Vaccine utilization project

Model and insights

Business understanding

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

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.

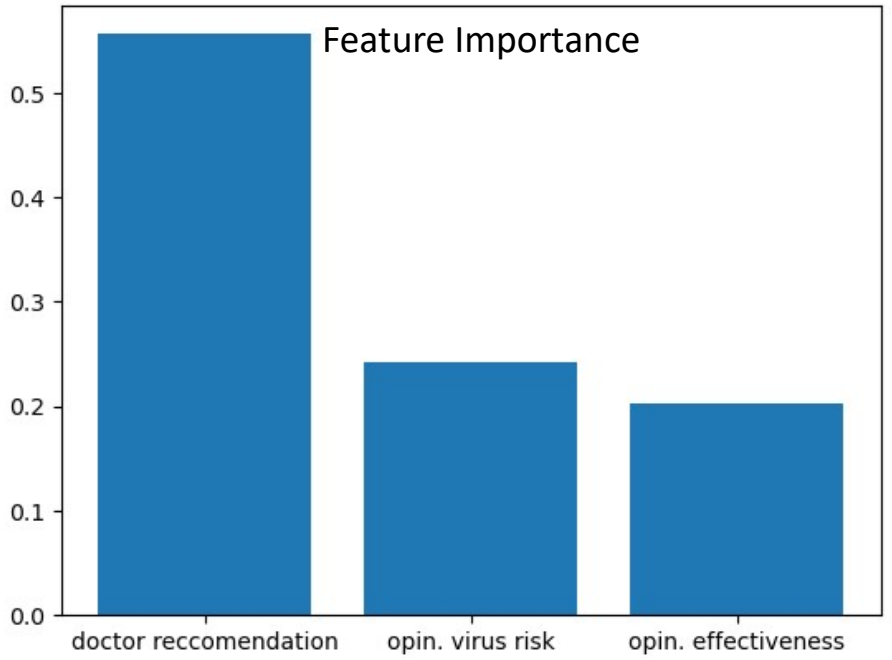
Modeling Overview: baseline, initial tree, second iteration

- 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

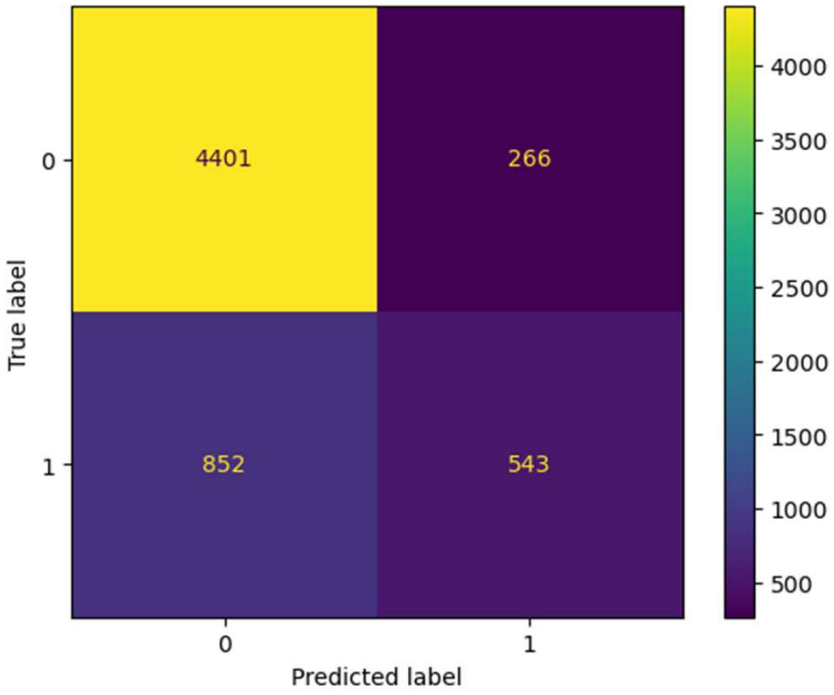


Results comparison (accuracies)

	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



	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?