Patient Treatment Classification Project

BY SHARON CHELANGAT.

SCOPE:

- Overview
- Business and Data Understanding
- Modeling
- Evaluation
- Recommendations
- Next Steps
- Thank you

OVERVIEW:

- **Project:** Analyzing patient treatment data for improved healthcare outcomes.
- **Objective:** Predict optimal treatment strategies using machine learning.
- **Goal:** Provide actionable recommendations for healthcare professionals.
- **Presentation Agenda:** Business and data understanding, modeling, evaluation, recommendations, next steps.
- **Data:** Patient treatment data.

Business and Data Understanding

Business Context:

Importance of accurate treatment predictions in healthcare.

Need for personalized treatment strategies.

Data Understanding

Dataset:
Electronic Health Record Predicting
Data Source:
Private Hospital in Indonesia
Purpose:
Determine patient classification (in-care or out-care) based on laboratory test results and attributes

Data Understanding

Attribute Information:

Variables

Haematocrit, Haemoglobins, Erythrocyte, Leucocyte, Thrombocyte, MCH, MCHC, MCV, Age, Sex, Source.

Measurement Unit

Continuous (for laboratory test results), Nominal-Binary (for sex), Nominal (for source).

Note: *Patient privacy and data protection regulations have been strictly adhered to, ensuring anonymity and compliance.*

Modeling

Utilizing Patient Treatment Data for Predictive Modeling

Machine Learning Models Used:

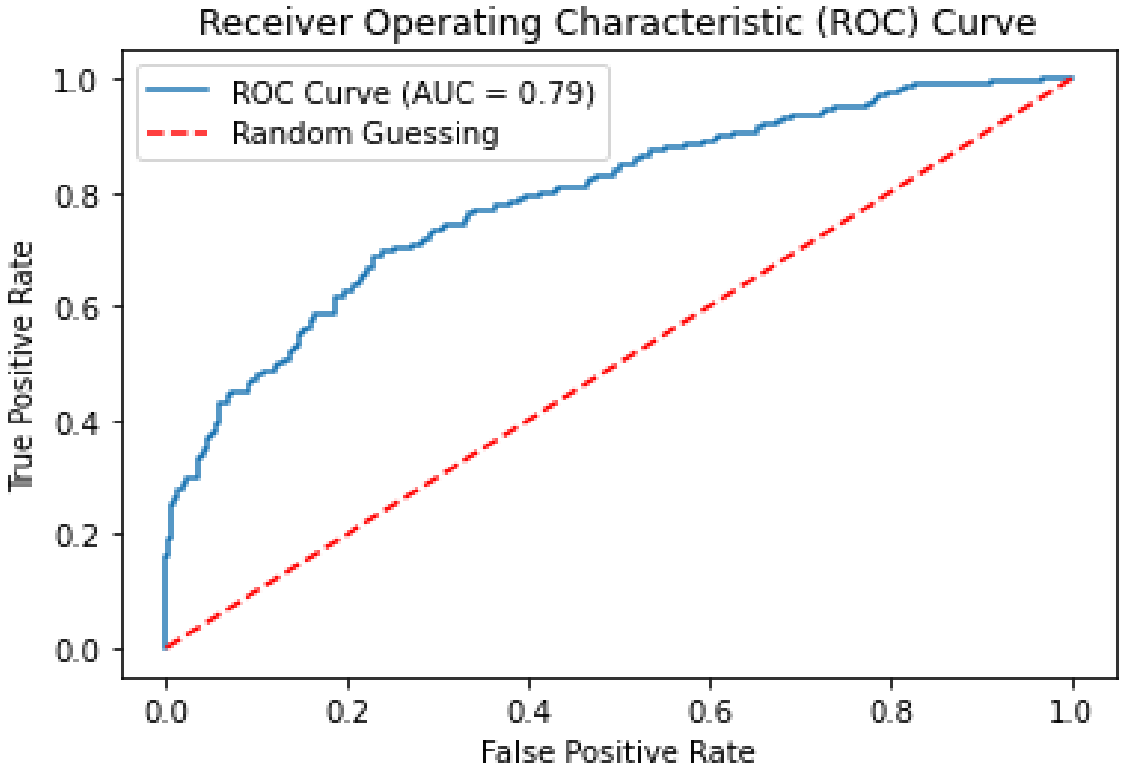
- K-Nearest Neighbors (KNN)
- Decision Trees
- Random Forest

Model Training and Evaluation:

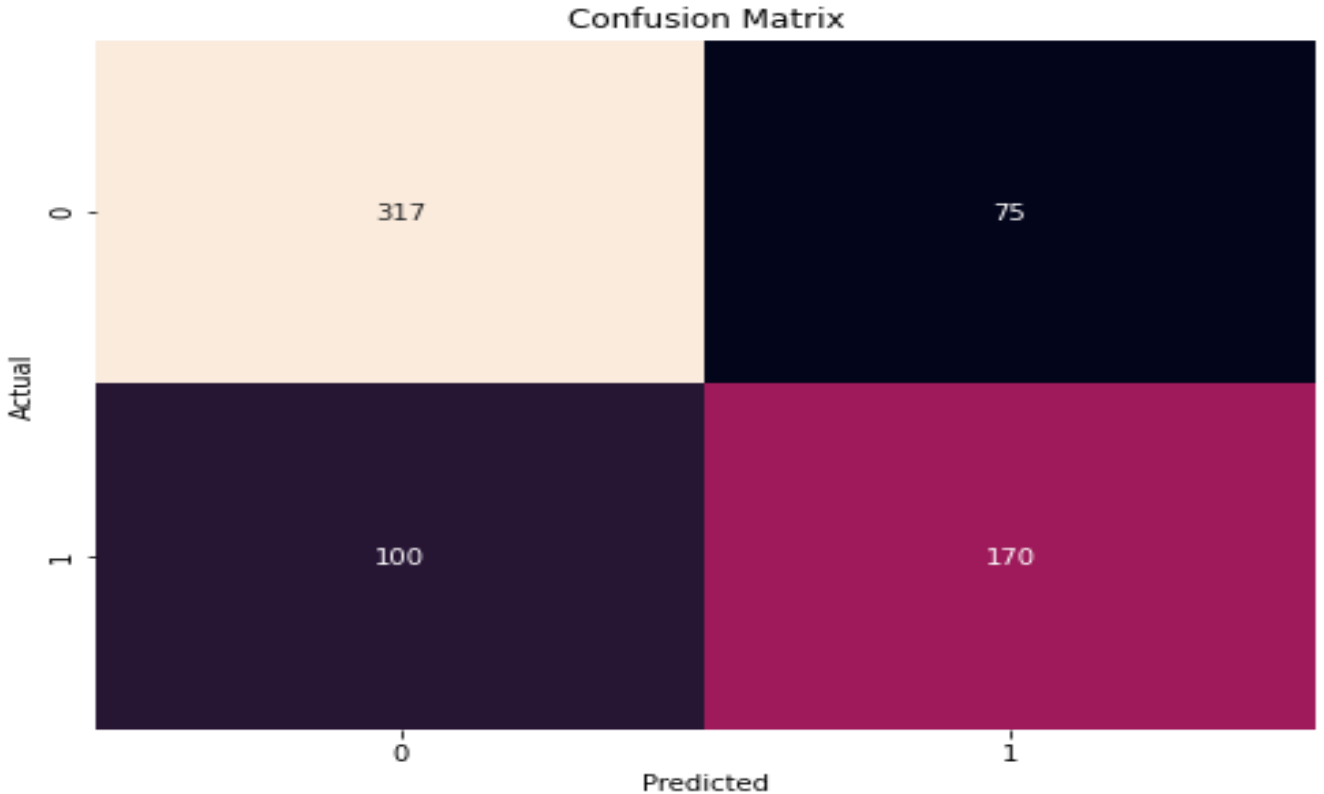
- Data preprocessing and feature engineering
- Splitting data into training and testing sets
- Training models on the training set
- Evaluating model performance using appropriate metrics

Modeling

Accuracy curve



Confusion matrix



Evaluation

Model Performance Evaluation:

- Accuracy: Measures the overall correctness of the model's predictions.
- Precision: Measures the proportion of true positive predictions among all positive predictions.
- Recall: Measures the proportion of true positive predictions among all actual positive instances.

Evaluation

KNN: Accuracy = 71%, Precision = 64%, Recall = 63%

Decision Trees: Accuracy = 69%, Precision = 62%, Recall = 62%

Random Forest: Accuracy = 72%,Precision = 65%, Recall = 63%

Recommendations

- ENSURE DATA QUALITY: VALIDATE AND CLEAN THE INPUT DATA TO IMPROVE THE MODEL'S PERFORMANCE AND RELIABILITY.
- SELECT RELEVANT FEATURES: CONTINUOUSLY EVALUATE AND UPDATE THE SELECTED FEATURES BASED ON THEIR IMPORTANCE AND RELEVANCE TO IMPROVE PREDICTION ACCURACY.
- MONITOR AND RETRAIN THE MODEL: REGULARLY MONITOR THE MODEL'S PERFORMANCE AND RETRAIN IT WITH NEW DATA TO KEEP IT UP TO DATE AND ACCURATE.

Next Step

- Collect more data for improved model accuracy.
- Explore alternative machine learning algorithms.
- Collaborate with healthcare professionals for validation and refinement.
- Continuously update and monitor the model.
- Incorporate additional factors for better predictions.

THANK YOU!!

Thank you for your time and attention.
We appreciate your feedback.