Vironix Group Presentation

Presented by Victoria McGraw, Hanna Reed, and Hunter Walt
6/18/2021

What is COPD?

Chronic Obstructive Pulmonary Disease refers to a group of Diseases that cause airflow blockage and breathing problems

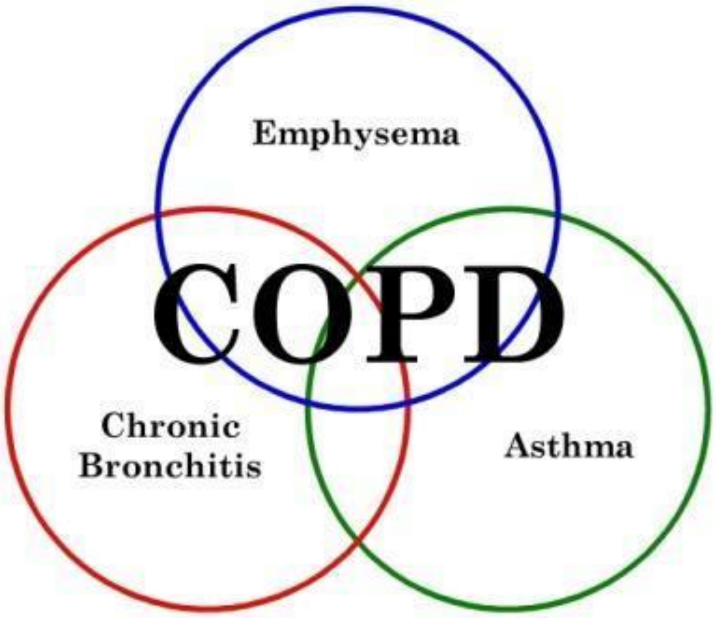
3rd leading cause of death worldwide

12-16 million cases in the US

Main risk factors: tobacco smoking, chemical inhalation, burning biofuels

There can be acute worsening of symptoms: exacerbations

COPD is associated with significant other chronic illnesses



Symptoms of COPD

Patients may also experience different symptoms in different stage:

Most common respiratory symptoms are dyspnea, cough, sputum production, wheezing

Early symptoms include: occasional shortness of breath, mild but recurrent cough, and the need of clearing throat often

Worsening symptoms include: shortness of breath, wheezing, chest tightness, chronic cough, frequent respiratory infections such as cold or flu, lack of energy, fatigue, swelling of the feet, ankles, or legs, and possible weight loss.

Emphysema

Wall fibers in alveoli sacs damaged

Macrophages and neutrophils secrete elastase

Destroys wall elasticity needed for full exhalation

Labored breathing, cough, mucus

Heart overwork & damage

Damage irreversible

Chronic Bronchitis

Airway inflammation - excess viscous mucus

Impaired cilia function - impairs clearance

Result: Airway blockage

Too little O_2 in, CO_2 out

High risk of exacerbations

High risk of mortality from all causes

Exacerbations

Sudden increase in airway resistance

Life threatening

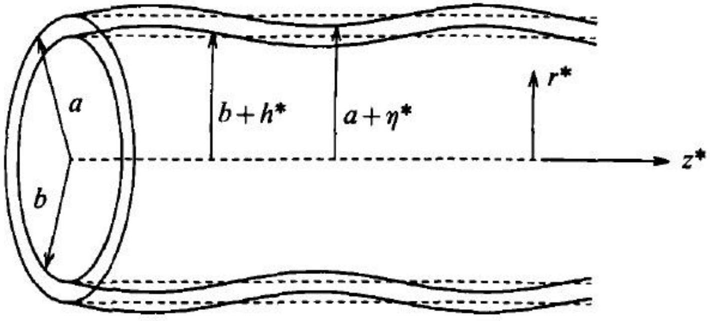
Severe outflow limitation

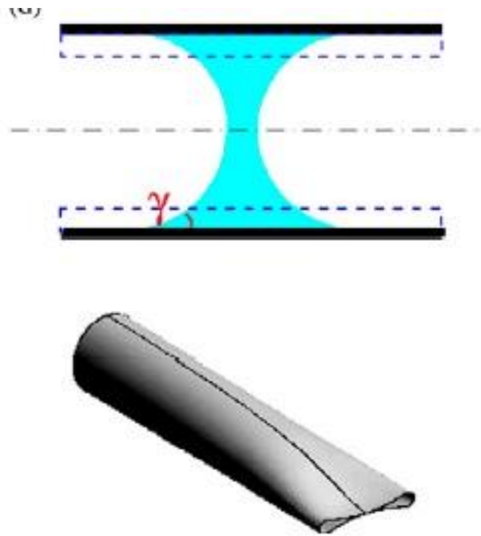
Dynamic lung hyperinflation

Pressure after exhalation positive, not negative -> hard to inhale

Surface tension and wall elasticity instabilities collapse airways

Critical thickness **h*** of fluid lining can trigger sudden liquid lens formation or wall collapse:





h* down as: Surface tension up (dysfunctioning surfactants)

Wall elasticity up

Emphysema has stiffer walls - **Chronic bronchitis has higher exacerbation risk**

Vironix Goal

- Create an app that can take real-time patient data at home and predict exacerbation severity
- Can we establish baseline symptoms/features of COPD that can help predict acute exacerbation?
- Can we generate realistic simulated data?
- Helpful for personalized care and early detection of acute exacerbations

Concerns

Definitions of exacerbations vary in the literature

Hospital patient data is hard to get access to

Data bias by cohorts

Many GOLD Standard symptoms indicative of severe exacerbation are either hard or unreliable to measure at home

Features/Comorbidities of Interest

- Features:
  - Age (Categories: 10 years from 40-90))
  - Previous Exacerbations (Categories: 0-4)
  - Sex
  - Smoking Status
  - Wheezing
  - Coughing
  - Sore Throat
  - Rhinorrhea
  - Sputum Production
  - Headache
- Comorbidities:
  - Asthma
  - Chronic Heart Failure
  - Chronic Kidney Disease

Data

Standardizing Measurement of Chronic Obstructive Pulmonary Disease Exacerbations

PLOS | ONE

RESEARCH ARTICLE

Predicting hospital admission at emergency department triage using machine learning

Woo Suk Hong[1], Adrian Daniel Haimovich[1], R. Andrew Taylor[2]*

1 Yale School of Medicine, New Haven, Connecticut, United States of America, **2** Department of Emergency Medicine, Yale School of Medicine, New Haven, Connecticut, United States of America

a case-control study

G Rohde, A Wiethege, I Borg, M Kauth, T T Bauer, A Gillissen, A Bufe, G Schultze-Werninghaus

Thorax 2003;**58**:37–42

Real Hospital Data Set

- Emergency Room admission data from 560,486 patients with 972 columns of data
- 44,343 had COPD
- 16,347 had COPD and asthma
- 9,425 had COPD and heart failure
- 5,885 had COPD and chronic kidney disease
- 773 had COPD, asthma, heart failure, and chronic kidney disease
- Issues with data: Incomplete and only provides chief complaint, but has valuable vitals/Rx information

Real Hospital Data Set

Condition	Admitted	Discharged	Percent Admitted
COPD	24,243	20,100	54.6%
COPD and Asthma	7,962	8,475	48.4%
COPD and Heart Failure	6,458	2,967	68.5%
COPD and Kidney Disease	4,004	1,881	68.0%
COPD, Heart Failure, and Kidney Disease	531	242	68.7%

Generating Patient Data - Features

We wish to generate patients with the following information:

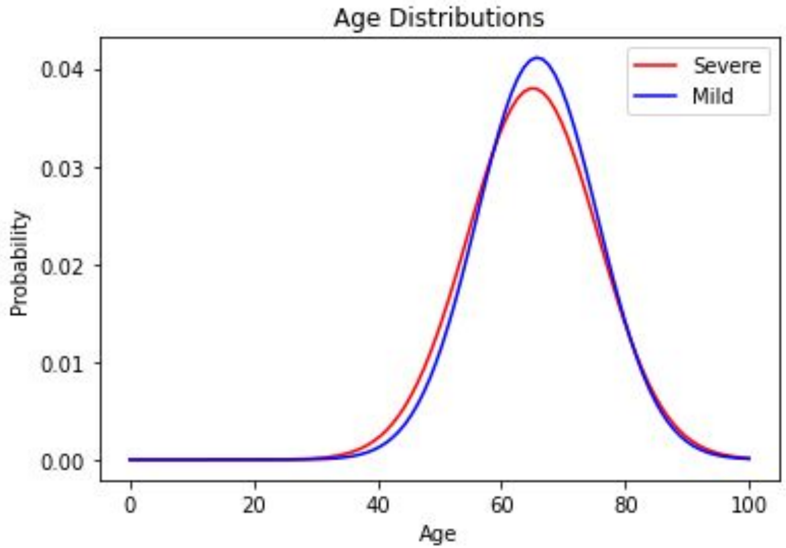
Age	{0, 1, 2, 3, 4}	**Congestion**	{0, 1}
Previous Exacerbations	{0, 1, 2, 3, 4}	**Sore Throat**	{0, 1}
Wheezing	{0, 1}	**Headache**	{0, 1}
Smoking	{0, 1}	**Rhinorrhea**	{0, 1}
Sex	{0, 1}	**Sputum**	{0, 1}

Generating Patient Data - Simple Method

We assume a normal distribution for age [3]

We generate ages according to these distributions and categorize them .

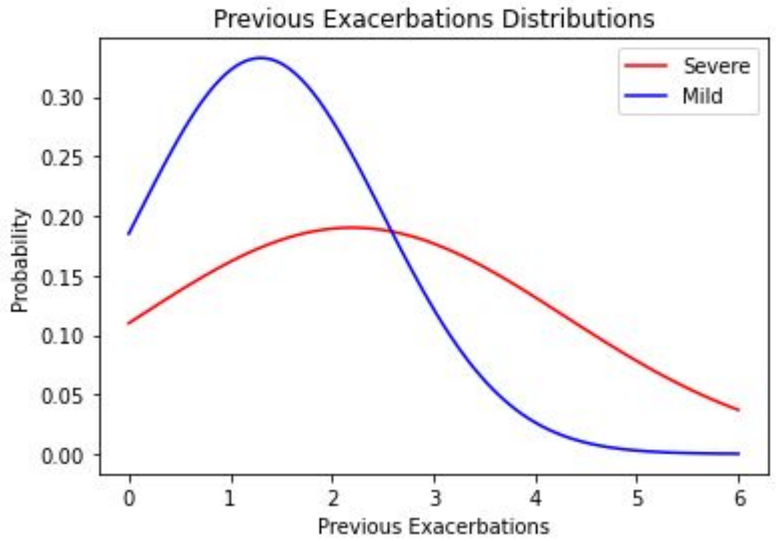
Group	0	1	2	3	4
Age	<50	50-60	60-70	70-80	>80



Generating Patient Data - Simple Method

Similarly, we assume a normal distribution for previous exacerbations with means and standard deviations provided by [3]

We generate a number of previous exacerbations for each patient and round to the nearest integer between 0 and 4.



Generating Patient Data - Simple Method

For binary features, we calculate the probability that an individual with a severe exacerbation has a given feature:

$$P(F|S) = \frac{P(S|F)P(F)}{P(S)}$$

We assume that this probability will be the proportion of severe patients that have this feature.

Generating Patient Data - Results

We generated 1000 mild patients and 1000 severe patients

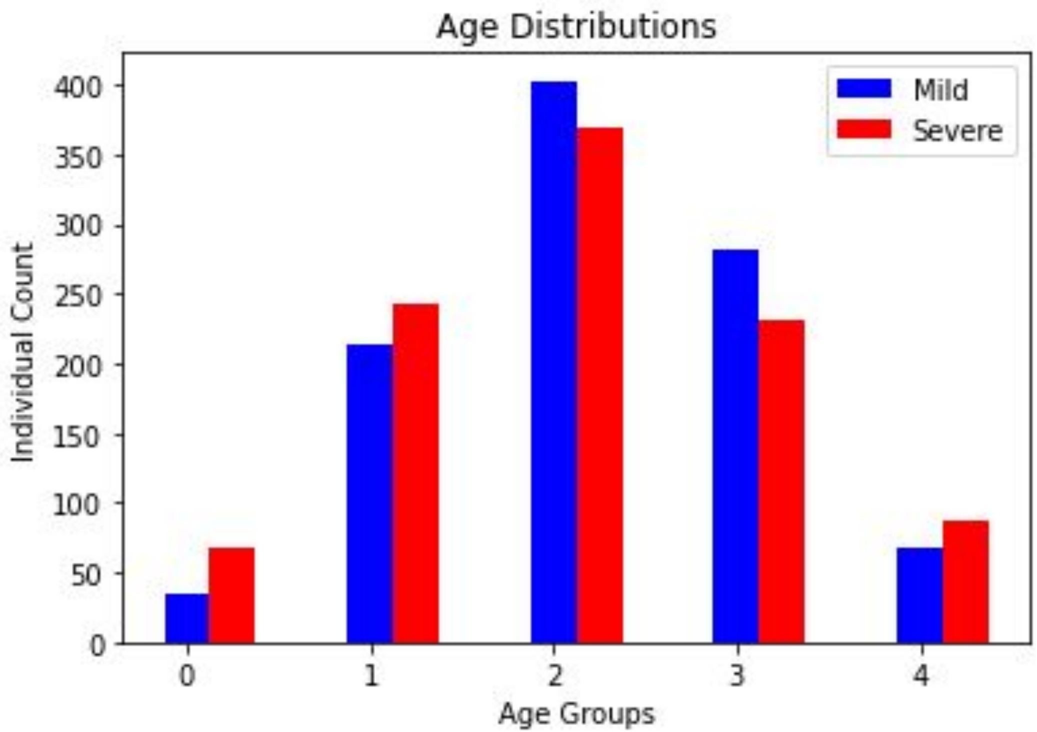
In the following slides, we will discuss:

- Distributions of features among patients
- Correlations between features found in the generated data

Age

- Mean and standard deviation for age was very similar from our source

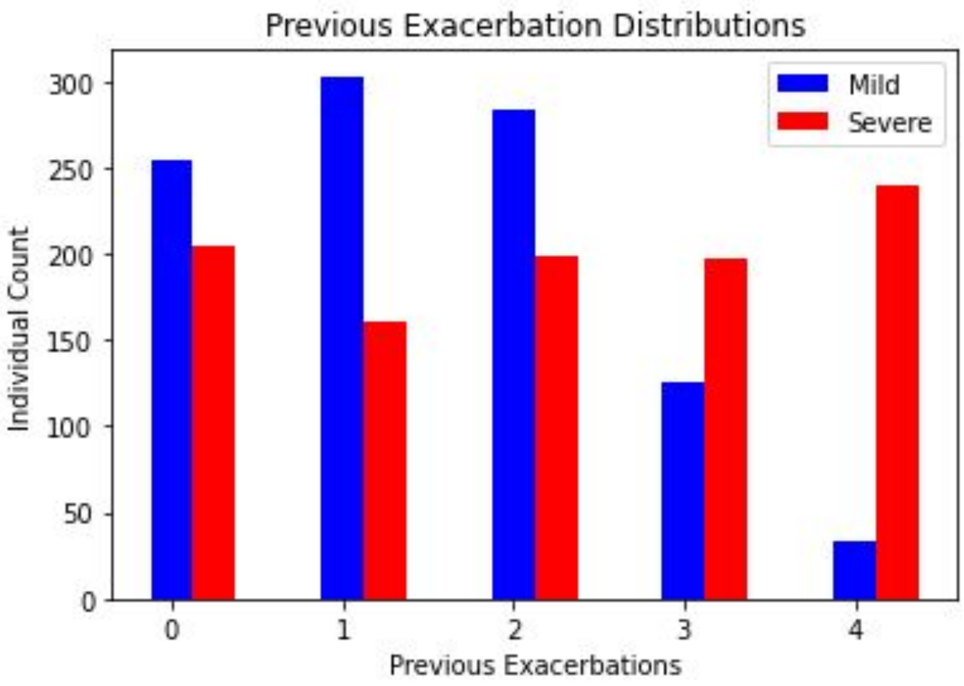
- Would expect mild to have younger individuals



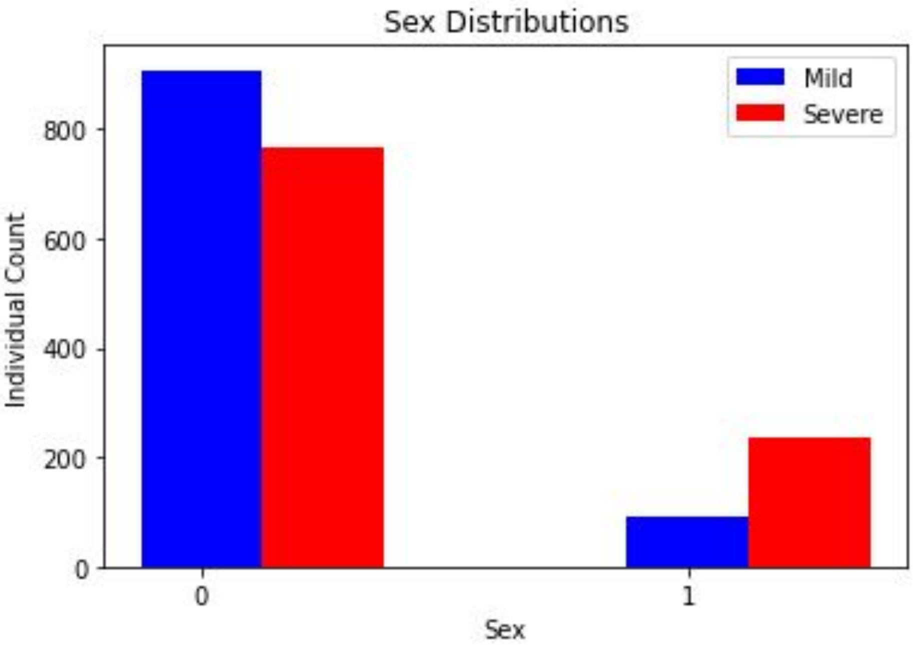
Previous Exacerbations

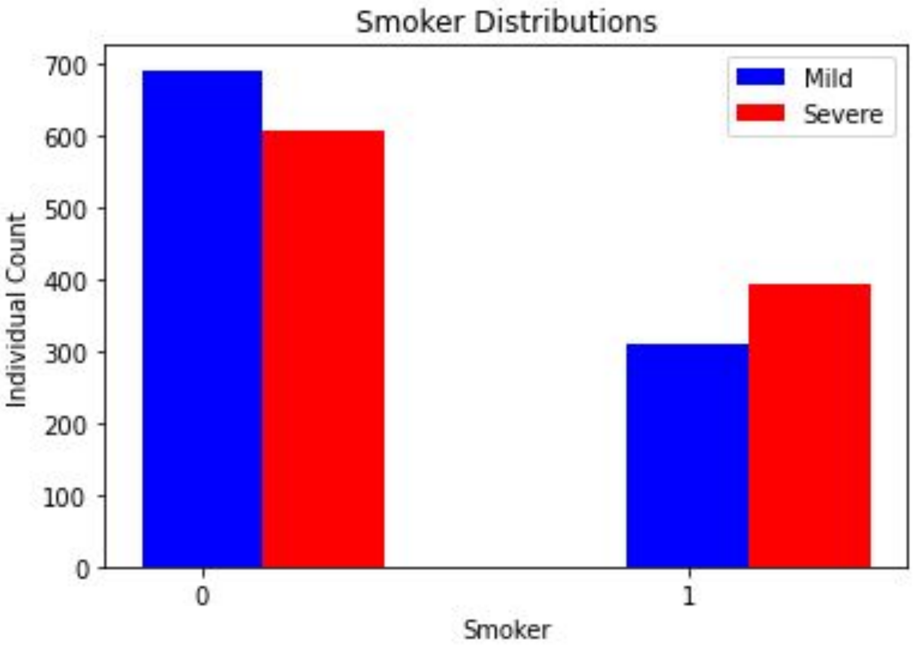
- Severe had a mean of 2.2 and standard deviation of 2.1

- Mild had a mean of 1.3 and a standard deviation of 1.2

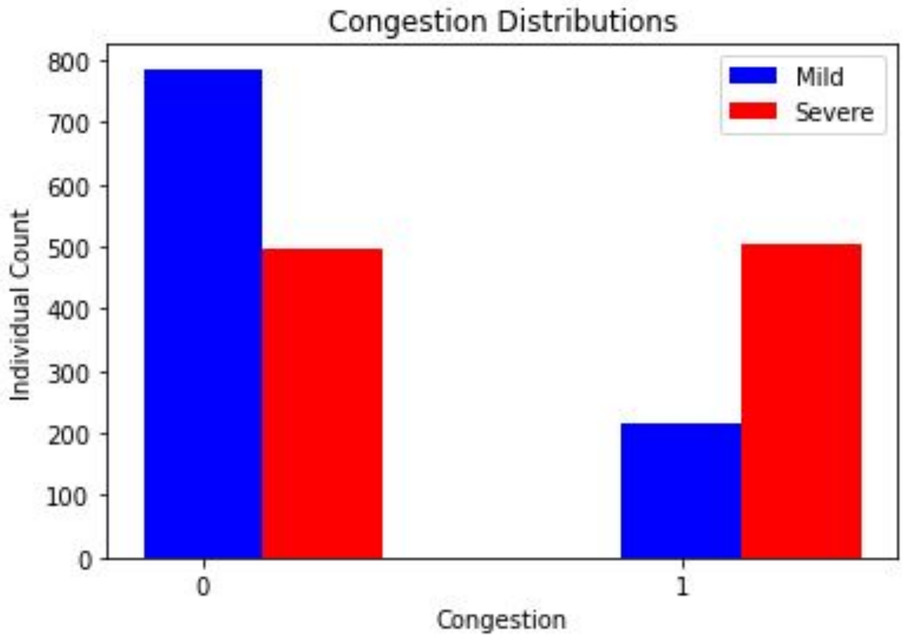


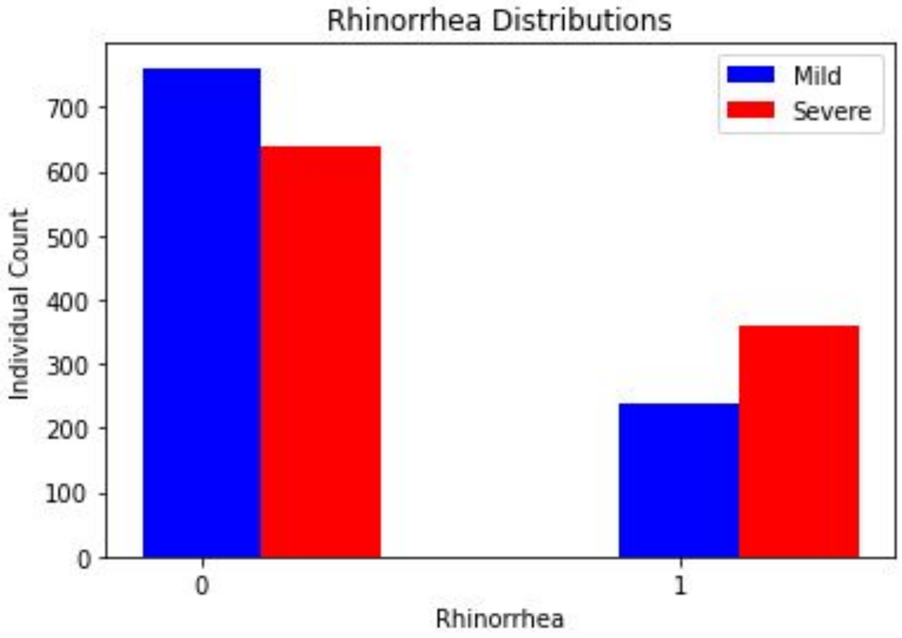
Sex and Smoking



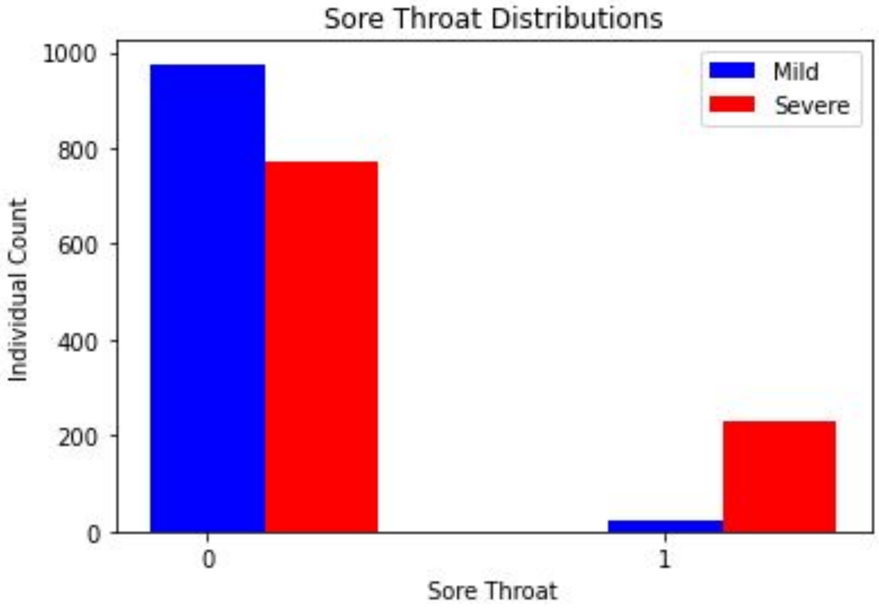


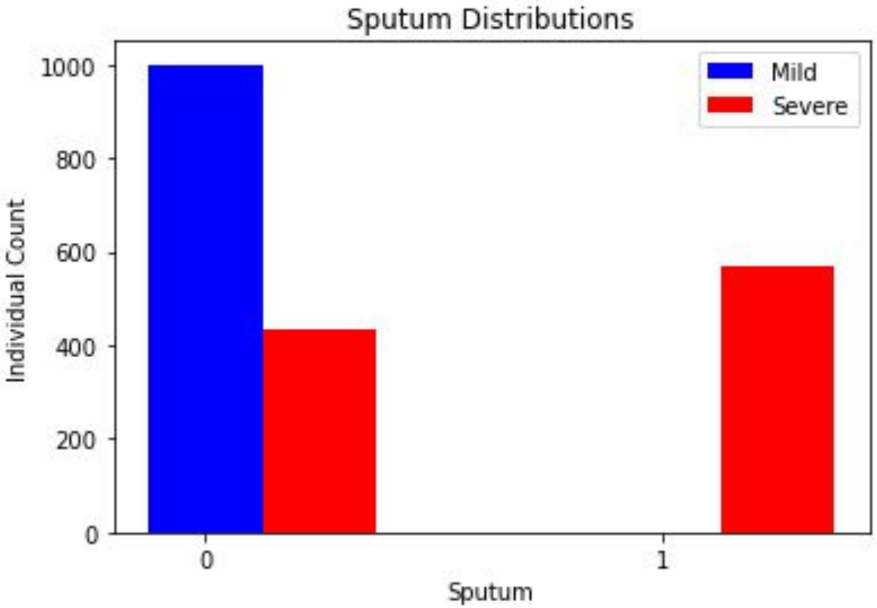
Congestion and Rhinorrhea



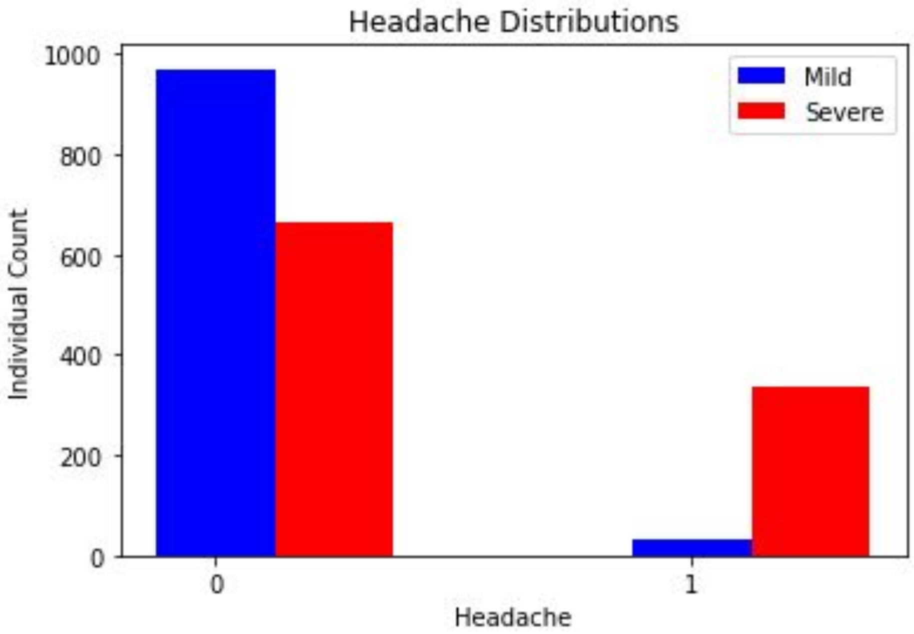


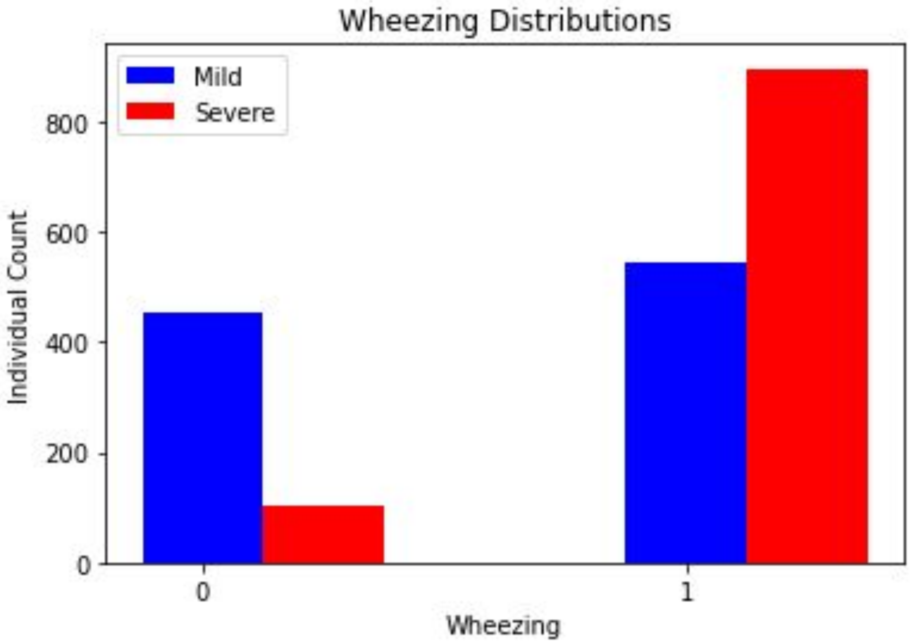
Sore Throat and Sputum



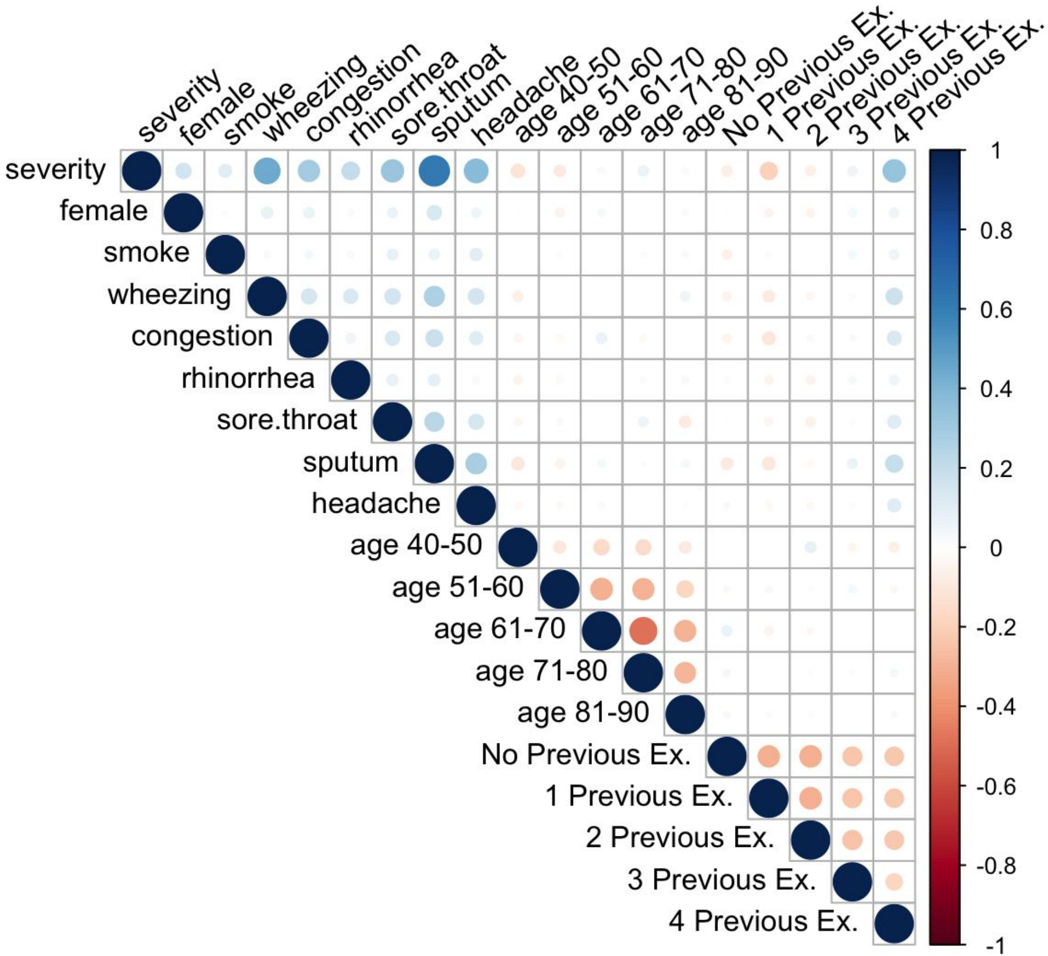


Headache and Wheezing





Feature Correlation



<u>**Notice:**</u>

- **Sputum** and **wheezing** are highly correlated with severity
- **1 previous exacerbation** is negatively correlated with severity
- **4 previous exacerbations** is positively correlated with severity
- **Age** does not seem to be correlated with severity

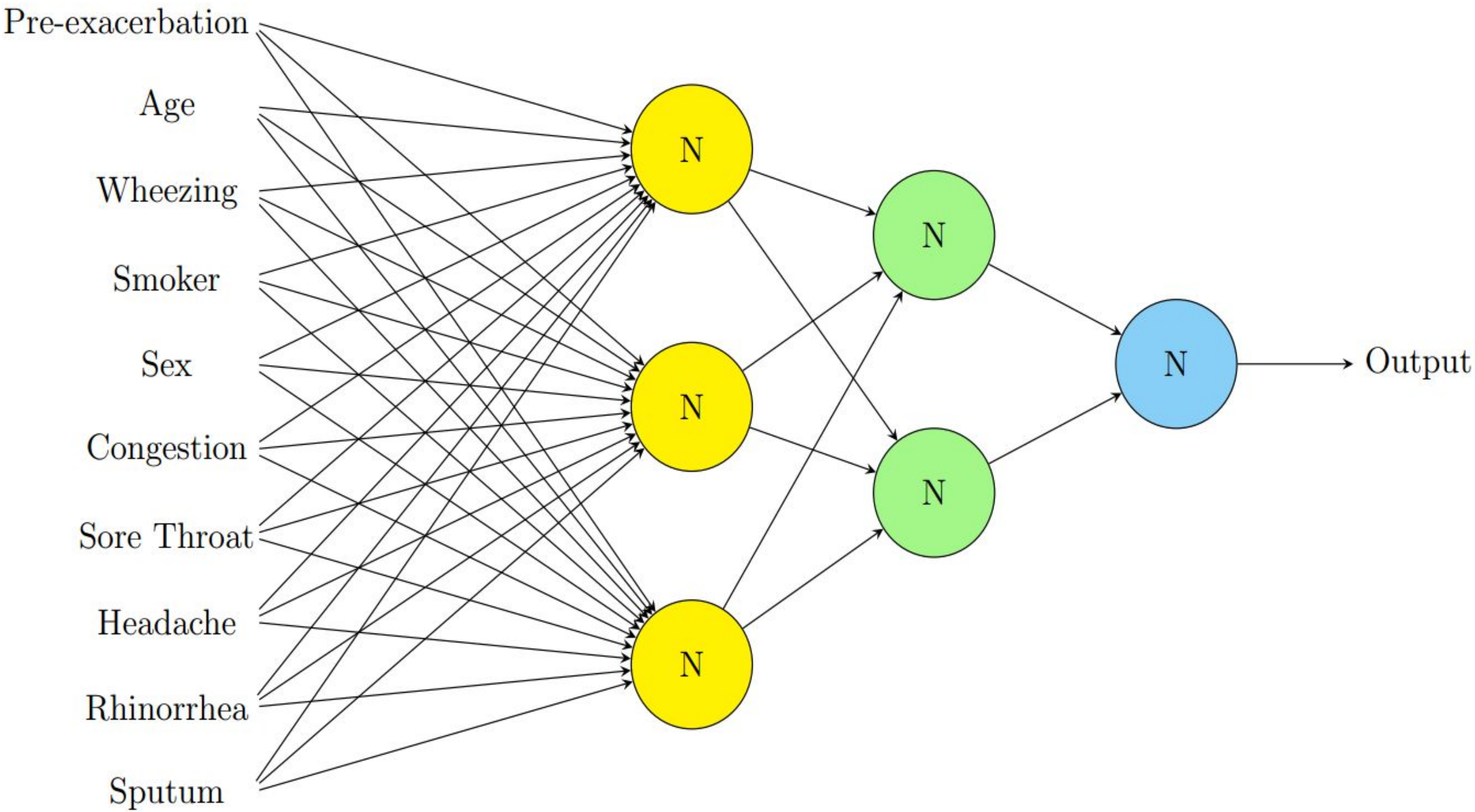
Do We Have Realistic Patients?

- We ran some data checks:
    - 86 / 1000 patients were severe and younger than 50 years old
    - 117 / 1000 patients were mild and had 3 or more binary symptoms
        - (wheezing, smoking, congestion, sore throat, headache, rhinorrhea, sputum)
    - 13 / 1000 patients were mild had 4 or more binary symptoms
- We should **not** have a negative correlation between severity and any number of previous exacerbation
- Older ages should be more highly correlated with severity

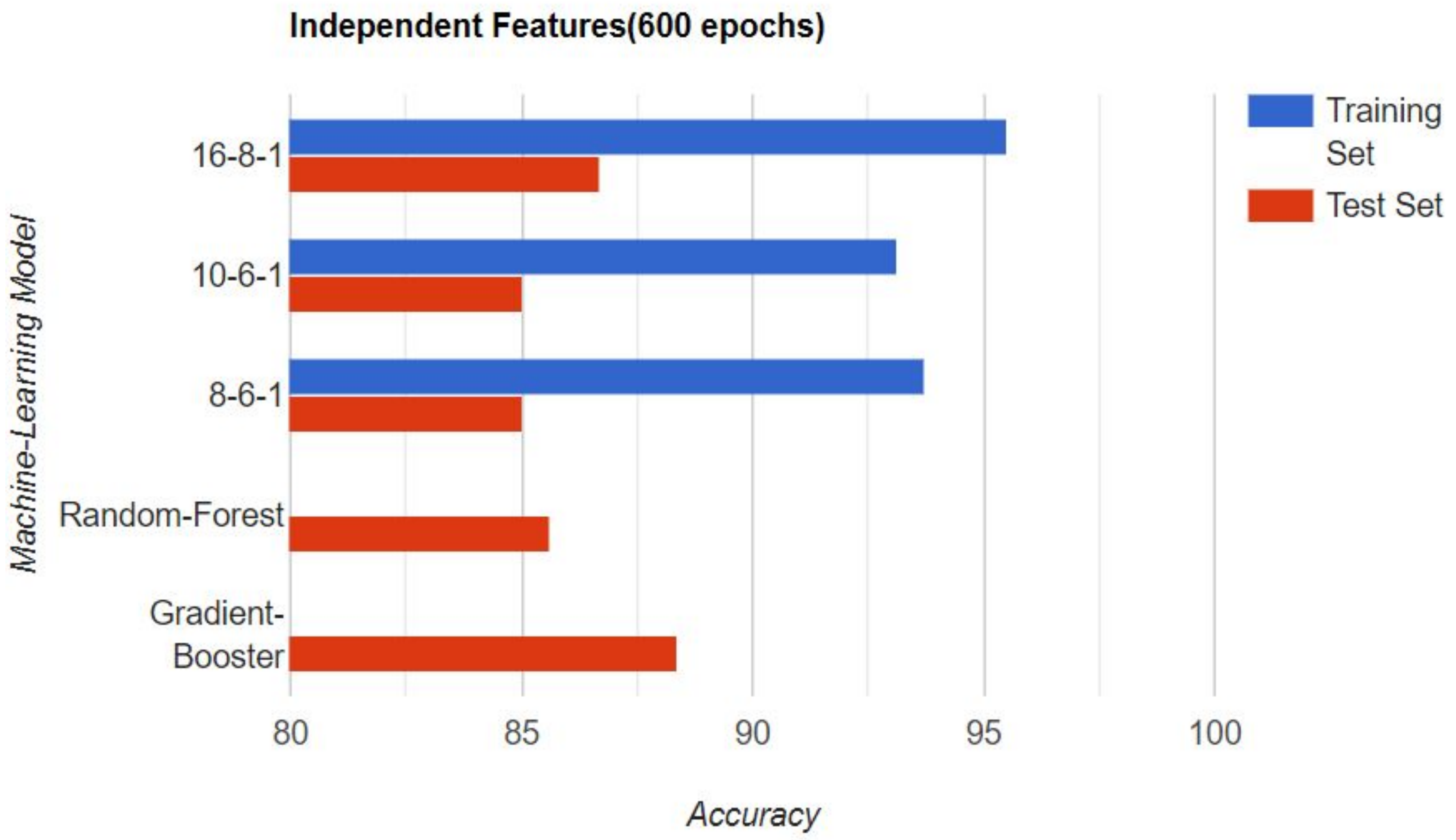
Implementing a Neural Network

- We implemented a neural network using python:
  - TensorFlow and Keras

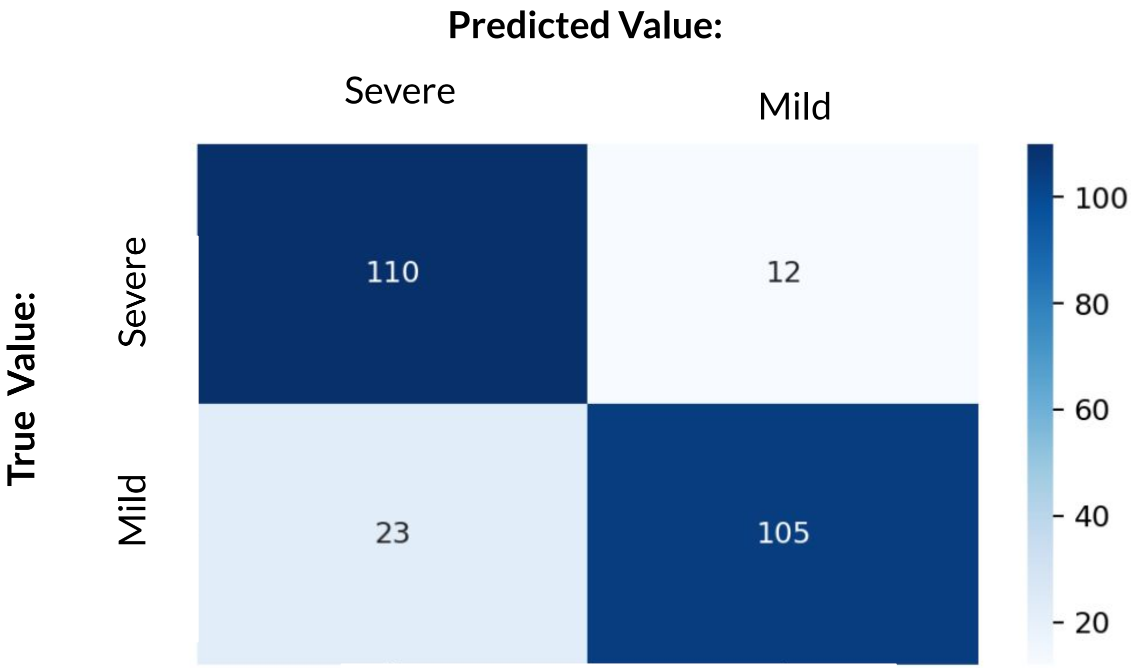
- We experimented with the architecture of the network to see the impact on prediction accuracy and the confusion matrix


Pre-exacerbation
Age
Wheezing
Smoker
Sex
Congestion
Sore Throat
Headache
Rhinorrhea
Sputum
N
N
N
N
N
N
Output

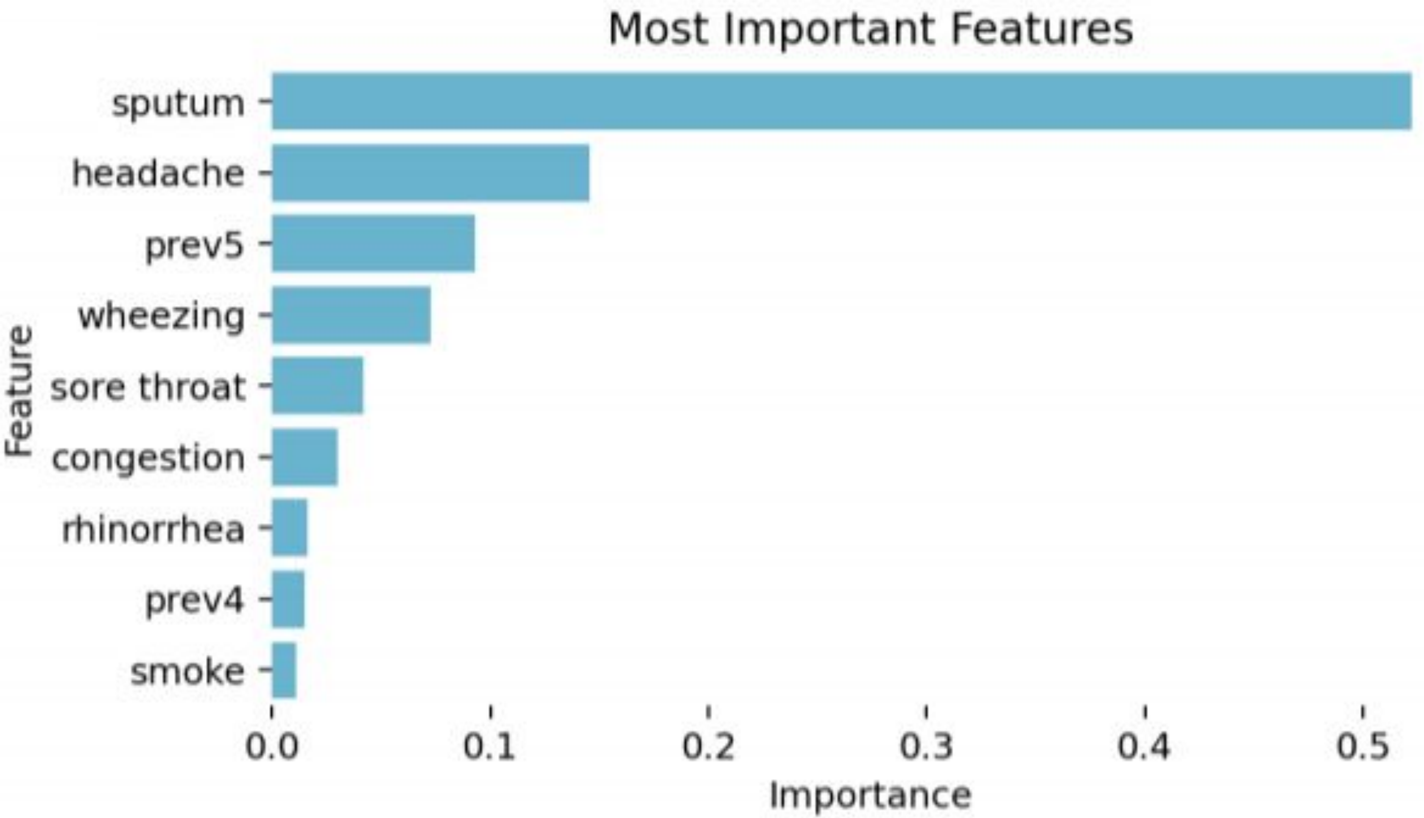
Model Comparison- Accuracy



Model Comparison - Confusion Matrix



Important Features



Generating Patient Data - Multivariate Method

- No longer assuming complete independence of all features.
- Uses a branching algorithm to indicate influence
- Utilizes a multivariate normal distribution for correlated features
- Use distributions based on the Simple Method’s parameters

Generating Patient Data - Multivariate Model

Age

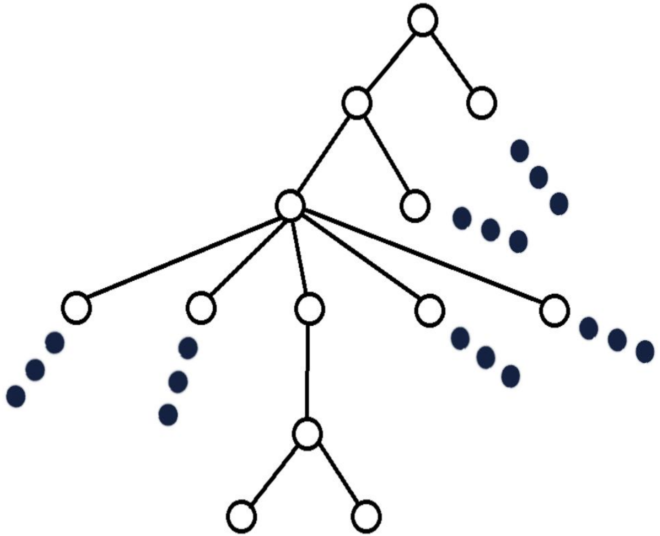
Sex

Smoking Status

Previous Number of Exacerbations

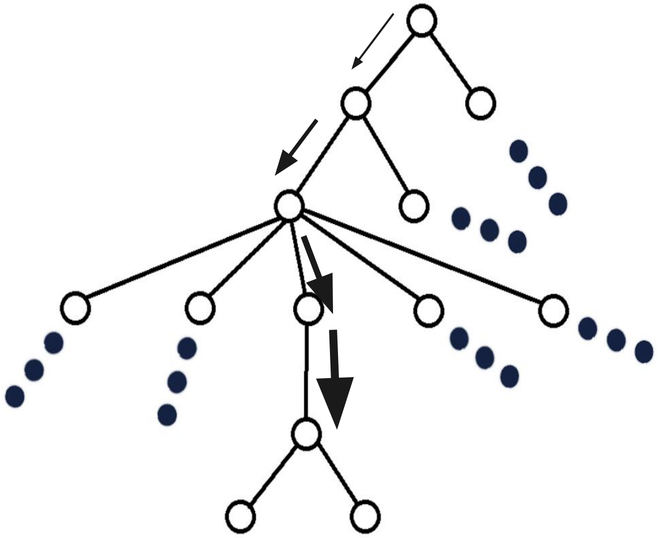
Correlated Features

Severity



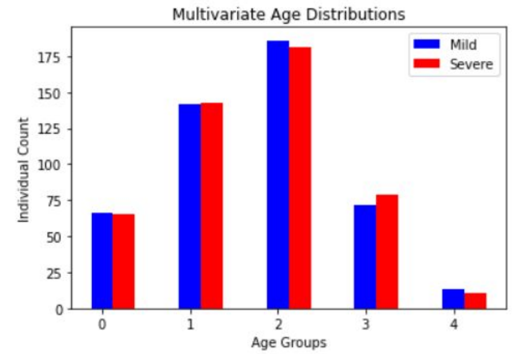
Generating Patient Data - Multivariate Method

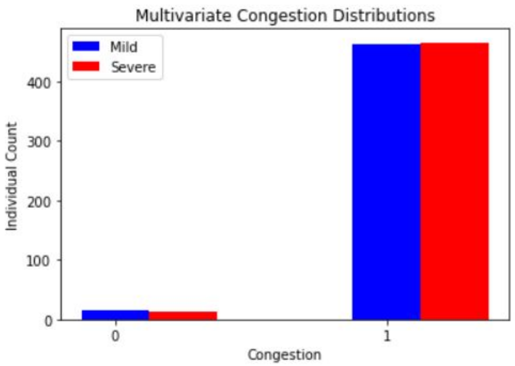
- Impact factor gives weights to the Simple Method's parameters
- Multivariate Normal Distribution takes into account influence on correlated values
- Branching Approach allows cumulative impact depending on path



Generating Patient Data - Results

- Main limitation current available data
- Framework is constructed
- This can be expanded and adapted on features of interest
- Algorithm is transferable to patient identification
- Conclusion: Not quite ready to replace Simple Method, but it has a lot of potential





Possible Expansions (Allergens and Pollutants)

- While there has been a link shown between allergens and exacerbation rates in patients with COPD the addition of this into a predictive model would be difficult given the individual nature of allergies.
- Pollutants on the other hand are a more generalizable trigger for exacerbation.
- There are six primary pollutants tracked O_3(Ozone), SO_2(Sulfur Dioxide), NO_2(Nitrogen Dioxide), CO(Carbon Monoxide), $PM_{2.5}$(Particulate Matter < 2.5 micrometers), and PM_{10}(Particulate Matter < 10 micrometers).
- The literature found that the primary pollutants were $PM_{2.5}$, PM_{10}, NO_2, and O_3, while there was no real correlation with SO_2 and CO.

Possible Expansions (Allergens and Pollutants)

- Given that there is a link between pollutants and rate of exacerbation we can work out a couple of possible approaches to how to implement this into the model and the app.
- The simplest way would be to pull location data from the clients device and acquire the air quality index (AQI) and assess a weight based off of the AQI score and a percentage based off of data.
- The more difficult method, however potentially, more accurate would be to pull pollutants individually and assign weight to each of them based off of the percentages found in the literature.
- One advantage to this is that users of the app would not have to worry about entering AQI data, the app would handle it for them.

Conclusion

- Completed an extensive literature review
- Identified a main collection of features that are indicative of severity in COPD
- Improved the current method of generating patients
- Constructed a neural network to output severity
- Constructed base framework of a more complex method of generating patients
- Began expansion of considering integration of additional features without patient input

Thank you!

References:

Rohde, G., et al. "Respiratory viruses in exacerbations of chronic obstructive pulmonary disease requiring hospitalisation: a case-control study." *Thorax* 58.1 (2003): 37-42.

Hong, Woo Suk, Adrian Daniel Haimovich, and R. Andrew Taylor. "Predicting hospital admission at emergency department triage using machine learning." *PloS one* 13.7 (2018): e0201016.

Leidy, Nancy K., et al. "Standardizing measurement of chronic obstructive pulmonary disease exacerbations: reliability and validity of a patient-reported diary." *American journal of respiratory and critical care medicine* 183.3 (2011): 323-329.

2021 GOLD report