Covid-19 Outbreak - Cross sectional Analysis of US Counties

Muhammad Zargham, Clark Granger, Zachery Carlson

May 8, 2020

Abstract

We conduct a cross sectional analysis of Covid-19 cases as on April 30, 2019 across the US counties based on their socio-economic and geographical factors that may be important in explaining the different pace of virus spread. Along with identifying these important factors, we try to fit negative binomial, KNN and Random forest models to predict the number of covid-19 cases and to see if these socio economic variables can lead to a better prediction of covid cases. We find that some important socio-economic features are correlated with the number of cases and can give further insight for studying this pandemic.