**WIL Project Progress and Future plan**

1. **WIL Project Progress**

* We have selected a business case “As an WHO, we are predicting that a person is susceptible to covid virus based on his/her age and sex. Moreover, we will also try to categorize it based on locations” for our project.
* We went through different datasets and chose one dataset which includes some features that we think may help us to get some solution for the business case.
* The dataset that we are working on is publicly available on GitHub.

< [*https://github.com/beoutbreakprepared/nCoV2019*](https://github.com/beoutbreakprepared/nCoV2019) *>*

* We are using R language to better interpret the statistics and for creating interactive web application.
* We are half way done with the Data Preprocessing as it takes considerable amount of processing time for Data Analysis.

1. **Future Plan**

* We will be completing the Data Preprocessing step in the upcoming week. Then we will move ahead with the descriptive analysis.
* We will further visualize the dataset to identify patterns and to extract meaningful insights in compliance with our business case.
* We will apply hill climbing algorithm to select features for building efficient machine learning models.
* We will train different models to then predict the unseen labeled data. The performance metrics such as classification report and confusion matrix will be used to identify the mislabelled data.
* The metrics will be used to measure the predictive power of different machine learning algorithms.
* Overall, the algorithm with the best performance measures will be used for solving our business case.