Who Gets Vaccinated?

Machine Learning Predictions of H1N1 and Seasonal Flu Vaccine Uptake

Why This Matters

➡ Vaccination saves lives, but not everyone chooses to get vaccinated.

➡ By understanding who is more or less likely to take vaccines, public health officials can:

• Predict and profile hesitant groups

• Improve outreach and messaging

• Make informed decisions in future pandemics

Project Overview

- • Predict the likelihood of receiving H1N1 and seasonal flu vaccines

- • Multi-label probabilistic classification

- • Supports targeted public health strategies

Who Benefits from This?

- Public health authorities (CDC, WHO)
- Healthcare providers & outreach programs
- Researchers in epidemiology and behavioral health
- Policy-makers targeting equitable healthcare access

Dataset and Tools

- Data: The data originates from the National 2009 H1N1 Flu Survey, conducted in the United States to understand vaccine behaviors during the H1N1 pandemic. Source: 2009 National H1N1 Flu Survey (N = ~26,000)

- Tools: Python, Pandas, scikit-learn, Numpy

- Models: Logistic Regression, Random Forest

- Metrics: ROC AUC, F1 Score, Accuracy

Modeling Pipeline

1. Data Cleaning & Imputation

2. One-Hot Encoding & Scaling

3. Model Training and Validation

4. Feature Selection and Final Prediction
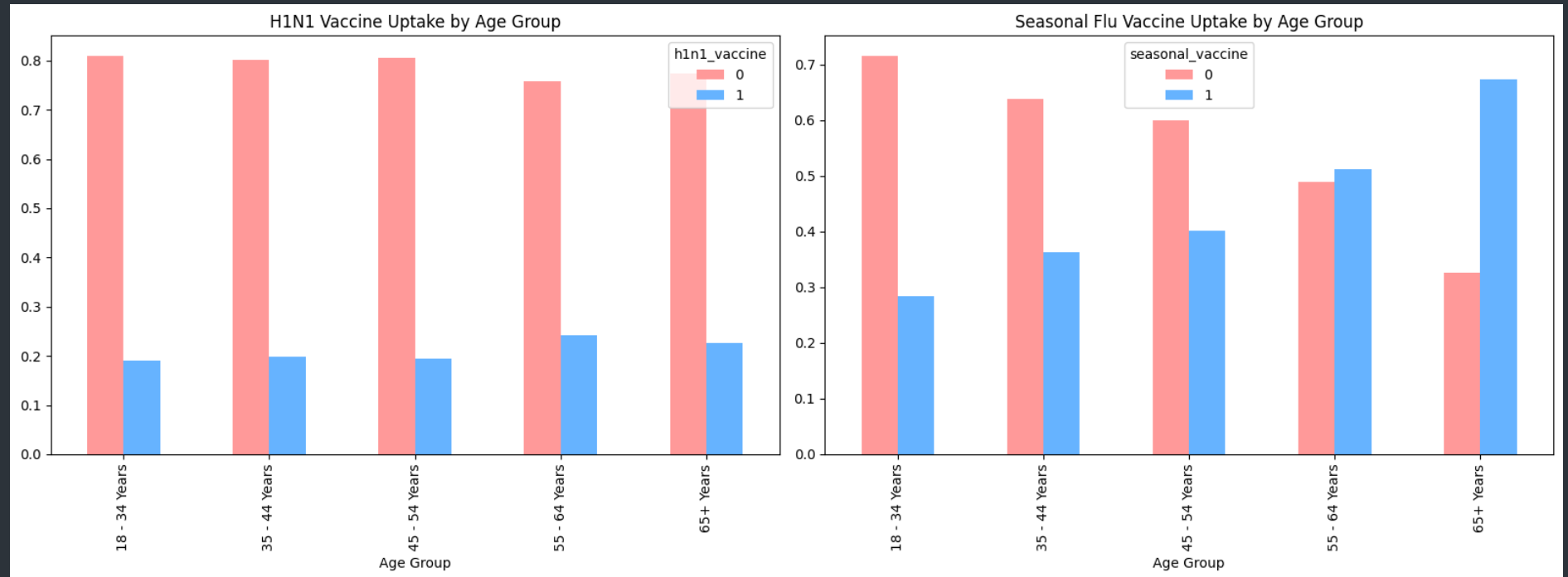
Final Model : Logistic Regression Chosen for interpretability and efficiency
                         Comparable performance to Random Forest

                         Simpler deployment and communication
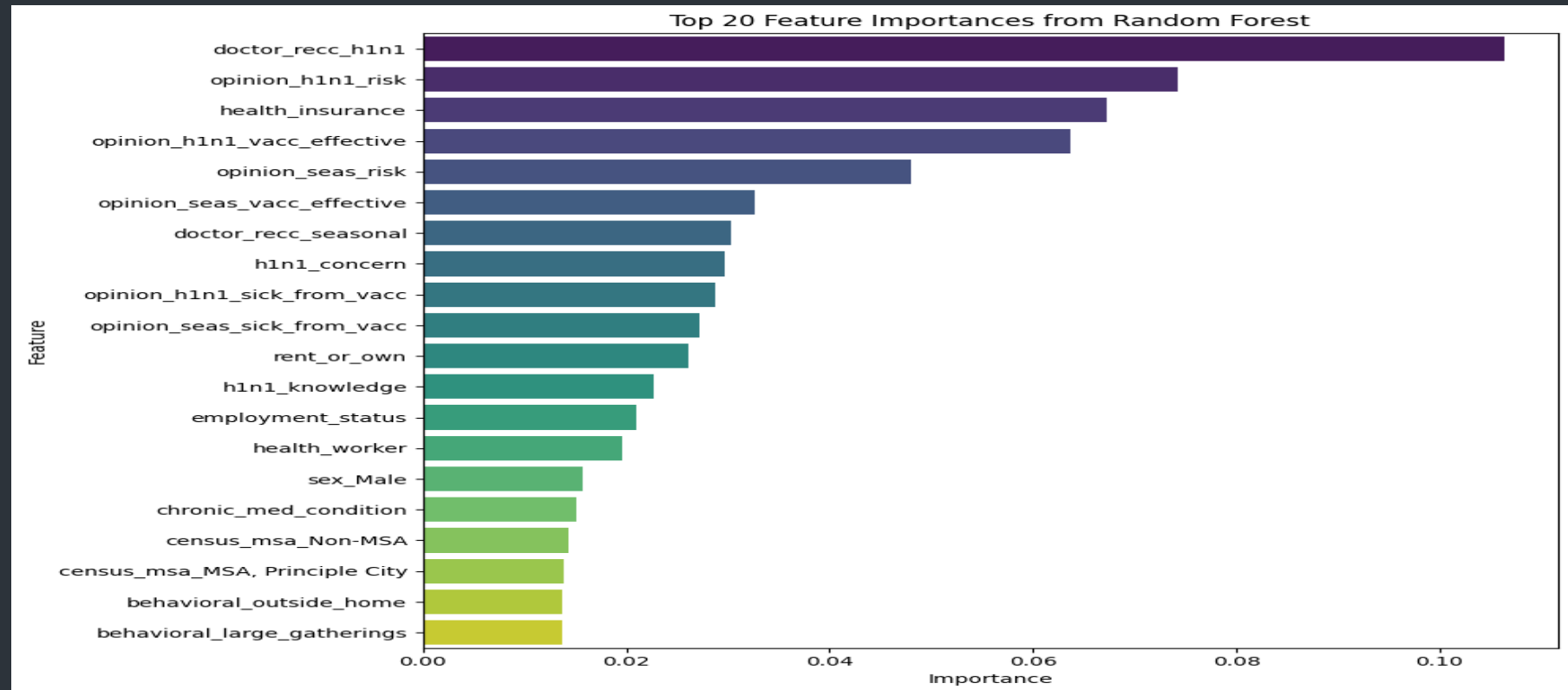
Model Performance (Validation Set)

- H1N1 Vaccine:

- Accuracy: ~77.7% | ROC AUC: 0.835 | F1: ~0.68

- Seasonal Flu Vaccine:

- Accuracy: ~77.4% | ROC AUC: 0.850 | F1: ~0.75

Vaccine uptake by age group

Top 20 Features Determining Vaccine Uptake

Top 20 Feature Importances from Random Forest

Key Findings

- Doctor recommendation is a strong driver of vaccine uptake.
- Older individuals are more likely to get vaccinated.
- Those with chronic illnesses and high health awareness vaccinate more.
- Many still miss vaccines due to lack of information or trust.

What Can Be Done

- Encourage doctors to recommend vaccines more actively.
- Target public campaigns at younger and healthy-feeling individuals.
- Focus resources on groups with low access or awareness.
- Use data to identify and support hesitant populations.
- Target outreach to under-advised populations
- Leverage behavioral indicators for messaging
- Involve healthcare providers more actively
- Use this model to forecast future vaccine hesitancy

Final Takeaways

- • Data-driven insights help identify who needs vaccine outreach.

- • Machine learning reveals behavioral and health-based patterns.

- • With the right strategy, we can improve public health outcomes.

- • These methods can support future pandemics and vaccine rollouts.

Conclusion

- Developed predictive models for vaccine uptake
- Logistic Regression selected as final model
- Models show strong generalization with interpretable insights
- Ready for public health deployment or further research

THANK YOU.

QUESTIONS?