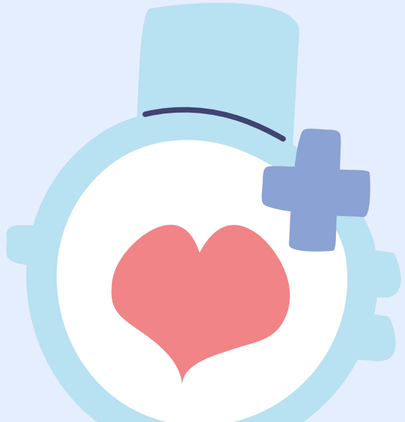
Predicting Heart Attack Risk

Don’t Guess the Signs.
Predict, Prevent, Protect.



By Nikhar Bhavsar

Heart Disease: A Leading Cause of Death in Canada

According to Public Health Canada

Heart disease is the second leading cause of death in Canada, accounting for over 50,000 deaths annually.

1 in 3 cases go undiagnosed until severe symptoms occur.

Why is Early Detection Challenging?

Doctor Shortages

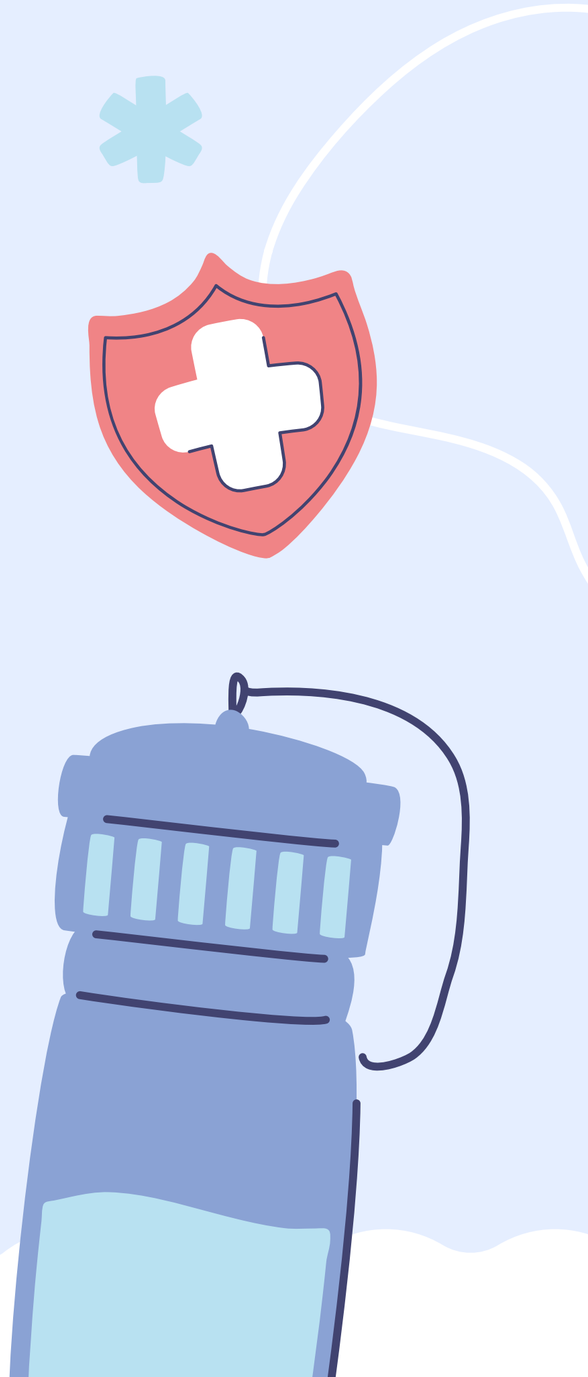
4.7 million Canadians lack a family physician (Statistics Canada 2023).

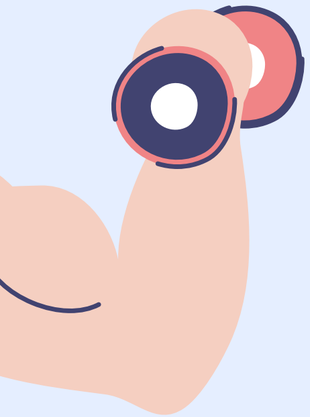
Limited Awareness

Early warning signs like high blood pressure or cholesterol are often ignored.

Testing Gaps

Rural and remote communities lack access to diagnostic facilities.



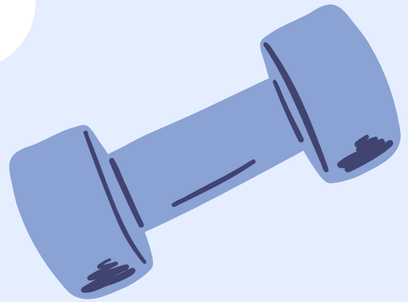


Proposed Solution

Early Risk Assessment with Machine Learning

- Analyzes health data to predict heart disease risk.
- Factors considered: Diabetic, Physical Active, BMI, etc.
- Alerts users to seek medical attention before severe symptoms arise.

Early detection can reduce healthcare costs and save lives.



Data Overview

Source

Kaggle

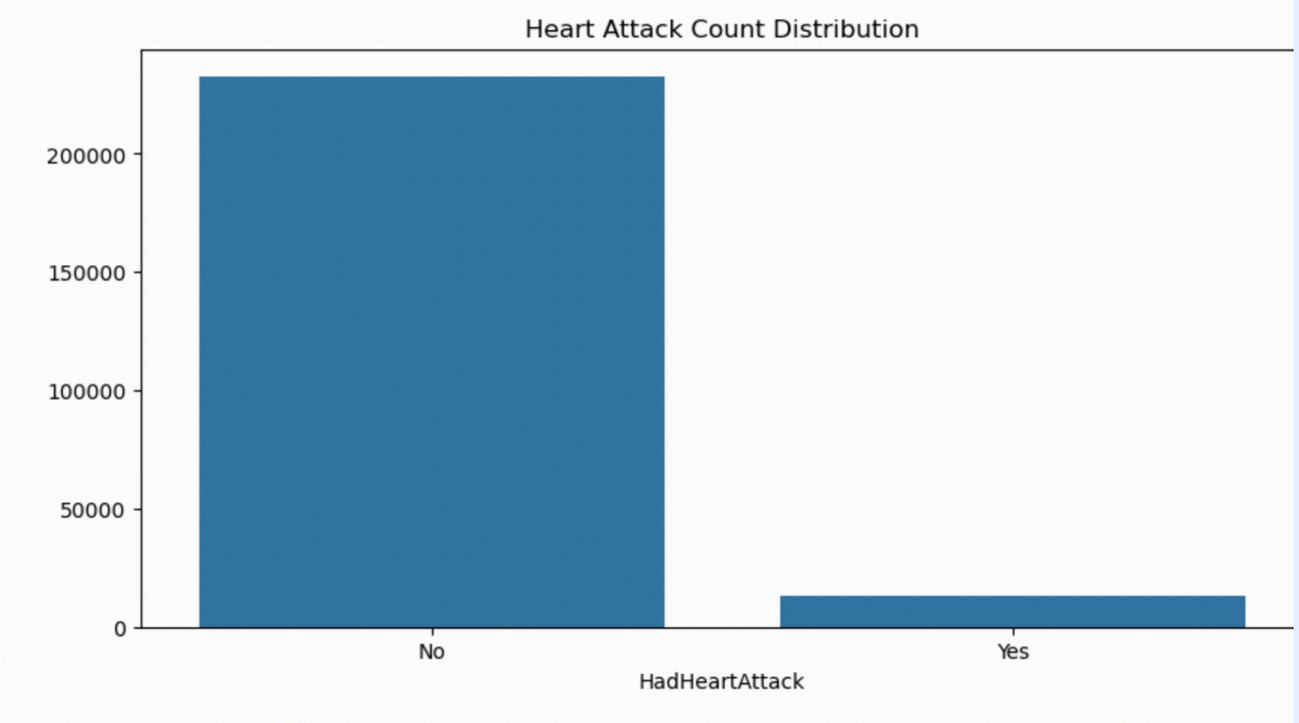
Shape

Rows - 246022
columns - 40

Factors

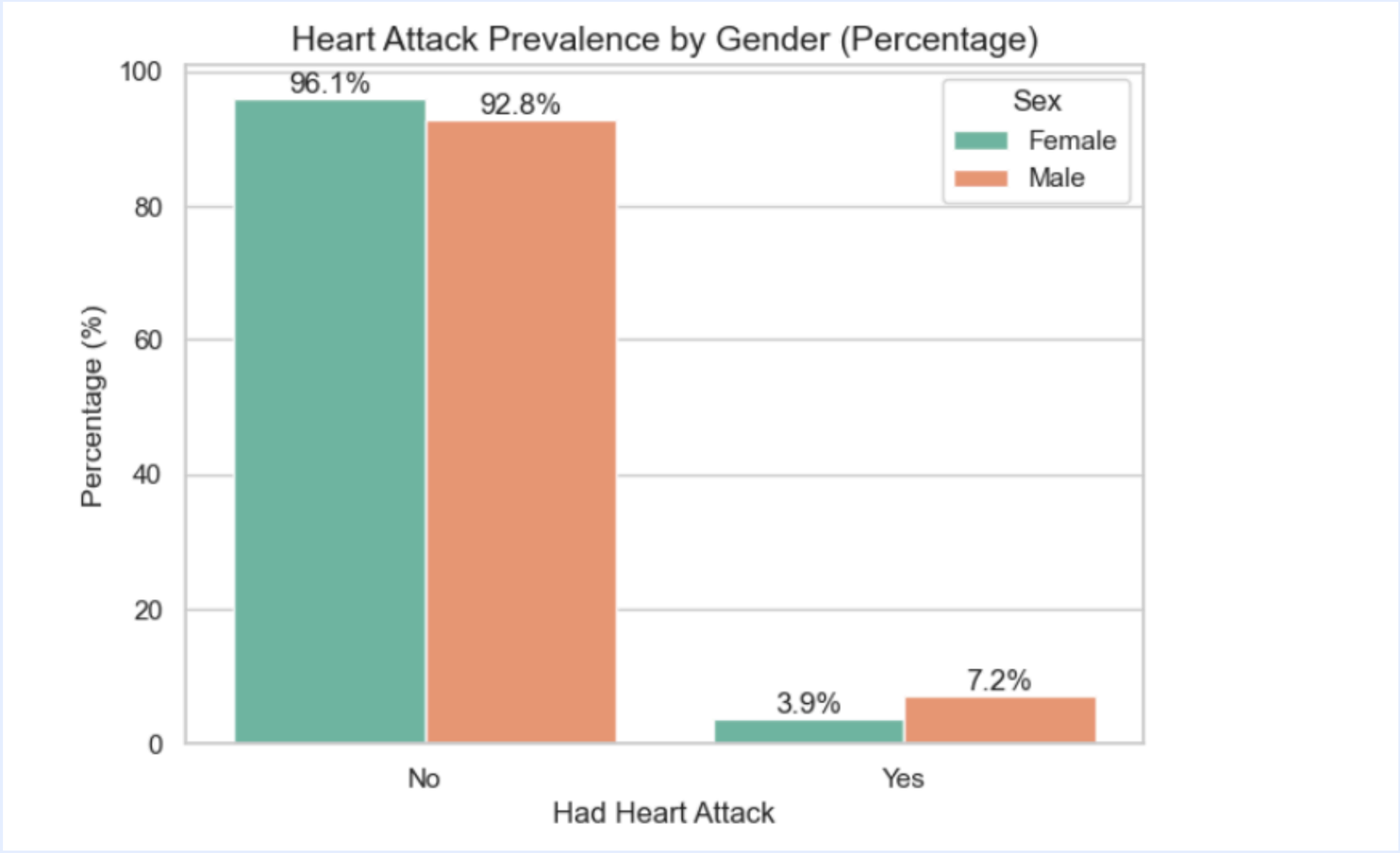
Age, Gender, Diabetic
Physically Active etc.

EDA Findings



No Heart Attack	232578	94.54%
Had Heart Attack	13435	5.46%

EDA Findings



PreProcessing

One-hot and Ordinal Encoding of the Categorical features

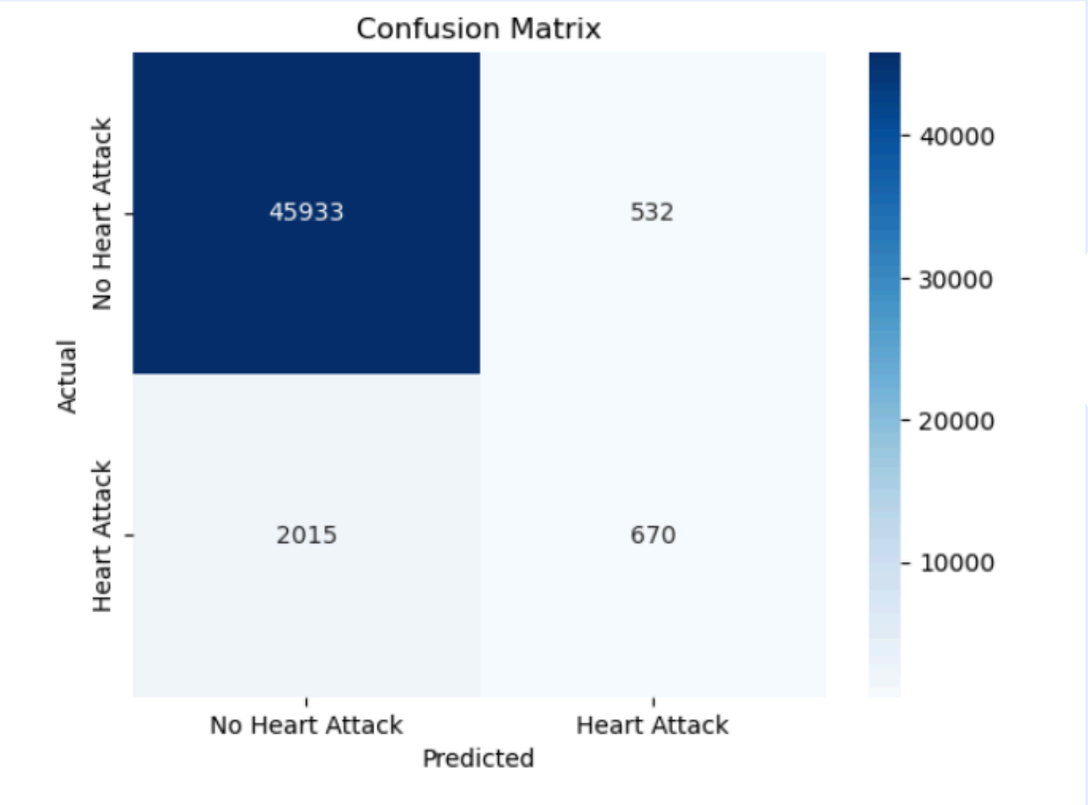
Scale the Numerical features

Add them into the ColumnTransformer

Logistic Model – Initial

Classification Report:

	precision	recall	f1-score	support
0	0.96	0.99	0.97	46465
1	0.56	0.25	0.34	2685
accuracy			0.95	49150
macro avg	0.76	0.62	0.66	49150
weighted avg	0.94	0.95	0.94	49150

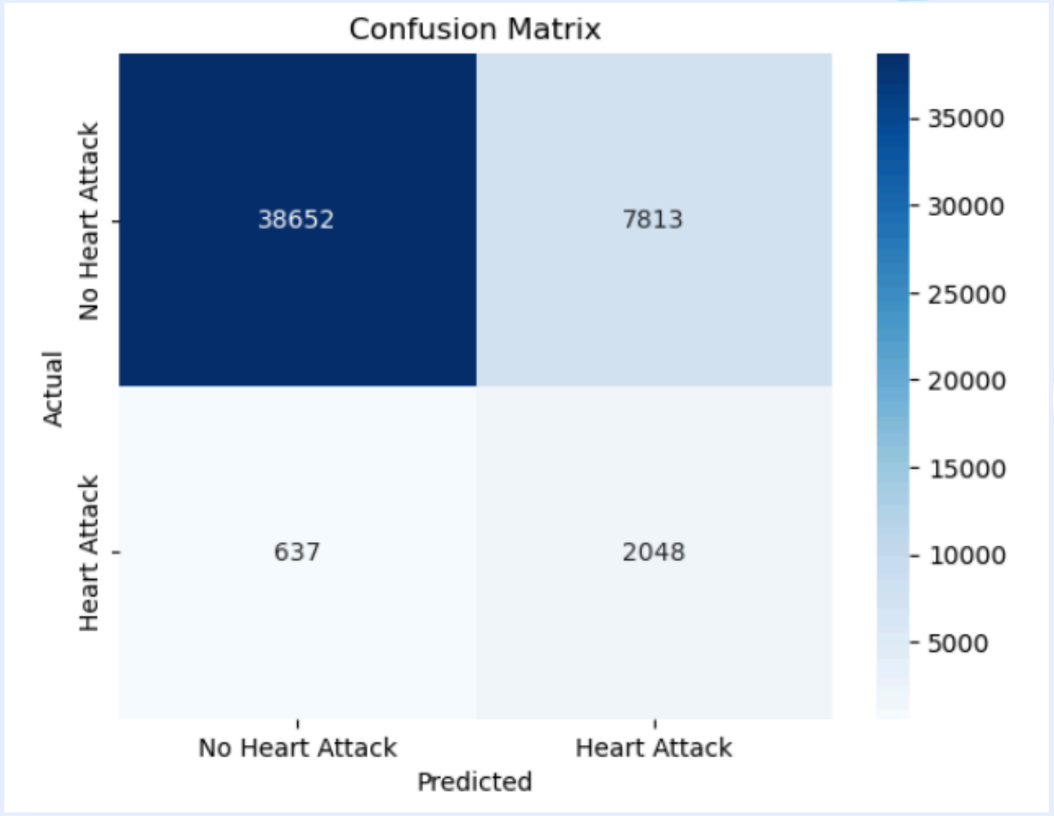


Without Handling the Class imbalance

Logistic Model – 2

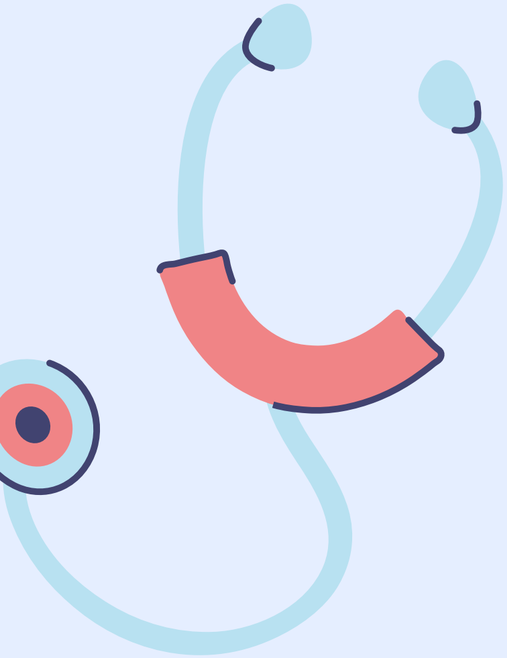
Classification Report:

	precision	recall	f1-score	support
0	0.98	0.83	0.90	46465
1	0.21	0.76	0.33	2685
accuracy			0.83	49150
macro avg	0.60	0.80	0.61	49150
weighted avg	0.94	0.83	0.87	49150



With Handling the Class imbalance

Next Steps:



1. Try Dimensionality Reduction - PCA
2. Try out other models like SVM - Random Forest - XGBoost
3. Compare All models and use GridSearch for hyperparameter tuning.

Thank You