Title: Heart Disease Detection Using Machine Learning Techniques: A Comparative Analysis

INTERNSHIP PROJECT REPORT

Submitted in partial fulfillment of the requirements for the award of the degree

Of

MASTER of TECHNOLOGY
DEPARTMENT OF ARTIFICIAL INTELLIGENCE & DATA SCIENCES

By

Nayan Saxena
00902212023

Guided by

Shailesh D Kamble
Associate Professor
Indira Gandhi Delhi Technical University For Women

INDIRA GANDHI DELHI TECHNICAL UNIVERSITY FOR WOMEN
NEW DELHI – 110006

June 2024-August 2024

CERTIFICATE

I, Nayan Saxena, certify that the Internship Project Report entitled Heart Disease Detection Using Machine Learning Techniques: A Comparative Analysis is done by me and it is authentic work carried out by me at Indira Gandhi Delhi Technical University For Women. For this project, no work has been submitted before for any degree or diploma of the award, to the best of my knowledge and belief.

Signature of the student
Date: 04/12/2024

Certified that the project report entitled Heart Disease Detection Using Machine Learning Techniques: A Comparative Analysis done by the above student is completed under my guidance.

Signature of Mentor
Date: 04/12/2024
Mentor Name: Dr.Shailesh D Kamble
Designation: Associate Professor
Compnay/Univeristy Name: Indira Gandhi Delhi Technical University For Women

UNDERTAKING REGARDING ANTI-PLAGIARISM

I, ***Heart Disease Detection Using Machine Learning Techniques: A Comparative Analysis***, hereby, declare that the material/ content presented in the report are free from plagiarism and is properly cited and written in my own words. In case, plagiarism is detected at any stage, I shall be solely responsible for it.A copy of the Plagiarism Report is also enclosed.

Signature of the Student

Nayan Saxena

00902212023

ACKNOWLEDGEMENT

I would like to acknowledge my mentor Dr.Shailesh D Kamble for his very helpful comments, support and encouragement.

Finally, I am grateful to Indira Gandhi Delhi Technical University For Women, for providing a healthy, supportive and understanding environment. They allowed me the freedom to explore innovative models to simplify a complex business problem. This made my project work possible without any hindrance.

Signature of the Student

Nayan Saxena

00902212023

DECLARATION

I, **Nayan Saxena** solemnly declare that the internship project report, ***Heart Disease Detection Using Machine Learning Techniques: A Comparative Analysis,*** is based on my own work carried out under the supervision of **Dr.Shailesh D Kamble.** I assert the statements made and conclusions drawn are an outcome of my work. I further certify that:

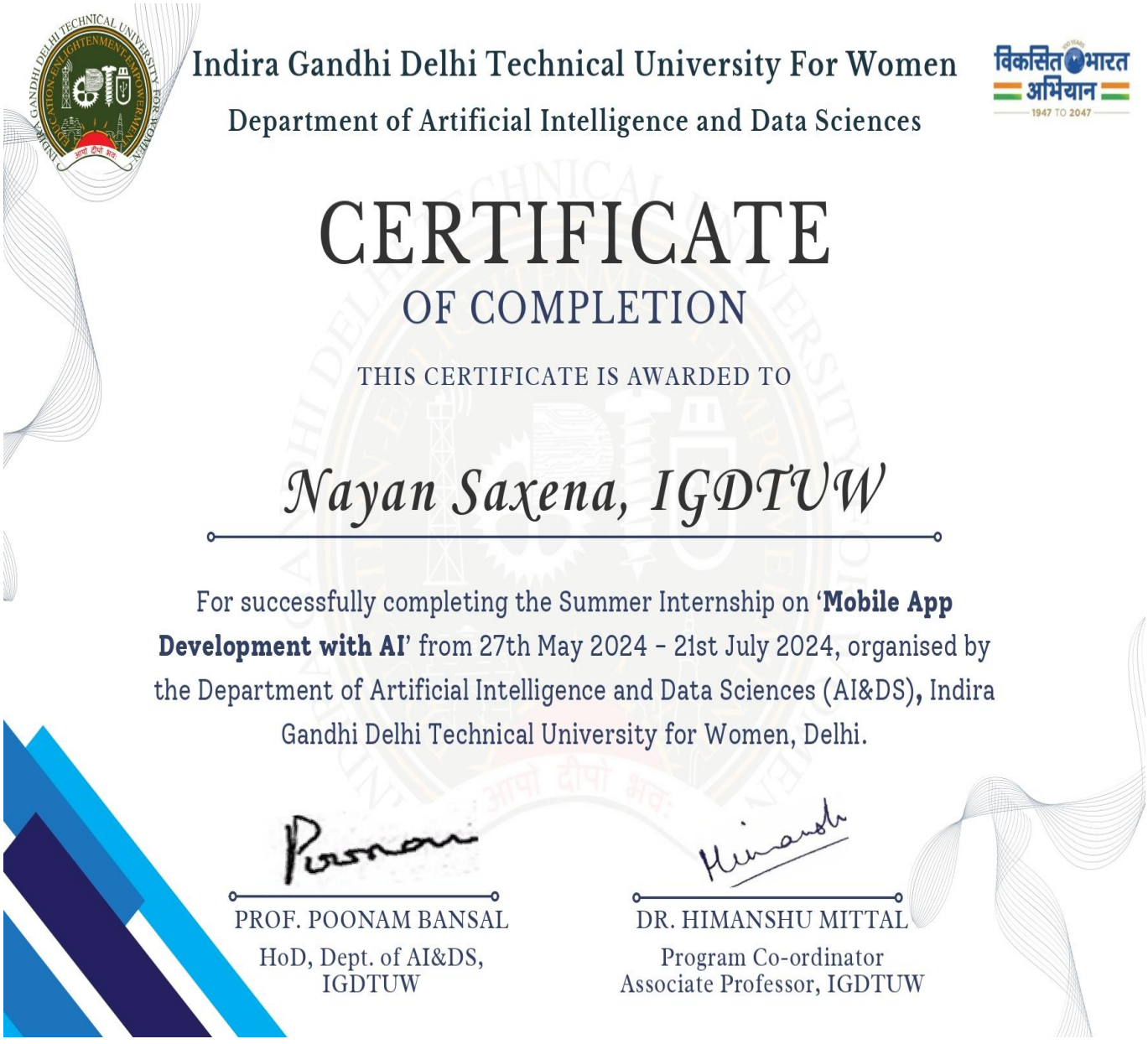
I. The work contained in the report is original and has been done by me under the supervision of my supervisor.
II. The work has not been submitted to any other Institution for any other degree/diploma/certificate in this university or any other University of India or abroad.
III. We have followed the guidelines provided by the university in writing the report.
IV. Whenever we have used materials (text, data, theoretical analysis/equations, codes/program, figures, tables, pictures, text etc.) from other sources, we have given due credit to them in the report and have also given their details in the references.

Signature of the Student

Nayan Saxena

00902212023

Internship Certificate



Indira Gandhi Delhi Technical University For Women

Department of Artificial Intelligence and Data Sciences

विकसित भारत अभियान 1947 TO 2047

CERTIFICATE

OF COMPLETION

THIS CERTIFICATE IS AWARDED TO

Nayan Saxena, IGDTUW

For successfully completing the Summer Internship on '**Mobile App Development with AI**' from 27th May 2024 - 21st July 2024, organised by the Department of Artificial Intelligence and Data Sciences (AI&DS), Indira Gandhi Delhi Technical University for Women, Delhi.

PROF. POONAM BANSAL
HoD, Dept. of AI&DS,
IGDTUW

DR. HIMANSHU MITTAL
Program Co-ordinator
Associate Professor, IGDTUW

TABLE OF CONTENTS

Introduction

Heart disease is a leading cause of death worldwide. Early detection plays a crucial role in reducing mortality rates and improving patient outcomes. Leveraging machine learning techniques, this project evaluates and compares the performance of several models—Support Vector Machine (SVM), Bagging, Boosting, Random Forest, k-Nearest Neighbors (KNN), and Logistic Regression—for heart disease detection. The study also highlights significant insights into demographic and clinical patterns using data visualization.

System Requirement Analysis

Software Requirements

- **Operating System:** Windows 10 or higher / Linux / macOS
- **Programming Language:** Python (version 3.7 or above)
- **Libraries and Frameworks:**
  - Scikit-learn
  - Pandas
  - NumPy
  - Matplotlib
  - Seaborn

Hardware Requirements

- **Processor:** Intel i5 or higher / AMD equivalent
- **RAM:** 8 GB minimum (16 GB recommended)
- **Storage:** 20 GB free disk space
- **GPU (Optional):** NVIDIA GTX 1050 or higher for accelerated computations

System Analysis

Feasibility Study

The feasibility study focuses on determining the project's practicality across different dimensions.

Economic Feasibility

- The project is cost-effective, relying on open-source tools like Python and Scikit-learn. No proprietary software or expensive hardware is required.

Implementation Feasibility

- The models were implemented using well-documented libraries and require basic machine learning expertise. The dataset used was publicly available, ensuring easy reproducibility.

Work Description

1. **Dataset Preparation:**
    - Cleaned and normalized clinical data.
    - Features included age, gender, chest pain type, blood pressure, cholesterol, fasting blood sugar, ST depression, and exercise-induced angina.
    - Data split: 70% for training, 30% for testing.

2. **Model Implementation:**
    - **SVM, Bagging, Boosting, Random Forest, KNN, Logistic Regression** were trained and evaluated.

- Ensemble techniques (Bagging, Boosting, and Random Forest) enhanced predictive accuracy.

3. **Data Visualization:**
   - Observed heart disease prevalence by demographics and symptoms.
   - Identified age-related trends in blood pressure and cholesterol.

4. **Performance Metrics:**
   - Models evaluated using accuracy scores and confusion matrices.

Work Outcome

Model Performance

Model	Accuracy	Comments
SVM	99%	Excellent for non-linear data and high-dimensional spaces.
Bagging	99%	Reduced variance, highly stable predictions.
Boosting	99%	Corrected weak learners for enhanced performance.
Random Forest	99%	Effective with high-dimensional and noisy data.
KNN	96%	Simple, struggled with large and noisy data.
Logistic Regression	85%	Interpretability was its key strength.

Visualization Insights

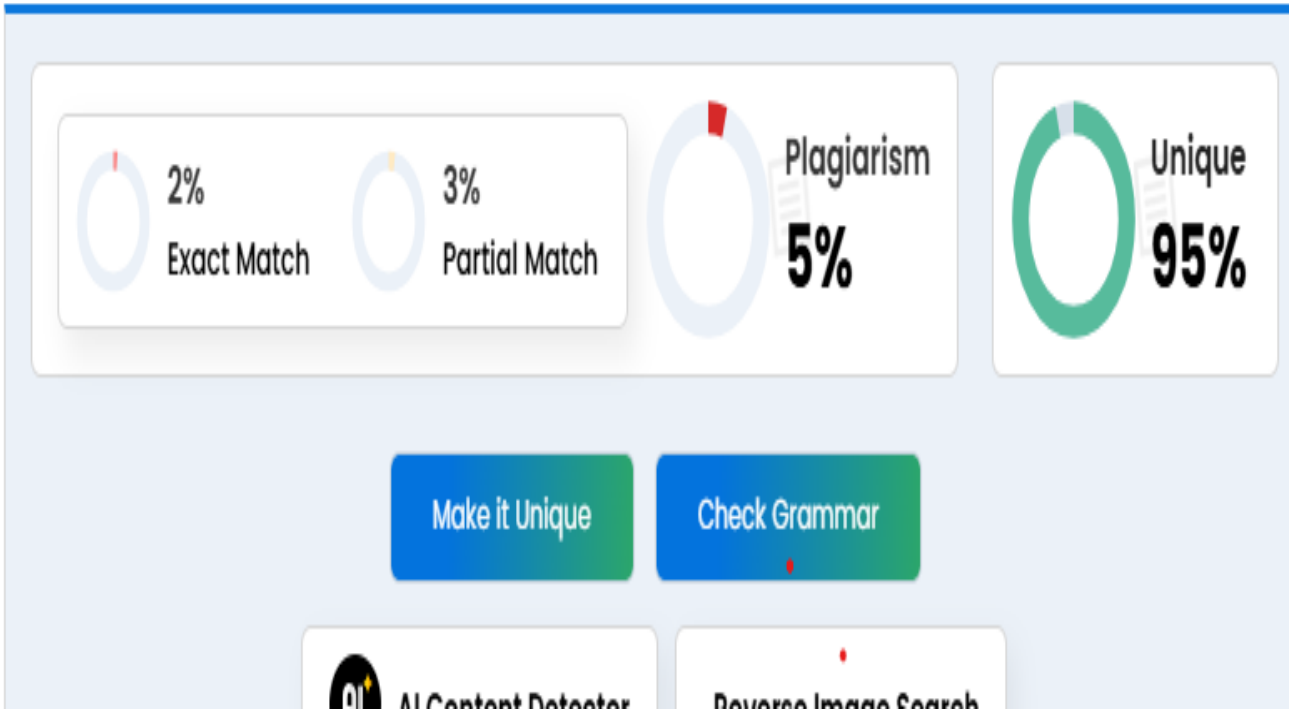
1. **Prevalence:**
    - 54% of the population had heart disease.
    - Males had a higher prevalence compared to females.
2. **Chest Pain Analysis:**
    - 85–90% of heart disease cases were asymptomatic.
    - Non-anginal chest pain was more common in males.
3. **Age Trends:**
    - Blood pressure and cholesterol peaked between ages 42–77.
    - Post 77, cholesterol and blood pressure declined.
4. **Gender-based Observations:**
    - ST Depression (Myocardial ischemia) and fasting blood sugar levels were more prevalent in males.
5. **Exercise-induced Angina:**
    - Exercising with angina significantly increased the risk of heart disease.

Conclusion

This study demonstrates the effectiveness of machine learning models in detecting heart disease, with SVM, Bagging, Boosting, and Random Forest achieving 99% accuracy. Data visualizations provided critical insights into demographic and clinical patterns, emphasizing the importance of machine learning tools in clinical decision-making. Future work will focus on improving model explainability and applying these findings to real-world datasets.

Plagiarism Report

95% of the content was found unique due to research findings solely being derived from the tasks assigned under the internship.



References

1. Bishop, C. M. (2006). Pattern Recognition and Machine Learning.
2. Breiman, L. (2001). Random Forests. Machine Learning.
3. Freund, Y., & Schapire, R. E. (1997). A Decision-Theoretic Generalization of On-Line Learning and an Application to Boosting.
4. Cortes, C., & Vapnik, V. (1995). Support-Vector Networks. Machine Learning.

Acceptance Notification and Review Result of Your Paper - JETIR551909 | JETIR (ISSN:2349-5162) | www.jetir.org | editor@jetir.org Your Email id: nayangm002@gmail.com

Jetir Journal <editor@jetir.org>
To: nayangm002@gmail.com

4 December 2024 at 15:56

International Journal of Emerging Technologies and Innovative Research(JETIR)
An International Scholarly Open Access Journal, Peer-Reviewed, Refereed Journal Impact Factor 7.95 Calculate by Google Scholar and Semantic Scholar | AI-Powered Research Tool, Multidisciplinary, Monthly, Multilanguage Journal Indexing in All Major Database & Metadata, Citation Generator, Peer-Reviewed, Refereed, Indexed, automatic Citation Open Access Journal

Acceptance Notification and Review Result of Your Paper - JETIR551909 | JETIR (ISSN:2349-5162) | www.jetir.org | editor@jetir.org Your Email id: nayangm002@gmail.com
Track Your Paper Link Track Your Paper https://www.jetir.org/trackauthorhome.php?a_rid=551909

Dear NAYAN SAXENA,
Your manuscript with Registration ID: JETIR551909 has been **Accepted** for publication in the Journal of Emerging Technologies and Innovative Research (www.jetir.org). Track Your Paper Link Track Your Paper https://www.jetir.org/trackauthorhome.php?a_rid=551909 Your Review Report is as follows:

An International Scholarly Open Access Journal, Peer-Reviewed, Refereed Journal Impact Factor 7.95
UGC Approved Journal No 63975(19)
Review Results:

Registration ID	JETIR551909
Email ID	nayangm002@gmail.com
Paper Title	Heart Disease Prediction using Machine Learning Techniques
Review Status	Accepted
Impact Factor & Licence:	Open Access, Peer-Reviewed, Refereed, Indexing,ISSN Approved,DOI and Creative Common Approved & 7.97 Calculated by Google Scholar
Unique Contents	88 %
Comments	Paper Accepted Complete Payment and documents Process. Complete Phase 2 Payment and Phase 3 Documents Process so within 1 to 2 day it will publish

Phase 2: Pay Publication Fees and Open Access Processing Charges
For Indian Author Payment Details
For Foreign/International Author payment details

For Publishing a Paper at JETIR, the Author needs to pay publication charges, Open Access Processing Charges as follows.

For Indian Authors ₹1570/- INR (online Payment) **Publication Charge Details**	**Best/Fast Payment Link 1:** ₹1570/- INR PAY **Payment Link 2:** Click Here to Pay **Payment Link 3:** Click Here to Pay **BANK Details**: Bank Account Details
For Foreign/International author $59/- (USD)	
Only for Foreign/International author $59/- (USD)	**Payment Link 1 :** Click Here to Pay **Payment Link 2 :** Click Here to Pay **Payment Link 3 :** Click Here to Pay **Payment Link 4 :** Click Here to Pay **BANK Details**: Bank Account Details

After Sending this all 3 Documents to editor@jetir.org Please wait for 2 to 3 Days for Our acknowledgment.

Phase 3: Submission of Documents
Track Your Paper Link https://www.jetir.org/trackauthorhome.php?a_rid=551909

Please complete payment and send payment proof along with below documents

On **Acceptance of Paper** Kindly submit following Documents to editor@jetir.org
Visit Track Your Paper Link Track Your Paper https://www.jetir.org/trackauthorhome.php?a_rid=

551909 or AUTHORHOME using your registration ID (551909) and email ID (nayangm002@gmail.com) to download Undertaking Form.

1. Final Paper in word file if no changes than send same word file of paper as you send | or Final Paper with all corrections and as per JETIR format in doc/docx format Download
2. Undertaking of Authors (Must be manually signed , only one author sign required) Download
3. Payment Receipt or Payment proof or Screenshot of payment or account statement (any format accepted)

If all author not available than complete above process with single author all author signatures not compulsory complete process with single author only. Send all Documents to editor@jetir.org wait for 1 to 2 Day for Status updation. If all author not available than complete below process with single author

Publication of Paper	Within 01-02 Days after Submitting all required 3 documents to editor@jetir.org.

This is System Generated Email. Kindly Mention your registration Id for any informaiton retrival send query to editor@jetir.org .

Web Admin
JETIR - International Journal of Emerging Technologies and Innovative Research
www.jetir.org | editor@jetir.org
A complete Eco-Friendly Journal
ePaperSubmission | eDocuments | ePayment | eReviewCard | eCertificates | ePaymentReceipt

Note: Please consider the environment before printing this email

Confidentiality Note: This e-mail and any attachments are confidential and may be protected by legal privilege. If you are not the intended recipient, be aware that any disclosure, copying, distribution or use of this e-mail or any attachment is prohibited. If you have received this e-mail in error, please notify me immediately by returning it and delete this copy from your system. Thank you for your cooperation.

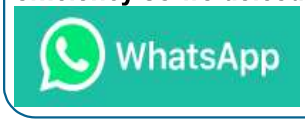
Communication Process:
Every Day we received very large number of call so it's decrease our work efficiency so we deiced to provide support on WhatsApp and Email only.

Option 1: **Save Our Contact No in your mobile phone with name "Jetir Journal'**	WhatsApp Contact: 9426033211 WhatsApp Support +919426033211 Note: Do not call on this number Only send your query to WhatsApp only
Option 2: **Send your query through Email. Our team will respond you within 2 days.**	Email id: editor@jetir.org . Note: Send all query to this email address editor@jetir.org
Option 3: **In the case when your query doesn't resolve using above two options then follow below procedure to contact Editor of JETIR Send me "Call Request " to WhatsApp and email With your mobile no so our editor sir will contact you soon.**	Send me "Call Request" to WhatsApp and email With your mobile no so our editor sir will contact you . WhatsApp Contact: 9426033211 Email: editor@jetir.org Send me "Call Request " to WhatsApp and email With your mobile no so our editor sir will contact you . WhatsApp Contact: 9426033211 editor@jetir.org .

NOTE: You can use any of above three options and we will solve your all problem within 2 to 3 day. Thank you for your kind support.

Thanks and Regards,

Communication Process:
Every Day we received very large number of calls so it's decreased our work efficiency so we deiced to provide support on WhatsApp and Email only.



Option 1: **Save Our Contact No in your mobile phone with name "Jetir Journal"**	WhatsApp Contact 1: 9426033211 Note: Do not call on this number Only send your query to WhatsApp only
Option 2: **Send your query through Email. Our team will respond you within 2 days.**	Email id: editor@jetir.org . Note: Send all query to this email address editor@jetir.org
Option 3: **In the case when your query doesn't resolve using above two options then follow below procedure to contact Editor of JETIR Send me "Call Request " to WhatsApp and email With your mobile no so our editor sir will contact you soon.**	Send me "Call Request" to WhatsApp and email With your mobile no so our editor sir will contact you . WhatsApp Contact 1: 9426033211 WhatsApp Contact 2: 7048421102 Email: editor@jetir.org Send me "Call Request " to WhatsApp and email With your mobile no so our editor sir will contact you . WhatsApp Contact 1: 9426033211 WhatsApp Contact 2: 7048421102 editor@jetir.org .

NOTE: You can use any of above three options and we will solve your all problem within 2 to 3 day. Thank you for your kind support.

Editor in Chief
JETIR - Journal of Emerging Technologies and Innovative Research
(UGC and ISSN Approved Journal)
www.jetir.org | editor@jetir.org
A complete Eco-Friendly Journal : ePaperSubmission | eDocuments | ePayment | eReviewCard | eCertificates | ePaymentReceipt

---

Note: Please consider the environment before printing this email

Confidentiality Note: This e-mail and any attachments are confidential and may be protected by legal privilege. If you are not the intended recipient, be aware that any disclosure, copying, distribution or use of this e-mail or any attachment is prohibited. If you have received this e-mail in error, please notify me immediately by returning it and delete this copy from your system. Thank you for your cooperation.