**Student Name:** Keyun Zhou

**Student ID:** 501147287

**Date of Submission:** September 29, 2022

**Predictive analysis for factors associated with mortality risk in COVID-19 patients**

Coronavirus disease 2019 (COVID-19) is a pathogenic viral infection that has been negatively affecting human health on a global scale, causing millions of deaths (WHO, 2022). With its sudden surge in 2019, healthcare facilities were well exceeding their functional capacities. Despite the effort to contain and delay the spread of the disease, the number of infected individuals continued to rise. With an increase in patient volume, hospital beds, medical equipment, and healthcare staffing becoming increasingly in demand, it soon caused great distress in the healthcare system. Furthermore, this lack of adequate response and control measures to mitigate the current stress could lead to further exposure and poor health outcomes.

COVID-19 is known to spread through respiratory droplets or direct contact with an infected individual, such as through saliva or sneezing, attacking the respiratory system of an individual (Han et al., 2020). However, the range of symptoms or the degree of severity after the contraction of CONVID-19 varies from case to case. Some patients do not display any symptoms, known as asymptomatic, and some experiences mild symptoms such as fever and cough to more severe cases such as organ damage and even death (Mohapatra et al., 2020). The underlying cause of such variation among individuals can be complex. Data showed that some groups could be at greater risk than others, depending on several factors. The study by Ma et al., 2020 suggested that more severe cases of COVID-19 were observed in males (59.7%) and in patients of older age (54.5 vs 44.5 years of age) (Ma et al., 2020). In the current project, a predictive analytics theme will be used to examine factors, such as age, sex, and geographical location that are associated with detrimental outcomes resulting from COVID-19, to facilitate emergency response planning, and relieve potential stress on the healthcare system of Ontario. This will help to prioritize and target patients to increase efficiency in patient care, and also to raise public awareness of the prevention of COVID-19. The questions are: what types of factors predict COVID-19 mortality, whether there is a specific link between gender and severity, dataset is obtained from an Open data catalog, by the Government of Toronto, providing confirmed cases of COVID-19. The project will predict the mortality of COVID-19 patients using Python and R, with Random Forest, logistic regression, and XGBoost.

**Data set link:** <https://data.ontario.ca/en/dataset/confirmed-positive-cases-of-covid-19-in-ontario>

References

Han, Q., Lin, Q., Jin, S., & You, L. (2020). Coronavirus 2019-nCoV: A brief perspective from the front line. *Journal of Infection, 80*(4), 373-377.

Ma, C., Gu, J., Hou, P., Zhang, L., Bai, Y., Guo, Z., Wu, H., Zhang, B., Li, P., & Zhao, X. (2020). Incidence, clinical characteristics and prognostic factor of patients with COVID-19: a systematic review and meta-analysis. *medRxiv.* <https://doi.org/10.1101/2020.03.17.20037572>

Mohapatra, R. K., Pintilie, L., Kandi, V., Sarangi, A. K., Das, D., Sahu, R., & Perekhoda, L. (2020). The recent challenges of highly contagious COVID-19, causing respiratory infections: Symptoms, diagnosis, transmission, possible vaccines, animal models, and immunotherapy. *Chemical Biology & Drug Design, 95*(5), 1187-1208.

World Health Organization. (2022, September 16). *WHO Coronavirus (COVID-19) Dashboard*. <https://covid19.who.int/>