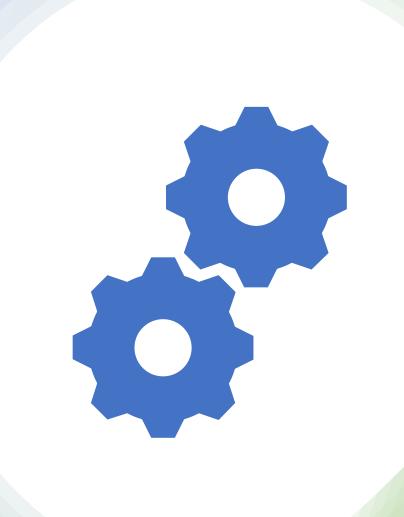
Technical Assessment (Part 1)



## Research Objectives



Predicting the risk of severe illness or death from COVID-19

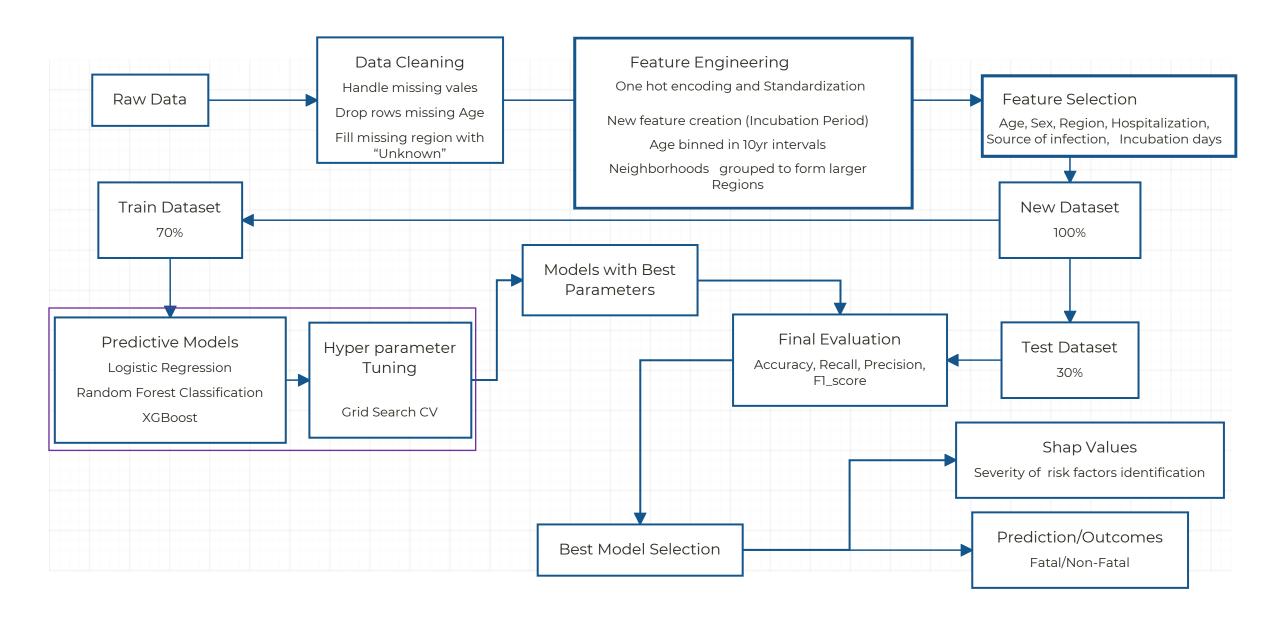


Determining the severity of factors on COVID-19 fatalities

## Data Structure and Exploratory Data Analysis (EDA)

Key insights derived from the analysis of **Independent Features** Dataset: COVID-19 ML models selected Tools/Libraries used data exploration(Fatal cases in Toronto selected Outcomes): EDA (Pandas, Supervised – Higher likelihood Age, Sex, Seaborn (Catplot, Patient level data Labeled training Classification in older age group Distplot, Boxplot, data is available NumPy Classification -Hospitalization Males have a Last updated: Machine Learning Target variable is higher probability history, Source of February 2023 categorical (Fatal (Scikit-learn) than women Infection vs. Non-Fatal) Patients with **Evaluation/metrics** 397.089 entries. Region ,Incubation Base model hospitalization (Scikit-learn) Logistic Regression 15 columns history are more days likely Target column Source of infection Random Forest Identified: less likely due to Classification, "Outcome" XGBoost close contact

## Supervised Machine Learning Classification Workflow



## Conclusion



The utilization of labeled training data holds promise in identifying valuable patterns and correlations



The model's potential to prioritize healthcare resources by predicting COVID-19 patient risk levels is significant



The exploration of influential factors may provide crucial insights into COVID-19 fatalities, guiding informed intervention strategies.



Continuous monitoring and proactive model updates are imperative for adapting to evolving COVID-19 patient outcomes