Industry Oriented Mini Project Report
on

INSTANT HEALTH SCAN USING ML & DL

Submitted in partial fulfillment of the requirements
for the award of degree of

BACHELOR OF TECHNOLOGY
in
Information Technology
by

Kolli Renu Sreeja(20WH1A1201)
Alane Shivani(20WH1A1240)
Kolusu Sahithi(20WH1A1256)

Under the esteemed guidance of
Mr. B.Srinivasulu
Assistant Professor

Department of Information Technology
BVRIT HYDERABAD College of Engineering for Women
Rajiv Gandhi Nagar,Nizampet Road, Bachupally,Hyd–500090
(Affiliated to Jawaharlal Nehru Technological University, Hyderabad)
(NAAC 'A' Grade & NBA Accredited- ECE, EEE, CSE & IT)

Dec 5,2023

BVRIT HYDERABAD
College of Engineering for Women
Rajiv Gandhi Nagar, Nizampet Road, Bachupally, Hyderabad – 500090
(Affiliated to Jawaharlal Nehru Technological University Hyderabad)
(NAAC 'A' Grade & NBA Accredited- ECE, EEE, CSE & IT)

CERTIFICATE

This is to certify that the Project report on **" Instant Health Scan using ML & DL"** is a bonafide work carried out by **K.Renu Sreeja (20WH1A1201), A.Shivani (20WH1A1240) and K.Sahithi (20WH1A1256)** in the partial fulfillment for the award of B.Tech degree in **Information Technology , BVRIT HYDERABAD College of Engineering for Women, Bachupally, Hyderabad** affiliated to Jawaharlal Nehru Technological University, Hyderabad, under my guidance and supervision. The results embodied in the project work have not been submitted to any other university or institute for the award of any degree or diploma.

Internal Guide	**Head of the Department**
Mr. B. Srinivasulu	**Dr. Aruna Rao S L**
Assistant Professor	**Professor & HoD**
Department of IT	**Department of IT**

External Examiner

DECLARATION

We hereby declare that the work presented in this project entitled **"Instant Health Scan using ML & DL"** submitted towards completion of in IV year I sem of B.Tech IT at "BVRIT HYDERABAD College of Engineering for Women",Hyderabd is an authentic record of our original work carried out under the esteemed guidance of **Mr. B. Srinivasulu, Assistant Professor**, Department of Information Technology.

K.Renu Sreeja (20WH1A1201)

A. Shivani (20WH1A1240)

K.Sahithi (20WH1A1256)

This project report is dedicated to my beloved Family members and supervisor for their limitless support and encouragement and to you as a reader

ACKNOWLEDGMENTS

We would like to express our profound gratitude and thanks to **Dr. K. V. N. Sunitha, Principal, BVRIT HYDERABAD College of Engineering for Women** for providing the working facilities in the college.

Our sincere thanks and gratitude to **Dr. Aruna Rao S L, Professor & Head, Department of IT, BVRIT HYDERABAD College of Engineering for Women** for all the timely support, constant guidance and valuable suggestions during the period of our project.

We are extremely thankful and indebted to our internal guide, **B. Srinivasulu, Assistant Professor, Department of IT, BVRIT HYDERABAD College of Engineering for Women** for his constant guidance, encouragement and moral support throughout the project.

Finally, we would also like to thank our Project Coordinators **K.S. Niraja, Assistant Professor and A.Aruna Jyothi, Assistant Professor**, all the faculty and staff of Department of IT who helped us directly or indirectly, parents and friends for their cooperation in completing the project work.

K.Renu Sreeja (20WH1A1201)

A.Shivani (20WH1A1240)

K.Sahithi (20WH1A1256)

ABSTRACT

The healthcare sector has emerged as a key beneficiary of technological advancements. The Covid-19 pandemic highlighted the challenges of limited access to definitive outcomes. As industries rapidly embraced digital tools and transformative technologies, healthcare faced the hurdle of physical hospital visits and online consultations. Our project seeks to remedy this issue by replacing the need for doctor consultations with faster, more precise results. Our focus lies on implementing detection systems for seven diseases, encompassing infectious diseases, neurological conditions, chronic ailments, and cancer. This selection addresses a diverse range of health concerns, reflecting a comprehensive approach. To achieve this, we are integrating both machine learning and deep learning techniques. This hybrid strategy leverages the strengths of each technique, promising a robust disease detection framework. Ultimately, our initiative aligns with the evolving healthcare landscape, harnessing technology to enhance access and accuracy while alleviating the burdens posed by traditional healthcare practices.

Keywords: Machine Learning, Convolutional Neural Networks

Contents

Declaration **i**

Acknowledgement **iii**

Abstract **i**

1 Introduction **1**
1.1 Machine Learning . . . 2
1.1.1 Random Forest . . . 2
1.1.2 XGBoost . . . 3
1.2 Deep Learning . . . 3
1.2.1 Artificial Neural Networks . . . 4
1.2.2 Deep Neural Networks . . . 5
1.2.3 Convolutional Neural Networks (CNNs) . . . 5
1.3 Motivation . . . 6
1.4 Objective . . . 6
1.5 Problem Definition . . . 6
1.6 Aim . . . 6

2 Literature Survey **8**

3 System Design **10**
3.1 Proposed System . . . 10
3.2 Architecture . . . 11
3.2.1 System Architecture . . . 11

4 System Modules **12**
4.1 Imaging Diagnostics (Alzheimer's, Brain Tumor, Pneumonia, COVID-19) . . . 12
4.2 Data-Driven Disease Prediction (Heart, Diabetes, Breast Cancer) . . . 12

5 Implemention **13**
5.1 Steps involved 13
5.2 Model Architecture 14

6 Results and Discussions **17**

7 Conclusions and future works **26**
7.1 Conclusion 26
7.2 Future scope 26

List of Figures

1.1 Instant Health Scan 2
1.2 A Random Forest Classifier 2
1.3 Flow Diagram of XGBoost Algorithm 3
1.4 ML as a subset of AI and DL as a subset of ML 4
1.5 An Artificial Neural Network with layer coloring 4
1.6 A generic CNN architecture 5

3.1 Architecture 11

5.1 COVID-19 14
5.2 Brain Tumor 15
5.3 Alzheimer 15
5.4 Firestore Storage 16

6.1 Home Page of Instant Health Scan 17
6.2 Loss and Accuracy Plot for Training of the COVID-19 detection model 18
6.3 Confusion Matrix for predictions of the COVID-19 detection model 18
6.4 Loss and Accuracy Plot for Training of the Brain Tumor detection model 19
6.5 Confusion Matrix for predictions of the Brain Tumor detection model 19
6.6 Loss and Accuracy Plot for Training of the Alzheimer detection model 20
6.7 Confusion Matrix for predictions of the Alzheimer detection model 20
6.8 Loss and Accuracy Plot for Training of the Pneumonia detection model 21
6.9 Confusion Matrix for predictions of the Pneumonia detection model 21

6.10 Loss and Accuracy Plot for Training of the Diabetes detection model 22
6.11 Loss and Accuracy Plot for Training of the Breast Cancer detection model 23
6.12 Loss and Accuracy Plot for Training of the Heart Disease detection model 24

List of Tables

6.1 Accuracy Table 25

Chapter 1

Introduction

Healthcare organizations, irrespective of their size, type, or specialty, are increasingly recognizing the potential of artificial intelligence (AI) to enhance patient care, reduce costs, and improve operational efficiency. The rapid growth in the availability and sophistication of AI has provided stakeholders, including providers and payers, with a plethora of tools and strategies to consider. Understanding the intricacies of how data is ingested, analyzed, and presented to end-users is crucial for managing expectations regarding accuracy and reliability. It also plays a pivotal role in determining the investments required to optimize an organization's data assets. To make informed choices between vendor products or when hiring data science staff to develop in-house algorithms, healthcare organizations must have a solid understanding of the various facets of artificial intelligence and their application to specific use cases. A logical starting point is exploring deep learning, a branch of artificial intelligence that has rapidly become transformative for healthcare. Deep learning enables the analysis of data with unprecedented speed and precision. While many headlines in the industry currently focus on small-scale pilots or research projects, deep learning is progressively making its way into practical tools with high-value applications in real-world clinical settings. Several promising use cases for deep learning include innovative patient-facing applications and strategies for enhancing the health IT user experience. Convolutional Neural Networks (CNNs), a type of deep learning, excel in analyzing images such as MRI results or X-rays. According to experts in computer science at Stanford University, CNNs are specifically designed for efficient image processing, allowing them to handle larger images with increased effectiveness. Some CNNs are even approaching or surpassing the accuracy of human diagnosticians in identifying crucial features in diagnostic imaging studies.

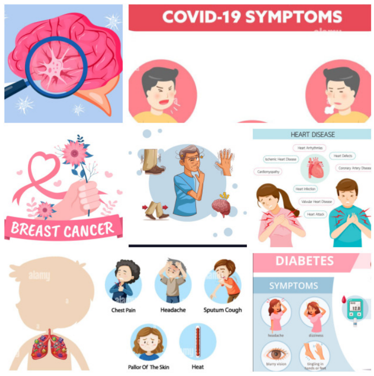

Figure 1.1: Instant Health Scan

1.1 Machine Learning

Machine learning (ML) is a subset of artificial intelligence focused on creating methods that leverage data to enhance task performance. ML algorithms, using training data, build models for making predictions or decisions without explicit programming.

1.1.1 Random Forest

Random forests, an ensemble learning method, construct multiple decision trees during training. For classification, the output is the majority class, and for regression, it's the average prediction. Random forests mitigate overfitting through bootstrap aggregating (bagging), where trees are trained on random samples with replacement. "Feature bagging" is employed to address correlated trees, enhancing accuracy under various conditions.

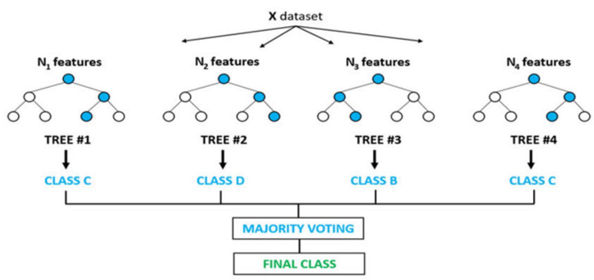

Figure 1.2: A Random Forest Classifier

1.1.2 XGBoost

XGBoost, a decision-tree-based ensemble ML algorithm, operates within a gradient boosting framework. It excels in structured/tabular data scenarios. Unlike traditional gradient boosting, XGBoost uses a second-order Taylor approximation in the loss function, resembling the Newton-Raphson method.
XGBoost features system optimization, regularization (Lasso and Ridge Regression), parallelization, cache block optimization, tree pruning, cache-awareness, out-of-score computation, sparsity awareness, weighted quantile sketch, and built-in cross-validation. These aspects optimize computational performance, handle diverse data types efficiently, and prevent overfitting. The algorithm's flexibility is evident in its ability to handle sparse data, different sparsity patterns, and perform well on various problem sizes.

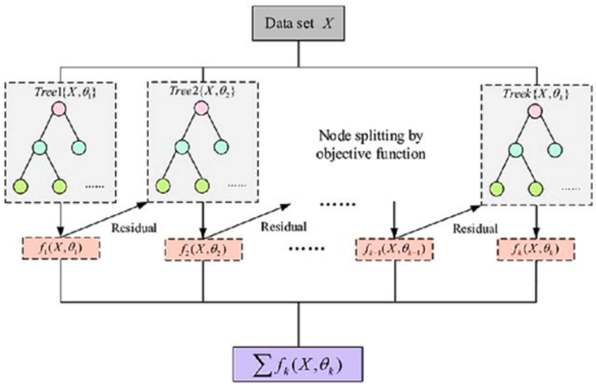

Figure 1.3: Flow Diagram of XGBoost Algorithm

1.2 Deep Learning

Deep learning, also referred to as deep structured learning, is a subset of machine learning methods centered on artificial neural networks with representation learning. Learning can be supervised, semi-supervised, or unsupervised. Various deep-learning architectures, including deep neural networks, deep belief networks, deep reinforcement learning, recurrent neural networks, and convolutional neural networks, have demonstrated remarkable success across diverse fields.

Deep learning's "deep" designation originates from the utilization of multiple layers in the network. Early findings revealed the limitations of linear perceptrons, leading to the development of deep learning concerned with an unbounded number of layers of bounded size. This approach allows practical application and optimized implementation while maintaining theoretical universality.

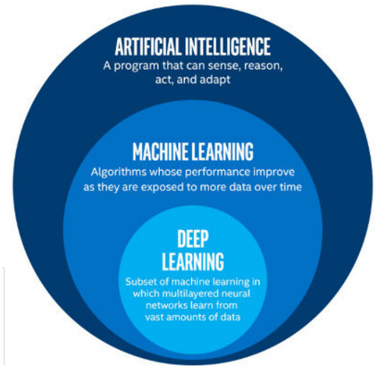

Figure 1.4: ML as a subset of AI and DL as a subset of ML

1.2.1 Artificial Neural Networks

Artificial neural networks (ANNs), also known as neural networks (NNs), are computing systems inspired by biological neural networks. ANNs consist of interconnected nodes, or artificial neurons, modeling the neurons in a biological brain. Each connection transmits a signal, and the output of each neuron is determined by a non-linear function of its inputs. Neurons are organized into layers, with signals traveling from the input layer to the output layer, possibly traversing layers multiple times.

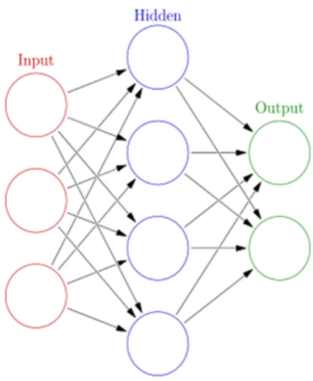

Figure 1.5: An Artificial Neural Network with layer coloring

1.2.2 Deep Neural Networks

A deep neural network (DNN) extends the concept of ANNs with multiple layers between the input and output layers. Despite different types of neural networks, they share components such as neurons, synapses, weights, biases, and functions, functioning similarly to the human brain and trainable like other machine learning algorithms.

1.2.3 Convolutional Neural Networks (CNNs)

In deep learning, a convolutional neural network (CNN) is a specialized artificial neural network applied extensively in visual imagery analysis. CNNs utilize convolutional layers, pooling layers, ReLU layers, fully connected layers, and loss layers to transform input volumes into output volumes through differentiable functions.

Key CNN layers include:

- **Convolutional Layer:** Core building block using learnable filters to detect specific features in the input.
- **Pooling Layer:** Non-linear downsampling, often implemented as max pooling.
- **ReLU Layer:** Applies the rectified linear unit activation function to introduce non-linearities.
- **Fully Connected Layer:** Performs final classification through connections to all activations in the previous layer.
- **Loss Layer:** Specifies how training penalizes the deviation between predicted and true data labels.

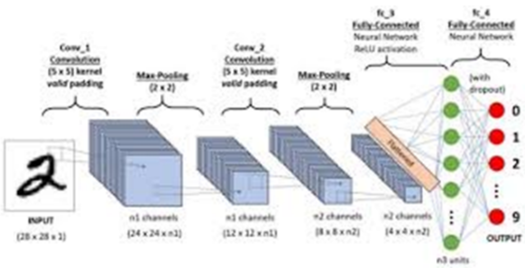

Figure 1.6: A generic CNN architecture

1.3 Motivation

In the midst of COVID-19 disruptions, the motivation for this project is to use technology to make disease detection more accessible. With limited access to hospitals and challenges in online consultations, the project aims to bridge healthcare gaps. By leveraging technology, it enables easier and more widespread detection of diseases. This approach becomes crucial during a pandemic, where traditional healthcare systems may face limitations. The goal is to use innovative solutions to overcome barriers, ensuring that people can still receive timely and effective healthcare even in challenging circumstances, ultimately contributing to better public health outcomes.

1.4 Objective

The objective of this text is to convey the growing recognition within healthcare organizations of the potential benefits offered by artificial intelligence (AI) in terms of enhancing patient care, reducing costs, and improving operational efficiency.[1] It aims to highlight the rapid evolution and increasing sophistication of AI tools, providing stakeholders with various options to consider. The text emphasizes the importance of understanding the intricacies of data processing in AI applications and the critical role it plays in managing expectations and making informed investment decisions.[2]

1.5 Problem Definition

The problem addressed is the need for healthcare organizations to navigate the complex landscape of AI tools and strategies. With the influx of options, there is a challenge in understanding how data is processed and presented, influencing decisions related to accuracy, reliability, and resource investments. [3] The text suggests that organizations may face difficulties in choosing between vendor products or deciding whether to develop in-house algorithms. This highlights the need for a solid understanding of AI's different facets and their specific applications in healthcare.

1.6 Aim

The aim of the text is to guide healthcare organizations in making informed decisions regarding the adoption of AI. It encourages organizations to develop a robust understanding of AI, particularly focusing on the transformative potential of deep learning. By exploring deep learning as a logical starting point, the text aims to equip healthcare organizations with the knowledge necessary to evaluate the practical applications of AI, especially in real-world clinical settings. The mention of Convolutional Neural Networks (CNNs) underscores the specific aim of showcasing the promising applications of deep learning in analyzing medical images and potentially outperforming human diagnosticians.[4]

Chapter 2

Literature Survey

The paper "Health Care Application using Machine Learning and Deep Learning" offers a comprehensive examination of the pivotal role played by machine learning in revolutionizing healthcare practices. Through a meticulous analysis of extensive patient data, machine learning algorithms prove instrumental in predicting disease likelihoods, enabling early diagnoses, and facilitating more effective treatments. This emphasis on early disease prediction not only enhances patient outcomes but also substantially increases survival rates.[5] The paper delves into a range of machine learning techniques, including Decision Trees, Logistic Regression, Random Forest, and Naive Bayes, illustrating their effectiveness in disease prediction models. Real-world applications underscore the tangible impact of these models, with instances like Support Vector Machine's remarkable accuracy in heart disease prediction and Logistic Regression's impressive performance in diabetes prediction. As these machine learning techniques become integral to web-based healthcare applications, considerations around algorithm selection, ensuring data quality, and meticulous hyperparameter tuning take center stage. The adoption of machine learning in healthcare, as illuminated by the insights in this paper, holds immense potential for reshaping patient care, reducing costs, and elevating overall healthcare efficiency to new heights.[6]

The paper titled "Prediction of Disease Using Machine Learning" explores a range of machine learning algorithms for disease prediction, including Naïve Bayes, Decision Tree, Random Forest, and Support Vector Machine (SVM). Naïve Bayes, a probabilistic algorithm, assumes feature independence for simplicity and speed. Decision Tree recursively partitions datasets, offering interpretability. Random Forest, an ensemble algorithm, enhances accuracy and robustness. SVM finds the optimal hyperplane in a high-dimensional space, catering to linear and nonlinear tasks with high precision.[7] The document underscores the importance of feature selection and data preprocessing, employing techniques like correlation-based feature selection and recursive feature elimination. Data preprocessing methods, such as normalization and standardization, contribute to model refinement. The paper delves into the dataset used for training and testing, elucidating its features, attributes, and evaluation metrics like accuracy, precision, recall, and F1 score. [8]These machine learning algorithms, presented in the context of disease prediction, play pivotal roles, each with its strengths and weaknesses. The choice of algorithm depends on specific problem requirements and dataset characteristics as outlined in "Prediction of Disease Using Machine Learning."

[9]The paper, "Disease Prediction using Machine Learning," introduces a predictive modeling system focused on disease prognosis through user-input symptoms. Utilizing the Naive Bayes Classifier, the system calculates disease probabilities, covering ailments like Diabetes, Malaria, Jaundice, Dengue, and Tuberculosis. With a blend of linear regression, decision tree, and machine learning techniques, the system aims to empower end-users for accurate, early-stage disease prediction, considering both structured and unstructured data. The study outlines the dataset details, evaluation metrics, and the potential of machine learning in revolutionizing healthcare outcomes, providing a comprehensive framework for predicting chronic diseases.[10] The proposed system not only streamlines existing machine learning algorithms for predicting chronic diseases but also incorporates Naive Bayes, KNN, and Logistic Regression to enhance accuracy. The architecture integrates structured and textual data, showcasing a holistic approach to disease prediction. The paper concludes by emphasizing the practical significance of the proposed system in predicting diseases based on symptoms, offering a potential breakthrough in personalized healthcare. [11]

Chapter 3

System Design

3.1 Proposed System

"Instant Health Scan" revolutionizes healthcare with a hybrid machine learning and deep learning framework for swift and precise diagnoses of seven diverse diseases. This tech-driven initiative streamlines diagnostics, challenging traditional timelines and ensuring accessibility and efficiency.[12] Emblematic of technology's transformative role in healthcare, it empowers individuals with timely health insights, fostering early intervention and improved outcomes. A comprehensive approach covers various health concerns, contributing to a forward-looking healthcare paradigm. Instant Health Scan signifies a commitment to innovation, informed decision-making, and ensuring well-being through advanced diagnostics.

3.2 Architecture

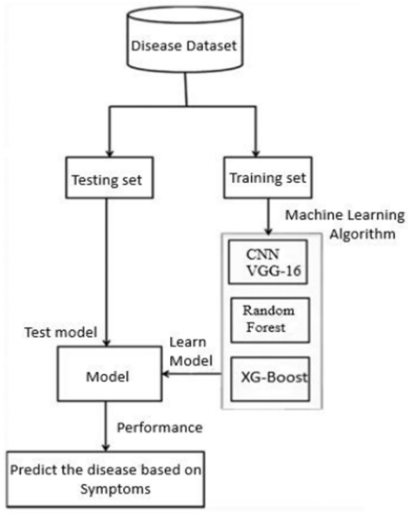

Figure 3.1: Architecture

3.2.1 System Architecture

Introducing an Advanced Healthcare Prediction Framework consisting of two distinct modules: Module 1 specializes in Imaging Diagnostics, leveraging sophisticated image analysis to predict conditions like Alzheimer's, Brain Tumor, Pneumonia, and COVID-19. This module excels in visually-driven disease identification.[13] On the other hand, Module 2 focuses on Data-Driven Disease Prediction for heart-related issues, diabetes, and breast cancer, utilizing diverse datasets and predictive modeling. By integrating both image-based and value-based insights, this framework offers a comprehensive approach, enhancing early detection and precise diagnosis in healthcare.

Chapter 4

System Modules

4.1 Imaging Diagnostics (Alzheimer's, Brain Tumor, Pneumonia, COVID-19)

This module employs cutting-edge image analysis techniques to predict diseases such as Alzheimer's, Brain Tumor, Pneumonia, and COVID-19. By processing medical images, it provides detailed insights into complex conditions, aiding in early detection and precise diagnosis. Whether identifying neural irregularities or recognizing respiratory infections, this module excels in image-centric disease prediction.

4.2 Data-Driven Disease Prediction (Heart, Diabetes, Breast Cancer)

In contrast, Module 2 focuses on diseases such as Heart conditions, Diabetes, and Breast Cancer using numerical data. By aggregating diverse datasets, this module employs predictive modeling and data analysis to discern patterns and forecast the likelihood of these diseases. It emphasizes the importance of numerical values, medical history, and clinical assessments to make informed predictions, contributing to a holistic healthcare approach. Together, these modules form a robust framework for comprehensive health prediction, incorporating both image-based and value-based insights."

Chapter 5

Implemention

5.1 Steps involved

- The system was developed in Python using Anaconda tools.
- Experiments focused on classification approaches and feature extraction techniques.
- The hardware used for experiments included an Intel Core i5-6200U processor and 8GB RAM.
- Disease datasets from various sources were divided into standard 70% training and 30% testing sets using Scikit-learn.
- Performance metrics such as TP, FP, TN, FN were utilized to calculate accuracy, precision, recall, and F1-score.
- Disease-specific parameters were employed to determine patient ailment predictions.
- Input validation was implemented to ensure proper user inputs.
- A Flask API was used to facilitate model interaction.
- Model interaction was demonstrated through a user-friendly web page designed for heart disease predictions.
- The system accommodated predictions for multiple diseases, allowing users to activate specific disease models based on inputs.

- Visual representation of disease predictions, including heart disease, provided prompt results for user queries.A user-friendly interface enabled patients to input disease-specific parameters, ensuring accurate and reliable predictions.
- A user-friendly interface enabled patients to input disease-specific parameters, ensuring accurate and reliable predictions.
- Standard machine learning algorithms were compared for accuracy and processing time.
- The study highlighted the superiority of Convolutional Neural Network (CNN) in terms of accuracy and processing time.
- Disease risk predictions were presented as accurate and general.
- The study emphasized the efficiency of the CNN algorithm in terms of both accuracy and processing time.
- The conclusion summarized the findings and implications of the study.
- The overall system showcased the practical application of machine learning in disease prediction with a user-friendly interface and robust performance metrics.

5.2 Model Architecture

COVID-19 Detection Model

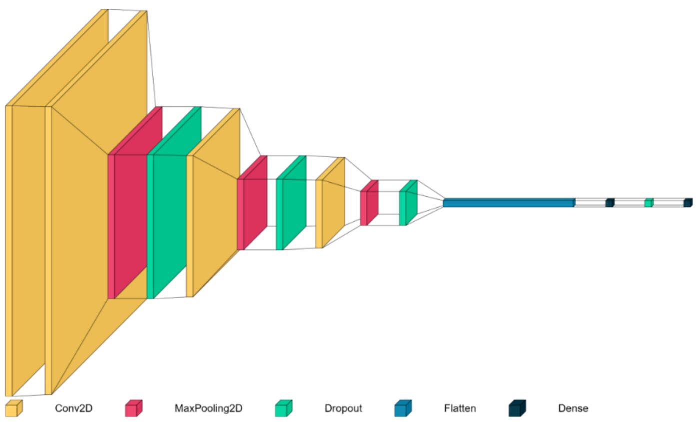

Figure 5.1: COVID-19

Brain Tumor Detection Model

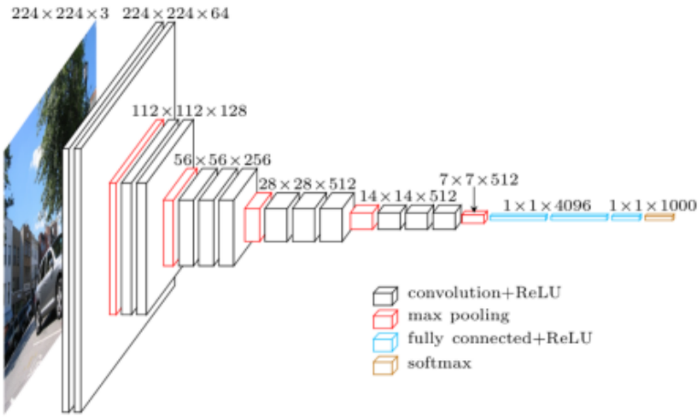

Figure 5.2: Brain Tumor

Alzheimer Detection Model

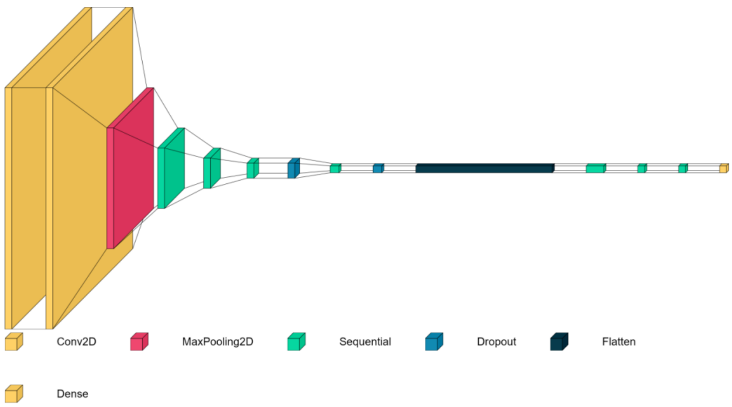

Figure 5.3: Alzheimer

Cloud Firestore: Cloud Firestore is a flexible, scalable NoSQL cloud database to store and sync data for client- and server-side development. It is a flexible, scalable database for mobile, web, and server development from Firebase and Google Cloud. Like Firebase Realtime Database, it keeps your data in sync across client apps through realtime listeners and offers offline support for mobile and web so you can build responsive apps that work regardless of network latency or Internet connectivity. Cloud Firestore also offers seamless integration with other Firebase and Google Cloud products, including Cloud Functions. Cloud Firestore is also available in native Node.js, Java, Python, Unity, C++ and Go SDKs, in addition to REST and RPC APIs. Following Cloud Firestore's NoSQL data model, you store data in documents that contain fields mapping to values. These documents are stored in collections, which are containers for your documents that you can use to organize your data and build queries. Documents support many different data types, from simple strings and numbers, to complex, nested objects. You can also create subcollections within documents and build hierarchical data structures that scale as your database grows. The Cloud Firestore data model supports whatever data structure works best for your app.

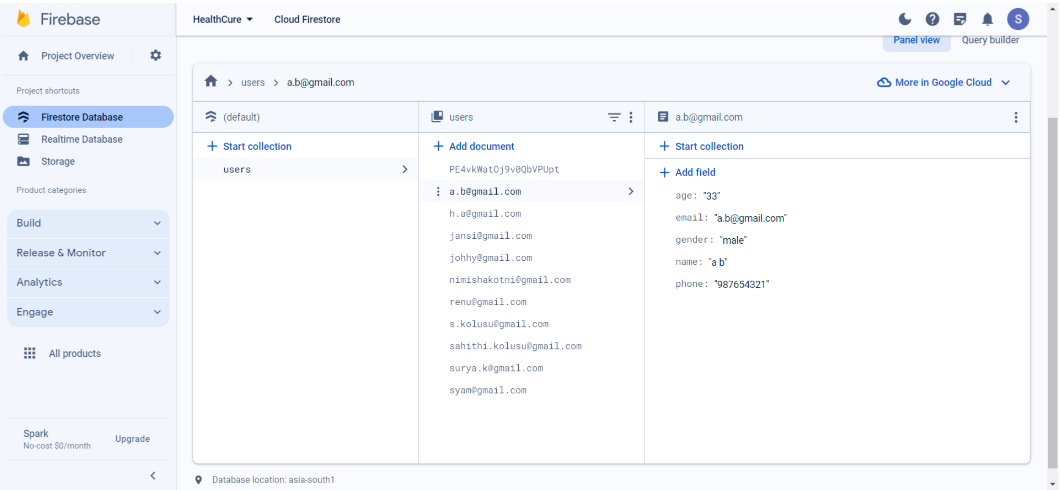

Figure 5.4: Firestore Storage

Chapter 6

Results and Discussions

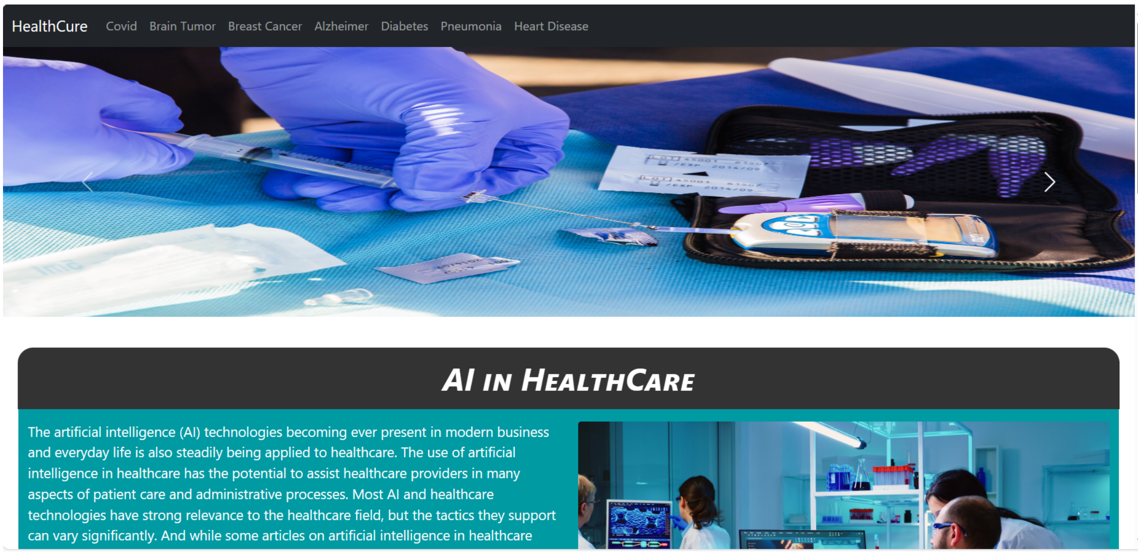

Figure 6.1: Home Page of Instant Health Scan

COVID-19: Coronavirus disease 2019 (COVID-19) is a contagious disease caused by a virus, the severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2). The first known case was identified in Wuhan, China, in December 2019. The disease spread worldwide, leading to the COVID-19 pandemic. COVID-19 transmits when people breathe in air contaminated by droplets and small airborne particles containing the virus. The risk of breathing these in is highest when people are in close proximity, but they can be inhaled over longer distances, particularly indoors.

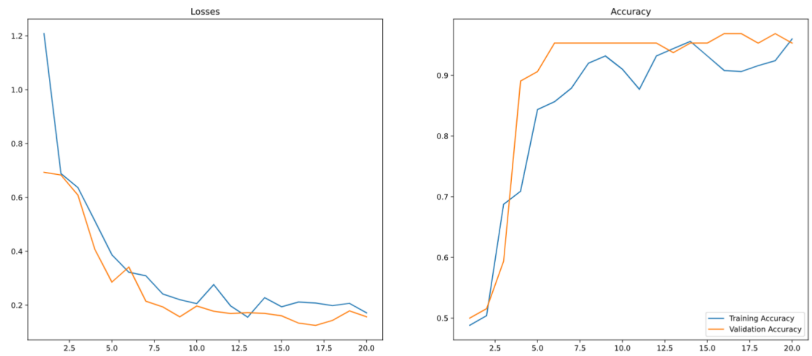

Figure 6.2: Loss and Accuracy Plot for Training of the COVID-19 detection model

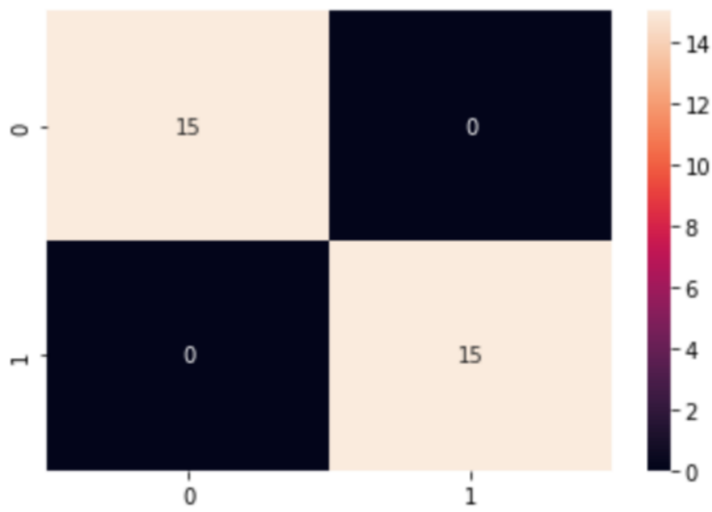

Figure 6.3: Confusion Matrix for predictions of the COVID-19 detection model

Brain Tumor: A brain tumor occurs when abnormal cells form within the brain. There are two main types of tumors: malignant tumors and benign (non-cancerous) tumors.These can be further classified as primary tumors, which start within the brain, and secondary tumors, which most commonly have spread from tumors located outside the brain, known as brain metastasis tumors

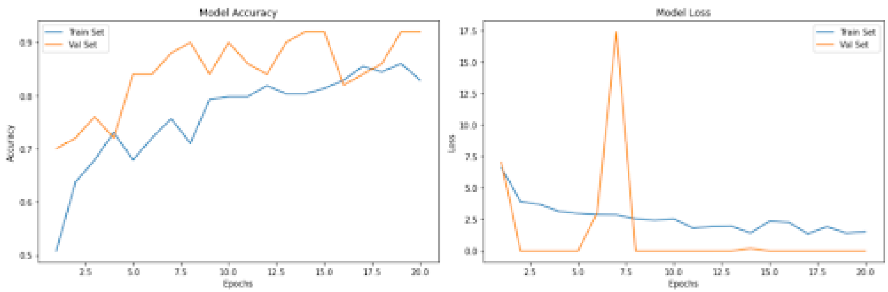

Figure 6.4: Loss and Accuracy Plot for Training of the Brain Tumor detection model

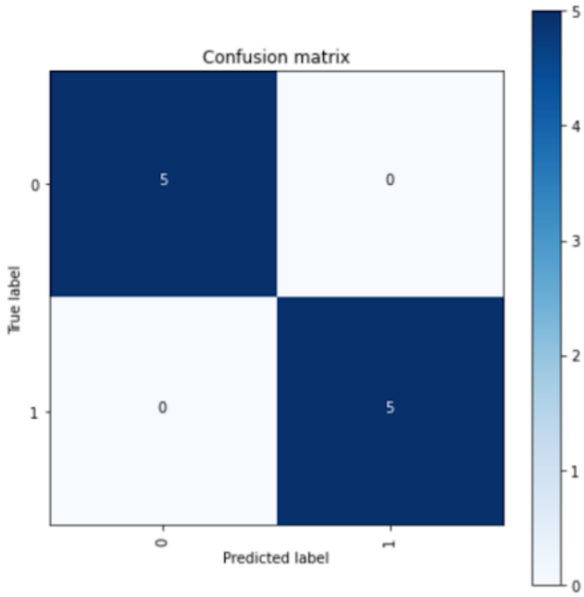

Figure 6.5: Confusion Matrix for predictions of the Brain Tumor detection model

Alzheimer: Alzheimer's disease (AD) is a neurodegenerative disease that usually starts slowly and progressively worsens. It is the cause of 60–70% of cases of dementia. The most common early symptom is difficulty in remembering recent events. As the disease advances, symptoms can include problems with language, disorientation (including easily getting lost), mood swings, loss of motivation, self-neglect, and behavioral issues. As a person's condition declines, they often withdraw from family and society.

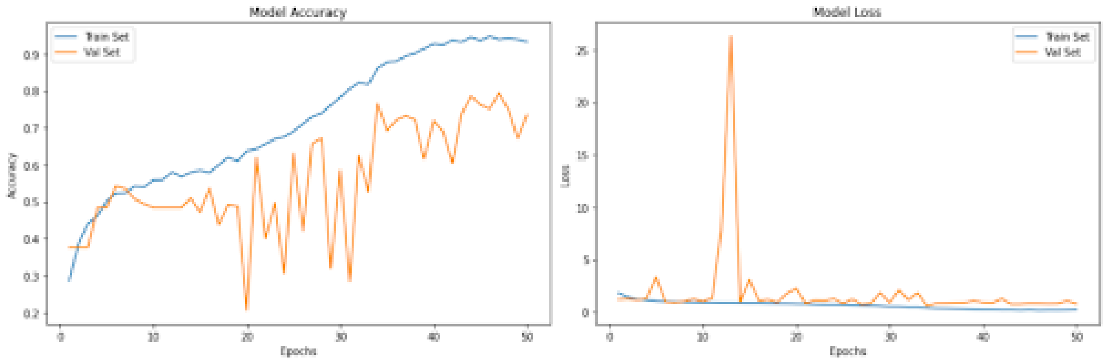

Figure 6.6: Loss and Accuracy Plot for Training of the Alzheimer detection model

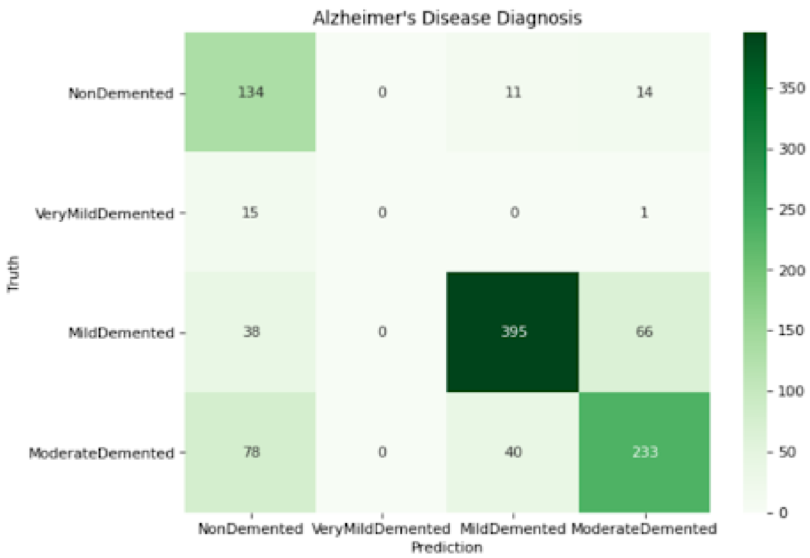

Figure 6.7: Confusion Matrix for predictions of the Alzheimer detection model

Pneumonia: Pneumonia is an inflammatory condition of the lung primarily affecting the small air sacs known as alveoli. Symptoms typically include some combination of productive or dry cough, chest pain, fever, and difficulty breathing. The severity of the condition is variable. Pneumonia is usually caused by infection with viruses or bacteria, and less commonly by other microorganisms.

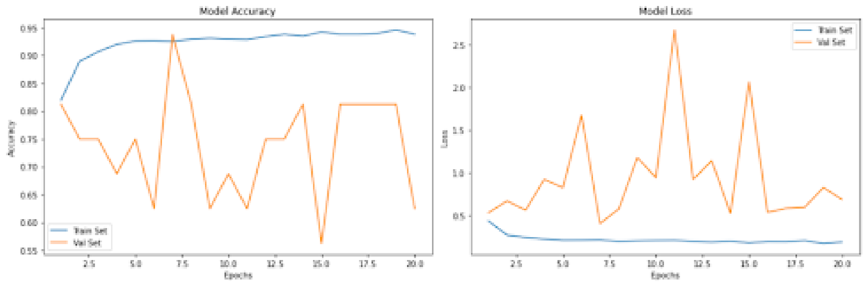

Figure 6.8: Loss and Accuracy Plot for Training of the Pneumonia detection model

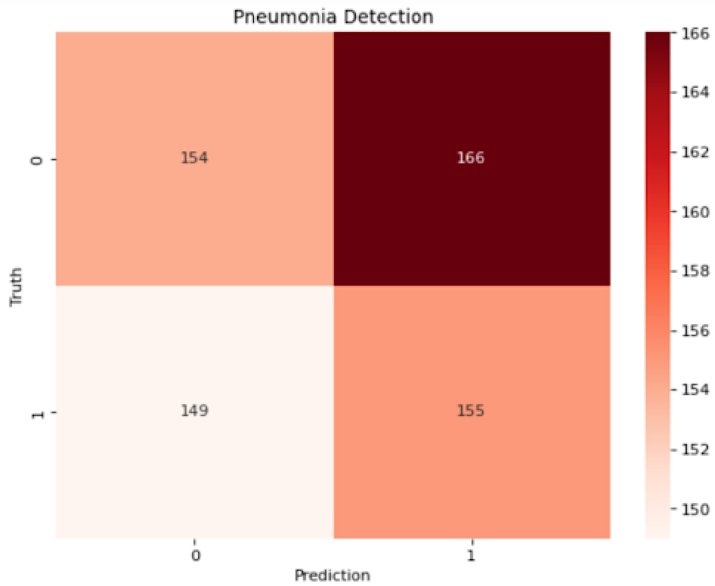

Figure 6.9: Confusion Matrix for predictions of the Pneumonia detection model

Diabetes: Diabetes mellitus, commonly known as diabetes, is a group of metabolic disorders characterized by a high blood sugar level (hyperglycemia) over a prolonged period of time. Symptoms often include frequent urination, increased thirst and increased appetite. If left untreated, diabetes can cause many health complications. Acute complications can include diabetic ketoacidosis, hyperosmolar hyperglycemic state, or death. Serious long-term complications include cardiovascular disease, stroke, chronic and kidney disease.

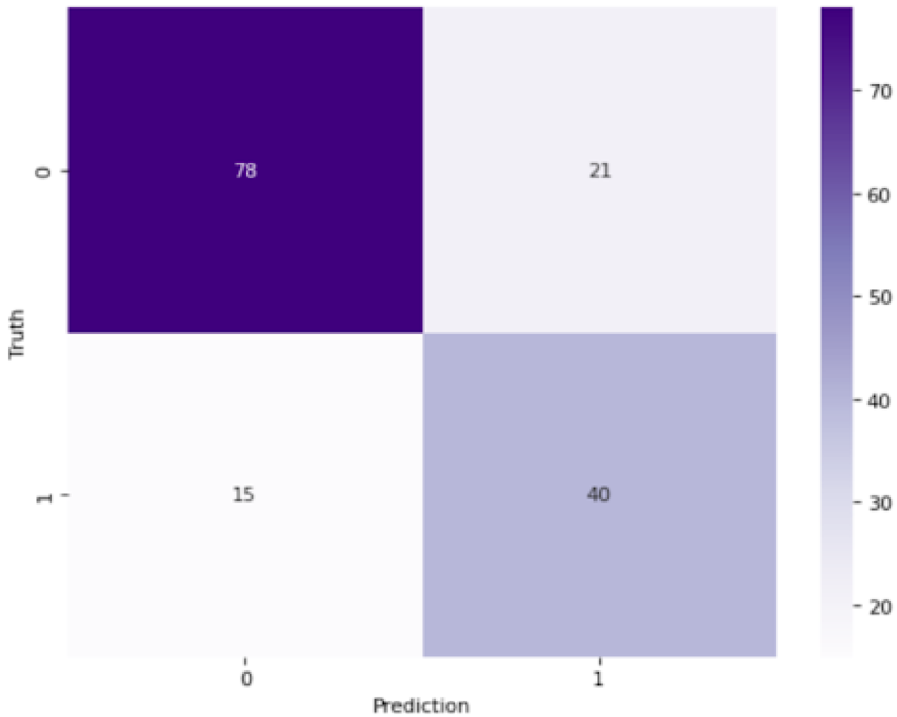

Figure 6.10: Loss and Accuracy Plot for Training of the Diabetes detection model

Breast Cancer: Breast cancer is cancer that develops from breast tissue. Signs of breast cancer may include a lump in the breast, a change in breast shape, dimpling of the skin, fluid coming from the nipple, a newly inverted nipple, or a red or scaly patch of skin. In those with distant spread of the disease, there may be bone pain, swollen lymph nodes, shortness of breath, or yellow skin. Breast cancer most commonly develops in cells from the lining of milk ducts and the lobules that supply these ducts with milk. The diagnosis of breast cancer is confirmed by taking a biopsy of the concerning tissue.

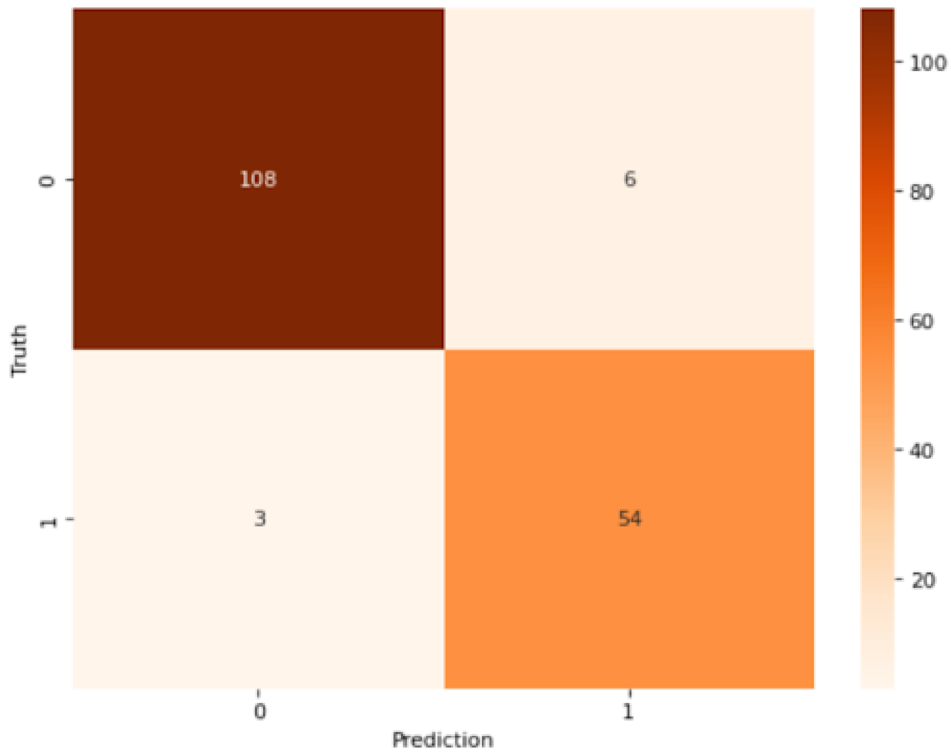

Figure 6.11: Loss and Accuracy Plot for Training of the Breast Cancer detection model

Heart Disease: Cardiovascular disease (CVD) is a class of diseases that involve the heart or blood vessels. CVD includes coronary artery diseases (CAD) such as angina and myocardial infarction (commonly known as a heart attack). Other CVDs include stroke, heart failure, hypertensive heart disease, rheumatic heart disease, cardiomyopathy, abnormal heart rhythms, congenital heart disease, valvular heart disease, carditis, aortic aneurysms, peripheral artery disease, thromboembolic disease, and venous thrombosis.

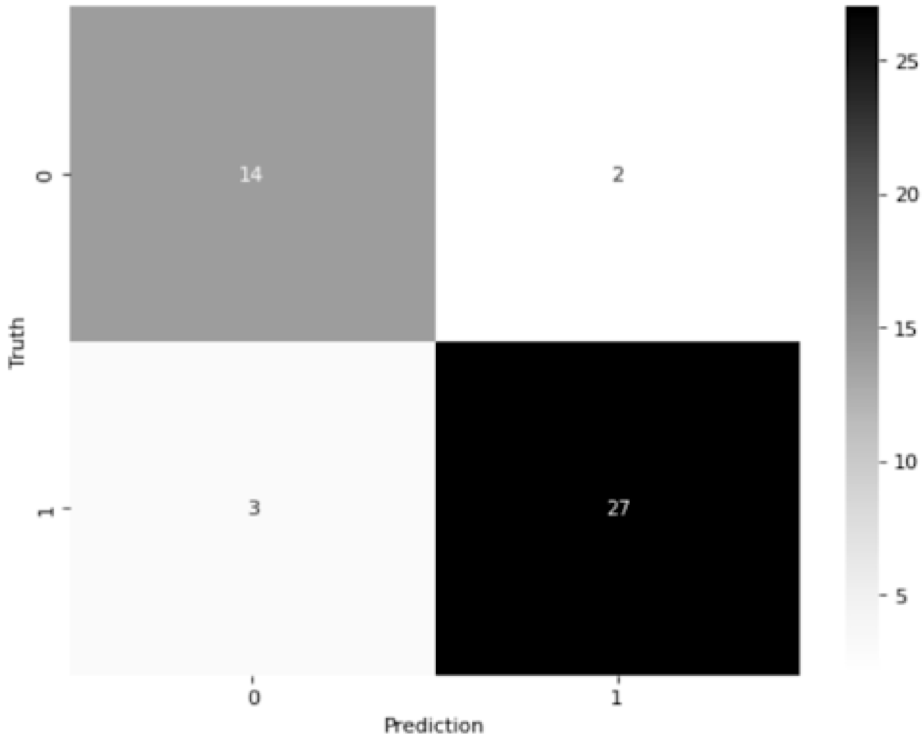

Figure 6.12: Loss and Accuracy Plot for Training of the Heart Disease detection model

Accuracy Table

Disease	**Model Architecture**	**Accuracy**
Alzheimer	CNN	73.54%
Brain Tumor	CNN with pre-trained VGG16 weights	100%
Breast Cancer	Random Forest	91.81%
COVID-19	CNN	93%
Diabetes	Random Forest	66.8%
Pneumonia	CNN	83.17%
Heart Disease	XGBoost	86.96%

Table 6.1: Accuracy Table

Chapter 7

Conclusions and future works

7.1 Conclusion

This project effectively utilizes AI to detect seven diseases, merging CNNs for image analysis and classical ML algorithms (Random Forest and XGBoost). It provides a convenient, accurate, and efficient disease detection method, showcasing AI's prowess in medical diagnostics. The project's application of CNNs in image analysis and XGBoost for structured data underlines AI's versatility in healthcare, promising faster and more accessible disease detection, reducing report reliance and enhancing diagnostic.

7.2 Future scope

- **Expanded Disease Detection:** The project can be expanded to detect a broader spectrum of diseases and health conditions, including rare and complex disorders, enabling early diagnosis and treatment.

- **AI Explainability:** Enhancing the interpretability and transparency of AI models, making their decision-making processes more understandable to healthcare professionals and patients, thus building trust and facilitating acceptance.

Bibliography

[1] Q. M. S. X. e. a. Gai, K., “Smart data deduplication for telehealth systems in heterogeneous cloud computing,” *J. Commun. Inf. Netw. 1*, 2016.

[2] M. A. MD. Atikur Rahman, Tania Ahmed Nipa, “Predicting disease from several symptoms using a machine learning approach,” *IRJET*, 2023.

[3] A. P. A. S. J. G. Ajinkya Padule, Aman Patel, “Health care application using machine learning and deep learning,” *IRJET*, 2022.

[4] L. Sharmila, “A comparative study of neural network and fuzzy neural network for classification,” 2022.

[5] S. . V. H. . S. S. S, Vinitha S, “Disease prediction using machine learning over big data,” *Computer Science Engineering: An International Journal*, 2018.

[6] M. J. J. Mrs. K. Uma Maheswari, “Neural network based heart disease prediction,” *IRJET*, 2017.

[7] H. . S. N. Kunjir, Ajinkya Sawant, “Data mining and visualization for prediction of multiple diseases in healthcare,” 2017.

[8] Y. L. D. W. Y. Z. M. Chen, Y. Ma and C. H. Youn, “Wearable 2.0: Enabling human-cloud integration in next generation healthcare systems,” *IEEE*, 2017.

[9] X. . C. N. . L. H. . J. Y.-G. Qian, Buyue Wang, “A relative similarity based method for interactive patient risk prediction,” *Data Mining and Knowledge Discovery*, 2017.

[10] K. H. L. W. M. Chen, Y. Hao and L. Wang, “Disease prediction by machine learning over big data from healthcare communities,” *IEEE*, 2017.

[11] C. W. T. M. M. H. Y. Zhang, M. Qiu and A. Alamri, "Health-cps: Healthcare cyber-physical system assisted by cloud and big data," *IEEE*, 2016.

[12] I. Sarker, "Machine learning: Algorithms, real-world applications and research directions," *SN COMPUT. SCI. 2*, 2021.

[13] V. K. S. S. P. A. K. Kedar Pingale, Sushant Surwase, "Disease prediction using machine learning," *IRJET*, 2019.