Developing and Analyzing Deep Learning and Natural Language Processing Systems in the Context of Medical Information Processing

[1]Emmanuel, Victoria Nkemjika*, [2]Enyinnaya Chibunna Victor, [3]Ezeoha Bright Uzoma, [4]Ogbonna Tochukwu Loveday

[1, 4]Imo State University

[2]Abia State College of Health Sciences and Management Technology

[3]Abia State Polytechnic

***Corresponding Author**

DOI: https://doi.org/10.51584/IJRIAS.2024.908065

Received: 09 August 2024; Accepted: 14 August 2024; Published: 21 September 2024

ABSTRACT

This work aims to develop and analyze deep learning and natural language processing systems in the context of medical information processing. The amount of data created about patients in the healthcare system is always increasing. The human review of this enormous volume of data derived from numerous sources is expensive and takes a lot of time. Additionally, during a patient visit, doctors write down the patient's medical encounter and send it to nurses and other medical departments for processing. Often, the doctor doesn't have enough time to record every observation made while examining the patient and asking about their medical history which takes time for a medical diagnosis to be made. The manual review of this vast amount of data generated from multiple sources is costly and very time-consuming. It brings huge challenges while attempting to review this data meaningfully. Therefore, the goal of this research is to create a system that will address the aforementioned issues. The suggested method extracts voice data from medical encounters and converts it to text using Deep Learning (DL) and Natural Language Processing (NLP) techniques. More so, the system developed will improve medical intelligence processing by using deep learning to analyze medical datasets and produce results of a diagnosis, assisting medical professionals at various levels in making realistic, intelligent decisions in real-time regarding crucial health issues. The system was designed using the Object-Oriented Analysis and Design Methodology (OOADM), and the user interfaces were put into place utilizing Natural Language Processing techniques, particularly speech recognition and natural language comprehension. Speech recognition allows for the taking of free text notes, which can drastically cut down on the amount of time medical staff spends on labor-in the tensive clinical recording. By extracting different pieces of data for medical diagnosis and producing results in a matter of seconds, a deep learning algorithm demonstrates a significant capacity to construct clinical decision support systems. The system's results demonstrate that the deep learning algorithm enabled medical intelligence to be 96.7 percent accurate.

INTRODUCTION

In underdeveloped nations like Nigeria, doctors document patient visits on paper and pass them along to nurses and other medical departments for processing. These documents, which are used to process medical data for the healthcare industry, must be error-free. In hospitals, patient record errors are a major issue. During a patient visit, 81 percent of doctors did not have access to all the information they needed, according to observational research done in a university clinic (Tang, 2014). The amount of patient-focused data created in the healthcare system is continually increasing. Numerous departments and/or units are being added to hospitals. Medical equipment, lab findings, electronic prescriptions, therapeutic decisions, and clinically observed values by doctors and nurses could all produce patient-oriented data. These data are dispersed, so getting access to them requires requesting one of the many hospital departments. The human review of this enormous volume of data derived from numerous sources is expensive and takes a lot of time. When attempting to review this data meaningfully, it

ISSN No. 2454-6194 | DOI: 10.51584/IJRIAS |Volume IX Issue VIII August 2024

poses enormous obstacles. Additionally, during a patient visit, doctors write down the patient's medical encounter and send it to nurses and other medical departments for processing. These documents, which are used to process medical data for the healthcare industry, must be error-free. Often, the doctor doesn't have enough time to record every observation made while examining the patient and asking about their medical history. In hospitals, patient record errors are a major issue (Ana and Almeida, 2016). Any omission-related recording inaccuracy in a patient's medical investigation may have an impact on how well a patient responds to therapy. To create a medical information system, deep learning and natural language processing must be used. Therefore, this work to develop and analyze deep learning and natural language processing system in the context of medical information processing.

LITERATURE REVIEW

Medical Case-Based Reasoning Systems

In many fields, case-based reasoning (CBR) has proven to be an effective strategy for knowledge-based systems; however, using this technique in the medical field presents additional challenges. Case-based reasoning refers to the process of understanding and resolving new problems using examples from the past. A case-based reasoning approach tries to adapt its solutions to the present case by recalling previous cases that were similar to the current issue. The fundamental presumption is that similar issues will have comparable solutions. This presumption holds for many practical domains, albeit it is not always true. CBR involves two primary tasks: The first step is retrieval, which involves looking for or tallying up the most comparable cases. A sequential calculation is feasible if the case base is relatively small; otherwise, quicker non-sequential indexing or classification techniques (such as nearest neighbor match) should be used. Since a lot of research has been done in this area recently, it is now simple to identify powerful case-based reasoning (CBR) retrieval algorithms that are suitable for almost every application situation.

The second step, adaptation (reuse and revision), is adapting solutions from earlier, analogous circumstances to meet a more recent situation. A straightforward solution transfer is sufficient if there are no significant variations between the current situation and a comparable scenario. Sometimes only a few changes are needed, while other times the transition is a very difficult process (Benotti *et al.*, 2014). The adaptation is still entirely domain-dependent because no universal adaptation algorithms or procedures have yet been created.

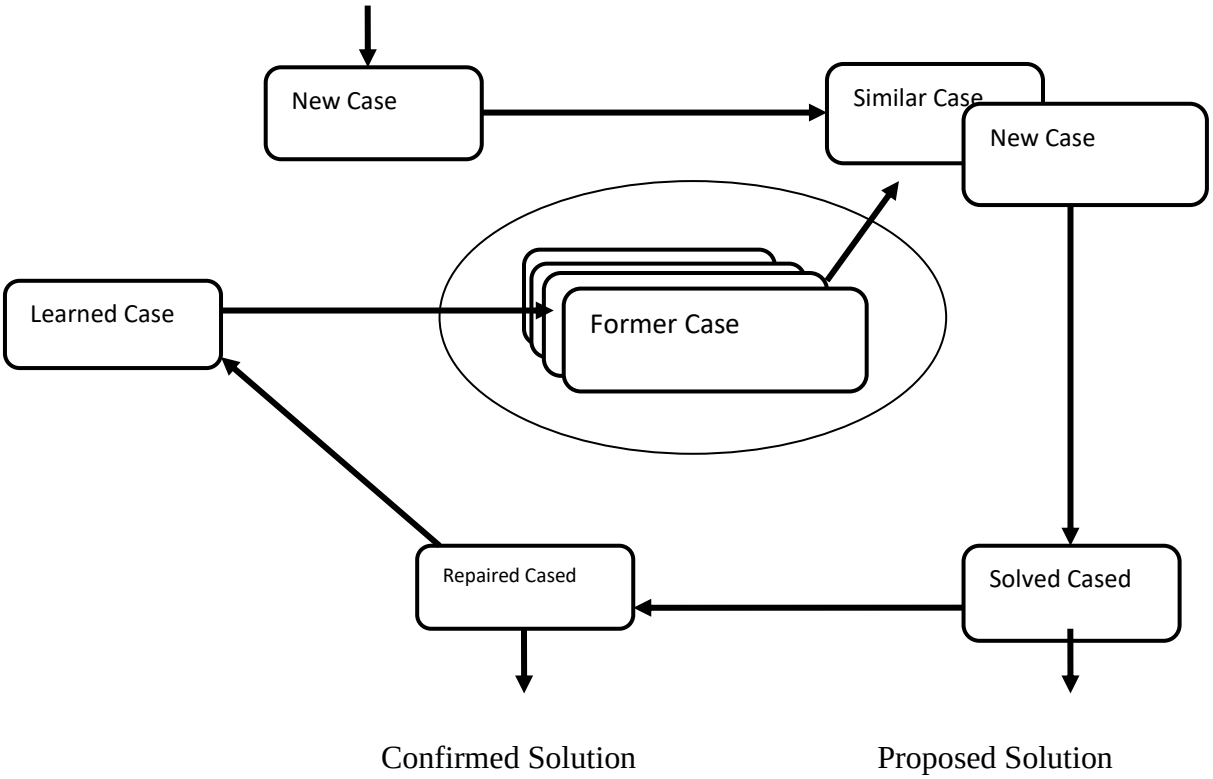

Figure 1: Cycle of Case-Based Reasoning (Alexander, 2017)

ISSN No. 2454-6194 | DOI: 10.51584/IJRIAS |Volume IX Issue VIII August 2024

Case-based reasoning (CBR) has primarily been used in medicine for diagnostic and in certain cases for therapeutic functions. Case-oriented strategies for tutoring and retrieval methods to look for similar photos are similar techniques that have been employed in other fields.

CASEY is one of the first medical expert systems to employ CBR methods (Koton, 1988). It addresses the diagnosis of heart failure.

Long-Short Term Memory Networks (LSTMs)

Recurrent neural networks (RNNs) with LSTMs can learn and remember long-term dependencies. Long-term memory retention is the default mode of operation (Avijeet, 2022). Over time, LSTMs preserve information. Due to their ability to recall prior inputs, they are helpful in time-series prediction. In LSTMs, four interacting layers connect in a chain-like structure to communicate especially. LSTMs are frequently employed for voice recognition, music creation, and drug research in addition to time-series predictions. An illustration of how LSTMs work is shown in Figure 5 below. The following are the steps for LSTMs.

1. They start by forgetting pointless details from the previous state.

2. After that, they update specific cell-state values.

3. The output of specific cell state components

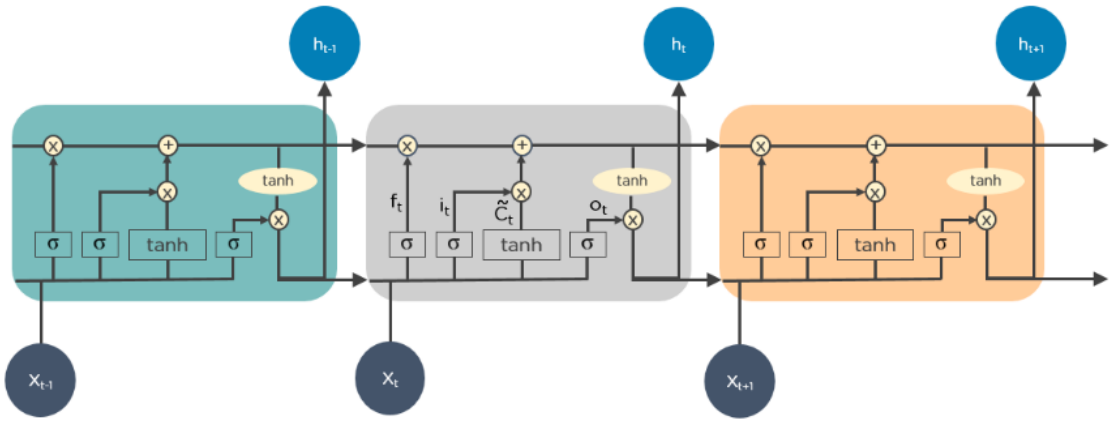

Figure 2: Diagram of how LSTMs operate (Avijeet, 2022)

METHODOLOGY

The object-oriented analysis and design methodology (OOADM) for software development. The analysis models are developed further using object-oriented design (OOD) to create implementation requirements. The deep learning and natural language processing system was developed in the context of processing medical information using the PHP-MySQL program and Java Script. We will be able to accomplish the goals mentioned by using PHP-MySQL. The user-friendly nature of PHP-MySQL makes it possible to create interfaces that can be changed programmatically. A reliable database that can ensure database integrity, database protection, and support huge databases is MYSQL. Before deciding on PHP and JavaScript, several issues were taken into account, including deep learning, online database access, data transmission via networks, database security, voice data capture, etc.

RESULTS

The Proposed System

Figure 3 illustrates the user interface (UI) and backend of the proposed NLP-driven application. The backend receives text or speech input from the user via the user interface (UI), processes it using NLP models, and then returns the results to the user by offering certain services via the UI. Backend knowledge bases are also necessary

ISSN No. 2454-6194 | DOI: 10.51584/IJRIAS |Volume IX Issue VIII August 2024

for applications that primarily rely on knowledge. Through speech, writing, and other means, the UI enables information interaction, for improving the user experience with intelligent systems and accomplishing smart medical information processing, easily accessible user interfaces are essential. Utilizing NLP techniques, particularly speech recognition and natural language comprehension, one can develop such user interfaces. The fundamental design of NLP-driven applications is depicted in Figure 3.

Speech recognition can be used in the proposed system to take free text notes, which can cut down on the amount of time medical staff spends on the labor-intensive clinical recording. Additionally, the suggested system included clinical decision support (CDS) systems, which can offer doctors recommendations for diagnoses and treatments utilizing deep learning. Heart illness, often known as cardiovascular disease, will be the focus of the CDS (CVDs). The heart illness dataset (hears that log Cleveland and Hungary final) from kaggle.com/datasets will be used for deep learning (2022). between users and intelligent systems. The system was created with the utmost consideration for security, continuity, and accessibility of health information across space and time.

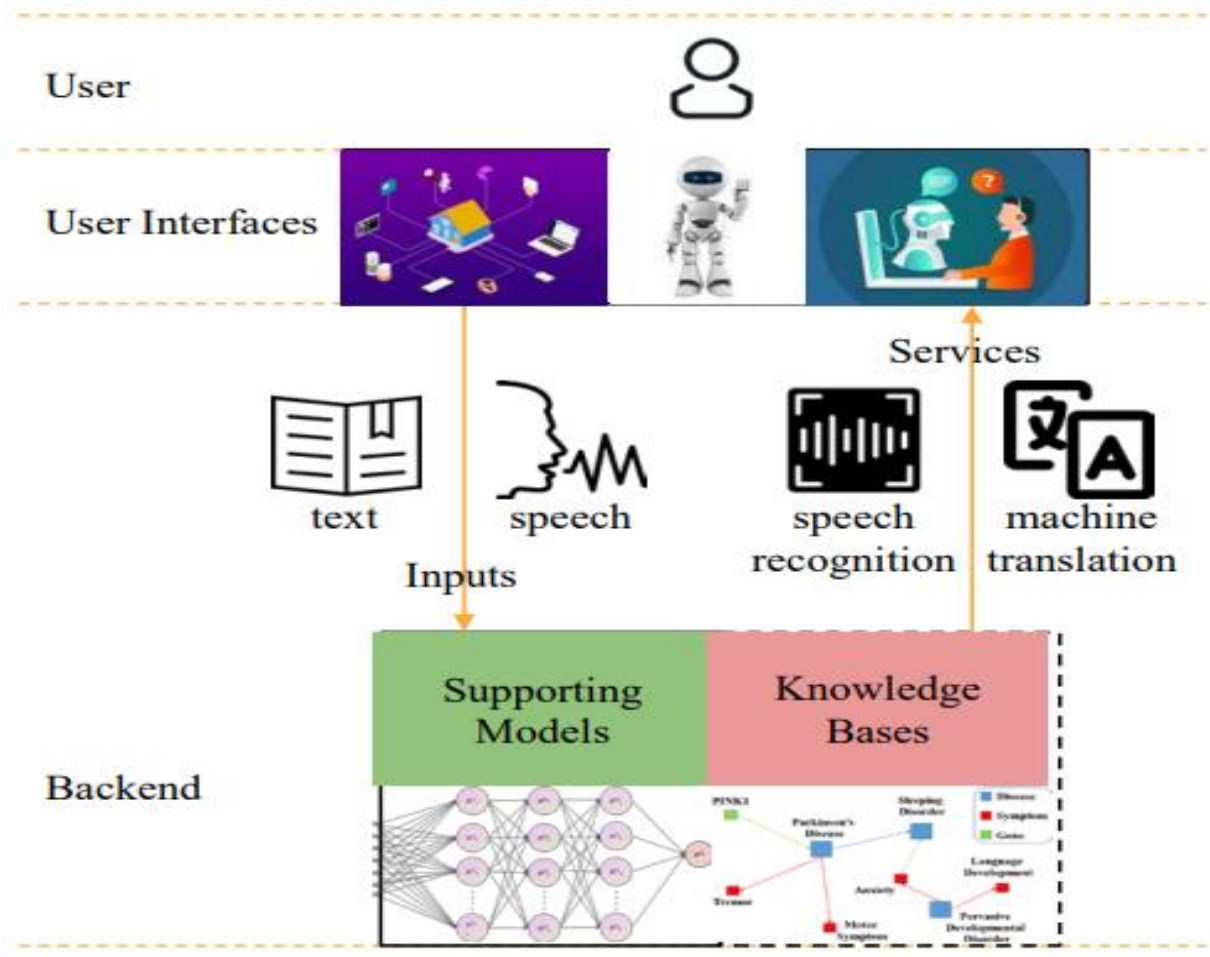

Figure 3; illustrates the fundamental design of NLP-driven apps.

Advantage of the Proposed System

The suggested system will have the following benefits:

1. Voice-to-text conversion that is automatic
2. The system would ensure real-time processing, analyzing, and accessing of data.
3. The system would include a convenient platform for communication between medical professionals and the general public.
4. The system would be knowledgeable, trustworthy, adaptive, adaptable, flexible, and agile.

HIGH-LEVEL MODEL OF THE PROPOSED SYSTEM

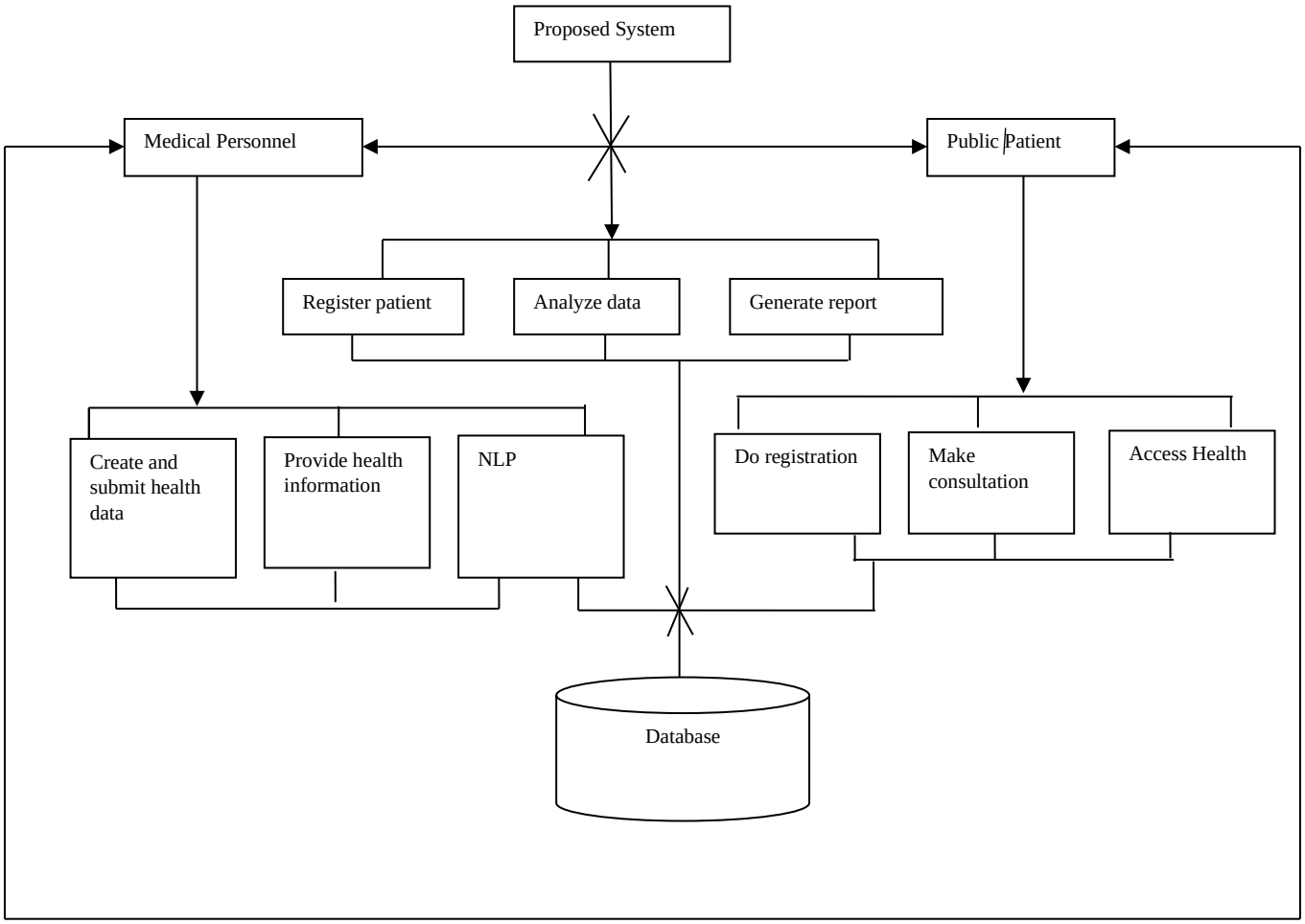

Figure 4: High-Level Model of the Proposed System

Following is a detailed discussion of the high-level model's functions;

1. The proposed system has nine processes, including the creation of profiles, data analysis, report generation, creation and submission of health data, provision of health information and consultation, registration, consultation, and access to health, with three objects to be included: medical personnel, the proposed system, and the general public/patient.

2. The user (public/patient) created a page with his or her profile, which includes medical history, after registering with the medical unit.

CHOICE OF PROGRAMMING ENVIRONMENT

Many different programming languages were taken into account when creating this software. Numerous variables, such as online database access, data transmission via networks, database security, online database retrieval, multiuser network access, online data capture, etc., were taken into account. This work made use of PHP-MySQL and JavaScript to accomplish the aforementioned goals. Additionally, PHP-MySQL allows for the creation of an interface that can be changed programmatically and is very user-friendly. A reliable database that can ensure database integrity, database protection, and support huge databases is MYSQL.

Justification of Language

Java Script and the PHP-MySQL programming language were selected because they offer the benefits of simple development, flexibility, the capacity to offer the developer potential clues, and the ability to construct a graphical user interface.

ISSN No. 2454-6194 | DOI: 10.51584/IJRIAS |Volume IX Issue VIII August 2024

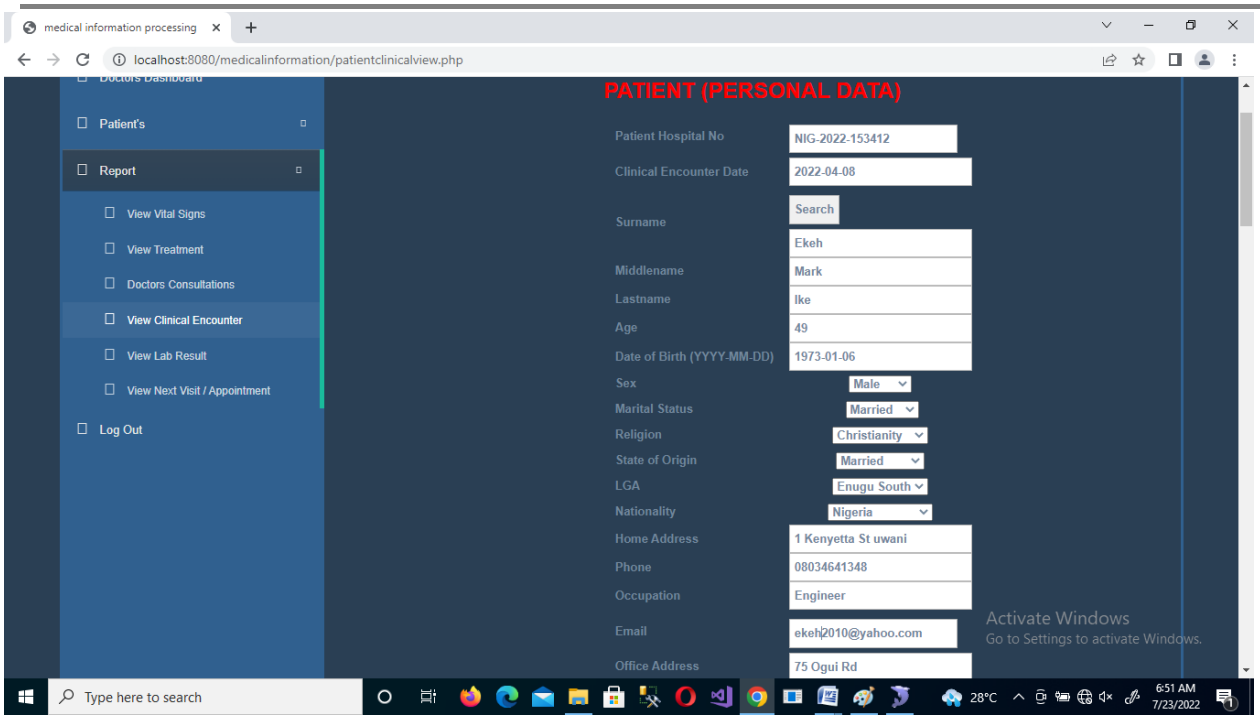

Figure 5: Patient Clinical Encounter / Voice data capturing page 1 (patient personal data)

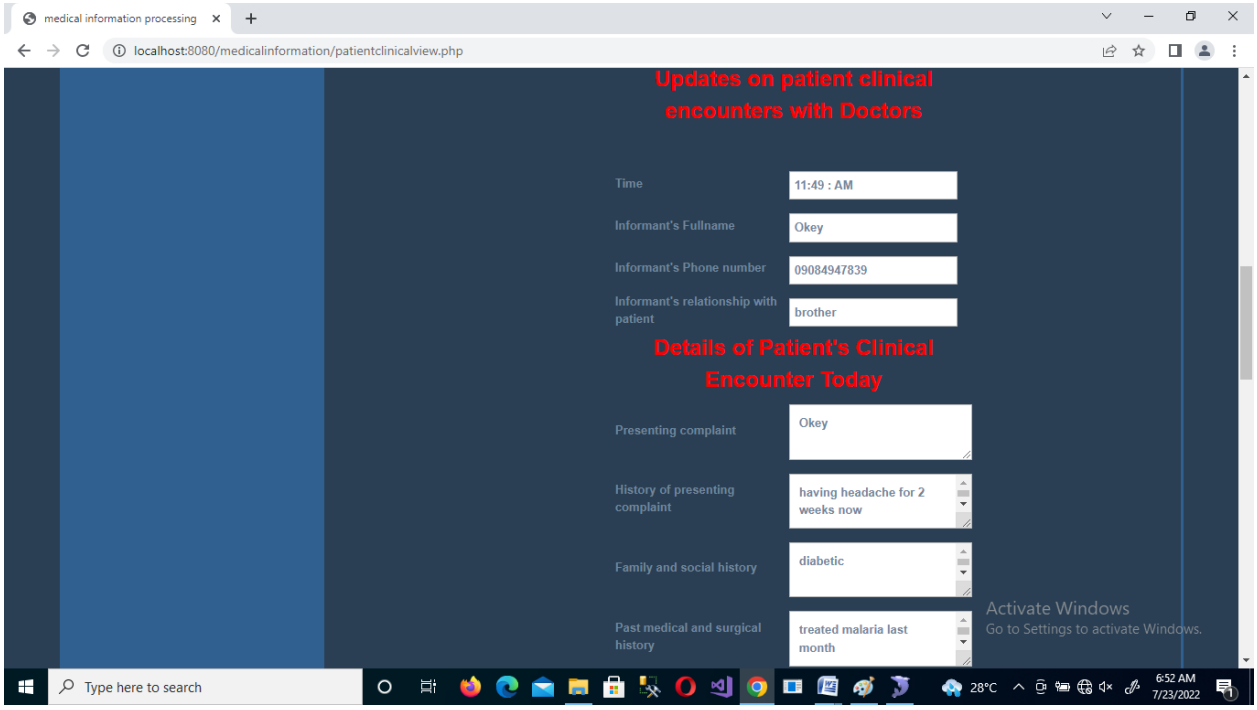

Figure 6: Patient Clinical Encounter / Voice data capturing page 3 (patient clinical data)

CONCLUSION

Therefore, the research project devised a sophisticated system capable of converting speech into text during medical interactions using advanced natural language processing (NLP). This technology facilitates the seamless transcription of clinicians' verbal observations made during patient examinations into written records swiftly and accurately. Consequently, patient medical information is documented promptly and without delay. Deep learning and NLP play crucial roles in the analysis of medical data, empowering healthcare providers to deliver timely, high-quality care to patients. Medical datasets contain extensive historical data on health conditions, diagnostic criteria, and treatment outcomes, enabling clinicians to glean valuable insights from vast collections of records. Such learning processes cannot be effectively undertaken manually; however, the application of deep learning algorithms enhances efficiency, accuracy, and practical utility in clinical settings.

ISSN No. 2454-6194 | DOI: 10.51584/IJRIAS |Volume IX Issue VIII August 2024

REFERENCES

1. Alexander, W. F. (2017). Improving Clinical Decision Making with Natural Language Processing and Machine Learning. Computer Science and Engineering at the Massachusetts Institute of Technology. 13: 45
2. Ana, C., and Almeida, B. (2016). A systematic literature review on Natural Language Processing in Business Process Identification and Modeling. Computer Science. 24: 45 - 35
3. Avijeet, B. (2022). Top 10 Deep Learning Algorithms You Should Know in 2022. https://www.simplilearn.com/tutorials/deep-learning-tutorial/deep-learning- algorithm
4. Avijeet, B. (2022). Top 10 Deep Learning Algorithms You Should Know in 2022. https://www.simplilearn.com/tutorials/deep-learning-tutorial/deep-learning- algorithm
5. Benotti, L., Martínez, M., and Schapachnik, F. (2014). Engaging high school students using chatbots. In: Proceedings of the 2014 Conference on Innovation & Technology in Computer Science Education, pp. 63-68.
6. Tang, P.C., Faf champs, D. and Short liffe, E.H. (2014). Traditional medical records as a source of clinical data in the outpatient setting. Proceedings of the Annual Symposium on Computer Application in Medical Care. American Medical Informatics Association. pp 575
7. Tian, S., Yang, W., Grange, J. M. L., Wang, P., Huang, W. and Ye, Z. (2019). Smart healthcare: Making medical care more intelligent. Global Health Journal, 3(3), 62–65.