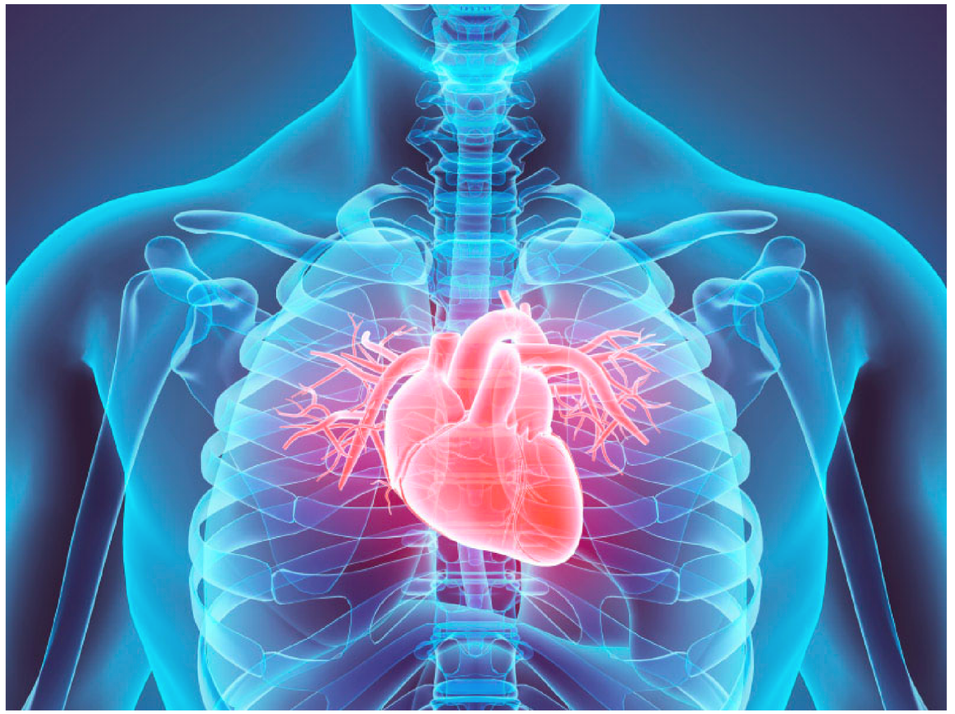
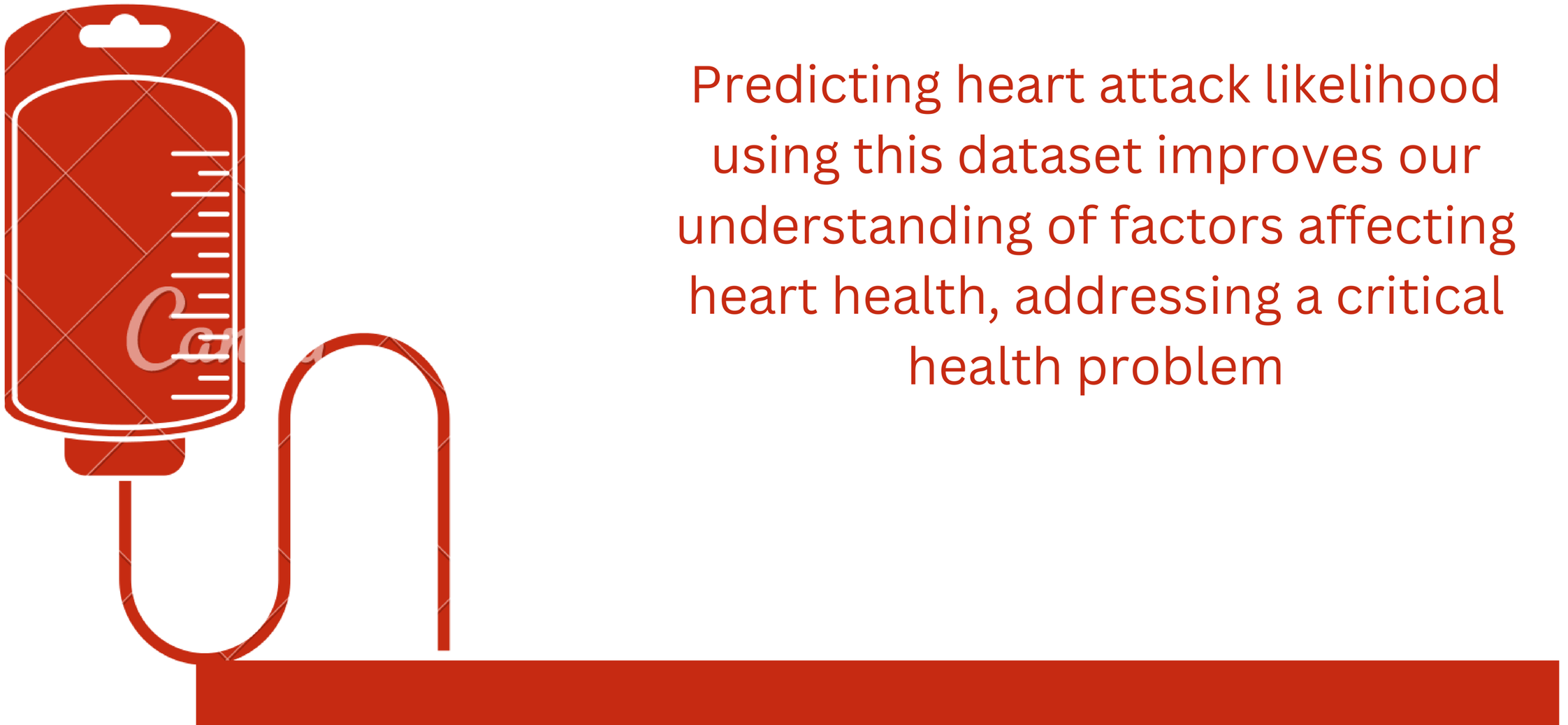
HEART ATTACK PREDICTIVE MODEL



Aiden Lee
Heather Robson
Angelica Hussar
Daniel Andreou
Enrique Bouche

PROBLEM WE ARE TRYING TO SOLVE



Predicting heart attack likelihood using this dataset improves our understanding of factors affecting heart health, addressing a critical health problem

THE DATA

Heart Attack Analysis & Prediction Dataset | Kaggle

Age of the patient (34 - 77)
Gender of patient (0 = F, 1 = M)
Chest Pain Type
Typical angina (0): Chest pain or discomfort that is provoked by discomfort/emotional stress relieved by rest. Squeezing like sensation in chest.
Atypical angina (1): Chest pain that doesn't meet the criteria for typical angina, abnormal chest pain.
Non-anginal (2): Chest pain in patients without heart disease.
Asymptomatic (3): Chest pain with no symptoms of heart attack.
Resting Blood Pressure
Cholesterol level
Fasting blood sugar: Measures the amount of glucose in the blood when one hasn't eaten anything.
Resting electrocardiographic results: Electrical activity of your heart while you are at rest. It provides information about your heart rate and rhythm, and can also show if there is enlargement of the heart, or evidence of a previous heart attack.
Max heart rate achieved: Greatest number of beats per minute your heart can possibly reach.
Exercise induced angina (1 = Y, 0 = N): Pain in the chest that comes on with exercise.
Oldpeak: Exercise induced segmented depression, used for reliable ECG finding.
Output (0 = less chance of HA, 1 = high chance of HA)

VISUALS

MACHINE LEARNING SOLUTION

Supervised Learning - Classification
Model - Logistic Regression

Our goal was to develop a predictive model that assesses an individual's risk for a heart attack based on health-related features.

MODEL DEVELOPMENT

- **Split data** - Output column was used to train the model by dropping column for test data
- **Feature Selection** of the top 5 features based on the chi-squared test & Regularization
- **Hyperparameter tuning:** Best hyperparameters: {'C': 0.1, 'penalty': 'l2'} Best cross-validation accuracy score: 0.8150724637681158

MODEL RESULTS

Pre-Hypertuning

```
# Create and save the testing classification report
target_names = ['Less Likely', 'More Likely']
testing_report = classification_report(y_test, testing_predictions, target_names=target_names)
print(testing_report)
```

```
              precision    recall  f1-score   support

 Less Likely       0.91      0.83      0.87        35
 More Likely       0.86      0.93      0.89        41

    accuracy                           0.88        76
   macro avg       0.88      0.88      0.88        76
weighted avg       0.88      0.88      0.88        76
```

MODEL RESULTS

Post Hypertuning

```
# Create and save the testing classification report
target_names = ['Less Likely', 'More Likely']
testing_report = classification_report(y_test, testing_predictions, target_names=target_names)

print(testing_report)
```

```
              precision    recall  f1-score   support

 Less Likely       0.91      0.86      0.88        35
 More Likely       0.88      0.93      0.90        41

    accuracy                           0.89        76
   macro avg       0.90      0.89      0.89        76
weighted avg       0.90      0.89      0.89        76
```

MAIN TAKEAWAYS

- Why should the business care about your project?
  - A predictive model for heart attack risk empowers patients to take preventive action by identifying high-risk individuals and enabling targeted interventions.

- If you were to continue working on this project in the future, what features would you add? What changes would you make?
  - Expanding the sample set and continuously training the predictive model with more diverse and representative data improves accuracy and performance. Updating and retraining the model with new data ensures that it remains relevant and useful over time.

MAIN TAKEAWAYS

- What are our main conclusions?
  - Without a model, there is no single data point that can be used to accurately predict a heart attack
  - Our model was ran and optimized and can predict with 90% accuracy the chance of a heart attack (f1-score)
  - A ten percent margin of error is considered wide for a serious matter like predicting a heart attack. While having a model can be a useful tool, it should not be the only test.

THANK YOU!