Prediction of Seasonal Flu Vaccinations

Name: Odhiambo Mandela

Overview

- Vaccines are a pillar to preventive healthcare.
- They are key to preventing the spread of dangerous diseases.
- Dissent exists against vaccines.
- Data science (DS) can be a tool that helps to reveal why people do not take vaccines.
- Also, DS can be used to forecast people's receptiveness to vaccines.

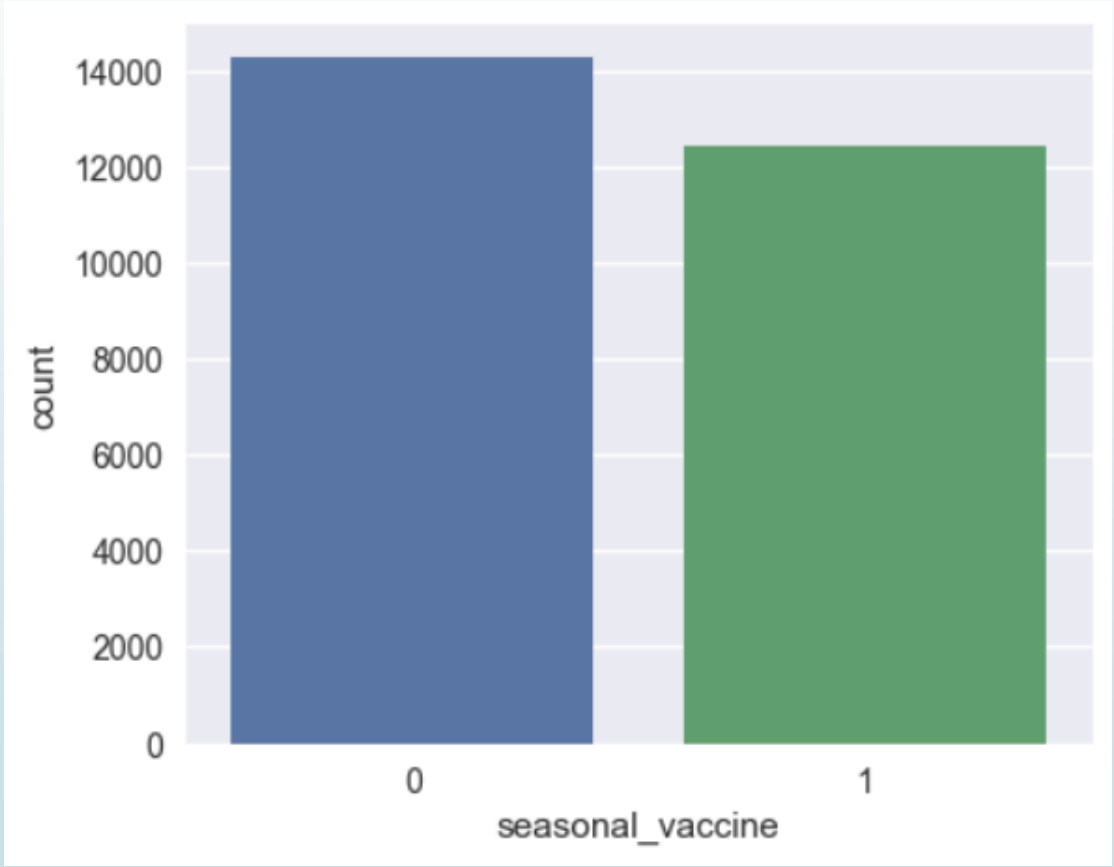
Business Understanding

- Planning successful immunization campaigns and advancing public health depend on accurately predicting vaccine uptake.
- Targeted treatments can be created by precisely identifying people who are most likely to accept or reject immunizations.
- With the help of data science models, this study aims to shed light on the variables affecting vaccine acceptability.

Data Understanding

- Dataset used was downloaded from Driven Data.
- The dataset was part of a 2009 National H1N1 Flu Survey.
- The dataset contains 35 features and 2 target variables.
- For this project, the seasonal_vaccine variable was used as the target variable.
- The h1n1_vaccine had imbalance issues and hence dropped.

Preview of the Seasonal Vaccine Variable



- Less imbalance issues in this target variable.

Modeling

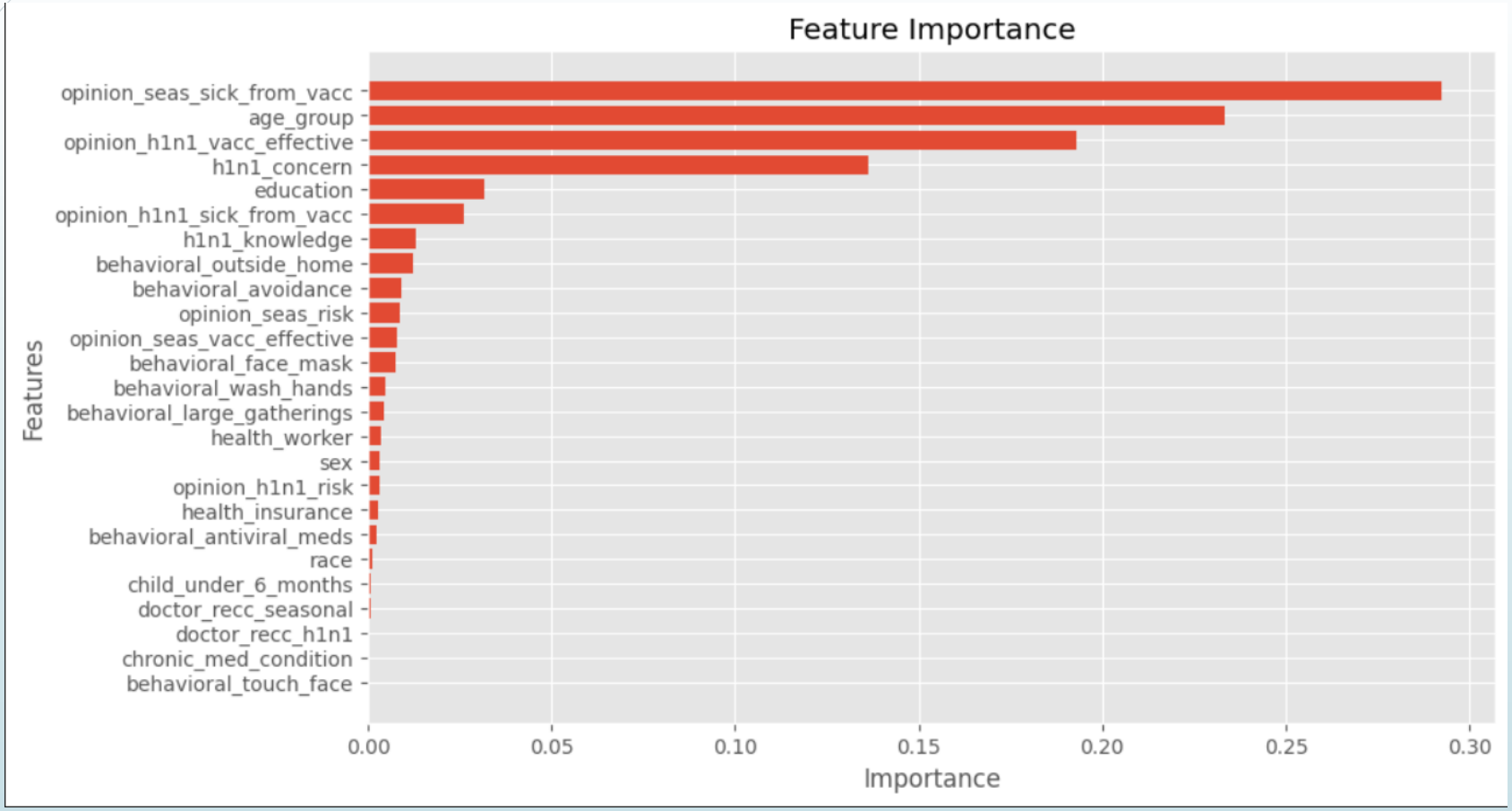
- Various machine learning algorithms, including logistic regression, random forest, gradient boosting, and XGBoost, will be employed.
- Pipelines and grid search will be utilized to optimize hyperparameters and improve model performance.
- The models will be evaluated using metrics such as accuracy, precision, recall, and F1-score.

Evaluation

	Cross_Validation_Accuracy	Model_Train_Accuracy	Model_Test_Accuracy
Logistic_regression	0.769	0.771	0.782
Random_forest	0.774	0.796	0.780
Gradient_boosting	0.778	0.786	0.785
XGBoost	0.778	0.786	0.785

- Gradient Boosting and XGBoost models posted the highest scores.
- The 80% accuracy target was not reached, but the models were able to post accuracies that were more than 76% successful with minimal issues of overfitting and underfitting.

Best Features



Recommendations

- Vaccination campaigns should target young people.
- The public health efforts should be geared towards shaping the public narrative regarding seasonal vaccines.
- Conduct the survey on an inclusive population. Survey participants were predominantly white.

Next Steps

- Tune the model even more.
- Reduce the number of features and compare the results.