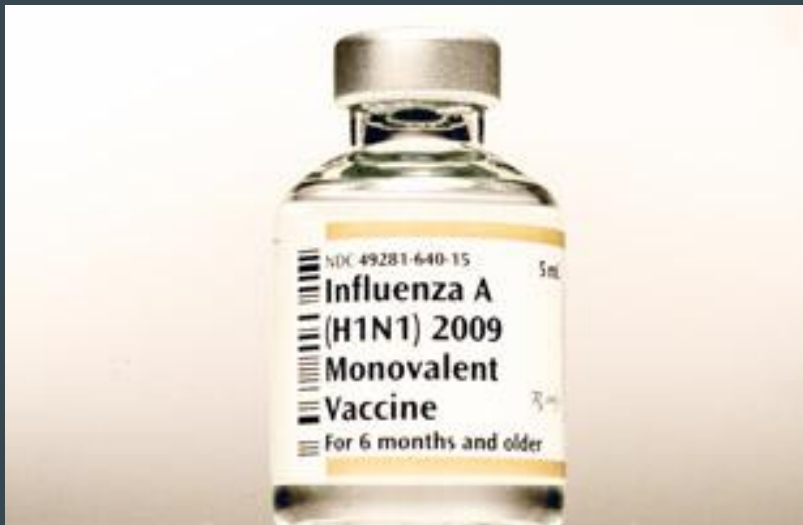


# H1N1 Vaccination Rates



Elizabeth Webster  
May 2022

# Introduction



This project explores the use of classification models in order to predict whether or not an individual will be willing to receive the H1N1 Vaccine.

We will be looking for the model that has the highest accuracy score for our test data set.

# Outline

- Business Problem
- Data
- Methods
- Results
- Feature Importance
- Recommendations
- Next Steps

# Business Problem

- Company that distributes vaccines
- Predict the number of vaccines for certain populations
- Primarily concerned with model accuracy scores

# Data

## National 2009 H1N1 Flu Survey

Includes features such as:

- H1N1 Concern
- H1N1 Knowledge
- Personal Behaviors
- Doctor Recommendations
- Age
- Education Level
- Race



# Methods - Data Cleaning

- Removing unnecessary columns
- Finding and filling missing data
- Encoding categorical columns
- Addressing the class imbalance

# Methods - Creating Models

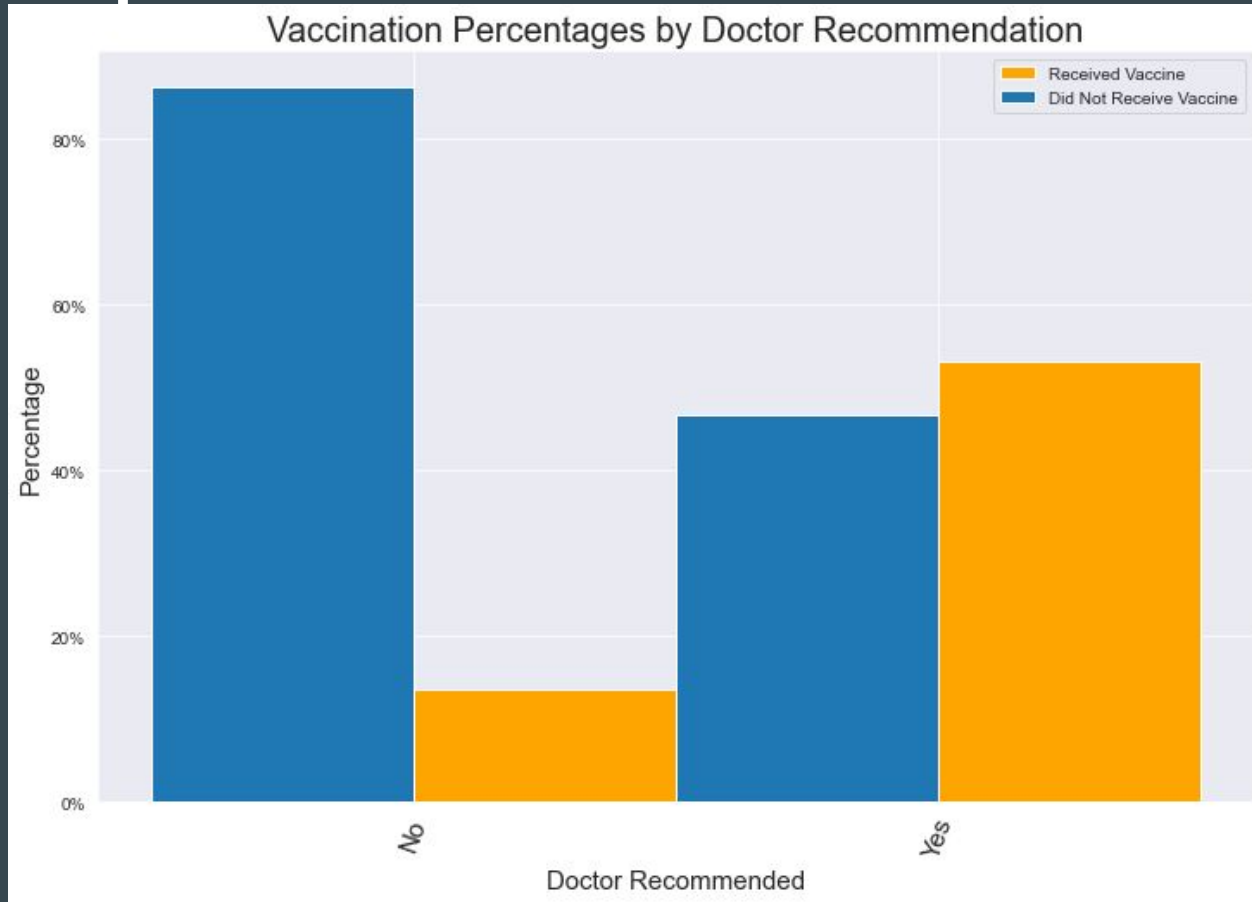
- Baseline Model - Decision Stump
- Advanced Models - Random Forests & XG Boost
- Parameter Tuning through Grid Searches

# Results

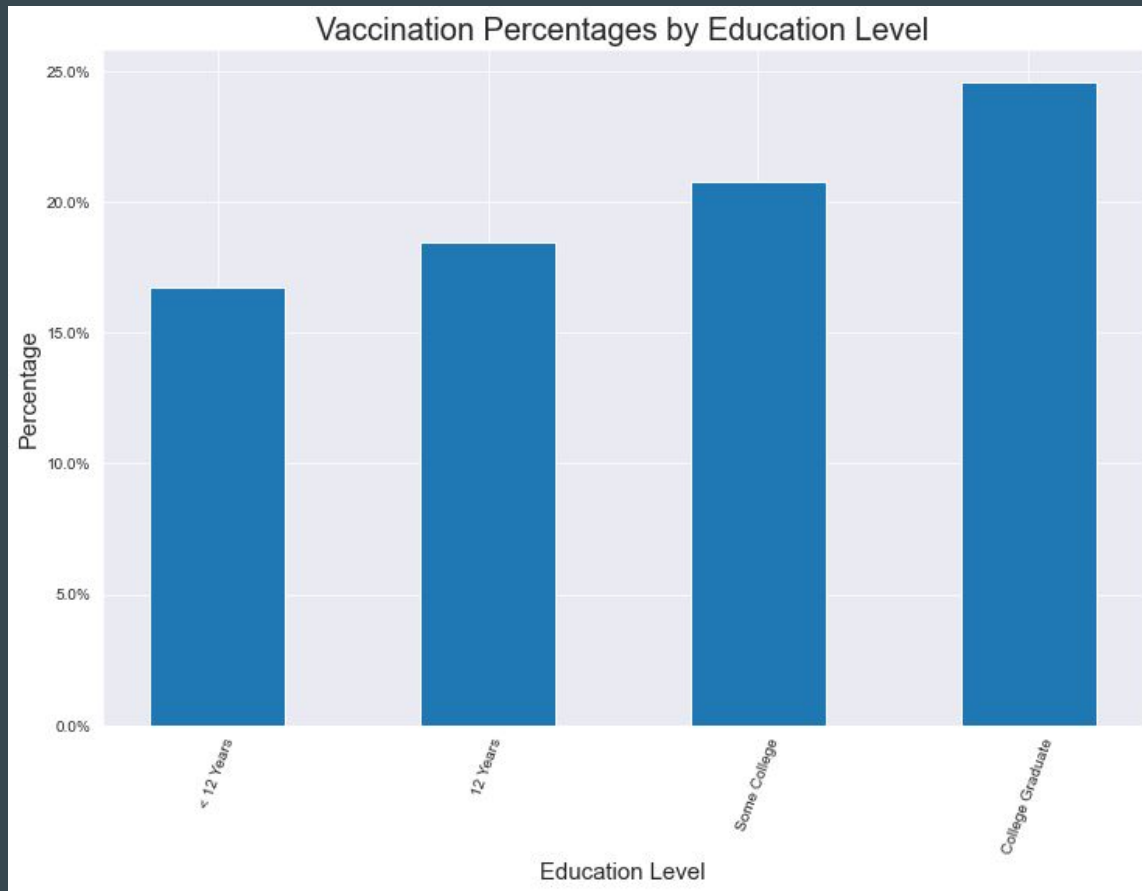
- Decision Tree - 81% accuracy score
- Random Forests - 83% accuracy score
  - Overfitting tendencies
- XG Boost - 84% accuracy score



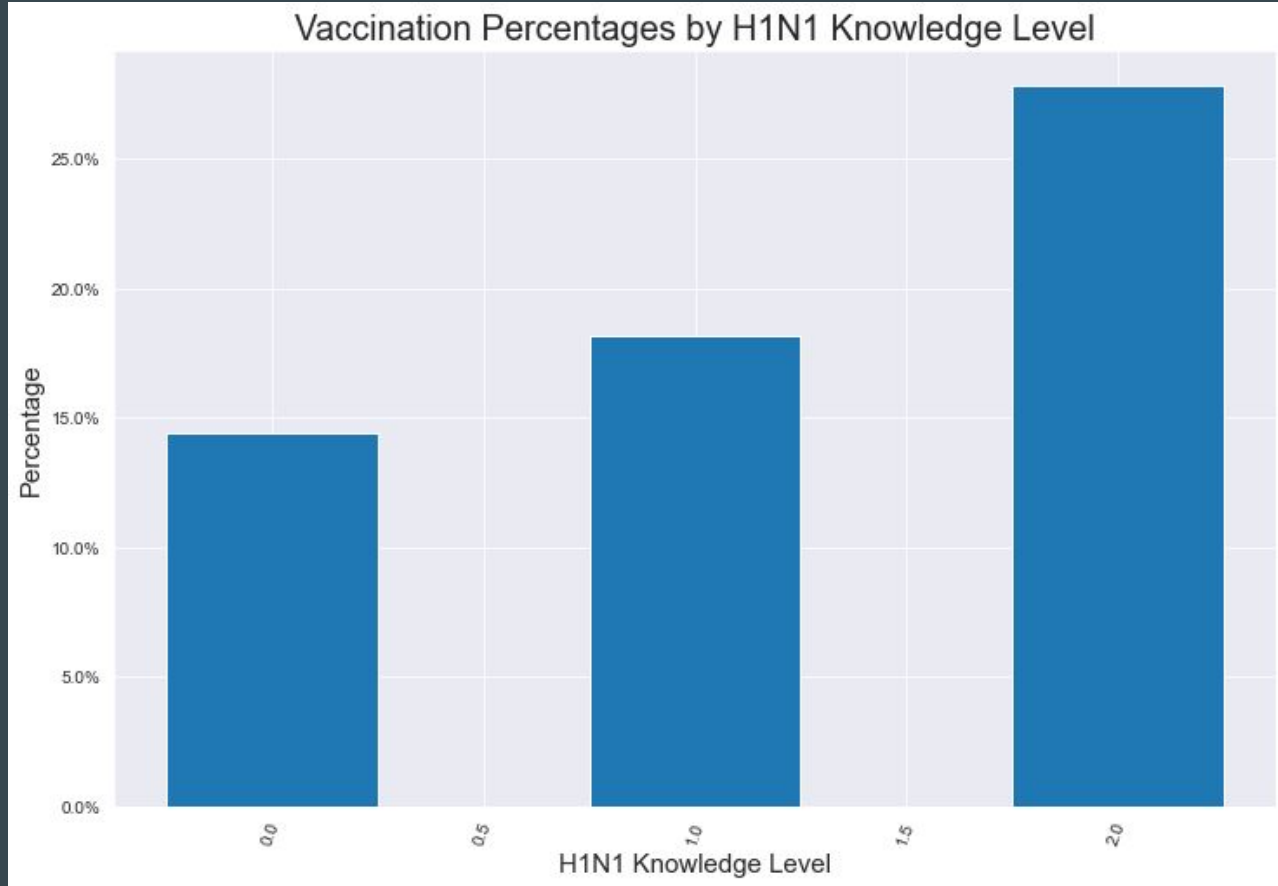
# Feature Importance - Doctor Recommendation



# Feature Importance - Education Level



# Feature Importance - H1N1 Knowledge



# Recommendations

- Use the XG Boost model for predicting vaccinations for a population
- Distribute more vaccines to populations with higher levels of college educated individuals
- Use information from clinics to understand the amount of patients they are recommending the vaccine to

# Next Steps

- Tackle the class imbalance
- Understand vaccination rates by group
- Promoting vaccines through clinics

# Thank You!

Email: [eaw524@gmail.com](mailto:eaw524@gmail.com)

GitHub: <https://github.com/elizabeth524>