Computer Science Department
CS675 – Introduction to Data Science (CRN: 73453)
Fall 2020

Project #2

Implement a Time Series Forecasting model in Python, by using the **TensorFlow/Keras** modules.

The forecasting model should be able to predict the Sunspots (see below) by using Neural Networks (NNs), specifically Recurrent Neural Networks (RNNs) and/or Long Short Term Memory (LSTMs).

Get the data from **Kaggle's** Repository:
https://www.kaggle.com/robervalt/sunspots

Here is a sample of the dataset (out of 3,234 records):

	Date	Monthly Mean Total Sunspot Number
0	1749-01-31	96.7
1	1749-02-28	104.3
2	1749-03-31	116.7
3	1749-04-30	92.8
4	1749-05-31	141.7
5	1749-06-30	139.2
6	1749-07-31	158
7	1749-08-31	110.5
8	1749-09-30	126.5
9	1749-10-31	125.8
10	1749-11-30	264.3

The dataset contains Monthly Mean Total Sunspot Number – from January 1749 to July 2018.

Sunspots are temporary phenomena on the Sun's photosphere that appear as spots darker than the surrounding areas. They are regions of reduced surface temperature caused by concentrations of magnetic field flux that inhibit convection. Sunspots usually appear in pairs of opposite magnetic polarity. Their number varies according to the approximately 11-year solar cycle.

Source: https://en.wikipedia.org/wiki/Sunspot

Write **Python** scripts in order to complete the following tasks along with their output. All work should be done and submitted in a single **Jupyter Notebook.**

1) Predict the Sunspots values from August 2018 to December 2019.
   << Output >>:
   - Build a NN Model with multiple (Keras) layers.
   - Tune Hyper-parameters (especially LR = Learning Rate)
   - Use various optimizers (Adam, Huber, …)
2) Evaluate all models by providing their respective MAE (Mean Absolute Error)