Seasonal Flu Vaccination Predictions

Amelia Dobronyi

Overview

- Business problem

- Data Overview

- Model Validation Statistics

- Model Results

- Business Recommendations

- Conclusion

Business Objective

- Public health officials need to increase the number of individuals who receive seasonal flu vaccines in the United States*

  - Current vaccination rate: 49.2%

  - Target vaccination rate: 70%

- Using results of classification model, suggest where marketing resources should be spent

Data Overview

- Approximately 26,000 instances of demographic information and flu vaccine decisions from 2009 National H1N1 Flu Survey

- Types of questions include:

  - Age and household demographic statistics

  - Doctor recommendations for a vaccine

  - Opinions on risks associated with both getting the flu and/or the vaccine

Model Overview and Process
Start with classification…

- Using the given data, build models to predict [read: *classify*] whether or not an individual will get a seasonal flu vaccine

- Use the models on untested data to 'score' the model and determine its effectiveness
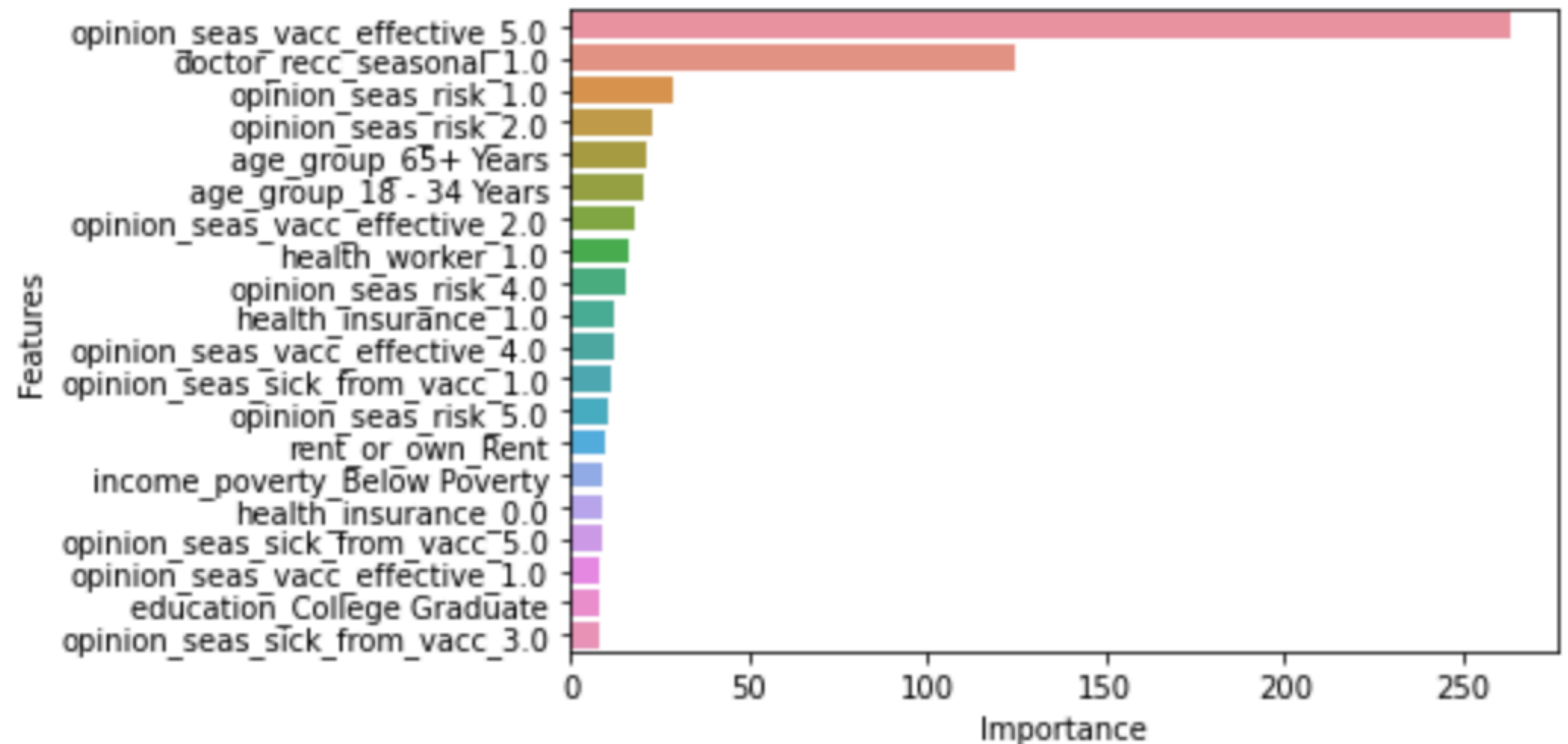
Model Scores
How to choose the best model?

- **Accuracy:** how good was the model as classifying correctly

- **F1 score:** combines precision and recall

  - We care about recall because we don't want false positives
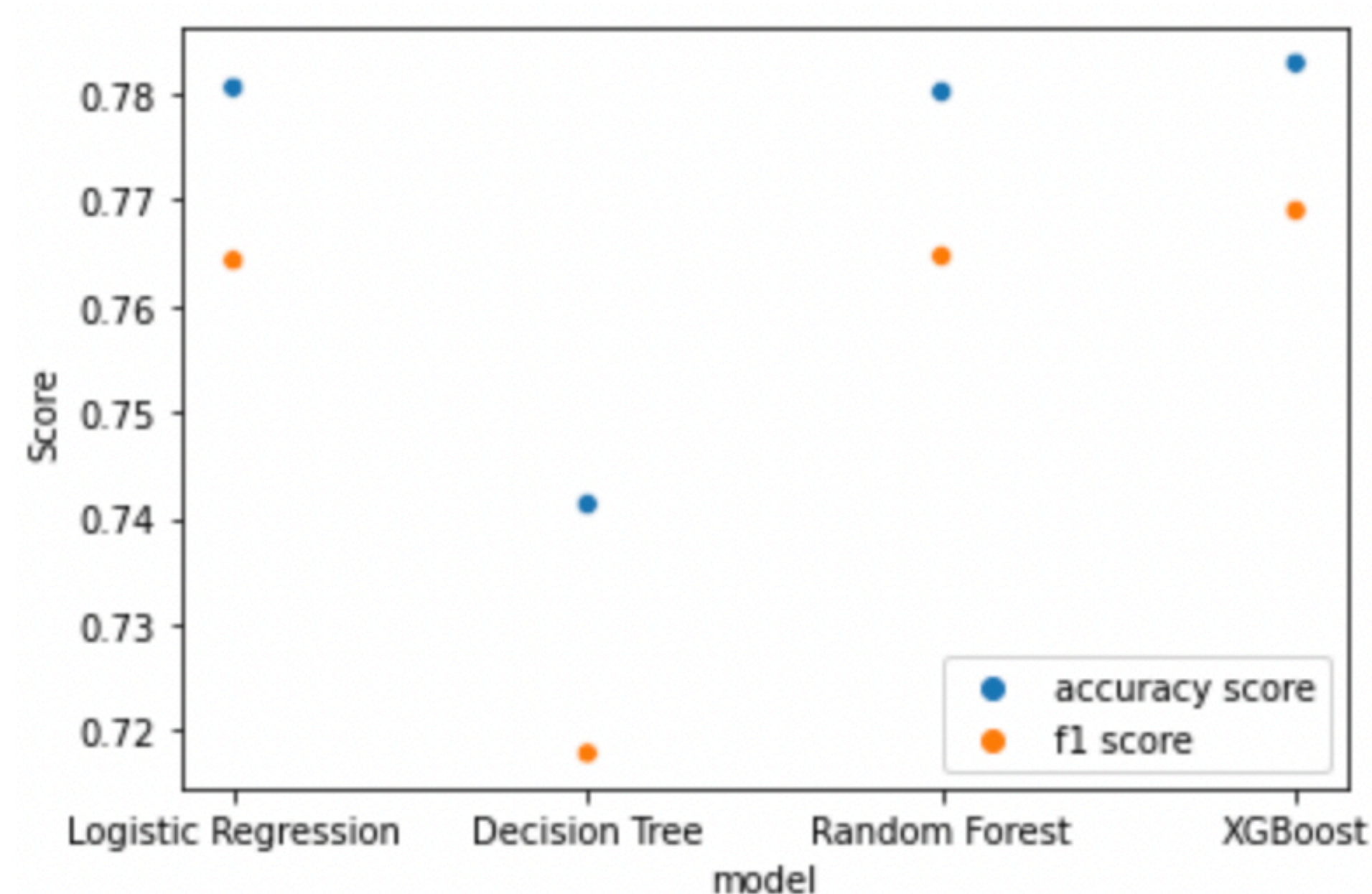
Feature Importance

- Opinion on vaccine effectiveness

- Doctor recommendation to get vaccine

- Opinion on flu risk

- Age group

1. Advertise effectiveness of vaccine and flu risk

Fig 1. Percentage of individuals
vaccinated increased with higher opinions
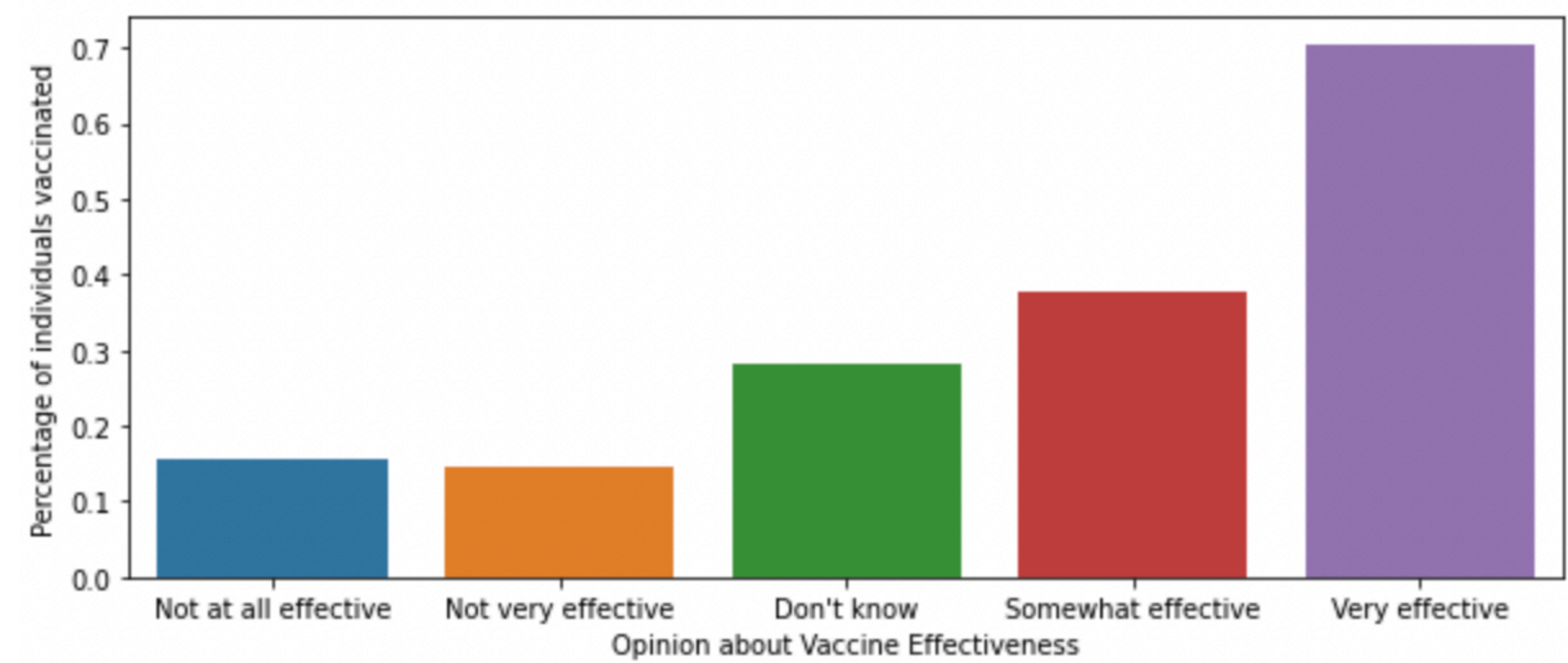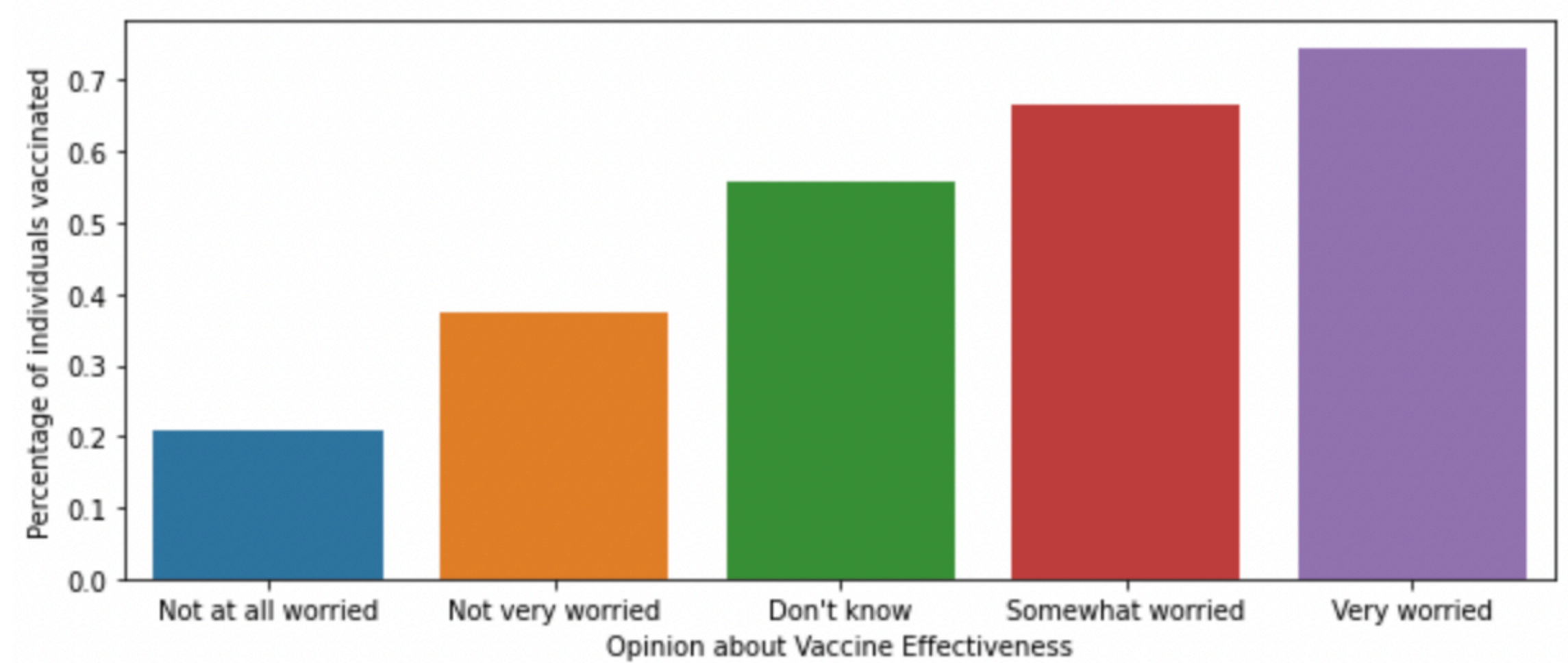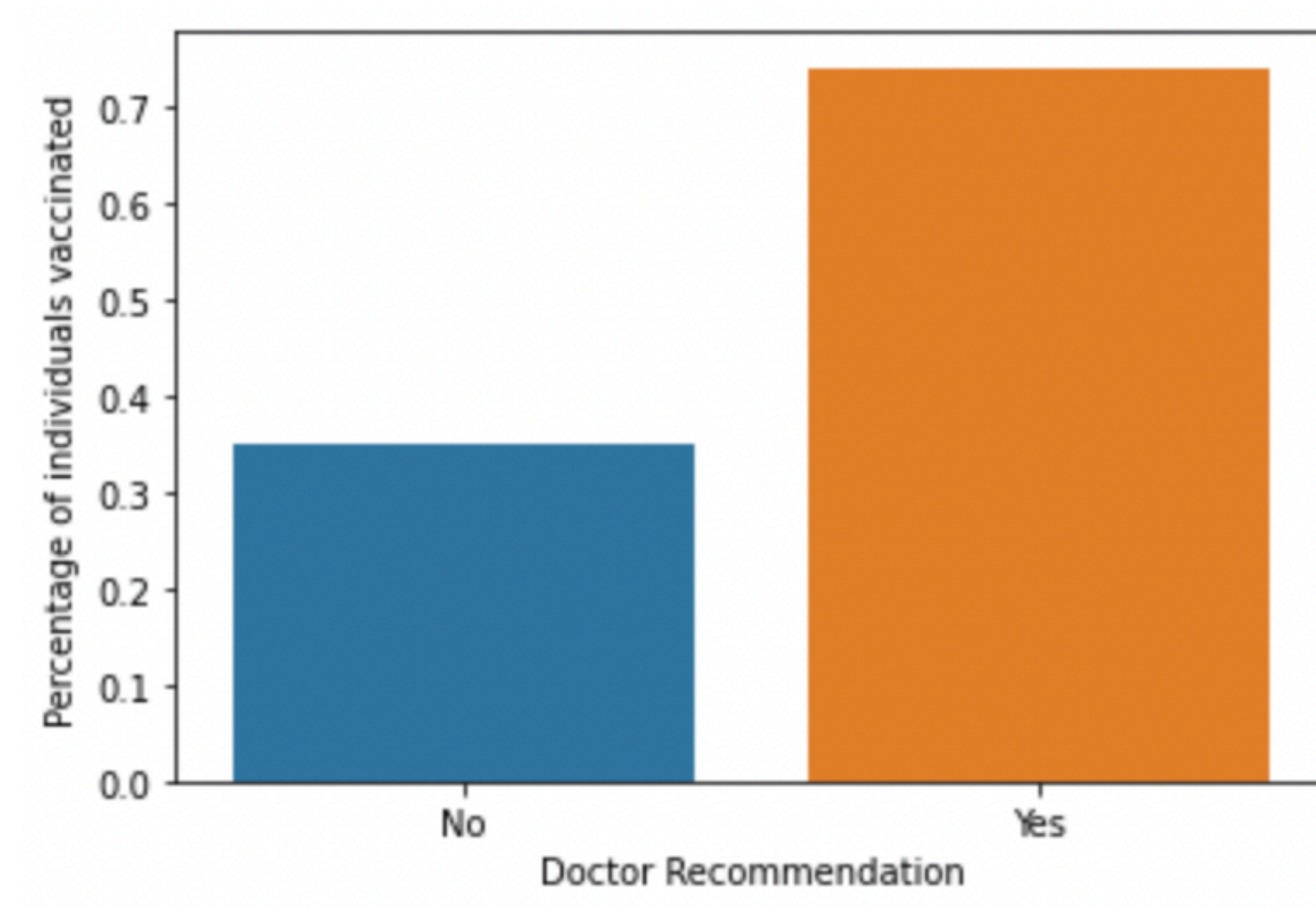of vaccine effectiveness

Fig 2. Percentage of individuals
vaccinated increased with greater fears
about potential risks of getting the flu

2. Encourage more doctors to recommend vaccination

Percentage of individuals vaccinated was about twice as high for those who received a doctor recommendation to get the flu vaccine

3. Target sub-55 age groups

- The only age group that is reaching target vaccination levels is 65+

- To increase vaccine uptake, target lower age groups in marketing
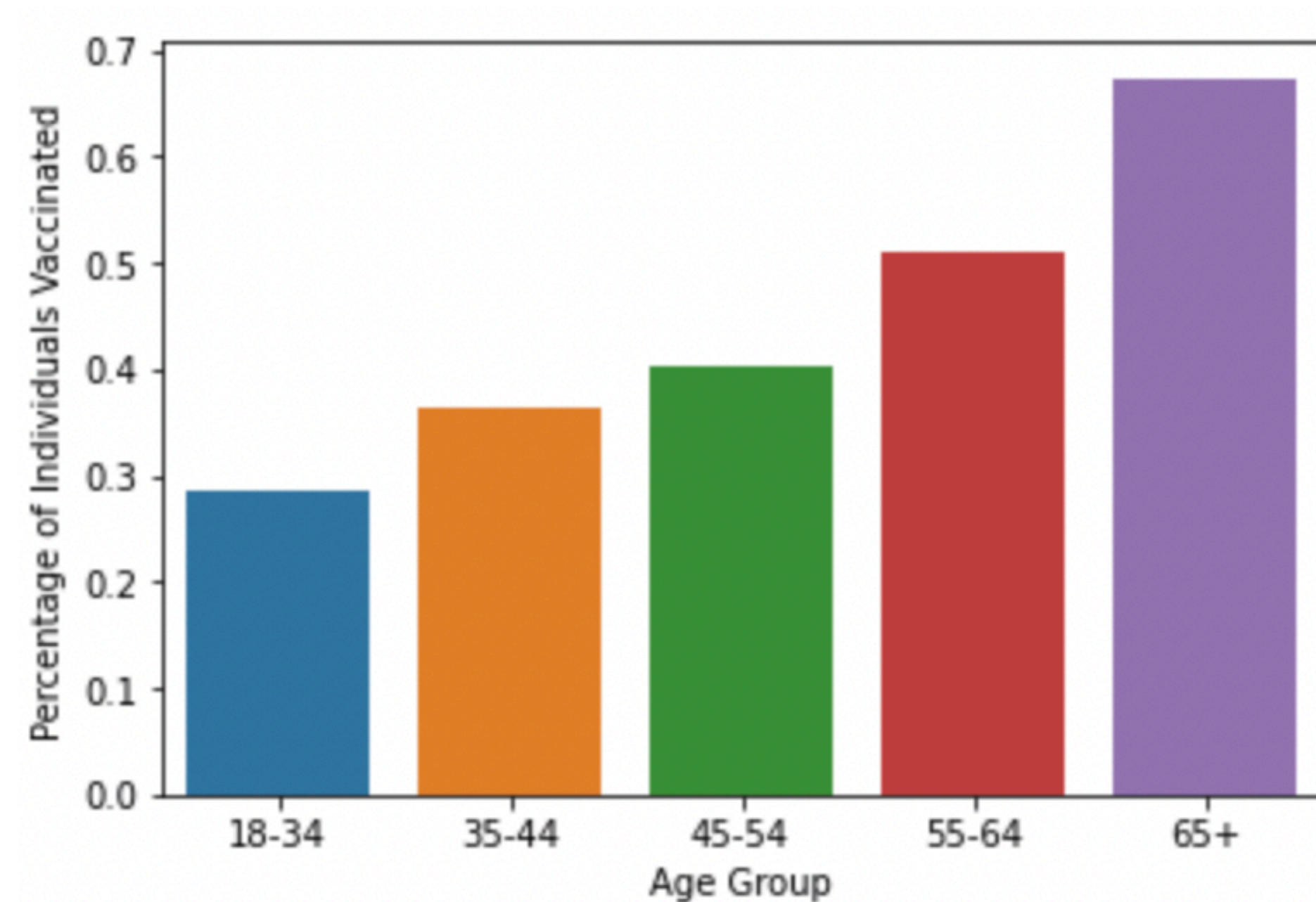
Conclusion

- XGBoost model chosen as best classifier on basis of accuracy

- Using model-important factors, solidify marketing efforts to increase vaccine uptake:

    - Increase awareness of vaccine effectiveness and flu risk

    - Encourage more doctors to directly recommend flu vaccine

    - Target lower (sub-55) age groups specifically

Future Improvements to Analysis

- Deal with missing data more accurately or via imputation

- Examine interactions between demographic data

- Use more advanced parameter tuning techniques for further classification gain

Contact

- Email: adobronyi@mac.com

- GitHub: @adobronyi

- LinkedIn: https://www.linkedin.com/in/amelia-dobronyi-cfa-3a12b96b/