

# COVID ML Program- A Concept Overview

## Machine Learning Algorithm:

The goal of this algorithm is to examine patient conditions and assign probabilities to each patient. By feeding the algorithm significant data, we can extrapolate from this how likely it is a patient will die from COVID-19. By presenting this data accurately, it will serve as a model to aid doctors in their fight against the Pandemic and save lives.

COVID-19 is an insidious disease. It can pry on the most vulnerable populations as well as cause a considerable impact on healthy individuals.

## Data Analysis

Based on current research the following conditions have been shown to have a significant impact on COVID Mortality. These will make up the basis of the program.

Dataset Summary

Blood Pressure - Ratio

Smoker - Y/N

O2 Saturation - %

Asthma - Y/N

Diabetes - Y/N

Obesity - Y/N

Lung Disease - Y/N

Heart Disease - Y/N

## Program Execution

How the program will execute its functions:

By applying a k-th nearest neighbor algorithm, I will be able to give within reasonable accuracy the change of mortality. This result will be derived from the Machine Learning Program as it will be able to assess the greatest risk factors and associate patient data with a higher likelihood of mortality.

To ensure accuracy, I will also be implementing a secondary ML program to verify/contradict the results. Any discrepancies will be obvious, and will allow for accurate assessment of reasoning behind the results.

## **Data Normalization:**

Given the broad range of COVID symptoms and the resulting effects (i.e O2 Saturation, BP, and disease history), it is important to normalize the data. This will ensure that outliers are included in the final analysis. It also establishes a baseline for the algorithm to work off of to ensure a similar picture and accurate data representation.