Project Plan | Prediction of deaths involves by COVID2019

Contributed by Nikolai Horozov, Jerome Sepin, Vithersan Somasundaram and Gaby Stout

# General topic

We use Long Short-Term Memory (LSTM) Models to forecast new COVID death. Model is trained in data from early COVID infections from Switzerland. The results from these models are tested with data from another Country, namely Germany to check the validity fo the method.

# Motivation

COVID-19 first appeared in the world in December 2019. Since then, it has caused more than thousands of deaths and millions of infections. With this massive number of deaths, it has become one of the biggest crises in this world.

In addition to the loss of human lives, this pandemic has also caused severe damage to the global economy. Due to lockdowns and distancing strategies, it has also had a negative impact on psychological and social spheres.

# Data

The data comes from an open dataset from COVID-19 Data Hub. It is compiled from various sources hourly crunched. Among other things, the data set contains:

* standard COVID-19 variables (total population, cumulative number of cases, tests, deaths, recovered, daily number of cases, . . . )
* policy measures
* geographic information
* external identifiers
* google and apple mobility reports

*Guidotti, E., Ardia, D., (2020), "COVID-19 Data Hub", Journal of Open Source Software 5(51):2376, doi:*[*10.21105/joss.02376*](https://doi.org/10.21105/joss.02376)*.*

## Data processing

* Data Cleaning
* Explorative Data Analysis (EDA)
* Predictive analysis

# Research Question

1. How is the data distributed
2. How accurately is the prediction of COVID19 deaths

# Analysis Techniques / algorithms

Time Series Deep Learning, Forecasting Covid deaths with Keras Stateful LSTM

1) What questions will you try to answer with the project?

2) What data will you be using (and what is the source of that data)?

3) What techniques will you use for the analyses?