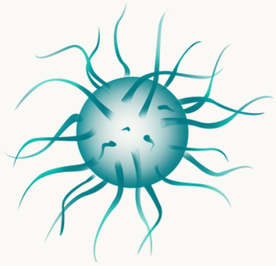
H1N1 VACCINE

STATUS PREDICTION



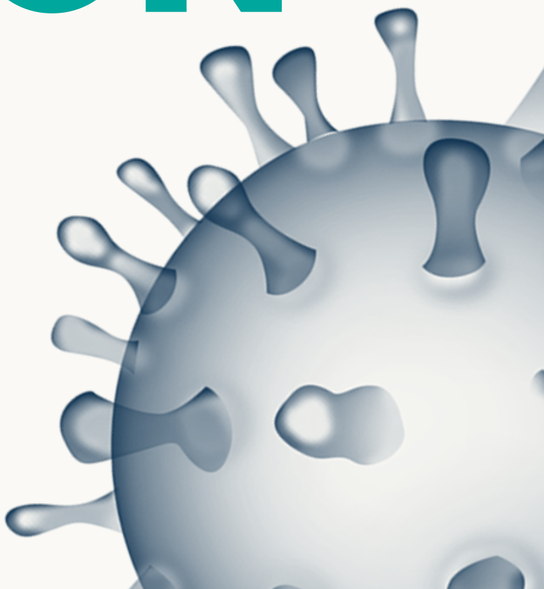
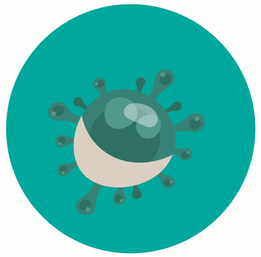


Table of Content



Introduction

- Project Overview
- Business Problem

Data Understanding

- Data Sources
- Features

Analysis

- Factors Influencing Vaccination Decisions:

Modelling

- Modelling Context
- Evaluation Metrics
- Model Comparison
- Feature Importance

Conclusion & Recommendation

- Finding
- Model Selection
- Recommendation

Introduction

01 Project Overview

- The project aims to predict individuals' likelihood of receiving H1N1 based on demographic, behavioral, and health factors

02 Business Problem

- The challenge lies in improving vaccination rates for both H1N1 and seasonal flu vaccines to mitigate the spread of influenza viruses. Vaccine hesitancy and refusal are significant challenges, necessitating an understanding of factors influencing vaccination decisions for targeted interventions.

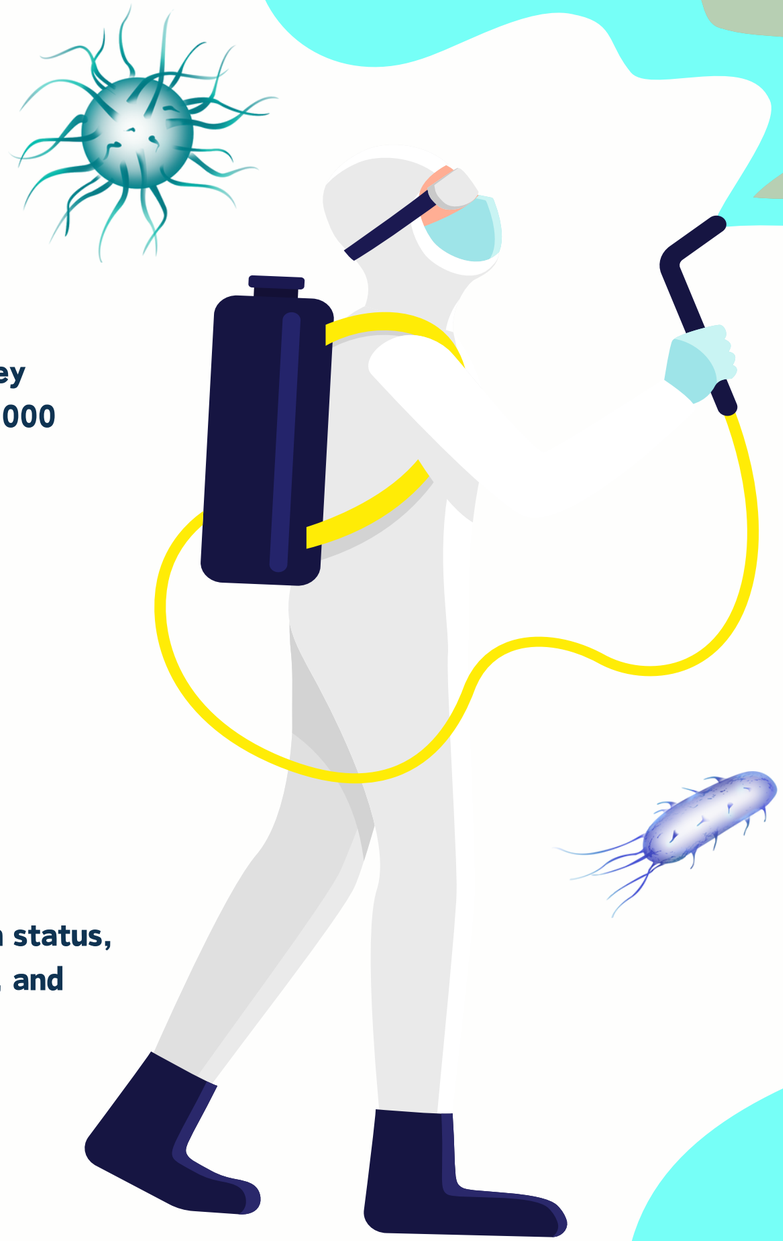
Data Understanding

Data Sources

The dataset originates from the NHFS National Flu Survey conducted in 2009, capturing responses from around 26,000 individuals regarding their vaccination status and various demographic, behavioral, and health factors.

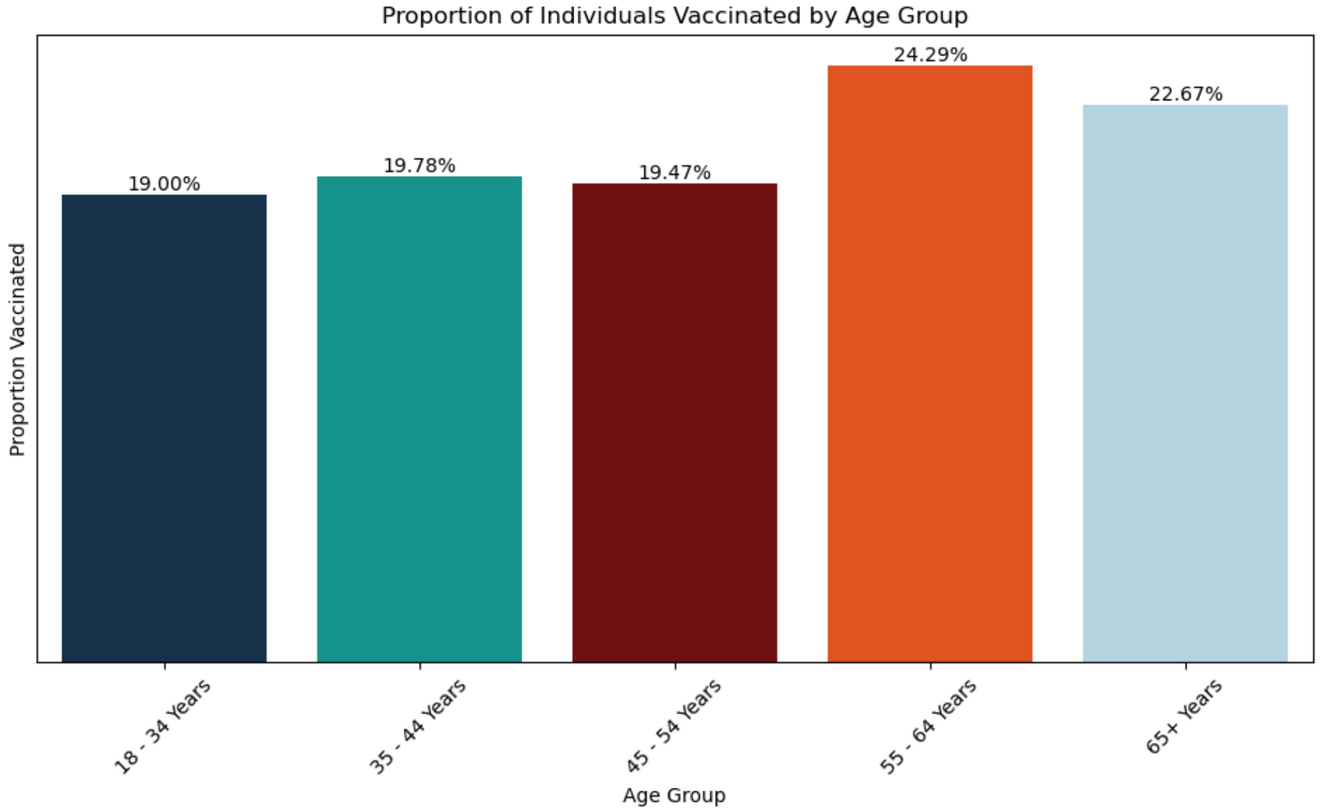
Features Overview

The dataset includes respondent information, vaccination status, health factors, behavioral factors, opinions and concerns, and knowledge levels related to H1N1 flu.



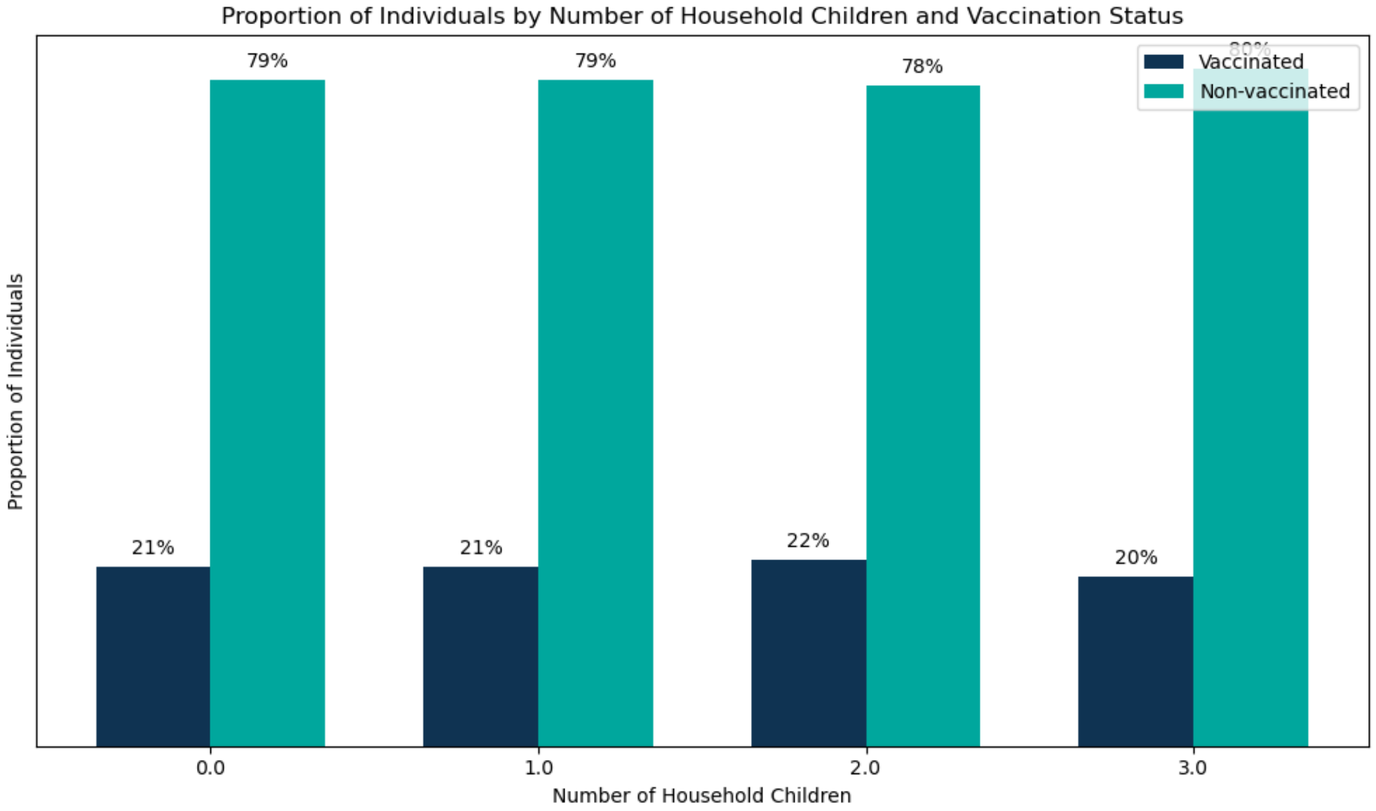
ANALYSIS

Age Group



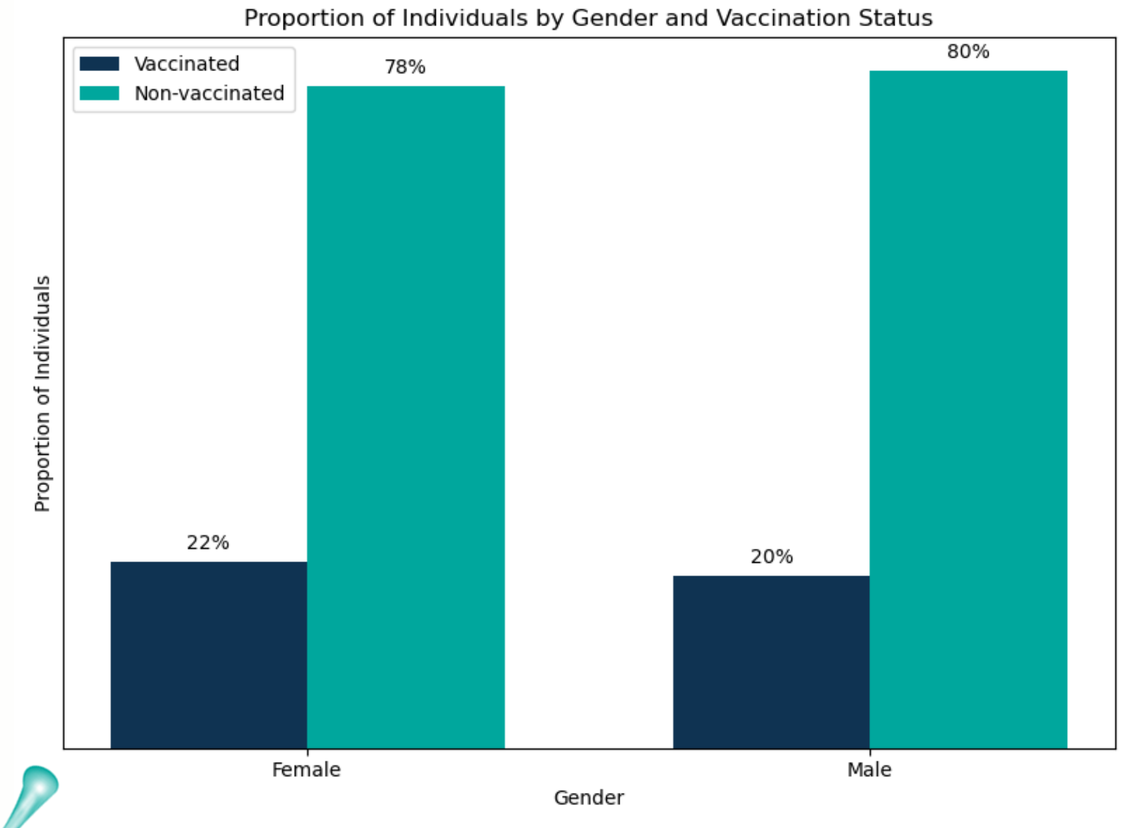
H1N1 vaccination rates vary by age, with higher rates among older individuals (55+) compared to younger groups (18–34), indicating age as a key factor influenced by susceptibility and awareness.

No.of children in a household



H1N1 vaccination rates slightly decrease as the number of children in the household increases, with around 21% vaccinated in households with no children or one child, compared to approximately 20% in households with two or three children.

Gender

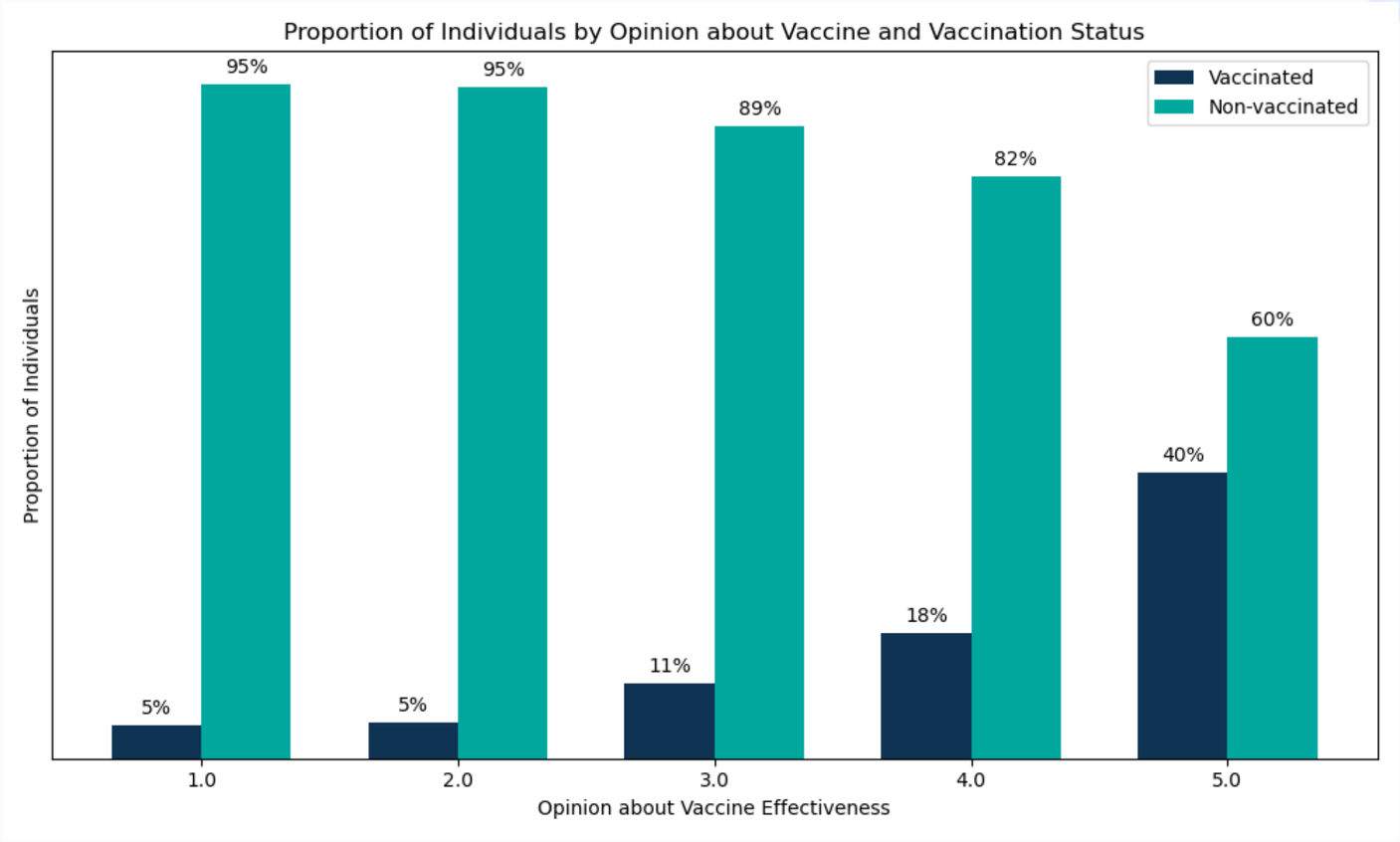


78%

80%

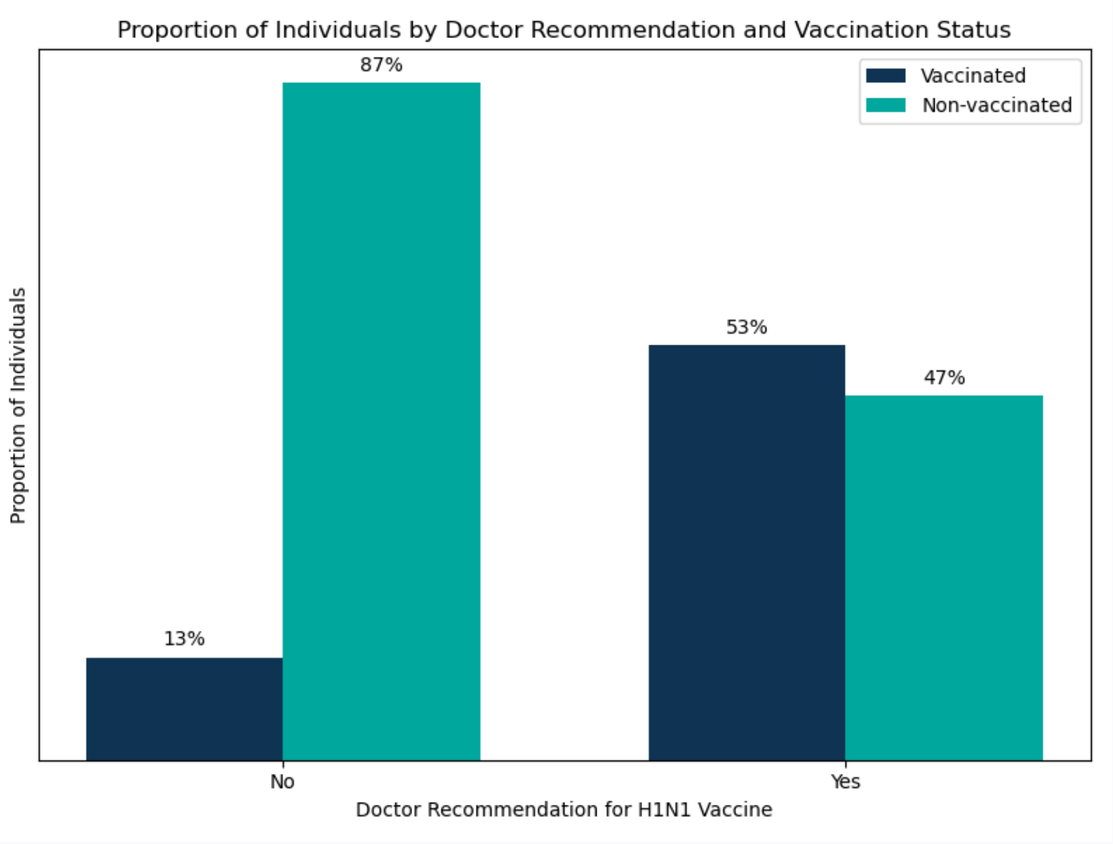
Both genders exhibit similar H1N1 vaccination rates, with slightly more females vaccinated. Any differences are minor and could be influenced by factors like healthcare access, perceived virus susceptibility, socioeconomic status, and vaccination attitudes.

Opinion



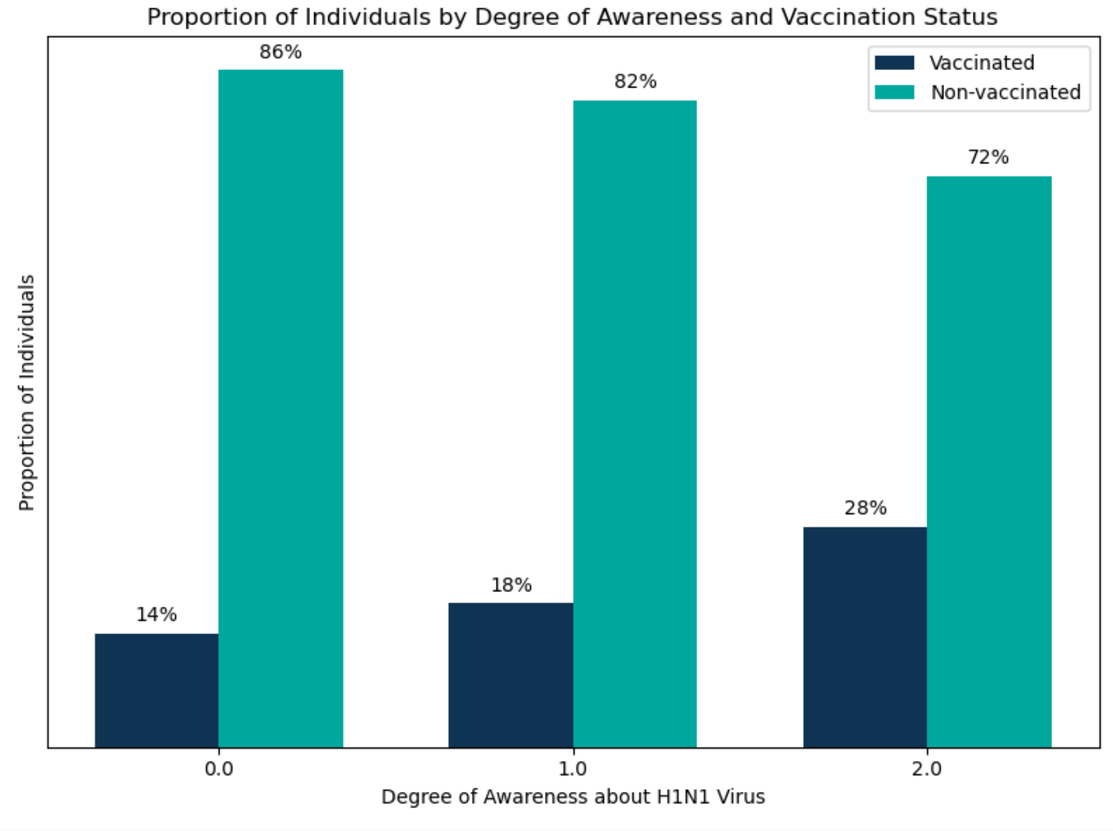
Perceptions of vaccine effectiveness strongly influence vaccination decisions: those perceiving low effectiveness show low vaccination rates, while moderate and high perception groups exhibit higher rates, with 17.70% and 40.48% respectively. Addressing public perceptions is crucial for promoting vaccination uptake and public health initiatives.

Doctors Recommendation



Doctor recommendation strongly influences H1N1 vaccination: 86.88% without it chose not to vaccinate, while 53.24% with it opted for vaccination. Healthcare professionals play a crucial role in promoting vaccination and shaping public health behaviors.

Knowledge awareness



Less knowledge about H1N1 flu means fewer vaccinations (85.59% for no knowledge, 81.82% for limited knowledge), while higher knowledge leads to more vaccinations (27.81%). This shows the importance of education in encouraging vaccinations during flu outbreaks.

Modelling

Modelling Context

Big Problem

False Positive

Predicitng that people got vaccine when they actually did not

NOT a Big Problem

False negative

Predicting people did not get vaccine when they actually did

FALSE POSITIVE

FALSE NEGATIVE

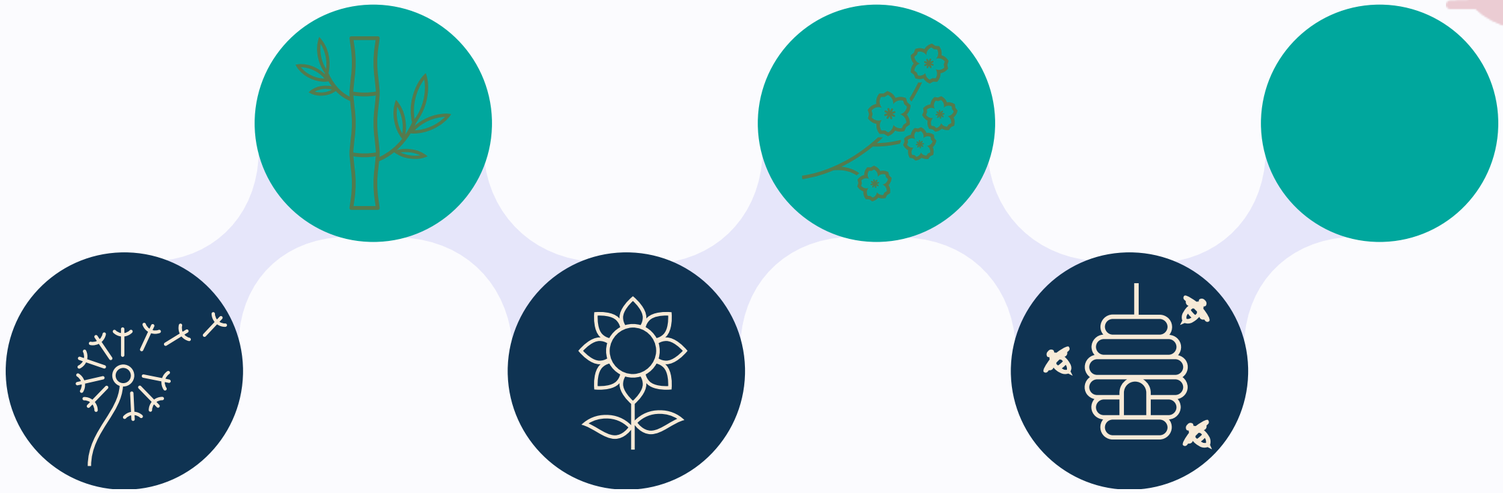
Evaluation Metrics

Precision

Precision measures how accurate the model is when it predicts someone is vaccinated. High precision means the model rarely makes mistakes, minimizing false positives.

F1 score

Harmonic mean of precision and recall, offering a balanced assessment of performance.



Accuracy

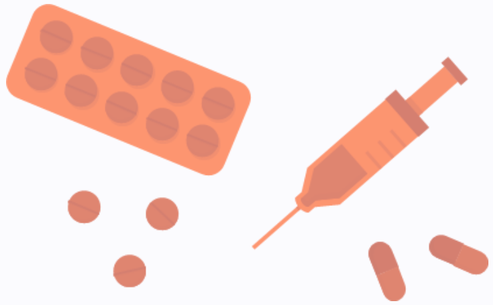
indicates how well each algorithm performs in predicting whether an individual is vaccinated against H1N1 or not.

Recall

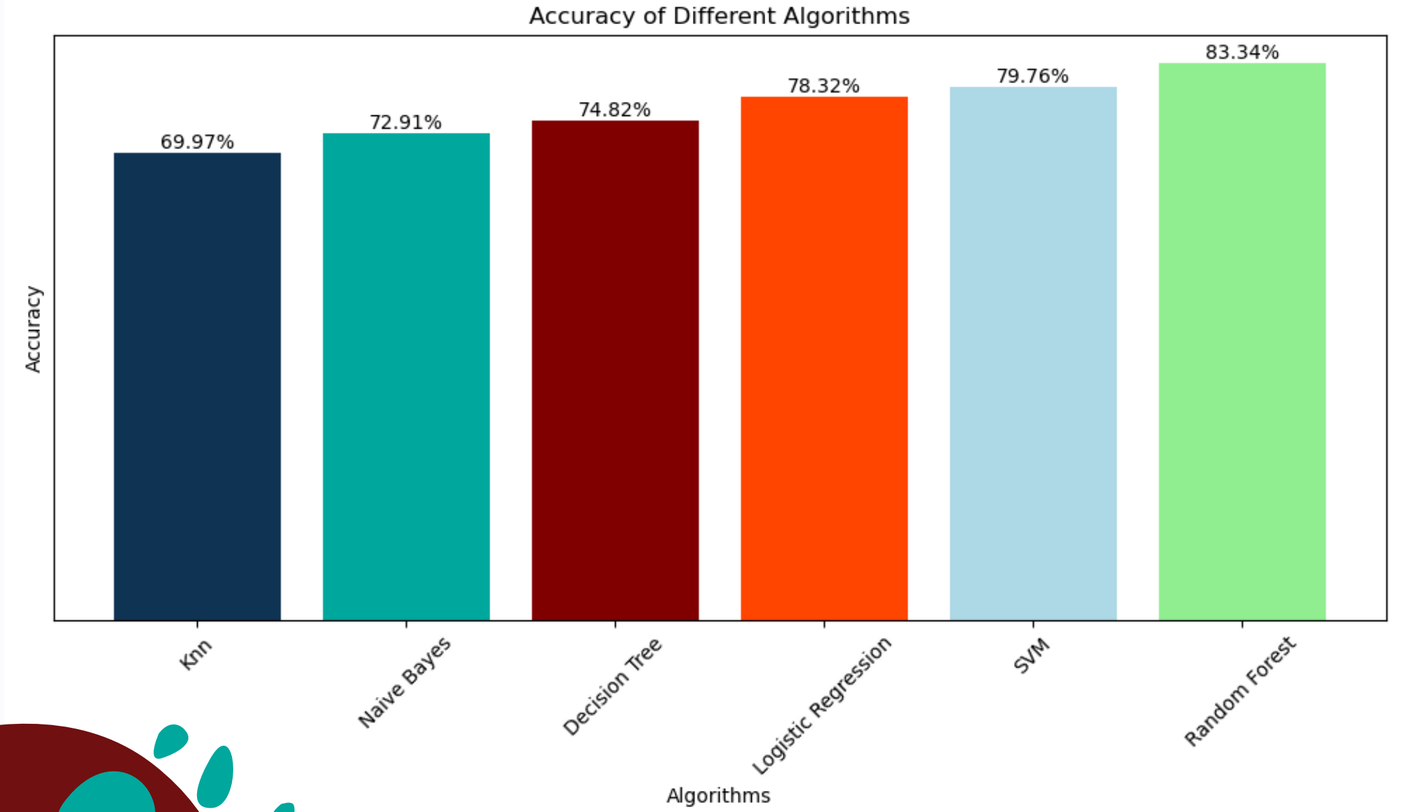
Proportion of actual positives correctly identified by the model, secondary to precision.

ROC Curve

AUC quantifies discrimination ability, with higher values indicating better performance.

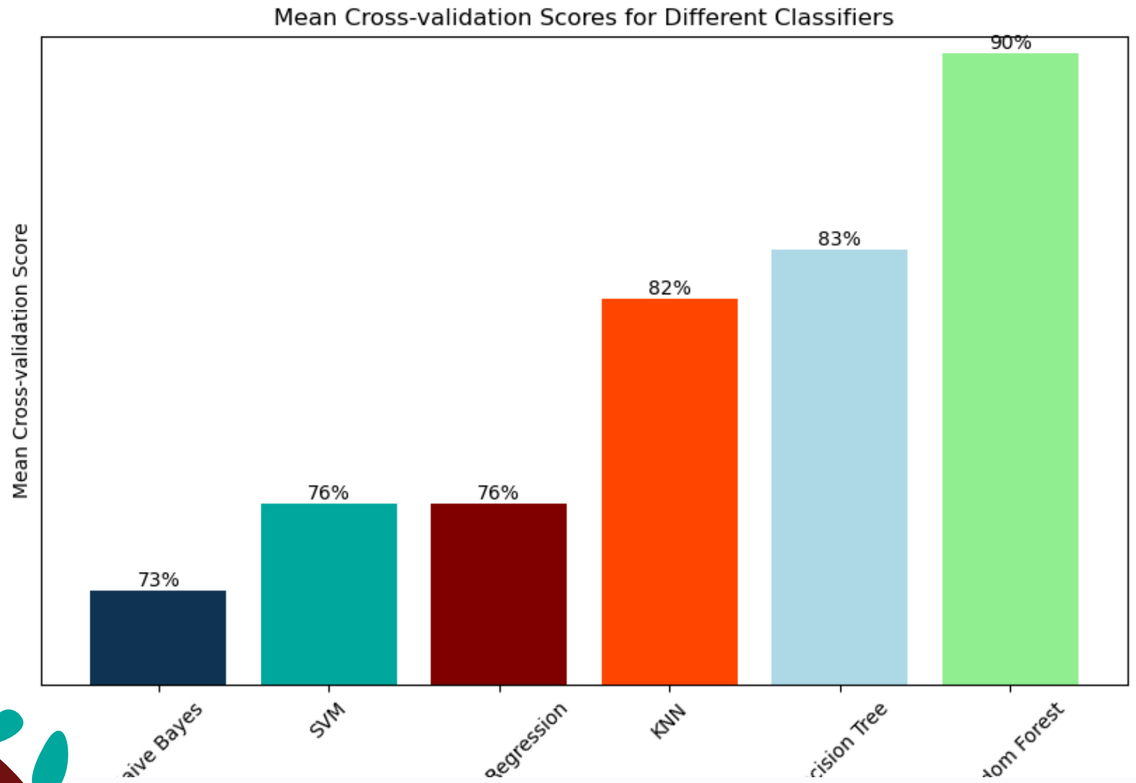


Baseline Model Accuracy



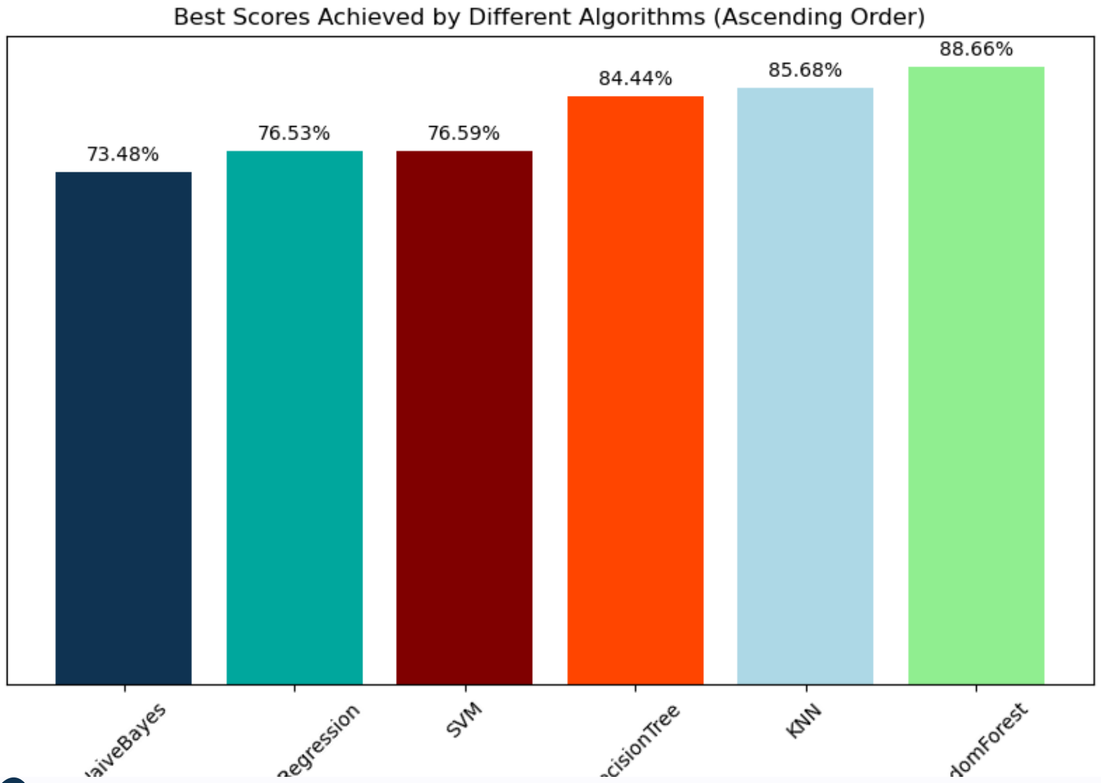
Random Forest outperforms the other algorithms with 83.34%, followed by SVM and Logistic Regression. Naive Bayes and Decision Tree have slightly lower accuracies, while KNN has the lowest accuracy among the listed algorithms.

Cross Validation Scores

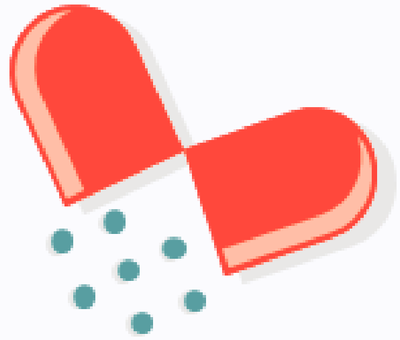


Random Forest demonstrates the highest performance among the models, followed by Decision Tree and KNN, while Naive Bayes, SVM, and Logistic Regression perform relatively close to the baseline accuracy.

Hypertuned model accuracy

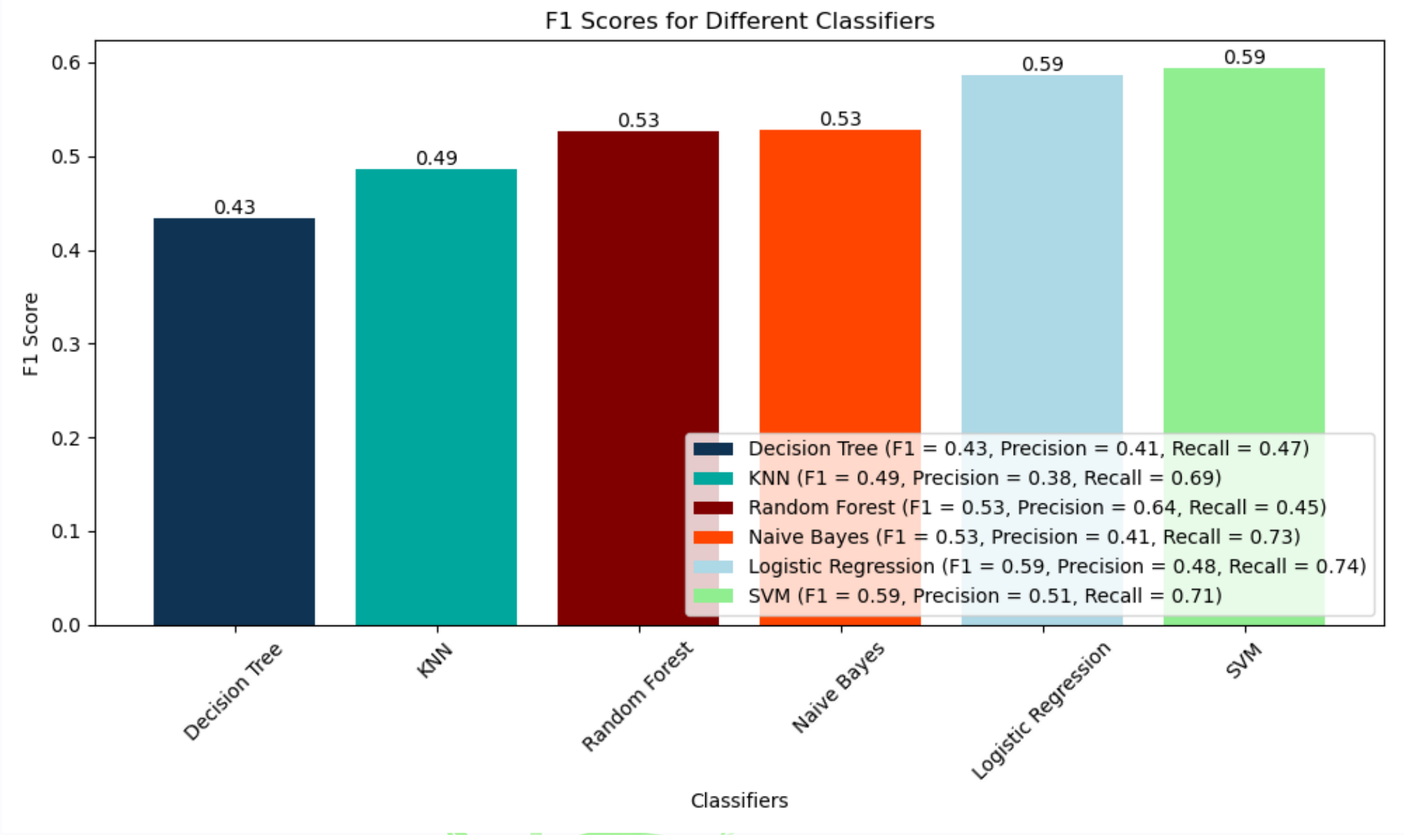


All models generally perform consistently between training and validation data, KNN stands out for its significant improvement in generalization performance, indicating its robustness in making accurate predictions on unseen data.



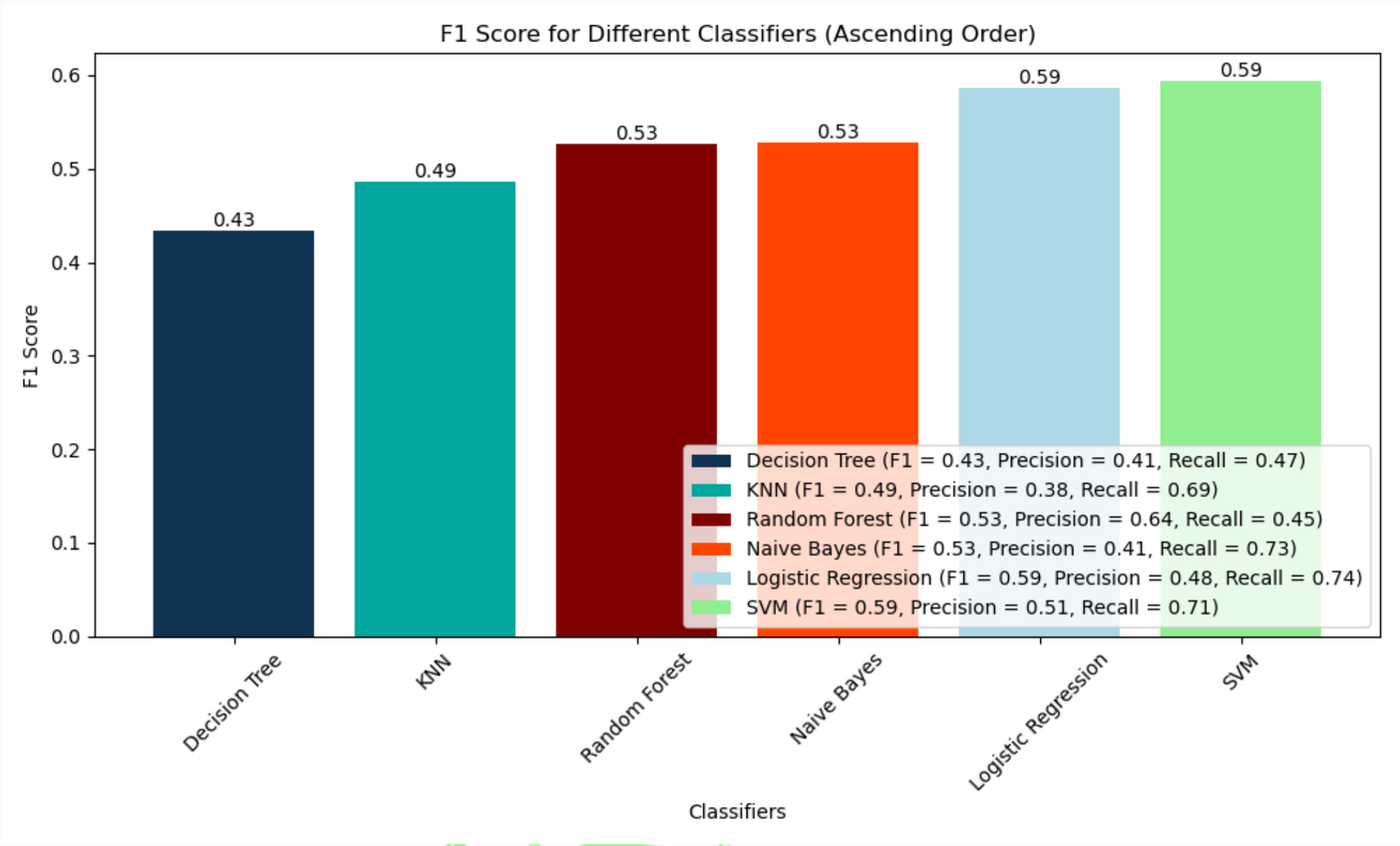
Baseline Models f1 score

Logistic Regression and SVM have the highest F1 scores among the listed classifiers. Therefore, they are deemed to have performed relatively better in terms of balancing precision and recall, indicating their effectiveness in correctly identifying positive instances while minimizing false positives.



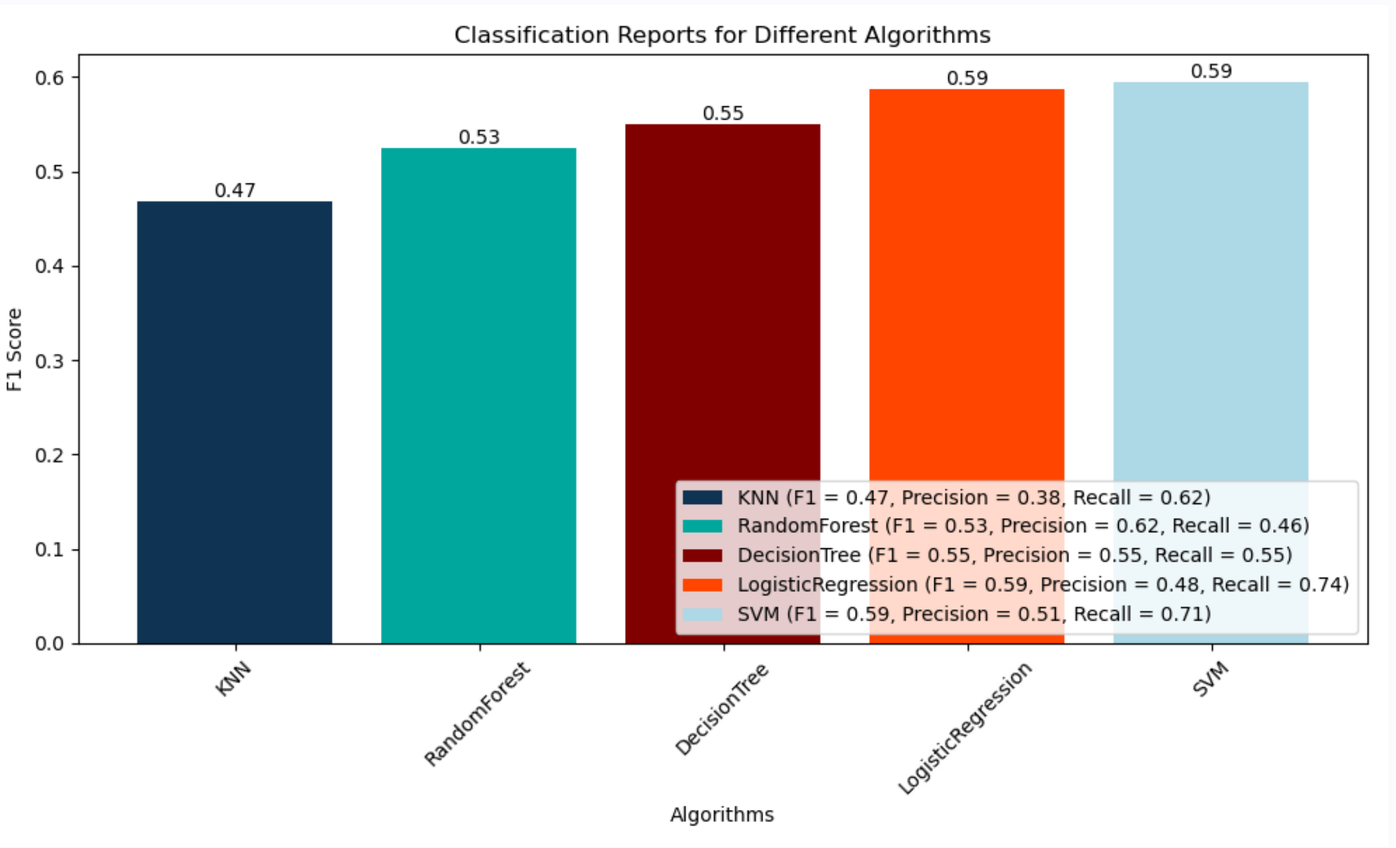
cross validation f1 score

Logistic Regression and SVM have the highest F1 scores among the listed classifiers. Therefore, they are deemed to have performed relatively better in terms of balancing precision and recall, indicating their effectiveness in correctly identifying positive instances while minimizing false positives.



Hypertuned model f1 score

Logistic Regression, SVM, and Decision Tree have the highest F1 scores among the listed algorithms. Therefore, they are considered to perform relatively better in terms of balancing precision and recall, suggesting their effectiveness in correctly identifying positive instances while minimizing false positives.



Feature Importance

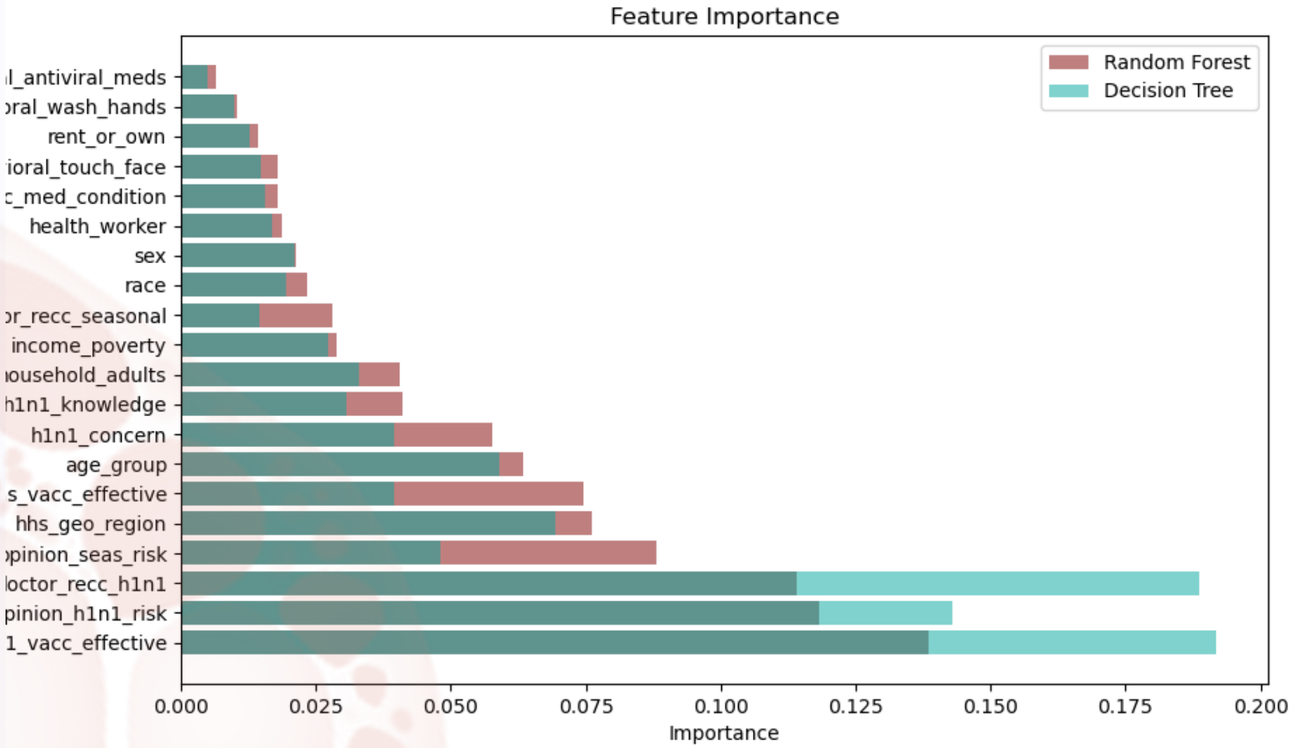
Decision tree & Random forest

Overall, both classifiers seem to prioritize "opinion_h1n1_vacc_effective" as the most influential feature, suggesting its strong predictive power in determining vaccination behavior.

opinion_h1n1_vacc_effective

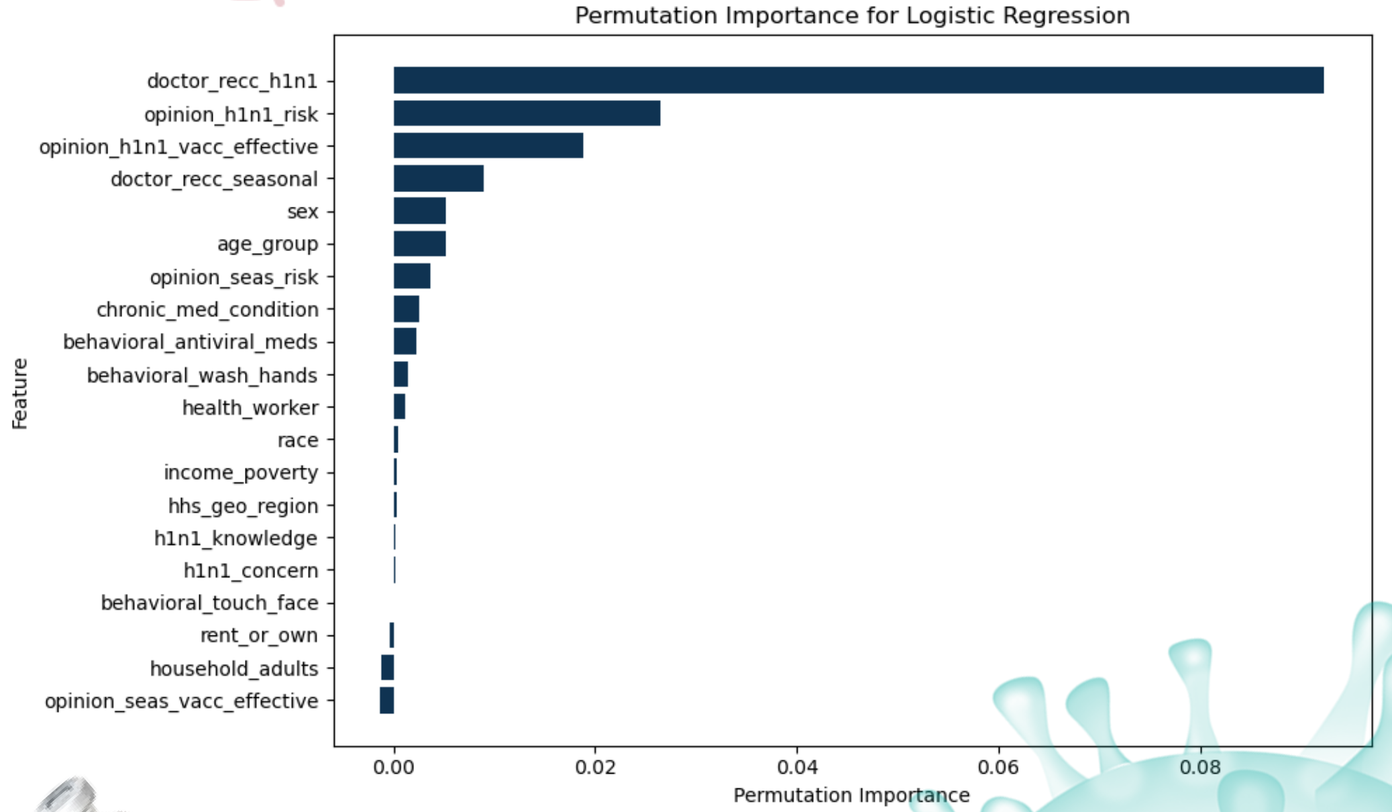
opinion_h1n1_risk

doctor_recc_h1n1



Logistic Regression

Overall, this indicates the relative importance of each feature according to the Logistic Regression model, with "doctor_recc_h1n1" being the most influential feature in predicting vaccination behavior.



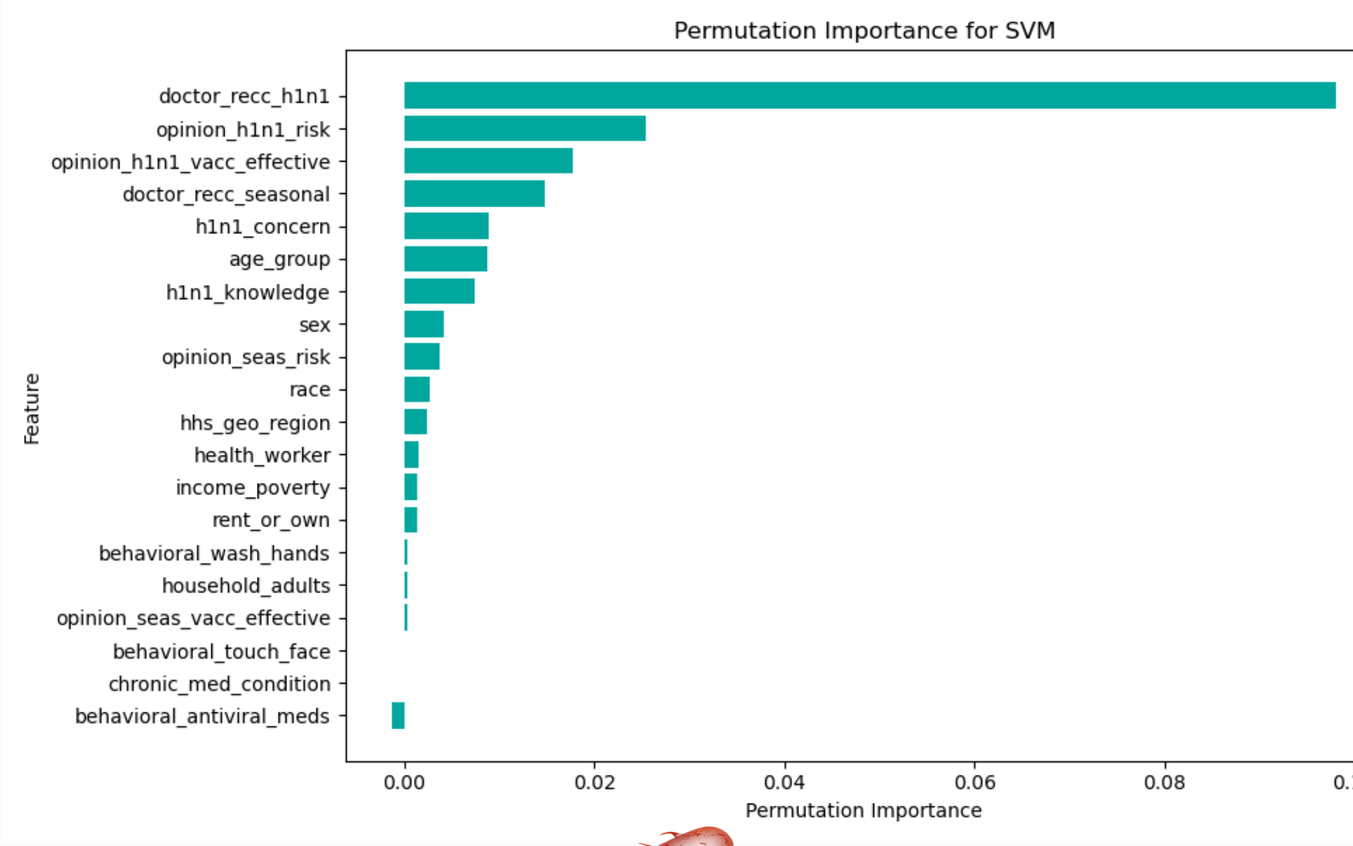
doctor_recc_h1n1

opinion_h1n1_risk

opinion_h1n1_vacc_effective

Support Vector Machine

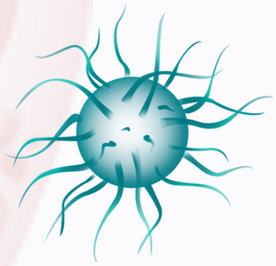
Overall, this indicates the relative importance of each feature according to the Support Vector Machine model, with "doctor_recc_h1n1" being the most influential feature in predicting vaccination behavior.



doctor_recc_h1n1

opinion_h1n1_risk

opinion_h1n1_vacc_effective

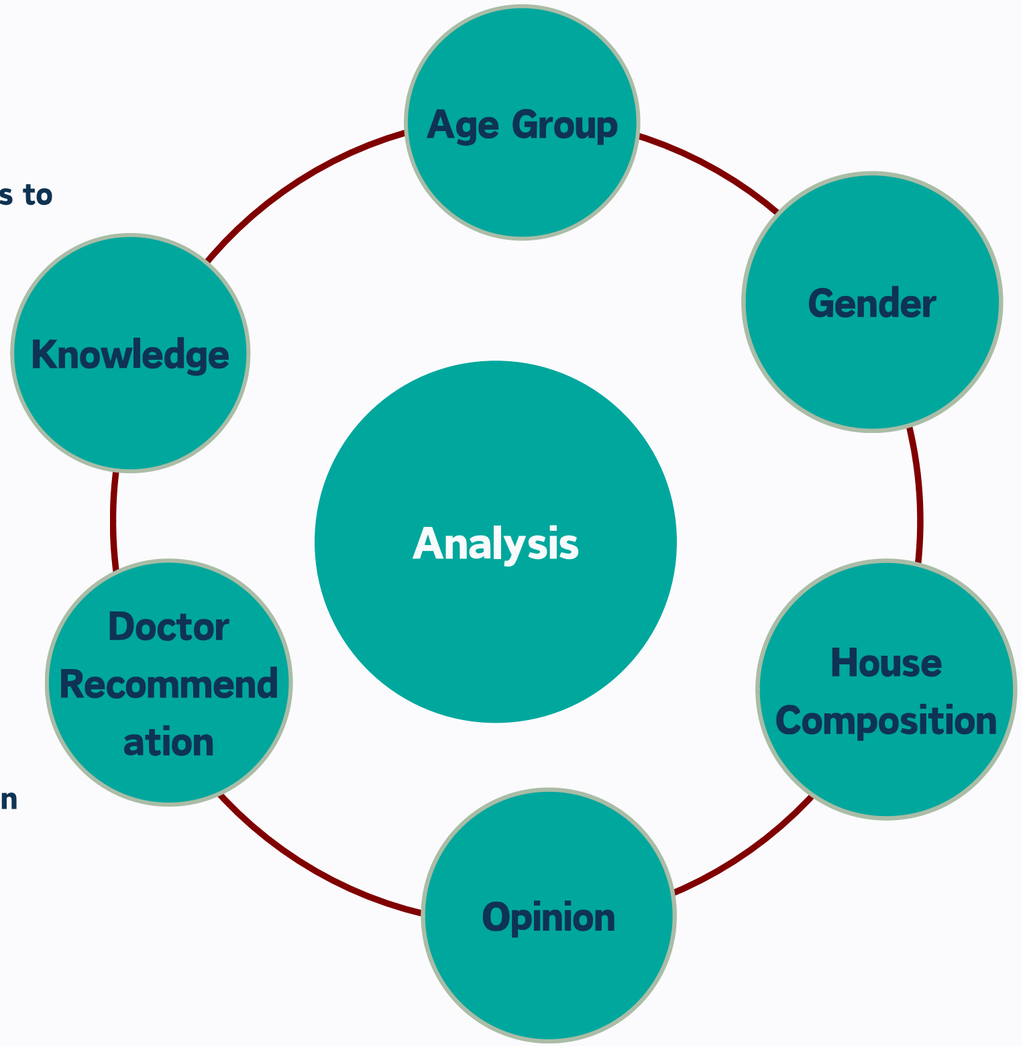


Conclusion & Recommendation

Findings

Older age groups are more vaccinated.

Greater flu knowledge leads to higher vaccination rates.



Both genders have similar vaccination rates.

Vaccination rates slightly decrease with more children in the household.

Doctor recommendations significantly boost vaccination rates.

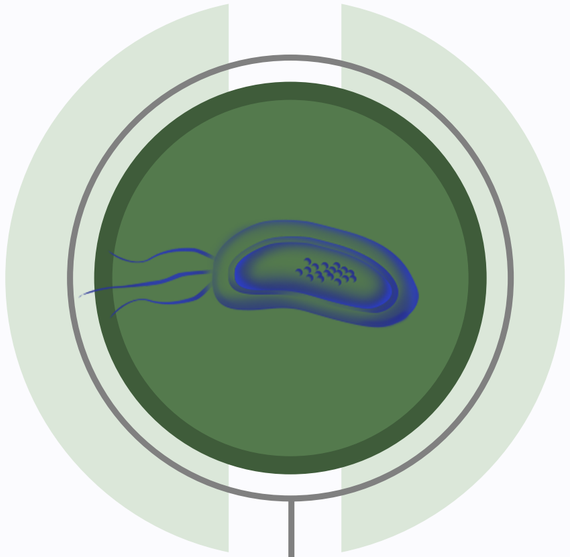
Perceived vaccine effectiveness influences vaccination likelihood.

Model Selection

Both Logistic Regression and SVM consistently demonstrated the highest F1 scores among the listed algorithms. This suggests that they excel in achieving a balance between precision and recall, which is crucial for accurately identifying positive instances while minimizing false positives.

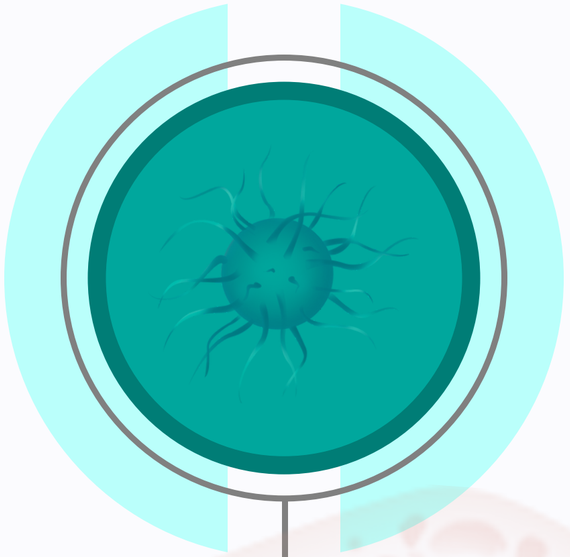
If interpretability and simplicity are paramount considerations, Logistic Regression may be the preferable option. On the other hand, if sacrificing some interpretability in favor of potentially higher accuracy is acceptable, SVM could be the more suitable choice.

Recommendation



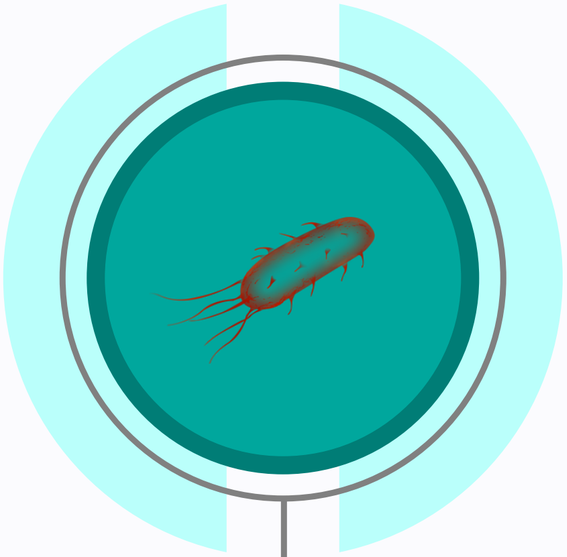
Targeted Education Campaigns

Develop campaigns to raise awareness of vaccination benefits, especially targeting younger age groups and individuals with limited knowledge of H1N1 flu.



Healthcare Provider Training

Train healthcare professionals to advocate for H1N1 vaccination, enhancing uptake among eligible individuals.



Public Messaging

Address misconceptions about vaccine effectiveness through public messaging, emphasizing vaccination's role in preventing disease spread.

Research and Monitoring

Continuously monitor vaccination rates and factors influencing behavior to tailor interventions effectively.

Community Engagement

Involve communities in vaccination promotion strategies, considering their concerns and preferences.

Thank you
Doctors for
Saving lives

The end