



Mathematical Biology Seminar

Monday, March 11, 2024

3 pm MDT - 457 CAB (in person)

Join Zoom Meeting

<https://ualberta-ca.zoom.us/j/98497695684?pwd=SG5pcUVRS0xucW5xd0xBTm1VVctEUT09>

Meeting ID: 984 9769 5684

Passcode: 32123



Amit Chakraborty

Department of Mathematical and Statistical Sciences

University of Alberta

From policy to prediction of infectious diseases: All machine learning models do well; Bayesian networks do better.

To improve the forecasting accuracy of the spread of infectious diseases, a hybrid model was recently introduced where the commonly assumed constant disease transmission rate was actively estimated from enforced mitigating policy data by a machine-learning model and then fed to an extended susceptible-infected-recovered (SIR) model to make forecasts. Testing only one machine-learning (ML) model, i.e., gradient boosting model (GBM), the work left open whether other ML models would perform better. Here, we tested Bayesian networks, gradient boosting models, k-nearest neighbors, and linear regressions to forecast the number of COVID-19-infected cases in the US and Canadian provinces based on policy indices of future 35 days. There was no significant difference in the mean absolute percentage errors (MAPEs) of these ML models over the combined dataset. However, in two provinces, a significant difference was observed, with Bayesian networks outperforming the other models. Moreover, Bayesian networks significantly outperformed the other models in the training datasets. Although performing similarly, the results put forward Bayesian networks as the best candidate, should there be only one ML model used.

COLLABORATIVE MATHEMATICAL BIOLOGY GROUP

***MATHEMATICAL & STATISTICAL SCIENCES
UNIVERSITY OF ALBERTA***