International Research Journal of Engineering and Technology (IRJET) e-ISSN: 2395-0056
Volume: 11 Issue: 06 | Jun 2024 www.irjet.net p-ISSN: 2395-0072

Multi-Disease Prediction Web Application: A Step Towards Integrated Healthcare Diagnosis and Prevention.

Dr. Subha Subramaniam
HoD, Assistant Professor Department of Electronics and Computer Science Shah & Anchor Kutchhi Engineering college Chembur, India

Biswajit Mohanty
Department of Electronics and Computer Science Shah & Anchor Kutchhi Engineering college Chembur, India

Divyesh Pahurkar
Department of Electronics and Computer Science Shah & Anchor Kutchhi Engineering college Chembur, India

Pranav Jha
Department of Electronics and Computer Science Shah & Anchor Kutchhi Engineering college Chembur, India

Roshan Bakka
Department of Electronics and Computer Science Shah & Anchor Kutchhi Engineering college Chembur, India

---

***Abstract*— Early disease identification and correct diagnosis are extremely difficult for the healthcare industry to achieve, which frequently leads to treatment delays, higher medical expenses, and worse patient outcomes. In order to solve these issues, this project proposes the creation of a disease prediction model that makes use of convolutional neural networks (CNNs) and machine learning techniques.**

In order to accurately diagnose six prevalent diseases using medical data, this research outlines the construction of a disease prediction model that makes use of convolutional neural networks (CNNs) and machine learning techniques. Targeted illnesses include brain tumors, diabetes, Alzheimer's disease, heart disease, pneumonia, and breast cancer.

The work centers on two main approaches:

1. Conventional machine learning algorithms for illnesses with structured data (like diabetes, heart disease), and

2. CNN architectures for illnesses identified by imaging tests (such brain tumors, pneumonia, and breast cancer).

***Keywords*— Multi-disease detection, Diagnosis, Healthcare, treatment, Cancer, Disease, CNN, ConvNet, XGboost.**

I. INTRODUCTION

Advances in technology and the growing accessibility of healthcare data are driving constant change in the healthcare industry. Despite these progress, global healthcare systems continue to face obstacles such postponed illness identification, imprecise diagnosis, and rising medical expenses. Accurate diagnosis and early detection are essential for managing diseases effectively and enhancing patient outcomes. This project aims to use machine learning techniques and convolutional neural networks (CNNs) to construct a robust disease prediction model in response to these obstacles.

In order to diagnose six major diseases—pneumonia, breast cancer, diabetes, Alzheimer's disease, heart disease, and brain tumors—this research aims to develop a flexible and precise illness prediction tool.

Our goal is to improve the effectiveness and precision of disease detection by utilizing various datasets that include medical images and related metadata. This will allow us to fully utilize the power of machine learning algorithms and CNNs.

II. RESEARCH METHODOLOGY

The research methodology for a Multi-Disease Prediction Web Application involves data collection, preprocessing, and model training. Data from medical records and imaging are curated, anonymized, and cleaned to ensure quality. Various machine learning algorithms, particularly deep learning models like CNNs and RNNs, are trained on labeled datasets for multiple diseases. Model performance is validated using cross-validation and fine-tuned for accuracy. The final model is deployed in a web app for real-time disease prediction and integrated healthcare diagnostics.

CNN and machine learning algorithms were utilized in this project to train the model on the dataset and forecast the disease. When MRI scans or chest X-rays are used as inputs for disorders like pneumonia, brain tumors, and alzheimer's, CNN is used to determine whether or not the patients have these conditions. To train the model for additional diseases, machine learning algorithms such as

random forest and XG Boast are utilized. The dataset is preprocessed and filtered to get the highest level of accuracy.

A **Convolutional Neural Network (ConvNet/CNN)** is a Deep Learning algorithm that can take in an input image, assign importance (learnable weights and biases) to various aspects/objects in the image, and be able to differentiate one from the other.

The pre-processing required in a ConvNet is much lower as compared to other classification algorithms. While in primitive methods filters are hand-engineered, with enough training, ConvNets have the ability to learn these filters/characteristics.

The architecture of a ConvNet is analogous to that of the connectivity pattern of Neurons in the Human Brain and was inspired by the organization of the Visual Cortex.

The **Random Forest algorithm** is an ensemble learning technique that constructs multiple decision trees and combines their predictions to improve accuracy and reduce overfitting. It operates by creating multiple bootstrap samples from the original training data and growing a separate decision tree for each sample. During the tree-building process, the algorithm randomly selects a subset of features at each node, introducing randomness and reducing the correlation between individual trees. To make a prediction, the new instance is passed through each tree in the forest, and the final prediction is determined by majority vote (for classification) or by averaging the predictions (for regression).

XGBoost (Extreme Gradient Boosting) is a highly efficient and optimized implementation of the gradient boosting algorithm. It is an ensemble learning technique that combines multiple weak decision tree models in an iterative manner to create a strong predictive model. XGBoost builds upon the concept of gradient boosting by introducing several enhancements and optimizations, such as regularization techniques to prevent overfitting, parallel processing capabilities for efficient handling of large datasets, and advanced features like missing value handling and built-in cross-validation.

III. FLOWCHART

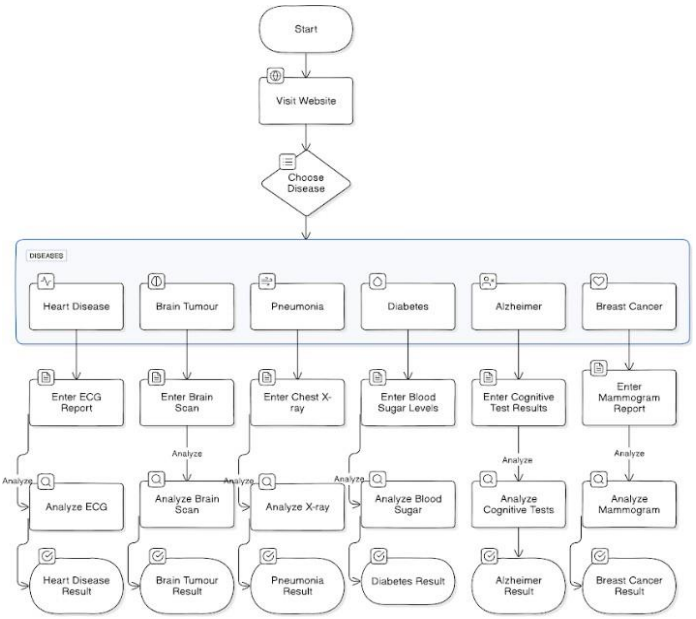

Fig. 1. Model Flow.

IV. RESULT & DISCUSSION

Brain Tumour Detection

A brain tumor detection web app using machine learning employs advanced algorithms to analyze medical images (such as MRI scans) for early diagnosis. The app leverages convolutional neural networks (CNNs) to identify tumor presence, type, and location with high accuracy. Users upload scans, which the system processes to highlight abnormal regions. This technology enhances diagnostic speed and accuracy, assisting healthcare professionals in making informed decisions. The app often includes features like real-time analysis, detailed reporting, and integration with medical databases, ensuring comprehensive and efficient tumor detection and monitoring.

Used VGG-16 for feature extraction. Used custom- made CNN .The accuracy achieved was around 100% (just tested on 10 images).

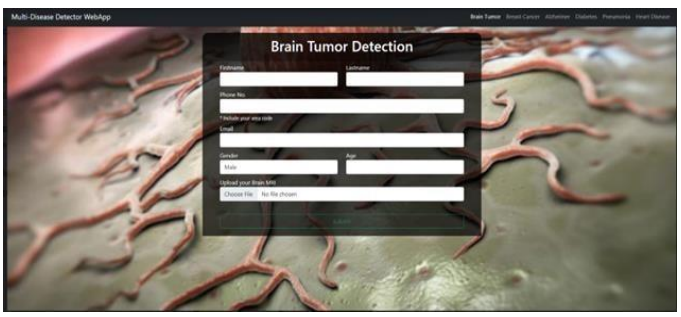

Fig. 2. Brain Tunour Detection

Breast Cancer Detection

A breast cancer detection web app leverages machine learning to assist in early diagnosis. It uses trained models, often involving deep learning and convolutional neural networks (CNNs), to analyze mammogram images or histopathological data. Users upload medical images, which the app processes to identify patterns indicative of cancerous cells. The system provides probability scores or classifications (benign or malignant), aiding radiologists in decision-making. Incorporating features like user-friendly interfaces, secure data handling, and integration with medical databases enhances its utility. The app aims to improve diagnostic accuracy, reduce workload for healthcare professionals, and ultimately enhance patient outcomes through early detection.

Used Random Forest for this use case.

The accuracy achieved was around 91.81%

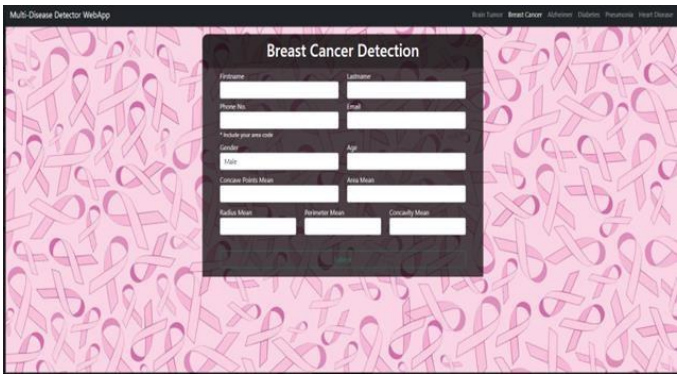

Fig. 3. Breast Cancer Detection

Alzimer Detection

An Alzheimer's detection web app uses machine learning to identify early signs of the disease from medical data such as MRI scans, genetic profiles, and cognitive test results. By analyzing patterns and anomalies in this data, the app can predict the likelihood of Alzheimer's, providing early intervention opportunities. The web app typically features user-friendly interfaces for data input, results visualization, and reporting. It leverages algorithms like convolutional neural networks (CNNs) for image analysis and various statistical models for other data types. This tool aids healthcare professionals in making informed decisions, potentially improving patient outcomes and managing the disease more effectively.

Trained CNN architecture for this use case. The accuracy achieved was around 73.54%

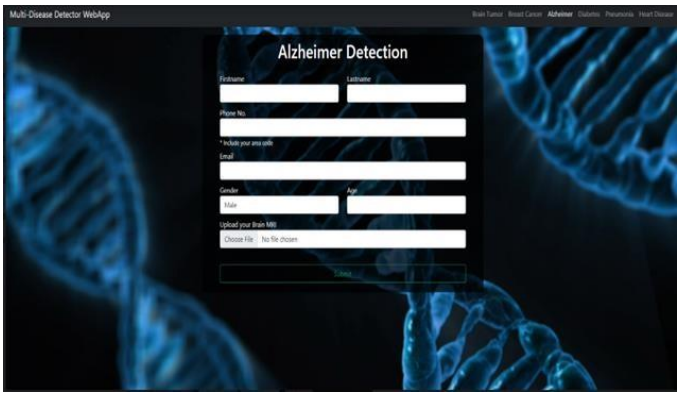

Fig. 4. Alzimer Detetion

Heart Disease Detection

A machine learning webapp for heart disease detection leverages algorithms to analyze patient data, such as age, blood pressure, cholesterol levels, and ECG results. It predicts the likelihood of heart disease by training on historical medical records and identifying patterns associated with the condition. Users input their health metrics, and the app provides a risk assessment, offering personalized insights and potentially early detection. Such tools aim to support healthcare professionals in decision-making and improve patient outcomes through data-driven predictions.

Used XGBoost for this use case.

The accuracy achieved was around 86.96%

Fig. 5. Heart Disease Detection

Pneumonia Detection

A Pneumonia Detection web app leverages machine learning to analyze chest X-ray images for signs of pneumonia. The app uses a trained convolutional neural network (CNN) model to classify images as either normal or indicative of pneumonia. Users upload X-ray images, and the app processes them to provide a diagnosis within seconds, significantly aiding in early detection and treatment. This technology enhances diagnostic accuracy,

reduces the workload for radiologists, and ensures timely medical intervention. The app is particularly useful in areas with limited access to healthcare professionals, providing an accessible and reliable diagnostic tool.

Used custom CNN architecture for this use case. The accuracy achieved was around 83.17%.

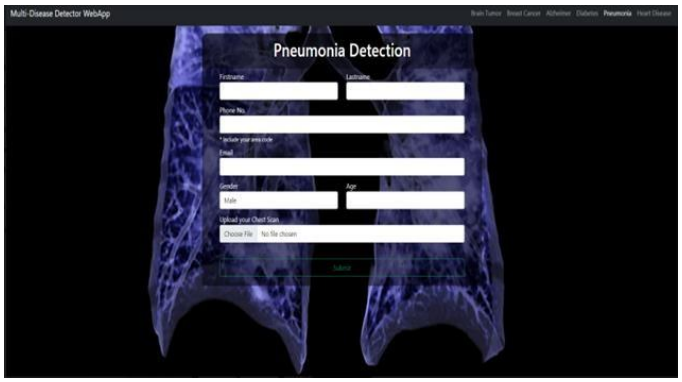

Fig.6. Pneumonia Detection

V. CHALLENGES AND FUTURE DIRECTIONS

Data Quality and Quantity: Collecting big and diverse datasets for each disease is critical for developing accurate prediction models. However, gathering such data while maintaining its quality, privacy, and representativeness can be difficult.

Complexity of Algorithms: Developing robust machine learning algorithms capable of reliably predicting numerous diseases necessitates extensive model and feature engineering. Each disease may have distinct diagnostic criteria and prediction signs, which adds to the complexity.

Interpretability: To acquire the trust of users and healthcare professionals, prediction models must be transparent and interpretable. Complex machine learning models may lack interpretability, making it difficult to understand the logic behind predictions.

Integration with Healthcare Systems: Integrating the online application with existing healthcare systems and electronic health records (EHRs) while meeting regulatory criteria.

VI. CONCLUSION

While the disease prediction system we developed has shown promising results in detecting and predicting various critical diseases, we acknowledge that there is still room for improvement and further research. The integration of advanced machine learning techniques has enabled accurate predictions, but the true impact lies in responsible and ethical deployment in real-world healthcare settings.

We hope that our efforts in developing this system will contribute to the broader goal of enhancing patient care and improving clinical outcomes. However, we recognize that this is merely a step in a continuous journey towards leveraging technology for the betterment of human health and well-being. Collaboration with domain experts, regulatory bodies, and the wider medical community is crucial to ensure that our system aligns with best practices and meets the highest standards of safety and efficacy.

VII. REFERENCES

1. Litjens, G., Kooi, T., Bejnordi, B. E., Setio, A. A. A., Ciompi, F., Ghafoorian, M., ... & Sánchez, C. I. (2017). A survey on deep learning in medical image analysis. Medical image analysis, 42, 60-88.

2. Esteva, A., Kuprel, B., Novoa, R. A., Ko, J., Swetter, S. M., Blau, H. M., & Thrun, S. (2017). Dermatologist-level classification of cancer with deep neural networks. Nature, 542(7639), 115-118.

3. Ren, S., He, K., Girshick, R., & Sun, J. (2015). Faster r-cnn: Towards real-time object detection with region proposal networks. Advances in neural information processing systems, 28.

4. Simonyan, K., & Zisserman, A. (2014). Very deep convolutional networks for large-scale image recognition. arXiv preprint arXiv:1409.1556.

5. Breiman, L. (2001). Random forests. Machine learning, 45(1), 5-3.

6. Chen, T., & Guestrin, C. (2016). Xgboost: A scalable tree boosting system. Proceedings of the 22n