Title

Analyzing the Impact of Different Countries' Approaches to the COVID-19 Pandemic on Their Cumulative Infection Curves By Using Nonparametric Density Regression and Clustering Methods

200-Word Abstract

This project aims to study patterns in different national responses to the pandemic and their effect on controlling the outbreak through a nonparametric density and regression technique and various clustering tools. Seven different countries, United States of America, Italy, South Korea, Taiwan, United Kingdom, Brazil, and India were chosen for analysis due to their varied methods and successes in handling COVID-19 infection. First, a novel nonparametric density technique was utilized to accurately and consistently partition each country's COVID-19 cumulative growth curve into different waves and phases, and each phase was modeled using linear, quadratic, or logarithmic regression. Every country's model variables, which are the slope, r-square value, duration, and model type of each phase, were then connected to real-life factors, such as cluster outbreaks, government regulations, and the availability of healthcare resources. Multivariate correlation was conducted to uncover the relationships between model variables, and running the variable clustering algorithm showcased which model variable from a previous phase would be a good predictor for the infection situation in the following phase. Finally, based on the multivariate correlation and variable clustering results, the most important variables were used in hierarchical clustering to identify and explain the most similar and most different countries.

150-Word Abstract

This project aims to study patterns in national responses to the pandemic and their effect on controlling the outbreak through a nonparametric density and regression technique and various clustering tools. A novel nonparametric density technique was utilized to consistently partition seven different countries' COVID-19 cumulative growth curves into waves and phases, and each phase was modeled using linear regression. Every country's model variables (the slope, r-square value, duration, and model type of each phase) were then connected to real-life factors, such as cluster outbreaks and government regulations. Multivariate correlation was conducted to uncover the relationships between model variables, and variable clustering showcased which model variable would be a good predictor for the infection situation in a certain phase. Finally, based on the multivariate correlation and variable clustering results, the most important variables were used in hierarchical clustering to identify and explain the most similar and most different countries.

Keywords

COVID-19, Phase Modeling, Nonparametric Density, Regression, Cluster Variables, Multivariate Correlation, Hierarchical Clustering