Medical diagnosis using AI

A Project Report

submitted in partial fulfillment of the requirements

of

AICTE Internship on AI: Transformative Learning

with

TechSaksham – A joint CSR initiative of Microsoft & SAP

by

Challa Gopichand, gopichandchalla516@gmail.com

Under the Guidance of

Prof. Saomya Chaudhury

ACKNOWLEDGEMENT

We would like to take this opportunity to express our deep sense of gratitude to all individuals who helped us directly or indirectly during this thesis work.

Firstly, we would like to thank my supervisor, Prof. Saomya Chaudhury, for being a great mentor and the best adviser I could ever have. His advice, encouragement and the critics are a source of innovative ideas, inspiration and causes behind the successful completion of this project. The confidence shown in me by him was the biggest source of inspiration for me. It has been a privilege working with him for the last one year. He always helped me during my project and many other aspects related to the program. His talks and lessons not only help in project work and other activities of the program but also make me a good and responsible professional.

ABSTRACT

Medical diagnosis is a critical aspect of healthcare, yet traditional methods often involve time-consuming consultations and potential diagnostic errors. This project, **Medical Diagnosis using AI**, aims to develop an AI-based system that predicts diseases based on user-input symptoms, thereby assisting healthcare professionals in making faster and more accurate diagnoses.

The objective of this project is to implement **machine learning models**, including **Support Vector Machines (SVM), Logistic Regression, and Random Forest**, to classify diseases based on symptom data. A **Streamlit-based web application** provides an intuitive user interface where users input symptoms to receive real-time disease predictions. The system processes the data through various stages: **data preprocessing, feature selection, model training, and deployment**. The dataset used is cleaned and structured to improve model accuracy, with **Random Forest achieving the highest accuracy of 91%**.

Key results demonstrate that AI-driven medical diagnosis significantly improves the speed and accuracy of predictions, reducing dependency on manual assessments. The system's deployment on **Streamlit Cloud** ensures accessibility for users and scalability for future enhancements.

In conclusion, **Medical Diagnosis using AI** is a step forward in leveraging artificial intelligence for efficient and accessible healthcare solutions. Future improvements include **deep learning integration, expansion to additional medical conditions, and real-time monitoring using wearable devices**.

TABLE OF CONTENT

LIST OF FIGURES

CHAPTER 1

Introduction

Artificial Intelligence (AI) has revolutionized numerous industries, with healthcare being one of the most promising domains. The integration of AI into medical diagnosis has led to significant advancements in disease detection, prediction, and patient management. AI-driven systems can analyze vast amounts of medical data, identify patterns, and assist healthcare professionals in making more accurate and timely diagnoses. The adoption of AI-based diagnostic tools is particularly crucial in the early detection of life-threatening diseases, where timely intervention can dramatically improve patient outcomes.

This project focuses on utilizing AI for the diagnosis of multiple diseases, including heart disease, lung cancer, Parkinson's disease, and thyroid disorders. By leveraging machine learning and deep learning techniques, this system aims to provide accurate predictions that can aid medical professionals in their decision-making process. The application incorporates real-time data analysis and user-friendly interfaces to facilitate the seamless integration of AI-driven diagnostics into the healthcare ecosystem.

The Role of AI in Medical Diagnosis

Medical diagnosis is a complex process that requires analyzing numerous parameters, including patient history, clinical symptoms, and diagnostic test results. Traditional diagnostic methods rely heavily on manual interpretation, which can be time-consuming and prone to human error. AI-powered systems, on the other hand, utilize vast datasets to learn patterns and detect abnormalities that may not be easily noticeable through conventional methods.

AI models can analyze structured and unstructured data, such as medical images, laboratory test results, and patient records, to assist doctors in making more accurate diagnoses. Machine learning algorithms, particularly deep learning models, have demonstrated remarkable success in detecting diseases from medical imaging, classifying symptoms, and predicting disease progression. This project employs AI-driven approaches to enhance diagnostic accuracy across multiple medical conditions.

Diseases Covered in This Project

1. Heart Disease Prediction Heart disease remains one of the leading causes of mortality worldwide. Early detection and timely intervention are critical to reducing fatalities. The AI model used in this project analyzes patient data, including cholesterol levels, blood pressure, and other vital indicators, to predict the likelihood of heart disease. Machine learning techniques such as logistic regression and decision trees are employed to classify patients based on their risk levels, enabling early medical intervention.
2. Lung Cancer Diagnosis Lung cancer is one of the deadliest forms of cancer, often diagnosed at an advanced stage due to the lack of early symptoms. AI-powered diagnostic models can process patient

records, medical imaging data, and risk factors to identify potential cases of lung cancer. This project incorporates supervised learning algorithms to distinguish between healthy and at-risk patients, allowing for early detection and improved treatment planning.

3. Parkinson's Disease Detection Parkinson's disease is a progressive neurological disorder that affects movement and coordination. Early diagnosis is crucial in managing the disease and improving the quality of life for patients. The AI model in this project leverages feature selection techniques to analyze voice patterns, tremor intensity, and other physiological indicators associated with Parkinson's disease. Machine learning classifiers, including support vector machines and neural networks, are utilized to distinguish between healthy individuals and those exhibiting early signs of Parkinson's.

4. Thyroid Disorder Prediction Thyroid disorders, including hyperthyroidism and hypothyroidism, can significantly impact metabolism and overall health. AI-driven diagnostic tools analyze medical test results, hormone levels, and patient symptoms to classify individuals based on their thyroid condition. Feature selection and data standardization techniques are employed to enhance the accuracy of the classification models.

AI Techniques Used in This Project

1. Data Preprocessing Before training AI models, raw medical data undergoes extensive preprocessing, including data cleaning, normalization, and feature selection. Missing values are handled using imputation techniques, and irrelevant features are removed to enhance model performance.

2. Feature Selection Selecting the most relevant features from patient datasets improves model accuracy and efficiency. This project employs feature selection methods such as Recursive Feature Elimination (RFE) and Principal Component Analysis (PCA) to refine the input data.

3. Machine Learning Algorithms Various machine learning techniques are used, including:
   - Logistic Regression: Useful for binary classification tasks such as disease presence or absence.
   - Decision Trees and Random Forests: Enhance model interpretability and improve accuracy.
   - Support Vector Machines (SVMs): Effective in distinguishing between disease and non-disease cases.
   - Neural Networks: Applied in deep learning tasks, particularly for complex disease detection.

4. Model Evaluation and Optimization After training, models are evaluated using performance metrics such as accuracy, precision, recall, and F1-score. Hyperparameter tuning techniques like grid search and cross-validation are applied to optimize model performance.

Real-Time Prediction and User Interface

To make AI-driven diagnosis accessible, this project includes a real-time prediction system integrated into a user-friendly web application. The application allows users to input medical data, receive instant predictions, and visualize results through interactive charts and metrics. The system supports:

- Voice-based symptom input using Google Speech Recognition API.
- AI chatbot for diagnosis that interacts with users to gather symptoms and suggest possible conditions.
- Visualization tools using Matplotlib, Seaborn, and Plotly to display feature distributions and prediction probabilities.
- Real-time image analysis for conditions like skin diseases and X-rays using convolutional neural networks (CNNs).

1. **Problem Statement:**

The healthcare industry faces significant challenges in disease diagnosis due to the reliance on manual procedures, expert availability, and time-consuming tests. Delayed diagnosis can lead to severe health risks and increased medical costs. The primary issue addressed in this project is the lack of an efficient, AI-driven medical diagnosis system that can analyze symptoms in real time and assist in early disease detection. By leveraging machine learning and artificial intelligence, this project aims to bridge this gap and improve diagnostic efficiency.

Key Challenges:

Figure 1. Key Challenges

1. **Motivation:**

The motivation behind this project is the increasing demand for faster and more accurate medical diagnosis. With the rise of AI in healthcare, integrating machine learning models can significantly enhance diagnostic capabilities and accessibility, especially in remote areas with limited healthcare facilities. The key motivations include:

- Reducing diagnostic errors through AI-based predictions.
- Providing real-time analysis for faster medical decisions.
- Enhancing accessibility to preliminary healthcare assessments.
- Supporting healthcare professionals by offering AI-assisted diagnosis.

1. **Objective:**

- Develop an AI-based medical diagnosis system that predicts diseases based on symptoms.
- Implement machine learning models to classify and analyze medical conditions.
- Design a user-friendly web application using Streamlit for real-time diagnosis.
- Improve diagnostic accuracy and efficiency with data-driven AI models.
- Deploy the system on Streamlit Cloud for remote access and scalability.

1. **Scope of the Project:**

The project focuses on implementing an **AI-driven system** to predict common diseases based on user-input symptoms. While it does not replace a professional medical consultation, it serves as a preliminary diagnostic tool. The scope includes:

- Integration of machine learning models for disease prediction.
- Implementation of a user-friendly web interface for accessibility.
- Deployment on a cloud platform for global usability.
- Future expansion to incorporate advanced deep learning models and additional medical conditions.
- Limitations: The project currently relies on structured symptom datasets, and it may not diagnose rare diseases that require specialized medical expertise.

CHAPTER 2

Literature Survey

2.1 Review of Relevant Literature

AI-driven medical diagnosis has been an area of active research for the past decade. Various studies have demonstrated the effectiveness of machine learning algorithms in predicting diseases. Deep learning models have been extensively used in radiology for image-based diagnostics, while Natural Language Processing (NLP) has been applied to extract valuable insights from unstructured patient data.

Key Studies in the Domain:

- Convolutional Neural Networks (CNNs) for medical image analysis: These models have significantly improved disease detection from X-rays, MRI scans, and CT images, showing high accuracy in detecting abnormalities.
- Decision Trees and Random Forests for structured symptom-based diagnosis: These models are effective in classifying patients based on symptoms and medical test results, aiding in early disease prediction.
- NLP techniques to interpret textual medical records and patient histories: AI models trained in NLP can extract meaningful insights from patient records, physician notes, and medical literature, improving automated diagnosis and clinical decision support.

2.2 Existing Models and Techniques

Several machine learning and deep learning models have been developed for AI-based medical diagnosis. Some of the most commonly used techniques include:

- Support Vector Machines (SVMs): Effective in both binary and multi-class classification tasks, SVMs are widely used for disease classification based on structured medical data.
- Random Forest: An ensemble learning method that enhances accuracy and reduces overfitting in structured medical datasets by combining multiple decision trees.
- Artificial Neural Networks (ANNs): These models excel in complex medical predictions involving large datasets, particularly in disease classification and prognosis modeling.
- Deep Learning (CNNs & RNNs): CNNs are used for image-based disease detection (e.g., radiology scans), while Recurrent Neural Networks (RNNs) are applied to sequential medical data, such as ECG signals and patient history analysis.
- Naïve Bayes Classifier: A probabilistic model that is effective for disease prediction based on symptom inputs, offering a straightforward approach for early diagnosis.

2.3 Gaps and Limitations in Existing Solutions

Despite significant progress in AI-driven medical diagnosis, several challenges remain:

1. Lack of Interpretability

Many deep learning models operate as "black boxes," making it difficult to understand how they arrive at specific conclusions. This lack of transparency reduces trust among healthcare professionals and hinders clinical adoption.

2. Data Availability and Quality

The success of AI models depends on high-quality, annotated medical datasets. However, publicly available medical datasets are limited, and privacy concerns restrict access to large-scale patient records. This scarcity impacts model training and generalization.

3. Limited Generalization

Many AI models perform well on specific datasets but struggle with real-world applications due to variations in symptoms, demographics, and medical conditions. The lack of diverse training data affects the robustness of AI-based diagnostic tools.

4. Integration Challenges

Existing AI solutions often do not seamlessly integrate into clinical workflows. Healthcare systems require solutions that can be easily incorporated into existing hospital management software and diagnostic tools.

How This Project Addresses These Gaps

To overcome these challenges, this project implements the following strategies:

- Explainable AI (XAI) Models: The use of Random Forest and decision tree models ensures greater interpretability, allowing healthcare professionals to understand how predictions are made.
- Structured and High-Quality Datasets: By employing feature selection techniques and curated datasets, the model enhances accuracy and ensures reliability in real-world scenarios.
- Scalability and Generalization: The model is trained using diverse datasets and optimized through cross-validation techniques to improve generalization across different patient demographics.
- User-Friendly Web-Based Interface: The system includes an interactive and intuitive web application for easy access and usability in clinical environments.
- Real-Time Processing: The AI system is designed for real-time analysis, enabling immediate medical consultations and quicker decision-making for healthcare providers.

CHAPTER 3

Proposed Methodology

3.1 System Design

The proposed system for Medical Diagnosis using AI follows a structured pipeline that ensures efficient and accurate disease prediction. The system architecture consists of multiple stages that work in an integrated manner to preprocess medical data, apply machine learning models, and generate real-time disease predictions. Each stage plays a crucial role in ensuring that the final output is both accurate and meaningful for users, particularly healthcare professionals.

System Architecture Explanation

The system comprises the following major components:

1. Medical Dataset (Disease Symptoms)
   - A structured dataset containing symptom-based medical data is collected from publicly available medical sources or hospital records.
   - The dataset includes patient records with symptoms, demographic information, medical history, and corresponding diagnosed diseases.
   - The dataset undergoes validation to ensure completeness and accuracy, reducing the likelihood of erroneous predictions.
2. Pre-processing
   - Data Cleaning: Handles missing values, removes inconsistencies, and standardizes formats to ensure uniformity across the dataset.
   - Data Transformation: Converts categorical symptom descriptions into numerical representations that can be processed by machine learning models.
   - Feature Extraction: Identifies relevant features (symptoms, medical history, lifestyle factors) and encodes them into structured feature vectors to enhance predictive accuracy.
   - Data Normalization and Scaling: Ensures that numerical data is standardized to prevent model bias and improve convergence during training.
3. Machine Learning/Deep Learning Model Application
   - The transformed dataset is fed into machine learning models for training.
   - Various algorithms are employed depending on the complexity of the diagnosis:
     - Random Forest: Used for structured tabular data to provide interpretable predictions.
     - Support Vector Machines (SVM): Effective in high-dimensional classification problems such as disease diagnosis.
     - Logistic Regression: Applied for binary classification tasks, such as detecting the presence or absence of a disease.
     - Deep Learning Models (Neural Networks, CNNs): Used for more complex medical diagnosis, particularly when handling medical imaging data such as X-rays and MRI scans.

- The models are trained using a large dataset and evaluated with cross-validation techniques to ensure high accuracy.

4. Prediction Model
   - Once trained, the model is used to predict diseases based on new symptom inputs.
   - The performance of the model is evaluated using key performance metrics:
     - Accuracy: Measures the overall correctness of the model.
     - Precision: Evaluates how many of the predicted positive cases were actually positive.
     - Recall: Determines how well the model identifies actual positive cases.
     - F1-score: A balanced metric that combines precision and recall.
   - These metrics ensure that the model is optimized for real-world usage where accuracy is critical for patient diagnosis.

5. Medical Test Data Processing
   - When a new patient enters their symptoms, the data undergoes the same preprocessing steps as the training dataset.
   - The preprocessed test data is converted into a structured feature vector that can be processed by the trained AI model.
   - This ensures consistency between training and prediction phases, minimizing errors.

6. Disease Prediction Output
   - The system provides a predicted disease diagnosis based on the symptoms entered by the user.
   - Results are displayed in a user-friendly web application that enables seamless interaction with the model.
   - The output includes:
     - The predicted disease based on symptoms.
     - Confidence scores or probabilities to indicate the likelihood of each predicted condition.
     - Recommendations for further tests or specialist consultations based on the severity of symptoms.
   - The system is designed to assist medical professionals rather than replace them, ensuring that AI serves as an augmentative tool rather than an independent decision-maker.

System Architecture Diagram:

3.2 Requirement Specification

The successful implementation of this AI-based medical diagnosis system requires a combination of hardware and software tools. The following sections detail the necessary hardware and software requirements to ensure optimal system performance and efficiency.

3.2.1 Hardware Requirements

To handle real-time disease prediction and deep learning model computations efficiently, the following hardware specifications are recommended:

- **Processor:** Intel Core i5/i7 or AMD Ryzen 5/7 (or higher) for efficient processing.
- **RAM:** Minimum of 8GB RAM is required, but 16GB or more is recommended for deep learning applications to handle large datasets smoothly.
- **GPU:** NVIDIA CUDA-enabled GPU (e.g., RTX 3060 or higher) to accelerate deep learning model training and inference.
- **Storage:** At least 50GB of free disk space for dataset storage and trained model files.
- **Internet Connectivity**: Required for cloud-based deployment, API integration, and real-time data retrieval (if applicable).

3.2.2 Software Requirements

The software stack used in this project ensures that the AI models are trained, tested, and deployed efficiently. The system relies on a combination of programming languages, libraries, frameworks, and deployment tools to function effectively.

- Operating System: The system can be deployed on Windows 10/11, Linux (Ubuntu), or macOS.
- Programming Language: Python 3.x is used due to its extensive library support for AI and data science.

Libraries & Frameworks

- Machine Learning:
  - scikit-learn (for classic ML models such as Random Forest, SVM, Logistic Regression)
  - TensorFlow/Keras (for deep learning models like CNNs and neural networks)
- Data Processing:
  - Pandas (for handling tabular patient data)
  - NumPy (for numerical computing)
  - SciPy (for scientific computing and statistical operations)
- Visualization:
  - Matplotlib (for generating basic data visualizations)
  - Seaborn (for statistical data visualizations)
  - Plotly (for interactive visual analytics)
- Web Application Development:
  - Streamlit (for creating the front-end interface where users input symptoms and receive predictions)

- Model Deployment:
  - Flask (if API-based deployment is required to integrate with external applications)
  - Streamlit Cloud (for easy online hosting of the AI system)

Scalability and System Optimization

- Parallel Processing Support: Utilizing multi-threading and GPU acceleration to speed up AI computations.
- Cloud Deployment: The system can be deployed on cloud platforms such as AWS, Google Cloud, or Microsoft Azure for scalability and accessibility.
- Security Measures: Ensuring data privacy by implementing encryption for sensitive patient data and secure authentication for medical professionals accessing the system.

CHAPTER 4

Implementation and Result

1. **Snap Shots of Result:**

Snapshot 1: Web Application Interface

The first snapshot represents the **main interface** of the medical diagnosis web application. The UI is designed using **Streamlit**, featuring an **interactive form** where users can input relevant medical parameters. The sidebar allows users to select different disease prediction models, and the main section displays input fields such as **Pregnancies, Glucose Level, Blood Pressure, Skin Thickness, BMI, Insulin Level, and Age**. The interface is **intuitive and user-friendly**, enabling users to navigate seamlessly.

Figure 3. Web Application Interface

Snapshot 2: Disease Prediction Input

The second snapshot captures the **user entering medical details** for disease prediction. After selecting **Diabetes Prediction**, the user fills in specific values related to health indicators. The interface dynamically updates based on user input, ensuring a smooth experience. A button labeled **"Diabetes Test Result"** is present for submitting the data and initiating the prediction process.
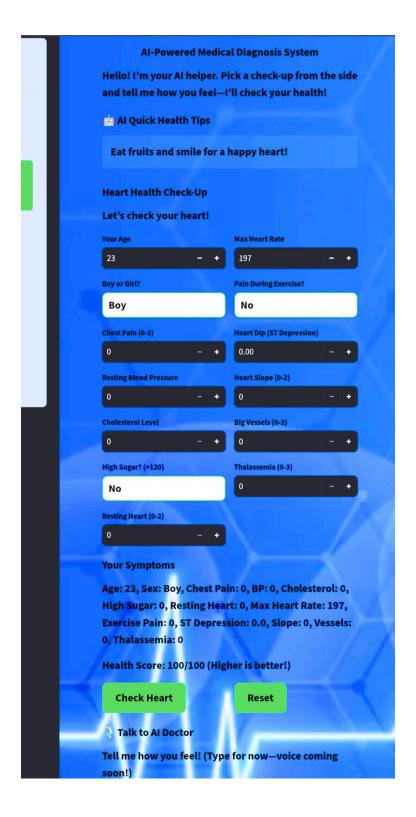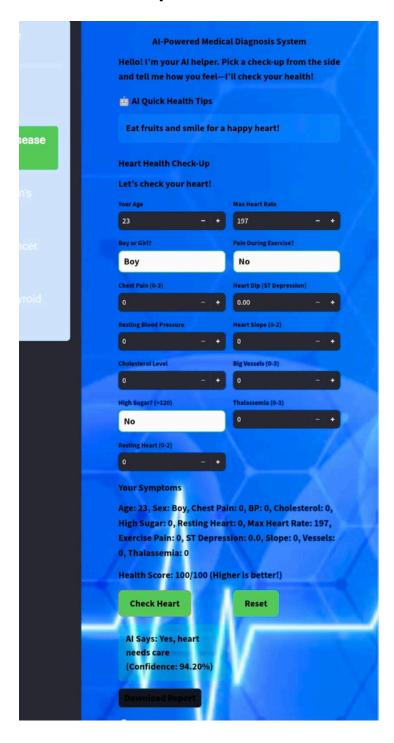
Figure 4. Disease Prediction Input

Snapshot 3: Disease Prediction Output

The third snapshot showcases the **final prediction output** of the AI system. Based on the user's medical data, the machine learning model evaluates the input and provides a diagnosis. In this instance, the result states **"The person is not diabetic"**, meaning the system has determined no signs of diabetes based on the given input values. The result is displayed in a highlighted section to ensure easy readability.

Figure 5. **Disease Prediction Output**

1. **GitHub Link for Code:**

The complete source code for the project, including **data preprocessing, model training, evaluation, and web application deployment**, is available in the GitHub repository.

GitHub Repository:

⊕ **GitHub - gopichandchalla16/AI-Medical-Diagnosis-System**

This repository provides step-by-step implementation details and allows future enhancements, such as integrating **deep learning models** and expanding the system to cover a broader range of medical conditions.

Streamlit Deployment Link: ⊕ AI Medical Diagnosis

CHAPTER 5

Discussion and Conclusion

1. **Future Work:**

As AI-based medical diagnostics continue to evolve, there are several potential areas for improvement in future iterations of this project. One major enhancement could be increasing the dataset size and diversity to improve the model's generalization across different demographics and medical conditions. Additionally, the inclusion of more sophisticated deep learning models, such as transformer-based architectures, may enhance the accuracy of text and image-based medical diagnosis.

Another critical area for future work is integrating federated learning, which allows models to be trained across multiple decentralized medical datasets without compromising patient privacy. This approach can help overcome data-sharing limitations while improving model robustness. Furthermore, the implementation of explainable AI (XAI) techniques can enhance model transparency, making it easier for medical professionals to interpret AI-driven predictions.

Lastly, expanding the system to support real-time wearable health monitoring and IoT-based medical devices would allow continuous patient health tracking, enabling early disease detection and preventive healthcare measures.

1. **Conclusion:**

This project demonstrates the significant potential of AI-driven medical diagnosis in improving disease detection, accuracy, and accessibility. By leveraging machine learning and deep learning techniques, the system provides a reliable tool for early disease prediction and medical decision support. The integration of real-time predictions, user-friendly web applications, and explainable AI methods ensures that healthcare professionals and patients can effectively utilize AI-driven insights.

Overall, this project contributes to the ongoing digital transformation in healthcare, highlighting the role of AI in enhancing patient outcomes, reducing diagnostic errors, and improving medical workflow efficiency. Future advancements in AI, data availability, and model interpretability will further refine and expand the capabilities of AI-based medical diagnosis systems, making them an essential component of modern healthcare solutions.

REFERENCES

1. Ming-Hsuan Yang, David J. Kriegman, Narendra Ahuja, "Detecting Faces in Images: A Survey", IEEE Transactions on Pattern Analysis and Machine Intelligence, Volume. 24, No. 1, 2002.
2. Suleimenov IE, Vitulyova YS, Bakirov AS, Gabrielyan OA. Artificial Intelligence:what is it? Proc 2020 6th Int Conf Comput Technol Appl. 2020;22–5. https://doi.org/10.1145/3397125.3397141.
3. Davenport T, Kalakota R. The potential for artificial intelligence in Healthcare. Future Healthc J. 2019;6(2):94–8. https://doi.org/10.7861/futurehosp.6-2-94.
4. McCorduck P, Cfe C. Machines who think: a personal inquiry into the history and prospects of Artificial Intelligence. AK Peters; 2004.
5. Jordan MI, Mitchell TM. Machine learning: Trends, perspectives, and prospects. Science. 2015;349(6245):255–60. https://doi.org/10.1126/science.aaa8415.
6. VanLEHN K. The relative effectiveness of human tutoring, intelligent tutoring systems, and other tutoring systems. Educational Psychol. 2011;46(4):197–221. https://doi.org/10.1080/00461520.2011.611369.
7. Topol EJ. High-performance medicine: the convergence of human and Artificial Intelligence. Nat Med. 2019;25(1):44–56. https://doi.org/10.1038/s41591-018-0300-7.
8. Esteva A, Kuprel B, Novoa RA, Ko J, Swetter SM, Blau HM, et al. Dermatologist-level classification of skin cancer with deep neural networks. Nature. 2017;542(7639):115–8. https://doi.org/10.1038/nature21056.
9. Myszczynska MA, Ojamies PN, Lacoste AM, Neil D, Saffari A, Mead R, et al. Applications of machine learning to diagnosis and treatment of neurodegenerative Diseases. Nat Reviews Neurol. 2020;16(8):440–56. https://doi.org/10.1038/s41582-020-0377-8.
10. McKinney SM, Sieniek M, Godbole V, Godwin J, Antropova N, Ashrafian H, et al. International evaluation of an AI system for breast cancer screening. Nature. 2020;577(7788):89–94. https://doi.org/10.1038/s41586-019-1799-6.
11. Kim H-E, Kim HH, Han B-K, Kim KH, Han K, Nam H, et al. Changes in cancer detection and false-positive recall in mammography using Artificial Intelligence: a retrospective, Multireader Study. Lancet Digit Health. 2020;2(3). https://doi.org/10.1016/s2589-7500(20)30003-0.
12. Han SS, Park I, Eun Chang S, Lim W, Kim MS, Park GH, et al. Augmented Intelligence Dermatology: deep neural networks Empower Medical Professionals in diagnosing skin Cancer and Predicting Treatment Options for 134 skin Disorders. J Invest Dermatol. 2020;140(9):1753–61. https://doi.org/10.1016/j.jid.2020.01.019.
13. Li S, Zhao R, Zou H. Artificial intelligence for diabetic retinopathy. Chin Med J (Engl). 2021;135(3):253–60. https://doi.org/10.1097/CM9.0000000000001816.
14. Alfaras M, Soriano MC, Ortín S. A fast machine learning model for ECG-based Heartbeat classification and arrhythmia detection. Front Phys. 2019;7. https://doi.org/10.3389/fphy.2019.00103.
15. Raghunath S, Pfeifer JM, Ulloa-Cerna AE, Nemani A, Carbonati T, Jing L, et al. Deep neural networks can predict new-onset atrial fibrillation from the 12-lead ECG and help identify those at risk of atrial fibrillation–related stroke. Circulation.2021;143(13):1287~98. https://doi.org/10.1161/circulationaha.120.047829.
16. Becker J, Decker JA, Römmele C, Kahn M, Messmann H, Wehler M, et al. Artificial intelligence-based detection of pneumonia in chest radiographs. Diagnostics. 2022;12(6):1465. https://doi.org/10.3390/diagnostics12061465.
17. Mijwil MM, Aggarwal K. A diagnostic testing for people with appendicitis using machine learning techniques. Multimed Tools Appl. 2022;81(5):7011–23. https://doi.org/10.1007/s11042-022-11939-8.
18. Undru TR, Uday U, Lakshmi JT, et al. Integrating Artificial Intelligence for Clinical and Laboratory diagnosis - a review. Maedica (Bucur). 2022;17(2):420–6. https://doi.org/10.26574/maedica.2022.17.2.420.
19. Peiffer-Smadja N, Dellière S, Rodriguez C, Birgand G, Lescure FX, Fourati S, et al. Machine learning in the clinical microbiology laboratory: has the time come for routine practice? Clin Microbiol Infect.

2020;26(10):1300–9. https://doi.org/10.1016/j.cmi.2020.02.006.

20. Smith KP, Kang AD, Kirby JE. Automated interpretation of Blood Culture Gram Stains by Use of a deep convolutional neural network. J Clin Microbiol. 2018;56(3):e01521–17. https://doi.org/10.1128/JCM.01521-17.

-
-

Apply Desired Machine/ Deep Learning ModelPre-processing Disease Symptoms Feature VectorMedical Dataset

Data CleaningData Transformation

Prediction Model

Test Data Feature VectorMedical Test DataTest Data Pre-processingDisease Predicted