Related Work (Rough Draft)

The related and previous work considered in completing this project falls broadly under two categories. The first is the related work that characterizes how demographics are differently affected by Covid-19 infection and how that research educates the direction of our project. The second is how our selected data mining algorithm has been applied previously and how and if our selected changes have been implemented in similar cases.

While the study of Covid-19 only began in earnest this year for obvious reasons, there is significant research already completed that we may observe to direct our own goals for this project. One such preliminary study briefly compares the mortality rate between Chinese and Italian Covid-19 patients and reveals significant disparity in fatality rate between equivalent populations [s1]. The fact that distributions of similar people in different geographic locations have different resilience to Covid-19 prompts inspection of this question on a finer regional level. While two countries as geographically and ethnically distinct as China and Italy may have considerable differences it may be worth examining if such regional changes in Covid-19 recovery rates exist regionally in Canada.

Examining how demographics react to Covid-19 across Canada has already been attempted in a paper entitled “Demographic Profile of COVID-19 Cases, Fatalities, Hospitalizations and Recoveries Across Canadian Provinces” [s2]. This study compares the rates of hospitalization, fatality, and recovery of Canadians across the country by several different demographic characteristics. The authors break risk populations up by sex and age in a given region and found significant differences in the rate of hospitalization and death between provinces. However, this study was published in May of 2020 and is therefore extremely limited to a few select provinces that had substantial outbreaks at the time, namely British Columbia, Ontario, and Quebec. As of the time of writing (November 22, 2020), most provinces have substantial outbreaks ongoing in addition to significant daily hospitalizations and fatalities. With more up-to-date data and a larger sample size, it’s likely that we can find more similar results across the country.

A broader study done in Sweden examines finer demographics utilizing data about a subject’s age, net income, education, civil status, and country of birth [s3]. The additional personal information allows for a finer examination of exactly what factors might influence a patient’s chances of surviving Covid-19. So while a simple examination of age and gender yields valuable information, a greater knowledge of an individual’s health and socio-economic position will provide much more detailed information on what risk factors are most important to patient death.

Finally, while examining how Covid-19 affects demographics differently is not a novel idea, it may provide novel results given that new data is being received on a daily basis. Moving into our method of mining we can take the lessons learned in previous studies to drive the direction of our data mining efforts.

Sources:  
s1: https://onlinelibrary.wiley.com/doi/pdfdirect/10.1002/jmv.25860

s2: <https://www-deslibris-ca.uml.idm.oclc.org/ID/10103979>

S3: https://www.nature.com/articles/s41467-020-18926-3