

PREDICTING HOW LIKELY INDIVIDUALS ARE TO RECEIVE H1N1 VACCINES

Overview

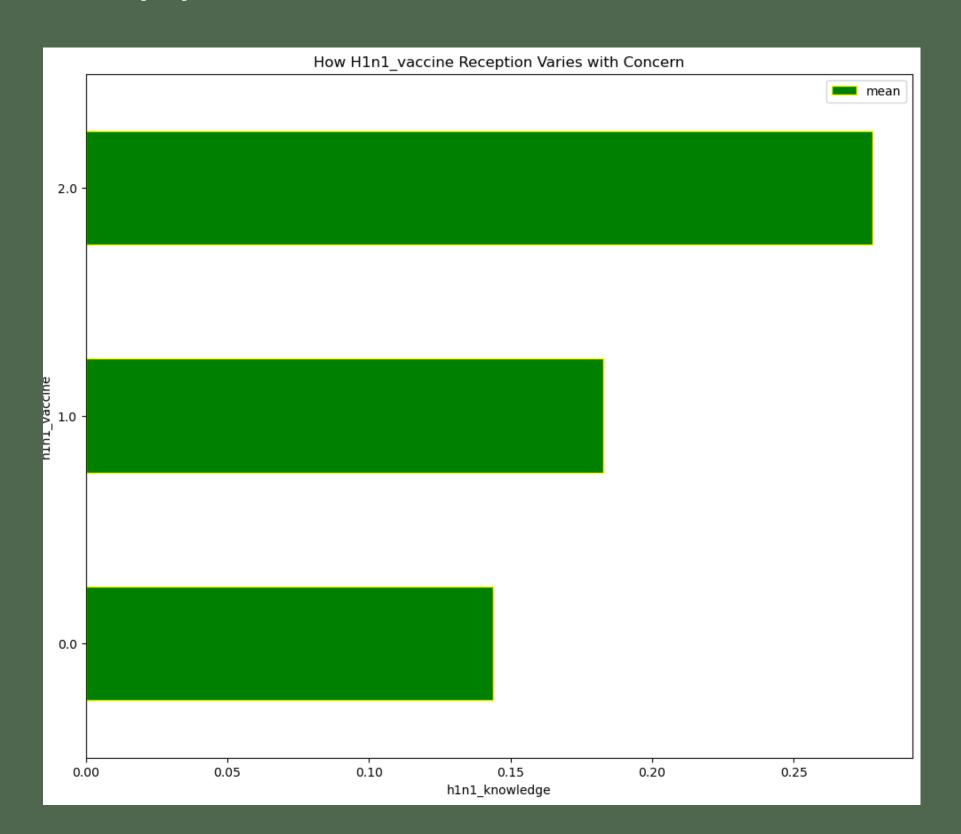
The role vaccination plays in combating diseases, especially infectious ones, by leveraging immunization to minimize disease transmission and achieve herd immunity is crucial. This project aims to forecast individuals' likelihood of receiving the H1N1 flu vaccine using data sourced from the 2009 United States National H1N1 flu survey. The respondents of this survey were queried about their vaccination status alongside personal details, including factors like health insurance coverage and occupation as healthcare workers, to gather diverse data points. This data points were then used to help in the prediction.

Business and Data Understanding

During late 2009 and early 2010, the United States conducted the National 2009 H1N1 Flu Survey with the objective of determining individuals' vaccination status. This project aims to ascertain whether an individual received the vaccine, based on specific characteristics. This study holds potential for future use by stakeholders in the public health sector, aiding in the identification of target populations for vaccination effort and how they could manage the administration of vaccines.

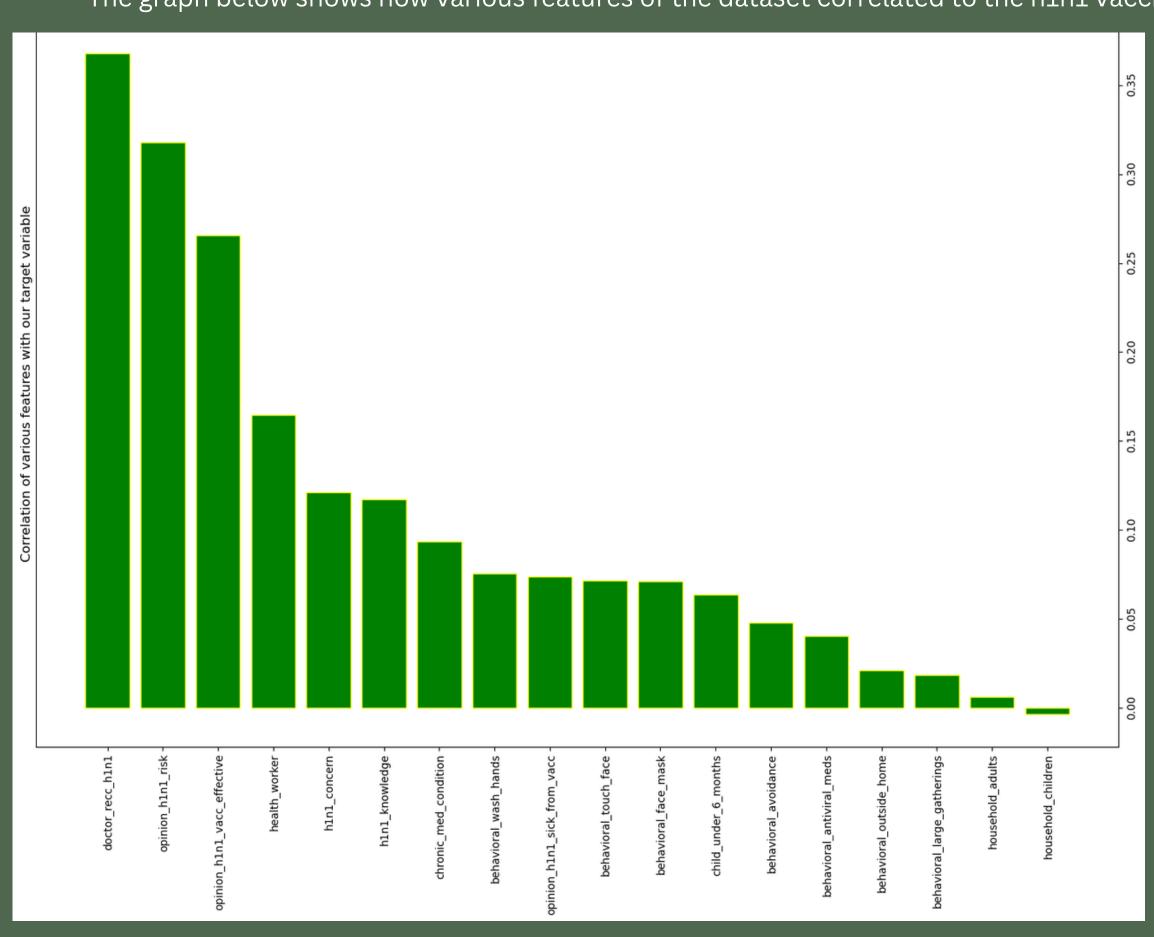
Unviariate analysis

In the univariate analysis below it was evident that only about 5000 from the population of about 20000 took the vaccine



Multivariate analysis

The graph below shows how various features of the dataset correlated to the h1n1 vaccine



Modelling

- Our Baseline model was a simple logistic regression model fitted on the training set without feature selection it had an f1 score of 43%.
- The second model used was the K-Nearest Neighbors (KNN) model. It was built after feature selection was performed using the most important features. It had an f1 score of 36%
- The third model we used was the Decision Tree model. It was built without hyperparameter tuning, resulting in poor performance. With an f1 score of 37%.
 - The final model was a Decision Tree model with hyperparameter pruning. This Improved model after hyperparameter tuning, leading to a higher F1 score of 78%.

Evaluation

The final model had an f1 score of 78%. This surpassed the previously set success metrics of 75%.

Recommendations and next steps

- It would be advisable to conduct future interview in person so as to obtain honest opinions
- The interviewers should attempt to find more data as to the particular reasons that one does not want to take the vaccines and address this issues.
- There should be more interviews being conducted to avoid the reliance on old datasets such as this one.
- The findings of this research acknowledged the importance of public awarness therefore the administrators should act on this.

THANK YOU