Wales
Data Nation
Accelerator

Rare Clinical Event Modelling and Prediction for Covid Patients

Summary

The COVID-19 pandemic has generated a widespread strain on NHS resources that has prompted an urgent need for prognostic tools enabling the practitioners to assess patient outcome effectively and accurately. In the early stages of the pandemic, when there was scarcity of data, mathematical modelling has provided robust scenarios that allowed health board to effectively plan their response. Currently, enough data on patient outcomes have become available, enabling effective deployments of data-driven approaches based on methods of machine learning.

In this project, we implemented a novel demonstrator of clinical outcome predictors based on a synergy between a mathematical model and a data-driven approach, specifically a 1-Dimension Convolutional Network (CNN). The identified predictors are fed into a compartmental model to assess the risks probabilities of the population developing possible outcomes due to COVID-19; therefore, the model has the added capability of identifying those at risk of post-infection complications and rare outcomes.

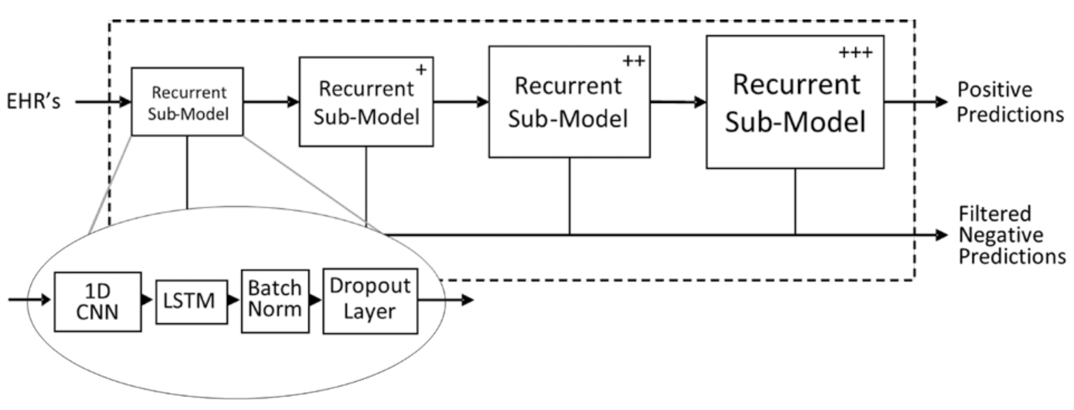

Figure 1: Cascading convolutional deep Learning model, including architecture of each cascade

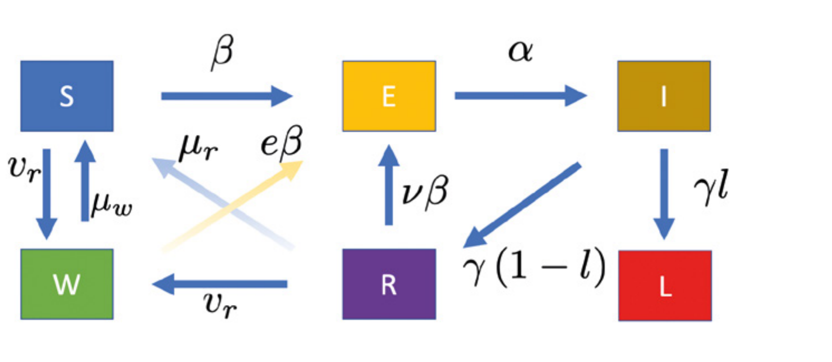

Figure 2: The Mathematical model, including the stages of the infections (Susceptible, Exposed, Infected and Removed), the vaccination rate W and the rate of rare clinical outcomes L

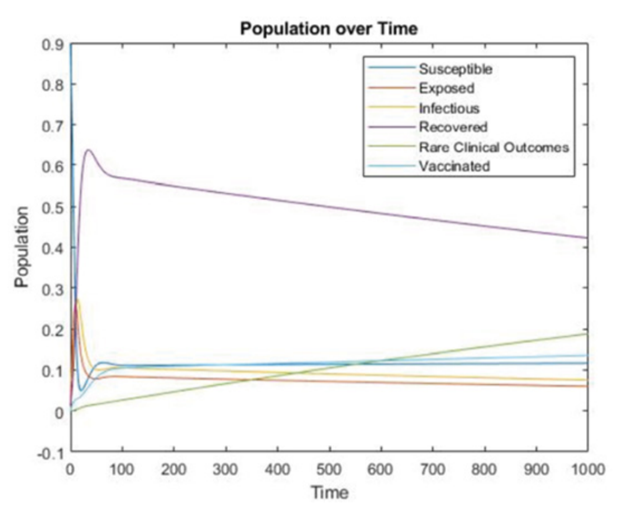

Figure 3: Mathematical projection over 1000 days of the fraction of individuals in each disease category

	Full Feature set		50 top features	19 top features
	Cascading LSTM	Convolutional LSTM	Cascading LSTM	Cascading LSTM
Sensitivity	55.97%	46.66%	36.75%	38.39%
Specificity	88.92%	73.25%	93.18%	87.87%
Positive Predictive Value (PPV)	40.60%	19.11%	42.17%	30.00%
AUROC	72.00%	59.96%	64.96%	63.13%
F1 Score	47.00%	27.11%	39.27%	33.69%

Table 1: Performance of Deep Learning Model variants on complete and subsets of clinical features

Benefits

A combination of the breakthroughs in rare event predictions and interpretations of the biomarkers obtained through deep learning with the simple and yet realistic disease outcome derived from mathematical modelling provides a robust support to clinicians in their need to assess long-term effects of COVID-19 on individuals.

The model achieved a breakthrough in rare event predictions with 89% specificity and 56% sensitivity in predicting COVID-19 hospitalisations. We found that high concentrations (above 10mg/L) of serum C-Reactive Protein were strongly associating with COVID-19 hospitalisations. Moreover, high CRP levels identified that 9% of the population had elevated risks of hospitalisation from COVID-19.

Further exploitation/next steps

Future use of these models is not limited to its application on COVID-19 but could offer insight into key predictors for other conditions within the EHRs. Enhanced versions of the model will allow us to understand indicators of rare outcomes. Thus, future exploration will be aligned with improving performance and explainability of the model.

Authors:

Prof. Xianghua Xie and Prof. Biagio Lucini, Gavin Tsang, Suraj Ramchand, Sam Oliver, Tabitha Lewis. Computational Foundry, Swansea University and Dr Duncan Cole, School of Medicine, Cardiff University.

Project group

Cardiff University / Prifysgol Caerdydd
Prifysgol Abertawe Swansea University