Test Person 2022-05-12

Email: testperson@example.com
Class: E1a
Teacher: Professor Teacher

DLAI (18 February 2022)

Number of words: 507

Question 1 (7 points)

Consider an uncharacteristic architecture which takes as input vectors $\mathbf{x} \in R^n$ and is trained through

optimizing a *vector-valued* loss $\mathcal{L}(\mathbf{x}) \in R^m$ with $m \gg n$.

Assume moreover that you observe particularly slow training time during the optimization.
How would you modify the backpropagation procedure to speed it up?
Please explain your reasoning.

Question 2 (7 points)

Assume you want to adversarially attack an image classifier. Please write down the iterations that are used to find the adversarial attack, explaining them in detail.
Assume now that you prefer having the distortion concentrated on very few pixels rather than many. How do you enforce this constraint?

Question 3 (6 points)

Consider a single layer of a graph neural network.
You want to choose a neighborhood aggregation scheme to derive the representation of each node as a function of its neighborhood.
Intuitively, an expressive aggregation scheme should be able to associate different neighborhoods to different values to avoid losing information at each aggregation step.

Assuming input node features E to be one-hot encodings, i.e. $E \in \{0, 1\}^k$ and $\sum_i E_i = 1$, which

aggregation scheme would you choose among mean, max and sum?
Please explain your reasoning.

Question 4 (5 points)

When operating with very little data, supervised samples are not sufficient to train a typical deep architecture. In this case, you can either choose to act on the (limited) data, on the model, or on the optimization procedure. Assume you want to classify a set of rare classes (e.g. images of pangolins and wombats) for which supervised data is extremely scarce. You can assume that you have lots of data for other more common classes (e.g. dogs and cats). How would you proceed?

If this wasn't enough, you are asked to design a model that is now expected to detect Bronterocs, an unknown alien species, for which no visual representation of any kind exists. How can you tackle the problem?

Question 5 (3 points)

How many learnable parameters has a 3-layer MLP with bias (input → hidden1 → hidden2 → output) with 40-20-10-3 neurons, with batch normalization in the first hidden layer, and dropout (p=0.7) in the second hidden layer?

Question 6 (6 points)

You want to predict the trajectory of human bodies. To do this, you design a network that given a time istant t

and a point $\mathbf{x}$predicts its next position $\mathbf{x}_{t+1}$. For some reason, you observe that your network fails to predict

when humans stand still, even on the training set.

Your network is composed of a set of blocks you implemented yourself, each composed of a convolutional layer

followed by a ReLU activaction $f\left(\mathbf{x}\right) = \max\left(0, \mathbf{x}\right)$, dropout and batch normalization

$$\mathbf{x} \mapsto \frac{\mathbf{x} - \mathbb{E}_{\mathbf{x} \in \mathcal{X}}\left[\mathbf{x}\right]}{\sqrt{\mathrm{var}_{\mathbf{x} \in \mathcal{X}}\left[\mathbf{x}\right]}}.$$

You start to suspect you may have introduced a bug in your implementation. What may be causing the problem? Please explain your reasoning.

Question 7 (6 points)

Assume you are training a deep architecture composed of five convolutional layers, each followed by a ReLU, batch norm, dropout and mean pooling. Assume moreover that you observe that some neurons are always giving 0 as output independently of the input.

What could be happening? How would you mitigate the issue?