

HEART DISEASE PREDICTION



A Project report submitted in partial fulfilment of requirements for the award of degree of

BACHELOR OF TECHNOLOGY IN COMPUTER SCIENCE AND ENGINEERING (AI & ML)

by

ADIKE ARUN KUMAR (219X1A3333)

GOSULA SASI KUMAR REDDY (219X1A33147)

S.VEERA JAYAPRATHAP REDDY (219X1A3359)

Under the esteemed guidance of

Sri. Syed Nadeem
Assistant Professor
Department of ECS.

Department of Emerging Technologies in Computer Science

G. PULLA REDDY ENGINEERING COLLEGE (Autonomous): KURNOOL
(Affiliated to JNTUA, ANANTAPURAMU)

2024 – 2025

Department of
EMERGING TECHNOLOGIES IN COMPUTER SCIENCE
G. PULLA REDDY ENGINEERING COLLEGE (Autonomous): KURNOOL
(Affiliated to JNTUA, ANANTAPURAMU)



CERTIFICATE

This is to certify that the Project Work entitled 'Heart Disease Prediction' is a bonafide record of work carried out by

ADIKE ARUN KUMAR (219X1A3333)

GOSULA SASI KUMAR REDDY (219X1A33147)

S.VEERA JAYAPRATHAP REDDY (219X1A3359)

Under my guidance and supervision in partial fulfilment of the requirements for the
award of degree of

BACHELOR OF TECHNOLOGY
IN
COMPUTER SCIENCE AND ENGINEERING
(AI & ML)

Dr. S. Shabana Begam

Associate Professor,
Department of ECS.,
G. Pulla Reddy Engineering College,
Kurnool.

Dr. R. Praveen Sam

Professor & Head of the Department,
Department of ECS.,
G. Pulla Reddy Engineering College,
Kurnool.

Signature of the External Examiner :

DECLARATION

We hereby declare that the project titled “HEART DISEASE PREDICTION” is an authentic work carried out by us as the students of **G. PULLA REDDY ENGINEERING COLLEGE (Autonomous) Kurnool**, during 2024-2025 and has not been submitted elsewhere for the award of any degree or diploma in part or in full to any institute.

ADIKE ARUN KUMAR
(219X1A3333)

GOSULA SASI KUMAR REDDY
(219X1A3347)

S.VEERA JAYA PRATHAP REDDY
(219X1A3359)

ACKNOWLEDGEMENT

We wish to express our deep sense of gratitude to our project guide **Sri. Syed Nadeem, Assistant Professor** of Emerging Technologies in Computer Science Department, G. Pulla Reddy Engineering College, for her immaculate guidance, constant encouragement and cooperation which have made possible to bring out this project work.

We are grateful to our project in charge **Sri. U. U. Veerendra, Assistant Professor** of Emerging Technologies in Computer Science Department, G. Pulla Reddy Engineering College, for helping us and giving the required information needed for our project work.

We are thankful to our Head of the Department **Dr. R. Praveen Sam Garu**, for his whole hearted support and encouragement during the project sessions.

We are grateful to our respected Principal **Dr. B. Sreenivasa Reddy Garu** for providing requisite facilities and helping us in providing such a good environment.

We are to convey our acknowledgements to all the staff members of the Emerging Technologies in Computer Science department for giving the required information needed for our project work.

Finally, We wish to thank all our friends and well wishers who have helped us directly or indirectly during the course of this project work.

ABSTRACT

TITLE: Heart Disease Pre

Heart disease is a leading global health concern, requiring accurate and timely diagnosis to reduce mortality rates and improve patient outcomes. This project harnesses the power of machine learning to predict the likelihood of heart diseases by analyzing critical patient data, such as age, blood pressure, cholesterol levels, heart rate, and lifestyle habits. Utilizing advanced algorithms like Support Vector Machines (SVM), Random Forests, and Neural Networks, the system processes and models structured data after thorough preprocessing steps, including cleaning, feature selection, and normalization. The machine learning models provide personalized risk scores and actionable diagnostic insights, aiding healthcare professionals in early detection and preventive interventions. Coupled with intuitive visualizations and user-friendly dashboards, this solution enhances clinical decision-making while reducing diagnostic errors and resource strain on healthcare systems. Designed for adaptability and scalability, the system seeks to empower healthcare providers, improve patient care, and pave the way for the integration of AI-driven diagnostics into medical practice

TABLE OF CONTENTS

	Page. No.
1. Introduction	01
1.1 Introduction	01
1.2 Motivation	02
1.3 Problem Definition	02
1.4 Objectives of the Project	04
1.5 Limitations of the Project	05
1.6 Organization of the Report	06
2. System Specifications	08
2.1 Specifications	08
2.1.1 Programming Languages	08
2.1.2 Frameworks	08
2.1.3 Libraries and Tools	09
2.1.4 Data Visualization	09
2.1.5 Development Tools	10
2.2 Hardware and Software Requirements	10
3. Literature Survey	13
3.1 Introduction	13
3.2 Existing Tools and Approaches	14
3.2.1 Sentiment and Emotion Analysis	14
3.2.2 Topic Modelling and Trend Identification	15
3.2.3 Data Retrieval and Preprocessing	15
3.2.4 Visualization Tools	15
3.2.5 Bot Detection and Misinformation Filtering	15
3.3 Gaps in Existing Solutions	16

3.3.1 Limited Sentiment and Emotions Granularity	16
3.3.2 Lack of User-Centric Customization	16
3.3.3 Poor Accessibility and Interpretability	16
3.3.4 Incomplete Noise and Spam Filtering	16
3.3.5 Static and Non-Scalable Architecture	16
3.4 Proposed System	16
4. System Design & Architecture	19
4.1 Introduction	19
4.2 Overall System Workflow	20
4.3 Frontend Architecture	21
4.4 Backend Architecture	23
4.5 Modules and Functionalities	26
4.5.1 Data Collection via Twitter API	26
4.5.2 Preprocessing and Storage	28
4.5.3 Sentiment Analysis using VADER	30
4.5.4 Visualization Components	31
5. Implementation	34
5.1 Frontend Code	34
5.1.1 User Interaction and Form Handling	34
5.1.2 Enhanced Visual Analytics	35
5.2 Backend Code	36
5.2.1 Tweet Extraction	36
5.2.2 Sentiment Classification with VADER	36
5.2.3 Building the Flask Endpoint	37
5.3 Integration and Deployment	37
6. Results and Discussion	40
6.1 Visualization Outputs	40
6.2 Sentiment Distribution Analysis	41

6.2.1 Positive Sentiment Analysis	41
6.2.2 Negative Sentiment Analysis	41
6.2.3 Neutral Sentiment Analysis	42
6.3 Case Study	42
6.3.1 Data Collection Implementation	43
6.3.2 Conclusion of Case Study	44
7. Testing and Validation	47
7.1 Data Collection Testing	47
7.2 Sentiment Analysis Testing	47
7.3 Data Preprocessing Testing	48
7.4 Scalability and Performance Testing	48
7.5 Frontend and User Experience Testing	48
8. Conclusion and Future Enhancements	50
8.1 Conclusion	50
8.2 Future Enhancements	51
9. References	54

LIST OF FIGURES

FIGURE NO	FIGURE NAME	PAGE NO.
4.1	Architectural Diagram	19
4.2	Component Diagram	33
5.1	Use-Case Diagram	39
6.1	Initial Visualization of the Website	45
6.2	Analysing the File	45
6.3	The Output with Bar Graph	46
6.4	Individual Positive Tweets	46
7.1	Checking to give no CSV File	49

LIST OF TABLES

TABLE NO	TABLE NAME	PAGE NO.
2.2.1	Hardware Requiremennts	11
2.2.2	Software Requirements	12
6.3.1	Sentiment Distribution of a Case Study	43