Progress Report.

Main changes: We are no longer investigating factors that put specific people at risk, instead analyzing counties instead. Additionally we are pivoting from neural networks to a random forest

Model, given that the quantity of dependent variables makes it hard to scale.

So far in this project, we have constructed a random forest regressor model in which we are analyzing the death rate of counties based on certain data in hopes of predicting more at risk counties. We are training our random forest on counties that have a high rate of infection and mortality rate. We can then use our trained random forest on counties with little to no infection to predict what might happen in case of infection in those counties. Currently, our trained random forest has a high rate of mean absolute error and we believe this might be to our data not being thoroughly cleaned as well as including training on counties that do not have a long enough exposure to the virus. We believe that cleaning further and making some modifications will allow us to bring down this error and further increase the predictive power of our trained random forest.

Josh Cordova

Erik Tarango

Kobe Rico

Aech Loar