

The background features a series of overlapping, wavy lines in shades of red, orange, green, and blue, creating a sense of depth and movement. A bright, glowing light source is positioned in the upper center, casting a soft, pinkish-purple glow across the entire scene. The lines appear to be part of a larger, unseen structure, possibly a wireframe model of a landscape or a complex data visualization.

Analysis of housing prices




Business
understanding


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



A large orange circle is positioned on the left side of the slide, partially cut off by the edge.

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 series of four yellow curved dashes are located in the bottom right corner of the slide.

A large orange circle is positioned on the left side of the slide, partially cut off by the edge.

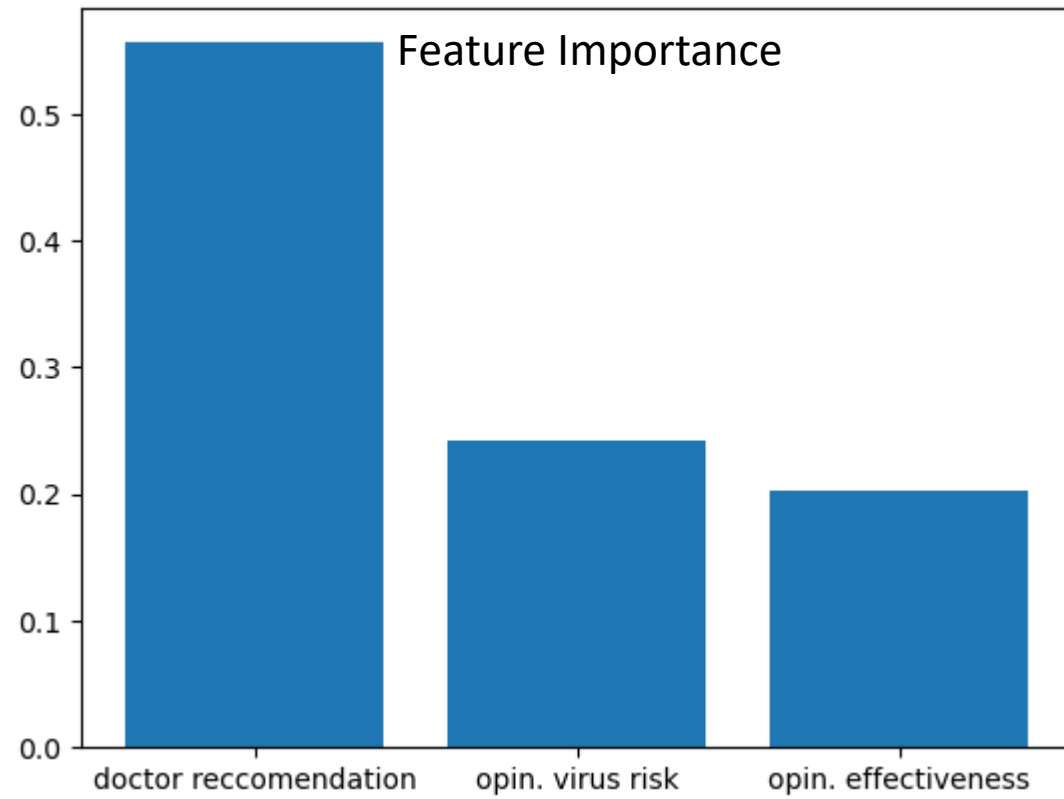
Modeling Overview: baseline, initial tree, second iteration

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



Final model

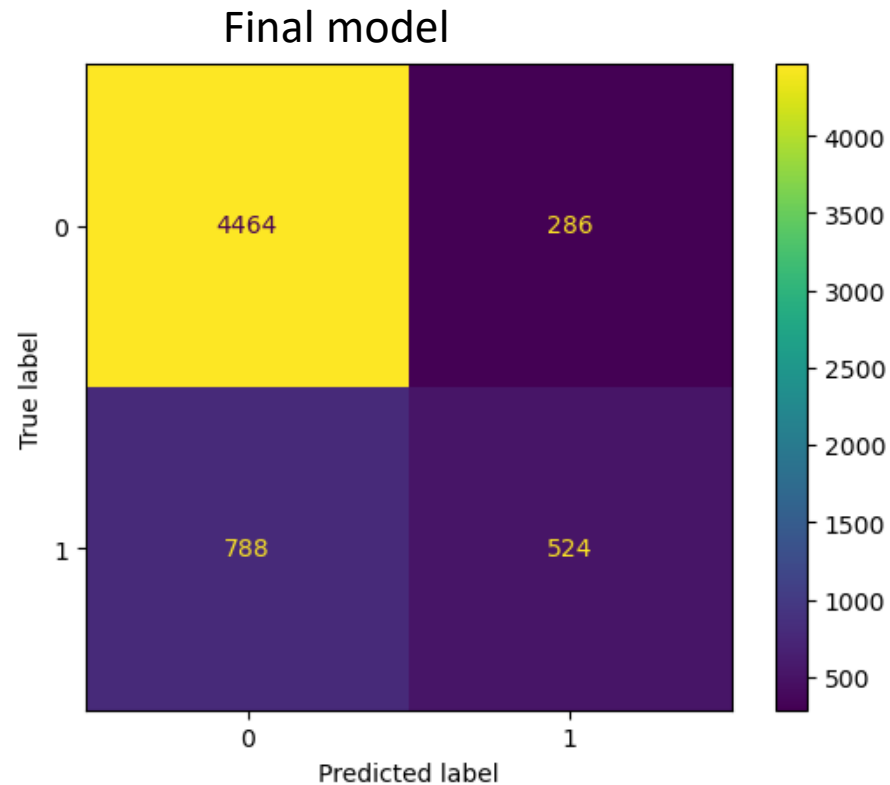
- Features: presence of a doctor recommendation, opinion of virus risk, and opinion of vaccine effectiveness
- Hyperparameters:
 - maximum depth of the tree is increased from 5 to 10



Final model & Results

	Logistic Regression (baseline)	First tree model	Final tree model
Training accuracy	81.67%	81.73%	81.79%
Test accuracy	82. 27%	82. 25%	82. 28%

Confusion matrix & False positive comparison



	Final model	first tree	Baseline
false positives	286	293	259
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?