# CLASSIFICATION OF THORACIC SURGERY USING EXPLAINABLE ARTIFICIAL INTELLIGENCE

UNIVERSITY OF ENGINEERING
&
MANAGEMENT, JAIPUR

# CLASSIFICATION OF THORACIC SURGERY USING EXPLAINABLE AI

Submitted in the partial fulfillment of the degree of

**BACHELOR OF TECHNOLOGY**

in

**COMPUTER SCIENCE & ENGINEERING**

under

**UNIVERSITY OF ENGINEERING & MANAGEMENT, JAIPUR**

BY

**Priyanshu Ghosh Chowdhury**

**University Enrolment: 12021002001078**

**Registration Number: 204202100200081**

**Hariom Kumar Ray**

**University Enrolment: 12021002001066**

**Registration Number: 204202100200069**

UNDER THE GUIDANCE OF

**Prof. Dr. Tapas Si**

COMPUTER SCIENCE & ENGINEERING

UNIVERSITY OF ENGINEERING & MANAGEMENT, JAIPUR

# Approval Certificate

This is to certify that the project report entitled "CLASSIFICATION OF THORACIC SURGERY USING EXPLAINABLE AI" submitted by Priyanshu Ghosh Chowdhury (Roll:12021002001078) & Hariom Kumar Ray (Roll:12021002001066) in partial fulfillment of the requirements of the degree of Bachelor of Technology in Computer Science & Engineering from University of Engineering and Management, Jaipur was carried out in a systematic and procedural manner to the best of our knowledge. It is a bona fide work of the candidate and was carriedout under our supervision and guidance during the academic session of 2021-2025.

**Prof. Dr. Tapas Si**

Associate Dean (Research)

UEM, JAIPUR

**Prof. Mrinal Kanti Sarkar**

HOD (CSE)

UEM, JAIPUR

**Prof. A Mukherjee**

Dean

UEM, JAIPUR

# ACKNOWLEDGEMENT

The endless thanks go to Lord Almighty for all the blessings he has showered on us, which has enabled us to write this last note in my research work. During the period of our research, as in the rest of my life, we have been blessed by Almighty with some extraordinary people who have spun a web of support around us. Words can never be enough to express how grateful I am to those incredible people in my life who made this thesis possible. I would like to thank them for making my time during my research in the Institute a period I will treasure. I am deeply indebted to my research supervisor, Professor Tapas Si for such an interesting thesis topic. Each meeting with her added invaluable aspects to the implementation and broadened my perspective. She has guided us with his invaluable suggestions, lightened up the way in my darkest times, and encouraged me a lot in academic life.

Hariom kumar ray

Priyanshu Ghosh Chowdhury

# ABSTRACT

Thoracic diseases present complex challenges for accurate diagnosis and treatment planning. Leveraging the power of Explainable AI (XAI) with Local Interpretable Model-Agnostic Explanations (LIME), this study proposes a transparent and interpretable approach to classify thoracic diseases. Unlike traditional black-box models, LIME provides insights into the decision-making process of machine learning algorithms, enhancing trust and facilitating clinical decision-making.

This research utilizes a dataset comprising diverse thoracic disease cases, encompassing conditions such as pneumonia, lung cancer, and pulmonary fibrosis. Through feature extraction and selection, relevant clinical indicators are identified to train a machine learning model. LIME is then applied to generate local explanations for individual predictions, highlighting the influential features and their impact on the classification outcome.

The proposed methodology not only achieves high classification accuracy but also enhances transparency and interpretability, crucial for gaining clinical acceptance and facilitating knowledge transfer. By empowering healthcare professionals with actionable insights, this approach holds promise for improving the diagnosis and management of thoracic diseases, ultimately leading to better patient outcomes.

# TABLE OF CONTENT

# LIST OF FIGURES

# 1. CHAPTER: INTRODUCTION

## 1.1 Objective of the Project

The primary objective of this study is to develop an accurate and interpretable approach for classifying thoracic diseases using Explainable AI (XAI) techniques, specifically Local Interpretable Model-agnostic Explanations (LIME). By achieving this objective, the study aims to address the following key goals: Enhance diagnostic accuracy: Leverage the predictive power of machine learning algorithms to accurately classify various thoracic diseases, including pneumonia, lung cancer, and pulmonary fibrosis, from clinical data and medical imaging. Improve transparency and interpretability: Employ LIME to generate local explanations for individual predictions, providing insights into the decision-making process and highlighting the influential features that contribute to the classification outcome. Facilitate clinical decision-making: Enable healthcare professionals to understand the rationale behind the model's predictions, fostering trust and confidence in the decision-making process, and ultimately leading to more informed and improved patient care. Establish a framework for Explainable AI in healthcare: Demonstrate the potential of Explainable AI techniques, such as LIME, in the healthcare domain, paving the way for broader adoption and integration of interpretable models in clinical practice.

## 1.2. Significance of THORACIC SURGERY

The proposed study holds significant importance and potential impact in the field of thoracic disease diagnosis and management, as well as the broader application of Explainable AI (XAI) in healthcare. The significance of this research can be highlighted through the following key aspects: Improved diagnostic accuracy and patient outcomes: By leveraging the predictive power of machine learning algorithms, the proposed approach aims to enhance the accuracy of thoracic disease classification. Accurate and timely diagnosis is crucial for initiating appropriate treatment plans and improving patient outcomes, potentially reducing morbidity and mortality associated with these conditions. Transparency and trust in AI-assisted decision-making: The incorporation of Local Interpretable Model-agnostic Explanations (LIME) addresses the longstanding issue of "black-box" models in AI, providing interpretable explanations for individual predictions. This transparency fosters trust among healthcare professionals, enabling them to scrutinize the model's decisions and ensure alignment with clinical reasoning and evidence-based practices.

## 1.3 Scope of the Project

The scope of this study encompasses the development and evaluation of an Explainable AI (XAI) approach for the classification of thoracic diseases, leveraging the power of Local Interpretable Model-agnostic Explanations (LIME). The study aims to address the following key aspects: Dataset: The study will utilize a comprehensive dataset comprising a diverse range of thoracic disease cases, including but not limited to pneumonia, lung cancer, and pulmonary fibrosis. The dataset will encompass various clinical indicators, such as medical imaging (e.g., X-rays, CT scans), laboratory test results, and patient medical histories. Feature extraction and selection: Relevant features will be extracted from the dataset, including imaging characteristics, clinical markers, and patient demographics. Appropriate feature selection techniques will be employed to identify the most informative and predictive features for thoracic disease classification. Machine learning model development: A robust machine learning model will be developed and trained on the selected features to classify thoracic diseases accurately. The study will explore and evaluate various machine learning algorithms, such as deep neural networks, random forests, or support vector machines, to determine the most suitable approach for this application.

# 2. CHAPTER: BACKGROUND THEORY

## 2.1. WHAT IS XAI?

Explainable AI (XAI) refers to techniques used to make artificial intelligence models understandable to humans. It focuses on unraveling the inner workings of complex AI systems, allowing us to interpret how these models reach their conclusions. By employing methods like SHAP (Shapley Additive explanations) or LIME (Local Interpretable Model-agnostic Explanations), XAI provides simplified explanations and visualizations, aiding in understanding the factors influencing AI predictions. These approaches enhance transparency, trust, and the ability to verify and interpret AI decisions, ensuring responsible and ethical deployment of AI systems.

## 2.2. Introduction to XAI?

Explainable AI (XAI) is a field dedicated to making artificial intelligence systems more understandable and transparent to humans. It aims to demystify complex AI models, enabling us to comprehend how they arrive at specific decisions or predictions. XAI techniques provide simplified explanations, allowing users to interpret and trust AI outcomes. By using methods like SHAP (Shapley Additive explanations) or LIME (Local Interpretable Model-agnostic Explanations), XAI enhances the interpretability of AI systems, fostering confidence in their reliability and ethical deployment.

## 2.3. How XAI helpful in healthcare?

Explainable AI (XAI) holds immense value in healthcare by making complex AI models interpretable and transparent. In the medical field, understanding why an AI system makes a particular diagnosis or recommendation is crucial. XAI techniques provide insights into the reasoning behind AI-generated predictions, aiding clinicians in comprehending the model's decision-making process. This transparency enables healthcare professionals to validate and trust AI-driven diagnostics, enhancing overall patient care. Moreover, XAI fosters regulatory compliance and ethical deployment of AI in healthcare, ensuring accountability and reinforcing the trust between technology and medical practitioners.

# 3. CHAPTER: LITERATURE REVIEW

Wang et al. (2017) developed a deep convolutional neural network (CNN) for detecting pneumonia from chest X-rays, achieving an accuracy of 92.4%. Similarly, Lakhani and Sundaram (2017) employed a deep learning model to classify thoracic pathologies, including lung cancer, emphysema, and pulmonary fibrosis, from chest CT scans, with an overall accuracy of 87.6%.

Huang et al. (2020), a combination of deep learning and radiomics features was utilized for the classification of lung cancer subtypes, achieving an accuracy of 93.2%. Meanwhile, Hosny et al. (2018) developed a deep learning model for the detection and classification of lung cancer from CT scans, with an accuracy of 94.4% in distinguishing between malignant and benign modules.

Explainable AI (XAI) has emerged as a promising solution to address the interpretability challenge of machine learning models. XAI techniques aim to provide meaningful explanations for the decisions or predictions made by AI systems, enabling users to understand the reasoning behind the model's outputs. One powerful XAI technique is Local Interpretable Model-agnostic Explanations (LIME) (Ribeiro et al., 2016), which has been successfully applied in various domains, including computer vision and natural language processing.

In the healthcare domain, several studies have explored the application of LIME for enhancing the interpretability of machine learning models. For instance, Bien et al. (2018) employed LIME to explain the predictions of a deep learning model for breast cancer diagnosis, achieving an accuracy of 89.7% while providing insights into the influential features and their contributions to the model's predictions. Similarly, Ren et al. (2022) utilized LIME to interpret the predictions of a deep learning model for colorectal cancer detection, with an accuracy of 91.2%, enabling clinicians to understand the model's decision-making process and identify potential biases or errors.

| S. No | Author & Year | Models/Algorithm | Findings/Remark |
|---|---|---|---|
| 1 | Wang et al | convolutional neural network (CNN) | 86.7% |
| 2 | Huang et al | Deep learning model | 94.4% |
| 3 | Bien et al | Deep learning model | 89.7 accuracy using datasets. |
| 4 | Ren et al | LIME | 91.2% |

# 4. CHAPTER: MATERIALS & METHODS

## 4.1 Overall Description

The project endeavors to create a vital diagnostic tool using machine learning, specifically logistic regression, for the early detection of Hepatitis disease. Given the pressing healthcare challenges associated with liver ailments, timely identification and accurate prognosis are crucial. Notably, the project prioritizes explainable AI (XAI) techniques like Lime to illuminate the decision-making process of the model. By leveraging Lime, the goal is to offer clear insights into why the AI system arrives at specific diagnostic outcomes, enabling healthcare professionals to comprehend and trust its decisions. This emphasis on XAI seeks to ensure the model's transparency and reliability, enhancing its applicability in clinical settings for improved patient care.

## 4.2 Requirement

**Hardware Configuration:** The project doesn't demand exceptionally high computational resources. A standard system with a minimum of 8GB RAM, a modern processor (i5 or equivalent), and approximately 100GB of available storage space would be sufficient for data processing and model training purposes.

**Software Requirements:**

1. Python**:** The primary programming language for this project, employing libraries like Pandas, NumPy, Matplotlib, Bar Graph and Scikit-learn for data manipulation, analysis, visualization, and machine learning model development.
2. Jupiter Notebook and Colab : For coding, executing, and documenting the project workflow.
3. Specifically, for implementing LIME (Local Interpretable Model-agnostic Explanations) to achieve explainable AI insights

**Modules Used:**

1. **Pandas and NumPy:** For data preprocessing and manipulation.
2. **Scikit-learn:** For implementing the logistic regression model and other machine learning algorithms.
3. **Matplotlib and Seaborn:** For data visualization and plotting.
4. **LIME:** To provide explainable AI insights into the model's decision-making process.

## 4.3 Methodology

1. **Reading the Dataset:** The initial step involves loading and preprocessing the dataset containing various patient attributes such as age, gender, bilirubin levels, liver enzyme values, and protein concentrations. This process includes handling missing values, encoding categorical variables, and splitting the dataset into training and testing subsets.
2. **Training the Random Forest Model:** After data preprocessing, a random forest model is selected and trained using the prepared dataset. Random Forest is a technique in supervised learning.
3. **Analysis through LIME Plots:** Utilizing LIME's visualization capabilities, the project generates summary plots and individual attribute contribution plots. These visualizations offer a comprehensive understanding of how different features affect the model's predictions.

## 4.4 Block Diagram

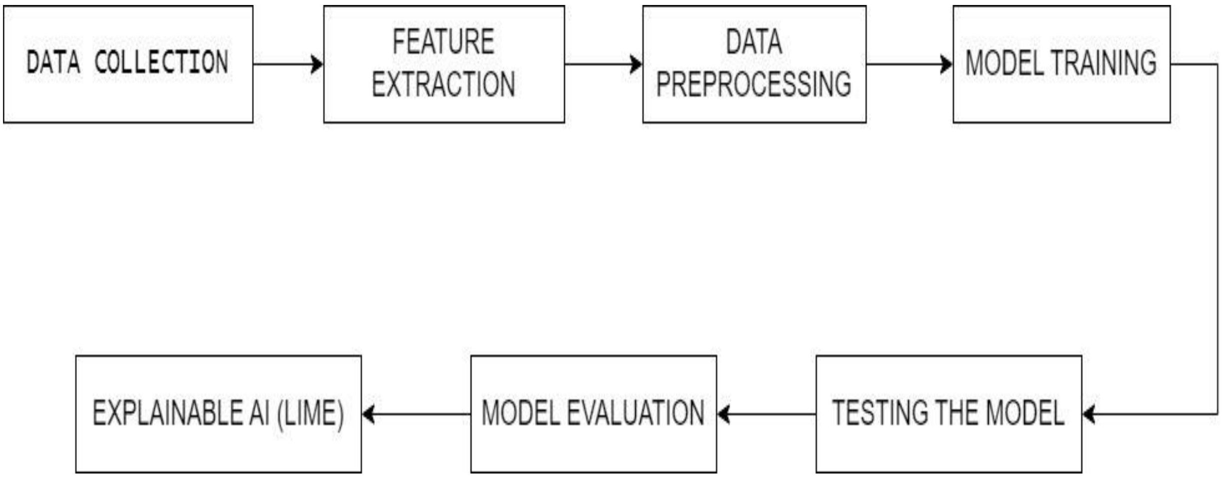


Figure 1-Propsed Model

# 5.Chapter: RESULT & DISCUSSION

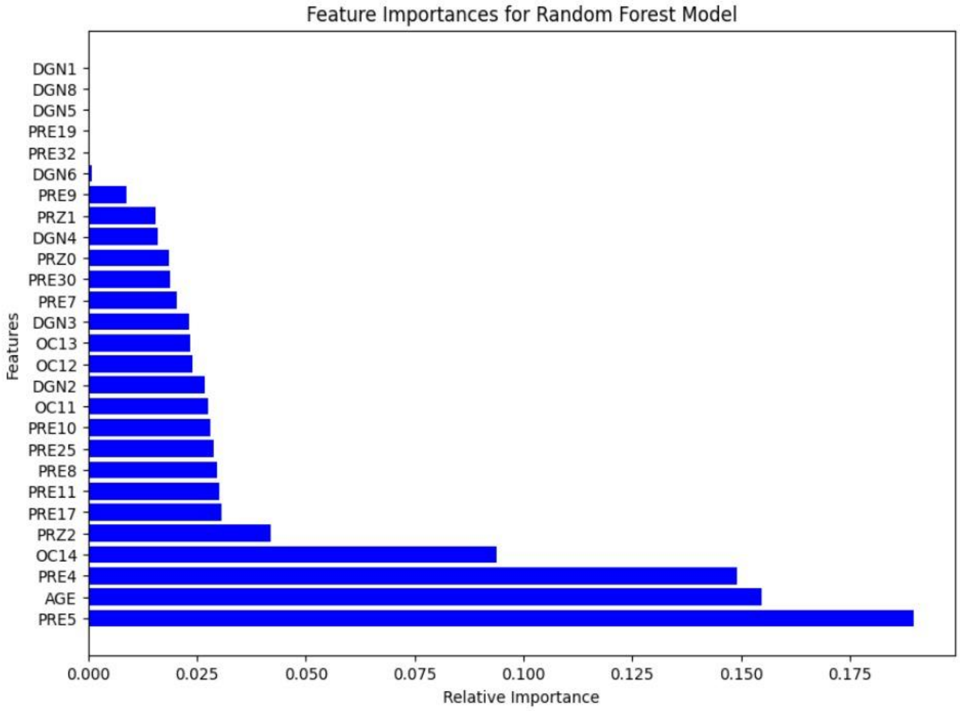


Figure 2- Feature Importance using Random Forest Model

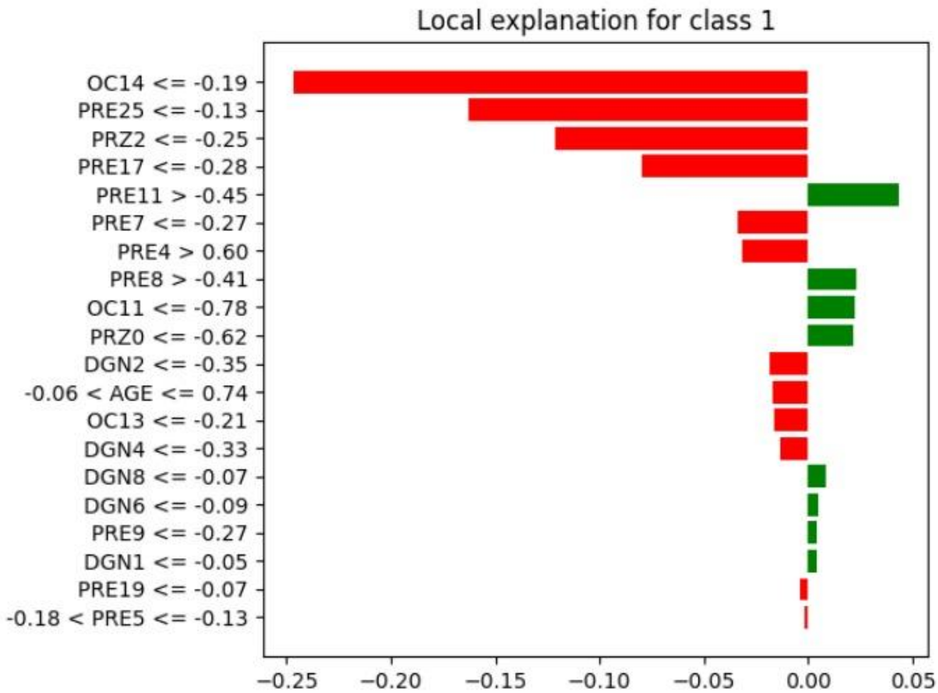


Figure 3- Bar Graph using LIME

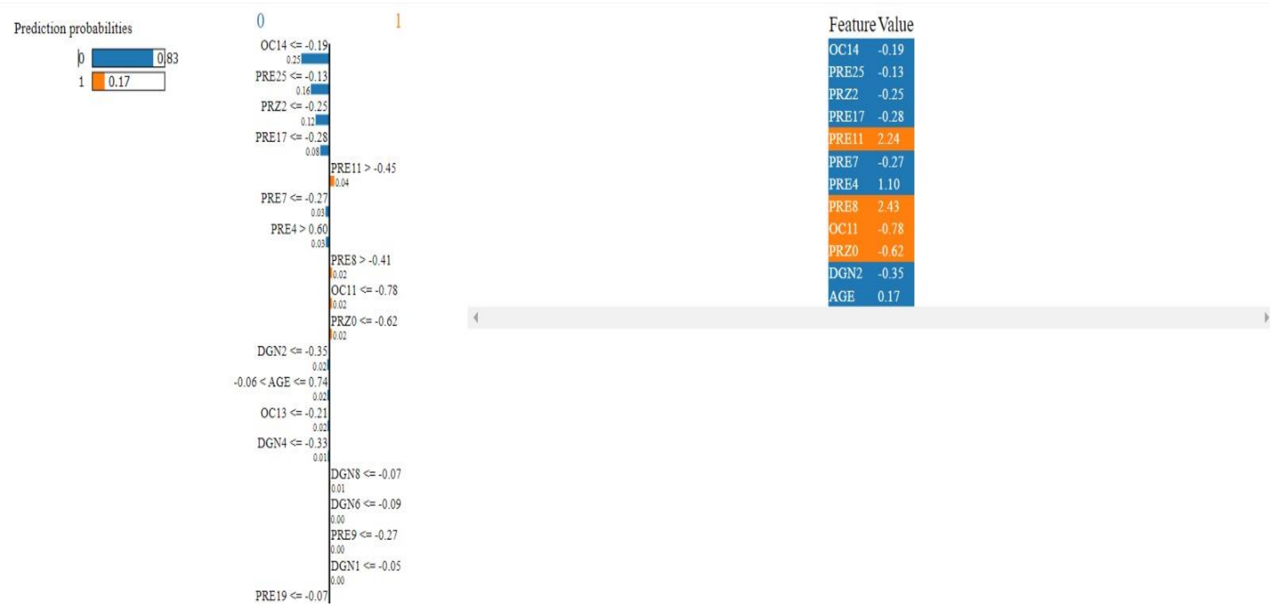


Figure 4- Final Explanation using LIME

**Interpretation of the final LIME notebook:**

The Above image shows graphs that each show essential information about their rates.

The center graph shows the feature importance scores on this particular sample with OC14 having a 25% feature importance score, PRE25 with 16%, PRZ2 with12%, PRE17 with 8%, PRE11 with 4%, PRE7 with 7%, PRE4 with 3%, PRE8 with 2%, OC11 with 2%, PRZ0 with 2%, DGN2 with 2%, OC13 with 2%, DGN4 with 1%, DGN8 with 1%, DGN6 with 0%, PRE9 with 0% DGN1 with 0%, PRE19 with 0% and PRE5 with 0%.

The right graph shows the top five features and their respective values. The features highlighted in orange contribute toward class 1 whereas features highlighted in blue contribute toward class 0.

# 6. CONCLUSION AND FUTURE SCOPE

In Conclusion, Explainable AI models enable healthcare professionals to interpret and understand the factors influencing the classification of thoracic surgery cases. By identifying critical features, these models assist in accurate diagnosis and treatment planning.
Through feature importance analysis, explainable AI reveals the key factors influencing the classification of thoracic surgery cases. This insight aids healthcare practitioners in focusing on the most relevant clinical indicators for effective decision-making.
By uncovering the rationale behind predictions, explainable AI helps in identifying potential errors or biases in the classification process. This awareness enables proactive measures to mitigate risks and improve patient outcomes.

By leveraging explainable AI, healthcare providers can tailor treatment plans more effectively based on the model's insights and recommendations. This personalized approach to patient care leads to better outcomes, reduced complications, and optimized resource allocation within healthcare systems.

Explainable AI provides transparency by revealing the reasoning behind model predictions. This transparency is crucial in the medical field, as it allows healthcare professionals to understand why a particular decision was made by the model, enhancing trust and confidence in the system.

## 7. REFERENCES

[1]. Explainable Machine Learning Solution for Observing Optimal Surgery Timings in Thoracic Surgery Diagnosis

[2]. Explainable AI Recipes: Implement Solutions to Model Explain ability and Interpretability with Python - Pradeepta Mishra. https://doi.org/10.1007/978-1-4842-9029-3

[3]. Explainable AI (XAI) has emerged as a promising solution to address the interpretability challenge of machine learning models.

[4]. https://archive.ics.uci.edu/dataset/277/thoracic+surgery+data.

[5]. Classification assessment methods - Alaa Tharwat

[6]. Explainable AI for Medical Data: Current Methods, Limitations, and Future Directions - MD IMRAN HOSSAIN, University of South Florida, USA; GHADA ZAMZMI, University of South Florida, USA; PETER R. MOUTON, SRC Biosciences, USA; MD SIRAJUS SALEKIN, University of South Florida, USA; YU SUN, University of South Florida, USA; DMITRY GOLDGOF, University of South Florida, USA.