**PLEASE NOTE:**

There are two code files in the zip folder

* The file ‘**Covid-19\_pred\_project.ipynb’** is the notebook code file for the model predicting covid\_19 cases in a US county with **4** algorithms (**Random Forest, Linear Regression, Bayesian Ridge regression, Gradient Boosting regression) and k-fold cross validation.**
* The file ‘**Bayes\_County(1).ipynb**’ is the file implementing **Naïve Bayes for each county.**

The file **‘covid-19\_pred\_proj\_code.pdf’** is the pdf version of our Jupyter notebook code file.

The file **‘project-paper-updated-draft.docx’** is an incomplete draft of the paper with updates changes

The file **‘Covid-19\_prediction\_merin\_tianhao\_ruiling.pptx’** is the presentation slide (incomplete) that will be updated soon.

The file **‘covid19cases\_test.csv’** is the dataset file taken from the the California Health and Human Services Open Data Portal (CHHS) website – [click here](https://data.chhs.ca.gov/dataset/covid-19-time-series-metrics-by-county-and-state/resource/046cdd2b-31e5-4d34-9ed3-b48cdbc4be7a).