Forecasting Virus Outbreaks With Social Media Data Via Neural Ordinary Differential Equations

2025-01-01

Table of contents

1 Abstract

[R] Claims: social media data as early predictor of epidemic waves (1); online polls can be used as predictor (2); neural ODE can capture the dynamics and estimate new infections well (3); consequences of change in infections can be predicted with neural ODEs (4).

[T] Define COVID-19, neural ODE, social media, forecast, prediction, …

2 Introduction

[R] Pandemic → parameter estimates, not the other way around.

[R] That is a nice quote that was included.

[>] "You tell me what numbers to put in my equations, and I'll give you the answer …But you can't tell me the numbers, because nobody knows them…"

[Q] How is forecasting vital for health during epidemics and pandemics?

[Q] What health surveillance systems have been established across the globe?

[Q] What are example information sources in health surveillance?

[T] Add the other two entries from PAPER pg. 1

[T] Define weight adjustment and sample bias.

[R] What I am getting is that digital surveillance (old) and "late indicators" together as predictors outperform either predictor alone.

[T] Add Mermaid model for part with "M" on pg. 2. This describes the data available.

[T] Describe the tasks with "T" as a Mermaid model as well.

[T] Describe parts in [ ] using mathematics. What is "this object"?

[Q] How do "these phase space methods" (why is it called this) allow "the prediction of potential future … region."

3 COVID-19 Symptom Survey Through Facebook

[Q] What are ll the numerical indicators?

[R] The most meaningful part here is (1) how do the sruvey responses yield the numerical indicators and (2) what are the numerical indicators? (pg. 2)

[R] From Facebook (with public health officials) as the data providers. (pg. 2)

[Q] Why is this study a "non-formal" investigation of the indicators' recall? (pg. 3)

4 Models: First Principls And Data Driven

[R] Need a model that "relates the rate of variation if the different indicators to the model's state variables" and "relates the new cases as a function of the different signals extracted from the surveys."

[R] Claim is need a data drive over parameter driven model?

[R] For a region, $\vec{y}(t)$ is a vector of indicators (and new cases); the model is a "function that approximates the vector's temporal resolution".

[Q] "Nowt clear how to characterize the link between them from first principles." How much though went into this? "…absence of known functional form that links the variables."

[R] Rate of variation in indicators / variables as $\frac{\triangle \vec{y}}{\triangle t}$ with sufficiently small $\triangle t$ as $\frac{d\vec{y}}{dt}$.

[T] Define parameterized function, NN, …

[Q] What other parametric / non-parametric options exist for the task, why a neural ODE?

[R] $\frac{d\vec{y}}{dt} = NN(\vec{y}, t, \theta)$ with θ as the weights; also, this depends on t; the forward pass solves the initial value problem, i.e. gets the value of $\vec{y}(t_0)$.

[R] So neural-ODEs are time continuous so non-uniform data and predictors are available (unlike RNNs and LSTMs).

[T] Claude to use `forecasttools` or `forecasttools-py` (first with adding data access options) to get the data then set up MLflow comparison for neural ODEs with LSTMs and RNNs.