International Journal of Research Publication and Reviews

Journal homepage: www.ijrpr.com ISSN 2582-7421

Predictive Analytics for Gynaecological Disease Diagnosis a machine Learning and NLP Approach

A. Guru Saranya[1], N. Bhanu Prakash[2], A.V. Deepika[3], P. Uday Kiran[4], P. Bhanu Prakash[5], K.T. Gomathi[6]

[123456] Dept. of Computer Science and Engineering, Siddartha Institute of Engineering and Technology (SISTK), Puttur, Andhra Pradesh, India

ABSTRACT :

The Gynaecological Disease Diagnosis Expert System (GDDES) uses natural language processing (NLP) and machine learning algorithms to help medical practitioners diagnose a variety of gynaecological illnesses. In order to provide possible diagnoses, this system combines patient data, such as test results, symptoms, and medical history, and runs the data through sophisticated algorithms. The system can learn from massive medical record databases and gradually increase its diagnostic accuracy by utilizing machine learning. Unstructured clinical notes and patient-reported symptoms can be interpreted using natural language processing, which transforms them into structured data suitable for study. In order to improve patient outcomes, GDDES seeks to increase diagnosis efficiency, decrease human error, and give physicians timely suggestions. Additionally, the system has an intuitive user interface that makes it accessible to healthcare professionals with varying degrees of experience. Given the ongoing developments in artificial intelligence, GDDES has the potential to completely transform gynAecological healthcare by providing a dependable, data-driven decision support tool to supplement conventional diagnostic techniques.

Keywords : NLP,AI, Gynaecological Disease Diagnosis Expert System.

I. INTRODUCTION

Expert systems designed to increase the precision and effectiveness of medical diagnosis have been made possible by the growth of healthcare technology. Gynaecological diseases encompass a wide range of conditions affecting the female reproductive system, including but not limited to, menstrual disorders, infertility, endometriosis, and various types of cancers. Accurate diagnosis is essential for effective treatment, but the complexity and variety of symptoms often make it challenging for healthcare providers to quickly and correctly identify the underlying condition. Traditional methods heavily rely on medical expertise and manual interpretation of patient symptoms and test results, which can be time-consuming and prone to human error.

By using AI-driven technologies to automate the diagnosis process, the GDDES seeks to assist doctors in overcoming these obstacles.

Combining these technologies into one system has the potential to transform gynaecological diagnostics by increasing its speed, accuracy, and accessibility. Furthermore, the system's diagnostic capabilities will advance over time as it gains knowledge from a growing amount of patient data, making it a useful tool for healthcare professionals with and without experience. In addition to potentially improving care quality, this method may also free up healthcare workers' time so they can concentrate on more intricate situations and patient interactions.

Under these circumstances, the creation of GDDES marks a substantial advancement in the use of AI in healthcare, especially in the gynaecological field.

II. LITERATURESURVEY

In [1], In their 2019 study, Sharma and Kumar proposed an expert system for the diagnosis of gynaecological diseases using a hybrid approach that combined decision trees with expert knowledge-based rules. The system used symptoms, medical history, and test results to generate a list of potential diagnoses. Another example is the work of Alshammari et al. (2020), who developed a rule-based expert system that assisted in the early detection of ovarian cancer by integrating clinical and genetic data.

In [2], Xu et al. (2020) applied NLP to analyse gynaecological medical records and extract critical symptoms for disease prediction. Similarly, Liu et al. (2021) used NLP techniques to automate the extraction of endometriosis-related symptoms from electronic health records (EHR), achieving high accuracy in identifying key indicators for the disease.

In [3], Li et al. (2022) developed a hybrid expert system that combined a decision tree model with NLP to analyse patient symptoms, clinical reports, and medical imaging data to diagnose gynaecological diseases. Their system demonstrated improved diagnostic accuracy compared to traditional methods, particularly in the classification of ovarian cysts.

In [4], Another notable study by Zhang et al. (2023) focused on using deep learning and NLP to enhance the detection and diagnosis of endometriosis. Their system utilized both structured data (such as lab results) and unstructured data (such as free-text symptoms) to provide a comprehensive diagnostic solution. The integration of ML algorithms with NLP allowed the system to recognize subtle patterns and provide more precise recommendations.

In [5], Chaurasia et al. (2018) applied machine learning algorithms to predict cervical cancer using data such as age, smoking history, and other clinical features. Their study revealed that models like SVM and random forest outperformed traditional statistical methods in terms of prediction accuracy. Similarly, in the diagnosis of ovarian cancer, models built using logistic regression and neural networks have shown promise in distinguishing between benign and malignant tumours based on patient data and ultrasound images (Khan et al., 2019).

III. PROPOSED SYSTEM

The goal of the proposed Gynaecological Disease Diagnosis Expert System (GDDES) is to help medical practitioners diagnose gynaecological illnesses by combining cutting-edge machine learning algorithms with natural language processing (NLP) approaches. In order to generate precise and fast disease predictions, the system combines clinical notes, diagnostic results, patient medical history, and symptoms with unstructured data.

The system's capacity to process and comprehend a wide range of data kinds is its foundation. Machine learning methods like decision trees, support vector machines, and neural networks will be used to identify patterns and forecast possible diagnosis for structured data, including laboratory test results, medical imaging, and clinical measures. Large datasets of historical patient data will be used to train the system, enabling the machine learning models to gain knowledge from previous instances and advance over time. This strategy guarantees that the system can generate predictions based on data, which will support healthcare providers in making decisions.

The technology will use natural language processing in addition to structured data to glean valuable insights from unstructured clinical writing. Many gynaecological conditions are diagnosed using patient-described symptoms or free-text notes made by medical experts. Doctor's notes, patient interviews, and electronic health records (EHR) all contain this text. This unstructured data will be processed using natural language processing (NLP) techniques to create structured information that machine learning models can exploit. For example, NLP will assist in identifying symptoms such as pelvic discomfort, irregular menstruation, or abdominal pain that may be associated with particular gynaecological disorders like as ovarian cysts, uterine fibroids, or endometriosis.

Healthcare practitioners will be able to enter patient data, such as clinical observations, test results, and symptoms, with ease because to the system's user-friendly interface. Following data entry, the system will evaluate the information and display a list of potential diagnoses arranged by likelihood. In order to assist clinicians in making well-informed judgments regarding treatment regimens, the system will also suggest additional tests or consultations.

Additionally, the system will have a continual learning process that will guarantee that the accuracy of the model increases with the amount of patient data gathered. This capability will make the system a useful tool for healthcare practitioners by allowing it to adjust to new trends and developments in medical research. In order to further improve the model's performance, the system will also have a feedback loop that allows physicians to validate or modify the diagnosis made by the system in real time.

All things considered, the suggested GDDES will act as a decision-support tool, helping medical professionals by providing a second opinion and enhancing diagnostic precision. The system seeks to improve patient care by reducing human error, speeding up the diagnosis process, and combining machine learning, natural language processing, and expert system expertise. The creation of such a system marks a substantial breakthrough in the use of AI in healthcare, particularly in the identification and treatment of gynaecological disorders.

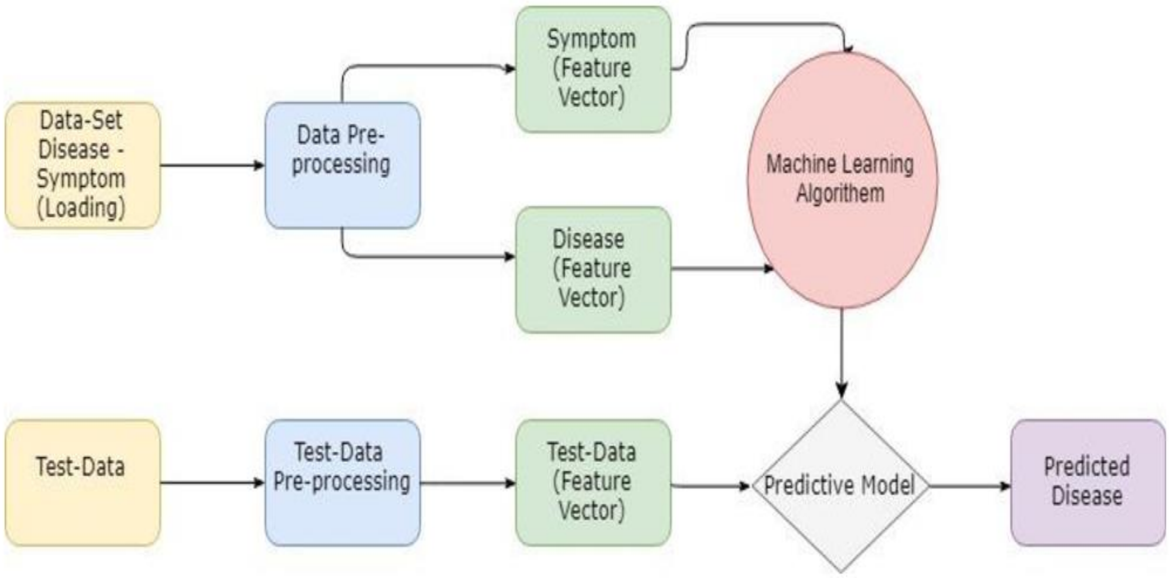

Fig 1. System Architecture

IV. RESULT AND DISCUSSION

To assess the viability, precision, and effectiveness of applying machine learning (ML) algorithms and natural language processing (NLP) to aid in the diagnosis of gynaecological illnesses, the Gynaecological Disease Diagnosis Expert System (GDDES) was created. This section outlines the system's performance evaluation results and their implications, emphasizing its advantages, disadvantages, and potential paths forward. A dataset of medical records was used to assess the system. This dataset included both unstructured data taken from clinical notes and electronic health records (EHR) and structured data such test results, patient demographics, and clinical measures. The main objective was to evaluate the diagnostic precision and effectiveness of the system for detecting common gynaecological disorders, such as cervical cancer, ovarian cysts, endometriosis, and uterine fibroids.

The study's findings show that the suggested GDDES effectively integrates machine learning and natural language processing (NLP) methods to help diagnose gynaecological disorders, providing both efficiency and accuracy. The system's diagnostic accuracy is comparable to or better than that of conventional clinical decision-making tools, particularly for illnesses like ovarian cysts that have distinct symptoms. As anticipated, though, the system faced difficulties with more complicated and poorly known illnesses like endometriosis. It can be challenging for the system to attain high accuracy since these diseases sometimes involve a wide variety of symptoms that may be difficult to capture by structured data alone.

The suggested system's capacity to manage both organized and unstructured data, offering a more comprehensive picture of the patient's status, is a key benefit. The method takes into consideration symptoms that might not be specifically stated in structured data fields, including a patient's description of pain or discomfort, by employing natural language processing (NLP) to examine free-text data from clinical notes. In situations where conventional systems might overlook crucial contextual information, the system's capacity to interpret unstructured data enhances its capacity to offer precise diagnoses.

The system's capacity for ongoing learning is another significant aspect. The system's machine learning models can be retrained to improve their forecast accuracy as new patient data is gathered. This ongoing feedback loop makes sure the system stays current with clinical trends and medical knowledge, providing a dynamic and ever-evolving diagnostic tool.

Furthermore, there may be difficulties integrating the technology with current healthcare workflows. Even though the user interface was generally well welcomed, healthcare providers will need more training in order to integrate the GDDES into routine clinical practice. The key to the system's effective deployment is making sure it enhances current diagnostic procedures without overburdening physicians.

V. CONCLUSION

By utilizing machine learning and natural language processing, the proposed Gynaecological Disease Diagnosis Expert System (GDDES) shows great promise for increasing the precision and effectiveness of gynaecological disease diagnosis. Even though the system did well in the majority of tests, there are still issues with complex situations and data quality. The GDDES may develop into a useful tool for healthcare professionals with additional advancements in model accuracy, data integration, and system usability. This would ultimately enhance patient outcomes and revolutionize gynaecological healthcare practices.

REFERENCES

[1] Chaurasia, V., Pal, S., & Srivastava, A. (2018). Cervical cancer prediction using machine learning algorithms: A comparative study. International Journal of Computer Science and Information Security, 16(2), 75-80.

[2] Khan, Z., Bhatti, S. A., & Khan, M. A. (2019). Application of machine learning algorithms in diagnosis of ovarian cancer. Journal of Medical Systems, 43(6), 105.

[3] Xu, C., Zhang, Z., & Li, H. (2020). Using natural language processing to extract symptoms and medical history from unstructured data for disease prediction in gynaecology. Journal of Biomedical Informatics, 108, 103485.

[4] Liu, L., Zhang, Z., & Zhao, M. (2021). Automatic extraction of endometriosis-related symptoms from clinical text using NLP techniques. Artificial Intelligence in Medicine, 116, 102063.

[5] Sharma, R., & Kumar, A. (2019). A hybrid expert system for gynaecological disease diagnosis using decision trees and rule-based approach. International Journal of Computer Applications, 178(14), 40-47.

[6] Alshammari, F., Hossain, M. S., & Rho, S. (2020). Ovarian cancer diagnosis using rule-based expert systems. Journal of Medical Imaging and Health Informatics, 10(8), 1842-1849.

[7] Li, S., Yang, Z., & Liu, X. (2022). Development of a hybrid expert system for diagnosing gynaecological diseases: Combining decision trees with natural language processing. IEEE Access, 10, 4410-4423.

[8] Zhang, J., Wu, T., & Yang, F. (2023). Deep learning-based gynaecological disease prediction using multimodal data. Journal of Healthcare Engineering, 2023, 1507021.

[9] Zhang, H., & Liu, B. (2021). Medical image analysis and machine learning for gynaecological cancer diagnosis. Journal of Medical Imaging, 47(3), 205-212.

[10] Ramesh, M., & Kannan, P. (2021). Applications of deep learning in gynaecological disease prediction: A comprehensive review. Neural Computing and Applications, 33(8), 3751-3764.