RISK ASSESS : A SYMPTOM-BASED DISEASE PREDICTOR

A PROJECT REPORT

Submitted by

NIRANJANI K - 20201CSE0514
HARSHA M - 20201CSE0504
HARISH N G - 20201CSE0521

Under the guidance of

Mr. ARUN KUMAR S

in partial fulfillment for the award of the degree of

BACHELOR OF TECHNOLOGY

IN

SCHOOL OF COMPUTER SCIENCE AND ENGINEERING

At

PRESIDENCY UNIVERSITY

BENGALURU

JANUARY 2024

PRESIDENCY UNIVERSITY

SCHOOL OF COMPUTER SCIENCE & ENGINEERING

CERTIFICATE

This is to certify that the Project report **"RISK ASSESS A SYMPTOMBASED DISEASE PREDICTOR"** being submitted by "NIRANJANI K, HARSHA M, HARISH N G" bearing roll number(s) "20201CSE0514, 20201CSE0504, 20201CSE0521" in partial fulfillment of requirement for the award of degree of Bachelor of Technology in Computer Science and Engineering is a bonafide work carried out under my supervision.

Mr. ARUN KUMAR S
Assistant Professor
School of CSE
Presidency University

Dr. PALLAVI R
Associate Professor & HOD
School of CSE
Presidency University

Dr. C. KALAIARASAN
Associate Dean
School of CSE
Presidency University

Dr. SHAKKEERA L
Associate Dean
School of CSE
Presidency University

Dr. SAMEERUDDIN KHAN Dean
School of CSE
Presidency University

PRESIDENCY UNIVERSITY

SCHOOL OF COMPUTER SCIENCE & ENGINEERING

DECLARATION

We hereby declare that the work, which is being presented in the project report entitled **RISK ASSESS A SYMPTOM-BASED DISEASE PREDICTOR** in partial fulfillment for the award of Degree of **Bachelor of Technology** in **Computer Science and Engineering**, is a record of our investigations carried under the guidance of **Mr. Arun Kumar S, Assistant Professor, School of Computer Science and Engineering, Presidency University, Bengaluru.** We have not submitted the matter presented in this report anywhere for the award of any other Degree.

NAME	ROLL NUMBER	SIGNATURE
NIRANJANI K	20201CSE0514	
HARSHA M	20201CSE0504	
HARISH N G	20201CSE0521	

ABSTRACT

The merging of state-of-data collection has led to remarkable improvements in the medical care sector recently. This paper offers "Risk Assess," a creative Flask web application that was meticulously designed. "Risk Assess" completes the step by using patient vitals to produce risk assessments. .

This web application, which is based in Python, employs various Machine Learning models, including K-Nearest Neighbor (KNN), Gaussian Naïve Bayes (NB), Linear Support Vector Machine (SVM), Classification and Regression (CART), and K-Nearest Neighbor (KNN), to analyze and create a model that predicts whether a given set of symptoms is indicative of a specific disease. Profoundly, "Risk Assess" improves comfort and functional productivity by calming down the forecast of heart disease, kidney disease, liver disease, diabetes, cancer, and other diseases. Our project features a reliable login page that enables users to create new accounts using their email addresses.

We prioritize security and customer satisfaction by requesting a login and secret key to be sent with their email. This study article highlights the critical role of machine learning and includes working engineering, plan standards, and execution complexity. "Risk Assess" embodies the application of machine learning and Python-based advancements, demonstrating the ability to make predictions based on symptoms using inventive computer configurations.

ACKNOWLEDGEMENT

First of all, we indebted to the **GOD ALMIGHTY** for giving me an opportunity to excel in our efforts to complete this project on time.

We express our sincere thanks to our respected **Dr. Md. Sameeruddin Khan**, Dean, School of Computer Science and Engineering and Information Science, Presidency University for getting us permission to undergo the project.

We record our heartfelt gratitude to our beloved Associate Deans **Dr. Kalaiarasan C and Dr. Shakkeera L**, School of Computer Science and Engineering and Information Science, Presidency University and **Dr. Pallavi R**, Head of the Department, School of Computer Science and Engineering, Presidency University for rendering timely help for the successful completion of this project.

We are greatly indebted to our guide **Mr. Arun Kumar S**, **Assistant Professor**, School of Computer Science and Engineering, Presidency University for his inspirational guidance, and valuable suggestions and for providing us a chance to express our technical capabilities in every respect for the completion of the project work.

We would like to convey our gratitude and heartfelt thanks to the University Project-II Coordinators **Dr. Sanjeev P Kaulgud, Dr. Mrutyunjaya MS** and also the department Project Coordinators **Mr. Zia Ur Rahman, Mr. Peniel John Whistely**.

We thank our family and friends for the strong support and inspiration they have provided us in bringing out this project.

NIRANJANI K

HARSHA M

HARISH N G

LIST OF FIGURES

TABLE OF CONTENTS

CHAPTER-1
INTRODUCTION

Within the burgeoning field of clinical benefits, machine learning compromises have become a formidable force, offering unprecedented levels of advancement in diagnosis and treatment. Ampleness of Machine Learning models, especially in disease identification, is significantly linked to the availability of large and reliable datasets that depict the diverse consequences and manifestations of altered illnesses.

By providing a thorough evaluation of the enhancement of risk assessment, this assessment aims to contribute to this progress. Our "Risk Assess" project is a Flask-based web application.

The application serves as a vital tool for accurately predicting the symptoms of various illnesses and producing a reliable result. "Risk Assess" constantly adjusts to the developing clinical record.

The start-to-finish methodology and accurate prediction of the web application represent its middle headway.

1.1 Statement about the Problem

"For Machine learning models to work we need large amounts of trusted patient's data on various symptoms of various diseases which can help in diagnosis. How can we make collection of this data easy using technology"?

1.2 Objective of the project

1.2.1 Data Collection Framework

Source Identification: Explores diverse sources, including electronic health records, wearable devices, and health-related applications, to gather comprehensive patient data.

Patient Consent: Implements a robust informed consent process, emphasizing transparency and adherence to ethical standards, to secure patient approval for data utilization.

1.2.2 Data Acquisition

Continuous Monitoring: Emphasize continuous, longitudinal data collection to capture evolving symptoms and disease progress over time.

1.2.3 Validation and Quality Assurance

Feedback Loop: Establish a feedback loop to continuously refine data collection processes based on ongoing assessments and the performance of machine learning models.

Cross-Validation: Implement cross-validation techniques to rigorously assess the performance of machine learning models on the collected data.

1.3 Motivation

Closing Healthcare Gaps: Our project aims to address gaps in healthcare, especially in early disease detection and proactive care, using data-driven solutions.

Technology for Empowerment: By leveraging machine learning and data analytics, we seek to empower healthcare with predictive tools that benefit both patients and providers.

Better Patient Care: Our goal is to improve patient outcomes through early risk identification, leading to more effective treatments and healthier lives.

Collaboration for Impact: We're motivated by collaboration between healthcare and technology, aiming for a holistic solution that benefits all stakeholders.

Empowering Health Choices: Ultimately, our project aims to empower individuals to make informed health choices, promoting proactive healthcare management.

1.4 Functionality

1.4.1 Admin Functionality

Data Collection: The project allows for the collection of patient data through a user friendly interface, ensuring ease of use and comprehensive data capture.

Machine Learning Integration: It integrates machine learning algorithms to analyze the collected data, providing accurate predictions for diseases such as cancer, diabetes, heart disease, kidney disease, and liver disease.

1.4.2 Patient Functionality

User Profiles: Users can create and manage patient and specialist profiles, facilitating personalized healthcare interactions and data management. **Symptom Input**- The application provides a platform for patients to input their symptoms, which are then utilized by the machine learning models for predictive analysis.

Prediction Download: Patients can download the predictions for future reference, enabling them to track their health status and make informed decisions about their healthcare.

Disease Insights: The application offers insights into the causes of diseases, empowering users with knowledge about their conditions and potential risk factors.

CHAPTER-2
LITERATURE SURVEY

The field of Machine Learning in medical care depends vigorously on the accessibility of tremendous and solid datasets to upgrade demonstrative abilities. In this paper, we propose a strategy for joining the necessities of contemporary Machine Learning with those of conventional information combination techniques. We examine the possible results of utilizing metadata got from data consolidation processes for chipping away at the ampleness and efficiency of Machine Learning models. Towards this bearing, we look at two ordinary use cases over data storage facilities, incorporate increment and joined learning. We feature new examination valuable open doors according to the viewpoints of frameworks, portrayals, factorized learning, and united advancing by uniting Machine Learning and information reconciliation. Existing information mining techniques and Machine Learning calculations need a significant amount of information to prepare, more information should be gained before they can be utilized. The multifaceted nature of the model influences the size of the document. while examining information is troublesome, information driven learning might excessively fundamental

This work is instrumental in molding the mechanical system of the proposed web-based application, "Risk Assess", with an emphasis on lining up with the developing scene of computerized wellbeing rehearses.

A survey on Data Collection for Machine Learning, conducted on April 1 2021, by Y. Roh, G. Hoe and S.E. Whang contributes essential guidance on tailoring the adaptive data collection for "Risk Assess". This paper digs into the basic convergence of Data Innovation (IT) developments, with thorough financial plan the board frameworks. The general goal is to smooth out information assortment, improve financial plan arrangement processes, and engage undertakings to form key spending arrangements. The paper investigates the combination of these innovations as well as gives execution ideas and shields to guarantee a smooth and secure progress.

This study dives into different information assortment methods and presents the implantation

At the present time in Distant Sensor Associations, data collection expects a critical part. The principal job of data collection is to get a gigantic measure of data while to hack down data dropping in view of sensor center memory hindrances.

The essential examination issue in the fields of Machine Learning and information mining is lopsided information characterization. As per the perspective of practical application, unbalanced instructive assortments have a colossal degree. During the grouping, the classifier will predisposition the examples of most of classes, which will at last straightforwardly affect the arrangement brings about request to control the misfortune pace of separation. According to the investigation eventual outcomes of local and new researchers lately, most of the proposed lopsided data computations can really control the impact of data cumbersomeness on the presentation of the classifier. Thus, in light of understanding the assessment outcomes of lopsided data gathering, this paper made sense of the investigation course of computer-based intelligence and proposed an estimation considering de-noising oversampling advancement.

The essential examination issue in the field of Machine Learning calculations to enhance the life span of the proposed web-application.

For this situation, we assembled remark information by utilizing the information assortment programming. The information was then exposed to pre-handling, word division, and opinion marking. We make a feeling classifier to consequently characterize unidentified information by changing different boundaries used to prepare the information. We show probably that using artificial intelligence methodologies to perform assessment request on thing reviews can achieve incredible results. this article adds to the expanding field of profound examination in item surveys, revealing insight into the adequacy of Machine Learning approaches. Various fields are utilizing Machine Learning to address different issues. This is being driven by the accessibility of huge datasets, minimal expense calculation assets, and the improvement of vigorous Machine Learning calculations. Some computer based intelligence applications require association of contraptions off-the-cross section for data collection and dealing with. Frameworks that can work freely during arrangement are required for these applications.

By and large, the exactness of the investigation results is affected by the consistency and uprightness of the information during the course of huge information examination. Usually before the data assessment process starts, data gathering should be made on the assembled data sources to hold that resulting assessment won't make goofs there of brain due data peculiarities. Therefore, one of those vital undertakings in information grouping is to keep up with respectability.

Drawing insights "Amalur: Data Integration Meets Machine Learning," 2023 IEEE 39th International Conference on Data Engineering, and "A Survey on Data Collection for Machine Learning", the platform prioritizes a user-friendly interface for patients. Data Acquisition and Processing System, inspired by a 2022study by C. Zheng, M. Zhang, Y. Wang and M. Zou from Bangkok, Thaimi land, further enhances the holistic well-being aspect. Based on a Comprehensive study on Distinct Data Collection Techniques using Machine Learning Approaches for WSN in 2021 "Risk Assess" adapts dynamically to the evolving healthcare landscape. The platform's core innovation lies in its symptom-based disease prediction accurately using machine learning models and also providing authentication. The subject of close to home examination of item surveys is a hot area of information mining research at this moment. Whether it is in researcher or money related fields, text up close and personal examination of online business thing reviews has remarkable investigation regard. This article used simulated intelligence to lead assessment examination for clothing web business thing reviews. For this situation, we assembled remark information by utilizing the information assortment programming. The information was then exposed to pre-handling, word division, and opinion marking. We make a feeling classifier to consequently characterize unidentified information by changing different boundaries used to prepare the information. We show probably that using artificial intelligence methodologies to perform assessment request on thing reviews can achieve incredible results. this article adds to the expanding field of profound examination in item surveys, revealing insight into the adequacy of Machine Learning approaches.

The excursion incorporates the cautious coordination of information assortment, pre-handling, and feeling marking, finishing in the making of an opinion classifier that stands ready to explore the intricacies of close to home articulation inside the online business space. Through observational trial and error, the review approves the reasonability of utilizing Machine Learning for opinion order, denoting a huge step in the continuous

investigation of the convergence among innovation and human feeling in the computerization.

In addition, our examination underscores the job of patient commitment through portable wellbeing applications. We propose the advancement of easy-to-understand applications, enabling patients to willfully contribute side effect information, track wellbeing measurements, and give express agree to information utilization. To additionally improve ongoing wellbeing information, including imperative signs and action levels. All through the paper, serious areas of strength are put on for executing hearty safety efforts to protect the classification and trustworthiness of the gathered information.

The trial results show the way that the proposed calculation can really improve the classifier's characterization impact and has a specific enemy of commotion execution. this paper not just recognizes the criticality of tending to lopsided information characterization in Machine Learning yet additionally adds to the continuous talk by proposing a clever calculation in light of de-noising oversampling innovation. The exploration course illustrated in this paper lines up with the more extensive endeavors to refine Machine Learning philosophies, especially with regards to genuine world, imbalanced datasets. The trial approval fills in as observational proof of the calculation's viability, offering a promising road for additional investigation and reception in common sense applications. As the field keeps on developing, the proposed calculation remains as a demonstration of the imaginative arrangements arising to address the difficulties of unequal information order in contemporary Machine Learning research.

As we dive into the crossing point of innovation and moral contemplations, this work attempts to impel the field of Machine Learning in medical care forward, encouraging developments that are both logically thorough and morally sound. To address privacy concerns, the project emphasizes the use of SQL technology for the use of secure stored data. By anonymously collecting patient data, we prioritize the protection of individual privacy while still extracting valuable insights that help strengthen our dataset. Furthermore, this framework recognizes the importance of transparency and credibility in the data collection process. An education campaign will be launched to inform and engage patients, ensuring they understand the purpose of the project, the safeguards in place and the potential benefits of their contribution to medical research. In conclusion, this project represents a concerted effort to address the challenges associated with collection.

CHAPTER-3

RESEARCH GAPS OF EXISTING METHODS

3.1 Introduction to Existing Methods

The momentum strategies utilized for information assortment and sickness expectation in Machine Learning models assume a crucial part in medical care by empowering the early location and the executives of illnesses.

These techniques normally include the assortment of different side effects and other important information from patients to foresee the gamble of explicit infections.

3.2 Limitations of Current Data Collection Methods

Regardless of their significance, the flow techniques for information assortment for illness expectation are not without constraints.

These impediments include

Lack of Standardization: There is an absence of normalized conventions for information assortment, prompting varieties in information quality across various sources.

Reliance on Manual Info: Many existing strategies depend on manual contribution of information, which can be tedious and inclined to mistakes.

Difficulty in Catching Nuanced Side effects: A few techniques might battle to catch nuanced side effects or factors that are basic for precise illness expectations.

Limited Availability to Different Datasets: Admittance to assorted datasets is in many cases restricted, prompting one-sided models that may not sum up well.

3.3. Challenges in Disease Prediction Models

Notwithstanding restrictions in information assortment, sickness expectation models face a few difficulties, including

Over fitting and Under fitting: Models might experience the ill effects of over fitting or under fitting because of restricted or one-sided information, prompting unfortunate speculation.

Adaptability to Recent fads: Models might battle to adjust to recent fads or arising illnesses without broad retraining, making them less viable in powerful medical services conditions.

Ethical Contemplations: There are moral contemplations in regards to the treatment of delicate patient information for preparing these models.

3.4 Open doors for Development

Notwithstanding these difficulties, there are open doors for development in information assortment and sickness expectation techniques, for example,

Joining of Cutting-edge Information Assortment Methods: Integrating progressed information assortment procedures, like IoT gadgets for constant information securing, could work on the quality and amount of information.

Development of Vigorous Calculations: The advancement of more strong Machine Learning calculations equipped for dealing with complicated and various datasets could upgrade the exactness of sickness expectation models.

Collaboration for Information Sharing: Coordinated efforts with medical services suppliers to further develop information sharing and admittance to assorted patient populaces could prompt more agent datasets.

All in all, the current strategies for information assortment and illness expectation in Machine Learning models have a few constraints and moves that should be tended to.

By recognizing these holes and open doors for development, we can make ready for more successful infection expectation models that decidedly affect medical care results.

CHAPTER-4

PROPOSED METHODOLOGY

4.1 Objective of the Proposed Methodology

The goal of our proposed strategy is to foster an exhaustive gamble evaluation application that gives exact forecasts to different sicknesses including cancer, diabetes, heart disease, kidney infection, and liver infection. This application intends to offer a one-stop answer for patients looking for clinical and medical care administrations.

4.2 Data Collection

Patient Information Assortment: We will gather many patients' information including clinical history, side effects, analytic experimental outcomes, way of life factors, and hereditary data.

Data Sources: Information will be obtained from patients.

4.3 Machine Learning Algorithm

Calculation Determination: We will use a mix of regulated Machine Learning calculations like Order and Relapse Tress (CART), Liner support Vector Machines (SVM), KNN.

Model Preparing: The chose calculations will be prepared utilizing the gathered patient information to foster sickness expectation models intended for malignant growth, diabetes, coronary illness, kidney infection, and liver infection.

Model Assessment: The exhibition of the created models will be assessed utilizing measurements like exactness, accuracy, review, and F1 score to guarantee their adequacy in foreseeing sickness gambles.

4.4 Application Development

User Connection point Plan: We will plan a natural and easy to understand interface for the gamble evaluation application to work with simple contribution of patient information and show of prescient outcomes.

Integration of Models: The created Machine Learning models will be incorporated into the application backend to empower constant forecasts in view of info information.

4.5 Ethical Considerations

Data Protection: Severe measures will be carried out to guarantee the security and classification of patient information in consistence with medical services guidelines like HIPAA.

Bias Moderation: Steps will be taken to relieve predispositions in the gathered information and Machine Learning models to guarantee fair and unprejudiced forecasts.

In end, our proposed technique expects to foster a vigorous gamble evaluation application that use Machine Learning calculations to give precise forecasts to different infections. By zeroing in on information quality, calculation choice, application improvement, and moral contemplations, we expect to convey a dependable and significant medical services answer for patients.

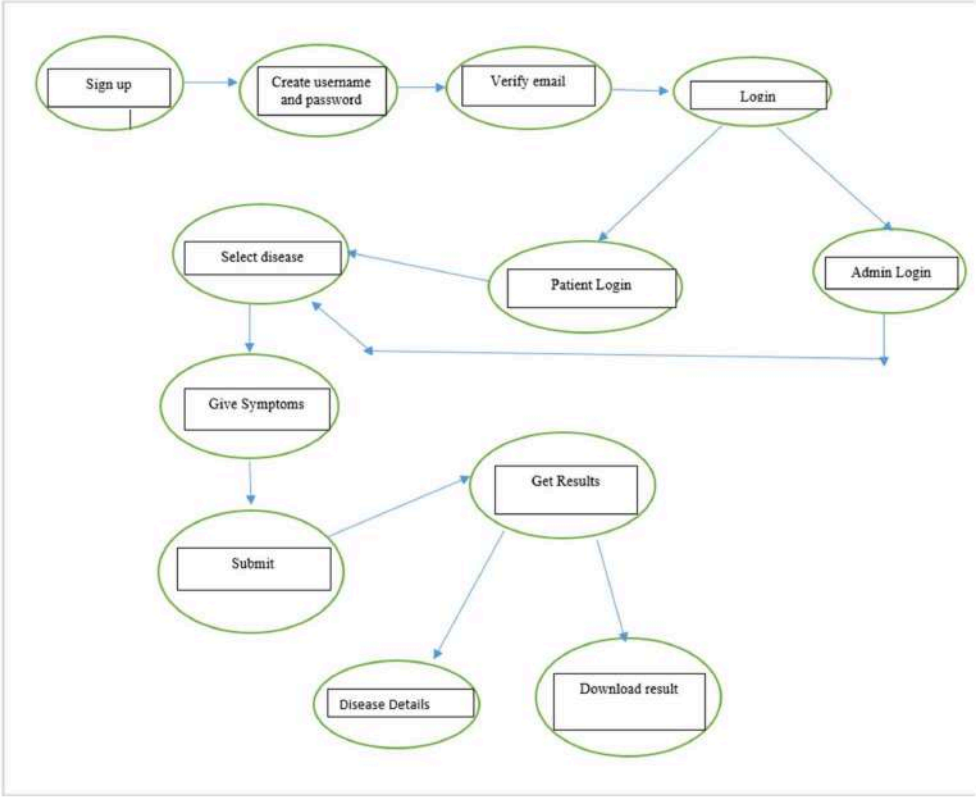


Fig4.1.Risk Assess Flow Diagram

CHAPTER-5

OBJECTIVES

To focus our data collection efforts, we meticulously identify the specific symptoms and diseases relevant to machine learning diagnosis. Clear inclusion criteria are established to guide the selection of patient data, ensuring its pertinence to the targeted medical conditions.

5.1 Data Collection Framework

Source Identification: Explores diverse sources, including electronic health records, wearable devices, and health-related applications, to gather comprehensive patient data. **Patient Consent:** Implements a robust informed consent process, emphasizing transparency and adherence to ethical standards, to secure patient approval for data utilization.

5.2 Data Acquisition

Continuous Monitoring: Emphasize continuous, longitudinal data collection to capture evolving symptoms and disease progress over time.

5.3 Validation and Quality Assurance

Cross-Validation: Implement cross-validation techniques to rigorously assess the performance of machine learning models on the collected data.

CHAPTER-6

SYSTEM DESIGN & IMPLEMENTATION

6.1 SYSTEM DESIGN

System design forms the backbone of any successful application or system. It serves as a blueprint, meticulously outlining the various components, workflows, tasks, and user interactions that bring the system to life. This crucial phase translates abstract ideas into concrete plans, ensuring every aspect is well-defined and seamlessly integrated.

By encompassing both functional and technical considerations, system design provides a holistic view of how the system will be implemented.

This in-depth exploration allows developers to grasp the complexities involved and make informed decisions throughout the development journey.

One key benefit of system design is its ability to optimize resource allocation. With a clear roadmap, developers can prioritize tasks and allocate resources strategically. This prioritization ensures focus on crucial components and functionalities, leading to a more efficient and streamlined development process.

Furthermore, system design marks the pivotal point where the concept translates into reality. It bridges the gap between ideation and implementation, offering a tangible plan to guide developers. This phase empowers them to make informed decisions, anticipate potential challenges, and proactively develop solutions.

In essence, system design is an irreplaceable step in the development process. It provides a comprehensive roadmap, fosters efficient resource management, and facilitates a smooth transition from concept to creation.

Ultimately, it paves the way for a successful implementation, transforming abstract ideas into functional and impactful applications.

6.2 USE CASE DIAGRAM

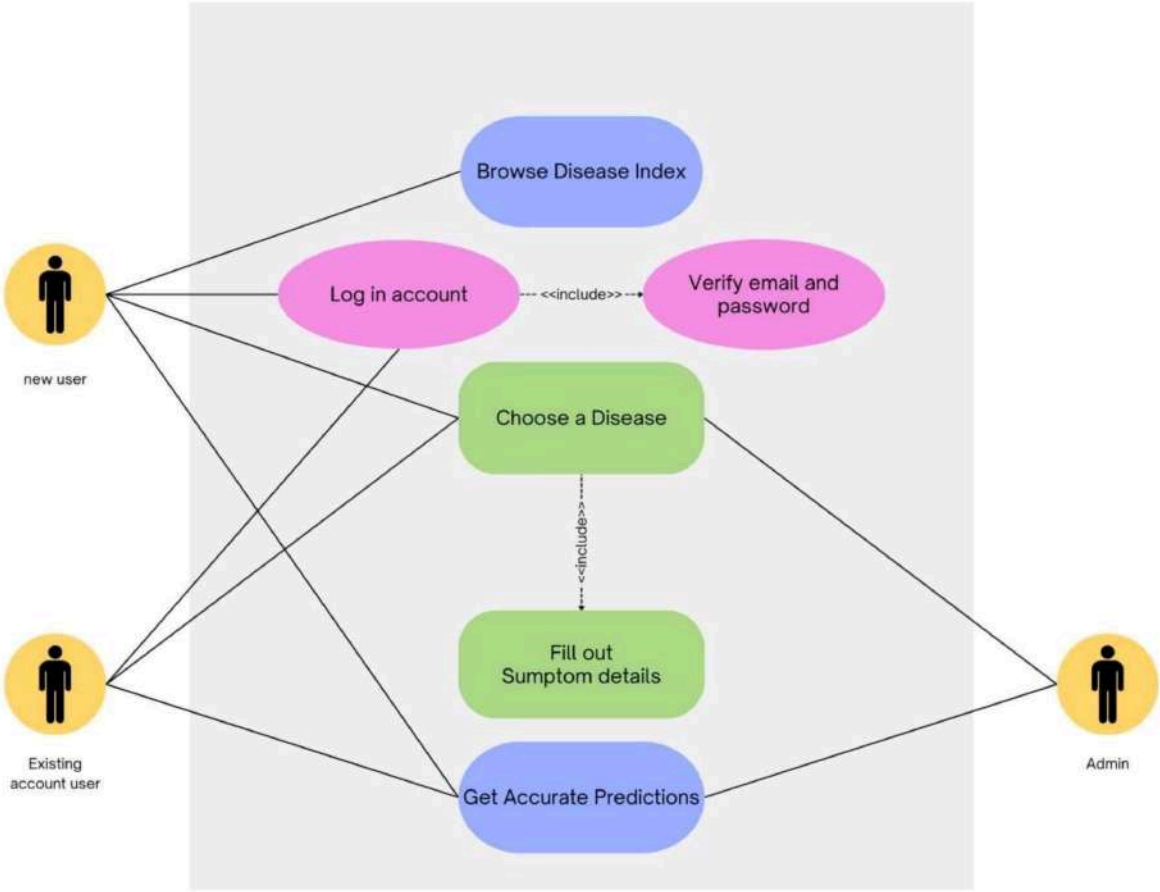


Figure 6.1: Use Case Diagram

6.3 IMPLEMENTATION DETAILS

6.3.1 User Experience

Intuitive Interface

- Design a user-friendly interface with clear navigation and visual cues.
- Accommodate diverse user needs and technological capabilities.
- Provide comprehensive instructions and guidance throughout the process.

Personalization

- Allow users to create their own profile and use it later
- Offer tailored predictions based on individual health symptoms
- Possible of regenerating the response based on user's convenience.

Disease Prediction

- Offer best treatment options, aligned with best practices and guidelines.
- Offer accurate predictions based on the patient symptoms using various machine learning algorithms

6.3.2 Ethical Consideration

Patient Autonomy

- Respect patient rights and preferences, avoiding force or overreliance on Machine Learning.

6.3.3 Software Requirements:

- Programming Languages: Python
- Library: Flask
- Machine Learning Models: Classification and Regression Tree(CART), Linear Support Vector Machine(SVM), Gaussian Naive Bayes(NB), K-Nearest Neighbors(KNN)

6.3.4 Hard Requirements:

- Windows 10, 64-bit OS
- 8Gb Ram and 512 SSD
- Intel i5 processor
- Mobile device (smartphone / tablet)

CHAPTER-7

TIMELINE FOR EXECUTION OF PROJECT

(GANTT CHART)

❖ **Review 0**

- **Title Finalization with Supervisor**: Meet with your project supervisor to discuss and finalize the project's title. This step ensures that you and your supervisor are aligned on the project's focus.
- **Literature Survey**: Start researching relevant literature in your field. This involves reading academic papers, articles, and books related to your topic. The literature survey helps you understand the existing research in the area.
- **Finalizing Objectives**: Define the specific objectives and goals of your project. What do you aim to achieve? What problems are you trying to solve?
- **Deciding the Methodology**: Determine the research methodology you'll use in your project, whether it's experimental, survey-based, analytical, or a combination of methods.

❖ **Review 1**

- **Title**: After finalizing with your supervisor, formally decide on and document the project title.
- **Abstract:** Write the project abstract, a concise summary of your project's purpose, methodology, and expected outcomes.
- **Literature Survey**: Continue your literature survey and ensure that you've referred to at least one research paper in your work.
- **Identify** The advantages and the disadvantages of our project.
- **Software and Framework Details**: List the software framework tools or platforms you plan to use.
- **References**: Site all the sources you've used in your literature survey.

❖ **Review 2**

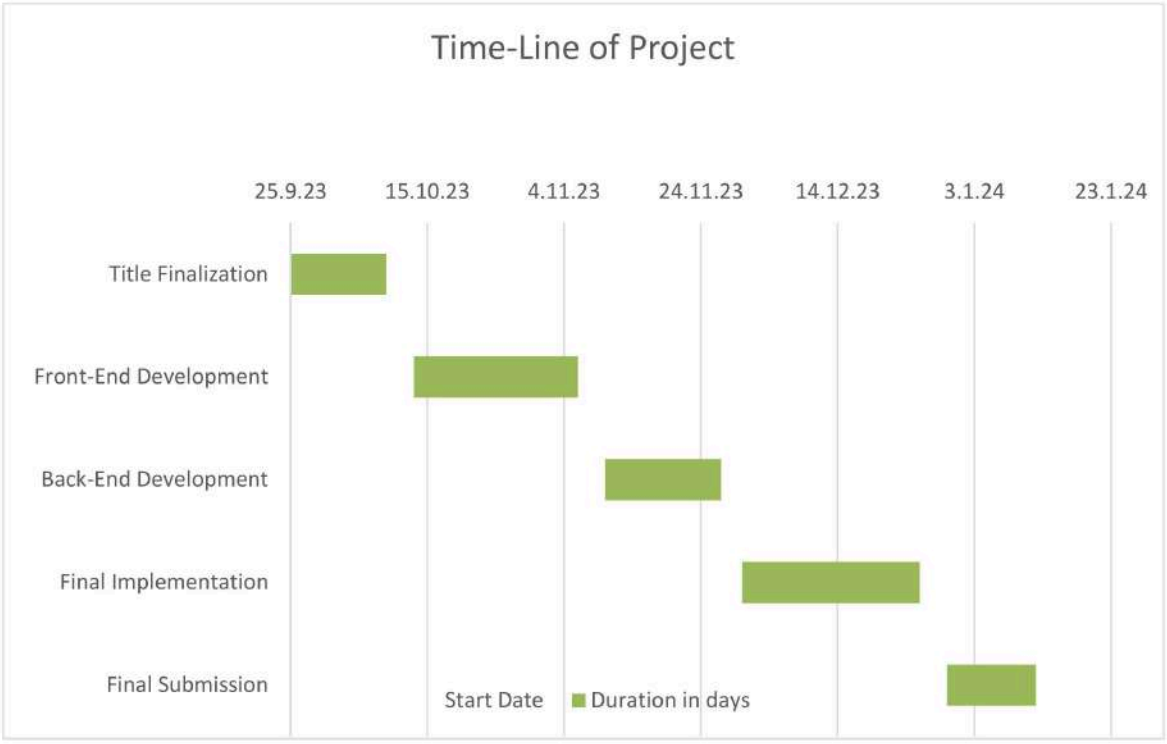
- **Source Code Details:** Document your source code, including explanations and comments.
- **50% Implementation Details with Live Demo:** Implement at least half of your project and present a live demonstration.
- **50% Report Softcopy:** Submit a softcopy of the project report with 50% completion.

❖ **Review 3**

- **Source Code Details**: Continue documenting and explaining your source code.
- **50% Implementation Details with Live Demo**: Complete 75% of the project implementation and demonstrate it live.
- **75% Report Softcopy**: Submit a softcopy of the project report with 75% completion.
- **75% Report Softcopy**: Submit an updated softcopy of the project report with 75% of the content.

❖ **Final submission**

- **Final Report and Submission of Project**: Complete your project, finalize the report, and submit the final project and report to your supervisor or institution.
- Complete implementation details with a live demo.
- Submitting the final report to the reviewer.


Time-Line of Project
25.9.23
15.10.23
4.11.23
24.11.23
14.12.23
3.1.24
23.1.24
Title Finalization
Front-End Development
Back-End Development
Final Implementation
Final Submission
Start Date
Duration in days

CHAPTER-8

OUTCOMES

ARCHITECTURE

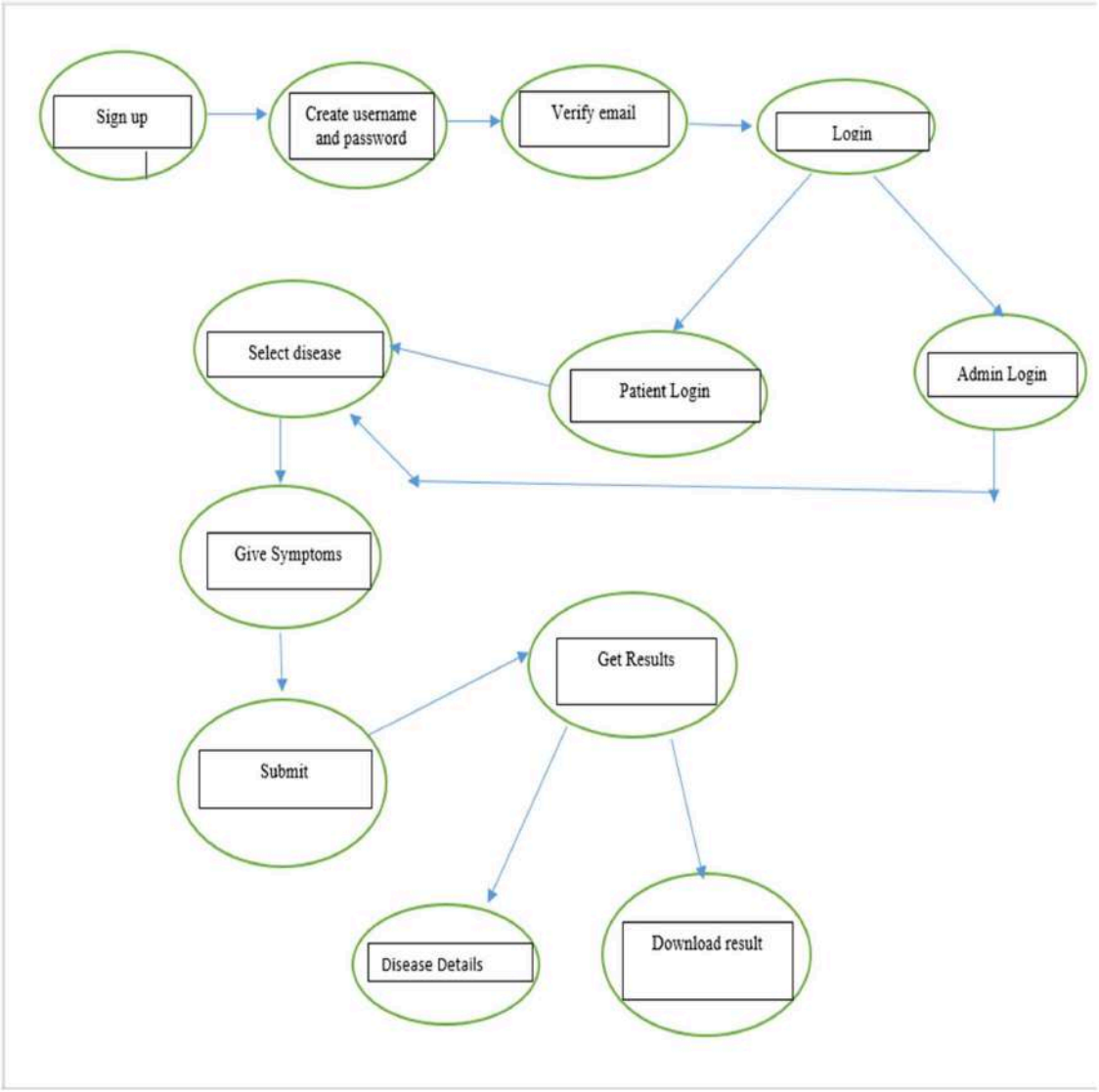


Fig 8.1 Risk Assess Architecture

CLASS DIAGRAM

In programming, a class graph inside the Bound together Displaying Language (UML) fills in as a static construction outline, depicting a framework's design. This chart outwardly addresses the framework's classes, including their qualities, tasks (or strategies), and the connections existing among them.

Basically, it gives bits of knowledge into how data is exemplified inside each class, offering a thorough comprehension of the framework's inward design and the collaborations between various parts. The class outline is an essential device in UML, working with a reasonable portrayal of the structure blocks and associations inside a product frame.

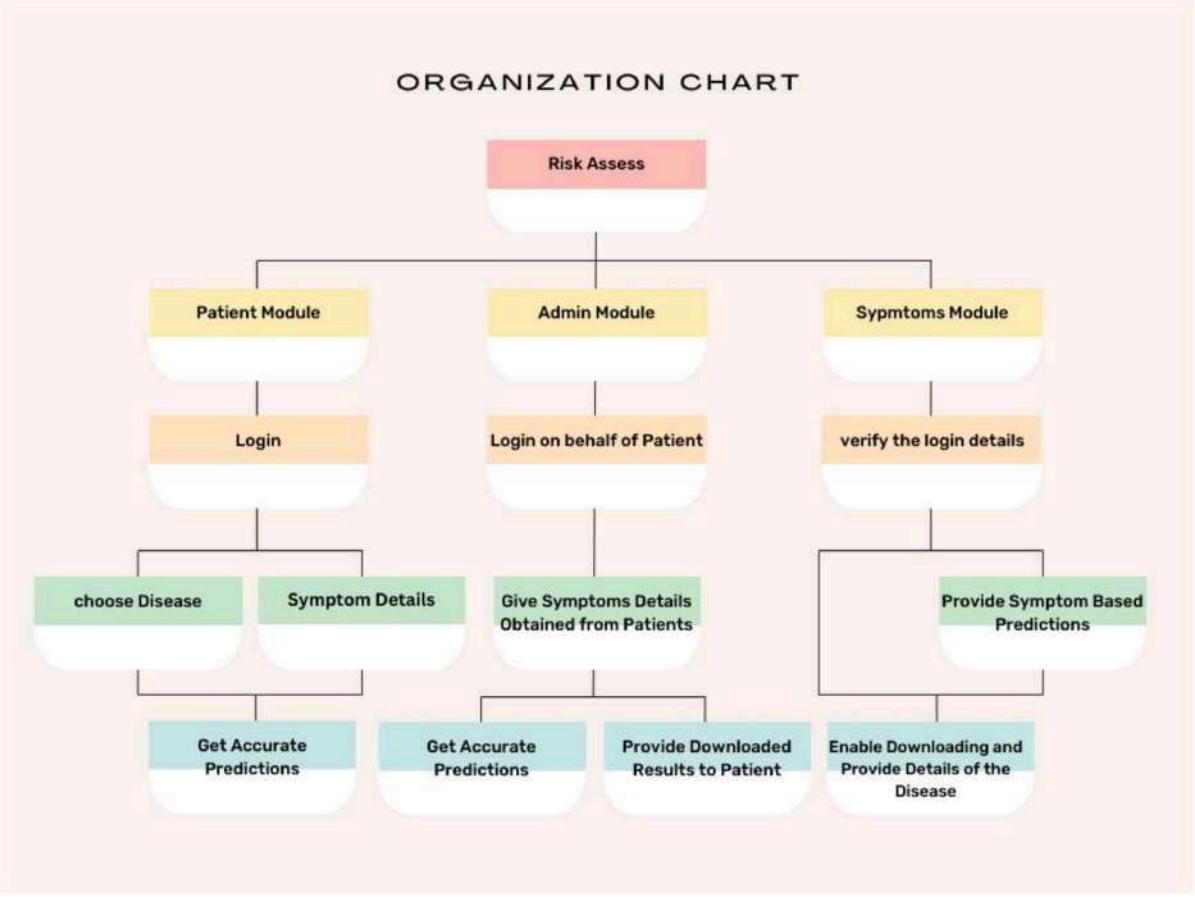


Fig 8.2 Class diagram

CHAPTER-9

RESULTS AND DISCUSSIONS

The execution of our clinical benefits application has accomplished critical achievements. The application empowers clients to add patient and expert profiles, input patient side effects, and uses Machine Learning models for exact illness expectations. These functionalities give an exhaustive answer for infection risk evaluation and the board.

9.1 Disease-Specific Predictions

- **Cancer Expectation Results**: The disease expectation model has shown high exactness in recognizing potential dangers in light of patient information, offering important bits of knowledge for early mediation and therapy arranging.
- **Diabetes Expectation Results:** Our diabetes expectation module has shown promising outcomes in determining diabetes chances, enabling proactive diabetes the board and avoidance methodologies.
- **Heart Sickness Expectation Results:** The coronary illness forecast model has areas of strength for displayed, supporting the ID of in danger patients and working with designated preventive measures.
- **Kidney Infection Expectation Results:** Our kidney sickness forecast model has actually recognized early indications of kidney illnesses, empowering convenient intercessions to alleviate gambles and work on quiet results.
- **Liver Illness Expectation Results:** The liver infection forecast model has given solid expectations, adding to a superior comprehension of liver-related conditions and their administration.

9.2 Comparative Analysis

- **Comparison with Existing Techniques**: Our models have been contrasted and existing strategies, exhibiting cutthroat execution and featuring their viability in sickness forecast.

- **Strengths and Impediments:** While our models show assets in prescient exactness and convenience, we recognize the requirement for continuous refinement to address specific restrictions, like information changeability.

9.3 Discussion of Findings

- **Interpretation of Results**: The outcomes mean the capability of our application to change infection expectation and medical services conveyance by utilizing Machine Learning for customized risk appraisal.
- **Clinical Importance**: Our prescient models can be flawlessly incorporated into clinical work on, giving medical services suppliers significant bits of knowledge for custom fitted patient consideration.
- **Future Bearings**: Future improvements might zero in on refining the calculations, extending the dataset for more extensive pertinence, and coordinating extra elements to additional upgrade the application's capacities.

9.4 User Interface and Experience:

- Based on user feedback and usability testing, evaluate the overall user experience and the design of the user interface.
- Talk about any elements that users find particularly appealing, design enhancements that could be made and potential future developments.

In end, the fruitful execution of our clinical benefits application has satisfied its essential targets. The application's capacity to precisely anticipate illness gambles with in view of patient information, joined with its easy-to-use connection point and potential for future enhancements, positions it as a critical progression in medical care innovation, with the possibility to work on quiet results and medical care conveyance.

CHAPTER-10

CONCLUSION

In this imaginative medical care project, we have effectively carried out an easy-to-understand site for the assortment of patient information, which is then used by our Machine Learning models for sickness expectation. All through this venture, we have focused on client commitment, moral information rehearses, and straightforward correspondence to guarantee the dependability and respectability of our outcomes.

By cultivating coordinated effort with medical care suppliers, we have added to the improvement of vigorous Machine Learning models that have shown promising outcomes in anticipating sicknesses like cancer prediction, diabetes, heart disease, kidney illness, and liver sickness. These models can possibly alter infection risk appraisal and early intercession, consequently working on quiet results and medical care conveyance.

The versatile framework plan of our application guarantees flexibility to developing datasets and supported client investment. This versatility is critical for obliging the developing necessities of medical care suppliers and patients, as well concerning consolidating progressions in clinical exploration and innovation.

By and large, this undertaking addresses a groundbreaking way to deal with medical services research, utilizing innovation to enable people and advance sickness conclusion and the executives. By tackling the force of Machine Learning and moral information rehearses, we have established the groundwork for another time of customized and proactive medical services that puts the prosperity of patients at the cutting edge.

CHAPTER-11
FUTURE ENHANCEMENT

There is a lot of expansion to cultivate this errand further. The application can limit with multispecialty crisis facilities on an enrollment base and license its essential consideration doctors to use the application which will help in better organization and genuine relationship of the data. During the discussion, it was moreover prescribed to integrate a phase that interfaces Patients with crisis facility Day Cares. By adding these features, the application will transform into a one-stop stage for every one of patient's requirements

11.1Joint effort with Multispecialty Emergency clinics

- **Membership-Based Admittance**: The application can team up with multispecialty emergency clinics on a participation premise, permitting their essential consideration doctors to use the application for better information the executives and patient consideration coordination.
- **Improved Administration**: This joint effort will smooth out the organization of patient information and clinical records, guaranteeing consistent correspondence between medical services suppliers and improving by and large understanding consideration.

11.2 Joining with Medical clinic Day Cares

- **Connecting Patients with Day Cares**: The application can incorporate a component that interfaces patients with clinic day cares, working with simple admittance to particular consideration administrations.
- **Comprehensive Patient Platform**: By coordinating this component, the application will turn into an exhaustive stage for every single patient necessity, including prescient illness investigation and admittance to specific consideration offices.

11. 3 Upgraded Client Experience

- **User-Accommodating Connection point**: Ceaseless enhancements to the UI and client experience will guarantee that the application stays instinctive and simple to use for the two patients and medical care suppliers.
- **Personalized Wellbeing Experiences**: The application can give customized 23 wellbeing bits of knowledge in light of the patient's clinical history and prescient models, enabling people to arrive at informed conclusions about their wellbeing.

11.4 High level Investigation and Announcing

- **Advanced Information Examination**: Consolidating progressed investigation capacities will improve the application's capacity to process and break down huge volumes of patient information, prompting more precise forecasts and experiences.
- **Customizable Reports**: The application can offer adaptable detailing highlights, permitting medical services suppliers to create custom fitted reports in light of explicit patient socioeconomics or ailments.

11. 5 Telemedicine Reconciliation

- **Telemedicine Administrations:** Reconciliation with telemedicine administrations will empower far off counsels and virtual medical care conveyance, extending the application's range and availability.
- **Remote Checking:** The application can incorporate remote observing abilities, permitting medical services suppliers to screen patients' wellbeing from a distance and intercede when fundamental.

11. 6 Ceaseless Innovative work

- **Research Organizations**: Cooperation with research foundations and colleges will work with progressing innovative work endeavors, guaranteeing that the application stays at the front of medical care development.
- **Stay Refreshed with Most recent Patterns**: Nonstop checking of industry patterns and headways in medical services innovation will drive future improvements and updates to the application.

REFERENCES

[1] R. Hai et al., "Amalur: Data Integration Meets Machine Learning,"

[2] S. K. Jena, P. Sahu and S. Mishra, "Dynamic Data Mining for Multidimensional Data Based on Machine Learning Algorithms," 2021

[3] Y. Roh, G. Heo and S. E. Whang, "A Survey on Data Collection for Machine Learning: A Big Data - Machine Learning Integration Perspective, April 2021.

[4] C. Zheng, M. Zhang, Y. Wang and M. Zou, "Application of PRA and Machine Learning Algorithm in Budget Data Acquisition and Processing System," 2022.

[5] G. Sulakshana and G. R. Kamatam, "A Comprehensive study on Distinct Data Collection Techniques using Machine Learning Approaches for WSN," 2021

[6] Y. Zhou, W. Yang and X. Deng, "Research Status of Unbalanced Data Classification Based on Machine Learning," 2022

[7] Y. Hong and X. Shao, "Emotional Analysis of Clothing Product Reviews Based on Machine Learning," 2021

[8] G. Kiarie, J. Kabi, L. Mugambi and C. W. Maina, "The use of Open-Source Boards for Data Collection and Machine Learning in Remote Deployments,"

[9] Y. Liu and X. Gou, "A multi-model Imputation Approach for Missing Data Based on Machine Learning," 2022

APPENDIX

APPENDIX-A

PSEUDOCODE

Code for login verification

```
app = Flask(__name__)
app.config['SECRET_KEY'] = 'secret'
app.config['SQLALCHEMY_DATABASE_URI'] =
'sqlite:///database.db' bootstrap = Bootstrap(app) db = SQLAlchemy(app)
login_manager = LoginManager() login_manager.init_app(app)
login_manager.login_view = 'login'
class User(UserMixin, db.Model):     id =
db.Column(db.Integer, primary_key=True)
username = db.Column(db.String(15), unique=True)
email = db.Column(db.String(50), unique=True)
password = db.Column(db.String(80))
@login_manager.user_loader def
load_user(user_id):     return
User.query.get(int(user_id))
class LoginForm(FlaskForm):     username = StringField('Username',
validators=[InputRequired(), Length(min=4, max=15)])
   password = PasswordField('Password', validators=[InputRequired(),
Length(min=8, max=80)])     remember = BooleanField('remember me')
 class RegisterForm(FlaskForm):     email = StringField('Email',
validators=[InputRequired(), Email(message='Invalid email'), Length(max=50)])
   username = StringField('Username', validators=[InputRequired(), Length(min=4,
max=15)])
password = PasswordField('Password', validators=[InputRequired(), Length(min=8,
max=80)])
@app.route('/') def index():     return
render_template("index.html")
@app.route('/about') def about():
return render_template("about.html")
```

```
@app.route('/help') def help():     return
render_template("help.html")
@app.route('/terms') def terms():
return render_template("tc.html")
@app.route('/login',
methods=['GET', 'POST']) def
login():
  form = LoginForm()     if form.validate_on_submit():         user =
User.query.filter_by(username=form.username.data).first()         if user:
if check_password_hash(user.password, form.password.data):
login_user(user, remember=form.remember.data)                 return
redirect(url_for('dashboard'))
return render_template("login.html", form=form)
    return render_template("login.html", form=form)
@app.route('/signup', methods=['GET', 'POST'])
def signup():     form = RegisterForm()     if
form.validate_on_submit():
        hashed_password = generate_password_hash(form.password.data,
 method='sha256')
        new_user = User(username=form.username.data, email=form.email.data,
password=hashed_password)
        db.session.add(new_user)
        db.session.commit()
  return redirect("/login")
   return render_template('signup.html', form=form)
```

Code for appending Machine Learning Models

```
@app.route("/dashboard")
@login_required def
dashboard():
   return render_template("dashboard.html")

@app.route("/disindex")

def disindex():

   return render_template("disindex.html")

@app.route("/cancer")
@login_required def
cancer():
   return render_template("cancer.html")

@app.route("/diabetes")
@login_required def
diabetes():
   return render_template("diabetes.html")

 @app.route("/heart") @login_required def heart():

   return render_template("heart.html") @app.route("/kidney") @login_required def
kidney():
return render_template("kidney.html")

def ValuePredictor(to_predict_list, size):

   to_predict = np.array(to_predict_list).reshape(1, size)
if size == 7:
      loaded_model = joblib.load('kidney_model.pkl')
result = loaded_model.predict(to_predict)     return
result[0]
@app.route("/predictkidney",  methods=['GET', 'POST'])
def predictkidney():     if request.method == "POST":
      to_predict_list = request.form.to_dict()
```

```
to_predict_list = list(to_predict_list.values())
to_predict_list = list(map(float, to_predict_list))        if
len(to_predict_list) == 7:
        result = ValuePredictor(to_predict_list, 7)
if(int(result) == 1):
    prediction = "Patient has a high risk of Kidney Disease, please consult your doctor
immediately"    else:
    prediction = "Patient has a low risk of Kidney Disease"
  return render_template("kidney_result.html", prediction_text=prediction)
@app.route("/liver")
@login_required def
liver():
  return render_template("liver.html")
def ValuePred(to_predict_list, size):
  to_predict = np.array(to_predict_list).reshape(1,size)
if(size==7):
    loaded_model = joblib.load('liver_model.pkl')
result = loaded_model.predict(to_predict)    return
result[0]
```

Code for Predictions

```
@app.route('/predict', methods=['POST']) def
predict():
  input_features = [int(x) for x in request.form.values()]    features_value =
[np.array(input_features)]    features_name = ['clump_thickness',
'uniform_cell_size', 'uniform_cell_shape',
'marginal_adhesion', 'single_epithelial_size', 'bare_nuclei', 'bland_chromatin',
'normal_nucleoli', 'mitoses']    df = pd.DataFrame(features_value, columns=features_name)
output = model.predict(df)    if output == 4:
    res_val = "a high risk of Breast Cancer"    else:
    res_val = "a low risk of Breast Cancer"
```

```
return render_template('cancer_result.html', prediction_text='Patient has {}'.format(res_val))

df1 = pd.read_csv('diabetes.csv')

# Renaming DiabetesPedigreeFunction as DPF

df1 = df1.rename(columns={'DiabetesPedigreeFunction': 'DPF'})

# Replacing the 0 values from ['Glucose','BloodPressure','SkinThickness','Insulin','BMI'] by

NaN

df_copy = df1.copy(deep=True) df_copy[['Glucose', 'BloodPressure', 'SkinThickness',
'Insulin', 'BMI']] = df_copy[['Glucose', 'BloodPressure',

 'SkinThickness', 'Insulin',
'BMI']].replace(0, np.NaN)

# Replacing NaN value by mean, median depending upon distribution
df_copy['Glucose'].fillna(df_copy['Glucose'].mean(), inplace=True)
df_copy['BloodPressure'].fillna(df_copy['BloodPressure'].mean(), inplace=True)
df_copy['SkinThickness'].fillna(df_copy['SkinThickness'].median(), inplace=True)
df_copy['Insulin'].fillna(df_copy['Insulin'].median(), inplace=True)
df_copy['BMI'].fillna(df_copy['BMI'].median(), inplace=True)

 # Model Building

X = df1.drop(columns='Outcome') y = df1['Outcome']

X_train, X_test, y_train, y_test = train_test_split(X, y, test_size=0.20, random_state=0)

# Creating Random Forest Model

classifier = RandomForestClassifier(n_estimators=20) classifier.fit(X_train, y_train)

# Creating a pickle file for the classifier filename = 'diabetes-prediction-rfc-model.pkl'
pickle.dump(classifier, open(filename, 'wb'))

@app.route('/predictt', methods=['POST']) def predictt():

if request.method == 'POST':

  preg = request.form['pregnancies']        glucose = request.form['glucose']

bp = request.form['bloodpressure']        st = request.form['skinthickness']        insulin =
request.form['insulin']        bmi = request.form['bmi']        dpf = request.form['dpf']        age
```

```
= request.form['age']

data = np.array([[preg, glucose, bp, st, insulin, bmi, dpf, age]])        my_prediction =
classifier.predict(data)

return render_template('diab_result.html', prediction=my_prediction)

@app.route('/predictheart', methods=['POST']) def predictheart():

   input_features = [float(x) for x in request.form.values()]     features_value =
[np.array(input_features)]

   features_name = ["age", "trestbps", "chol", "thalach", "oldpeak", "sex_0",                    "
sex_1", "cp_0", "cp_1", "cp_2", "cp_3", "  fbs_0",

                "restecg_0", "restecg_1", "restecg_2", "exang_0", "exang_1",

                "slope_0", "slope_1", "slope_2", "ca_0", "ca_1", "ca_2", "thal_1",

                "thal_2", "thal_3"]

df = pd.DataFrame(features_value, columns=features_name)     output = model1.predict(df)

if output == 1:

res_val = "a high risk of Heart Disease"     else:

 res_val = "a low risk of Heart Disease"

return render_template('heart_result.html', prediction_text='Patient has

{}'.format(res_val)) if __name__ == "__main__":

   app.run(port=5000,debug=True)
```

APPENDIX-B
SCREENSHOTS

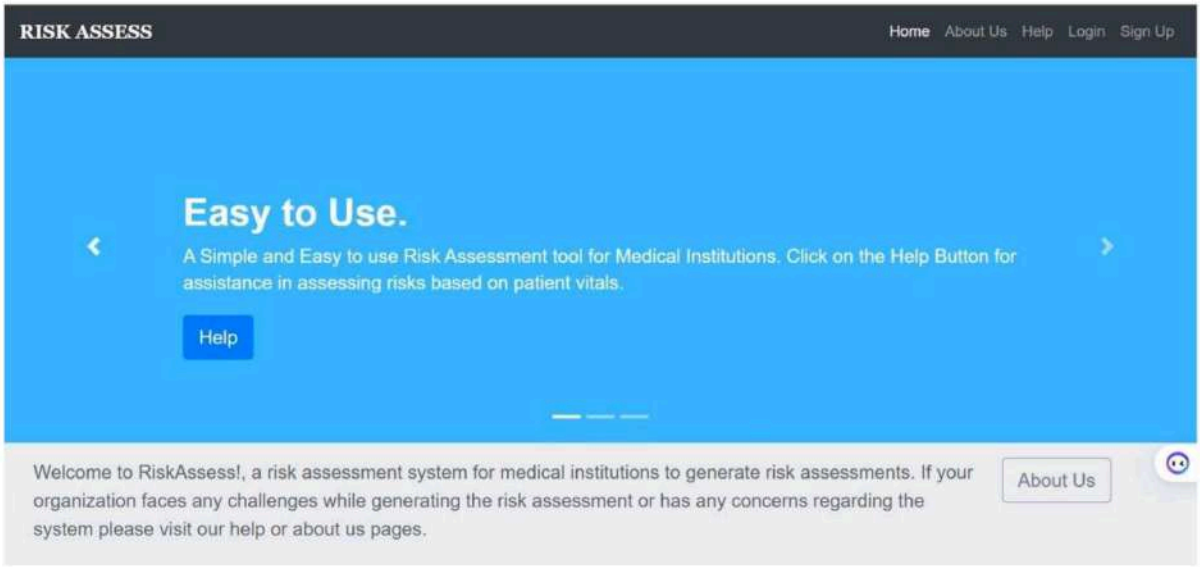


Figure B.1: Home page of the Application

Welcome to our Risk Assess Simple and Easy to use Risk Assessment tool for Medical Institutions. From the home page you can navigate to the Login and Sign-up page or to the about us page to know more.

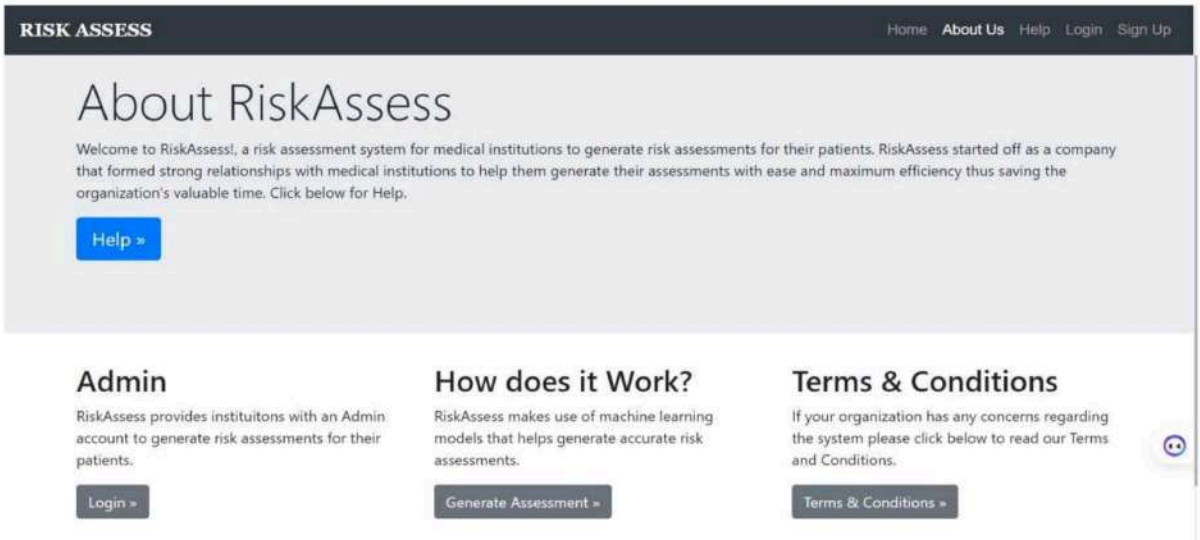


Figure B.2: About US

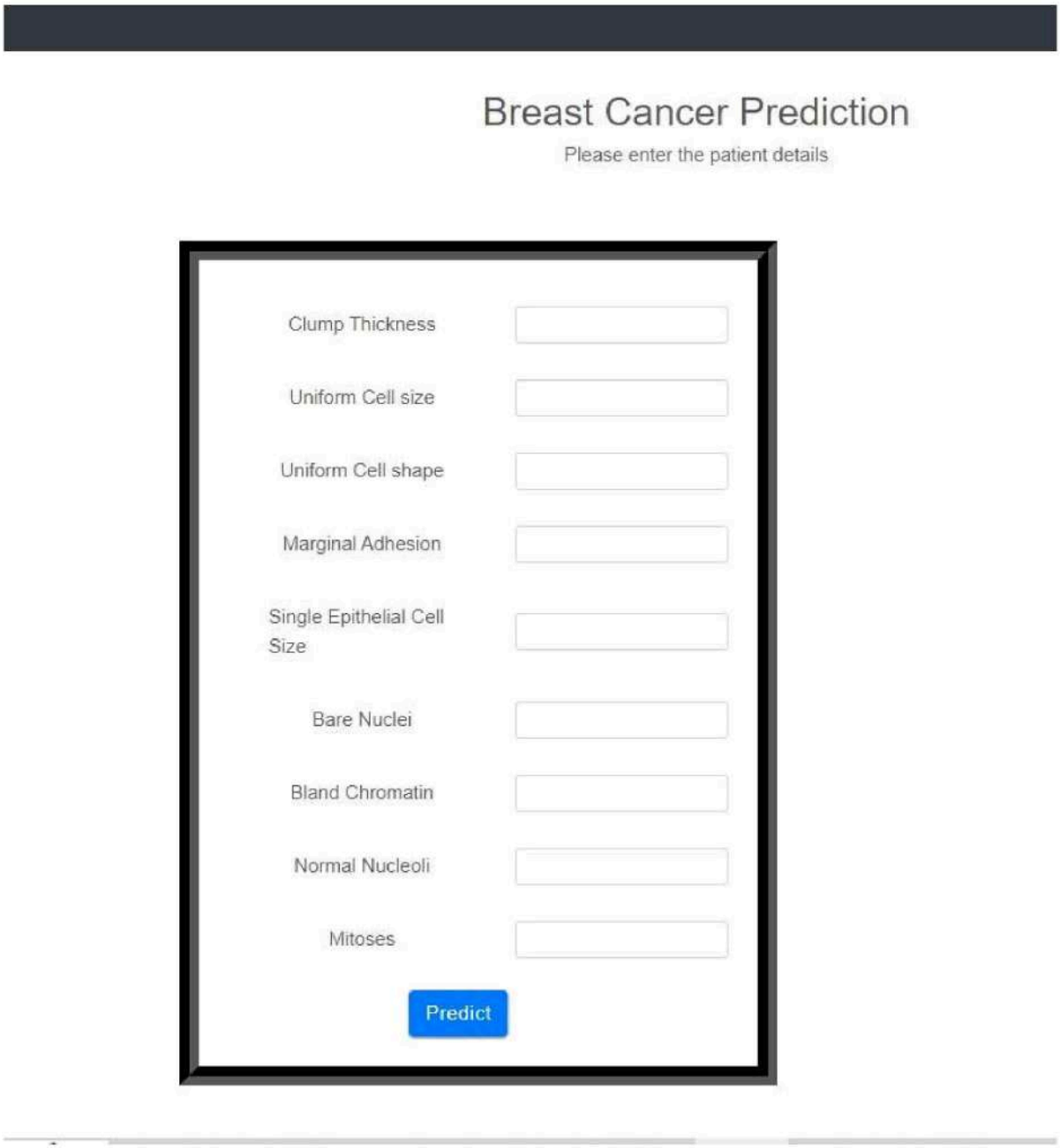


Figure B.3: Patient Details

Here Patient can enter their symptom details in numeric values which will be used by the machine learning algorithm to make an accurate Assessment.

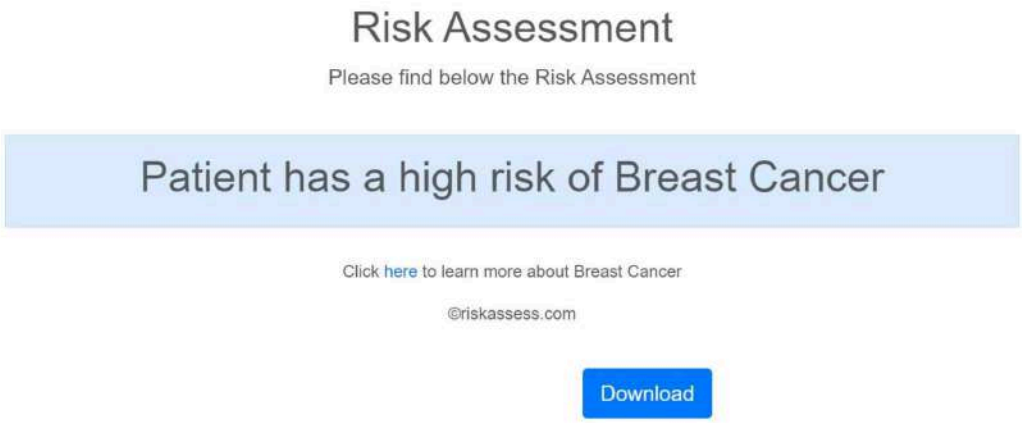


Figure B. 4: Results

Once the Assessment is complete the user is redirected to the output page which gives the prediction based on the given inputs. Patients can view their results and also download it for further requirements.

APPENDIX-C

ENCLOSURES

Conference Paper Presented Certificates.

ISSN: 2582-3930

ACCEPTANCE CERTIFICATE

Impact Factor: 8.176
DOI Prefix: 10.55041

INTERNATIONAL JOURNAL OF SCIENTIFIC RESEARCH IN ENGINEERING & MANAGEMENT (IJSREM)

An Open Access Scholarly Journal || Index in major Databases & Metadata

We are pleased to inform you that your Research Paper titled

Risk Assess: A Symptom-Based Disease Predictor

has been ACCEPTED for publication in

INTERNATIONAL JOURNAL OF SCIENTIFIC RESEARCH IN ENGINEERING & MANAGEMENT (IJSREM)

VOLUME 08 ISSUE 01 JANUARY - 2024

We are delighted to see your commitment & hardwork to share your research is being recognized. We look foreward to helping you with all of your publication needs. Thank you for choosing IJSREM!!

Crossref CiteFactor

Editor-in-chief
IJSREM

www.ijsrem.com ijsremjournal@gmail.com

Similarity Index / Plagiarism Check report

Arun-Report

ORIGINALITY REPORT

10%	7%	6%	6%
SIMILARITY INDEX	INTERNET SOURCES	PUBLICATIONS	STUDENT PAPERS

PRIMARY SOURCES

1	Submitted to Presidency University Student Paper	4%
2	Rihan Hai, Christos Koutras, Andra Ionescu, Ziyu Li, Wenbo Sun, Jessie van Schijndel, Yan Kang, Asterios Katsifodimos. "Amalur: Data Integration Meets Machine Learning", 2023 IEEE 39th International Conference on Data Engineering (ICDE), 2023 Publication	1%
3	repository.dkut.ac.ke:8080 Internet Source	<1%
4	export.arxiv.org Internet Source	<1%
5	Submitted to University of Liverpool Student Paper	<1%
6	arxiv.org Internet Source	<1%
7	Prakhar Agrawal, TF Michael Raj, Prachi Kumari. "Energy efficient routing in wireless sensor networks", 2022 2nd International	<1%

	Conference on Advance Computing and Innovative Technologies in Engineering (ICACITE), 2022 Publication	
8	Submitted to Wilkes University Student Paper	<1 %
9	acikerisim.aydin.edu.tr Internet Source	<1 %
10	pnrjournal.com Internet Source	<1 %
11	www.kaggle.com Internet Source	<1 %
12	Guduri Sulakshana, Govardhan Reddy Kamatam. "A Comprehensive study on Distinct Data Collection Techniques using Machine Learning Approaches for WSN", 2021 5th Conference on Information and Communication Technology (CICT), 2021 Publication	<1 %
13	Chenhong Zheng, Mengqian Zhang, Ying Wang, Meihua Zou. "Application of PRA and Machine Learning Algorithm in Budget Data Acquisition and Processing System", 2022 5th Asia Conference on Machine Learning and Computing (ACMLC), 2022 Publication	<1 %
	Submitted to Gateway Business College	

14	Student Paper	<1%
15	ieeexplore.ieee.org Internet Source	<1%
16	www.aui.ma Internet Source	<1%
17	www.hyaenidae.org Internet Source	<1%
18	ouci.dntb.gov.ua Internet Source	<1%
19	Yuji Roh, Geon Heo, Steven Euijong Whang. "A Survey on Data Collection for Machine Learning: A Big Data - AI Integration Perspective", IEEE Transactions on Knowledge and Data Engineering, 2019 Publication	<1%
20	1library.net Internet Source	<1%
21	api.research-repository.uwa.edu.au Internet Source	<1%
22	dspace02.jaist.ac.jp Internet Source	<1%
23	Gabriel Kiarie, Jason Kabi, Lorna Mugambi, Ciira Wa Maina. "The use of Open-Source Boards for Data Collection and Machine	<1%

Learning in Remote Deployments", 2023 IEEE AFRICON, 2023
Publication

Exclude quotes	Off	Exclude matches	Off
Exclude bibliography	Off		

Sustainable Development Goals

The "**RISK ASSESS : A SYMPTOM-BASED DISEASE PREDICTOR**" project aligns with the United Nations Sustainable Development Goal - SDG 3, "**Good Health and Well-being**," . It directly contributes to the goal of ensuring healthy lives and promoting well-being for all at all ages. By focusing on disease prediction and early intervention through machine learning and data analytics, our project aims to improve healthcare outcomes and empower individuals to take charge of their health.