Seasonal Flu Vaccine: Predictive Model

Caroline Surratt

Phase 3 Project

Flatiron Data Science Flex (June 2023)

Business Understanding

"Immunizations are among the most successful and cost-effective health interventions ever devised." -World Health Organization

This model can be used to identify features that are the strongest predictors of vaccination status for the seasonal flu.

Healthcare providers can use this knowledge to:

- identify individuals unlikely to get vaccinated
- provide appropriate interventions
- improve overall vaccination rates

Data Understanding

National 2009 H1N1 Flu Survey

- 26,707 responses
- conducted via telephone
- responses are anonymous
- features include individuals' behavior, demographics, and opinions/knowledge
- target variable indicates vaccination status (binary)

Modeling

Strategy:

- 2 preprocessing strategies were used
- 10 moderately-generic models were created
- Model with the highest overall accuracy was fine-tuned

Evaluation

Primary goal: identify individuals who have not gotten seasonal flu vaccine in order to provide interventions.

False positive: predicting that a person ***is unvaccinated*** for seasonal flu when they ***actually are vaccinated.***

False negative: predicting that a person ***is vaccinated*** for seasonal flu when they ***actually are not vaccinated.***

For the context of this model, it will be more beneficial to avoid false negatives.

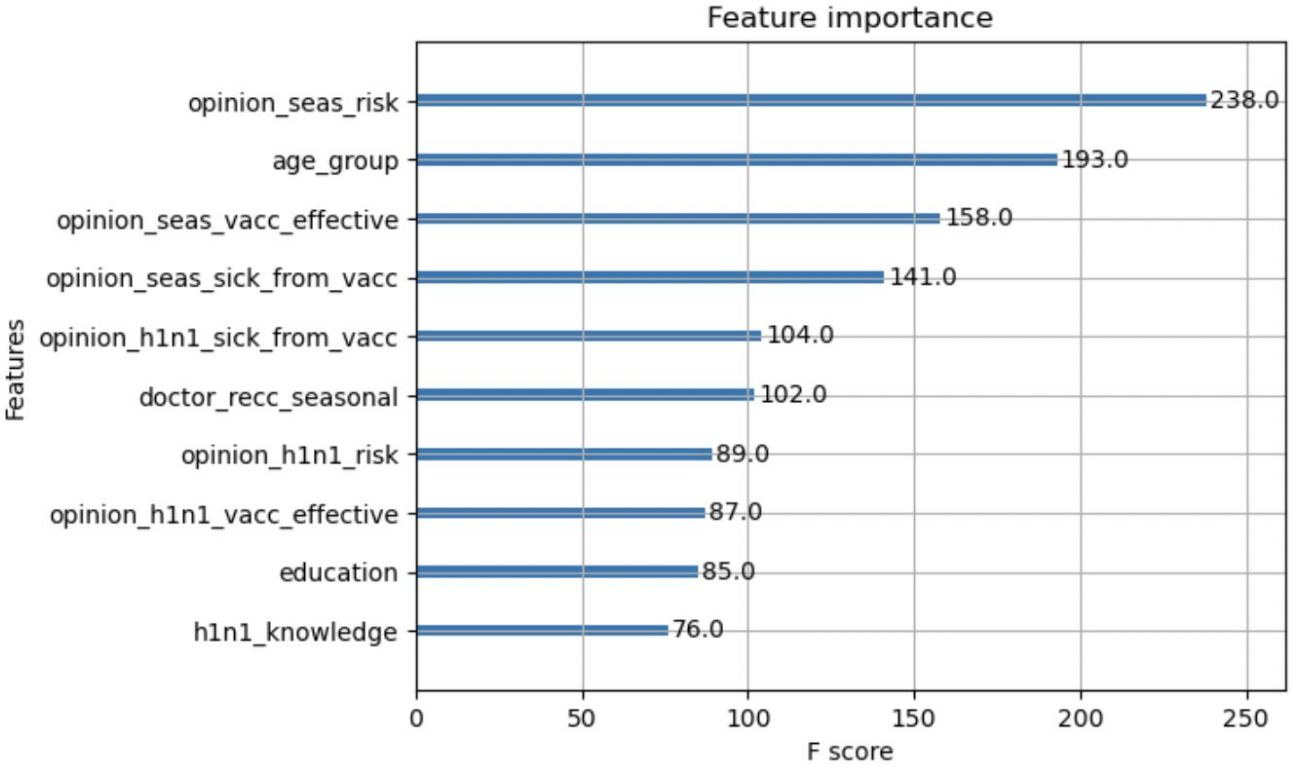
Therefore, recall score is utilized for model tuning.

Evaluation

After tuning, the best model had the following metrics:

- Accuracy: 79.66%
- Recall: 81.84%

Findings


Feature importance
opinion_seas_risk
age_group
opinion_seas_vacc_effective
opinion_seas_sick_from_vacc
opinion_h1n1_sick_from_vacc
doctor_recc_seasonal
opinion_h1n1_risk
opinion_h1n1_vacc_effective
education
h1n1_knowledge
238.0
193.0
158.0
141.0
104.0
102.0
89.0
87.0
85.0
76.0
Features
0
50
100
150
200
250
F score

Recommendations

Healthcare providers can take the following actions in order to influence the features that are the most important to predictive ability:

- Display informative/educational materials about the risks of the flu.
- Talk to individuals about the risks of flu and vaccines during *all* routine/preventative appointments.
- Directly recommend vaccination to all patients.

Next Steps

- Utilize more recent data (consider COVID implications)
- Consider additional strategies for reducing noise in dataset
- Controlled experiment/interventions

Thank you!

Caroline Surratt

Email: carolinecsurratt@gmail.com

GitHub: @ccsurratt

LinkedIn: linkedin.com/in/carolinesurratt/