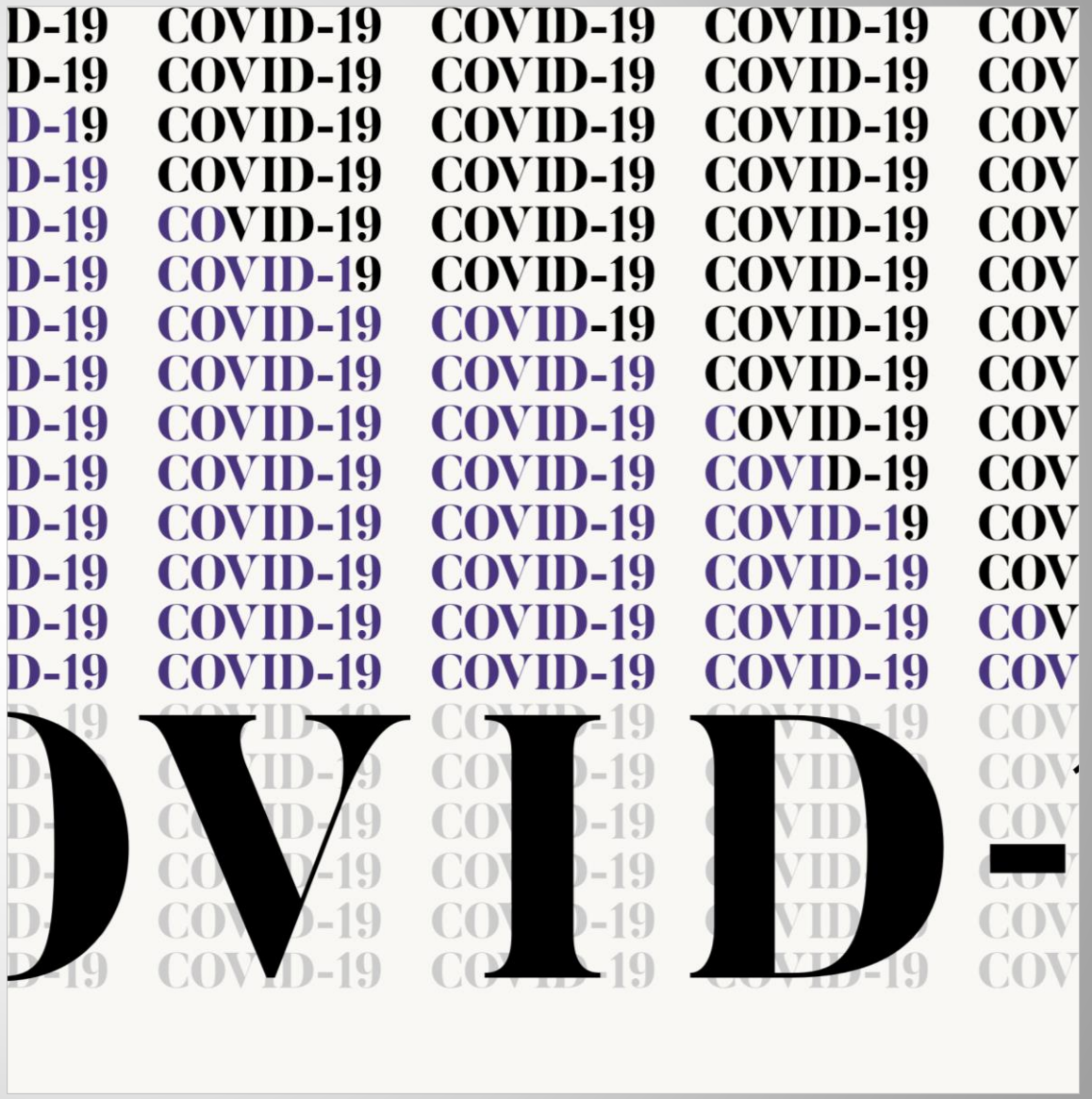
DATA MODELLING AND ANALYZING CORONAVIRUS SPREAD USING DATA SCIENCE

REVIEW – 2
TARP – ECE3999
FACULTY – Dr. ILAVARASAN T

TEAM 11

V SAI YASHWANTH – 17BEC0369
K C S S D CHAKRADHAR – 17BEC0873
M SAI PRANAV – 17BEC0708
B K HEM CHARAN – 17BEC0189
K SURYA TEJA – 17BEC0138
KONG KON J – 17BEC0076



INTRODUCTION

- Presently, there are so many dashboards and statistics around the Coronavirus spread available all over the internet.
- With so much information and expert opinions, to see different nations adopting different strategies, from complete lockdown to social distancing to herd immunity, one is left thinking as to what the right strategy is for them.

IS THERE ANY BASIS TO THESE OPINIONS AND ADVICE?

- So this is an attempt of data modelling and analyzing Coronavirus (COVID-19) spread with the help of machine learning and literature survey.
- This analysis will help us to find the basis behind common notions about the virus spread from purely a dataset perspective.

WORK COMPLETED

- Successfully collected all the Data Sets needed for our Project.
- Came to know what are the data factors that are correlated with each other in one particular dataset.
- Created a Linear Regression model to predict how many people will die based on the people infected.
- Created a multi Linear Regression by doing some feature selection and tested the model with the prediction.

CHANGES CARRIED TO REVIEW – 2

Previously, We created a model that takes only confirmed cases into consideration and predicted the deaths but this is not the right procedure.

Now we created the model by taking all the factors into consideration and predicted the model.

By using the model we can proceed further with the model and also work on with the other datasets and analyze the corona virus spread.

ALGORITHM

- Install Anaconda on system for working on with machine learning.
- In the Anaconda tool use the tool jupyter notebook for working.
- Search for the resources and collect the required datasets for making the model.
- Now first understand what all are the datasets contain and do the required feature selection on the data.
- Create the model and work on increasing the accuracy or decreasing the error.

DATASETS

- time_series_covid19_confirmed_global.csv
- time_series_covid19_deaths_global.csv
- time_series_covid19_recovered_global.csv
- cases_country.csv

REFERENCES

- https://www.emeraldgrouppublishing.com/journal/idd/using-data-science-understand-coronavirus-pandemic
- https://www.ncbi.nlm.nih.gov/pmc/articles/PMC7289234/
- https://in.springboard.com/blog/data-modelling-covid/

THANK YOU