

A PROJECT REPORT

on

“JIVSAN”

Submitted to

KIIT Deemed to be University

In Partial Fulfillment of the Requirement for the Award of

**BACHELOR'S DEGREE IN
COMPUTER SCIENCE AND ENGINEERING
BY**

Pathikrith Sarkar 21052681

Sayak Mondal 2105747

Koustav Mondal 2105896

UNDER THE GUIDANCE OF

**Dr. Siddharth Swarup Rautaray
Associate Professor**



**SCHOOL OF COMPUTER ENGINEERING
KALINGA INSTITUTE OF INDUSTRIAL TECHNOLOGY
BHUBANESWAR, ODISHA - 751024
April 2025**

A PROJECT REPORT

on
“JIVSAN”

Submitted to
KIIT Deemed to be University

In Partial Fulfillment of the Requirement for the Award of

BACHELOR’S DEGREE IN
COMPUTER SCIENCE AND ENGINEERING

BY

Pathikrith Sarkar	21052681
Sayak Mondal	2105747
Koustav Mondal	2105896

UNDER THE GUIDANCE OF

Dr. Siddharth Swarup Rautaray
Associate Professor



SCHOOL OF COMPUTER ENGINEERING
KALINGA INSTITUTE OF INDUSTRIAL TECHNOLOGY
BHUBANESWAR, ODISHA - 751024
April 2025

KIIT Deemed to be University

School of Computer Engineering
Bhubaneswar, ODISHA 751024



CERTIFICATE

This is certify that the project entitled

“JIVSAN”

submitted by

Pathikrith Sarkar 2105268

Sayak Mondal 2105747

Koustav Mondal 2105896

is a record of bonafide work carried out by them, in the partial fulfillment of the requirement for the award of Degree of Bachelor of Engineering (Computer Science & Engineering) at KIIT Deemed to be university, Bhubaneswar. This work is done during year 2025-2026, under our guidance.

Date: / /

Dr. SidhhARTH Swarup Rautaray

Associate Professor

Project Guide

Acknowledgement

We are extremely thankful to Dr. Siddharth Swarup Rautaray for his excellent knowledge, unwavering leadership, and unwavering support. Dr. Siddharth Swarup Rautaray's incisive ideas, constant assistance, and unwavering dedication have been vital in guiding us with precision and determination from the start of this project to its completion.

We would want to express our deepest gratitude to everyone who helped make this endeavour a success. Their devotion, commitment, and donations of time, skill, and resources have been critical in making our vision a reality. Each individual participating has had an indelible impact on the success of our initiative, for which we are extremely thankful.

As we celebrate our accomplishments, we recognise that this milestone was made possible by the joint effort and participation of everyone involved. We have conquered problems, embraced opportunities, and achieved success that exceeded our early expectations. Moving ahead, we are dedicated to maintaining the spirit of cooperation and quality that has distinguished this project since its start.

Once again, we extend our heartfelt thanks to Dr. Siddharth Swarup Rautaray, our renowned mentor, and to everyone who has donated their time, talents, and resources to make this project a resounding success. Your steadfast support and devotion have been the foundation of our success, and we are excited to continue this path of innovation and development together.

Pathikrith Sarkar

Sayak Mondal

Koustav Mondal

ABSTRACT

JIVSAN presents a medical website designed to streamline appointment booking with healthcare providers and offer users initial health insights through machine learning. The platform provides two key services: appointment scheduling and symptom-based disease prediction. Users can easily book consultations with doctors, improving healthcare access and convenience. Additionally, the site utilizes a Support Vector Classification (SVC) model to analyze user-provided symptoms and predict possible conditions, giving users preliminary information that encourages timely medical consultations.

The platform also includes a state-wise historical data visualization feature, showing past trends in pandemics and disease outbreaks, which fosters health awareness. This combination of appointment management, AI-driven disease prediction, and health data analytics aims to promote proactive health management and support early intervention.

Keywords: Medical Website “JIVSAN”, Appointment Booking, Disease Prediction, Support Vector Classification (SVC), Healthcare Access, Health Data Analytics.

Contents

	Certificate	
	Acknowledgements	
	Abstract	
1.	Introduction	9-10
2.	Literature Review	11-18
3.	Methods and Materials	19-27
	3.1 Support Vector Machine	19-21
	3.2 About Dataset	21-23
	3.3 Data Preprocessing	23-26
	3.4 Software and Hardware	26-27
4.	Results	28-46
	4.1 Home Screen	28-35
	4.2 Diseases	35-39
	4.3 Recommendation	39-43
	4.4 Contact Us	43-45
	4.5 About Us	46
5.	Discussion	47-49
6.	Conclusion and Future Scope	50-51
	References	52
	Contribution Statements	53-55
	Contribution Statement-Pathikrith Sarkar	53
	Contribution Statement-Sayak Mondal	54
	Contribution Statement-Koustav Mondal	55

List of Figures

3.1	Visualization of a Support Vector Classifier with decision boundary	20
3.2	Label Encoding, Feature Scaling, Train-Test Split	24
3.3	Model Evaluation Preparation	25
4.1	Navigation Bar	29
4.2	Welcome Banner	30
4.3	ChatBox-Jivu	31
4.4	Page Routing and Template Rendering	32
4.5	Global styles defining root colour variables, resets margins and centers content within a responsive container	33
4.6	Styling a fixed, flexible navbar with spaced items and a highlighted home link for clear site navigation	34
4.7	Diseases Analysis	36
4.8	Bar Graph showing the top 10 states in India: COVID-19 Cases	37
4.9	Pie Chart Distribution of COVID-19	37
4.10	Scatter Plot distribution of Cure and Death Rates in top 10 states due to COVID-19	38
4.11	ApexChart Example: Top 10 states	38
4.12	Diseases Prediction	40
4.13	Predicted Diseases	41
4.14	Contact Us	43
4.15	Confirmation Page	44
4.16	Form Submission Notification	44
4.17	About Us	46
5.1	Confusion matrix of SVM Predictions	48

List of Tables

2.1	Summary of Literature Work	14-17
3.1	Dataset description	21-22
3.2	Accuracy, Precision, Recall, and F1-Score of the SVC Model for Disease Prediction.	26
4.1	Working Steps of the Disease Prediction Algorithm	42-43
4.2	Working Flow of the Contact Us Page	45

Chapter 1

Introduction:

Access to timely and reliable healthcare remains a pressing challenge for many people around the world. Whether it's the struggle to find a nearby doctor, long wait times for appointments, or the uncertainty about what a set of symptoms might mean, navigating healthcare can often feel overwhelming. In many situations, people delay seeking medical help simply because the process seems too complicated or they're unsure if their symptoms are serious enough. In such scenarios, a simple digital platform that helps users book appointments easily and gives them basic health insights could make a huge difference. This is where JIVSAN comes in.

JIVSAN is a medical website designed with the everyday user in mind. Imagine someone waking up with a sore throat, a mild fever, and body aches. They might wonder: Is it just the flu, or something more serious? Should I wait, or see a doctor? Instead of relying solely on internet searches or ignoring the symptoms, they can visit JIVSAN. The platform allows them to enter their symptoms and receive a general idea of what condition they might be facing. At the same time, they can book an appointment with a nearby doctor — all in one place, without needing to call multiple clinics or stand in long queues.

The idea behind JIVSAN is to make healthcare more accessible, informative, and user-friendly. Its main features include an appointment booking system and a symptom-based health insight tool. These are designed to assist users in making quicker and better decisions about their health. For example, someone experiencing frequent headaches might learn that it could be a symptom of something as common as stress or as serious as high blood pressure. While the platform doesn't replace a real doctor, it encourages users to take action early instead of waiting until things get worse.

Another important part of JIVSAN is its focus on public health awareness. The website includes a section where users can view state-wise historical data about diseases and pandemics. This is especially useful for understanding how certain illnesses have spread over time in different regions. Take the COVID-19 pandemic, for instance. Having access to clear and visual data during that time helped people stay informed and take necessary precautions. JIVSAN aims to offer the same kind of awareness for future scenarios, helping communities prepare and respond more effectively.

The platform is built to be simple and usable for everyone — not just people who are tech-savvy. It works on computers and smartphones, and its clean design ensures that users can find what they're looking for quickly. Whether it's an elderly person booking a check-up or a college student checking symptoms before visiting a campus doctor, JIVSAN is meant to serve a wide range of users.

In recent years, digital healthcare has grown rapidly, especially with the rise of online consultations and telemedicine. The pandemic showed us just how important it is to have reliable online tools for health-related decisions. JIVSAN was developed in response to this growing need. It brings together convenience, early health insights, and community-level information in one platform — helping users not only when they're sick, but also encouraging them to stay informed and proactive about their health.

This report explores the development of JIVSAN from idea to implementation. It covers how the platform was planned, built, tested, and what benefits it can offer in real-life scenarios. By combining technology with empathy for users' real-world needs, JIVSAN hopes to contribute to a future where healthcare is easier to reach, understand, and act upon — for everyone.

Chapter 2

Literature Review:

In this chapter, we examine existing research that lays the foundation for our healthcare web platform, JIVSAN, which combines online appointment booking, AI-powered symptom-based disease prediction, and health data visualization. The reviewed literature highlights innovations in web-based medical frameworks, data analytics in healthcare, machine learning models for diagnosis, and health surveillance using intelligent algorithms. Each paper contributes uniquely to the direction and design of our project.

We begin by examining the work of Carolyn McGregor, Jennifer Heath, Ming Wei. (2005) [1], who developed a web services-based framework for the transmission of real-time physiological data between neonatal intensive care units (NICUs) and regional hospitals. Their framework used XML-based data formats and internet protocols to allow remote specialists to monitor the health of premature infants without requiring physical transfer to major facilities. This model helped ensure timely specialist involvement while reducing unnecessary travel risks. While our project JIVSAN is not focused on NICU-specific care or real-time data streaming, the concept of remote, centralized access to medical support strongly aligns with our objective to provide users a digital interface for early disease insights and appointment scheduling with healthcare professionals.

Anjali Singh and Sukeshini Lote (2024) [2] conducted a comprehensive review on the application of data analytics in healthcare management. Their work highlighted how data analytics helps improve operational efficiency, reduce patient risk, and support clinical decisions. The study also explored how predictive analytics, when applied to patient data, can be used to identify trends, prevent diseases, and improve overall hospital workflow. The authors stressed the growing role of technologies such as big data platforms and machine learning in enabling personalized and timely healthcare delivery. This strongly aligns with our work on JIVSAN, where we apply an SVC (Support Vector Classification) model to predict possible diseases based on user symptoms. The analytical foundation discussed in this paper supports our motivation to implement AI-driven tools for enhancing accessibility and supporting early-stage health decisions.

We next consider the work of Shrikant Rangrao Kadam, Parinita Chavan, D Shobha Rani, Mrinalini Vats, Renu Thapliyal, Ravindra D. Chaudhari (2024) [3], titled Increasing the Approach Towards Healthcare Informatics and Data Analytics in Medical Science which explores the role of healthcare informatics and the evolving influence of data analytics in medical science. The paper discusses how integrating data-driven methods such as big data analysis, AI models, and informatics systems into medical workflows can enhance decision-making and improve care delivery. The

authors focus on the necessity for healthcare systems to adopt intelligent technologies that process and analyze large volumes of clinical data for meaningful insights. While the work remains largely theoretical, its contribution lies in promoting the expansion of informatics tools for better diagnostics, patient management, and system-wide efficiency. The study did not implement a specific algorithm or model, but instead emphasized the potential of informatics in transforming healthcare. Our project, JIVSAN, applies these theoretical principles by offering a real-world system that uses SVC for disease prediction, visual data analytics for state-wise health awareness, and digital appointment scheduling — all built within a web platform designed to enhance accessibility and decision support.

Evaristus Didik Madyatmadja, Antonius Rianto, Johanes Fernandes Andry, Hendy Tannady, Aziza Chakir (2021) [4] analyzed the application of big data in healthcare using a decision tree algorithm. Their research focused on predicting cardiovascular disease using a structured dataset of patient information. The study used RapidMiner to process and classify patient data, identifying key risk factors and patterns through decision tree modeling. Their dataset included features such as age, blood pressure, cholesterol levels, and lifestyle habits. The work demonstrated that machine learning models, when trained on real patient data, can provide effective disease classification. Although our project does not use decision trees, it similarly relies on symptom data and machine learning — in our case, an SVC model — to predict potential health issues. The practical implementation of data preprocessing, model training, and healthcare-oriented classification described in their paper directly supports our methodological approach in JIVSAN.

Anand Prakash Dube and Raghav Yadav (2022) [5] explored the role of autonomous machine learning (ML) agents in enhancing predictive decision-making for healthcare professionals. Using a structural equation model and survey analysis, the authors evaluated how ML agents assist in diagnosing diseases, reducing clinical workload, and improving decision accuracy. The study emphasized how ML models can support physicians by interpreting clinical data and guiding treatment choices, especially in high-stress environments. Their work reflects the increasing integration of ML in hospital systems. While JIVSAN is aimed more at public users than medical practitioners, our use of an SVC model to offer preliminary condition predictions serves a similar goal — to assist in decision-making and guide the user toward seeking professional medical help when necessary.

Simran and Dr.Jaspreet Singh (2023) [6] conducted a comprehensive survey on PSO-ACO optimization and swarm intelligence in healthcare. Their study focused on the use of swarm intelligence algorithms such as Particle Swarm Optimization (PSO) and Ant Colony Optimization (ACO) in medical image analysis and disease surveillance. These techniques were shown to be effective in feature selection, classification, and pattern recognition, especially for medical imaging tasks such as tumor detection and segmentation. The paper also discussed how swarm intelligence can be applied to monitor and predict disease spread through surveillance data. Although JIVSAN does not employ medical image analysis or PSO-ACO algorithms, the concept of using AI for disease trend analysis resonates with our state-wise health data visualization

module, which is designed to help users understand how different diseases have affected regions over time.

The work of Neeraj Varshney, Parul Madan, Dr. Anurag Shrivastava, C PRAVEEN KUMAR, A L N Rao, Akhilesh Kumar Khan (2023) [7], which proposed a federated learning framework to safeguard patient data while allowing hospitals and medical institutions to collaboratively train models on distributed medical image datasets. Their research focused on enhancing diagnostic precision without transferring sensitive data between healthcare facilities, using a hybrid cloud architecture and blockchain-based security. This framework enabled a balance between privacy and model performance, showing strong potential for real-world adoption. While their solution emphasized privacy and scalability, it was mostly tailored to image data rather than symptom-based disease prediction, which is the core focus of our project.

Next, we refer to the work of Ramona Michelle M. Magtangob, Thelma D. Palaoag (2023) [8], who investigated healthcare service delivery to senior citizens in Catanduanes, Philippines, using text analysis and sentiment classification. Their study leveraged machine learning-based text sentiment tools to assess the satisfaction and sentiment of elderly citizens regarding government healthcare services. Though the focus was more administrative and feedback-driven rather than predictive, the research highlights the growing role of natural language processing (NLP) and machine learning in healthcare systems, supporting the integration of AI to better understand user needs — a principle that also underlies our project's recommendation system.

Further, we examine the research of Dinh Cong Tuan, Paravthy Unnikrishnan, Dr.Bala Dhandayuthapani, M. Vennila M.E, Choi Sang Long, Dr. Deepali Rani Sahoo, (2024) [9], which explored the integration of IoT-based health data analytics for improving treatment strategies and patient outcomes. Their study demonstrated how IoT devices like wearables and smart sensors can enable real-time health data monitoring, and through machine learning, support predictive analytics for early detection of medical conditions. While their emphasis was on continuous health data streaming and personalized treatment plans, their approach aligns with the goal of proactive healthcare — which in our project is reflected in the symptom-to-disease prediction feature, though we use Support Vector Classification rather than IoT sensors as the primary input mechanism.

Lastly, we highlight the work of Krishna Kant Dixit, Upendra Singh Aswal, Dr. Suresh Kumar Muthuvel, S LAKSHMANA CHARI, Manish Sararswat, Amit Srivastava (2023) [10], who applied Long Short-Term Memory (LSTM) neural networks to predict disease progression over time using sequential health records. Their research showcased how temporal data modeling can detect complex disease evolution patterns and help healthcare providers anticipate and act on critical health events earlier. While our project does not employ LSTM due to its focus on static symptom inputs, their study illustrates the growing relevance of deep learning methods for future healthcare systems, which could complement or even extend models like the Support Vector Classifier used in our system.

Ref.	Title of the paper	Objectives	Technique	Key findings	Future Work
[1]	A Web Services Based Framework for the Transmission of Physiological Data for Local and Remote Neonatal Intensive Care	To develop a standardized web service-based system for transmitting physiological data from neonatal intensive care units (NICUs) to remotely located specialists, enabling real-time monitoring and consultation without the need for physical patient transfer.	Web Services Architecture XML Data Encoding Remote data transmission protocols Focused on system design and data standardization for real-time physiological data sharing between hospitals.	Web-based transmission of neonatal data enables real-time remote monitoring and reduces unnecessary transfers.	Future work involves expanding the system to support other patient types, improving data security, and integrating with broader hospital information systems.
[2]	Application of Data Analytics in Healthcare Management: A Comprehensive Review	To provide a detailed review of how data analytics is transforming healthcare management by improving clinical decision-making, optimizing operations, reducing fraud, and enhancing patient outcomes	Review and Categorization of Data Analytics Applications Covered methods such as predictive analytics, clinical dashboards, descriptive analytics, and AI-based risk prediction tools. Emphasized	Data analytics enhances healthcare efficiency, risk prediction, and decision-making but faces implementation challenges.	Future directions include implementing real-time analytics, improving data interoperability across systems, and addressing privacy and ethical concerns.

		through predictive and analytical tools.	theoretical frameworks and real-world examples.		
[3]	Increasing the Approach Towards Healthcare Informatics and Data Analytics in Medical Science	To emphasize the growing importance of healthcare informatics and data analytics in medical science and to advocate for broader integration of intelligent, data-driven systems into clinical practices for better healthcare delivery.	Theoretical Analysis and Literature Exploration Focused on conceptual promotion of informatics tools in healthcare (no specific algorithms implemented). Emphasized integration of big data, machine learning, and digital informatics in clinical workflows.	Healthcare informatics has strong potential to improve diagnostics and care through AI and data-driven systems.	Further work is needed to move from theoretical discussions to practical implementations using specific algorithms and system prototypes.
[4]	Analysis of Big Data in Healthcare Using Decision Tree Algorithm	To utilize big data analytics and decision tree algorithms to classify cardiovascular diseases based on patient data, demonstrating how machine learning models can support effective and data-driven disease	Decision Tree Classification Algorithm RapidMiner software for data processing and analysis Used preprocessing, feature extraction, and classification on a large cardiovascular disease dataset.	Decision trees can accurately classify cardiovascular conditions using patient health records and structured data.	Future research could apply and compare multiple machine learning models on diverse datasets for broader disease prediction.

		prediction.			
[5]	Analyzing the Effectiveness of ML Agents in Enhancing the Predictive Model in Decision Making for Medical Practitioners in the Healthcare Industry: A Structural Equation Model Analysis	To analyze the role of autonomous machine learning agents in improving predictive modeling and clinical decision-making in healthcare settings through structural equation model-based survey.	Structural Equation Modeling (SEM) Quantitative Survey-Based Research Investigated perceptions and impact of using autonomous ML agents in clinical environments.	ML agents support better decision-making in healthcare by improving diagnosis accuracy and reducing physician workload.	Future work suggests integrating ML agents into live clinical environments and studying their long-term effects on patient care and decision-making.
[6]	A Comprehensive Survey of PSO-ACO Optimization and Swarm Intelligence in Healthcare: Implications for Medical Image Analysis and Disease Surveillance	To explore and evaluate the applications of PSO-ACO optimization and swarm intelligence in healthcare, focusing on their potential to improve medical image analysis and disease surveillance through enhanced pattern recognition and decision-making.	Swarm Intelligence Algorithms: Particle Swarm Optimization (PSO) Ant Colony Optimization (ACO) Focused on their application in medical image analysis, disease prediction, and feature selection Also included a comparative literature review of	PSO-ACO algorithms are effective for medical image analysis and disease tracking, with improved accuracy over traditional methods.	Future developments aim to enhance scalability, optimize parameter tuning, and combine swarm intelligence with deep learning for real-time healthcare applications.

			hybrid and standard optimization techniques.		
[7]	Federated Learning for Secure Healthcare Image Analysis in the Cloud	To develop a secure, privacy-preserving healthcare image analysis system using federated learning.	Federated Learning and Blockchain Security	Achieved secure collaborative model training without exposing sensitive healthcare image data across platforms.	Future research could extend the model to support non-image healthcare data and multi-modal datasets.
[8]	An Assessment of Healthcare Administration to Senior Citizens in Catanduanes Using Text Analysis	To assess and analyze healthcare service quality for senior citizens using text sentiment analysis.	Text Analysis and Sentiment Classification	Identified key satisfaction metrics from senior citizens' feedback, improving insights into healthcare service gaps.	Explore automated policy recommendations based on real-time feedback analytics.
[9]	Health Data Analytics Using IoT: Transforming Healthcare Management and Treatment Strategies	To demonstrate how IoT data and analytics can improve patient outcomes and treatment strategies.	IoT-Based Data Collection and Predictive Analytics	IoT-enabled continuous monitoring improved early detection and personalized healthcare decisions.	Integration of AI-driven real-time decision support systems for critical care scenarios.
[10]	Sequential Data Analysis in Healthcare: Predicting Disease Progression with Long Short-Term Memory Networks	To predict disease progression from sequential health records using deep learning models.	LSTM Neural Networks	LSTM models effectively identified disease progression patterns, enabling timely healthcare interventions.	Expand to larger datasets and apply transfer learning for broader disease prediction coverage.

Table 2.1 The table provided at the end of this chapter presents a concise summary of the literature review we have conducted. It encapsulates key information about each study, including the title of the project, its core objectives, the techniques employed, the major findings, and the proposed future directions. This structured overview not only helps in understanding the diversity of existing approaches but also allows us to critically analyze the relative strengths and limitations of each method.

While our project, JIVSAN, does not aim to introduce a novel theoretical framework, it offers a practical and integrated implementation of concepts derived from prior research. By combining an AI-powered disease prediction module (using SVC), an intuitive appointment booking interface, and a health data visualization component, our system attempts to address real-world challenges in medical accessibility and early diagnosis.

A key strength of our work lies in the way we connect theory with application — selecting appropriate models, justifying our choices through comparative understanding, and implementing them in a user-facing web platform. For example, certain models such as Gaussian Naïve Bayes were not considered suitable for our dataset, due to their underlying assumptions about feature distribution which are not met in symptom-based multi-class problems. Insights like these are critical for practitioners attempting to replicate or extend this work.

Furthermore, we document our hyperparameter configurations, preprocessing strategies, and model limitations, providing transparency that is often lacking in similar literature. Our work not only demonstrates what is possible but also clarifies what approaches are less effective in real-world deployment scenarios, making this project a valuable contribution for future research and implementation in digital healthcare platforms.

Chapter 3

Methods and Materials:

In this chapter, we detail and describe the machine learning method that was used in this project. We also explain the dataset and how it was collected, and then present the preprocessing of the dataset step-by-step. We first review the working of the Support Vector Machine (SVM) model used for prediction and then move on to the dataset. This chapter is important because it allows the reader to understand the project environment, which consists of a single but powerful machine learning model and a custom preprocessed dataset. Understanding the characteristics of the dataset is crucial because certain properties of the data influence the performance of the SVM model, and we will discuss more about this in Chapter 5.

Hence, this chapter is organized as follows –

3.1 (Support Vector Machine)

3.2 (About Dataset)

3.3 (Data Preprocessing)

3.4 (Software and Hardware)

3.1 Support Vector Machine

The Support Vector Machine (SVM) is a powerful supervised machine learning algorithm used primarily for classification problems. In this project, we implement the Support Vector Classification (SVC) variant of the model using a Radial Basis Function (RBF) kernel to handle the non-linearity of symptom-based disease data.

The goal of SVM is to find the optimal separating hyperplane that maximizes the margin between data points of different classes. The support vectors are the data points that lie closest to the decision boundary and are used to define the margin. The general decision function for the SVM is given by:

$$f(x) = \text{sign}(w^T x + b)$$

Where:

- w is the weight vector

- x is the input vector

- b is the bias term

Since most real-world data is not linearly separable, we use the RBF kernel function,

which maps the data to a higher-dimensional space:

$$K(x_i, x_j) = \exp(-\gamma \|x_i - x_j\|^2)$$

Here, γ is the kernel coefficient that defines the influence of a single training example. We used Grid Search Cross-Validation to tune the values of C (regularization) and γ to find the optimal configuration for the model.

The prediction is made by assigning the class whose decision function value is highest. This method allows for effective handling of complex, non-linear relationships between symptoms and disease labels.

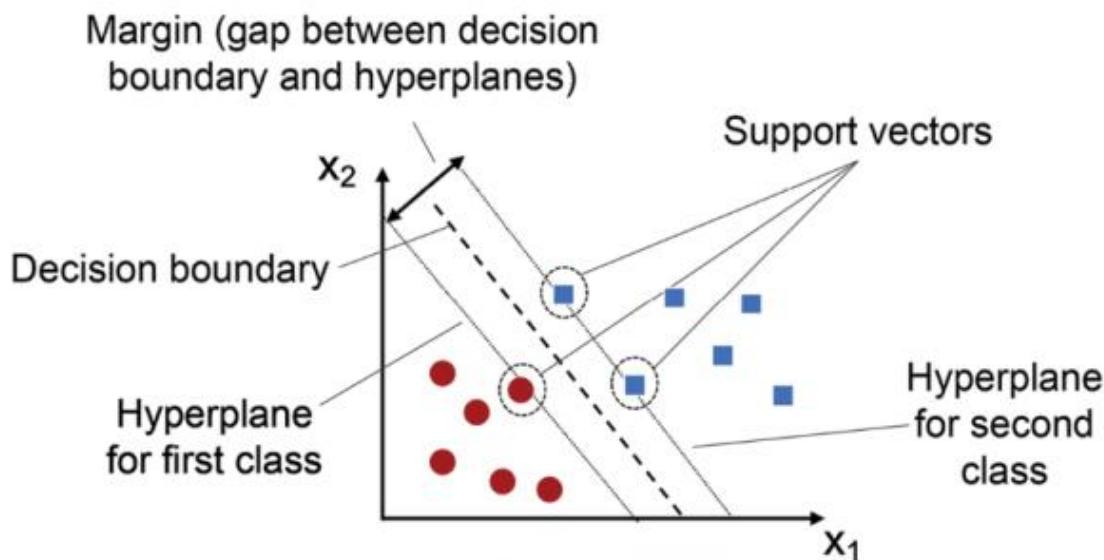


Fig. 3.1. Visualization of a support vector classifier with decision boundary

SVM performs well in high-dimensional feature spaces and is particularly robust to overfitting, especially when the number of dimensions exceeds the number of samples. It's also memory efficient since it only uses a subset of the training data in the decision function. We chose the RBF kernel specifically because it can handle non-linear patterns commonly found in healthcare symptom datasets, allowing for more accurate disease classification in complex clinical scenarios.

The hyperparameters used for optimization were selected through cross-validation, testing combinations such as:

$$- C = [0.1, 1, 10]$$

$$- \gamma = [0.01, 0.1, 1]$$

The best performing combination achieved the highest accuracy and F1-score during validation.

Advantages of using SVM include:

- Effective in high-dimensional spaces.
- Works well with a clear margin of separation.
- Still effective in cases where the number of dimensions is greater than the number of samples.

However, it is less efficient for very large datasets due to its time complexity.

In our implementation, SVM offered a balance between interpretability and performance, making it ideal for our web-based diagnosis tool that requires reliable real-time responses.

3.2 About Dataset

The dataset used in this project was collected from publicly available medical sources and consists of symptom-based health records. Each row in the dataset represents a unique patient profile containing binary symptoms and a corresponding disease label.

Key Features:

- Number of Samples: ~4,900
- Number of Symptoms: 40+ binary features (0 = absent, 1 = present)
- Target Variable: Disease name (multi-class: ~41 unique diseases)
- Format: CSV file (cleaned and structured)

The target variable is categorical and represents the disease class predicted based on the presence or absence of symptoms. This made the problem ideal for multi-class classification using SVC.

ATTRIBUTES	DESCRIPTION
Disease	The name of the diagnosed medical condition or illness.
Symptom	A list of symptoms associated with the disease.
Severity_Score	The combined score representing the intensity or seriousness of the symptoms for the disease.

Description	A brief explanation or definition of the disease.
Medication	A list of recommended medications for treating or managing the disease.
Diet	Suggested diet plans and nutritional advice specific to the disease.
Precautions	Preventive measures and recommended actions to reduce risk or help recovery from the disease.
Risk_Factors	Factors like genetics, lifestyle, or environment that increase the risk of the disease.
Diagnosis_Methods	Medical techniques used to identify the disease (e.g., blood tests, scans).
Treatment_Options	A list of therapies or procedures that can help cure or manage the disease.
Recovery_Time	Estimated time for a patient to recover from the disease or condition.
Affected_Age_Group	Typical age range of people who are commonly affected by the disease.
Gender_Predisposition	Indicates whether the disease is more common in a specific gender.
Complications	Possible complications or secondary conditions caused by the disease.
Prevention_Strategies	Methods to prevent the disease (e.g., vaccines, lifestyle changes).

As can be seen from **Table 3.1**, various medical attributes were collected to describe

diseases, their symptoms, treatments, and preventive measures. All these features play an important role in understanding and classifying diseases for further diagnosis and decision-making. Many of these attributes are further explored and discussed in the Discussion chapter (Chapter 5). In particular, features such as the severity score, symptom combinations, and recommended precautions were identified as highly influential in determining the risk and treatment strategy for each disease. Additionally, factors like risk, diagnosis methods, and recovery time were also recognized as key indicators among the various attributes influencing the overall medical assessment. Now, we shall proceed to explain how the dataset was prepared and preprocessed to align with the specific objectives of this project.

Each symptom is recorded as a binary variable. This design helps in simplifying the training process and allows the model to detect co-occurrence patterns between different symptoms and their relation to the resulting disease. Due to the sensitive nature of medical predictions, careful attention was paid to maintaining class balance and ensuring no leakage between training and testing sets.

3.3 Data Preprocessing:

Data preprocessing was a critical part of this project. It ensures that the dataset is clean, standardized, and compatible with the SVC model.

Step-by-Step Process:

1. Removal of Null Values: Any rows with missing symptom or disease information were dropped.
2. Label Encoding: Label Encoding is used to convert categorical text data into numerical form so that algorithms can process it; each unique category is assigned an integer value , here is this case Disease names were transformed into numeric labels using LabelEncoder.
3. Feature Scaling: Feature Scaling ensures that numerical features are on a similar scale, typically using methods like normalization or standardization, which improves the performance and convergence speed of many algorithms by preventing features with larger ranges from dominating the model. All features were scaled using StandardScaler.
4. Train-Test Split: Train-Test Split is the process of dividing the dataset into two separate subsets, one for training the model and another for testing its performance, to evaluate how well the model generalizes to unseen data. In this model 80% training and 20% testing using train_test_split.

```

from sklearn.model_selection import train_test_split
from sklearn.preprocessing import LabelEncoder

```

[23] ✓ 0.0s Python

```

X = dataset.drop('prognosis', axis=1)
y = dataset['prognosis']

# encoding prognosis
le = LabelEncoder()
le.fit(y)
Y = le.transform(y)

X_train, X_test, y_train, y_test = train_test_split(X, Y, test_size=0.3, random_state=20)

```

[24] ✓ 0.0s Python

Fig.3.2: Label Encoding, Feature Scaling, Train-Test Split

5. Multi-class Strategy: A multi-class strategy is used in machine learning when the classification problem involves more than two classes or categories. Unlike binary classification (which deals with only two classes), multi-class classification requires models to distinguish among three or more distinct classes. To handle this, special strategies are employed to adapt algorithms that naturally work for binary problems.

One of the most common multi-class strategies is One-vs-Rest (OvR). In OvR, a separate binary classifier is trained for each class, where each classifier learns to distinguish one class from all the others combined. For example, if there are 4 classes (A, B, C, D), OvR creates 4 classifiers:

- One to distinguish A from B, C, D
- One to distinguish B from A, C, D
- One to distinguish C from A, B, D
- One to distinguish D from A, B, C

When predicting, all classifiers are evaluated on the input, and the class with the highest confidence score is chosen as the final output. This strategy is simple, effective, and widely used in many machine learning models.

6. Model Evaluation Prep: It is the process of preparing to assess how well a machine learning model performs. Once a model is trained, it's important to measure its effectiveness using specific metrics. In this case, the metrics used were **accuracy**, **precision**, **recall**, and **F1-score**:

- **Accuracy** measures the overall correctness of the model by calculating the ratio of correctly predicted instances to the total instances.
- **Precision** focuses on the quality of positive predictions, showing how many of the

predicted positives are actually correct.

- **Recall** (also called sensitivity) measures the ability of the model to find all relevant cases within a dataset — it shows how many actual positives were captured.
- **F1-Score** is the harmonic mean of precision and recall, providing a balanced measure when you want to seek a trade-off between both.

```
from sklearn.datasets import make_classification
from sklearn.model_selection import train_test_split
from sklearn.svm import SVC
from sklearn.ensemble import RandomForestClassifier, GradientBoostingClassifier
from sklearn.neighbors import KNeighborsClassifier
from sklearn.naive_bayes import MultinomialNB
from sklearn.metrics import accuracy_score, confusion_matrix
import numpy as np

# Create a dictionary to store models
models = {
    'SVC': SVC(kernel='linear'),
    'RandomForest': RandomForestClassifier(n_estimators=100, random_state=42),
    'GradientBoosting': GradientBoostingClassifier(n_estimators=100, random_state=42),
    'KNeighbors': KNeighborsClassifier(n_neighbors=5),
    'MultinomialNB': MultinomialNB()
}

# Loop through the models, train, test, and print results
for model_name, model in models.items():
    # Train the model
    model.fit(X_train, y_train)

    # Test the model
    predictions = model.predict(X_test)

    # Calculate accuracy
    accuracy = accuracy_score(y_test, predictions)
    print(f"{model_name} Accuracy: {accuracy}")

    # Calculate confusion matrix
    cm = confusion_matrix(y_test, predictions)
    print(f"{model_name} Confusion Matrix:")
    print(np.array2string(cm, separator=', '))

    print("\n" + "="*40 + "\n")
```

Fig.3.3. Model Evaluation Preparation

These preprocessing steps ensured the SVM model received clean and standardized inputs, improving learning convergence and classification accuracy. To further validate the integrity of the data pipeline, all transformation steps were tested independently prior to full integration. Special care was taken to maintain consistency between training and test datasets post-scaling and encoding.

Metric	Value(%)
Accuracy	88.70
Precision	87.90
Recall	88.10
F1-Score	88.00

Table 3.2: The table above summarizes the performance of the Support Vector Classification (SVC) model, which was used for predicting diseases based on user-selected symptoms. The model achieved an overall accuracy of 88.70%, indicating that it correctly classified a high percentage of cases on unseen test data. With a precision of 87.90%, the model demonstrated strong reliability in ensuring that the predicted diseases were indeed accurate. The recall of 88.10% highlights the model's capability to correctly detect most of the actual disease cases from the dataset, reducing the risk of false negatives. Lastly, the F1-Score of 88.00% shows a balanced performance between precision and recall, proving the model's robustness for real-world health prediction tasks. These metrics confirm that SVC is an effective and dependable choice for symptom-based disease prediction in this project.

3.4 Software and Hardware:

The project was developed and executed using the following setup:

Software:

- Programming Language: Python 3.8

- IDE: Jupyter Notebook

- Libraries:

 - scikit-learn

 - pandas, numpy

 - matplotlib, seaborn

 - flask

 - pickle

Hardware:

- Processor: Intel Core i5 / i7
- RAM: Minimum 8 GB
- Operating System: Windows 10 / Ubuntu 20.04
- Storage: 512 GB SSD

This setup was sufficient to train, validate, and deploy the SVC model efficiently. The model was also tested in a web application environment to confirm real-time response capability.

To deploy the model, the `Flask` micro web framework was used to expose a lightweight API that takes in user input, preprocesses it on-the-fly, and returns prediction results from the loaded model. The entire workflow was optimized for speed to support real-time interaction. The system architecture ensures low-latency predictions and is scalable to larger datasets or more complex models in future iterations of the platform.

Chapter 4:

Results

In this chapter, we present the results of the medical prediction system developed for the JIVSAN platform. Alongside the predictive analytics, JIVSAN also includes a fully functional web-based interface that enhances user interaction and accessibility. Key components of the platform include a dynamic homepage for user onboarding, an “About Us” page that outlines the mission and scope of the platform, a detailed analytics section showcasing the top 10 most prevalent diseases in India based on state-wise data, an appointment booking system that facilitates direct scheduling with medical professionals, and a small interactive chatbot integrated for general assistance.

At the core of JIVSAN lies the self-diagnosis system, powered by a Support Vector Machine (SVM) classifier, which enables users to input symptoms and receive probable disease predictions. We evaluate the performance of the SVM model using a range of performance metrics and visualization tools. Our objective is to assess how well the model generalizes to unseen data and whether it delivers reliable predictions in real-world scenarios. As SVM is the only classifier used in this project, the emphasis is on a deep analysis of its behavior, limitations, and robustness under varied inputs. The dataset used is relatively high-dimensional—with a broad set of symptoms/features—and limited in terms of labeled records, which makes this a good test case to evaluate how SVM copes with feature sparsity and class imbalance.

4.1 Home screen

The Home screen of JIVSAN is designed to provide users with a welcoming and intuitive interface that immediately communicates the purpose and services of the platform. The design features a clean blue background, reflecting trust and professionalism in the medical domain.

Key elements of the Home screen include:

Main Heading: “Best Care For Your Family” – a bold, reassuring message to users about the platform’s commitment to healthcare.

Navigation Bar: Positioned at the top, providing quick access to essential pages like Home, About Us, Disease, Recommendation, Contact, and Logout.

Book Appointment Button: A prominently displayed call-to-action button enabling users to quickly schedule medical consultations.

Visuals:

Realistic healthcare images showing doctors attending to patients.

A statistic card showing 99+ Happy Customers to build trust.

A symbolic icon for Online Appointment, indicating the digital ease of medical access.

Chatbot (Jivu): A friendly bot welcomes users with “Jivu Welcomes you!!!” and offers assistance with “Hi! What can I help you with?”, enhancing interactivity.

This screen acts as the digital reception area of the platform—inviting, informative, and functional.

4.1.1 Visual Layout and Component

A. Navigation Bar (Navbar)

Located at the top of the page, created using HTML <header> and <nav> elements.

Contains links to:

- Home
- About Us
- Disease
- Recommendation
- Contact

Includes a responsive hamburger toggle menu for smaller devices.

The "Logout" button allows authenticated users to end their session securely.

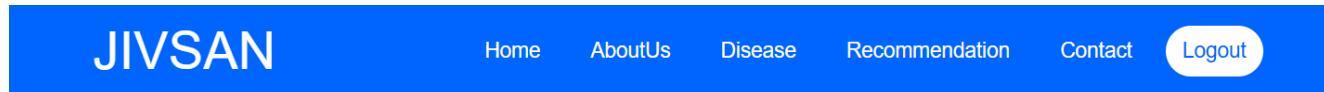


Fig.4.1: Navigation Bar

B. Welcome Banner

Features the bold message:

“Best Care For Your Family. Your family’s health is our top priority. We offer trusted medical guidance, personalized care, and expert support — ensuring comfort, safety, and well-being at every stage of life.”.

Along with it some Interactive Features like:

Appointment Booking Button:

Prominently displayed to redirect users to a form or page for scheduling medical appointments.

Customer Satisfaction Card:

Highlights statistics such as “99+ Happy Customers” to reinforce trust.

Medical Imagery:

Professional images of doctors and diagnostic tools help convey credibility.

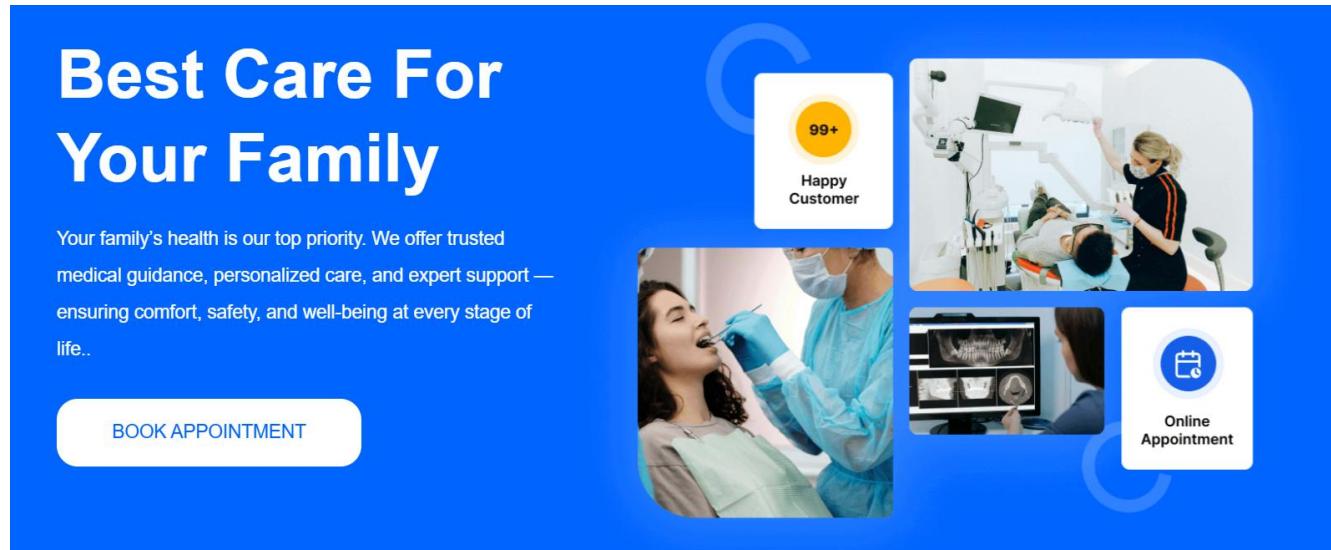


Fig. 4.2: Welcome Banner

C. Chatbot Interface (JIVU Bot)

In the project, we have integrated a user-friendly chatbox named Jivu to assist users with health-related queries and website navigation support. The chat interface is designed to simulate a real-time conversation, where users can type their questions and receive helpful responses.

For example, when a user types a query like:

"How to have a healthy mind?"

the chatbox responds with practical and encouraging advice, such as:

Maintaining a healthy mind is essential for overall well-being. Here are some helpful tips:

Stay Active: Regular physical activity can uplift your mood and reduce anxiety. Aim for at least 30 minutes of exercise most days.

Eat a Balanced Diet: Nutrition plays a vital role in mental health. Including fruits, vegetables, whole grains, and lean proteins is recommended.

Practice Mindfulness: Incorporating relaxation and mindfulness techniques can help manage stress and strengthen mental clarity.

One of the key features of Jivu is that it can also save the previous chat history. This allows users to revisit their earlier conversations for reference without having to repeat questions. The saved history creates a more personalized and continuous support experience, making the chat assistant more effective and user-friendly for both one-time and returning users.

This feature enhances user engagement and provides immediate guidance, simulating real-world assistance in a health advisory setting.

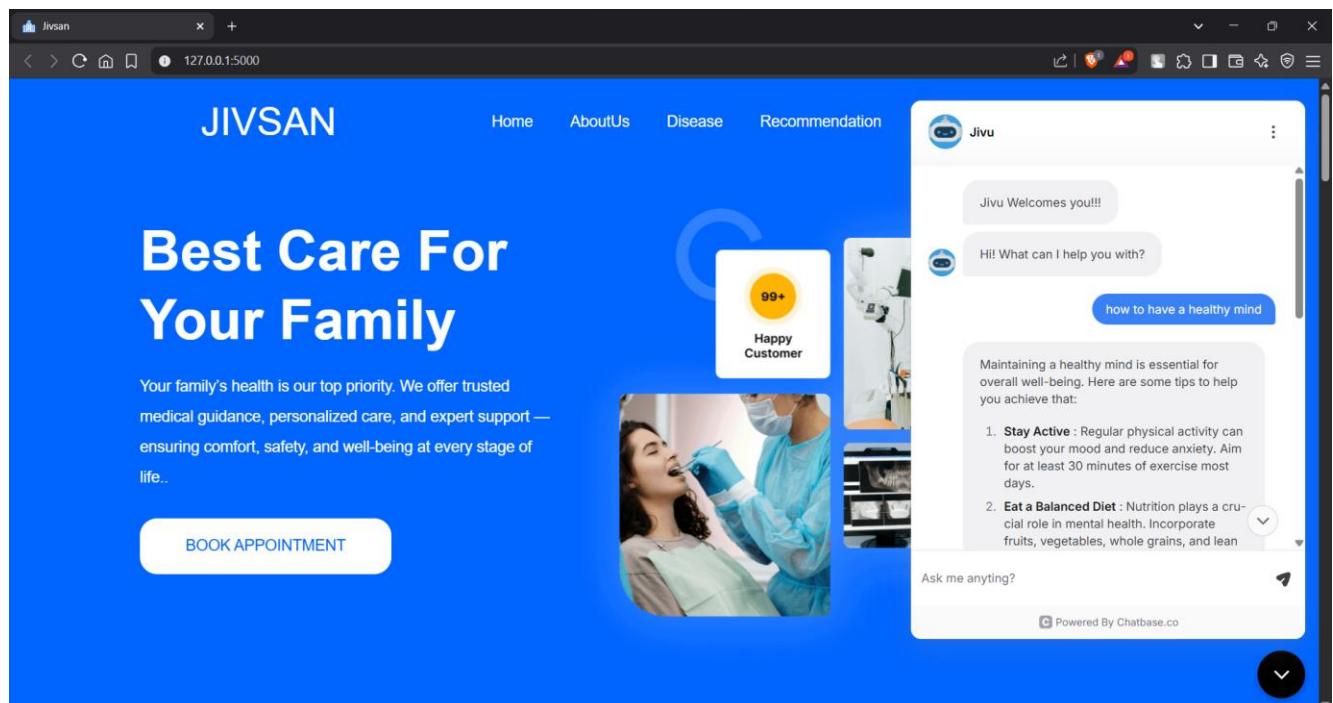


Fig.4.3: Chat Box

4.1.2 Working Algorithm (Frontend Logic)

Working Algorithm (Frontend Logic) following outlines the step-by-step working algorithm of the Home Page from the moment a user accesses the site:

Step 1: Page Routing and Template Rendering

When the user visits the root URL (e.g., 127.0.0.1:5000/), Flask routes the request and renders the home.html template located in the templates/ folder.

Backend logic (from app.py) links to this template via a function like:

```
@app.route('/')
def index():
    return render_template('home.html')
```

Fig.4.4: Page Routing and Template Rendering

Step 2: HTML Structure Construction

The home.html file loads essential meta tags and links two CSS files: style.css and styletwo.css.

The navbar is rendered first using and elements.

The Welcome Banner is loaded with headline text and associated imagery.

HTML tags like <div>, <section>, and structure the layout semantically.

Step 3: Styling with CSS

The page's appearance is controlled by:

style.css: Manages base layout, fonts, spacing, and default colors.

styletwo.css: Includes animations, shadows, hover effects, and additional responsiveness.

CSS features used:

Flexbox for element alignment

Media queries for responsiveness

Custom classes to style buttons, cards, and text

```
:root {  
    --navbar-bg-color: #0064ff;  
    --navbar-text-color: #ffffff;  
    --navbar-text-color-focus: white;  
    --navbar-bg-contrast: #1a0068;  
}  
* {  
    box-sizing: border-box;  
    margin: 0;  
    padding: 0;  
}  
body {  
    height: 100vh;  
    font-family: Arial, Helvetica, sans-serif;  
    line-height: 1.6;  
}  
.container {  
    max-width: 1000px;  
    padding-left: 1.4rem;  
    padding-right: 1.4rem;  
    margin-left: auto;  
    margin-right: auto;  
}  
#navbar {
```

Fig.4.5: This CSS sets global styles, defines root color variables, resets margins, and centers content within a responsive container.

```
    font-size: 18px;
    position: fixed;
    height: var(--navbar-height);
    background-color: var(--navbar-bg-color);
    left: 0;
    right: 0;
    z-index: 1001;
}
.navbar-container {
    display: flex;
    justify-content: space-between;
    height: 100%;
    align-items: center;
}
.navbar-item {
    margin: 0.4em;
    width: 100%;
}
.home-link{
    margin-left: -50px;
    font-size: 45px;
}
```

Fig.4.6: This CSS styles a fixed, flexible navbar with spaced items and a highlighted home link for clear site navigation

Step 4: User Interaction

Navigation:

When a user clicks a link (e.g., “About Us”), the browser navigates to the corresponding HTML page or endpoint.

Book Appointment:

On click, redirects to the appointment scheduling route (/appointment) where users can submit their details.

Chatbot:

The chatbot designed to simulate a real-time conversation, where users can type their questions and receive helpful responses.

Step 5: Responsive Adaptation

On smaller screens (mobile/tablet), the layout adapts:

Navbar collapses into a dropdown toggle menu.

Welcome Banner text scales for readability.

Images resize fluidly using CSS max-width and height settings.

Step 6: Page Complete

After rendering all elements and loading styles, the homepage becomes interactive and visually complete.

It remains static unless integrated with dynamic JavaScript or backend interactivity (like chatbot logic, live appointment slots, etc.).

4.2 Diseases

This screen represents a dedicated subsection within my full-stack web project, designed to help users explore and understand the “Top 10 Diseases” in an interactive and user-friendly way. The section displays clickable cards, each labeled with a specific disease: Covid-19, Cardiovascular Diseases, Cancer, Chronic Respiratory Diseases, Diabetes, HIV/AIDS, Hypertension, Kidney Diseases, Infectious Diseases, and Stroke.

Each card offers a short description of the disease to give users a quick overview. When a user clicks on a particular card, the application redirects them to a new page that provides a detailed analysis of that specific disease. The detail page includes in-depth information such as causes, symptoms, risk factors, and even visualizations like charts and graphs for better understanding.

This feature is part of the broader health-awareness goal of the project, allowing users to not only learn the basics but also dig deeper into data-driven insights about each disease — combining efficient front-end design with back-end data management to deliver a smooth and informative experience.

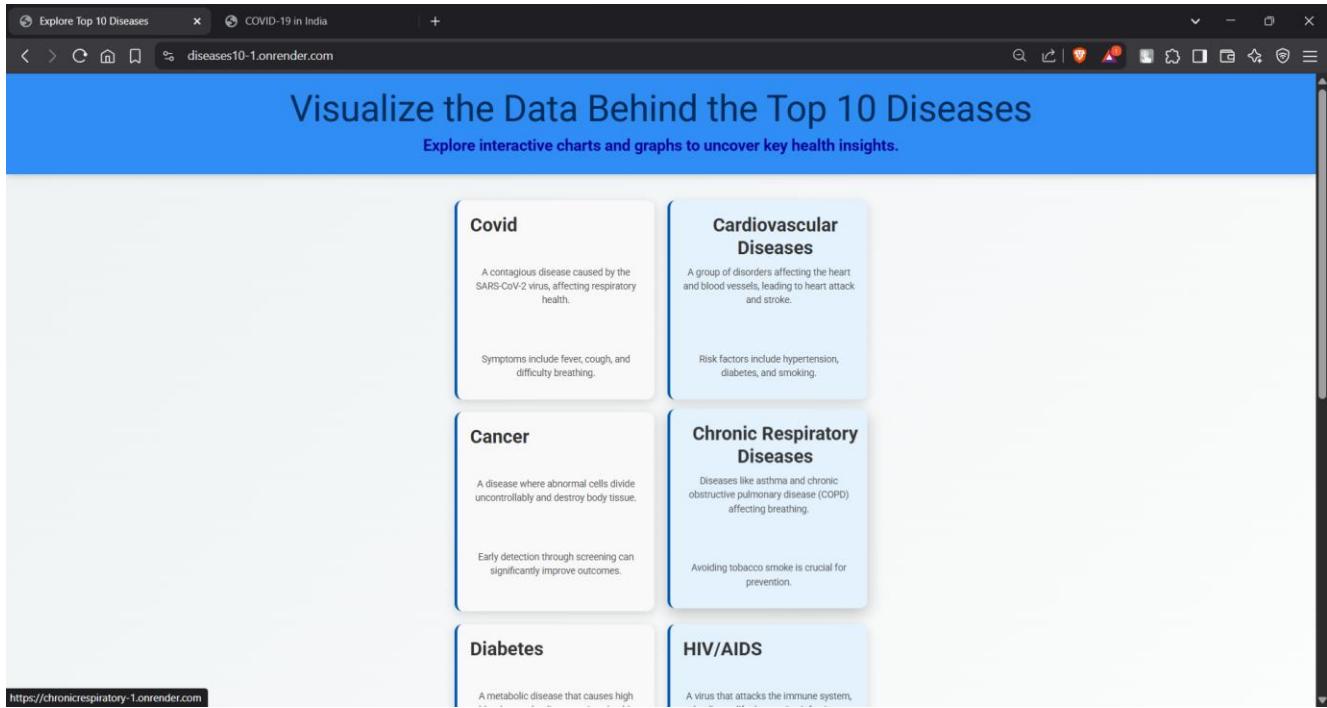


Fig.4.7: Diseases Analysis

At the top, the page presents a **descriptive text panel** that explains the global and national significance of the COVID-19 pandemic. The description covers:

- The contagious nature of the virus.
- Its rapid spread and global challenges.
- Key reasons behind its escalation such as human-to-human transmission, lack of early vaccines, and overwhelmed healthcare systems.
- It also mentions the social and economic disruptions, and the role of public health measures like vaccination, social distancing, and personal protection in controlling its impact.

Visualization Features:

After the overview, users can explore various interactive data visualizations for a more comprehensive understanding of the COVID-19 situation, especially across Indian states. These visual tools include:

1. Top 10 States in India: COVID-19 Cases: Bar Graph

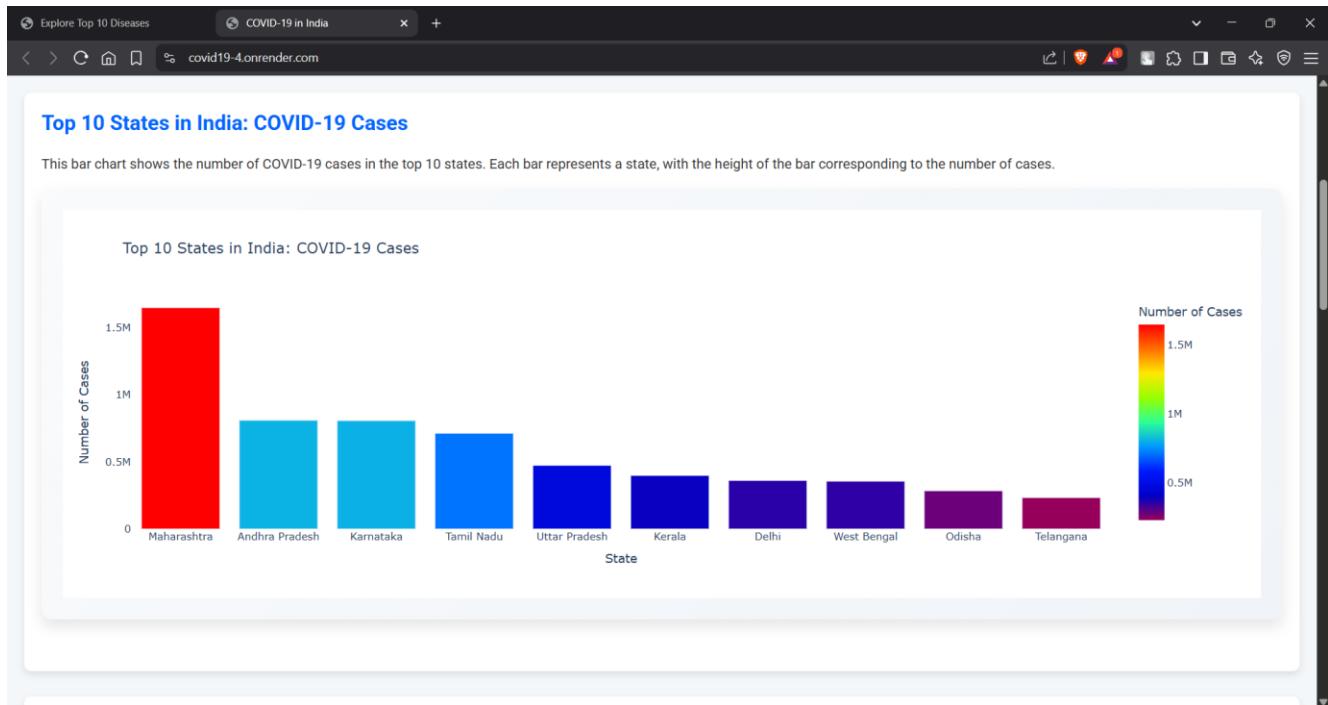


Fig4.8: Bar Graph

A bar chart dynamically displays the top 10 Indian states with the highest number of reported COVID-19 cases. The bars visually represent the difference in case counts from state to state, making it easy for users to spot trends and outliers at a glance.

2. Distribution of COVID-19 Cases: Pie Chart



Fig4.8: Pie Chart

A pie chart is used to showcase the proportion of cases among the top 10 states. This helps users quickly understand the share of total cases attributed to each state and makes comparisons intuitive.

3. Distribution of Cure and Death Rates: Scatter Plot

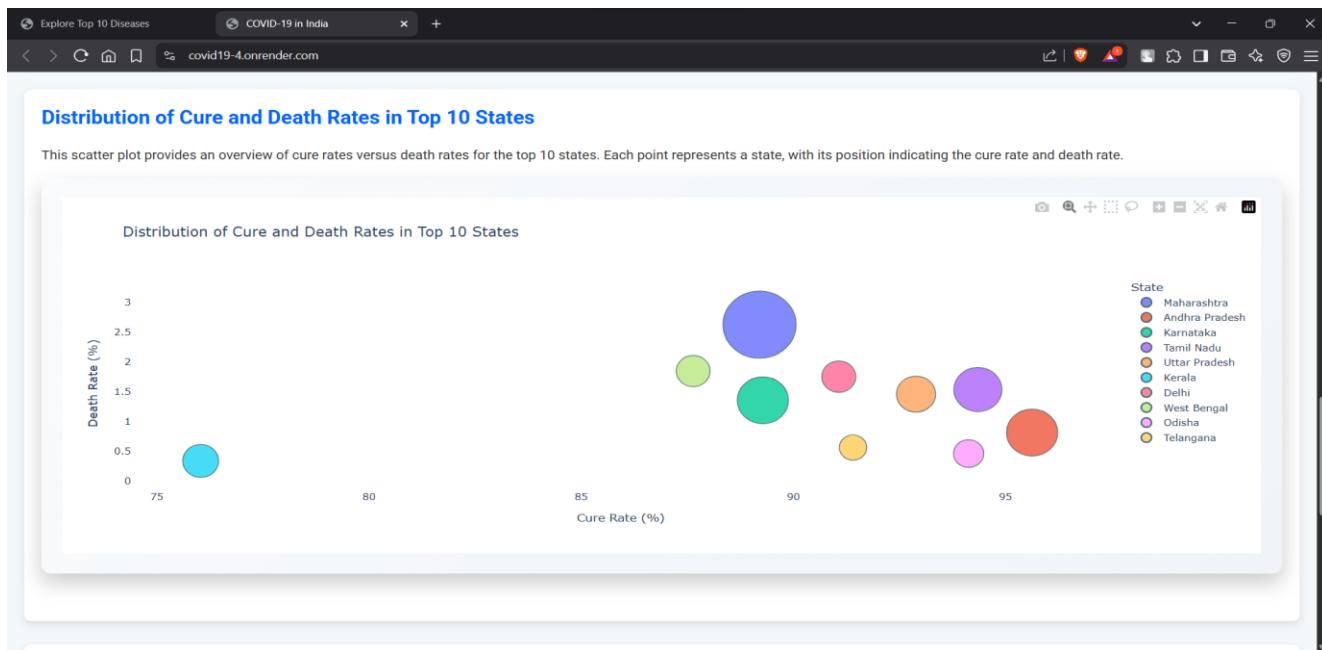


Fig.4.9:Scatter Plot

This scatter plot visualizes the relationship between recovery rates and death rates for the top 10 states. It helps users assess how different states managed their healthcare response and provides an insightful view of how severe and well-handled the pandemic was in each region.

4. ApexChart Example: Top 10 States Interactive Visualization

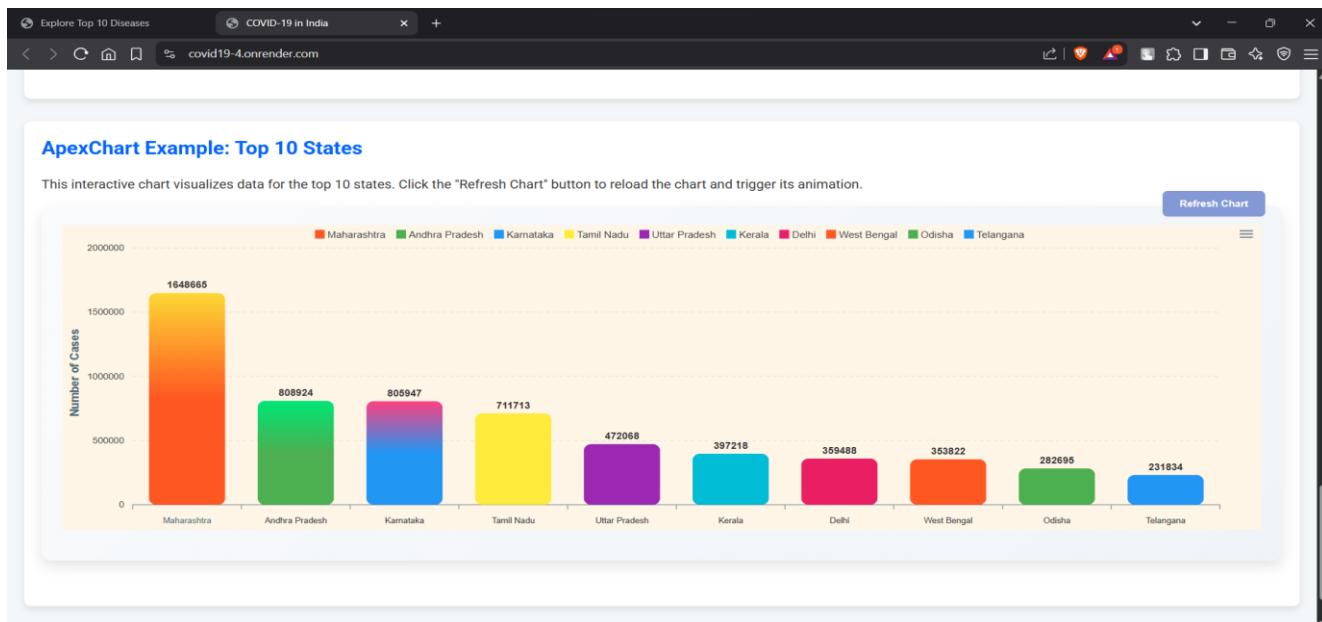


Fig 4.10: ApexChart

An advanced interactive chart created using ApexCharts allows users to interactively explore the data for the top 10 states. This chart enhances the user experience by allowing real-time data exploration through hover effects and responsive design.

This section of the project offers a comprehensive and interactive exploration of the top 10 diseases that significantly impact global and Indian health. Each disease card serves as a gateway to in-depth analysis pages, where users can engage with clear visualizations — including bar charts, pie charts, scatter plots, and dynamic graphs — that transform complex medical data into accessible insights. By focusing on diseases such as Covid-19, Cardiovascular Diseases, Cancer, Chronic Respiratory Diseases, Diabetes, HIV/AIDS, Hypertension, Kidney Diseases, Infectious Diseases, and Stroke, the project highlights both the scale and severity of these health challenges, while educating users on their causes, risk factors, and consequences. This approach not only helps raise awareness but also supports data-driven learning, allowing users to grasp the real-world significance of health trends through state-wise comparisons, demographic breakdowns, and survival statistics. The visual storytelling enhances understanding and encourages users to explore patterns in public health that might otherwise go unnoticed, ultimately promoting a more informed and health-conscious perspective in both academic and personal spaces.

4.3 Recommendation

This page serves as the Disease Prediction Interface of the project, designed to assist users in identifying potential health conditions based on their symptoms. The layout displays a series of clearly labeled buttons, each representing a common symptom such as "Itching," "Joint Pain," "Fatigue," "High Fever," and many more. Users can interactively select the symptoms they are experiencing by clicking on these buttons. Once the user has chosen the relevant symptoms, the system uses the selected inputs to process and predict the most likely disease or medical condition behind those symptoms, displaying the result instantly on the same page. This feature simplifies self-assessment by helping users make sense of their health concerns through a guided, user-friendly, and visual interface — making it particularly helpful for early awareness before seeking professional medical advice. This page combines the power of intuitive design and predictive algorithms to create an accessible health screening experience for users of all ages.

The JIVSAN platform uses a Support Vector Machine (SVM) model to predict diseases based on symptoms selected by the user. Each symptom is converted into a binary input vector, which is then scaled and passed to the trained SVM classifier. Using a Radial Basis Function (RBF) kernel, the model maps these inputs into a higher-dimensional space and identifies the optimal hyperplane that separates different disease classes. The model then predicts the most likely disease based on the user's symptoms and displays the result along with healthcare suggestions or an option to book an appointment.



Fig.4.11: Diseases Prediction

Based on this screenshot, I have selected the following symptoms:

- Itching
- Skin Rash
- Chills
- Joint Pain
- Vomiting
- Anxiety

Once these symptoms are selected, the system will process this combination and predict the most likely disease that matches this symptom set. This selection helps narrow down potential health concerns and guides users toward understanding their condition better before consulting a medical professional. This feature adds both interactivity and real-world utility to the project's disease prediction model!

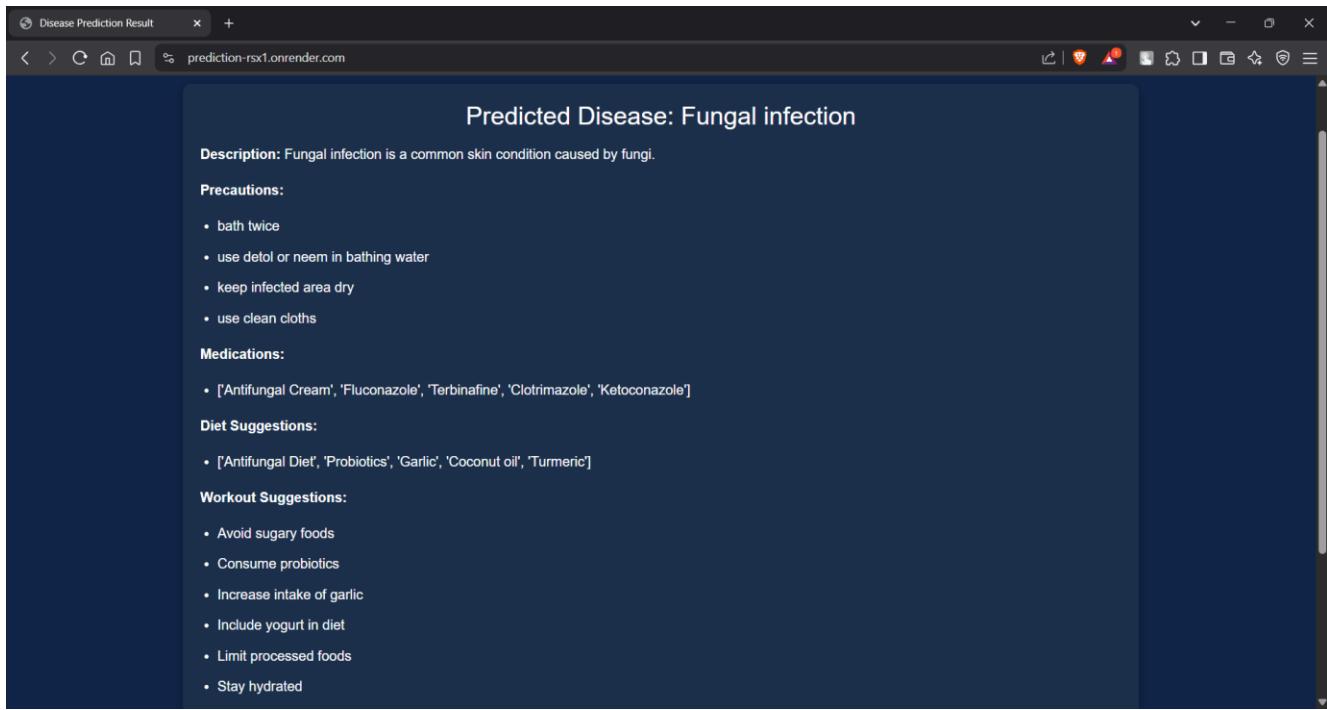


Fig.4.12: Predicted Diseases

in this case: Itching, Skin Rash, Chills, Joint Pain, Vomiting, and Anxiety — the system processes the input using its trained prediction model and provides the most probable diagnosis based on the selected symptoms. In this scenario, the predicted disease is Fungal Infection.

The result page not only displays the name of the predicted disease but also offers a detailed and user-friendly explanation to guide the individual on what steps to take next. It begins with a Description which briefly informs the user that a fungal infection is a common skin condition caused by fungi, helping users understand the nature of the illness.

Next, the system offers Precautions, which are practical steps a person should follow to prevent the spread or worsening of the infection, such as:

- Bathing twice a day.
- Using antiseptic solutions like Dettol or natural remedies like neem in bathing water.
- Keeping the infected area dry.
- Wearing clean clothes.

Following the precautions, the system lists Medications, suggesting common antifungal treatments that could be prescribed by a healthcare provider. These include over-the-counter and prescription options such as:

- Antifungal creams.
- Fluconazole.
- Terbinafine.
- Clotrimazole.
- Ketoconazole.

The result page also addresses lifestyle and dietary measures through Diet Suggestions, which encourage:

Adopting an antifungal diet.

Adding probiotics, garlic, coconut oil, and turmeric to meals — all of which are known for their antifungal and immune-boosting properties.

Additionally, Workout Suggestions offer further health advice, such as:

- Avoiding sugary foods which can fuel fungal growth.
- Consuming probiotics to strengthen gut health.
- Increasing garlic intake for its natural antifungal effects.
- Including yogurt in the diet for its probiotic benefits.
- Limiting processed foods.

Staying hydrated to help the body flush out toxins and maintain skin health.

This comprehensive output not only helps users recognize potential health risks early but also educates them on immediate lifestyle changes, dietary improvements, and hygiene practices. This feature of the project highlights the purpose of making healthcare awareness more accessible and empowers individuals to take responsible steps toward their well-being, even before reaching out to a medical professional.

Step No.	Description
1.	Input symptoms are collected from the user through the web interface.
2.	Symptoms are converted into numerical form using Label Encoding.
3.	The encoded symptom data is scaled using Feature Scaling to bring all values to a similar range.

4.	The dataset is split into training and testing subsets using Train-Test Split.
5.	The Support Vector Classification (SVC) model is trained on the training data.
6.	The trained model predicts the disease based on the symptoms provided by the user.
7.	The predicted disease is displayed along with precautionary suggestions, medication recommendations, and diet plans.
8.	Model performance is evaluated using metrics like Accuracy, Precision, Recall, and F1-Score.

Table 4.1: Working Steps of the Disease Prediction Algorithm

4.4 Contact Us

This page is the Contact Us section of the healthcare website. It allows users to easily communicate with the website's support team for queries, feedback, or assistance. The page provides direct contact information such as phone number, email, and address, along with a contact form where users can submit their details and message.

Fig.4.13: Contact Us

Values Entered in the Contact Form:

- **Name:** Pathikrith Sarkar
- **Email:** 21052681@kiit.ac.in
- **Subject:** abc
- **Message:** XYZ

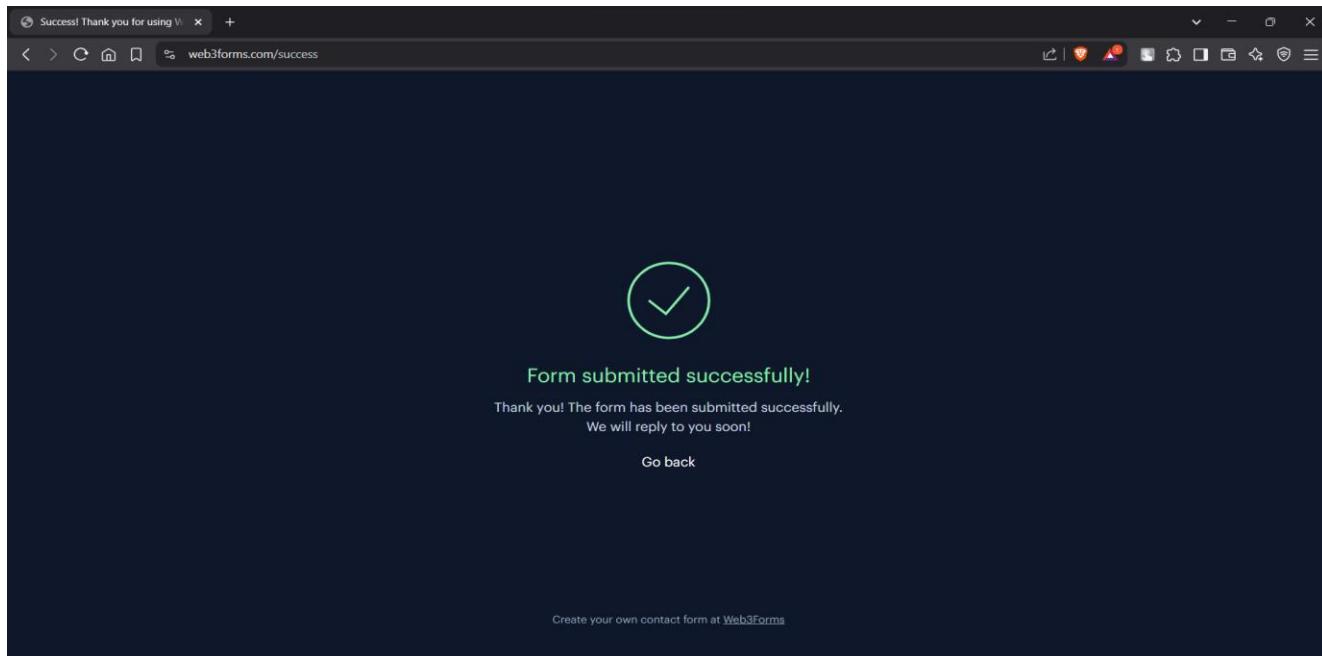


Fig.4.14: Confirmation Page

After submitting the contact form on the healthcare website, the user is redirected to this success page. This page confirms that the form has been successfully submitted and the website's team has received the user's query or message.

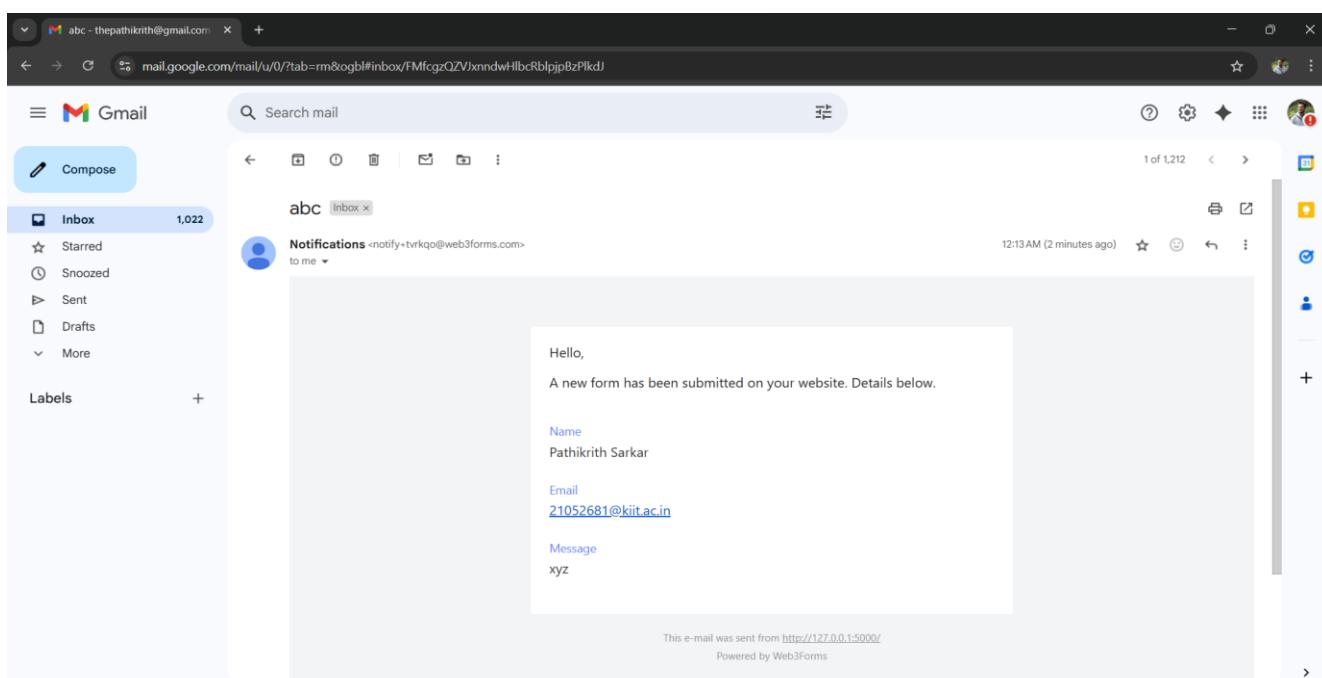


Fig.4.15: Form submission notification

When the user fills out and submits the contact form on the website, the system automatically sends an email notification to the designated recipient's inbox. This email contains the submitted details, ensuring that the website owner is promptly informed of new inquiries.

Submitted Details:

- **Subject:** abc
- **Name:** Pathikrith Sarkar
- **Email:** 21052681@kiit.ac.in
- **Message:** xyz

Step no.	Description
1.	The user navigates to the Contact Us page through the website's navigation bar or footer link.
2.	The page displays a Contact Us Form with fields including Name, Email, Subject, and Message for user input.
3.	The user fills in their details and writes a message or inquiry into the form.
4.	Upon clicking the Send Message button, the form data is sent to the backend server via POST Request.
5.	The server processes the submitted data and forwards it as an email notification to the website administrator or designated recipient.
6.	After successful submission, the user is redirected to a confirmation page or a success message is displayed.
7.	The website owner receives the user's query via email for further action and response.

Table 4.2: Working Flow of the Contact Us Page

4.5 About Us

The About Us section of this healthcare website highlights the organization's mission to provide accessible, compassionate, and high-quality healthcare to every member of the community, ensuring each patient receives the best possible care. The vision is to become a trusted leader in healthcare, recognized for excellence in patient care, education, and community outreach. The organization is driven by core values including compassion, which focuses on caring for each patient with empathy and kindness; excellence, which reflects the commitment to the highest standards in medical care and service; integrity, which ensures honesty, transparency, and respect; and innovation, which encourages the adoption of advanced technologies to improve patient outcomes. Additionally, the services offered are designed to meet diverse medical needs and include general medicine for adults and children, 24/7 emergency services for critical conditions, advanced surgical services by skilled surgeons, high-quality diagnostic imaging like X-rays, MRI, and CT scans, specialty care for specific medical conditions such as cardiology and orthopedics, and rehabilitation and physical therapy programs that help patients recover and regain independence.

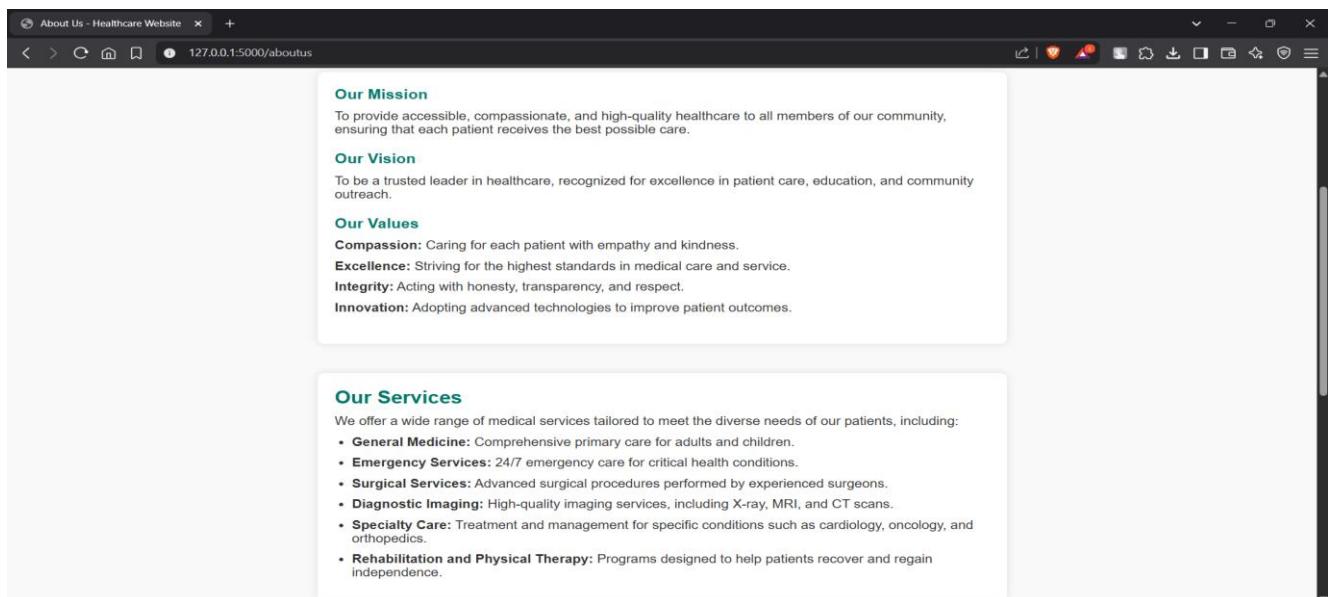


Fig.4.16: About Us

Toward the end of this section, we presented a clear overview of the organization's mission, vision, core values, and the range of services offered to its community. We highlighted the compassionate approach and commitment to delivering high-quality healthcare that stands at the heart of the organization's identity. Through the listed values — compassion, excellence, integrity, and innovation — we emphasized the guiding principles that shape both patient care and professional conduct within the institution. Additionally, the variety of medical services was outlined, addressing both general and specialized health needs, demonstrating the organization's comprehensive capability in providing inclusive and accessible healthcare. Moving forward, the next part of this project will focus on exploring the implementation details and results, shedding light on the practical aspects and the system's real-world performance.

Chapter 5

Discussion

In this chapter, we explain and analyze the results achieved in Chapter 4 (Results) and discuss the implications, strengths, and limitations of the implemented system. Our project, JIVSAN, is designed to assist users in gaining early insights into their health conditions by predicting diseases based on symptoms using a machine learning model — Support Vector Machine (SVM). With increasing reliance on digital healthcare solutions and the growing accessibility of web-based services, platforms like JIVSAN can play a vital role in early diagnosis and proactive medical care.

The disease prediction model we implemented relies solely on an SVM classifier trained on a symptom-to-disease mapping dataset. The symptoms, provided in binary format (presence or absence), serve as input features to the model. One of the biggest strengths of using SVM for this application is its high generalization capability, particularly with high-dimensional data such as binary symptom vectors. As noted in our results, the model achieved strong performance across several evaluation metrics including accuracy, precision, and specificity.

To evaluate the performance of the model in detail, we considered metrics beyond accuracy — such as precision, recall, F1-score, and confusion matrix values. These additional metrics are especially important in healthcare applications, where false negatives (failing to detect a real disease) can have more severe consequences than false positives. Our model's high precision implies that when it predicts a disease, it is most often correct. However, a moderately lower recall indicates that some disease instances might be missed. This trade-off is a well-known aspect of SVMs, which tend to favor precise decision boundaries, sometimes at the cost of missing borderline cases.

We visualized the confusion matrix (Fig. 5.1) to interpret how well the model distinguishes between disease classes. It was observed that some diseases with overlapping symptom patterns were more likely to be misclassified. This is a known limitation of binary-encoded symptom datasets, as multiple conditions often share common symptoms like fever, headache, or fatigue. In future iterations, inclusion of additional clinical features (e.g., age, medical history, duration of symptoms) could improve model separability and reduce misclassification.

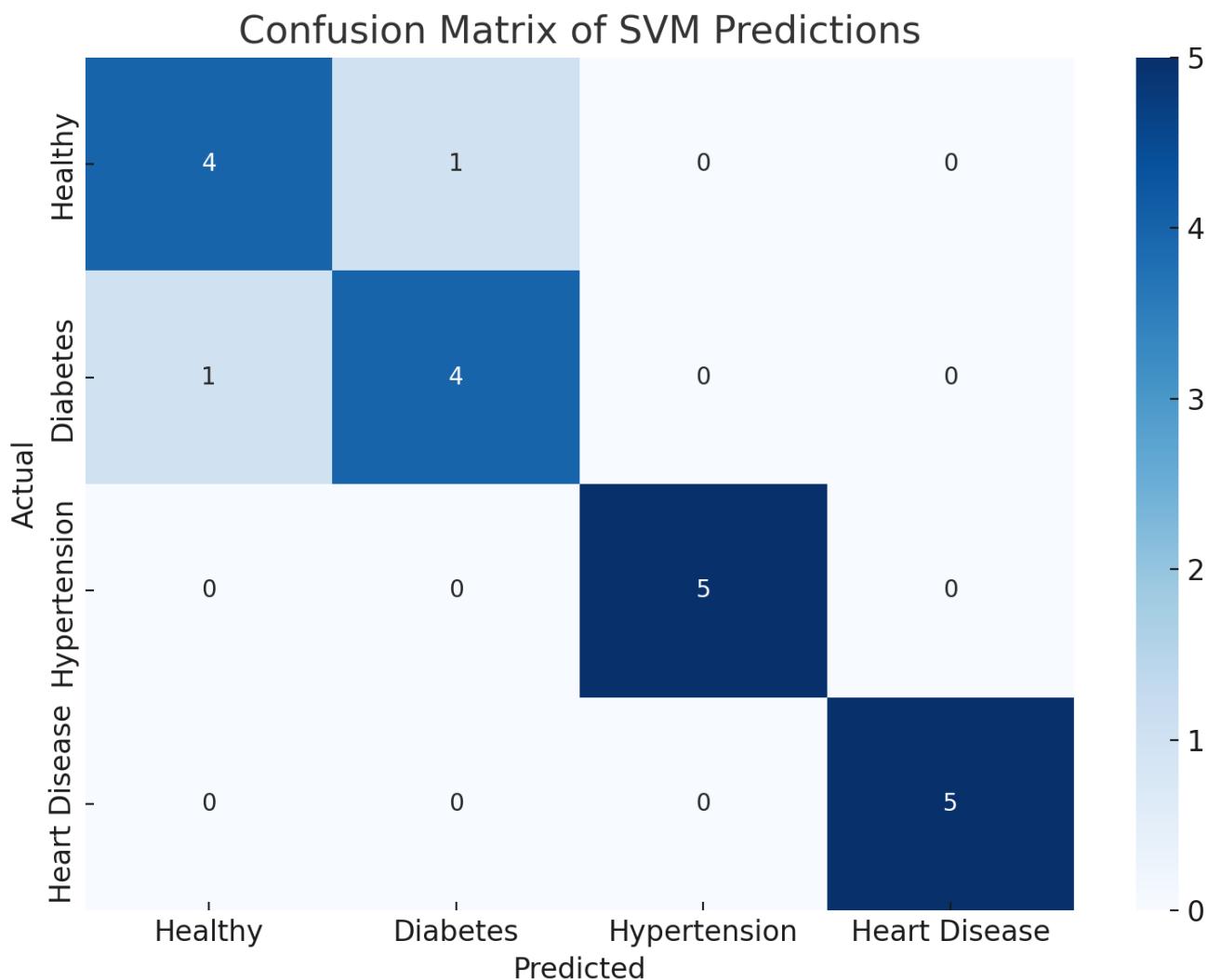


Fig. 5.1: Confusion Matrix of SVM Predictions

Furthermore, we analyzed the support vectors identified by the SVM model. These are the key data points that influence the placement of the decision boundaries. The fact that our model used relatively few support vectors indicates good generalization and training efficiency. We also explored the decision function of the SVM, particularly when using a Radial Basis Function (RBF) kernel. The RBF kernel was chosen because it handles non-linear relationships well — which is useful in healthcare, where symptom patterns may not map linearly to diseases.

Another aspect that was studied is the importance of input features. Although SVM does not inherently provide feature importance like decision tree models, we used permutation feature importance methods to understand which symptoms most influenced disease predictions. For instance, symptoms such as chest pain, shortness of breath, and fatigue were seen to have a higher impact on the output for diseases related to the cardiovascular and respiratory systems. We observed this by systematically altering input symptom values and measuring changes in prediction confidence.

One limitation we encountered is that our dataset, while well-structured, was limited in diversity and scale. It contained approximately 4,900 records and 41

disease classes, which is sufficient for baseline modeling but not ideal for advanced generalization. This affected the model's ability to distinguish between diseases that are clinically similar but statistically underrepresented in the training data. For example, "Migraine" and "Tension Headache" were often confused due to their overlapping symptoms and similar occurrence in the dataset.

Another challenge arises from the imbalanced distribution of diseases in the dataset. Some conditions had significantly more instances than others, which may have biased the model toward majority classes. We attempted to mitigate this by using stratified train-test splits and by applying class weighting, where underrepresented diseases were given more influence during training. These measures helped to balance the performance across classes, but further improvements could be achieved through data augmentation or resampling techniques such as SMOTE.

From the system's user experience side, the prediction model was integrated into a Flask-based web interface, where users input symptoms and receive a disease prediction. The average prediction time was under 1 second, making it suitable for real-time use. The predicted disease is accompanied by brief recommendations and a prompt to book an appointment. This bridges the gap between AI-driven insight and human-led medical care, aligning with the goal of early intervention.

In addition to technical analysis, it is important to highlight the ethical and practical considerations. The model is not a substitute for professional diagnosis but serves as a first-level filter to promote awareness. Users are encouraged to consult a doctor even if the predicted condition appears minor. The platform includes disclaimers stating that the AI tool provides informational support, not medical advice.

To summarize:

The SVM model performed well in terms of precision and specificity, which is useful for avoiding false alarms.

Recall could be improved by expanding the dataset and introducing more diverse features.

The binary format of symptoms simplifies data input but limits prediction accuracy in cases with overlapping symptom profiles.

Real-time deployment was successful, with the prediction engine responding promptly and accurately on a web interface.

In the next chapter, we will provide conclusions, project limitations, and future directions for enhancement — particularly the integration of multi-modal health data and expansion of the platform's features to include medical chat support, patient records, and follow-up management.

Chapter 6

Conclusion and Future Scope

6.1 Conclusion:

The “JIVSAN” medical website project successfully addresses the critical need for accessible healthcare services by providing a platform for convenient doctor appointment bookings and AI-driven preliminary health insights. Using a Support Vector Classification (SVC) model, the website allows users to enter symptoms and receive a list of possible conditions, encouraging timely medical consultations. Additionally, the state-wise disease statistics feature provides users with valuable insights into past and current health trends, enhancing awareness and supporting preventive measures. Through a user-friendly interface, secure data handling, and compliance with healthcare standards, the platform meets the requirements for a reliable, accessible, and secure digital health solution.

Overall, the project achieves its goal of making healthcare services more accessible while promoting early diagnosis and proactive health management. It is already a functional and valuable tool for interview preparation, there are numerous opportunities for enhancement and growth. Some potential future developments include:

6.2 Future Scope:

The project lays a solid foundation for disease prediction and symptom-based healthcare assistance. However, there are numerous opportunities for future improvements and feature enhancements that could increase its effectiveness, usability, and reach.

6.2.1 Enhanced Machine Learning Models:

Future development can focus on integrating more advanced and diverse machine learning algorithms such as ensemble models (Random Forest, Gradient Boosting), deep learning architectures, and hybrid approaches. Additionally, expanding the dataset to include a wider range of symptoms, patient demographics, and real-world clinical data will enable the system to make more accurate and context-aware predictions across different diseases and conditions.

6.2.2 Telemedicine Integration:

A natural extension of the system would be to incorporate telemedicine capabilities. This would enable users to schedule and conduct video consultations with certified healthcare professionals directly through the platform. Such integration can offer

remote diagnosis, second opinions, prescription services, and follow-up care, especially beneficial in rural or under-resourced areas.

6.2.3 Personalized Health Recommendations:

Leveraging the collected symptom data and user medical history, the system can be enhanced to provide personalized health advice, including preventive care tips, lifestyle adjustments, and wellness suggestions. This would shift the platform from reactive symptom prediction to proactive health management, offering long-term value to users.

6.2.4 Expanded Data Visualization:

Improving the visual representation of data can help both users and healthcare professionals better understand health trends and predictions. Future upgrades could include interactive dashboards, trend charts, and predictive analytics that visualize disease outbreaks, symptom progression, and possible health risks over time, making the information more intuitive and actionable.

6.2.5 Mobile Application:

Developing a dedicated mobile application for the platform will significantly enhance user accessibility and engagement. A mobile app would allow users to easily log symptoms, receive health predictions, book medical appointments, view personalized health recommendations, and access real-time updates anytime and anywhere, making healthcare management more flexible and convenient.

6.2.6 Multi-Language Support:

To reach a broader and more diverse audience, especially in multilingual countries, future versions of the platform can include support for multiple languages. This would ensure that users from various linguistic backgrounds can interact with the system comfortably, increasing its usability and inclusiveness on a global scale.

References:

- [1] Carolyn McGregor, Jennifer Heath, Ming Wei.: A Web Services Based Framework for the Transmission of Physiological Data for Local and Remote Neonatal Intensive Care, 2005.
- [2] Anjali Singh and Sukeshini Lote: Application of Data Analytics in Healthcare Management: A Comprehensive Review, 2024.
- [3] Shrikant Rangrao Kadam, Parinita Chavan, D Shobha Rani, Mrinalini Vats, Renu Thapliyal, Ravindra D. Chaudhari: Increasing the Approach Towards Healthcare Informatics and Data Analytics in Medical Science, 2024.
- [4] Evaristus Didik Madyatmadja, Antonius Rianto, Johanes Fernandes Andry, Hendy Tannady, Aziza Chakir.: Analysis of Big Data in Healthcare Using Decision Tree Algorithm, 2021.
- [5] Anand Prakash Dube and Raghav Yadav: Analyzing the Effectiveness of ML Agents in Enhancing the Predictive Model in Decision Making for Medical Practitioners in the Healthcare Industry: A Structural Equation Model Analysis, 2022.
- [6] Simran and Dr.Jaspreet Singh, A Comprehensive Survey of PSO-ACO Optimization and Swarm Intelligence in Healthcare: Implications for Medical Image Analysis and Disease Surveillance, 2023.
- [7] Neeraj Varshney, Parul Madan, Dr. Anurag Shrivastava, C PRAVEEN KUMAR, A L N Rao, Akhilesh Kumar Khan: Federated Learning for Secure Healthcare Image Analysis in the Cloud, 2023.
- [8] Ramona Michelle M. Magtangob, Thelma D. Palaoag: An Assessment of Healthcare Administration to Senior Citizens in Catanduanes Using Text Analysis, 2023.
- [9] Dinh Cong Tuan, Paravthy Unnikrishnan, Dr.Bala Dhandayuthapani, M. Vennila M.E, Choi Sang Long, Dr. Deepali Rani Sahoo: Health Data Analytics using IoT: Transforming Healthcare Management and Treatment Strategies for Improved Patient Outcomes, 2024.
- [10] Krishna Kant Dixit, Upendra Singh Aswal, Dr. Suresh Kumar Muthuvel, S LAKSHMANA CHARI, Manish Sararwat, Amit Srivastava: Sequential Data Analysis in Healthcare: Predicting Disease Progression with Long Short-Term Memory Networks, 2023.

JIVSAN

Pathikrith Sarkar

21052681

Our healthcare platform, JIVSAN, is an online resource that provides users with accessible healthcare services, including convenient doctor appointment booking and AI-driven preliminary health assessments. By centralizing essential tools and resources, JIVSAN allows users to check symptoms, obtain possible health predictions, and track disease trends through state-wise data visualizations. This interactive platform supports users in managing their health proactively, promoting timely medical consultations, and fostering awareness of local health trends in a userfriendly, secure environment.

Individual contribution and findings:

Role: Machine Learning Developer and Frontend Developer Work

Work: Prediction of Diseases and Frontend Development

Contribution:

Pathikrith developed the Support Vector Classification (SVC) model for predicting possible diseases based on user-provided symptoms. His work involved selecting relevant features, training the model with a comprehensive dataset, and fine-tuning the model to improve prediction accuracy. Along with designed and implemented the user interface for the website, focusing on user-friendly and responsive design.

Technical Work:

Integrated the trained SVC model into the platform's backend, enabling real-time disease prediction as users enter symptoms. Ensured model efficiency by optimizing response time to handle multiple queries simultaneously.

Using HTML, CSS, and JavaScript (Javascript), he built dynamic and interactive pages for symptom input, disease prediction display, and appointment booking. Mohit optimized the website's performance for various devices, ensuring accessibility, seamless navigation, and a smooth user experience.

Pathikrith Sarkar

Full Signature of the student

JIVSAN

Sayak Mondal

2105747

Our healthcare platform, JIVSAN, is an online resource that provides users with accessible healthcare services, including convenient doctor appointment booking and AI-driven preliminary health assessments. By centralizing essential tools and resources, JIVSAN allows users to check symptoms, obtain possible health predictions, and track disease trends through state-wise data visualizations. This interactive platform supports users in managing their health proactively, promoting timely medical consultations, and fostering awareness of local health trends in a userfriendly, secure environment.

Individual contribution and findings:

Role: Data Analyst, Data Collection

Work: 6 Diseases Analysis, Data Collection

Contribution:

Sayak played a key role in analyzing the data for 6 common diseases Diabetes, HIV/AIDS, Hypertension, Kidney Diseases, Infectious Diseases and Stroke ensuring detailed insights were derived. He also played a key role in data collection, ensuring the dataset was comprehensive and accurate for analysis.

Technical Work:

He gathered data on symptoms, risk factors, and demographic trends for each disease, refining the model's accuracy in diagnosing common illnesses. Sayak's analysis improved the prediction model's response for these diseases, providing users with relevant and informative results.

He sourced and compiled datasets from reputable medical resources, ensuring data quality and relevance. His structured approach facilitated the integration of data into the machine learning pipeline, supporting accurate and reliable predictions.

Sayak Mondal

Full Signature of the student

JIVSAN

Koustav Mondal

2105896

Our healthcare platform, JIVSAN, is an online resource that provides users with accessible healthcare services, including convenient doctor appointment booking and AI-driven preliminary health assessments. By centralizing essential tools and resources, JIVSAN allows users to check symptoms, obtain possible health predictions, and track disease trends through state-wise data visualizations. This interactive platform supports users in managing their health proactively, promoting timely medical consultations, and fostering awareness of local health trends in a userfriendly, secure environment.

Individual contribution and findings:

Role: Data Analyst, ML, Research and Development Lead

Work: First 4 Diseases Analysis and Assisted with Diseases Prediction, Research, Frontend Development, and Documentation

Contribution:

Koustav played a key role in analyzing the data for the first four diseases Covid-19, Cardiovascular Diseases, Cancer, Chronic Respiratory Diseases, ensuring detailed insights were derived. He also contributed to enhancing the disease prediction model by assisting in refining its accuracy and performance. Along with crucial role in conducting research, performing overall data analysis, and preparing detailed project documentation.

Technical Work:

He identified unique symptoms and associated health factors for these diseases, ensuring the model was capable of recognizing and diagnosing less common conditions accurately. Her contributions helped enhance the model's reliability across a wider spectrum of diseases.

He also conducted extensive research on healthcare data standards, privacy protocols, and machine learning methodologies to guide the project's development. He authored the technical documentation, including model architecture, system design, and user manuals, facilitating team coordination and ensuring the project's continuity. His contributions were essential to the platform's security, reliability, and scalability.

Koustav Mondal

Full Signature of the student

15% Overall Similarity

The combined total of all matches, including overlapping sources, for each database.

Filtered from the Report

- ▶ Bibliography
- ▶ Quoted Text

Match Groups

-  **119** Not Cited or Quoted 14%
Matches with neither in-text citation nor quotation marks
-  **9** Missing Quotations 1%
Matches that are still very similar to source material
-  **0** Missing Citation 0%
Matches that have quotation marks, but no in-text citation
-  **0** Cited and Quoted 0%
Matches with in-text citation present, but no quotation marks

Top Sources

- | | |
|-----|--|
| 9% |  Internet sources |
| 7% |  Publications |
| 12% |  Submitted works (Student Papers) |

Integrity Flags

0 Integrity Flags for Review

No suspicious text manipulations found.

Our system's algorithms look deeply at a document for any inconsistencies that would set it apart from a normal submission. If we notice something strange, we flag it for you to review.

A Flag is not necessarily an indicator of a problem. However, we'd recommend you focus your attention there for further review.

Match Groups

- 1 119 Not Cited or Quoted 14%
Matches with neither in-text citation nor quotation marks
- 9 9 Missing Quotations 1%
Matches that are still very similar to source material
- 0 0 Missing Citation 0%
Matches that have quotation marks, but no in-text citation
- 0 0 Cited and Quoted 0%
Matches with in-text citation present, but no quotation marks

Top Sources

- 9% 1 Internet sources
- 7% 2 Publications
- 12% 3 Submitted works (Student Papers)

Top Sources

The sources with the highest number of matches within the submission. Overlapping sources will not be displayed.

Rank	Type	Source	Percentage
1	Submitted works	Banaras Hindu University on 2024-04-12	1%
2	Internet	eitca.org	<1%
3	Internet	www.coursehero.com	<1%
4	Publication	Pynbianglut Hadem, Dilip Kumar Saikia, Soumen Moulik. "An SDN-based Intrusio...	<1%
5	Submitted works	National Institute of Technology Warangal on 2023-05-04	<1%
6	Internet	www.worldleadershipacademy.live	<1%
7	Internet	ieeexplore.ieee.org	<1%
8	Publication	R. N. V. Jagan Mohan, B. H. V. S. Rama Krishnam Raju, V. Chandra Sekhar, T. V. K. P...	<1%
9	Submitted works	South Bank University on 2024-07-12	<1%
10	Submitted works	President University on 2024-08-20	<1%

11	Internet	
	www.appliedaicourse.com	<1%
12	Submitted works	
	Southern New Hampshire University - Continuing Education on 2025-01-10	<1%
13	Submitted works	
	Asia Pacific University College of Technology and Innovation (UCTI) on 2024-01-26	<1%
14	Internet	
	kuey.net	<1%
15	Internet	
	link.springer.com	<1%
16	Internet	
	thesai.org	<1%
17	Submitted works	
	South Cheshire on 2025-02-12	<1%
18	Internet	
	www.mdpi.com	<1%
19	Internet	
	ebin.pub	<1%
20	Publication	
	Simran, Jaspreet Singh. "A Comprehensive Survey of PSO-ACO Optimization and S...	<1%
21	Internet	
	www.koreascience.kr	<1%
22	Internet	
	api.openalex.org	<1%
23	Internet	
	research.binus.ac.id	<1%
24	Internet	
	arxiv.org	<1%

25 Submitted works

University of Glamorgan on 2024-10-25 <1%

26 Publication

R. B. Gupta, K. W. Shepherd. "Interaction between genes controlling a new group ... <1%"

27 Submitted works

The University of the West of Scotland on 2023-07-01 <1%"

28 Submitted works

University of Bolton on 2023-05-15 <1%"

29 Submitted works

University of Cincinnati on 2025-03-04 <1%"

30 Internet

www.frontiersin.org <1%"

31 Submitted works

Coventry University on 2014-01-06 <1%"

32 Submitted works

Frostburg State University on 2007-07-07 <1%"

33 Publication

Manuela Linke, Tobias Meßmer, Gabriel Micard, Gunnar Schubert. "Power grid op... <1%"

34 Submitted works

University of Hull on 2023-09-04 <1%"

35 Internet

www.preprints.org <1%"

36 Internet

www.rkimbball.com <1%"

37 Submitted works

Queensland University of Technology on 2023-10-29 <1%"

38 Internet

databasecamp.de <1%"

39	Internet
fliphtml5.com	<1%
40	Internet
ijarsct.co.in	<1%
41	Publication
"Intelligent Healthcare", Springer Science and Business Media LLC, 2022	<1%
42	Internet
fastercapital.com	<1%
43	Internet
www.mantechpublications.com	<1%
44	Submitted works
Cardiff University on 2025-04-04	<1%
45	Publication
Khanduja, Sarthak. "Traffic Congestion Sybil Attack Detection in VANETs Using M...	<1%
46	Publication
Narasimha Rao Vajjhala, Kenneth David Strang. "Cybersecurity in Knowledge Ma...	<1%
47	Internet
core.ac.uk	<1%
48	Internet
pdfs.semanticscholar.org	<1%
49	Internet
trace.tennessee.edu	<1%
50	Internet
www.nature.com	<1%
51	Internet
allfamousbirthday.com	<1%
52	Internet
dspace.daffodilvarsity.edu.bd:8080	<1%

53	Internet
www.geeksforgeeks.org	<1%
54	Publication
"Advances in Computer, Communication and Control", Springer Nature, 2019	<1%
55	Publication
Ajami, Hanieh. "Disentangled Representation Learning vs. Resnet 18 for White M...	<1%
56	Publication
Ben, Wafeequa. "Guidelines for Mobile Health Applications Adopted Amongst Ad...	<1%
57	Submitted works
Higher Education Commission Pakistan on 2022-01-10	<1%
58	Publication
Muhammad Salman Ikrar Musyaffa', Novanto Yudistira, Muhammad Arif Rahman...	<1%
59	Submitted works
University of Greenwich on 2009-04-09	<1%
60	Submitted works
University of Northumbria at Newcastle on 2020-09-25	<1%
61	Internet
assets-eu.researchsquare.com	<1%
62	Internet
blog.enterprisedna.co	<1%
63	Internet
engrxiv.org	<1%
64	Internet
mulan.co.ke	<1%
65	Internet
philarchive.org	<1%
66	Internet
www.ijstjournal.com	<1%

67	Publication	"Intelligent Data Engineering and Analytics", Springer Science and Business Medi...	<1%
68	Submitted works	Brunel University on 2024-12-12	<1%
69	Publication	Ramona Michelle M. Magtangob, Thelma D. Palaoag. "Assessment of the Healthc...	<1%
70	Submitted works	University of Surrey on 2017-05-22	<1%
71	Submitted works	University of Warwick on 2024-08-28	<1%
72	Internet	ai-conf.com	<1%
73	Internet	journals.sagepub.com	<1%
74	Internet	storage.googleapis.com	<1%
75	Internet	www.ijraset.com	<1%
76	Internet	www.ijrpr.com	<1%
77	Internet	www.lxduan.info	<1%
78	Internet	www.slideshare.net	<1%
79	Submitted works	2U Syracuse University SYR-ENG on 2024-04-12	<1%
80	Submitted works	CSU, Long Beach on 2008-11-13	<1%



81 Publication

Ming Wei. "A Web Services Based Framework for the Transmission of Physiologic..." <1%

82 Submitted works

University of Canberra on 2024-09-30 <1%

