Final Project

NAME: DHARAN.N
DEPT.: ARTIFICIAL INTELLIGENCE AND DATA SCIENCE 3RD YEAR.
NAAN MUDHALVAN ID: au821721243013.
COLLEGE: SIR ISSAC NEWTON COLLEGE OF ENGINEERING AND TECHNOLOGY.

HEART DISEASE PREDICTION USING LOGISTIC REGRESSION

AGENDA

- PROBLEM STATEMENT.
- PROJECT OVERVIEW.
- WHO ARE THE END USERS?
- SOLUTION AND ITS VALUE PROPOSITION.
- THE WOW IN THE SOLUTION.
- MODELLING.
- DATA VISUALIZATION.
- RESULT.

PROBLEM STATEMENT

- Heart disease remains one of the leading causes of death worldwide, emphasizing the importance of early detection and prevention.

- Machine learning techniques offer promising avenues for predicting heart disease based on various risk factors, aiding healthcare professionals in making informed decisions and improving patient outcomes.

PROJECT OVERVIEW

- The primary objective of this project is to build a reliable and accurate predictive model capable of assessing an individual's risk of heart disease using logistic regression analysis.

- By analyzing various risk factors such as age, gender, cholesterol levels, and blood pressure, the model aims to provide healthcare professionals with valuable insights for early intervention and personalized care.

WHO ARE THE END USERS?

1.Healthcare Providers: Physicians, cardiologists, and nurses who use the model to assess patients' risk of heart disease and guide treatment decisions.

2.Medical Researchers: Researchers analyzing heart disease trends and risk factors within populations.

3.Healthcare Administrators: Hospital administrators and policymakers using the model to allocate resources and plan public health initiatives.

4.Patients: Individuals benefiting from personalized care and early interventions to reduce their risk of heart disease.

5.Health Insurance Companies: Insurers leveraging the model to assess policyholders' risk profiles and develop tailored wellness programs.

SOLUTION AND ITS VALUE PROPOSITION

- Our solution involves developing a predictive model using logistic regression to assess an individual's risk of heart disease based on demographic and clinical data.

- By analyzing factors such as age, gender, cholesterol levels, and blood pressure, our model predicts the likelihood of heart disease occurrence, enabling timely interventions and personalized care.

Value Proposition:

1.Early Detection: Identifying heart disease risk early, allowing timely interventions.

2.Personalized Care: Tailoring treatment plans based on individual risk profiles.

3.Resource Optimization: Efficient allocation of resources for targeted interventions.

4.Data-Driven Insights: Informing research and public health policies.

5.Cost Savings: Reducing healthcare costs associated with heart disease.

THE WOW IN THE SOLUTION

1.Precision Early Detection: Pinpointing heart disease risk with unparalleled accuracy, ensuring timely interventions and saving lives.

2.Personalized Proactive Care: Crafting bespoke treatment plans tailored to individuals' unique risk profiles, empowering patients to take proactive control of their health.

3.Resource Efficiency Redefined: Maximizing healthcare resources by targeting interventions precisely where they're needed most, optimizing outcomes and minimizing waste.

4.Insights that Drive Impact: Unveiling deep insights into heart disease trends and risk factors, guiding policymakers and researchers to enact transformative public health strategies.

5.Cost-Effective Innovation: Revolutionizing healthcare economics by curbing costs associated with heart disease treatment through proactive prevention and early intervention.

MODELLING

1.We load the heart disease dataset and preprocess it by dropping any rows with missing values.

2.We select relevant features (e.g., age, gender, cholesterol levels, blood pressure) and the target variable (heart disease status).

3.We split the dataset into training and testing sets to evaluate the model's performance.

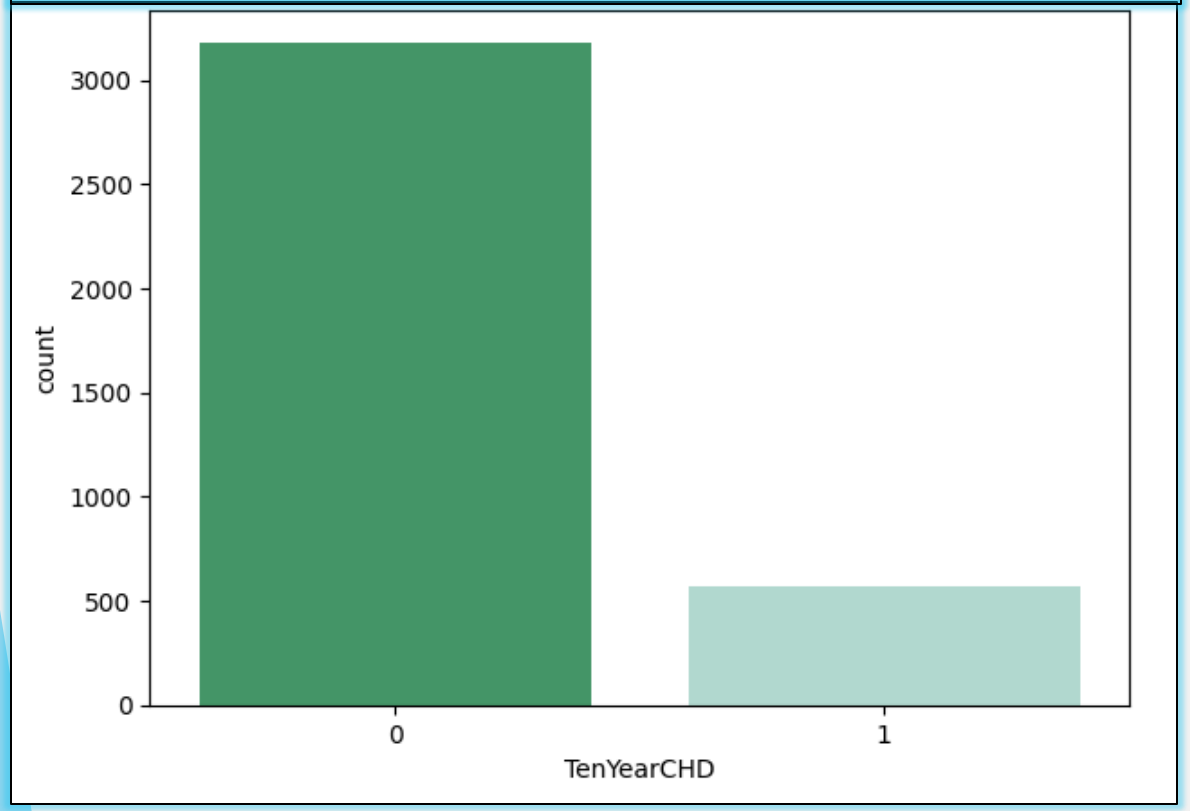
4.We scale the features using StandardScaler to ensure they have a similar scale, which is important for logistic regression.

5.We train a logistic regression model using the training data.

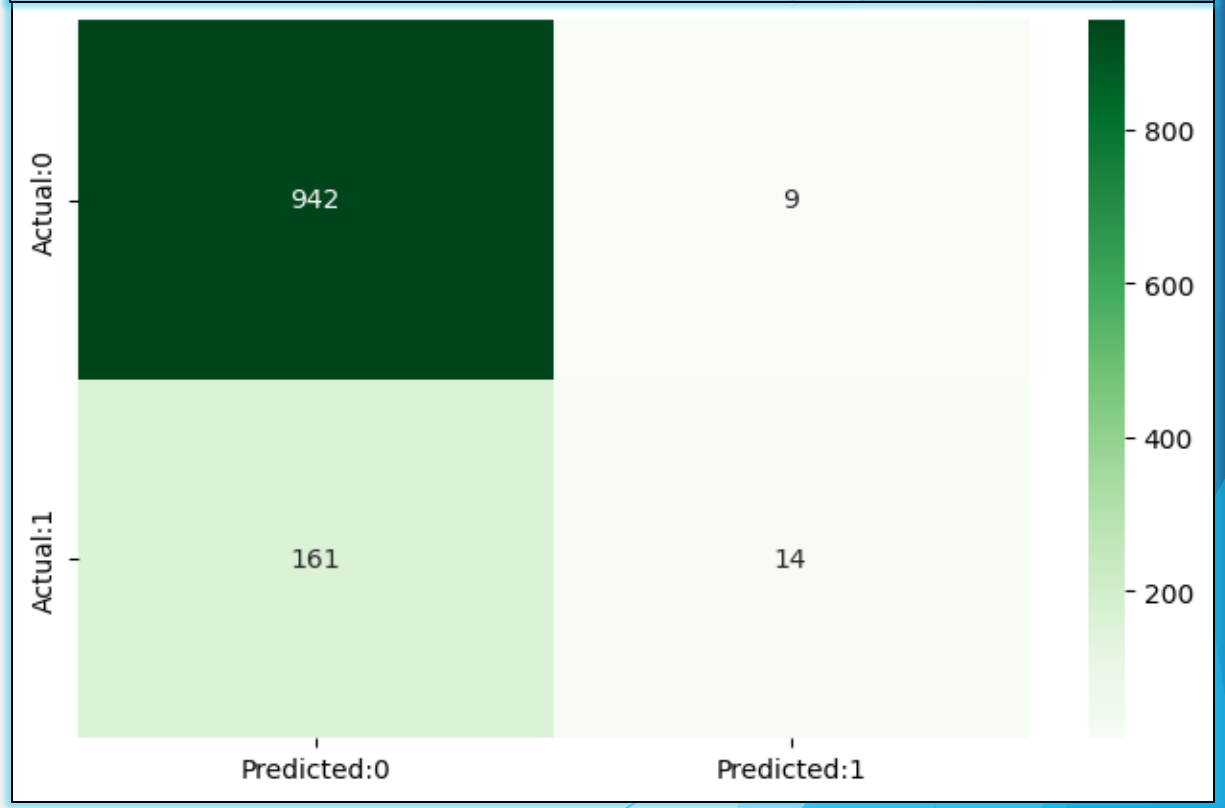
6.We evaluate the trained model's performance using accuracy and classification report metrics on the testing set.

DATA VISUALIZATION

COUNTING NO. OF PATIENTS AFFECTED WITH CHD



CONFUSION MATRIX HEATMAP



RESULTS

1.**Class 0 (Positive Class)**:

1. **Precision**: 0.85
2. **Recall**: 0.99
3. **F1-score**: 0.92
4. **Support**: 951

2.**Class 1 (Negative Class)**:

1. **Precision**: 0.61
2. **Recall**: 0.88
3. **F1-score**: 0.14
4. **Support**: 175

Therefore, The overall accuracy of the model is **0.85**.

The details for confusion matrix is =

	precision	recall	f1-score	support
0	0.85	0.99	0.92	951
1	0.61	0.08	0.14	175
accuracy			0.85	1126
macro avg	0.73	0.54	0.53	1126
weighted avg	0.82	0.85	0.80	1126