

Data Analytics Assignment 3

Anshul Shivhare (Sr no. 17828)

October 6 2021

1 Code Architecture

1. Calculated running average for tested and confirmed cases. running average of confirmed cases is used to find $CIR(t)$. Running average $c(t)$ will be used in loss function.
2. A function take input $S(t)$, $E(t)$, $I(t)$, $R(t)$, $B(t)$, $V(t)$, $W(t)$ and return the parameters at $t+1$ (apart from W , B and V). $V(t)$ is No. of First Dose Administered on date t .
3. Note: while taking average of 7 days, i am considering the current day also.
4. Later computing loss. Using running seven day average of confirmed cases and running seven day average of $e(t)$, and them multiplying it by Alpha.
5. $e(t)$ we get from model and $c(t)$ is from data.
6. Now to optimize our parameters, we use gradient descent.
7. After getting the optimized parameters, i am using the model(equations function) to predict for future.
8. Using the optimized parameters, and then updating beta, 3 new data frames are generated.
9. To deal with Nan or negative values, i decreased the learning rate in gradient descent for few parameters.
10. In prediction file, there are still Negative values. This can be handle by updating them close to 0 (not 0). and Later scaling all other parameters in such a way that total will become 70 M.
11. Due to time constraint i didn't plotted the new cases predicted on each day for different beta. but the dataframes is generated in Predicting_data_using_beta file.
12. Optimized value returned after traning is : Beta= 0.4311, CIR_O=15.172, S_t= 0.6362, E_O=0.0021, I_O=0.0016, R_O=0.3601, Loss= 0.0078 , N= 70,000,000
13. The above values are multiplied by N to get correct Population value.