

Flu Shot Learning



Wesley Dickens, Jason Mix, Kris Rubiano

Data

Data is from the United States CDC's National Center for Health Statistics. This information was collected for the National 2009 H1N1 Flu Survey via phone calls.

Aim

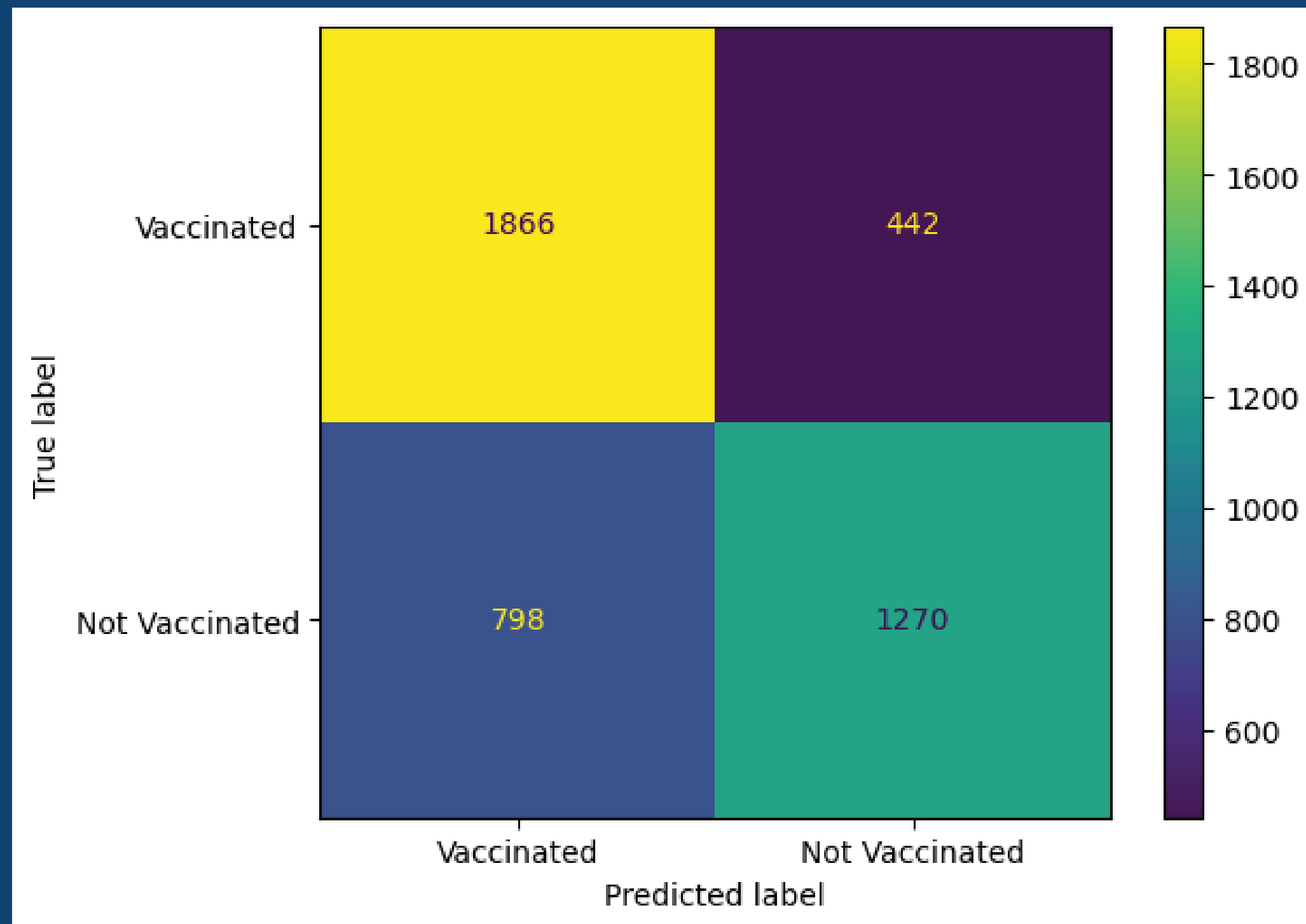
Our aim is to build an optimized model utilizing behavioral and demographic data to pinpoint individuals less likely to receive a flu shot, enhancing marketing efficiency.

Focused on optimizing flu shot campaign budget by targeting traditionally under-vaccinated population segments.

Models

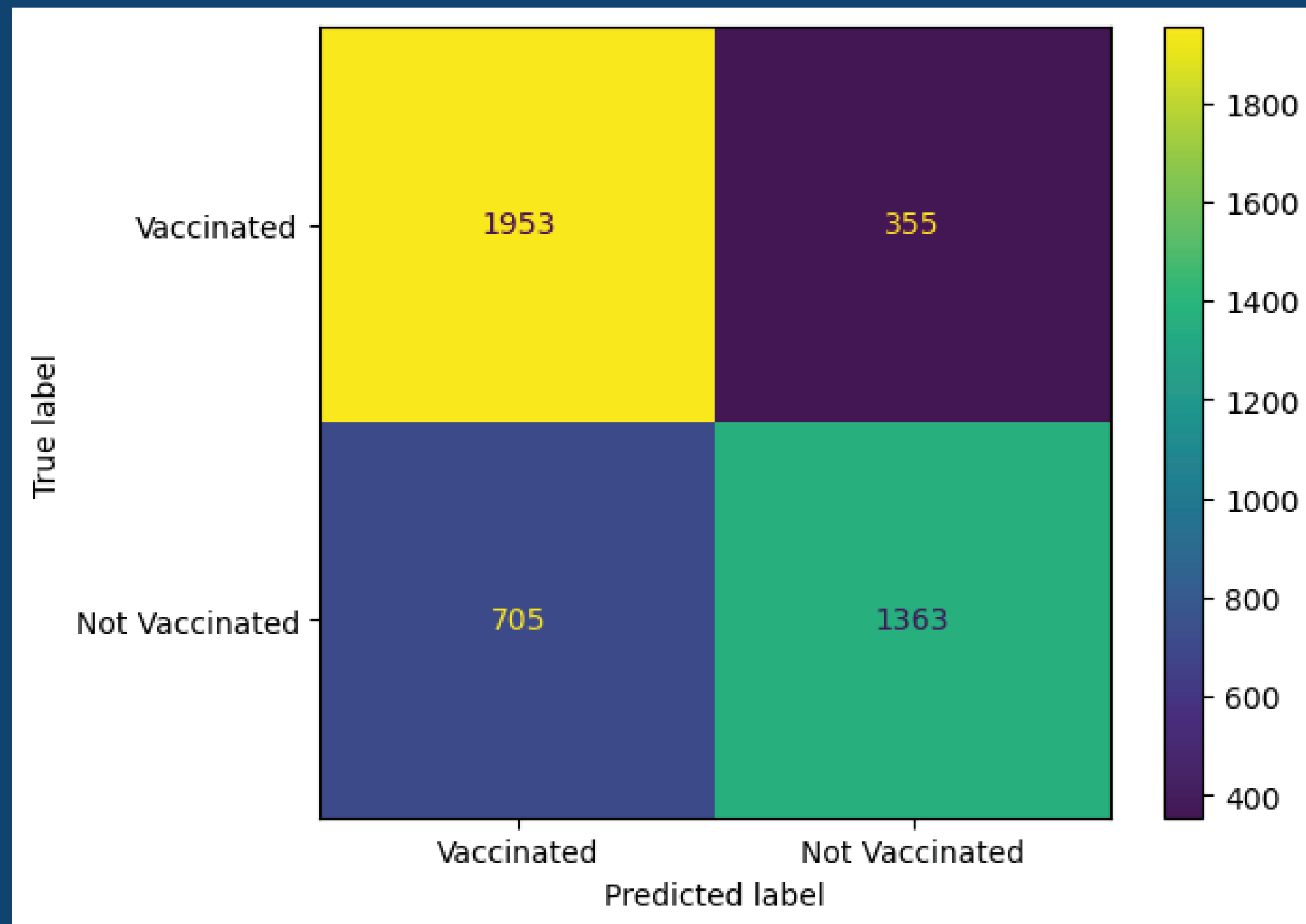
1. KNN
2. Decision Tree
3. Logistic Regression

KNN



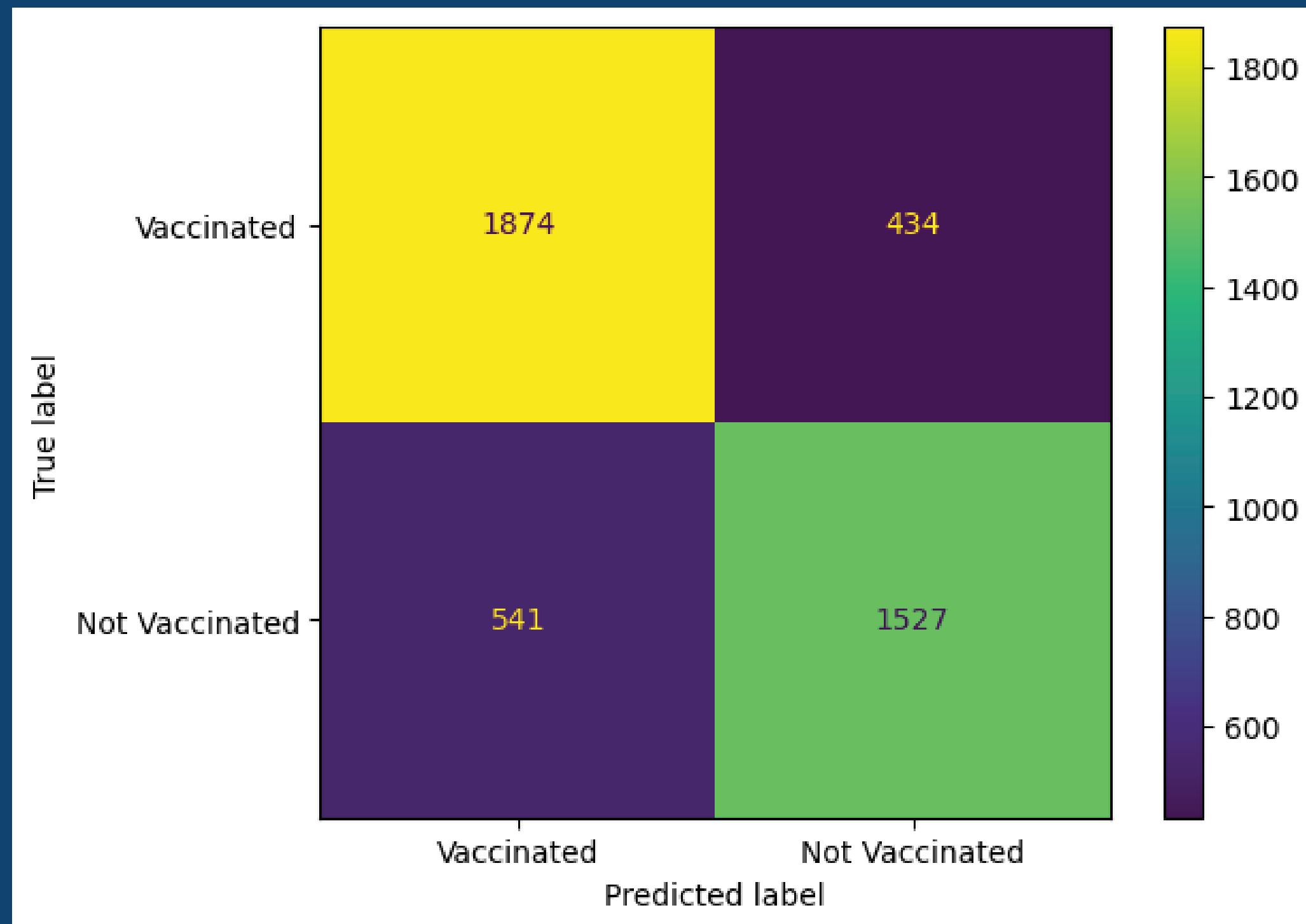
- Predicts with 53% accuracy, our lowest scoring model.
- This predictor likely exhibited lower predictive power compared to other models due to KNN's sensitivity to high-dimensional feature spaces and noise, which can impair its ability to generalize as effectively.

Decision Tree



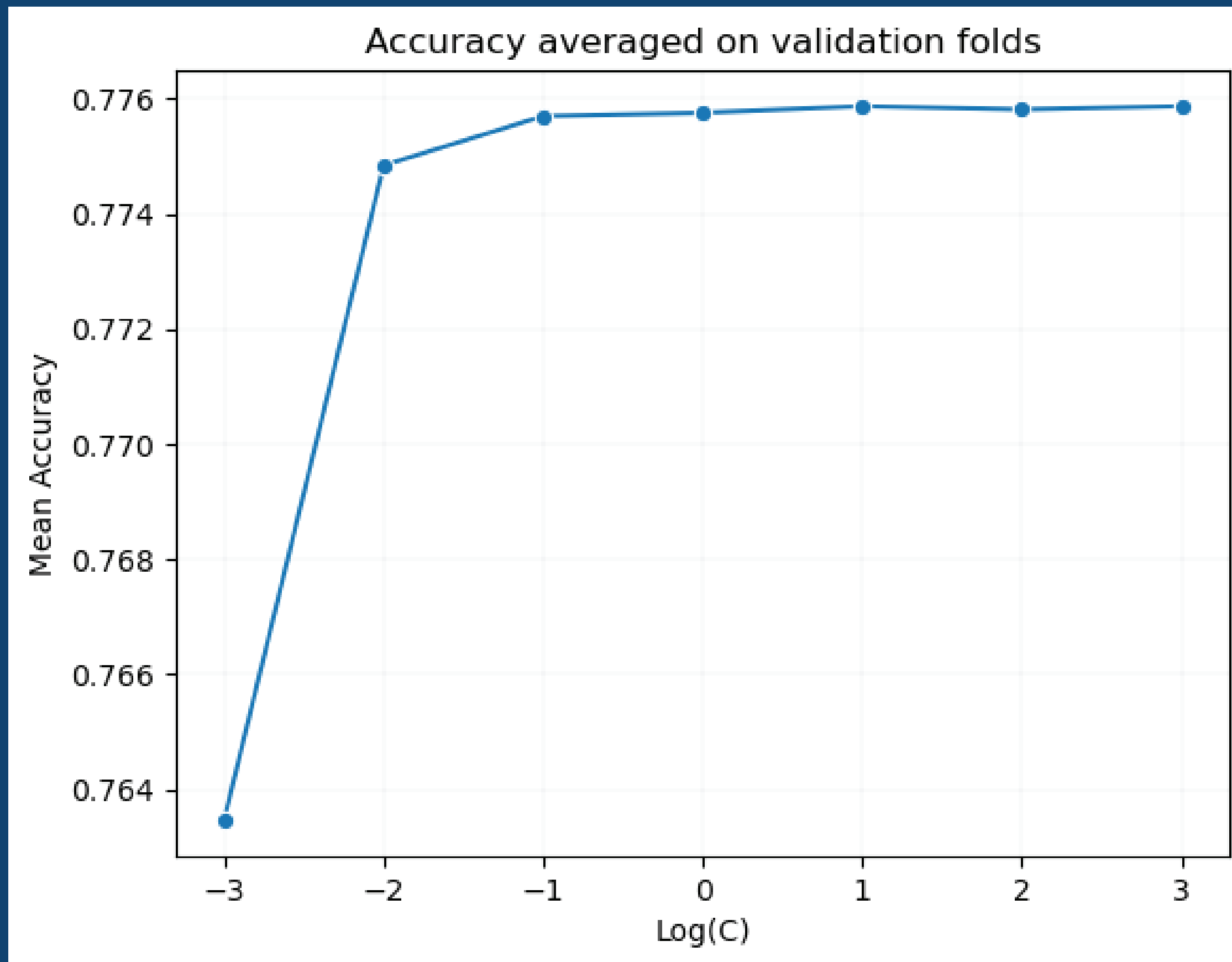
- Predicts with 76% accuracy.
- This model may have had lower predictive power compared to Logistic Regression due to its tendency to overfit complex datasets, leading to poor generalization.

Logistic Regression



- Predicts with 77% accuracy, with a False Positive rate at 12%.

Parameter Tuning



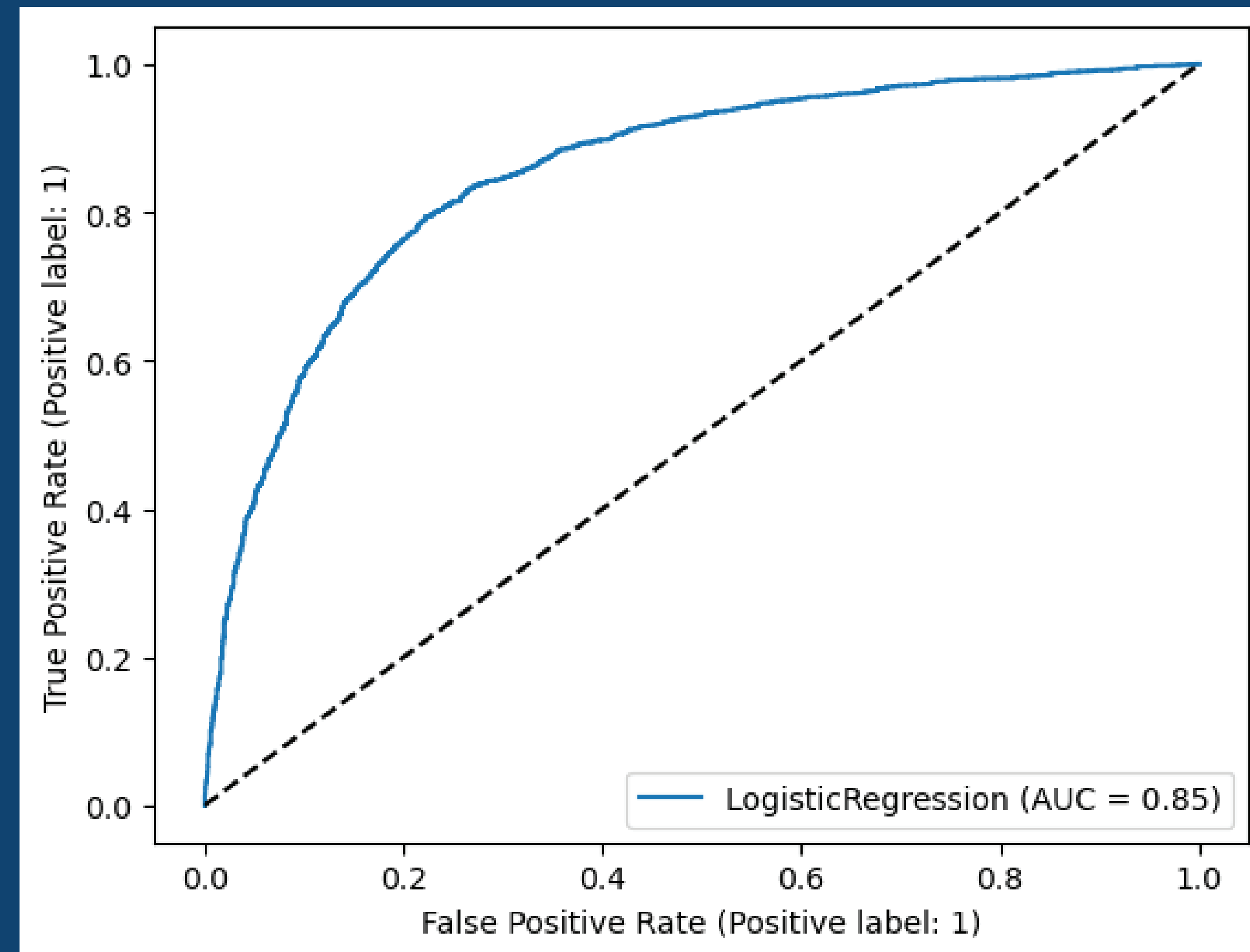
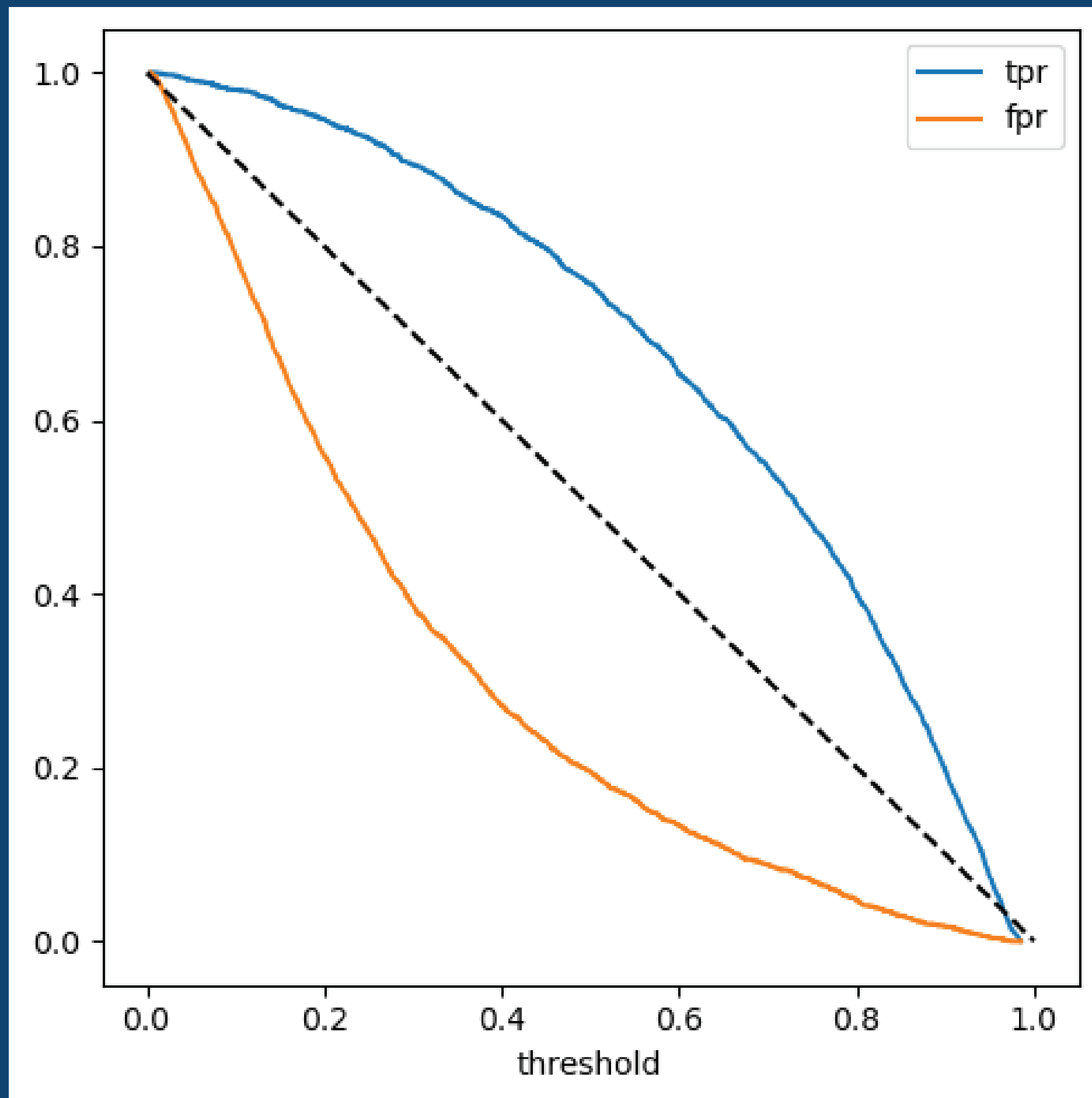
- To tune the hyper-parameters in our Logistic Regression model, we used a Grid Search approach and found that a C value of 1 with an L2 regularization penalty and a Newton-CG solver resulted in the highest accuracy.

Classification Report

	precision	recall	f1-score	support
0	0.78	0.81	0.79	2282
1	0.78	0.76	0.77	2094
accuracy			0.78	4376
macro avg	0.78	0.78	0.78	4376
weighted avg	0.78	0.78	0.78	4376

- Score post Logistic Regression increased by 1%. The tuned model is 78% accurate.

Threshold and ROC-AUC



Top 10 Features

Feature	Coefficient	Absolute Coefficient
Doctor Recommended Seasonal Vaccine	1.28	1.28
Age: 65+	0.83	0.83
Health Worker	0.81	0.81
Opinion of Seasonal Vaccine Effectiveness	0.58	0.58
Age: 18-34	-0.55	0.55
Opinion of Seasonal Flu Risk	0.55	0.55
Has Health Insurance	0.43	0.43
Age: 35-44	-0.27	0.27
Behavior: Touches Face	0.25	0.25
Race: Black	-0.22	0.22

Top 10 Features

Feature	Coefficient	Absolute Coefficient
Doctor Recommended Seasonal Vaccine	1.28	1.28
Age: 65+	0.83	0.83
Health Worker	0.81	0.81
Opinion of Seasonal Vaccine Effectiveness	0.58	0.58
Age: 18-34	-0.55	0.55
Opinion of Seasonal Flu Risk	0.55	0.55
Has Health Insurance	0.43	0.43
Age: 35-44	-0.27	0.27
Behavior: Touches Face	0.25	0.25
Race: Black	-0.22	0.22

Summary

- Our best model was Logistic Regression with tuned parameters at 78% prediction accuracy.
- Most effective predictor in determining if someone will get a seasonal flu vaccine is a doctor recommendation.
- Most effective inverse predictor in determining if someone will not get a flu vaccine is their age. People ages 18-34 tend to not be vaccinated when not considering other factors.