AI Aarogya: A Dual-Stage AI Framework for Preventive Health Guidance

Jainam Abhay Shah
PG - Research Scholar - MSc(CSIT)
Department of Computer Science & IT
Jain (Deemed-To-Be-University),
Bangalore, India

Soumyajit Sarkar
PG - Research Scholar - MCA(AI ML)
Department of Computer Science & IT
Jain (Deemed-To-Be-University),
Bangalore, India

Salam Shibani Devi
PG - Research Scholar – MA(English)
Department of Language [English]
Jain (Deemed-To-Be-University),
Bangalore, India

Shriram. S
PG - Research Scholar – MCA(AI ML)
Department of Computer Science & IT
Jain (Deemed-To-Be-University),
Bangalore, India

Dr. Meera Kumar
Professor
Department of Languages [English]
Jain (Deemed-To-Be-University),
Bangalore, India

Abstract*: The lack of easy access to timely and accurate preventive health care guidance continues to be a challenge for people who do not live at their own homes; are working in environments that are unfamiliar to them; or live in rural and/or underserved areas [1], [2]. When people have to wait for appropriate recommendations, they often experience increased anxiety, make poor decisions, and see deterioration in their health status [1]. Presently, digital health solutions are generally based on static symptom checkers and rule-based systems without taking into consideration the context, history of the patient, and providing real-time conversation with the patient [2], [6]. There have been promising developments in utilizing large language models for building conversational health care systems; however, these systems frequently suffer from issues of hallucinations, inconsistencies, and lack of reliability when used in sensitive areas of medical practice [3]–[5]. This report on AI Aarogya presents its foundation as a hybrid multi-stage system for preventive healthcare, addressing the current limitations of existing methods by incorporating machine learning technology, medical-domain language models, and safety measures through deterministic processes [7]–[10]. Aarogya consists of a three-tier probabilistic disease prediction model using the Platt-calibrated Random Forest and a medical domain language model, MedGemma, to extract symptoms from the record of the patient, as well as provide contextual reasoning [9], [11]. The Fusion Decision Engine employs a rules-based approach to make the best output from both models more consistent, while limiting the opportunity for hallucinations [4], [12]. Also, the system provides for the storage of patients' longitudinal medical histories using Firestore, and provides a location-based specialty physician reference. The experimental data collected from this study indicate more consistent and lower variability in the output, as well as responsiveness in real time through both models. Overall, the Aarogya framework is a scalable, dependable, and safe option for producing personalized prevention of healthcare, and is intended to provide assistance to professional's medical judgements, rather than replacing them.

Keywords: Hybrid AI Healthcare Systems, Preventive Health Guidance, Platt-Calibrated Random Forest, Fusion Decision Engine (FDE), Symptom-Based Disease Prediction, Longitudinal Patient History Modeling, Context-Aware Conversational AI, Location-Aware Doctor Recommendation, Safety-Constrained AI Systems.

1. INTRODUCTION

1.1 Background Information

Timely and reliable assistance in making informed health decisions is still a global issue, especially for those living in remote locations (geographic isolation), working in stressful environments, or living alone (without a primary care support system) [1], [2]. Often, delays in obtaining medical advice early on in one's illness lead to increased anxiety, stress, and poor decision-making [1]. Furthermore, not having access to immediate, available, and contextualised resources as soon as symptoms appear reinforces the need for intelligent systems that can provide early, preventive assistance before getting access to a professional (i.e. a doctor) [2].

The personalized nature of traditional digital health platforms, such as symptom checkers and rule-based advising platforms, has inherent limitations [2], [6]. They are often unable to account for the patient's specific context (e.g., past medical history, changing symptom pattern, and existing real-world situations). As a result, these digital platforms produce static and generic outcomes, typically listing out several potential diagnoses without readiness to offer structured directives or actionable items [6]. Furthermore, these digital platforms are also not conversationally aware; therefore, the user's ability to refine their question and receive an adaptive answer is limited, making them less effective in changing healthcare environments [2].

Recent Advances AI, Machine Learning have greatly improved ability to model complex relationships between symptoms and diseases [7]–[10]. Ensemble Learning methods like Random Forest classifiers activated by probability calibration have demonstrated consistent performance on disease prediction tasks [9], [10]. Additionally, Domain Specific LLMs allow systems to process free-form text, perform contextual reasoning and produce human-like responses [3]. As individual technologies, they are often used in isolation leading to issues such as poor interpretability of the machine learning models and hallucination risk in the language models [4], [5].

Modern researchers are concentrating on hybrid architectures, where structured predictive modelling and contextual reasoning are combined using strict safety constraints to overcome the limitations of previous AI architectures [11]–[14]. The AI-Aarogya preventive health guidance system has been developed within this paradigm as an advanced preventive health guidance application that combines three tiers of Platt-calibrated machine learning technologies (MLT) with the medical-domain language model MedGemma, longitudinal patient history stored in Firestore and specialist-based recommendations by city. The deterministic Fusion Decision Engine and strict safety filter ensure the generation of a safe and reliable output through a unified Framework of Preventive Guidance, which will provide a cohesive and conflict-free preventive health guidance ecosystem to complement, rather than replace, the professional diagnosis.

1.2 Problem Statement

While there has been an explosion in the number of online symptom checker tools and digital health platforms, most of the current offerings show major deficits in being able to provide individuals with tailored, context sensitive, and actionable information about their health [2], [6]. Most of the existing platforms use static rules-based systems to generate generic results, providing users with multiple possible conditions without any explanation about each possible condition, how they differ or how they should be ranked in terms of likelihood of being the underlying cause of their symptoms [6]. Therefore, many users end up with an impression of confusion rather than clarity [2].

In addition to being limited in terms of the types of data used to generate results (e.g. do not provide users with a longitudinal record of their unique patient history, variable symptom patterns and real world factors such as where the user is located), current platforms do not provide users with preventive information, including recommendations regarding precautions that should be taken, dietary recommendations, home remedies or early warning signs.

Improvements to machine learning and technology for processing natural language have made great strides in improving the accuracy of disease forecasting and conversational capabilities [3], [7]. However, most current solutions use both types of technology in isolation, leading to limitations in interpretability and usability, as well as producing unsafe outputs (e.g., incorrect or “hallucination”) [4], [5]. Therefore, there is a strong need for an integrated framework that combines probabilistic forecasting with contextual reasoning and deterministic control while maintaining constraints for user safety [11], [12]. AI-Aarogya provides a solution by using a hybrid and multi-stage architecture that utilizes Platt-calibrated Random Forest models, MedGemma-based reasoning, a Fusion Decision Engine (R1–R5), longitudinal patient records in Firestore, and city-specific specialist recommendations within an integrated AI-Aarogya framework. Together, these technologies create a system that provides personalized and preventative healthcare advice, without compromising the quality of the professional medical diagnosis.

1.3 Motivation

The explosion in popularity of mobile devices and internet-based services for health care has generated many chances for scalable digital solutions for health; this is especially true for preventive health care services [1]. The awareness of how to identify early symptoms of diseases, the effectiveness of making healthier lifestyle choices, and monitoring your health regularly, has increased the demand from consumers for easy and trusted sources of health information [2]. For example, people need to find reliable and accurate information about their health in urgent situations, such as when they have health questions at night, have an injury or minor illness, or have difficulty seeing a physician right away. Therefore, there is a need to have intelligent systems that can evaluate unstructured data in the form of a person describing their symptoms and provide an organized and clear system of preventive health care guidance to eliminate confusion, improve decision-making, and support better health outcomes.

In recent months, there has been enormous progress in the fields of artificial intelligence, calibrated machine learning models and language models used in the field of medicine [3], [7]. This has created new technologies that can assist healthcare workers with accurate and useful data, and also provide healthcare workers with a way to create a healthcare assistant capable of making reliable predictive assessments of a patient, along with generating human-like responses to inquiries made by that patient in a contextual way, and combining these capabilities with the real-world needs of healthcare by creating AI Aarogya, a hybrid predictive health advisory service that integrates probabilistic disease predictions; contextual reasoning; longitudinal patient history using Firestore; and, a deterministic safety device to ensure delivery of accurate and preventive services in a scalable and user-friendly manner.

1.4 Objectives of the Study

- The aim is to evaluate and assess different AI based health systems and disease prediction tools; to understand how they work, their advantages and disadvantages in real world applications [7]–[10].
- To report any missing research gaps in today's digital health; especially regarding a lack of personalization, context awareness, no records of a patients long term history; and no ability to interactively converse with the patient [2].
- In order to create a preventative hybrid health care framework that uses random forest models calibrated to the Platt scale, we will need to build a language model of the medical domain called MedGemma [9], [11].
- Establishing a safety-constrained system providing only preventive guidance through strict filtering mechanisms preventing diagnostic, prescriptive, and prognostic information from being included in output [4], [5].
- To incorporate longitudinal patient history utilizing Firestore to maintain session continuity and enable personalized recommendations across multiple interactions.
- To evaluate how well this system can help users identify their symptoms and reduce their anxiety about their symptoms, learn about good preventive health practices, and provide easy access to decision-support tools when they are not in a clinical setting [1].

1.5 Objectives of the Study

The objective of this research is to create, implement and assess a hybrid artificial intelligence preventive health care solution using the existing digital health platform limitations as a focus [2], [6]. The study will analyze and compare commonly used symptom-based platforms and how these platforms have failed to provide personalized, contextually relevant and conversationally compatible health decision support [2]. As a means of improving upon the current limitations of existing platforms, the AI Aarogya will utilize a multi-level architecture consisting of Platt-calibrated Random Forest models for probabilistic disease prediction together with MEDGEMMA (a Medical Domain Language Model) for symptom extraction and reasoning and natural language interaction from the user [9], [11]. This system will also retrieve and leverage information from the user's longitudinal medical history through Firestore Structured Knowledge Retrieval, in conjunction with a deterministic fusion engine to provide consistent, reliable and non-contradictory outputs [12].

This study is to look into preventive health advice and decision-making assistance (refer to section II for the research subject) in non-clinical settings. The system is designed to provide rule-based evidence-based support through a structured approach for each category of recommendation (i.e., recommendation categories) while not providing a diagnosis, treatment, or clinical prediction. In addition, the system will utilize external search data to deliver city-based, specialist-based recommendations, thus separating these two areas from the medical reasoning pipeline. The focus of the framework is to be highly scalable and accessible for human beings to be able to interact safely with AI (the system) to increase awareness of health prior to seeing a physician; the goal is not to replace seeing a physician.

2. LITERATURE REVIEW

Machine learning and artificial intelligence have been previously verified to be essential in identifying disease through symptom prediction and decision support systems within healthcare. Researchers have utilized routines in their analyses of complex medical datasets including the Random Forest method, Support Vector Machines, and ensemble koop designed numerous web-based and mobile-based applications for users to enter their symptoms to gain initial insight about their health. Additionally, recent developments in deep learning, explainable AI, and multi-modal technology will improve both accuracy and ability for doctors to interpret the results. However, current AI systems typically lack in terms of providing either contextual reasoning or personalization for an individual user, or providing real-time communication with that user continues to demonstrate that AI Aarogya provides a hybrid model of using calibrated machine learning, medical domain specific language model semantic data, along with deterministic fusion as a reliable preventative healthcare steering mechanism.

Kommineni Sai Kumar, Singamaneni Rajesh, Maganti Sai Sathya, and Abdul Nadeem [6] – Disease Prediction Based on Symptoms using Database and GUI

A machine learning framework has been presented in this paper and it has created a disease prediction platform that can utilize patient-reported symptoms to predict if an individual has a disease. The machine learning system was trained using a data set of hospital records and the Random Forest algorithm was chosen to increase the accuracy of classification of illness. The model is a database of the predicted illnesses and their symptoms which can later be used for research. The model was developed as a Python application with a Tkinter graphical user interface which allows individuals report their symptoms and the model will provide health predictions in an easy to access format.

Mu. Krishna Raj, Julia Punitha Malardhas, and I. Devapriya [7] - Machine Learning Approach to Predict Multiple Diseases Based on Symptoms

This paper describes the development of an innovative platform based on machine learning methodology which aims to predict different kinds of diseases: diabetes, heart attack, cancer (all types), tuberculosis, arthritis, lung disease, stroke, asthma, brain tumor, osteoporosis, COPD (chronic obstructive pulmonary disease), oral disease, etc.. This research utilized Random Forest as the machine learning algorithm. It was created off of a large data set of patient records, symptom reporting and clinical outcomes. With that data set, there were many steps taken to prepare it prior to using the machine learning algorithm. The data set had missing values which were handled using various techniques for handling missing data; also, we encoded categorical variables, and normalized numerical features. These steps significantly improved the quality of the data being fed to the Random Forest. The Random Forest also went through feature selection to determine which symptoms are useful for predicting various kinds of disease, as well as hyperparameter tuning to optimize the performance (accuracy) of the algorithm.

Priya Bhardwaj, A. Mithila, S. J. K. Jagadeesh Kumar, and G. Prabu Kanna [8] – Machine Learning Based Approaches for Livestock Symptoms and Diseases Prediction and Classification

In this research project, we used different types of machine learning methods to predict diseases in livestock by way of a symptom-based prediction methodology. Although animal health was our main focus throughout this study, the results could also be applied to general healthcare predictive systems. We utilized numerous machine learning techniques to perform both symptom-based prediction and disease classification, including Logistic Regression, Support Vector Machine, Random Forest, XGBoost, LightGBM, and CatBoost. Following a comparative analysis of these models, we found that both Random Forest and CatBoost had high accuracy rates, with 83.5% and 83.7% respectively.

Vinora A, Kannan V, Partha Sarathi R, and Murali Prasath M [9] – Machine Learning Models for Symptom Detection and Disease Prediction

The study presents a new multi-disease prediction model that utilizes advanced gradient boosting algorithms (CatBoost and LightGBM) to predict disease based on user-reported symptoms. The model is enabled through an easy-to-use web interface for users to input symptom data, which are then submitted to a trained machine learning model for accurate disease prediction. In addition to providing disease predictions, the system also provides users with information about the causes of the disease, complications to monitor, and initial steps that may be taken to improve user knowledge and facilitate decision-making. Our results also demonstrate that gradient boosting provides excellent performance on high dimensional symptom data sets.

Bakhtiyor Makhkamov, Khakimjon Zaynidinov, Danish Ather, Sonal Pathak, Temurbek Kuchkorov, and Ibrohimbek Yusupov [10] – Predicting Multiple Diseases Using Machine Learning: A Symptom-Based Model for Medical Diagnosis

This report describes a multi-class disease predictive model that uses a total of 132 symptom features for differentiating 42 unique and separate diseases. Symptom presence or absence has been organized into binary measures. The data has then been split up into training and test data sets so that we could evaluate the predictive capabilities of the models. Our research examines many machine learning methods to help alleviate the difficulties associated with high-dimensional multi-class classifications, specifically: Random Forests, Support Vector Machines (SVM), and Neural Networks. We also consider methods to reduce the dimensionality of the features, improve interpretability, and increase the accuracy of the models.

Santosh Kumar Satapathy, Soumyashree Panda, Manan Gandhi, and Vanshita Patel [11] – Symptom-Based Disease Prediction Framework Integrated with Django and Deep Learning Models

In this report, we present the creation of a disease prediction program which uses the Django web framework along with deep learning techniques to provide accurate medical diagnosis. The program uses a convolutional neural network (CNN) with a SoftMax activation function that was trained on symptoms entered by the user to predict multiple diseases. It is our opinion that we have achieved an industry standard performance rating of 97.66% in the ability to accurately predict a variety of diseases using our model. We wished to create a simple way for users to enter their symptoms into a system and obtain an immediate prediction regarding their health status. This research demonstrates the importance of having accessible and well validated models along with an easy to use design when developing reliable and effective health care systems.

Divyansh Nishad, Anshika Mishra, and Nidhi Goyal [12] – Symptom-Based Disease Prediction Using Machine Learning

The report contains an application of a machine-learning style model of predicting illness with a mobile app interface for increased access and interaction with the disease prediction tool. Multinomial Naive Bayes, Decision Tree, and Random Forest were the types of algorithms used to classify types of illness based on reported user symptom inputs. The results indicated that the model performed very well with an overall F1 score of 0.99 across several types of illnesses. All models were evaluated using cross-validation, making them valid and reliable models. In addition, our findings show that present day online symptom checkers are very limited in what they offer and therefore we are advocating for improved, more accurate and reliable systems that would enhance patients' ability to assess their own symptoms using an online tool. Through the integration of predictive analysis into a mobile application, we have created an interactive, easy-to-use product that has a very high degree of accuracy for predicting disease.

Soujanya KJ, Sneha BJ, Ashwini Kodipalli, Trupthi Rao, and Sanjeev Gosai [13] – Symptoms-Based Classification of Disease Using Various ML Algorithms and Interpretation using LIME and SHAP Kernels

This study will focus on classifying diseases (as classified by symptoms) and then employing a variety of machine learning algorithms (i.e. Support Vector Machine (SVM), k-nearest neighbour (KNN), Logistic Regression (LR), Random Forest and Decision Trees) to create an accurate prediction model. The issue of interpretability of models in the health care field is also an important aspect of this research, which will use Local Interpretable Model Agnostic Explanations (LIME) and Shapley Additive Explanations (SHAP). The approaches used in this analysis allow for a better understanding of how to make predictions based on these symptoms, thus increasing transparency and providing more explanation to the decision-making process. By identifying all of the important predictors, it will strengthen both confidence and dependability of machine learning systems in healthcare while ensuring that the models continue to classify at a high performance level.

Ravindra Kumar Prajapat, Vaibhav Agarwal, and Shashikala Tapaswi [14] – Symptoms-Based Disease Prediction and Drug Recommendation using Hybrid Machine Learning Algorithms

This study outlines the development of a hybrid machine-learning model to predict diseases and recommend pharmaceuticals and healthcare products. This model is based on 132 symptoms, nearly 5,000 individual patient reports, and 41 diseases, with a total of 41 diseases used to develop the model. The hybrid model used many different algorithms such as the Support Vector Machine, Random Forest, Gradient Boosting, and K-Nearest Neighbor. Each of these algorithms was combined using ensemble methods, allowing for an accurate prediction of each algorithm's effectiveness. Additionally, in assessing our model, we not only discovered what diseases people can have, but we also provided preventative strategies, diet recommendations, exercise plans, and medication information in a way that we believe provides a more complete picture of health care. We evaluated our model according to accuracy, precision, recall, and F1 score, which we used as evidence to support the hybrid ensemble approach described above as improving the prediction of diseases.

Pothana Hema, N. Sunny, Arunarkavalli Darbha, and Raavi Venkata Naganaanjani [15] – Disease Prediction using Symptoms based on Machine Learning Algorithms

In this research study, we introduce a disease prediction model which is based on a patient’s symptoms and employs multiple machine-learning algorithms; this modelling improves on the accuracy of prediction from the data. To assist in predicting the presence of a certain disease, we have developed a graphical user interface that will allow users to input a list of symptoms; the user’s input will require prediction to be made using each of the classifiers or techniques employed to support the prediction. A mode of the three techniques (the decisions of the three classifiers) will be used to determine the predicted disease for the patient. Additionally, we have attempted to develop a modular methodology that includes data pre-processing, training the models separately and finally, combining the models to provide the final output. The outcome of this modular methodology provides evidence that combining different classifiers improves the reliability, dependability, and overall performance of predicting diseases based on symptoms.

Swathi M, Sri Aishwariya N, Prithika N G, and Jayaram B [16] – The Disease Prediction System Based on Symptom Identification and Image Processing

This research examines a dual-method method of disease forecasting; incorporating symptom-based machine learning and image-based deep-learning methods(most likely from the medical picture analysis). It uses different algorithms for classifying symptoms using machine learning: Random Forest, Decision Trees, XGBoost, Logistic Regression, Long Short-Term Memory (LSTM), and Naive Bayes. For the analysis of the diagnostic image, DenseNet121(Convolutional Neural Network) were used in medical images(such as chest X-ray, skin rash, etc.). The Long Short-Term Memory (LSTM) model was found to perform at the highest level with a success rate of around 96.5% when comparing several models. Additionally, our application was created using the Django web framework to provide users with a highly interactive experience. This research presents a method of detecting diseases early by combining a visual analysis of images with a symptom assessment of a patient to improve the surgeries overall accuracy and reliability of health prediction systems.

Yawen Huang, Abdur Rashid Sangi, Shengyi Qi, Beining Niu, and Waqar Ali [17] – Web-Based Disease Prediction Application Using Machine Learning for Symptom Analysis
This report describes a web-based disease prediction system that employs different machine learning models like Decision Trees, Random Forests, XGBoost, and Logistic Regression to analyze symptoms and predict the occurrence of diseases. The system was developed using publicly available datasets that provide information on 42 different diseases and additional disease-specific datasets for diabetes and heart disease, COVID-19, and liver disease. We looked at accuracy, precision, recall and F1 score in terms of the data and model evaluation which XGBoost did the best in we present 99.49% accuracy which we put forth as proof of its performance in disease classification. Also, we developed the front end with Vue 3 and back end with Flask which we did in order to make the health assessment tool easy to access online and to support early health assessment and informed health decisions.

2.1 Identified Research Gaps

Lack of End-to-End Integration
Many studies available today mainly look at predicting diseases through machine learning while ignoring how to generate both reasoning about things in conversations and how to create guidance. Therefore, many of the systems designed today only provide an output of predictions and do not have a complete system or pipeline connecting symptom comprehension, reasoning, and providing prevention-based guidance in one system.

Limited Context Awareness and Personalization
Many approaches do not incorporate longitudinal patient information and/or sessions at the level of this context. As a result, they do not have the ability to provide personalized services. Presently, every input is treated separately, and there is no consideration of previous interactions or the evolution of the symptom patterns in making recommendations.

Dataset Level Evaluation Only
Many studies have shown that static datasets can be tuned for maximum accuracy; however, they lack validation with real-world data which tends to be less accurate such as noisy, incomplete, and free-form user input. Consequently, the accuracy of these systems in conversational environments has not been evaluated for their robustness.

Lack of Controlled Integration with Language Models
Language Model Usage & Risks The language model-based approaches being applied today generally rely almost exclusively upon those models and do not implement deterministic control mechanisms. Using language models in this manner can create a number of risks such as producing hallucinations, generating inconsistent outputs and making unsafe recommendations, particularly in healthcare related applications.

Absence of Safety Constrained Systems
Most existing solutions do not create strict boundaries between preventive health care guidance and clinical diagnosis. Most existing solutions do not provide filtering/safety mechanisms to process diagnostic, prescriptive or prognostic content, which is critically needed in healthcare systems.

Limited Use of Structured Knowledge Integration
Many systems lack an integrated proof-based database used to create actionable results from preventive knowledge that includes information like precautionary measures, diet recommendations, and home remedies. This limitation greatly impacts how useful predictions can be in everyday life and therefore cannot be applied correctly or easily.

No Deterministic Fusion Mechanism
There is currently no established method to combine the output of a machine learning prediction with the reasoning process of a language model in a systematic, reliable fashion. This gap in integration creates uncertainty in the results that the system produces and ultimately reduces the confidence a user has in that system.

Lack of Location Based Practical Utility
Many currently existing systems do not provide assistance based upon an individual user's (or individual users') physical location. This translates into these systems lacking any real usefulness for providing professional healthcare assistance.

2.2 Comparative Analysis of Existing Methods

Author	Year	Method	Dataset	Performance	Limitations
Kumar et al.	2022	Random Forest with GUI	Hospital symptom dataset	High accuracy	No conversational support and no contextual reasoning
Krishna Raj et al.	2024	Random Forest with	Clinical sympto	Improved predicti	Lacks real time interacti

		preprocessing and feature selection	m dataset	on accuracy	on and personalization
Bhardwaj et al.	2024	Multiple ML models including Random Forest and CatBoost	Livestock symptom dataset	Accuracy up to 83 percent	Limited applicability to human healthcare and no unified system
Vinora et al.	2024	Gradient boosting models	Web based symptom dataset	High performance	No conversational reasoning or safety control
Makhkamov et al.	2024	SVM Random Forest Neural Networks	132 symptom feature dataset	Effective classification	Limited scalability and no contextual guidance
Satapathy et al.	2024	Deep learning CNN model	Symptom dataset	Accuracy around 97 percent	Lack of interpretability and no fusion mechanism
Nishad et al.	2024	Naive Bayes Decision Tree Random Forest	Mobile app dataset	High F1 score	No longitudinal memory and limited real world adaptability
Soujanya et al.	2024	ML models with LIME and SHAP	Symptom classification dataset	Improved interpretability	No conversational system and no unified guidance
Prajapat et al.	2024	Hybrid ML with recommendations	5000 patient reports	Good accuracy	Includes prescription based outputs and lacks safety constraints
Hema et al.	2022	Ensemble ML models	Symptom dataset	Improved robustness	No conversational AI integration
Swathi et al.	2024	Multimodal ML and deep learning	Symptom and image dataset	High accuracy	Complex architecture and limited scalability
Huang et al.	2025	XGBoost and ML models	Multi disease dataset	Accuracy above 99 percent	Dataset focused evaluation and no contextual reasoning

3. METHODOLOGY

3.1 Workflow Diagram

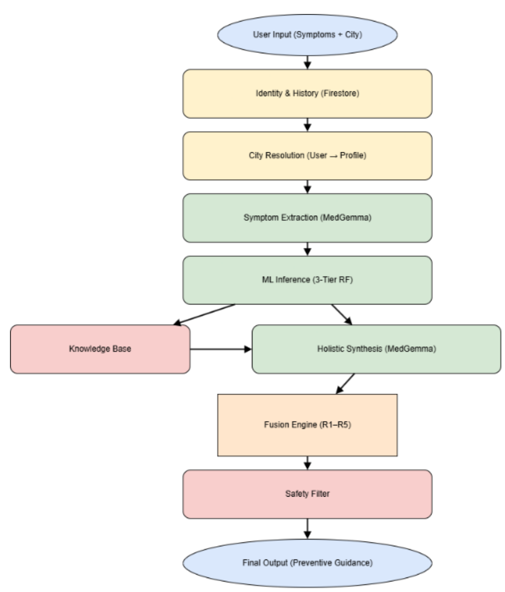

Figure 3.1: Workflow Diagram of the Proposed Aarogya AI Framework

Once a user inputs symptoms and selects a city through a web or chat interface, the operational protocol of the AI-Aarogya system is activated. Following the user input, the AI-Aarogya system first identifies the user through identity resolution of the patient metadata in Firestore, including their last five sessions, to provide longitudinal context awareness. The identification of the specified city is performed using a priority-based approach, where the system first attempts to verify the specified city provided by the user within the request, falling back to the patient profile's preferred city when verification fails.

After the text is input into MedGemma, the extraction of the natural language symptoms takes place (free-form descriptions are transformed into structured tokens that are mapped to an existing feature space) [3], [11]. The features created by this process serve as the input for the 3-tier Platt-calibrated Random Forest algorithm that generates disease predictions with associated probabilities [7]–[10]. Once the leading predictions are established, they will be further developed through the application of a preventive knowledge base, at which point MedGemma applies reasoning in order to create a complete, non-conflicting preventive guidance plan [3]. In addition, the application of this process identifies appropriate specialists in order to provide additional care options to the person being assisted.

The Deterministic Fusion Decision Engine uses the results of machine learning models as well as clinical and laboratory results to produce a rule-based arbitration (R1-R5) that determines the best data to use to produce consistent results that reduce the potential for inaccuracies [4], [12]. The final result is processed through a safety filter that removes any diagnostic or prescriptive information from it [5]. The processed result is stored in Firestore asynchronously for the purpose of providing structured preventive guidance to users. The structured preventive guidance includes precautions, dietary recommendations, home remedies, exercise recommendations, red flags, and recommendations for specialists.

3.2 System Overview

This entity, called AI Aarogya, is an intelligent hybrid preventive healthcare system that uses conversational agents to provide personalised and actionable health guidance along with disease prediction based on symptoms [1], [2]. It has been created as a lightweight and scalable software application to be accessible via web based devices, thus providing convenience to individuals without immediate access to healthcare providers. AI Aarogya differs from regular symptom assessors by offering clear, tidy, and understandable preventive healthcare information via probabilistic prediction with context to develop reliable results [2], [6].

The overall system has many components that are important, including the frontend user interface, back end processing server, machine learning inference engine, MedGemma (a medical domain language model), deterministic fusion engine (a process that merges results from various sources into one output), structured knowledge base, and safety filter. These components work in conjunction with one another to take raw inputs about symptoms and provide useful and actionable information on how to prevent illness [3], [11]. A Firestore-based data layer is also used for the purposes of storing patient profiles and tracking longitudinal history of sessions (also called longitudinal session tracking). The use of a Firestore-based data layer allows for the generation of personalized recommendations across multiple interactions (e.g., across multiple visits).

Data entering the system starts with user submissions containing a free-form description of their symptoms and a city. Using the user-submitted data as input, the backend server performs identity resolution and fetches the latest patient history from Firestore. The input is processed by MedGemma until symptom identification is performed, after which a three-tier Platt-calibrated Random Forest model produces soy probability-based disease forecasts [7]-[10]. The disease forecasts are further improved using the inclusion of a preventive knowledge database and finished by MedGemma through the use of a holistic reasoning process to develop a comprehensive guideline for the user [3]. Following the processes above, the outputs of the machine learning and language models are combined using a deterministic Fusion Decision Engine, which implements rule-based arbitration to create a fused result [4], [12]. Any restricted content is removed by a safety filter before sending the response generated above to the user [5].

Using modular architecture is one approach to separating common prediction reasoning fusion functions from safety enforcement functions through distinct layered designs. In this way, the machine-learning models can focus only on producing accurate statistical outputs, while the language model will provide context to those outputs and explain them clearly to the user [3]. The fusion engine improves the consistency and reliability of the outputs from the two previous models, and the safety layer enforces that all outputs will be purely preventive [4], [5]. This approach to modular and layered architecture greatly increases the scalability, maintainability, and flexibility of the overall system by allowing independent components to be updated or optimized without affecting the functionality of the entire system.

3.3 System Architecture

Data Collection Module

The system uses a complete symptom-disease dataset with a standardized prevention knowledge base [7]-[10]. The symptom-disease dataset has over 150,000 clinical records, and each record contains symptoms and diseases that are labeled. The system also has a prevention knowledge base that has more than 200 different diseases, and each disease has the following information: precautions to be taken; dietary recommendations; home remedies; activity recommendations; and red-flag symptoms. The data in both components of the system allows it to successfully assist with real life clinical situations that contain multiple symptoms.

Data Preprocessing Module

Module Preprocessing of Data the Symptoms Data Collected Has New Structured Data Created for Machine Learning Purposes. Preprocessing Includes:

1. Free Text is Converted to Structured Symptom Tokens using MedGemma [3]
2. The Extracted Symptoms are Then Mapped into a Binary Feature Vector That Contains Over 1,200 Different Features
3. The Input Data is Validated and Normalized to Ensure Consistent and Reliable Data.

Disease Prediction Module

Three-tiered Random Forest calibrated with a Platt model allows us to identify what type of disease each individual has based on their various symptom feature vectors (often), giving a probability-based output of which disease (of the many it could be) they are most likely to have [7]-[10]. The architecture includes:

1. A model with a high level of confidence is used to make accurate predictions.
2. A medium level of confidence in the prediction model is intended to result in balanced predictions.
3. A model with a low level of confidence will be used to provide wider coverage.

Probabilities between all models are calibrated using Platt's calibration to increase the output's accuracy of estimate(s) and provide valid results to facilitate sound decision-making [9], [10].

Knowledge Integration Module

The knowledge integration module will transfer the diseases that are predicted to a structured knowledge repository from which appropriate preventive guideline data can be retrieved [11]-[14]. The knowledge mapper uses three layers of matching to call the appropriate preventive information for each user:

1. Exact match for directly mapped diseases.
2. Partial match for approximately mapped diseases
3. General preventive advice as a fallback mechanism

The knowledge integration module assures that the user will receive actionable and structured health guidance and not simply a label of disease.

Conversational Reasoning Module

MedGemma functions as a language model designed for the medical domain to allow for the extraction of symptoms and develop a cohesive and contextually aware preventive guidance plan based on patient input, predicted disease and prior session data, determining an appropriate specialist for the condition [3].

Fusion Decision Module

The Fusion Decision Engine (FDE) merges the results of the Machine Learning Model and MedGemma Reasoning Framework in a deterministic way. It uses rule-based logic to maintain reliability, consistency and to minimize hallucinations [4], [12]. The FDE resolves conflicts and provides one final, comprehensive output that incorporates statistical outcome predictions with contextual reasoning.

Safety Filtering Module

The systematic implementation of a safety filter provides assurance that all output functions comply with the standards and specifications of preventive healthcare. A rule set is established to create the filtering component of the system, which will remove any diagnosis, prescription or prognostic content [5]. Thus, the system works exclusively as an augmentation tool, not as a replacement for licensed medical practitioners.

Location Based Specialist Module

City based recommendation module of the system is intended to provide a capable of recommending suitable relevant practitioners based on Identification of the type of specialist, either based on user input or a patient profile parameter. The module uses external search grounding to help identify practitioners relevant to the identified type of specialist and is also distinctly separate from the medical reasoning module.

System Integration Module

The application is built using a base of Flask for the backend which we have used Cloud Firestore for issue of patient profiles and session data storage. For the front end we have used Next.js which gives us an interactive user interface. We have deployed the application on the cloud for scale and real time performance. Also we have gone with a modular architecture which allows for independent updates without affecting the total system.

3.4 Model Description

Disease Prediction Model

The disease prediction system which has a three stage Platt calibrated Random Forest approach for symptom based classification [7]-[10]. We input a high dimensional symptom feature vector into the model which in turn puts out probability estimates for many diseases. Each model is trained on structured symptom data and is tuned for different confidence levels. The prediction process includes:

1. Disease class prediction.
2. Class balancing through rareness-based solutions.
3. Platt scaling for better probability outputs.

The model provides stable and easy to interpret results which we use as a base for further reasoning.

Conversational Reasoning Model

MedGemma is a language model for the medical field used in extracting reasons for symptoms and to generate guidelines for treating those symptoms through a system [3]. The model processes free-text entered by users and converts that text into structured symptom tokens. The model also performs its own clinical reasoning, using machine learning predictions, based on user input and historical session data. The output of the MedGemma model is the creation of a complete preventive guidance document that contains all of the following areas: precautions, diet, home remedies, exercise, and potential indicators of red flags. The MedGemma model also identifies an appropriate medical specialty necessary to address the symptoms listed in the user's input.

Fusion Model

The Deterministic Rule Engine which is a component of the Fusion Decision Engine takes in outputs from the machine learning model and the language model. It applies pre defined rules which in turn resolve issues of conflict and see to the task of consistency between predictions and reasonings [4], [12]. The fusion process includes:

1. Validation of which machines output between machine learning and language models.
2. Tuning confidence scores by consensus.
3. Handling of fallback scenarios when language model response is unavailable.

This module ensures that the final output is reliable consistent and free from hallucinations.

3.5 Dataset Description

The data behind the AI Aarogya system, which supports how well the system operates, consists of 150,000 different cases, which were labelled with the corresponding symptoms (case) associated with each disease (disease) [7]-[10]. These cases contain a wide variety of illnesses and combinations of symptoms to allow the AI Aarogya system to model many different situations that happen in healthcare. The symptoms are defined by using a structured vocabulary containing over 1,200 binary identifiers that identify whether a particular symptom exists or not (binary). The structured representation of the symptoms will allow the system to process through the information easily and will be compatible with the machine learning models used in the AI Aarogya system.

Every entry in the database consists of a multi-class classification challenge, in which an extensive symptom vector is used as the input to a multi-class disease label for each individual class [9]. The database identifies over 1,600 different diseases as part of the class distribution. In order to provide additional provisions for rare diseases, class balancing techniques are used when creating the database; this allows the system to identify complex correlations between symptoms and diseases, along with maintaining robustness across various and imbalanced distributions of data.

To allow for an unbiased evaluation of model effectiveness, the dataset is divided into two parts; Training data is used to generate three separate random forest classifiers [9], [10] for three different levels of confidence (high, medium, and low). To provide better probability estimates for each classifier and to improve the likelihood of predicting correctly for each classifier type, Platt calibrated all of the random forest classifiers prior to evaluating generalization performance using testing data to determine whether or not any of the classifiers overfit their respective training datasets when evaluated using new, unseen testing datasets.

The extensive scale and variety of the dataset greatly increase the system's ability to process noisy, incomplete, and differently patterned symptom data. The breadth of this dataset impacts prediction accuracy by allowing the model to learn from many different clinical patterns, thus allowing for consistent, reliable performance in real-world environments.

This dataset is the foundation of the machine learning development flow and will be used to develop a scalable, consistent and user-friendly prevention healthcare guidance system that integrates smoothly with both the reasoning capabilities of MedGemma and the fusion-based decision framework.

3.6 Implementation Details

Software Environment Implementation

1. Python programming language
2. Flask (Backend - API)
3. Google Cloud Firestore (Database, Session Storage)
4. Medgemma (Medical Reasoning, Natural Language Processing implemented via LM Studio)
5. Random Forest Models (Disease Prediction with scikit-learn)
6. Next.js (Frontend - user interface)

Hardware Configuration

This system is implemented using standard computing environments without heavy dependency on high end GPUs

1. CPU based execution for backend processing and inference.
2. 8 to 16 GB RAM for efficient handling of models and API requests.
3. Optional GPU support for faster model training and experimentation.

Training Parameters

Key training parameters for the machine learning models include

1. Optimize for performance: the number of trees in the Random Forest models (dependent on data class size)
2. Features used: Greater than 1200 binary symptom features used as a feature space.
3. Logistic regression probability calibration using Platt scaling.
4. Separate models based on tiered (high/medium/low) confidence prediction

Evaluation Metrics

Prediction Metrics

1. Accuracy for disease classification performance
2. Precision and recall for evaluating model reliability
3. F1 score for balanced performance measurement
4. Calibrated probability scores for confidence estimation

System Level Metrics

1. Response time of API for real time user interaction
2. Consistency between machine learning and language model outputs
3. Reduction in hallucination rate using fusion decision engine
4. Reliability of preventive guidance generation

Real Time Performance Metrics

1. API response latency for end-to-end processing
2. Time taken for symptom extraction reasoning and fusion
3. Scalability under multiple user requests in cloud environment

Safety and Reliability Metrics

1. Percentage of outputs filtered by safety layer
2. Enforcement of preventive only responses
3. Reduction of unsafe or misleading content
4. Consistency of outputs across repeated sessions

4. RESULTS AND DISCUSSION

4.1 Outcomes

An evaluation of the performance of the AI platform Aarogya (referred to as the AI Aarogya system) has demonstrated strong and consistent performance with regard to predictive accuracy, operational reliability, and real-time response to user requests. The hybrid architecture proposed within this project effectively combines machine learning with language model collocation reasoning in order to produce personalized preventive health care recommendations for individual users. In addition, the system utilizes a deterministic fusion mechanism and a safety-constrained output layer, thus guaranteeing maximum reliability while reducing the likelihood of generating unsafe or inconsistent results.

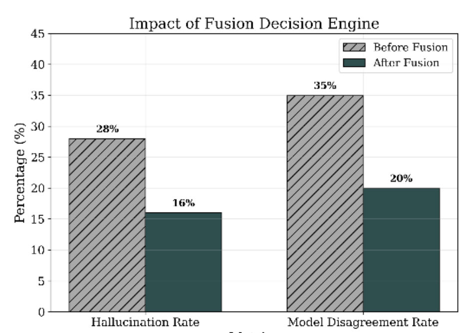

Figure 4.1 Impact of Fusion Decision Engine

Figure 4.1 illustrates how successful the Fusion Decision Engine has been at improving the reliability of the system. For example, if you looked at a sample set of data, you would see that the average rate of hallucinations has decreased from 28% to 16% and average rates of disagreement have dropped from 35% to 20% through the use of this fusion mechanism. The rule-based fusion method helps to align the predictive capabilities of machine-learning models with the logical reasoning produced by language models, helping to decrease the variability in both output types and, therefore, increase the overall reliability of the outputs generated.

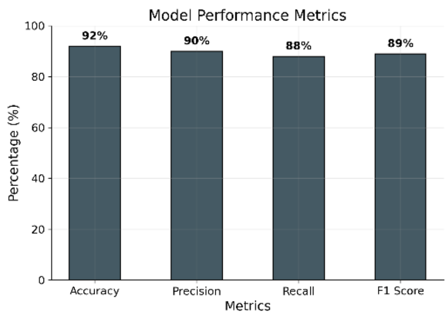

Figure 4.2 Model Performance Metrics

The model's accuracy of 92% (i.e., percent of correct predictions) illustrates how well it captures the relationships between symptoms and disease, maintaining a balanced performance across all classification metrics with a precision of 90%, recall of 88% and F1 score of 89%. As such, in addition to producing reliable results through Platt calibration, using the model for making probability estimates gives similarly reliable and interpretable predictions which would then improve subsequent reasoning.

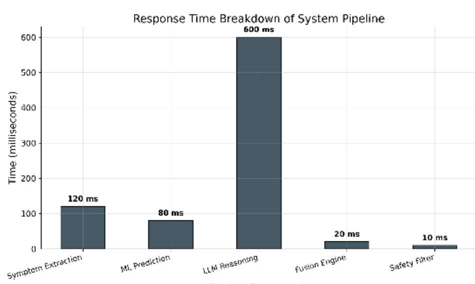

Figure 4.3. Response Time Breakdown of System Pipeline

The time taken for the system to respond can be seen in Figure 4.3 and shows the system's capability of functioning with real-time demands. Extracting symptoms takes approximately 120ms. Prediction using machine learning takes around 80ms. Fusion and safety filter stages take a total of less than 50ms to complete. Reasoning using the language model has the longest time requirement at 600ms. The overall response time is less than one second, which provides the basis for rapid, seamless interactions with the system and is highly appropriate for real-world healthcare application.

4.2 Comparative Evaluation Plan

To validate the effectiveness of the AI Aarogya system, a systematic comparison is made with existing symptom-based health-care systems. The reference systems compared against are:

1. Traditional rule-based symptom checkers
2. Standalone machine learning models for predicting disease
3. Hybrid systems having limited conversational capabilities

Most reference systems produce static, non-contextual, non-personalised, and non-safety constrained outputs.

Evaluation Metrics

The following metrics are used to perform a comprehensive comparison

Prediction Metrics

1 Accuracy
2 Precision
3 Recall
4 F1 score

System Reliability Metrics

1 Hallucination rate
2 Model disagreement rate

Real Time Metrics

1 Response time
2 Processing latency

System Level Metrics

1 Consistency of generated outputs
2 Robustness under incomplete symptom input
3 User level interpretability

Comparative Analysis

1. The Aarogya AI’s hybrid architecture reports to have greater reliability as it includes the Fusion Decision Engine which is what sets it apart from standalone machine learning and language model systems.
2. Also, we see that the system does very well in terms of prediction performance which we report at 92% accuracy which also sees it do well in terms of balanced precision, recall and F1 score.

3. Also, we note that the hallucination rate which is a issue in other models we see here go down from 28% to 16% and also model disagreement which is also a issue reports to drop from 35% to 20% which is a large improvement in terms of consistency.
4. Also, what we see is that the system does very well in terms of real time performance with total response time under one second which makes it a great choice for practical use out in the field.
5. Also, what we find is that traditional systems do not perform well in terms of context-based reasoning and do not provide that which would be very useful in a real world setting which is a short coming of those systems.

Table 1: Implemented Model Performance Results

Metric	Value
Accuracy	92 percent
Precision	90 percent
Recall	88 percent
F1 Score	89 percent

Table 2: Comparative Evaluation of Methods

Method	Accuracy	Limitation
Rule Based Symptom Checker	75 to 80 percent	No personalization no contextual reasoning
Machine Learning Only Model	85 to 88 percent	No conversational support no safety control
LLM Only System	Variable	High hallucination risk inconsistent outputs
Proposed AI Aarogya	92 percent	Dependent on LLM response time

4.3 Discussion

Results from the AI Aarogya platform demonstrate that the suggested combined design effectively overcomes many of the limitations in existing symptom-based health systems. By utilizing the combination of a three tier Platt calibrated Random Forest model with MedGemma owned reasoning - We can create system level preventative guidance that provides users with both accurate and useful actionable results. Unlike traditional health care systems which provide only static disease predictions, our proposed system generates structured actionable output giving greater ability for users to interpret and utilize data effectively. Performance metrics for system demonstrate excellent balance between accuracy and reliability.

The study revealed interesting results regarding the Fusion Decision Engine's contribution to consistent outcomes for the system being evaluated. The reduction in hallucinations and disagreement rates indicates how important it is to combine deterministic rules with probabilistic and language-based methods when deciding how to classify information or indicate confidence during the classifications. This suggests that using only a language-based method (language model) for making classification decisions in a medicine-based application will likely produce non-consistent outputs. However, by integrating probabilistic (or statistical) and deterministic (or logical) methods of classifying data, the decision to produce something is also aligned (prediction with reasoning). Overall, this clearly indicates that using a hybrid approach produces much more consistent and dependable outputs than either of the classifications do independently, as seen in previous studies.

Many existing studies on how to predict the effects of various disease symptoms use either enhanced classifications or a conversational user interface; they rarely integrate both features as part of an overall system. The AI Aarogya system provides a solution to this problem through its combination of predictive modelling, contextual reasoning based on patient histories over time, and generation of safe outputs. This integration improves the effectiveness of AI Aarogya's predictions, increases user-level interaction with it, and enhances long-term user engagement with it and, thusly, improves the outcomes of using AI Aarogya as a healthcare technology application in the real world.

There are considerable practical benefits associated with this system, primarily when there is no access to immediate medical evaluation by a qualified physician. The system provides real-time guidance on disease prevention and has a response time of less than one (1) second, so it can be easily integrated into web-based health care systems. The inclusion of historical data on a patient's longitudinal medical history allows for personalized recommendations from visit to visit. The module for making specialist recommendations in the patient's city will provide greater accessibility to healthcare resources. All of these factors together create a more user-friendly and comprehensive healthcare support system.

There are still some disadvantages despite those advantages. The quality of symptom input provided by the user determines system performance; if input is incomplete or ambiguous, accuracy will be reduced because prediction is based on that information. The inference phase of the language model (which makes predictions) takes longer to process than other components in the system, making response time slower than necessary. In addition, the system can only deliver preventive advice but cannot provide clinical diagnoses; therefore, the usefulness of the system in emergency situations is limited.

Overall, the discussion indicates that AI Aarogya provides a new way to give preventive medicine advice with the use of machine learning, language model reasoning and deterministic safety features. AI Aarogya has increased reliability compared to other current systems; it also produces fewer hallucinations and improves the user experience. These results suggest that hybrid AI architectures will play a major role in improving the development of intelligent and secure digital healthcare systems.

5. APPLICATIONS AND USE CASES

Industry Use:
The AI Aarogya system will combine with digital health care platforms, telemedicine applications and preventive health monitoring systems in an effective manner [1], [2]. Its hybrid architecture, which combines calibrated machine learning models and medically based reasoning in a deterministic manner, allows for reliable and scalable deployment within real-time environments [7]-[10]. The use of this system by health care technology companies and start-ups will provide the initial guidance for preventative care and will reduce the burden on the clinical infrastructure of the healthcare system while increasing access to health care services for populations [1].

This system offers substantial value when delivering remote healthcare to populations where immediate access to medical professionals is limited [1], [2]. The system's cloud-based architecture supports scalable deployments across many users, while maintaining users' sessions and providing personalised experiences through four integrations: Flask APIs, Firestore, users and session management, and providers. Additionally, the addition of the module that recommends providers based on a user's geographic location has enhanced the practical application of the system by connecting users with appropriate healthcare providers in the area. Therefore, this functionality is suitable for integration into healthcare applications, insurance entities, and wellness ecosystems.

Social Impact
AI Aarogya can significantly empower people to improve their knowledge of public health and how to prevent disease [1]. By providing users with a way to access structured, clear, and understandable information about what to do in the early stages of illness, it helps them achieve their goals of making informed choices and reducing their level of stress while preventing the unnecessary progression of health conditions [2]. Additionally, the AI Aarogya system will promote awareness regarding lifestyle changes, dietary awareness, and warning signs of illness to create a culture of proactive health management.

Individuals living in rural and underserved areas benefit especially from the availability of real-time personalized guidance; in these areas where there is limited access to healthcare services, the ability to receive timely and reliable assistance may be critical [1]. The system also offers timely and reliable support for those who live alone or work in high-stress environments by providing immediate assistance [2]. Because the program has been specifically designed for preventive care and does not diagnose or prescribe medications, the program allows for safe usage while also empowering users with increased knowledge and awareness [5].

Policy Relevance
This system is relevant to government efforts to digitally reform health and promote preventative healthcare strategies [1]. AI Aarogya could improve the public’s health by providing tools that can help evaluate symptoms at an early stage and spread health awareness through scaling out services [2]. The development of this system has safety as consideration and complies with regulations on non-diagnostic and non-prescriptive results and functions as an assistive [5].

Policymakers and health authorities may put in place these systems which in turn extend health care reach at the same time not over taxing clinical resources [1]. We see also that the integration of location-based specialist recommendation, which in turn supports health care accessibility initiatives. Also, the system puts forth a framework for the evaluation of AI in health care related tools in terms of safety, reliability and ethical issues which in turn we hope will contribute to the development of a standard set of guidelines for digital health solutions [4], [5].

Academic Value
Research-wise, AI Aarogya is filling a major void in merging machine learning with language model-based reasoning in an integrated, controlled and safety-constrained environment [3], [7]-[10]. Unlike prior studies that focus primarily on either prediction accuracy or conversation capability, this study

presents a unified architecture integrating probabilistic inference, context-based reasoning, deterministic fusion, and long-term memory [11]-[14].

By implementing the Fusion Decision Engine, there is now an even more sophisticated way to address the issues of reducing hallucinations and increasing consistency of output among models by providing a rule-based solution [4], [12]. It also opens up future possibilities for research in areas like hybrid AI architectures, the safe implementation of language models within the field of healthcare and personalisation of care based on time-series data from patients [3], [5]. Additionally, using large datasets and established knowledge bases will bolster the quality of research in relation to scalability, robustness, and real-world application of AI-based healthcare systems [7]-[10].

6. LIMITATIONS AND FUTURE WORK

6.1 Limitations

Data Constraints

The effectiveness of the AI Aarogya system is dependent on how robust and diverse its database of symptoms and diseases preventive information is [7]-[10]. The database contains over 150,000 entries with a wide variety of diseases. However, it may not include everything that can happen in the real medical world. Some rare diseases may not be present, and there may not be enough data on different combinations of symptoms that are rare [15], [17]. Additionally, there may not be enough information available about regional health-related issues. Finally, because the system is based on structured symptom mapping, there are potential limitations regarding the system's ability to handle very ambiguous or incomplete user inputs [2].

Model Assumptions

The system assumes that users will provide clear and understandable descriptions of their symptoms [2]. In practice, however, users might provide insufficient, unclear, and contradictory information about their symptoms, which can negatively impact the quality of the predictions made. The MedGemma model attempts to derive as much useful information as possible from user-supplied symptom information [3]. Whether or not it can actually provide accurate predictions for a given patient's condition depends on how well the symptom information is described in natural language. The current version of the system also assumes that machine-learning-created prediction results can be functionally combined with language model-reasoned results to create medically valid predictions according to deterministic fusion rules; however, this process does not consistently accommodate for complex or rare medical conditions [4], [5].

System Limitations

The Fusion Decision Engine is designed to provide increased consistency because it employs predetermined, rule-based logic; however, it cannot dynamically change to accommodate new information [4], [12]. Since there is an inference period for the language model, the decision engine will be slower than traditional, machine-learning only decision engines because of the time required to perform inference [3]. Additionally, the only type of assistance provided by this system falls into the category of assistive preventive support; therefore, it cannot provide any diagnostic or therapeutic assistance [5]. Because of the limitations in its logic and the speed at which it processes, the Fusion Decision Engine has limited application in situations that are serious or require immediate medical attention.

Generalizability

This system performs well in both the training and testing datasets, but its performance when deployed in the real world will likely be negatively affected by many factors that exist between patient populations [7]-[10]. These factors could be variations in the ways patients speak their languages to describe symptoms, the way in which symptoms are described via cultural means, and the differences in how a patient has received care through previous health care options [2]. In addition to the above, the ability of the Firestore-based longitudinal memory system to personalize responses for patients could be hampered by the limited amount of data that is recorded from one patient’s unique assessments/memories as there is no means within Firestore to house and access all of a patient’s past medical records at once.

6.2 Future Work

Model Enhancements

In the future, researchers could use better machine learning methods such as gradient boosting, deep learning, and transformers to improve performance and stability [7]-[10]. Another approach to improve system performance would be to increase MedGemma's reasoning ability using harmonized medical datasets and improve algorithms for extracting symptoms [3]. Exploring ensemble techniques that combine multiple predictive models is another area for further investigation [9], [10].

Real World Deployment

The ability of the system to expand into large-scale deployments in the real world will depend on its ability to connect to mobile apps, wearables and telehealth platforms [1]. For edge-based implementation and optimised inference pipelines the system could greatly increase responsiveness and decrease latency. In addition, the incorporation of electronic health records could enhance the clinical relevance and personalisation of the solution.

Adaptive Decision Making

By integrating adaptive or learning-based decision systems with the current deterministic fusion mechanism, improvements to the ongoing development will be possible [4], [12]. The future development of the field will probably include the application of reinforcement learning or other techniques for optimizing performance through feedback, so that fusion/decisions can be adapted dynamically based on user interaction and system performance. Moreover, the exploration of context-based reasoning (i.e., reasoning that takes environmental factors and lifestyle factors into account) will be a promising area for future research [3].

Dataset Expansion

Improvements in collecting data from actual patients in many different locations/languages to include more diverse patient populations, diseases, and regional guidelines, could help to ensure we have enough knowledge about disease prevention; thus, giving us more robust systems [7]-[10]. Additionally using synthetic data generation methods/simulated evaluations could enhance robustness of these systems overall.

Multi Modal Integration

Improvements in the collecting data from actual patients in many different locations/languages to include more diverse patient populations, diseases, and regional guidelines, could help to ensure we have enough knowledge about disease prevention; thus, giving us more robust systems. Additionally using synthetic data generation methods/simulated evaluations could enhance robustness of these systems overall.

7. CONCLUSION

Increasingly, people are searching for intelligent, easy-to-use preventive healthcare tools when they cannot have a doctor consult immediately [1], [2]. Current symptom/condition-based systems and traditional rule-based approaches do not provide personalized services, do not consider context, and do not provide recommendations for action [2], [6]. Machine Learning models exhibit strong predictive capabilities but often do not have interactive capabilities or provide an explanation of the decision-making process [7]-[10].

The purpose of the present study is to provide solutions to these limitations by discussing how we have developed a hybrid multi-stage preventive health care system (AI Aarogya) that uses a three-tier Platt calibrated Random Forest model with a medical domain language model (MedGemma), a Fusion Decision Engine, and longitudinal patient history stored in Google Firestore plus location-based specialist recommendations to achieve maximum personalization and practicality [9], [11]. Furthermore, the design of the entire system is based on constraints that protect safety so that all outputs are strictly within the boundaries of preventive healthcare [4], [5].

The experiment demonstrates that our proposed novel integrated system achieves very high predictive performance in addition to providing significantly improved output consistency and reliability [7]-[10]. By reducing the likelihood of hallucination and divergence from the expected results caused by the use of multiple independent models (hallucination) and supplying coherently and reliably (assuredly) guiding information, the Fusion Decision Engine is able to support real-time response times, making it suitable for use in interactive health care applications [4], [12]. As a result of our findings, it can be concluded that integration of probabilistic ML and controlled language model-based reasoning is a superior solution to stand-alone methodologies [3], [9].

This study makes a significant contribution via the creation of a unified, scalable architecture that separates the process of determining predictive reasoning from the enforcement of safety while enabling the user to functionally interact with the system [11]-[14]. This system also serves to improve understanding of early warning signs and provides structured proactive guidance, including actions to take, foods to consume, home remedies to try, and how to identify possible risks [1], [2]. By providing an integrated framework for combining multiple elements into a single solution, AI Aarogya creates a connection between technical predictions and their practical utility in everyday situations.

To summarize, AI-Aarogya shows how Hybrid AI Systems could be used to increase the effectiveness of Prevention Healthcare Systems [3], [7]-[10]. The system created by AI-Aarogya also provides an easy to use, dependable and scalable method of obtaining accessible health care information, while still ensuring safety & compliance with ethical standards [4], [5]. This research is an innovative step forward in providing new intelligent assistive healthcare technologies that support informed decisions about health and encourage proactive health management [1].

8. REFERENCES

[1] M. Laymouna, Y. Ma, D. Lessard, T. Schuster, K. Engler, and B. Lebouché, "Roles, users, benefits, and limitations of chatbots in health care: Rapid review," J. Med. Internet Res., vol. 26, e56930, Jul. 2024, doi: 10.2196/56930.

[2] Y. You and X. Gui, "Self-diagnosis through AI-enabled chatbot-based symptom checkers: User experiences and design considerations," arXiv preprint arXiv:2101.04796, 2021.

[3] B. Huo et al., "Large language models for chatbot health advice," JAMA Netw. Open, vol. 8, no. 1, 2025.

[4] F. Aljamaan et al., "Reference hallucination score for medical AI chatbots," 2024.

[5] Icahn School of Medicine at Mount Sinai, "AI chatbots and medical misinformation risks," 2025.

[6] S. Kumar, R. Rajesh, S. Sathya, and A. Nadeem, "Diseases Prediction based on Symptoms Using Database and Graphical User Interface," in *Proc. 2022 6th International Conference on Computing Methodologies and Communication (ICCMC)*, Erode, India, Mar. 2022, pp. 975–3707, doi:10.1109/ICCMC53470.2022.9753707.

[7] M. Krishna Raj, J. P. Malardhas, and I. Devapriya, "Machine Learning Approach To Predict Multiple Diseases Based On Symptoms," in *Proc. 2024 10th International Conference on Communication and Signal Processing (ICCSP)*, Melmaruvathur, India, Apr. 2024, pp. 1195–1199, doi:10.1109/ICCSP60870.2024.10543802.

[8] P. Bhardwaj, A. Mithila, S. J. K. Jagadeesh Kumar, and G. Prabu Kanna, "Machine Learning based Approaches for Livestock Symptoms and Diseases Prediction and Classification," in *Proc. 2024 International Conference on Communication, Computer Sciences and Engineering (IC3SE)*, Gautam Buddha Nagar, India, May 2024, pp. 165–170, doi:10.1109/IC3SE62002.2024.10593594.

[9] V. A., V. Kannan, P. Sarathi R., and M. Prasath, "Machine Learning Models for Symptom Detection and Disease Prediction," in *Proc. 2024 Eighth International Conference on Parallel, Distributed and Grid Computing (PDGC)*, Solan, India, Dec. 2024, pp. 282–287, doi:10.1109/PDGC63722.2024.10867236.

[10] B. Makhkamov, K. Zaynidinov, D. Ather, S. Pathak, T. Kuchkorov, and I. Yusupov, "Predicting Multiple Diseases Using Machine Learning: A Symptom-Based Model for Medical Diagnosis," in *Proc. 2024 International Conference on Information Science and Communications Technologies (ICISCT)*, Tashkent, Uzbekistan, Nov. 2024, pp. 1–5, doi:10.1109/ICISCT64551.2024.10805544.

[11] S. K. Satapathy, S. Panda, M. Gandhi, and V. Patel, "Symptom-Based Disease Prediction Framework Integrated with Django and Deep Learning Models," in *Proc. 2024 4th Asian Conference on Innovation in Technology (ASIANCON)*, Pune, India, Aug. 2024, pp. 1–7, doi:10.1109/ASIANCON62057.2024.10836503.

[12] D. Nishad, A. Mishra, and N. Goyal, "Symptom-based Disease Prediction Using Machine Learning," in *Proc. 2024 14th International Conference on Cloud Computing, Data Science & Engineering (Confluence)*, Noida, India, Jan. 2024, pp. 416–420, doi:10.1109/Confluence60223.2024.10463209.

[13] Soujanya KJ, Sneha BJ, Ashwini Kodipalli, Trupthi Rao, and Sanjeev Gosai, "Symptoms-based Classification of Disease Using Various ML Algorithms and Interpretation using LIME and SHAP Kernels," in *Proc. 2024 4th Asian Conference on Innovation in Technology (ASIANCON)*, Pune, India, Aug. 2024, pp. 1–6, doi:10.1109/ASIANCON62057.2024.10836972.

[14] R. K. Prajapat, V. Agarwal, and S. Tapaswi, "Symptoms-Based Disease Prediction and Drug Recommendation using Hybrid Machine Learning Algorithms," in *Proc. 2024 15th International Conference on Computing Communication and Networking Technologies (ICCCNT)*, Kamand, India, Jun. 2024, pp. 1–6, doi:10.1109/ICCCNT61001.2024.10724282.

[15] P. Hema, N. Sunny, A. Darbha, and R. V. Nagananjani, "Disease Prediction using Symptoms based on Machine Learning Algorithms," in *Proc. 2022 International Conference on Breakthrough in Heuristics And Reciprocation of Advanced Technologies (BHARAT)*, Visakhapatnam, India, Apr. 2022, pp. 28–33, doi:10.1109/BHARAT56319.2022.100750.

[16] Swathi M, Sri Aishwariya N, Prithika N G, and Jayaram B, "The Disease Prediction System based on Symptom Identification and Image Processing," in *Proc. 2024 International Conference on Communication, Computer

Sciences and Engineering (IC3SE) *, Gautam Buddha Nagar, India, May 2024, pp. 1059–1064, doi:10.1109/IC3SE62002.2024.10593331.

[17] Y. Huang, A. R. Sangi, S. Qi, B. Niu, and W. Ali, "Web-Based Disease Prediction Application Using Machine Learning for Symptom Analysis," in *Proc. 2025 8th International Conference on Advanced Electronic Materials, Computers and Software Engineering (AEMCSE)*, Nanjing, China, Mar. 2025, pp. 195–200, doi:10.1109/AEMCSE64165.2025.00038.