Heart Attack Analysis & Prevention Data

Mentors: Dr. Donald. S Williamson and Dr. Jeticia Sistrunk

&

Thank You National Science

Foundation

Who we are



Alexis Perkins

Health Sciences

Incoming Freshman Rincon, Georgia



Elikem Kalitsi
Computer Science and Electrical
Engineering

Incoming Freshman Washington DC and Accra, Ghana

Hypothesis & Research Question

- "Those with high levels of cholesterol have a higher chance of having a heart attack."
- Is there a correlation between cholesterol, chest pain, age?

Our Data

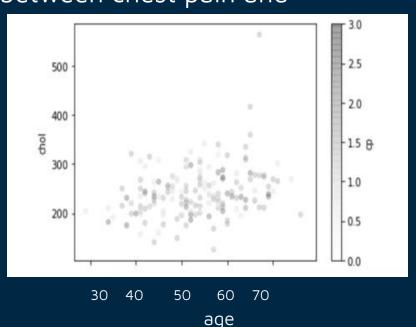
 Making predictions about chest pain (cp), cholesterol (chol), & age

Predicting if there's a correlation between chest pain and

cholesterol levels

High cholesterol: 240 mg/dl & up

- 303 patients
- Average cholesterol: 246 mg/dl



Why It is Important

By randomly choosing 3 variables, and focusing on 1 we can further analyze the variables that are important for concluding heart attacks, and remove the factors that prove little to nothing.

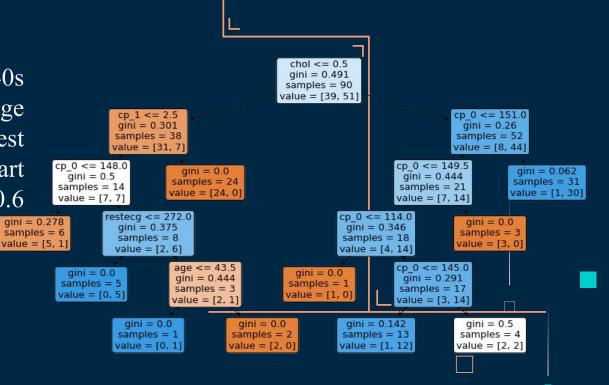
Description of Methods

- Python programming
- How we analyzed the data
 - Random Sampling
 - Decision Tree Classifier
 - Confusion Matrix
 - Data Cleansing
 - Scatter Plots

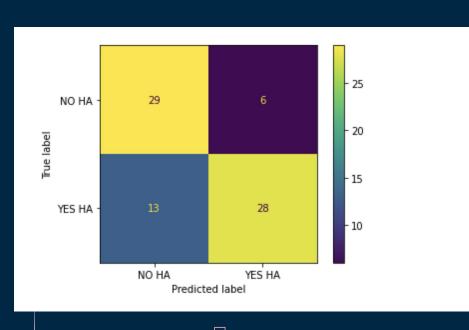
Analysis & Results

test subjects in their 30 - 40s with heart attacks have average 1.5 chest pain comparing to test subjects who did not have heart attacks with a chest pain of 0.6 and 0.4

Accuracy of 63%



Accuracy Dataset



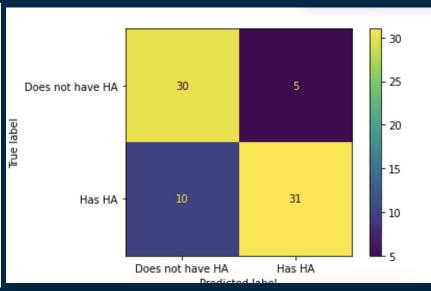


Chart 1: Test DataSet

Chart 2: Trained DataSet

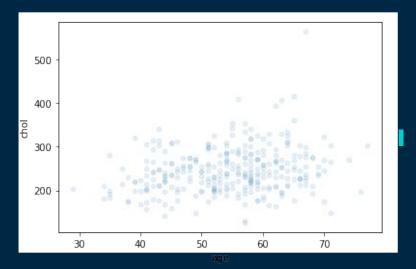
Predictions Results & Conclusion

- Cholesterol higher than 227 has a 46% chance of getting a heart attack
- Chest Pain higher than
 2.5 has a 55%
 chance of a heart attack.

```
chol <= 0.5
                 gini = 0.499
                samples = 90
               value = [43, 47]
               class = Yes HA
                              cp 0 <= 227.5
  gini = 0.43
                                qini = 0.363
samples = 48
                               samples = 42
value = [33, 15]
                              value = [10, 32]
class = No HA
                               class = Yes HA
                  gini = 0.0
                                                gini = 0.48
                samples = 17
                                              samples = 25
               value = [0, 17]
                                             value = [10, 15]
               class = Yes HA
                                              class = Yes HA
```

Conclusion

- Reject Hypothesis
- People aged 50-70
 have higher
 cholesterol levels.



Limitations

Apply to Dataset

Small Sample Size

Predictions Only

Uncertain of How Data was collected

Mentors

Thank you to...

Dr. Donald Williamson

Dr. Jeticia Sistrunk National Science Association

