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 fullfillment of the requirements of the degree of

Postgraduate Diploma in Advanced Studies and Research (DipASR)

in Computer Scinece 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 09, 2024 and accepted by the Department of Computer Science on December 09, 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 09, 2024

Gautam Ahuja

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

Dr. Rintu Kutum
Department of Computer Science
Ashoka University

Contents

List of Figures

List of Tables

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 (AI) study [1]. From the very inception of reasoning foundations in medical diagnosis and Clinical Decision Support (CDS) 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 (AIIMS), New Delhi, aims to develop a department-specific Clinical Decision Support System (CDSS) tailored to address the unique needs in the Outpatient Department (OPD) 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 2.2 (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.

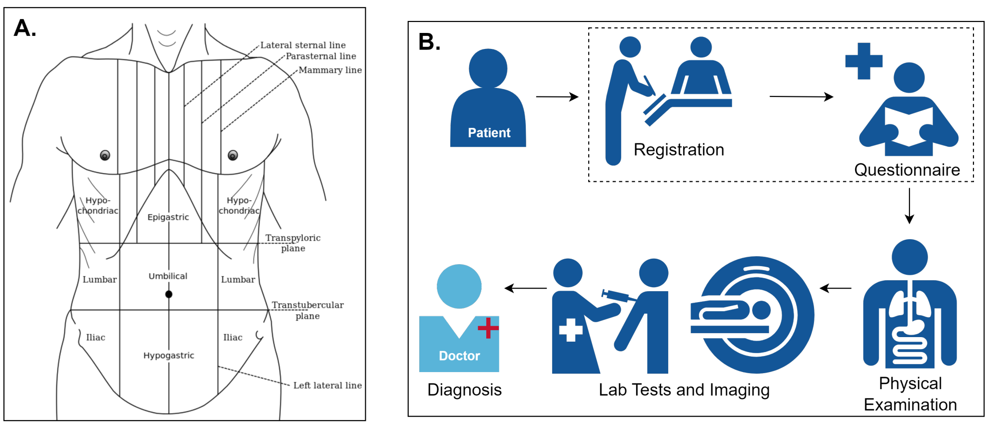
- **Surgical History**: Prior surgical procedures on the gallbladder, intestines, kidneys, or uterus.

- **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 2.2 (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.



A. Nine Anatomical Regions of the Abdomen.
B. Workflow of Abdominal Examination (in non emergency). The Dotted Box is the Step Where the CDSS is Integrated.

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

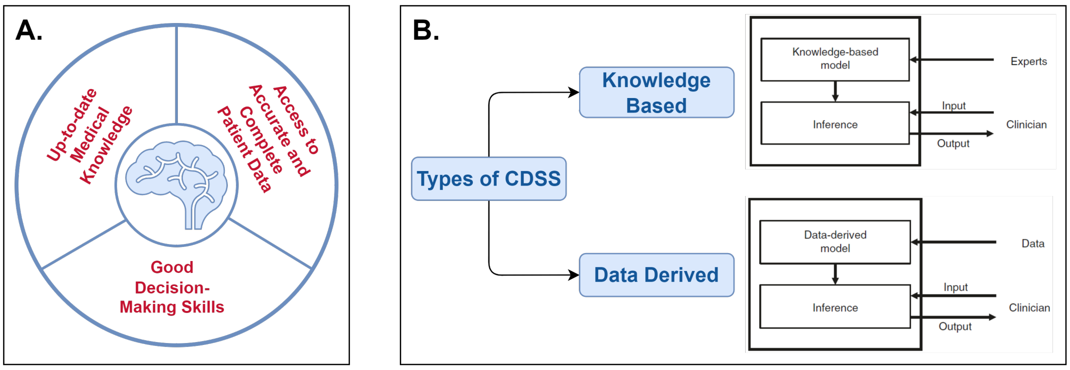
- Clinical Data Capture and Documentation Burden

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..



A. Effective Decision Making principles.
B. Different Types of Clinical Decision Support System (CDSS) as per Visweswaran et. al.

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

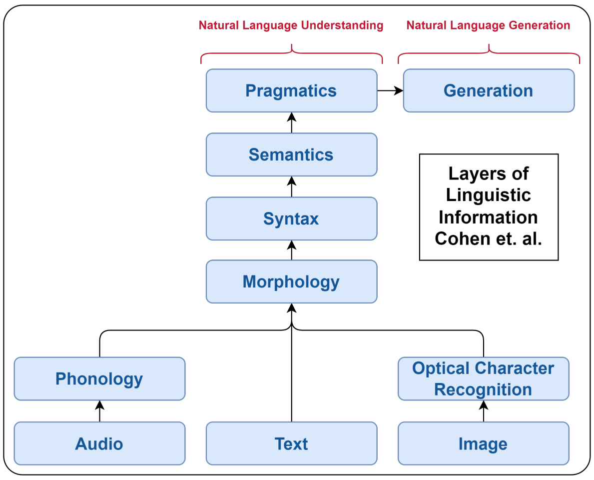


Figure 2.3: Linguistic Levels

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 2022-2023 Annual Report [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

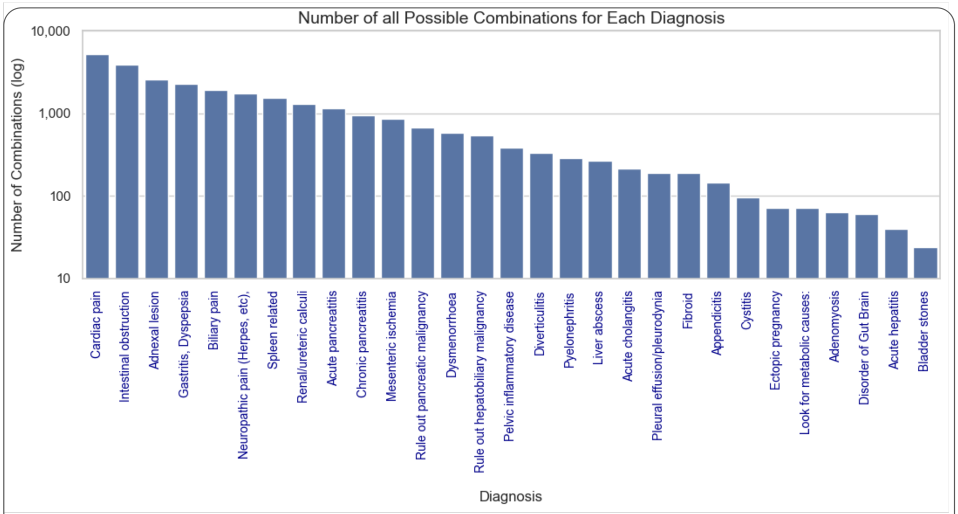
The Routine Gastroenterology OPD operates from Monday to Friday, 8:30 a.m. to 1:00 p.m. [17]. Given this limited time frame of 270 minutes daily over 5 working days, approximately **8,500 new cases are handled per day**, with multiple physicians addressing various chief complaints.

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 ¡ref¿, 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.



The number of possible combination of answers for each diagnosis. The total number of questions asked are 17. Based on a set of possible answer, a probable diagnosis is assigned.

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 pie chart below 2.5 [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.

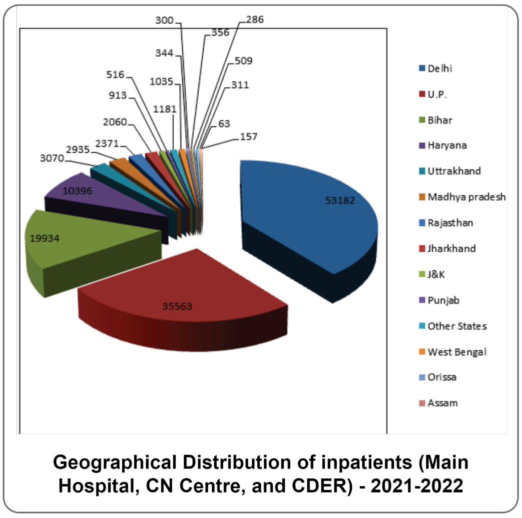


Figure 2.5: Inpatient Distribution

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 decision-support server [22]. 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. 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, multimodal 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 the 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.

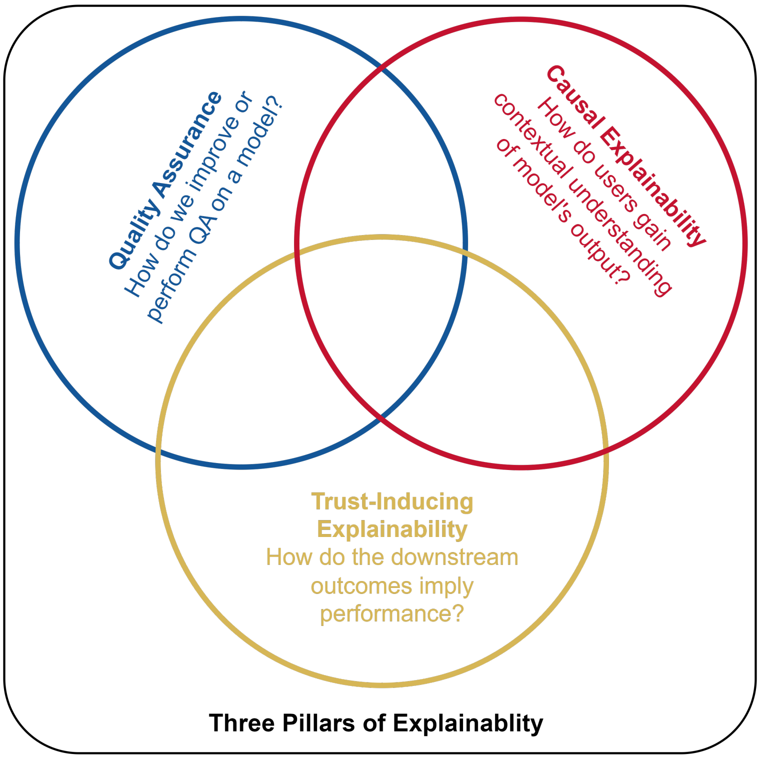


Figure 3.1: Three pillars of explainable AI

4 Content

4.1 Problem Statement / Objectives

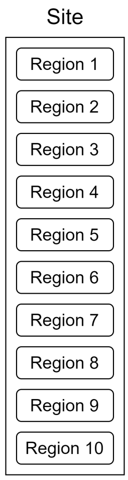


Figure 4.1: Example Image

4.2 Scope, Methodology, and Design

4.3 Work Done

5 Results & Discussions

6 Conclusion & Future Work

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`.

Appendix