Test Person 2022-05-12

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

DLAI (13 January 2022)

Number of words: 363

Question 1 (3 points)

How many learnable parameters has a 3-layer MLP with bias (input → hidden1 → hidden2 → output) with 10-20-10-30 neurons, with batch normalization in the first hidden layer, and dropout (p=0.7) in the second hidden layer?

Question 2 (8 points)

Consider an *untargeted* adversarial attack on an image. First, write down the iterations that are used to find the adversarial attack, explaining them in detail. Now, imagine you want to use the same technique to adversarially perturb an *undirected graph*. Would it work? If yes, why? If not, how would you modify the technique?

Question 3 (8 points)

Assume you trained a VAE for gray-scale images of human faces. Once the model is trained, taking a linear interpolation between latent codes results in a smooth transformation of a face into another. Now you are given an additional requirement: the interpolation in the latent space should correspond to a smooth interpolation of skin color. How can you achieve this?

Question 4 (4 points)

The encoder-decoder model of an AE is strictly related to the notion of *manifold hypothesis* in machine learning. In what way? Explain the manifold hypothesis, and show how it relates to the deterministic AE model.

Question 5 (5 points)

Assume you have a learning model that can do semantic segmentation of a 2D image; for example, given a street view image, it can correctly identify pedestrians, vehicles, and road signals. What are the main difficulties in applying this model to a video, as opposed to a single image?

Question 6 (4 points)

Mathematically, convolution is a *linear* operation and, therefore, it can be expressed using matrix notation. (1) Write down this expression, explaining in detail all the quantities involved. (2) Explain what "linearity" means mathematically. (3) Write a mathematical formula expressing the fact that convolution is translation-equivariant, and explain it in detail.

Question 7 (8 points)

Consider a simple self-attention layer (i.e. with no trainable parameters). We observe the following phenomenon.

Given an input sequence of length n, where each vector is drawn uniformly from the unit d-dimensional sphere with $d \gg n$, in the output we observe almost exactly the same sequence as the input. Why? Please explain your reasoning.