

CS594, Python Programming Lab

(<https://www.iitg.ac.in/asahu/cs594/>)

Assignment III : File Reading and List Manipulation

Deadline : 11.55 PM IST, 11th October 2020

- **Part (a)** : Given BSE Sensex data from 1st January 2015 to till today in CSV (comma separated values) file format. You need to read the data line by line, when reading a line, you are supposed to predict the next value using (a) average values of the last five days, (b) the EWMA method. The current line and the predicted values need to be written to another file line by line in CSV format.
 - Average values of the last five days: The prediction for the next value is done based on the average values of the last five days. If the value of the last five days is not available then it can be assumed to be zero.
 - EWMA method: The Exponentially Weighted Moving Average (EWMA) is a statistic for monitoring the process that averages the data in a way that gives less and less weight to data as they are further removed in time.

$$\mathbf{PRED}_t = \lambda \mathbf{OBS}_t + (1-\lambda) \mathbf{PRED}_{t-1} \quad \text{for } t=1, 2, \dots, n.$$

$0 < \lambda \leq 1$ is a constant that determines the depth of memory of the EWMA, \mathbf{OBS}_t is the observed value for the day t , \mathbf{PRED}_t is the predicted value for the day t . Initial observation (\mathbf{OBS}_0) and prediction (\mathbf{PRED}_0) may be assumed to be 0, and when at $t=1$, prediction error will be $\mathbf{OBS}_1 - \mathbf{PRED}_1$.

Report the average prediction error of both the prediction methods for the whole duration of calculation. The output CSV file should have two extra columns (PredAVG and PredEWMA) as compared to input CSV file. Also, you need to plot the observed value, predicted values for the whole duration of the calculation.

Test input file: CSVForDate.csv

(Available at <https://www.iitg.ac.in/asahu/cs594/>)

- **Part (b)** : Given an un-directed complete graph represented using adjacency matrix where each entry represents the weight of the edges. You need to write Python code to find a minimum spanning tree (MST) for the complete graph and report the minimum spanning weight and the tree. You are allowed to use any standard MST algorithms (either Prim's or Kruskal's MST or any easy Algorithm). You are allowed to use set, dictionary of the python library.

Test input files : tai100b.dat tai20b.dat (Available at <https://www.iitg.ac.in/asahu/cs594/>)

Submission procedure:

- Send your assignments code in compressed folder (tgx/zip/gz) to asahu < at > iitg < dot > ac < dot > in with "CS594: Assignment<III> , < RollNo > " as subject before the deadline
- Please embed comments, how to run and required inputs properly in the code, or a separate readme file.