(\sum_{i=1}^{D} w_{ji}^{(1)} x_i + w_{j0}^{(1)} \right) + w_{k0}^{(2)} \right)$$

where h is a layer with a linear activation function, and $\sigma(a) = \{1 + \exp(-a)\}^{-1}$ is the sigmoid activation function. Your classmate tells that she can derive an equivalent network which computes the same function, but with the final layer activation functions replaced with $\texttt{tanh}(a)(= \frac{e^a - e^{-a}}{e^a + e^{-a}})$. Is she correct? Can you prove and support her argument? *(Hint: Think about relating $\sigma(a)$ and $\texttt{tanh}(a)$, and then showing that the two networks' parameters differ only by linear transformations.)*

5. *(3 marks)* Consider a quadratic error defined by $E(\mathbf{w}) \approx E(\mathbf{w}^*) + \frac{1}{2}(\mathbf{w} - \mathbf{w}^*)^T \mathbf{H}(\mathbf{w} - \mathbf{w}^*)$, in which the Hessian matrix $\mathbf{H}$, evaluated at $\mathbf{w}^*$ has an eigenvalue equation given by $\mathbf{H}\mathbf{u}_i = \lambda_i \mathbf{u}_i$. Show that the contours of constant error are ellipses whose axes are aligned with the eigenvectors $\mathbf{u}_i$, with lengths that are inversely proportional to the square root of the corresponding eigenvalues λ_i. *(Hint: Note that it is possible to write $\boldsymbol{w} - \boldsymbol{w}^*$ as $\sum_i \alpha_i \boldsymbol{u}_i$ for some α_is (you can assume this, if you like!). How does this help you?)*

6. *(3 marks)* As part of improving customer service, Kaziranga National Park is designing a mobile application where a visitor can take a photo of an animal, and the application identifies the same. They have collected 20 images of all 200 species of animals in the park. They have access to a deep learning model deployed at Olympic National Park, Washington trained on 1 million images from 1000 classes. Can you help the digitization team at Kaziranga National Park to develop their deep learning model? What exact steps would you suggest?

2 Programming (35 marks)

- The programming questions are shared in "Assignment_3.zip". Please follow the instructions in the notebook. Turn-in the notebook via Google Classroom once you finish your work.
- Marks breakdown is as follows:
  - Question 1: 17 marks (15 marks for code completion and 2 marks for reporting train & test loss, final accuracies and the plots as requested.)
  - Question 2: 4 marks
  - Question 3: 6 marks
  - Question 4: 8 marks