The "ASK" Step Overview

Sarah Warda

The Problem

Covid-19 resulted in millions of deaths during the pandemic, where patient mortality was difficult to predict. Over time, we began to narrow down risk factors for poor patient outcomes to help the public make educated decisions about their health and well-being. Although medical professionals could narrow down risk factors, some patients without these risk factors died after hospitalization unexpectedly. In addition, hospitals struggled with PPE, hospital beds, medicine, staff, and various medical supplies. As a result, medical professionals would attempt to predict patient outcomes based on research, personal experience, and patient symptoms to effectively save as many lives as possible.

The goal of this project is to answer the question:

> *What risk factors and diagnostic features can be used to classify patient outcomes as a result of Covid-19, and can we use this model to predict patient death?*

The Data

The data set, which was obtained from Kaggle, is a cohort of patients that were used in research applications during the pandemic to find trends and important diagnostic features for patient mortality after hospital admission due to Covid-19. The original data set contains 85 variables and 4,711 rows, where each row is an individual patient. The 85 variables include basic patient information, pre-existing condition status, and blood panel results during their hospital stay. We will only start with 43 of the variables since half of the variables were transformations of existing variables. The original data set, which was later posted to Kaggle, is from a publication in 2021, although the publicly available data has been modified to protect patient confidentiality.