Global Journal of Engineering and Technology [GJET].
ISSN: 2583-3359 (Online)
Frequency: Monthly
Published By GSAR Publishers
Journal Homepage Link- https://gsarpublishers.com/journal-gjet-home/

Explainable deep learning models for healthcare decision support

BY

Zafarul Hasan[1], Mohammad Serajuddin[2*], Syed Ahad Murtaza Alvi[3], Amjad Khan[4], Rashid Ayub[5]

[1] Researcher, College of Dentistry, King Saud University, Riyadh, Saudi Arabia
[2] Researcher, College of Nursing, King Saud University, Riyadh, Saudi Arabia
[3] College of Applied Computer Sciences, King Saud University, Riyadh, Saudi Arabia
[4] Researcher, Prince Sultan Bin Abdulaziz College for Emergency Medical Services
King Saud University, Riyadh, Saudi Arabia
[5]Rashid Ayub, College of Computer and Information Sciences, King Saud University, P. O. Box-2454, Riyadh-11451, Saudi Arabia

Article History

Received: 07/02/2024
Accepted: 13/02/2024
Published: 15/02/2024

Vol – 3 Issue – 2

PP: - 01-10

Abstract

Deep learning has recently become a potent tool in healthcare, excelling at tasks like analyzing medical images, diagnosing diseases, and predicting patient outcomes. The research discusses about Deep Learning in Healthcare: Current Applications and Challenges or Interpretable Techniques in Deep Learning along with Performance Evaluation of Deep Learning Models. It has also done Clinical Standards and Decision Support Systems. The Methodology discusses comprehending the subjective perceptions and interpretations of healthcare professionals regarding Explainable Deep Learning Models (EDLMs) in clinical decision-making, this study employs an interpretivist research philosophy.

Index Terms- *Deep learning models AI-driven, Healthcare prediction and modeling, clinical standards, transparency, interpretability, healthcare practitioners*

1. Introduction

1.1. Research background

Deep learning has recently become a potent tool in healthcare, excelling at tasks like analyzing medical images, diagnosing diseases, and predicting patient outcomes. The opacity of these models, however, prevents their widespread use in clinical practice. Because conventional deep learning models function as "black boxes," it might be difficult to comprehend how they make decisions. Interpretability is crucial in complex healthcare scenarios because physicians need to be able to understand model predictions in order to feel confident and trusted about their recommendations [1-3]. Therefore, the creation of Explainable Deep Learning Models (EDLMs) that may offer clear, intelligible, and clinically significant insights into their reasoning is urgently needed. By concentrating on strategies and methodologies for improving the interpretability of deep learning models for healthcare, this research intends to close this gap and eventually enable their incorporation into decision support systems for more efficient and dependable clinical decision-making [4-5].

1.2. Research aim and objectives

1.2.1. Aim

The aim of this study is to advance the field of healthcare decision support by creating and testing Explainable Deep Learning Models (EDLMs), which offer open and understandable insights into their decision-making process, this work

1.2.2. Objectives

- To create EDLMs that are specifically designed for healthcare applications while incorporating interpretability methods.
- To objectively assess EDLM performance in comparison to established models and clinical benchmarks.
- To convert model results into knowledge that practitioners may use in the clinic.
- To increase the EDLM's robustness and generalization for use in actual clinical settings.

1.3. Research Rationale

An important barrier in the incorporation of computer science (AI) into clinical practice is being addressed through studies on Explainable Deep Learning Models (EDLMs) for medical decision assistance. Although deep learning has shown tremendous promise in healthcare applications, the inherent opacity of traditional models makes it difficult to implement widely [6]. Understanding the overall reasoning behind AI-driven suggestions is very much essential for clinicians to trust as well as accept in high-stakes medical circumstances. Through offering clear along with understandable facts about the way they make accurate decisions, EDLMs seek to close this gap. This research is very necessary to make sure that clinical judgments made by AI are not only correct but also clear as well as actionable [7-9]. Healthcare providers can take advantage of cutting-edge AI technologies while still having a clear understanding of the thinking behind each advice by creating EDLMs.

2. LITERATURE REVIEW

2.1. Deep Learning in Healthcare: Current Applications and Challenges

Deep learning has become a game-changing technique in the healthcare industry, transforming a number of research and clinical practices. It has properly demonstrated astounding success in the interpretation of medical images, enabling precise detection of ailments ranging from cancers to fractures. Deep learning models have also been used to extract useful data from patient records, reports, and academic papers using natural language processing tasks. Deep learning skills have also been helpful for disease progression and patient outcome prediction modeling. Despite these developments, deep learning for healthcare has yet to be widely adopted [10-15]. The "black box" problem, or the intrinsic opacity of traditional models, is one of the main issues. In high-stakes medical circumstances, understanding the logic behind an algorithm's predictions is essential. Additionally, a major barrier is still the requirement for sizable, varied, and annotated datasets, particularly for uncommon diseases or niche medical fields and other fields of data mining [16-29].

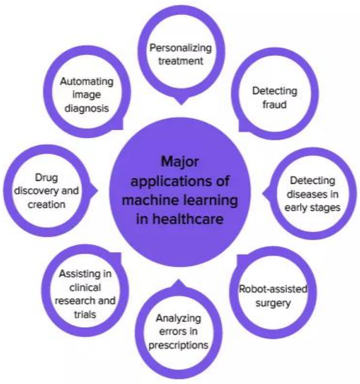

Fig. 2.1.1: Machine Learning in Healthcare

Additionally, continuing research and development is being done in the areas of guaranteeing model durability, privacy, and regulatory compliance. To fully utilize deep learning's promise for improving healthcare outcomes, these issues must be resolved.

2.2. Interpretable Techniques in Deep Learning

By addressing the opaqueness of deep learning models, interpretable techniques play a crucial role in enhancing their usability and reliability in crucial applications, particularly in the field of medicine. A popular technique called attention mechanisms enables models to concentrate on particular areas or features of the input data, giving insight into the process of making decisions. Saliency maps highlight key areas in a picture or sequence, making it easier to identify the factors that influence the model's predictions [30]. The contribution of each feature is quantified by its significance score, revealing the features that have the greatest impact on the model's output. Additionally, post-hoc explanations are provided by methods like LIME (Local Interpretable Model-agnostic Explanations) and SHAP (Shapley Additive Explanations), which approximate the behavior of complicated models.

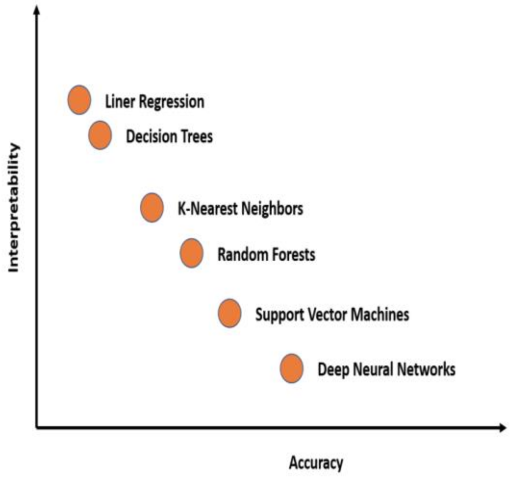

Fig. 2.2.1: Interpretable Machine Learning

They give features relevance ratings, offering insightful information about how particular inputs affect the model's output. Deep learning models' choices may be trusted and understood by researchers and doctors thanks to these interpretable methodologies, ensuring a seamless incorporation into healthcare systems that support decisions [31].

2.3. Performance Evaluation of Deep Learning Models in Healthcare

In order to successfully integrate deep learning models into clinical practice, it is essential to assess their efficacy in the healthcare industry. These models are evaluated according to their specific healthcare application-related parameters like as accuracy, sensitivity, and specificity. For tasks like tumor diagnosis or anomaly identification in medical imaging, measures like sensitivities (true positive rate) and specificity (true negative rate) are crucial. AUC-ROC, which measures the area under the receiver's operating characteristic curve, also offers a thorough evaluation of model performance [31]. Additionally, calibration assesses how closely expected probability match actual results, assuring accurate forecasts. To evaluate the generalizability of the model across various patient populations, cross-validation techniques are used.

*Corresponding Author: Mohammad Serajuddin

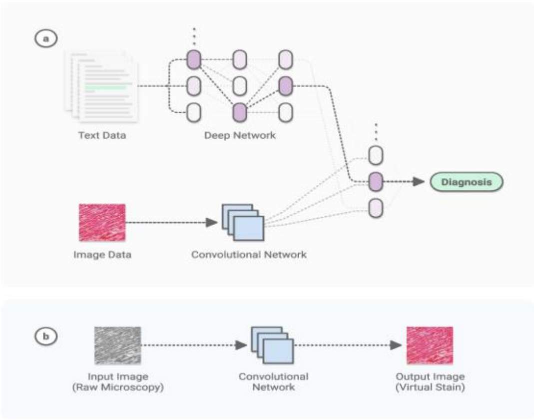

Fig. 2.3.1: Deep Learning Enabled Medical computer

Furthermore, the model's resilience and suitability for use in real-world scenarios are confirmed by external validation on other datasets. Studies that contrast deep learning models with more traditional approaches and accepted clinical standards also help to clarify the added value of these models in the healthcare industry [32].

2.4. Clinical Standards and Decision Support Systems

Modern healthcare procedures are supported by systems for supporting decisions and clinical standards. These standards include accepted protocols, best practices, and recommendations that direct healthcare personnel to deliver the best possible patient care. They are founded on a blend of clinical knowledge, evidence-based medicine, and patient choices. Clinical standards guarantee uniformity, security, and quality in the provision of healthcare in a variety of contexts. Healthcare professionals can use automated tools called decision support platforms (DSS) to help them make well-informed judgments [33]. They offer suggestions for evaluation, treatment planning, and evaluation using a combination of particular to the patient data, expertise in medicine, and algorithms. By providing timely, research-based advice, DSS can greatly improve clinical decision-making.

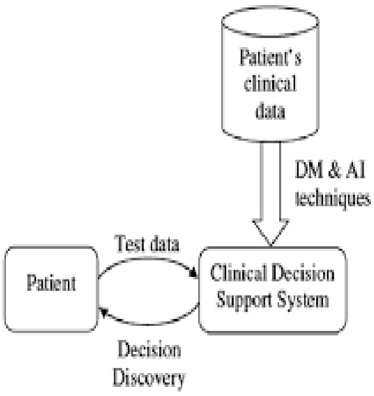

Fig. 2.4.1: Clinical Decision Support System

The potential to improve clinical standards by incorporating models of deep learning into systems that facilitate decisions is enormous. To retain trust and confidence in the advice these models offer, it is essential to guarantee their transparency and interpretability. This integration is a promising first step in the direction of achieving better and more individualized healthcare results [36-44].

2.5. Literature Gap

The material that is currently available on Explainable Deep Learning Models (EDLMs) in healthcare is mostly concerned with model building and performance assessment. Comprehensive studies that rigorously examine the conversion of the model's results into clinically significant insights and address the real-world applicability of EDLMs, nevertheless are noticeably lacking. This crucial component is still mostly unexplored and is a crucial subject for additional study in this domain.

3. METHODOLOGY

In order to comprehend the subjective perceptions and interpretations of healthcare professionals regarding Explainable Deep Learning Models (EDLMs) in clinical decision-making, this study employs an interpretivist research philosophy. With its emphasis on context or the social construction of reality, interpretivism is in line with the complex viewpoints of healthcare practitioners [45]. A deductive methodology is used, beginning with a conceptual structure built from prior literature and then testing hypotheses through the study of empirical data. Evaluating the applicability and efficacy of EDLMs in an environment of healthcare decision assistance is appropriate for this method [46]. To give a thorough picture of the situation and perceptions surrounding the use of EDLMs in healthcare settings, the study uses a descriptive methodology. Information is gathered by using secondary data sources. Peer-reviewed papers, proceedings of conferences, medical reports, and legislation fall under this category. Relevance, recency, and reliability will be used as the determining factors for these sources. In order to find pertinent studies, a thorough literature review is done. Databases are searched using keywords like "Explainable Deep Learning Models," "The medical field decision support," and related topics [47]. A detailed content analysis is performed on the data that was extracted.

Studies that concentrate on EDLM deployments in healthcare and offer insights into interpretation and usability for doctors and nurses must have been published during the last five years in order to meet the inclusion requirements. Important themes are noted about the interpretability of EDLMs, how to use them in making clinical choices, and any difficulties experienced by healthcare practitioners [48-64]. A thorough grasp of the state of research in this field is possible thanks to the identification of patterns and contradictions in the literature.

4. Results and Discussion

4.1. Perceptions of Interpretability in EDLMs

Concerning the ability to interpret Explainable Deep Learning Models (EDLMs) in clinical decision-making, healthcare professionals have a range of opinions. Some people have a lot of faith in how transparent these models are, emphasizing the ability to identify and comprehend the factors that affect the model's predictions. They believe that interpretability,

*Corresponding Author: Mohammad Serajuddin

particularly in situations where the model's advice directly affects patient care, is a crucial component in fostering confidence and acceptance among physicians [65]. On the other hand, certain experts might view EDLMs with some suspicion, voicing worries about the complicated nature of the algorithms that underlie them and the possibility of incorrect results. They see interpretability as a way to clarify these models' decision-making procedures and confirm their therapeutic applicability [66]. They stress how crucial it is to be able to question the outcomes of the model to make sure they line up with their individual clinical knowledge and intuition. Additionally, other medical professionals might have a more nuanced perspective, understanding that while accessibility is important, it shouldn't always be the only factor considered when assessing the effectiveness of EDLMs. They acknowledge that in some high-stakes situations, an appropriate balance between interpretability and accomplishment must be found, with a stronger focus on making accurate forecasts.

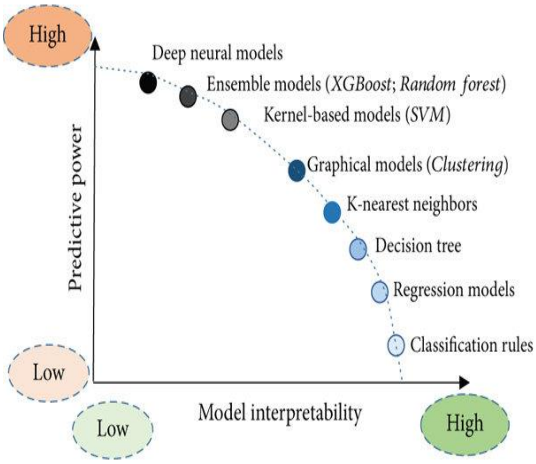

Fig. 4.1.1: Predictive Power VS. Interpretability

Along with this, some practitioners could see interpretability as one of several significant elements while still acknowledging its importance. To make sure that EDLMs are both comprehensible and practically usable in a variety of healthcare contexts, they underline the need for thorough validation studies, actual clinical trials, and incorporation feasibility assessments [67]. Overall, the various interpretations of accessibility in EDLMs highlight how difficult it is to apply cutting-edge machine learning models to clinical practice [68-72].

4.2. Usability in Clinical Decision-Making

Explainable Deep Learning Models (EDLMs) are thought to be most useful when they can be easily incorporated into clinical decision-making processes, according to healthcare practitioners. Many practitioners favor EDLMs that easily fit into their current workflows and cause little disturbance to established procedures or further training. They stress the importance of user-friendly interfaces that provide model outputs in a clear and understandable way. In addition, EDLMs that offer timely and useful advice are frequently preferred by healthcare professionals. They appreciate models that provide information in a clear, succinct manner so they may act fast and with knowledge [73]. This is especially important in settings for acute care where prompt treatments are necessary. Additionally, EDLMs that exhibit a high degree of adaptability to various patient demographics and clinical circumstances are valued by practitioners. Variations in patient characteristics, histories of illness, and particular clinical situations should be supported by the models. This adaptability makes sure that the suggestions made by the EDLMs are in line with the unique treatment requirements of each patient. Additionally, healthcare workers appreciate EDLMs that support group decision-making [74]. They are looking for models that can be a helpful tool for multidisciplinary teams, enabling for conversations and clinical consensus-building.

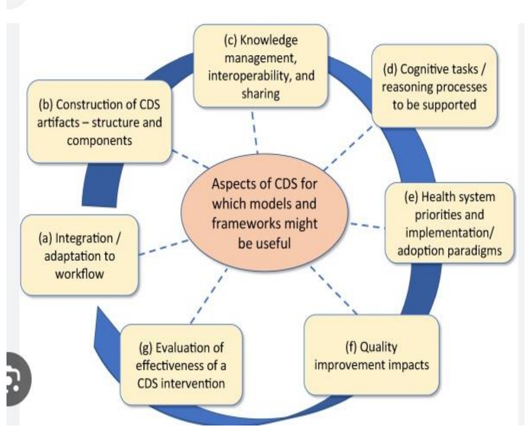

Fig. 4.2.1: Clinical Decision Support Model

The final treatment plan incorporates the knowledge and opinions of all team members thanks to this collaborative approach. Overall, it is believed that the effectiveness of EDLM uptake and its incorporation into clinical practice is directly related to their usefulness in medical decision-making [75-80].

4.3. Challenges and Concerns with EDLM Implementation

Healthcare practitioners face a number of significant difficulties and worries when adopting Explainable Deep Learning Models (EDLMs) in their clinical settings. The intricacy of these models is one of the main worries [81]. The complex workings underlying deep learning algorithms may be difficult for some practitioners to understand, especially when compared to more conventional, rule-based systems for supporting decisions. Due to their complexity, EDLMs may not be fully trusted or adopted in high-stakes healthcare situations. Furthermore, the deployment of EDLM raises questions regarding data security and privacy. Healthcare professionals are understandably concerned about the integrity and confidentiality of patient data [82]. They are concerned about potential flaws in the model distribution procedure that can expose private medical information. This issue highlights the importance of strong data protection procedures and adherence to strict regulatory standards [83-88]. The need for enormous computational resources presents another important challenge. For efficient conditioning and real-time inference in EDLMs, particularly those utilizing deep neural networks, strong hardware is required. Many healthcare facilities may experience difficulties purchasing and maintaining the essential computational infrastructure, especially for smaller facilities or those in environments with limited resources. Continuing education and training are also necessary for healthcare practitioners to use EDLMs properly.

Fig. 4.3.1: Issue and Challenging of Training Implementation

Critical components that require support and instruction include familiarity with understanding model outputs, comprehension of potential limitations, and incorporating the model into their current clinical processes. Last but not least, a major worry is the possibility of biases in the information used to train EDLMs. Biases in training information may result in skewed predictions, which may cause disparities in the provision of healthcare [89].

4.4. Desired Features and Improvements

In order to increase the usability and efficacy of Explainable Deep Learning Models (EDLMs) for clinical decision-making, healthcare practitioners articulate a number of desired features and prospective enhancements. The incorporation of real-time feedback methods is one important aspect. The ability to engage with the model and offer suggestions or clarifying queries to improve the recommendations is highly valued by practitioners [90]. This dynamic connection encourages the model and medical personnel to make decisions together. Practitioners also look for improved visualization methods within EDLM interfaces. Heatmaps and attention maps are excellent examples of simple and understandable graphical representations of model results [91]. These visual aids increase the trust of healthcare professionals in the recommendations by assisting them in immediately grasping the key variables impacting the model's predictions. Additionally, practitioners stress how crucial it is for the model's decision-making method to take contextual information into account. This comprises details about the patient, such as their medical background, comorbidities, and preferred course of therapy [92]. The relevance and usefulness of the predictions are increased by using this contextual information to guarantee that the model's proposals are tailored to the particular conditions of each patient [93]. The need for EDLMs to offer not only recommendations but also based on proof of reasons for those predictions is frequently expressed by practitioners.

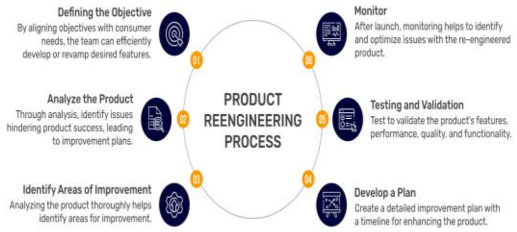

Fig. 4.4.1: Product Reengineering Process

This includes citations to case studies, recommendations, or other clinical literature that supports the model's conclusions. Greater assurance in the biological reliability of the model's results is provided by this feature. Additionally, models that allow for continual learning and adaptability are sought after by practitioners. The EDLM can be updated in response to fresh information and mounting evidence, ensuring that it stays current and correct as clinical practices and recommendations change over time [94]. The overall need for EDLMs to be shifting, interactive, and context-aware is reflected in these anticipated characteristics and enhancements [95].

5. Conclusion

The interpretability, usefulness, and practical application of Explainable Deep Learning Models (EDLMs) in healthcare support for choices have been the main topics of this study. The study has shed important light on the needs and perspectives of healthcare professionals through an interpretive lens and a deductive method. The results highlight the crucial role that interpretability plays in EDLMs, with clinicians rewarding models that provide open-and-shut views into their method of decision-making. Usability evolved as a crucial element, emphasizing the necessity for intuitive user interfaces, prompt advice, and flexibility to accommodate various clinical settings. However, issues with the complexity of the model, confidentiality of data, and required resources were all noted as major obstacles. Practitioners acknowledged a definite demand for features including real-time engagement, enhanced visualization methods, and contextual inclusion of patient data in response to these worries. Additionally, the need of ongoing education and the requirement for supporting predictions with facts was underlined [96]. This study emphasizes the significance of matching these models with the real-world demands and expectations of healthcare workers, laying the groundwork for expanding the introduction of EDLMs into clinical practice. In the end, EDLMs have the ability to greatly improve clinical decision-making by taking into account these factors, which will result in better patient outcomes and more efficient healthcare delivery [97].

5.1. Research recommendation

Continued Model Interpretability Research: To ensure that healthcare professionals can rely on and understand model outputs, additional research should explore cutting-edge methods for improving the usability of Explainable Deep Learning Models (EDLMs).

*Corresponding Author: Mohammad Serajuddin

User-Centric Interface Design: Designers as well as programmers should place a high priority on creating intuitive user interfaces that allow for easy communication between doctors and nurses and EDLMs, facilitating quick decision-making [98].

Strong Data Security Controls: To protect patient information and foster trust in the use of EDLM in healthcare settings, strict security and confidentiality of data measures must be implemented.

Continuous Training and Education: To guarantee that healthcare professionals are competent in utilizing and successfully understanding EDLM results, ongoing instruction should be made available to them.

5.2. Future work

The development of better interpretability strategies for Explainable Deep Learning Models (EDLMs), handling specific difficulties in complex medical contexts, should be the primary goal of future research. It would also improve user participation and decision-making to look for ways to seamlessly incorporate current feedback loops and adaptive communication capabilities within EDLM interfaces. Further research should examine how federated learning strategies might support decentralized healthcare information while protecting patient privacy [99-100]. Furthermore, in order to confirm the continued value and advantages of EDLMs in clinical practice, longitudinal studies evaluating their long-term effectiveness and effect are crucial.

Reference

1. Petch, J., Di, S. and Nelson, W., 2022. Opening the black box: the promise and limitations of explainable machine learning in cardiology. Canadian Journal of Cardiology, 38(2), pp.204-213.
2. Khan, S. (2020). Artificial Intelligence Virtual Assistants (Chatbots) are Innovative Investigators. *International Journal of Computer Science Network Security, 20*(2), 93-98.
3. Khan, S. (2021). Data Visualization to Explore the Countries Dataset for Pattern Creation. *International Journal of Online Biomedical Engineering, 17*(13), 4-19.
4. Bhatt, U., Xiang, A., Sharma, S., Weller, A., Taly, A., Jia, Y., Ghosh, J., Puri, R., Moura, J.M. and Eckersley, P., 2020, January. Explainable machine learning in deployment. In *Proceedings of the 2020 conference on fairness, accountability, and transparency* (pp. 648-657).
5. Khan, S. A., & Muneer, R. (2023). A Novel Thresholding for Prediction Analytics with Machine Learning Techniques. International Journal of Computer Science and Network Security, 23(1), 33.
6. Amann, J., Blasimme, A., Vayena, E., Frey, D. and Madai, V.I., 2020. Explainability for artificial intelligence in healthcare: a multidisciplinary perspective. *BMC Medical Informatics and Decision Making*, *20*(1), pp.1-9.
7. Rahman, M.A., Hossain, M.S., Alrajeh, N.A. and Guizani, N., 2020. B5G and explainable deep learning assisted healthcare vertical at the edge: COVID-I9 perspective. *IEEE Network*, *34*(4), pp.98-105.
8. Khan, S. (2021). Visual Data Analysis and Simulation Prediction for COVID-19 in Saudi Arabia Using SEIR Prediction Model. *International Journal of Online Biomedical Engineering, 17*(8).
9. Khan, S., & Alfaifi, A. (2020). Modeling of Coronavirus Behavior to Predict Its Spread. *International Journal of Advanced Computer Science Applications, 11*(5), 394-399.
10. Singh, A., Sengupta, S. and Lakshminarayanan, V., 2020. Explainable deep learning models in medical image analysis. *Journal of imaging*, *6*(6), p.52.
11. Ahmad, S., Khan, S., AlAjmi, M. F., Dutta, A. K., Dang, L. M., Joshi, G. P., & Moon, H. (2022). Deep Learning Enabled Disease Diagnosis for Secure Internet of Medical Things. *Computers, Materials Continua*, *73*(01), 965-979.
12. Khan, S., & AlAjmi, M. F. (2013). Impact of medical technology on expansion in healthcare expenses. *International Journal of Advanced Computer Science Applications, 4*(4).
13. Khan, S. (2021). Study Factors for Student Performance Applying Data Mining Regression Model Approach. *International Journal of Computer Science Network Security, 21*(2), 188-192.
14. R. M. Alotaibi and S. Khan, "Big Data and Predictive Data Analytics in the Smes Industry Using Machine Learning Approach," 2023 6th International Conference on Contemporary Computing and Informatics (IC3I), Gautam Buddha Nagar, India, 2023, pp. 2212-2217, doi: 10.1109/IC3I59117.2023.10397688.
15. Khan, S., Alqahtani, S. Hybrid machine learning models to detect signs of depression. Multimed Tools Appl (2023). https://doi.org/10.1007/s11042-023-16221-zYousef, R., Khan, S., Gupta, G., Albahlal, B. M., Alajlan, S. A., & Ali, A. (2023). Bridged-U-Net-ASPP-EVO and deep learning optimization for brain tumor segmentation. *Diagnostics*, *13*(16), 2633.
16. Khan, S., & Alshara, M. (2019). Development of Arabic evaluations in information retrieval. *International Journal of Advanced Applied Sciences*, *6*(12), 92-98.
17. Yousef, R., Khan, S., Gupta, G., Siddiqui, T., Albahlal, B. M., Alajlan, S. A., & Haq, M. A. (2023). U-Net-Based Models towards Optimal MR Brain Image Segmentation. *Diagnostics, 13*(9), 1624.
18. Khan, S., & Alshara, M. (2019). Fuzzy Data Mining Utilization to Classify Kids with Autism. *International Journal of Computer Science Network Security*, *19*(2), 147-154.

*Corresponding Author: Mohammad Serajuddin

19. Akram, A., Zafar, K., Mian, A. N., Baig, A. R., Almakki, R., AlSuwaidan, L., & Khan, S. (2023). On Layout Optimization of Wireless Sensor Network Using Meta-Heuristic Approach. *Computer Systems Science Engineering, 46*(3), 1-17.
20. Al Ajmi, D. M. F., Khan, S., & Khan, I. (2014). Cloud Computing Utilization for E-Learning Pharmaceutical System. *International Journal of Science & Technology Research*, *3*(3).
21. Alajmi, M., & Khan, S. (2011). *EFFECTIVE USE OF WEB 2.0 TOOLS IN PHARMACY STUDENT'S SKILLS PRACTICE DURING FIELD TRAINING.* Paper presented at the iceri2011 proceedings.
22. AlAjmi, M., & Khan, S. (2012). *Data Mining–Based, Service Oriented Architecture (SOA) In E-Learning.* Paper presented at the Iceri2012 Proceedings.
23. AlAjmi, M., & Khan, S. (2012). *The Utility of New Technologies in Enhancing Learning Vigilance in Educationally Poor Populations.* Paper presented at the EDULEARN12 Proceedings.
24. AlAjmi, M., & Khan, S. (2013). *Mobile Community Networks Information Investigation for Additional Significance.* Paper presented at the ICERI2013 Proceedings.
25. Thorsen-Meyer, H.C., Nielsen, A.B., Nielsen, A.P., Kaas-Hansen, B.S., Toft, P., Schierbeck, J., Strøm, T., Chmura, P.J., Heimann, M., Dybdahl, L. and Spangsege, L., 2020. Dynamic and explainable machine learning prediction of mortality in patients in the intensive care unit: a retrospective study of high-frequency data in electronic patient records. *The Lancet Digital Health*, *2*(4), pp.e179-e191.
26. AlAjmi, M., & Khan, S. (2015). *PART OF AJAX AND OPENAJAX IN CUTTING-EDGE RICH APPLICATION ADVANCEMENT FOR E-LEARNING.* Paper presented at the INTED2015 Proceedings.
27. AlAjmi, M., & Shakir, K. (2013). *Data Mining in Learning Management System utilizing Moodle.* Paper presented at the INTED2013 Proceedings.
28. AlAjmi, M. F., & Khan, S. (2011). *Effective Use Of Web 2.0 Tools Complex Pharmatical Skills Teaching And Learning.* Paper presented at the ICERI2011, 3rd International Conference on Education and New Learning Technologies, Spain.
29. Reddy, S., Allan, S., Coghlan, S. and Cooper, P., 2020. A governance model for the application of AI in health care. *Journal of the American Medical Informatics Association*, *27*(3), pp.491-497.
30. Sendak, M., Elish, M.C., Gao, M., Futoma, J., Ratliff, W., Nichols, M., Bedoya, A., Balu, S. and O'Brien, C., 2020, January. " The human body is a black box" supporting clinical decision-making with deep learning. In *Proceedings of the 2020 conference on fairness, accountability, and transparency* (pp. 99-109).
31. Magesh, P.R., Myloth, R.D. and Tom, R.J., 2020. An explainable machine learning model for early detection of Parkinson's disease using LIME on DaTSCAN imagery. *Computers in Biology and Medicine*, *126*, p.104041.
32. Angelov, P. and Soares, E., 2020. Towards explainable deep neural networks (xDNN). *Neural Networks*, *130*, pp.185-194.
33. Ghassemi, M., Oakden-Rayner, L. and Beam, A.L., 2021. The false hope of current approaches to explainable artificial intelligence in health care. *The Lancet Digital Health*, *3*(11), pp.e745-e750.
34. AlSuwaidan, L., Khan, S., Almakki, R., Baig, A. R., Sarkar, P., & Ahmed, A. E. (2022). Swarm Intelligence Algorithms for Optimal Scheduling for Cloud-Based Fuzzy Systems. *Mathematical Problems in Engineering, 2022*(Article ID 4255835), 11 pages. doi:10.1155/2022/4255835
35. Antony, M. J., Sankaralingam, B. P., Khan, S., Almjally, A., Almujally, N. A., & Mahendran, R. K. (2023).Brain–Computer Interface: The HOL–SSA Decomposition and Two-Phase Classification on the HGD EEG Data. *Diagnostics*, *13*(17), 2852.
36. Azrour, M., Mabrouki, J., Guezzaz, A., Ahmad, S., Khan, S., & Benkirane, S. (2024). IoT, Machine Learning, and Data Analytics for Smart Healthcare. In: CRC Press.
37. Chopra, P., Junath, N., Singh, S. K., Khan, S., Sugumar, R., & Bhowmick, M. (2022). Cyclic GAN Model to Classify Breast Cancer Data for Pathological Healthcare Task. *BioMed Research International*, *2022*(Article ID 6336700), 12 pages. doi:10.1155/2022/6336700
38. Fazil, M., Khan, S., Albahlal, B. M., Alotaibi, R. M., Siddiqui, T., & Shah, M. A. (2023). Attentional multi-channel convolution with bidirectional LSTM cell toward hate speech prediction. *IEEE Access, 11*, 16801-16811.
39. Gupta, G., Khan, S., Guleria, V., Almjally, A., Alabduallah, B. I., Siddiqui, T., . . . Al-Subaie, M. (2023). DDPM: A Dengue Disease Prediction and Diagnosis Model Using Sentiment Analysis and Machine Learning Algorithms. *Diagnostics, 13*(6), 1093.
40. Khan, S. (2016). How Data Mining Can Help Curb City Crime. *International Journal of Control Theory Applications, 9*(23), 483-488.
41. Khan, S., AlAjmi, D. M., Sarwar, A., & Akhtar, A. (2013). Keeping Data on Clouds: Cloud Computing Significance. *International Journal of Engineering Science Research IJESR, 3*(3).
42. Khan, S., AlAjmi, M., & Sharma, A. (2012). Safety Measures Investigation in Moodle LMS. *Special Issue of International Journal of Computer Applications, ICNICT*(4), 41-44.
43. Khan, S., & AlAjmi, M. F. (2017). *The Role of Open Source Technology in Development of E-Learning Education*. Paper presented at the

*Corresponding Author: Mohammad Serajuddin

Edulearn17 Proceedings.

44. Linardatos, P., Papastefanopoulos, V. and Kotsiantis, S., 2020. Explainable ai: A review of machine learning interpretability methods. *Entropy*, *23*(1), p.18.
45. Ghassemi, M., Oakden-Rayner, L. and Beam, A.L., 2021. The false hope of current approaches to explainable artificial intelligence in health care. *The Lancet Digital Health*, *3*(11), pp.e745-e750.
46. Longo, L., Goebel, R., Lecue, F., Kieseberg, P. and Holzinger, A., 2020, August. Explainable artificial intelligence: Concepts, applications, research challenges, and visions. In *International cross-domain conference for machine learning and knowledge extraction* (pp. 1-16). Cham: Springer International Publishing.
47. Zihni, E., Madai, V.I., Livne, M., Galinovic, I., Khalil, A.A., Fiebach, J.B. and Frey, D., 2020. Opening the black box of artificial intelligence for clinical decision support: A study predicting stroke outcome. *Plos one*, *15*(4), p.e0231166.
48. AlAjmi, M. F., Khan, S., & Sharma, A. (2014). *Collaborative learning outline for mobile environment.* Paper presented at the 2014 International Conference on Issues and Challenges in Intelligent Computing Techniques (ICICT).
49. Alfaifi, A. A., & Khan, S. G. (2022). Utilizing Data from Twitter to Explore the UX of "Madrasati" as a Saudi e-Learning Platform Compelled by the Pandemic. *Arab Gulf Journal of Scientific Research, 39*(3), 200-208. doi:10.51758/AGJSR-03-2021-0025
50. Almakki, R., AlSuwaidan, L., Khan, S., Baig, A. R., Baseer, S., & Singh, M. (2022). Fault Tolerance Byzantine Algorithm for Lower Overhead Blockchain. *Security Communication Networks, 2022*(Article ID 1855238), 9 pages. doi:10.1155/2022/1855238
51. Khan, S., & AlAjmi, M. F. (2019). A Review on Security Concerns in Cloud Computing and their Solutions. *International Journal of Computer Science Network Security, 19*(2), 10.
52. AlAjmi, M., Khan, S., & Zamani, A. S. (2012). Using instructive data mining methods to revise the impact of virtual classroom in e-learning. *International Journal of Advanced Science and Technology*, *45*(9), 125-134.
53. Khan, S., Al-Mogren, A. S., & AlAjmi, M. F. (2015). *Using cloud computing to improve network operations and management*. Paper presented at the 5th National Symposium on Information Technology: Towards New Smart World (NSITNSW).
54. Alrashed, F. A., Alsubiheen, A. M., Alshammari, H., Mazi, S. I., Al-Saud, S. A., Alayoubi, S., . . . Ahmad, T. (2022). Stress, Anxiety, and Depression in Pre-Clinical Medical Students: Prevalence and Association with Sleep Disorders. *Sustainability, 14*(18), 11320.
55. Khan, S., & Sharma, A. (2012). Moodle-Based LMS and Open Source Software (OSS) Efficiency in E-Learning. *International Journal of Computer Science Engineering Technology, 3*(4), 50-60.
56. Khan, S., Sharma, A., Zamani, A. S., & Akhtar, A. (2012). Data Mining for Security Purpose & Its Solitude Suggestions. *International Journal of Scientific Technology Research, 1*(7), 1-4.
57. Sharma, C., Khan, S., Mahajan, S., Alsagri, H. S., Almjally, A., Alabduallah, B. I., & Ansari, A. A. (2023). Lightweight Security for IoT. *Journal of Intelligent Fuzzy Systems*(Preprint), 1-17.
58. Khan, S., Siddiqui, T., Mourade, A., Alabduallah, B. I., Alajlan, S. A., almjally, A., . . . Alfaifi, A. (2023). Manufacturing industry based on dynamic soft sensors in integrated with feature representation and classification using fuzzy logic and deep learning architecture. *The International Journal of Advanced Manufacturing Technology*, 1-13.
59. AlAjmi, M. F., & Khan, S. (2014). Collaborative Pharmacy Student Learning Outline for Mobile Atmosphere. *International Journal of Advanced Computer Science Applications, 5*(3).
60. AlAjmi, M. F., Khan, S., & Alaydarous, A. (2014). Data Protection Control and Learning Conducted Via Electronic Media IE Internet. *International Journal of Advanced Computer Science and Applications, 5*(11), 85-91.
61. AlAjmi, M. F., Khan, S., & Sharma, A. (2013). Studying Data Mining and Data Warehousing with Different E-Learning System. *International Journal of Advanced Computer Science and Applications, 4*(1), 144-147.
62. Khan, S., & Alshara, M. A. (2020). Adopting Open Source Software for Integrated Library System and Digital Library Automation. *International Journal of Computer Science Network Security, 20*(9), 158-165.
63. Khan, S., & AlSuwaidan, L. (2022). Agricultural monitoring system in video surveillance object detection using feature extraction and classification by deep learning techniques. *Computers and Electrical Engineering, 102*, 108201.
64. Roscher, R., Bohn, B., Duarte, M.F. and Garcke, J., 2020. Explainable machine learning for scientific insights and discoveries. *Ieee Access*, *8*, pp.42200-42216.
65. Meske, C., Bunde, E., Schneider, J. and Gersch, M., 2022. Explainable artificial intelligence: objectives, stakeholders, and future research opportunities. *Information Systems Management*, *39*(1), pp.53-63.
66. Moncada-Torres, A., van Maaren, M.C., Hendriks, M.P., Siesling, S. and Geleijnse, G., 2021. Explainable machine learning can outperform Cox regression predictions and provide insights in breast cancer survival. *Scientific reports*, *11*(1), p.6968.
67. Tayyab, M., Hussain, A., Alshara, M. A., Khan, S.,

*Corresponding Author: Mohammad Serajuddin

Alotaibi, R. M., & Baig, A. R. (2022). Recognition of Visual Arabic Scripting News Ticker from Broadcast Stream. *IEEE Access*, *10*, 59189 - 59204. doi:10.1109/ACCESS.2022.3179366

68. Zamani, A. S., Miandad, M. J., & Khan, S. (2013). Data Center–Based, Service Oriented Architecture (SOA) in Cloud Computing. *International Journal of Computing Science Information Technology, 1*(1).
69. Agbley, B. L. Y., Li, J. P., Haq, A. U., Bankas, E. K., Mawuli, C. B., Ahmad, S., . . . Khan, A. R. (2023). Federated Fusion of Magnified Histopathological Images for Breast Tumor Classification in the Internet of Medical Things. *IEEE Journal of Biomedical Health Informatics*.
70. Khan, S., & Altayar, M. (2021). Industrial Internet of things: Investigation of the applications, issues, and challenges. *International Journal of Advanced Applied Sciences, 8*(1), 104-113. doi:10.21833/ijaas.2021.01.013
71. Khan, S., Ch, V., Sekaran, K., Joshi, K., Roy, C. K., & Tiwari, M. (2023). *Incorporating Deep Learning Methodologies into the Creation of Healthcare Systems.* Paper presented at the 2023 International Conference on Artificial Intelligence and Smart Communication (AISC).
72. Rudin, C. and Radin, J., 2019. Why are we using black box models in AI when we don't need to? A lesson from an explainable AI competition. *Harvard Data Science Review*, *1*(2), pp.1-9.
73. Zhou, J., Gandomi, A.H., Chen, F. and Holzinger, A., 2021. Evaluating the quality of machine learning explanations: A survey on methods and metrics. *Electronics*, *10*(5), p.593.
74. Chang, V., Bailey, J., Xu, Q.A. and Sun, Z., 2023. Pima Indians diabetes mellitus classification based on machine learning (ML) algorithms. *Neural Computing and Applications*, *35*(22), pp.16157-16173.
75. Khan, S., Fazil, M., Imoize, A. L., Alabduallah, B. I., Albahlal, B. M., Alajlan, S. A., . . . Siddiqui, T. (2023). Transformer Architecture-Based Transfer Learning for Politeness Prediction in Conversation. *Sustainability, 15*(14), 10828.
76. Khan, S., Moorthy, G. K., Vijayaraj, T., Alzubaidi, L. H., Barno, A., & Vijayan, V. (2023). *Computational Intelligence for Solving Complex Optimization Problems.* Paper presented at the E3S Web of Conferences.
77. Dr. Mohamed F AlAjmi, Dr. Arun Sharma Head and Shakir Khan, " Growing Cloud Computing Efficiency" International Journal of Advanced Computer Science and Applications(IJACSA), 3(5), 2012.
78. Nikolaidis, P., Ismail, M., Shuib, L., Khan, S., & Dhiman, G. (2022). Predicting Student Attrition in Higher Education through the Determinants of Learning Progress: A Structural Equation Modelling Approach. *Sustainability, 14*(20), 13584. doi:10.3390/su142013584
79. Rao, M. S., Modi, S., Singh, R., Prasanna, K. L., Khan, S., & Ushapriya, C. (2023). *Integration of Cloud Computing, IoT, and Big Data for the Development of a Novel Smart Agriculture Model.* Paper presented at the 2023 3rd International Conference on Advance Computing and Innovative Technologies in Engineering (ICACITE).
80. Angelov, P.P., Soares, E.A., Jiang, R., Arnold, N.I. and Atkinson, P.M., 2021. Explainable artificial intelligence: an analytical review. *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery*, *11*(5), p.e1424.
81. Islam, M.R., Ahmed, M.U., Barua, S. and Begum, S., 2022. A systematic review of explainable artificial intelligence in terms of different application domains and tasks. *Applied Sciences*, *12*(3), p.1353.
82. Van der Velden, B.H., Kuijf, H.J., Gilhuijs, K.G. and Viergever, M.A., 2022. Explainable artificial intelligence (XAI) in deep learning-based medical image analysis. *Medical Image Analysis*, *79*, p.102470.
83. Saleem, R. M., Bashir, R. N., Faheem, M., Haq, M. A., Alhussen, A., Alzamil, Z. S., & Khan, S. (2023). Internet of things-based weekly crop pest prediction by using deep neural network. *IEEE Access.*
84. Shakir Khan, M. F. A. (2013). The Open Source Software (OSS) Utilization in Project Scattered Computing Environments. *INTERNATIONAL JOURNAL OF SCIENTIFIC RESEARCH, 2*(2).
85. Singh, A. K., Khan, I. R., Khan, S., Pant, K., Debnath, S., & Miah, S. (2022). Multichannel CNN model for biomedical entity reorganization. *BioMed Research International*, *2022*(Article ID 5765629), 11 pages.
86. Solomon, D. D., Khan, S., Garg, S., Gupta, G., Almjally, A., Alabduallah, B. I., . . . Abdallah, A. M. A. (2023). Hybrid Majority Voting: Prediction and Classification Model for Obesity. *Diagnostics, 13*(15), 2610.
87. Sultan Ahmad, S. J., Abubaker E. M. Eljialy, Shakir Khan. (2021). A Systematic Review on e-Wastage Frameworks. *International Journal of Advanced Computer Science Applications*, *12*(12), 701-709.
88. Xu, F., Uszkoreit, H., Du, Y., Fan, W., Zhao, D. and Zhu, J., 2019. Explainable AI: A brief survey on history, research areas, approaches, and challenges. In *Natural Language Processing and Chinese Computing: 8th CCF International Conference, NLPCC 2019, Dunhuang, China, October 9–14, 2019, Proceedings, Part II 8* (pp. 563-574). Springer International Publishing.
89. Tran, K.A., Kondrashova, O., Bradley, A., Williams, E.D., Pearson, J.V. and Waddell, N., 2021. Deep learning in cancer diagnosis, prognosis

*Corresponding Author: Mohammad Serajuddin

and treatment selection. *Genome Medicine*, *13*(1), pp.1-17.

90. Holzinger, A., Malle, B., Saranti, A. and Pfeifer, B., 2021. Towards multi-modal causability with graph neural networks enabling information fusion for explainable AI. *Information Fusion*, *71*, pp.28-37.
91. Ahmed, I., Jeon, G. and Piccialli, F., 2022. From artificial intelligence to explainable artificial intelligence in industry 4.0: a survey on what, how, and where. *IEEE Transactions on Industrial Informatics*, *18*(8), pp.5031-5042.
92. Vassiliades, A., Bassiliades, N. and Patkos, T., 2021. Argumentation and explainable artificial intelligence: a survey. *The Knowledge Engineering Review*, *36*, p.e5.
93. Burkart, N. and Huber, M.F., 2021. A survey on the explainability of supervised machine learning. *Journal of Artificial Intelligence Research*, *70*, pp.245-317.
94. Hossain, M.S., Muhammad, G. and Guizani, N., 2020. Explainable AI and mass surveillance system-based healthcare framework to combat COVID-19-like pandemics. *IEEE Network*, *34*(4), pp.126-132.
95. Montani, S. and Striani, M., 2019. Artificial intelligence in clinical decision support: a focused literature survey. *Yearbook of medical informatics*, *28*(01), pp.120-127.
96. Grote, T. and Berens, P., 2020. On the ethics of algorithmic decision-making in healthcare. *Journal of Medical Ethics*, *46*(3), pp.205-211.
97. Amann, J., Blasimme, A., Vayena, E., Frey, D. and Madai, V.I., 2020. Explainability for artificial intelligence in healthcare: a multidisciplinary perspective. *BMC Medical Informatics and Decision Making*, *20*(1), pp.1-9.
98. Khan, S., & AlAjmi, M. F. (2014). Cloud Computing Safety Concerns in Infrastructure as a Service. Research Journal of Recent Sciences
99. Xiang, A., Sharma, S., Weller, A., Taly, A., Jia, Y., Ghosh, J., Puri, R., Moura, J.M. and Eckersley, P., 2020, January. Explainable machine learning in deployment. In *Proceedings of the 2020 conference on fairness, accountability, and transparency* (pp. 648-657).
100. Antony MJ, Sankaralingam BP, Khan S, Almjally A, Almujally NA, Mahendran RK.Brain–Computer Interface: The HOL–SSA Decomposition and Two-Phase Classification on the HGD EEG Data. *Diagnostics*. 2023; 13(17):2852.

*Corresponding Author: Mohammad Serajuddin