Assignment 1

D/L Method

Sangeet M

M.Tech AI, 2021-23

04-04-03-10-51-21-1-19322

Instructions

- The project folder contains 2 folders and this file
- The screenshots showing the final values of the variables Z and L, and of the loss variable are stored inside the **Output** folder
- The code folder has a **pre_process_mod.py** file which performs the pre-processing of data, and is imported into the main program at the time of execution
- The **main.py** uses the pre_process function from the pre_process_mod.py program and gets the data cleaned. The analysis is carried out on the new data and the outputs are shown
- the **Files** folder contains the input csv file

Program

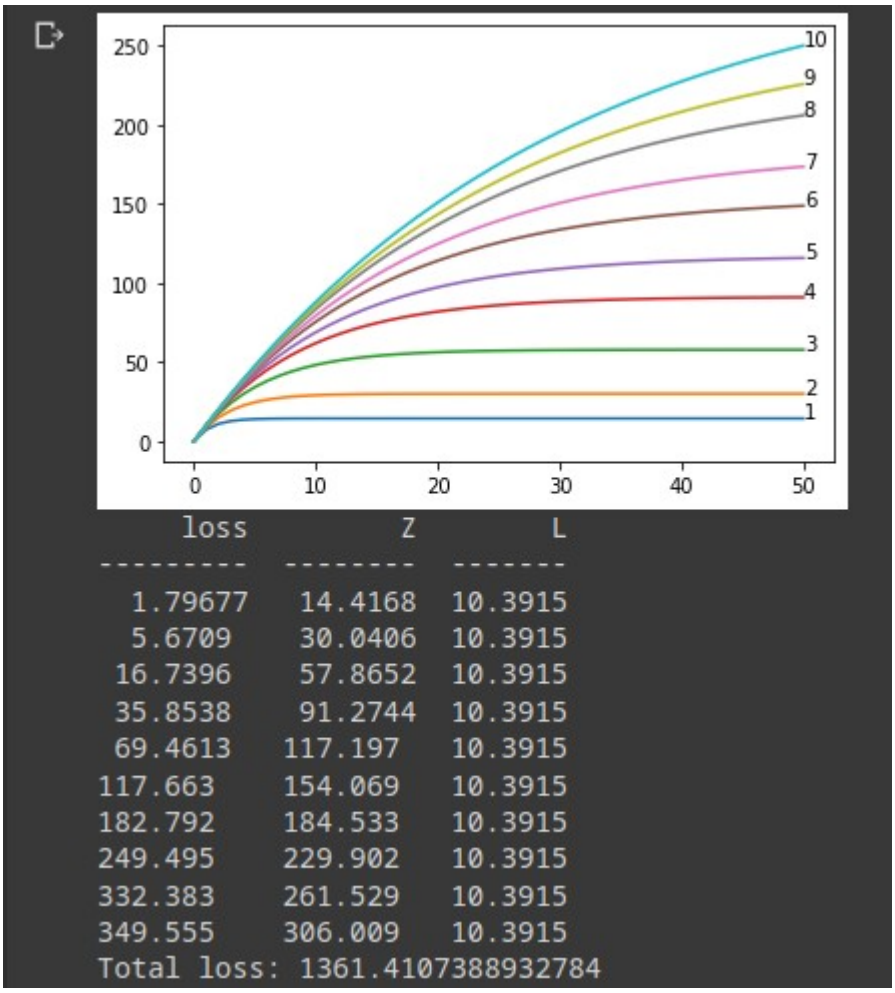
- pre_process:
  - This function takes the full csv file as input
  - the entries with innings as 1 are kept and rest are discarded
  - Then the fields with wrong data, that are of use for the analysis, are cleaned, after confirming the data is completely available. i.e the data are not cut in between the 50 overs, unless the team is fully out
  - Code doing this work is highlighted in the program and can be commented at execution to use the full data

- the cleaned data is cropped and returned to the caller program (Cropping is selecting the required attributes)

- main.py
  - Calls the **pre_process( )** function with the path to the input file and gets the data cleaned
  - the **loss_func( )** function is defined which takes parametres and data as inputs and calculates the mean square error on the given parameters
  - The **duck_lew_mod( )** function finds the required parameter values Z, using **optimize( )** from scipy library
  - The output, parameter values and the graph, are shown using the **disp( )** function

Output

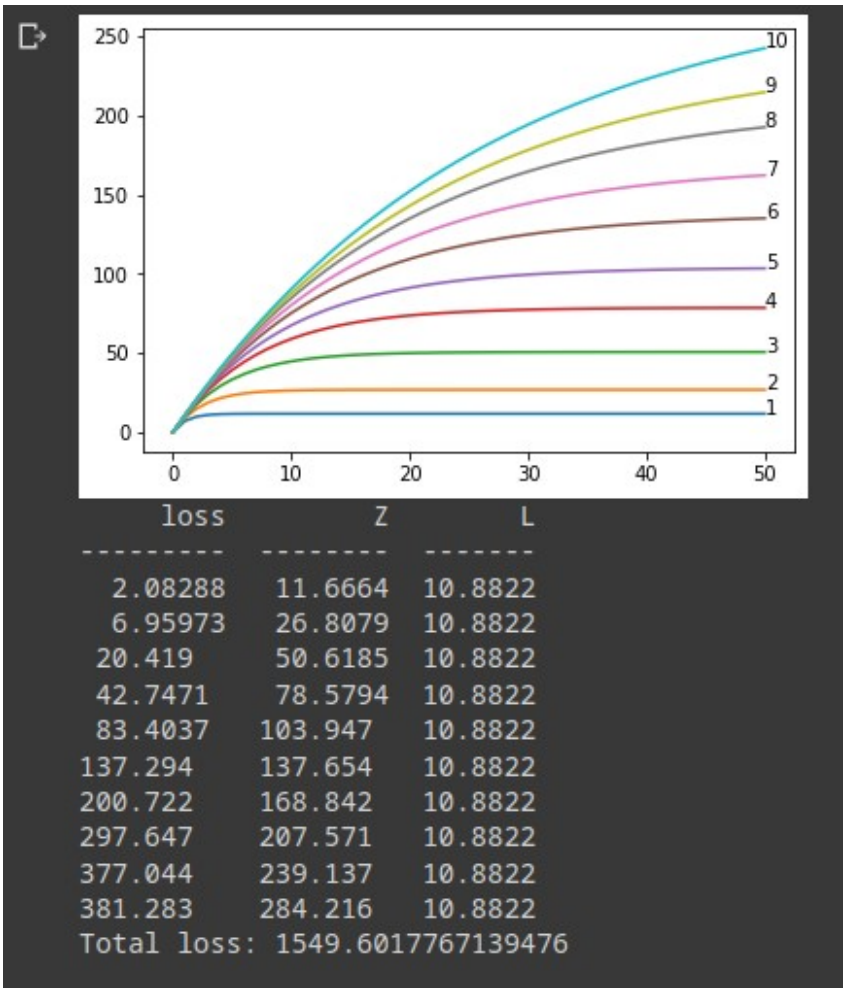
- The output for the analysis done on data without the incomplete entries mentioned in the pre_process part, under the heading: Program



loss	Z	L
1.79677	14.4168	10.3915
5.6709	30.0406	10.3915
16.7396	57.8652	10.3915
35.8538	91.2744	10.3915
69.4613	117.197	10.3915
117.663	154.069	10.3915
182.792	184.533	10.3915
249.495	229.902	10.3915
332.383	261.529	10.3915
349.555	306.009	10.3915

Total loss: 1361.4107388932784

- The ouput from the analysis done on the dataset having the incomplete data (50 overs' data not present and team is not fully out)



loss	Z	L
2.08288	11.6664	10.8822
6.95973	26.8079	10.8822
20.419	50.6185	10.8822
42.7471	78.5794	10.8822
83.4037	103.947	10.8822
137.294	137.654	10.8822
200.722	168.842	10.8822
297.647	207.571	10.8822
377.044	239.137	10.8822
381.283	284.216	10.8822

Total loss: 1549.6017767139476