

Predicting Death with COVID-19

Sunna Jo

Disclaimer

The results of this project should NOT be used to guide clinical decision making, personal decisions regarding seeking medical care or treatment, and/or for any reason(s) other than for educational purposes.

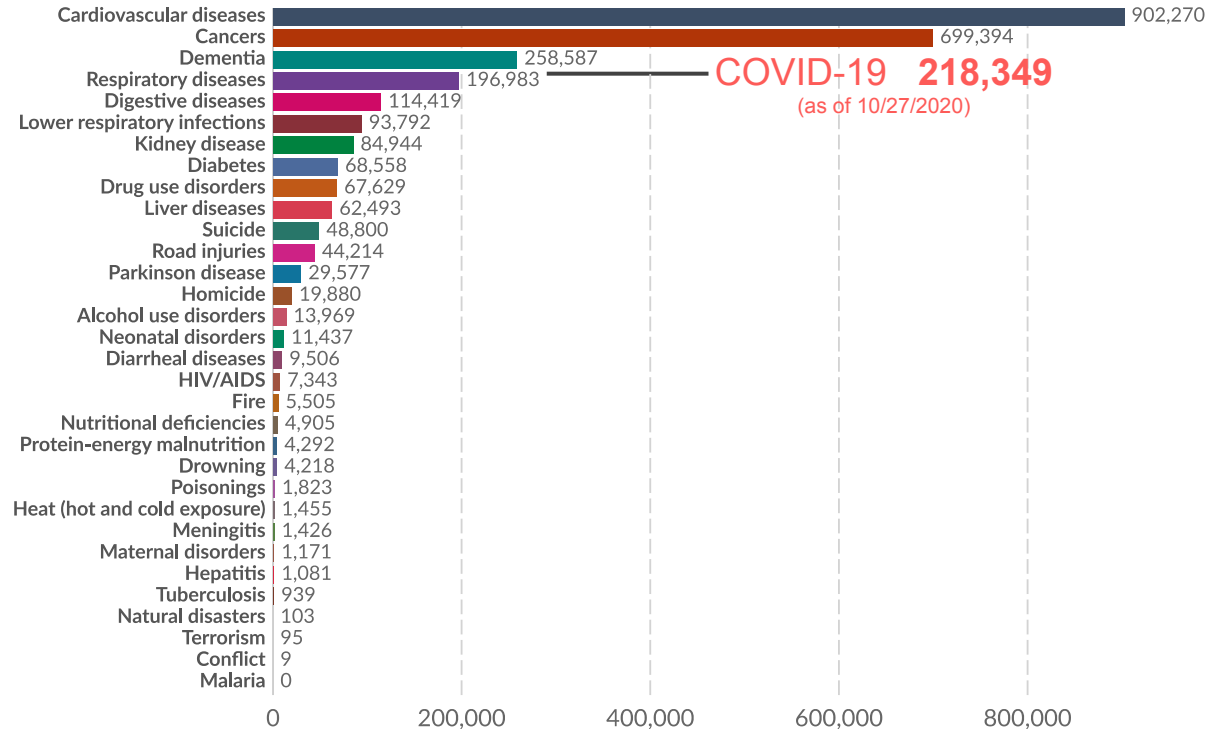
Overview

- Background & Objective
- Process
- Final Model
- Conclusions
- Application
- Future Work

How bad is it?

Number of deaths by cause, United States, 2017

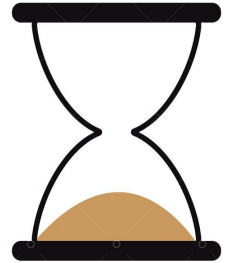
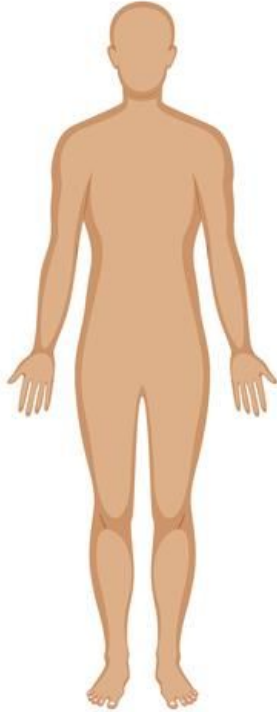
Our World
in Data



Source: IHME, Global Burden of Disease

OurWorldInData.org/causes-of-death • CC BY

What contributes to death?



**Exploratory
Data Analysis**
Data set variations
Visualizations

**Feature Engineering &
Selection**
40 features
Correlations
Feature analysis



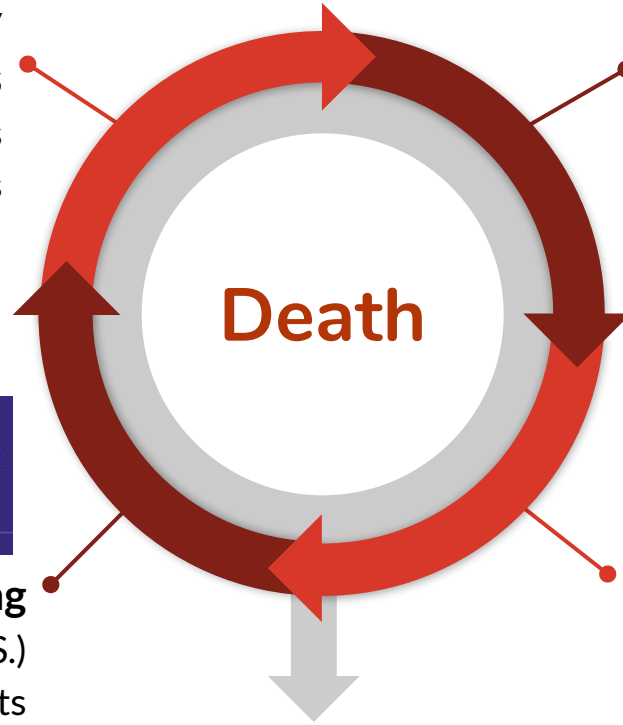
Data Collection & Cleaning
Patient-level, national-level (U.S.)
3 data sets
671,435 data points
4/1/20-9/28/20



Death

Model Testing
Metric: Recall
Different models →→ decision
tree-based methods
Hyperparameter tuning

Final Model
XGBoost
6 features

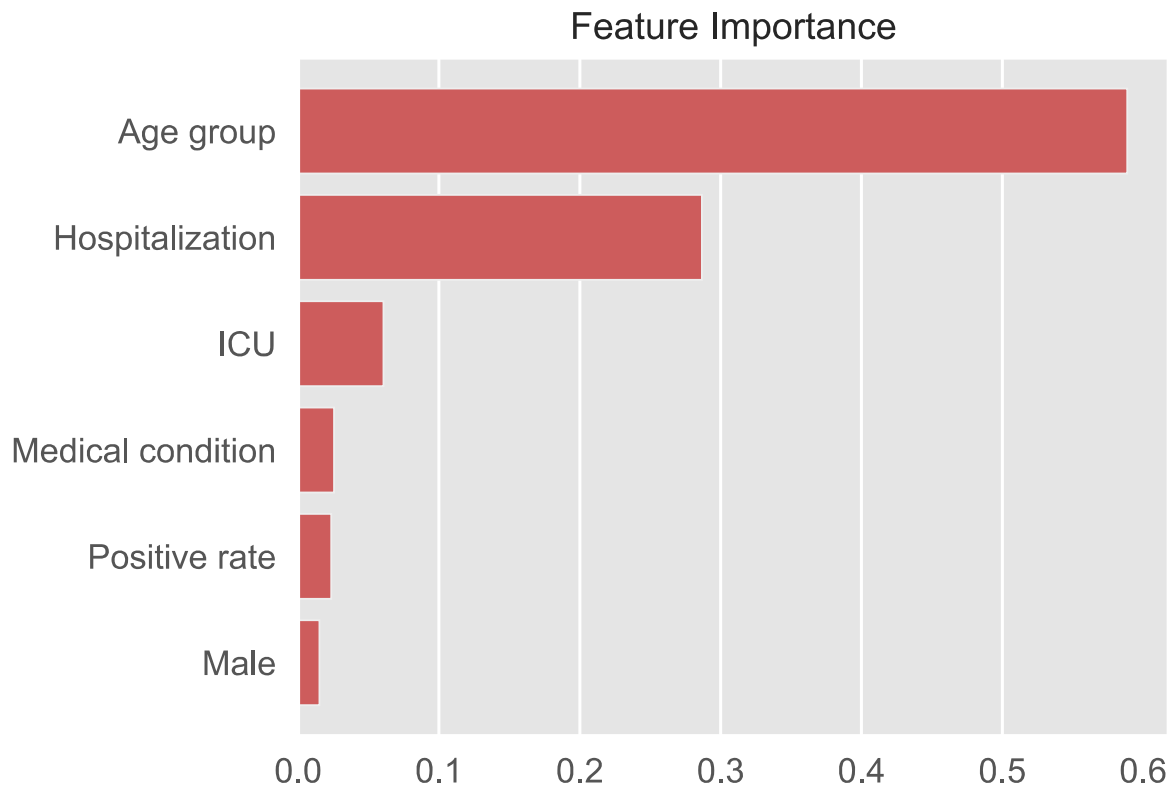


Final Model: XGBoost

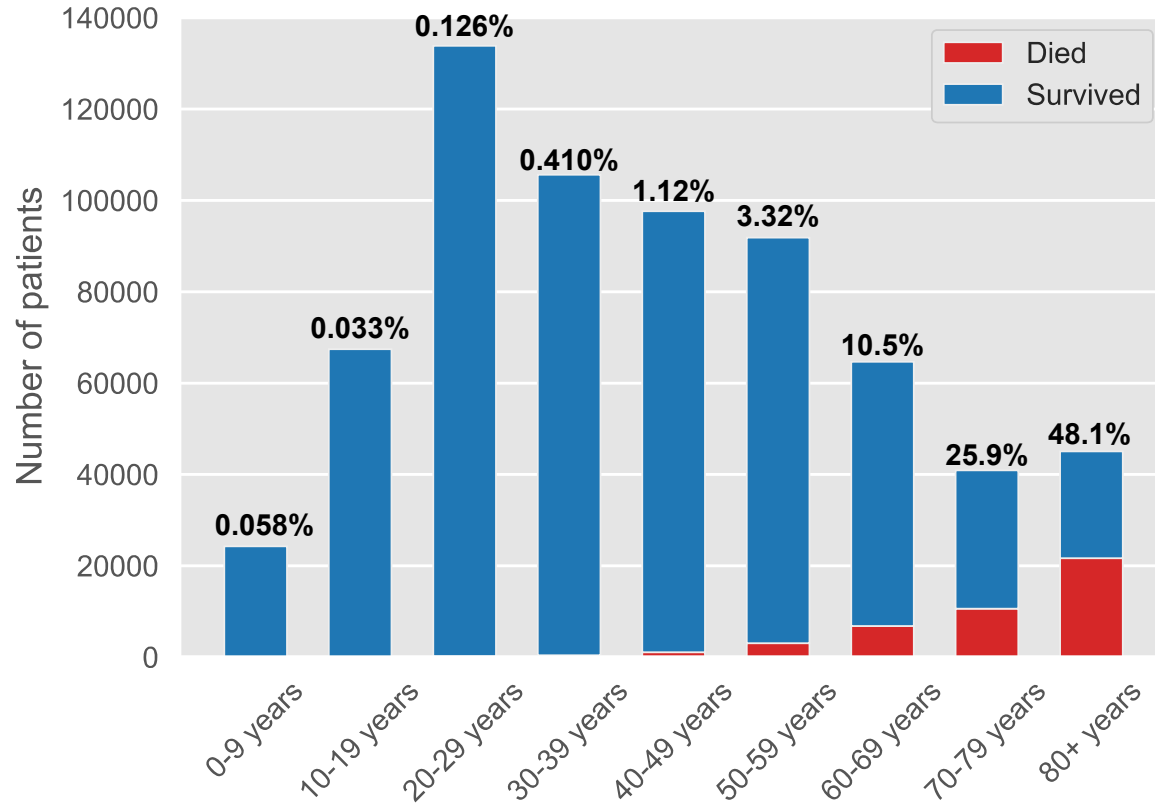
Metric	Test	Train
<i>Recall</i>	0.939	0.946
<i>Precision</i>	0.416	0.419
<i>F-beta (beta=2)</i>	0.756	0.756

		Confusion Matrix	
		No Death	Death
Actual	No Death	113855	11661
	Death	477	8294
		No Death	Death
		Predicted	

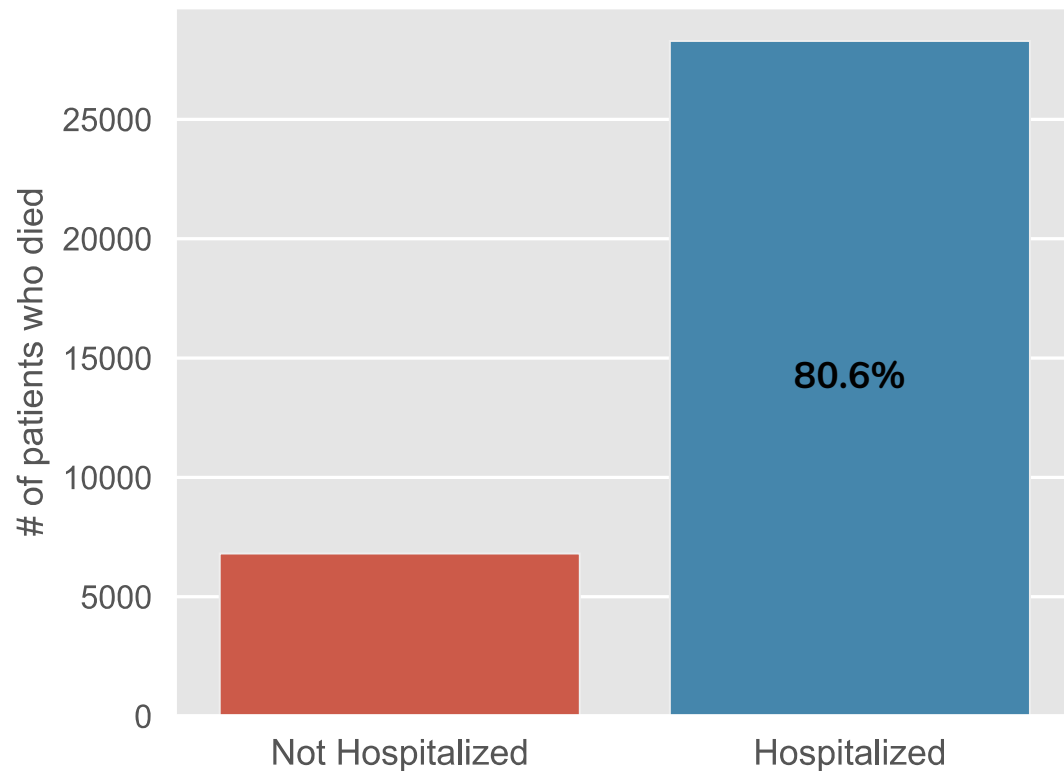
Conclusions



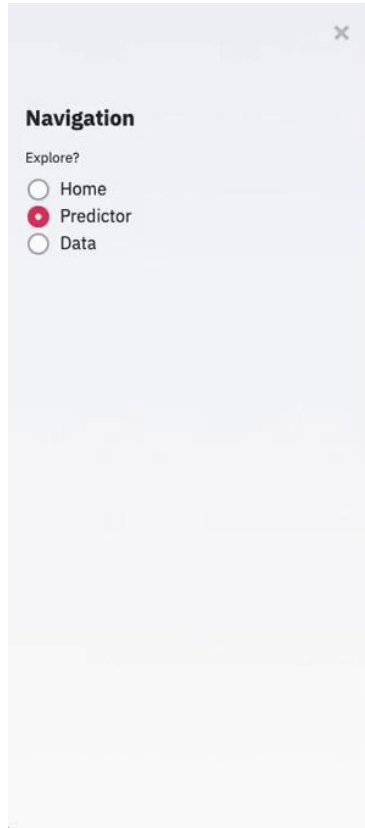
Outcome by Age Group



Hospitalizations Among Deaths



Application



Does the patient have an underlying medical condition?

- ☒ No
☐ Yes

What is the current positivity rate? (as a percentage)

6.30

- +

Look it up

What is the patient's biological sex?

- ☐ Female
☒ Male
☐ Other

Predict

The patient has a higher risk of death

 RUNNING... Stop 

Application

Navigation

Explore?

- ☐ Home
- ☐ Predictor
- ☒ Data

Interactive Data

Data Sources

CDC Public Use Surveillance Data

Choose variables

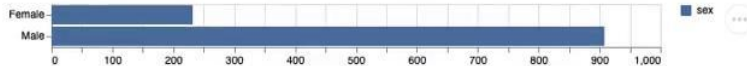
Choose an option

Compare patients who died to those who did not:

☒ Death

Choose variable

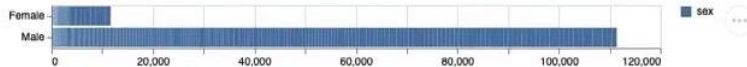
Sex



☒ No death

Choose variable

Sex



Future Work



- Improve precision
 - More data
 - More features
- Geography
 - State/local level
- Application

The background of the slide features a 3D rendering of several COVID-19 virus particles. These particles are depicted as spherical, grey, textured structures with numerous red, crown-like spikes protruding from their surfaces. Some particles also show small yellow hexagonal shapes. The particles are scattered across the frame, with some in sharp focus and others blurred in the background, creating a sense of depth.

Questions?

Thank you!

Appendix

Metric	Test	Train
<i>Accuracy</i>	0.910	0.910
<i>Recall</i>	0.939	0.946
<i>Precision</i>	0.416	0.419
<i>F1</i>	0.580	0.576
<i>F-beta (beta=2)</i>	0.756	0.756

Classification report

	precision	recall	f1-score	support
0	1.00	0.91	0.95	125516
1	0.42	0.95	0.58	8771
accuracy			0.91	134287
macro avg	0.71	0.93	0.76	134287
weighted avg	0.96	0.91	0.93	134287