# RESPONSIBLE AI BASED ML MODEL FOR THE DIAGNOSIS OF DISEASE DIAGNOSIS USING CHEST XRAYS

**A report submitted to**

**RAMAIAH INSTITUTE OF TECHNOLOGY**

**Bengaluru**

**MINI PROJECT (IS65)**

as partial fulfillment of the requirement for the award of degree of

**Bachelor of Engineering (B.E) in Information Science and Engineering**

**By**


| | |
|---|---|
| MITHUN R | 1MS20IS072 |
| MOHAMAD FARHAN | 1MS20IS073 |
| PUTTA KRISHNA BABU | 1MS20IS090 |
| R JAYANTH JADHAV | 1MS20IS091 |

Under the Guidance of

SURESH KUMAR K R

ASSISTANT PROFESSOR

DEPARTMENT OF INFORMATION SCIENCE AND ENGINEERING

RAMAIAH INSTITUTE OF TECHNOLOGY

(Autonomous Institute, Affiliated to VTU)

BANGALORE - 54

MAY 2023


## Department of Information Science and Engineering

## Ramaiah Institute of Technology

## Bengaluru - 54

# CERTIFICATE

his is to certify that ARYAN (USN- 1MS20IS098), NAME (USN- 1MSXXXXXX), NAME (USN- 1MSXXXXXX) AND NAME (USN- 1MSXXXXXX) who were working for their MINI PROJECT under my guidance, have completed the work as per my satisfaction with the topic TITLE OF THE PROJECT. To the best of my understanding, the work to be submitted in the dissertation does not contain any work, which has been previously carried out by others and submitted by the candidates for themselves for the award of any degree anywhere.

Name and Signature of the Guide

Signature of the HOD

**Other Examiners**

Name of the examiners

Signature with date

1.

2.

RAMAIAH
Institute of Technology

# DECLARATION

We hereby declare that the entire work embodied in this MINI PROJECT (IS65) report has been carried out by us at Ramaiah Institute of Technology under the supervision of Guide_Name. This project report has not been submitted in part or full for the award of any diploma or degree of this or any other University.

Name1 USN1
Name2 USN2
Name3 USN3
Name4 USN4

Place: Bangalore
Date:

# ABSTRACT

The purpose of this research project is to develop a cognitive solution for X-ray image analysis that can accurately and reliably diagnose diseases while addressing concerns related to responsible AI, including issues related to bias, privacy, and ethical considerations. X-ray image analysis is critical for the diagnosis and treatment of various diseases, but the accuracy and reliability of the analysis can be challenging, especially in resource-limited settings. Moreover, the use of AI models for X-ray image analysis raises concerns about bias, privacy, and ethical considerations.

# ACKNOWLEDGEMENT

We take this opportunity to express our gratitude to the people who have been instrumental in the successful completion of this project. We would like to express our profound gratitude to the Management and Dr. N.V.R Naidu Principal, M.S.R.I.T, Bengaluru for allowing us to explore our potential.

We sincerely thank our beloved Dr. Sanjay H A, HOD, Information Science and Engineering, for his constant support and guidance.

We wholeheartedly thank our project guide Prof. Suresh Kumar K R, for providing us with the confidence and strength to overcome every obstacle at each step of the project and inspiring us to the best of our potential. We also thank her for her constant guidance, direction, and insight during the project.

Finally, we would like to express sincere gratitude to all the teaching and non-teaching faculty of ISE Department, our beloved parents, seniors, and my dear friends for their constant support during the course of work.

# Contents

# List of Figures

# List of Tables

# Chapter 1

# INTRODUCTION

The problem is that while X-ray image analysis is critical for the diagnosis and treatment of various diseases, the accuracy and reliability of the analysis can be challenging, especially in resource-limited settings. The use of AI models for X-ray image analysis raises concerns about bias, privacy, and ethical considerations, which can undermine the trust in the diagnosis and treatment decisions made using these models. Therefore, there is a need to develop a cognitive solution for X-ray image analysis that can accurately and reliably diagnose diseases while addressing concerns related to responsible AI, including issues related to bias, privacy, and ethical considerations.

The topic of responsible AI and cognitive X-ray image analysis was chosen because of the potential benefits it can bring to the medical field. X-ray image analysis is critical for the diagnosis and treatment of various diseases, but the accuracy and reliability of the analysis can be challenging, especially in resource-limited settings

Therefore, developing a cognitive solution for X-ray image analysis that can accurately and reliably diagnose diseases while addressing concerns related to responsible AI, including issues related to bias, privacy, and ethical considerations, is an important research area. Such a solution can improve patient outcomes and reduce healthcare costs while maintaining the trust of patients and healthcare providers in the diagnosis and treatment decisions made using these models. This research topic is highly relevant in the current healthcare landscape, where the use of AI models in medical decision-making is becoming increasingly common.

## 1.1 Motivation and Scope

The requirement for accuracy, efficiency, and accessibility in healthcare is the main driver behind the development of responsible AI-based ML models for illness diagnosis utilising chest X-rays. One of the most often used diagnostic imaging tests is a chest X-ray, particularly for respiratory diseases like pneumonia and tuberculosis. Even for seasoned radiologists, precisely interpreting these pictures can be difficult and time-consuming.

It is conceivable to create models that can correctly diagnose diseases based on chest X-rays using machine learning techniques, potentially increasing diagnostic efficiency and accuracy. To make sure that these models are impartial, open, and ethical, it is crucial to create them ethically.

Figure 1.1 presents present-day challenges call for solutions that require the utmost precision and accuracy in the developed solutions as shown in Fig 1.1.

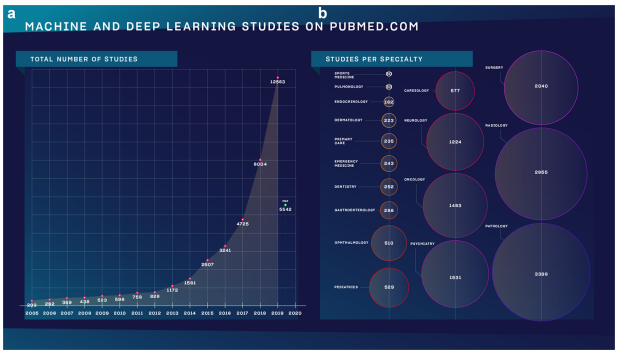


Figure 1.1: Exponential increase in application of AI-ML in medics

This however has led to many research challenges published by various journals.

### 1.1.1 SAMPLE

## 1.2 Issues and Challenges

## 1.3 Problem Statement

### 1.3.1 Sample_Section

## 1.4 Proposed Model

Explain the same with limited words. Not more than three pages.

1. All Table Caption should be in Title Case, TNR 10 Pt. It should be of the Format:
   - Table **??** Results of the Experiment ....(Centered)
2. It should be cited as Table **??**.
3. Caption should appear above the Table.
4. Table Header and the entries should be of Font TNR 10 Pt, Justified.
5. For wider Table, the page orientation can be Landscape.
6. For Larger Table, it can run to pages and the header should be repeated for each page of the Table.
7. Table must be adjusted to fit in the page and no single row is left out for a new page.

## 1.5 Organization of the Report

This will be the last section in the introduction section explaining how the report is organized in the remaining chapters.

# Chapter 2

# LITERATURE REVIEW

A literature review is a text of a scholarly paper, which includes the current knowledge including substantive findings, as well as theoretical and methodological contributions to a particular topic. Literature reviews are secondary sources and do not report new or original experimental work. Most often associated with academic-oriented literature, such reviews are found in academic journals and are not to be confused with book reviews that may also appear in the same publication. Literature reviews are a basis for research in nearly every academic field [1]. A narrow-scope literature review may be included as part of a peer-reviewed journal article presenting new research, serving to situate the current study within the body of the relevant literature and to provide context for the reader. In such a case, the review societally precedes the methodology and results sections of the work.

The work conducted by [2] has shown several limitations.

Producing a literature review may also be part of graduate and post-graduate student work, including in the preparation of a thesis, dissertation, or a journal article. Literature reviews are also common in a research proposal or prospectus (the document that is approved before a student formally begins a dissertation or thesis) [3, 4].

This was suggested by [5] that recognition is accurate.

## 2.1 Review Types

The main types of literature reviews are: evaluative, exploratory, and instrumental [4, 6]. A fourth type, the systematic review, is often classified separately but is essentially a

literature review focused on a research question, trying to identify, appraise, select, and synthesize all high-quality research evidence and arguments relevant to that question. A meta-analysis is typically a systematic review using statistical methods to effectively combine the data used on all selected studies to produce a more reliable result.

### 2.1.1 Process And Product

[7] distinguish between the process of reviewing the literature and a finished work or product known as a literature review. The process of reviewing the literature is often ongoing and informs many aspects of the empirical research project. All of the latest literature should inform a research project. Scholars need to be scanning the literature long after a formal literature review product appears to be completed.

## 2.2 Page Limitation

A careful literature review is usually 15 to 30 pages and could be longer. The process of reviewing the literature requires different kinds of activities and ways of thinking [4] and [8] link the activities of doing a literature review with Benjamin Bloom's revised taxonomy of the cognitive domain.

hello this is my paper .he quoted a word [9].

The introduction to imbalanced data is mentioned by [10].

# Chapter 3

# FRAMEWORK AND SYSTEM DESIGN

Flow chart, Activity diagram, State Diagram, Sequence diagram, Framework diagram, Use case diagram, Data flow diagram, ER diagram, Schema diagram any other diagram and explanation specific to your project.

## 3.1 Design Overview

## 3.2 Figure

- All Figure Caption should be in the Title Case, TNR 10 pt, and it should be of the Format: Fig Chapter Number.Figure Number Figure Caption
- It should be cited as Fig. **??** Caption must appear below the Figure.
- For Smaller: (4 Figures arranged in Two Columns) / Page; Portrait Mod.
- For Medium: (2 Figures arranged one below the other / Page; Portrait Mode.
- For Larger: 1 Figure / Page; Landscape Mode.
- Figure Label should be in Font TNR 10 pt, Bold.
- Figure Resolution should be minimum of 300 DPI.

Figure 3.1: Sample Picture of Universe

For example, The universe is immense and it seems to be homogeneous, in a large scale, everywhere we look at.

There's a picture of a galaxy above in Fig. 3.1

## 3.3 Stake-holder Design

### 3.3.1 Definition and Notations

Let $DS$ denote a binary imbalanced data set. The binary data set comprises Majority and Minority data sets indicated by $DS_{ma}$ and $DS_{mi}$ respectively. Let $|DS_{ma}|$ and $|DS_{mi}|$ be $p$ and $q$ respectively where, $DS_{ma} \cup DS_{mi} = DS$. The class set $C$ consists of majority class and minority class, $C = \{C_{ma}, C_{mi}\}$.

## 3.4 Architectural Design

## 3.5 Workflow

# Chapter 4

# Implementation

Algorithm 4.1 begins with minority training instances, $TR_{mi}$. This minority training set undergoes the partition synthesis process to arrive at an updated minority training data set with synthetic instance, $DS_{Smi}$.

**Algorithm 4.1:** Nearest Neighbor Classifier

**Input:**

- Minority Training data, $TR_{mi}$
- Majority Training Data, $TR_{ma}$
- Test Instance, $T$

**Output:** Predicted Class, C, where $C \; \varepsilon \; \{C_{mi}, \; C_{ma}\}$

1 Using $TR_{mi}$, generate synthetic_minority data set, $DS_{Smi}$;

2 Find $k$ greedy neighbors for the test pattern $T$ from $DS_{Smi}$. Let it be denoted by $\phi mi$;

3 Find $k$ greedy neighbors of the test pattern $T$ from the majority set $TR_{ma}$. Let it be denoted by $\phi ma$;

4 Find $k$ global greedy neighbors of the test tuple $T$ from the union of $\phi mi \cup \phi ma$;

5 Classify T to the class based on the majority vote among the $k$ global neighbors

## 4.1 Formats

Other formats are available. Students are allowed to proceed with the other formats.No programs can be included within the implementation chapter.

# Chapter 5

# Experiments and Results

The results in the form of graphs and tables can be illustrated in this chapter. Sample of writing a mathematical equation is as follows.

Let A be an attribute and $C_i$ be the class predictor variable representing the minority class data. Taking this into account, the confidence for the rule $A \rightarrow C_i$ is as follows, as shown in Eqn. 5.1

$$Confidence(A \rightarrow C_{it}) = \frac{Support(A \cup C_i)}{Support(A)} = \frac{m}{m+p} \tag{5.1}$$

The confidence of a given value of attribute A that it belongs to majority class, $\hat{C}_i$ is as given below as in Eqn. 5.2

$$Confidence(A \rightarrow \hat{C}_i) = \frac{Support(A \cup \hat{C}_i)}{Support(A)} = \frac{p}{m+p} \tag{5.2}$$

# Chapter 6

# Conclusion and Future Scope

## 6.1 Conclusion

## 6.2 Future Scope

# Bibliography


[1] A. Maedche and S. Staab, "Ontology learning for the semantic web," *IEEE Intelligent systems*, vol. 16, no. 2, pp. 72–79, 2001.

[2] H. He and E. A. Garcia, "Learning from imbalanced data," *IEEE Transactions on knowledge and data engineering*, vol. 21, no. 9, pp. 1263–1284, 2009.

[3] A. Doan, J. Madhavan, P. Domingos, and A. Halevy, "Learning to map between ontologies on the semantic web," in *Proceedings of the* $11^{th}$ *international conference on World Wide Web*. ACM, 2002, pp. 662–673.

[4] D. Duzdevich, S. Redding, and E. C. Greene, "Dna dynamics and single-molecule biology," *Chemical reviews*, vol. 114, no. 6, pp. 3072–3086, 2014.

[5] S. Kumar, S. Singh, and J. Kumar, "Multiple face detection using hybrid features with svm classifier," in *Data and Communication Networks*. Springer, 2019, pp. 253–265.

[6] M. Thiruganam, S. M. Anouncia, and S. Kantipudi, "Automatic defect detection and counting in radiographic weldment images," *International Journal of Computer Applications*, vol. 10, no. 2, pp. 1–5, 2010.

[7] D. Gilbarg and N. S. Trudinger, *Elliptic partial differential equations of second order*. Springer Publications, 2015.

[8] C. R. Kothari, *Research methodology: Methods and techniques*. New Age International Publications, 2004.

[9] S. Samarani, O. Allam, P. Sagala, Z. Aldabah, M.-A. Jenabian, V. Mehraj, C. Tremblay, J.-P. Routy, D. Amre, and A. Ahmad, "Imbalanced production of il-18 and its antagonist in human diseases, and its implications for hiv-1 infection," *Cytokine*, vol. 82, pp. 38–51, 2016.

[10] L. M. Mathews and H. Seetha, "On improving the classification of imbalanced data," *Cybernetics and Information Technologies*, vol. 17, no. 1, pp. 45–62, 2017.