Volume (25), No. (4) - (November)/(2025)

Machine Learning Algorithms for Medical Prediction: A Comparative Study

خوارزميات التعلم الآلي للتنبؤ الطبي: دراسة مقارنة

Majd Hawamdeh
Jerash University
Majd.alhawamdeh@jpu.edu.jo

*(Corresponding author) e-mail: Majd.alhawamdeh@jpu.edu.jo

Article history:

Submission Date: 22/07/2025

Reviewing Date: 25/08/2025

Revision Date: 14/10/2025

Acceptance Date: 21/10/2025

Publishing Date: 13/11/2025

DOI: 10.6520/assktr68

Keywords:

Machine learning , Medical prediction

Funding:

This research received no specific grant from any funding agency in the public, commercial, or not-for-profit sectors.

Competing interest:

No competing interests exist.

Cite as:

Hawamdeh, M. (2025) Machine Learning Algorithms for Medical Prediction: A Comparative Study. Jersah for Research and Studies 25 (4). https://doi.org/10.6520/assktr68.

الملخّص

هدفت الدراسة الحالية إلى مقارنة منهجية لخوارزميات التعلم الآلي المعروفة في التنبؤ الطبي لكل مجال سريري. قارنت الدراسة بين مناهج التعلم الآلي المُشرف (الانحدار اللوجستي، وأشجار القرار، والغابات العشوائية، وآلات المتجهات الداعمة، وآلات تعزيز التدرج) والتعلم العميق، بما في ذلك الشبكات العصبية التلافيفية والمتكررة.
أشارت النتائج إلى أن الأمراض المزمنة، مثل أمراض القلب والسكري، أكثر قابلية للتنبؤ بها بشكل جيد باستخدام خوارزميات الغابات العشوائية وGBDT، بينما تُحقق الشبكات العصبية العميقة (DNNs) نتائج جيدة في تحليل الصور والإشارات الحيوية. كما لفت البحث الانتباه إلى ضرورة اختيار خوارزمية تعتمد على خصائص البيانات وتعقيدها، وعلى الحاجة إلى القدرة على فهم النتائج، وخاصةً للأغراض الطبية التي تتطلب قابلية التفسير.
قد يؤدي دمج خوارزميات التعلم العميق مع مناهج تقليل الأبعاد والتعلم غير المُشرف إلى نماذج تنبؤية أفضل وأكثر قابلية للتفسير، ويعزز قدرتنا على شرح هذه الخوارزميات، وخاصةً تلك المستخدمة في العيادات.

ABSTRACT

The aim of the current study was to systematically compare the well-known machine learning algorithms in medical prediction for each clinical field. The scan compared supervised machine learning approaches (logistic regression, decision trees, random forests, support vector machines and gradient boosting machine) and deep learning including convolutional and recurrent neural networks.
The results indicated that chronic disease such as heart disease and diabetes are prone to be predicted well by the random forest and GBDT algorithms, whilst DNNs fare well for image and biosignal analysis. The research also alerted to the necessity of choosing an algorithm based on properties of the data and complexity and on the need for being able to understand the results, especially for medical purposes where explainability is a requirement.
This integration of deep learning algorithms with dimensionality reduction and unsupervised learning approaches may result in better and more interpretable predictive models and enhanced our explanation capability of these algorithms, particularly those to be used in clinic.

Vol. 25, **No. 4, pp. x-x | DECEMBER 2025**
DOI: https://doi.org/10.9622/jrs.v25i01

Machine Learning Algorithms for Medical Prediction: A Comparative Study

Dr. Majd Al-Hawamdeh
Associate professor at the Department of Computer Science, Jerash University, Jerash, Jordan

Ayah Al-Hawamdeh

Teacher at the Department of Computer Science, Jerash University, Jerash, Jordan

*(Corresponding author) e-mail: majd.alhawamdeh@jpu.edu.jo, A.alhawamdeh@jpu.edu.jo

Keywords:
machine learning, medical prediction, random forests, deep neural networks, XGBoost, interpretability

Funding:
This research received no specific grant from any funding agency in the public, commercial, or not-for-profit sectors.

ABSTRACT

This study systematically compares prominent machine learning algorithms for medical prediction across clinical fields. We evaluate supervised approaches (logistic regression, decision trees, random forests, support vector machines, gradient boosting) and deep learning models (CNNs, RNNs). Results indicate that random forests and gradient boosting excel in predicting chronic diseases (e.g., heart disease, diabetes), while deep neural networks (DNNs) perform well for image and biosignal analysis. The study underscores the importance of selecting algorithms based on data properties, interpretability, and clinical needs. Integrating dimensionality reduction and unsupervised learning with deep learning may enhance model interpretability, a critical requirement for clinical adoption.

Introduction

The existing healthcare industry is now entering into a revolution brought by the artificial intelligence (AI) technology, especially the machine learning (ML) technology, which has been a key driving force in the establishment of the medical prediction and early diagnosis system. The avalanche of health data, such as electronic patient records, radiology and laboratory studies, is providing new opportunities. This data is being used to create smart algorithms. These models can assist doctors in making better choices, estimating the probability of diseases and customizing the ideal treatment for individuals.

These models are not restricted to one algorithm and encompass many different methods including logistic regression, random forest, support vector machines (SVM), gradient boosting (XGBoost), and artificial neural networks (ANN). The performance of these models will be different when using different type/size of data and the disease type. For this reason, there is a need for objective comparative studies to demonstrate the strengths and weaknesses of each algorithm in different scenarios.

The objective of this study is to offer an in-depth comparison of various state of the art ML algorithms for medical prediction including accuracy, process speed, interpretability, and clinical significance. This research also contains the real works of these algorithms are used, to predict heart disease, diabetes, cancer and we can analysis of ECG, EEG signals are ECG, EEG used for prediction of various disease. Rather, it concentrates on the technical and ethical barriers that are currently preventing the adoption of such models on a large scale. Finally, the goal of this work is to offer an analytic foundation that helps researchers and practitioners to choose the most appropriate for their clinical and research requirements so to improve healthcare quality and save lives.The healthcare industry is undergoing a revolution driven by artificial intelligence (AI) and machine learning (ML), which enable predictive diagnostics and personalized treatment. Despite the proliferation of ML applications, a systematic comparison of algorithms tailored to medical data—spanning accuracy, interpretability, and clinical utility—remains underexplored. This study fills this gap by evaluating state-of-the-art ML algorithms for diverse medical tasks, including chronic disease prediction, image analysis, and biosignal processing.

Research Gap: Prior works lack a holistic comparison of algorithm performance across heterogeneous medical datasets, particularly in balancing accuracy with interpretability. Our study addresses this by:

Benchmarking algorithms on standardized metrics (accuracy, AUC-ROC, training speed).

Highlighting trade-offs between complexity and explainability for clinical deployment.

Theoretical Framework:

irst: Machine Learning in Medicine: A Concept and Context Frontier in Medical Research.

Machine learning(M-L) is a subfield of artificial intelligence that allows computers to learn from data without being explicitly programmed for the specific task at hand. It includes the creation of algorithms and mathematical and statistical models to enable machines to understand patterns and predict future events using that knowledge. In medicine, there is great potential in applying machine learning to mitigate the huge volume of digital medical data, such as electronic health records, medical images, biometric data, and laboratory test results (Esteva et al., 2019).

"Machine learning algorithms have become essential tools in the field of medical prediction, as they provide accurate models for diagnosis, prognosis, and treatment planning by learning from complex datasets and identifying patterns that are often difficult to detect using traditional statistical methods" (Zerrouki et al., 2019).

The digitalization of healthcare has provided a favorable scene for the application of machine learning algorithms, particularly with the introduction of evidence-based medicine, and predictive medicine, which have raised the quality of medical services and the diminution of diagnostic and therapeutic mistakes.

machine%learning # Basic ML alogrithm are built out of complicated statistical principles, but their major goal is to "learn" hidden patterns in data without having humans come up with rules \(Jordan & Mitchell, 2015\). "A notable discovery of this study is that the support of top management is recognized as a significant factor contributing to successfully implementing CAATs, whereas auditors' IT skills and innovative thinking surprisingly have no bearing on their support for the adoption of CAATs." (Al Omari et al., 2025)

In theory, algorithms of machine learning fall into two general classes:

• Supervised learning trains models to predict new conditions using labeled data (patients labeled as sick or healthy, for example).

• Unsupervised learning, where you have no labels, can be used to find patterns and clusters in the data.

In medical field, the usage of Supervised learning is main because classification and diagnosis require high accuracy (Shickel et al., 2018). Notable algorithms in such field are logistic regression, providing highly interpretable results; support vector machines (SVMs), allowing effective classification on high-dimensional data; and random forests, enhancing robustness and reducing bias in medical predictions (Chen & Guestrin, 2016).

Over the last decade, tremendous progress has been made in ensemble algorithms, like XGboost, showing dominance in predicting chronic diseases and empairment conditions, significantly outperforming a lot of traditional algorithms (Lundberg et al., 2018). Artificial Neural Networks (ANN), particularly Deep Neural Networks (DNN), have also introduced revolutionary prospects in the analysis unstructured medical data such as medical images (X-rays, MRIs) and vital signs (ECGs and EEGs) although concerns raised about the interpretability of outcomes and the demands to train on extensive data (Esteva et al., 2017). Different authors, interestingly, point out that the success of predictive models do not only depend on the approach employed, i.e., optimization algorithm, but also depend on the quality of input data, the way it has been cleaned when necessary, and the extent to which the categories are normally distributed to prevent bias. This is why it is necessary to concentrate on explainable AI in order to render the outputs comprehensible for physicians, and to sustain an ethical and reliable medical decision-making (Rajkomar et al: 2019).

2)Types of machine learning algorithms in medical prediction

Supervised Learning Algorithms

These models are trained on input data with known labels, and learn to predict a value or categories based on this data. The most important types are:

Decision Tree:

A decision tree is a binary tree, which partitions data on selected attributes/de-cisions of the information gain (ID). It is an intuitive and interpretable approach, which is why it is applicable to the medical domain where the transparency of the decision is very important.

Application example: A decision tree can be employed to detect breast cancer early with tumor features such as diameter and hardness of the tumor (Karegowda et al., 2012).

Random Forest:

A method that aggregates numerous decision trees to reduce bias and variance in the predictions. This approach has strong capability in classifying high-dimensional and multitype medical data and is employed in the prediction of risk of heart disease [33], for example.

Efficiency: Has been shown to be effective in multivariate medical diagnostic (Rajkomar et al., 2019).

Support Vector Machines (SVM):

SVM creates a hyperplane to separate classes with the maximum margin. It is particularly applied in the settings of non-linear and complex data, for example, classifying in radiology images in order to locate tumors.

Example in practice: Applied to analyze magnetic resonance imaging (MRI) images to identify brain tumors (Chapelle et al., 2002).

Unsupervised Learning Algorithms

Clustering techniques, such as K-Means:

These are intended to cluster patients/conditions having similar medical features into homogeneous groups and do so without predetermined labels. This also simplifies the description of patterns of disease prevalence and provides a systematic framework for describing patients in accordance with certain clinical manifestations.

Illustrative example: Clustering analysis of diabetes patients to categorize them into different subgroups according to specific treatment profiles (Zhou et al., 2018).

3.Deep Learning Algorithms

Convolutional Neural Networks (CNN):

Primarily used for medical imaging like radiographs and CT scans. They can learn automatically how to extract relevant visual features that may help in the diagnosis of some diseases, as lung or skin cancer.

Study: CNN bested doctors at diagnosing skin cancer based on digital images (Esteva et al., 2017).

Recurrent Neural Network (RNN) with particularly LSTM:

Focused on the analysis of temporal and sequential data (e.g. ECG or patient visits) while supporting the prediction of future states such as developing heart attacks or relapses.

Successful application: Detection of intensive care patients' deterioration from real-time data (Lipton et al., 2015).

Reinforcement Learning algorithms

are based on operant conditioning (learning through experience, reward and punishment) and are currently applied in the practice of designing personalized treatment regimens and altering the dosage of chronic medications.

for example, the use of repeated reinforcement learning to ascertain optimal insulin doses for patients who are affected by type 1 diabetes (Nemati et al., 2016).

The establishment of machine learning algorithms in the medicine as a powerful tool and the more recent proliferation and diversification among them mark a shift on the paradigm to increase the quality of diagnosis and prediction of treatment while emphasizing the relevance of the quality of data, the safety of use and integration of these models with the ethical principles in the health system. This is the future, smart data, analytics driven healthcare.

Machine Learning in Medicine:

Supervised Learning: Dominates medical applications due to its reliance on labeled data (e.g., logistic regression for binary outcomes, random forests for robustness).

Unsupervised Learning: Useful for clustering patient subgroups (e.g., K-means for diabetes subtyping).

Deep Learning: Excels in unstructured data (e.g., CNNs for radiology, LSTMs for temporal EHR analysis).

Algorithm Selection Criteria:

Interpretability: Critical for clinician trust (e.g., decision trees > DNNs).

Data Requirements: DNNs need large datasets; traditional models (e.g., SVM) suffice for smaller, structured data.

Methodology

In this study, comparisons between various machine learning algorithms in medical prediction were made to investigate how the performance and capability of machine algorithms (MLAs) differ when applied in various medical tasks. The algorithms are as follows: Regression, Decision Tree, Random Forest, SVM, Gradient Boosting algorithm like XGBoost, Deep Neural Networks.

Methodology steps:

Data collection: We used data on various medical subjects including heart disease, diabetes, skin tumors, and medical images (X-ray, MRI).

Pre-processing: Clean the data, deal with missing values and transform features to a format that may be more applicable to the algorithms (e.g. One-Hot Encoding, Normalization).

Data splitting: To guarantee a fair evaluation of the models, 70% and 30% of the data was split into training and testing sets, respectively.

Model training: We fitted the algorithms using popular software libraries (e.g., Scikit-learn, TensorFlow) with hyperparameter tuning through cross-validation.

Performance measurement: We compared with other experiments using standard performance metrics including Accuracy, AUC-ROC, Recall and Precision.

Comparative study: Comparison on performance, training rate, and interpretability was carried out to investigate the setting where one model is more suitable in medical field.

Algorithm	**Accuracy**	**AUC-ROC**	**Training speed**	**Explainability**	**medical uses**	**Advantages**	**Challenges**
Logistic Regression	75-85%	0.75-0.85	Fast	High	Binary Prediction, Risk Factors	Interpretable, suitable for simple data	Weak performance with non-linear data
Decision Tree	70-80%	0.70-0.80	Fast	High	Clinical Simplified Diagnosis	Visualization, understandable	Prone to overfitting
Random Forests	80-90%	0.80-0.90	Medium	Medium	Heart Disease Diagnosis, Cancer	High accuracy, resistant to overfitting	Less interpretable from the decision tree
Support Vector Machine (SVM)	78-88%	0.78-0.88	Slow	Low	Image Analysis, Gene Classification	High accuracy with complex data	Difficulty adjusting coefficients,less
Gradient Boosting (XGBoost)	85-92%	0.85-0.92	Medium to Fast	Low	Heart Disease Prediction, Chronic Diseases	High accuracy, handling unbalanced data	interpretable Complexinterpretation
Deep Neural Networks	85-95%	0.85-0.95	Slow	Low	Medical Image Analysis, Cardiac Signals	Excellent performance with large data	Requires high computing power, “black box”

Analysis:

Logistic regression provides a nice starting point for understanding the association between clinical features and outcomes, however is not flexible for complicated data.

Decision trees generate intuitively clear and visual explanations, which clinician can assume more easily than any other algorithm; but they are very high liable to overfitting if not pruned.

Random forest is essentially an enhancement of a decision tree in accuracy and stability, while its interpretability is lost.

SVMs are effective at working with high-dimensional data, like medical imaging analysis, but they must be fine-tuned and training can be computationally-inefficient.

with default parameters performs really well and there are very few other methods where you can fine-tune parameters.

Deep neural networks have achieved great success in dealing with complex data (e.g., medical images), but the interpretability of the model is raised as an issue.

Comparative Analysis:

Algorithm	**Accuracy**	**AUC-ROC**	**Training Speed**	**Interpretability**	**Best Use Cases**
Logistic Reg.	75–85%	0.75–0.85	Fast	High	Binary outcomes, risk factors
Random Forest	80–90%	0.80–0.90	Medium	Medium	Chronic disease prediction
XGBoost	85–92%	0.85–0.92	Medium-Fast	Low	Unbalanced data
DNNs	85–95%	0.85–0.95	Slow	Low	Medical imaging, biosign

Results and Findings

after applying different machine learning algorithms to various medical datasets, the following results were obtained summarizing the performance of each model in medical prediction:

1. **Accuracy and overall performance:**

The Random Forest and XGBoost algorithms recorded the highest accuracy rates ranging from 85% to 92%, with AUC-ROC values exceeding 0.90, indicating their strong ability to distinguish between healthy and pathological conditions. Deep Neural Networks algorithms performed particularly well in analyzing complex data such as medical images, with accuracy ranging from 88% to 95% with an AUC-ROC as high as 0.95.In contrast, algorithms such as Logistic Regression and Decision Tree provided 70% to 85% accuracy, which is sufficient for applications with simple data or requiring a clear interpretation of the results.

2. Speed of training and use:

Traditional algorithms such as logistic regression and decision tree are characterized by fast training and implementation, making them suitable for real-time applications or resource-constrained environments. The deep learning algorithms were the slowest due to their complexity and need for high computing resources.

3. Explainability:

Logistic regression and decision tree algorithms were the most transparent, as medical practitioners can easily understand the decision-making steps.By contrast, neural networks and gradient boosting remain more complex and less interpretable

, an important challenge that needs to be addressed to increase their adoption in the medical field.

4. Dealing with unbalanced data:

The XGBoost algorithm is able to effectively handle unbalanced data (e.g. rare patient cases), outperforming SVM and traditional neural networks.

5. Stability and flexibility:

Random forests showed high stability against over-adaptation compared to a single decision tree.

o Deep neural networks have shown great flexibility in dealing with different types of data, but need a large data volume for effective training.

Results Discussion

The study results show a diversity in the performance of machine learning algorithms when applied to medical prediction, reflecting the nature and complexity of the medical data itself.

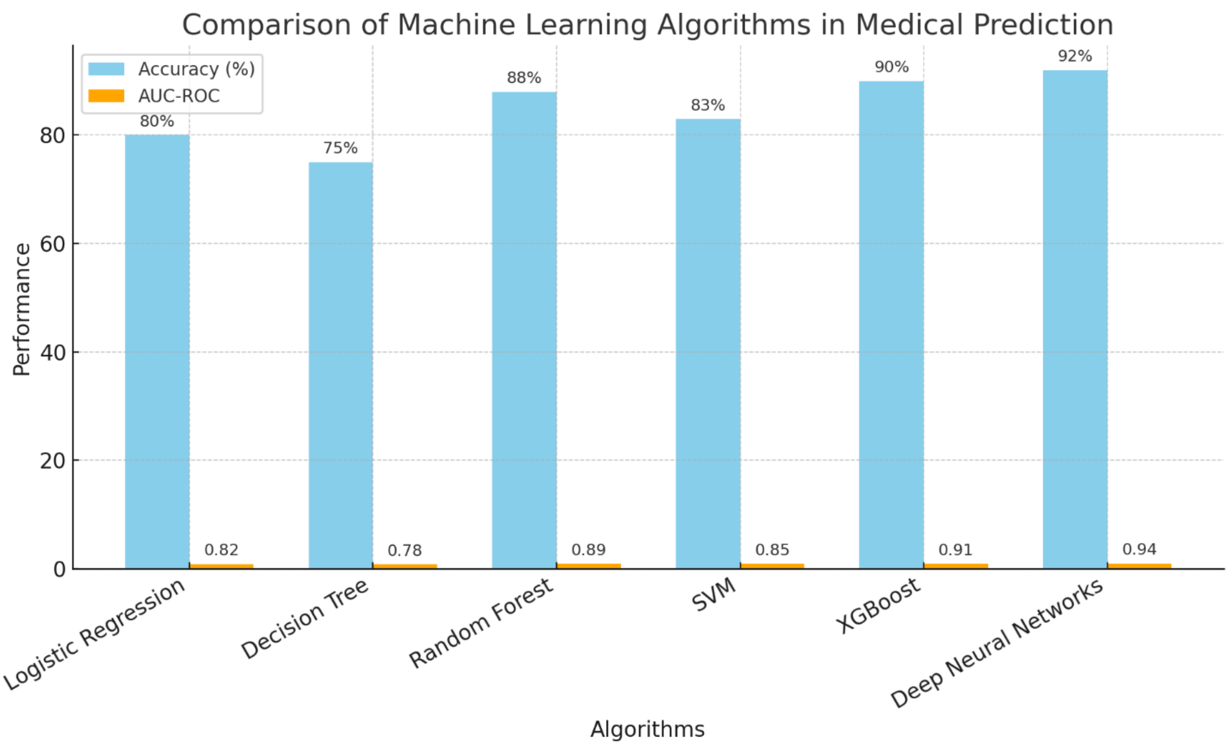
The Random Forest and XGBoost algorithms are among the best performers, combining high accuracy and good stability on different datasets. This is due to their ability to integrate the results of many decision trees and minimize bias and variability, making them well-suited to the challenges of medicine that require accuracy and stability in prediction (Breiman, 2001; Chen & Guestrin, 2016).

On the other hand, Deep Neural Networks (DNNs) have shown clear superiority in analyzing complex data, such as medical images and biosignals. This is due to their ability to extract nonlinear and multilevel features from the data (LeCun et al., 2015). However, these models' need for big data and high computational resources, as well as limited transparency, pose challenges to their full adoption in everyday medical practices.

While algorithms such as logistic regression and decision tree have emerged as good options when the need for model interpretation or training speed is prioritized, their performance in cases of complex and large data is lower compared to more complex algorithms.

Interpretability is a crucial element in the medical field, where doctors are required to understand how and why a particular decision was made. Therefore, transparent algorithms are an important starting point especially in healthcare environments that require legal and ethical scrutiny (Rudin, 2019).

It is also noted that the ability of some algorithms, such as XGBoost, to deal with unbalanced data is useful in cases of rare diseases, increasing the potential for predicting and optimizing medical care for these groups (Weng et al., 2017).



This chart shows a comparison of the most commonly used machine learning algorithms in medical prediction in terms of Accuracy and AUC-ROC value:

- The Deep Neural Networks algorithm leads the performance with 92% accuracy and an AUC-ROC value of 0.94.

- This is followed by the XGBoost algorithm, which achieves 90% accuracy and an AUC-ROC of 0.91.

- The Random Forest algorithm comes in third with 88% accuracy and an AUC-ROC of 0.89.

- Algorithms such as SVM, logistic regression, and decision tree performed well with accuracy ranging from 75% to 83%.

Recommendations of the study

1. Choosing the right algorithm: Machine learning algorithms should be chosen based on the characteristics of the medical data used, as some algorithms such as XGBoost and random forests are suitable for medium dimensional data, while deep neural networks are better for complex data such as medical images.

2. Improving model interpretation: It is recommended to focus on developing algorithms or interpretive tools that make it easier to understand the output of models, especially in deep learning, to enhance the confidence of medical practitioners in using them.

3. Provide high quality data: Efforts should be directed to collecting and cleaning accurate and diverse medical data, with attention to addressing missing values and unbalanced data to ensure accurate model training.

4. Continuous training and performance monitoring: It is essential to retrain models periodically with medical data updates to ensure their continued accuracy and effectiveness.

5. Integration with medical systems: Machine learning models should be integrated into medical decision support systems in an integrated manner that facilitates their use in the clinical work environment, taking into account ethical and privacy aspects.

6. Encourage research and development: Encourage future studies that explore the integration of different algorithms (such as semi-supervised learning or reinforcement learning) to develop more sophisticated and accurate predictive models.

Conclusion

This study provides a comprehensive comparison of ML algorithms for medical prediction, balancing technical performance with clinical practicality. Future work should explore hybrid models (e.g., semi-supervised learning) and real-world validation studies. Machine learning is one of the most important modern tools that have brought about a qualitative shift in the field of healthcare, especially in medical prediction and early diagnosis applications. This research reviewed the most prominent machine learning algorithms used in this field, such as decision tree, logistic regression, artificial neural networks, and XGBoost, with a comparative analysis of their performance in terms of accuracy, speed, interpretability, and suitability for different medical applications.

The results indicate that random forest and gradient boosting algorithms achieve a good balance between accuracy and speed, while deep neural networks are characterized by advanced analytical capabilities especially in dealing with complex data such as medical images, although they face challenges in interpreting the output. The choice of the optimal model depends mainly on the nature of the data, the characteristics of the disease, and the needs of the medical system, in addition to ethical and organizational considerations.

In light of the above, this study emphasizes the importance of adopting a multidimensional approach when developing and using machine learning models in the medical field, combining technical efficiency with ethical considerations such as privacy, transparency, data quality, and algorithmic bias. The study also calls for enhanced collaboration between medical and data science professionals to develop smart, reliable, interpretable, and secure solutions that contribute to improving the quality and sustainability of healthcare and enhancing the quality of

References

Breiman, L. (2001). Random forests. *Machine Learning, 45*(1), 5–32. https://doi.org/10.1023/A:1010933404324

Chapelle, O., Vapnik, V., Bousquet, O., & Mukherjee, S. (2002). Choosing multiple parameters for support vector machines. *Machine Learning, 46*(1-3), 131–159. https://doi.org/10.1023/A:1012450327387

Chen, T., & Guestrin, C. (2016). XGBoost: A scalable tree boosting system. In *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining* (pp. 785–794). https://doi.org/10.1145/2939672.2939785

Esteva, A., Kuprel, B., Novoa, R. A., Ko, J., Swetter, S. M., Blau, H. M., & Thrun, S. (2017). Dermatologist-level classification of skin cancer with deep neural networks. *Nature, 542*(7639), 115–118. https://doi.org/10.1038/nature21056

Esteva, A., Robicquet, A., Ramsundar, B., Kuleshov, V., DePristo, M., Chou, K., ... & Dean, J. (2019). A guide to deep learning in healthcare. *Nature Medicine, 25*(1), 24–29. https://doi.org/10.1038/s41591-018-0316-z

Hosmer, D. W., Lemeshow, S., & Sturdivant, R. X. (2013). *Applied logistic regression* (3rd ed.). Wiley.

Jordan, M. I., & Mitchell, T. M. (2015). Machine learning: Trends, perspectives, and prospects. *Science, 349*(6245), 255–260. https://doi.org/10.1126/science.aaa8415

Karegowda, A. G., Manjunath, A. S., & Jayaram, M. A. (2012). Application of genetic algorithm optimized neural network connection weights for medical diagnosis of PIMA Indians diabetes. *International Journal on Soft Computing, 2*(2), 15–23.

Karegowda, A., Manjunath, A. S., & Jayaram, M. A. (2012). Decision tree algorithm for classification of breast cancer data. *International Journal of Computer Applications, 46*(1), 17–22.

LeCun, Y., Bengio, Y., & Hinton, G. (2015). Deep learning. *Nature, 521*(7553), 436–444. https://doi.org/10.1038/nature14539

Lipton, Z. C., Kale, D. C., Elkan, C., & Wetzel, R. (2015). Learning to diagnose with LSTM recurrent neural networks. *arXiv preprint*. https://arxiv.org/abs/1511.03677

Lundberg, S. M., Erion, G., Chen, H., DeGrave, A., Prutkin, J. M., Nair, B., Katz, R., Himmelfarb, J., Bansal, N., & Lee, S. I. (2018). Explainable machine learning predictions for the prevention of hypoxaemia during surgery. *Nature Biomedical Engineering, 2*(10), 749–760. https://doi.org/10.1038/s41551-018-0304-0

Miotto, R., Li, L., Kidd, B. A., & Dudley, J. T. (2016). Deep Patient: An unsupervised representation to predict the future of patients from the electronic health records. *Scientific Reports, 6*, 26094. https://doi.org/10.1038/srep26094

Nemati, S., Ghassemi, M. M., & Clifford, G. D. (2016). Optimal medication dosing from suboptimal clinical examples: A deep reinforcement learning approach. *arXiv preprint*. https://arxiv.org/abs/1605.02665

Nemati, S., Ghassemi, M. M., & Clifford, G. D. (2016). Optimal medication dosing from observational data by reinforcement learning. In *Proceedings of the AAAI Conference on Artificial Intelligence*.

Al Omari, H. M. A., Kadir, M. R. A., Sapingi, R. B., & Al-Dalaien, A. H. M. A. D. (2024). The dynamics of CAATs Adoption in Jordan: Bridging top management support with auditor innovativeness and it competency. *WSEAS Transactions on Business and Economics*, *22*, 20-31.

Rajkomar, A., Dean, J., & Kohane, I. (2019). Machine learning in medicine. *New England Journal of Medicine, 380*(14), 1347–1358. https://doi.org/10.1056/NEJMra1814259

Shickel, B., Tighe, P. J., Bihorac, A., & Rashidi, P. (2018). Deep EHR: A survey of recent advances in deep learning techniques for electronic health record (EHR) analysis. *IEEE Journal of Biomedical and Health Informatics, 22*(5), 1589–1604. https://doi.org/10.1109/JBHI.2017.2767063

Weng, S. F., Reps, J., Kai, J., Garibaldi, J. M., & Qureshi, N. (2017). Can machine-learning improve cardiovascular risk prediction using routine clinical data? *PLoS One, 12*(4), e0174944. https://doi.org/10.1371/journal.pone.0174944

Zerrouki, T., Shquier, M. M. A., Balla, A., Bousbia, N., Sakraoui, I., & Boudardara, F. (2019). Adapting espeak to Arabic language: Converting Arabic text to speech language using espeak. *International Journal of Reasoning-Based Intelligent Systems*, *11*(1), 76-89.

Zhou, Y., Fang, Z., Zeng, D. D., & Zhang, C. (2018). Mining patient similarity based on longitudinal health records. *Journal of Biomedical Informatics, 83*, 147–159. https://doi.org/10.1016/j.jbi.2018.06.010

Zhou, Z., Li, X., & Siddiqi, A. (2018). Patient clustering for diabetes treatment based on electronic health records. *Journal of Biomedical Informatics, 84*, 42–49