PREDICTING VACCINE UPTAKE

INSIGHTS AND RECOMMENDATIONS

OVERVIEW

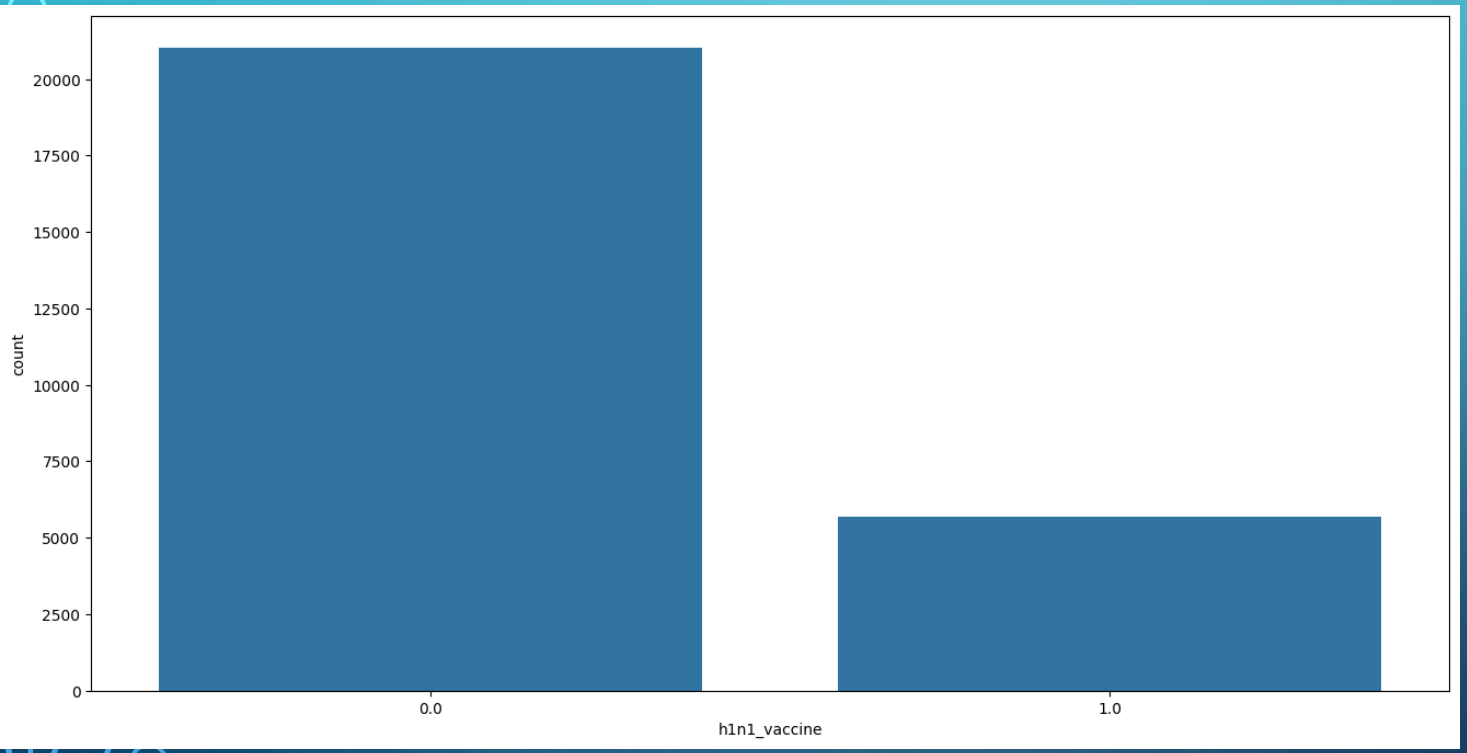
- This analysis aims to understand and predict vaccination behavior for H1N1 and seasonal flu using data from the National 2009 H1N1 Flu Survey. Findings will help public health organizations design targeted campaigns to increase vaccination rates.

BUSINESS AND DATA UNDERSTANDING

- - Stakeholder: Public health organizations and policymakers
- - Objective: Predict vaccine uptake based on demographic, social, and behavioral factors
- - Dataset Overview:
- - Targets: H1N1 vaccine uptake, seasonal flu vaccine uptake (binary: vaccinated or not)
- - Predictors: Demographic (age, gender), social (education, marital status), behavioral (health conditions, vaccine awareness)

EXPLORATORY DATA ANALYSIS

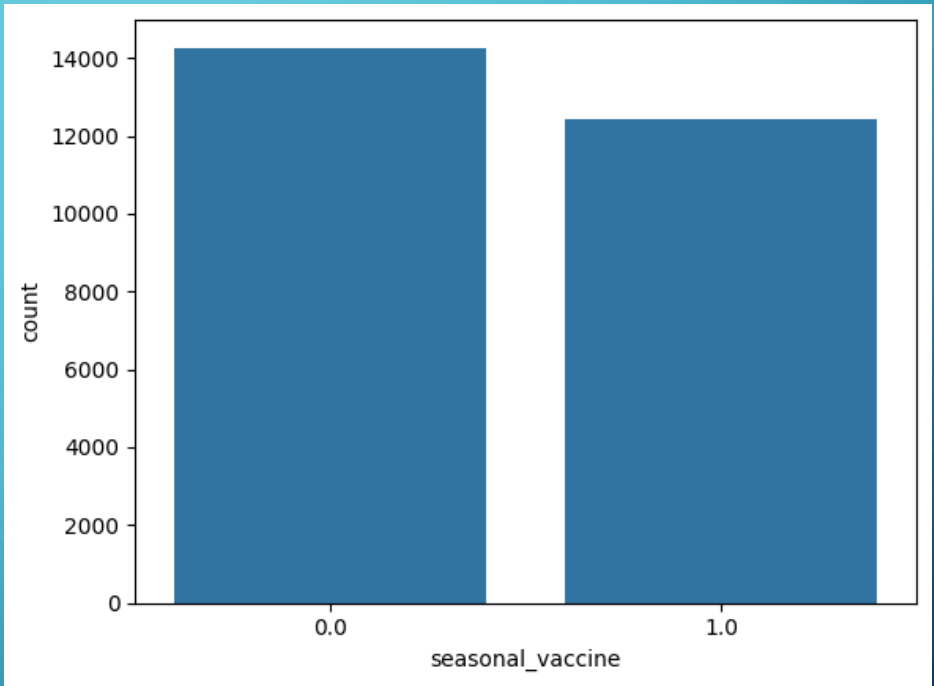
H1N1 VACCINE DISTRIBUTION



- Distribution of people who took the h1n1_vaccine
- Slightly more than 2000 didn't take the vaccine and 5000 did.

SEASONAL VACCINE DISTRIBUTION

- Slightly more than 14000 did not take the vaccine and slightly more than 12000 took it.



AUDIENCE DISTRIBUTION

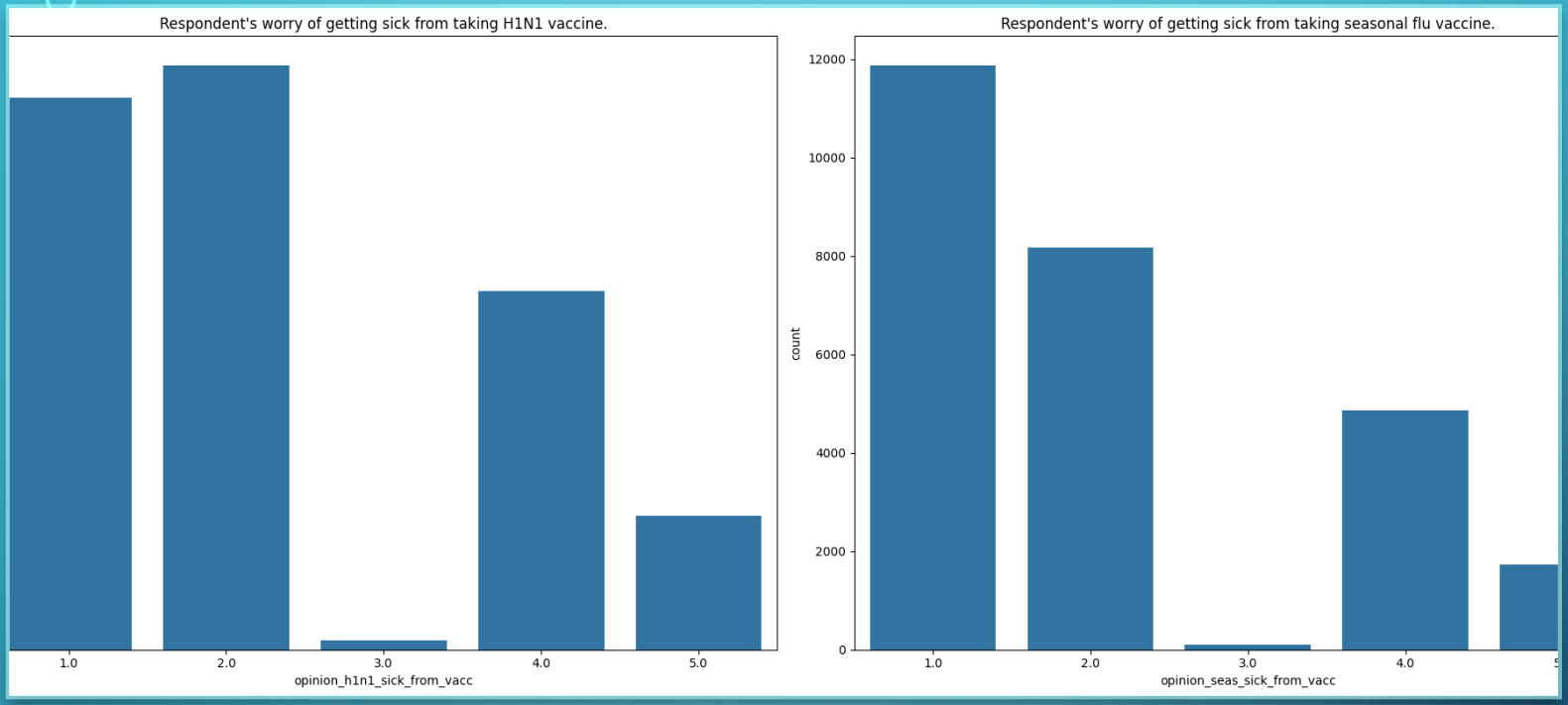
- The distribution of the audience that responded to the survey is as follows.

```
race
White                 21222
Black                  2118
Hispanic               1755
Other or Multiple      1612
```

```
sex
Female    15858
Male      10849
Name: count, dtype: int64
```

```
age_group
65+ Years          6843
55 - 64 Years      5563
45 - 54 Years      5238
18 - 34 Years      5215
35 - 44 Years      3848
```

```
education
College Graduate    11504
Some College         7043
12 Years             5797
< 12 Years           2363
```



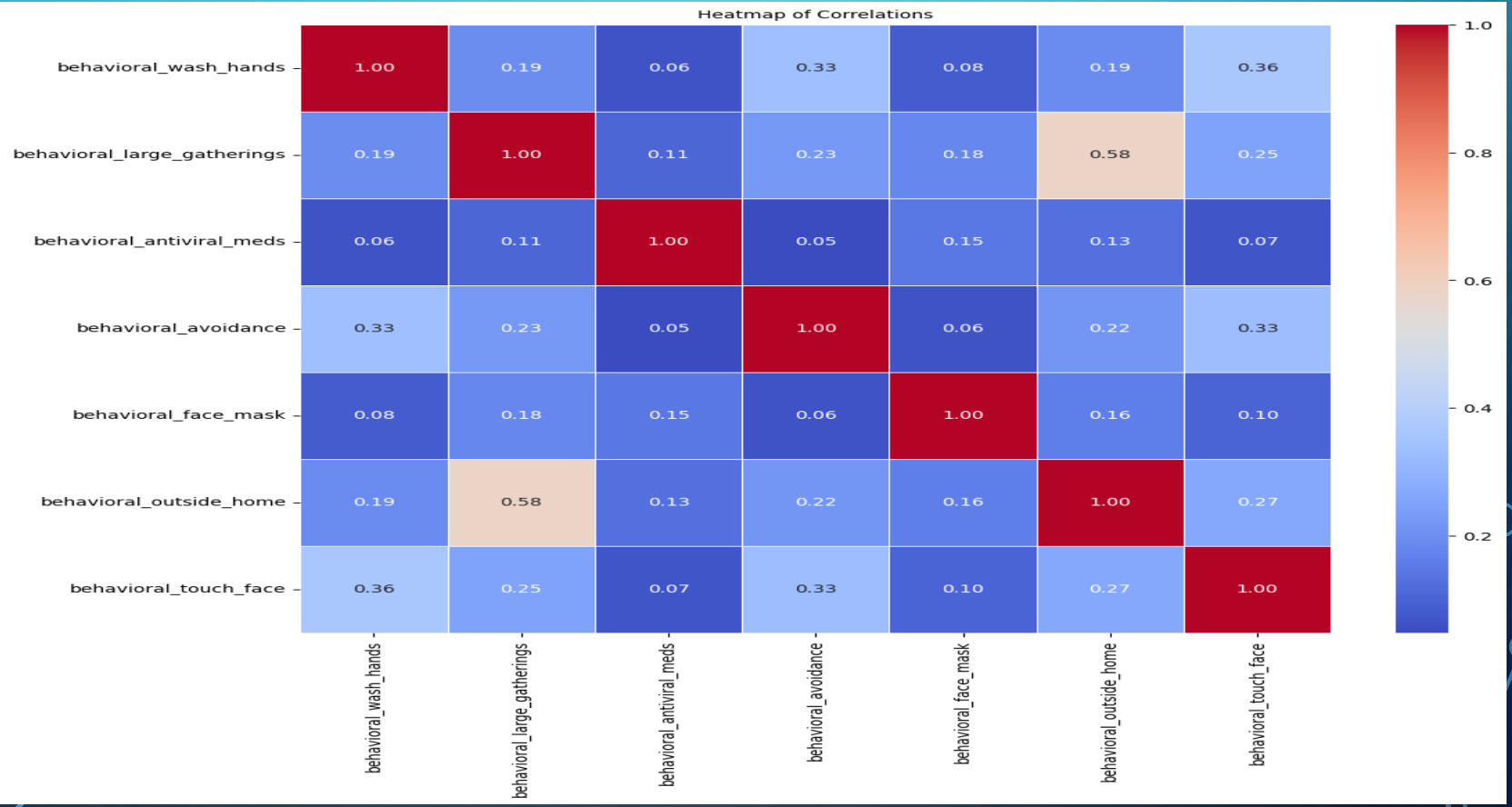
THE ABOVE IS THE CONCERN FOR TAKING THE VACICNE

- 1 Represents not all worried and 5 very worried.

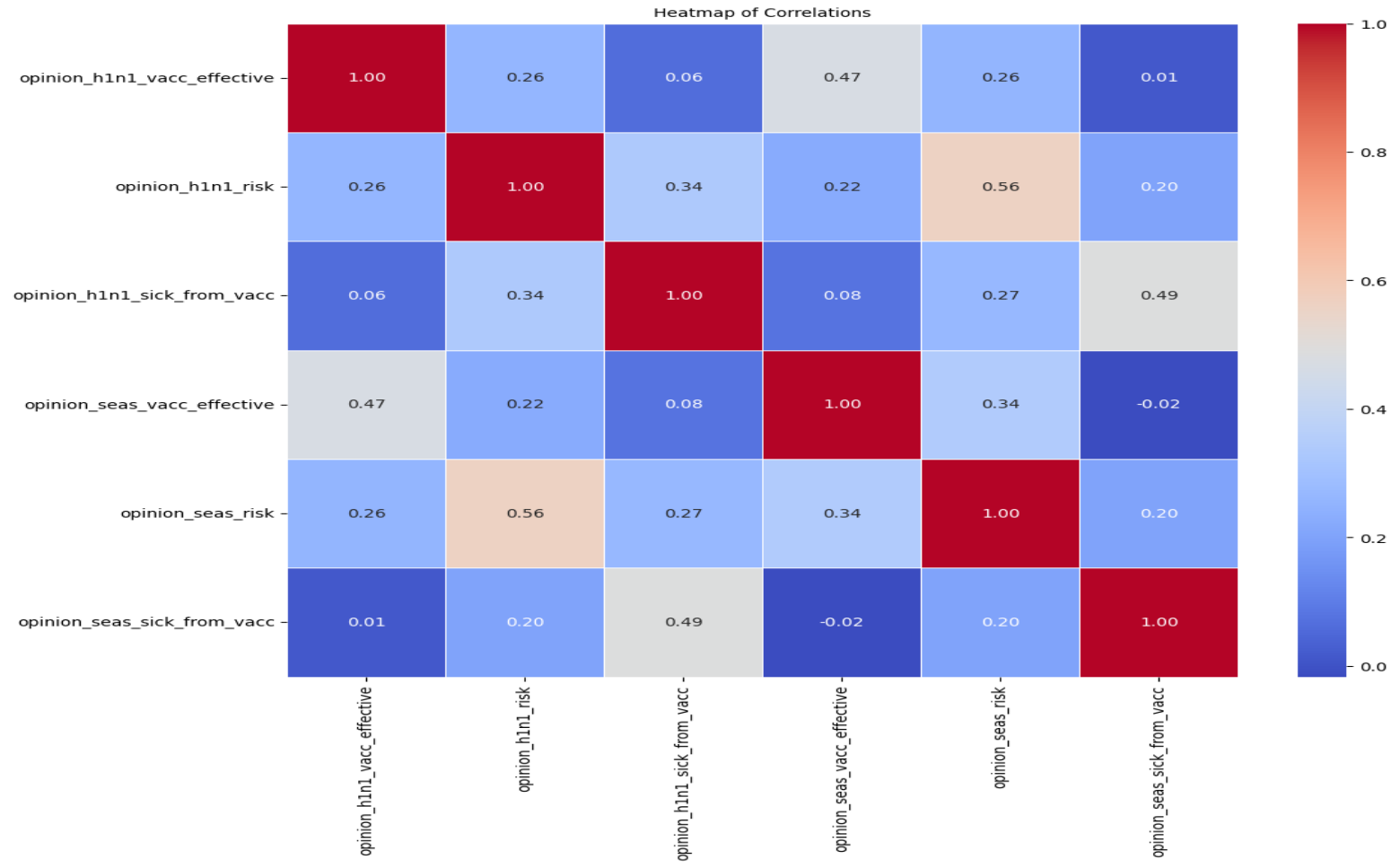
MEDICAL UNDERSTANDING FROM DATA

- **Conclusion:**
- **Public Awareness and Trust**: Seasonal flu vaccines are more familiar and accepted by the general public because they are offered regularly every year, whereas the H1N1 vaccine may be perceived as specific to a past pandemic.
- **Target Audience**: Seasonal flu vaccines protect against multiple influenza strains predicted to circulate in a given year, making them broadly applicable, while the H1N1 vaccine targets a specific strain, leading to a more limited perceived need.
- **Epidemiology**: The seasonal flu is a recurring issue, affecting millions annually, so its vaccine is prioritized for routine public health efforts. The H1N1 vaccine is targeted at a specific outbreak and has less uptake outside of outbreak periods.

RELATIONSHIP OF DIFFERENT PREVENTIVE METHODS



RELATIONSHIP OF OPINION ON VACCINE SAFETY



MODELING

- - Approach: Classifier chain Logistic regression to predict vaccine uptake
- - Why Classification: Identifies patterns in vaccination behavior for targeted interventions
- - Model Inputs: Age, education level, marital status, health conditions, opinions, behaviour

BASELINE MODEL VS HIGHEST PERFORMING MODEL

BASELINE

- Classifier chain Logistic regression

```
from skmultilearn.problem_transform import ClassifierChain
from sklearn.linear_model import LogisticRegression
from sklearn.metrics import accuracy_score
```

✓ 0.0s Python

```
chain1 = ClassifierChain(classifier=LogisticRegression())
chain1.fit(X_train_scaled,y_train)
predict_lr = chain1.predict(X_test_scaled)
```

✓ 0.2s Python

```
lr_cc= accuracy_score(y_test,predict_lr)
lr_cc
```

✓ 0.0s Python

```
0.6761512542119057
```

- The prediction is at 68%

HIGHEST PERFORMING

- Label powerset + Gradient Boosting

+ Code + Markdown

```
from skmultilearn.problem_transform import LabelPowerset
```

✓ 0.0s Python

```
model = LabelPowerset(GradientBoostingClassifier())
model_1 = model.fit(X_train_scaled, y_train)
predictions_nb_ps = model.predict(X_test_scaled)

nb_ps= accuracy_score(y_test,predictions_nb_ps)
nb_ps
```

✓ 13.4s Python

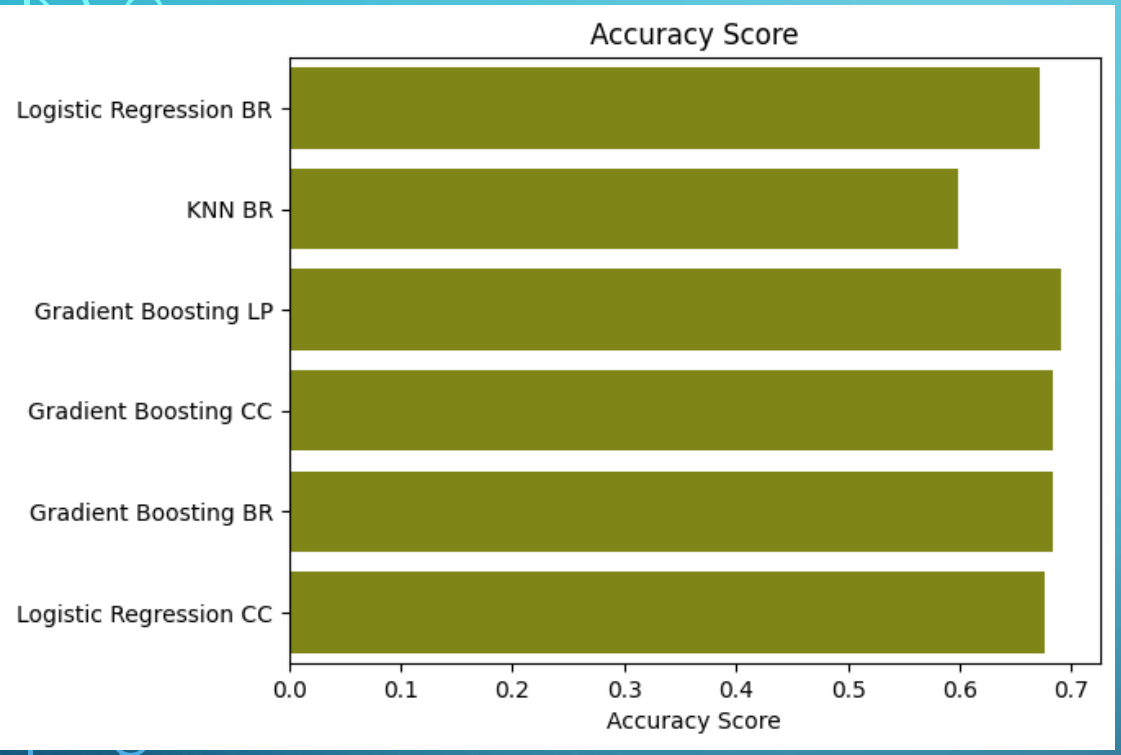
```
0.690565331336578
```

- The prediction is at 69%

EVALUATION

- - Key Metrics:
- - Accuracy: Overall correct predictions
- - Precision: Correctly predicted vaccinations
- - Recall: Effectiveness in identifying vaccinated individuals
- - Feature Importance:
- - Demographic factors like age and gender are critical
- - Behavioral aspects such as vaccine awareness influence decisions

EVALUATION




```
Accuracy: 0.690565331336578
Precision: 0.7503850457154486
Recall: 0.6467467187936331
```

- I compared the accuracy score with other models to see the best performing model which is the Gradient Boosting using Label PowerSet

RECOMMENDATIONS

- 1. Targeted Campaigns:
- - Focus on demographic groups less likely to vaccinate
- 2. Education Programs:
- - Increase awareness of vaccine benefits
- 3. Accessibility Initiatives:
- - Address barriers like cost and access to vaccination centers

NEXT STEPS

- - Expand the dataset with recent vaccination campaigns
- - Enhance the model with new predictors (e.g., social media influence)
- - Pilot targeted campaigns and measure their effectiveness

THANK YOU

- I look forward to discussing how these findings can improve public health outcomes.