

The background of the slide is an abstract composition of numerous thin, wavy lines in shades of red, orange, green, and blue. These lines flow across the frame, creating a sense of movement and depth. A bright, circular light source is positioned in the upper center, casting a soft glow and creating a lens flare effect. The overall color palette is warm and vibrant, with the lines appearing to glow against a darker, hazy background.

Vaccine utilization project

Model and insights

A large orange shape on the left side of the slide, consisting of a rectangle with a quarter-circle cutout on its right side.


Business
understanding


- Leader in the field of health care
- Likelihood to be vaccinated
- Person's background, views and behaviors



A large orange shape on the left side of the slide, consisting of a rectangle with a quarter-circle cutout on its right side.

Data understanding

- From the National 2009 H1N1 Flu Survey
 - Target: vaccinated against the H1N1 virus
 - Features: one's background, views and behaviors.
 - 27,000 observations (rows of data)
 - 21% of respondents received the H1N1 vaccine.
 - significant correlations: doctor recommendation, opinion of risk, and opinion of effectiveness.
- 
- A yellow dashed line in the bottom right corner, consisting of four curved segments that curve upwards and to the right.

A large orange shape on the left side of the slide, consisting of a rectangle with a quarter-circle cutout on its right side.

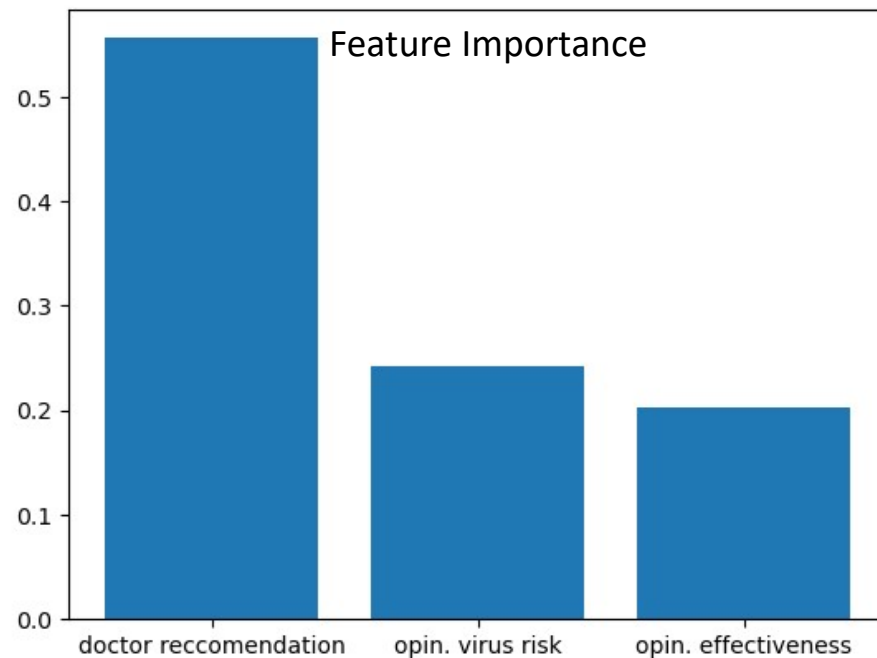
Modeling Overview: baseline, initial tree, second iteration

- Baseline model
 - logistic regression
- Decision tree model
- Final tree model:
 - reduce underfitting
 - more complex tree model
 - increase maximum depth



Final model

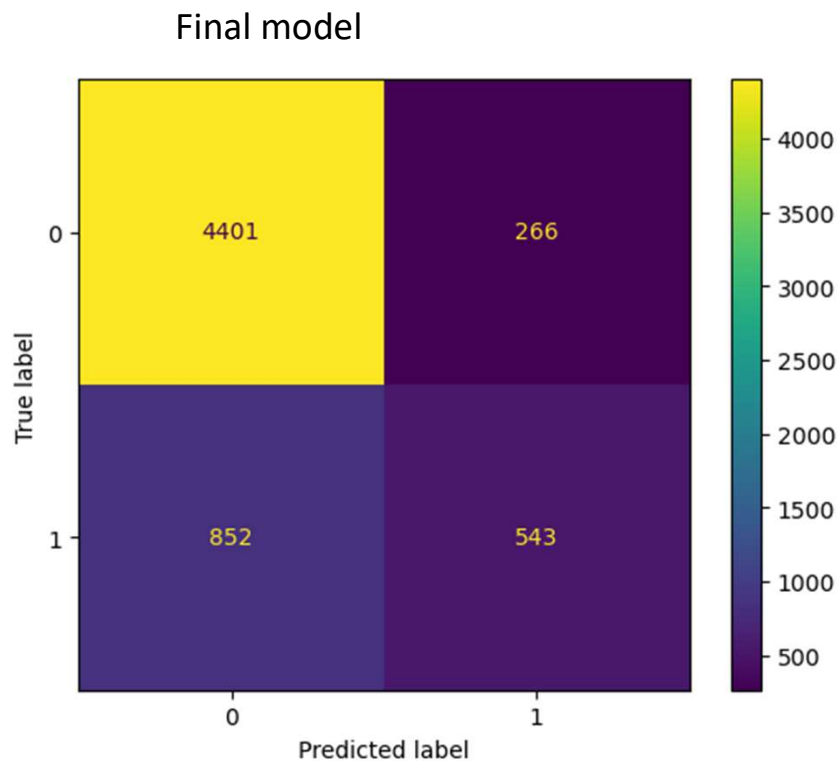
- Features: presence of a doctor recommendation, opinion of virus risk, and opinion of vaccine effectiveness
- Hyperparameters:
 - maximum depth of the tree is increased from 5 to 10



Results comparison (accuracies)

	Logistic Regression (baseline)	First tree model	Final tree model
Training accuracy	81.98%	82.00%	82.02%
Test accuracy	81.33%	81.05%	81.56%

Confusion matrix & False positive comparison



	Final model	first tree	Baseline
false positives	266	269	252
Total Observations	6,060	6,060	6,060
Actual negatives	4,740	4,740	4,740

Recommendations

- Model implementation for similar outbreaks
- Most important feature:
 - Doctor recommendation
- **Limitations**
 - slight improvement over iterations
 - accuracy (82%)

Thank you/Questions?