PROJECT OVERVIEW

Business Problem

Through the use of machine learning models, can heart failure be accurately predicted using only demographic and baseline pre-stress test data without the need to conduct exercise stress tests?

The Heart Failure Prediction dataset consists of 12 features (variables):

	Demographic		Baseline / Pre-Stress Test					Stress Test				Heart Disease
1	Age	Sex	ChestPainType	RestingBP	Cholesterol	FastingBS	RestingECG	MaxHR	ExerciseAngina	Oldpeak	ST_Slope	HeartDisease
2	40	M	ATA	140	289	0	Normal	172	N	0.0	Up	0
3	49	F	NAP	160	180	0	Normal	156	N	1.0	Flat	1
4	37	M	ATA	130	283	0	ST	98	N	0.0	Up	0
5	48	F	ASY	138	214	0	Normal	108	Y	1.5	Flat	1
6	54	M	NAP	150	195	0	Normal	122	N	0.0	Up	0
7	39	M	NAP	120	339	0	Normal	170	N	0.0	Up	0
8	45	F	ATA	130	237	0	Normal	170	N	0.0	Up	0
9	54	M	ATA	110	208	0	Normal	142	N	0.0	Up	0
10	37	M	ASY	140	207	0	Normal	130	Y	1.5	Flat	1

The dataset has been categorized to appropriately address our business problem:

- Age and Sex are Demographic data
- ChestPainType, RestingBP, Cholesterol, FastingBS, and RestingECG are Baseline / Pre-Stress Test data
- MaxHR, ExerciseAngina, Oldpeak, and ST_Slope are Stress Test data
- HeartDisease is the clinical finding/result for the patient subjects

Why Address this Problem? Potential Value to Predict Heart Failure Without Stress Tests

Early screening using only basic patient data could help identify high-risk patients who should be prioritized for stress tests. This could make heart disease screening more accessible, especially in resource-limited settings

Value for the Patient

Stress tests can be physically demanding, especially for patients who are at greater medical risk to develop arrythmias or a symptomatic cardiac event during the test. Predicting heart failure using baseline data reduces the physical and emotional burden associated with the investigation of a heart condition, and also shorten the time to diagnose the problem; thereby, improving the efficiency of the process.

Earlier identification of at-risk individuals using baseline data enables proactive lifestyle changes or treatments, which potentially reduces or avoids the progression of heart disease through timely interventions.

Furthermore, for Canadians living in remote, rural, or other underserved areas where exercise stress testing might not be readily available, patients could receive initial diagnoses through their primary care providers or other similar services, thereby, potentially improving healthcare access by virtue of identifying cardiac disease through routine physical examination.

Value for the Healthcare Provider

For healthcare providers – physicians in particular – a reduction in the need for NITs improves their efficiency in the management of patients. Instead, physicians can use baseline data models as a fast and efficient preliminary screening tool, which saves times for them to focus on more acute cases. Diagnosing with straightforward physical exam data reduces the need for resource-intensive testing for lower-risk patients.

Developing of an ML model for heart failure prediction offers a layer of data-driven insights to support clinical decision making. Finding the best and most responsive models could further stratify patients based on risk features, which subsequently, improves quality of care provided by physicians.

Due to the accessibility of physical exam data of patients, the application of ML models to demographic, baseline blood work and ECG data could be integrated into electronic health records (EHRs), which enables the potential for notable scalability and risk predictions.

Value for the Healthcare System

In Ontario alone, approximately 500,000 NITs are performed annually with direct costs totaling C$300M per year. More than half of these tests are exercise stress tests. Despite this cost, exercise stress tests are an effective low-cost way to assess heart disease with an approximate cost ranging from C$100-300 per test. In comparison to the acute management of coronary artery disease (CAD) and other cardiac conditions, performing these noninvasive cardiac diagnostic tests (NITs) were associated with a 12% reduction in downstream costs when compared to cases in which no NITs were performed.

As such, uncovering opportunities to reduce the overuse of NITs for the diagnosis of CAD without eroding the quality of cardiac care services would offer notable financial and logistical benefits as Canadian provinces and territories address ongoing health system pressures.

Data Collection

Data Cleaning

EDA

Age and Sex are included in the baseline and stress test subsets

Models

Logistical Regression
KNN
Neural Network (simple NN)

Findings

- We found that demographic (Age, Sex) and baseline pre-stress test data (ChestPainType, RestingBP, Cholesterol, FastingBS, RestingECG) provides a reasonable foundation for predicting heart disease *……….insert blurb with numbers*
- The output of our models show that logistical regression provided the best accuracy in comparison to KNN and NN *…….insert blurb with numbers*
- However, we found that running stress tests improves the accuracy to predict heart failure in comparison to using demographic and baseline data alone *………insert blurb w/ numbers*
- Features such as ChestPainType, ST_Slope, and ExerciseAngina were highly predictive; therefore, reaffirming their clinical relevance.

REFLECTIONS

This section outlines our team's collective reflections regarding our experience from this project. These insights are also detailed in our video presentation.

1. What did you learn?

Stress Testing is Beneficial and Recommended in the Investigation of Cardiac Issues

Although we found the sole use of demographic and baseline pre-stress test data provides a reasonable foundation for predicting heart disease, identifying heart disease is best done by incorporating exercise stress testing

Model Behaviour

Logistic Regression excels in interpretability and performed well on this dataset and shows robust predictive ability*……insert blurb, recap the numbers*

KNN depends heavily on hyperparameter tuning (k) and is sensitive to imbalanced data, requiring additional preprocessing.

Neural Networks can model complex relationships but require careful tuning and computational resources for optimal performance.

Metrics Interpretation

While we certainly want to ensure accuracy by identify True Positives, we recognized the risk of False Negatives would be most significant because this subset would be misdiagnosed, and therefore, would not receive treatment (i.e. patients who are informed they do not have heart disease when they actually have heart disease). Consequently, minimizing false negatives was top-of-mind

What Else Should We Include Here???

2. What challenges did you face?

Include some insightful conclusions from the EDA HERE…………

Class Imbalance

The dataset has slightly more samples in one target class (more Males than Females), which can bias model training

Feature Encoding

Transforming categorical data, such as ChestPainType such that the data can run through the ML models while still preserve interpretability required additional steps, namely one-hot encoding

What challenges occurred while working through the ML models??? Insert something HERE………

3. How did you overcome those challenges?

Explainability

Apply SHAP values to enhance the model interpretability of the NN model

Feature Engineering

We employed one-hot encoding for categorical variables and normalized numerical features to ensure compatibility across the 3 models we deployed

What OTHER SPECIAL STUFF DID WE DO FOR THE 3 MODELS? INSERT SOMETHING HERE….

4. If you had more time, what would you add?

Incorporate Additional Features

Perhaps the incorporation of additional demographic and/or baseline pre-stress test data in the ML models could improve predictability of heart disease. Variables such as family history, BMI, geography, socioeconomic status, etc. has significant influence on cardiac health.

Explore More ML Models

Given more time, using ensemble methods like Random Forest or Gradient Boosting, XGBoost for comparison might yield better results than other models as these models often outperform simpler models in structured datasets

Generalize through the Use of External Datasets

Testing the best models on an external datasets would validate the ML models we deployed and establish greater credibility for physicians to experiment in the real world

More Neural Network Tuning????

Experiment with more sophisticated architectures (e.g. dropout layers, batch normalization) to improve generalization

5. What strengths do you bring to a team environment?

Holistic Problem-Solving Approach

We combined our domain knowledge, technical expertise, and the individual learning we gained through the DSI program to address the many problems from multiple angles, which ensured that project solutions are technically sound, strategically aligned with business goals, and feasible for implementation

Collaborative Team Approach

We discussed any differences in the business and technical approaches effectively. Our conversations were always respectful and collegial. Any disagreements we experienced were resolved constructively, and overall, we had a positive team environment.

Proactivity and Initiative

We took the initiative to identify potential roadblocks reasonably early and suggest innovative solutions to keep projects on track. We took steps to proactively manage our project so that delays are minimized and team productivity is maximized.

Mindful to Address the Business Problem

Although we were ambitious to apply the new ML skills we learned from the program, our focus was centred on addressing the business problem. Part of this process includes being able to communicate these complex results clearly and effectively to stakeholders that, in this case, includes administrative managers, researchers, and healthcare providers