

Linguistic-agnostic Intelligent Support Systems for Probable Diagnosis to Aid in Clinical Settings

Gautam Ahuja

Capstone Project Report submitted to the Department of Computer Science
for the fulfillment of the requirements of the degree of

Postgraduate Diploma in Advanced Studies and Research (DipASR)

in Computer Science at Ashoka University

Supervisor:

1. Dr. Anurag Agarwal, Dean, Trivedi School of BioSciences, Ashoka University
2. Dr. Govind K Makharia, Professor, Department of Gastroenterology and Human Nutrition Adjunct Faculty, AIIMS - New Delhi
3. Dr. Rintu Kutum, Faculty Fellow, Department of Computer Science, Ashoka University

Project Assessment Committee:

1. Dr. Partha Pratim Das, Professor of Computer Science, Ashoka University
2. Dr. Ayush Agarwal, Department of Gastroenterology and Human Nutrition, AIIMS, New Delhi

The capstone project report was submitted to the Ashoka University on December 08, 2024 and accepted by the Department of Computer Science on December 08, 2024.

Acknowledgment

I want to thank my advisor, Dr. Rintu Kutum. I am grateful for his guidance, support, and encouragement throughout my project. I would also like to thank Dr. Anurag Agarwal and Dr. Govind K Makharia for their valuable suggestions and feedback on my project. I want to acknowledge the help of Sanjana Ahuja, who helped me with the initial design of my project, bringing in her expertise in the field of creative design. I want to thank my colleagues at the Augmented Health Systems Lab for their support and feedback. I want to thank my family and friends for their constant support and encouragement. Lastly, I would like to thank the Koita Centre for Digital Health - Ashoka, Department of Computer Science, Ashoka University, and AIIMS, New Delhi, for providing me with the necessary guidelines, resources and infrastructure to carry out my project.

Certificate

I, **Gautam Ahuja**, declare that the work entitled **Linguistic-agnostic Intelligent Support Systems for Probable Diagnosis to Aid in Clinical Settings** is an authentic record of my work carried out at the **Department of Computer Science, Ashoka University**, under the guidance of my supervisor, **Dr. Rintu Kutum** for the fulfillment of the requirements of the degree of **Postgraduate Diploma in Advanced Studies and Research (DipASR)** in the **Department of Computer Science** at **Ashoka University**.

I further declare that this thesis has not been submitted, in part or in full, for the award of any other degree, diploma, or certificate at this or any other institution. All sources of information, ideas, or quotations that are not my own have been duly acknowledged in the bibliography.

This declaration is made with the full knowledge and understanding of the university's regulations on academic honesty and integrity.

This thesis has been approved by the thesis advisor and the examination committee, and it fulfills the requirements for the award of the degree.

Date: December 08, 2024



Gautam Ahuja
ASP' 25, 1020211395
Department of Computer Science
Ashoka University

Dr. Rintu Kutum
Department of Computer Science
Ashoka University

Contents

| | |
|--|------|
| Acknowledgment | iii |
| Certificate | v |
| Contents | vii |
| List of Figures | ix |
| List of Tables | xi |
| Acronyms | xiii |
| 1 Introduction | 1 |
| 2 Background and Motivation | 3 |
| 2.1 Background | 3 |
| 2.1.1 Evaluation of Abdominal Pain | 3 |
| 2.1.2 Clinical Decision Support Systems (CDSS) | 5 |
| 2.1.3 Linguistics and Conversational Agents in Healthcare | 6 |
| 2.2 Motivation | 8 |
| 2.2.1 Department of Gastroenterology and Human Nutrition, AIIMS, New Delhi | 8 |
| 2.2.2 Source of Data | 9 |
| 2.2.3 Broader Impact | 10 |
| 3 Literature Review | 11 |
| 3.1 Literature Survey | 11 |
| 3.1.1 A brief history of the field | 11 |
| 3.1.2 The Era of Knowledge-Based Systems (KBS) | 11 |
| 3.2 State of the Art | 12 |
| 3.3 Gap Analysis | 13 |
| 4 Content | 15 |
| 4.1 Problem Statement | 15 |
| 4.2 Objectives | 15 |
| 4.3 Scope & Boundaries | 16 |
| 4.3.1 Scope | 16 |
| 4.3.2 Boundaries | 16 |
| 5 Methodology | 17 |
| 5.1 Protocol | 17 |
| 5.2 Probable Diagnosis and Organ of Origin | 17 |
| 5.2.1 Diagnosis Mapping | 18 |
| 5.2.2 Organ-of-Origin Identification | 18 |

| | | |
|---------------------|---|-----------|
| 5.2.3 | Workflow | 19 |
| 5.3 | Implementation of the CDSS | 20 |
| 5.3.1 | Frontend Development and User Interface | 20 |
| 5.3.2 | Backend Development | 20 |
| 5.3.2.1 | Data Dictionary | 20 |
| 5.3.2.2 | Conversational Agent | 21 |
| 5.3.3 | System Design | 21 |
| 5.3.4 | Augmented Workflow | 22 |
| 5.3.5 | Challenges & Mitigations | 22 |
| 5.4 | Evaluation | 23 |
| 6 | Results & Discussions | 25 |
| 6.1 | Results | 25 |
| 6.2 | Discussion | 25 |
| 7 | Conclusion & Future Work | 27 |
| 7.1 | Conclusion | 27 |
| 7.2 | Future Work | 27 |
| Bibliography | | 29 |
| Appendix | | 33 |
| 1 | Probable Diagnosis and Organ of Origin | 33 |
| 2 | Questionnaire | 34 |
| 3 | Frontend Screenshots for the Implemented CDSS | 36 |

List of Figures

| | | |
|-----|--|----|
| 2.1 | Evaluation of Abdominal Pain | 4 |
| 2.2 | Clinical Decision Support System | 6 |
| 2.3 | Linguistic Levels | 7 |
| 2.4 | Possible Combinations of Answers | 9 |
| 2.5 | Inpatient Distribution | 10 |
| 3.1 | Three pillars of explainable AI | 14 |
| 5.1 | Breakdown of the 17 Questions into Categories | 18 |
| 5.2 | Graphical Representation of the Mapping between Questions, Diagnoses, and Organs of Origin | 19 |
| 5.3 | Workflow of the Patient-Questionnaire Interaction | 19 |
| 5.4 | Creation of Conversational Agent | 21 |
| 5.5 | High-Level Design of the CDSS | 22 |
| 5.6 | CDSS Augmented Workflow for Abdominal Pain Evaluation | 22 |
| 1 | Regions of the Abdomen and Patterns of Pain. | 34 |
| 2 | Screenshots of the Android Application. | 36 |
| 3 | Screenshots of the Web Application. | 36 |
| 4 | Result Page of the Web Application. | 37 |

List of Tables

| | | |
|-----|--|----|
| 2.1 | AIIMS, New Delhi OPD Cases for The Department of Gastroenterology and Human Nutrition | 8 |
| 1 | Mapping of Organs of Origin to Probable Diagnoses. | 33 |
| 2 | Questionnaire for the Diagnostic Protocol. | 35 |

Acronyms

| | |
|-------|--|
| AI | Artificial Intelligence. |
| AIIMS | All India Institute of Medical Sciences. |
| AIM | Artificial Intelligence in Medicine. |
| CDS | Clinical Decision Support. |
| CDSS | Clinical Decision Support System. |
| LLM | Large Language Model. |
| NLP | Natural Language Processing. |
| OPD | Outpatient Department. |
| osLLM | Open Source Large Language Model. |

1 Introduction

Computer-aided clinical decision-making and reasoning have long been considered a model of human behavior. For many years, it has influenced and been the subject of Artificial Intelligence study [1]. From the very inception of reasoning foundations in medical diagnosis and Clinical Decision Support nearly 65 years ago—where the reasoning was based on symbolic logic and probability understanding and optimum treatment was calculated via value theory—to current advancements in AI, decision systems are being designed to model human knowledge and augment the work of clinicians [2, 3].

This project, in collaboration with the Department of Gastroenterology and Human Nutrition at the All India Institute of Medical Sciences, New Delhi, aims to develop a department-specific Clinical Decision Support System tailored to address the unique needs in the Outpatient Department setting, specifically for **abdominal pain**. The system introduces an intelligent CDSS, where a linguistic-agnostic conversational agent engages directly with patients, collecting key responses and extracting information from a protocol-driven questionnaire. This information is then presented along with a probable diagnosis and organ of origin.

Contemporary AI-based CDSS tends to operate as generalized, multi-purpose diagnostic tools that operate autonomously, often assuming full responsibility for diagnosis. By introducing a **department-specific, protocol-based deterministic approach** with AI-based linguistically agnostic conversational agents for providing probable diagnosis and organ of origin, the system ensures transparency, and accountability as the final decision remains with the physician.

The remainder of this report is organized as follows:

- **Section 2** outlines the background and motivation, including abdominal pain diagnosis, clinical decision support systems, linguistics, and reasoning-based medical diagnosis.
- **Section 3** provides a detailed literature survey, identifying the state-of-the-art in CDSS and gaps.
- **Section 4** outlines the problem statement, project objectives, scope, and boundaries of the system.
- **Section 5** explains the system's design, methodology, and implementation.
- **Section 6** presents results and discussions.
- **Section 7** concludes the report with future work and recommendations.

While the text-based conversational agent has been fully developed and tested as part of this thesis, the voice-based, linguistically-agnostic functionalities are planned for implementation in the following semester

2 Background and Motivation

2.1 Background

2.1.1 Evaluation of Abdominal Pain

Abdominal pain is one of the most common and diagnostically challenging chief complaints encountered in clinical practice [4]. This serves as a symptom of a wide range of underlying conditions, encompassing both gastrointestinal and non-gastrointestinal issues. This wide range of potential causes increases diagnostic complexity, requiring a systematic and comprehensive evaluation process.

The assessment of abdominal pain typically follows a structured framework that incorporates multiple clinical dimensions. These dimensions provide crucial clues that aid in narrowing down possible diagnoses. The key aspects considered during an evaluation include:

- **Location of Pain:** The abdomen is anatomically divided into nine regions (as shown in Figure 2.1 (A) below) — epigastric, umbilical, hypogastric, bilateral hypochondriac, bilateral lumbar, and bilateral iliac regions [5]. The specific region where pain is reported often serves as an essential diagnostic clue [4].
- **Presence of Danger Signs:** Symptoms like lightheadedness, altered sensorium, and respiratory distress may signal critical underlying conditions requiring immediate attention.
- **Severity of Pain:** Pain intensity (1-10) is subjectively reported by the patient but often correlates with the urgency of the clinical situation.
- **Onset of Pain:** The onset of pain, whether acute or gradual (insidious).
- **Character of Pain:** Pain can be classified as burning, stabbing, pin-pricking, constricting, throbbing, dull aching, or non-specific. Each pain type is linked to a distinct set of possible diagnoses.
- **Duration of Pain:** Short, self-limiting pain or prolonged or recurring pain.
- **Radiation of Pain:** The direction of pain radiation (to the back, shoulder, groin/inner thigh, arms, or neck) provides vital diagnostic insight.
- **Aggravating Factors:** Activities, such as eating, bending forward or sideways, passing stool, passing urine, menstruation, deep inspiration, or walking and exercise, can exacerbate abdominal pain.
- **Associated Symptoms:** Symptoms like fever, nausea, vomiting, constipation, diarrhea, jaundice, and changes in bowel habits serve as crucial diagnostic adjuncts. Additionally, systemic symptoms like weight loss, loss of appetite, and signs of anxiety or depression may signal specific gastrointestinal disorders.

- **Comorbidities:** Chronic illnesses such as diabetes, cardiovascular disease, kidney disease, or a history of gallstones can affect abdominal pain etiology.
- **Surgical History:** Prior surgical procedures on the gallbladder, intestines, kidneys, or uterus.
- **Gender-Specific Considerations:** Conditions related to the reproductive system, such as abnormal vaginal bleeding, menstrual irregularities, and foul discharge, play a role in the evaluation of abdominal pain, particularly in female patients.

The process of evaluating abdominal pain typically begins with patient history-taking to collect information on the above-mentioned aspects. This is followed by a physical examination, as shown in Figure 2.1 (B) below, where physicians use visual inspection and hands-on techniques such as palpation, percussion, and auscultation. The examination is often performed with the patient in a supine position with bent knees to relax the abdominal muscles and facilitate assessment [5, 6].

If the initial evaluation does not provide sufficient diagnostic clarity, physicians may order laboratory tests (e.g., blood work, urine analysis) or imaging studies (e.g., ultrasound, X-ray, or CT scan) to gather further evidence. Based on the collective information from history, physical examination, and diagnostic tests, physicians generate a differential diagnosis — a list of possible conditions that could explain the symptoms. This process continues until the most probable diagnosis is reached [7].

The CDSS developed as part of this thesis aims to streamline the information-gathering phase of abdominal pain evaluation. By integrating a linguistically-agnostic, text-based conversational agent, the system collects essential patient information regarding the dimensions mentioned earlier to identify a probable diagnosis and organ of origin using a deterministic approach. The final decision and subsequent physical examination remain under the physician's control.

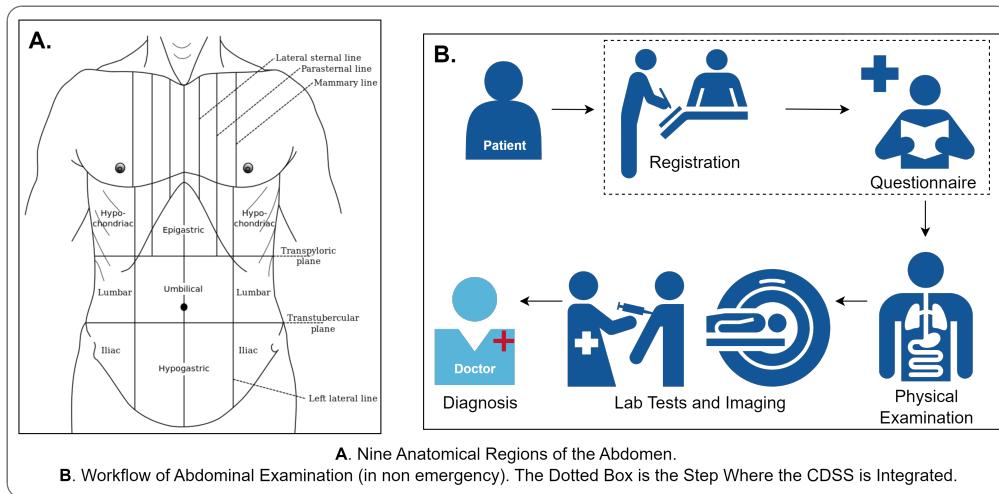


Figure 2.1: Evaluation of Abdominal Pain

2.1.2 Clinical Decision Support Systems (CDSS)

Clinical Decision Support (CDS)

Clinical Decision Support (CDS) refers to a broad set of tools, systems, and interventions aimed at enhancing clinical decision-making. By providing clinicians, healthcare workers, and patients with **situation-specific knowledge**, CDS supports a range of critical healthcare activities such as diagnosis, risk assessment, prognosis, and treatment selection [8]. This support can be delivered through various mediums, including reference guidelines, alerts, and evidence-based recommendations.

The goal of CDS is to **bridge the gap between clinical knowledge and patient care**. As medical knowledge expands and healthcare systems become more data-driven, CDS ensures that healthcare providers have access to timely and relevant information. This, in turn, enables more efficient and informed decision-making.

Clinical Decision Support System (CDSS)

A Clinical Decision Support System (CDSS) is a specialized type of CDS that employs **software-based tools** to directly aid healthcare providers in making clinical decisions. Unlike general CDS, which may include static guidelines or references, CDSS is dynamic, **patient-specific**, and interactive. By integrating patient-specific information with a computerized knowledge base, CDSS generates recommendations, alerts, and assessments that assist clinicians in diagnosis, treatment, and care planning [9].

“Any software designed to aid in clinical decision-making by matching patient characteristics to a computerized knowledge base to generate tailored patient-specific assessments or recommendations”

- Hunt et al. [9]

Unlike fully autonomous AI models, CDSS systems are designed to work **in collaboration with healthcare professionals**, offering them interpretive insights and enabling human oversight.

Effective clinical decision-making requires a clinician to have:

- Up-to-date Medical Knowledge
- Access to Accurate and Complete Patient Data
- Good Decision-Making Skills

However, clinicians face several significant challenges in meeting these requirements. Some of the prominent challenges include:

- Exponential Growth of Medical Knowledge
- Rapid Accumulation of Patient Data
- Increased Complexity in Clinical Inference

- Clinical Data Capture and Documentation Burden

Given the challenges and the rapid advancements in health makes it impossible for a clinician to remember and apply them in clinical care without some form of assistance, thus the role of Clinical Decision Support Systems (CDSS) becomes essential in modern healthcare [10]. With the adoption of Artificial Intelligence in Medicine, existing physicians can handle more cases as systems become more automated [11].

Types of Clinical Decision Support System (CDSS)

As per [10] CDSS can be broadly classified into two categories:

- **Knowledge-based CDSS:** The key components of a knowledge-based AI-CDS system include a knowledge base such as expert-derived rules and an inference mechanism for clinical application such as chained inference for rules.
- **Data Derived CDSS:** The key components of a data-derived AI-CDS include a model, such as a data-derived neural network, and an inference mechanism for clinical application, such as forward propagation in a neural network model..

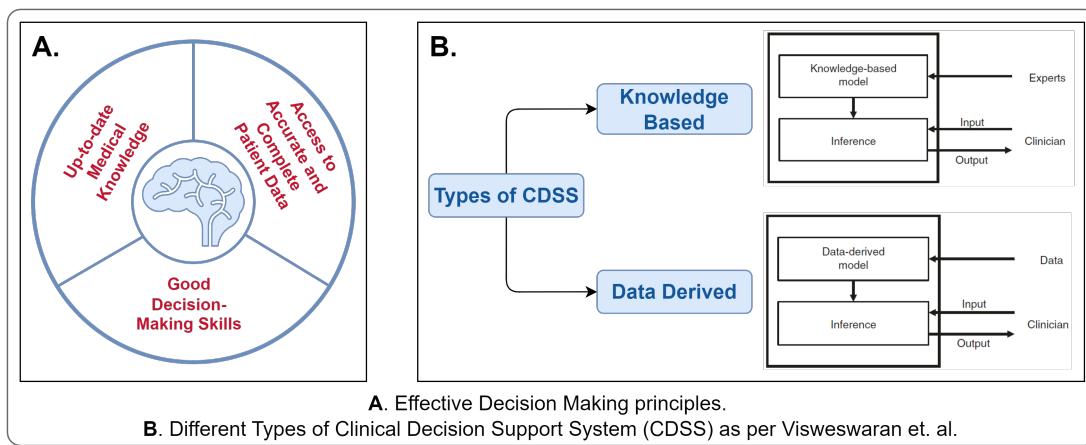


Figure 2.2: Clinical Decision Support System

2.1.3 Linguistics and Conversational Agents in Healthcare

Computational linguistics focuses on **modeling human language** using computational techniques, enabling machines to understand and process human language. The process of language understanding can be broken down into multiple hierarchical levels, each representing a critical step in the language comprehension pipeline [12]. These levels, illustrated in Figure 2.3, are as follows:

- **Phonology:** This level focuses on sound patterns in language. One of the most notable linguistic tasks at this level is **Automatic Speech Recognition (ASR)**, also known as speech-to-text, where speech waveforms are converted into textual data. The concept of **linguistic-agnostic** systems relates to the ability of ASR models to understand diverse accents, dialects, and speech variations without being restricted to a particular language or pronunciation. For example, a linguistic-agnostic system can process English spoken by people from different regions (e.g., American, British, and Indian accents) with equal accuracy.

- **Morphology:** Morphology focuses on how words are formed by combining **morphemes**, the smallest units of meaning. For example, in the word "unhappiness," the morphemes are "un-", "happy", and "-ness".
- **Syntax:** Syntax refers to the grammatical structure of language and focuses on how words are arranged in a sentence. It identifies the **relationships between words** and builds a hierarchical representation of the sentence, allowing for proper interpretation
- **Semantics:** Semantics deals with **word meanings** and their relationships within sentences. It enables systems to understand the meaning of words, phrases, and sentences after receiving inputs from the phonology, morphology, and syntax layers
- **Pragmatics:** This layer focuses on how context influences meaning by going beyond the literal meaning of words.
- **Generation:** While the above layers focus on language understanding, this final layer focuses on **realistic language generation** given a computational representation

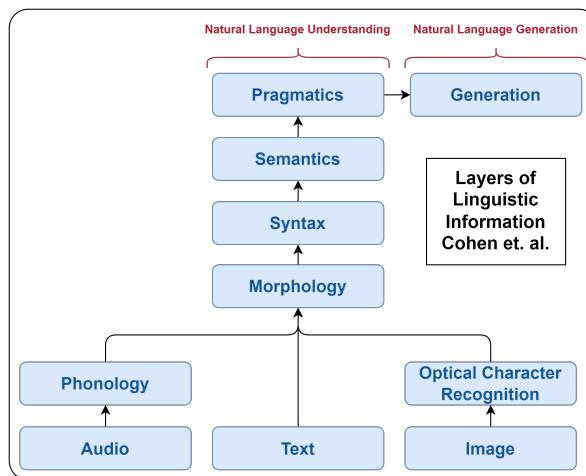


Figure 2.3: Linguistic Levels

Accurate interpretation of human language—especially spoken language—is one of the most critical factors that influence the success of **human-computer interaction**. To achieve a **linguistic-agnostic system**, a large, diverse dataset is required to capture regional, phonetic, and dialectal variations in speech.

Conversational agents, also known as **chatbots**, have a long history in healthcare, dating back to the development of ELIZA. Joseph Weizenbaum developed one of the first medical chatbots in the mid-1960s at the Massachusetts Institute of Technology (MIT) [13, 14]. Although ELIZA operated on simple pattern-matching rules, it demonstrated the potential of conversational agents to simulate human interaction.

Since then, NLP has progressed to the current state of Large Language Models (LLMs) are advanced **AI models** specifically trained to process, understand, and generate text.

One of the popular ways is how LLMs are used as conversational agents. Since the release of ChatGPT, numerous LLMs—both open-source and proprietary—have been developed at unprecedented speed [15].

2.2 Motivation

2.2.1 Department of Gastroenterology and Human Nutrition, AIIMS, New Delhi

The **Department of Gastroenterology and Human Nutrition** at the All India Institute of Medical Sciences (AIIMS), New Delhi, was established in 1971 to create qualified gastroenterologists for the country. It is one of the 49 teaching departments and centers at AIIMS, New Delhi.

According to the 67th AIIMS Annual Report (2022-2023) [16], the department managed a total of **1,35,944 outpatient department (OPD) cases**, of which **42,586 were new cases** and **93,358 were follow-up cases**. This is a significant increase from the previous year's figures of 17,790 new cases and 35,622 follow-up cases as shown in Table 2.1 below. In total, the Main Hospital at AIIMS catered to **10,39,523 patients** through its General OPD, Specialty Clinics, and Emergency Department.

| Year | New Cases | Follow-up Cases | Total Cases |
|-----------|-----------|-----------------|-------------|
| 2020-2021 | 7,920 | 11,956 | 19,876 |
| 2021-2022 | 17,790 | 35,622 | 53,412 |
| 2022-2023 | 42,586 | 93,358 | 1,35,944 |

Table 2.1: AIIMS, New Delhi OPD Cases for The Department of Gastroenterology and Human Nutrition

The Routine Gastroenterology OPD operates from Monday to Friday, 8:30 a.m. to 1:00 p.m. [17]. Given this limited time frame of 270 minutes daily over 5 working days, approximately **8,500 new cases are handled per day**, with multiple physicians addressing various chief complaints.

As discussed in Section 2.1.1, one of the most diagnostically challenging chief complaints is **abdominal pain**. A complete protocol of structured questions helps filter the probable diagnosis and the organ of origin. However, the high influx of patients leaves little room for physicians to conduct comprehensive questionnaire-based evaluations before physical examinations. Consequently, many patients do not undergo the complete protocol, increasing the cognitive burden on physicians and the risk of diagnostic errors.

This challenge forms the motivation for the development of a Clinical Decision Support System (CDSS). The CDSS aims to streamline the process of collecting responses to structured questionnaires and assist physicians in formulating probable diagnoses and identifying the organ of origin before physical examination.

Patient-side motivation: When deployed, the system can handle multiple patients simultaneously via multiple devices. It can conduct the initial evaluation and generate a probable diagnosis and organ of origin. The output will be printed for the patient, who can then present it to the physician. This approach reduces the consultation time for each patient and allows for a more focused interaction during the physical examination.

Physician-side motivation: Physicians no longer need to ask every patient the same set of questions repeatedly, allowing them to concentrate on higher-order tasks such as diagnosis and treatment. Additionally, the inferred probable diagnosis and identified organ of origin significantly aid physicians in their decision-making. As depicted in the bar chart in Figure 2.4, for a set of 29 probable diagnoses, the top diagnosis can have upwards of **5,000 possible combinations of answers** to the questionnaire. By automating this process, the physician's cognitive load is significantly reduced, thereby improving efficiency and reducing diagnostic errors.

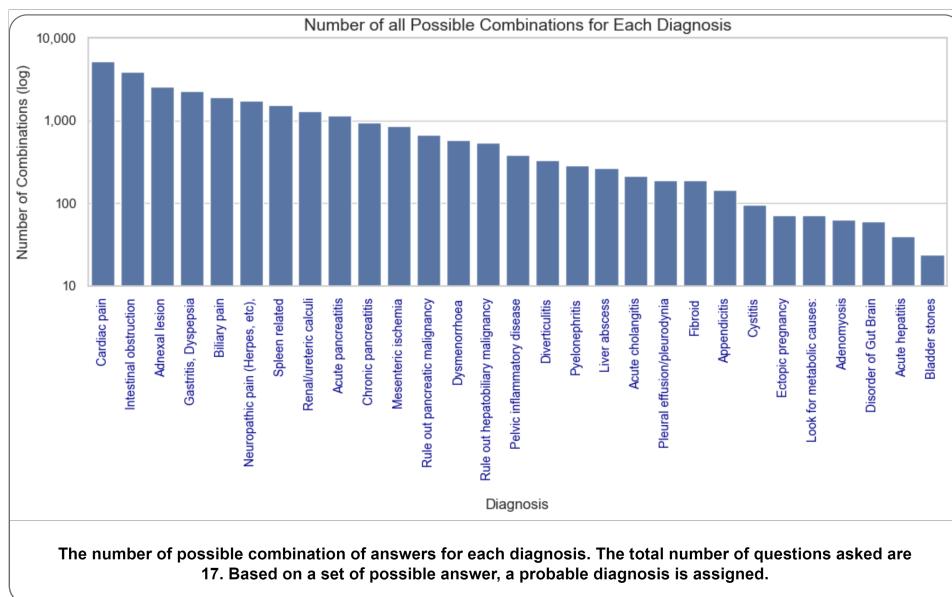
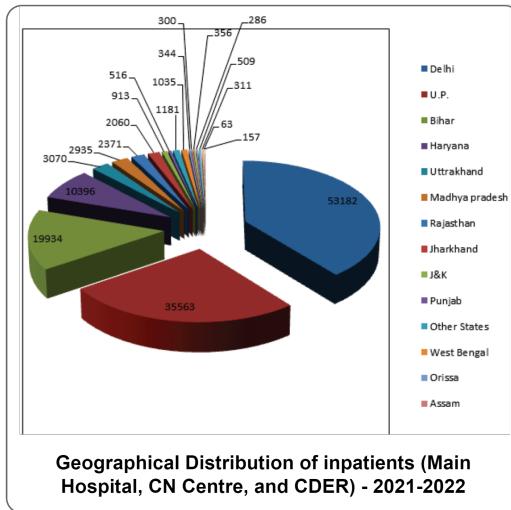


Figure 2.4: Possible Combinations of Answers

2.2.2 Source of Data

Another significant motivation is the **collection of diverse voice-based data** from patients. AIIMS serves as a **melting pot for linguistically diverse populations** from across India. This is reflected in the geographical distribution of inpatients at AIIMS during the year 2021-2022, as shown in the Figure 2.5 below [16].

This diversity in linguistic backgrounds presents a unique opportunity to collect and analyze **voice-based data** from patients. Such data can be used to develop **robust linguistic-agnostic conversational models** for the Indian population. A linguistic-agnostic system can recognize and process multiple accents, regional pronunciations, and dialectical variations. Including these better and more robust models will increase patient care and inclusivity.

**Figure 2.5:** Inpatient Distribution

2.2.3 Broader Impact

The broader impact of this system extends beyond the Department of Gastroenterology. It addresses several challenges posed by general-purpose large language models (LLMs) such as ChatGPT, which are not well-suited for department-specific protocols [15]. The issues are discussed in detail in the Literature Review section.

Given these limitations, the proposed approach can be extended to other departments and target chief complaints other than abdominal pain. This would allow the development of multiple CDSS subsystems for various medical specialties. These subsystems can work in coordination, forming a broader, integrated, AI-assisted healthcare ecosystem.

3 Literature Review

3.1 Literature Survey

3.1.1 A brief history of the field

The field of diagnostic reasoning in medicine became an early focus of AI, demonstrating that AI methods could approximate human performance in tasks requiring extensive domain knowledge. Early systems were designed to model human cognition explicitly, prioritizing interpretability over mere optimization of accuracy. Such systems were particularly adept at explaining their reasoning, contrasting with modern AI systems optimized solely for prediction accuracy, often at the cost of transparency [1].

One of the landmark systems in this domain was **MYCIN**, introduced in the 1970s [12, 18, 19]. MYCIN was developed to assist physicians in selecting appropriate antimicrobial therapies for severe infections. Its key components included:

- **Consultation Program:** Acquired patient data and provided treatment recommendations.
- **Explanation Program:** Generated English-language explanations, detailing why certain questions were asked and how conclusions were reached.

MYCIN's knowledge of infectious diseases was represented as production rules—conditional statements linking observations to inferred outcomes. It introduced backward chaining, a reasoning strategy that began with a hypothesis and worked backward to validate it using available evidence. This allowed MYCIN to answer "WHY" questions, making its decision-making process interpretable and user-friendly [19, 20].

Before MYCIN, the Leeds Abdominal Pain System, developed in the late 1960s, marked another early attempt at diagnostic reasoning. F. T. de Dombal and his colleagues at the University of Leeds created decision aids for diagnosing abdominal pain based on Bayesian probability theory. This work laid the foundation for probabilistic reasoning in medical decision-making [20].

The initial wave of AI-driven diagnostic systems emphasized the trade-off between accuracy and interpretability, often prioritizing the latter. Over time, the focus shifted to integrating AI into a larger system involving human decision-makers, aiming to improve the quality, efficiency, and safety of clinical practice.

3.1.2 The Era of Knowledge-Based Systems (KBS)

Following systems like MYCIN and the Leeds Abdominal Pain System, the next phase of AI in healthcare saw the rise of Knowledge-Based Systems (KBS). These systems

sought to replicate human reasoning in complex medical scenarios by formalizing knowledge into computational representations. The core ideas of KBS were:

1. **Representing knowledge using:**
 - Formal methods: Mathematical frameworks for precise reasoning.
 - Ontological commitments: Hierarchies of concepts organized logically.
 - Fragmentary theories of reasoning: Integrating logic, psychology, biology, statistics, and economics.
2. **Ensuring** that representations were computationally efficient and intuitive for human practitioners to understand and modify.

Prominent approaches to knowledge representation included Rules and Patterns—Logical or heuristic rules for decision-making. Probabilistic Models—Methods like Naive Bayes, Bayesian Networks, and Influence Diagrams. Causal Mechanisms—Explaining outcomes through cause-and-effect relationships. Fuzzy Logic—Handling uncertainty and imprecision in medical reasoning.

The success of KBS depended on robust methods for acquiring and organizing medical knowledge. Techniques included:

- **Taxonomic Ontologies:** Organizing concepts into hierarchical structures, specifying their super- and sub-categories.
- **Knowledge Graphs:** Capturing relationships between concepts for intuitive reasoning.
- **Textual Co-occurrence Analysis:** Identifying relationships from sentences, paragraphs, or articles.
- **Unified Medical Language Systems (UMLS):** Leveraging the Metathesaurus to bridge semantic gaps without constructing exhaustive ontologies.
- **Triple-Based Models:** Representing knowledge as ([concept] - [relation] - [concept]) triples, enabling systems to address "which," "why," and "does" questions [21].

Some other methods included conception and prototypical design of a internet based decision-support server [22] with a option-based input method and a knowledge-based, terminologically driven data processing system for decision support. These methods were observed to have a significant impact on user-friendliness and performance of the system.

3.2 State of the Art

The evolution of AI technologies in healthcare has driven a shift toward more dialogue-based approaches, taking advantage of intelligent agents to facilitate interactive conversations. These systems are designed to interact with humans using dialog systems, which are computational frameworks for managing dialog. The development of such technologies is particularly critical in healthcare, where the delivery of complex information must be accessible. These systems can play a transformative role for individuals with low health literacy or limited familiarity with technology, helping them navigate complex medical advice and

treatment options effectively [12].

Recent advancements in conversational LLMs like GPT-4 have demonstrated remarkable capabilities in reasoning, planning, and using contextual information. While GPT-4 is trained on vast amounts of text data, and primarily developed for general domain, it lacks the specialized medical knowledge required for clinical decision-making. Clinical data tends to be proprietary, sensitive, and subject to strict privacy regulations, making it challenging to access and use for training AI models. Because of the interactive nature of the system, the user can request more detail regarding the response by asking follow-up questions or asking for more concise responses in order to get “to the point” more rapidly [23].

On the other hand, systems like Google AMIE: A research AI system for diagnostic medical reasoning and conversations [24] represent a new generation of AI tailored specifically for medical diagnostics and reasoning. Google AMIE is a research-focused AI designed to:

- **Perform Diagnostic Medical Reasoning:** Simulating primary care physician (PCP) consultations and Objective Structured Clinical Examinations (OSCE).
- **Modular Architecture:** AMIE utilizes a stage and agents-based architecture, allowing for easy integration of new modules and agents.

A notable emerging trend also being observed where graphs-based models and graph representation learning [25] is being used in medicine to manage and analyze the vast, multi-modal datasets generated in healthcare. Graph-based models are particularly well-suited for capturing complex relationships between entities, such as diseases, symptoms, treatments, and patient demographics.

3.3 Gap Analysis

A significant limitation of current approaches is their black-box nature and full autonomy in diagnostic decisions [15]. There is a pressing need for auditable and traceable systems where AI assists rather than replaces clinical decision-making [12]. The ideal solution should provide explainable recommendations while keeping diagnostic control firmly in the hands of healthcare professionals, ensuring accountability, and maintaining the critical role of human expertise in patient care as explained by the three pillars in Figure 3.1 [26]. This is particularly important in specialized medical departments where standardized protocols and human oversight are essential for patient safety [12].

While recent advancements in LLMs have shown impressive capabilities in medical domains, there remains a critical gap in **department-specific** Clinical Decision Support Systems. Current general-purpose AI systems lack the specialized knowledge and protocols required for specific medical departments, potentially leading to misinformation and compromised patient care. Additionally, most existing systems operate as standalone solutions rather than taking a system-level perspective that considers integration with existing clinical workflows and protocols, incorporating both **rule-based** and **data-driven** approaches while also ensuring **explainability** and **transparency** in decision-making.

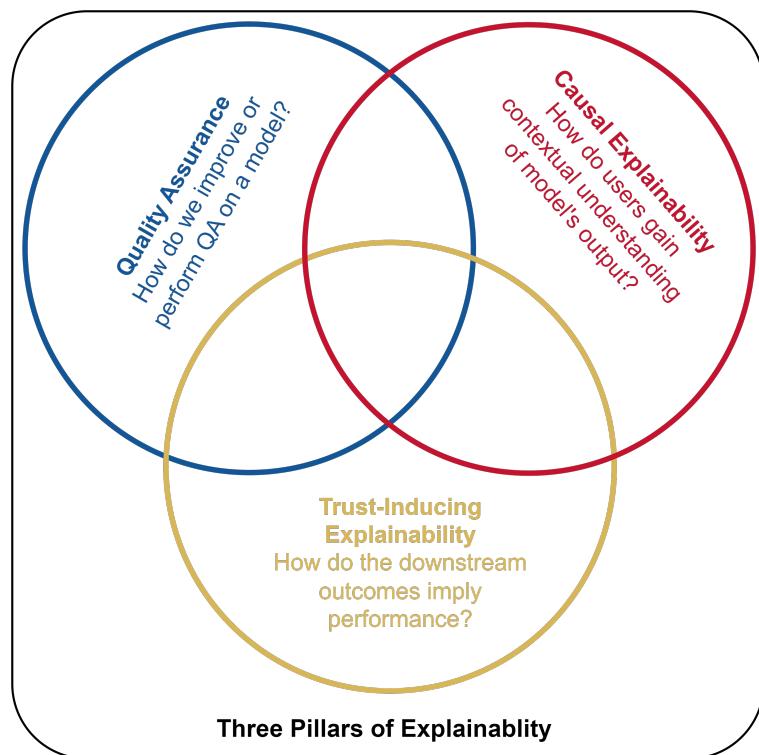


Figure 3.1: Three pillars of explainable AI

4 Content

4.1 Problem Statement

The Department of Gastroenterology and Human Nutrition at AIIMS, New Delhi, handles a high influx of outpatient cases daily. Due to time constraints and the high volume of patients, physicians have limited capacity to conduct comprehensive, protocol-driven evaluations for each patient. For the chief complaint of abdominal pain, a structured diagnostic approach is required. This process typically begins with identifying the pain rating and location, followed by questions on aggravating factors, associated symptoms, and the patient's medical history. However, given the volume of patients, it becomes difficult for physicians to conduct a complete assessment for every individual.

To ensure patient care quality, there is a need for a support system that can streamline the diagnostic process by guiding the patients through initial screening and decreasing the cognitive burden on healthcare providers.

The secondary problem is the need for linguistically agnostic systems. Since patients at AIIMS come from diverse linguistic backgrounds, it is essential to create a system that can interact with patients regardless of their language proficiency.

4.2 Objectives

The primary objectives of this project are as follows:

- **Development of a Department-Specific CDSS:** To design and develop a department-specific, intelligent clinical decision support system for the screening of abdominal pain. This system will identify the probable diagnosis and organ of origin based on a structured, protocol-driven questionnaire to aid physicians.
- **Design of a Linguistic-Agnostic Conversational Agent:** To design and develop a conversational agent that is linguistic-agnostic to drive above mentioned CDSS.
- **Deployment and Integration:** To deploy and integrate the system in the OPD setting at the Department of Gastroenterology and Human Nutrition, AIIMS, New Delhi.

The project also has a set of secondary objectives, which are as follows:

- **Collection of Linguistic Data:** To collect linguistic data from patients at AIIMS, representing diverse accents, dialects, and language variations. This data will be used to improve the robustness of linguistic-agnostic models for future use.

- **Evaluation of a Rule-Based and Generative Hybrid System:** To develop and evaluate a hybrid approach that combines a rule-based, deterministic decision-making engine with a generative AI conversational agent for more robust and explainable clinical support.

4.3 Scope & Boundaries

4.3.1 Scope

This project focuses on the development of a department-specific Clinical Decision Support System (CDSS) for the Department of Gastroenterology and Human Nutrition at AIIMS, New Delhi. The primary goal of the system is to support the initial evaluation of patients presenting with abdominal pain.

The system does not autonomously diagnose patients. Instead, it provides probable diagnoses and suggests a possible organ of origin. Final diagnostic authority remains with the physician, who can audit and modify the system's suggestions based on clinical judgment. It has Autonomous Decision-Making and follows a human-in-the-loop approach.

4.3.2 Boundaries

The system is limited to the initial evaluation and probable diagnosis of abdominal pain and does not include physical examination components. It is tailored to the Department of Gastroenterology and Human Nutrition's protocol for abdominal pain at AIIMS, New Delhi, and does not generalize to other medical departments or conditions. Additionally, the system does not make final diagnostic decisions, as the ultimate responsibility for diagnosis rests with the physician.

5 Methodology

5.1 Protocol

The development of the Clinical Decision Support System (CDSS) was guided by a structured protocol provided by the physicians of the Department of Gastroenterology and Human Nutrition at AIIMS, New Delhi. This protocol defines the flow of patient interaction through a set of questions designed to capture essential diagnostic information related to abdominal pain.

The initial protocol consisted of a 12-question diagnostic questionnaire that addressed key clinical dimensions of abdominal pain, as described in Section 2.1.1. These questions explored aspects like pain location, intensity, aggravating factors, associated symptoms, etc. However, for a more controlled and streamlined workflow, the questionnaire was later broken down to include 17 refined questions. By refining these questions, a categorization of questions was made possible. For instance, questions related to menstrual cycle health were separated from the general set of symptoms to provide more focused and relevant insights for female patients.

Through refinement, the 17 questions were grouped into the following categories:

1. **Discriminators:** Critical questions are used to differentiate between emergency cases and regular OPD cases. If the answer to any of these questions indicates a potential emergency, the patient is automatically redirected to the emergency department instead of continuing through the normal OPD process.
2. **Demographic:** Questions related to the patient's demographic information, such as age and gender.
3. **Female-Specific:** Questions specific to female health issues — menstrual health — that play a significant role in the diagnostic process for women but are irrelevant to male patients.
4. **General:** These are common questions applicable to all patients, regardless of gender, and form the core of the diagnostic questionnaire.

A detailed breakdown of the 17 questions into the four categories is presented in Figure 5.1. A tabulated version of the 17 questions and their respective answers can be found in Appendix 2.

5.2 Probable Diagnosis and Organ of Origin

The core function of the CDSS is to map patient responses to probable diagnoses and the organ of origin. The process for determining these outcomes relies on a pre-

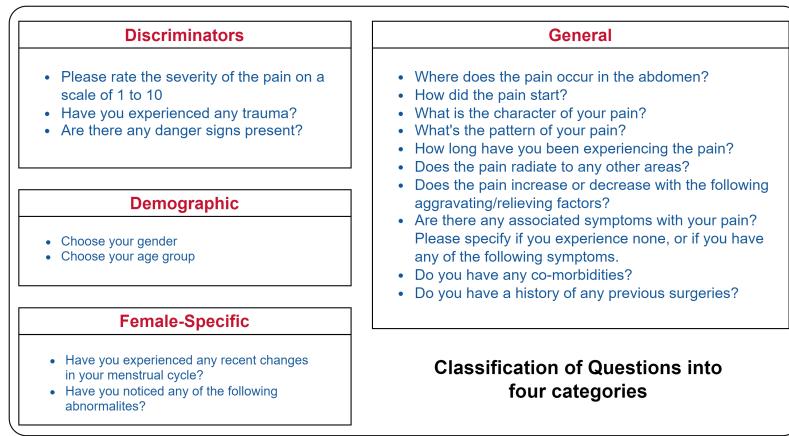


Figure 5.1: Breakdown of the 17 Questions into Categories

defined list of possible diagnoses and corresponding organs of origin, as provided by the Department of Gastroenterology and Human Nutrition at AIIMS.

5.2.1 Diagnosis Mapping

The Physician provided a list of 29 possible diagnoses related to abdominal pain. Each diagnosis is associated with a specific set of responses to the 17 questions from the protocol. To identify a potential diagnosis, the patient's responses are compared against this set of known answer combinations. If the patient's responses align with one of these pre-defined answer sets, a match is established for the probable diagnosis.

5.2.2 Organ-of-Origin Identification

The CDSS performs the identification of the organ of origin for abdominal pain. This identification is done via the mapping established by linking 29 diagnoses to a set of 19 possible organs of origin. Once a probable diagnosis is identified, the system determines the corresponding organ of origin using a deterministic mapping approach. This allows the system to be interpretable.

The relationships between questions, diagnoses, and organs of origin are represented as a graph structure. Each diagnosis is connected to potential answers for the 12 core questions, forming distinct paths in the graph. The graphical representation in Figure 5.2 provides a visual explanation of how diagnoses are reached based on patient responses. A tabulated mapping of organs of origin to diagnoses can be found in Appendix 1.

In Figure 5.2, the pink nodes (center) correspond to the 29 diagnoses, while the yellow nodes represent the 19 possible organs of origin (center below). The nodes representing answers to the 12 original questions are displayed in 12 distinct colors, each corresponding to a specific question.

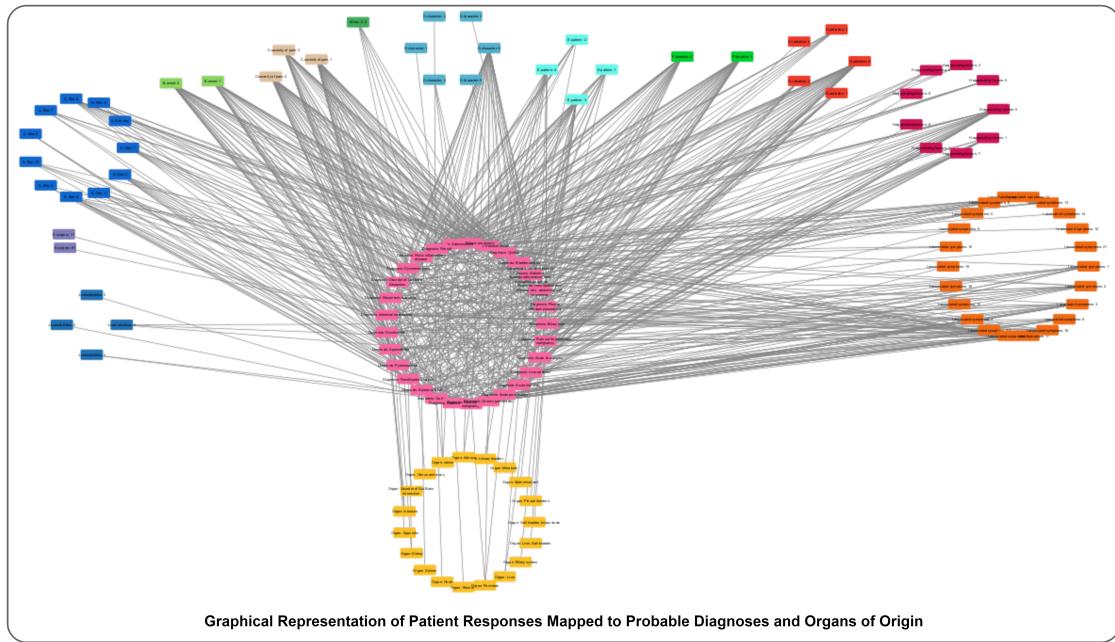


Figure 5.2: Graphical Representation of the Mapping between Questions, Diagnoses, and Organs of Origin

5.2.3 Workflow

The workflow of the questionnaire begins with the patient's interaction with the system through a structured questionnaire. As shown in Figure 5.3, it starts with discriminator questions to identify emergencies; if flagged, the patient is directed to the emergency department. Otherwise, the patient proceeds to demographic questions to gather basic information like age and gender, followed by general questions covering universal symptoms and history. Female patients are asked additional questions related to menstrual health before proceeding to the general questions. The patient's responses are then mapped to probable diagnoses and organs of origin, as described in the previous section. The system generates a report based on the patient's responses, which contains all the answers provided by the patient, the probable diagnosis, and the organ of origin for the physician's reference.

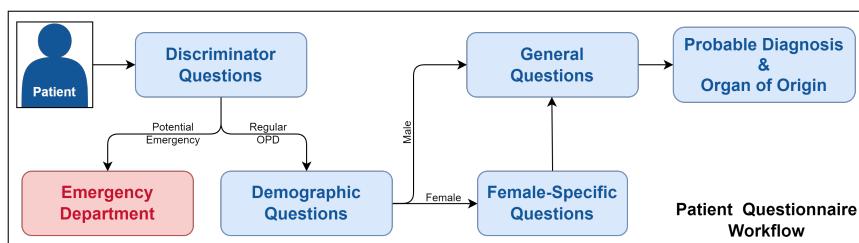


Figure 5.3: Workflow of the Patient-Questionnaire Interaction

5.3 Implementation of the CDSS

The CDSS was implemented using a comprehensive technology stack. The aim was to create a system that is user-friendly, efficient, and easily accessible to physicians. The system was designed to be integrated into the hospital's existing infrastructure, allowing for seamless interaction with the CDSS.

The implementation involved the creation of both an Android application and a web application for patient interaction. The initial version Android application was developed using Java and Kotlin in Android Studio, while the web application was built using Streamlit in Python. Streamlit is an open-source Python framework for data scientists and AI/ML engineers to deliver dynamic data apps [27].

5.3.1 Frontend Development and User Interface

The frontend development involved creating a user-friendly interface for patient interaction. The Android application was designed to be intuitive and easy to navigate, with a clean and simple layout. The web application was developed to provide a seamless experience for patients interacting with the system on a desktop or mobile device. The frontend was designed to be responsive, ensuring that the system is accessible across a wide range of devices.

The initial prototype of both the Android application was designed using Figma, a web-based design tool. Both applications were designed to follow the workflow defined in Figure 5.3. Both applications followed a click-through option for patient interaction, where the patient could navigate through the questionnaire by answering each question sequentially. The responses were stored in the system and used to generate the probable diagnosis and organ of origin. As seen in [22], even a simple interface with a well-defined workflow can be effective and have a significant impact on user-friendliness.

The demonstration of the Android application design can be found on our website [here](#). Another demonstration of the web application for a successful probable diagnosis and organ of origin can be found [here](#). The web application for an emergency department visit can be found [here](#). The screenshots and complete URLs of the same are shown in the Appendix 3. A future version of the CDSS will include the interface for voice-based interaction

5.3.2 Backend Development

The backend of the CDSS has two major components: the data dictionary and the conversational agent. The two components work together via a Python script. The entire backend is containerized using Docker for easy deployment and management.

5.3.2.1 Data Dictionary

Multiple data dictionaries contain the mapping between patient responses and probable diagnoses and organs of origin. These dictionaries are used to identify the probable di-

agnosis and organ of origin based on the patient’s responses. Dictionaries are stored in JSON format for easy access and retrieval.

5.3.2.2 Conversational Agent

The conversational agent is responsible for interacting with the patient and guiding them through the questionnaire. The agent asks questions based on the patient’s responses and stores the answers for further processing. The design of the conversational agent is shown in Figure 5.4.

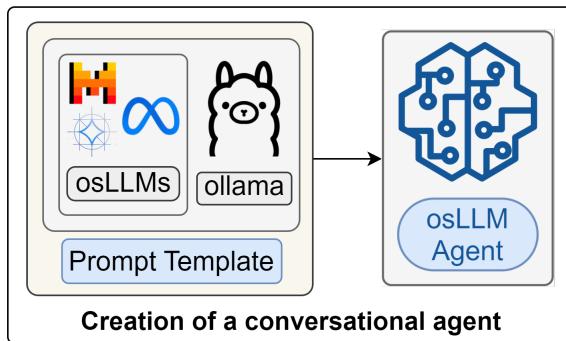


Figure 5.4: Creation of Conversational Agent

We create a conversational agent using the open source Large Language Models (osLLMs) like Mistral (Mistral 7B)[28], Llama (Llama 3.1 8B) [29], Gemma (Gemma 2 9B) [30], and others. The models were selected based on their performance and parameter size. From the above-mentioned models, Gemma 2 9B was selected for the conversational agent based on its performance and parameter size. We use prompt engineering [31] to provide the conversational agent with the necessary context to extract the patient’s responses as accurately as possible after asking each question with its respective answer choices.

The conversational agent is hosted and managed by using Ollama [32], an open-source platform that allows for the deployment and management of large language models. Ollama provides an efficient environment for running these models locally, ensuring privacy and faster processing. Ollama provides an efficient environment for running these models locally, ensuring privacy and faster processing. The model is queried with the patient’s responses using a REST API. We used Nvidia RTX A5000 GPU for loading the model and processing the patient’s responses.

5.3.3 System Design

The system design of the CDSS is shown in Figure 5.5. The system consists of two main components: the frontend and the backend. The frontend includes the Android and web applications, while the backend includes the data dictionary and the conversational agent. The containerized backend can be deployed on a local server or cloud platform such as AWS or Google Cloud.

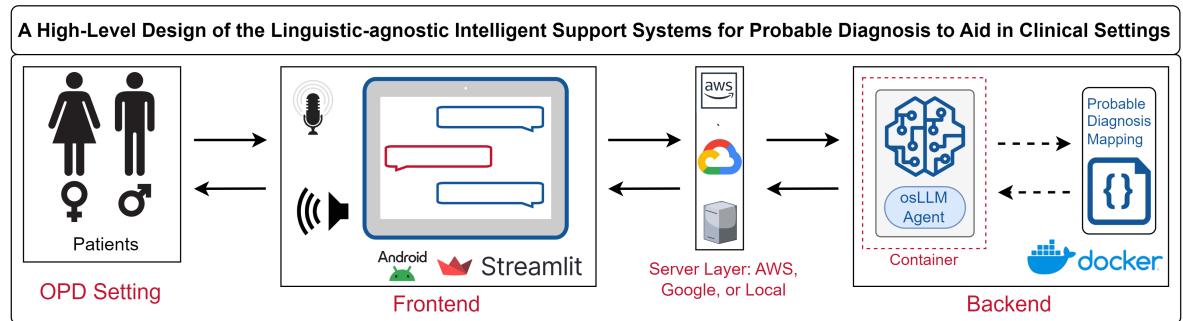


Figure 5.5: High-Level Design of the CDSS

5.3.4 Augmented Workflow

The high-level augmentation of CDSS into the existing workflow of evaluation of abdominal pain is shown in Figure 2.1 (B), depicted by the dotted box. A more detailed view of the augmented workflow is shown in the Figure 5.6.

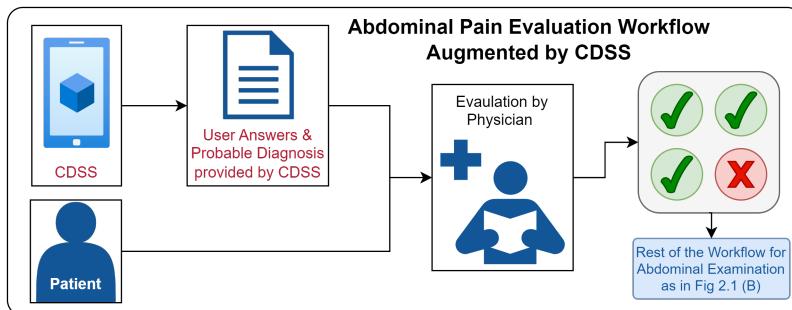


Figure 5.6: CDSS Augmented Workflow for Abdominal Pain Evaluation

The patient interacts with the CDSS through the Android or web application, answering a set of structured questions. The responses are processed by the conversational agent, which identifies the probable diagnosis and organ of origin based on the patient's responses. The system generates a report containing the patient's responses, probable diagnosis, and organ of origin. The report is then reviewed by the physician, who uses the information to further the diagnostic process. The auditing of responses and the probable diagnosis is done by the physician to ensure the accuracy of the system. In case of any discrepancies, the physician can ask the patient for additional information to refine the diagnosis.

5.3.5 Challenges & Mitigations

One of the Challenges of any AI-based CDSS is the issue of explainability. The compartmentalization of the system ensures that the components can be updated independently without affecting the overall functionality of the system. It also ensures that the system can be easily audited and the output of each component can be tracked providing transparency, explainability, and interpretability. In case of any conflicting diagnosis, the output of the system can be easily traced back to the patient's responses and the mapping between the responses and the probable diagnosis.

Another challenge is the issue of privacy and security. The system is designed to be deployed on a local server or cloud platform, ensuring that patient data is stored securely and complies with all relevant data protection regulations. The open-source nature of the system allows for easy customization and mitigates the issue of privacy and security. The system can be easily adapted to meet the specific requirements of different departments or hospitals.

5.4 Evaluation

Currently, the CDSS is in the final stages of development and is yet to be deployed for clinical evaluation. The evaluation process will involve a beta test with a small group of physicians from the Department of Gastroenterology and Human Nutrition at AIIMS. The physicians will interact with the system and provide feedback on the system's usability, accuracy, and overall performance. The feedback will be used to refine the system further before a full-scale deployment.

6 Results & Discussions

6.1 Results

The Clinical Decision Support System (CDSS) implementation successfully addresses the challenge of conducting comprehensive protocol-driven evaluations in high-volume OPD settings. The system integrates a 17-question diagnostic protocol (categorized into discriminators, demographic, female-specific, and general questions) with a linguistic-agnostic conversational agent powered by Google's Gemma 2 9B model. The system's compartmentalized architecture features both Android and web-based frontends alongside a containerized backend, enabling deterministic mapping between patient responses and probable diagnoses with their corresponding organs of origin. This approach ensures transparency and traceability in the decision-making process while maintaining system modularity and interpretability.

6.2 Discussion

The department-specific CDSS demonstrates potential in streamlining the systematic evaluation of abdominal pain, which typically requires assessment across multiple clinical dimensions including pain location, intensity, character, duration, and associated symptoms. By automating the initial information-gathering phase, the system supports physicians in following the structured diagnostic framework while managing high patient volumes. The hybrid approach, combining rule-based decision-making with AI-powered conversation handling, maintains human oversight in the diagnostic process - a crucial factor in clinical settings. Unlike autonomous AI systems, this human-in-the-loop approach ensures that final diagnostic decisions remain with physicians while providing them with organized patient responses and probable diagnoses. While the current text-based implementation shows promise, clinical evaluation at AIIMS will assess its impact on patient flow and diagnostic efficiency. The system's foundation also supports future expansion into voice-based linguistic-agnostic functionality, potentially improving accessibility for India's linguistically diverse patient population.

7 Conclusion & Future Work

7.1 Conclusion

The evaluation of abdominal pain remains one of the most diagnostically challenging tasks in clinical practice, requiring systematic assessment across multiple dimensions. In high-volume settings like the Department of Gastroenterology and Human Nutrition at AIIMS, New Delhi, where physicians handle over 1,35,000 OPD cases annually, conducting comprehensive protocol-driven evaluations for each patient becomes increasingly difficult. While AI-driven Clinical Decision Support Systems offer potential solutions, their black-box nature and autonomous decision-making raise concerns about explainability and accountability in healthcare settings.

This project presents a novel approach that combines the benefits of structured diagnostic protocols with modern AI capabilities. Through collaboration with AIIMS, we developed a department-specific CDSS that uses a deterministic mapping approach to link patient responses from a 17-question diagnostic protocol to probable diagnoses and organs of origin. The system's compartmentalized architecture, featuring both Android and web-based interfaces backed by a containerized backend with Gemma 2 9B model-powered conversational agent, ensures transparency and traceability in the decision-making process.

The system demonstrates how AI can augment rather than replace clinical decision-making by maintaining human oversight while streamlining initial patient evaluation. By following a human-in-the-loop approach, the CDSS supports physicians in managing high patient volumes while ensuring that final diagnostic authority remains with healthcare professionals. The successful implementation of the text-based system, designed to handle diverse patient populations at AIIMS, establishes a foundation for more accessible and efficient clinical support tools.

7.2 Future Work

Future development of the system will focus on several key areas. First, the Android and web applications will be enhanced with additional features including user registration, authentication, and integration with hospital information systems to enable seamless patient data management. The interface will be optimized based on user feedback from the initial deployment phase.

A crucial advancement will be the development and implementation of the linguistic-agnostic conversational agent's voice-based functionality. This will leverage the diverse patient population at AIIMS to create robust models capable of handling various Indian languages, accents, and dialects. The voice-based system will significantly improve acces-

sibility, particularly for patients with limited literacy or technological familiarity.

The final phase will involve comprehensive deployment, integration, and evaluation at AI-IMS. This will include rigorous testing of the system's impact on patient flow, diagnostic efficiency, and physician cognitive load. The evaluation will also assess the accuracy of probable diagnoses and organ-of-origin predictions, with feedback from healthcare professionals being incorporated into system improvements. Long-term studies will measure the system's effectiveness in supporting clinical decision-making while maintaining high standards of patient care.

Bibliography

- [1] T. A. Cohen, V. L. Patel, and E. H. Shortliffe. Introducing ai in medicine. In *Intelligent Systems in Medicine and Health: The Role of AI*, pages 3–20. Springer, 2022.
- [2] R. S. Ledley and L. B. Lusted. Reasoning foundations of medical diagnosis: symbolic logic, probability, and value theory aid our understanding of how physicians reason. *Science*, 130(3366):9–21, 1959.
- [3] A. Rajkomar, J. Dean, and I. Kohane. Machine learning in medicine. *New England Journal of Medicine*, 380(14):1347–1358, 2019.
- [4] S. L. Gans, M. A. Pols, J. Stoker, M. A. Boermeester, and E. S. Group. Guideline for the diagnostic pathway in patients with acute abdominal pain. *Digestive surgery*, 32(1):23–31, 2015.
- [5] C. A. Mealie, R. Ali, and D. E. Manthey. Abdominal Examination [online]. 2024. Last Updated: 2024-05-25. URL: <https://www.ncbi.nlm.nih.gov/books/NBK459220/>.
- [6] S. L. Cartwright and M. P. Knudson. Evaluation of acute abdominal pain in adults. *American family physician*, 77(7):971–978, 2008.
- [7] Differential Diagnosis [online]. Last Updated: 2022-01-27. URL: <https://my.clevelandclinic.org/health/diagnostics/22327-differential-diagnosis>.
- [8] J. A. Osheroff, J. M. Teich, B. Middleton, E. B. Steen, A. Wright, and D. E. Detmer. A roadmap for national action on clinical decision support. *Journal of the American medical informatics association*, 14(2):141–145, 2007.
- [9] D. L. Hunt, R. B. Haynes, S. E. Hanna, and K. Smith. Effects of computer-based clinical decision support systems on physician performance and patient outcomes: a systematic review. *Jama*, 280(15):1339–1346, 1998.
- [10] S. Visweswaran, A. J. King, and G. F. Cooper. Integration of ai for clinical decision support. In *Intelligent Systems in Medicine and Health: The Role of AI*, pages 285–308. Springer, 2022.
- [11] T. Panch, P. Szolovits, and R. Atun. Artificial intelligence, machine learning and health systems. *Journal of global health*, 8(2), 2018.
- [12] T. A. Cohen, V. L. Patel, and E. H. Shortliffe. *Intelligent systems in medicine and health: The role of AI*. Springer, 2022.
- [13] C. J. Haug and J. M. Drazen. Artificial intelligence and machine learning in clinical medicine, 2023. *New England Journal of Medicine*, 388(13):1201–1208, 2023.
- [14] J. Weizenbaum. Eliza—a computer program for the study of natural language communication between man and machine. *Communications of the ACM*, 9(1):36–45, 1966.

- [15] J. Clusmann, F. R. Kolbinger, H. S. Muti, Z. I. Carrero, J.-N. Eckardt, N. G. Laleh, C. M. L. Löffler, S.-C. Schwarzkopf, M. Unger, G. P. Veldhuizen, et al. The future landscape of large language models in medicine. *Communications medicine*, 3(1):141, 2023.
- [16] All India Institute of Medical Sciences (AIIMS). 67th annual report 2022-2023, February 2024. Last updated on: 02 Feb 2024. URL: https://www.aiims.edu/images/pdf/annual_reports/english1.pdf.
- [17] All India Institute of Medical Sciences (AIIMS). AIIMS OUT PATIENT DEPARTMENT (OPD) SERVICES [online]. Last Updated: 2024-05-25. URL: <https://www.aiims.edu/aiims/hosp-serv/citizen-charter/opd-services.htm>.
- [18] E. Shortliffe. *Computer-based medical consultations: MYCIN*, volume 2. Elsevier, 2012.
- [19] E. H. Shortliffe and B. G. Buchanan. A model of inexact reasoning in medicine. *Mathematical biosciences*, 23(3-4):351–379, 1975.
- [20] M. A. Musen, B. Middleton, and R. A. Greenes. Clinical decision-support systems. In *Biomedical informatics: computer applications in health care and biomedicine*, pages 795–840. Springer, 2021.
- [21] D. Demner-Fushman, W. W. Chapman, and C. J. McDonald. What can natural language processing do for clinical decision support? *Journal of biomedical informatics*, 42(5):760–772, 2009.
- [22] H.-P. Eich and C. Ohmann. Internet-based decision-support server for acute abdominal pain. In *Joint European Conference on Artificial Intelligence in Medicine and Medical Decision Making*, pages 103–112. Springer, 1999.
- [23] P. Lee, S. Bubeck, and J. Petro. Benefits, limits, and risks of gpt-4 as an ai chatbot for medicine. *New England Journal of Medicine*, 388(13):1233–1239, 2023.
- [24] A. Karthikesalingam and P. Natarajan. Amie: A research ai system for diagnostic medical reasoning and conversations. *Google Research*, 2024.
- [25] R. Johnson, M. M. Li, A. Noori, O. Queen, and M. Zitnik. Graph artificial intelligence in medicine. *Annual Review of Biomedical Data Science*, 7, 2024.
- [26] K. Miller. Should AI Models Be Explainable? That depends. [online]. 2021. Last Updated: 2021-05-16. URL: <https://hai.stanford.edu/news/should-ai-models-be-explainable-depends>.
- [27] Streamlit [online]. URL: <https://streamlit.io/>.
- [28] A. Q. Jiang, A. Sablayrolles, A. Mensch, C. Bamford, D. S. Chaplot, D. d. l. Casas, F. Bressand, G. Lengyel, G. Lample, L. Saulnier, et al. Mistral 7b. *arXiv preprint arXiv:2310.06825*, 2023.
- [29] A. Dubey, A. Jauhri, A. Pandey, A. Kadian, A. Al-Dahle, A. Letman, A. Mathur, A. Schelten, A. Yang, A. Fan, et al. The llama 3 herd of models. *arXiv preprint arXiv:2407.21783*, 2024.

- [30] G. Team, M. Riviere, S. Pathak, P. G. Sessa, C. Hardin, S. Bhupatiraju, L. Hussenot, T. Mesnard, B. Shahriari, A. Ramé, et al. Gemma 2: Improving open language models at a practical size. *arXiv preprint arXiv:2408.00118*, 2024.
- [31] P. Liu, W. Yuan, J. Fu, Z. Jiang, H. Hayashi, and G. Neubig. Pre-train, prompt, and predict: A systematic survey of prompting methods in natural language processing. *ACM Computing Surveys*, 55(9):1–35, 2023.
- [32] Ollama [online]. URL: <https://ollama.com/>.

Appendix

1 Probable Diagnosis and Organ of Origin

Given below is a table mapping the 19 organs of origin to the 29 probable diagnoses. The Table 1 is structured such that each row corresponds to a organ of origin, and each value in column **Probable Diagnoses** corresponds to a (list of) probable diagnosis(es) that can originate from that organ.

| No. | Organ of Origin | Probable Diagnoses |
|-----|-----------------------------------|--|
| 1. | Gall Bladder, Biliary Ducts | Biliary Pain |
| 2. | Liver & Gall Bladder | Rule Out Hepatobiliary Malignancy |
| 3. | Biliary System | Acute Cholangitis |
| 4. | Liver | 1. Liver Abscess 2. Acute Hepatitis |
| 5. | Pancreas | 1. Acute Pancreatitis 2. Chronic Pancreatitis 3. Rule Out Pancreatic Malignancy |
| 6. | Stomach | Gastritis, Dyspepsia |
| 7. | Heart | Cardiac Pain |
| 8. | Spleen | Spleen Related |
| 9. | Kidney | 1. Renal/Ureteric Calculi 2. Pyelonephritis |
| 10. | Appendix | Appendicitis |
| 11. | Intestine | 1. Diverticulitis 2. Intestinal Obstruction 3. Mesenteric Ischemia |
| 12. | Disorder of Gut Brain Interaction | Disorder of Gut Brain Interaction |
| 13. | Uterus & Ovary | Dysmenorrhoea |
| 14. | Uterus | 1. Pelvic Inflammatory Disease 2. Fibroid 3. Adenomyosis |
| 15. | Adnexa | 1. Ectopic Pregnancy 2. Adnexal Lesion |
| 16. | Urinary bladder | 1. Cystitis 2. Bladder stones |
| 17. | Metabolic | Look for Metabolic Causes: Diabetes, Hyperparathyroidism, Lead Intoxication, Porphyria |
| 18. | Abdominal Wall | Neuropathic Pain (Herpes, etc), Abdominal Wall Hematoma |
| 19. | Pleural Disorders | Pleural Effusion/Pleurodynia |

Table 1: Mapping of Organs of Origin to Probable Diagnoses.

2 Questionnaire

The 17 questions used in the diagnostic questionnaire are listed below along with the possible answers. The Table 2 is structured such that each row corresponds to a question, and each value in column **Possible Answers** corresponds to a (list of) possible answer(s) to that question.

The Figure 1 below shows the 9 regions of the abdomen and 4 patterns of pain that are referred to in the questionnaire.

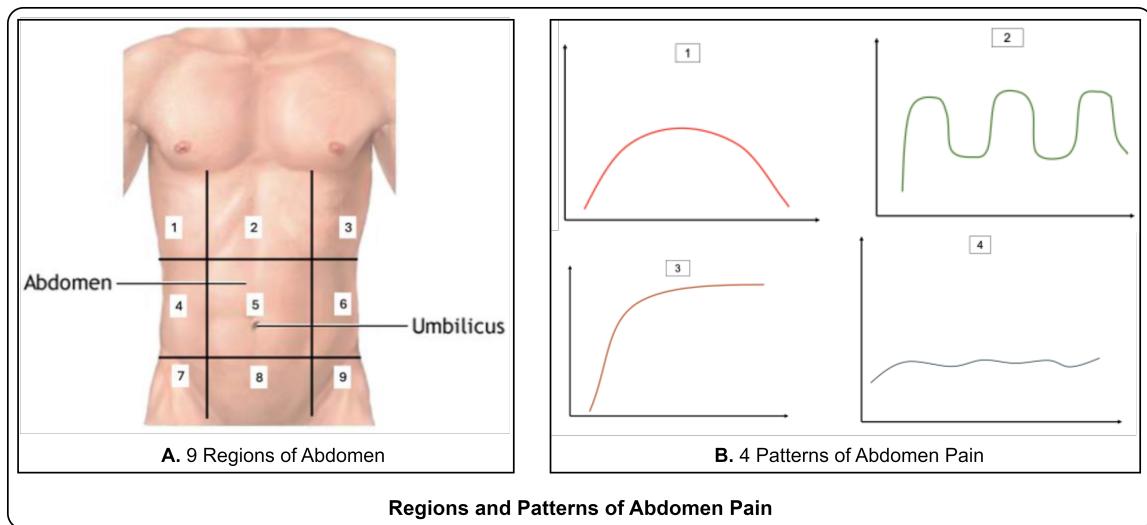


Figure 1: Regions of the Abdomen and Patterns of Pain.

| No. | Question | Possible Answers (seperated by semi-colon) |
|-----|--|---|
| 1 | Please rate the severity of the pain on a scale of 1 to 10 | 1-3; 4-7; 8-10 |
| 2 | Have you experienced any Trauma? | Yes; No |
| 3 | Are there any danger signs present? | Light Headedness; Altered Sensorium; Respiratory Distress; None |
| 4 | Choose your gender | Male; Female |
| 5 | Choose your age group | 0-15; 15-25; 25-45; 45-60; 60+ |
| 6 | Where does the pain occur in the abdomen? | <Image of 9 regions of Abdomen> |
| 7 | How did the pain start? | Over Minutes to Hours (Acute); Over Hours to Days (Insidious) |
| 8 | What is the character of your pain? | Burning Pain; Stabbing Pain; Pin Pricking Pain; Constricting Pain; Throbbing Pain; Dull aching/non-specific Pain |
| 9 | What's the pattern of your pain? | <Image of 4 patterns of pain> |
| 10 | How long have you been experiencing the pain? | Less than 3 months; More than 3 months |
| 11 | Does the pain radiate to any other areas? | No Radiation; To Back; To Shoulder; To Groin/Inner Thigh; To Arms/Neck |
| 12 | Does the pain increase or decrease with the following aggravating/relieving factors? | No Aggravating/Relieving Factors; With Food Intake; Bending Forward; Passing Stool; Passing Urine; Menstruation; Bending Sideways; Deep Inspiration; Walking and Exercise |
| 13 | Are there any associated symptoms with your pain? Please specify if you experience none, or if you have any of the following symptoms. | None; Lump; Fever or Chills; Nausea and/or Vomiting; Abdominal Bloating; Constipation; Diarrhoea; Blood in Stools/Black Stools; Jaundice; Burning Micturition; Blood in Urine; Weight Loss; Loss of Appetite; Stress, Anxiety, Depression, Palpitation; Shortness of Breath; Swelling in Neck, Axilla; Skin Changes Over Abdominal Wall |
| 14 | Do you have any comorbidities ? | None; Diabetes; Heart Disease; Kidney Disease; Gall Stones |
| 15 | Do you have a history of any previous surgeries? | None; Gall Bladder; Intestine; Kidney; Uterus |
| 16 | Have you experienced any recent changes in your menstrual cycle? | Changes in Periods; Absence of Periods; None |
| 17 | Have you noticed any of these abnormalities? | Abnormal Vaginal Bleeding; Foul Smelling Discharge; None |

Table 2: Questionnaire for the Diagnostic Protocol.

3 Frontend Screenshots for the Implemented CDSS

The following are the screenshots of the frontend of the implemented CDSS. The screenshots are taken from the web application that is developed for the CDSS. The website for the CDSS is hosted at <https://kutumlab.github.io/abdominal-pain-cdss/>.

Android Application: Four screenshots of the Android application are shown in the Figure 2. The screenshots show the question and their respective options.

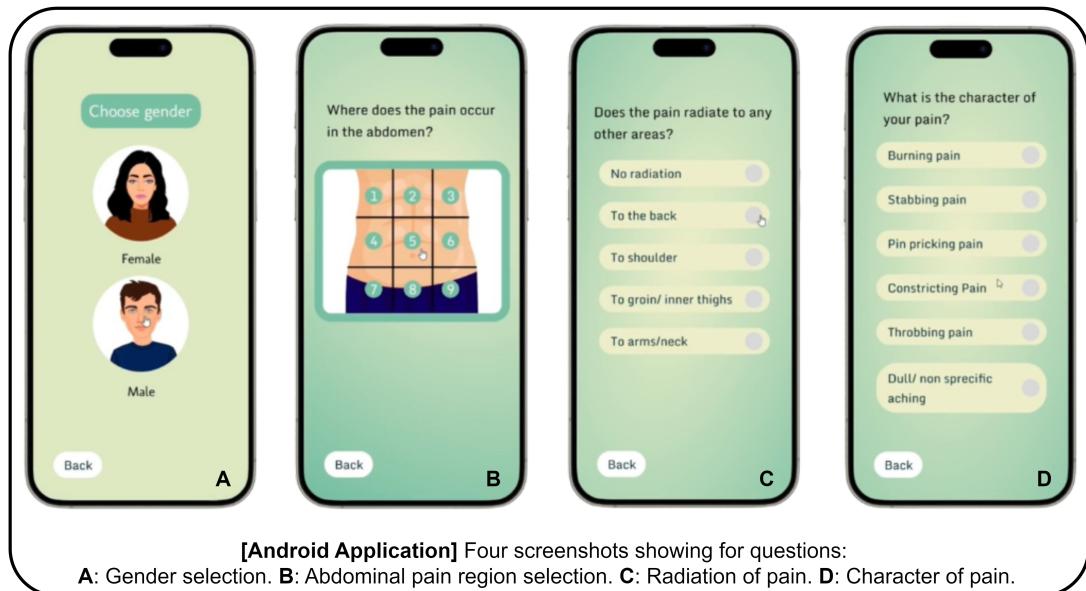


Figure 2: Screenshots of the Android Application.

Web Application: The screenshots of the web application are shown in the Figure 3. The screenshots for the same questions and their respective options as in the Android application are shown.

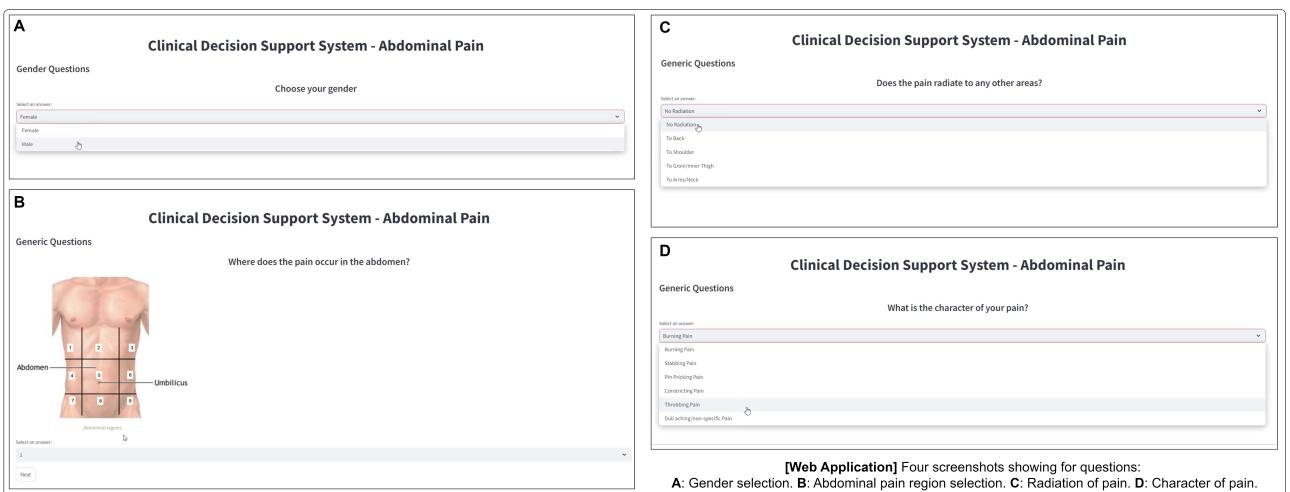


Figure 3: Screenshots of the Web Application.

Result Page: The result page of the web application is shown in the Figure 4. The result page shows two screenshots, one with a successful probable and organ of origin diagnosis, and the other with a prompt to visit emergency department, triggered by the presence of danger signs in discriminator section of the questionnaire.

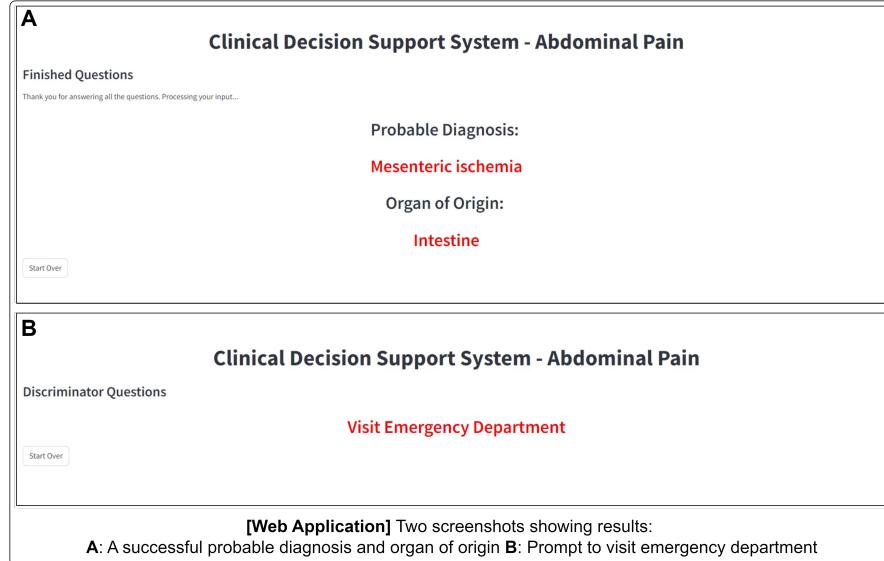


Figure 4: Result Page of the Web Application.