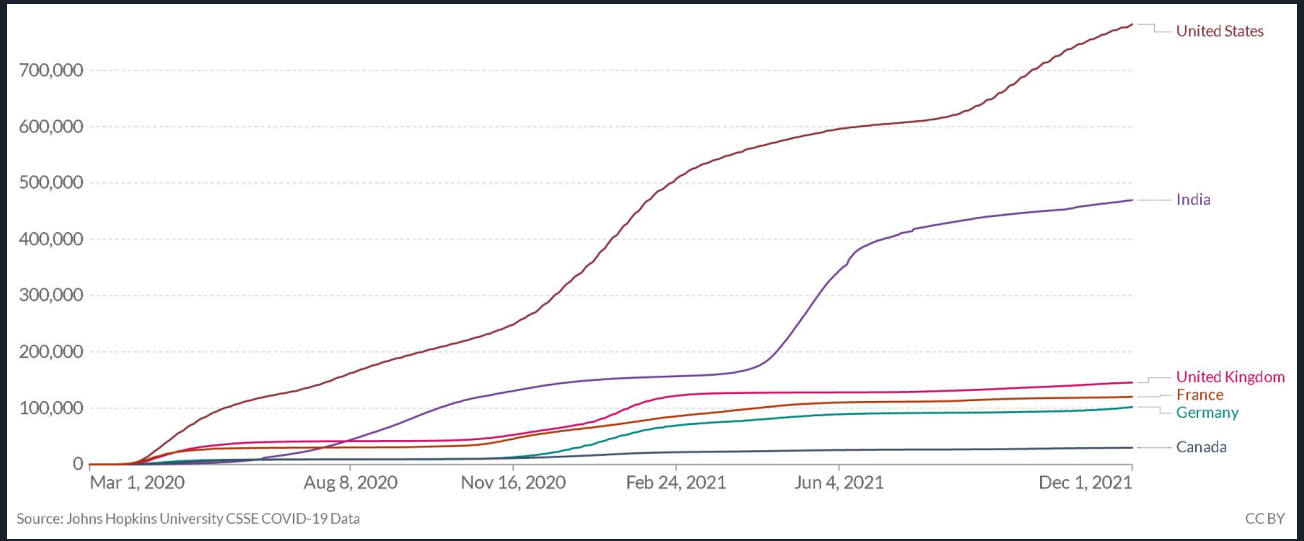
Forecasting COVID-19 Cases and Mortality using SIR Model with Time-Varying Parameters Forecasted via Ensemble Learning

Whitaker Chu, Kai Ouyang, Nicholas Saney

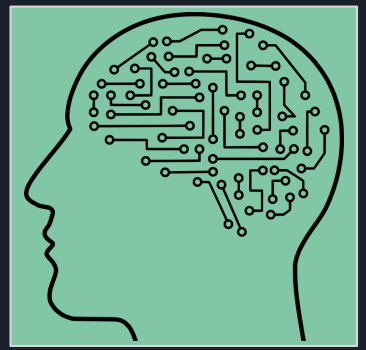
Motivation

- COVID-19 has killed 5.22M People Worldwide
- Modern data science is effective, but there is always room for improvement

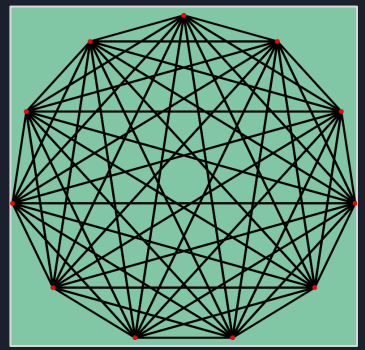


Introduction

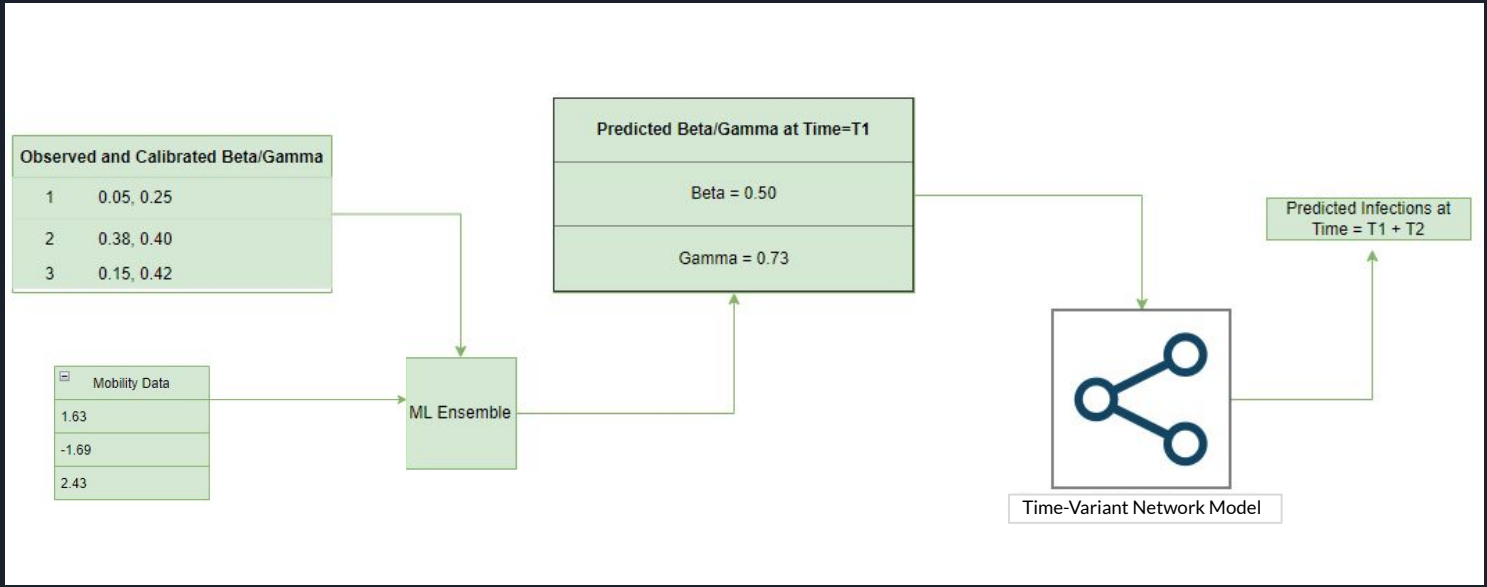
- Network based models are loose estimators with no external factors
- Deep Learning to predict deaths/infections with external factors is incredibly difficult
- Perhaps there is a way to use external factors to tune a network based model



VS



Our Model


Observed and Calibrated Beta/Gamma
1 0.05, 0.25
2 0.38, 0.40
3 0.15, 0.42
Mobility Data
1.63
-1.69
2.43
ML Ensemble
Predicted Beta/Gamma at Time=T1
Beta = 0.50
Gamma = 0.73
Time-Variant Network Model
Predicted Infections at Time = T1 + T2

Data Sourcing

- Mobility data from google and apple devices at statewide granularity
- Death and infection counts from Johns Hopkins and the CDC at statewide granularity
- Population data by state from the US Census (2019)
- ~93MB of raw data

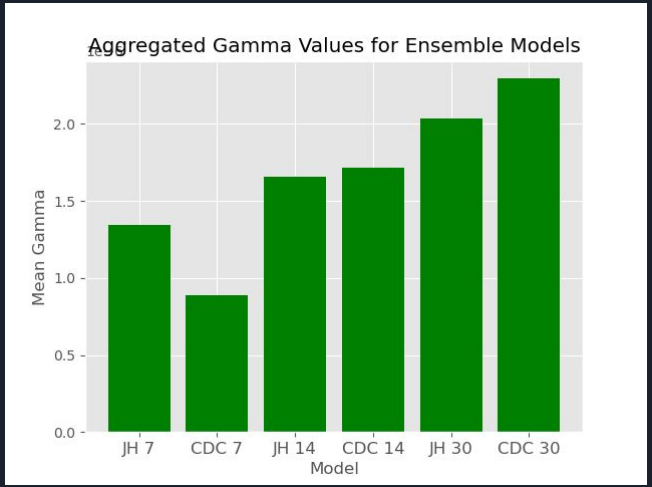
Walking	Driving	Transit	Δ Recreation	Δ Grocery	Δ Parks	ΔTransit Stations	Δ Workplace	Δ Residential	New Infections
109.04	119.22	52.58	-18.0	-6.0	31.0	-36.0	-32.0	12.0	9806.0

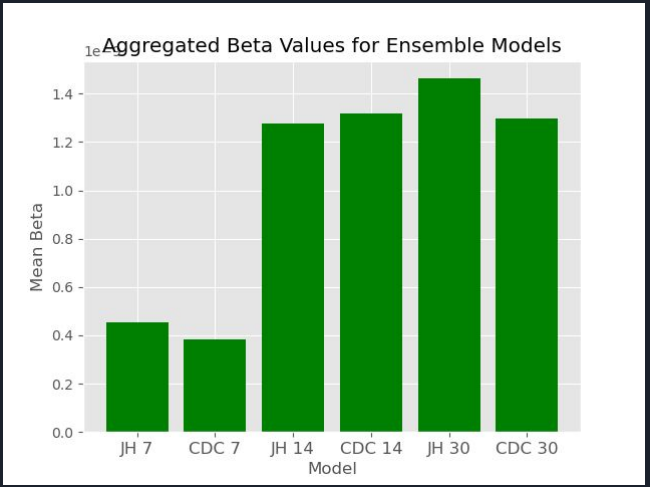
Calibrating Historic Beta and Gamma

- Historic Beta and Gamma were estimated using recorded daily infection/death counts from Johns Hopkins University
- The counts were then normalized based on state populations to produce Beta and Gamma values for each state for Jan-August of 2021

Predicting Future Beta and Gamma

- 6 models were chosen for and trained to predict beta and gamma using mobility data and historic Beta/Gamma
- These 6 models were then ensembled to create a single stronger model





Predicting Future Beta and Gamma

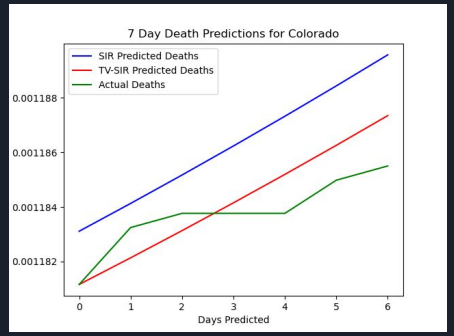
time range	cdc_beta_ARIMA	cdc_beta_OLS	cdc_beta_DT	cdc_beta_RF	cdc_beta_Bagged_DT	cdc_beta_Bagged_RT	cdc_beta_ensemble
7	1.73E-06	2.64E-09	6.80E-10	6.67E-10	1.38E-09	6.53E-10	3.82E-10
14	1.50E-06	5.61E-09	1.81E-09	1.79E-09	1.83E-09	1.73E-09	1.32E-09
30	2.10E-06	1.98E-09	1.91E-09	1.89E-09	1.91E-09	1.90E-09	1.30E-09
avg	1.78E-06	3.41E-09	1.47E-09	1.45E-09	1.71E-09	1.43E-09	1.00E-09
time range	jh_beta_ARIMA	jh_beta_OLS	jh_beta_DT	jh_beta_RF	jh_beta_Bagged_DT	jh_beta_Bagged_RT	jh_beta_ensemble
7	1.34E-06	5.28E-09	8.28E-10	7.95E-10	3.12E-09	9.58E-10	4.52E-10
14	2.38E-06	6.98E-09	1.75E-09	1.74E-09	1.64E-09	1.58E-09	1.28E-09
30	1.85E-06	1.85E-09	2.19E-09	2.17E-09	2.18E-09	2.17E-09	1.46E-09
avg	1.86E-06	4.70E-09	1.59E-09	1.57E-09	2.31E-09	1.57E-09	1.06E-09

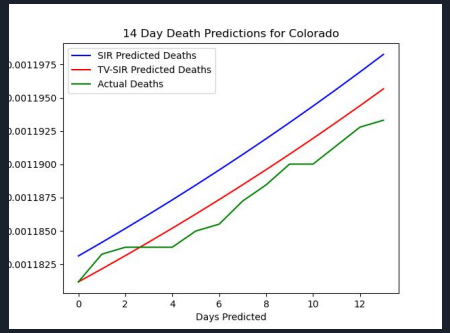
time range	cdc_gamma_ARIMA	cdc_gamma_OLS	cdc_gamma_DT	cdc_gamma_RF	cdc_gamma_Bagged_DT	cdc_gamma_Bagged_RT	cdc_gamma_ensemble
7	6.16E-06	1.11E-05	3.12E-06	2.75E-06	3.23E-06	2.74E-06	8.89E-07
14	1.63E-05	4.87E-05	4.17E-06	4.40E-06	4.93E-06	4.40E-06	1.72E-06
30	1.90E-05	1.13E-05	5.25E-06	5.07E-06	5.23E-06	5.18E-06	2.29E-06
avg	1.38E-05	2.37E-05	4.18E-06	4.07E-06	4.46E-06	4.11E-06	1.63E-06
time range	jh_gamma_ARIMA	jh_gamma_OLS	jh_gamma_DT	jh_gamma_RF	jh_gamma_Bagged_DT	jh_gamma_Bagged_RT	jh_gamma_ensemble
7	7.16E-06	1.22E-05	3.36E-06	3.43E-06	6.53E-06	3.42E-06	1.34E-06
14	1.97E-05	3.20E-05	5.33E-06	3.36E-06	5.83E-06	3.63E-06	1.65E-06
30	1.87E-05	9.61E-06	4.11E-06	4.26E-06	4.16E-06	4.16E-06	2.04E-06
avg	1.52E-05	1.79E-05	4.26E-06	3.68E-06	5.51E-06	3.74E-06	1.68E-06

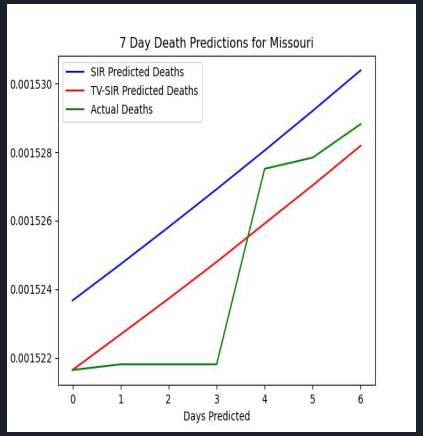
Discrete Time-Variant SIR (TV-SIR) Model

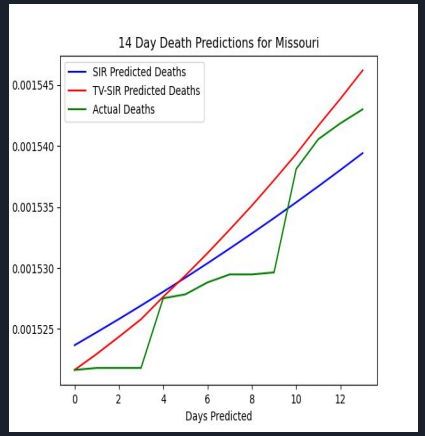
- Operates on the assumption that Beta and Gamma are non-constant
- Works similarly to a standard discrete-type SIR model, but factors in a new Beta and Gamma at each time step
- We use the forward-looking Beta/Gamma predictions as input, and allow TV-SIR model to base its own predictions on those predictions

Results

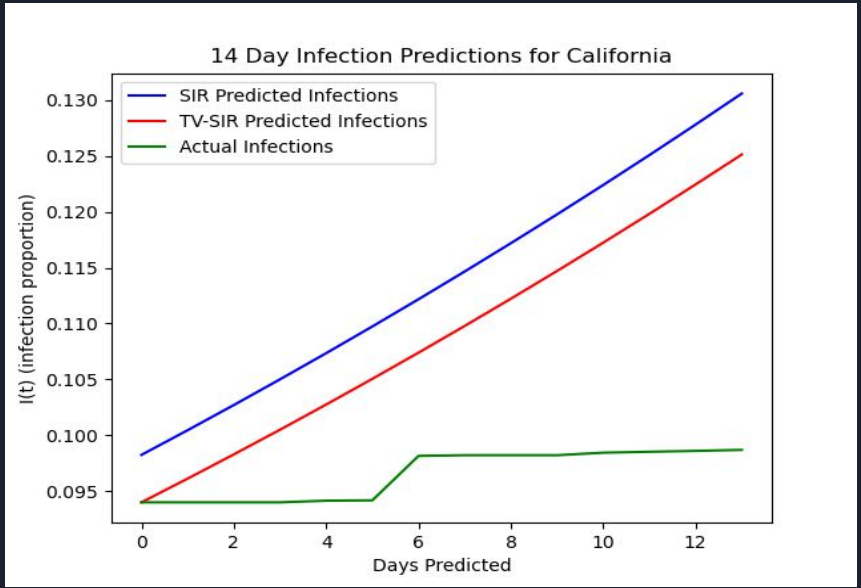

7 Day Death Predictions for Colorado
SIR Predicted Deaths
TV-SIR Predicted Deaths
Actual Deaths
0.001188
0.001186
0.001184
0.001182
0 1 2 3 4 5 6
Days Predicted

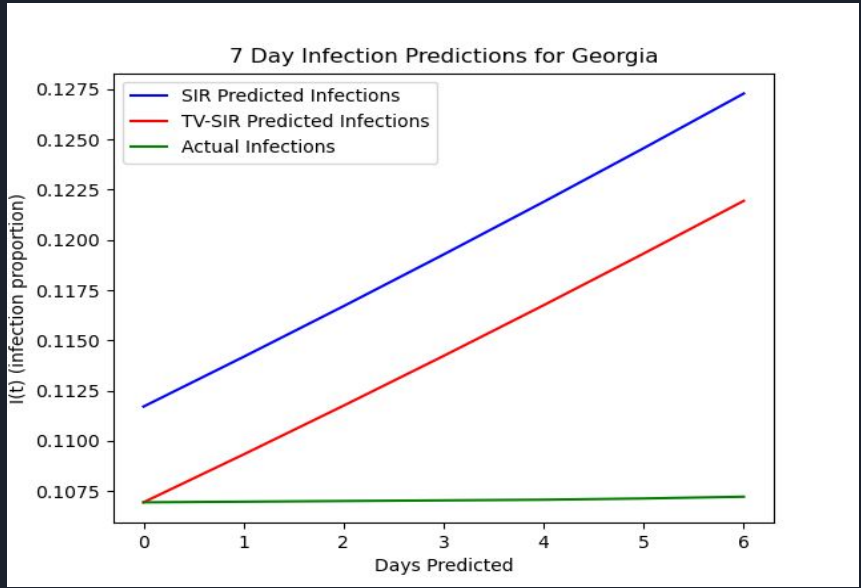

14 Day Death Predictions for Colorado
SIR Predicted Deaths
TV-SIR Predicted Deaths
Actual Deaths
0.0011975
0.0011950
0.0011925
0.0011900
0.0011875
0.0011850
0.0011825
0 2 4 6 8 10 12
Days Predicted


7 Day Death Predictions for Missouri
SIR Predicted Deaths
TV-SIR Predicted Deaths
Actual Deaths
0.001530
0.001528
0.001526
0.001524
0.001522
0 1 2 3 4 5 6
Days Predicted


14 Day Death Predictions for Missouri
SIR Predicted Deaths
TV-SIR Predicted Deaths
Actual Deaths
0.001545
0.001540
0.001535
0.001530
0.001525
0 2 4 6 8 10 12
Days Predicted

Results Continued





Discussion

- Confirmed cases ≠ Actual cases
- TV-SIR *can* be used to estimate deaths more reliably than standard-SIR
- Estimating Beta/Gamma could be used to validate future models