H1N1 Vaccine Model Analysis

By: Jack Locke

Summary

Predicting whether someone received the H1N1 vaccine or not

Cost Trade-off:

- Missed Opportunity
- Sending survey/medical record request to the wrong person

Outline

- Business Problem
- Data Understanding
- Limitations
- Features
- Model/Evaluation
- Conclusion
- Next Steps

Business Problem

- Private Insurance Company; send follow-survey/request medical records
- Goal: predicting who received the H1N1 vaccine
    - Doctor recommendation, health worker, opinion on H1N1 risk
- Increased knowledge → better insurance policies → company growth

Data understanding

- National 2009 H1N1 Flu Survey Dataset
- Social, economic, and demographic background, opinions on risks of illness and vaccine effectiveness, and behaviors towards mitigating transmission
- 26,000 rows and 38 columns

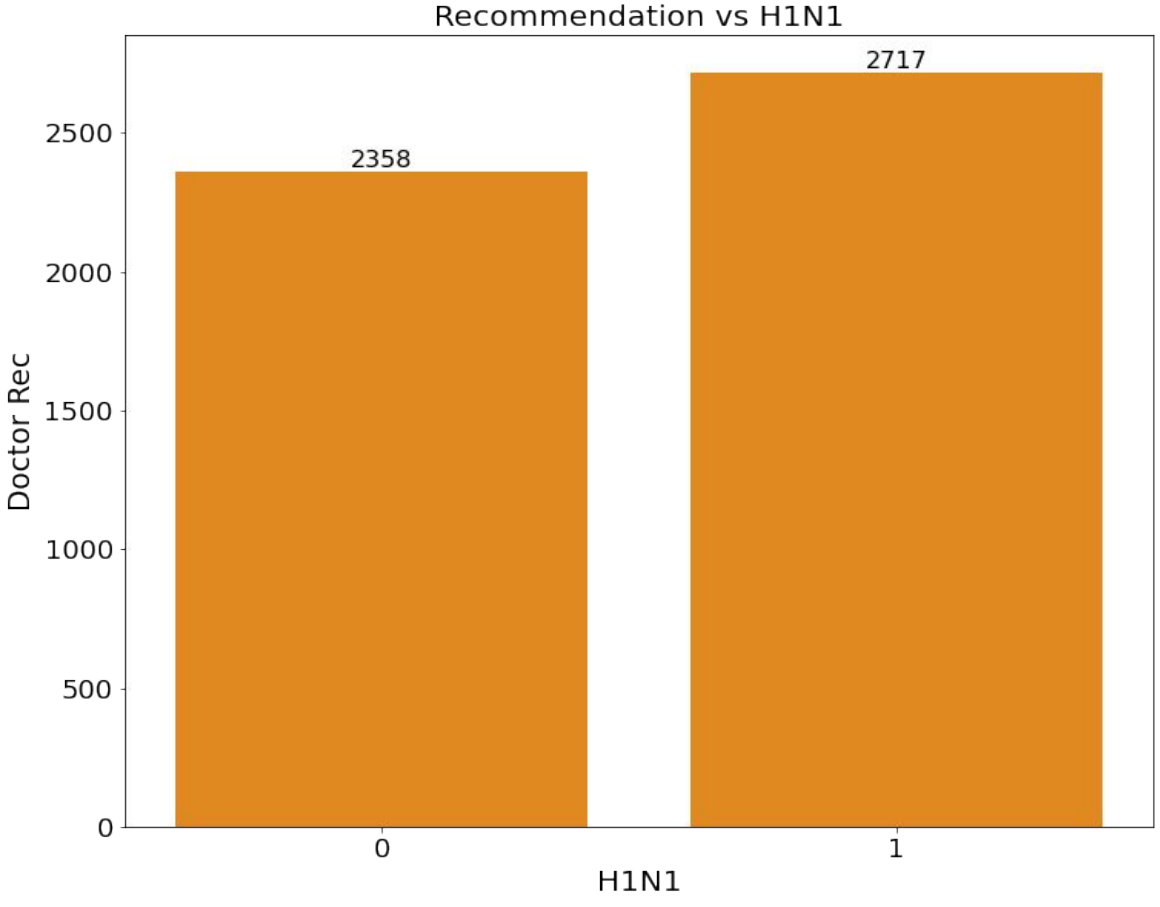
Limitations

- Account for all people that received H1N1 vaccines
- Account for all the factors that lead to receiving H1N1 vaccines
- Missing health insurance and employment data
- Survey Data:
  - Biased, Inaccurate, Unreliable

- Patterns/trends
- Follow-up Survey, Medical Records

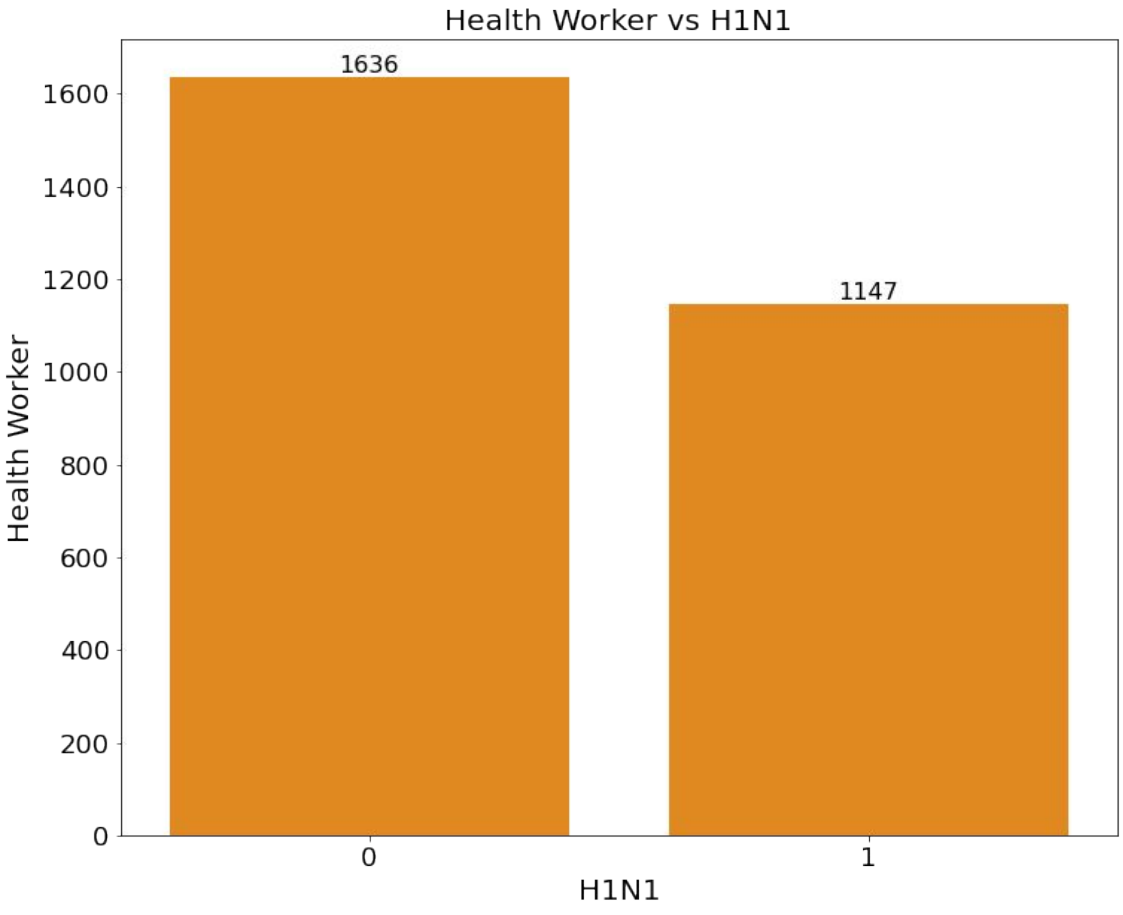
Feature 1:

- Doctor recommended vaccine
- Doctor rec vs. H1N1 vaccine



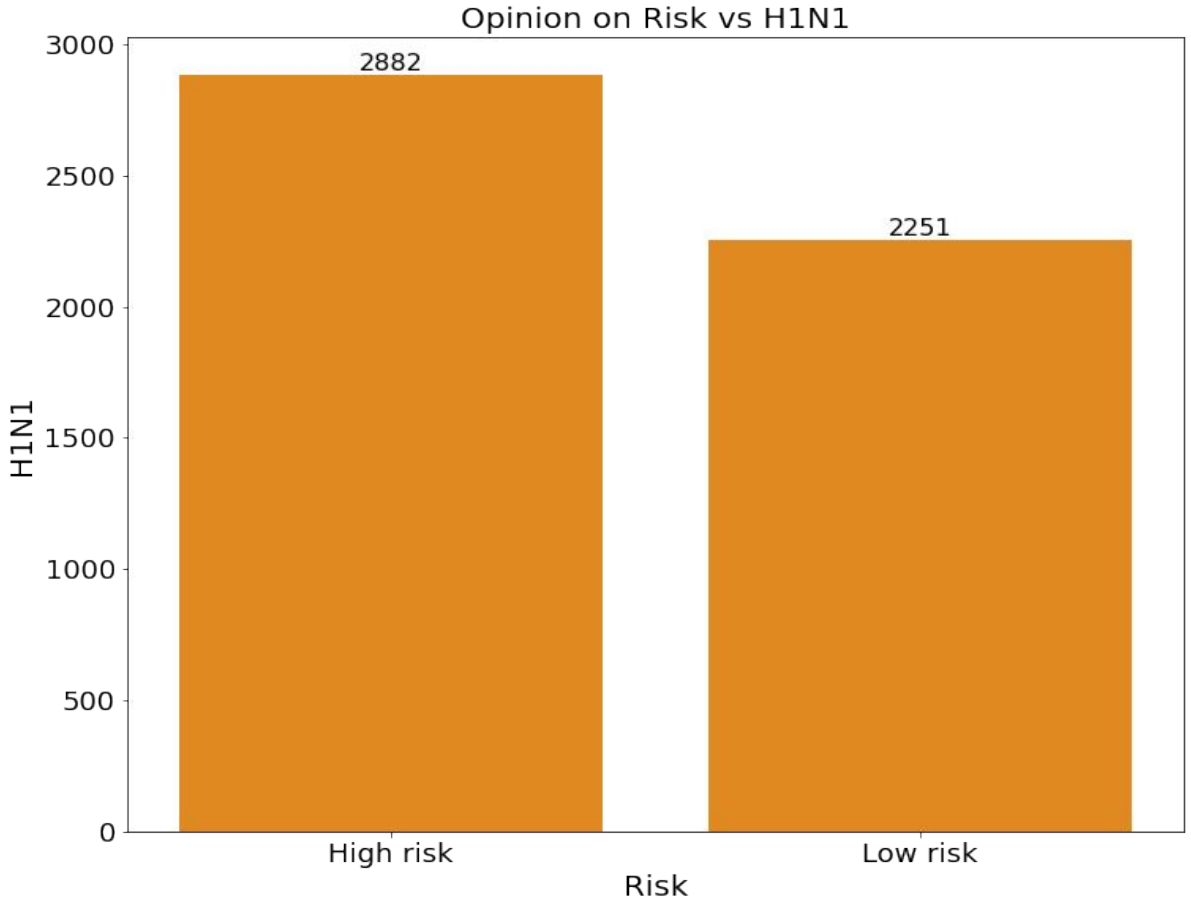
Feature 2:

- Is a healthcare worker
- Health worker vs H1N1 vaccine



Feature 3

- H1N1 vaccine received
- Opinion on risk vs H1N1 vaccine



Model/Evaluation

- Logistic Regression
- Cost Trade-off
  - Missed opportunity vs sending to wrong person
- Three Scenarios based on hypothetical: Surveying 10,000 people with a cost of $1 to create/ship
  - Model = ~72%, cost = $1,800
  - Model = ~81%, cost = $3,000
  - Model = ~91%, cost = $4,800

Conclusion

- Predict who received the H1N1 vaccine
  - Send follow-up survey, request medical records
- Increased knowledge → better insurance policies → company growth
- Three features:
  - Doctor recommendation, health worker, opinion on H1N1 risk
- Logistic Regression

Next Steps

- Cost trade-off
- Send follow-up survey/request medical records → reduce limitations
- Survey focus:
  - Three features, insurance/employment information
- More reliable/accurate model

Questions

Thank you

Email: jackdlocke@gmail.com

Github: https://github.com/johnlocke333/h1n1_flu_analysis

LinkedIn: www.linkedin.com/in/john-l-276395142