

Unveiling the Black Box: Explainable AI in Diabetes and Cancer Care

Rishi Jain¹, N. K. Tripathi¹

1. Asian Institute of Technology, Thailand

*Corresponding email: rishijain@ait.asia

Abstract: The integration of artificial intelligence (AI) into healthcare has yielded significant advancements in diagnostics, prognostics, and personalised treatment pathways, particularly in the domains of diabetes and cancer care. However, the opaque “black-box” nature of many AI models hinders their clinical adoption due to limited interpretability and trust. This review explores the transformative potential of Explainable Artificial Intelligence (XAI) in enhancing transparency and accountability within AI-driven medical systems. Focusing on diabetes and oncology, the study systematically synthesises literature published between 2021 and 2025, highlighting how XAI techniques—such as SHAP, LIME, Grad-CAM, and federated learning—bridge the gap between complex algorithms and clinical interpretability. In diabetes care, XAI enables the visualisation of critical predictors like HbA1c levels, BMI, and glucose trends, thereby improving diagnostic precision, insulin dosing, and complication forecasting. In cancer research, XAI supports early detection, subtype classification, and personalised drug recommendations across various malignancies including breast, lung, colorectal, cervical, and ovarian cancers. The review also delves into bibliometric analyses, revealing a global surge in XAI-related healthcare research, with India leading the publication count. Additionally, the paper addresses ethical, regulatory, and equity considerations, emphasising how XAI can detect bias, support FDA and EMA compliance, and enhance model generalizability. By showcasing use cases across multi-modal data applications—spanning genomics, imaging, and metabolomics—the study underlines XAI’s critical role in fostering clinician trust, patient safety, and transparent AI governance. Recommendations for future research include developing clinically integrated XAI tools, advancing real-world validations, and supporting collaborative, privacy-preserving data ecosystems. Ultimately, XAI is positioned not merely as a technical add-on but as a foundational enabler for trustworthy, human-centric AI in healthcare.

Keywords: Explainable AI (XAI), Diabetes Prediction, Cancer Diagnosis, Model Interpretability, Healthcare AI Trust

1. Introduction

Artificial intelligence (AI) has become a transformative force in healthcare, driving advancements in disease diagnosis, treatment optimisation, and personalized medicine. Machine learning and deep learning models are now extensively used in clinical decision-making, offering unprecedented accuracy in detecting diseases, predicting patient outcomes, and recommending targeted therapies. However, despite these impressive capabilities, the “black-box” nature of AI systems poses a significant barrier to widespread adoption in clinical settings. Healthcare professionals and patients alike struggle to trust AI-driven recommendations when the underlying decision-making process is opaque and unexplainable.

Explainable AI (XAI) emerges as a crucial solution to this challenge, providing transparency into how AI models arrive at specific conclusions. By making AI-generated decisions interpretable, XAI fosters trust, accountability, and regulatory compliance, ensuring that clinicians can validate AI-driven insights before applying them to patient care. In high-stakes medical applications such as diabetes and cancer management, where early detection and precise treatment are paramount, XAI plays an essential role in bridging the gap between cutting-edge AI and real-world clinical practice.

2. The Global Burden of Diabetes and Cancer

Diabetes and cancer are among the most pressing global health challenges today. According to the World Health Organization (WHO) [1] ([who.int+1nypost.com+1](https://www.who.int+1nypost.com+1)), the number of people living with diabetes rose from 200 million in 1990 to 830 million in 2022, with prevalence increasing more rapidly in low- and middle-income countries. In the United States, the Centers for Disease Control and Prevention (CDC)[2] (cdc.govdiabetes.org) reported that approximately 38 million people have diabetes, with 1 in 5 unaware of their condition. The American Diabetes Association (ADA) [3] further highlights that 11.6% of the U.S. population, or 38.4 million Americans, had diabetes in 2021, with 8.7 million cases undiagnosed.

Cancer remains a leading cause of death worldwide, accounting for nearly 10 million deaths in 2020, or nearly one in six deaths, as reported by the WHO [4] ([who.int](https://www.who.int)). In the United States, the CDC [2] (cdc.gov+1gis.cdc.gov+1) documented 1,777,566 new cancer cases in 2021 and 608,366 cancer deaths in 2022. Projections by the American Cancer Society estimate that in 2025, there will be approximately 2,041,910 new cancer cases and 618,120 cancer deaths in the U.S. cancerstatisticscenter.cancer.org [5]

3. XAI in Diabetes Care

Diabetes is a chronic metabolic disorder affecting millions worldwide, requiring continuous monitoring and management to prevent complications. AI has significantly improved diabetes care through predictive analytics, personalized treatment plans, and automated insulin delivery systems. However, without transparency, clinicians may hesitate to rely on AI-driven recommendations, fearing misdiagnoses or inaccurate insulin dosage predictions. XAI addresses this issue by offering clear, interpretable insights into the factors influencing AI's decisions, enabling informed medical interventions.

One of the key areas where XAI has shown promise is in predicting diabetic complications, such as diabetic retinopathy (DR). AI-driven models analyze clinical, biochemical, and metabolomic data to assess DR progression, identifying critical biomarkers like HbA1c, triglyceride levels, and retinal vessel abnormalities. XAI techniques, such as SHAP (SHapley Additive Explanations) and LIME (Local Interpretable Model-agnostic Explanations), highlight which biomarkers contribute most significantly to a diagnosis, allowing physicians to validate AI's recommendations and adjust treatment plans accordingly.

Similarly, in continuous glucose monitoring (CGM) systems, AI models predict blood sugar fluctuations and suggest insulin dosage adjustments based on real-time data. XAI enhances these systems by explaining how factors like recent carbohydrate intake, physical activity, and stress levels influence glucose predictions. This interpretability is crucial for clinicians and patients, ensuring that insulin dosing decisions are not blindly followed but are understood and verified based on medical rationale.

Beyond disease management, XAI also contributes to diabetes prevention by analyzing genetic, lifestyle, and behavioural data to identify individuals at high risk of developing Type 2 diabetes. AI models trained on large-scale datasets can detect subtle patterns in dietary habits, physical activity levels, and genetic predisposition, providing personalized recommendations for lifestyle modifications. XAI ensures that these predictions are not only accurate but also interpretable, enabling individuals and healthcare providers to take proactive steps toward diabetes prevention.

3.1 Explainable AI Models for Diabetes Prediction and Management

Recent studies highlight the role of XAI in diabetes prediction and management, with various researchers developing models that enhance interpretability. For instance, [6] introduced a diabetes prediction model that integrates XAI to visualize decision-making processes, using methods like SHAP or LIME to clarify the influence of individual features (e.g., fasting glucose or BMI) on the model's predictions. This is important because it ensures that clinicians understand which factors contribute to the prediction of diabetes risk, making the model more transparent and less of a "black box."

Similarly, [7] developed an interpretable ML model for diabetes detection, emphasizing the importance of feature explainability in medical AI applications. Moreover, [8] proposed an ensemble approach combining multiple models for diabetes management. The integration of XAI allows for more robust predictions and explains how individual patient data influences treatment recommendations, thus enhancing model interpretability for clinical decision-making.

3.2 Explainable AI in Handling Imbalanced Data and Early Detection

Handling imbalanced datasets is another significant challenge in diabetes prediction, as imbalanced data can lead to biased predictions. [9] addressed this issue using techniques like SMOTE oversampling and Near Miss under sampling, integrating XAI to gain insights into how the model handles imbalanced data and which features drive the predictions. This is critical in medical applications where dataset imbalance is a common concern, and understanding the model's handling of such data is key to ensuring trust and transparency. In the realm of early detection, [10] explored the potential of XAI in the early detection and progression analysis of Type 2 Diabetes Mellitus (T2DM) using publicly available datasets. They employed machine learning algorithms alongside XAI techniques like Shapley values and Decision Trees to identify key biomarkers for T2DM detection.

Additionally, [11] used Random Forest and XGBoost models, integrating SHAP values for feature selection, to predict the risk of Gestational Diabetes Mellitus (GDM), enhancing early detection and intervention strategies. Similarly, [12] proposed an ensemble local explainable agnostic model for diabetes prediction, achieving 81% accuracy with the Pima Indian diabetes dataset.

[13] also contributed by focusing on the integration of XAI with deep learning models for early detection of diabetic retinopathy. Their work incorporated explainability into neural networks, providing visual explanations that helped ophthalmologists understand the decision-making process behind retinopathy diagnoses, thus enhancing trust in automated systems.

These advancements collectively underscore the growing importance of Explainable AI in transforming diabetes prediction, diagnosis, and management, offering a clearer path toward personalized and transparent healthcare. Future work should aim at incorporating real-world clinical data, improving model generalizability across diverse populations, and developing user-friendly AI-driven monitoring systems to support patient care.

4. XAI in Cancer Diagnosis and Treatment

In oncology, AI has revolutionized cancer diagnosis, prognosis, and treatment planning by analyzing medical imaging, genomic profiles, and histopathological slides with remarkable precision. However, the complexity of deep learning models in cancer detection often results in opaque decision-making, where clinicians struggle to understand why an AI model classifies a lesion as malignant or benign. XAI is critical in addressing this issue, offering transparency into AI-driven cancer diagnostics and ensuring that predictions align with clinical expertise.

[14] systematically reviewed the role of Explainable AI (XAI) in cancer care using 69 studies from 2020 to 2024 following PRISMA guidelines. Key findings revealed limited clinician involvement (83% of studies) and a lack of rigorous XAI evaluation (87%), impacting clinical relevance. The dominance of post-hoc methods (SHAP, LIME, Grad-CAM) highlights challenges in reliability and real-world application. The study underscores the need for standardized metrics, clinician-centric interfaces, and integration frameworks to enhance AI-driven decision-making in oncology.

For instance, in breast cancer detection, AI models analyze mammograms, and pathology slides to identify malignancies. XAI techniques, such as saliency maps and attention mechanisms, highlight the specific regions in an image—such as microcalcifications or irregular tumor margins—that influence AI’s classification. This allows radiologists and oncologists to verify AI’s findings, reducing false positives and negatives and improving diagnostic confidence.

Similarly, in precision oncology, AI models analyze genomic data to recommend targeted therapies for cancer patients. By ranking genetic mutations (e.g., BRCA1/2 in breast cancer or EGFR mutations in lung cancer), AI can suggest personalized treatment plans tailored to a patient’s molecular profile. XAI ensures that these recommendations are not only accurate but also understandable, allowing oncologists to make informed decisions about the most effective therapies.

Explainable AI (XAI) is vital for trustworthy decision-making in smart healthcare. [15] explore XAI models for cancer image classification, addressing interpretability challenges and proposing a novel framework integrating deep learning with explainability techniques. Their model achieves 97.72% accuracy, 90.72% precision, and 96.72% F1-score, enhancing AI reliability in healthcare.

Another crucial application of XAI in oncology is in cancer prognosis and recurrence prediction. AI models trained on vast datasets assess tumour genomics, patient medical histories, and imaging data to estimate the likelihood of cancer recurrence. XAI techniques clarify which factors—such as tumour size, lymph node involvement, or specific biomarkers—contribute most significantly to recurrence risk. This transparency is invaluable for oncologists and patients, facilitating shared decision-making and personalized follow-up strategies.

4.1 Explainable AI in Lung and Colorectal Cancer Diagnosis

In recent years, Convolutional Neural Networks (CNNs) have become widely utilized for cancer detection due to their ability to effectively learn spatial hierarchies from medical images. [16] proposed a Federated Learning (FL)-based model for detecting colorectal cancer using histopathological images. This approach not only enhances privacy by utilizing decentralized datasets but also improves the accuracy of cancer classification. Their use of Explainable AI (XAI) methods, such as visualizing features using super pixels, helps in understanding model decisions. They experimented with multiple CNN architectures, finding that ResNeXt50 provided the highest accuracy (99.53%), which was further optimized using FL, resulting in a 96.04% accuracy and an F1 Score of 0.96.

In the context of lung cancer, [17] applied XAI to a stack-ensemble machine learning model to investigate spatial variations in the contributions of risk factors to lung and bronchus cancer (LBC) mortality rates in the United States. Their model outperformed traditional base learners like Random Forest (RF) and Gradient Boosting Machine (GBM). Using a permutation-based feature importance technique, they identified smoking prevalence as the most significant factor, followed by poverty and elevation. They also showcased the spatial heterogeneity of LBC mortality, providing valuable insights into how these risk factors vary across counties.

Further advancing lung cancer diagnosis, [18] explored the effectiveness of various machine learning models, including Support Vector Machines (SVM), Random Forest (RF), and decision trees. They employed XAI techniques, such as LIME and SHAP, to understand model predictions and their underlying logic. This approach enhanced the interpretability of the models, shedding light on how specific features, such as symptoms and lifestyle choices, contributed to lung cancer predictions.

In a similar vein, [19] introduced a multi-scale CNN model for lung and colon cancer classification. Enhanced with XAI techniques like Grad-CAM, the model enables visualization of critical regions in medical images, thus providing clinicians with valuable insights into the classification process. By integrating XAI, their model ensures transparency, helping healthcare professionals understand the rationale behind predictions and improving trust in AI-based diagnoses.

This body of work underscores the significant potential of AI and XAI in enhancing the accuracy, transparency, and interpretability of lung and colorectal cancer detection systems, fostering trust among clinicians and advancing early detection methods

4.2 Explainable AI in Breast Cancer Biomarker Discovery

Breast cancer is a leading cause of cancer-related deaths, and recent advancements in multi-omic data have helped unravel its molecular complexity. [20] proposed XAI-CNVMarker, an explainable AI framework that identifies 44 CNV biomarkers linked to breast cancer subtypes. Using deep learning for classification, the model achieved 0.712 accuracy with 95% confidence. Gene set analysis identified subtype-specific pathways and druggable genes, validated on the METABRIC dataset, demonstrating its clinical relevance.

Additionally, XAI-MethylMarker, also by [21], uses methylation data to classify breast cancer subtypes and explain key methylation patterns driving decisions. This interpretability is crucial for personalized medicine, as understanding these patterns aids in tailoring patient treatments. Both frameworks highlight the role of explainable AI in providing clinically reliable and interpretable biomarkers for breast cancer diagnosis and treatment.

4.3 Explainable AI in Cancer Drug Recommendation and Diagnosis

The integration of Explainable AI (XAI) in cancer treatment and diagnosis is crucial for enhancing transparency and trust in AI-driven decisions. [22] proposed a modular XAI-based drug recommendation system leveraging cancer omics data to improve precision oncology. Their approach enhances traceability, achieving a 59.24% traceability rate across 70,211 drug recommendations, making AI-driven therapy decisions more interpretable and reliable for clinicians and patients.

Similarly, [23] introduced an XAI model for cancer diagnosis trained on a large dataset of cancer images. The model enhances interpretability while maintaining high predictive accuracy, ensuring transparent and reliable medical decision-making. Both studies highlight the growing role of XAI in bridging the gap between complex AI models and their practical application in cancer research, fostering trust and improving clinical decision support systems.

4.4 Explainable AI in Breast Cancer Detection and Diagnosis

Breast cancer remains one of the most prevalent and lethal diseases worldwide, emphasizing the need for early detection and improved diagnostic accuracy. Recent studies highlight the application of Explainable Artificial Intelligence (XAI) in enhancing breast cancer detection, offering more interpretable and transparent models that clinicians can trust.

[24] proposed a novel hybrid approach, XAI-RACapsNet, to improve breast cancer detection using mammogram images. By incorporating image pre-treatment techniques like noise reduction and histogram equalization, and combining them with an advanced explainable AI-based Region of Interest (ROI) segmentation method, this approach demonstrated superior performance in terms of accuracy and interpretability. The integration of Relevance Aware Capsule Network (RACapsNet) further enhanced the model's explainability, providing clearer insights into feature relevance through heat maps.

[25] introduced a unique combination of the CatBoost classification model with a multi-layer perceptron (MLP) network to address breast cancer diagnosis. This model utilized Shapley additive explanations (SHAP) for interpretability, emphasizing the importance of feature engineering in identifying critical variables. The study's findings reinforced the significance of model transparency in healthcare applications.

[26] presented a framework combining image data and numerical data for breast cancer detection. By employing the U-NET transfer learning model for image-based predictions and integrating a custom CNN with Random Forest (RF) and Support Vector Machine (SVM) ensemble models, they achieved an accuracy rate of 99.99%. This study demonstrated the effective use of XAI for interpreting machine learning model predictions, thereby enhancing the diagnosis of breast cancer.

[27] proposed a framework that combines Random Forest (RF) classifiers with SHAP and LIME to improve both the accuracy and transparency of breast cancer detection. The model achieved a validation accuracy of up to 98% with SHAP and LIME helping to explain key features like menopause and tumor size. This approach demonstrated the critical role of explainable AI in ensuring that AI-based models are trusted and adopted in clinical settings.

[28] utilized machine learning models such as SVM, Decision Tree, KNN, and Logistic Regression, with an emphasis on feature selection using bagging boosting techniques. The incorporation of XAI methods, including SHAP and LIME, helped explain feature importance and individual predictions. Their study, using the BCGENES dataset, reached an accuracy of 88%, demonstrating the potential of XAI in improving feature interpretability in breast cancer diagnostics.

[29] combined deep learning and classical machine learning techniques into a custom lightweight neural network model for breast cancer detection. Their model demonstrated an impressive accuracy of 97.54%, and the integration of XAI ensured that the model was both accurate and interpretable. This approach provided clinicians with a tool that offered transparency, improving their ability to make informed decisions based on AI-driven predictions.

[30] explored AI in breast cancer detection, highlighting the superiority of deep learning over machine learning with a 2% accuracy improvement. Their review focused on various imaging modalities like mammograms and ultrasound, emphasizing the importance of Explainable AI (XAI) techniques to enhance model interpretability for clinicians, though challenges like image quality remain.

[31] proposed the Breast Cancer Causal XAI Diagnostic Model (BCCXDM), combining TabNet with causal graphs and Graph Neural Networks (GNN). Their approach improved model interpretability and produced results that closely aligned with clinical standards.

[32] used XGBoost and SHAP values to identify key biomarkers for breast cancer, focusing on immune gene expression in PBMCs. They identified genes like SVIP and BEND3 as non-invasive diagnostic biomarkers, offering potential for early detection.

[33] utilized Local Interpretable Model-agnostic Explanations (LIME) to enhance transparency in breast cancer prediction using LightGBM, CatBoost, and XGBoost on the SEER dataset. SMOTE and Tomek Links addressed class imbalance, with LightGBM achieving the best performance. LIME explanations highlighted key clinical features like T Stage and cancer grade, ensuring model interpretability and trust in AI-driven diagnostics.

Further, in the context of gene expression, [34] developed GeneXAI, a multi-modal framework that identifies influential genes in different stages of breast cancer. The integration of XAI techniques enables the identification of key genes involved in cancer progression, providing valuable insights into potential therapeutic targets. This model is important because it offers both biological insights and practical guidance for treatment strategies.

Similarly, [35] proposed a hybrid machine learning-XAI algorithm for breast cancer prevention, offering clinicians a decision-support system that helps explain the reasoning behind risk assessments.

4.5 Explainable AI in Cervical Cancer Screening and Diagnosis

Cervical cancer screening and diagnosis face both socio-cultural and technological challenges that impact early detection and prevention efforts. [36] explored the influence of emotional attachment to institutional logics on women's participation in cervical cancer screening in Mozambique. Their study identified two competing logics: *preservation logic*, shaped by cultural norms and emotions like shame and fear, and *prevention logic*, focused on screening adoption. Women with greater autonomy were more likely to detach from traditional norms and accept screening, highlighting the need for culturally adapted interventions to improve screening uptake.

On the technological front, [37] developed a deep-learning-based classifier using liquid cytology images for cervical cancer severity classification. Their 4-phase optimization process resulted in a 97% accuracy for 4-class classification and 100% accuracy for 2-class classification, with execution times under 1 second. The study emphasized the importance of Explainable AI (XAI) to assist pathologists in diagnosis verification, enabling confidence assessment and incremental learning systems.

Together, these studies underscore the dual importance of socio-cultural adaptation and advanced AI-driven diagnostic tools in improving cervical cancer detection and prevention strategies.

4.6 Explainable AI in Metabolomics and Ovarian Cancer Detection

Metabolomics data is inherently complex, requiring advanced computational approaches for meaningful interpretation. [38] proposed a unified pipeline that integrates Automated Machine Learning (AutoML) with Explainable AI (XAI) for optimized metabolomics analysis. Their study tested renal cell carcinoma (RCC) urine metabolomics and ovarian cancer (OC) serum metabolomics, demonstrating superior classification performance of AutoML (Auto-sklearn) over standalone ML models like SVM and k-NN. AutoML achieved AUC scores of 0.97 (RCC) and 0.85 (OC), highlighting its effectiveness. SHAP analysis provided biological insights by identifying key discriminatory metabolites such as dibutylamine for RCC and ganglioside GM(d34:1) for OC.

[39] focused on early detection of ovarian cancer using ML and XAI. Their study applied ensemble learning techniques (Boosting, Bagging, Stacking) to a dataset of 349 patients and found that SVM achieved 85% accuracy, which improved to 89% with ensemble stacking. They leveraged Shapley values to interpret model decisions, offering clinicians a transparent view of key predictors like tumor markers, imaging characteristics, and family history.

Expanding on ovarian cancer detection, [40] introduced DeepResVit, a hybrid deep learning model combining ResNet-152 and Vision Transformer (ViT) for classification. Image preprocessing techniques, including data augmentation and L1 regularization, enhanced model robustness. XAI techniques like Grad-CAM++ improved transparency in decision-making.

DeepResVit achieved 98.65% accuracy, precision, and F1-score, demonstrating its effectiveness in early ovarian cancer detection.

These studies emphasize the critical role of XAI in improving AI model interpretability and clinical applicability in oncology, advancing precision medicine and early cancer detection.

4.7 Explainable AI in Skin Cancer Detection and Risk Prediction

Skin cancer, particularly melanoma, presents diagnostic challenges due to heterogeneous skin images and the need for both accuracy and interpretability in AI-driven diagnostics. [41] introduced SmartSkin-XAI, an AI-based diagnostic system integrating DenseNet121 with Explainable AI (XAI) to improve early detection and patient management. Evaluated on the ISIC and Kaggle datasets, the model achieved 97% and 98% classification accuracy, respectively, outperforming benchmark models like InceptionV3 and ResNet50. By incorporating XAI techniques, SmartSkin-XAI enhances clinical trust and decision-making, making AI diagnostics more transparent and applicable in real-world healthcare settings.

Beyond diagnosis, [42] explored AI's potential in skin cancer risk prediction using 2D facial images and data from the Rotterdam Study. Their XAI-based deep-learning survival analysis demonstrated a higher predictive accuracy (c-index = 0.72) than traditional risk factor-based models (c-index = 0.59), suggesting that facial image analysis could serve as an alternative screening tool. This approach offers an accessible, non-invasive method for identifying high-risk individuals, potentially personalizing early screening strategies.

Both studies highlight XAI's transformative role in skin cancer diagnostics and risk assessment, ensuring interpretability, accuracy, and clinical relevance, thereby bridging the gap between AI-driven insights and medical decision-making.

4.8 Explainable AI in Liver and Pancreatic Cancer Detection

In the field of liver cancer detection, [43] utilized deep learning models integrated with XAI techniques to explain how specific medical imaging features influence detection decisions. Similarly, [44] proposed a ResNet-based liver tumor detection method, incorporating LIME (Local Interpretable Model-agnostic Explanations) to enhance interpretability and trustworthiness. Their approach achieved state-of-the-art performance, reinforcing the clinical applicability of AI in liver disease diagnosis.

Both studies highlight XAI's essential role in medical imaging, ensuring AI-driven models are transparent, interpretable, and clinically reliable for improved diagnostic accuracy and decision-making.

The use of XAI in cancer research enhances model transparency and trustworthiness, which is critical in medical applications where decisions can directly impact patient outcomes. These models are especially valuable in precision oncology, where understanding the contributions of specific features—such as genetic markers or imaging data—is essential for personalized treatment strategies. Future research should focus on integrating multi-modal data sources, refining interpretability techniques, and developing clinically deployable models for improved diagnostic and prognostic outcomes.

Despite extensive research, pancreatic cancer remains a highly fatal disease with minimal improvement in survival rates. While deep learning has advanced predictive analytics in oncology, its increasing complexity has led to "black box" concerns, limiting acceptance in healthcare. This has driven renewed interest in Explainable AI (XAI) to enhance model interpretability. [45] leveraged XAI with CT images and clinical data to improve pancreatic cancer detection and survival analysis. By marking tumor regions and identifying key predictive features, XAI aids clinicians in making informed decisions. Their study focuses on XAI strategies in deep and machine learning rather than solely on prediction and survival methods.

Addressing Bias, Ethics, and Regulatory Compliance

Beyond improving diagnostic accuracy and treatment planning, XAI plays a fundamental role in addressing biases and ethical concerns in AI-driven healthcare. AI models trained on biased datasets can inadvertently perpetuate disparities in medical care, leading to misdiagnoses or suboptimal treatment recommendations for underrepresented populations. For example, diabetes risk prediction models may be less accurate for certain ethnic groups if the training data lacks sufficient diversity. XAI helps detect such biases by exposing how different variables influence AI's decisions, allowing researchers and clinicians to refine models for more equitable healthcare outcomes.

Additionally, regulatory bodies such as the U.S. Food and Drug Administration (FDA) and the European Medicines Agency (EMA) emphasize the need for transparency in AI-driven medical technologies. XAI supports regulatory compliance by generating audit trails for AI-based decisions, ensuring that AI-powered diagnostic tools and treatment recommendations meet safety and efficacy standards. For example, AI-driven insulin delivery systems must comply with FDA's AI/ML Software as a Medical Device (SaMD) guidelines, and XAI provides the necessary explainability to satisfy these regulatory requirements.

5. Bibliometric Analysis

This study follows the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) guidelines, ensuring a structured and transparent approach to literature selection and analysis. The identification phase involved retrieving relevant documents from the Scopus database using the keywords "Explainable AI," "Cancer," and "Diabetes." In the screening phase, duplicate records were removed, and only publications from 2021 to March 2025 in English were considered. The eligibility criteria included selecting journal articles, conference proceedings, and review papers that focused on Explainable AI applications in healthcare, searching for their titles. After applying these filters, 40 documents were finalized for bibliometric analysis. The analysis phase involved a comprehensive examination of trends in Explainable AI for cancer diagnosis, diabetes prediction, and decision support systems. The study covers publications from 2021 to 2025, reflecting recent advancements in the field. The findings provide insights into research trends, keyword co-occurrence, and source contributions, highlighting the growing significance of Explainable AI in medical decision-making and predictive analytics.

The graph shown in Fig.1 presents the number of documents published per year from 2021 to March 2025, based on Scopus data. It highlights significant growth in research publications over the years, particularly from 2022 onward. The number of documents saw a sharp rise, peaking in 2024 with around 25 publications, indicating increasing interest in the topic. The decline in 2025 is due to data being collected only until March, suggesting that the full-year count may be higher. This trend underscores the growing importance of research in this domain, with a notable surge in contributions leading up to 2024.

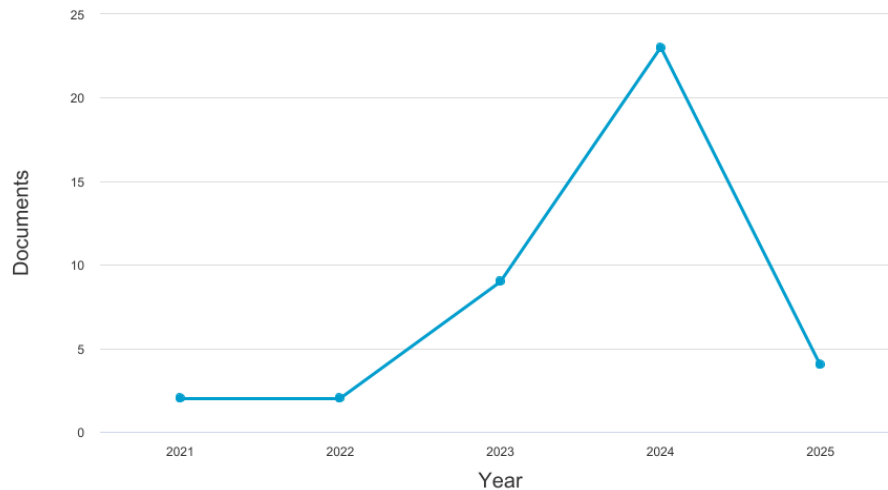


Figure 1: Number of documents published per year from 2021 to March 2025

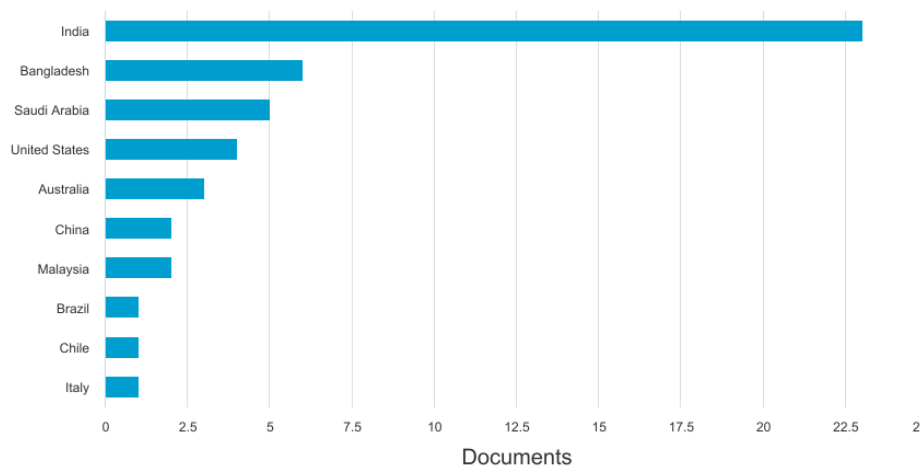


Figure 2: Documents published by country or territory

The graph in Fig. 2 presents a comparative analysis of document counts by country or territory, highlighting research contributions indexed in Scopus. India leads with a significantly higher number of documents, surpassing all other countries by a wide margin. Following India, Bangladesh and Saudi Arabia show moderate research contributions, though their document counts are notably lower. The United States, Australia, and China also contribute to the dataset.

but with relatively fewer publications. Meanwhile, Malaysia, Brazil, Chile, and Italy have the lowest document counts among the listed countries. The data suggests that research output in the analyzed domain is particularly strong in India, with notable contributions from Bangladesh and Saudi Arabia. The presence of Western countries, including the United States and Australia, indicates a global research effort, though with varying levels of contribution.

6. Discussion: Reasons for the Use of XAI Models in Both Cancer and Diabetes Research

The primary reason for integrating XAI into AI models in both cancer and diabetes research is to improve the interpretability and transparency of decision-making processes. In healthcare, particularly in medical diagnostics and treatment planning, trust in AI systems is crucial. Healthcare professionals must be able to understand how AI models arrive at their decisions to ensure that they can rely on these systems to guide clinical decisions.

XAI is also essential in personalized medicine, where individual patient characteristics (such as genetic markers, imaging data, or lifestyle factors) need to be carefully considered in treatment plans. By explaining which features influence a model's predictions, XAI helps clinicians tailor interventions to the specific needs of each patient.

Furthermore, in regulatory contexts, where AI systems are subject to scrutiny and approval, XAI methods provide the necessary transparency to demonstrate that the models operate in a clear, understandable manner. This is particularly important in areas like cancer and diabetes, where treatment decisions can have significant impacts on patient health.

Additionally, XAI helps in handling challenges such as imbalanced datasets, feature selection, and data privacy concerns. In diabetes, for example, XAI enables a deeper understanding of how certain features (like blood sugar levels or lifestyle factors) contribute to diabetes prediction, while in cancer, XAI helps identify critical biomarkers and genetic factors that influence cancer diagnosis and treatment planning.

Key Points:

- XAI plays a crucial role in enhancing the interpretability and transparency of AI models in cancer and diabetes research.
- In cancer research, XAI helps improve diagnostic accuracy, identify biomarkers, and support personalized treatment strategies.
- In diabetes research, XAI models assist in predicting diabetes risk, managing treatment, and personalizing care plans.
- XAI techniques such as SHAP, LIME, and Grad-CAM are widely used to visualize the decision-making process and explain model predictions.
- The integration of XAI improves trust in AI systems, which is critical for clinical adoption and decision-making.
- The use of federated learning and privacy-preserving models ensures that sensitive patient data is secure while collaborating across institutions.

7. Conclusion and Recommendations for Future Research:

Explainable AI is not just a technological advancement—it represents a paradigm shift in healthcare, ensuring that AI-driven decisions are interpretable, trustworthy, and clinically relevant. By making AI models transparent, XAI bridges the gap between complex algorithms and human expertise, empowering healthcare professionals to make informed decisions with confidence. In diabetes management, XAI enhances patient safety by explaining insulin dosage predictions and disease progression risks. In oncology, it provides clarity in cancer diagnostics and treatment recommendations, fostering trust in AI-driven healthcare solutions.

As AI continues to evolve, the demand for explainability will only grow, particularly in high-stakes medical applications. Future research should focus on developing more intuitive and user-friendly XAI methods that seamlessly integrate into clinical workflows. By prioritizing transparency and interpretability, XAI will play a pivotal role in shaping the future of AI in healthcare, ultimately improving patient outcomes and advancing medical innovation.

1. **Integration of Multi-modal Data:** Future studies should focus on combining various data sources (such as genomic, imaging, and clinical data) to improve the robustness and generalizability of AI models, particularly in personalized cancer and diabetes treatment.
2. **Improving XAI Techniques:** There is a need to refine existing XAI methods, making them more interpretable and clinically actionable. These techniques should be tailored for specific clinical applications to ensure their practical utility.
3. **Real-world Validation:** Models and frameworks incorporating XAI should be validated in real-world clinical settings to ensure that they provide reliable insights for healthcare professionals and improve patient outcomes.
4. **Collaboration and Data Sharing:** Researchers should explore collaborative, privacy-preserving AI frameworks like federated learning to facilitate data sharing across institutions without compromising patient confidentiality.
5. **Patient-Centric Models:** Future AI models should prioritize the inclusion of patient feedback and real-world health data to create more personalized and adaptive treatment strategies. This will foster greater engagement and trust in AI-driven healthcare tools.

In conclusion, the integration of XAI into AI models has the potential to revolutionize cancer and diabetes research by enhancing transparency, improving model trustworthiness, and supporting personalized treatment approaches. Continued advancements in this area will be crucial for the widespread adoption of AI in clinical decision-making.

References:

- [1] “Diabetes.” <https://www.who.int/news-room/fact-sheets/detail/diabetes> (accessed Mar. 28, 2025).
- [2] “Cancer Data and Statistics | Cancer | CDC.”

- <https://www.cdc.gov/cancer/data/index.html> (accessed Mar. 28, 2025).
- [3] “Diabetes in America: Prevalence, Statistics, and Economic Impact.” <https://diabetes.org/about-diabetes/statistics/about-diabetes> (accessed Mar. 28, 2025).
- [4] “Cancer.” <https://www.who.int/news-room/fact-sheets/detail/cancer> (accessed Mar. 28, 2025).
- [5] “Cancer Statistics Center - American Cancer Society.” <https://cancerstatisticscenter.cancer.org/> (accessed Mar. 28, 2025).
- [6] Y. Zhao, J. K. Chaw, M. C. Ang, M. M. Daud, and L. Liu, “A Diabetes Prediction Model with Visualized Explainable Artificial Intelligence (XAI) Technology,” in *Advances in Visual Informatics*, 2024, pp. 648–661.
- [7] A. Panda and A. Behera, “Interpretable Machine Learning Model for Diabetes Disease Detection using Explainable Artificial Intelligence (xAI),” *Prospect. Sci. Technol. Appl.*, pp. 81–88, Jan. 2024, doi: 10.1201/9781003489443-10/INTERPRETABLE-MACHINE-LEARNING-MODEL-DIABETES-DISEASE-DETECTION-USING-EXPLAINABLE-ARTIFICIAL-INTELLIGENCE-XAI-ABHILASHA-PANDA-ANUKAMPA-BEHERA.
- [8] R. Ganguly and D. Singh, “Explainable Artificial Intelligence (XAI) for the Prediction of Diabetes Management: An Ensemble Approach,” *Int. J. Adv. Comput. Sci. Appl.*, vol. 14, no. 7, pp. 158–163, 2023, doi: 10.14569/IJACSA.2023.0140717.
- [9] N. M. Nayan, A. Islam, M. U. Islam, E. Ahmed, M. M. Hossain, and M. Z. Alam, “SMOTE Oversampling and Near Miss Undersampling Based Diabetes Diagnosis from Imbalanced Dataset with XAI Visualization,” *Proc. - IEEE Symp. Comput. Commun.*, vol. 2023-July, 2023, doi: 10.1109/ISCC58397.2023.10218281.
- [10] R. Ahuja and G. Indra, “Leveraging XAI for Discovering Crucial Demographic, Clinical and Pathological Diabetes Mellitus Biomarkers,” *Proc. 2024 3rd Ed. IEEE Delhi Sect. Flagsh. Conf. DELCON 2024*, 2024, doi: 10.1109/DELCON64804.2024.10866967.
- [11] A. Maaloul, M. Jemel, and N. Ben Azzouna, “XAI based feature selection for gestational diabetes Mellitus prediction,” *10th 2024 Int. Conf. Control. Decis. Inf. Technol. CoDIT 2024*, pp. 1939–1944, 2024, doi: 10.1109/CODIT62066.2024.10708408.
- [12] V. Aelgani, S. K. Gupta, and V. A. Narayana, “Local Agnostic Interpretable Model for Diabetes Prediction with Explanations Using XAI,” *Lect. Notes Networks Syst.*, vol. 606, pp. 417–425, 2023, doi: 10.1007/978-981-19-8563-8_40.
- [13] P. Nagaraj, V. Muneeswaran, A. Dharanidharan, K. Balanathanan, M. Arunkumar, and C. Rajkumar, “A Prediction and Recommendation System for Diabetes Mellitus using XAI-based Lime Explainer,” *Int. Conf. Sustain. Comput. Data Commun. Syst. ICSCDS 2022 - Proc.*, pp. 1472–1478, 2022, doi: 10.1109/ICSCDS53736.2022.9760847.
- [14] Y. Abas Mohamed, B. Ee Khoo, M. Shahrimie Mohd Asaari, M. Ezane Aziz, and F. Rahiman Ghazali, “Decoding the black box: Explainable AI (XAI) for cancer

- diagnosis, prognosis, and treatment planning-A state-of-the art systematic review,” *Int. J. Med. Inform.*, vol. 193, p. 105689, Jan. 2025, doi: 10.1016/J.IJMEDINF.2024.105689.
- [15] A. Singhal, K. K. Agrawal, A. Quezada, A. R. Aguiñaga, S. Jiménez, and S. P. Yadav, “Explainable Artificial Intelligence (XAI) Model for Cancer Image Classification,” *Comput. Model. Eng. Sci.*, vol. 141, no. 1, pp. 401–441, Aug. 2024, doi: 10.32604/CMES.2024.051363.
- [16] N. T. Arthi *et al.*, “Decentralized Federated Learning and Deep Learning Leveraging XAI-Based Approach to Classify Colorectal Cancer,” *Proc. IEEE Asia-Pacific Conf. Comput. Sci. Data Eng. CSDE 2022*, 2022, doi: 10.1109/CSDE56538.2022.10089344.
- [17] Z. U. Ahmed, K. Sun, M. Shelly, and L. Mu, “Explainable artificial intelligence (XAI) for exploring spatial variability of lung and bronchus cancer (LBC) mortality rates in the contiguous USA,” *Sci. Reports 2021 111*, vol. 11, no. 1, pp. 1–15, Dec. 2021, doi: 10.1038/s41598-021-03198-8.
- [18] N. Darshan, S. R. Darshan, A. Kodipalli, and T. Rao, “Comparison of Various Computational Models for the Accurate Detection of Lung Cancer and Interpreting Using XAI,” *2024 Int. Conf. Knowl. Eng. Commun. Syst. ICKECS 2024*, 2024, doi: 10.1109/ICKECS61492.2024.10616715.
- [19] J. Gabriel Avina-Cervantes *et al.*, “An End-to-End Lightweight Multi-Scale CNN for the Classification of Lung and Colon Cancer with XAI Integration,” *Technol. 2024, Vol. 12, Page 56*, vol. 12, no. 4, p. 56, Apr. 2024, doi: 10.3390/TECHNOLOGIES12040056.
- [20] S. Rajpal *et al.*, “XAI-CNVMarker: Explainable AI-based copy number variant biomarker discovery for breast cancer subtypes,” *Biomed. Signal Process. Control*, vol. 84, p. 104979, Jul. 2023, doi: 10.1016/J.BSPC.2023.104979.
- [21] S. Rajpal *et al.*, “XAI-MethylMarker: Explainable AI approach for biomarker discovery for breast cancer subtype classification using methylation data,” *Expert Syst. Appl.*, vol. 225, p. 120130, Sep. 2023, doi: 10.1016/J.ESWA.2023.120130.
- [22] P. Sahoo, D. P. Naidu, M. V. S. Samartha, S. Palei, B. Jena, and S. Saxena, “Drug Recommendation System for Cancer Patients Using XAI: A Traceability Perspective,” *Commun. Comput. Inf. Sci.*, vol. 2010 CCIS, pp. 278–287, 2024, doi: 10.1007/978-3-031-58174-8_24/TABLES/2.
- [23] S. Kothari *et al.*, “Cancer XAI: A Responsible Model for Explaining Cancer Drug Prediction Models,” *Int. J. Intell. Syst. Appl. Eng.*, vol. 11, no. 4, pp. 472 – 484, 2023, [Online]. Available: <https://www.scopus.com/inward/record.uri?eid=2-s2.0-85174798620&partnerID=40&md5=03cf384a1ffd060830ee17f4b02333fc>
- [24] A. Alhussen, M. Anul Haq, A. Ahmad Khan, R. K. Mahendran, and S. Kadry, “XAI-RACapsNet: Relevance aware capsule network-based breast cancer detection using mammography images via explainability O-net ROI segmentation,” *Expert Syst. Appl.*, vol. 261, p. 125461, Feb. 2025, doi: 10.1016/J.ESWA.2024.125461.
- [25] P. N. Srinivasu *et al.*, “XAI-driven CatBoost multi-layer perceptron neural network for

- analyzing breast cancer,” *Sci. Reports* 2024 141, vol. 14, no. 1, pp. 1–19, Nov. 2024, doi: 10.1038/s41598-024-79620-8.
- [26] R. M. Munshi, L. Cascone, N. Alturki, O. Saidani, A. Alshardan, and M. Umer, “A novel approach for breast cancer detection using optimized ensemble learning framework and XAI,” *Image Vis. Comput.*, vol. 142, p. 104910, Feb. 2024, doi: 10.1016/J.IMAVIS.2024.104910.
- [27] N. Hariprasad, M. Annal Priyanga, S. Duraimurugan, M. Sushmitha, and S. Vedha Shri, “A Comparative Study of Machine Learning Classifiers and Ensemble Method for Breast Cancer Detection Using XAI Technique,” *Proc. 2024 3rd Ed. IEEE Delhi Sect. Flagsh. Conf. DELCON 2024*, 2024, doi: 10.1109/DELCON64804.2024.10866753.
- [28] D. T. Avlani, M. B. Abhijna, G. S. Sai Disha, A. Kodipalli, and T. Rao, “Comprehensive Methodologies for Breast Cancer Classification: Leveraging XAI LIME, SHAP Bagging Boosting, and Diverse Single Classifiers,” *2024 4th Asian Conf. Innov. Technol. ASIANCON 2024*, 2024, doi: 10.1109/ASIANCON62057.2024.10837893.
- [29] S. A. Tanim, G. M. I. Alam, T. E. Shrestha, M. Islam, F. Jahan, and K. Nur, “Breast Cancer Diagnosis with XAI-Integrated Deep Learning Approach,” *2024 Int. Conf. Innov. Intell. Informatics, Comput. Technol. 3ICT 2024*, pp. 659–665, 2024, doi: 10.1109/3ICT64318.2024.10824574.
- [30] R. Karthiga, K. Narasimhan, T. V. H. M, and R. Amirtharajan, “Review of AI & XAI-based breast cancer diagnosis methods using various imaging modalities,” *Multimed. Tools Appl.*, vol. 84, no. 5, pp. 2209–2260, Oct. 2024, doi: 10.1007/S11042-024-20271-2/TABLES/8.
- [31] D. Chen, H. Zhao, J. He, Q. Pan, and W. Zhao, “An Causal XAI Diagnostic Model for Breast Cancer Based on Mammography Reports,” *Proc. - 2021 IEEE Int. Conf. Bioinforma. Biomed. BIBM 2021*, pp. 3341–3349, 2021, doi: 10.1109/BIBM52615.2021.9669648.
- [32] S. Kumar and A. Das, “Peripheral blood mononuclear cell derived biomarker detection using eXplainable Artificial Intelligence (XAI) provides better diagnosis of breast cancer,” *Comput. Biol. Chem.*, vol. 104, p. 107867, Jun. 2023, doi: 10.1016/J.COMPBIOLCHEM.2023.107867.
- [33] V. Gupta and R. Sharma, “Enhancing Breast Cancer Prediction with XAI-Enabled Boosting Algorithms,” *2024 15th Int. Conf. Comput. Commun. Netw. Technol. ICCCNT 2024*, 2024, doi: 10.1109/ICCCNT61001.2024.10723843.
- [34] S. Manna, S. Mistry, and D. De, “GeneXAI: Influential gene identification for breast cancer stages using XAI-based multi-modal framework,” *Med. Nov. Technol. Devices*, vol. 25, p. 100349, Mar. 2025, doi: 10.1016/J.MEDNTD.2024.100349.
- [35] F. Silva-Aravena, H. Núñez Delafuente, J. H. Gutiérrez-Bahamondes, and J. Morales, “A Hybrid Algorithm of ML and XAI to Prevent Breast Cancer: A Strategy to Support Decision Making,” *Cancers* 2023, Vol. 15, Page 2443, vol. 15, no. 9, p. 2443, Apr. 2023, doi: 10.3390/CANCERS15092443.

- [36] G. Fulane, M. Major, C. Lorenzoni, and K. Munguambe, “The Influence of Institutional Logics and Emotions on the Uptake of Cervical Cancer Screening: A Case Study From Xai-Xai, Mozambique,” *Heal. Serv. Insights*, vol. 17, Jan. 2024, doi: 10.1177/11786329231224619/SUPPL_FILE/SJ-DOCX-1-HIS-10.1177_11786329231224619.DOCX.
- [37] J. Civit-Masot, F. Luna-Perejon, L. Muñoz-Saavedra, M. Domínguez-Morales, and A. Civit, “A lightweight xAI approach to cervical cancer classification,” *Med. Biol. Eng. Comput.*, vol. 62, no. 8, pp. 2281–2304, Aug. 2024, doi: 10.1007/S11517-024-03063-6/FIGURES/11.
- [38] O. O. Bifarin and F. M. Fernández, “Automated Machine Learning and Explainable AI (AutoML-XAI) for Metabolomics: Improving Cancer Diagnostics,” *J. Am. Soc. Mass Spectrom.*, vol. 35, no. 6, pp. 1089–1100, Jun. 2024, doi: 10.1021/JASMS.3C00403/ASSET/IMAGES/LARGE/JS3C00403_0006.JPEG.
- [39] S. Lavanya J M and S. P, “Innovative approach towards early prediction of ovarian cancer: Machine learning- enabled XAI techniques,” *Heliyon*, vol. 10, no. 9, p. e29197, May 2024, doi: 10.1016/J.HELIYON.2024.E29197.
- [40] A. R. Aurnob, T. E. Shrestha, M. S. Al Huda, S. A. Tamim, R. A. Arman, and M. A. Ali, “DeepResVit: A Hybrid Deep Learning Approach for Ovarian Cancer Classification with XAI,” *2nd Int. Conf. Inf. Commun. Technol. ICICT 2024*, pp. 229–233, 2024, doi: 10.1109/ICICT64387.2024.10839719.
- [41] S. A. Hamim, M. U. I. Tamim, M. F. Mridha, M. Safran, and D. Che, “SmartSkin-XAI: An Interpretable Deep Learning Approach for Enhanced Skin Cancer Diagnosis in Smart Healthcare,” *Diagnostics 2025, Vol. 15, Page 64*, vol. 15, no. 1, p. 64, Dec. 2024, doi: 10.3390/DIAGNOSTICS15010064.
- [42] X. Liu *et al.*, “Predicting skin cancer risk from facial images with an explainable artificial intelligence (XAI) based approach: a proof-of-concept study,” *eClinicalMedicine*, vol. 71, p. 102550, May 2024, doi: 10.1016/J.ECLINM.2024.102550.
- [43] H. Arjun Kumar, R. Aswin, and U. Rahul Varma, “Deep Learning Approach in Detecting Liver Cancer from Medical Images with Explainable AI (XAI) Technique,” *15th Int. Conf. Adv. Comput. Control. Telecommun. Technol. ACT 2024*, vol. 2, pp. 1934–1941, Jan. 2024.
- [44] R. Aswin, H. Arjun Kumar, and U. Rahul Varma, “ResNet-Based Deep Learning Framework for Liver Cancer Detection with Explainable AI (XAI) Technique,” *2024 15th Int. Conf. Comput. Commun. Netw. Technol. ICCCNT 2024*, 2024, doi: 10.1109/ICCCNT61001.2024.10724461.
- [45] B. Srinidhi and M. S. Bhargavi, “An XAI Approach to Predictive Analytics of Pancreatic Cancer,” *2023 Int. Conf. Inf. Technol. Cybersecurity Challenges Sustain. Cities, ICIT 2023 - Proceeding*, pp. 343–348, 2023, doi: 10.1109/ICIT58056.2023.10225991.

