# Multiple Disease Prediction System using Machine Learning and Flask web Framework Application

Challa Varshitha[1]
Dept. of Computer Science And Engineering
Parul University
Vadodara, Gujarat -391760

***Abstract*— Multi-health prognosis plays a critical role in early disease detection, personalized healthcare, and preventive medicine. This project focuses on predicting the presence or absence of specific diseases, including Diabetes, Breast Cancer, Heart conditions, Kidney disorders, and Liver diseases. Users can input their health data through forms, and the system utilizes machine learning algorithms such as Random Forest and Support Vector Machines (SVM) to generate predictions. While traditional methods often rely on individual symptom analysis, they may face challenges like false positives due to overlapping symptoms or noisy data. To overcome these limitations, we propose a comprehensive framework that analyzes the broader context of a user's health data, capturing key patterns and traits associated with each disease. Predictions are presented as clear diagnostic outcomes, enabling users to understand their health status effectively.**

**To enhance the model's performance, we incorporate a data augmentation technique designed to diversify the training dataset, improving both training speed and prediction accuracy. Additionally, we introduce a streamlined prediction network called the Dual Correlation Attention-Guided Prognosis Detector, which efficiently identifies potential health issues in a single step. Once a prediction is made, the system provides personalized recommendations, including dietary guidelines, exercise routines, and lifestyle adjustments, tailored to the user's predicted condition. This approach aims to deliver accurate, efficient, and actionable health insights, empowering users to take proactive steps toward better health management.**

**Keywords: health prediction, multi-disease prognosis, deep learning, machine learning, personalized healthcare, Random Forest, Support Vector Machines (SVM), data augmentation, Dual Correlation Attention-Guided Prognosis Detector, early disease detection, dietary recommendations, exercise planning, lifestyle modifications.**

## I. Introduction

Health prediction is the process of identifying and analyzing potential health conditions or diseases based on an individual's medical data over time. It plays a crucial role in healthcare, where the goal is to extract relevant health insights for specific diseases of interest while filtering out irrelevant or noisy data. This capability is essential for early disease diagnosis, personalized treatment planning, and preventive care. The availability of large-scale health datasets, including electronic health records (EHRs), medical imaging, and patient-reported data, has enabled researchers to monitor health trends, predict disease risks, and evaluate treatment outcomes across different timeframes. This makes health prediction one of the most significant challenges in modern healthcare.

With advancements in medical technology and data collection methods, the quality and detail of health data have improved significantly. However, the increased volume and complexity of data also introduce challenges such as noise, missing values, and data imbalance, which complicate the prediction process. Despite these challenges, high-quality data presents new opportunities for more accurate and detailed health predictions, enabling better decision-making in areas such as disease prevention, treatment optimization, and patient management.

Health prediction techniques are generally classified based on the granularity of analysis, with symptom-level, disease-level, and multi-disease-level approaches being the most common. Symptom-level methods focus on analyzing individual symptoms or biomarkers to predict specific health conditions, often generating binary outcomes to indicate the presence or absence of a disease. Disease-level approaches detect or classify specific diseases by extracting features from patient data, such as medical history, lab results, or imaging scans. Multi-disease-level methods analyze broader health patterns to predict the risk of multiple diseases simultaneously, providing a comprehensive view of an individual's health status.

Each of these techniques offers unique advantages and challenges. Symptom-level analysis provides detailed insights into specific health indicators but may be sensitive to noise or irrelevant data. Disease-level methods offer a more structured approach by focusing on meaningful health conditions, while multi-disease-level techniques can capture broader health trends but may require more complex data integration. As health prediction tasks continue to evolve, the development of advanced techniques that can handle diverse data types, reduce noise, and provide accurate results remains a key focus of research.

In this project, we explore recent advancements in health prediction methodologies, particularly focusing on machine learning and deep learning-based approaches that leverage multi-disease datasets. By examining the strengths and limitations of existing methods, we aim to highlight the potential of modern techniques to address challenges such as data quality, feature extraction, and prediction accuracy in healthcare applications. Our goal is to build a robust system that predicts the risk of diseases like Diabetes, Breast Cancer, Heart conditions, Kidney disorders, and Liver diseases, while providing personalized recommendations for diet, exercise, and lifestyle modifications to improve patient outcomes.

## II. LITERATURE REVIEW

Machine learning plays a fundamental role in our multi-health prognosis system, enabling accurate disease prediction and personalized healthcare recommendations. Traditional diagnostic methods often rely on symptom-based assessments, which can lead to false positives or overlooked conditions due to overlapping symptoms. Our approach overcomes these limitations by leveraging machine learning models like Random Forest and Support Vector Machines (SVM) to analyze complex health data and detect potential diseases with high accuracy. These models are particularly effective at handling structured medical datasets, extracting meaningful patterns, and reducing the impact of noisy or incomplete information.

A major challenge in health prediction is the variability and imbalance in medical data. Health records often contain missing values, inconsistent measurements, and imbalanced class distributions, which can affect model performance. To mitigate these issues, we incorporate data augmentation techniques that enhance dataset diversity, improving both training stability and predictive accuracy. Feature engineering plays a crucial role in this process, as selecting the most relevant health indicators—such as blood glucose levels, cholesterol, kidney function metrics, and liver enzyme levels helps refine the decision-making process of the models.

The Dual Correlation Attention-Guided Prognosis Detector introduced in this project optimizes predictions by capturing relationships between different health indicators, ensuring a comprehensive analysis of user data. By integrating feature correlation and attention mechanisms, this system enhances detection accuracy while maintaining efficiency. Unlike rule-based diagnostic systems, which rely on predefined medical knowledge, our machine learning approach dynamically adapts to new data, making it scalable and robust for evolving healthcare needs.

Beyond diagnosis, our project extends to personalized health recommendations. Once a prediction is made, the system generates tailored dietary plans, exercise routines, and lifestyle modifications based on the detected condition. This personalized approach empowers users to take proactive steps toward improving their health rather than merely reacting to medical diagnoses.

In summary, our multi-health prognosis system harnesses the power of machine learning to provide accurate disease predictions, mitigate the limitations of traditional diagnostic methods, and offer actionable health insights. By leveraging advanced models, feature selection techniques, and data augmentation, we create a scalable and efficient solution for early disease detection and preventive healthcare. Future research could explore integrating ensemble learning methods, semi-supervised techniques, or explainable AI models to further enhance transparency and trust in automated health prediction systems.

## III. METHODOLOGY

This section outlines the methodology followed in the development of the Change Detection in Remote Sensing Images system, detailing the technology stack, system architecture, frontend and backend implementation, database design, and testing procedures.

### *A. Technology Stack*

The Change Detection in Remote Sensing Images system employs a modern tech stack designed to ensure scalability, efficiency, and a seamless user experience. The technology stack comprises:

**Frontend**: FLASK, HTML,CSS (integrated templates)
**Backend**: FLASK , PKL Files
**DataScience Key Skills** : Machinr Learning, DeepLearning
**Tools**: MongoDB, Git, Vscode

### *B. System Architecture*

The architecture of the Multiple Disease Prediction system is structured to enable seamless integration between the frontend, backend, and Machine Learning,deep learning model. It ensures efficient data flow across the system, facilitating real-time processing of user input in forms and predicts outputs. The architecture is designed to support scalability, allowing the system to handle increasing volumes of image data and users without sacrificing performance.

### *C. Frontend Implementation*

The frontend of the Change Detection system is developed using Flask's templating engine,HTML,CSS which allows for smooth integration with the backend. Key features include:

**Register Page:** The Register page allows users to give details of user to access their accounts by entering their credentials in login page.


fig 1 Register Page

**Login Page:** The login page allows users to securely access their accounts by entering their credentials.


fig 2 Login Page

**User Interface (UI)**: The UI is designed for user-friendliness, enabling users to upload image pairs, view detected changes, and manage results efficiently.

**Home Page**: The home page serves as the main entry point, providing an overview of the system and easy navigation to other features.

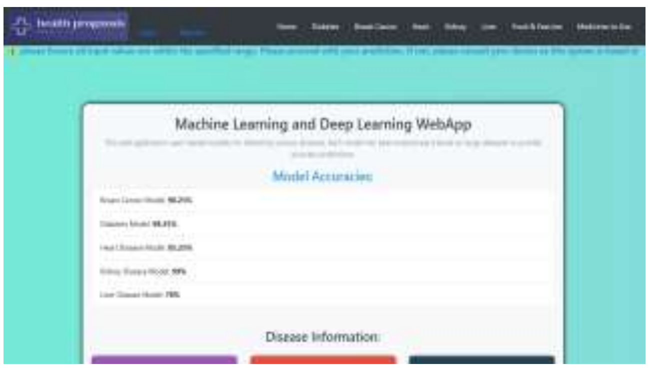


fig 3 Home Page1


fig 4 Home Page2


Fig 5 Home Page3

**Prediction Pages :** The prediction page allows users to give input values pairs and view the results after submit the predict button in all pages of the health prediction process.


fig 6 Diabetes Predictor Interface


fig 7 Breast Cancer Predictor Interface


fig 8 Heart Disease Predictor Interface


fig 9 Kidney Disease Predictor Interface


fig 10  Liver Disease Predictor Interface

**Results Page :** This page tells whether user has disease or not  if disease is predicted then it navigates to food ,exercise and medicine recommendation pages as below.

fig 11 Results Interface

**Food and Exercise Recommendation Page:**

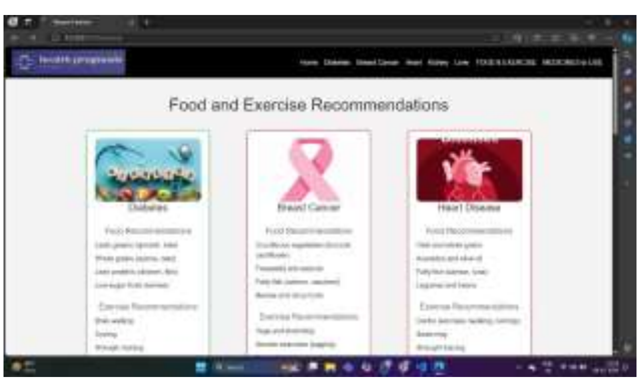


fig 12 Food and Exercise Recommendation Interface

**Medicine Recommendation Page:**

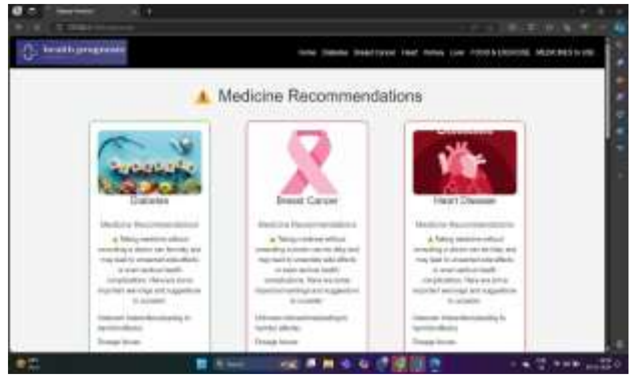


fig 12 Medicine Recommendation Interface

### *D. Backend Implementation*

The backend is powered by Flask, responsible for managing operations, API route handling, and interaction with the Machinelearning model and ML Algo's for detection. Key functionalities include:

**API Routes**: The backend includes RESTful API routes and Pickle Files which stores the input fields that handle requests related touser input uploads, health prediction processing, and result retrieval.

**Machine Learning Model Integration**: The backend interacts with the model to process input  pairs, perform prediction , and generate the result.

**Database Interaction**: It uses MangoDB's ORM with an database to perform CRUD operations for managing user data, image history, and change detection results.

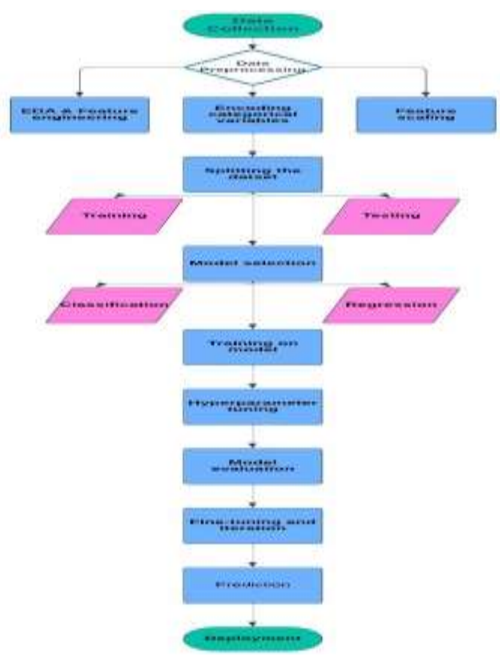


fig 13 system architecture


fig 6 class diagram

## IV. SYSTEM DESIGN

### A. *UML Diagrams*

**1)Use Case Diagram:**

The use case of the Multiple Disease Prediction System using Machine Learning and Flask web Framework Application outlines the key functionalities and interactions between users and the system. Users can perform various actions such as uploading input pairs, initiating change detection processing, and viewing the resulting difference inputs. The diagram effectively captures how users interact with the system, providing a clear view of user requirements and core functionalities.

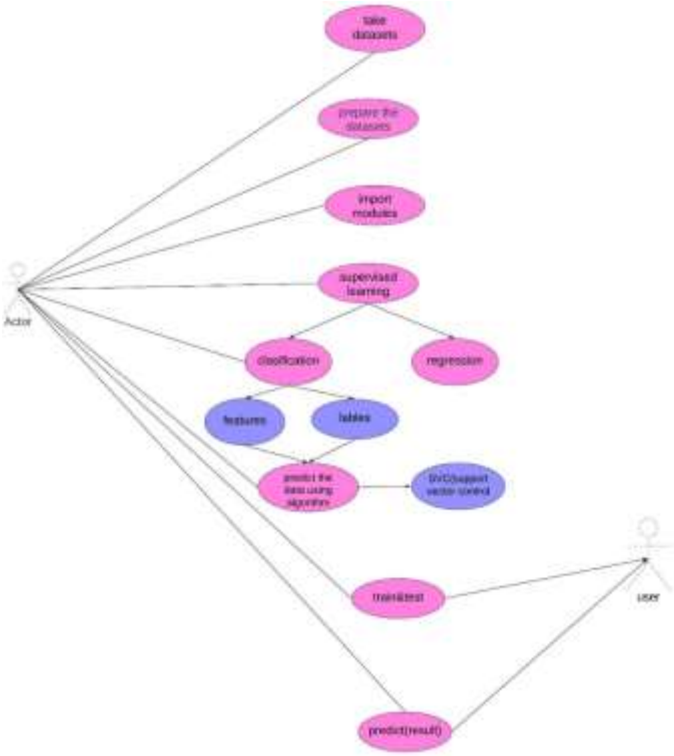


fig 14 use case diagram

### B. *Class Diagram:*

A class diagram in the Unified Modeling Language (UML) is a static structure diagram in software engineering that outlines the system's classes, attributes, operations, and relationships. In the Change Detection in Remote Sensing Images system, the class diagram represents key components such as the user interface, image handling, deep learning model, and database interactions. It captures the relationships between classes, such as how users interact with the system, how images are processed by the CNN model

### C. *Sequence Diagram:*

Sequence diagrams in UML, also known as event diagrams or timing diagrams, are interaction diagrams that illustrate the order and flow of operations. In the Change Detection in Remote Sensing Images system, the sequence diagram demonstrates the process flow from when a user uploads image pairs to when the CNN model processes these images and returns the detected changes. The diagram outlines the interactions between the user interface, backend API, deep learning model, and the SQLite database.

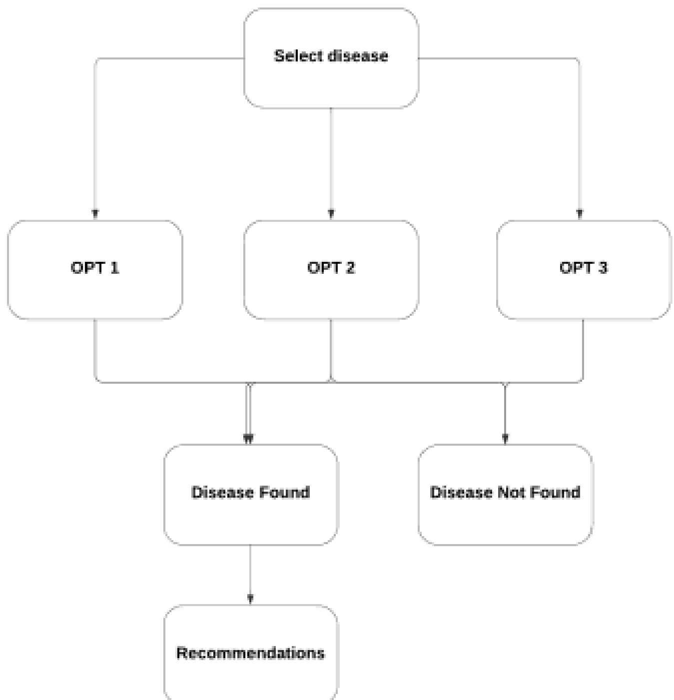


fig 7 sequence diagram

### D. *ER Diagram:*

An Entity-Relationship (ER) diagram describes a database's structure using entity sets and relationship sets, illustrating the logical relationships among tables and their attributes. In the Change Detection in Remote Sensing Images system

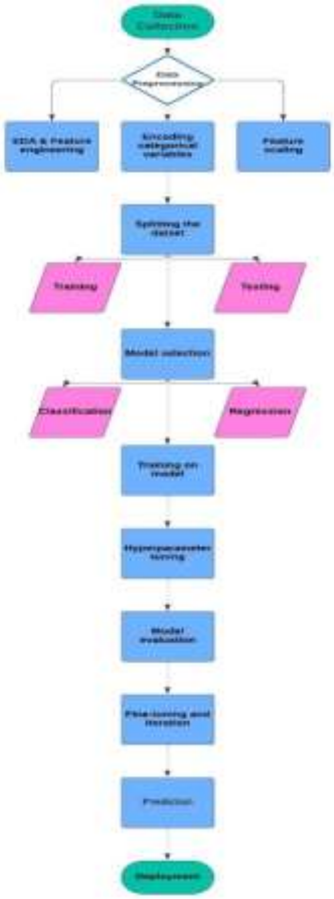


fig 8 ER diagram

## V. CONCLUSION

In conclusion, object-level health prediction focuses on identifying and analyzing changes in an individual's health status over time using advanced techniques such as deep learning and machine learning. This approach aims to detect specific health conditions or risks, such as Diabetes, Breast Cancer, Heart diseases, Kidney disorders, and Liver diseases, by analyzing patterns in medical data. Object-level health prediction has a wide range of applications across various fields, including personalized healthcare, preventive medicine, and treatment optimization. By identifying changes in health indicators and predicting potential risks, this approach can enhance safety, improve efficiency, and support informed decision-making for patients and healthcare providers. As a rapidly evolving field, object-level health prediction holds significant potential to transform industries and professions, making it a critical area for ongoing research and development. By leveraging advanced algorithms and data-driven insights, this approach can contribute to better health outcomes and more effective healthcare systems.

## ACKNOWLEDGMENT

I WOULD LIKE TO EXPRESS MY SINCERE GRATITUDE TO MY GUIDE, ASSISTANT PROFESSOR KIRAN MACWAN , ARUNESH PRATHAP SINGH, FOR HIS INVALUABLE GUIDANCE AND SUPPORT THROUGHOUT THIS RESEARCH PROJECT. HIS EXPERTISE AND CONSTRUCTIVE FEEDBACK WERE INSTRUMENTAL IN SHAPING MY WORK.

I ALSO WISH TO THANK MY TEAM MEMBERS FOR THEIR COLLABORATION AND DEDICATION IN THE DEVELOPMENT OF MULTIPLE DISEASE PREDICTION SYSTEM USING MACHINE LEARNING AND FLASK WEB FRAMEWORK APPLICATION OUR TEAMWORK MADE THIS PROJECT BOTH ENJOYABLE AND SUCCESSFUL.

## REFERENCES

[1] M. Yang, "Object Recognition," in *Encyclopedia of Database Systems*, pp. 1936-1939, 2009..
[2] A. Memo, "Review of Machine Learning in an Artistic Context," *Medium*, 2016.
[3] Y. LeCun, Y. Bengio, and G. Hinton, "Deep Learning," *Nature*, vol. 521, no. 7553, pp. 436-444, 2015.
[4] D. C. Cirean, U. Meier, J. Masci, L. M. Gambardella, and J. Schmidhuber, "Flexible, High Performance Convolutional Neural Networks for Image Classification," 2011.
[5] S. Saha, F. Bovolo, and L. Bruzzone, “Destroyed-buildings detection from VHR SAR images using deep features,” Proc. SPIE, vol. 10789, Oct. 2018, Art. no. 107890Z.
[6] F. Bovolo, L. Bruzzone, and S. Marchesi, “A multiscale technique for reducing registration noise in change detection on multitemporal VHR images,” in Proc. Int. Workshop Anal. Multi-temporal Remote Sens. Images, Jul. 2007, pp. 1–6.
[7] P. Zhang, M. Gong, L. Su, J. Liu, and Z. Li, “Change detection based on deep feature representation and mapping transformation for multi-spatial-resolution remote sensing images,” ISPRS J. Photogramm. Remote Sens., vol. 116, pp. 24–41, Jun. 2016.
[8] J. Geng, J. Fan, H. Wang, X. Ma, B. Li, and F. Chen, “High-resolution SAR image classification via deep convolutional autoencoders,” IEEE Trans. Geosci. Remote Sens., vol. 12, no. 11, pp. 2351–2355, Nov. 2015.
[9] A. Pomente, M. Picchiani, and F. Del Frate, “Sentinel-2 change detection based on deep features,” in Proc. IEEE Int. Geosci. Remote Sens. Symp. (IGARSS), Jul. 2018, pp. 6859–6862.
[10] N. Otsu, “A threshold selection method from gray-level histograms,” IEEE Trans. Syst., Man, Cybern., vol. SMC-9, no. 1, pp. 62–66, Jan. 1979.
[11] Chaurasia A, Culurciello E. LinkNet: Exploiting encoder representations for efficient semantic segmentation[C]// 2017 IEEE Visual Communications and Image Processing (VCIP). IEEE, 2018.
[12] Wang Q, Zhang X, Chen G, et al. Change detection based on Faster R-CNN for high-resolution remote sensing

images[J]. Remote Sensing Letters, 2018, 9(10-12):923-932.

[13] T. A. Łabuz, "Environmental impacts-coastal erosion and coastline changes," In Second Assessment of Climate Change for the Baltic Sea Basin, pp. 381-396, 2015, doi: 10.1007/978-3-319-16006-1_20.

[14] W. Xin, T. Can, W. Wei and L. Ji, "Change Detection of Water Resources via Remote Sensing: An LV-NSCT Approach," Applied Sciences, vol. 9, no. 6, 2019, Art. No. 1223, doi: 10.3390/app9061223.

[15] J. P. Mondejar and A. F. Tongco, "Near infrared band of Landsat 8 as water index: A case study around Cordova and Lapu-Lapu City, Cebu, Philippines," Sustainable Environment Research, vol. 29, no. 1, 2019, Art. No. 16, doi: 10.1186/s42834-019-0016-5.

[16] W. Chen, Y. Wang, X. Li, Y. Zou, Y. Liao and J. Yang, "Land use/land cover change and driving effects of water environment system in Dunhuang Basin, northwestern China," Environmental earth sciences, vol. 75, no. 12, 1027, 2016, doi: 10.1007/s12665-016-5809-9.

[17] S. J. Russell and P. Norvig, Artificial Intelligence: A Modern Approach, vol. 9, no. 2.

[18] S. Majumdar, Deep Columnar Convolutional Neural Network, vol. 145, no. 12, pp. 2532, 2016.

[19] B. Coppin, Artificial Intelligence Illuminated. 2008.

[20] S. J. Russell and P. Norvig, Artificial Intelligence: A Modern Approach, vol. 9, no. 2