MedWise: AI/ML Powered Healthcare Web Application

**Abstract—Seeking healthcare in an appropriate amount of time is often costly and time-consuming, requiring people to see doctors even for minor health issues. This becomes inconvenient, especially when it is not an emergency condition and requires a long wait. This paper presents a smart application based on Artificial Intelligence (AI) and Machine Learning (ML), called MedWise, which aims to enhance healthcare services by making them more efficient and available.  
  
The responsive AI chatbot is one of the most crucial components of MedWise, enabling users to receive health tips on request, get instructions tailored to their issues, analyze their symptoms, and better understand health and diseases. The chatbot also provides possible solutions, allowing people to address health concerns without needing to visit a physician every time.  
  
To improve the disease modeling capabilities of the web application, it is complemented with advanced machine learning techniques for disease modeling and Convolutional Neural Networks (CNNs) for medical imaging diagnosis. This integrated approach bridges self-health management with professional healthcare through the MedWise application, making it more optimal and versatile.  
  
These factors transform healthcare by making it more accessible, affordable, and faster, enabling users to manage their health and acquire relevant health advice when needed.**

***Keywords: Machine learning, Deep learning, Artificial intelligence, Healthcare system***

I. INTRODUCTION

Right to good health is a basic right of every individual as reiterated by the World Health Organization of the United Nations. It is a right that every person should have regardless of where the person is located, or their social and economic status, and so on. But in the case of India, a country whose population today has gone beyond one point four two billion, the healthcare system still has a number of deficiencies that makes it unable to attend to the health care needs of all the citizens in the society. India is populated by a large mix of people who live in different regions, have different cultural practices and come from different income groups. This mix is a reason for the unequal distribution of healthcare services, and the wide gap existing between urban and rural areas in terms of structures, types of medical professionals, and their levels of income. Even majority of people in different places are faced with the challenge of getting appropriate healthcare services in the right time. The COVID-19 pandemic highlighted these systemic healthcare gaps. During the crisis, millions of people faced challenges such as misinformation about the virus, delays in accessing accurate medical advice, and shortages of healthcare resources.

One of the prominent challenges which India’s healthcare system has to resolve is the relatively scarce number of doctors and other health care personnel in rural communities. Despite urban areas having better infrastructure and more healthcare facilities, urban residents face their own set of challenges, such as busy schedules and logistical difficulties, which deter timely medical consultations. Financial constraints exacerbate the situation, with many individuals unable to afford regular medical check-ups or treatments. These problems in large part tend to afflict and accentuate chronic health issues among the rural poor. The unavailability of local health experts compels the residents to seek medical help from distant towns. This travel is not only lengthy but expensive and as such not many poor rural folks get the health care. More so, non-availability of essential primary health care services such as lack of proper diagnostic facilities to aid in early detection of illnesses to well equipped hospitals lead ordinary illnesses to develop into life threatening conditions. These were existing challenges before COVID-19, which makes me want to reiterate that these require immediate fixes.

MedWise is presented as a transformative solution intended to bridge the gaps and address the systemic issues in this challenge. By integrating artificial intelligence (AI) and machine learning (ML), MedWise provides a radical change in healthcare as it improves the ease of access, convenience and cost effectiveness to everyone. MedWise has potential of the patients or target groups engaging user-friendly interaction and health assistance that incorporate new technologies capable of addressing a wide range of issues including advanced support for early diagnosis of health conditions and anticipating disease outlook for preventive health care.

MedWise has an intelligent chat bot integrated within the app whereby the user can interact with the health app quite naturally with guidance of RAG technology and conversational AI models developed by OpenAI. The chat bot acts as the first port of call for people looking for medical help articulating symptoms or asking general health questions. The Max of the device is updated with what is termed relevant into the RAG model which assists in real time responses making the bot or assistant auto relevant and contextually relevant at all times. This enables users to receive timely advice and valuable information about the disease, which, in turn, enables the people to make informed choices with the information.

The application incorporates advanced AI-based tools, such as disease and image analysis predictive models. Such functions allow users to gather information on possible threats to their health and make decisions regarding treatment at an earlier stage. For example, disease prediction instruments can make use of the data provided by users to establish the risks of complications of diabetes and also primary and secondary heart disease as well as liver disease. Also, using same models, image analysis features help in detection of critical medical conditions such as brain tumours and arthritis through assessment of various medical images. These diagnostic features make MedWise an all-encompassing platform that transits through the prediction of diseases, prompt diagnosis, and prognostic evaluation, all with online convenience.

II. PROPOSED WORK

The proposed system, utilizes advanced AI and ML technologies to revolutionize healthcare delivery. By providing immediate diagnostic insights, symptom analysis, and customized recommendations, it lessens reliance on conventional healthcare resources and is intended to deliver accurate, rapid, and individualized medical support. The solution ensures a smooth and efficient healthcare experience by addressing important issues including consultation delays and inefficient resource use.  
  
In order to close gaps in healthcare delivery and accessibility, the system focuses on time-saving and resource-efficient alternatives. Because of its creative methodology, the system is positioned as a useful and expandable instrument to eliminate systemic inefficiencies and enhance healthcare results, especially in underprivileged areas.

The Proposed system involves following Key Operation:

1. User profile management
2. Disease predictions
3. Image analysis
4. Medical chatbot

CLASSIFICATION ALGORITHMS

The system utilizes a variety of classification algorithms to forecast diseases, assessing each one's performance to determine which model produces the most accurate and trustworthy findings. When the goal is to identify or classify data according to specified criteria, classification algorithms are essential to machine learning tasks. These algorithms work by examining a labeled dataset, in which every data item has a target label or category attached to it. The model can correctly categorize fresh, unknown examples because it learns patterns, relationships, and underlying structures from the training data.

The system evaluates the algorithms on multiple datasets related to different diseases to identify the algorithm that consistently delivers the best predictive results. This is what ensures that the system makes use of the best possible model for disease prediction and, therefore, optimizes the overall effectiveness in making timely and accurate medical inferences.

1. **KNN:** KNN is simple algorithm that classifies based upon the majority class of those nearest neighbours. It calculates the distance between points in the feature space and assigns the class of the closest ones.
2. **SVC:** SVC works by finding a hyperplane that maximizes the margin between different classes. It maps data to higher dimensions to separate them efficiently. SVC is highly effective for binary classification tasks and performs well with complex datasets.
3. **Naïve Bayes Classifier**: Naive Bayes applies Bayes' theorem and assumes that features are conditionally independent given the class. It calculates the probability of each class based on the input features and selects the class with the highest probability. It's fast and works well with high-dimensional data.
4. **Decision Tree:** A Decision Tree splits the data into subsets based on feature values. It forms a tree-like structure, with every node representing a decision, while branches lead to possible outcomes. It is very intuitive to interpret but could overfit easily if it isn't kept in check properly.
5. **Random Forest:** Random Forest uses lots of decision trees combined and gives the result for high accuracy. Each tree learns in a random subset of the data, reducing overfitting and increasing robustness. In both classification and regression problems, it is very efficient.

For image analysis of brain tumour classification the system uses a Convolutional Neural Network algorithm. Initially input images are pre-processed. Each image is resized into a specific dimensional size and normalized to ensure stable and efficient training. Subsequently, the CNN extracts features from this image with various convolutional layers. The first layer of applies 32 filters over the input data for detecting low-level features like edges and textures, which is followed by ReLU activation for nonlinearity. MaxPooling layers are used for downscaling the image spatially while preserving significant dimensions of the feature. Other convolution layers learn much higher-level complex patterns such as shapes of the tumor and unique structures; when flattened into a one-dimensional vector, this vector is fed into fully connected (dense) layers to extract high-level relations such as tumor and non-tumor. Dropout helps in avoiding overfitting through random switching off neurons from training. Finally, the output activation is through a sigmoid function which will output a probability value classification over "tumor" or "no tumor". This model is developed to optimize the binary cross-entropy loss function using Adam algorithm for better error minimization in the classification accuracy. Built to be an efficient and reliable tool for healthcare professionals, especially those in underserved areas, this enables the system to accurately identify brain tumors through medical images.

**Key components:**

**ReLU:** The Rectified Linear Unit, is an activation function that adds non-linearity to the input so the model can recognize intricate patterns.

**MaxPooling:** A method that makes the model more effective by reducing the image's spatial dimensions while maintaining key features.

**Dropout:** A regularization technique that ensures the model generalizes better by randomly deactivating neurons during training, hence preventing overfitting.

The output layer's sigmoid activation function transforms the model's output into a probability value (ranging from 0 to 1) for binary classification (tumor vs. no tumor).

In binary classification tasks, the error between predicted probabilities and true labels is computed using the Binary Cross-Entropy Loss loss function.

**Adam Optimizer**: An optimization algorithm used to adjust the model’s weights during training to minimize the loss and improve accuracy.

The system's chatbot is based on retrieval-augmented generation (RAG), in which the process begins by converting the information into smaller, meaningful chunks. These chunks represent distinct segments of information that can be easily retrieved and processed. After dividing the information into chunks, each chunk is embedded using the Alibaba-NLP/gte-base-en-v1.5 model, which transforms the text into vectors that capture its semantic meaning. The embedded chunks are stored in a vector database, which enables efficient retrieval based on similarity. When a user gives a query, it is also embedded into a vector. The system then searches the database to find the 3 most similar chunks to the query. These relevant chunks are combined with the query and passed to OpenAI's GPT model, which produces a response by combining the query with the retrieved information, providing a more context-aware and accurate answer. This RAG-based approach enhances the ability of the chatbot to provide more precise and relevant responses.

Flow Chart







