Human disease detection

Harshit kumar
AIT-CSE,
Chandigarh University,
Gharuan, India
harshit2003dec@gmail.com

Yudhishter kumar
AIT-CSE,
Chandigarh University,
Gharuan, India
yudhishtergurjar44@gmail.com

Akansha rohilla
AIT-CSE,
Chandigarh University,
Gharuan, India
rohillaakansha90@gmail.com

Shubham sharma
AIT-CSE,
Chandigarh University,
Gharuan, India
Shubhamsharma781806@gmail.com

Kirti
AIT-CSE,
Chandigarh University,
Gharuan, India
Supervisor

Abstract - The evolving landscape of health information needs is reshaping how people seek and access medical knowledge worldwide. Many individuals encounter challenges when searching online for information on diseases, diagnoses, and various treatments. Developing a recommendation system that utilizes review mining could greatly streamline this process for healthcare professionals and patients alike. However, users often struggle to navigate the diverse and complex medical terminology found across various platforms. The goal of such a system is to tailor recommendations to meet the unique needs of users within the health domain.

Keywords: *Health information, Recommendation system, Healthcare, Patients, Disease, Treatment accessibility.*

I.INTRODUCTION

Healthcare for humans is one of the most crucial topics in society. The identification of the nature of the illness or other problems by examination of the symptoms i.e. diagnosis always stands first in the overall curing procedure of disease. Thus, we can say that diagnosis and prediction of disease are the most crucial aspects to be considered before thinking about the exact procedure of curing the disease. The diagnostic process demands considerable time and financial investment. As a result, people with poor financial backgrounds cannot get accurate diagnoses of disease, which may create life or death. So WeCare is basically a Human Disease Detection System which is supposed to input some data related like sugar levels etc. and in response to that it gives the type of disease that person is having as an output. So, this System will be salutary to people who are not able to pay huge amounts to get diagnosed as well as the ones who require an immediate diagnosis.

There are several reasons why a client use a human disease detection system. These include:

- **Early detection:** Detecting a disease in its early stages can greatly increase the chances of successful treatment and recovery. Early detection can also prevent the disease from progressing to more severe stages, which can be more difficult and expensive to treat.
- **Prevention:** Detection of certain diseases can enable healthcare professionals to identify individuals who may be at high risk of developing a disease in the future. With this information, preventive measures can be taken to reduce the risk of disease occurrence, such as changes in lifestyle or medications.
- **Treatment planning:** Accurate disease detection can aid healthcare professionals in determining the best course of treatment for an individual. Different diseases may require different treatment approaches, and early detection can enable healthcare professionals to develop a personalized treatment plan that is most effective for each individual.
- **Public health:** Detecting and monitoring diseases is important for public health, as it can help identify outbreaks and trends in disease occurrence. This information can help create plans to stop diseases from spreading in communities.
- **Economic impact:** Disease detection and treatment can have a significant economic impact, both on individuals and society as a whole. Early detection and treatment can reduce healthcare costs, improve productivity, and improve quality of life for individuals and their families.
- **Cost Savings:** Remote disease detection can lead to cost savings for both individuals and healthcare systems. By avoiding unnecessary hospital visits and diagnostic tests, users can reduce healthcare expenses, while healthcare facilities can allocate resources more efficiently, potentially lowering overall healthcare costs.

II.BACKGROUND

This project focuses on machine learning (ML), a subset of artificial intelligence that simulates human behavior using data and algorithms. ML is broadly categorized into four branches:

Supervised ML: Supervised learning involves mapping input data to output labels, typically used in classification or regression tasks. Common algorithms in this domain include SVM, Naïve bayes, random forest, ANN, logistic regression. The objective is to identify patterns or structures in the input data that enable accurate prediction of output data. It's important to note that the correctness of output labels is determined solely from the training data, which may not always reflect real-world accuracy due to noisy or incorrect labels. Model complexity and the bias-variance tradeoff are key considerations in supervised learning, with both factors influencing each other.

Unsupervised ML: Unsupervised learning encompasses tasks such as clustering, representation learning, and density estimation, where the goal is to discover the inherent structure of the data without labeled guidance. Common algorithms in unsupervised learning include k-means clustering, principal component analysis, and autoencoders. Since no labels are provided, evaluating model performance in unsupervised learning is more challenging compared to supervised learning methods.

Semi-supervised learning: It utilizes both labeled and unlabeled data to enhance model performance, particularly beneficial in scenarios where obtaining labeled data is restricted or expensive.

Reinforcement learning: It entails an agent acquiring optimal actions through interaction with an environment, receiving feedback in the form of rewards or penalties.

III.EXISTING SYSTEM APPROACH:

The existing system approach to human disease detection encompasses a variety of diagnostic methods and tests aimed at identifying the presence of specific diseases or conditions. These methods often involve a combination of medical history assessment, physical examinations, laboratory tests, and imaging studies. Here are some examples of common diagnostic approaches in the existing system:

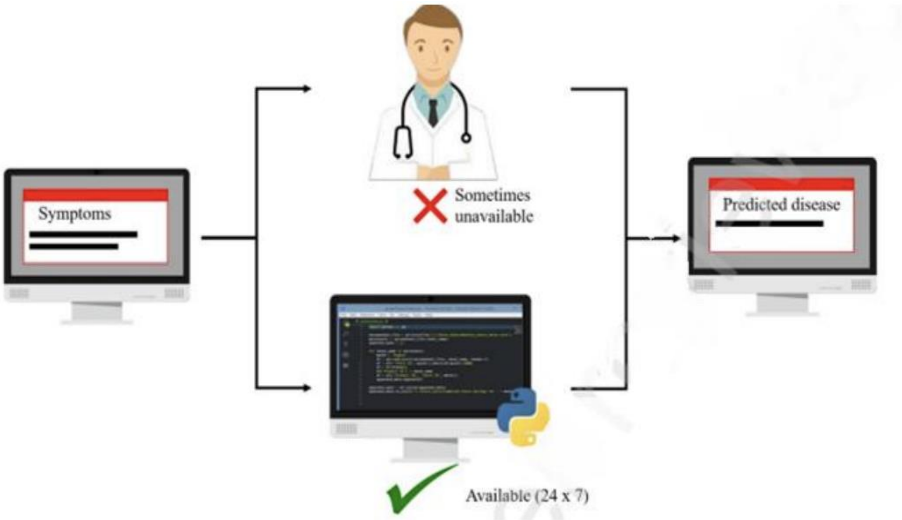
- **Medical History Assessment:** Healthcare professionals typically begin the diagnostic process by collecting information about the patient's medical records, including past illnesses, family history of diseases, medication usage, and lifestyle factors. This information helps in identifying potential risk factors and guiding further diagnostic investigations.
- **Physical Examination:** A comprehensive physical assessment is performed to assess the patient's general health condition and detect any observable indications or symptoms of illness. Healthcare providers may examine vital signs, palpate specific areas of the body, listen to heart and lung sounds, and perform other assessments relevant to the suspected condition.
- **Laboratory Tests:** Laboratory tests play a crucial role in disease detection by analyzing biological samples such as blood, urine, saliva, or tissue specimens. These tests can detect abnormalities in various parameters, including blood cell counts, biochemical markers, hormone levels, infectious agents, genetic mutations, and tumor markers. Examples of common laboratory tests include complete blood count (CBC), blood chemistry panels, urinalysis, microbiological cultures, and genetic testing.
- **Imaging Studies:** Various imaging modalities, such as X-rays, computed tomography (CT), magnetic resonance imaging (MRI), and ultrasound, are employed to visualize internal structures and detect abnormalities indicative of specific diseases or conditions.

IV.STRUCTURE OF PROPOSED METHODOLOGY:

The proposed methodology aims to revolutionize disease detection by eliminating the need for hospital visits through a streamlined and accessible approach. The structure of the proposed methodology consists of several key components designed to facilitate remote symptom assessment and disease prediction. Here's an outline of the proposed methodology structure:

- **Symptom Selection Interface:** The methodology begins with the development of a user-friendly interface accessible via web or mobile platforms. This interface allows individuals to input their symptoms conveniently from the comfort of their homes. The symptom selection interface should be intuitive, featuring a comprehensive list of symptoms organized in a user-friendly format.
- **Data Collection and Processing**: Upon symptom input, the selected data are collected and processed using advanced algorithms to prepare them for disease prediction. This stage involves cleaning and preprocessing the data to ensure accuracy and consistency. Additionally, any missing or incomplete information may be addressed through data imputation techniques to enhance the reliability of the predictive model.
- **Machine Learning Model Integration:** The preprocessed symptom data are then fed into a machine learning model trained on a diverse dataset of known disease-symptom associations. Various machine learning algorithms such as neural networks, support vector machine, random forest or decision tree can be employed to develop the predictive model. The model is trained to recognize patterns and correlations between symptoms and diseases, enabling accurate disease prediction.
- **Disease Prediction and Output:** Using the trained machine learning model, the methodology predicts the likelihood of various diseases based on the input symptoms. The output provides a ranked list of potential diseases along with corresponding probabilities or confidence scores. This information empowers individuals to make informed decisions about seeking medical attention or pursuing further diagnostic evaluation
- **Accessibility and Scalability:** The proposed methodology prioritizes accessibility and scalability to reach a wide audience and accommodate varying healthcare needs. It should be accessible across diverse devices and internet connectivity levels, catering to individuals in remote or underserved areas. Scalability ensures that the methodology can handle increasing volumes of users and adapt to

evolving healthcare demands.



In Fig. 1, we present our proposed disease prediction system, which can be used as a reliable substitute for doctors who may not always be available. This system can be used at any time to predict diseases based on an individual's symptoms, name and gender, which are fed into the ML model for further processing. After preliminary data processing, the ML model uses the input data to train and test the algorithm, resulting in the predicted disease with high accuracy.

Functioning of ML models:

In our approach to remote disease detection, we utilize multiple machine learning models, including Random Forest, K-Nearest Neighbors (KNN), Naive Bayes, and Decision Tree. Each of these models undergoes training and evaluation to determine which one exhibits the best accuracy. The model with the highest accuracy is then selected for disease prediction.

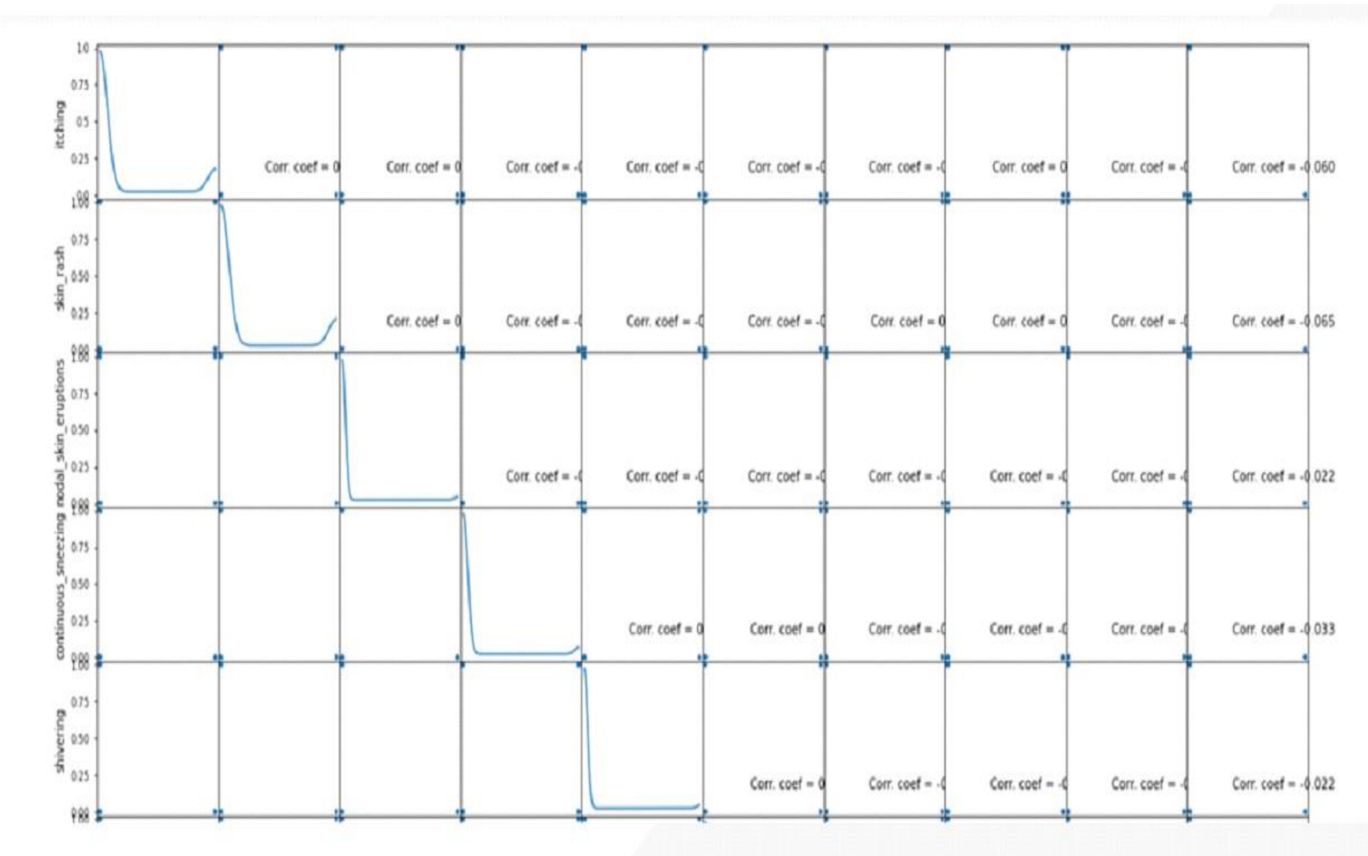


Fig.2: Scatter and Density plot

V. DRAWBACKS:

Accuracy Concerns: While our model strives for high accuracy, there's always a margin of error inherent in predictive algorithms. False positives or false negatives could occur, leading to misdiagnosis or unnecessary worry for users.

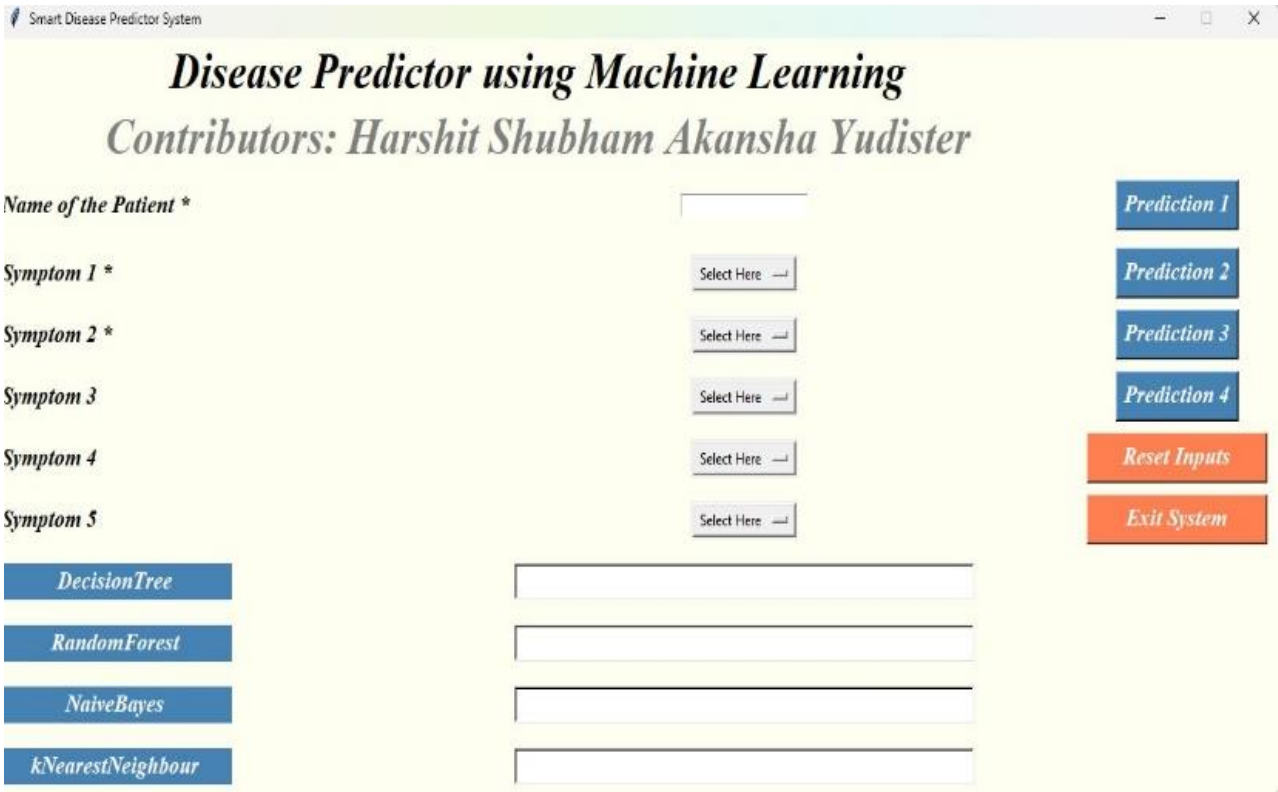
Dependency on Data Quality: The quality and representativeness of our model depend significantly on the data used for training. Biases or inaccuracies in the training data may influence the model's predictions, possibly resulting in biased outcomes.

Privacy Concerns: Despite robust privacy measures, there's always a risk of unauthorized access or data breaches, especially when handling sensitive health information. Maintaining strict security protocols and compliance with data protection regulations is crucial to mitigate these risks.

Digital Accessibility Barriers: Not all individuals may have access to the technology required to use our model effectively, such as internet connectivity or compatible devices. This could exacerbate existing healthcare disparities, leaving certain populations underserved or excluded from the benefits of remote disease detection.

VI. RESULTS:

Our findings highlight the importance of utilizing symptom-based models in disease detection, particularly in scenarios where access to medical expertise may be limited or where early detection is critical for timely intervention. By allowing individuals to input their symptoms directly into the model, we empower them to take a proactive role in managing their health while also providing valuable insights to healthcare providers. By harnessing the power of machine learning and human-centric design, we can pave the way for a future where disease diagnosis is more accessible, accurate, and personalized, ultimately improving health outcomes and quality of life for all. The implementation of our GUI marks a significant milestone in bridging the gap between advanced technology and everyday healthcare needs. By offering a user-friendly platform for symptom input and disease prediction, we have facilitated seamless interaction between individuals and predictive models, thereby empowering users to make informed decisions about their health with ease and confidence.
The GUI designed for this project is a basic Tkinter interface featuring labels, a message box, a button, text fields, a title, and a dropdown menu.



VII. FUTURE WORK:

In our future work, we aim to advance our remote disease detection system by integrating Advanced machine learning methods like deep learning and ensemble techniques are utilized to enhance predictive accuracy. Additionally, we plan to personalize disease predictions by incorporating individual patient data and lifestyle factors. Real-time monitoring capabilities will enable users to track their health status continuously, while enhanced user engagement and education will empower individuals with knowledge about their health. Collaborations with healthcare providers will facilitate seamless integration into clinical practice, supported by rigorous validation studies and adherence to ethical and regulatory standards. Our overarching goal is to create a transformative healthcare solution that improves accessibility, empowers individuals, and enhances health outcomes globally.

VIII. REFERENCES:

[1] https://www.researchgate.net/publication/349054979_Human_Diseases_Detection_Based_On_Mach

[2] Comparison of Machine Learning Approaches for Prediction of Advanced Liver Fibrosis in Chronic

[3] Cinarer, G., & Emiroglu, B. G. (2019). Classification of Brain Tumors by Machine Learning Algorithms. 3rd International Symposium on Multidisciplinary Studies and Innovative Technologies, ISMSIT 2019 -Proceedings. https://doi.org/10.1109/ISMSIT.2019.8932878

[4] Acharya, U. R., Fujita, H., Oh, S. L., Hagiwara, Y., Tan, J. H., & Adam, M. (2017). Application of deep convolutional.

[5] Hashem, S., Esmat, G., Elakel, W., Habashy, S., Raouf, S. A., ElHefnawi, M., Eladawy, M., & ElHefnawi, M. (2018).

[6] Parkinson's disease. Computer Methods and Programs in Biomedicine, 113(3), 904–913. 35. https://doi.org/10.1016/j.cmpb.2014.01.004

[7] hybrid intelligent system for accurate detection of human diseases.

[8] Grimson, J., Stephens, G., Jung, B., Grimson, W., Berry, D., & Pardon, S. (2001). Sharing healthcare records over Internet IEEE Internet Computing, 5(3), 49–58.

[9] Biology Society, EMBS, 7334–7337. https://doi.org/10.1109/EMBC.2013.6611252

[10] Fan, C. H., Hsu, Y., Yu, S. N., & Lin, J. W. (2013). Detection of myocardial ischemia episode using morphological

[11] productivity and outcomes of care. Medical Care, 15(6), 482–487. https://doi.org/10.1097/00005650-26.197706000-00004

[12] Daniels, M., & Schroeder, S. A. (1977). Variation among physicians in the use of laboratory tests II. Relation to clinical

[14] Disease Detection. 1st International Informatics and Software Engineering Conference: Innovative Technologies for Digital Transformation, IISEC 2019 - Proceedings, 2–5. 12. https://doi.org/10.1109/UBMYK48245.2019.8965542

[16] https://www.nibusinessinfo.co.uk/content/five-stages-project-management

[17] https://www2.deloitte.com/us/en/pages/life-sciences-and-health-care/articles/future-of-artificial-intelligence-in-health-care.html

[18] https://www.brookings.edu/research/risks-and-remedies-for-artificial-intelligence-in-health-care/

[19] https://www.analyticsinsight.net/top-10-ai-breakthroughs-in-healthcare-to-note-down/

[20] https://www.ncbi.nlm.nih.gov/pmc/articles/PMC8%20950225/#:~:text=While%20the%20individual's%20ability%20%20restricts,health%20care%20may%20be%20developed