**CS-025**

**Project Title**

B.S. (CS)/(SE) Final year project proposal

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Sr.no | Name | Roll no | ST/CT No | Sec |
| 1 | Syed Hashir Ali Hussaini | 20B-051-CS | CT-20084 | B |
| 2 | Sayed Muhammad Abdullah | 20B-063-CS | CT-20077 | B |
| 3 | Muhammad Huzaifa Khan | 20B-039-CS | CT-20049 | B |

Batch 20B

Date: 30-10-2023

Department of Computer Science  
 Usman Institute of Technology

**Usman Institute of Technology**

**Department Of Computer Science**

**FINAL PROJECT APPROVAL FORM**

The Head of Department, Date: 30-10-2023

Computer Science Department, Batch: \_\_\_\_\_\_\_\_\_

UIT

Karachi.

Subject: **Bachelor of Science in Computer Science/Software Engineering Final Year Project**

Respect Sir,

We, the below listed students of final Year BS \_\_\_\_\_\_\_ class, desire to undertake work on the following project.

We request you to kindly grant approval for undertaking the work on the above-cited project. I abide by all terms and conditions mentioned below.

1. I have selected this project on my own.
2. I have no objection working under the supervision of male/female supervisor, or if my project work is evaluated by male/female externals.
3. I am sure I can complete this project till \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_.
4. I am eager to work under the supervision of advisor assigned to this project.
5. I understand that FYP committee can modify the scope of the project as and when required.
6. I know that if do not appear in regular project progress presentations/milestones my project will be disqualified.
7. I know that if I do not appear in mid project presentation, whenever it is scheduled, I will not be eligible for final project viva
8. I fully understand that “***cheating***”**\*** may lead to cancelation of my project.
9. I understand that the decision of the FYP evaluation committee, for all issues, would be final, and no objections will be accepted.
10. I have no objection presenting my project to external or internal examiner assigned by the Head of the Department.
11. Project and Product deliverables at the time of submission of final year project every group is responsible to submit complete running system along with printed reports, source code, hardware (if any) etc to the project coordinator.
12. It would be the responsibility of Project coordinator to keep record of all projects in a system (in running form) as it would help to continue next project in continuation, depends upon the scope and application of project.
13. Proper dressing and way of presentation should be in English during proposal defend session, milestones and final presentations.
14. Marking of milestones and final presentation should be based on individual evaluation of each faculty members and marks would be granted during session.
15. When we go for proposal defends session a list of all previous projects with their brief introduction must be available during session for our reference. (Introduction, Scope of project, tools and technology and batch must be available).
16. I understand that it is my responsibility to update my advisor and FYP committee members with the status of my project and submit reports on time.

* Copying code from any resources
* Using off the shelf components without prior permission
* Outsourcing your project
* Hire a resource for the completing the FYP code or any part of the project.

Yours sincerely,

S.No. Roll No. Name Email Cell NO. Signature

1. 20B-051-CS S. Hashir Ali Hussaini 20b-051-cs@students.uit.edu 0330-2134872

2. 20B-063-CS Sayed M. Abdullah 20b-063-cs@students.uit.edu 0333-2512070

3. 20B-039-CS M. Huzaifa Khan 20b-039-cs@students.uit.edu 0345-4669034

**TABLE OF CONTENTS**

Project executive summary 1

Project overview 1

Project objectives 1

Project scope 2

In scope: 2

Out of scope: 2

Deliverables produced 2

Project estimated effort/cost/duration 3

Estimated effort hours: 3

Estimated duration: 3

Project assumptions 4

Project risks 4

Project approach 5

Tools and technologies 5

Outstanding Issues 5

Expected Final product 5

Similar products available 5

List down similar products available online with URL.Project approvals: 5

Project approvals: 6

# Project executive summary:

The executive summary of the project is a concise yet comprehensive overview of the key elements of the project definition. It covers the following high-level aspects:

* **Project Objectives:** The project aims to develop a cutting-edge medical image analysis system using deep learning models. These models will be trained on a private dataset to diagnose various diseases from chest X-ray and CT scan images. The system will prioritize patients based on disease severity and generate automated diagnostic reports.
* **Project Scope:** The project's scope includes data collection, model training, efficiency comparison, disease diagnosis, severity prioritization, and automated report generation. It encompasses hardware and software requirements, ensuring that the infrastructure and tools necessary for successful project execution are in place.
* **Assumptions:** Several assumptions have been made to facilitate project planning and execution. These include the availability of data, hardware, software, and skilled resources, as well as compliance with data privacy regulations.
* **Project Risks:** Various risks have been identified, ranging from data availability and hardware/software issues to regulatory compliance and data security. Mitigation strategies are in place to minimize these risks and ensure project success.
* **Costs:** While not explicitly outlined in this summary, the project involves costs related to hardware procurement, software licensing, and potential cloud services. These costs will be included in the detailed project plan.
* **Timeline:** A detailed project timeline has not been provided in this summary but will be developed as part of the project plan. The timeline will include milestones and deliverable deadlines.
* **Approach:** The project approach involves utilizing high-performance GPUs, deep learning frameworks like TensorFlow, and modern development tools to create an efficient and user-friendly medical image analysis system.
* **Organization:** The project team consists of skilled individuals specializing in deep learning, software development, and system administration. Effective collaboration and communication within the team are key to project success

# 

# Project overview

**Background and Context:**

The project focuses on the development of a deep learning-based medical image analysis system with the primary objective of detecting multiple diseases from CT scans and chest X-rays. This project is initiated due to the increasing importance of leveraging advanced technologies to enhance the accuracy, efficiency, and automation of disease diagnosis in the field of radiology.

Medical imaging, including CT scans and chest X-rays, plays a pivotal role in the early detection and diagnosis of a wide range of diseases, such as pulmonary conditions, cardiovascular disorders, and various types of cancers. The traditional approach to interpretation of medical images is highly dependent on the expertise of radiologists, which can be time-consuming, subject to human error, and may not always be readily available in remote or underserved areas. This project aims to address these challenges by harnessing the power of deep learning, which has demonstrated significant potential in automating and improving the accuracy of disease detection in medical images.

**Business Value:**

The business value of this project is multifaceted and holds great significance in the healthcare industry:

1. Improved Accuracy: The implementation of deep learning models for disease detection can significantly enhance the accuracy of diagnosis, reducing false positives and false negatives. This, in turn, leads to more effective treatment planning and better patient outcomes.

2. Efficiency: Automation of the diagnostic process through AI-driven models can accelerate the interpretation of medical images. This means quicker diagnosis and treatment decisions, reducing patient waiting times and the workload of healthcare professionals.

3. Accessibility: By integrating AI-powered disease detection into medical facilities, especially in remote or underserved areas, access to prompt and accurate medical diagnosis can be extended to a broader population.

4. Resource Optimization: This project can help optimize the utilization of healthcare resources by prioritizing patients based on the severity of their conditions. It ensures that the most critical cases receive immediate attention.

5. Cost Reduction: Accurate and timely diagnosis can lead to cost savings in the long run by avoiding unnecessary procedures, late-stage treatments, and hospitalization.

6. Research and Development: This project contributes to the advancement of AI technology in healthcare, fostering innovation and research in the field of medical image analysis.

# 

# Project objectives

Our project will meet the following objectives:

**Objective #1: Data Set Training**

* Specific: Train a private medical image dataset, consisting of chest X-ray and CT scan images, using deep learning models.
* Measurable: Achieve a specific accuracy or performance metric in disease detection through the trained models.
* Achievable: The resources, including the dataset and deep learning infrastructure, are available for this objective.
* Realistic: Given the advancements in deep learning and the availability of medical imaging data, this objective is realistic.
* Time-Based: Complete the data training process within a specified timeframe.

**Objective #2: Model Efficiency Comparison**

* Specific: Evaluate and compare the efficiency and accuracy of various deep learning models for disease detection using medical images.
* Measurable: Measure the performance of different models using quantitative metrics (e.g., accuracy, precision, recall).
* Achievable: Resources for model training and evaluation are accessible.
* Realistic: Comparative analysis of deep learning models is a well-established practice in the field.
* Time-Based: Complete the model efficiency comparison within a defined timeframe.

**Objective #3: Disease Diagnosis and Severity Prioritization**

* Specific: Develop a system that can accurately diagnose multiple diseases from chest X-ray and CT scan images.
* Measurable: Assess the system's performance in terms of disease identification and prioritize patients based on the severity of their conditions.
* Achievable: Access to medical imaging data and expertise in deep learning is available.
* Realistic: The integration of AI for disease diagnosis and prioritization is a feasible application.
* Time-Based: Implement and validate the disease diagnosis and severity prioritization system within a predetermined timeframe.

**Objective #4: Automated Report Generation**

* Specific: Create an automated reporting system that generates diagnostic reports based on the results of disease detection and severity assessment.
* Measurable: Assess the accuracy and completeness of the generated reports.
* Achievable: The necessary data and AI infrastructure are in place for report generation.
* Realistic: Automated reporting is a valuable extension of the disease diagnosis system.
* Time-Based: Develop and validate the automated reporting system within a specified timeframe.

# Project scope

The project scope statement concisely outlines the project's objectives and boundaries. It includes data collection, model development, disease diagnosis, prioritization, automated reporting, and evaluation within scope. Key exclusions involve infrastructure, regulatory compliance, and clinical decision-making. This statement maintains a clear project focus.

The scope of this project includes and excludes the following items.

## In scope:

**1. Deliverables:**

* Business Requirements: Defining the specific needs and objectives of the project, including disease detection and prioritization.
* Current State Assessment: Analyzing the current state of medical image analysis, identifying challenges, and assessing existing technologies.
* Data Collection and Preparation: Collection of a private medical image dataset, including chest X-ray and CT scan images. Preprocessing and data augmentation to ensure the quality and diversity of the dataset.

**2. Life-Cycle Processes:**

* Analysis: The project encompasses the analysis of the current state, requirements, and data collection methods.
* Design: Designing the deep learning models, diagnostic algorithms, and automated reporting system.
* Testing: Testing, validating, and evaluating the deep learning models and system for accuracy and reliability.

**3. Data:**

* Types of Data: Medical image data, including chest X-ray and CT scan images.
* Data Preparation: Data collection, preprocessing, and augmentation are part of the scope to ensure dataset quality.

**4. Data Sources:**

* Data collected from private medical image datasets.
* Data augmentation may involve external data sources for diversity.

**5. Organizations:**

* Healthcare organizations, including hospitals and clinics where the system will be deployed for medical image analysis.

**6. Functionality:**

* Disease Diagnosis: The core functionality includes automated disease detection from medical images, prioritization based on severity, and automated reporting.
* Information Retrieval: Retrieving patient information for reporting is in scope.
* Image Processing: Deep learning-based image processing is central to the project.

## Out of scope:

**1. Deliverables:**

* Business Strategy: The project does not include the development of a business strategy; it is focused on technical aspects.
* Legal Compliance Documentation: While regulatory compliance is important, the project does not involve creating legal compliance documentation.

**2. Life-Cycle Processes:**

* Implementation: The implementation of the developed system into healthcare facilities is not within the scope.
* Maintenance and Support: Ongoing maintenance and support of the system are out of scope.

**3. Data:**

* Financial Data: The project does not involve financial data.
* Employee Data: Employee-related data is not part of the project's scope.

**4. Data Sources:**

* Databases unrelated to medical imaging are out of scope.

**5. Organizations:**

* Human Resources: HR processes are not in scope.
* Manufacturing: Manufacturing processes and organizations are unrelated to the project.

**6. Functionality:**

* Decision Support: While the system aids in diagnosis, it does not make clinical decisions.
* AR/VR and Blockchain: These technologies are not part of the project's scope; it is primarily focused on medical image analysis.

## 

## Deliverables produced

**Project Deliverable 1: Deep Learning Models Comparative Study**

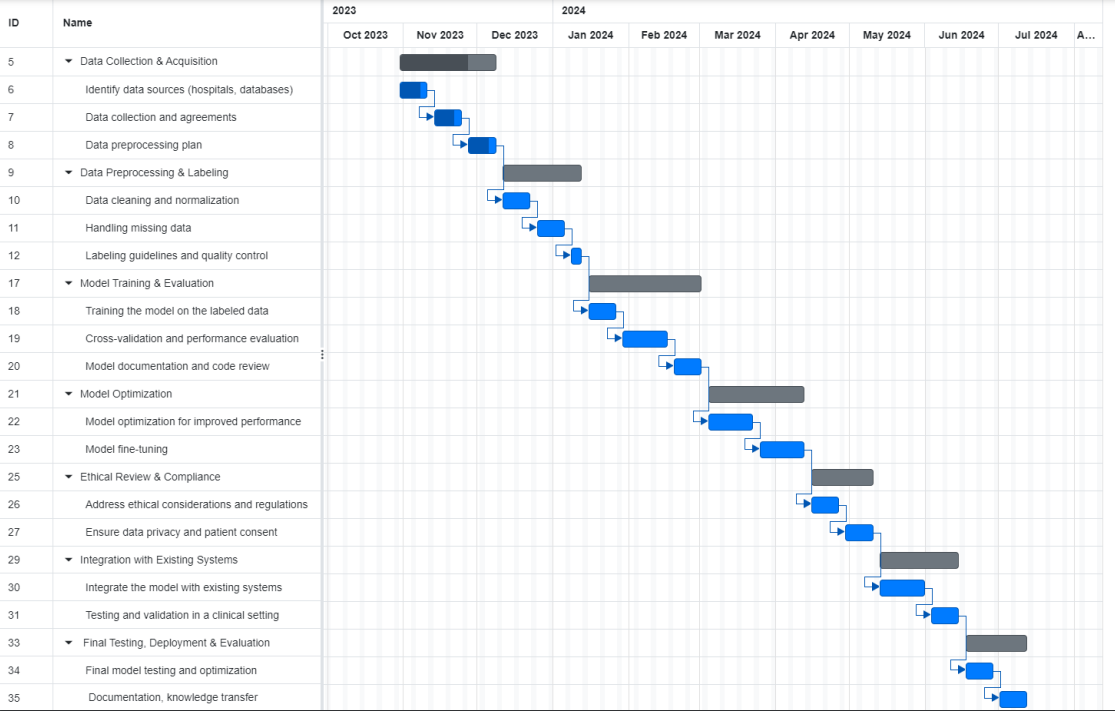
* **Description:** This deliverable involves conducting a comprehensive comparative study of various deep learning models for medical image analysis. The study aims to evaluate the efficiency and complexity of these models when trained on our private medical image dataset. The results will inform the selection of the most effective model for disease diagnosis, prioritization based on severity, and automated diagnostic report generation.
* **Milestone:** Completion of the comparative study and the generation of a report summarizing the findings.

**Product Deliverable 2: Disease Diagnosis System**

* **Description:** The disease diagnosis system is a key product deliverable. It will be built based on the findings from the deep learning models comparative study. This system will utilize the selected deep learning model to diagnose diseases from medical images, prioritize patients according to the severity of their conditions, and generate automated diagnostic reports.
* **Milestone:** Deployment of the disease diagnosis system for use in healthcare facilities.

## Project estimated effort/cost/duration

**GANTT chart/ Increments/ sprints:**

.

## Estimated effort hours:

An estimated effort of each member per week to be mentioned, and estimated time with your supervisor per week.

|  |  |
| --- | --- |
| **Members** | **Estimated hrs per Week** |
| Syed Hashir Ali Hussaini | 11 Hrs |
| Sayed Mohammad Abdullah | 12 Hrs |
| Muhammad Huzaifa Khan | 11 Hrs |
| Discussion/Meeting with Supervisor | 2 Hrs |

## Estimated duration:

An Estimated completion date of each milestone to be mentioned along with its deliverables. You can take help from FYP 2023 -2024 Tentative Plan

|  |  |  |
| --- | --- | --- |
| **Milestone** | **Date completed** | **Deliverable(s) completed** |
| Project planning | Mm/dd/yy | 1. Project definition 2. Workplan |
| Milestone 1 | Mm/dd/yy | 1. All Diagrams giving overview of project 2. First four chapters 3. Selecting Algorithms |
| Milestone 2 |  | 1. POC of the project |
| Milestone 3 |  | 1. Fully functional project for demonstration |
| Milestone 4 |  | 1. Hosted project with all functionalities and polished edges |
| Project conclusion |  |  |

# Project assumptions

1. Hardware and Software Accessibility: The assumption is that the necessary hardware, including high-performance GPUs, and software tools for deep learning model development will be obtainable as required.
2. Collaborative Team: The assumption is that the project team members will effectively collaborate and communicate to meet project goals, despite potential geographic or time zone differences.
3. Regulatory Compliance: The assumption is that the project will adhere to all relevant data privacy and healthcare regulations, and any necessary approvals or permissions will be obtained.
4. Stable Deep Learning Frameworks: The assumption is that the deep learning frameworks, such as TensorFlow, will remain stable and supported throughout the project, ensuring compatibility and functionality.
5. Cloud Services Reliability: If cloud services like AWS are utilized, the assumption is that these services will provide reliable performance and availability.
6. Data Security: The assumption is that measures for safeguarding patient data and images will be effective, ensuring data security and privacy.
7. User Interface Development Tools: The assumption is that the tools and technologies used for frontend and backend development will continue to be available and well-maintained.
8. Project Management Tools: The assumption is that the project management tools like Trello will continue to support project organization, task management, and collaboration.

# Project risks

|  |  |
| --- | --- |
| **Risk Description** | **Mitigation Strategy** |
| **Hardware and Software Unavailability Risk:** | * Maintain close contact with hardware and software suppliers. * Consider alternative vendors or cloud services to ensure access to essential resources. |
| **Data Security Risk:** | * Implement robust data security protocols, encryption, and access controls. * Regularly audit and test the security measures to identify vulnerabilities. |
| **Regulatory Compliance Risk:** | * Stay updated on relevant regulations and maintain a legal consultant's support for compliance. |
| **Cloud Service Reliability Risk:** | * Diversify cloud service providers or have a contingency plan to move to another provider. |
| **Project Team Collaboration Risk:** | * Implement clear communication protocols, schedule regular team meetings, and use collaboration tools. * Facilitate effective communication and cooperation. |
| **Deep Learning Framework Stability Risk:** | * Continuously monitor updates and changes in the framework. * Ensure compatibility testing when updates are applied. |
| **User Interface Development Tools Risk:** | * Regularly update and adapt development tools to stay current with technology advancements. * Ensure compatibility with other project components. |

# 

# Project approach

* **Project Structuring:**
  + **Phases:** The project will be structured into several key phases, including data collection, deep learning model training, model evaluation, and system development. Each phase will have well-defined objectives and deliverables.
* **Data Collection and Preparation:**
  + **Data Sources:** Private medical image datasets, including chest X-ray and CT scan images, will be collected.
  + **Preprocessing:** Data preprocessing and augmentation techniques will be applied to enhance dataset quality and diversity.
* **Deep Learning Model Comparative Study:**
  + **Model Selection:** Various deep learning models, such as Convolutional Neural Networks (CNNs), will be selected for the comparative study.
  + **Training:** The selected models will be trained on the private dataset, with a focus on optimizing their efficiency and performance.
  + **Evaluation:** The comparative study will include rigorous evaluation using appropriate metrics to determine the best-performing model.
* **Disease Diagnosis System Development:**
  + **Model Integration:** The most effective deep learning model will be integrated into the disease diagnosis system.
  + **Severity Prioritization:** Algorithms for severity prioritization will be implemented to classify patients based on the severity of their conditions.
  + **Automated Reporting:** The system will generate automated diagnostic reports based on disease detection and prioritization results.
* **Testing and Validation:**
  + **Quality Assurance:** Rigorous testing and validation will be conducted to ensure the system's accuracy and reliability.
  + **User Acceptance:** Healthcare professionals will participate in user acceptance testing to refine the system.
* **Project Management:**
  + **Resource Planning:** Adequate resources, including data, computing infrastructure, and personnel, will be allocated for each phase.
  + **Timeline:** A project timeline will be established with clear milestones for each phase.
  + **Risk Management:** A proactive approach to identifying and mitigating risks will be implemented throughout the project.
* **Documentation and Reporting:**
  + Comprehensive documentation will be maintained at each project stage, including the deep learning model comparative study results, system design, and testing reports.
  + Regular progress reports will be provided to stakeholders.
* **Stakeholder Collaboration:**
  + Ongoing collaboration with healthcare providers and professionals will ensure the system aligns with clinical requirements and is user-friendly.

## Tools and technologies

**Hardware:**

* **Graphics Processing Unit (GPU):**
  + NVIDIA high-performance GPU with a focus on CUDA compatibility for deep learning model acceleration.
  + Specifications: (Include GPU model and specifications here, if available)
* **Central Processing Unit (CPU):**
  + Robust CPU with multiple cores, e.g., Core i7 8th Gen, to support various project software processes.
  + Specifications: (Include CPU model and specifications here, if available)
* **Random Access Memory (RAM):**
  + A minimum of 16GB RAM for memory-intensive tasks, such as deep learning model training and image manipulation.
  + Specifications: (Include RAM specifications here, if available)

**Software:**

**1. Data Gathering and Storage:**

* **NumPy:** For fundamental numerical operations and data handling.
* **Pandas:** For data manipulation and structured data handling.
* **OpenCV:** Essential for image preprocessing, augmentation, and handling.

**2. Medical Imaging Data Handling:**

* **PyDicom:** Specifically, for working with DICOM format medical images.

**3. Data Preprocessing and Image Augmentation:**

* **imgaug:** To perform data augmentation and generate new samples from existing

images.

* **scikit-image**: Provides additional image processing functions.
* **Albumentations:** Versatile, deep learning-focused image augmentation.
* **Augmentor:** Simplifies image augmentation for machine learning.

**4. Imbalanced Dataset Handling:**

* **imbalanced-learn:** Useful for addressing imbalanced datasets.

**5. Model Development and Comparative Analysis:**

**Deep Learning Frameworks:**

* **PyTorch:** A popular framework known for flexibility and dynamic computation

graphs.

* **TensorFlow:** A widely used framework known for scalability and production

readiness.

* **Keras:** A high-level API that can run on top of both PyTorch and TensorFlow.

**6. Model Architecture and Training Tools:**

* **PyTorch Lightning (for PyTorch):** Simplifies PyTorch model training and

experiment management.

* **TensorFlow Addons (for TensorFlow):** Offers additional tools for TensorFlow

models.

**7. GPU Support (Optional but Recommended):**

* **NVIDIA CUDA Toolkit:** Required if using NVIDIA GPUs for acceleration.

**8. Hyperparameter Tuning:**

**Optimization Libraries (Optional):**

* **Optuna:** A Python library for optimizing machine learning model parameters.
* **Keras Tuner:** Specifically designed for tuning hyperparameters in Keras models.

**9. Model Evaluation:**

* **scikit-learn:** For model evaluation, including metrics, cross-validation, and data

splitting.

**10. Web Application Frameworks (Optional):**

* **Flask, Django, or FastAPI:** For building web-based applications for model

access and interaction.

**11. Model Deployment Tools (In Production):**

* **TensorFlow Serving, TensorFlow Lite, ONNX:** For deploying deep learning

models to production.

* **AWS SageMaker, Azure ML:** Cloud-based deployment platforms.

**12. Monitoring and Logging (In Production):**

* **Prometheus, Grafana:** cloud-specific monitoring services.

**14. Security and Compliance (In Production):**

* Tools and practices to ensure data security, encryption, access control, and

compliance with healthcare regulations (e.g., HIPAA).

**15. Testing and CI/CD (In Production):**

* **Jenkins, Travis CI, GitLab CI/C:** testing and continuous integration/continuous

deployment tools.

**16. Medical Imaging Libraries (Throughout):**

* **SimpleITK:** A high-level image processing library for medical images.
* **ANTsPy:** A library for image registration, normalization, and analysis of

medical images.

* **Nibabel:** For reading and writing neuroimaging data, which can be adapted for

medical image formats.

* **Medical Image Analysis Libraries (As Needed):** Libraries like MIALab that

offer tools for feature extraction, image registration, and segmentation of

medical images.

# Algorithms

* Resnet
* VGG
* Chexnet
* DenseNet

# References

# [CheXNet: Radiologist-Level Pneumonia Detection on Chest X-Rays with Deep Learning](https://arxiv.org/abs/1711.05225)

* [Densely Connected Convolutional Networks](https://arxiv.org/abs/1608.06993)

# [Very Deep Convolutional Networks for Large-Scale Image Recognition](https://arxiv.org/abs/1409.1556)

# [Deep Residual Learning for Image Recognition](https://arxiv.org/abs/1512.03385)

# Expected Final product

For the project demonstration:

* We'll set up a special place to show how our project works. It'll be like a test environment.
* We need powerful computers to make our project run smoothly. We'll have those computers ready for the demo.
* We want to show that our project can recognize diseases by looking at X-ray pictures. It can also figure out how bad the diseases are and write a report about them.
* During the demo, we'll pretend to use our project just like in a real hospital. We'll show how it looks at X-ray pictures, recognizes diseases, and writes a report.

# Similar products available

**1. Functional Activities:**

* **ChexNet:** ChexNet is a notable deep learning model designed to identify common thoracic diseases in X-ray images. It focuses on specific diseases like pneumonia, edema, and pneumothorax

**URL:** [ChexNet on arXiv](https://arxiv.org/abs/1711.05225)

.

**2. GUI (Graphical User Interface):**

* **Aidoc:** Aidoc is a platform that offers an AI-powered radiology assistant with a user-friendly interface. It helps radiologists in detecting abnormalities in various medical images, including CT scans

**URL:** [Aidoc Website](https://www.aidoc.com/)

.

**3. Reporting:**

* **Zebra Medical Vision:** This platform provides AI-driven medical imaging insights. It can identify various conditions in medical images and generate reports.

**URL:** [Zebra Medical Vision Website](https://www.zebra-med.com/)

**4. General Medical Image Analysis:**

* **Google Health's DeepMind:** DeepMind's AI algorithms have been used for analyzing medical images, including eye scans and other diagnostic images, showcasing the potential for deep learning in healthcare.

**URL:** [Google Health - DeepMind](https://health.google/health-research/)

# Project approvals:

Name:  
Project Supervisor Signature

Member, FYP Committee, UIT Signature