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

[Extended Abstract]

Konstantinos Georgiou
The University of Tennessee, Knoxville
Knoxville, TN, United States
kgeorgio@vols.utk.edu

ABSTRACT

The global distribution of the COVID-19 vaccines is one of the most challenging tasks the modern health industry has ever faced. Predicting the number of a country's daily vaccinations can be vital for its protecting its economy and adapting its policy response. We present a pipeline that trains on all countries' historical vaccination data and attempts to predict their next ten days of daily vaccinations by utilizing their underlying relationships. We use Encoder-Decoder Long Short-Term Memory Networks with walk-forward validation and evaluate the results using mean, per-country, and per-date RMSE.

CCS Concepts

•Computing methodologies → Neural networks; •Applied computing → Health care information systems;

Keywords

Multivariate Time Series; Time Series Forecasting; Recurrent Neural Networks; Long-Short Term Memory Networks; COVID -19

1. MOTIVATION

The COVID-19 pandemic is one of the most impactful events of recent human history with profound effects on the health of billions and the economies of almost all of the world's countries. The imperative need for the rapid development and global distribution of the COVID-19 vaccines generated one of the most complex tasks of modern public health history. The urgency and significance of this task have attracted the interest of a large portion of the scientific community as there are many insights that can be obtained from these data.

One insight that can be found very useful, is the prediction of a country's future daily vaccinations. By doing so, we could identify potential future reductions of the rate of some countries' vaccinations, which can give their governments the opportunity to control possible complications.

2. CONTRIBUTIONS

Many attempts have been made to extract useful insights from these data, but most of them are comparative analyses between two or more countries' historical vaccination records, and some of them attempt to predict the number of future vaccinations of one country at a time. We present a method that utilizes all countries' historical vaccination data in order to identify causal relationships between them and make simultaneous predictions on their next ten days of daily vaccinations. We do heavy preprocessing to fix the null values of the dataset, we then use an external dataset to recalculate the per hundred people values of the vaccination columns, we train an Encoder-Decoder Long Short-Term Memory Network, and finally evaluate the predictions using the average, per-country, and per-date RMSE.

The rest of the paper is organized as follows: Section 3 describes the dataset, preprocessing methodology, and evaluation techniques used in this study; Section 4 describes and discusses the results; Section 5 summarizes our findings and proposes future improvements.

3. METHODOLOGY

In this section, we describe the datasets used and the framework we developed to predict the countries' future daily vaccinations.

3.1 Dataset

Text

3.2 Preprocessing

Text

3.3 Training

Text

3.4 Evaluation

Text

4. RESULTS

Text

5. CONCLUSIONS AND FUTURE WORK

Text

6. REFERENCES