Test Person 2021-07-13

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

2021-10-19

DLAI (08 July 2021)

Word count: 522

Question 1 (4 points)

How many learnable parameters has a 2-layer CNN with bias and max pooling after each convolution (input → hidden → output), where the input has 100x300 pixels with 20-dimensional features, hidden is composed of 100x300 pixels with 30-dimensional features, and the output has 100x300 pixels and 3-dimensional features? Assume both CNN kernels have size 5x5.

Question 2 (4 points)

What is the manifold hypothesis, and how does it relate to the mechanism underlying an autoencoder?

Question 3 (6 points)

Suppose you want to use Transformers for the following applications:

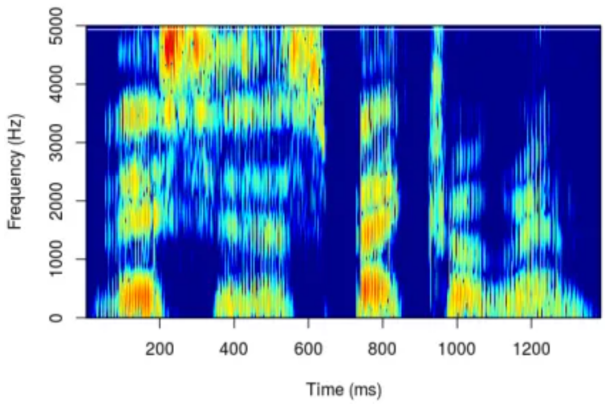
1. Weather forecast
2. Translation of ancient books in the Vatican archive
3. 3D shapes classification
4. Live translation of videocalls
5. Sentiment analysis of text

In which applications would you use a triangular mask for the attention weights at training time? Why?

Question 4 (6 points)

Suppose you trained a deep network comprised of several linear layers, batchnorm, ReLU activations and dropout regularization. You now want your trained model to behave non-deterministically at inference time, i.e., the model should yield slightly different outputs when given the same input multiple times.What is the simplest strategy you can think of to achieve this?

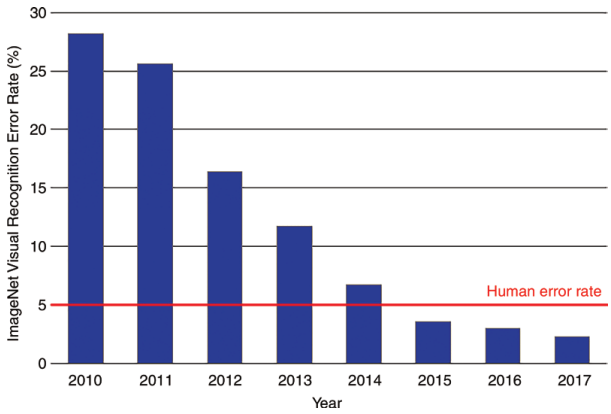
Question 5 (8 points)
In anticipation of the upcoming Olympic games, you have recorded with your band the hymn of your nation. To impress your friends, you would like to generate many versions of your recording, each with a different style: a Rock version, a Jazz version, a Blues version, etc. Now suppose you have a pre-trained neural network with 420,000 parameters, capable of identifying music genres from spectrograms (Classical, Rock, Jazz, Blues...). Your recording at 48kHz can be represented through a spectrogram of size 48000x5000. What optimization problem do you set up to achieve your goal? How many parameters will the optimization algorithm update at each iteration of gradient descent?



Example of a spectrogram.

Question 6 (6 points)
Suppose you want to implement a graph VAE, so that you can generate new graphs by decoding samples from its latent space. How would you phrase the reconstruction loss? Consider two settings: 1) all the graphs in the training set have the same number of nodes and the same number of edges; 2) the graphs in the training set have a different number of nodes and edges. Motivate your answer, and discuss the strengths and weaknesses of your loss definition.

Question 7 (6 points)
You are invited for an interview about the state of AI by a friend of yours running a popular youtube channel.
Your friend: Is it really true that machines are nowadays better than humans at computer vision tasks? I heard this claim in a talk accompanied by this plot:



You: I would like to answer by showing another image. It is fun, I promise:

Your friend: I see your point! What would it take for a computer to understand this image as you or I do?
You: [Write explicitly all the pieces of knowledge that have to fall in place for the image above to make sense. Then discuss what we can already do with present AI techniques and what we are missing.]

Test Person