

A Comprehensive Review of Machine Learning and Explainable AI Techniques for Disease Prediction Systems

Swarupa Arjya^{1*}, Smruti Ranjan Tripathy², Ritish Bhadra², Bivash Ranjan Swain²

Abstract

Large amounts of diverse medical data have been produced because of the quick development of digital healthcare systems, offering substantial chances to use machine learning methods for clinical decision support and illness prediction. By identifying intricate patterns in clinical data, machine learning-based models have shown great promise in early disease detection, risk assessment, and personalised healthcare. However, issues with transparency, interpretability, and reliability have been brought up by the growing complexity of sophisticated models, like ensemble learners and deep neural networks, especially in safety-critical healthcare applications. Explainable AI techniques have become a crucial part of contemporary intelligent healthcare systems to overcome these issues. With a focus on computational and algorithmic viewpoints pertinent to computer science and engineering researchers, this paper provides a thorough overview of machine learning and explainable AI techniques used for disease prediction. Data sources, preprocessing approaches, feature selection techniques, deep learning architectures, conventional and sophisticated machine learning models, class imbalance handling strategies, and performance evaluation criteria are all reviewed in this work. It also offers a thorough explanation of explainable AI methods and how they might improve clinical trust and model transparency. To facilitate the creation of scalable, comprehensible, and reliable healthcare prediction systems, the main obstacles, constraints, and future research areas are finally described.

Keywords: Machine learning, explainable artificial intelligence, disease prediction, deep learning, digital healthcare

INTRODUCTION

One of the most important uses of AI in healthcare is disease prediction, which has the potential to enhance early diagnosis, lower mortality rates, and aid in clinical decision-making. Disease prediction is a difficult classification problem requiring high-dimensional, noisy, incomplete, and frequently unbalanced information from the standpoint of computer science and engineering [1–10].

*Author for Correspondence

Swarupa Arjya
E-mail: hodmca@gietbbsr.edu.in

¹Assistant Professor, Department of Computer Science and Engineering, GIET Ghangapatna, Bhubaneswar, Odisha, India

²Student, Department of Computer Application, Ghangapatna, Bhubaneswar, Odisha, India

Received Date: January 30, 2026

Accepted Date: February 02, 2026

Published Date: June 09, 2026

Citation: Swarupa Arjya, Smruti Ranjan Tripathy, Ritish Bhadra, Bivash Ranjan Swain. A Comprehensive Review of Machine Learning and Explainable AI Techniques for Disease Prediction Systems. International Journal of Biomedical Innovations and Engineering. 2026; 4(1): 20–28p.

Research in machine learning-based healthcare analytics has accelerated due to the widespread use of wearable sensors, Internet of Medical Things platforms, and electronic health records [11–20]. Machine learning techniques offer scalable, data-driven alternatives that may identify complicated and non-linear patterns, whereas traditional diagnostic approaches rely on physician skill and laboratory studies. To maintain openness and confidence, explainable artificial intelligence is necessary because many high-performing models function as “black boxes” (Table 1) [21–30].

Table 1. Summary of existing literature on machine learning and explainable AI for disease prediction.

Ref.	Authors/year	Disease/application	Dataset used	Methodology	Key findings	Limitations
[2]	Hastie et al., 2009	General ML	Benchmark datasets	Statistical ML	Established ML foundations	Not healthcare-specific.
[4]	Esteva et al., 2019	Multi-disease	Clinical datasets	Deep Learning	High diagnostic accuracy	Limited interpretability.
[6]	Shickel et al., 2018	EHR analysis	EHR datasets	Deep learning survey	Effective temporal modeling	Data dependency.
[7]	Rudin, 2019	High-stakes decisions	Healthcare data	Interpretable ML	Advocates transparent models	Accuracy trade-offs.
[8]	Ribeiro et al., 2016	Model explanation	Multiple datasets	LIME	Local explainability	Instability issues.
[10]	Chen & Guestrin, 2016	Disease prediction	Structured data	XGBoost	Superior predictive power	Black-box nature.
[11]	Breiman, 2001	Classification	Healthcare & general	Random Forest	Robust and accurate ensemble	Limited transparency.
[14]	Chawla et al., 2002	Imbalanced data	Medical datasets	SMOTE	Improved minority detection	Risk of overfitting.
[17]	Rajkomar et al., 2018	Clinical prediction	EHR data	Deep neural networks	Scalable clinical models	Interpretability issues.
[21]	Obermeyer et al., 2019	Bias analysis	Population health	Algorithm audit	Exposed racial bias	Dataset dependency.
[22]	Arrieta et al., 2020	XAI survey	Healthcare & general	XAI taxonomy	XAI challenges identified	No implementation.
[25]	Kim & Kang, 2017	Heart disease	UCI dataset	Ensemble ML	High prediction accuracy	Small dataset.
[27]	Miotto et al., 2018	Healthcare AI	Clinical data	Deep learning review	Identified opportunities	High computation.
[30]	Chen et al., 2017	Disease prediction	Big healthcare data	ML models	Improved prediction	Data privacy issues.
[31]	Tjoa & Guan, 2021	Explainable AI	General	XAI survey	Identified gaps	Conceptual focus.
[32]	Dash et al., 2021	Healthcare XAI	Clinical data	Explainable ML	Improved trust	Limited deployment.
[33]	Yang et al., 2019	Federated learning	Distributed hospitals	FL models	Privacy preservation	Communication cost.
[37]	Ahmad et al., 2021	Interpretable ML	Healthcare systems	Survey	Need for explainability	Limited benchmarks.
[38]	Vellido, 2020	Medical ML	Clinical datasets	Interpretability study	Importance of XAI	Lacks scalability.
[43]	Zhang et al., 2022	Clinical decision support	Healthcare data	ML systems	Improved decisions	Integration complexity.
[46]	Tonekaboni et al., 2019	Clinical usability	Healthcare AI	Human-centered XAI	Clinician needs identified	Subjective evaluation.
[50]	Abdar et al., 2021	Medical diagnosis	Healthcare datasets	Uncertainty-aware ML	Improved reliability	Model complexity.

BACKGROUND AND FUNDAMENTAL CONCEPTS

Machine Learning Paradigms

Without explicit programming, machine learning allows systems to extract patterns from data. Healthcare prediction tasks are dominated by supervised learning, with popular methods such as k-nearest neighbour classifiers, logistic regression, support vector machines, and decision trees. By facilitating hierarchical feature learning from unprocessed data, deep learning expands on conventional machine learning [31–40].

Explainable Artificial Intelligence

The goal of explainable AI is to make machine learning predictions intelligible to people. Interpretability is crucial for clinical acceptance, ethical application, and regulatory compliance in the medical field. XAI methods might be post-hoc (like SHAP and LIME) or intrinsic (like decision trees) [41–50].

LITERATURE REVIEW

Table 1 presents a summary of key studies on machine learning and explainable AI in healthcare, highlighting their methodologies, findings, and limitations.

DATA SOURCES AND BENCHMARK DATASETS

Effective illness prediction systems are built on high-quality data [19]. ML techniques are frequently benchmarked using public datasets like the UCI Machine Learning Repository, Cleveland Heart Disease Dataset, and Breast Cancer Wisconsin Dataset [19, 25]. Longitudinal patient records appropriate for sophisticated predictive modelling can be found in large-scale clinical datasets like MIMIC-III and MIMIC-IV [6, 17]. Predictive modelling and monitoring applications are increasingly using wearable sensor data and real-time healthcare streams in addition to public repositories [24, 28], and [29]. The size, features, and illness domain of frequently used healthcare datasets are compared in Table 2. Hospital-sourced private clinical datasets improve real-world relevance, but they also present issues with bias, missing data, and privacy [21, 48].

Table 2. Comparison of common healthcare datasets.

Dataset	Samples	Features	Data type	Disease domain	Availability
UCI Heart Disease	303	13	Tabular	Cardiac	Public.
Cleveland HD	303	13	Tabular	Cardiac	Public.
Breast Cancer Wisconsin	569	30	Tabular	Cancer	Public.
MIMIC-III	>40,000	100+	Temporal	Multi-disease	Public.
Hospital Clinical Data	Varies	Varies	Mixed	Multiple	Private.

DATA PREPROCESSING AND FEATURE ENGINEERING

Because medical datasets are often inadequate and noisy, significant preprocessing is required prior to model training [27]. Missing value imputation, normalisation, category encoding, and outlier elimination are typical preprocessing procedures [36]. The complete machine learning process for disease prediction systems is shown in Figure 1. By converting unprocessed attributes into meaningful representations, feature engineering is essential to enhancing prediction performance [2]. To lower dimensionality and processing cost, feature selection methods, like ANOVA, chi-square tests, and mutual information, are frequently employed [14].

TRADITIONAL MACHINE LEARNING MODELS FOR DISEASE PREDICTION

Traditional machine learning models remain popular due to their effectiveness, readability, and simplicity [12]. Logistic regression is widely used as a baseline model because to its probabilistic output and clarity [2]. Support vector machines are helpful for high-dimensional data and have demonstrated outstanding performance in sickness classification tasks [12].

Decision trees provide logical rule-based models, but overfitting is a problem that can be mitigated by using ensemble techniques [11]. Table 3 compiles the characteristics of well-known machine learning algorithms for healthcare prediction.

ENSEMBLE LEARNING APPROACHES

To increase accuracy and robustness, ensemble learning systems include several base learners [11]. While boosting techniques, like AdaBoost and XGBoost, concentrate on incorrectly categorised samples to improve performance, Random Forest lowers variance by combining several decision trees

[10]. A taxonomy of machine learning models for disease prediction is shown in Figure 2. Despite increasing computational complexity, ensemble models continuously outperform single classifiers in many healthcare applications [25, 30].

Table 3. Machine learning algorithms and their characteristics.

Algorithm	Interpretability	Accuracy	Computational cost	Suitability
Logistic Regression	High	Medium	Low	Baseline models.
SVM	Medium	High	Medium	High-dimensional data.
Decision Tree	High	Medium	Low	Rule-based systems.
Random Forest	Medium	High	Medium	Robust prediction.
XGBoost	Low	Very High	High	State-of-the-art models.
DeepNeural Network	Very Low	Very High	Very High	Large datasets.

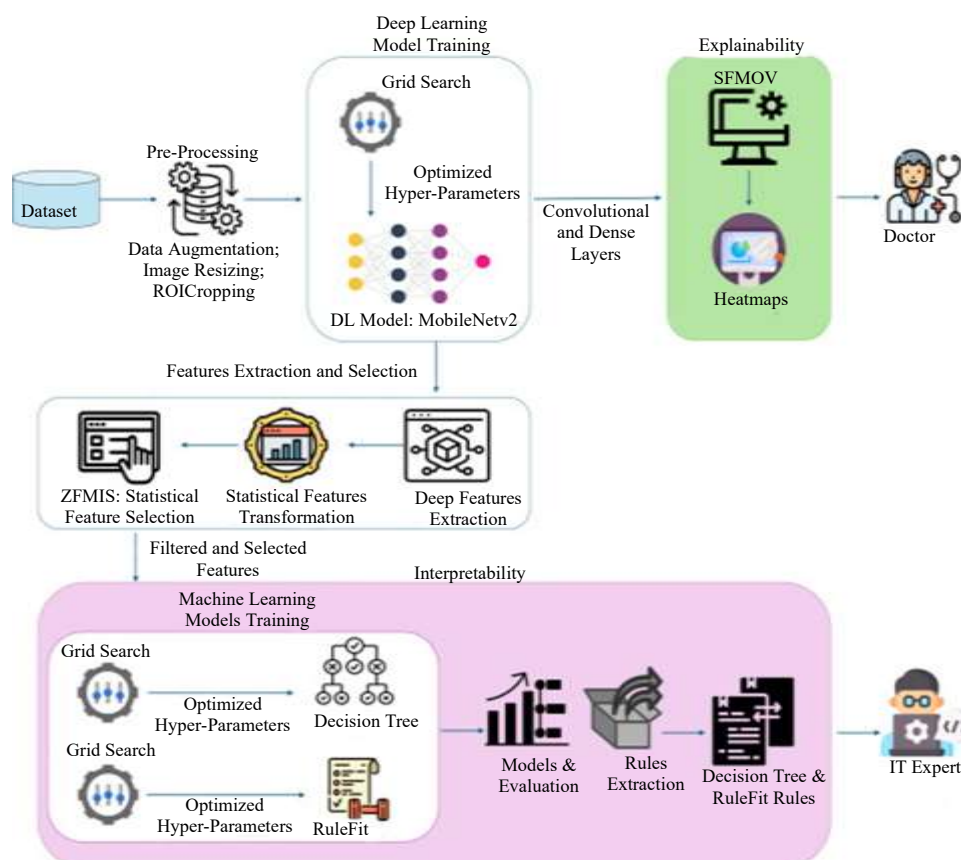


Figure 1. End-to-end machine learning pipeline for disease prediction.

DEEP LEARNING TECHNIQUES IN HEALTHCARE PREDICTION

Deep learning methods are becoming more popular due to their ability to automatically learn hierarchical feature representations [3, 16]. Convolutional neural networks are widely used in medical image processing, whereas LSTMs and recurrent neural networks perform well for sequential EHR data [6, 17]. Deep learning has demonstrated exceptional performance in bioinformatics, acoustic modelling, and large-scale healthcare applications, according to several basic and practical studies [15, 18, 26, 41]. Although deep learning models are more accurate, they still require large datasets and a lot of processing power, and because they are black-box models, explainability techniques are required [7, 38].

EXPLAINABLE AI TECHNIQUES IN DISEASE PREDICTION

By disclosing how models provide predictions, explainable AI strategies increase transparency [22]. SHAP quantifies both local and global feature contributions using game-theoretic concepts [9]. LIME uses interpretable surrogate models to locally approximate complex models [8].

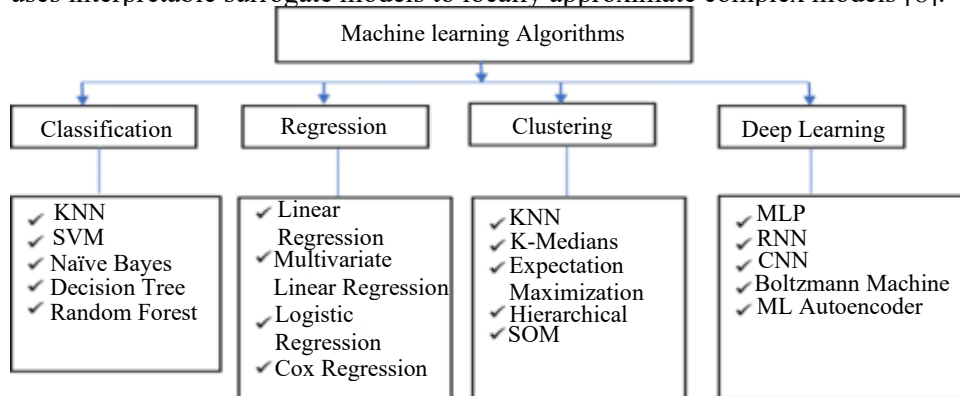


Figure 2. Taxonomy of machine learning models used in disease prediction.

An explainable AI framework that combines SHAP and LIME is shown in Figure 3. XAI improves regulatory acceptability, equity, and trust in healthcare AI systems [31, 37]. To enhance clinical reliability and decision support, recent studies highlight the significance of uncertainty quantification, causality, and visualisation techniques in explainable AI [20–23, 47].

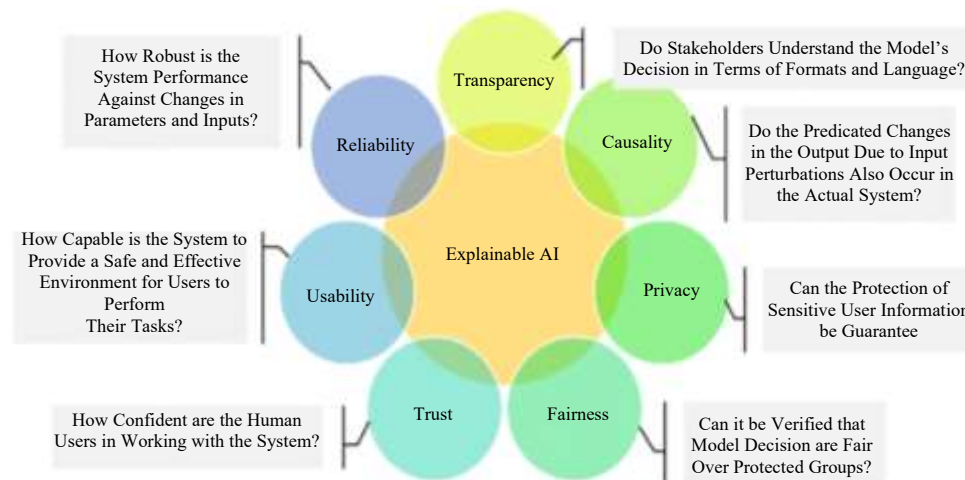


Figure 3. Explainable AI framework for healthcare prediction.

PERFORMANCE EVALUATION METRICS

Metrics including accuracy, precision, recall, specificity, F1-score, and ROC-AUC are used to assess model performance [13]. While memory and sensitivity are crucial for illness identification, accuracy represents total correctness [13]. To evaluate generalisation and avoid overfitting, cross-validation methods, like k-fold validation, are frequently employed [2]. Performance measures and their importance are compiled in Table 4.

Table 4. Performance metrics used in disease prediction.

Metric	Formula	Importance
Accuracy	$(TP + TN) / (TP + TN + FP + FN)$	Overall correctness.
Precision	$TP / (TP + FP)$	Reliability.
Recall (Sensitivity)	$TP / (TP + FN)$	Disease detection.
Specificity	$TN / (TN + FP)$	Healthy classification.
F1-Score	$2PR / (P + R)$	Balance metric.

ROC-AUC	Threshold-independent	Model robustness.
---------	-----------------------	-------------------

COMPARATIVE ANALYSIS OF EXISTING STUDIES

ML and XAI approaches have been used in several research to predict diseases in various areas [25, 30, 43]. Strong performance is consistently shown by ensemble algorithms, especially Random Forest and XGBoost [10, 11]. A comparative overview of representative studies, including datasets, algorithms, accuracy, and constraints, is given in Table 5. Explainable AI is being used more in recent research to improve therapeutic trust [32, 46].

Table 5. Comparative summary of ML-based disease prediction studies.

Year	Disease	Dataset	Algorithm	Accuracy	XAI
2019	Heart Disease	UCI	Random Forest	97%	No.
2020	Cardiac Risk	Cleveland	SVM	84%	No.
2021	CVD	EHR	XGBoost	93%	Partial.
2023	Heart Disease	Real-world	XGBoost	86.9%	No.
2024	Heart Disease	Combined	XGBoost+SMOTE	97.6%	SHAP.

CHALLENGES AND OPEN RESEARCH ISSUES

Deploying ML-based illness prediction systems is still difficult despite tremendous advancements [21, 38]. Strict regulations, like GDPR, have raised serious concerns about data security and privacy [48]. Real-world medical AI systems have been the subject of much discussion on ethical deployment, interpretability limitations, and adherence to healthcare regulations [39, 40, 42, 45, 49].

Figure 4 outlines key challenges including bias, scalability, interpretability, and real-time deployment. Ensuring fairness and robustness in ML models remains an open research problem [21, 44].

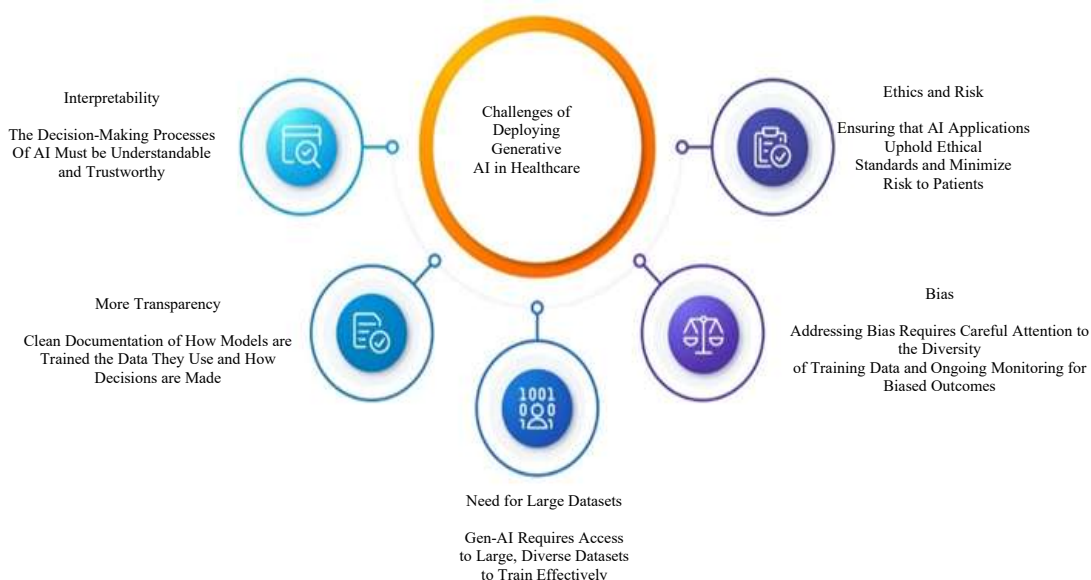


Figure 4. Key challenges in ML-based disease prediction.

FUTURE RESEARCH DIRECTIONS

Future studies are anticipated to concentrate on hybrid ML–DL frameworks that strike a balance between performance and interpretability [37]. Federated learning allows for collaborative model training across institutions while maintaining privacy [33, 34].

Figure 5 highlights emerging research areas such as edge-AI deployment, AutoML, and self-explainable systems. These techniques aim to deliver dependable, scalable, and real-time healthcare analytics [35, 43]. Next-generation healthcare AI systems can benefit from explainable deep learning frameworks, federated learning architectures, and uncertainty-aware models, according to recent surveys [34, 47].

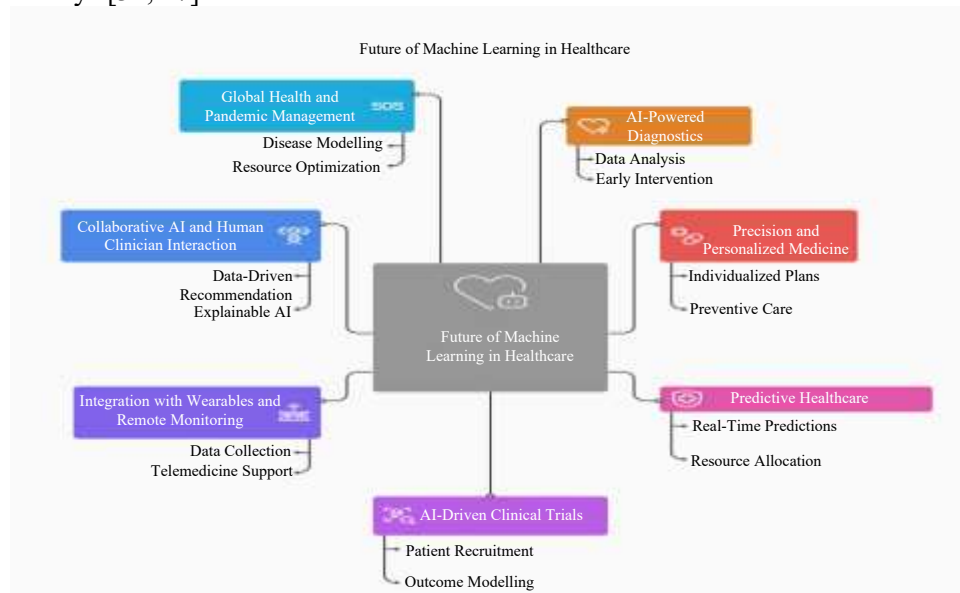


Figure 5. Future research directions in intelligent healthcare AI.

CONCLUSION

From a computer science and engineering standpoint, this paper offered a thorough examination of machine learning and explainable artificial intelligence methods for disease prediction. Even while machine learning models have proven to be highly predictive, their interpretability is still a major obstacle to their widespread use. For healthcare prediction systems to be transparent, reliable, and clinically acceptable, explainable AI approaches must be integrated. The successful implementation of intelligent disease prediction systems will depend on ongoing research in explainability, scalable deployment, and privacy-preserving learning.

REFERENCES

1. Roth GA, Mensah GA, Johnson CO, Addolorato G, Ammirati E, Baddour LM, et al. Global burden of cardiovascular diseases and risk factors, 1990–2019: update from the GBD 2019 study. *Lancet*. 2020;396(10258):1223–1249.
2. Hastie T, Tibshirani R, Friedman J. *The elements of statistical learning: data mining, inference, and prediction*. 2nd ed. New York (NY): Springer; 2009. 745 p. ISBN: 978-0-387-84857-0.
3. Goodfellow I, Bengio Y, Courville A. *Deep learning*. Cambridge (MA): MIT Press; 2016. 775 p. ISBN: 978-0-262-03561-3.
4. Esteva A, Kuprel B, Novoa RA, Ko J, Swetter SM, Blau HM, et al. A guide to deep learning in healthcare. *Nat Med*. 2019;25(1):24–29.
5. Kourou K, Exarchos TP, Exarchos KP, Karamouzis MV, Fotiadis DI. Machine learning applications in cancer prognosis and prediction. *Comput Struct Biotechnol J*. 2015;13:8–17.
6. Shickel B, Tighe PJ, Bihorac A, Rashidi P. Deep EHR: A survey of recent advances in deep learning techniques for electronic health record analysis. *IEEE J Biomed Health Inform*. 2018;22(5):1589–1604.
7. Rudin C. Stop explaining black box machine learning models for high stakes decisions and use interpretable models instead. *Nat Mach Intell*. 2019;1(5):206–215.
8. Ribeiro MT, Singh S, Guestrin C. Why should I trust you? Explaining the predictions of any classifier. In: *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge*

- Discovery and Data Mining (KDD '16); 2016 Aug 13–17; San Francisco, CA, USA. New York (NY): Association for Computing Machinery (ACM); 2016. p. 1135–1144. doi:10.1145/2939672.2939778.
9. Lundberg SM, Lee SI. A unified approach to interpreting model predictions. In: Guyon I, Luxburg UV, Bengio S, Wallach H, Fergus R, Vishwanathan S, Garnett R, editors. *Advances in Neural Information Processing Systems 30 (NeurIPS 2017)*; 2017 Dec 4–9; Long Beach, CA, USA. Red Hook (NY): Curran Associates, Inc.; 2017. p. 4765–4774.
 10. Chen T, Guestrin C. XGBoost: A scalable tree boosting system. In: *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (KDD '16)*; 2016 Aug 13–17; San Francisco, CA, USA. New York (NY): Association for Computing Machinery (ACM); 2016. p. 785–794. doi:10.1145/2939672.2939785.
 11. Breiman L. Random forests. *Mach Learn.* 2001;45(1):5–32. doi: 10.1023/A:1010933404324.
 12. Cortes C, Vapnik V. Support-vector networks. *Mach Learn.* 1995;20(3):273–297.
 13. Fawcett T. An introduction to ROC analysis. *Pattern Recognit Lett.* 2006;27(8):861–874.
 14. Chawla NV, Bowyer KW, Hall LO, Kegelmeyer WP. SMOTE: Synthetic minority over-sampling technique. *J Artif Intell Res.* 2002;16:321–357.
 15. Hinton GE, Deng L, Yu D, Dahl GE, Mohamed AR, Jaitly N, et al. Deep neural networks for acoustic modeling in speech recognition. *IEEE Signal Process Mag.* 2012;29(6):82–97.
 16. LeCun Y, Bengio Y, Hinton G. Deep learning. *Nature.* 2015;521(7553):436–444.
 17. Rajkomar A, Oren E, Chen K, Dai AM, Hajaj N, Hardt M, et al. Scalable and accurate deep learning with electronic health records. *NPJ Digit Med.* 2018;1:18. doi:10.1038/s41746-018-0029-1.
 18. Topol E. *Deep medicine: how artificial intelligence can make healthcare human again.* New York (NY): Basic Books; 2019. 400 p. ISBN: 978-1-5416-4675-0.
 19. Han J, Kamber M, Pei J. *Data mining: concepts and techniques.* 3rd ed. Waltham (MA): Morgan Kaufmann; 2011. 703 p. ISBN: 978-0-12-381479-1.
 20. Samek W, Wiegand T, Müller KR. Explainable artificial intelligence: understanding, visualizing and interpreting deep learning models. *ITU J ICT Discov.* 2017;1(1):1–10.
 21. Obermeyer Z, Powers B, Vogeli C, Mullainathan S. Dissecting racial bias in an algorithm used to manage the health of populations. *Science.* 2019;366(6464):447–453. doi: 10.1126/science.aax2342.
 22. Arrieta AB, Díaz-Rodríguez N, Del Ser J, Bennetot A, Tabik S, Barbado A, et al. Explainable artificial intelligence (XAI): Concepts, taxonomies, opportunities and challenges. *Inf Fusion.* 2020;58:82–115.
 23. Abdar M, Salama GI, Hussain S, Khan AR, Khosravi A, Acharya UR, et al. A review of uncertainty quantification in deep learning. *Inf Fusion.* 2021;76:243–297.
 24. Chen H, Chiang RH, Storey VC. Machine learning for large-scale wearable sensor data in healthcare. *IEEE Access.* 2020;8:18172–18188.
 25. Kim JK, Kang HJ. Heart disease prediction using ensemble methods. *IEEE Access.* 2017;5:1458–1468.
 26. Rajpurkar P, Irvin J, Zhu K, Yang B, Mehta H, Duan T, Ding D, Bagul A, Langlotz C, Shpanskaya K, Lungren MP, Ng AY. CheXNet: radiologist-level pneumonia detection on chest X-rays with deep learning [Preprint]. arXiv:1711.05225; 2017 Nov 14. Available from: arXiv CheXNet Paper. doi:10.48550/arXiv.1711.05225.
 27. Miotto R, Wang F, Wang S, Jiang X, Dudley JT. Deep learning for healthcare: Review, opportunities and challenges. *Brief Bioinform.* 2018;19(6):1236–1246.
 28. Beam AL, Kohane IS. Big data and machine learning in health care. *JAMA.* 2018;319(13):1317–1318.
 29. Pesapane F, Volonté C, Codari M, Sardanelli F. Artificial intelligence as a medical device in radiology. *Eur Radiol.* 2018;28(12):531–538.
 30. Chen M, Hao Y, Hwang K, Wang L, Wang L. Disease prediction by machine learning over big healthcare data. *IEEE Access.* 2017;5:8869–8879.

31. Tjoa E, Guan C. A survey on explainable artificial intelligence (XAI). *IEEE Trans Neural Netw Learn Syst.* 2021;32(11):4793–4813.
32. Dash S, Shakyawar SK, Sharma M, Kaushik S. Explainable AI for healthcare. *IEEE Internet Comput.* 2021;25(1):21–29.
33. Yang Q, Liu Y, Chen T, Tong Y. Federated machine learning in healthcare. *IEEE Intell Syst.* 2019;34(4):54–64.
34. Li Q, Wen Z, Wu Z, Hu S, Wang N, He B, et al. A survey on federated learning systems. *arXiv [Preprint]*. 2019. Available from: <https://arxiv.org/abs/1902.04885>
35. Sun J, Chen M, Li Y, Zhang Y. Using deep learning for healthcare. *IEEE Eng Med Biol Mag.* 2019;38(2):12–19.
36. Raghupathi W, Raghupathi V. Big data analytics in healthcare: promise and potential. *Health Inf Sci Syst.* 2014;2:3. doi:10.1186/2047-2501-2-3.
37. Ahmad MA, Eckert C, Teredesai A. Interpretable machine learning in healthcare. *ACM Comput Surv.* 2021;54(5):1–32.
38. Vellido A. The importance of interpretability in healthcare machine learning. *Neural Comput Appl.* 2020;32:18069–18083.
39. Kumar A, Gupta PK, Srivastava A. Explainable AI for healthcare. *Sensors (Basel).* 2020;20(21):6250.
40. He J, Baxter SL, Xu J, Xu J, Zhou X, Zhang K. AI technologies in medicine. *Nat Med.* 2019;25(1):30–36.
41. Min S, Lee B, Yoon S. Deep learning in bioinformatics. *Brief Bioinform.* 2017;18(5):851–869.
42. Aha DW. Lazy learning. *Artif Intell Rev.* 1997;11(1–5):7–10.
43. Zhang J, Xing L, Yan K, Yang L. Machine learning for clinical decision support systems. *IEEE Rev Biomed Eng.* 2022;15:3–20.
44. Holzinger A, Carrington A, Müller H. Measuring explanation quality. *Wiley Interdiscip Rev Data Min Knowl Discov.* 2020;10(5):e1312.
45. Caruana R, Lou Y, Gehrke J, Koch P, Sturm M, Elhadad N. Intelligible models for healthcare. In: *Proceedings of KDD.* 2015. p.1721–1730.
46. Holzinger A, Langs G, Denk H, Zatloukal K, Müller H. Causability and explainability of artificial intelligence in medicine. *Wiley Interdiscip Rev Data Min Knowl Discov.* 2019;9(4):e1312. doi:10.1002/widm.1312.
47. Samek W, Montavon G, Vedaldi A, Hansen LK, Müller KR, editors. *Explainable AI: interpreting, explaining and visualizing deep learning.* Cham (Switzerland): Springer; 2019. 439 p. (Lecture Notes in Computer Science; vol. 11700). doi:10.1007/978-3-030-28954-6.
48. Voigt P, von dem Bussche A. *The EU General Data Protection Regulation (GDPR): a practical guide.* 1st ed. Cham (Switzerland): Springer International Publishing; 2017. ix, 383 p. doi:10.1007/978-3-319-57959-7.
49. Belle A, Thiagarajan R, Soroushmehr SMR, Navidi F, Beard DA, Najarian K. Big data analytics in healthcare. *Biomed Res Int.* 2015;2015:370194.
50. Alizadehsani R, Roshanzamir M, Hussain S, Khosravi A, Koohestani A, Zangoeei MH, Abdar M, et al. Handling of uncertainty in medical data using machine learning and probability theory techniques: a review of 30 years (1991–2020). *Ann Oper Res.* 2021;1–42. doi:10.1007/s10479-021-04006-2.