RESEARCHERID
THOMSON REUTERS
[Pooja et al., 11(4), Apr 2026]

ISSN : 2455-9679
SJIF Impact Factor : 6.190

IJRTSM

INTERNATIONAL JOURNAL OF RECENT TECHNOLOGY SCIENCE & MANAGEMENT

"DEVELOPMENT OF HEALTHCARE DECISION SUPPORT SYSTEMS FOR DISEASE DIAGNOSIS AND PREDICTION USING EXPLAINABLE AI: A LITERATURE REVIEW"

Pooja Raikwar [1], Amlesh Singh [2]

[1] [1]Research Scholar, Department of Computer Science and Application, Rabindranath Tagore University, Raisen, Madhya Pradesh, India

[2]Assistant Professor, Department of Computer Science and Application, Rabindranath Tagore University, Raisen, Madhya Pradesh, India

pooja.raikwar1234@gmail.com

amlesh.singh@rntu.ac.in

Corresponding Author: amlesh.singh@rntu.ac.in

ABSTRACT

The growing convolution of healthcare data and the critical requirement for accurate disease diagnosis and prediction have catalysed the acceptance of Artificial Intelligence (AI) in clinical supervision. However, traditional AI and machine learning models habitually function as "black boxes," limiting interpretability and clinician trust. This literature review explores the development of Healthcare Decision Support Systems (HDSS) leveraging Explainable AI (XAI) approaches to provide accurate, transparent, and trustworthy disease diagnosis and prediction. It systematically examines recent research on explainable systems like Generalized Additive Models (GAMs), Explainable Boosting Machines (EBMs), and post-hoc explanation techniques including SHAP, LIME, and K-LIME across diverse healthcare datasets. The review highlights how XAI contributes to enhanced model transparency, clinician trust, and actionable insights, while addressing challenges related to scalability, integration into clinical workflows, and evaluation metrics. By synthesizing current findings, this work provides a comparative overview of the state-of-the-art in XAI-enabled HDSS and identifies research gaps for future development.

Keywords: *Artificial Intelligence, Decision Support System, GAM, XAI, Healthcare System.*

I. INTRODUCTION

Healthcare systems produce huge quantity of data from patient details, laboratory outcomes, disease images, and wearable devices. Leveraging this data for accurate disease diagnosis and prediction drastically enhance patient results [3]. AI-based HDSS have demonstrated strong predictive capabilities in detecting conditions ranging from diabetes to mental health disorders. However, conventional machine learning models often run as black boxes, offering slight vision into the reasoning behind extrapolations. This absence of explainability limits clinical assumption, as healthcare professionals require transparent and trustworthy explanations to make informed decisions [4, 7].

The XAI addresses these issues by offering interpretable and actionable insights while maintaining high predictive performance. Techniques like GAMs and EBMs allow modelling of feature contributions in a human-understandable way [13]. Post-hoc methods, including SHAP, LIME, and K-LIME, further enhance transparency by providing local and global explanations of model outputs. Recent studies have applied XAI to diverse healthcare domains, including

ISSN : 2455-9679
SJIF Impact Factor : 6.190

structured datasets like PIMA Diabetes, unstructured text datasets for mental health detection, and multimodal clinical data [11, 15].

Despite these advances, several challenges remain, including integrating XAI models seamlessly into existing HDSS, ensuring scalability across heterogeneous datasets, and developing robust evaluation metrics for explanation quality and clinical usefulness. This literature review critically examines current research on XAI in HDSS, highlighting methodologies, performance, and interpretability, while identifying gaps and future research directions for developing reliable, transparent, and clinically applicable decision support systems.

1.1. The Role of AI in Healthcare Decision Support Systems

1.1.1. Current State of Healthcare AI

The healthcare industry has witnessed a flow in the adoption of AI-driven technologies in recent years. AI utilizations in healthcare incorporate a broad range of areas, containing medical imaging, disease diagnosis, drug identification, patient monitoring, and treatment recommendation. These applications leverage ML algorithms to analyse huge datasets, extract patterns, and create predictions that can help healthcare specialist in decision-making procedure. One of the primary drivers of AI's adoption in healthcare is its potential to enhance the precision and effectiveness of medical tasks. For example, AI strategies have illustrated the strength to identify anomalies in medical figures, like X-rays and MRIs, with a level of precision comparable to or uniform superior that of human experts. In addition, AI-powered clinical decision support systems can offer real-time suggestions according to patient data, enabling more personalized and evidence-based care [4].

1.1.2. Challenges in Healthcare Decision Support

Despite the promising advancements in healthcare AI, several challenges persist:

1. Data Quality and Availability**:** AI models in healthcare heavily depend on high-quality, diverse, and well-labelled datasets. However, data quality and availability can be inconsistent, leading to issues such as bias and data insufficiency [5].

2. Interpretability and Transparency: Many AI models used in healthcare are complex and difficult to interpret. The opacity of these models hinders understanding, making it ambitious for clinicians to trust and act upon their recommendations.

3. Regulatory and Ethical Compliance: Healthcare AI systems must observe to rigorous regulatory standards, including data privacy procedures like the Health Insurance Portability and Accountability Act (HIPAA) in the United States. Ensuring ethical compliance and patient consent is also a complex issue [6].

4. Accountability and Liability: Establishing accountability and liability when AI systems are involved in medical decisions is a legal and ethical issue. Who is responsible when an AI-driven recommendation leads to an adverse outcome?

1.1.3. Requirement for Explainable AI in Healthcare

To address the challenges mentioned above and maximize the potential benefits of AI in healthcare, there is a enhancing requirement for XAI. XAI emphases on progressing AI models and systems that not only make accurate predictions but also provide clear explanations for their decisions. In healthcare, the ability to explain why an AI model recommends a precise treatment or diagnosis is crucial for several reasons:

1. Building Trust: Healthcare specialists and patients are more prospective to trust AI systems if they recognize the reasoning behind the recommendations.

2. Clinical Validation: Explanations can assist clinicians in confirming AI-driven visions and making informed decisions.

3. Regulatory Compliance: Transparent AI models are more likely to meet regulatory and ethical standards for healthcare applications [7].

4. Error Detection: Explanations can help identify errors or biases in AI models, improving patient safety.

ISSN : 2455-9679
SJIF Impact Factor : 6.190

1.2. Foundations of Explainable AI

1.2.1. Definition and Terminology

Explainable AI (XAI) is a branch of artificial intelligence that focuses on making machine learning and deep learning models more interpretable and transparent. It aims to enable humans, particularly domain experts, to recognize and trust AI-driven determinations. In the context of healthcare, XAI ensures that AI models not only offer predictions or recommendations but also offer explanations that clinicians and patients can comprehend [8].

To understand XAI better, it's essential to clarify some key terminology:

1. Interpretability: Interpretability refers to the level to which a human can recognize the internal schemes and decision-making procedures of an AI model. It measures how well a model's outputs can be explained.
2. Explain ability: Explain ability goes a step further by explicitly offering explanations for AI model decisions. It involves conveying the validation behind a model's predictions or recommendations in a comprehensible manner [9].

1.2.2. Importance of Interpretability

Interpretability is fundamental in healthcare because it permits healthcare specialists to:

Validate AI-generated insights: Clinicians can assess whether an AI model's recommendations align with their medical knowledge and clinical experience. Interpretability helps them identify potential errors or inconsistencies.

Make informed decisions: When healthcare providers understand why an AI system makes a particular recommendation, they can create more updated decisions about patient care, treatment plans, and interventions.

Establish trust: Transparent AI models can build trust among healthcare specialists, patients, and regulatory bodies. Trust is crucial for widespread recognition and assumption of AI in healthcare.

1.2.3. Types of Explainability Methods

1. Explainable AI employs various methods and techniques to enhance model interpretability and explainability. These techniques can be widely categorized into the following types:
2. Model-specific methods: Some AI models are essentially interpretable, like linear regression or decision trees. These models offer built-in explanations because their decision rules are easy to follow. In healthcare, interpretable models like logistic regression are commonly used for risk prediction and diagnosis.
3. Post-hoc explainability techniques: Post-hoc methods are processed after an AI model has made predictions. They aim to describe the model's decisions without modifying its architecture. Common post-hoc techniques include feature importance analysis, SHAP (SHapley Additive exPlanations) values, and LIME (Local Interpretable Model-agnostic Explanations).
4. Rule-based approaches: Rule-based systems generate descriptions in the form of if-then rules. These rules provide clear guidelines for decision-making. Rule-based XAI methods are particularly useful in medical expert systems, where medical knowledge is encoded into rule sets [10].
5. Visual explanations and heatmaps: Visualizations can make complex AI outputs more understandable. Heatmaps, for instance, offer the regions of an image or data that contributed most to a model's decision. In medical imaging, heatmaps can help radiologists pinpoint areas of concern.
6. Case-Based Reasoning (CBR): CBR systems provide explanations by finding and presenting similar cases from historical data that led to similar outcomes. This approach leverages past patient cases to justify current recommendations.
7. Natural Language Explanations: Translating AI outputs into natural language explanations allows for easy comprehension. AI systems can generate textual or spoken explanations that describe why a particular decision or recommendation was made.

In the successive sections of this paper, we will explore deeper into these explainability methods and explore their

 ISSN : 2455-9679
SJIF Impact Factor : 6.190

applications in healthcare decision support systems.

1.3. Benefits and Advantages of Explainable AI in Healthcare

XAI offers numerous benefits and advantages in healthcare decision support systems, making it an important tool for both healthcare specialists and patients. In this section, we explore the advantages of XAI and how it positively impacts the healthcare ecosystem.

1.3.1. Improved Clinical Decision-Making

One of the main advantages of XAI in healthcare is its strength to enhance clinical decision-making. By offering transparent explanations for AI-driven endorsements, XAI permits healthcare specialists to:

1. Understand Rationale: Clinicians can comprehend why a specific diagnosis, treatment, or prediction was made by an AI system. This understanding allows them to evaluate the recommendation in the context of their clinical expertise.
2. Validate AI Insights: XAI empowers clinicians to endorse AI-generated insights. They can evaluate whether the AI model's rationale aligns with the patient's medical history, symptoms, and available data, reducing the risk of erroneous decisions.
3. Consider Patient-Specific Factors: XAI systems can highlight which patient-specific factors influenced a recommendation. This information enables personalized and patient-cantered care, accounting for individual variations in treatment response.

1.3.2. Enhanced Trust and Acceptance

Trust is a important factor in the adoption of AI in healthcare. XAI helps build and maintain trust among healthcare specialists, patients, and regulatory bodies. Key factors contributing to enhanced trust include:

1. Transparency: XAI models offer clear and interpretable explanations for their decisions. This transparency demystifies AI and fosters trust by removing the "black-box" perception.
2. Accountability: When AI systems provide explanations, it becomes easier to attribute responsibility for recommendations. This accountability encourages responsible use of AI in healthcare.
3. Regulatory Compliance: Transparent AI systems are more likely to comply with healthcare regulations and data secrecy laws, like the HIPAA in the United States and the General Data Protection Regulation (GDPR) in Europe.
4. Ethical Considerations: Trust is also tied to ethical considerations. XAI enables clinicians and patients to assess the ethical implications of AI-driven decisions, ensuring that AI aligns with ethical standards and patient values.

1.3.3. Regulatory Compliance

Healthcare is a widely regulated industry, and AI systems used in healthcare must observe to stringent regulatory standards. XAI facilitates regulatory compliance in several ways:

1. Data Privacy: XAI models can provide explanations while preserving patient privacy. They do not disclose sensitive patient information but rather explain the decision-making process based on the data.
2. Auditability: Transparent AI systems are more amenable to audits and evaluations by regulatory bodies. Explanations make it easier to demonstrate compliance with regulatory requirements.
3. Informed Consent: Transparent AI systems enable notified consent processes. Patients have the right to understand how AI-driven endorsements affect their care and can make informed decisions about treatment options.
4. Accountability: Regulatory bodies often require clear accountability in healthcare decision support systems. XAI's transparency aids in identifying the responsible parties in case of issues or adverse outcomes.
5. By facilitating compliance with healthcare regulations, XAI paves the way for the responsible and ethical use of AI in healthcare [15].

ISSN : 2455-9679
SJIF Impact Factor : 6.190

II. Literature Review

Knapič et al. (2021) [1] explores XAI in human decision support systems for the medical domain. It emphasizes how interpretable machine learning models, including rule-based systems and visualization techniques, can enhance clinician understanding, trust, and usability. The paper demonstrates that combining predictive performance with explainability improves human-AI collaboration in medical diagnosis. Tarnowska et al. (2021) [2] propose an XAI-based clinical decision support system for diagnosing hearing disorders. The system uses interpretable models to highlight feature contributions, such as audiological measures, improving clinicians' trust and diagnostic efficiency. Iqbal et al. (2021) [3] focused on crime prediction and gun violence detection, this paper illustrates the potential of AI decision support systems in safety-critical applications. It underscores predictive modelling, while raising concerns about ethical use and interpretability, which are also relevant for healthcare applications.

Rehman et al. (2023) [4] presents methods for integrating XAI into clinical decision support systems to enhance disease prediction. It shows that interpretable models can maintain predictive accuracy while increasing transparency, facilitating adoption in real-world clinical environments. Bayer et al. (2022) [5] investigate the influence of domain expertise on trusting XAI-enabled decision support. Results show that expert clinicians critically assess AI explanations, highlighting the need for context-sensitive, user-centric explanation design. Antoniadi et al. (2021) [6] identifies challenges and opportunities in XAI for healthcare CDSS. It categorizes XAI methods, highlights lack of standard evaluation metrics, and emphasizes the need for integration into clinical workflows.

Panigutti et al. (2023) [7] proposes a co-design methodology for human-centred XAI in clinical decision support. Involving clinicians in the design loop ensures usability, trust, and practical applicability of AI systems. Pierce et al. (2022) [8] focuses on ethical and technical aspects of explainability in AI-based medicine. It stresses that transparency is crucial for trust, patient safety, and compliance, particularly in genomic and complex clinical data analysis. Wang et al. (2019) [9] presents a theory-driven framework for designing user-centric XAI. Principles such as cognitive load reduction, adaptive explanations, and human-comprehensible formats improve trust and decision-making in AI-assisted systems.

Polat Erdeniz et al. (2022) [10] compares explanation methods like SHAP and LIME for healthcare decision support systems. It demonstrates trade-offs between fidelity, usability, and transparency, emphasizing the need for interpretable communication of model predictions. Srinivasu et al. (2022) [11] reviews tools and case studies transitioning from black-box AI to XAI in healthcare. It highlights methods including rule extraction, attention mechanisms, and gradient-based explanations, while noting challenges in scalability and workflow integration. Holzinger et al. (2017) [12] discussing requirements for XAI in medicine. It argues that interpretability, human-AI collaboration, and rigorous evaluation are essential for adoption of AI in clinical practice.

Saraswat et al. (2022) [13] explore XAI's role in Healthcare 5.0, emphasizing patient-centred, intelligent care. They discuss opportunities in telemedicine, personalized treatment, and IoT-based monitoring, while highlighting privacy and real-time interpretability challenges. Niranjan et al. (2023) [14] presents an XAI-driven COVID-19 diagnosis system combining image segmentation and interpretable classification. The approach improves diagnostic accuracy while providing visual explanations for clinical decision-making.

Gerlings et al. (2022) [15] exploring "explainable to whom?" The paper demonstrates that different stakeholders (clinicians, patients, administrators) require tailored explanations, stressing context-dependent design. Yang et al. (2023) [16] describes A comprehensive survey of XAI approaches, limitations, and applications. It categorizes methods into intrinsic and post-hoc, highlighting challenges in computational cost, scalability, and evaluation metrics, with specific applications in healthcare. Velez & Kim (2017) [17] advocates a rigorous scientific framework for interpretable machine learning, introducing metrics such as fidelity, comprehensibility, and user trust, laying the foundation for structured XAI evaluation.

Pendyala & Kim (2023) [18] evaluates reliability of machine learning models in mental health using XAI. It shows that interpretability improves trust, uncovers biases, and enhances model robustness in sensitive domains. Shrikumar et al. (2017) [19] introduces DeepLIFT for propagating activation differences in deep learning models to identify important features. Applied to genomics, it provides fast, consistent, and interpretable feature attributions. Kerz et al. (2023) [20] applies XAI to mental health detection via language behaviour. Using NLP models and interpretability techniques, it

ISSN : 2455-9679
SJIF Impact Factor : 6.190

identifies linguistic markers associated with mental health issues, improving transparency and clinician trust.

Table 1 describes the comparative analysis of several prior research works on explainable AI on healthcare systems.

Table 1. Comparative Analysis

Paper	Proposed Methodology	Performance Factors	Advantages	Disadvantages
Pendyala, V., et al. (2023) [18]	Decision Trees, Random Forests, SVM, Neural Networks for mental health assessment. Implemented SHAP and LIME for interpretability of model predictions.	Accuracy, Precision, Recall, F1-Score, AUC-ROC	Comprehensive Performance Evaluation, Model Variety, Enhanced Trust, Feature Insights, Resilience to Noise, Generalizability	Increased computational complexity and difficulty in model selection, Model performance heavily depends on the quality and quantity of the mental health data used, require domain expertise.
Kerz, E. et al. (2023) [20]	Applied machine learning models to analyse language behaviour for mental health detection. Integrated XAI techniques such as SHAP and LIME to explain model predictions.	Accuracy, Precision, Recall, F1-Score: AUC-ROC1. Model Robustness, Scalability, Transparency, Feature Importance,	Rich Data Source, Non-Intrusive, Innovative Use of NLP, Scalable Solutions, Enhanced Trust, Insightful Explanations,	Handling and analysing language data raises privacy concerns, Computational Overhead, time-consuming and costly.
Rehman, A. et al. (2023). [4]	Deployment of XAI models to improve disease prediction accuracy and interpretability in clinical settings	Prediction accuracy, interpretability, user satisfaction	Better disease prediction, increased trust in AI recommendations, improved clinical outcomes	Integration challenges with existing systems, high computational requirements
Bayer, S. et al. (2022). [5]	Analysis of the impact of domain expertise on the trust and utilization of XAI-based decision support systems	Trust levels, decision adherence, system usability	Insights into the importance of domain expertise, improved trust and utilization of XAI systems	Potential bias if domain expertise is lacking, difficulty in standardizing expertise levels
Panigutti, C., et al. (2023). [7]	Collaborative design approach involving stakeholders to develop human-centered, explainable AI for clinical decision support	Stakeholder engagement, system usability, interpretability	Ensures system meets user needs, improves user satisfaction and trust, better clinical outcomes	Time-consuming and resource-intensive process, potential for conflicting stakeholder requirements
Pierce, R. L., et al. (2022). [8]	Examination of explainability requirements in modern AI-based clinical decision	Explainability, user trust, decision accuracy	Clarifies explainability requirements, enhances trust and acceptance of AI systems, supports regulatory compliance	Ambiguity in explainability standards, challenges in meeting diverse user needs

ISSN : 2455-9679
SJIF Impact Factor : 6.190

Paper	Proposed Methodology	Performance Factors	Advantages	Disadvantages
	support systems			
Polat Erdeniz, S., et al. (2022). [10]	Application of various XAI techniques to elucidate ML predictions in healthcare decision support systems	Prediction clarity, system usability, diagnostic accuracy	Improved clarity of predictions, enhanced trust in AI decisions, better clinical decision-making	High computational cost, potential for over-simplification of complex models
Srinivasu, P. N., et al. (2022). [11]	Review of existing XAI tools and case studies in healthcare, highlighting their application and effectiveness	Tool effectiveness, case study outcomes, interpretability	Provides practical insights into XAI tool usage, highlights successful case studies, aids in tool selection	Rapidly changing tool landscape, case studies might not be universally applicable
Saraswat, D., et al. (2022). [13]	Exploration of the opportunities and challenges presented by XAI in the context of Healthcare 5.0	System adaptability, user trust, technological integration	Identifies key opportunities and challenges, supports strategic planning, enhances future readiness	Broad exploration might lack depth, rapid technological changes can impact relevance
Niranjan, K., et al. (2023). [14]	Use of explainable AI techniques for COVID-19 diagnosis through fused classification and segmentation approaches	Diagnostic accuracy, model interpretability, response time	Enhances diagnostic accuracy, provides clear explanations, supports timely decision-making	High computational demands, need for extensive validation and testing

2.1. Research Gaps

Here are some research gaps identified from the above papers on explainable AI (XAI) for decision support systems in healthcare and safety domains:

1. Many studies highlight the development of XAI tools but do not delve deeply into how humans interpret or interact with these explanations in real-world clinical settings, leaving a gap in understanding the user experience.

2. The lack of focus on domain-specific challenges limits the generalizability of XAI tools across various medical fields, leaving room for research to address unique needs and constraints in different specialties.

3. Although many papers touch on the importance of trust, there is insufficient research on how transparency affects users' trust in XAI, especially for non-expert users such as patients and caregivers.

4. Ethical challenges and legal implications are underrepresented in most papers despite being crucial for widespread adoption in sensitive domains like healthcare and public safety.

These gaps highlight the need for further exploration of the interaction between AI models and users, balancing accuracy and interpretability, addressing ethical considerations, and ensuring the generalizability of XAI methods across domains and models.

ISSN : 2455-9679
SJIF Impact Factor : 6.190

III. Conclusions

This literature review demonstrates that Explainable AI has become a pivotal component in the development of Healthcare Decision Support Systems for disease diagnosis and prediction. XAI approaches, including GAMs, EBMs, and post-hoc explanation methods like SHAP, LIME, and K-LIME, enhance model transparency and clinician trust, enabling the adoption of AI in sensitive healthcare environments. The review highlights that while predictive accuracy has improved substantially, integrating interpretability, scalability, and workflow compatibility remains an ongoing challenge. Future research should focus on multimodal data integration, standardized metrics for evaluating explanation quality, and designing user-centric explanations that meet clinical requirements. Overall, XAI-enabled HDSS represents a promising pathway to reliable, transparent, and actionable AI-assisted healthcare, bridging the gap between predictive performance and clinical interpretability.

References

[1] Knapič, S., Malhi, A., Saluja, R., & Främling, K. (2021). Explainable artificial intelligence for human decision support system in the medical domain. Machine Learning and Knowledge Extraction, 3(3), 740-770.

[2] Tarnowska, K. A., Dispoto, B. C., & Conragan, J. (2021). Explainable AI-based clinical decision support system for hearing disorders. AMIA Summits on Translational Science Proceedings, 2021, 595.

[3] Iqbal, A., Zahid, S. B., & Arif, M. F. (2021). Artificial Intelligence for Safer Cities: A Deep Dive into Crime Prediction and Gun Violence Detection. International Journal Of Computer Science And Technology, 5(1), 547-552.

[4] Rehman, A., Farrakh, A., & Mushtaq, U. F. (2023). Improving Clinical Decision Support Systems: Explainable AI for Enhanced Disease Prediction in Healthcare. International Journal of Computational and Innovative Sciences, 2(2), 9-23.

[5] Bayer, S., Gimpel, H., & Markgraf, M. (2022). The role of domain expertise in trusting and following explainable AI decision support systems. Journal of Decision Systems, 32(1), 110-138.

[6] Antoniadi, A. M., Du, Y., Guendouz, Y., Wei, L., Mazo, C., Becker, B. A., & Mooney, C. (2021). Current challenges and future opportunities for XAI in machine learning-based clinical decision support systems: a systematic review. Applied Sciences, 11(11), 5088.

[7] Panigutti, C., Beretta, A., Fadda, D., Giannotti, F., Pedreschi, D., Perotti, A., & Rinzivillo, S. (2023). Co-design of human-centered, explainable AI for clinical decision support. ACM Transactions on Interactive Intelligent Systems.

[8] Pierce, R. L., Van Biesen, W., Van Cauwenberge, D., Decruyenaere, J., & Sterckx, S. (2022). Explainability in medicine in an era of AI-based clinical decision support systems. Frontiers in Genetics, 13, 903600.

[9] Wang, D., Yang, Q., Abdul, A., & Lim, B. Y. (2019, May). Designing theory-driven user-centric explainable AI. In Proceedings of the 2019 CHI conference on human factors in computing systems (pp. 1-15).

[10] Polat Erdeniz, S., Veeranki, S., Schrempf, M., Jauk, S., Ngoc Trang Tran, T., Felfernig, A., ... & Leodolter, W. (2022, September). Explaining machine learning predictions of decision support systems in healthcare. In Current Directions in Biomedical Engineering (Vol. 8, No. 2, pp. 117-120). De Gruyter.

[11] Srinivasu, P. N., Sandhya, N., Jhaveri, R. H., & Raut, R. (2022). From blackbox to explainable AI in healthcare: existing tools and case studies. Mobile Information Systems, 2022, 1-20.

[12] Holzinger, A., Biemann, C., Pattichis, C. S., & Kell, D. B. (2017). What do we need to build explainable AI systems for the medical domain?. arXiv preprint arXiv:1712.09923.

[13] Saraswat, D., Bhattacharya, P., Verma, A., Prasad, V. K., Tanwar, S., Sharma, G., ... & Sharma, R. (2022).

ISSN : 2455-9679
SJIF Impact Factor : 6.190

Explainable AI for healthcare 5.0: opportunities and challenges. IEEE Access.

[14] Niranjan, K., Kumar, S. S., Vedanth, S., & Chitrakala, S. (2023). An Explainable AI driven Decision Support System for COVID-19 Diagnosis using Fused Classification and Segmentation. Procedia computer science, 218, 1915-1925.

[15] Gerlings, J., Jensen, M. S., & Shollo, A. (2022). Explainable ai, but explainable to whom? an exploratory case study of xai in healthcare. Handbook of Artificial Intelligence in Healthcare: Vol 2: Practicalities and Prospects, 169-198.

[16] W. Yang, Y. Wei, H. Wei, Y. Chen, G. Huang, X. Li, R. Li, N. Yao, X. Wang, X. Gu, M. B. Amin and B. Kang, “Survey on Explainable AI: From Approaches, Limitations and Applications Aspects”, Human Centric Intelligent Systems, Vol. 3, pp. 161-188, 2023.

[17] F. D. Velez and B. Kim, “Towards a rigorous science of interpretable machine learning”, arXiv preprint arXiv:1702.08608, 2017.

[18] Pendyala, V., Kim, H. (2023). Assessing the Reliability of Machine Learning Models Applied to the Mental Health Domain using Explainable AI. Electronics, MDPI, 13, 1025, 1-21.

[19] A. Shrikumar, P. Greenside, and A. Kundaje, “Learning important features through propagating activation differences”, Proceedings of the 34th International Conference on Machine Learning, pp. 1-8, 2017.

[20] Kerz, E., Zanwar, S., Qiao, Y. and Wiechmann, D. (2023). Toward explainable AI (XAI) for mental health detection based on language behaviour. Front. Psychiatry, 14:1219479, 1-20.