Project Proposal

This repo is one of the T5 Bootcamp requirements.

The correlation between chronic diseases and Covid?

In a short period of time and under epidemiological conditions, many therapeutic methods have been proposed. One of the best ways to discover the causes of the disease is how it is related to chronic diseases.
The study aims to find out the disease most closely related to the Corona virus, by analyzing the mentioned data.

Dataset

To achieve the goal of this study, the COVID-19 dataset will be used together with the evidence of patients with other chronic diseases. This dataset can be found at Kaggle.

This dataset contains tweets for the following feature:

entry_date

date_symptoms

date_died

pneumonia

pregnancy

diabetes

asthma

hypertension

other_disease

cipher

obesity

renal_chronic

tobacco

contact_other_covid

covid_res

Features such as the graphical analysis, which contains linking data to a more clearly visible picture, and the date, which includes death and injury, are used to identify the most important features of this study.

Tools

There are tools that will be used to achieve the goal of this study, such as: `TensorFlow, matplotlib, pandas, nltk` for discovering the data and train a model. The work will be done through Jupyter notebook.
Furthermore, the Sentiment Analysis from `nltk` will be used to determine the target of the data in order to train the model in supervised manner.

TO DO:

- Explore the data and come up with EDA phases then use a model to fit the data.
- **NOTE:** the used features may be increased or changed and the model as well.