



EXPLAINABLE ARTIFICIAL INTELLIGENCE IN HEALTHCARE

Research Proposal

Aug 2026

Abdullah S. Basaad
Research Methodology In Computer Science (AI800)

Table of contents	
Chapter 1	
Abstract	3
Introduction	3
Background	4
Problem Statement	4
Research Objectives	5
Research Scope	5
Chapter 2	
Literature Review	6
Chapter 3	
Research Methodology	11
Expected Outcomes	12
Gantt Chart	13
References	14

CHAPTER 1

1.1 ABSTRACT

During a recent visit to a doctor, I was surprised to see ChatGPT used by him to assist in answering medical questions related to my diagnosis. This experience made me reflect on the growing role of Artificial Intelligence (AI) in healthcare. While AI has become a powerful technology that improves efficiency across many fields, I wondered whether I would fully trust an AI system to evaluate my health condition and support medical decisions. This question highlights one of the biggest challenges facing AI in healthcare: trust. Artificial Intelligence has become an important technology in modern healthcare, supporting disease diagnosis, medical imaging, patient monitoring, and clinical decision-making. Although advanced machine learning and deep learning models achieve high prediction accuracy, many operate as "black-box" systems that provide little explanation for their decisions. This lack of transparency reduces trust among healthcare professionals and limits the adoption of AI in clinical practice. Explainable Artificial Intelligence (XAI) aims to address this problem by providing understandable explanations for AI predictions using techniques such as SHAP, LIME, and Grad-CAM. This research proposes a literature review of Explainable AI in healthcare to examine existing explainability techniques, their applications, strengths, limitations, and current research challenges. The study also evaluates how XAI improves transparency, trust, and decision-making while identifying research gaps and recommending future directions for developing more reliable and clinically applicable AI systems.

1.2 INTRODUCTION

Artificial Intelligence has rapidly transformed many industries, with healthcare being one of its most promising application areas. AI technologies are now used for disease diagnosis, medical image analysis, patient monitoring, treatment recommendation, and predictive analytics. These technologies have demonstrated the ability to improve diagnostic accuracy, reduce healthcare costs, and assist clinicians in making faster and more informed decisions. Despite these advantages, many modern AI models, particularly deep learning models, function as black-box systems. Although they produce highly accurate predictions, they provide little information about how these predictions are made. This lack of transparency creates challenges for healthcare professionals who must understand and justify medical decisions before applying them in clinical practice. Explainable Artificial Intelligence (XAI) has emerged as a solution to this problem by providing human-understandable explanations for AI predictions. XAI techniques help clinicians identify the factors influencing model decisions, improve confidence in AI systems, and support ethical and regulatory requirements in healthcare. As AI becomes increasingly integrated into medical practice, explainability is becoming an essential requirement rather than an optional feature.

Keywords: XAI, SHAP, LIME, Grad-CAM, Transparency, Trust

1.3 BACKGROUND

Healthcare organizations are generating large amounts of clinical data from electronic health records, medical imaging systems, wearable devices, and laboratory tests. Machine learning and deep learning algorithms have successfully used these data to predict diseases, classify medical images, estimate patient risks, and support clinical decision-making.

Although these AI models often achieve excellent predictive performance, their complex internal structures make it difficult for clinicians to understand how decisions are generated. This lack of interpretability has become one of the major barriers preventing widespread adoption of AI in healthcare, where transparency, accountability, and patient safety are critical.

To address this challenge, Explainable Artificial Intelligence (XAI) techniques such as SHAP, LIME, Grad-CAM, and other explanation methods have been developed. These approaches provide insights into model behavior by identifying the features or image regions that contribute most to predictions. Previous research has shown that XAI can improve clinician trust, support decision-making, and satisfy ethical and regulatory requirements. However, many existing studies rely on limited datasets, lack standardized evaluation methods, and provide little validation using real clinical environments. Therefore, further research is needed to review current XAI techniques, evaluate their effectiveness, identify research gaps, and recommend future directions for transparent and trustworthy AI systems in healthcare.

1.4 PROBLEM STATEMENT

Artificial Intelligence (AI) became an essential technology in modern healthcare, supporting applications such as disease diagnosis, medical imaging, patient monitoring, and clinical decision-making. The recent improvements in machine learning and deep learning have improved the accuracy of healthcare prediction models. However, even though with high performance, many AI systems operate as black-box models, making it difficult for healthcare professionals to understand how predictions are generated. Aravindkumar et al. stated that numerous studies have shown that models can achieve accurate predictions despite being based on incorrect assumptions (2026). This lack of transparency reduces clinicians' confidence in AI systems and limits their adoption in real-world healthcare environments (Dave et al., 2020; Kenigson, 2025).

Existing Explainable Artificial Intelligence (XAI) techniques, including SHAP, LIME, and Grad-CAM, is proposed to improve the interpretability of AI models by providing explanations for their predictions. These methods enable clinicians to identify the factors influencing diagnostic decisions and improve trust in AI-assisted healthcare systems (Panda et al., 2023; Amin et al., 2024). Although these techniques have shown good results, several challenges remain. Most existing studies rely on publicly available datasets with limited patient diversity, while only a small number evaluate their models using real clinical data or involve healthcare professionals in assessing the usefulness of AI explanations (AlKhanbouli et al., 2025).

Developing trustworthy and interpretable AI systems is becoming important because healthcare decisions directly affect patient safety and cure outcomes. Explainable AI can improve transparency, support clinical decision-making, reduce diagnostic errors, and encourage healthcare professionals to adopt AI technologies with greater confidence. Furthermore, transparent AI systems help satisfy ethical, legal, and regulatory requirements for medical applications, making explainability an essential component of future healthcare AI systems (Kenigson, 2025).

Several studies have attempted to address the black-box problem by integrating explainability methods with machine learning and deep learning models. SHAP and LIME have become the most widely adopted techniques for explaining structured healthcare data, while Grad-CAM is commonly applied to medical imaging applications. Although these approaches improve model transparency, most studies still focus on prediction accuracy rather than evaluating the quality and clinical usefulness of the generated explanations. Additionally, there is currently no standardized framework for evaluating explainability across different healthcare applications (Aravindkumar et al., 2026).

Therefore, more researching is needed to investigate the current state of Explainable Artificial Intelligence in healthcare by reviewing existing explainability techniques, identifying their strengths and limitations, and examining how they can be effectively integrated into clinical practice. Such research will help identify current knowledge gaps and provide recommendations for developing more transparent, trustworthy, and clinically applicable AI systems that better support healthcare professionals in making informed medical decisions (AlKhanbouli et al., 2025; Aravindkumar et al., 2026).

1.5 RESEARCH OBJECTIVES

The primary objective of this research is to investigate the current applications of Explainable Artificial Intelligence in healthcare and evaluate how explainability techniques improve the transparency and trustworthiness of AI-based medical decision-making.

The specific objectives of this research are:

- **To identify the Explainable Artificial Intelligence techniques currently used in healthcare applications.**
 - Will help finding strengths and weakness for each technique, and therefore exploring ways to improve.
 - Differentiate between the techniques will help to compare between them.
- **To examine the machine learning and deep learning models integrated with Explainable AI methods.**
 - To help validating the outcome of XAI.
- **To compare the strengths and limitations of existing Explainable AI techniques used in healthcare.**
 - To help deciding which technique is most suitable in application cases.
 - To help avoiding the use of a particular technique in some cases and prefer others.

- **To evaluate the research methodologies and datasets employed in previous studies.**
 - To make sure results were actually generated from accurate assumption and valid data.
 - To ensure covering all possible population and variabilities.
- **To identify the current challenges and research gaps affecting the adoption of Explainable Artificial Intelligence in healthcare.**
 - To guide the way for future improvements.
- **To recommend future research directions for developing more transparent, interpretable, and trustworthy AI systems for healthcare applications.**
 - To provide practical solutions that could close gaps found.

1.6 RESEARCH SCOPE

This research focuses on reviewing the current state of Explainable Artificial Intelligence (XAI) in healthcare. The study examines published research on explainability techniques, including SHAP, LIME, Grad-CAM, and other commonly used XAI methods, and evaluates their applications in healthcare domains such as disease diagnosis, medical imaging, clinical decision-making, and patient monitoring.

The review includes studies that integrate XAI with machine learning and deep learning models to improve the transparency and interpretability of AI systems. It also compares the strengths, limitations, datasets, evaluation methods, and research methodologies used in previous studies. Furthermore, the research identifies current challenges, research gaps, and future directions for developing trustworthy and clinically applicable AI systems.

CHAPTER 2

2.1 LITERATURE REVIEW

1. Explainable AI Meets Healthcare: A Study on Heart Disease Dataset

The paper explains how Explainable Artificial Intelligence (XAI) can improve the accuracy of machine learning models used for heart disease prediction. It states that even though AI models can achieve high accuracy, their hidden way of work makes it difficult for healthcare professionals to understand and trust the predictions. To overcome this issue, the study applies several XAI techniques to a heart disease prediction model built using a heart disease dataset and a classification algorithm. The paper uses many explanation methods, such as LIME and SHAP, Anchors, Counterfactual Explanations, Counterfactuals with Prototypes, Contrastive Explanation Method (CEM), Kernel SHAP, and Integrated Gradients. Using these methods, the study shows how different patient features contribute to prediction outcomes and provides explanations for the model's decisions. The results show that factors such as the number of blocked blood vessels, chest pain type, cholesterol level, and thalassemia have the greatest influence on predicting heart disease. Integrating techniques with machine learning models can increase transparency, helps

clinicians to better understand AI predictions, and increase trust in using AI for medical decisions.

2. Explainable Artificial Intelligence for Healthcare Applications using Random Forest Classifier with LIME and SHAP

The paper provides an overview of Explainable Artificial Intelligence (XAI) and describes its application in healthcare using a Random Forest classifier combined with LIME and SHAP explainability techniques. The paper highlights the increasing adoption of AI in healthcare and emphasize that many machine learning and deep learning models operate as black boxes making it difficult for users to understand how predictions are generated. To address this issue, the paper reviews XAI concepts including explainability, transparency, interpretability, trustworthiness, and quality evaluation criteria. The experimental study uses a publicly available diabetes dataset from Kaggle, where a Random Forest classifier is employed to predict diabetes, and SHAP and LIME are applied to explain model behavior. The study demonstrates that integrating XAI techniques with machine learning models improves transparency, interpretability, and trust without affecting prediction quality. The authors conclude that explainable AI has significant potential to support healthcare decision-making by allowing clinicians to better understand AI-generated predictions. They also suggest future research should explore more explainability techniques, and enhance the trustworthiness and practical deployment of explainable AI systems in real-world healthcare environments.

3. Explainable AI in Healthcare: A Systematic Review of XAI Use Cases in Imaging, Diagnostics, and Rehabilitation

The paper answers the question of how Explainable Artificial Intelligence (XAI) is being used to increase transparency and trust in healthcare AI systems. The authors examined 36 studies in the last 5 years on three healthcare domains: medical imaging, clinical diagnosis, and rehabilitation. The study emphasizes that traditional machine learning and deep learning models operate mysteriously which leads to complicate the understanding of the reasoning behind predictions for healthcare professionals. To overcome this challenge, XAI techniques such as SHAP, LIME, and Grad-CAM can be used to describe model decisions. The findings have proofed that Grad-CAM is commonly used in medical imaging to generate visual heatmaps, but SHAP and LIME are primarily used in diagnosis and rehabilitation to explain individual patient predictions. Several studies have combined multiple explanation methods to improve reliability and clinician confidence. However, the paper also mentions several limitations such as limited external validation, inconsistent evaluation of explanation quality, reliance on single-center datasets, and the lack of standardized reporting and faithfulness metrics. It then suggests that future research should focus on developing standardized frameworks for evaluating XAI, generating dataset diversity and integrating explainability into real clinical workflows.

4. An Explainable AI Framework for Artificial Intelligence of Medical Things (AIoMT)



The paper presents a new Explainable Artificial Intelligence (XAI) framework intended to increase the transparency, reliability, and accuracy of AI-driven healthcare systems, especially for brain tumor detection using MRI images. The proposed framework combines three popular XAI techniques: LIME, SHAP, and Grad-CAM to deliver understandings for the model's predictions, allowing healthcare professionals to understand why an analysis was made instead of depending on a unexplained system. The architecture introduces cloud and edge computing within an Artificial Intelligence of Medical Things (AIoMT) environment, allowing efficient processing of large medical datasets while supporting simultaneous diagnosis. The ensemble model combines predictions from five convolutional neural networks, leading to improved performance and reduced prediction errors. Experimental evaluation using the Br35H brain tumor MRI dataset demonstrated outstanding performance, achieving 99% training accuracy, 98% validation accuracy, and high precision, while visual explanations generated by SHAP, LIME, and Grad-CAM emphasized the image regions and features influencing each prediction. These explanations increase clinician trust, improve interpretability, and support more informed medical decision-making. The paper showed that integrating explainable AI with AIoMT has significant potential to enhance healthcare outcomes by providing accurate, transparent, and trustworthy diagnoses. It also directs future work to focus on extending the framework to other diseases, improving explainability methods and incorporating real-time patient data.

5. The Role of Explainable Artificial Intelligence in Disease Prediction

This paper defines how Explainable Artificial Intelligence (XAI) is moving transparency and trust in AI-based disease prediction systems to higher levels. Using the PRISMA methodology, 30 studies published in the last years were analyzed, covering cardiovascular diseases among other diseases. In result, SHAP and LIME were found as the most used XAI techniques because they assist clinicians to understand how AI models get to their predictions, making diagnostic decisions more transparent and trustworthy. Traditional AI has proved its accuracy in predicting diseases, however, clinicians find it hard to give their full trust to it as they do not know how it really works. The paper numbers many key challenges, including limited dataset diversity, reliance on single data modalities, and rarity of user-friendly interfaces for healthcare professionals. To avoid these challenges, the paper suggests the following: standardized XAI frameworks, more diverse datasets, improved clinician training, and continued researching into transparent AI systems that can be integrated into everyday medical practice.

6. Explainable AI for Healthcare: Developing Transparent and Interpretable Models for Medical Diagnosis

This paper explores the growing importance of Explainable Artificial Intelligence (XAI) in making healthcare AI systems more transparent, trustworthy, and clinically useful. While advanced machine learning and deep learning models achieve high diagnostic accuracy in applications, their "black-box" nature limits physician trust and practical adoption. The paper reviews model-agnostic explanation methods, including LIME, SHAP, and LORE and Model-specific such as decision trees and Attention Mechanisms, explaining how these techniques help clinicians understand the reasoning behind AI predictions. Through a comprehensive review of existing studies and real-world healthcare applications, the authors demonstrate that explainable AI improves clinical confidence, patient safety while maintaining high accuracy diagnostic performance. Although XAI models may experience a slight reduction in predictive accuracy

compared to traditional black-box models, the increased transparency enhances their acceptance in clinical practice. The paper concludes that explainability should be considered a fundamental requirement rather than an optional feature in medical AI, as it enables healthcare professionals to validate AI-generated recommendations, supports ethical and legal standards, and promotes patient-centered decision-making. The authors recommend researching AI models that combine the predictive strength of deep learning with explainability techniques to accelerate the responsible adoption of AI in healthcare.

7. A Perspective on Explainable Artificial Intelligence Methods: SHAP and LIME

The paper compares SHAP and LIME, focusing on their ability to improve the transparency of machine learning models, particularly in healthcare. The authors emphasize that although artificial intelligence models can achieve high predictive accuracy, their "black-box" nature makes it difficult for clinicians to understand how predictions are generated. Explainability methods such as SHAP and LIME address this issue by identifying the features that make the model's predictions, therefore, increasing transparency and user trust. The paper compares the two techniques, explaining that SHAP provides both global and local explanations based on game theory, while LIME generates local explanations by approximating complex models with simpler interpretable models.

To demonstrate their practical application, the authors use biomedical case studies involving myocardial infarction prediction and analyze how SHAP and LIME explain the decisions of different machine learning models. The results show that both methods successfully identify influential clinical features, helping users understand AI predictions. However, the study also found out important limitations. It found that SHAP and LIME do not always produce the same explanations because their results depend on the machine learning model and the relationship between features. This can make the explanations less reliable and reduce clinicians' trust in AI-based healthcare decisions.

The paper concludes that while SHAP and LIME significantly improve the interpretability of healthcare AI systems, they should not be used without understanding their assumptions and limitations. The authors recommend presenting explanations alongside clear descriptions of how they are generated, comparing multiple machine learning models, and applying additional techniques to account for correlated features. They argue that explainability should be implemented carefully to ensure AI systems remain transparent, trustworthy, and suitable for supporting clinical decision-making in healthcare.

8. Evaluation of Explainable AI by Medical Experts: a Survey of the Existing Approaches

This paper reviews how Explainable Artificial Intelligence is evaluated by medical professionals in healthcare. Instead of focusing on developing new AI models, the authors examine whether doctors actually understand, trust, and benefit from the explanations provided by XAI systems. The study explains that AI is increasingly used in healthcare for disease diagnosis, medical imaging, treatment planning, and patient monitoring. However, many AI

models are difficult to understand because they work as "black boxes." XAI techniques such as SHAP, LIME, and Grad-CAM help solve this problem by explaining how AI reaches its decisions, making the results easier for healthcare professionals to interpret and trust.

The authors reviewed around 40 studies that evaluated XAI with medical practitioners. They found that most studies used deep learning models for tasks such as disease classification and medical image analysis, while post-hoc explanation methods like SHAP, LIME, and Grad-CAM were the most commonly used. However, many studies had important weaknesses. Most involved only a small number of doctors, did not compare different XAI methods, and often did not perform statistical testing to confirm their results. As a result, many papers reported positive findings without providing strong evidence that the explanations were truly effective in real clinical settings.

The paper concludes that explainability is essential for increasing trust and transparency in healthcare AI, but current evaluation methods need improvement. The authors recommend using larger studies, involving real clinical decision-making, comparing multiple XAI techniques, and applying better evaluation standards. These improvements would help ensure that XAI systems are more reliable, understandable, and useful for supporting healthcare professionals in making safe and informed medical decisions.

9. Explainable Artificial Intelligence for Human Decision-Support System in Medical Domain

This paper studies how Explainable Artificial Intelligence can improve decision support in the medical field. The authors focus on helping doctors trust AI systems that analyze medical images. They explain that deep learning models, such as CNNs, can detect diseases with high accuracy, but they often work as black boxes, meaning users cannot understand how the model makes its decisions. Because of this, doctors may hesitate to rely on AI for important medical decisions. The paper shows that XAI can solve this problem by making AI predictions more transparent and easier to understand.

The researchers trained a CNN model to detect bleeding in medical images. They then applied three explainable AI methods: LIME, SHAP, and CIU. These methods highlight the important parts of an image that influenced the AI's prediction, allowing users to understand why the model classified an image as normal or bleeding. The explanations were compared to see which method provided the clearest and most useful support for human decision-making.

To evaluate the methods, the authors conducted user studies where participants used the explanations to interpret medical images. The results showed that all three XAI methods helped users understand AI predictions, but CIU provided the clearest explanations and supported better decision-making than LIME and SHAP. It also generated explanations faster than the other methods. The study concludes that explainable AI can increase transparency, improve trust in AI systems, and make AI more useful for healthcare professionals.

10. Survey of Explainable AI Techniques in Healthcare

This paper reviews the use of Explainable Artificial Intelligence (XAI) in healthcare and explains how XAI can make AI models more understandable and trustworthy for doctors. The authors explain that many deep learning models can predict diseases with high accuracy, but they work as black boxes, meaning doctors cannot easily understand how the model reaches its decision. Since medical decisions are very important, healthcare professionals need clear explanations before trusting AI predictions. XAI helps solve this problem by showing why a model made a certain prediction, which improves transparency, trust, and confidence in AI systems.

The paper describes several popular XAI techniques used in healthcare, including LIME, SHAP, Grad-CAM, Saliency Maps, and Attention mechanisms. These methods explain which features of medical images are most important for predicting diseases. The authors also explain that XAI can be applied to medical images and health records to help doctors better understand AI predictions and support clinical decision-making.

The paper also discusses the current challenges of using XAI in healthcare. These include model bias, privacy, ethics, explanation quality, and the need for explanations that are easy for clinicians to understand. The authors conclude that XAI is essential for increasing trust in AI-based disease prediction systems. They recommend developing more reliable and user-friendly XAI methods that can be safely used in real clinical environments.

CHAPTER 3

3.1 RESEARCH METHODOLOGY

This study will adopt a qualitative research approach to investigate the role of Explainable Artificial Intelligence (XAI) in healthcare. The research aims to examine how XAI techniques improve the transparency, interpretability, and trustworthiness of AI systems used for medical diagnosis, disease prediction, medical imaging, and clinical decision-making.

The research will begin with an extensive review of recent academic publications obtained from databases such as PubMed, MDPI, and Google Scholar. The collected studies will help identify current XAI techniques, healthcare applications, existing challenges, and research gaps.

Following the literature review, the identified XAI techniques, including SHAP, LIME, and Grad-CAM, will be analyzed and compared based on several criteria, such as explainability, interpretability, accuracy, usability, computational complexity, and suitability for different healthcare applications. Their advantages, limitations, and practical use in clinical environments will also be evaluated.

The collected information will then be analyzed to identify current trends, common challenges, and opportunities for future research. Finally, recommendations will be proposed to support the development of more transparent, reliable, and trustworthy AI systems in healthcare.

The study is limited to published academic research and does not involve collecting patient data or conducting clinical experiments. Since only publicly available research articles are used, no ethical concerns related to human participants are expected. All referenced materials will be properly cited to maintain academic integrity.

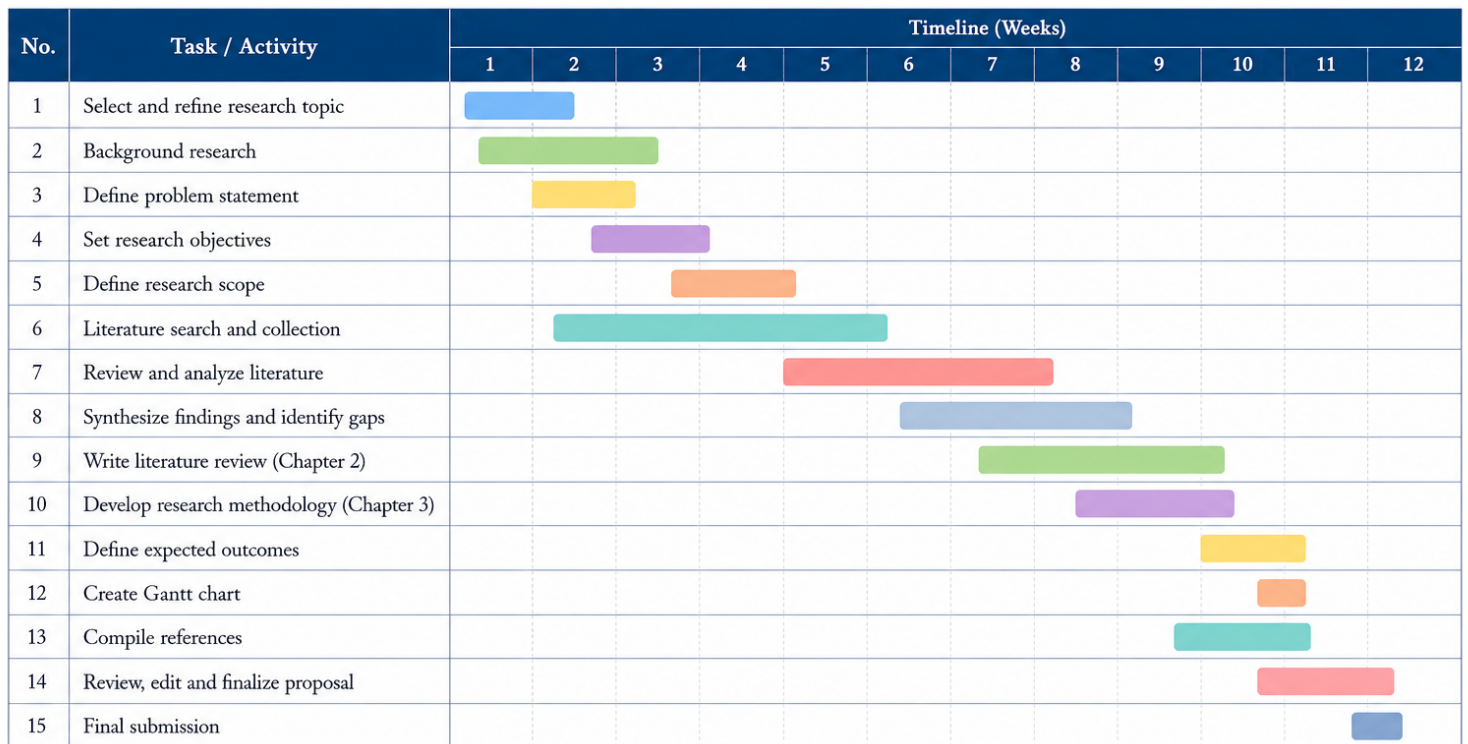
3.2 EXPECTED OUTCOMES

This research is expected to provide a comprehensive understanding of the current applications of Explainable Artificial Intelligence (XAI) in healthcare. It is expected to identify the most commonly used explainability techniques, such as SHAP, LIME, and Grad-CAM, and evaluate their role in improving the transparency and interpretability of AI systems.

The study is also expected to identify the strengths, limitations, and challenges of applying XAI in healthcare, particularly in areas such as disease diagnosis, medical imaging, and clinical decision-making. Furthermore, the research is expected to highlight existing research gaps and suggest possible directions for future studies.

The findings are expected to contribute to the existing body of knowledge by providing researchers and healthcare professionals with a clearer understanding of how XAI can improve trust, transparency, and the adoption of AI systems in healthcare. Ultimately, the study aims to support the development of more reliable, understandable, and trustworthy AI solutions for healthcare applications.

3.3 GANTT CHART



REFERENCES

- AlKhanbouli, R. R., Almadhaani, H. M., Alhosani, F., & Simsekler, M. C. E. (2025). The role of explainable artificial intelligence in disease prediction: A systematic review. *BMC Medical Informatics and Decision Making*, 25. <https://doi.org/10.1186/s12911-025-02944-6>
- Amin, A., Hasan, K., Zein-Sabatto, D., Chimba, S., Ahmed, I., & Islam, T. (2024). An explainable AI framework for Artificial Intelligence of Medical Things. *arXiv*. <https://arxiv.org/abs/2403.04130>
- Aravindkumar, A., Ramadoss, M., Ahmed, S., Ahmed, F., Sampath, V., & Lakshminarayanan, K. (2026). Explainable AI in healthcare: A systematic review of XAI use cases in imaging, diagnostics, and rehabilitation. *Healthcare*, 14(1), 92. <https://doi.org/10.3390/healthcare14010092>
- Babakov, N., Rezgova, E., Reiter, E., & Bugarín, A. (2025). Evaluation of Explainable AI by medical experts: A survey of the existing approaches. In *Proceedings of the Second Workshop on Explainable Artificial Intelligence for the Medical Domain (EXPLIMED 2025)* (Article 8). CEUR-WS. <https://ceur-ws.org/Vol-4059/paper8.pdf>
- Chaddad, A., Peng, J., Xu, J., & Bouridane, A. (2023). Survey of Explainable AI techniques in healthcare. *Sensors*, 23(2), 634. <https://doi.org/10.3390/s23020634>
- Dave, D., Naik, H., Singh, S., & Patel, P. (2020). Explainable AI meets healthcare: A study on heart disease dataset. *arXiv*. <https://arxiv.org/abs/2011.03195>
- Kenigson, J. (2025). Explainable AI for healthcare: Developing transparent and interpretable models for medical diagnosis. ResearchGate. https://www.researchgate.net/publication/394917443_explainable_ai_for_healthcare_developing_transparent_and_interpretable_models_for_medical_diagnosis
- Knapič, S., Malhi, A., Saluja, R., & Främling, K. (2021). Explainable Artificial Intelligence for human decision-support system in medical domain. *Machine Learning and Knowledge Extraction*, 3(3), 740–770. <https://arxiv.org/abs/2105.02357>
- Panda, M., Mahanta, S. R., & Mahanta, R. (2023). Explainable artificial intelligence for healthcare applications using Random Forest classifier with LIME and SHAP. *arXiv*. <https://arxiv.org/abs/2311.05665>
- Salih, A., Raisi-Estabragh, Z., Boscolo Galazzo, I., Radeva, P., Petersen, S. E., Menegaz, G., & Lekadir, K. (2023). A perspective on explainable artificial intelligence methods: SHAP and LIME. *arXiv*. <https://arxiv.org/abs/2305.02012>