

# Seasonal Flu Vaccine Prediction

Marcelo Scatena  
Anthony Warren  
Piotr Czolpik

U.S. Department of Health & Human Services

# Overview

- Business Problem
- Data & Methods
- Models
- Final Model Evaluation
- Conclusion



# Business Understanding

- Provide recommendations about what could be done to increase the number of vaccinated people
- Find out which characteristics lead people to be more or less prone to have taken the seasonal flu vaccine
- Train and run a model that can correctly identify the likelihood that a randomly chosen individual received the seasonal flu vaccine or not



# Data

- 2009 National H1N1 Flu survey
- Over 26000 samples
- 35 Features



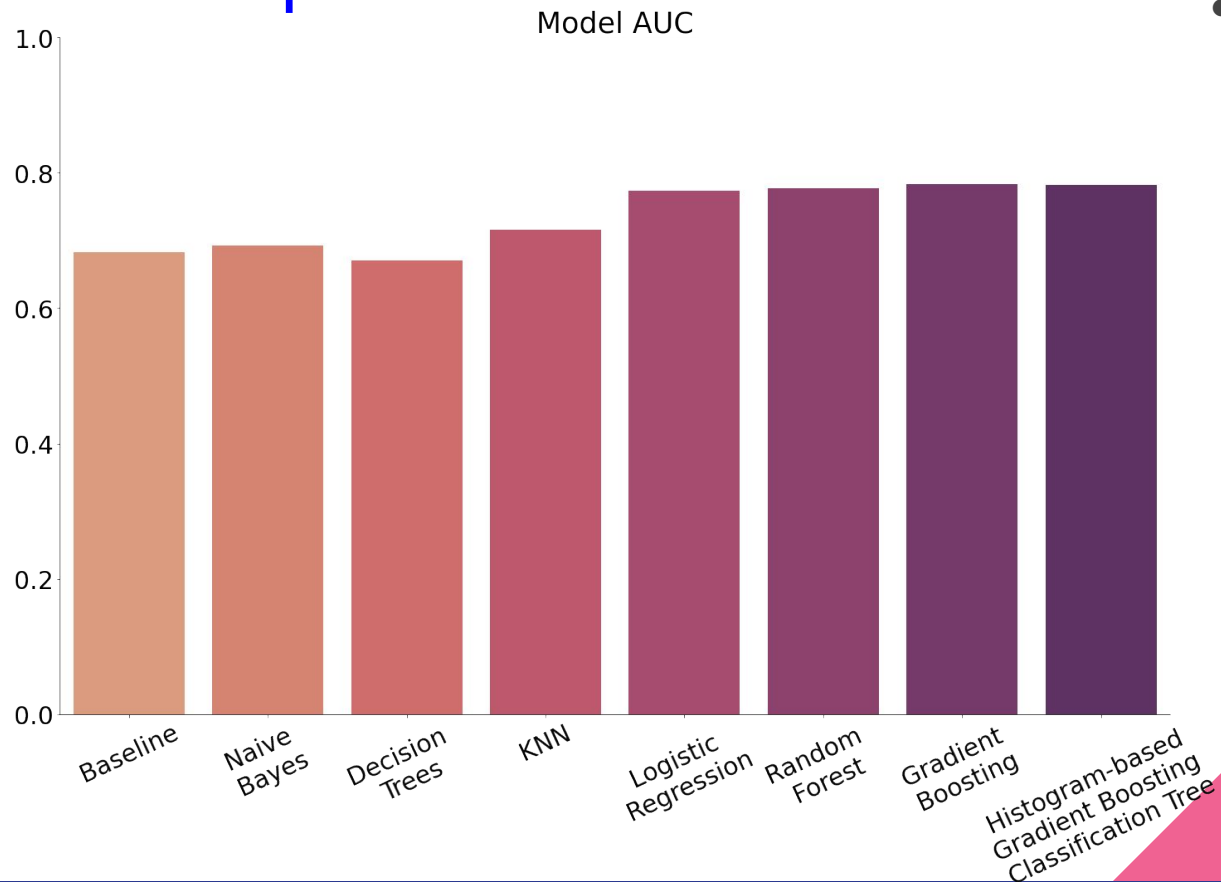
# Summary

- Final Model Accuracy & Precision
  - Model Performance & Meaningful Metric
  - Precision: ~79%
  - Accuracy: ~76%
  - Over 10% increase in accuracy and precision from our baseline
- Most important features to watch include opinions regarding the seasonal flu and age groups (Anthony to remove)
- Recommendations
  - Personal Opinions Matter
  - Age Matters
  - Professional Opinions Matter



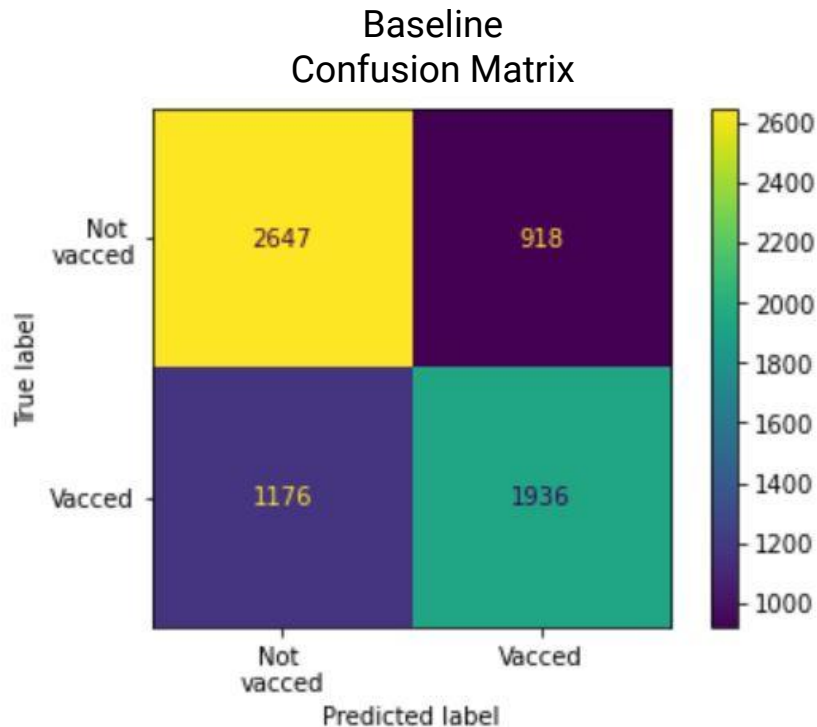
# Model Comparisons

- Metric to evaluate different models



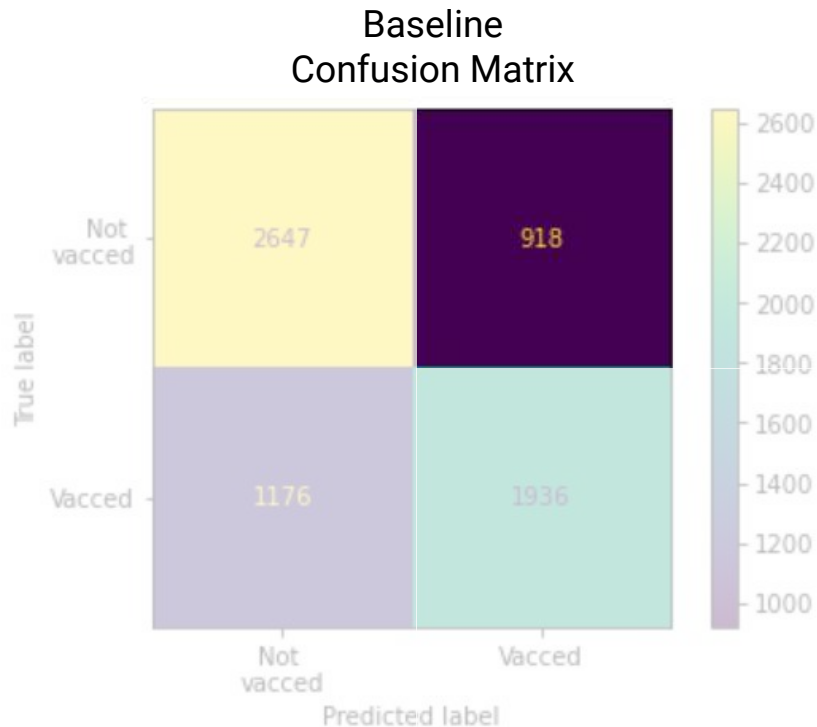
# Baseline Model

- We chose precision over other metrics for our models' evaluations because we'd rather focus on the individuals who have not received the vaccine but the model predicts they did (false positives)
- Accuracy & Precision
  - Model Performance & Meaningful Metric
  - Precision: ~67%; Accuracy: ~68%



# Baseline Model

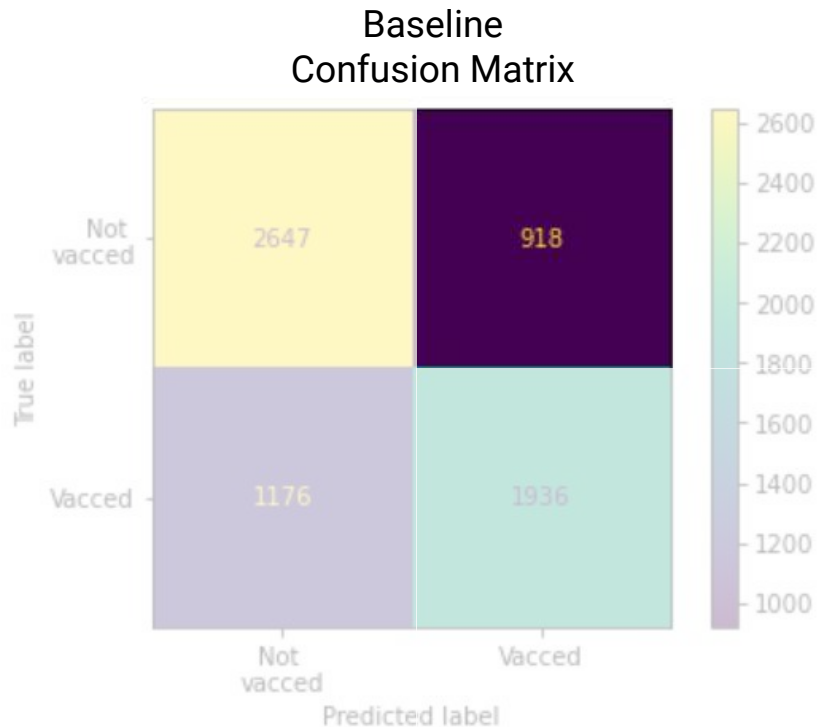
- We chose precision over other metrics for our models' evaluations because we'd rather focus on the individuals who have not received the vaccine but the model predicts they did (false positives)
- Accuracy & Precision
  - Model Performance & Meaningful Metric
  - Precision: ~67%; Accuracy: ~68%



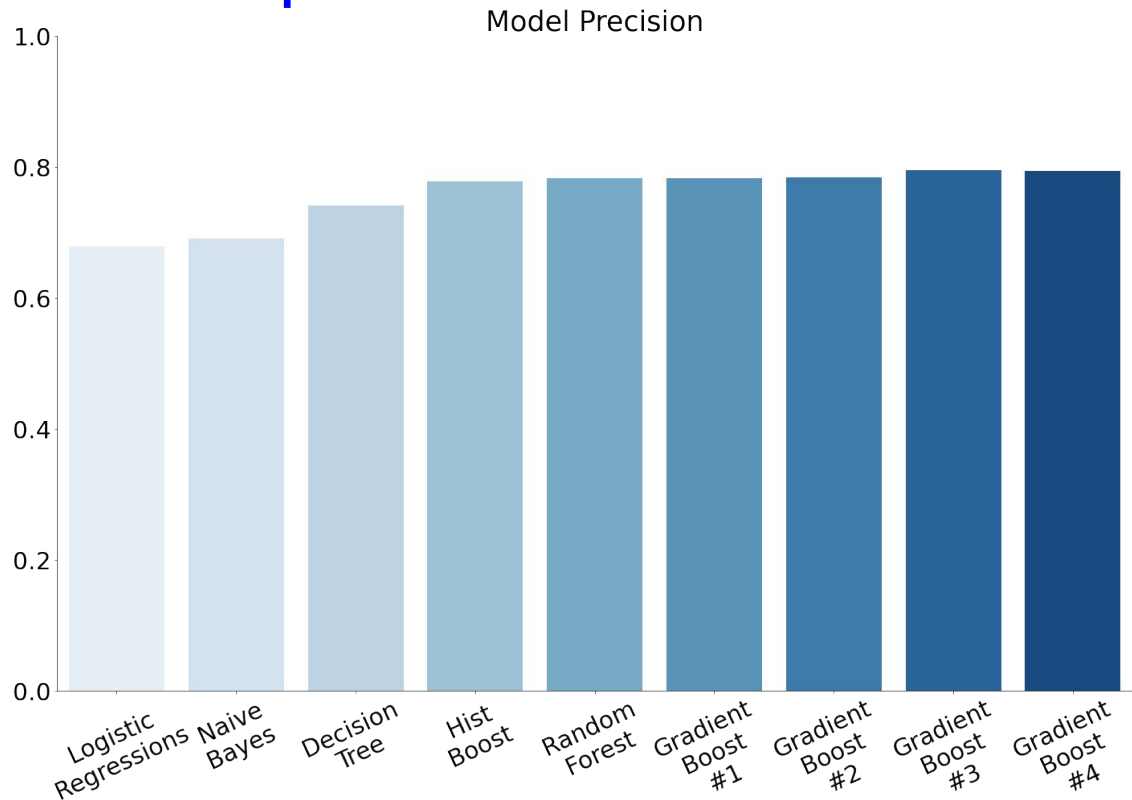


# Baseline Model

- We chose precision over other metrics for our models' evaluations because we'd rather focus on the individuals who have not received the vaccine but the model predicts they did (false positives)
- Accuracy & Precision
  - Model Performance & Meaningful Metric
  - Precision: ~67%; Accuracy: ~68%



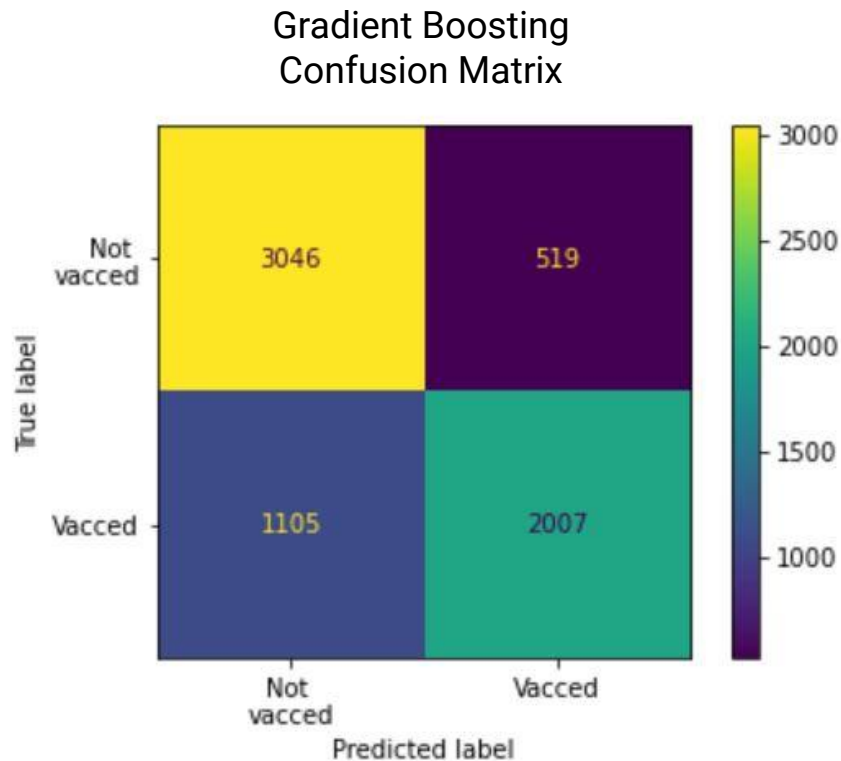
# Model Comparisons



- Baseline Model:  
~68%
- Tuned Gradient Boosting:  
~79%

# Final Model

- Precision: ~79%
- Accuracy: ~75%
- Over 17% increase in precision
- Over 10% increase in accuracy
- Over 40% decrease in False Positives



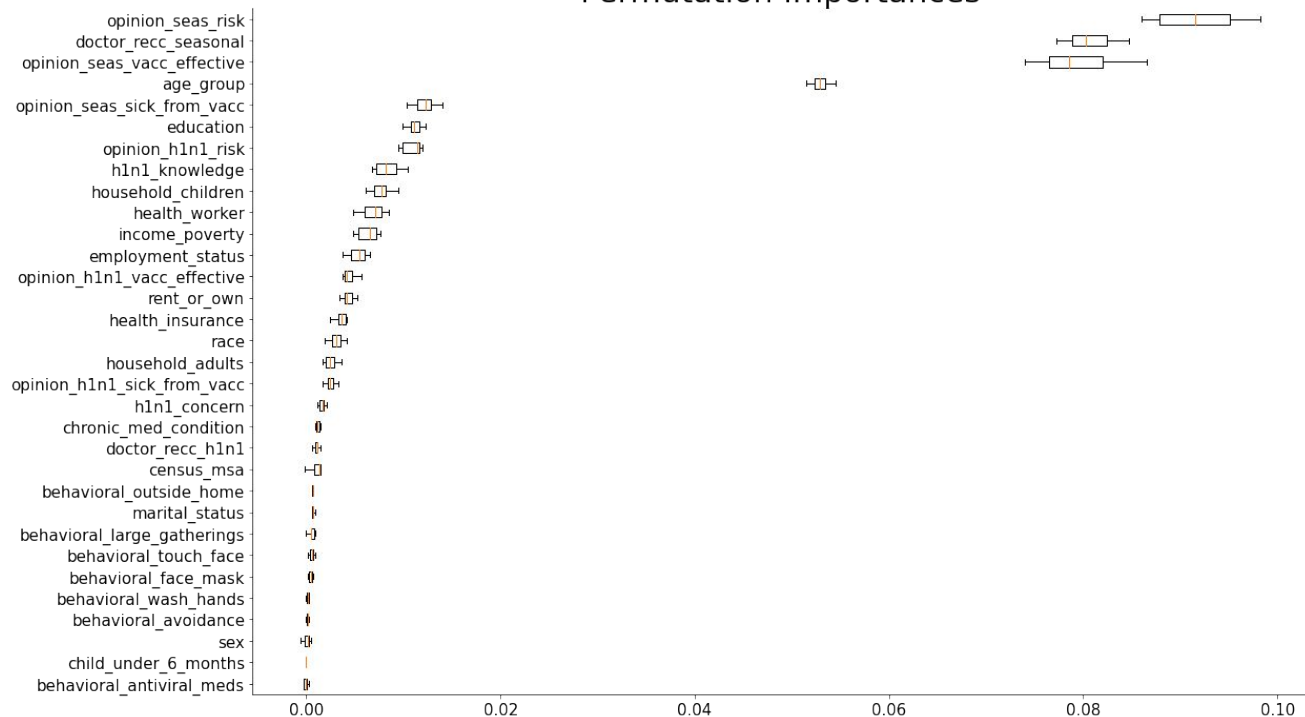
# Feature Permutation Importance

- Ranking metric that allows us to determine the features/variables that are the most impactful to the accuracy of our classifier model
- It does not predict the likelihood of vaccination, just the impact it has to our model's accuracy

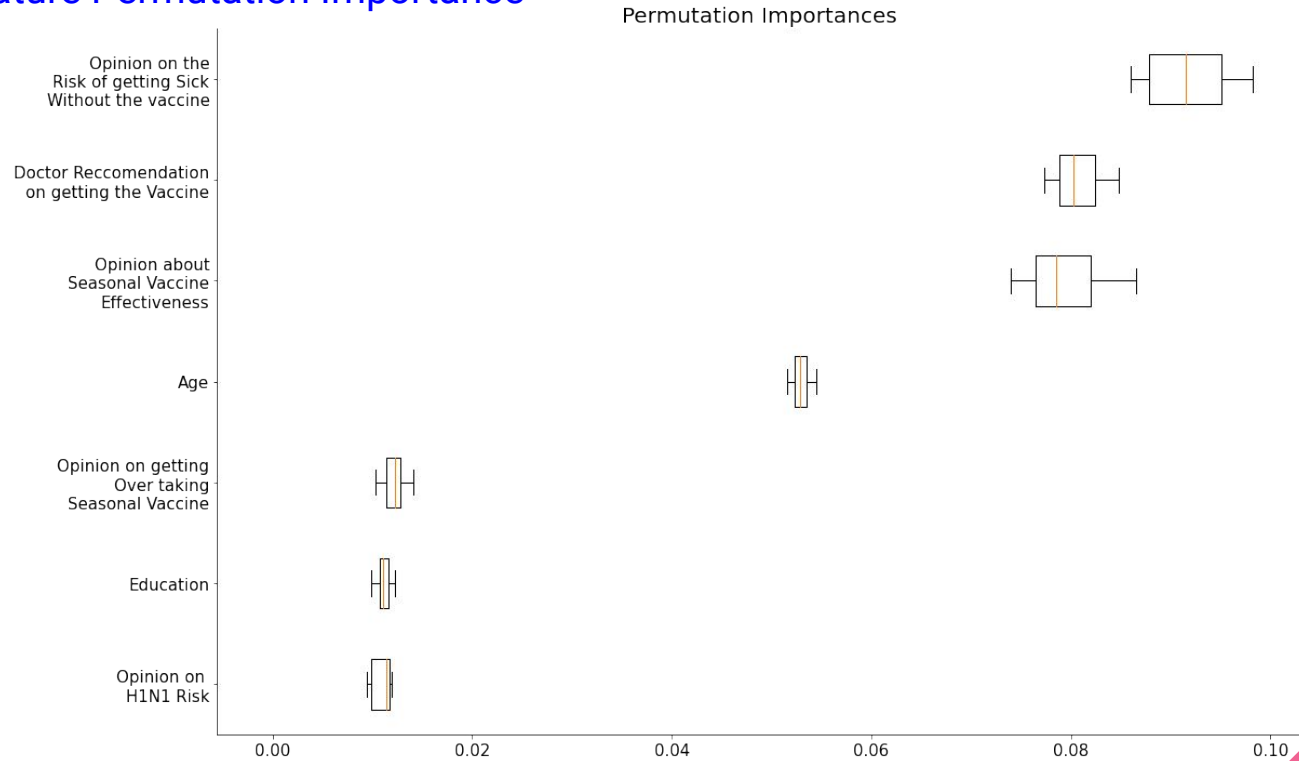


# Feature Permutation Importance

Permutation Importances

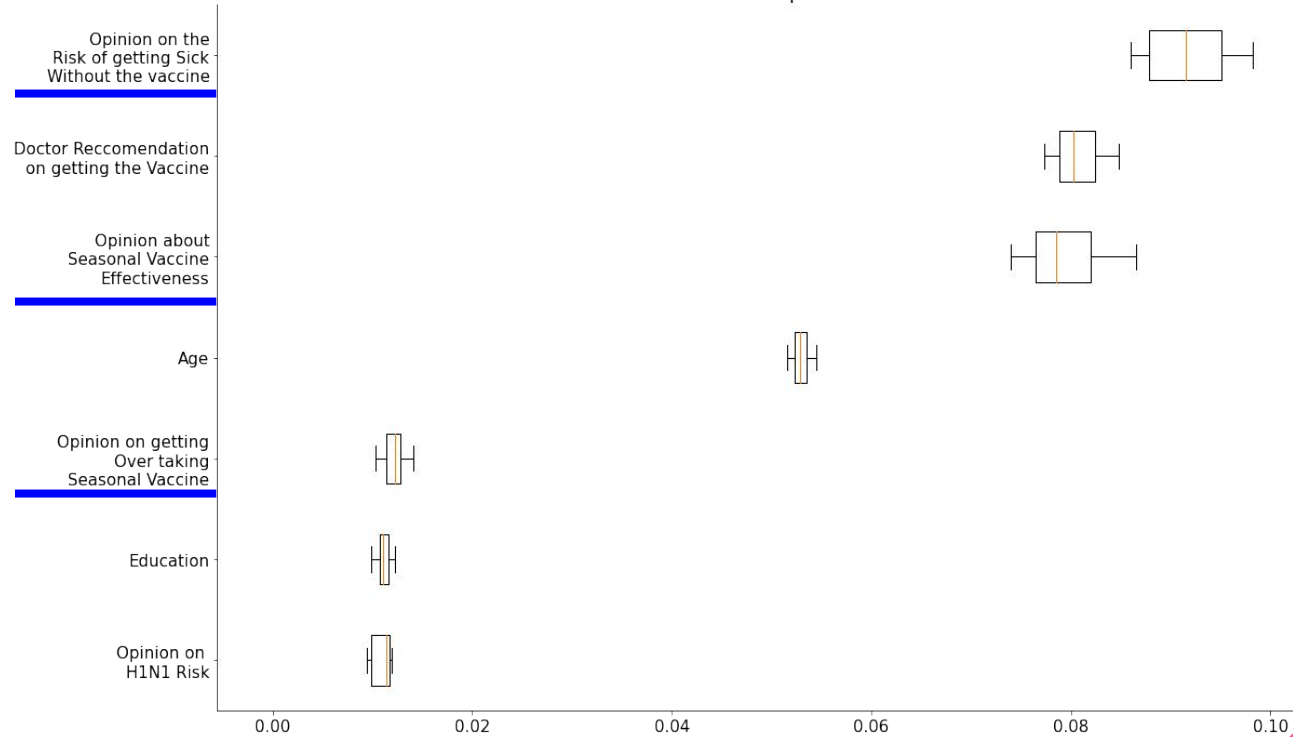


# Feature Permutation Importance

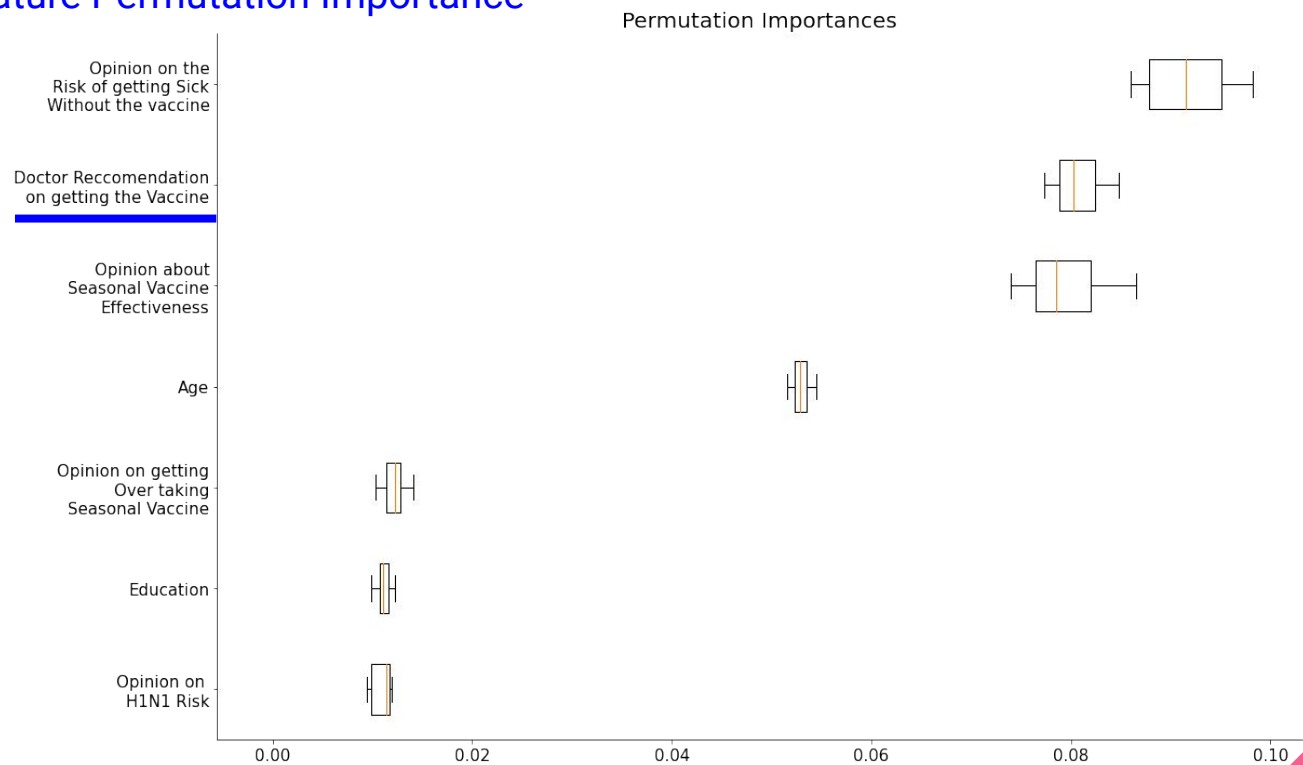


# Feature Permutation Importance

Permutation Importances

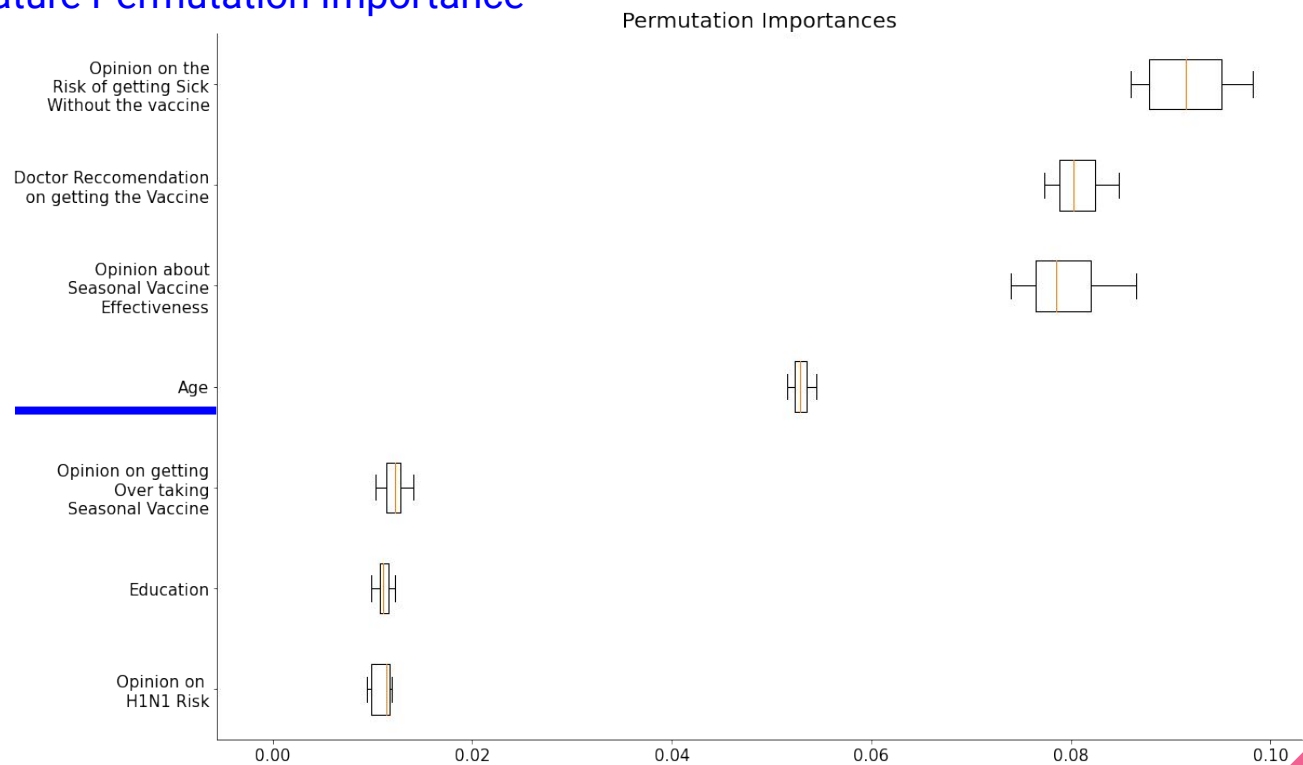


# Feature Permutation Importance

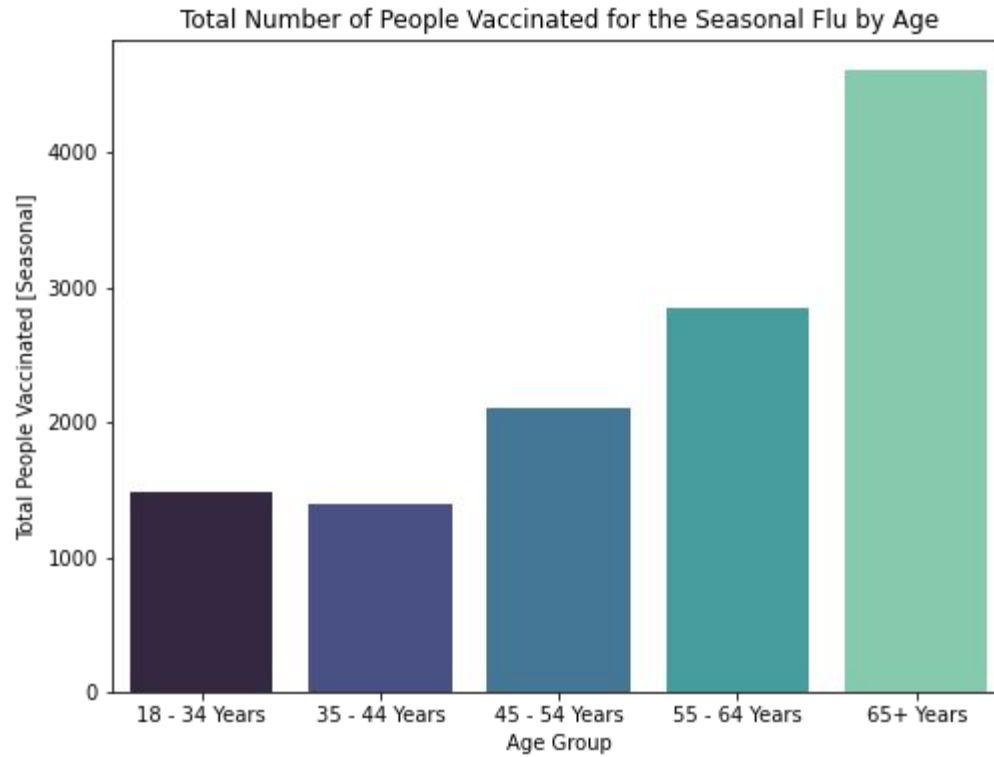




# Feature Permutation Importance



# Seasonal Flu Vaccination Rates by Age



# Recommendations

- Personal Opinions Matter
  - Raise awareness of dangers of the seasonal flu via new campaigns to everyone
- Age Matters
  - A more specific campaign towards folks aged below their forties could be more beneficial as they are the least likely age groups to be vaccinated
- Professional Opinions Matter
  - Have doctors and practitioners reach out to patients to take the vaccine



# THANK YOU



Marcelo Scatena

Anthony Warren

Piotr Czulplik