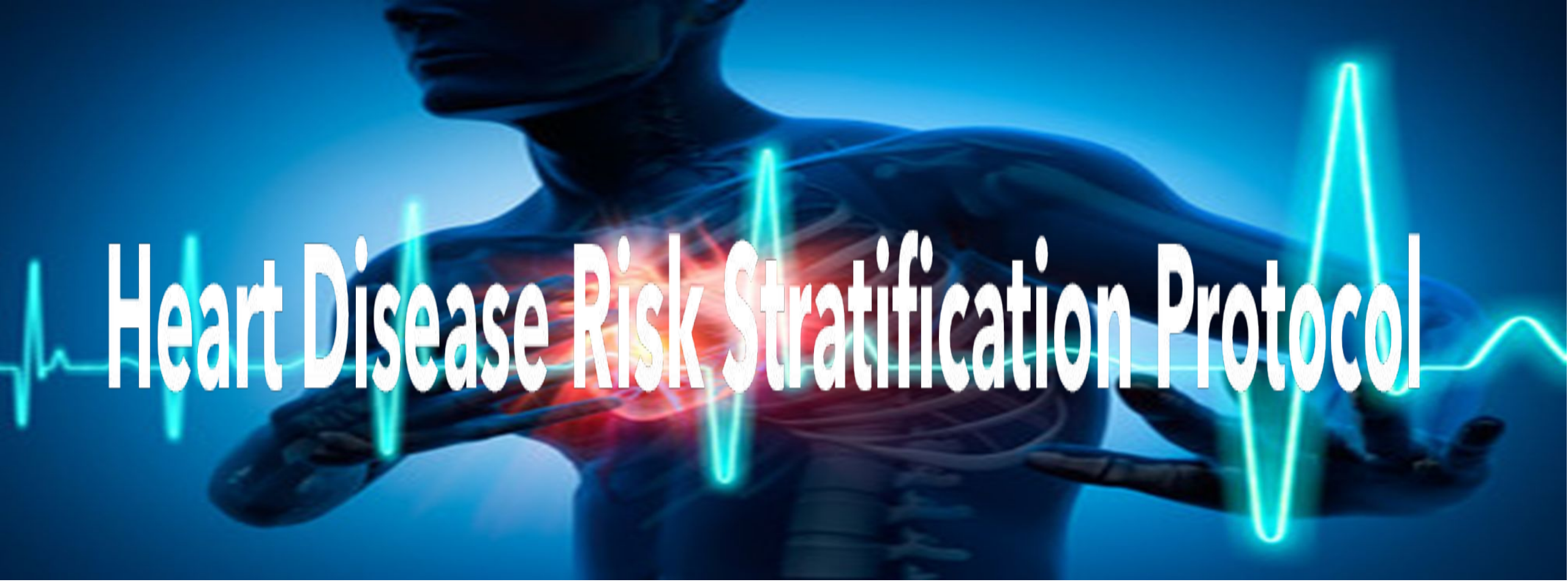


Using an Interpretable Decision Tree - Machine Learning in Healthcare

About Me

Dr. Samuel Israel

- Mb,bs | M.Sc Data Science & AI |
- Healthcare & Clinical Data Scientist

Outlines

- Introduction
- Purpose & Scope
- Interpretable Decision Tree Model
- Depth ≤ 3 for Interpretability
- Inputs Required
- Stepwise Risk Assessment
- Step 1: Baseline Risk Assessment
- Step 2A: Lower Baseline Risk Pathway
- Clinical Action
- Step 2B: Higher Baseline Risk Pathway
- Operating Modes (Threshold Selection)
- Clinical Disclaimer
- Conclusion
- Appreciation

Introduction

Clinical Decision-Support Protocol

This protocol presents an interpretable, machine-learning–based approach to heart disease risk stratification.

It is designed to support clinicians in prioritizing patients for further cardiac evaluation using routinely collected clinical data.

Purpose & Scope

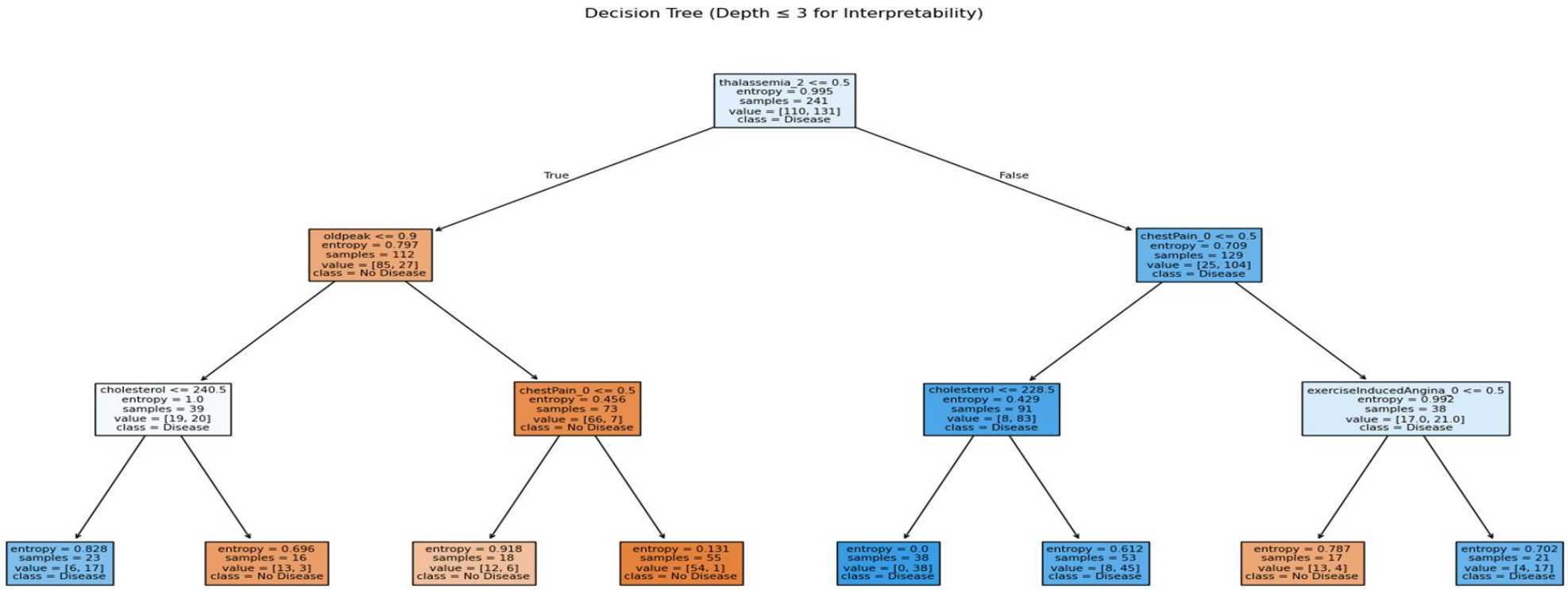
Purpose

- Supports clinicians in risk stratification
- Assists in prioritizing patients for further cardiac evaluation
- Uses routinely available clinical features

Merits

- Decision-support tool only
- Does not replace clinical judgment or diagnostic testing.

Interpretable Decision Tree Model



Depth ≤ 3 for Interpretability

Decision Tree:

- Tree depth is intentionally limited to maintain transparency, allowing each prediction to be explained through a small number of clinically meaningful steps.
- Key splits include thalassemia category, ST depression (oldpeak), chest pain type, cholesterol, and exercise-induced angina, all identified through exploratory analysis.
- Each decision path forms a clear IF–THEN rule to support clinical risk stratification rather than diagnosis

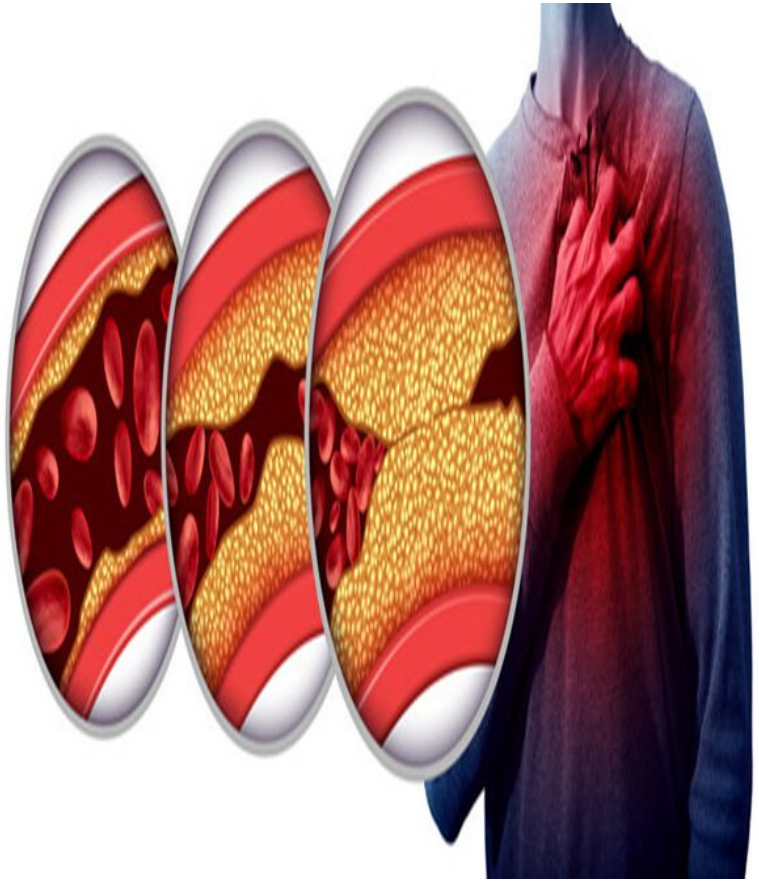
Inputs Required

Routine Clinical Data:

- Thalassemia category
- Chest pain type
- ST depression during exercise (Oldpeak)
- Total cholesterol
- Exercise-induced angina (Yes / No)

Stepwise Risk Assessment

- The first step establishes baseline risk using thalassemia category, which emerged as the strongest initial risk separator in the model.
- Patients are routed into lower or higher baseline risk pathways accordingly.



Step 1: Baseline Risk Assessment

Check Thalassemia Category

- Category = 2→ Higher baseline risk→ Proceed to Step 2B
- Category ≠ 2→ Lower baseline risk→ Proceed to Step 2A

Step 2A: Lower Baseline Risk Pathway

Thalassemia ≠ 2

Assess ST Depression (Oldpeak):

- Oldpeak ≤ 0.9
  - Chest pain = type 0 → Low Risk
  - Chest pain ≠ type 0 → Moderate Risk
- Oldpeak > 0.9
  - Cholesterol > ~240 mg/dL → Elevated Risk
  - Otherwise → Moderate Risk

Clinical Action

- High-risk patients should be prioritized for diagnostic evaluation, while moderate-risk patients require close follow-up.
- The protocol emphasizes early identification without replacing clinical discretion.

Clinical Action

Lower Baseline Risk

- Low Risk
  - Routine monitoring
- Moderate / Elevated Risk
  - Consider ECG or stress testing

Step 2B: Higher Baseline Risk Pathway

Thalassemia = 2

Assess Chest Pain Type:

- Chest pain ≠ type 0
  - Cholesterol ≤ ~228 mg/dL → High Risk
  - Otherwise → Moderate–High Risk
- Chest pain = type 0
  - Exercise-induced angina present → High Risk
  - Absent → Moderate Risk

Clinical Action

Higher Baseline Risk

- High Risk
  - Prioritize diagnostic evaluation
- Moderate Risk
  - Close follow-up and further assessment

Operating Modes

Use Case	Threshold	Goal
Screening / Triage	Low (≈ 0.05)	Minimize missed disease (Recall ≈ 97%)
Decision Support	Moderate (≈ 0.20)	Balance sensitivity & specificity

Clinical Disclaimer

- Reflects statistical associations, not causation
- Derived from a specific dataset
- External validation required before clinical deployment
- Intended to support, not replace, clinician decision-making

Conclusion

- A shallow Decision Tree was chosen to balance performance and transparency, producing clear IF–THEN rules aligned with clinical reasoning.
- Healthcare-focused evaluation and threshold optimization enabled both high-sensitivity screening and balanced decision-support use cases.
- Overall, the project emphasizes the importance of interpretability, appropriate metrics, and responsible modeling in healthcare machine learning.

THANK YOU!

Linkedin:
www.linkedin.com/in/dr-samuel-israel-90893b228