BMI-6015-Fall 2018
Assignment- Neural Network Training and Hyper-parameter Selections
Release Date: 12-03-2018
Due Date: 12-13-2018 (11:59 pm)
Marks: 100 points

The purpose of this assignment is to familiarize ourselves with convolutional neural networks CNN (basics, calculations, and implementation) and to experiment, using Keras, fully connected neural network training and hyper-parameter selection. In this assignment, you may use any python, scikit-learn, and Tensorflow/Keras libraries per question instructions. Any student collaboration and "copy & paste" from any other resource are totally prohibited. Also, for any question or inquiry, you may want to email the professor and the course TA. To submit the assignment, you will receive instructions from the professor for where you should upload your solutions. In this assignment, you may do part III only either alone or in a group of two students (if you wish); otherwise, you will be required to answer the other parts by yourself only. The course TA may only help you in any debugging issue. Your solution should be provided using Jupyter notebook and document file for Part I (if you wish).

Good Luck!
Samir Abdelrahman

A. Part I (30 points): CNN arithmetic calculation.

Requirements: You will be required to understand CNN convolution arithmetic and pooling calculations (11-12-2018 lecture & https://arxiv.org/pdf/1603.07285.pdf).

Question: Given the stride (s), the number of zero-padding (p), answer the following questions:

(a) If the input matrix = $\begin{pmatrix} 1 & 4 & 2 & 8 & 7 \\ 3 & 4 & 5 & 1 & 2 \\ 0 & 8 & 9 & 6 & 4 \\ 4 & 3 & 8 & 2 & 6 \\ 6 & 2 & 1 & 1 & 5 \end{pmatrix}$ and the convolution kernel is $\begin{pmatrix} 1 & 0 & 1 \\ 1 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$, calculate the size and the content of output matrix for each of the following case:

a.1) s=1 and p=0.
a.2) s=1 and p=1.
a.3) s=1 and p=2.
a.4) s=2 and p=2.

(b) If the input matrix = $\begin{pmatrix} 1 & 4 & 2 & 7 \\ 3 & 4 & 2 & 1 \\ 3 & 9 & 6 & 4 \\ 3 & 5 & 2 & 6 \end{pmatrix}$ and s=1, apply 3x3 average-pooling, show the calculations and the contents of the output matrix.

B. Part II (30 points): Multi-head Keras CNN development for multi-channel modeling:

Requirements: You will be required to study, understand, and modify the last notebook cell in 11-26-2018/ (multi-step-forecasting notebook).

Question: It is required to use the AU.csv data in the 11-26-2018/ folder to implement the following tasks:

(1) Applying all data pre-processing steps including loading, splitting, encoding, and scaling; when necessary.

(2) Developing your model using Keras CNN models libraries. You need to modify the last notebook cell in 11-26-2018/ (multi-step-forecasting notebook) to apply the model on AU.csv data.

(3) Calculating the total numbers of parameters of any sub-model or feature model. Use (https://stackoverflow.com/questions/42786717/how-to-calculate-the-number-of-parameters-for-convolutional-neural-network/42787467) to guide your calculation.

(4) Testing if the overall model is over-fitted.

C. Part III (40 points): Research question for improving neural network performance and hyper-parameter selection:

Requirements: You may want to solve it alone or with one of your course colleague only (cite your name(s) and any web link you have used in your answer). You need to use Fashion-MNIST dataset in Keras to answer the following questions:

Questions:

(a) Apply three methods (excluding L2 regularization) to improve neural network performance as shown in NN_Strategies_Examples notebook in the assignment 2 folder independently (i.e., each one alone) or with the others (i.e., all of the three methods together). Describe the concept of each method. Analyze the results for each model.

(b) Use different hyper-parameter settings and activation functions to find the most accurate model performance of the generated models. To find the most accurate one, you need to use either:

b.1) Grid-Search Keras/Sklearn wrapper methods that were described in the lecture 11-26-2018.

Or

b.2) Any tool that is described in (https://medium.com/@mikkokotila/a-comprehensive-list-of-hyperparameter-optimization-tuning-solutions-88e067f19d9)