

MAIS 202

FALL 2020

## COVID-19 X-RAY DETECTION

1. The **dataset** is already available on,  
<https://www.kaggle.com/bachrr/covid-chest-xray>  
I choose this dataset because the metadata fields make it quite ideal for the way I want to execute the data analysis on this project.
2. The **methodology** I plan to use is KNN model. The dataset has a specific field for the findings, which I can associate with the age, offset. And I can train the model to get to detect the infection. The dataset is very well detailed, already categorised so that the data analysis becomes simplified.

The **accuracy** of the model is I would say 60-70, because this is still a novel disease, the data can be centric to a few people who got infected by a single strain of the virus, while the virus is still mutating and every strain differs.