As per UGC guidelines an electronic bar code is provided to secure your paper

International Journal for Modern Trends in Science and Technology
Volume 11, Issue 05, pages 46-57.
ISSN: 2455-3778 online
Available online at: http://www.ijmtst.com/vol11issue05.html
DOI: https://doi.org/10.5281/zenodo.15252388

AI-Driven Healthcare System: Transforming Disease Detection and Prevention for a Healthy Future

L Venkata Sai Meghana[1] | Dr. K Venkatesh[2]

[1]PG Scholar, Department of Artificial Intelligence & Data Science, Ramachandra College of Engineering, Eluru, India.
[2] Professor, Department of Artificial Intelligence & Data Science Ramachandra College of Engineering, Eluru, India.
Corresponding author: meghanalingala23@gmail.com

To Cite this Article
L Venkata Sai Meghana & Dr. K Venkatesh (2025). AI-Driven Healthcare System: Transforming Disease Detection and Prevention for a Healthy Future. International Journal for Modern Trends in Science and Technology, 11(05), 46-57. https://doi.org/10.5281/zenodo.15252388

Article Info
Received: 22 March 2025; Accepted: 18 April 2025.; Published: 20 April 2025.

KEYWORDS

AI in Healthcare, Clustering, Deep Learning, Disease Detection, Machine Learning, Medical Imaging, Preventive Healthcare and Smartwatch Data

ABSTRACT

A rising number of chronic and acute illnesses requires the development of an advanced medical system with early detection and prevention skills. Traditional methods usually need expensive equipment and expert medical knowledge which limits accessibility and delays emergency treatment. In order to address these problems this study proposes an AI- driven healthcare platform that integrates several health parameters utilizing methods from deep learning as well as machine learning and uses data from watches to continuously monitor people's daily health. The ideal number of groups in the system which employs K-Means and Agglomerative Clustering, is ascertained using the Elbow Method. K-Means achieves a silhouette score of 0.85 indicating outstanding clustering effectiveness. Furthermore, the framework classifies illnesses into No Dementia, Very Mild Disease, Mild Dementia and Moderate Dementia using cutting-edge deep learning models like Inception V3, ResNet50, and VGG19 for brain disease detection. These models achieve test accuracies of 90.94%, 79.55% and 93.14% respectively. Similarly, CNN-based models use multiple layers of perception on textual data to detect skin disorders with 95% accuracy lung conditions with 92.15% accuracy during testing, blood-related cancer stages with 95% reliability and cardiovascular conditions with 99.17% accuracy. Additionally, a Decision Tree Classification uses textual data to efficiently classify liver disorders with 98% accuracy. Streamlit UI which offers an easy-to-use interface for disease diagnosis, real-time health monitoring and customized preventive interventions is used to implement the system. This system seeks to transform illness detection and prevention by fusing cutting-edge AI algorithms with actual health data, thereby improving accessibility, health outcomes and

1. INTRODUCTION

Over time healthcare has changed dramatically moving from antiquated methods to cutting-edge technologically advanced alternatives. In the past manual inspection, patient-reported symptoms and standard laboratory testing were the main methods used to diagnose diseases [1]. Even though these techniques have proven successful they frequently take a lot of time and skill which causes delays in the process of diagnosis and therapy. The advent of imaging techniques, electronic health records (EHRs) and personal health monitoring devices has improved detection accuracy and efficiency [2]. However, even these developments early detection of chronic illnesses remains difficult highlighting the necessity of data mining and Neural Networks to further improve healthcare systems. The main pillars of the current healthcare systems are expensive diagnostic procedures, doctor expertise and hospital visits. Rural and disadvantaged populations have less access to healthcare because many diagnostic procedures including blood tests, CT scans and MRI scans require specific devices and skilled personnel [3]. Furthermore, the majority of current systems are unable to efficiently use the massive amounts of real-time health data collected by wearable health monitoring devices such as smartwatches, for predictive analytics. Furthermore, conventional clustering techniques in health monitoring frequently fall short of producing useful patient groupings which results in ineffective tailored healthcare recommendations. These drawbacks draw attention to the shortcomings of traditional healthcare methods making the creation of an AI-powered system that can handle massive amounts of medical data, identify illnesses early and offer individualized health insights necessary.

For comprehensive illness prevention and diagnosis, this study proposes an AI-powered medical system that integrates deep learning, machine learning and health data collected from smartwatches. The full framework is implemented by Streamlit UI, which provides an intuitive and interactive user experience for measuring wellness in real time and providing guidance on preventative healthcare [4]. This study was motivated by the rising incidence of chronic diseases such as skin cancer, pneumonia, heart problems and Alzheimer's disease worldwide. Many of these illnesses can be managed with early identification, but existing approaches cannot provide real-time personalized insights based on individual health trends. An opportunity to use AI-driven methods for ongoing health surveillance and disease prediction has arisen with the popularity of personal health monitoring devices [5]. Furthermore, greater risk stratification is made possible by grouping people according to health characteristics which enables focused interventions before a problem develops. The goal of this project is to improve illness detection and prevention and general healthcare accessibility by bridging the gap between AI-based predictive analytics [6] and real-time health monitoring.

II.LITERATURE SURVEY

Over the past ten years, there has been a considerable advancement in the application of data mining techniques for sickness identification and prevention. Numerous AI-powered techniques for using medical imaging and textual health records to identify ailments have been investigated by researchers. For instance, convolutional neural networks (also known as CNNs) have demonstrated remarkable efficacy in recognizing skin cancer, lung disorders, and brain issues from medical images. Studies have shown that MRI scans, employing models such as Inception V3, ResNet50, and VGG19, may accurately detect neurological diseases including dementia and Alzheimer's [7]. Similarly, CNN- based models have demonstrated exceptional efficacy in distinguishing between cancer and benign skin lesions when used to diagnose skin disorders. Additionally, deep learning architectures like Recurrent Neural Networks (RNNs) and LSTM networks (Long Short-Term Memory) have been used to assess textual health data and health record data. This has allowed for the automatic diagnosis of heart and liver disorders. Despite these technological advancements, most of these systems focus on identifying a single disease rather than using an integrated approach that combines real-time health monitoring with AI-driven multi-disease diagnosis and therapy. Numerous methods for clustering

[8] have been studied for categorizing individuals based on health parameters in order to personalize healthcare solutions. Although deep learning-enhanced clustering

and hybrid clustering approaches have been used in some recent studies, these methods have not been successfully merged into an immediate smartwatch-based health tracking system. Furthermore, rather than continuously gathering health metrics from wearable technology, current healthcare AI models mostly rely on diagnostic data from hospitals. By integrating deep learning-driven disease identification, machine learning-driven clustering and wristwatch health monitoring this research seeks to close these gaps and offer a complete, real-time healthcare solution [9].Kim et al. develop a model that predicts the prevalence of cardiovascular disease using health-related data from the Korean National Health and Nutrition Examination Survey [18]. The model uses artificial neural networks, logistic regression, and support vector machines to categorize the prevalence of cardiovascular illness. The support vector machine approach was the most accurate. The study aims to examine the parameters or structure of the support vector system in order to increase accuracy and determine the importance of input factors. This method could be used to predict when cardiovascular disease will develop. Sengul et al. in order to forecast detection of falls from four everyday activities sitting, squatting, jogging and walking a smart watch-based technology is being created. Data from gyroscope and acceleration sensors is gathered by the system and sent to the cloud. Activities are grouped according to their classes using a deep learning algorithm [19]. The technique expands training data samples using Bica cubic Hermite interpolation. The system classifies activities using a simultaneous long short-term memory (BiLSTM). When all activities are taken into account the system achieves 99.59% and 97.35% accuracy whereas binary classification yields flawless accuracy. Tarafdar et al. by using of multimodal sensors included in smart devices like wearables, smartwatches and smartphones they are essential for activity-based wellness management. These gadgets have a wealth of data that can be combined to identify human activities. According to research boosting algorithms beat conventional machine learning algorithms [20] in identifying fundamental human actions like walking, sitting, standing, exercising and sleeping. They can also extract information for human activity identification in real-life contexts. In order to guide the selection of sensor elements for enhancement the study also contrasts the potential of smartwatches and cellphone in activity detection. The study has significance for activity-based wellness treatment both theoretically and practically.

III.METHODOLOGY

1.Data Collection

A wide range of medical and physical information gathered from smartwatches, imaging devices and textual health records are used in the suggested AI-driven healthcare system. Large volumes of real-time health information are produced by smartwatch devices which continuously track users' heart rates, oxygen levels in their blood, workouts, patterns of sleep and stress levels. This information is essential for spotting odd patterns and forecasting possible health hazards. To guarantee data quality, preprocessing methods including outlier detection, normalization and scaling of features are necessary because raw smartwatch data is frequently noisy and inconsistent. Mean imputation is used to manage missing values and methods for selecting features are used to keep only the most pertinent parameters for grouping people according to their health attributes. We categorize individuals into four groups no dementia, mild dementia, very mild dementia and moderate dementia using a big dataset of MRI scans to identify brain disorders. The dataset consists of 3,200 images of individuals without dementia, 2,240 images of individuals with mild to moderate dementia, 896 images of individuals with mild dementia, and 64 images of individuals with substantial dementia. Given the class imbalance, data augmentation techniques such as rotation, flipping and contrast modification are utilized to increase the number of examples in minority categories. MRI scans are normalized and reduced to 224x224 pixels before being fed into models of neural networks for feature extraction and categorization. The 5,000 labeled images used to categorize skin diseases include nine different skin conditions: Actinic Keratosis, Basic Cell Tumors, Dermatofibroma, Melanoma, Nevus tumors, Colored Benign Acne, Seborrheic Keratosis, Scaly Cell Carcinoma, and Vascular Lesions. Since some classes have fewer instances than others, techniques including image cropping, zooming, and random noise addition are utilized to balance the dataset. Each image is reduced to 224 by 224 pixels, normalized, and, if necessary, transformed to grayscale in order to enhance model learning. The dataset is then split into learning (80%) and evaluating (20%) sets to ensure a trustworthy evaluation.

A dataset of 3,242 blood smear pictures gathered from 89 patients suspect of having leukemia with acute lymphoblastic leukemia (ALL) is utilized to detect blood diseases. Four classes that correspond to various disease phases are used to categorize the dataset. Preprocessing techniques include enhancement of contrast, noise reduction with median filtering and architectural treatments to emphasize important characteristics because blood cell pictures frequently have poor contrast and variable illumination. Model generalization is further enhanced by data augmentation methods like elastic transformation and random rotation. A dataset of 5,216 chest X-ray pictures is utilized to detect lung diseases. 1,341 of the photos are classified as normal (healthy) while 3,815 are classified as pneumonia cases. 390 cases of pneumonia and 234 healthy cases make up the 624 images in the testing set. Because chest X-ray pictures frequently have uneven brightness and contrast preprocessing techniques such noise filtering, edge sharpening and histogram equalization are required. Better model learning is ensured by applying geometric transformations and synthetic minority oversampling (SMOTE) to overcome data imbalance. Textual information containing clinical indicators like blood pressure, cholesterol, variability in heart rate and ECG readings is used to forecast heart disease. To guarantee consistency the dataset is subjected to preprocessing procedures such feature encoding, value that is missing imputation and standardization. Since text medical data frequently includes redundant or irrelevant attributes, algorithms for choosing features such as PCA (principal component analysis) and RFE (Recursive Feature Elimination) are utilized to increase predictive accuracy. Liver disease is classified using a dataset that includes many biochemical features, such as bilirubin levels, enzyme counts and albumin levels. Examples regarding information preparation include outlier detection, scaling features to ensure uniformity across several attributes, and log transformations to address uneven distributions.

2.Data Preprocessing

In applications for healthcare where data originates from a variety of sources such as wearable technology, imaging services and textual records efficient data preprocessing [21] is an important part in guaranteeing accuracy of machine learning tasks. Model performance may be adversely affected by noise values that are missing, inconsistencies and variable feature distributions found in the raw data gathered from watches, MRI scans, X-ray pictures, blood sample slides and organized health records. A methodical data pretreatment pipeline that included feature extraction, dimensionality reduction, augmentation, normalization and data cleaning was put in place to solve these issues [22]. Whereas categories are represented using one-hot encoding or label encoding to guarantee compatibility with machine learning models missing values in the case of structured health data (such as records of heart and liver diseases) were dealt with using estimation methods like mean, median or k-nearest neighbor (KNN) imputation. Scaling techniques like Min-Max Scale and Standardization [23] were employed to bring all features to a uniform scale in order to avoid skewing models toward high-magnitude features that prevents bias caused on by significant magnitude differences by bringing the values of features into a specified range [0,1].

$$I_{normalized} = \frac{I_{original}}{255}$$

Where, $I_{original}$ is the original pixel value *(ranging from 0 to 255)*. $I_{normalized}$ is the normalized pixel value *(ranging from 0 to 1)*. 255 is the maximum pixel intensity in an 8-bit image.

A kernel (filter) is used to each convolution operation in CNNs (such as VGG-19, AlexNet, ZFNet, LeNet-5, and InceptionV3) in order to extract characteristics from input images.

$$O(i,j) = \sum_{m=0}^{M-1} \sum_{n=0}^{N-1} I(i+m, j+n) \cdot K(m,n)$$

where: $O(i,j)$ is the output feature map, $I(i+m, j+n)$ is the input image, $K(m,n)$ is the convolutional kernel (filter),

M, N are the filter dimensions.

X = $\underline{X - Xmin}$
$normalized \quad X - X$
$max\,min$

Where, X is the original feature value, $Xmax$ and $Xmin$ are the low and peak values of the main features. Preprocessing aimed to improve image quality and guarantee consistency across datasets for medical imaging applications such as blood smear slides, skin lesion images, lung X-rays and brain MRI scans. Changes in contrast, decision, conditions of lighting and artifacts had to be addressed because the photos came from a variety of medical devices and organizations. To improve the models, capacity to concentrate on structure patterns instead of color fluctuations, grayscale conversion was used to exclude extraneous color information from X-ray and MRI data. Histogram equalization [24] and context-limited adaptive equalization of histograms (CLAHE) were used to enhance the contrast of images and make important elements more visible which makes medical images more dynamic by distributing intensity values around the area known as the histogram.The clip factor is denoted by α. The maximum number of luminance bins and the overall number of bytes in the region are denoted by **B** and L, respectively. To ensure uniformity across the collection, photos were also resized to a preset pixel size of either 224 x 224 or 256 x 256. Data augmentation techniques [25] such as random shifts, zooming, flipping, brightness modifications, and elastic deformations, were used to ensure that the neural networks learnt robust and invariant representations, avoid overfitting, and enhance model generalization. By choosing the most pertinent traits and removing unnecessary or uninformative features, feature engineering significantly improved model performance [26]. In order to preserve just the most predictive features for textual health records (like those used to forecast liver and heart disease), feature selection was carried out utilizing mutual information, correlation analysis, and recursive feature elimination (RFE). PCA (principal component analysis) and t-Distributed Sequential Neighbor Embedding (t-SNE) were used to reduce dimensionality while preserving important data by eliminating features that are not needed.

where λ are the eigenvalues.

$$Cov(X)W = \lambda W$$

pij

$$KL(P\,||\,Q) = \sum_i \sum_j pij \log qij$$

Where, pij and qij are pairwise similarities in high and low dimensional spaces. By eliminating noise and emphasizing the most important health indicators, dimensionality reduction enhanced the effectiveness and understanding of the clusters in clustering models like K-Means and Agglomerative Clustering [27]. In order to extract excellent hierarchical features pertinent to illness classification, transfer learning was utilized for deep learning models. This required optimizing models that were already trained like VGG19, ResNet50 and Inception V3. Class disparity is a significant problem in medical datasets since some diseases (such uncommon skin cancers or specific leukemia stages) may have significantly fewer samples than more common conditions. Techniques such as the Synthetic Minority Oversampling Techniques (SMOTE) were employed to even out the classes in text data in order to keep models based on machine learning from becoming biased against the dominant class. By guaranteeing that each model was trained on high-quality, consistent and balanced data the preprocessing pipeline ultimately aided in early disease detection and personalized health analysis producing more reliable, accurate and useful healthcare forecasts.

3. *Clustering-Based Health Profiling*

The suggested approach (as shown in Fig.2) groups people according to their health attributes as determined by smartwatch data using K-Means and Agglomerative Clustering. Features including heart rate, oxygen levels in the blood, step count, sleeping patterns, consumption of calories, levels of stress and activity duration are all included in the dataset. These characteristics are essential to evaluating overall health trends and recognizing potential illnesses at an early stage. However, because raw watch data often contains noise and inconsistencies preprocessing techniques including scaling of features (Min-Max Normalization), identifying outliers and imputation of missing values are employed to normalize the data. PCA or principal component analysis and other methods for lowering dimensionality are used to boost clustering performance by reducing feature duplication. The well-known centroid-based clustering technique K-Means separates data into K clusters based on feature similarity. In order to help with risk estimation and personalized health monitoring, K-Means is utilized in this work to divide people into multiple health categories. Following a random setup of K centroids the algorithm allocates each data point to the nearest centroid by iteratively changing its position parameters until resolution. The Elbow Method determines the optimal number of clusters by comparing the within-cluster average of the sum of square (WCSS) [28] versus the overall number of clusters.The perfect K value is shown by the “elbow” the point at which WCSS begins to level out. By identifying healthy, moderate-risk and high-risk individuals the clustering method (as shown in Fig.3). A hierarchical clustering method called agglomerative clustering [29] is also used to evaluate how well it groups people according to their health metrics. Agglomerative Clustering begins with every point of data as a separate cluster and repeatedly merges the closest groups based on linking criteria (e.g. Ward’s approach, single connection or complete linkage) in contrast to K-Means which necessitates a predetermined number of clusters. This approach is computationally demanding and does not scale well with huge datasets despite the fact that it offers an organization of health profiles. Agglomerative Clustering is less effective than K-Means for actual time health tracking because of its hierarchical structure which makes it difficult to handle dynamic health data updates. The effectiveness of both approaches to clustering is evaluated using the Silhouette Score [30]. With the Silhouette Score of 0.85 for K-Means which measures cluster unity and division compact and distinct clusters are predicted. However, as seen by its higher Silhouette Score, agglomerative grouping performs worse as it comes to gathering complex health data.

4.Brain Disease Detection

Early detection is crucial for the successful treatment of complex neurological conditions like Alzheimer's illness and other forms of dementia. Brain MRI scans can be divided into four groups using the proposed system (Fig. 4): No Dementia, Moderate Dementia, Very Mild Dementia, and Moderate Dementia. It accomplishes this via the use of deep learning models, namely VGG19, ResNet50, and Inception V3 [31]. The dataset demonstrates the class difference in real medical imaging data with 3,200 images of individuals without dementia, 2,240 images of individuals with severe Alzheimer's disease, 896 images of individuals with mild Alzheimer's disease, and 64 images of individuals with moderate dementia. Data augmentation methods including picture flipping, motion, enhancing the contrast and zooming are used to increase model generalization in order to lessen this problem. Preprocessing techniques including picture normalization, noise reduction and scaling to 224×224 pixels are carried out prior to putting the images into the models used for deep learning. By taking these precautions, the input data is kept consistent and of high quality enabling the models to extract reliable characteristics from MRI scans [32]. Three powerful deep neural network models Inception V3, ResNet50, and VGG19 are utilized to precisely identify brain disorders. Inception V3, which is well-known for its efficient feature extraction through Inception modules is the most comprehensive model in terms of both learning and generalization. It attains 90.94% test accuracy and 98.94% training accuracy. ResNet50, designed to handle deep network training via residual learning, exhibits some overfitting with a training accuracy of 95.60% and a testing accuracy of 79.55%. However, because of its deep sequenced architecture of 19 layers, VGG19 surpasses both models (as illustrated in Fig. 4) in training with a precision of 99.42% and retains a great test efficiency of 93.14%.

These deep learning models offer a potent, automated and non-invasive approach for early diagnosis when used in brain disease detection. The method reduces the need for manual diagnosis which may be laborious and prone to human mistake, by empowering medical practitioners to evaluate MRI data with high precision. Additionally, the use of Grad-CAM technology (Gradient-weighted Class Activation Map) [33] facilitates visual interpretability by helping physicians identify which MRI scan regions impacted the model's conclusion. By being categorized according to No Dementia, Very Moderate Dementia, Mild Dementia or

Moderate Dementia patients can receive customized treatment planning. By recommending specific medications, lifestyle modifications and early activities, this enhances patient outcomes. These model's excellent accuracy makes them a valuable addition to clinical procedures enhancing the early detection and management of diseases of the brain.

5.Skin Disease Detection

One of the most prevalent health issues in the world is skin disease which includes both benign and cancerous illnesses. Effective therapy depends on an early and precise diagnosis particularly for diseases like squamous cell carcinoma and melanoma that can become fatal if not identified in a timely manner. The proposed method employs a model based on Convolutional Neural Networks (C to classify nine distinct types of skin diseases: actinic keratosis, basal cell carcinoma, skin fibroid, Melanoma, nevus, colored benign acne, seborrhea keratosis, squamous cell carcinoma and vascular lesions [34]. To ensure a fair training process, the dataset's 5,000 pictures are divided equally among these categories. Because skin lesion photos can vary in terms of illumination, resolution and skin tones, preprocessing techniques such boosting contrast, equalizing histograms and decreasing noise are employed. The photos are then resized to 224×224 pixels to ensure optimal feature extraction and set up the input dimensions for the CNN model. The CNN model used to categorize skin diseases consists of many layers of convolution, maximum pooling layers, and fully linked levels. Its goal is to efficiently recognize and remove intricate patterns from pictures of skin lesions. The initial layers gather fundamental texture and color information, while the deeper levels focus on disease-specific characteristics including irregular shapes, imbalance and border definitions all crucial components for distinguishing benign from illnesses like melanoma.

To prevent excess fitting and improve the model's generalization ability, strategies including dropout normalization and batch regularization are employed. The model is trained using an optimizer based on Adam with a classifying entropy crossing loss function because the task entails multi-class classification. The CNN model's ability to recognize complex skin patterns is demonstrated by its impressive 95% accuracy rate after extensive training and validation (Fig. 5). The validation accuracy remains constant, demonstrating the model's good generalization on unknown input. Dermatology and telemedicine will be significantly impacted if an AI-powered skin disease diagnosis technology is successfully implemented. The model is a useful tool for robotic, real-time skin condition detection because of its high accuracy and dependability which relieves dermatologists of some of their workload and allows for quicker second opinions. By integrating this approach into telemedicine platforms or mobile applications users can upload pictures of skin lesions and get immediate risk assessments and AI-driven diagnostics. Furthermore, by emphasizing the key areas of an image that affected the model's prediction Grad-CAM visualization [35] offers interpretability and facilitates medical practitioners' validation of AI-generated results. Early detection of skin illnesses guarantees prompt intervention and treatment, which may save lives especially in situations of malignant conditions like melanoma. This approach may become a common diagnostic tool in hospitals and online healthcare services with additional actual testing and medical validation improving patient access to dermatological care globally.

6.Lung Disease Detection

Pneumonia and other lung illnesses are major global health concerns that frequently need to be diagnosed and treated very once to avoid serious repercussions. Early identification is essential for successful medical management because pneumonia in particular induces irritation of the lungs and fluid accumulation. Conventional diagnostic techniques including clinicians, interpretation of chest X-rays [36], can be laborious and prone to error by humans. Among the 5,216 training images in the dataset 3,815 X-rays are classified as pneumonia and 1,341 are deemed normal (healthy). There are 390 cases of pneumonia and 234 normal cases among the 624 photos in the testing collection. Preprocessing methods including contrast enhancement, picture normalization, and scaling (to 224×224 pixels) are used to increase model accuracy and guarantee constant input quality because X-ray images frequently exhibit changes in contrast, brightness and noise. In order to extract and learn significant characteristics from chest

X-ray pictures the CNN model utilized for lung disease diagnosis includes numerous convolutional layers, maximum pooling layers and entirely connected layers. While deeper levels detect more intricate lung structures and anomalies linked to pneumonia, the first layers catch simple edges and textures. Dropout layers are utilized to avoid overfitting and batches of normalization is used to enhance generalization and accelerate convergence. Since the problem involves binary classification, the model is developed using an optimizer named Adam with a cross-entropy binary loss function.The CNN model's 98.16% training achievement rate, 100% validation accuracy, and 92.15% reliability in tests following prolonged training show how reliable and effective it is at diagnosing pneumonia. The model's strong real- world applicability is ensured by its high validation accuracy which indicates that it generalizes effectively to unknown data (as shown in Fig.6). Telemedicine, automated healthcare systems and early pneumonia diagnosis are all significantly impacted by the AI-powered lung condition detection system. Patients and physicians can instantly upload images of their chests and receive AI-driven diagnoses by incorporating this model into telemedicine platforms, smartphone applications or hospital workflows. This reduces diagnostic delays and speeds up medical intervention. This model's utility in the diagnosis of respiratory diseases [37] can be increased by further training it on a variety of datasets to identify more lung conditions like tuberculosis or pneumonia linked to COVID-19. With more clinical testing and approval from regulators, this artificial intelligence system could become an essential diagnostic tool, particularly in areas with little resources and a lack of qualified radiologists. This would ensure that lung diseases are identified more quickly and easily worldwide.Reducing mortality rates necessitates early identification and avoidance of heart disease, which remains one of the leading causes of death worldwide. Traditional methods of diagnosing heart illness include cardiac echocardiograms, electrocardiograms (ECGs), and blood tests [38]. Despite their benefits, these processes can be costly and time-consuming. The suggested system analyzes structured textual health data and makes very accurate predictions about the risk of heart disease by utilizing machine learning techniques particularly a Multilayer Perceptron (MLP) model (as shown in Fig.7). Numerous medical parameters, including age, cholesterol, blood pressure, heart rate change, glucose levels, and prior heart disorders, are included in the dataset to assess cardiovascular health. To guarantee consistency and precision in the model's predictions data preprocessing methods such handling missing values, scaling features and outlier removal are used. Methods like Synthetic Minority Over-sampling (SMOTE) are employed to balance the dataset because textual health data frequently contains unbalanced distributions. This ensures that the model does not favor any particular class. A feedforward neural network (ANN) known as the Multilayer Perceptron (MLP) architecture is used to classify cardiac disease. The design consists of an input layer, multiple hidden layers with ReLU activation, and an output layer with a function to activate that classifies using softmax. While the input layer receives organized health data, the hidden layers learn intricate correlations between medical parameters, enabling a thorough understanding of cardiac risk factors.The model is created using the optimizer developed by Adam with a loss of cross-entropy because determining whether cardiac disease is present or not is a binary task. The MLP model offers excellent predictive power and reliability when it involves diagnosing cardiovascular diseases from textual medical information, as evidenced by its astounding 99.17% accuracy following validation and training. The model's efficiency and generalization abilities are further improved by hyperparameter tweaking which includes maximizing training rates, dropout periodicity and batch sizes. Cardiovascular diagnostics could undergo a revolution if AI-driven heart disease diagnosis is included into clinical decision-making systems. Hospitals, wearable technology and telemedicine platforms can all use this paradigm to give patients immediate risk evaluations based on their medical records. The system may analyze daily health patterns and identify early warning symptoms of heart disease by evaluating real-time information from wearable technology, such as smartwatches. This enables users to promptly seek medical assistance. Medical practitioners can also utilize this AI model as a tool for decision-making to enable early intervention, customized treatment plans, and risk factor changes. To ensure wider applicability in international healthcare settings, further developments will incorporate deep learning techniques combine textual factors with ECG data and expand the dataset to encompass a wider range of demographics. This MLP-based model for predicting heart illness is a potent and effective diagnostic tool that could lessen the prevalence of cardiovascular disorders worldwide.The deadly blood malignancy known as acute lymphoblastic leukemia (ALL) mainly affects white blood cells and causes immature lymphoblasts to proliferate out of control. To guarantee prompt medical intervention and raise survival chances, early identification is crucial. Peripheral blood smear (PBS) images must be examined under a microscope in order to diagnose ALL, which is a laborious process that heavily relies on skilled pathologists. The suggested approach classifies blood sample images into three key ALL stages Early Pre-B, Pre-B and Pro-B using a deep learning-based Convolutional Neural Network (CNN) framework [39] in order to improve the precision and effectiveness of leukemia detection. With 3,242 blood smear photos taken from 89 people the dataset guarantees a representative and varied training sample. Preprocessing techniques including grayscale conversion, contrast enhancement, noise reduction, and scale (to 224×224 pixels) are employed to enhance image quality and standardization before being fed into the CNN model. The CNN model is designed to effectively extract complex morphological information from blood smear images in order to reliably discriminate across ALL stages. The design consists of many convolutional layers, followed by optimal pooling layers for reducing dimensionality, fully linked layers for classification, and ReLU activation features. The model is created using an Adam optimization algorithm with a class cross-entropy loss function because the categorization is divided into multiple categories. To enhance model generalization and avoid overfitting data augmentation methods including picture rotation, flipping and magnification are used. The CNN model's remarkable 95% accuracy after intensive training and validation shows how reliable it is at differentiating between leukemia stages (as shown in Fig.8). Strong generalization is indicated by the model's high validating and test accuracy which guarantees accurate forecasts even on unknown blood smear images. Hematology and oncology will be significantly impacted by the incorporation of AI-driven blood tumor diagnosis. Hospitals and testing facilities can use this CNN-based technology to diagnose leukemia automatically with little assistance from humans. The program assists pathologists and oncologists in making quicker more accurate diagnoses and treatment decisions by quickly assessing blood smear pictures. Additionally, remote leukemia screening can be made possible by integrating this model into telehealth platforms especially in developing nations where access to qualified hematologists is scarce. Future developments include integrating explainable AI approaches for model interpretability merging deep learning with biological biomarkers to further improve leukemia classification and strengthening model resilience through training on larger and more varied datasets. Our CNN-powered cancer detection method has a chance to revolutionize blood cancer diagnostics by ensuring early and accurate detection, improved therapy planning, and improved patient outcomes.

IV. RESULTS

The performance of the proposed AI-driven healthcare system was thoroughly evaluated using a number of deep learning as well as machine learning methods. Precision, recall, accuracy, F1-score, and silhouette scores were important performance metrics. The Elbow Method was used to examine the clustering frameworks using K-Means and Agglomerative Clustering figuring out the ideal number of clusters based on the health features of the individuals. When compared to Agglomerative Clustering, K-Means showed better clustering ability with a silhouette score of 0.85. The deep learning models showed impressive results in disease classification, VGG19 (99.42% train, 93.14% test), ResNet50 (95.60% train, 79.55% test) and Inception V3 (98.94% train, 90.94% test) for brain disease identification. Similarly, the Decision Tree Classification for liver illness achieved 98% accuracy using textual data, while CNN algorithms for skin, lung and plasma cancer classification obtained 95%, 98.16%, and accuracy levels of 95%, respectively (as shown in Fig.9). The suggested artificial intelligence system is made to function flawlessly in a Streamlit-based user interface (UI), making it easy for people to engage with the platform. Users can manually enter medical variables associated with certain disorders or enter their real-time medical information from wearables. Users submit MRI scans to diagnose brain diseases and deep learning models interpret the scans to categorize the illness into one of four groups, Moderate Dementia, Very Moderate Dementia, Mild Dementia or No Dementia. Clinical decision-making is aided by the system's comprehensive diagnostic insights which include highlighted areas of concern in MRI scans using Grad-CAM representations.Users can upload photos of skin lesions for the purpose of detecting skin diseases and a CNN model will analyze the photographs and categorize them into one of nine skin disease types. For every class the model produces probability ratings that show the most probable diagnosis and visual justifications for the model's reasoning. Similar to this, customers can upload chest X-ray photos to the lung cancer detection module which accurately classifies them as either healthy or pneumonia. The model's emphasis areas are further explained by heatmap visualizations which provide diagnostic transparency. Peripheral Blood Smear (PBS) images are analyzed by the blood disease detection system, which then categorizes them into cancer stages. After classification, the approach gives pathologists and physicians a summary of the disease's course allowing them to determine its severity and adjust treatment plans accordingly. By reducing human error and expediting the leukemia diagnosing process an AI-driven categorization guarantees that patients receive timely and efficient interventions. Users input important clinical information, including blood pressure, glucose, cholesterol, heart rate variation and other cardiovascular indicators, in order to diagnose heart disease. This information is processed by the Multilayer Perceptron (MLP) model, which determines a person's low, moderate or high risk of cardiovascular disease. The system creates a customized heart health report based on the risk group, offering dietary advice, lifestyle modifications and medical testing for additional assessment.A Decision Tree Classifier is used to analyze textual data in order to diagnose liver disease; it has a 98% classification accuracy for liver health status. The system categorizes users into healthy, at-risk, or unhealthy groups based on their input of bilirubin levels, hepatic enzyme values and other critical health markers. Preventive actions such as suggested dietary changes, exercise regimens and medical examinations for preserving liver function are displayed on the Streamlit interface after a diagnosis. A daily wellness tracking system is also included into the Streamlit dashboard which tracks smartwatch data continually and notifies users of health trends. Users can take preventative action before issues occur by receiving alerts and warnings if their medical parameters diverge from typical values. In order to help users and healthcare professionals examine long-term trends to make well-informed medical decisions the system also keeps track of past medical records (as shown in Fig.10). The system's capacity to provide preventive measures in addition to disease diagnosis makes it special. In order to stop future deterioration, the system recommends breathing exercises, vaccination notifications and pollution exposure measures, for example, if an individual is at risk of pneumonia. Similar to this, users are given recommendations for UV protection, dermatology therapy and home remedies for skin disease diagnosis depending on the skin condition classification (as shown in Fig.11).

Additionally, the system promotes behavioral changes for liver and heart health by offering individualized exercise regimens, food recommendations and stress-reduction strategies (as shown in Fig.12). AI-powered chatbot integration also makes healthcare more accessible and participatory by enabling users to ask health-related queries and get immediate, evidence-based suggestions. The AI-driven healthcare system's scalability allows for future extensions such as integration with wearable technology, hospital databases, and electronic medical records (EMRs). Using multimodal data (e.g., MRI scans and laboratory test results for a more thorough diagnosis) and further enhancing model interpretability using explainable AI approaches are potential future improvements. The platform may also be expanded to detect additional diseases add real-time patient data to its collection, and improve preventive actions with cutting-edge AI-driven health analytics. By providing early disease identification, real-time monitoring and customized prevention tactics this AI-powered system seeks to transform healthcare and ensure a better future for people everywhere through ongoing research and model improvements.

V. CONCLUSION

By utilizing cutting-edge neural networks and machine learning algorithms the AI-driven healthcare system described in this study offers a revolutionary approach to early disease identification, real-time health monitoring and customized preventative actions. The system effectively determines brain diseases (dementia and Alzheimer's), diseases of the skin, infections of the lungs, cardiac conditions, blood-related conditions (leukemia) and liver conditions with remarkable accuracy by integrating smartwatches health data, imaging analysis and narrative health records. Personalized health insights are made possible by the K-Means clustering methodology which guarantees that people are grouped according to their health characteristics. The deep learning techniques like CNN, Multilayer Perceptron, VGG19, ResNet50 and Inception V3 achieved high classification accuracies demonstrating their dependability in disease prediction. The effectiveness of textual data analysis in healthcare was further illustrated by the use of the Decision Tree Classifier for liver illness and the MLP for heart disease. Users may easily enter health information, upload medical photos and obtain AI-powered diagnostics and preventive health advice thanks to the system's interactive Streamlit user interface. This research provides a non-invasive, effective and easily accessible medical solution that enables people to take active decisions regarding their health lessens the strain on medical systems, and improves early disease intervention by fusing real-time monitoring of health with predictive analytics. In order to guarantee that this Intelligence-powered system keeps transforming preventative medicine and customized healthcare, future improvements will concentrate on increasing illness coverage integrating data from various sources fusion and enhancing model comprehension with explainable AI methodologies.

Conflict of interest statement

Authors declare that they do not have any conflict of interest.

References

[1] Kumar, Yogesh, et al. "Artificial intelligence in disease diagnosis: a systematic literature review, synthesizing frameworkandfutureresearchagenda."Journalofambientintelligenceandhumanizedcomputing14.7(2023):8459-8486.

[2] Krishna, BV Rama, L. L. S. Maneesha, and Arepalli Ramadevi. "An RFID based COVID Patient Health Care Monitoring System for Government Hospitals." Journal homepage:www.ijrpr. com ISSN 2582: 7421.

[3] Chaka,Brian,andM.Hardy."ComputerbasedsimulationinCTandMR Iradiographyeducation:currentrole and future opportunities." Radiography 27.2 (2021): 733-739.

[4] Mohan,GSSSSVKrishna,etal."HarnessingDeepNeuralNetworksfor AccuratePCOSDiagnosisfromMedical Images."2024InternationalConferenceonComputationalIntelligence forGreenandSustainableTechnologies (ICCIGST). IEEE, 2024.

[5] Paul-Chima,UgwuOkechukwu,etal."IntegratingWearableHealthMonitoringDeviceswithIoTforEnhanced Personal Health Management: A Comprehensive Review." Data Process 12 (2024): 13.

[6] Monfredi, Oliver J., et al. "Continuous ECG monitoring should be the heart of bedside AI-based predictive analytics monitoring for early detection of clinical deterioration." Journal of electrocardiology 76 (2023): 35-38.

[7] Shah, Syed Rehan, et al. "Comparing inception V3, VGG 16, VGG 19, CNN, and ResNet 50: a case study on early detection of a rice disease." Agronomy 13.6 (2023): 1633.

[8] Hassan,Mubashir,etal."Innovationsingenomicsandbigdataanalytics forpersonalized medicineandhealth care: A review." International journal of molecular Sciences 23.9 (2022): 4645.

[9] Elhadad,Ahmed,etal."Fogcomputingserviceinthehealthcaremonitoringsystemformanagingthereal-time notification." Journal of Healthcare Engineering 2022.1 (2022): 5337733.

[10] Dunn,Jessilyn,etal."Wearablesensorsenablepersonalizedprediction sofclinicallaboratorymeasurements." Nature medicine 27.6 (2021): 1105-1112.

[11] Gupta,VishanKumar,etal."PredictionofCOVID-19confirmed,death, andcuredcasesinIndiausingrandom forest model." Big Data Mining and Analytics 4.2 (2021): 116-123.

[12] Kavitha, C., et al. "Early-stage Alzheimer's disease prediction using machine learning models." Frontiers in public health 10 (2022): 853294.

[13] Dildar,Mehwish,etal."Skincancerdetection:areviewusingdeeplearningtechniques."Internationaljournal of environmental research and public health 18.10 (2021): 5479.

[14] Das, Pradeep Kumar, Biswajit Nayak, and Sukadev Meher. "A lightweight deep learning system for automatic detection of blood cancer." Measurement 191 (2022): 110762.

[15] Bhatt,ChintanM.,etal."Effectiveheartdiseasepredictionusingmachinelearningtechniques."Algorithms 16.2(2023):88.

[16] Goyal, Shimpy, and Rajiv Singh. "Detection and classification of lung diseases for pneumonia and Covid-19 usingmachineanddeeplearningtechniques."JournalofAmbientIntelligenceandHumanizedComputing14.4 (2023): 3239-3259.

[17] Banapuram, Chandini, et al. "A Comprehensive Survey of Machine Learning in Healthcare: Predicting Heart andLiverDisease,TuberculosisDetectioninChestX-Ray Images." SSRGInternational JournalofElectronics and Communication Engineering 11.5 (2024): 155-169.

[18] Kim, Min-Jeong. "Building a cardiovascular disease prediction model for smartwatch users using machine learning: Basedon the Korea national healthandnutrition examination survey." Biosensors11.7 (2021): 228.

[19] Şengül, Gökhan, et al. "Deep learning based fall detection using smartwatches for healthcare applications." Biomedical Signal Processing and Control 71 (2022): 103242.

[20] Tarafdar, Pratik, and Indranil Bose. "Recognition of human activities for wellness management using a smartphone and a smartwatch: a boosting approach." Decision Support Systems 140 (2021): 113426.

[21] Fan,Cheng,etal."Areviewondatapreprocessingtechniquestowardefficientandreliableknowledgediscovery from building operational data." Frontiers in energy research 9 (2021): 652801.

[22] Li,Boyi,etal."Onfeaturenormalizationanddataaugmentation."ProceedingsoftheIEEE/CVFconferenceon computer vision and pattern recognition. 2021.

[23] Sinsomboonthong, Saichon. "Performance Comparison of New Adjusted Min-Max with Decimal Scaling and Statistical Column Normalization Methods forArtificial Neural Network Classification." International Journal of Mathematics and Mathematical Sciences 2022.1 (2022): 3584406.

[24] Boddapati,MohanSaiDinesh,etal."CreatingaProtectedVirtualLearningSpace:AComprehensiveStrategy forSecurityandUserExperienceinOnlineEducation."InternationalConferenceonCognitiveComputingand Cyber Physical Systems. Cham: Springer Nature Switzerland, 2023.