Homework 3: Support Vector Machines and Neural Networks

Due 23:59 on Monday 31 October 2022

You will do this assignment individually and submit your answers via the Blackboard course website.

1. Margins [15pts]

Consider the hyperplane given by $\mathbf{w}^T\mathbf{x} + b = 0$. For an arbitrary data point $\mathbf{x}$, what is the distance between $\mathbf{x}$ and the hyperplane, in terms of $\mathbf{w}$ and b?

2. SVMs [15pts]

Suppose that we train two SVM's, the first containing all of the sample data points and the second trained on a dataset constructed by removing some of the support vectors from the first set. How does the size of the optimal margin change between the first and second SVM?

3. PyTorch [4pts]

Work through the "Introduction to PyTorch - YouTube Series" which can be found here: https://pytorch.org/tutorials/beginner/introyt.html. In particular, work through sections 1 (Introduction to PyTorch), 2 (Introduction to PyTorch Tensors), 3 (The Fundamentals of Autograd), 4 (Building Models with PyTorch), 5 (PyTorch TensorBoard Support), and 6 (Training with PyTorch). Each section of the tutorial has a video and code.

You do not need to submit anything for this problem other than any questions you may have on the tutorial materials. I will try to answer them in class after questions are aggregated.

4. More PyTorch [15pts]

Use scikit-learn to load the breast cancer Wisconsin dataset: https://scikit-learn.org/stable/modules/generated/sklearn.datasets.load_breast_cancer.html

Convert the data into a PyTorch TensorDataset, and then split it into 80% for training and 20% for testing. Try to build your own classification neural network for this data. Play around with different numbers of hidden layers and layer sizes, as well as different optimization learning rates. Submit your code for this problem as well as a write up of what you tried and the results.

5. Calibration [1pt]

Approximately how long did this homework take you to complete?