CSCE822 Homework 3

Problem 1: Describe how to do evaluation using hold-out and cross-validation (CV) methods. Discuss in what situation one should use hold-out or CV. Discuss the advantages of CV evaluation method.

Problem2:

Use 2 figure to illustrate the 3 key ideas of Support vector machines (SVM) algorithms: maximum margin classifier, map to high-dimension. Also discuss the kernel trick and explain why the SVM can be used for DNA sequence or graph/network classification using the kernel trick.

Problem 3: ensemble algorithms

3.1 Give 3 examples of famous ensemble machine learning algorithms

3.2 Explain why ensemble algorithms are usually much more powerful than standalone algorithms

3.3 Suppose a standalone algorithm A has accuracy of 70% and we built an ensemble algorithm out of 7 different instances of algorithm A. Calculate the accuracy of the ensemble algorithm (hint: check the slides)

3.4 What is the difference of bagging vs. boosting?

Problem 4: Dealing with unbalanced dataset for classification and regression

Read the following blog on how to deal with unbalanced dataset issue. And write a summary on 1) How does unbalanced dataset cause issue in our classifier performance evaluation?

2) What options should you have to deal with classification over unbalanced dataset? Discuss the advantage of each approach. (use Table).

- https://towardsdatascience.com/methods-for-dealing-with-imbalanced-data-5b761be45a18
- https://medium.com/james-blogs/handling-imbalanced-data-in-classification-problems-7de598c1059f

- https://elitedatascience.com/imbalanced-classes
- https://www.kdnuggets.com/2017/06/7-techniques-handle-imbalanced-data.html

Problem 5:

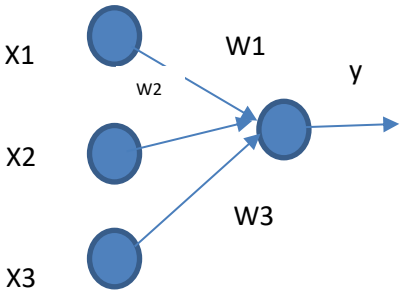
5.1 Traditional multi-layer feedforward perceptron ANN has a notorious difficulty to train for models with multiple hidden layers. Describe how the recent deep learning algorithms solve this problem.

5.2 backpropagation

Consider a 1-layer neural net with three input units, 1 output unit, no hidden units and no bias terms. Suppose that the output unit uses a sigmoid activation function, *i.e.*, $y = 1/(1+e^{-z})$, where z is the total input to the unit. Let y be the computed output of the neural net, let d be the desired output, and let $C = -d\log y - (1-d)\log(1-y)$ be the cross entropy error. Write down the equations for a single step of weight updates by gradient descent (based on a single data sample), and derive all the necessary derivatives. Simplify your answers, and be sure to clearly identify all the variables you use.

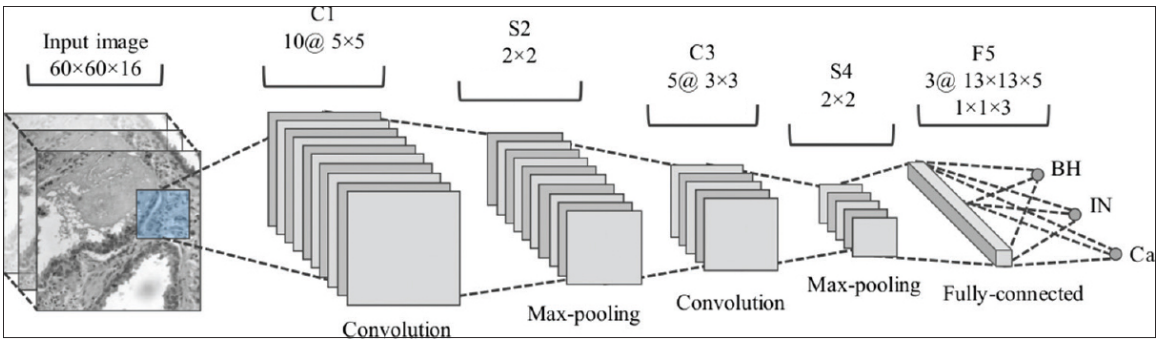
Hint: use the chain rule and recall the following results:

$$\frac{\partial y}{\partial z} = y(1-y) \qquad \frac{\partial \log u}{\partial u} = \frac{1}{u}$$



Only need to show $\frac{dE}{dw_1}$

5.3 Figure below shows a deep convolutional neural networks with a couple of convolution/max-pooling/and fully connected layers. The input is 16-channel 60x60 images. The convolution layer filter is represented by n@a x a where a is the number of filters and axa is the filter size. Max-pooling operator is represented by b x b e.g. 2x2. Calculate the feature map dimension of each of the layers: C1, S2, C3, S4.



Problem 6:

Describe the main difference of a recurrent neural network compared to a fully-connected feedforward neural networks. What types of problems are most suitable for recurrent neural networks?

Describe the main idea of auto-encoder neural network models. What types of problems Autoencoders are suitable for?

Submit all your code and reports to:

http://dropbox.cse.sc.edu