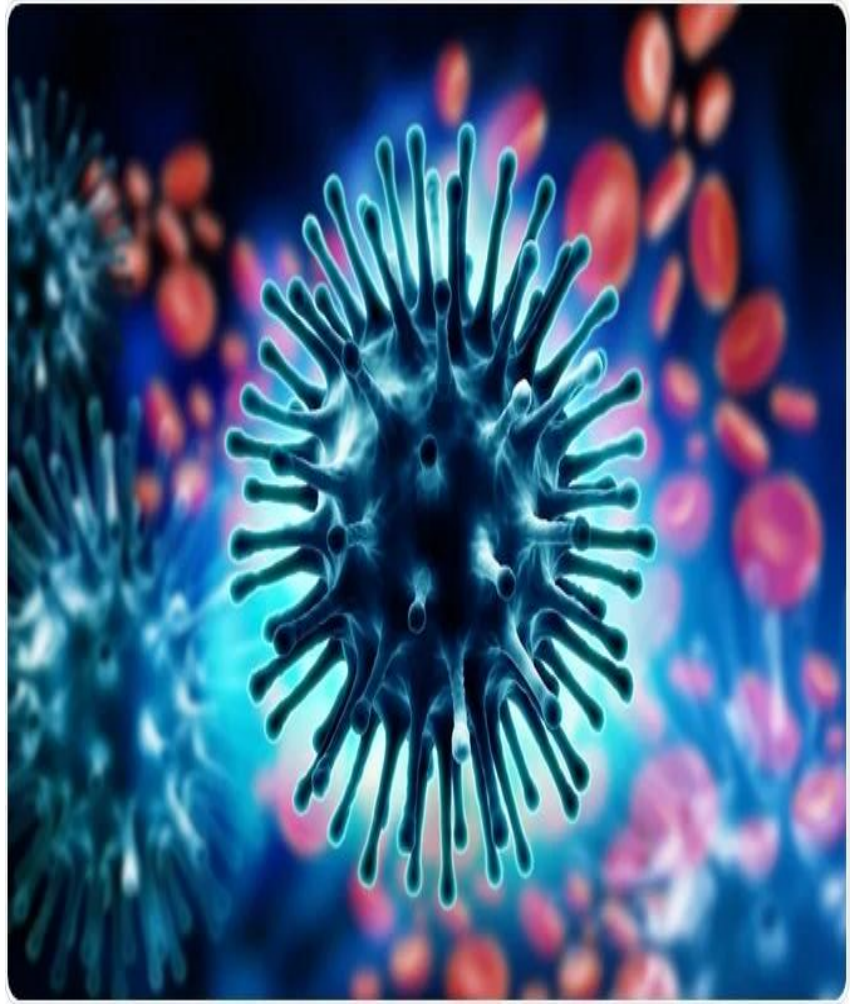


FLU VACCINE PREDICTION USING MACHINE LEARNING

A DATA-DRIVEN APPROACH TO ADDRESS VACCINE HESITANCY

DOREEN MURUGI, MORINGA SCHOOL

6/6/2025

OVERVIEW

- Predict H1N1 and Seasonal flu vaccine uptake
- Analyze individual factors influencing vaccination decisions
- Build and evaluate predictive models for actionable insights

BUSINESS UNDERSTANDING

- Vaccine hesitancy contributes to preventable disease spread
- Governments need data-driven tools to improve vaccine campaigns
- Identify population segments to target with awareness efforts

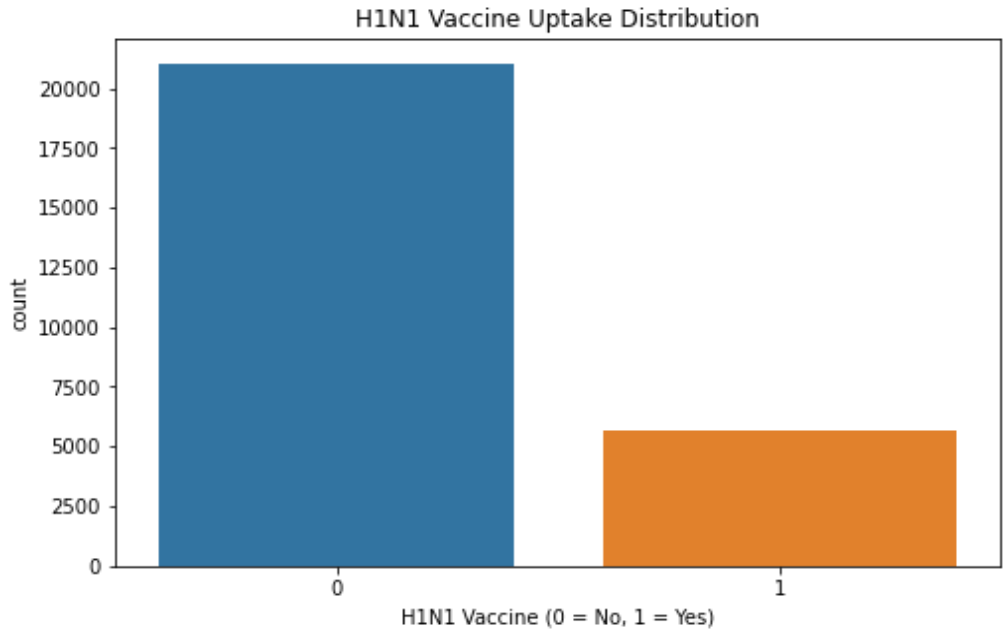
DATA UNDERSTANDING

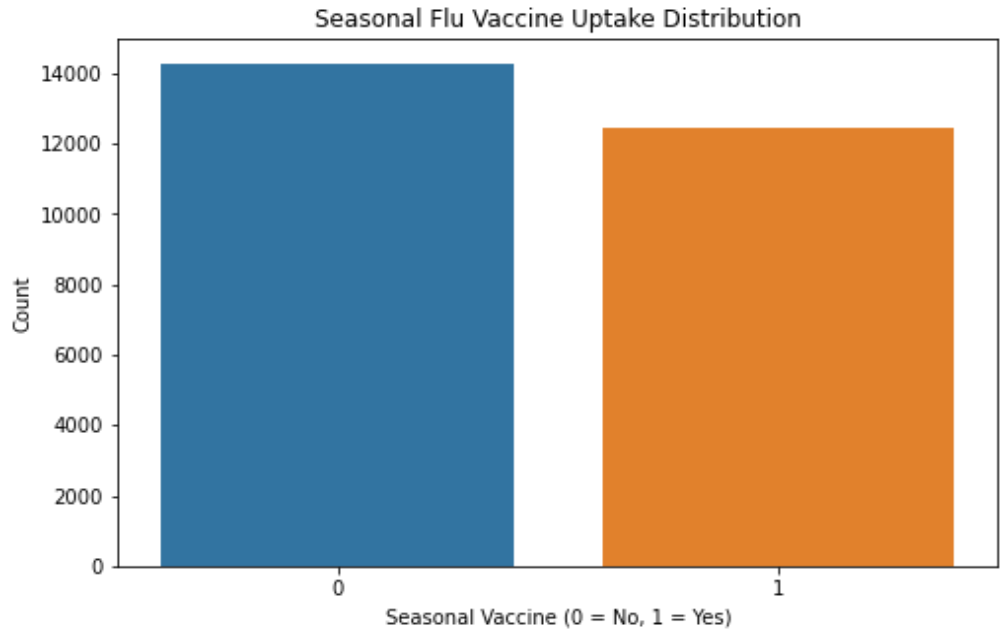
- Dataset: H1N1 Flu Vaccine Prediction
- Size: 26,707 rows, 32 features
- Features: Demographics, health behaviors, opinions
- Targets: H1N1 vaccine, Seasonal flu vaccine (binary)

opinion_h1n1_vacc_effective	opinion_h1n1_risk	opinion_h1n1_sick_from_vacc	opinion_seas_vacc_effective	opinion_seas_risk	opinion_seas_sick_from_vacc	age_group	education	race	sex
3	1	2	2	1	2	55 - 64 Years	< 12 Years	White	Female
5	4	4	4	2	4	35 - 44 Years	12 Years	White	Male
3	1	1	4	1	2	18 - 34 Years	College Graduate	White	Male

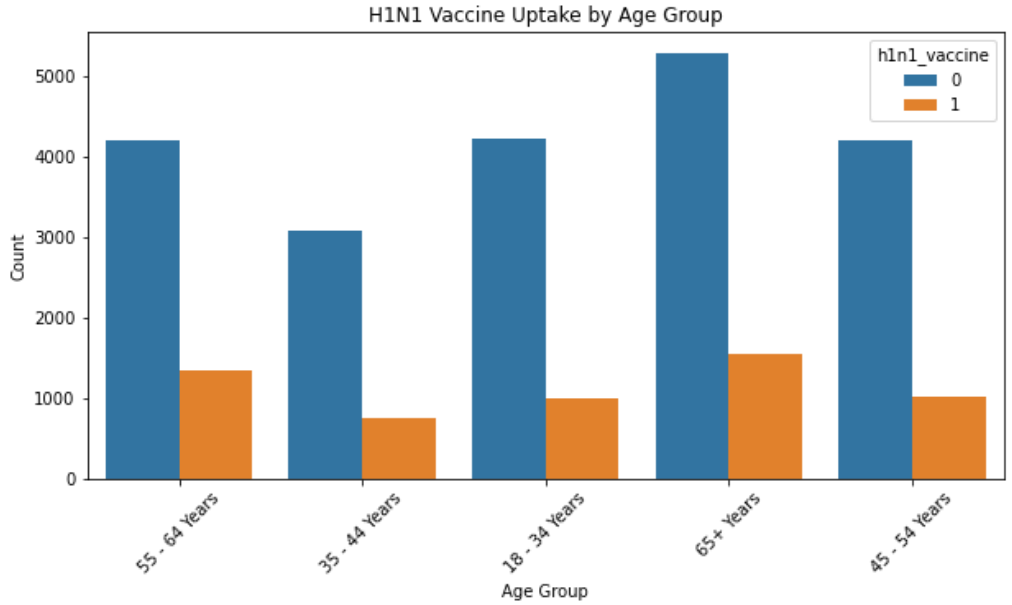
Sample of the training dataset

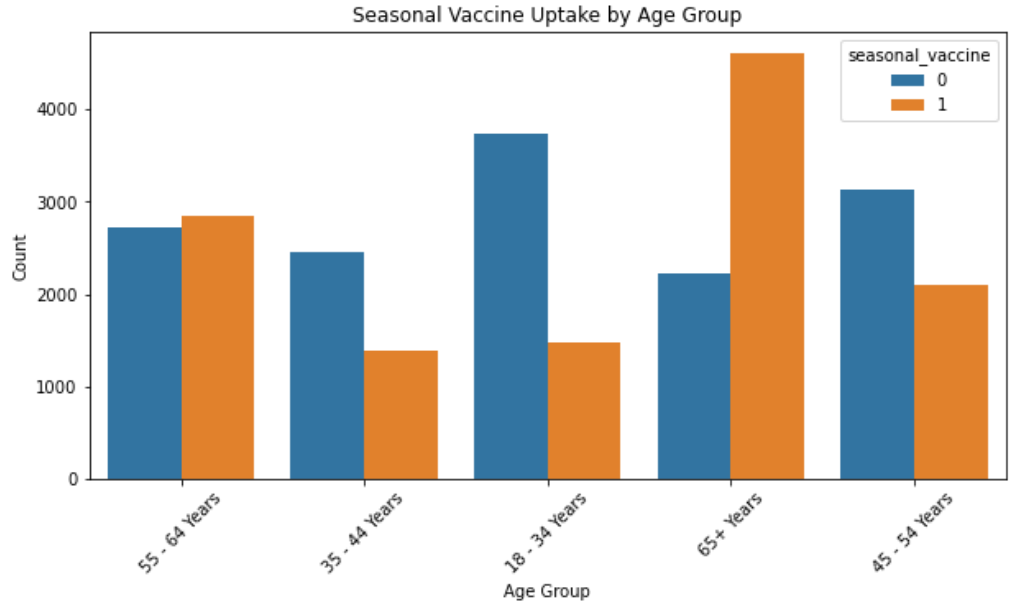
H1N1 AND SEASONAL FLU VACCINE UPTAKE DISTRIBUTION (0 = NOT VACCINATED, 1 = VACCINATED)





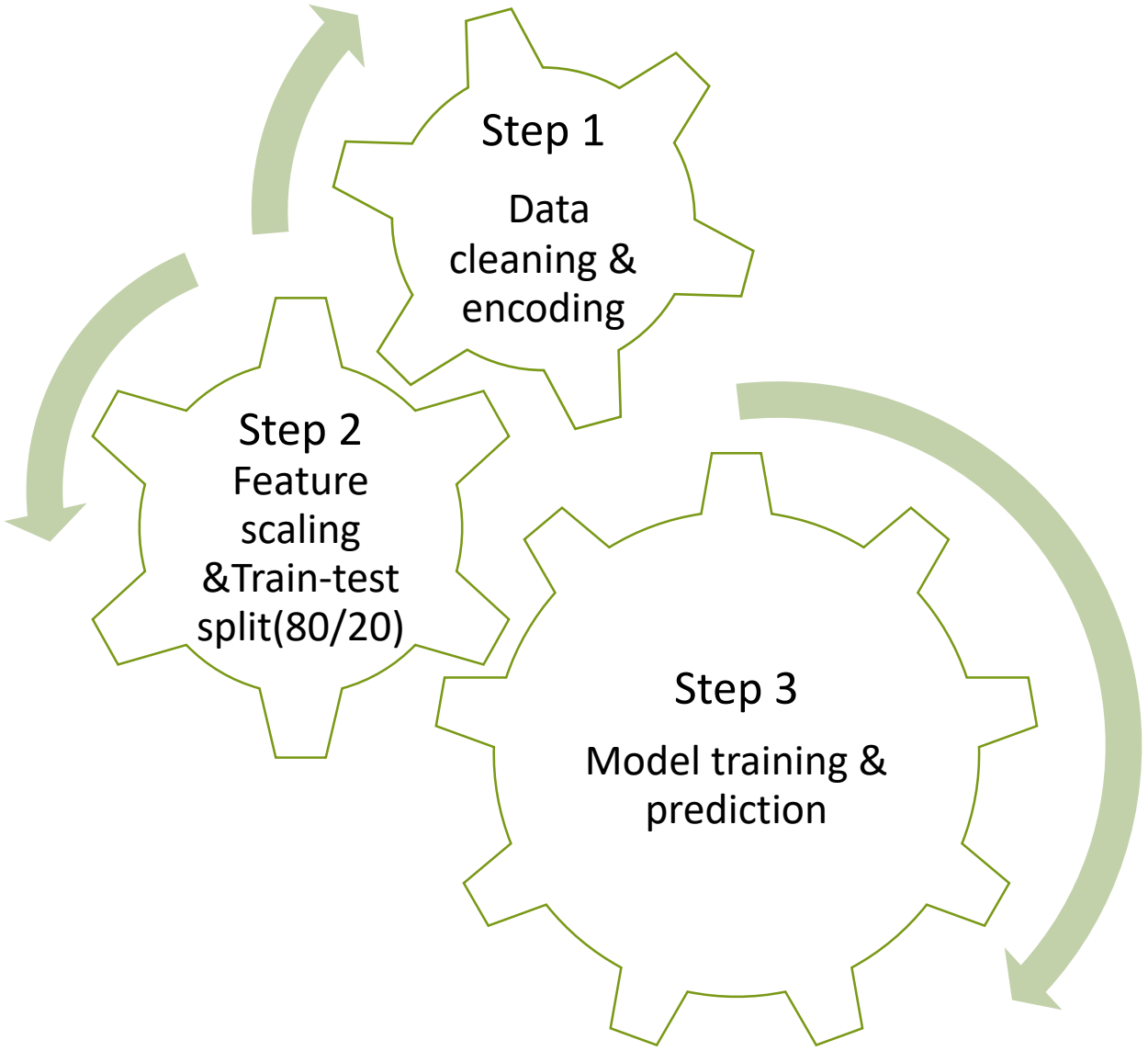
H1N1 AND SEASONAL FLU VACCINE UPTAKE DISTRIBUTION BY AGE GROUP





MODELLING

- Models Used:
  - Logistic Regression
  - Decision Trees

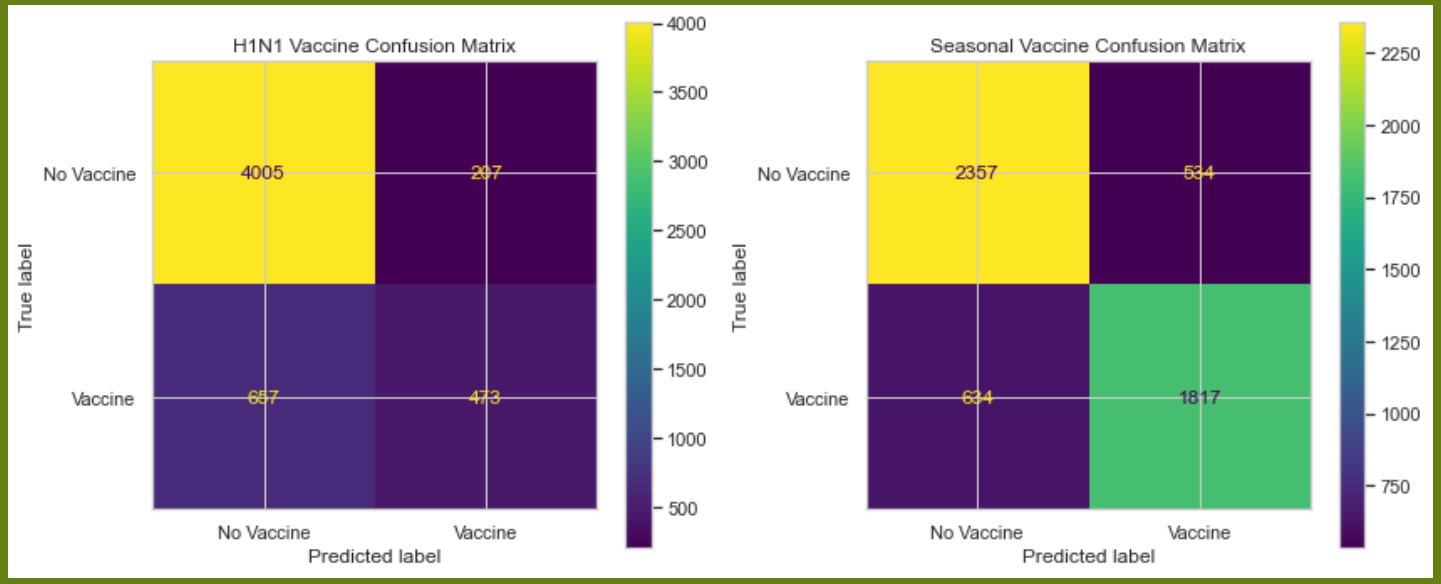


EVALUATION

METRIC: ROC AUC Score

Logistic Regression

H1N1: 0.826
Seasonal: 0.852

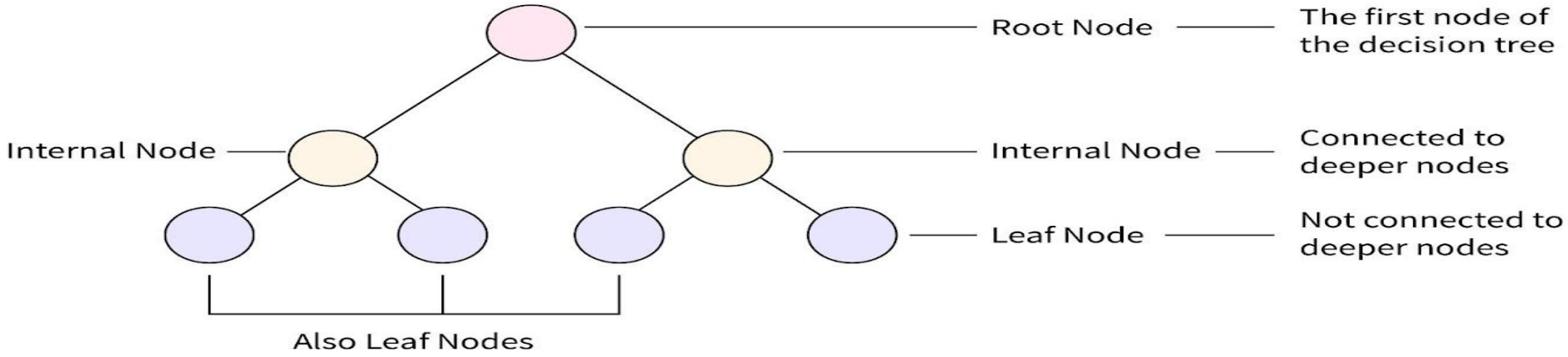


True Positives (bottom-right): Correctly predicted vaccinated individuals.

True Negatives (top-left): Correctly predicted non-vaccinated individuals.

False Positives (top-right): Predicted vaccinated but actually not.

False Negatives (bottom-left): Predicted not vaccinated but actually are.



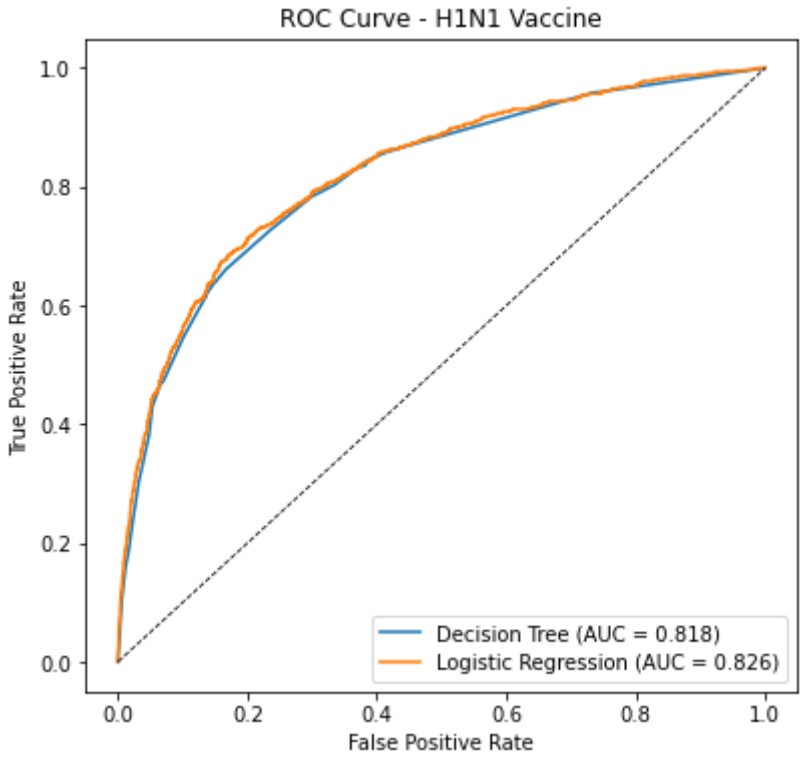
Decision Tree

H1N1: 0.818
Seasonal: 0.828

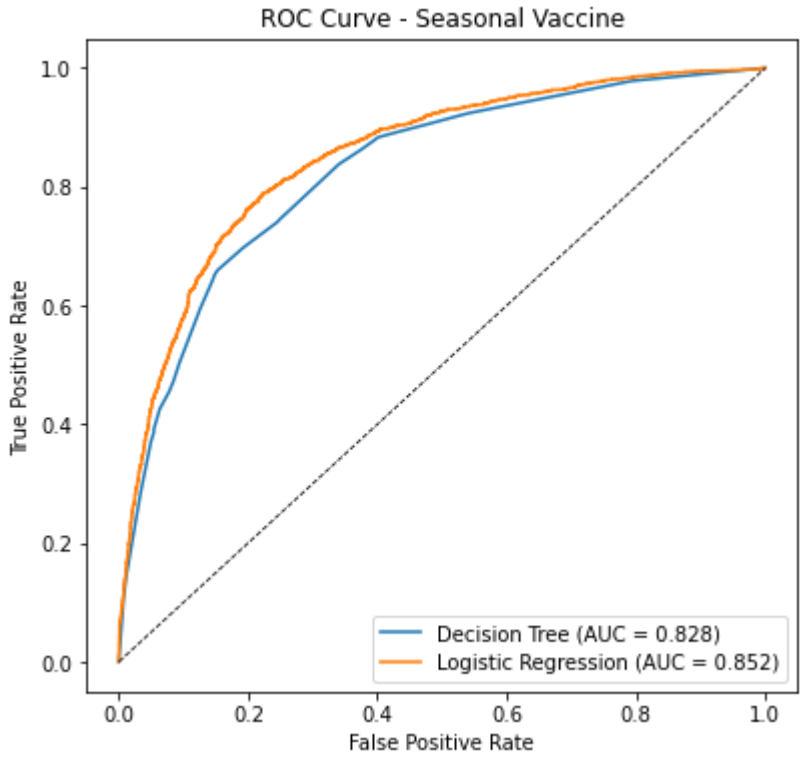
Comparison Between the Two Models

Logistic Regression slightly outperformed Decision Tree.

ROC – H1N1 VACCINE



ROC – SEASONAL VACCINE



KEY INSIGHTS

- H1N1 Predictors:
  - Doctor recommendation
  - Belief in vaccine effectiveness
  - Perceived personal risk
- Seasonal Predictors:
  - Risk perception
  - Trust in doctors
  - Age group (65+)

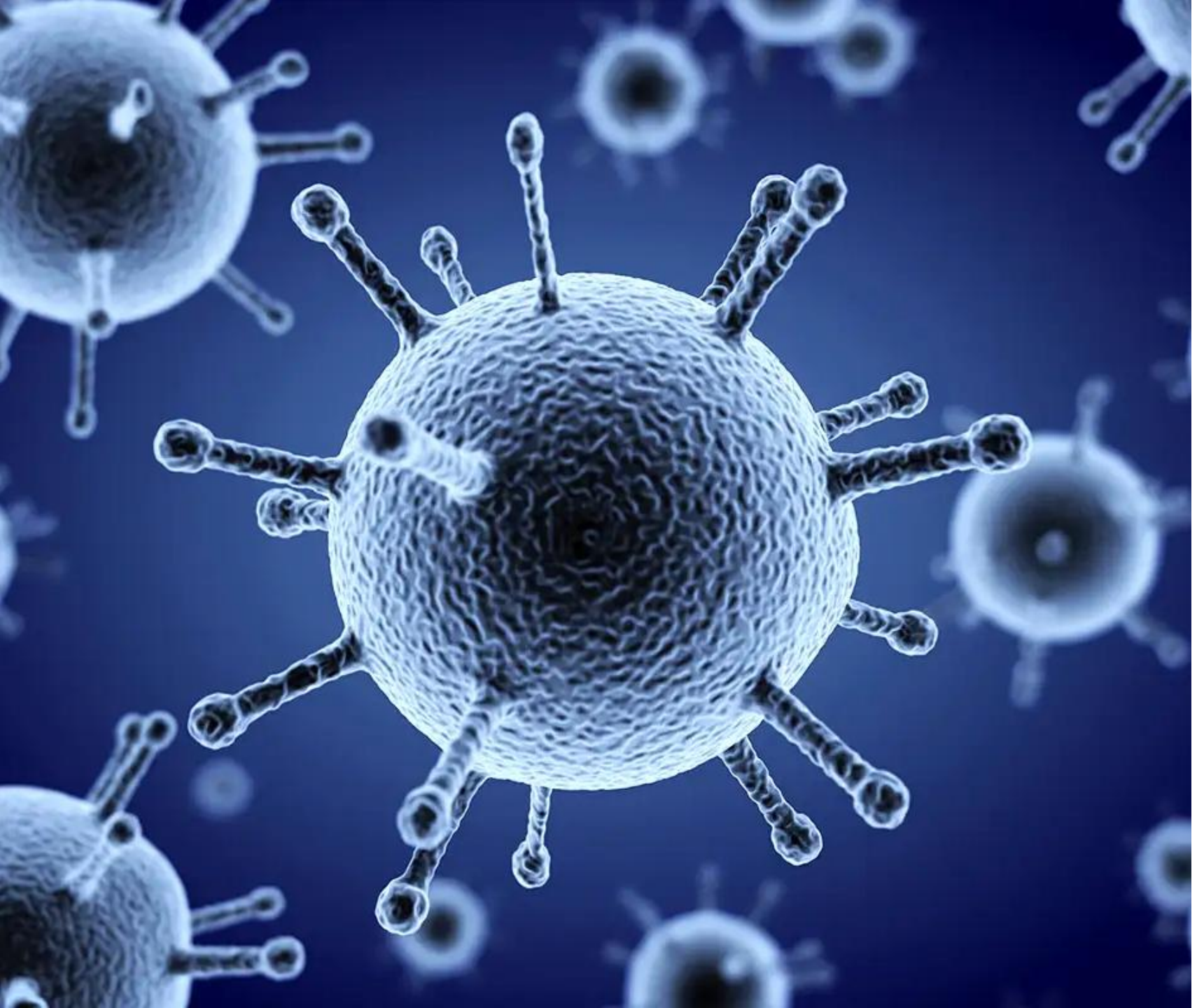
CONCLUSION

- Logistic Regression is effective for this task
- Behavior, trust in doctors, and risk perception drive vaccine decisions
- Predictive models can support public health targeting

- Train health workers to promote vaccine benefits
- Tailor messages for specific age groups
- Integrate models into health planning systems to forecast uptake

RECOMMENDATIONS

Made based on the predictions made by the model.



THANK YOU!

- Questions?
- Contact Info: doreen.murugi@student.moringaschool