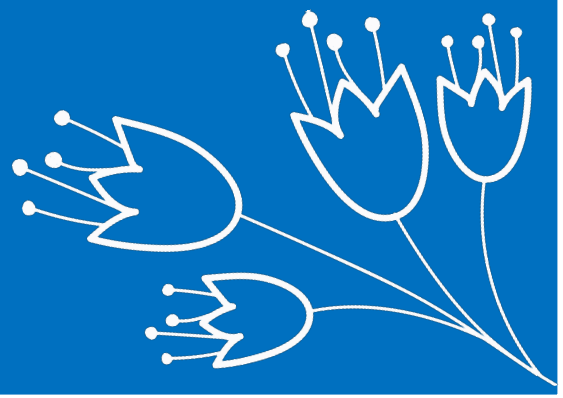
Predicting vaccination rate

Joshua Ko

Predictive model used

- ✔ Logistic Regression
- ✔ KNN
- ✔ Decision Tree



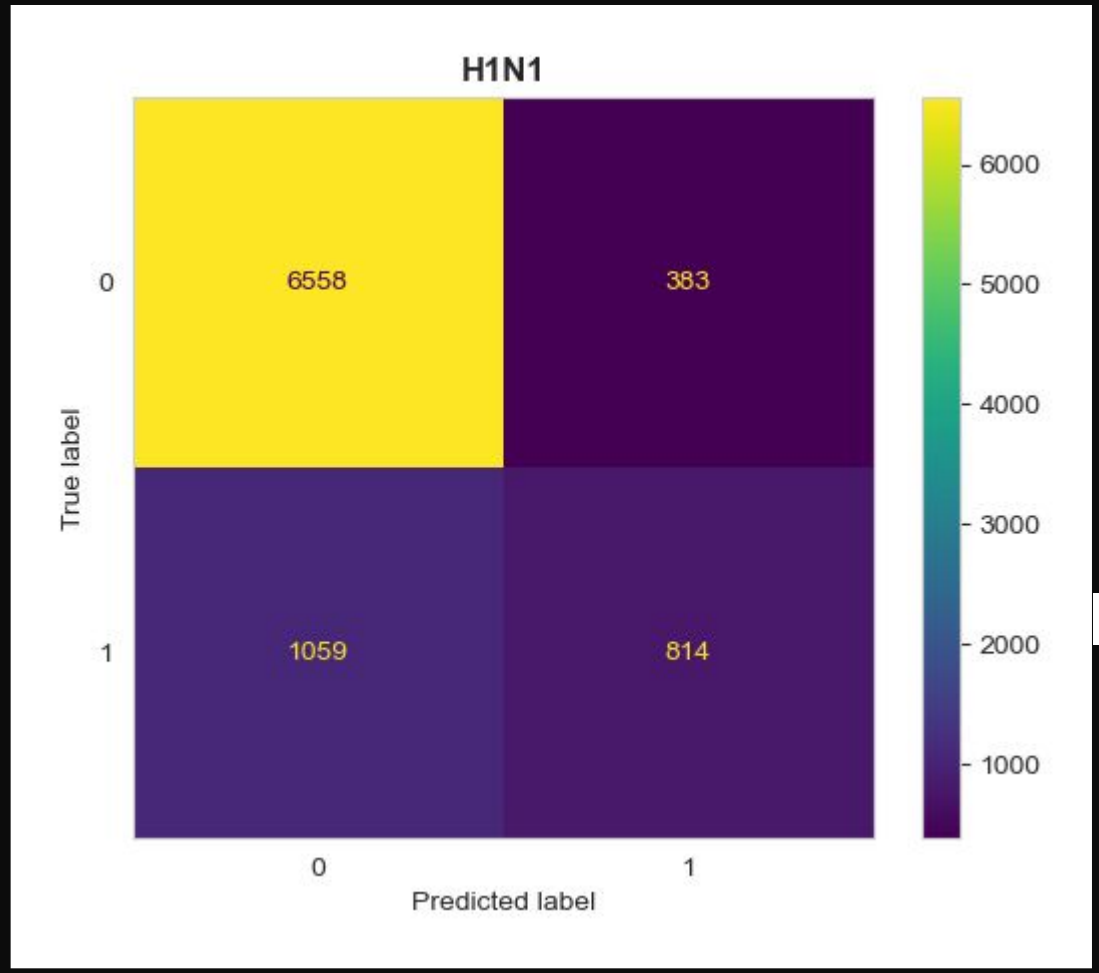
Accuracy of each model

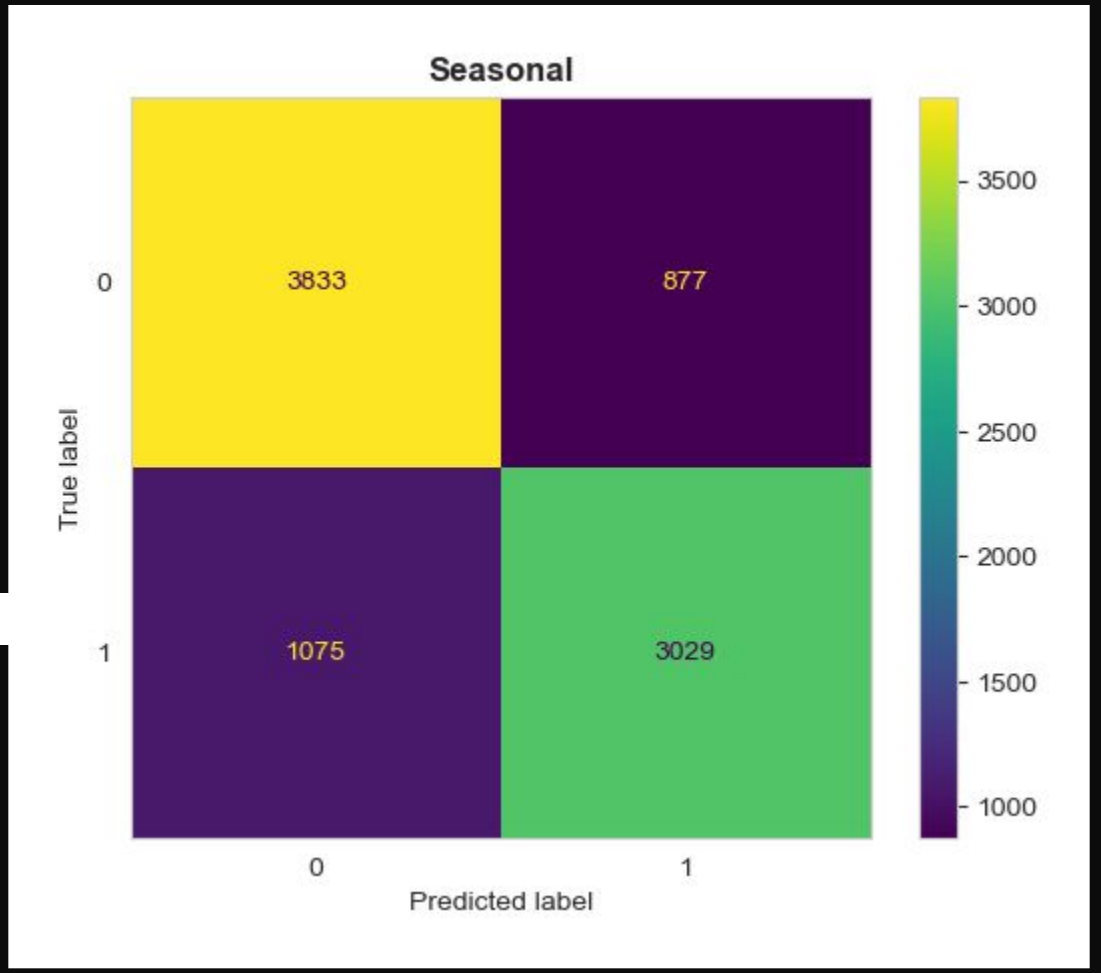
Model	*Avg. Accuracy*
Logistic	0.81
KNN	0.77
Decision Tree	0.78

Performance of the model

- Confusion Matrix
- Positive and negative predictive values
  - Precision
  - Recall
  - Accuracy
  - F1
- Receiver operator characteristic curve (ROC)

Confusion Matrix

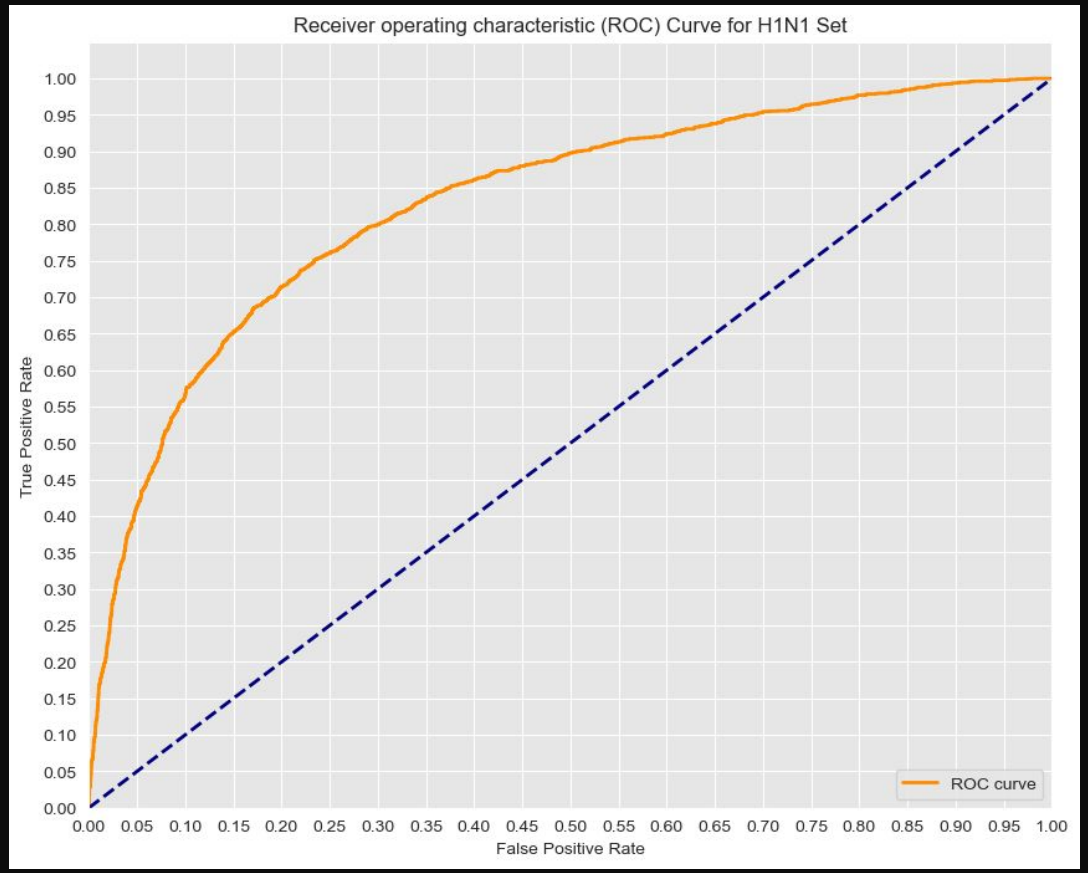


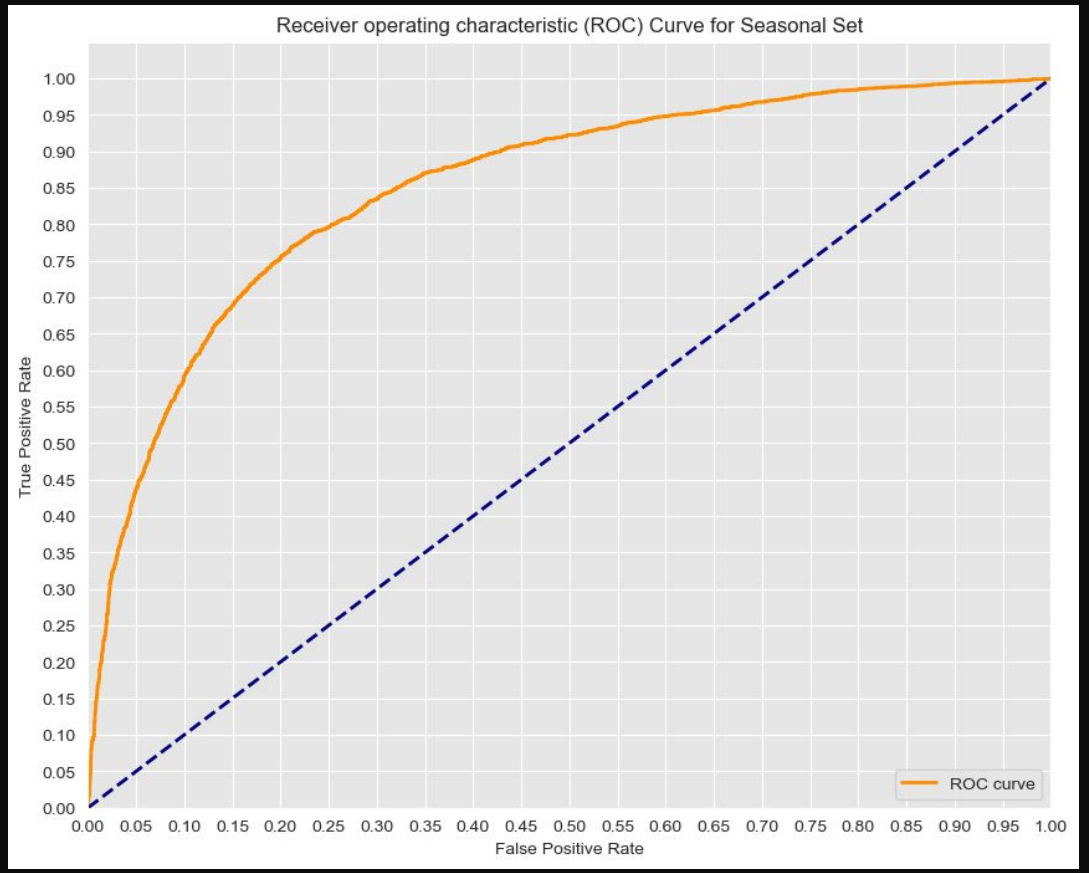


Prediction Scores

	H1N1	Seasonal
Precision	0.68	0.78
Recall	0.43	0.74
Accuracy	0.84	0.78
F1	0.53	0.76

ROC curve


Receiver operating characteristic (ROC) Curve for H1N1 Set
True Positive Rate
False Positive Rate
ROC curve


Receiver operating characteristic (ROC) Curve for Seasonal Set
True Positive Rate
False Positive Rate
ROC curve

Final predictions

respondent_id	h1n1_vaccine	seasonal_vaccine
26707	0.091963	0.277842
26708	0.036076	0.049654
26709	0.487698	0.582256
26710	0.479471	0.866802
26711	0.149953	0.504796

Improvements & Further developments

- **Making a more accurate model**
  - Working with less null values
  - Finding the best predictive model for each column
  - Trying different combinations of hyperparameters
- **Removing variables**
  - The dataset has multiple columns
  - Removing more irrelevant columns

Thank You!

This Presentation is Prepared by

Link:
https://github.com/Jko0425

Email:
joshuawko@gmail.com