Predicting Cardiovascular Disease

- Classification Project Presentation-

Ron Lodetti Jr.
September 19, 2023

Overview

- Business Problem
- Data & Methods
- Modeling
- Evaluation
- Conclusions

Business Problem

Stakeholder:
Private Hospital

Context:
New online program focused on heart disease prevention

Problem:
How can we identify the individuals who would benefit the program?

Summary

Source
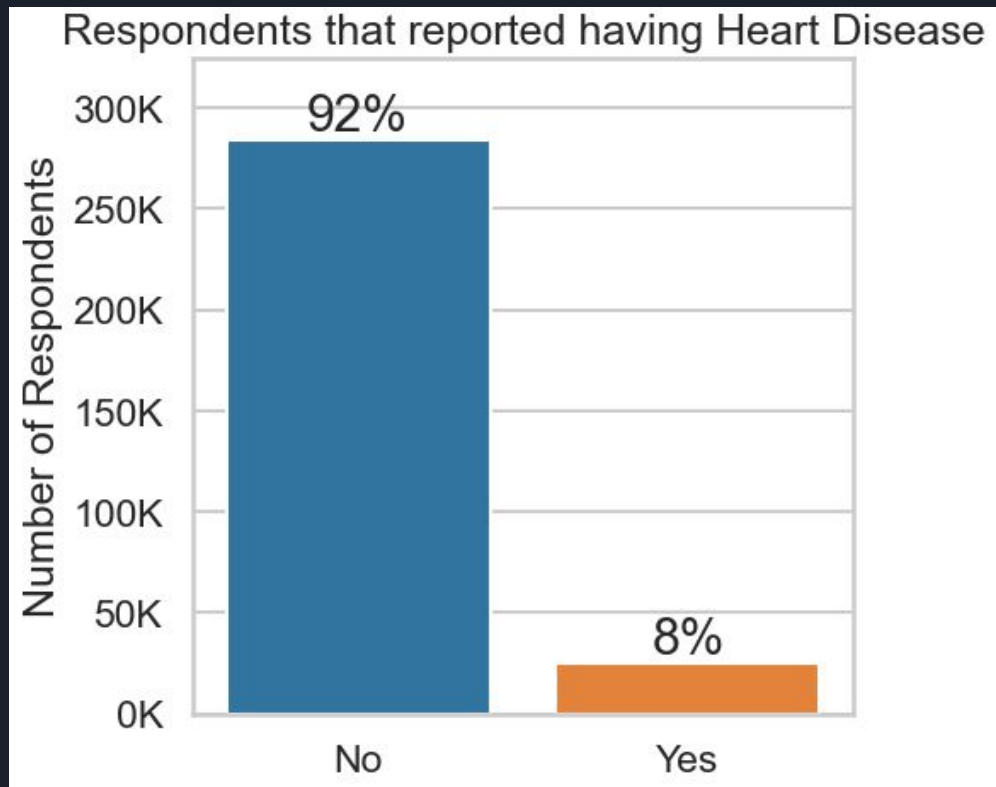
- CDC

Method

- ML with Classification Algorithms

Findings

- Correctly identify 81% of people with Heart Disease
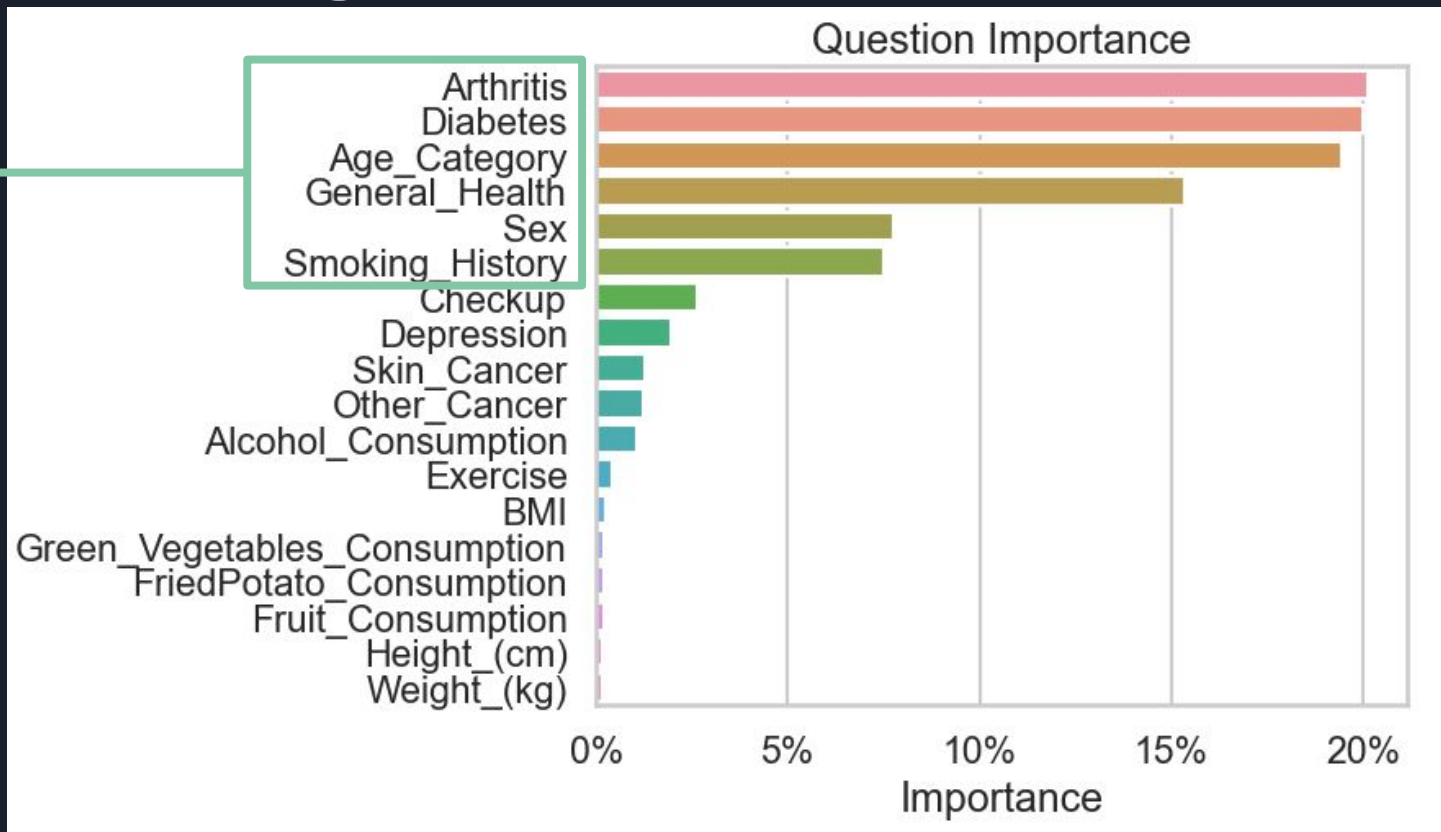
- 6-question survey

Data & Methods

CDC Survey

- +300K Respondents

- 18 Questions

- **Imbalanced target**

Respondents that reported having Heart Disease

Modeling

Important

Question Importance

	Importance
Arthritis	
Diabetes	
Age_Category	
General_Health	
Sex	
Smoking_History	
Checkup	
Depression	
Skin_Cancer	
Other_Cancer	
Alcohol_Consumption	
Exercise	
BMI	
Green_Vegetables_Consumption	
FriedPotato_Consumption	
Fruit_Consumption	
Height_(cm)	
Weight_(kg)	

Evaluation

# of questions	18	6	2
Correct	81%	81%	78%
Error	28%	29%	45%

Good model performance with only six questions!

Conclusion

Develop a 6-question survey:

- Arthritis
- Diabetes
- Age
- General Health
- Sex
- Smoking History

Correctly Identify **81%**

Limitations

- Not effective for high cost of mislabeling

Next Steps:

- **Gather more data** to include other factors that determine Heart Disease

- **Adjust threshold** based on stockholders

Thank you!

Email:
ron.lodetti@gmail.com

Github:
@rlodetti

Linkedin:
@ronlodetti