Predicting Seasonal Flu Vaccination Status

Presented by Jonah Devoy

Purpose

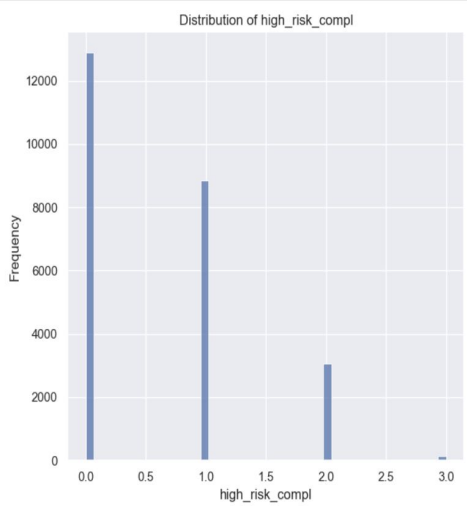
The goal of this project was to build a classifier to predict whether someone was vaccinated against the seasonal flu or not as accurately as possible, while maximizing recall of those that elect not to get the vaccine.

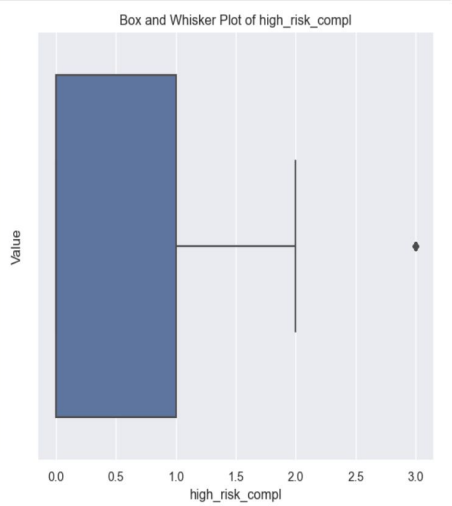
About the data

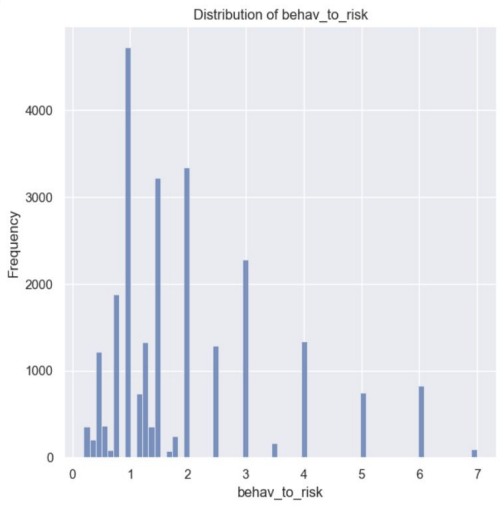
- The data used was obtained from DrivenData
- Flu Shot Learning: Predict H1N1 and Seasonal Flu Vaccines
- Phone survey from 2009 asked respondents whether they had received the H1N1 and seasonal flu vaccines, in conjunction with questions about themselves.
- These additional questions covered their social, economic, and demographic background, opinions on risks of illness and vaccine effectiveness, and behaviors towards mitigating transmission. A better understanding of how these characteristics are associated with personal vaccination patterns can provide guidance for future public health efforts.

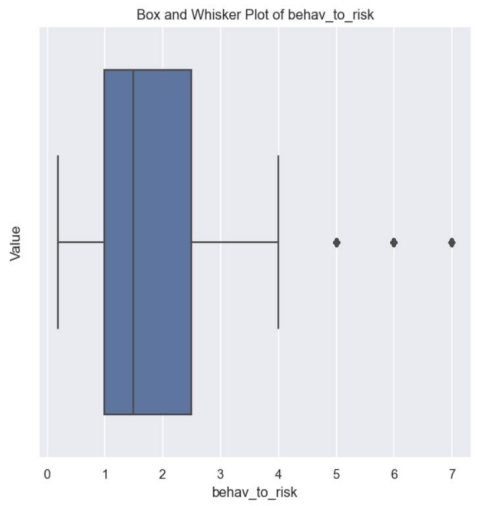
Feature Engineering:

- behav_score: a variable that represents how much an individual has done behaviorally to avoid the flu
- behav_to_risk: a variable that represents the ratio of how much an individual has done behaviorally to avoid the flu
- high_risk_cat - a variable that represents an individual's overall risk for developing flu-related complications.
- race - This column was originally made up of 4 categories: White, Black, Hispanic, and Other or Multiple.









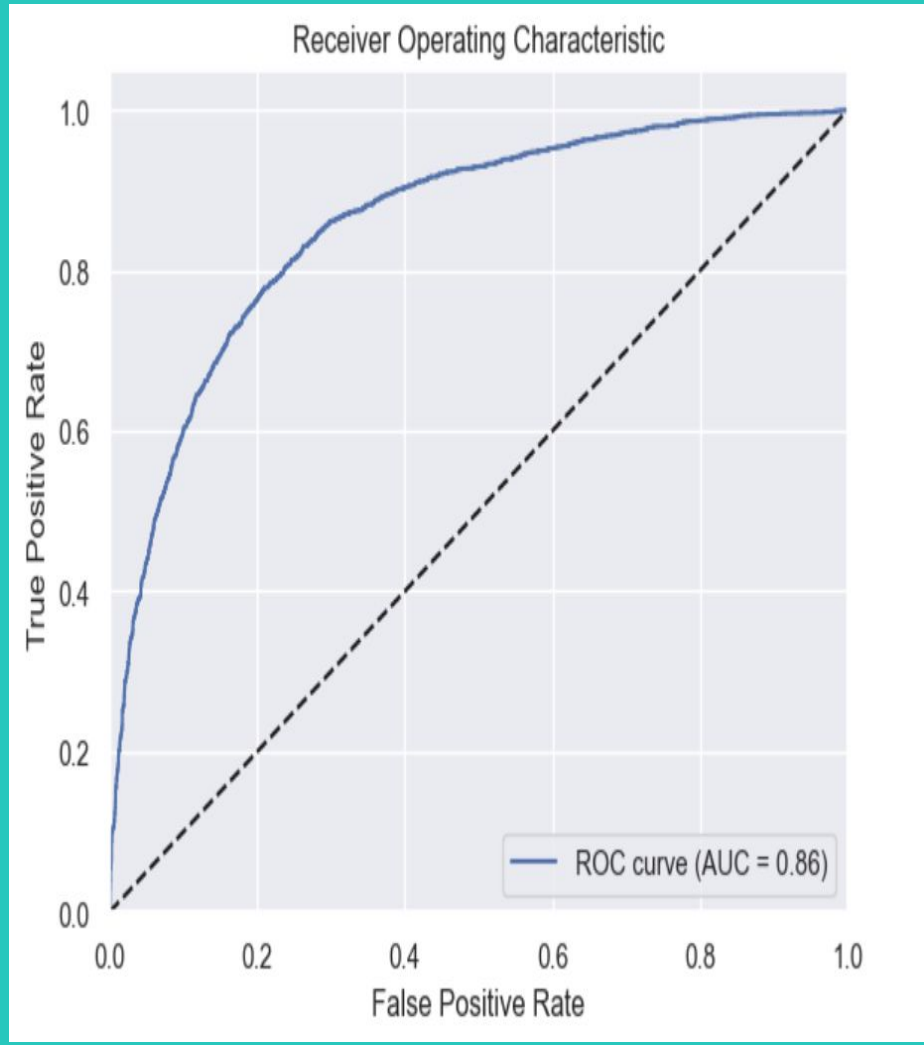
Modeling

LogisticRegression: 0.78

RandomForestClassifier: 0.77

KNeighborsClassifier: 0.69

DecisionTreeClassifier: 0.69



Feature Importance:

- opinion_seas_vacc_effective - how effective on a scale from 1 to 5 (5 being very effective) the respondent believes the vaccine to be at protecting against the flu
- doctor_rec_seasonal - whether or not the individual's doctor recommended they get the vaccine, specifically 1: their doctor did recommend getting vaccinated was the best predictor
- opinion_seasonal_risk - how concerned on a scale from 1 to 5 (5 being very effective) the respondent is about getting the seasonal flu without the vaccine
- age_group - 2 specific age groups were especially useful for predicting vaccination status: 65+ years and 18-34 years, and a third also made it into the top 20 predictive features (55 - 64 Years)
- health_worker - whether or not the individual is a health worker

Conclusion:

3 most influential factors in determining whether someone got vaccinated against the seasonal flu in 2009 were:

- How effective they believed the vaccine to be at protecting against the flu.
- Whether a doctor recommended they get the vaccine.
- Their perceived level of risk of getting sick with the flu without the vaccine.

Recommendations:

- Increase public awareness of the effectiveness of the vaccine at protecting against the flu.
- Doctors should regularly recommend that their patients get vaccinated against the seasonal flu each year.
- Increase health literacy

QUESTIONS?

Jonah Devoy -
jonah.devoy@yahoo.com