HACETTEPE UNIVERSITY
Department of Electrical and Electronics Engineering

ELE 492: Machine Learning
Course Project

Visualisation and Classification of Novel COVID-19 Dataset

Student Name: Baki Emre Bulut
Student ID: 21427779

Goals

- Visualization: Generating ease observable and interpretable outputs for understand process data.
- Classification: Using datasets, obtaining meaningful relations for pandemia between input countries.

Key Results

- Visualizations change graph and cummulative graph of death, recovered, confirmed and active patients in a user interface.
- Classifications countries with stage of epidemic by using slope of change graph of active patients.
- Classification using cummulative confirmed case per population.
- Clasification using difference between peak and start time of infection.

Main Datasets

- **Novel Corona Virus 2019 Dataset**
- Day level information on covid-19 affected cases
- Columns:
- Sno - Serial number
- ObservationDate - Date of the observation in MM/DD/YYYY
- Province/State - Province or state of the observation (Could be empty when missing)
- Country/Region - Country of observation
- Confirmed - Cumulative number of confirmed cases till that date
- Deaths - Cumulative number of deaths till that date
- Recovered - Cumulative number of recovered cases till that date

Secondary Dataset

- **countryinfo**
- Relevant variables that may be required in order to predict COVID's progression.
- This dataset is using to classifying countries depending their populations.

Metadata of Datasets

Main Dataset

- License: Data files © Original Authors
- Visibility: Public
- Dataset owner: SRK (at kaggle.com)
- Source: https://www.kaggle.com/sudalairajkumar/novel-corona-virus-2019-dataset

Secondary Dataset:

- License: Database: Open Database, Contents: Database Contents
- Visibility: Public
- Dataset owner: My Koryto (at kaggle.com)
- Source: https://www.kaggle.com/koryto/countryinfo