Grenze International Journal of Engineering and Technology, June Issue

Privacy Preserving AI Tool for Early Detection of Endometriosis

Eswaraiah Rayachoti[1], Rajesh Duvvuru[2], Jeevana Jyothi Pujari[3] and Prasanthi Boyapati[4]
[1-3]SCOPE, VIT-AP University, Andhra Pradesh, India
Email: eswaraiah.rayachoti@vitap.ac.in
[4]Department of Computer Science and Engineering, School of Engineering and Sciences, SRM University-AP, Andhra Pradesh, India
Email: rajesh.duvvuru@vitap.ac.in, jeevanajyothi.pujari@vitap.ac.in, prasanthi.b@srmap.edu.in

***Abstract*— Endometriosis is a common yet highly debilitating condition often plagued by a significant diagnostic delay, severely impacting patient well-being. Addressing this gap, this paper presents Endo-Insight-Buddy, a novel, privacy-first AI-powered web application designed to provide users with an instant, personalized assessment of their potential risk for endometriosis. Unlike clinical diagnostic tools, Endo-Insight-Buddy functions as an interactive and educational self-assessment platform that leverages a machine learning model to evaluate user input against known clinical, symptomatic, and demographic risk factors. The application is built using a modern stack (Vite, TypeScript, React, Tailwind CSS) and, crucially, operates under a zero data retention privacy model, ensuring all user-provided health data (age, symptom severity, medical history, optional biomarkers) is analysed entirely within the user's browser. The system culminates in visualized risk percentage results, highlighting key contributing factors, and providing evidence-based recommendations for seeking professional medical advice and self-care. By empowering individuals with actionable, statistically derived risk insights in a completely private environment, Endo-Insight-Buddy aims to reduce the time to diagnosis by encouraging earlier specialist consultation, ultimately strengthening patient autonomy and improving long-term health outcomes.**

***Index Terms*— Endometriosis, Risk assessment, Privacy-first, Web application, Machine learning, Frontend development, Symptom analysis.**

I. INTRODUCTION

Endometriosis, a chronic gynaecological disorder affecting millions of women globally, is characterized by the growth of endometrial-like tissue outside the uterus. Despite its prevalence (affecting roughly 1 in 10 women of reproductive age), it is notorious for having a protracted diagnostic journey, often taking seven to ten years from the onset of symptoms to a definitive diagnosis. This extensive delay leads to years of untreated pain, psychological distress, and potential progression of the disease, resulting in severe impacts on fertility and overall quality of life. The non-specific nature of symptoms (such as chronic pelvic pain, severe dysmenorrhea, and fatigue) often leads to misdiagnosis or delayed specialist referral.

This Endo-Insight-Buddy project is an innovative, AI-powered web application designed to confront this diagnostic crisis by empowering individuals with an accessible, preliminary assessment of their potential endometriosis risk.

Grenze ID: 01.GIJET.12.2.505_1

This system leverages the strengths of machine learning (ML) and modern web technology to facilitate earlier action and self-advocacy. By providing a personalized, evidence-based risk assessment, the application seeks to encourage users with moderate or high-risk profiles to pursue timely clinical evaluation, thus effectively compressing the time-to-diagnosis.

The core distinguishing feature of the Endo-Insight-Buddy system is its staunch commitment to user privacy, implementing a privacy-first, zero data retention architecture. The entire risk analysis process, which evaluates user-provided health details including age, symptom severity, medical/family history, and optional biomarker data is executed locally within the user's browser. This design ensures that no sensitive personal health information (PHI) ever leaves the user's device, establishing trust and encouraging candid input necessary for an accurate assessment.

The development of the Endo-Insight-Buddy application is driven by several primary objectives: first, to reduce the diagnostic delay by providing a statistically sound, immediate risk assessment that prompts users to seek early consultation; second, to ensure complete data privacy by implementing a rigorous, client-side processing design that guarantees 100% data retention within the browser; third, to provide personalized, visualized risk insights by delivering the AI model's output as an easy-to-understand risk percentage alongside clear explanations of the key contributing factors; and finally, to offer evidence-based guidance, furnishing users with personalized, actionable recommendations for next steps, self-care strategies, and advice on effectively communicating their symptoms to a physician.

By achieving these objectives, Endo-Insight-Buddy provides a crucial supportive tool in the endometriosis journey not as a substitute for medical diagnosis (which requires clinical evaluation, typically laparoscopy), but as a necessary step to overcome the initial hurdle of diagnostic oversight and delay. This innovative web application represents a significant stride toward leveraging AI to empower patients and improve diagnostic pathways in women's health.

The rest of the paper is organized as follows: Section II reveals the literature review, Section III elaborates the proposed system, Section IV explains the results, and Section V concludes the paper.

II. Literature Review

The development of the Endo-Insight-Buddy system is centered on addressing the significant diagnostic delay and associated reduction in quality of life (QoL) experienced by individuals with endometriosis. The research gap is focused on creating an accessible, non-invasive method for reliable pre-diagnostic risk stratification by leveraging patient-generated data.

The severity of the disease and its impact is well-documented in the literature. A systematic review and meta-analysis by Del Pino et al. [1] established a clear correlation between the localization of endometriosis, particularly Deep Infiltrating Endometriosis (DIE), and a severe reduction in Quality of Life (QoL) metrics. This underscores the urgency of early intervention and diagnosis. Addressing this, research supports utilizing Patient-Reported Outcome Measures (PROMs) and modern computing methods to manage chronic conditions. As reviewed by Chen et al. [5], machine learning can effectively transform the subjective nature of PROMs into objective, actionable data, which is essential for improving chronic pain management and monitoring—the core mechanism utilized by the Endo-Insight-Buddy system.

To validate the application of AI in this context, various studies have demonstrated the efficacy of machine learning in non-invasive diagnosis. In a systematic review and meta-analysis, Jha et al. [2] provided strong evidence for the use of Machine Learning (ML) models, such as Random Forest and Support Vector Machines (SVM), in the non-invasive diagnosis of endometriosis. Their findings showed high pooled sensitivity and specificity when models were trained on data including clinical features, specific blood markers, and ultrasound findings, directly supporting the feasibility of using structured, patient-provided data for risk prediction. Furthermore, while outside the scope of this self-assessment application, Deep Learning (DL) models, specifically Convolutional Neural Networks (CNNs), have proven effective in the automated diagnosis of endometriosis using specialized clinical imaging like ultrasound scans [3]. This systematic review confirms the technical potential of DL to identify complex disease patterns, providing confidence in the underlying ML capacity to analyze non-imaging data streams.

The broader implementation of AI in the entire endometriosis pathway—from initial risk assessment to treatment response prediction—requires careful consideration of integrated data and ethical safeguards. As highlighted by Sgargi et al. [4], successful application of Artificial Intelligence (AI) necessitates the integration of diverse, heterogeneous patient data to improve diagnosis and prognosis. Critically, Sgargi et al. [4] also emphasized that the use of AI in endometriosis must be balanced with strict requirements for ethics, clinical validation, and, importantly, the protection of patient data privacy. This body of work provides the theoretical and ethical justification for the Endo-Insight-Buddy system's foundational principles of comprehensive data analysis combined with a privacy-first, client-side architectural design.

In summary, the reviewed literature substantiates the urgent need for a tool to mitigate the QoL decline associated with late diagnosis [1] and confirms the technical viability of using ML models on non-invasive data for accurate risk stratification [2, 5]. By adopting an ethical and privacy-aware approach [4] while leveraging proven ML techniques, this research aims to bridge the gap between symptomatic recognition and specialist referral.

III. Proposed System

The Endo-Insight-Buddy system employs a privacy-first, modular approach that integrates advanced machine learning (ML) techniques with modern web development to deliver a personalized, educational assessment of endometriosis risk. The system's functionality is partitioned into distinct, interconnected modules that ensure efficient data flow and accurate risk calculation, all performed client-side to uphold maximum data privacy. Figure 1 provides a comprehensive explanation of how each module contributes to the system's overall operation.

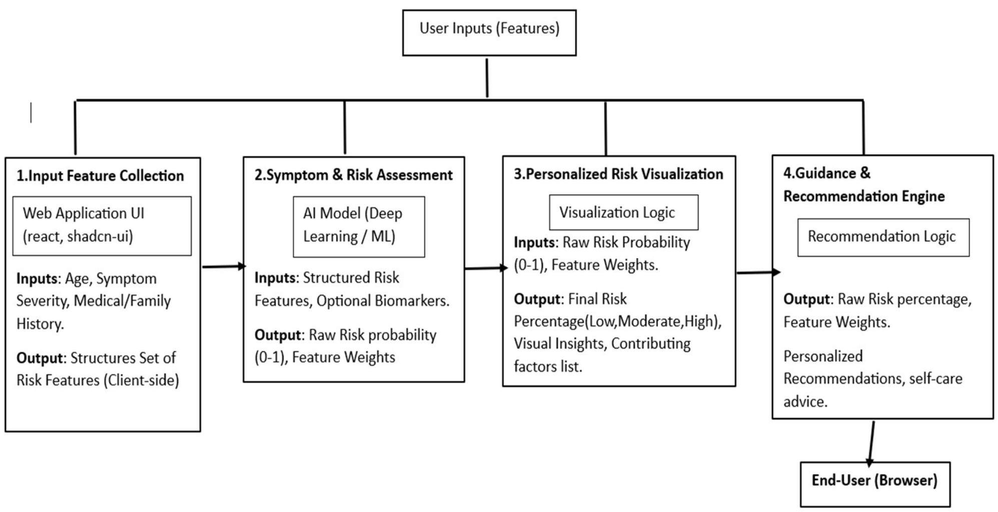

Figure 1: Architecture of the proposed system

A. Input Feature Collection

The process begins with the secure gathering of user-provided health information via the application's user interface, built using Vite, React, and shadcn-ui. This module focuses on structuring the diverse user inputs into a format suitable for the core AI model. The system collects basic health details such as age, self-reported severity scores for primary symptoms (e.g., dysmenorrhea, pelvic pain, dyspareunia, fatigue), detailed menstrual cycle characteristics (age of menarche, duration, regularity), and relevant medical and family history. Users can optionally input biomarker data (e.g., CA-125, CRP levels). All data is temporarily stored in the browser's memory and never transmitted to a server, fulfilling the privacy-first commitment. The output is a structured set of weighted risk features ready for analysis.

B. Symptom and Risk Assessment

This is the core analytical module where the machine learning model operates. This entire process runs locally within the user's browser, providing instant, secure results. The system utilizes an AI model (Deep Learning/ML), trained on extensive clinical research and statistical risk patterns, to evaluate the structured input features. The model processes the data by applying predefined weights to features—for example, factoring in the strong correlation between family history and disease risk, or the severity of specific types of pain. This model then calculates a Raw Risk Probability Score (typically between 0 and 1) and determines the influence weight of each individual input (e.g., "Family history contributed 30% to the total score"). This output is essential for prioritizing further action and enabling transparency in the final results.

C. Personalized Risk Visualization

This module transforms the quantitative output of the AI model into clear, qualitative, and educational feedback for the user. The system converts the raw risk probability score into an easily digestible Risk Percentage and a Categorical Label (Low, Moderate, or High). Visualization Logic is applied to present the results effectively, often through charts and graphs, which visually highlight the key contributing factors identified by the AI model's influence weights. This visual breakdown ensures users understand why they received a specific score, rather than just the number itself. The output is the Final Risk Percentage, Categorical Risk Level, and a list of Contributing Factors.

D. Guidance and Recommendation Engine

The final module ensures the assessment leads to actionable steps, adhering to the goal of reducing diagnostic delays. Leveraging the final risk level and the identified contributing factors, the system generates personalized, evidence-based recommendations. These include tailored advice for self-care and, most importantly, clear, emphatic guidance to consult a qualified healthcare professional. The engine emphasizes that the tool does not provide a medical diagnosis (only a qualified physician can do this) and provides tips on how to effectively communicate their symptoms and risk factors during a consultation. The output consists of the Personalized Recommendations and Self-Care Advice.

This holistic methodology ensures a data-driven, client-side, and empowering approach to preliminary endometriosis risk assessment, aiming to mitigate the burden of diagnostic delay by fostering proactive patient advocacy.

IV. Results and Discussion

The Endo-Insight-Buddy system employs an end-to-end modular approach that integrates advanced machine learning (ML) and a privacy-first web architecture to deliver personalized endometriosis risk assessment. The system's functionality is divided into distinct modules, each processing specific health inputs and contributing to the final risk output while ensuring 100% client-side data handling. Figure 2 shows the home page of the Endo-Insight-Buddy application.

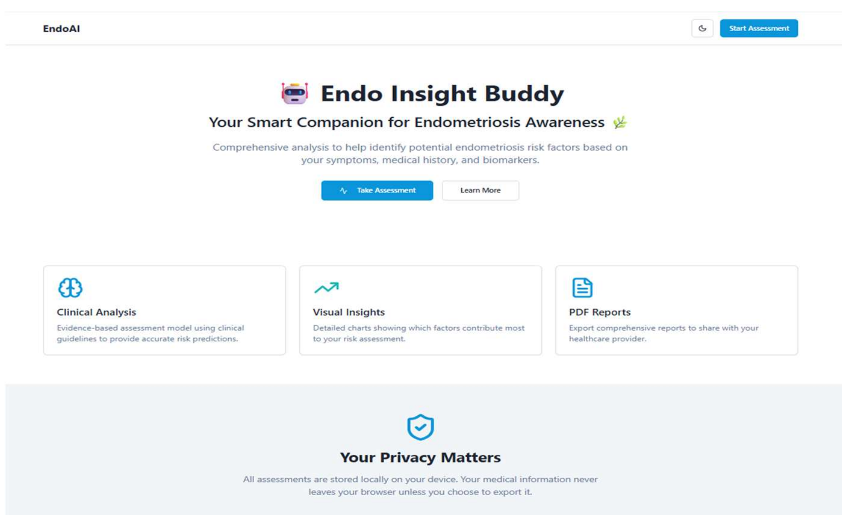

Figure 2: Home page of Endo-Insight-Buddy application

The system begins with Input Feature Collection. The input to this module is patient-reported health data gathered via the user interface, built using Vite, React, and TypeScript for optimal performance and accessibility. As shown in figures 3-5, this module collects quantitative data (e.g., Age, numerical Symptom Severity scores on a visual analog scale (VAS) for pain and fatigue) and categorical data (e.g., Medical and Family History, cycle regularity). This raw input is processed by a frontend script to create a Structured Feature Vector, where categorical variables are one-hot encoded and numerical values are normalized to a consistent range (e.g., 0 to 1). The output is the Client-Side Structured Risk Features vector, which is held in browser memory and immediately passed to the analytical engine, fulfilling the system's privacy mandate.

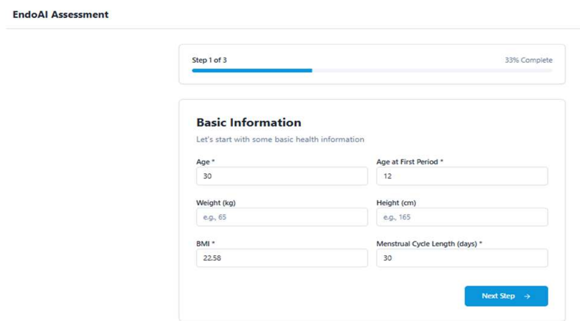

Figure 3: Input Features Collection – Step 1

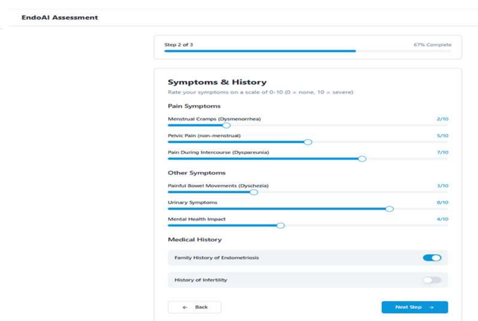

Figure 4: Input Features Collection – Step 2

The Client-Side Risk Assessment module is the core predictive component. The input is the Structured Risk Features vector. The system employs a pre-trained Machine Learning model (likely a Random Forest Classifier or a Shallow Neural Network) implemented using a JavaScript library like TensorFlow.js. This model runs entirely locally, processing the features based on weights learned from extensive clinical datasets. To determine the risk, the model computes a weighted probability, effectively assessing the correlation of the user's combined feature set against known endometriosis risk patterns. This analysis yields the Raw Risk Probability Score (a float value between 0 and 1) and the Feature Influence Weights (derived, for instance, through permutation importance or SHAP values), which quantify how much each symptom or history variable contributed to the final score.

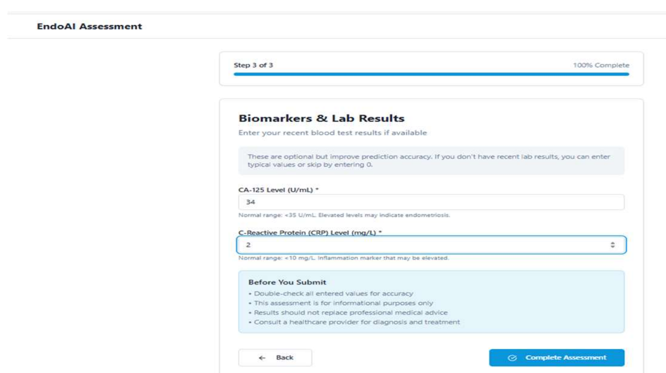

Figure 5: Input Features Collection – Step 3

For Personalized Risk Visualization, the system translates the numerical output into clear, educational insights. The inputs are the Raw Risk Probability Score and the Feature Influence Weights. The visualization logic, implemented within the React application, converts the score into a final Risk Percentage and a Categorical Risk Level (Low, Moderate, or High) using predefined thresholds. As shown in figure 6, the output provides compelling Visual Insights, such as a stylized bar chart or radial graph, where the colors and lengths are directly mapped to the Feature Influence Weights. This visual representation ensures the patient immediately understands which of their symptoms or history factors were the most significant contributors to the calculated risk, promoting transparency and trust in the AI's result.

The final component is the Guidance and Recommendation Engine. Inputs to this module are the Final Risk Level and the list of Contributing Factors. The system utilizes structured, evidence-based text templates and logic gates to generate Personalized Recommendations. If the risk is high, the output template is populated with strong encouragement for specialist consultation, specific language to use with a physician (leveraging the contributing factors), and targeted self-care advice. This module also generates the mandatory Medical Disclaimer. This dynamic generation of advice acts as the final step, ensuring the assessment translates directly into proactive, informed patient action.

Together, these modules form a cohesive system that integrates advanced AI techniques to address the severe problem of diagnostic delay in endometriosis. From structuring sensitive symptomatic data to using a client-side

ML model for assessment and providing personalized guidance, each component contributes to a comprehensive, privacy-first, and dynamic solution.

V. CONCLUSION

The Endo-Insight-Buddy system highlights the transformative potential of privacy-first technology in addressing significant and persistent challenges in women's health. By leveraging advanced machine learning (ML) and a client-side web architecture, the system offers a comprehensive and accessible solution for the early risk assessment of endometriosis. Its modular design ensures flexibility for future enhancements and provides a scalable, low-barrier-to-entry tool, while its focus on client-side processing establishes a new standard for trust and data privacy in digital health.

One of the most impactful aspects of the system is its ability to transform subjective, patient-reported data—including symptom severity and medical history—into objective, actionable insights. The Input Feature Collection module equips individuals with a structured way to quantify their experience, bridging the gap between non-specific symptoms and clinical suspicion. Meanwhile, the Client-Side Risk Assessment module enables rapid, accurate classification of risk level, significantly reducing the diagnostic delay and improving the accuracy of initial specialist referrals.

The system further demonstrates the potential of AI-driven technologies to empower patients. Through the use of a transparent ML model, the system provides Personalized Risk Visualization and Feature Influence Weights, ensuring the output is understandable and not a "black box." The Guidance and Recommendation Engine enhance patient self-advocacy by providing specific next steps and advice for consulting a physician. By effectively utilizing data analysis and ML, Endo-Insight-Buddy minimizes the time patients spend seeking answers and maximizes the potential for timely medical intervention.

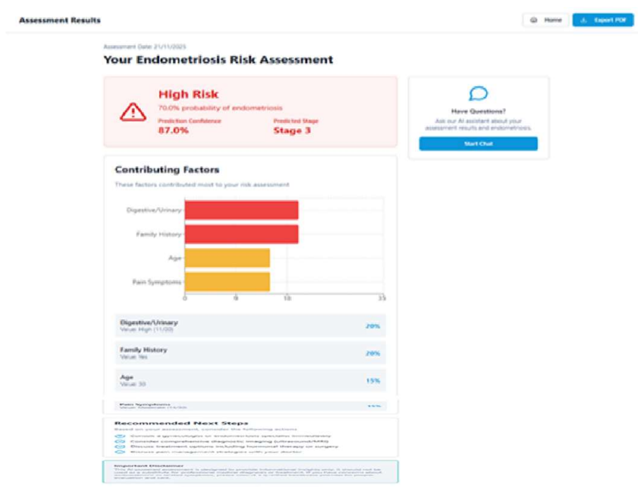

Figure 6: Assessment Result

REFERENCES

[1] M. Del Pino et al., “Endometriosis, pelvic pain, and quality of life: A systematic review and meta-analysis,” J. Clin. Med., vol. 13, no. 24, p. 7538, 2024.

[2] S. Jha et al., “Machine learning approaches for non-invasive diagnosis of endometriosis: A systematic review and meta-analysis,” PLoS One, vol. 19, no. 3, e0300186, 2024.

[3] A. Toglia et al., “Deep learning for endometriosis diagnosis in ultrasound images: A comprehensive survey,” Results Control Optim., vol. 15, p. 100411, 2024.

[4] T. Sgargi et al., “Artificial intelligence in clinical decision-making for endometriosis: A review on current landscape and future perspectives,” J. Clin. Med., vol. 13, no. 18, p. 5445, 2024.

[5] Y. Chen et al., “The power of symptom mapping: How patient-reported outcome measures and machine learning can advance clinical care in endometriosis,” Chronic Pain Manag.: Int. J., pp. 1–13, 2025.