Healthcare and Medical Devices, Vol. 171, 2025, 170–179
https://doi.org/10.54941/ahfe1006203

Integrating Explainable Machine Learning Techniques for Predicting Diabetes: A Transparent Approach to AI-Driven Healthcare

Amos Njeru[1], Ruth Wario[2], Rosa Njagi[3], Lucy Gitonga[4], and Casam Nyaga[5]

[1]Department of Data Science, Grand Canyon University Street Address. 3300W. Camelback Road Phoenix, AZ 85017; Mailing Address. P.O. Box 11097. Phoenix, AZ 85061, United States America
[2]Department of Computer Science and Informatics, University of the Free State, Private Bag X13, Kestell 9866, Republic of South Africa
[3]Department of STC 3 Renal Transplant, Urology Inpatient Baylor Scott & White Health Central Texas2401 south 31st Street temple, Texas 76508, United States America
[4 and 5]School of Nursing and Public Health, Chuka University P.O. Box 109 – 60400 Chuka, Kenya

ABSTRACT

Diabetes mellitus is a global health concern affecting millions worldwide, with profound medical and socioeconomic implications. The increasing adoption of machine learning (ML) in healthcare has revolutionized clinical decision-making by enabling predictive diagnostics, personalized treatment plans, and efficient resource allocation. Despite their potential, many ML models are often regarded as "black boxes" due to their lack of transparency, which raises significant challenges in critical fields like healthcare, where explain ability is crucial for ethical and accountable decision-making (Hassija et al., 2024). Explainable Artificial Intelligence (XAI) has emerged as a solution to address these challenges by making ML models more interpretable and fostering trust among healthcare practitioners and patients. This paper explores the integration of XAI techniques with ML models for diabetes prediction, emphasizing their potential to enhance transparency, trust, and clinical utility. We present a comparative analysis of popular XAI methods, such as SHAP (Shapley Additive Explanations), LIME (Local Interpretable Model-agnostic Explanations), and attention mechanisms, within the context of healthcare decision support. These techniques are evaluated based on interpretability, computational efficiency, and clinical applicability, highlighting the trade-offs between accuracy and transparency. The study underscores the critical role of interpretability in advancing trust and adoption of AI-driven solutions in healthcare, while addressing challenges such as balancing model performance with explain ability. Finally, future directions for deploying explainable ML in healthcare are outlined, aiming to ensure ethical, transparent, and effective AI implementation.

Keywords: Diabetes prediction, Explainable artificial intelligence (XAI), Machine learning in healthcare and transparency and interpretability

INTRODUCTION

Diabetes mellitus presents a significant and escalating global health challenge, affecting an estimated 463 million adults in 2019, a figure projected to rise sharply in the coming decades (Verma et al., 2021). Diabetes mellitus presents a significant and escalating global health challenge, affecting an estimated 463 million adults in 2019, a figure projected to rise sharply in the coming decades (Verma et al., 2021). Beyond its profound medical implications, diabetes has far-reaching socioeconomic impacts that underscore the need for comprehensive approaches to its management and prevention.

In the labor market, diabetes significantly affects productivity and economic output. Individuals with diabetes often experience reduced work capacity due to complications such as fatigue, neuropathy, and frequent medical appointments. This not only results in absenteeism but also in "presenteeism," where individuals are present at work but unable to perform optimally. These productivity losses impose substantial economic burdens on both employers and economies, further exacerbating the societal cost of the disease.

On a social level, the impact of diabetes extends beyond the affected individuals, particularly influencing family dynamics. Children of parents with diabetes may face emotional stress, financial hardships, and a compromised quality of life due to their parents' frequent illnesses and healthcare needs. These children may take on caregiving roles at a young age, affecting their educational and social development. Additionally, the stigma and challenges associated with managing a chronic illness can strain familial relationships and social interactions, underscoring the broader societal impact of diabetes.

Addressing these multidimensional challenges requires a holistic approach that combines medical intervention, workplace accommodations, and social support systems to mitigate the far-reaching consequences of diabetes on individuals, families, and societies (Mujumdar & Vaidehi, 2019; Hasan et al., 2020). Accurate and timely prediction of diabetes is crucial to guiding preventive measures and optimizing treatment outcomes (Gowthami et al., 2024; Jaiswal, Negi, & Pal, 2021). Traditional statistical models, while interpretable, often fail to capture the intricate relationships and nonlinearities in clinical datasets (Sonar & JayaMalini, 2019). In contrast, machine learning (ML) models excel in handling such complexities, offering superior predictive capabilities but at the expense of transparency, which has earned them the moniker "black boxes" (Hassija et al., 2024).

Explainable Artificial Intelligence (XAI) seeks to bridge this gap by enhancing the interpretability of ML models, particularly in high-stakes domains like healthcare, where trust, accountability, and ethical decision-making are paramount (Rudin, 2019; Sadeghi et al., 2024). By elucidating the inner workings of ML algorithms, XAI fosters confidence among healthcare professionals and patients, enabling informed clinical decisions while addressing regulatory and ethical requirements (Nasarian et al., 2024). For example, methods such as SHAP (Shapley Additive Explanations) and LIME (Local Interpretable Model-agnostic Explanations) offer insight into model predictions, while attention mechanisms provide a means to highlight clinically significant features (Ahmed et al., 2024; Kalusivalingam et al., 2021).

This paper examines the integration of XAI techniques with ML models for diabetes prediction, focusing on their ability to balance the trade-offs between accuracy and interpretability. By conducting a comparative analysis of popular XAI methods, we evaluate their strengths and limitations in terms of interpretability, computational efficiency, and clinical applicability (Guidotti et al., 2018; Abdullah, Zahid, & Ali, 2021). Moreover, the study underscores the critical role of interpretability in fostering the trust and adoption of AI-driven healthcare solutions, emphasizing the need to balance model performance with transparency (Petch, Di, & Nelson, 2022; Carabantes, 2020). Future directions for deploying explainable ML in healthcare are also outlined, aiming to ensure ethical, transparent, and effective implementation of AI in clinical practice.

MACHINE LEARNING FOR DIABETES PREDICTION

Machine learning (ML) models have shown exceptional promise in predicting diabetes, leveraging features such as glucose levels, BMI, age, and family history. These models have significantly enhanced the ability to identify individuals at risk, contributing to early intervention and better disease management.

Commonly used ML models for diabetes prediction include:

i. Logistic Regression (LR): A baseline model for binary classification, logistic regression is favored for its simplicity and interpretability. It provides a direct understanding of how features contribute to diabetes risk, making it ideal for initial assessments (Soni & Varma, 2020).
ii. Decision Trees (DTs): These models are easily interpretable due to their hierarchical structure. DTs perform well with a smaller dataset but may overfit if not pruned appropriately (Mujumdar & Vaidehi, 2019).
iii. Random Forests (RFs) and Gradient Boosting Machines (GBMs): Ensemble methods like RFs and GBMs combine the predictions of multiple weak learners to improve accuracy. However, they trade interpretability for performance, as their complexity often obscures the decision-making process (Hasan et al., 2020).
iv. Support Vector Machines (SVMs) and Neural Networks (NNs): These models are highly effective for complex data structures, offering superior predictive power. However, their black-box nature makes interpretability challenging, necessitating the integration of explainable AI (XAI) techniques to make their outputs actionable (Sarwar et al., 2018; Khanam & Foo, 2021).

Recent studies underscore the importance of integrating diverse models to enhance predictive performance. For instance, ensemble approaches combining multiple classifiers have achieved high accuracy and robustness in diabetes prediction (Hasan et al., 2020; Sonar & JayaMalini, 2019). Additionally, SVMs and NNs have proven effective for large, multidimensional datasets, but their adoption in healthcare settings has been limited by their complexity and lack of transparency (Jaiswal, Negi, & Pal, 2021).

Despite advancements, a key limitation of many ML models lies in their interpretability. The need for explainable and interpretable outputs is critical for healthcare applications, where clinicians must understand the reasoning behind predictions to make informed decisions (Mujumdar & Vaidehi, 2019). This challenge highlights the importance of explainable AI (XAI) to bridge the gap between predictive accuracy and actionable insights.

EXPLAINABLE AI (XAI) TECHNIQUES

EXPLAINABLE ARTIFICIAL INTELLIGENCE (XAI) METHODS

Explainable Artificial Intelligence (XAI) methods are essential for interpreting machine learning (ML) models, particularly in high-stakes fields like healthcare. These methods can be broadly categorized into two main approaches:

Intrinsic Interpretability

Some ML models, such as Decision Trees and Logistic Regression, are intrinsically interpretable. These models enable clinicians to directly understand the contribution of features to predictions, making them ideal for applications requiring transparency and trust (Dwivedi et al., 2023; Chaddad et al., 2023). Their simplicity allows for clear decision-making and supports straightforward clinical adoption.

Post-Hoc Explanation Techniques

Post-hoc XAI methods are designed to enhance the interpretability of complex ML models without altering their predictive mechanisms. Prominent techniques include:

- SHAP (SHapley Additive Explanations): SHAP quantifies the contribution of each feature to a model's prediction, offering a comprehensive explanation that aligns with game theory principles. It has been widely adopted in healthcare to elucidate feature importance in models such as Random Forests and Neural Networks (Ekanayake et al., 2022; Nohara et al., 2022).
- LIME (Local Interpretable Model-Agnostic Explanations): LIME approximates the behavior of a complex model locally by using a simpler surrogate model. This approach enables clinicians to understand individual predictions, making it particularly useful in case-based reasoning and anomaly detection (Zafar & Khan, 2021; Zhao et al., 2021).
- Feature Importance Scores: These scores highlight the relative significance of features, particularly in ensemble models like Gradient Boosting Machines, offering insights into the key drivers of model predictions (Speith, 2022; Antwarg et al., 2021).

- Visualization Techniques: Visualization tools, such as saliency maps for Neural Networks, provide a graphical representation of the regions or features contributing most to the model's decisions. These techniques are especially useful in domains like histopathology and energy systems (Graziani et al., 2021; Machlev et al., 2022).

Post-hoc XAI methods bridge the gap between the predictive accuracy of complex models and the need for actionable insights, fostering trust and reliability in their deployment (Abusitta et al., 2024; Parisineni & Pal, 2024). By employing these methods, practitioners can ensure that ML predictions are both effective and interpretable, aligning with the demands of critical applications like healthcare and energy systems.

INTEGRATION OF XAI IN DIABETES PREDICTION

The integration of Explainable Artificial Intelligence (XAI) in diabetes prediction has enhanced both accuracy and interpretability, providing valuable insights for clinical decision-making. Diabetes, particularly Type 2 diabetes mellitus (T2DM), requires effective predictive models that not only yield accurate results but also provide explainable outputs to ensure trust and applicability in healthcare settings.

Machine Learning Algorithms

Several machine learning (ML) techniques have been leveraged in diabetes prediction:

- **Tree-Based Models**: Models such as XGBoost, Kernel-Tree Boosting (KTBoost), and Natural Gradient Boosting (NGBoost) stand out due to their ability to handle complex datasets while maintaining high accuracy. These models are particularly effective in managing missing data and reducing overfitting (Arslan et al., 2024; Tasin et al., 2023).
- **Other Classifiers**: Decision trees, Support Vector Machines (SVM), Random Forests, and ensemble approaches also contribute to diabetes prediction, with XGBoost frequently showing accuracy rates between 81% and 98% (Tasin et al., 2022; Nagaraj et al., 2022).

Feature Selection and Data Processing

Feature selection techniques like LASSO regression enhance model interpretability by identifying significant biomarkers linked to T2DM. Additionally, robust preprocessing methods manage missing values and improve data quality, further boosting the reliability of ML models (Alsaleh et al., 2022; Naik et al., 2021).

Explainability Techniques

XAI tools such as SHAP (SHapley Additive exPlanations) and LIME (Local Interpretable Model-agnostic Explanations) are pivotal in understanding model decisions. These techniques clarify the influence of individual features on predictions, fostering transparency and increasing trust in AI-driven healthcare systems (Srinivasu et al., 2024; Vinh & Byeon, 2024).

Key Findings

- **Enhanced Accuracy**: Integrating metabolomics data with advanced ML models has significantly improved prediction outcomes. For instance, KTBoost paired with metabolomics profiling excels in early diagnosis and personalized treatment strategies (Arslan et al., 2024; Tasin et al., 2022).
- **Trustworthy Predictions**: By offering explainable outputs, XAI strengthens the reliability of AI-driven systems in clinical settings. This transparency is critical for understanding and implementing recommendations effectively (Alsaleh et al., 2022; Nagaraj et al., 2022).

CHALLENGES IN IMPLEMENTING EXPLAINABLE ARTIFICIAL INTELLIGENCE (XAI)

The implementation of Explainable Artificial Intelligence (XAI) presents numerous challenges across various domains. A significant obstacle lies in the **trade-off between accuracy and interpretability**. High-performing machine learning models, such as deep neural networks, often function as "black boxes," making their internal workings opaque. Attempts to enhance explainability can compromise their predictive accuracy, creating tension between interpretability and performance (Das & Rad, 2020; Antoniadi et al., 2021). This trade-off is particularly critical in applications where decision-making transparency is essential, such as healthcare and finance.

Another key challenge is the **lack of standardized evaluation metrics** to assess the quality of explanations. Current evaluation methods are often subjective, relying on human judgment to determine whether an explanation is comprehensible or actionable (Arrieta et al., 2020; Saeed & Omlin, 2023). The absence of universally accepted benchmarks complicates the comparison of different XAI approaches and hinders their adoption in high-stakes fields, such as digital pathology and cybersecurity (Evans et al., 2022; Senevirathna et al., 2024).

Moreover, there are challenges associated with **domain-specific requirements**. For instance, in healthcare, the need for explanations that align with clinical reasoning is critical. However, current XAI methods often fail to bridge the gap between algorithmic outputs and domain-specific knowledge, limiting their utility in clinical decision support systems (Antoniadi et al., 2021; Hulsen, 2023). Similarly, in security applications, the complexity of underlying systems, such as 5G networks, requires explanations that are both technically accurate and accessible to stakeholders (Senevirathna et al., 2024).

The **explainability paradox** further complicates XAI implementation. As explanations become more detailed, they may also become more complex, undermining their accessibility to non-expert users. This paradox poses significant challenges in ensuring that explanations are both comprehensive and user-friendly (Evans et al., 2022; Weber et al., 2023). Addressing this issue requires balancing the granularity of explanations with their interpretability.

Lastly, there is a growing concern about the **ethical and social implications** of XAI. Inadequate explanations can exacerbate biases and inequalities, particularly in applications where fairness is paramount. Researchers emphasize the need for XAI systems to not only explain decisions but also ensure they are unbiased and equitable (Adadi & Berrada, 2018; Longo et al., 2024). This challenge underscores the importance of integrating ethical considerations into the design and evaluation of XAI systems.

In summary, implementing XAI faces numerous challenges, including the accuracy-interpretability trade-off, lack of standardized evaluation metrics, domain-specific requirements, the explainability paradox, and ethical concerns. Addressing these issues requires interdisciplinary collaboration and innovation to realize the full potential of XAI across diverse applications.

FUTURE DIRECTIONS

The future of AI-driven healthcare lies in achieving a harmonious balance between predictive accuracy and interpretability. Explainable Artificial Intelligence (XAI) has demonstrated potential in enhancing transparency and fostering trust in clinical settings. To maximize its impact on diabetes prediction and management, several key areas warrant attention.

Hybrid Models for Enhanced Transparency

Future research should explore hybrid approaches that integrate interpretable models (e.g., linear regression or decision trees) with post-hoc XAI techniques, such as SHAP or LIME. This combination can ensure robust predictive performance while maintaining explainability, making the models more practical for clinical applications (Ann Jo & Deni Raj, 2023). These hybrids would allow clinicians to both trust and verify AI-driven recommendations.

Real-World Testing and Validation

Although XAI has proven effective in controlled experiments, its performance in real-world settings remains underexplored. Future efforts should involve deploying XAI-enhanced models in diverse clinical environments to evaluate their impact on patient outcomes, workflow efficiency, and acceptance by healthcare providers (Kong et al., 2024). Additionally, these studies should focus on addressing operational challenges such as integration with electronic health records (EHRs) and data privacy concerns.

Education and Training for Healthcare Professionals

Equipping clinicians and other healthcare stakeholders with the knowledge to interpret and trust XAI outputs is crucial for adoption. Training programs should focus on demystifying AI concepts and emphasizing how XAI can support evidence-based decision-making. Interdisciplinary collaboration between AI experts and healthcare educators can bridge the knowledge gap, fostering a culture of trust and reliance on AI tools (Abulibdeh et al., 2024).

Personalization and Equity

Future advancements in XAI for diabetes prediction should aim to improve personalization by accounting for individual patient factors such as genetics, lifestyle, and socio-economic background. Simultaneously, researchers must prioritize addressing health disparities by ensuring that models are inclusive and equitable, avoiding biases that could adversely impact underserved populations (Kalusivalingam et al., 2021).

Continuous Learning Systems

Developing XAI-driven systems capable of continuous learning from new data will enhance their adaptability and relevance in dynamic healthcare environments. Such systems should be designed to incorporate feedback from clinicians and patients, enabling iterative improvements and sustained accuracy over time (Javed et al., 2023).

Regulatory and Ethical Considerations

As XAI becomes integral to healthcare, regulatory frameworks must evolve to include guidelines for transparency, accountability, and ethical use. Future research should explore methods to ensure that XAI systems adhere to these principles while safeguarding patient data and respecting user autonomy (Rane & Paramesha, 2024).

CONCLUSION

The integration of Explainable Artificial Intelligence (XAI) techniques with machine learning (ML) models offers a transformative pathway for enhancing diabetes prediction and management. By addressing the "black box" nature of traditional ML models, XAI ensures transparency, accountability, and trust qualities that are indispensable in the healthcare domain. This paper has highlighted the importance of explainability in fostering the adoption of AI-driven solutions, offering a balance between predictive accuracy and interpretability.

The study demonstrates that XAI methods, such as SHAP and LIME, provide actionable insights, enabling clinicians to understand and trust AI-generated predictions. Moreover, the exploration of hybrid models and feature selection techniques underscores the potential for achieving both high performance and transparency. However, challenges such as the trade-offs between interpretability and complexity, ethical concerns, and the need for standardized evaluation metrics remain critical barriers to widespread adoption.

Future research must focus on refining XAI approaches to enhance their clinical applicability, scalability, and inclusivity. By aligning technological advancements with the ethical and practical demands of healthcare, XAI can serve as a cornerstone for building trustworthy, patient-centered AI systems. As the healthcare industry continues to evolve, embracing explainable solutions will be pivotal in driving better outcomes and fostering long-term confidence in AI technologies.

REFERENCES

Abulibdeh, A., Zaidan, E., & Abulibdeh, R. (2024). Navigating the confluence of artificial intelligence and education for sustainable development in the era of industry 4.0: Challenges, opportunities, and ethical dimensions. *Journal of Cleaner Production, 140527*.

Adadi, A., & Berrada, M. (2018). Peeking inside the black-box: A survey on explainable artificial intelligence (XAI). *IEEE Access, 6*, 52138–52160.

Alsaleh, M. M., Yeon, K. M., & SohailAkhtar, Q. M. S. J. (2022). XAI implementation on preliminary data analysis phase: Explainable output application with prediction of diabetes mellitus at early stage. *Turkish Journal of Computer and Mathematics Education (TURCOMAT), 13(2)*, 1070–1078.

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.

Arrieta, A. B., Díaz-Rodríguez, N., Del Ser, J., Bennetot, A., Tabik, S., Barbado, A.,... & Herrera, F. (2020). Explainable artificial intelligence (XAI): Concepts, taxonomies, opportunities and challenges toward responsible AI. *Information Fusion, 58*, 82–115.

Arslan, A. K., Yagin, F. H., Algarni, A., Karaaslan, E., Al-Hashem, F., & Ardigò, L. P. (2024). Enhancing type 2 diabetes mellitus prediction by integrating metabolomics and tree-based boosting approaches. *Frontiers in Endocrinology, 15*, 1444282.

Das, A., & Rad, P. (2020). Opportunities and challenges in explainable artificial intelligence (XAI): A survey. *arXiv preprint arXiv:2006.11371*.

Dauda, O. I., Awotunde, J. B., AbdulRaheem, M., & Salihu, S. A. (2022). Basic issues and challenges on explainable artificial intelligence (XAI) in healthcare systems. *Principles and methods of explainable artificial intelligence in healthcare*, 248–271.

Ding, W., Abdel-Basset, M., Hawash, H., & Ali, A. M. (2022). Explainability of artificial intelligence methods, applications and challenges: A comprehensive survey. *Information Sciences, 615*, 238–292.

Evans, T., Retzlaff, C. O., Geißler, C., Kargl, M., Plass, M., Müller, H.,... & Holzinger, A. (2022). The explainability paradox: Challenges for XAI in digital pathology. *Future Generation Computer Systems, 133*, 281–296.

Gowthami, S., Reddy, R. V. S., & Ahmed, M. R. (2024). Exploring the effectiveness of machine learning algorithms for early detection of Type-2 Diabetes Mellitus. *Measurement: Sensors, 31*, 100983.

Hassija, V., Chamola, V., Mahapatra, A., Singal, A., Goel, D., Huang, K.,... & Hussain, A. (2024). Interpreting black-box models: A review on explainable artificial intelligence. *Cognitive Computation, 16(1)*, 45–74.

Hasan, M. K., Alam, M. A., Das, D., Hossain, E., & Hasan, M. (2020). Diabetes prediction using ensembling of different machine learning classifiers. *IEEE Access, 8*, 76516–76531.

Hulsen, T. (2023). Explainable artificial intelligence (XAI): Concepts and challenges in healthcare. *AI, 4(3)*, 652–666.

Jaiswal, V., Negi, A., & Pal, T. (2021). A review on current advances in machine learning-based diabetes prediction. *Primary Care Diabetes, 15(3)*, 435–443.

Javed, A. R., Ahmed, W., Pandya, S., Maddikunta, P. K. R., Alazab, M., & Gadekallu, T. R. (2023). A survey of explainable artificial intelligence for smart cities. *Electronics, 12(4)*, 1020.

Kalusivalingam, A. K., Sharma, A., Patel, N., & Singh, V. (2021). Leveraging federated learning and explainable AI for advancing health equity: A comprehensive approach to reducing disparities in healthcare access and outcomes. *International Journal of AI and ML, 2(3).*

Khanam, J. J., & Foo, S. Y. (2021). A comparison of machine learning algorithms for diabetes prediction. *ICT Express, 7(4)*, 432–439.

Kong, X., Xing, Y., Tsourdos, A., Wang, Z., Guo, W., Perrusquia, A., & Wikander, A. (2024). Explainable interface for human-autonomy teaming: A survey. *arXiv preprint arXiv:2405.02583*.

Longo, L., Brcic, M., Cabitza, F., Choi, J., Confalonieri, R., Del Ser, J.,... & Stumpf, S. (2024). Explainable artificial intelligence (XAI) 2.0: A manifesto of open challenges and interdisciplinary research directions. *Information Fusion, 106*, 102301.

Mujumdar, A., & Vaidehi, V. (2019). Diabetes prediction using machine learning algorithms. *Procedia Computer Science, 165*, 292–299.

Naik, H., Goradia, P., Desai, V., Desai, Y., & Iyyanki, M. (2021). Explainable artificial intelligence (XAI) for population health management–an appraisal. *European Journal of Electrical Engineering and Computer Science, 5(6)*, 64–76.

Nagaraj, P., Muneeswaran, V., Dharanidharan, A., Balananthanan, K., Arunkumar, M., & Rajkumar, C. (2022). A prediction and recommendation system for diabetes mellitus using XAI-based lime explainer. In *2022 International Conference on Sustainable Computing and Data Communication Systems (ICSCDS)* (pp. 1472–1478). IEEE.

Rane, N. L., & Paramesha, M. (2024). Explainable artificial intelligence (XAI) as a foundation for trustworthy artificial intelligence. *Trustworthy Artificial Intelligence in Industry and Society*, 1–27.

Rawal, A., McCoy, J., Rawat, D. B., Sadler, B. M., & Amant, R. S. (2021). Recent advances in trustworthy explainable artificial intelligence: Status, challenges, and perspectives. *IEEE Transactions on Artificial Intelligence, 3(6)*, 852–866.

Saeed, W., & Omlin, C. (2023). Explainable AI (XAI): A systematic meta-survey of current challenges and future opportunities. *Knowledge-Based Systems, 263*, 110273.

Senevirathna, T., La, V. H., Marchal, S., Siniarski, B., Liyanage, M., & Wang, S. (2024). A survey on XAI for 5G and beyond security: Technical aspects, challenges, and research directions. *IEEE Communications Surveys & Tutorials*.

Soni, M., & Varma, S. (2020). Diabetes prediction using machine learning techniques. *International Journal of Engineering Research & Technology (IJERT), 9(09)*, 2278–0181.

Sonar, P., & JayaMalini, K. (2019, March). Diabetes prediction using different machine learning approaches. In *2019 3rd International Conference on Computing Methodologies and Communication (ICCMC)* (pp. 367–371). IEEE.

Swati, Vijarania, M., Jaglan, V., & Le, D. N. (2024). The need for XAI: Challenges and its applications. In *Reshaping Intelligent Business and Industry: Convergence of AI and IoT at the Cutting Edge*, 69–80.

Tasin, I., Nabil, T. U., Islam, S., & Khan, R. (2022). Diabetes prediction using machine learning and explainable AI techniques. *Healthcare Technology Letters, 10(1–2)*, 1–10. https://doi.org/10.1049/htl2.12039

Tasin, I., Nabil, T. U., Islam, S., & Khan, R. (2023). Diabetes prediction using machine learning and explainable AI techniques. *Healthcare Technology Letters, 10(1–2)*, 1–10.

Vinh, T. Q., & Byeon, H. (2024). Towards transparent diabetes prediction.