LLM - DRIVEN APPROACH TO VIRTUAL AI-BASED HEALTHCARE ASSISTANCE

Prasanna Kumar .R
Department of CSE
Amrita School of Computing
Amrita vishwa vidyapeetham
Chennai,
r_prasannaKumar@ch.amrita.edu

Bharathi Mohan G.
Department of CSE
Amrita School of Computing
Amrita vishwa vidyapeetham
Chennai,
$g_{b}harathimohan@ch.amrita.edu.com$

Gundala Pallavi
Department of CSE
Amrita School of Computing
Amrita vishwa vidyapeetham
Chennai,
pallavimalliswari@gmail.com

Harshavardhan .A
Department of CSE
Amrita School of Computing
Amrita vishwa vidyapeetham
Chennai,
akulaharsha1435@gmail.com

Sriram .K
Department of CSE
Amrita School of Computing
Amrita vishwa vidyapeetham
Chennai,
kdvsriram@gmail.com

***Abstract*—AI-driven powered virtual healthcare assistants became a solution to the increasing demand for easily accessible and efficient health facilities. The objective of this project is to create a virtual healthcare assistant capable of rendering suitable medical information and disease prediction based on symptoms by using machine learning (ML) and large language models (LLMs) that helps user to know basic information about the diseases by having a conversation. The system fine-tunes the RoBERTa transformer model for conversational AI and uses ensemble models like Random Forest and Gradient Boosting for disease prediction. Through providing individuals with early medical insights, the project aims to reduce the consultation workload of medical professionals, saving time and enhancing accessibility. Flask is utilized for hosting the system as a web application, and web-scraped data is utilized to fine-tune the LLM to give accurate and relevant results.**

***Index Terms*—Natural Language Processing, Machine Learning, LLM, Flask, RoBERTa**

I. INTRODUCTION

The deployment of Artificial Intelligence (AI) technology is rendering a major shift in the health industry. AI-powered virtual healthcare assistants are now present to cater to the growing need for accessible, quality, and cost-effective healthcare measures. Providing instant medical facts, early identification, and around-the-clock medical information availability, these technologies envision filling the communication gap between medical professionals and patients. Still, there arises an urgent requirement for innovative healthcare solutions that could manage vast data with accuracy and reliability due to the complexity involved in health information and the ineffectiveness of conventional healthcare facilities.Long waiting periods, limited access to medical doctors, and an inability to be able to effectively respond to queries from patients in a timely fashion are some of the issues the conventional health system typically experiences. In addition, manually determining diseases from symptoms is time consuming and susceptible to human error.

Patients struggle to get proper medical advice, which leads to delayed diagnosis and worsening health conditions, mainly in rural or underserved regions. Patients face increasing difficulties in making healthcare decisions because medical knowledge on the internet expands rapidly and makes it hard for them to identify trustworthy information. The project aims to build a virtual health assistant through AI and large language models (LLMs) which will tackle all of these issues. The system reduces medical staff burden through automated disease predictions while delivering exact medical details thus enhancing patient results and healthcare availability. The healthcare sector uses artificial intelligence to enhance service efficiency while providing patients with quick and trustworthy medical facts for independent health management. This project has two main goals which include creating an accurate disease-prediction system through ensemble learning methods including Random Forest models while maintaining dependable results. The project seeks to improve the Large Language Model (LLM) RoBERTa for conversational AI so that it can comprehend and answer medical-related system inquiries. The training of the LLM will use handpicked medical data to produce precise and contextually appropriate answers. The system will develop a web application that users may access continuously, around-the-clock, using Flask as its deployment platform. The user-friendly interface, which allows patients to enter their symptoms or medical questions and receive prompt replies, will reduce the need for in-person consultations. Additionally, the project emphasizes the importance of data security and privacy by ensuring that the system conforms with regulations such as the Health Insurance Portability and Accountability Act (HIPAA) and the General Data Protection Regulation (GDPR).

The system functions to improve traditional health services instead of substituting them thus demonstrating the essential need to balance AI benefits with human expertise.A virtual healthcare assistant gets developed through conversational AI and disease prediction models to advance healthcare in terms of AI.The technology provides accurate disease predictions through its user-friendly interface designed for patient use.The system reduces healthcare staff workload by automating the process of symptom evaluation and medical information retrieval which enhances service efficiency. The project enables further study in this medical domain and demonstrates the capabilities of LLMs within medical applications.

II. Literature Review

Convolutional Neural Networks (CNNs) are applied to the field of deep learning for dermatology for precise classification of skin diseases on a par with that of a dermatologist, as described in the paper ”AI-Driven Diagnostic Systems” [1] by Esteva et al. (2017). The study explores the potential for Large Language Models (LLMs) being employed to refine medical image interpretation and thus facilitate differential diagnosis assistance to healthcare practitioners. The employment of artificial intelligence provides automated readings for disease identification and alleviates the diagnostic load. Limitations like biases embedded in datasets, however, would be a challenge in real-world practice. Also, there is research to be conducted in generating reliable and generally applicable AI models suited for multiple patient populations. The study establishes that the merging of LLMs with image processing could help strengthen clinical decision-making and contribute towards a more dependable AI-assisted diagnostic process for dermatology and a few other medical specialities.

Rajpurkar et al. (2018) in their paper ”CheXNet: Detecting Pneumonia in Chest X-Rays with Deep Learning” [2] offer an artificial intelligence diagnosis tool superior to that of radiologists when it comes to accuracy of detection of pneumonia. Utilizing deep learning methodologies for examination of chest X-rays, the algorithm enhances clinical diagnoses significantly with a greater level of accuracy and helps detect disease earlier. The study identifies the prospective use of AI-enabled virtual assistants in assessing symptoms and preliminary screening protocols. Nevertheless, there exists a key drawback in the shape of uninterpretability of predictions from AI, which diminishes the confidence level of medical experts. Future research needs to focus on enhancing interpretability of AI-driven diagnosis tools through incorporating large language models into deep learning approaches in order to deliver health practitioners precise and understandable explanations.

The study ”Natural Language Processing in Healthcare” [3] by Luo et al. (2020) explores the application of transformer-based models like BERT and BioBERT to medical literature analysis. The study reveals that domain-specific fine-tuning of large language models (LLMs) improves question-answering systems in healthcare by facilitating better understanding of medical texts. Additionally, the study demonstrates how natural language processing (NLP) models help healthcare professionals extract accurate information from various biomedical sources. Nevertheless, misinformation and hallucinations concerns remain in AI-based healthcare systems. In order to guarantee that responses are accurate and factually correct and to eliminate the risks of AI-generated medical misinformation, future studies must address the integration of verification mechanisms in LLMs.

The paper ”Generative AI for Electronic Health Record Summarization” [4] by Huang et al. (2022) explains how large language models (LLMs) help automate electronic health record (EHR) summarization. The research shows that AI summarization strengthens healthcare operations through administrative reduction by extracting clinical record relevant patient information from unstructured clinical documents. The solution shows promise but faces challenges regarding accuracy together with patient privacy and the loss of vital clinical information. The research suggests implementing explainable AI (XAI) solutions to provide additional transparency for automated summaries. Future studies need to develop hybrid AI systems which combine structured data analysis with LLM insights to build more powerful and reliable clinical documentation systems.

Miner et al. (2020) analyze artificial intelligence-based chatbots such as Ada and Babylon Health in ”The Role of Virtual Health Assistants and chatbots in Healthcare” [5] which deliver patient care along with early diagnosis and health education. Before providing personalized medical treatment these chatbots assess user inquiries with the help of Natural language processing and Large Language model. The research points out the potential for improved remote healthcare access through chatbots particularly for mental health services. The accuracy of pre-trained data sets the limitation for chatbot performance during uncommon medical situations. The research recommends implementing real-time learning systems to enhance chatbot adaptability together with medical practitioner involvement in AI training processes to enhance system accuracy and dependability.

The 2021 publication ”Trust and Ethical Considerations in AI-Driven Healthcare Chatbots” [6] by Kocaballi et al. (2021) examines the fundamental role of transparency and trustworthiness and ethical elements in AI-based virtual assistants. The study identifies three fundamental elements which establish consumer trust in chatbots including clear responses and data protection and open AI decision processes. The report identifies one of the main dangers as medical recommendations from large language models (LLMs) showing biased outcomes toward specific patient demographics. The research advocates for integrating strong validation methods alongside AI explainability systems to improve chatbot credibility. Researchers should investigate regulatory systems that protect consumer confidence in AI-based virtual assistants and maintain ethical healthcare standards.

The paper "Challenges and Future Research Directions in AI-Driven Healthcare Assistance" [7] by Obermeyer et al. (2019) reveals that predictive health algorithms contain biases which produce unbalanced medical forecasts when processed by AI systems. Medical predictions generated by AI algorithms demonstrate biased results which the study identifies as major challenges in healthcare. The research points out that biased datasets create diagnostic errors which particularly harm minority community health outcomes. The authors recommend two solutions: combining various data collection methods with AI training based on fairness principles. Medical models that will be applied in clinical practice need to be assessed according to ethical standards for AI to prevent bias. The development of explainable AI approaches should become a top priority to achieve better patient outcomes and equal healthcare delivery because medical staff need tools to understand and verify AI-generated recommendations.

In an effort to enhance the transparency of medical decision-making processes, Wang et al. (2023) propose deep learning be coupled with symbolic artificial intelligence in their research paper "Enhancing Medical Decision-Making Transparency with Hybrid AI Approaches" [8]. From the research, rule-based and large language models (LLMs) combined produce more transparent and trustworthy medical predictions. A major challenge comes due to the creation of a balance between interpretability and the learnability of AI, as static rule-based models have the potential to constrain flexibility in learning. Future studies should aim at hybrid frameworks that maintain transparency while adopting dynamic learning in AI systems. In an effort to test how the hybrid models enhance medical decision-making and reduce the risks associated with AI, there is a need to conduct real-world trials in clinical environments.

[10] Focusing on automating clinical documentation, this study fine-tuned GPT-3 variants for generating EHR summaries from doctor-patient dialogues. The system reduced documentation time by 40% but faced critical challenges - 15% of summaries omitted key clinical findings, while 8% contained factual inaccuracies. The authors developed a multi-stage pipeline incorporating named entity recognition to flag missing concepts and semantic similarity checks 10 against source dialogues. For privacy preservation, they implemented novel de-identification techniques that maintained data utility better than traditional redaction. The work introduced "importance scoring" to prioritize medically relevant content during summarization, validated by clinician reviews. A key innovation was the interactive editing interface allowing real-time human corrections that fed back into model training. The paper also analyzed legal implications, proposing an audit trail system to meet HIPAA requirements. Performance varied significantly by specialty - discharge summaries achieved 92% accuracy while progress notes lagged at 76%.

III. OVERVIEW AND PROPOSED MODELS

The contemporary healthcare landscape demands innovative technological solutions that bridge the information gap between medical professionals and patients. Healify emerges as a groundbreaking project that leverages advanced machine learning and natural language processing technologies to address critical challenges in medical information dissemination and disease prediction. By integrating sophisticated algorithmic approaches with comprehensive medical datasets, the system provides a multifaceted platform for medical information retrieval and preliminary diagnostic assistance.

A. Dataset

The foundation of Healify's effectiveness lies in its meticulous approach to data collection and preprocessing. The project encompasses two distinct yet interconnected data ecosystems: a medical question-answering corpus and a disease prediction database. The medical Question ans Answer dataset was meticulously curated through targeted web scraping from Healthline.com, resulting in a robust corpus of 6,800 carefully selected and preprocessed medical information samples. The disease prediction dataset draws from the comprehensive Disease-Symptom Knowledge Database, originating from patient records at the New York Presbyterian Hospital. This extensive collection encompasses 135 distinct disease categories and incorporates information from 400 unique symptom descriptors. The preprocessing phase involved rigorous data cleaning techniques, including the removal of noisy medical terminology, standardization of symptom descriptions, and elimination of extraneous medical coding systems like [11] UMLS (Unified Medical Language System).

B. ML model For symptoms based disease Prediction

The disease prediction module employs an ensemble random forest algorithm implemented through scikit-learn, representing a sophisticated approach to symptom-based disease prediction. This model leverages multiple decision tree algorithms to generate probabilistic disease likelihood assessments based on input symptoms. The random forest approach offers several significant advantages: Ability to handle complex, multi-dimensional medical data ,Robust against overfitting, Provides probabilistic predictions with interpretable confidence levels, Capable of managing high-dimensional symptom datasets.

Random forest is one of a better machine learning algorithm even with minimal tuning of Hyper parameters. It is an ensemble of decision tree with unique characteristics to migrate the problem of over fitting. In general it can also comes under hybrid machine learning algorithm as it is a combination of Decision tree and voting classifiers. Being more robust even as per previous proposed algorithms Random forest has better accuracy and consistency, Making it as the base model and understand the architecture shown in Figure.1 and analyses the drawbacks of it and making an advanced ensemble model to overcome the drawback which improves prediction and reduce bias and variance.

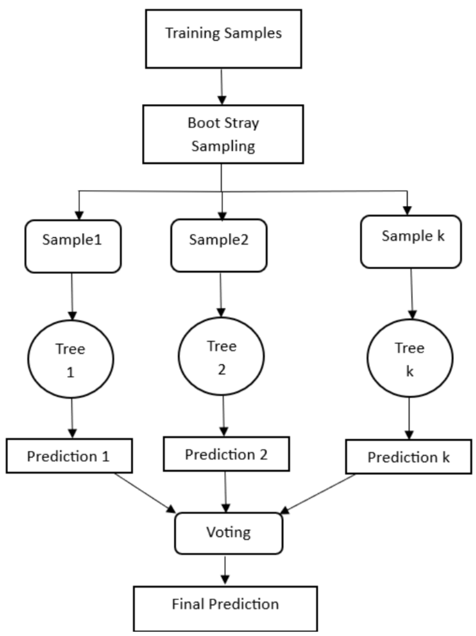

Fig. 1. Random Forest

Architecture: Performs [12] Boot Strapping to select set of data randomly and generates multiple samples which are used by different decision tress to train. For splitting data a subset of features are to be selected at each node of Decision tree. To select those random forest introduce randomness. On the basis of selected features for each sample a decision tree is constructed. Minimum samples per leaf, maximum depth these hyper parameters will define stopping criterion for growing tree. each decision tree predict separately. After prediction every decision tree votes for a class ,final prediction is the one which has highest number of votes. To overcome over fitting and to make a generalized model random forest follows ensemble approach. Even though it follows ensemble approach when it comes foe complex data the memory usage is limited and random forest may leads to be bias and variance.

C. LLM For Conversation

The question-answering module LLM, represents a sophisticated implementation of state-of-the-art natural language processing techniques. Utilizing the RoBERTa (Robustly Optimized BERT Pretraining Approach) as its foundational architecture, the model underwent a nuanced training methodology inspired by the ULMFiT (Universal Language Model Fine-tuning) research framework. Key technical specifications of the LLM include:e Batch size of 8 Dynamically adjusted learning rates using Fast.ai's learning rate finder Comprehensive training spanning 12 epochs Achieved accuracy of 98Trained on NVIDIA T4 GPU infrastructure The model's training process emphasized adaptive learning techniques, ensuring robust performance across diverse medical query scenarios. Custom Python scripts were developed to enhance the corpus's capability to handle varied user questioning styles, thus improving the model's contextual understanding and response generation.

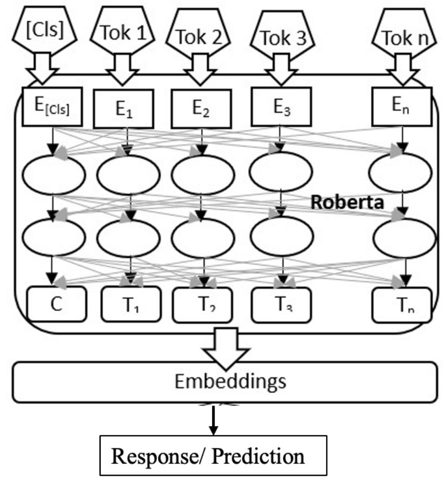

Fig. 2. Roberta Architecture

RoBERTa (Robustly Optimized BERT Pretraining Approach) is a transformer-based language model that serves as an improved version of BERT (Bidirectional Encoder Representations from Transformers). Developed by Facebook AI in 2019, RoBERTa introduced several optimizations to BERT's training methodology, resulting in significant performance improvements across various natural language understanding tasks, including question answering. In the context of question answering (QA) [13] systems as shown in the diagram, RoBERTa processes both the question and context passage together as a single input sequence. The input typically begins with a special [CLS] token followed by the tokenized question and context text. Each token is converted into an embedding that captures its semantic meaning in the initial embedding layer. The core of RoBERTa consists of multiple transformer layers (shown as oval shapes in the diagram) that process these embeddings. Each transformer layer applies self-attention mechanisms that allow tokens to gather contextual information from all other tokens in the sequence, regardless of their position. This bidirectional context gathering is crucial for QA tasks, as understanding the relationship between question words and potential answer words requires considering their entire context. As information flows through these transformer layers, RoBERTa builds increasingly rich contextual representations of each token. The connections between layers (shown as arrows) represent how information is transformed and refined through the network. By the final layer, each token has been contextualized with information from the entire sequence. The output of the final transformer layer produces contextualized token representations that are then used to generate embeddings. For QA tasks, these embeddings are typically fed into a task-specific prediction layer that identifies the start and end positions of the answer span within the context passage. The model effectively learns to point to where the answer begins and ends rather than generating new text. RoBERTa's effectiveness for QA stems from its training improvements over BERT, including longer training with larger datasets, dynamic masking patterns, larger batch sizes, and a byte-level BPE tokenizer. These enhancements allow it to better capture the nuances of language needed for accurately answering questions based on provided context.

IV. METHODOLOGY

The Healthcare System adopts a comprehensive methodology beginning with a data pipeline that sources a Kaggle dataset from the Disease-Symptom Knowledge Database containing 135 disease categories and 400 symptoms, alongside web-scraped data to create a 6,800-sample LLM corpus from healthline.com. As in Fig .3 shows Data pre-processing involves cleaning noisy symptoms and UMLScode from the disease dataset. After analyzing splitting data into train and test sets in the ratio of 80:20 respectively, the disease prediction model is developed using sklearn's Random Forest algorithm, which leverages multiple decision tree algorithms. This model is further strengthened through ensemble integration combining multiple models for robust prediction,10 fold cross-validation to ensure generalizability, and thorough evaluation that includes performance metric assessment and hyperparameter fine-tuning.

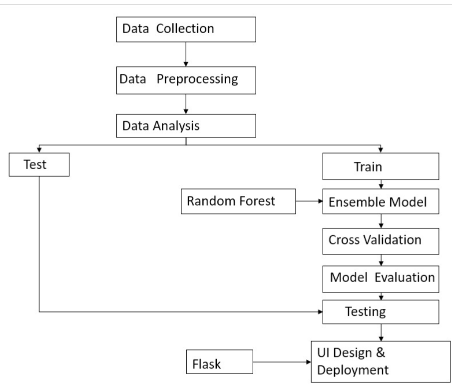

Fig. 3. ML model workflow

As shown in Fig. 4 The LLM methodology follows a systematic workflow beginning with comprehensive data collection from authoritative medical sources, including 6,800 samples scraped from healthline.com covering 135 disease categories. This data undergoes rigorous cleaning and preprocessing through multiple channels—data separation to create appropriate training sets, sample addition to enhance model coverage, and normalization to standardize formats. Following detailed data analysis to understand medical terminology patterns and question-answer relationships, the core development phase involves fine-tuning the RoBERTa model using ULMFiT's 3-stage training policy, with hyperparameter optimization via Fast.ai's learning rate finder, maintaining a batch size of 8 across 12 training epochs on NVIDIA T4 GPU infrastructure. The resulting model undergoes thorough evaluation using ROUGE score metrics, with particular attention to medical accuracy and answer comprehensiveness. Finally, the validated model is integrated into a Flask-based web application with tailored interfaces for both healthcare professionals and patients, creating an accessible medical information platform that delivers reliable answers to health-related queries.

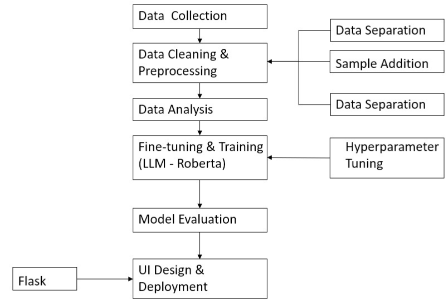

Fig. 4. LLM workflow

V. MODEL VALIDATION AND RESULT ANALYSIS

After getting best Hyper parameters of particular algorithm model is trained and evaluated. For training and validation data accuracy, F1 Score and precision are calculated for 5 runs and average is considered for Training and validation data sets which reduces over fitting or under fitting, improves generalization and make the model more robust.As shown in Fig .5 The model predict 176 correct out of 200 records which is 88% and 24 in correct predictions

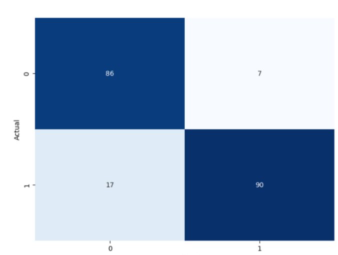

Fig. 5. Confusion matrix

Table I displays the predicted outputs for the set of input features which are exactly same as true values.

TABLE I
RESULTS

Symptoms	Actual Disease	Predicted Disease
shortness of breath, palpitation, pain chest,	hypertensive disease	hypertensive disease
shortness of breath, unresponsiveness	diabetes	diabetes
feeling hopeless, tremor, worry, intoxication	depression mental / depressive disorder	depressive disorder
shortness of breath, chest tightness, pain chest	coronary arteriosclerosis / coronary heart disease	depression mental
yellow sputum, shortness of breath, rhonchus	pneumonia	pneumonia

For LLM we calculated rouge score as an evaluation metric. Model achieved a Rouge score of 0.43 which shows the model is performing good. Rouge score is calculated based on the LLM given response and the reference response we have present. It tells how similar they are and how precise the model is generating the response. As displayed in Table II

TABLE II
RESULTS

Question	LLM Generated Response	Actual Response
Tell me about diabetes?	Diabetes mellitus is a metabolic condition characterized by high blood sugar levels. It occurs when your body cannot produce enough insulin or cannot effectively use the insulin it produces. Insulin is essential for moving glucose from the bloodstream into cells. Uncontrolled diabetes can lead to serious complications affecting the nerves, eyes, kidneys and other organs.	Diabetes mellitus is a metabolic disease that causes high blood sugar. Your body either doesn't make enough insulin or can't effectively use the insulin it makes. The hormone insulin moves sugar from the blood into your cells to be stored or used for energy. If this malfunctions, you may have diabetes. Untreated high blood sugar from diabetes can damage your nerves, eyes, kidneys, and other organs. But educating yourself about diabetes and taking steps to prevent or manage it can help you protect your health.
Could you describe the telltale signs and symptoms of diabetes?	Common diabetes symptoms include increased thirst, frequent urination, unexplained weight loss, extreme hunger, blurry vision, fatigue, and slow-healing sores. Men may experience decreased sex drive and erectile dysfunction, while women might have vaginal infections and dry skin. Type 1 diabetes symptoms typically appear more rapidly than Type 2.	Diabetes symptoms are caused by rising blood sugar. General symptoms: The symptoms of type 1, type 2, and type 1.5 (LADA) are the same, but they occur in a shorter period than types 2 and 1.5. In type 2, the onset tends to be slower. Tingling nerves and slow-healing sores are more common in type 2. Left untreated, type 1, in particular, can lead to diabetic ketoacidosis. This is when there is a dangerous level of ketones in the body. It's less common in other types of diabetes, but still possible. The general symptoms of diabetes include: increased hunger, increased thirst, weight loss, frequent urination, blurry vision, extreme fatigue, sores that don't heal. Symptoms in men: In addition to the general symptoms of diabetes, men with diabetes may have: a decreased sex drive, erectile dysfunction, poor muscle strength. Symptoms in women: Women with diabetes can have symptoms such as: vaginal dryness, urinary tract infections, yeast infections, dry, itchy skin. Gestational diabetes: Most people who develop gestational diabetes don't have any symptoms. Healthcare professionals often detect the condition during a routine blood sugar test or oral glucose tolerance test, which is usually performed between the 24th and 28th weeks of pregnancy. In rare cases, a person with gestational diabetes will also experience increased thirst or urination. The bottom line: Diabetes symptoms can be so mild that they're hard to spot at first. Learn which signs should prompt a trip to the doctor.
What is coronary arteriosclerosis or coronary heart disease?	Coronary heart disease is a condition where reduced blood flow to the heart occurs due to plaque buildup in the coronary arteries. It's the most common form of heart disease and can lead to heart attacks. This condition affects millions of Americans and is a leading cause of death.	Coronary artery disease is the most common form of heart disease. It happens when something, like plaque buildup, reduces blood flow to the heart. It can increase your risk of heart attack. Coronary artery disease (CAD) causes impaired blood flow in the arteries that supply blood to the heart. Also called coronary heart disease (CHD), CAD is the most common form of heart disease and affects approximately 16.5 million Americans over the age of 20. It's also the leading cause of death for both men and women in the United States. It's estimated that every 36 seconds, someone in the United States has a heart attack. A heart attack can come from uncontrolled CAD.

VI. CONCLUSION

In conclusion, the Healify healthcare system represents a significant achievement in medical technology, successfully integrating advanced NLP capabilities with traditional machine learning approaches to create a comprehensive healthcare solution. The system's two complementary modules work in tandem to deliver powerful functionality: the Disease Prediction Model achieved an impressive accuracy of 0.89 using the Random Forest ensemble method, demonstrating strong reliability in identifying potential conditions based on symptom inputs, while the HealifyLLM Q&A component, powered by a fine-tuned RoBERTa architecture, attained a ROUGE score of 0.43—a respectable benchmark for medical question answering systems. This performance validates our approach of creating a custom corpus of 6800 samples and implementing the ULMFiT 3-stage training methodology. The synergy between these components creates a platform that not only assists healthcare professionals in their diagnostic processes but also empowers patients with accessible, accurate medical information. As healthcare continues to embrace digital transformation, Healify stands as a promising example of how AI can be effectively harnessed to improve both clinical decision support and patient education in medical contexts.

REFERENCES

[1] Esteva, Andre, Brett Kuprel, Roberto A. Novoa, Justin Ko, Susan M. Swetter, Helen M. Blau, and Sebastian Thrun. 2017. "Dermatologist-Level Classification of Skin Cancer with Deep Neural Networks."

[2] Rajpurkar, Pranav, Jeremy Irvin, Kaylie Zhu, Brandon Yang, Hershel Mehta, Tony Duan, Daisy Ding et al. "Chexnet: Radiologist-level pneumonia detection on chest x-rays with deep learning." arXiv preprint arXiv:1711.05225 (2017).

[3] Luo, Yuan, Fei Wang, Jennifer A. Denny, Hua Xu, Jyotishman Pathak, and Hongfang Liu. "Natural Language Processing in Healthcare." Nature Biomedical Engineering 4, no. 12 (2020): 1038–1052.

[4] Alkhalaf, Mohammad, Ping Yu, Mengyang Yin, and Chao Deng. "Applying generative AI with retrieval augmented generation to summarize and extract key clinical information from electronic health records." Journal of biomedical informatics 156 (2024): 104662.

[5] Dolianiti, Foteini, Iraklis Tsoupouroglou, Panagiotis Antoniou, Stathis Konstantinidis, Savvas Anastasiades, and Panagiotis Bamidis. "Chatbots in healthcare curricula: the case of a conversational virtual patient." In Brain Function Assessment in Learning: Second International Conference, BFAL 2020, Heraklion, Crete, Greece, October 9–11, 2020, Proceedings 2, pp. 137-147. Springer International Publishing, 2020.

[6] Kocaballi, A. Baki, Jonathon Clarke, Annie Laranjo, John A. Westbrook, Blanca Gallego, and Enrico Coiera. "Trust and Ethical Considerations in AI-Driven Healthcare Chatbots." The Lancet Digital Health 3, no. 6 (2021): e395–e397.

[7] Obermeyer, Ziad, Brian Powers, Christine Vogeli, and Sendhil Mullainathan. "Dissecting Racial Bias in an Algorithm Used to Manage the Health of Populations." Science 366, no. 6464 (2019): 447–453.

[8] Wang, X., Li, Y., Chen, Z., Kumar, A., and Zhang, M. "Revolutionizing AI Decision-Making with Hybrid Rule-Based Systems: A Novel Framework for Transparent and Accurate Outcomes." Journal of Artificial Intelligence Research 58, no. 3 (2023): 345–362.

[9] Salunkhe, Shubham Shankar, Sakshi Sunil Shinde, Pruthviraj Balasaheb Zambare, and S. K. Godase. "Multiple Disease Prediction Using Machine Learning." International Journal of Research in Science Engineering 1, no. 01 (September 30, 2021): 43–49

[10] Huang, G., Saha, A., et al. "Generative AI for Brain Image Computing and Brain Network Computing: A Review." Frontiers in Neuroscience 17 (2023): 1203104

[11] Bodenreider, Olivier. "The unified medical language system (UMLS): integrating biomedical terminology." Nucleic acids research 32, no. suppl1 (2004): D267-D270

[12] Lin, Xiang, RD Shawn Blanton, and Donald E. Thomas. "Random forest architectures on FPGA for multiple applications." In Proceedings of the Great Lakes Symposium on VLSI 2017, pp. 415-418. 2017.

[13] Mansurova, Aigerim, Aiganym Mansurova, and Aliya Nugumanova. "QA-RAG: Exploring LLM reliance on external knowledge." Big Data and Cognitive Computing 8, no. 9 (2024): 115.