AI-Driven Chest Radiograph Analysis System

Intelligent Reasoning Systems Practice Module Project Proposal

Group 8

A0327973X Gu ShuoHan
A0329069U Pan Linyu
A0329955L Wang Siwei
A0329709R Yang Minyue
A0326794N Yang Ziyu

1. Project Overview

1.1 Project Background

Lung diseases are causing a global health emergency. In 2023 alone, tuberculosis (TB) reached 10.8 million new cases worldwide—the highest number ever recorded by the World Health Organization. Beyond TB, pneumonia kills over 2.5 million people each year. These diseases disproportionately affect developing countries, where access to proper medical care is often limited.

While chest X-rays are the primary tool doctors use to detect these diseases , accurate analysis is extremely difficult. Radiologists spend years learning to spot disease patterns, identify subtle changes, and differentiate between conditions that look very similar. The diagnostic process itself requires four key steps: analyzing lung areas, identifying small patterns, separating similar-looking diseases, and connecting findings with patient symptoms. Each step demands specialized expertise that can take decades to master. A tiny shadow could be pneumonia, TB, or cancer, and a missed diagnosis could cost a life.

The current chest X-ray diagnostic system faces three major problems:

- Specialist Shortage: There are not enough trained radiologists globally. The U.S. alone expects a shortage of over 30,000 radiologists by 2030, which causes delays of 24-48 hours in developed countries and leaves millions without access to radiology services in developing regions.
- Geographic Barriers: Rural and remote areas often lack radiologists, requiring images to be sent to distant cities, which adds further delays.
- Communication Problems: A significant number of chest X-ray abnormalities are initially missed or misinterpreted, and most patients do not understand their results even when they receive them.

This project aims to solve these critical issues by delivering expert-level AI-driven analysis where it is needed most, converting chest X-rays into clinical reports with a focus on early and accurate detection.

1.2 Project Introduction

Our AI-Driven Chest Radiograph Analysis System is an automated diagnostic platform that converts chest X-rays into clinical reports using integrated ML models, medical knowledge graphs, and large language models. The system addresses critical gaps in global diagnostic accessibility by delivering expert-level analysis without requiring on-site radiological expertise.

The system's workflow is designed for speed and efficiency. When healthcare providers upload a chest X-ray through our intuitive interface, the system immediately begins an automated analysis and delivers results within minutes, rather than the typical 24-48 hour delays. This rapid diagnosis capability enables faster clinical decision-making and patient care, which is especially critical when treating infectious diseases like tuberculosis or pneumonia. The system also generates standardized reports that follow established medical protocols, ensuring consistency in documentation and reducing the workload for healthcare professionals. Furthermore, its cloud-based deployment and 24/7 availability eliminate human resource constraints, making it valuable for emergency departments, remote clinics, and facilities across different time zones..

What sets our system apart is its dual-output design. For healthcare providers, it delivers detailed diagnostic reports with technical findings. For patients, it creates visual explanations that highlight key findings on the X-ray images using simple annotations, helping them understand their condition. By offering both outputs, our system directly tackles major challenges: access barriers are eliminated through cloud deployment, time delays are reduced from days to minutes, and communication gaps are bridged with patient-friendly explanations.

1.3 Proven Demand in Growing Market

Our AI-driven chest radiograph analysis system enters a rapidly expanding market with proven demand and strong financial returns. The AI medical imaging market demonstrates robust growth potential, expanding from $5.86 billion in 2024 to an expected $20.40 billion by 2029, while other sources report varying figures ranging from $1.44 billion in 2024 to $1.65 billion in 2024 projected to reach $4.54 billion by 2029. Despite differences in exact figures, all reports confirm a compound annual growth rate (CAGR) of 30% or higher, reflecting urgent global demand for automated diagnostic solutions.

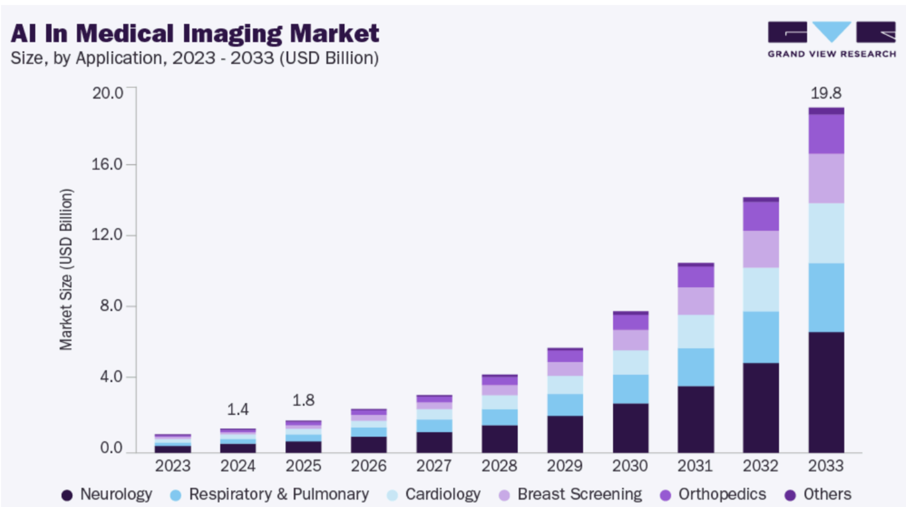


Figure 1.3 AI In Medical Imaging Market Size chart

The financial opportunity is compelling, with healthcare AI investments delivering an average return of $3.20 for every dollar invested, with ROI realized within 14 months according to a March 2024 Microsoft-IDC study. This strong ROI is driven by significant cost savings from reduced diagnostic delays and improved operational efficiency. Market readiness is demonstrated by the fact that 79% of healthcare organizations currently use AI technology, indicating widespread acceptance and adoption readiness.

Our system directly targets the 4+ billion underserved people who currently lack access to adequate radiological services, representing a massive addressable market that traditional healthcare delivery models cannot reach cost-effectively. The clinical need has never been more urgent, with the WHO Global TB Report 2024 emphasizing critical need for improved diagnostic tools and post-COVID recovery efforts highlighting healthcare system vulnerabilities. Additionally, close to 30% of players in the AI medical imaging market already offer software solutions for processing X-ray images, validating our specific focus area while creating opportunities for differentiation through our dual-output patient communication approach.

2.Literature Review

2.1 CheXpert

The CheXpert dataset, developed by a team from the Massachusetts Institute of Technology (MIT), stands as a highly influential large-scale resource in the chest X-ray imaging domain. It encompasses 224,316 chest X-rays from 65,240 patients. A core contribution of this dataset is the introduction of an uncertainty label mechanism, along with comparative studies against expert diagnoses.

For label annotation, the research team employed a rule-based automatic labeling approach, enabling efficient processing of large-scale image data. Critically, this method assigns uncertainty attributes to labels well-aligned with real-world clinical practice, where radiologists' assessments of some imaging findings are not always definitive. In model training and uncertainty management, the DenseNet-121 Convolutional Neural Network (CNN) was utilized to predict 14 chest pathology categories. To tackle label uncertainty, four technical strategies were proposed: U-Ignore (disregarding uncertain labels), U-Zeros/U-Ones (assigning uncertain labels to 0 or 1 respectively), U-SelfTrained (a semi-supervised learning strategy leveraging uncertain labels for model training), and U-MultiClass (a multi-classification strategy optimizing model output to fit uncertainty-laden label distributions).

In terms of intelligent reasoning techniques, CheXpert-related research spans several key areas: at the computer vision level, it enables the detection of pathological signs in chest X-rays; via the U-SelfTrained semi-supervised learning strategy, it effectively addresses training data with uncertain labels; the U-MultiClass multi-classification strategy refines model output, boosting adaptability to complex pathological classifications; and probabilistic reasoning is used to calibrate predictions, enhancing the reliability of model outputs.

The CheXpert dataset and its associated findings provide a vital foundation for subsequent work. For our research, we build upon this foundation in two primary ways: first, we aim to reproduce and optimize the research results, leveraging the dataset's rich resources and uncertainty-handling methodologies; second, we extend the application scope by connecting to report generation tasks, advancing chest radiological image analysis from single pathological detection toward more comprehensive clinical workflows.

2.2 MLLMs in Radiology Report Generation

In the realm of medical case analysis, Multimodal Large Language Models (MLLMs) have emerged as a promising avenue, integrating multimodal data (such as medical images, textual case records, and clinical metrics) to facilitate comprehensive diagnostic and analytical processes. Existing research can be categorized into several technical paradigms, each with distinct characteristics in addressing clinical challenges.

(1) Single-modal Image

Analogous to the “Pure CT image” strategy in radiology report generation, some early MLLM applications in medical case analysis relied heavily on single - modality medical imaging data (e.g., CT scans, MRI images). These approaches employed traditional computer vision techniques like Convolutional Neural Networks (CNNs) and DenseNets to extract pathological features from images. The core advantage lies in achieving pathological localization in scenarios where textual annotations of medical cases are scarce or unavailable. For instance, in tumor - related case analysis, such models can segment tumor regions from CT images with high precision. However, a significant limitation is that the outputs, often in the form of segmentation masks or pathological probability scores, lack clinical contextual support and they fail to integrate essential textual information from medical cases, such as patients' medical histories, symptoms, and prior diagnostic records, thus limiting the clinical interpretability of the results.

(2) Image - Text Fusion Paradigms

Building upon the “Combination of image - text” methodology, MLLMs for medical case analysis leverage multimodal fusion techniques to associate visual features from medical images with semantic information from textual case descriptions. This integration aims to resolve the “image - text misalignment” issue prevalent in single - modal approaches. By fusing image features (e.g., pathological manifestations in X - rays) with textual data (e.g., clinical findings in case reports), these models enhance the clinical relevance of case analysis. They can generate more contextually appropriate interpretations, such as correlating an observed lung opacity in an X - ray with corresponding symptoms described in the case text. Nevertheless, these models rely heavily on high - quality image - text paired data from medical cases. Moreover, despite the fusion, some models still suffer from hallucinations, where the generated analyses may include non - existent pathological findings or incorrect correlations between images and text.

(3) Step-by-Step Reasoning Driven Methods

Inspired by the “Step - by - step Reasoning based method”, certain MLLMs are designed to emulate the step - by - step diagnostic reasoning process of clinicians when analyzing medical cases. These models first mine step - by - step reasoning chains from expert - annotated medical cases. Then, they employ supervised learning to imbue the model with reasoning logic and reinforcement learning to optimize the factuality of reasoning. This paradigm addresses the defect of many existing models that “only provide results without transparent reasoning”, thereby enhancing the interpretability of case analysis. For example, when diagnosing a complex cardiovascular case, the model can outline steps like "first assess heart size from imaging → then check vascular calcification → finally correlate with symptom descriptions in the case text". However, this approach depends on high - quality, expert - curated clinical case reports for mining reasoning chains, and the training process involving both supervised and reinforcement learning is relatively complex, increasing the technical barrier for implementation.

(4) Multi-Agent Collaborative Solutions

Drawing from the "Multi - agent solution" framework, some MLLM systems adopt a modular agent collaboration approach for medical case analysis. The process typically involves agents for retrieval (to find similar historical cases), drafting (to generate initial case analyses), refinement (to extract key clinical findings), vision (to analyze medical images), and synthesis (to integrate all information into a comprehensive report). This modular collaboration resolves the local information loss issue that occurs due to global optimization in single - agent models, as each agent specializes in a specific aspect of case analysis. However, this paradigm requires coordinating compatibility between different agent modules. Additionally, the retrieval agent, which sources similar cases, is prone to introducing noise if the retrieved historical cases are of low quality or irrelevant, impacting the overall accuracy of case analysis.

Building upon these advancements, our research intends to implement an image - text fusion strategy to ensure that image features extracted from medical cases are accurately aligned with real world pathological findings described in textual case records. Furthermore, we will leverage domain knowledge to enable step - by - step reasoning during medical case analysis, aiming for more explainable and accurate diagnostic and analytical outcomes.

3.Project Scope

3.1 Project Definition

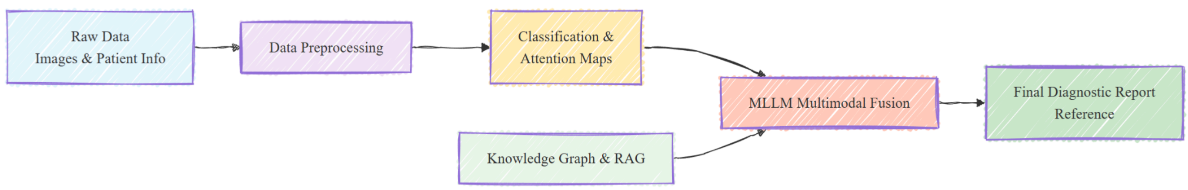


Figure 3.1 Project architecture flow chart

This project aims to develop a multi-modal AI-powered clinical decision support system. The system will leverage patient chest X-ray images and demographic information (e.g., age, gender, BMI) to automatically generate preliminary diagnostic reports and treatment plan suggestions. Our goal is to ease the workload of radiologists and clinicians and provide robust AI assistance for their decision-making.

- **Academic Aspect:** We will explore a novel method that combines traditional explainable AI (XAI) techniques with cutting-edge multi-modal large language models (MLLMs). By integrating attention maps from our classification model, we aim to improve the accuracy and

interpretability of the multi-modal model's output, pushing the boundaries of AI applications in specific, high-stakes domains.
- **Market Aspect:** The system can serve as an invaluable diagnostic aid, integrating seamlessly into existing hospital Picture Archiving and Communication Systems (PACS). In the future, this system has the potential to shorten diagnosis times and improve diagnostic consistency, especially in remote areas or regions with limited medical resources, by offering expert-level support to primary care facilities.

3.2 Key Technical Focus

This project focuses on the following intelligent reasoning systems and techniques:

- **Explainable AI (XAI):** We will first train a classification model on the CheXpert dataset to identify various thoracic diseases. More critically, we will focus on extracting attention maps from the model's decision-making process. These maps will act as a bridge, visually conveying the model's "points of focus" to the subsequent generation model. This enhances the overall transparency and trustworthiness of the system.
- **Multi-modal Large Language Models (MLLM) and Retrieval-Augmented Generation (RAG)**: We will feed the raw images, attention maps, classification results, and patient metadata as a multi-modal input to a large language model. To ensure the medical accuracy of the generated reports, we will specifically integrate Retrieval-Augmented Generation (RAG). We plan to utilize a pre-built medical knowledge graph to retrieve and inject reliable medical knowledge in real-time. This helps to combat model "hallucinations" and ensures the rigor of the generated content.

3.3 Limitations & Constraints

Academic Research Limitations:

- **Data Dependency:** While the CheXpert dataset is large, its annotations may contain inconsistencies or limitations. The model's ability to generalize to new or rare diseases it hasn't seen before may be challenged.
- **Interpretability Bottleneck:** While attention maps provide a basis for the decision, they are merely approximations of the model's focus, not a direct reflection of human clinical reasoning. Quantifying and evaluating the effectiveness of this "explanation" remains a significant academic challenge.

Market Applicability Limitations:

- **Legal and Regulatory Compliance:** Any medical AI product must undergo rigorous regulatory scrutiny, such as HIPAA in the United States. Data privacy and security are core challenges, requiring us to ensure that patient data is protected at the highest level throughout the entire process.
- **Liability Issues:** If the AI system provides an incorrect diagnostic suggestion, how is liability determined? This involves complex legal, ethical, and professional issues regarding the doctor-AI collaboration model. Our system is designed purely as an assistive tool, and the final diagnostic and decision-making authority must always remain with the human physician.

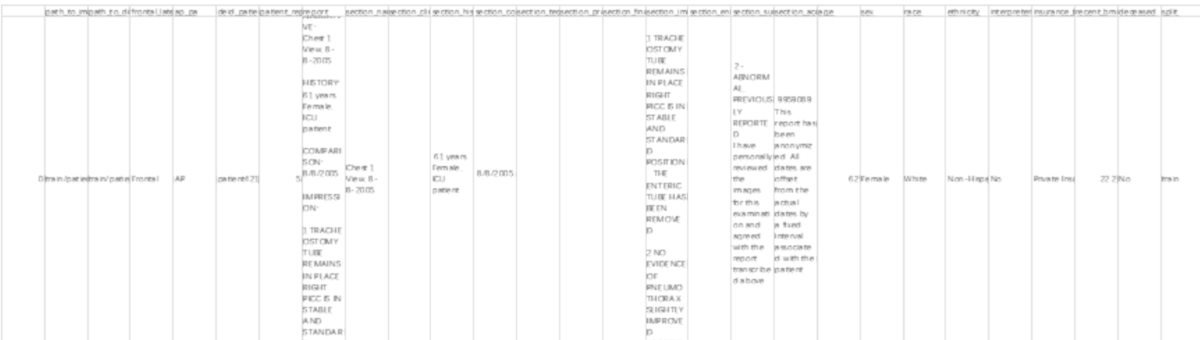
- **User Adoption:** We need to account for the slow adoption rate of new technologies in clinical practice. The system must prove its reliability, usability, and efficiency gains in a real-world setting to earn the trust and acceptance of doctors and medical institutions.

4.Data Collection and Preparation

4.1 Core dataset

CheXpert Plus is a large-scale dataset that pairs chest X-ray images with their corresponding radiology reports, designed to advance research in medical AI and multimodal learning. It comprises a total of 223,228 meticulously curated report-image pairs, derived from 187,711 independent studies involving 64,725 unique patients. This ensures considerable demographic diversity and clinical representativeness.

The dataset includes annotations for a wide range of thoracic conditions and findings. Common labels typically encompass pathologies such as Atelectasis, Cardiomegaly, Consolidation, Edema, Pleural Effusion, as well as support devices like Support Devices. Additionally, it often features observations such as Lung Lesion, Lung Opacity, Pneumonia, Pneumothorax, and Fracture. These structured annotations facilitate a variety of supervised learning tasks, including disease classification, localization, and automated radiology report generation. With its large scale, high-quality pairing, and rich structured labels, CheXpert Plus serves as a critical resource for developing and evaluating models in medical image analysis, natural language processing, and joint vision-language applications.



	path_to_im	path_to_di	frontal_lat	ap_pa	deid_patie	patient_re	report	section_na	section_cli	section_his	section_co	section_te	section_pr	section_fin	section_im	section_en	section_su	section_ac	age	sex	race	ethnicity	interpreter	insurance_	recent_bmi	deceased	split
0	train/patie	train/patie	Frontal	AP	patient21	5	VE: Chest 1 View: 8-8-2005 HISTORY: 61 years Female ICU patient COMPARISON: 8/8/2005 IMPRESSION: 1 TRACHEOSTOMY TUBE REMAINS IN PLACE RIGHT PICC IS IN STABLE AND STANDAR	Chest 1 View: 8-8-2005		61 years Female, ICU patient	8/8/2005				1 TRACHEOSTOMY TUBE REMAINS IN PLACE RIGHT PICC IS IN STABLE AND STANDARD POSITION. THE ENTERIC TUBE HAS BEEN REMOVED. 2 NO EVIDENCE OF PNEUMOTHORAX SLIGHTLY IMPROVED		2 - ABNORMAL PREVIOUSLY REPORTED I have personally reviewed the images for this examination and agreed with the report transcribed above	9858089 This report has been anonymized. All dates are offset from the actual dates by a fixed interval associated with the patient	62	Female	White	Non-Hispa	No	Private Insu	22.2	No	train

Figure 4. 1 Partial data of the Chexpert-plus dataset

4.2 Dataset labeling

The CheXpert Plus dataset was preprocessed and partitioned into structured sections using an automated label extraction system, as established in prior research. This system leverages natural language processing (NLP) techniques, informed by methodologies similar to those described in Irvin et al. ("CheXpert: A Large Chest Radiograph Dataset with Uncertainty Labels and Expert Comparison," 2019) and enhanced by subsequent models like CheXbert (Smit et al., 2020), to identify disease mentions from free-text radiology reports with high precision. The accuracy of the automated labels has been rigorously validated by a panel of board-certified radiologists, ensuring the clinical reliability and annotation consistency of the entire dataset.

The dataset comprehensively covers 14 clinically significant cardiopulmonary conditions. These labels include Atelectasis, Cardiomegaly, Consolidation, Edema, Pleural Effusion, Pneumonia, Pneumothorax, Lung Lesion, Lung Opacity, Fracture, and the presence of Support Devices, among others. This rich set of annotations provides a foundational resource for training and evaluating deep learning models in various medical artificial intelligence tasks, including thoracic abnormality classification, lesion localization, and automated radiology report generation.

A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S
Path	Sex	Age	Frontal/Lat	AP/PA	No Finding	Enlarged C	Cardiomeg	Lung Opac	Lung Lesio	Edema	Consolidat	Pneumoni	Atelectasis	Pneumoth	Pleural Eff	Pleural Oth	Fracture	Support Devices
CheXpert-v1.0-small/train/patient00001/study1/view1_frontal.jpg	Female	68	Frontal	AP	1									0				1
CheXpert-v1.0-small/train/patient00002/study2/view1_frontal.jpg	Female	87	Frontal	AP			-1	1		-1	-1		-1		-1		1	
CheXpert-v1.0-small/train/patient00002/study1/view1_frontal.jpg	Female	83	Frontal	AP				1			-1						1	
CheXpert-v1.0-small/train/patient00002/study1/view2_lateral.jpg	Female	83	Lateral					1			-1						1	
CheXpert-v1.0-small/train/patient00003/study1/view1_frontal.jpg	Male	41	Frontal	AP						1				0				
CheXpert-v1.0-small/train/patient00004/study1/view1_frontal.jpg	Female	20	Frontal	PA	1	0					0				0			
CheXpert-v1.0-small/train/patient00004/study1/view2_lateral.jpg	Female	20	Lateral		1	0					0				0			
CheXpert-v1.0-small/train/patient00005/study1/view1_frontal.jpg	Male	33	Frontal	PA	1		0				0				0			1
CheXpert-v1.0-small/train/patient00005/study1/view2_lateral.jpg	Male	33	Lateral		1		0				0				0			1

Figure 4. 2 Example of Label Dataset

4.3 Data Processing and Utilization Pipeline

The flowchart below delineates the comprehensive data processing pipeline employed in this study, which bifurcates into two distinct pathways to leverage the CheXpert-Plus dataset for different computational objectives.

(1) Classification Task Pipeline (Left Branch):

This branch is designed for developing supervised image classification models. It commences with the Raw Data Sources. The relevant data is then structured into a Labeled Dataset comprising approximately 220,000 entries (e.g., image-label pairs). To mitigate potential class imbalance and ensure a representative sample for model generalization, Stratified Sampling is performed on this dataset. This critical step yields a Balanced, Sampled Dataset. The unique identifiers (IDs) from this curated sample are subsequently used to retrieve the corresponding images and labels, forming the final dataset specifically allocated for Classification Model Training. The ultimate goal of this branch is Model Application for diagnostic prediction on new, unseen data.

(2) Knowledge Graph & LLM Pipeline (Right Branch):

This parallel branch is architected to support knowledge-driven and natural language processing tasks. It directly utilizes the CheXpert-Plus Dataset, which contains paired radiology reports and images. The process involves Filtering Corresponding Reports & Images based on specific criteria relevant to the downstream tasks. The filtered textual reports are then utilized for KG Construction, building a structured representation of medical knowledge embedded in the radiology literature. Concurrently, these reports serve as a high-quality corpus for LLM Tasks, such as fine-tuning large language models for report generation, summarization, or question-answering in the cardiopulmonary domain.

In summary, this integrated pipeline facilitates a dual approach: one for data-driven discriminative modeling via supervised classification, and another for knowledge-intensive generative and reasoning tasks leveraging modern NLP techniques.

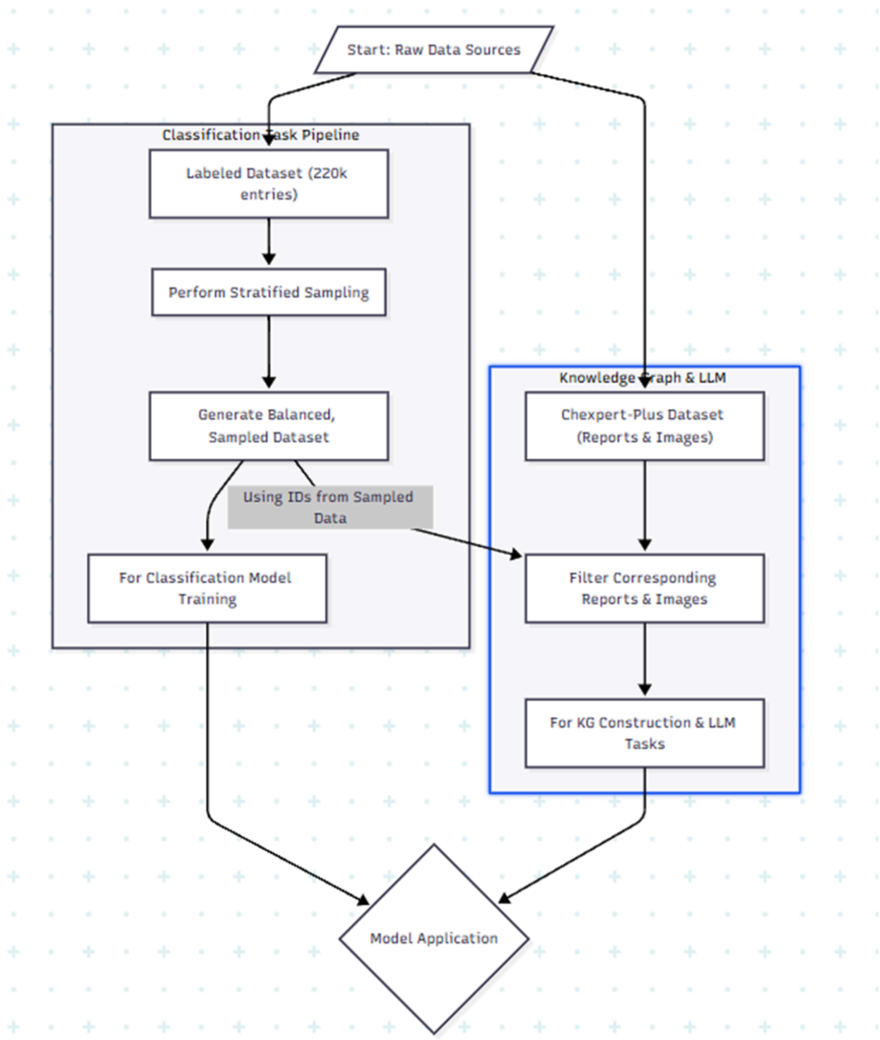


Figure 4. 3 Data processing flowchart

4.4 Challenges in Data processing

Challenge 1: Data Quality and Noise in Textual Reports

The initial raw textual data extracted from radiology reports presents significant challenges for computational analysis. These reports are inherently unstructured, containing numerous sources of noise. These include non-standard abbreviations, spelling variations, typographical errors, and inconsistent formatting. A critical prerequisite for processing is the identification and removal of Protected Health Information (PHI) to ensure patient confidentiality and comply with regulations such as HIPAA. Furthermore, the narrative style of reports often lacks a machine-parsable structure, making it difficult to isolate specific clinical findings, impressions, or recommendations automatically. This noise and heterogeneity can severely degrade the performance of downstream Natural Language Processing (NLP) models if not adequately addressed.

Mitigation Strategy: Robust Text Preprocessing Pipeline

To remediate these issues, we implemented a comprehensive text preprocessing pipeline. The initial step involves the automated redaction of Personally Identifiable Information (PII) using named entity recognition models tuned for clinical text. Subsequently, we perform spelling correction and normalization using specialized medical dictionaries. A key component is abbreviation expansion, where domain-specific acronyms (e.g., "PTX" for pneumothorax) are mapped to their canonical forms to resolve ambiguity. Finally, section parsing is applied to segment reports into structured fields (e.g., "Findings," "Impression"), which facilitates the targeted extraction of clinically relevant information and improves the efficiency of subsequent modeling tasks.

Challenge 2: Complexity in Knowledge Graph Construction

The transformation of unstructured medical text into a structured Knowledge Graph (KG) is a complex endeavor. The primary difficulty lies in the accurate and granular extraction of medical entities (e.g., diseases, anatomy, devices) and the precise semantic relationships that exist between them (e.g., "located_in," "suggests," "is_absence_of"). Clinical language is highly contextual, often expresses uncertainty, and relies on implicit knowledge, making automated information extraction prone to error. The construction of a high-fidelity KG requires not only recognizing entities but also disambiguating them and inferring their correct interrelations, which is a non-trivial NLP task.

Mitigation Strategy: Advanced Entity and Relation Extraction Framework

Our approach to KG construction leverages state-of-the-art Natural Language Processing techniques. We employ a domain-specific pre-trained language model, namely BioBERT, which is pre-trained on large-scale biomedical corpora, as our foundational encoder. This provides a robust contextualized understanding of medical terminology. To further enhance performance and adapt to the specific nuances of our radiology report dataset, we fine-tune this model on a carefully annotated subset of our data. This fine-tuning process is specifically designed for the joint task of named entity recognition (NER) and relation extraction (RE), significantly improving the accuracy of both identifying medical concepts and classifying the relationships between them, thereby enabling the automatic population of a high-quality knowledge graph.

5.System Design

5.1 Label Prediction

The CheXpert dataset provides image–label pairs that allow for direct supervised training, and models such as DenseNet121 combined with the AUC margin loss have established strong baselines in this setting. This supervised approach is effective because it explicitly optimizes performance against curated labels, yielding robust diagnostic accuracy across medical imaging tasks. On the other hand, CheXZero introduces a complementary methodology inspired by CLIP, but specifically adapted for the clinical domain. Instead of relying on predefined labels, it employs a ViT-B/16 vision backbone in combination with BioClinicalBERT as the text encoder, and the two modalities are aligned through contrastive loss. This design allows the system to learn from unstructured clinical text and imaging data, enabling zero-shot and multimodal inference without requiring direct supervision. Taken together, the supervised DenseNet-based methods from CheXpert and the multimodal ViT-B/16 approach from CheXZero highlight two complementary strategies: one grounded in robust labeled datasets, and the other leveraging flexible, label-free learning from clinical narratives and images. Integrating insights from both could lead to more powerful and generalizable medical AI systems.

5. 2 Report Generation:Knowledge Integration And Reasoning

The project leverages a locally deployed multimodal large language model (MLLM) powered by Ollama, which serves as the core reasoning engine for medical image analysis and report generation. To enhance the model's interpretability and domain specificity, it is integrated with a locally constructed medical knowledge graph, enabling the incorporation of prior medical knowledge and supporting structured reasoning over clinical concepts. This hybrid design ensures that the model does not simply rely on pattern recognition from raw data, but can also ground its outputs in established medical relationships and guidelines. In addition, the system employs the LangChain API as a middleware layer, allowing for flexible query handling and the orchestration of external reasoning workflows. Through LangChain, the model can connect seamlessly with auxiliary tools, data sources, and other large-scale language models, thereby extending its reasoning capabilities beyond the core deployment and ensuring adaptability for diverse diagnostic and decision-support scenarios.

5.3 Front-end And Back-end Interaction Design Of The Website

In the proposed online medical X-ray diagnostic system, once the uploaded image is processed by the AI model, the system automatically generates a structured diagnosis report containing both predicted findings and interpretive notes. This report is then delivered back to the user through a secure web-based interface, ensuring that patients and clinicians can conveniently view the diagnostic outcome in real time. At the same time, the report is stored in a dedicated history record database, which not only preserves the results for future clinical reference but also enables longitudinal tracking of patient conditions. Such a design facilitates continuous monitoring, comparative analysis across multiple visits, and supports research efforts by allowing aggregated data retrieval for broader medical insights.

5.4 FlowChart

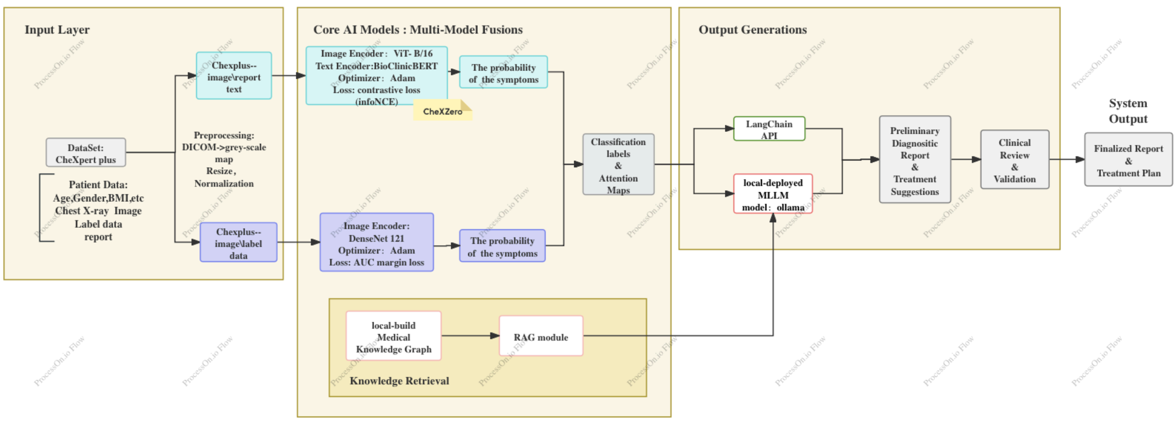


Figure 5.4 System Design Flowchart

6.Implementation

6.1 Front-end Website Implementation

The front-end of the system is implemented as a web-based interface that provides an accessible and user-friendly environment for clinicians and researchers to interact with the diagnostic platform. The design emphasizes modularity and clarity, and is structured into four functional modules: **Total Scans**, which displays the cumulative number of X-ray images processed; **Normal Cases**, which organizes and presents cases that have been classified as free of abnormalities; **Abnormal Cases**, which highlights scans identified with potential pathological findings; and **Analysis History**, which stores all previously generated diagnostic reports for subsequent review and longitudinal analysis.

The workflow begins with an **image upload** component, through which users submit chest X-ray images for analysis. Upon upload, the request is transmitted to the backend server, where the AI diagnostic system processes the image and produces an automated diagnostic report. The resulting report is then rendered dynamically within the web interface and categorized under the appropriate module (normal or abnormal). At the same time, the report is archived into the history database, ensuring both immediate access to the diagnostic outcome and long-term availability for future retrieval, comparative studies, and patient monitoring. This modular design not only enhances usability but also supports scalability, allowing new functionalities to be incorporated into the front-end as the system evolves.

6.2 Back-End Database Implementation

The back-end database serves as the foundational layer of the system, providing secure and efficient storage, indexing, and retrieval of all diagnostic data. It is organized into four main modules: **Total Scans**, which records the overall number of X-ray images processed by the system; **Normal Cases**, which stores cases identified as healthy or without abnormalities; **Abnormal Cases**, which logs findings with potential pathologies for further medical review; and **Analysis History**, which archives detailed diagnostic reports along with associated metadata for long-term tracking and follow-up analysis. When a user uploads an image through the front-end interface, the system processes the scan, generates a structured diagnostic report, and automatically stores the results in the appropriate modules within the database. This architecture ensures both systematic categorization of medical cases and seamless retrieval for clinicians and researchers, enabling case comparisons, and improved decision support.

6.3 Label-Prediction Implementation

This part of the project is currently under active implementation. We are setting up the complete workflow for label prediction, including collecting and preprocessing datasets, implementing the predictive models, and preparing evaluation metrics.We will refer to the implementation form of the paper to reproduce this result. Instructions for acquiring the datasets are provided on the official websites, ensuring transparency and accessibility. Additionally, relevant articles and references are being reviewed and incorporated to guide model design, optimization, and benchmarking. This ongoing work ensures that the label-prediction module is fully functional and aligned with current research standards.