Simultaneous Time Series Forecasting on the World's COVID-19 Daily Vaccinations

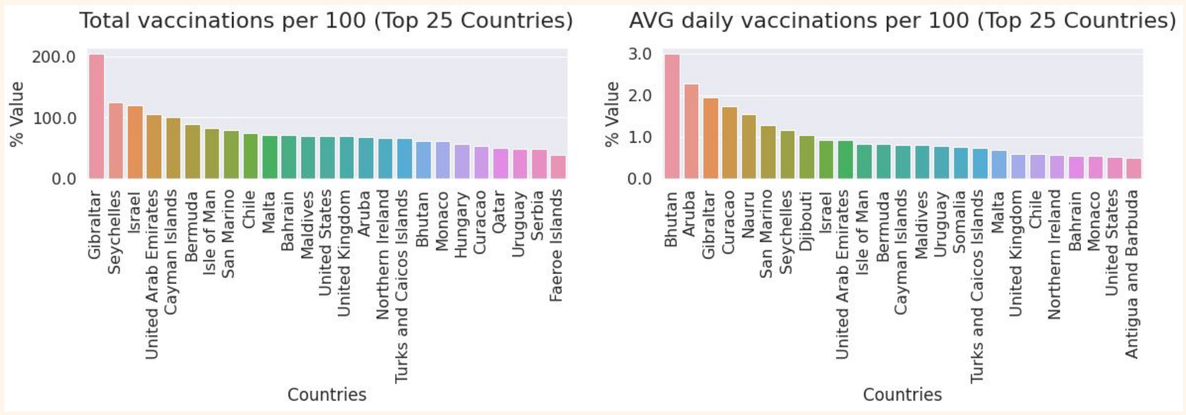
Konstantinos Georgiou

Course Advisor: Michela Taufer

The University of Tennessee, Knoxville

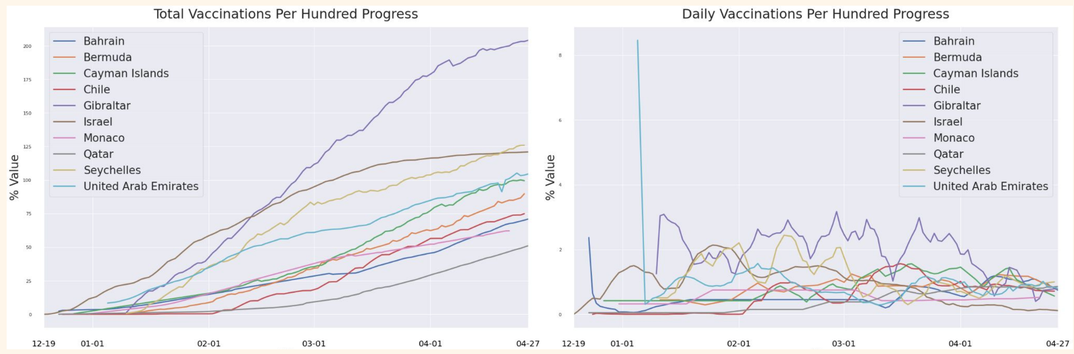
Motivation

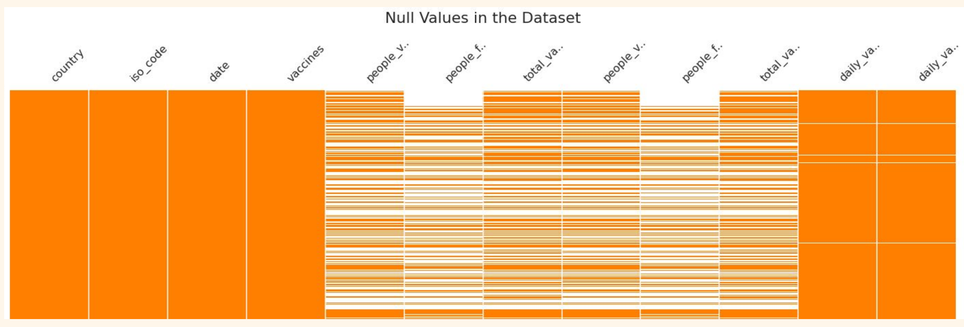
- The distribution of COVID-19 vaccines affects the health of billions of people as well as the state of the world's economies
- Many efforts have been made to extract useful insights from these data, but most of them are comparative analysis between two or more countries
- As of today, no method attempted to simultaneously predict the number of daily vaccinations of all the countries by utilizing the correlations between them
- We introduce a method that uses *Encoder-Decoder Long Short-Term Memory Networks* With *Multivariate Inputs* and *Walk-Forward Validation* of 10 days.



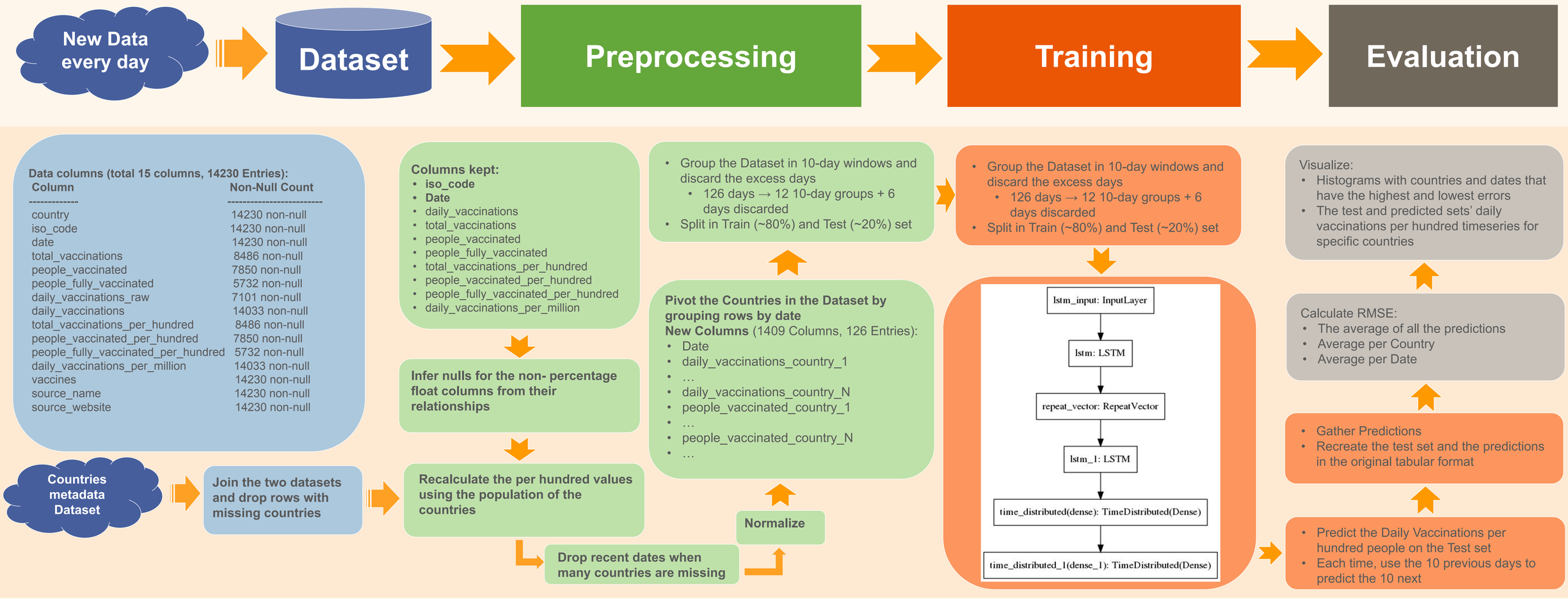
Dataset

- Contains daily vaccination data for 193 different countries and 135 dates
- 14230 15-dimensional data from which 8 dimensions where used
- The dataset has many null values, most of which can be inferred from other values





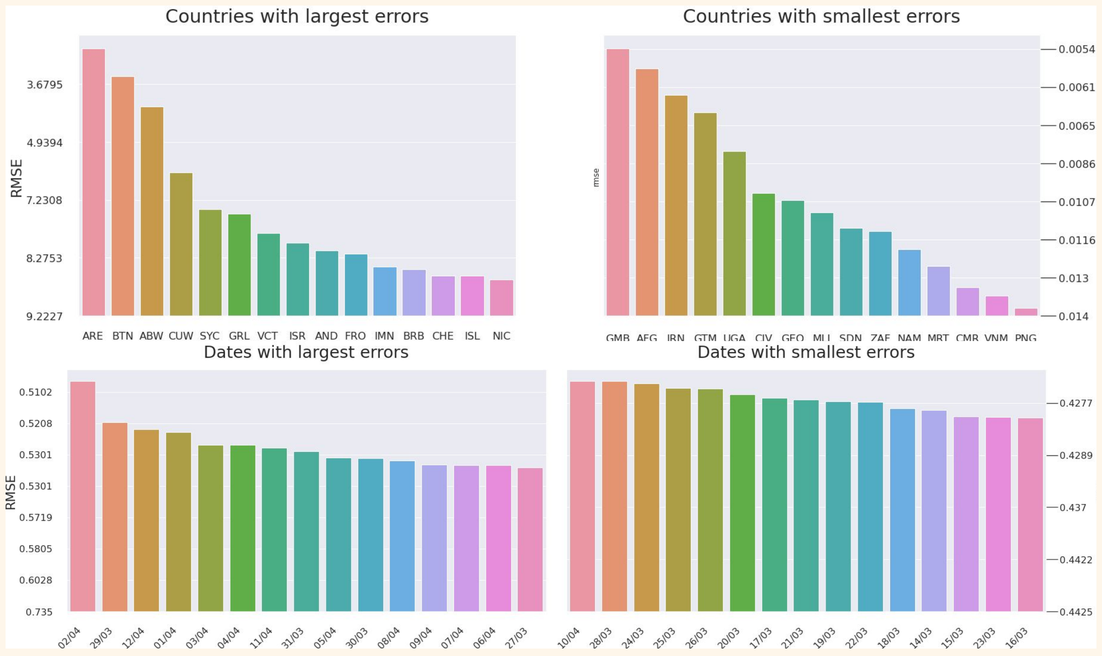
Workflow



Evaluation

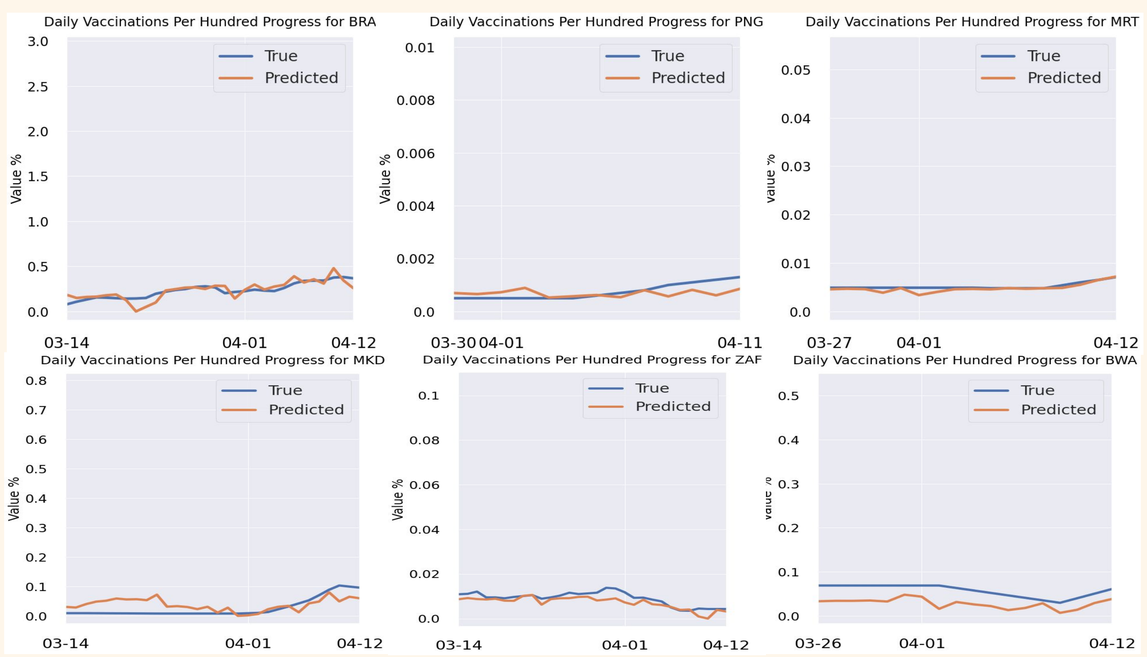
Average RMSE: 0.246174

- Out of the *170* countries, *164* had mean RMSE less than *1.0*, *153* less than *0.5*, and *73* less than *0.1*
  - Min: *0.00023*, Max: *5.5133*
- Out of the *30* dates, *25* had mean RMSE less than *0.30*, *15* less than *0.25*, and *11* less than *0.20*
  - Min: *0.16928*, Max: *0.35180*



Results

True VS Predicted Daily Vaccinations per million for 6 countries



Conclusions

Proved that it's possible to simultaneously predict the number of daily vaccinations of all the countries by finding correlations between their historical data

Future Work

Incorporate static features such as the *Vaccines* used, *Health Expenditure per GDP* and the *Number of Physicians per Million*, all of which are included in the metadata Dataset