CPSC/AMTH 453/553 GENE/CBB 555 Final Exam

Date: 12-15-2021 Total points: 120 Points for 100%: 100 points

Multiple choice (30 points, multiple answers allowed):

1. (5 points) Which of the following are self-supervised tasks for a neural network?
    a. Reconstruction
    b. Discrimination (as in a GAN)
    c. Classification
    d. Regression
    e. Context prediction (for a word or object in an image)
2. (5 points) Which of the following is a solution for mode collapse in a GAN?
    a. Early stopping
    b. Training the generator more than the discriminator
    c. A batch-level statistical penalty
    d. Weight clipping in the discriminator
    e. Non-sigmoidal final layer
3. (5 points) How are vanishing gradients combatted in deep neural networks?
    a. Sparser connectivity between layers
    b. Skip connections
    c. Non-saturating activation functions
    d. Pooling
    e. Non-saturating loss functions
4. (5 points) What are some advantages of PCA over non-linear dimensionality reduction methods?
    a. The ability to relate points in PC space back to original features
    b. Interpretability of the axes
    c. Ability to separate clusters
    d. Ability to denoise data
    e. Computational efficiency
5. (5 points) Kernel PCA uses the Kernel Trick to turn which linear method into a non-linear method?
    a. PCA
    b. Classic MDS
    c. Diffusion maps
    d. tSNE
    e. UMAP
6. (5 points) When is the dimensionality reduction in an autoencoder similar to PCA?
    a. With linear activations
    b. With PCA initialization
    c. When trained to minimize variance
    d. With sigmoidal activations
    e. With a 2-D bottleneck

Short answer (40 points):

1. (10 points) For each of the following tasks, describe an appropriate dimensionality reduction algorithm and justify your decision.
    a. Clustering
    b. Visualization
    c. Detecting trajectories in data
    d. Understanding directions of high variation
    e. Denoising

2. (10 points) For each of the following tasks, describe an appropriate neural network architecture and justify your decision.
    a. Visualizing high-dimensional data
    b. Generating random images of cats
    c. Predicting the next element of a sequence
    d. Converting images of cats to “equivalent” images of dogs
    e. Generating poetry

3. (10 points) You hired an awesome photographer to take pictures at the launch party of your new start-up, HypertronAI. Unfortunately someone drinking too much kool-aid runs into the photographer, spilling their drink on the camera lens and leaving all your pictures a slight shade of pink. Luckily your new intern Alex claims to be a deep learning expert who can correct this mistake. How could Alex design/train a neural network to do this?

4. (10 points) Suppose you found out that political affiliation is a low frequency feature on a friend-graph (i.e., a graph with people as vertices and edges between friends) in the friendbook social media app. How would you use this information to impute political affiliation where the entry is missing from a profile with a graph neural network? What kind of aggregation would you use?

Long answer (30 points):

1. (15 points) Suppose you come across some books in a language you have never heard of before and want to separate out the basic parts of speech. How can you train a neural network to do this? What would be the architecture, loss function and training procedure? Suppose someone then told you the word for “banana” in the language. How do you find other likely fruits?

2. (15 points) Suppose you had access to a dataset of graphs that represented molecular structure. How would you generate another molecule-like graph? What would be the architecture of your network and what would be the loss function? How would you train the network?

Bonus question (20 points):

How would you train a neural network to have an embedding layer that is the same as tSNE? Specify the architecture, loss function and training.