VISVESVARAYA TECHNOLOGICAL UNIVERSITY

'JNANA SANGAMA' BELAGAVI-590 018, KARNATAKA

MINI-PROJECT REPORT

ON

"DISEASE PREDICTION USING MACHINE LEARNING"

**SUBMITTED IN PARTIAL FULFILLMENT OF THE REQUIREMENT
FOR THE VI SEMESTER, BE, MINI-PROJECT-21CSMP67**

**BACHELOR OF ENGINEERING
IN
COMPUTER SCIENCE & ENGINEERING**

Submitted By

1. **OSAMA MIKRANI**
   [1CG22CS079]

2. **YOGESH REBARI**
   [1CG22CS126]

3. **RAGHU GR**
   [1CG22CS090]

4. **PAVAN U**
   [1CG22CS083]

Under the guidance of:

Mr. Harish Athreya , MTech.,
Asst. Prof., Dept. of CSE,
CIT, Gubbi, Tumakuru.

HOD:

Dr. Shantala C P , Ph.D
Prof & Head, Dept. of CSE,
CIT, Gubbi, Tumakuru.

Channabasaveshwara Institute of Technology

(NAAC Accredited & ISO 9001:2015 Certified Institution)
NH 206 (B.H. Road), Gubbi, Tumakuru – 572 216. Karnataka.

(Affiliated to Visvesvaraya Technological University, Belagavi & Recognized by AICTE New Delhi)

2024-25

Channabasaveshwara Institute of Technology
(NAAC Accredited & ISO 9001:2015 Certified Institution)
NH 206 (B.H. Road), Gubbi, Tumakuru – 572 216. Karnataka.

(Affiliated to Visvesvaraya Technological University, Belagavi & Recognized by AICTE New Delhi)

2024-25

DEPARTMENT OF COMPUTER SCIENCE & ENGINEERING

<u>CERTIFICATE</u>

This is to certify that the mini project work entitled **"DISEASE PREDICTION USING MACHINE LEARNING"** has been successfully carried out by **OSAMA MIKRANI [1CG22CS079], PAVAN U 1CG22CS083], RAGHU GR [1CG22CS090], YOGESH REBARI [1CG22CS0126],** bonafide students of **CHANNABASAVESHWARA INSTITUTE OF TECHNOLOGY, GUBBI, TUMAKURU,** under our supervision and guidance and submitted in partial fulfillment for VI Semester BE, Mini-project-21CSMP67 by **Visvesvaraya Technological University, Belagavi** during the academic year of 2024–2025. It is certified that all corrections/suggestions indicated for internal assessment have been incorporated in the report deposited in the departmental library. The mini project report has been approved as it satisfies the academic requirements for the above said degree.

Guide: H.O.D:

Mr. Harish Athreya M.Tech.,   **Dr. Shantala C P**, Ph.D
Asst. Prof., Dept. of CSE,   Prof & Head, Dept. of CSE,
CIT, Gubbi, Tumakuru.   CIT, Gubbi,Tumakuru.

Principal:

Dr. Suresh D S Ph.D
Director & Principal
CIT, Gubbi, Tumakuru.

ABSTRACT

This study explores the development and evaluation of predictive models for disease diagnosis using a symptom-based dataset sourced from Kaggle. The dataset includes 132 binary symptom features and a categorical target variable representing 42 distinct diseases. Four machine learning algorithms were employed to build predictive models: Random Forest, XGBoost, Support Vector Machines (SVM), and Logistic Regression. Each algorithm was trained and tested on separate datasets to ensure robust model performance and generalizability. Random Forest and XGBoost, as ensemble methods, excelled in handling complex feature interactions, while SVM effectively classified diseases by maximizing the decision boundary. Logistic Regression provided an interpretable baseline for comparison. Model evaluation focused on accuracy, precision, and recall, demonstrating the potential of machine learning to enhance disease prediction and streamline diagnostic processes in healthcare.

Table of Contents