Online-Journals.org

iJOE International Journal of Online and Biomedical Engineering

iJOE | eISSN: 2626-8493 | Vol. 21 No. 6 (2025) | OPEN ACCESS

https://doi.org/10.3991/ijoe.v21i06.54211

PAPER

A Systematic Literature Review on Machine Learning in Healthcare Prediction

Nur Farah Afifah Ahmad Sukri[1](✉), Wan Mohd Amir Fazamin Wan Hamzah[1,2], Mohd Kamir Yusof[1], Ismahafezi Ismail[1], Harmy Mohamed Yusoff[3], Azliza Yacob[4]

[1]Faculty of Informatics and Computing, Universiti Sultan Zainal Abidin, Terengganu, Malaysia

[2]Artificial Intelligence Research Centre for Islam and Sustainability (AIRIS), Universiti Sultan Zainal Abidin, Terengganu, Malaysia

[3]Faculty of Medicine, Universiti Sultan Zainal Abidin, Terengganu, Malaysia

[4]University College TATI, Terengganu, Malaysia

sl4722@putra.unisza.edu.my

ABSTRACT

Rapid technological advancement will continue to create new values and transform experiences in many sectors, including healthcare. Several key trends are shaping today's healthcare system, including the use of machine learning (ML). This systematic literature review (SLR) explores the application of ML in healthcare, particularly in predictive analytics. The SLR also includes a few papers on machine learning operations (MLOps) in healthcare, reflecting limited studies on the topic. This suggests significant potential for further exploration in MLOps. The review compares findings from various studies, many of which agree that ML enhances the scalability and reliability of predictive models. This study aims to assess the most effective ML algorithms and methodologies used in healthcare prediction. It also attempts to identify features influencing the outcomes of ML applications in healthcare predictions. Findings suggest that ML can improve prediction accuracy using the appropriate dataset, optimal feature selection model, and a tailored ML algorithm for specific tasks. The literature highlights challenges, including the need for specialised skills and the complexity of integrating MLOps into existing healthcare systems.

KEYWORDS

machine learning (ML), machine learning operations (MLOps), machine learning models, machine learning algorithms, healthcare prediction, healthcare

1 INTRODUCTION

Technological innovation has taken place in every aspect of today's life [1]. In healthcare, both providers and patients are significantly impacted by technological advancements. Given the current global population, healthcare systems are encountering substantial challenges [2], which influenced the rise of chronic diseases and the demand for more efficient, personalised, and proactive healthcare solutions. In the old healthcare system, treatment is provided only when symptoms appear, along with general treatment plans. This approach is incompetent in today's

Ahmad Sukri, N.F.A., Wan Hamzah, W.M.A.F., Yusof, M.K., Ismail, I., Yusoff, H.M., Yacob, A. (2025). A Systematic Literature Review on Machine Learning in Healthcare Prediction. *International Journal of Online and Biomedical Engineering (iJOE)*, 21(6), pp. 155–177. https://doi.org/10.3991/ijoe.v21i06.54211

Article submitted 2025-01-01. Revision uploaded 2025-02-10. Final acceptance 2025-02-25.

complex environment that requires more personalised care. Traditional statistical methods used to predict outcomes often struggle to consider the complex interactions among various clinical, demographic, and molecular factors impacting the progression of disease [3].

To overcome these problems, the healthcare system is undergoing a significant transformation, driven by the integration of new-wave technologies, such as machine learning (ML), artificial intelligence (AI), and big data analytics [4]. Specifically, ML is changing various aspects of healthcare—from predictive analytics and diagnostic accuracy to treatment optimisation and patient monitoring [4]. One significant advantage of ML in healthcare is its potential to improve diagnostic accuracy and facilitate early disease detection and prediction [5]. For example, ML algorithms can analyse medical imaging data, including X-rays and MRIs, to detect anomalies that may show signs of cancer or neurological disorders [6]. This ability not only supports early diagnosis but also enables the development of personalised treatment strategies for each patient's unique requirements, thereby improving overall care quality [7].

However, current healthcare prediction systems face substantial challenges such as data privacy and security, data diversity and generalisation, model personalisation, scalability and efficiency [8], [10]. ML models' complexity requires substantial training and deployment resources [12], [13], and frequent updates are necessary to maintain accuracy as new data becomes available. The integration of machine learning operations (MLOps) presents a promising solution to these challenges. A fundamental goal of implementing MLOps in healthcare is to enhance the use of ML in clinical workflows [9].

According to [11], MLOps ensures efficient implementation of ML solutions while adhering to high standards of data management and compliance. This approach accelerates predictive analytics deployment and promotes continuous improvement through real-time monitoring [14]. A key element of MLOps is the implementation of continuous integration (CI) and continuous deployment (CD) practices. CI/CD pipelines facilitate automated testing, building, and deployment of ML models, reducing manual intervention and minimising error risks. This agile approach is essential in fast-paced clinical environments [11]. Despite these advancements, there remains limited literature regarding the effective integration of MLOps with ML in healthcare prediction models.

This systematic literature review (SLR) aims to explore ML applications within the healthcare sector, focusing on prediction models' effectiveness. By analysing existing studies, this review will highlight how MLOps can enhance ML model accuracy while identifying commonly used algorithms and methodologies in healthcare. It will also find the most influential features of ML models. Furthermore, it will examine key challenges and opportunities in deploying MLOps in healthcare settings, providing insights into leveraging these technologies to improve global health outcomes.

2 RESEARCH METHODOLOGY

This systematic review provides a thorough overview of the application of ML algorithms and methods in healthcare prediction. To conduct the review, this paper adhered to the guidelines and procedures recommended by [15]. One approach was to follow the preferred reporting items for systematic reviews and meta-analyses

(PRISMA) protocol [16], [17]. The methodology used in conducting the review was divided into three stages:

- Planning the review
- Conducting the review
- Reporting the review

2.1 Planning the review

In this initial stage, it is essential to define the research objectives and establish keywords for paper selection, as well as to formulate inclusion and exclusion criteria. Since this paper aims to identify the application of ML in healthcare prediction, the keywords used for paper selection included healthcare prediction, healthcare, ML and ML operations.

The objectives of this study are:

1. To summarise data and features used in ML in healthcare prediction.
2. To evaluate common methods used by researchers in healthcare prediction.
3. To identify the ML algorithm used in healthcare prediction.

Once the research objective is clearly defined, the next step is to construct the research question [18]. The research questions of this study were mapped from the objectives stated above. Referring to guidelines by [15], three aspects were considered as the study criteria: population, intervention and outcomes of the study. Here are the specifics for each study criteria:

- Population: high school, higher educational institute and industrial sector
- Intervention: methods, algorithms and prediction techniques.
- Outcome: best features for prediction and successful prediction techniques or approaches.

The criteria outlined above then directed the formulation of the research questions (RQs) for this study:

- RQ1: What features are most commonly utilised by researchers in the healthcare sector for ML applications?
- RQ2: How do ML methods used by researchers in healthcare prediction influence the evaluation of model accuracy and effectiveness?
- RQ3: Which algorithms are considered the most effective in ML models within the healthcare sector?

2.2 Conducting the review

This stage is where a screening process is performed to identify the most relevant papers to the research objectives. This study utilised the PRISMA guidelines [19] to conduct the screening process, which was divided into three phases.

Search strategy. Papers were collected from four different databases: IEEEXplore, ScienceDirect, SpringerLink and SAGE Publication, focusing on articles published between the years 2020 and 2024. To obtain the most relevant dataset, a search strategy using related keywords was used. Papers related to healthcare prediction, healthcare, ML and MLOps keywords were selected before going through the next screening process. Table 1 summarises the databases used in this study and the number of papers collected from each database.

Table 1. Databases used for research findings

No.	Database Name	No. of Papers
1	IEEEXplore	52
2	ScienceDirect	61
3	SpringerLink	63
4	SAGE Publication	41

Selection criteria. Selection criteria are a guideline to ensure each selected paper is the best fit for a study. It is to guarantee that the selected articles clearly define their methods and findings, can be consistently replicated by other researchers, and do not rely on preconceived ideas or biases before the research is conducted [12, 13]. The criteria included were as follows:

- Studies that used ML, ML algorithms, ML techniques, or MLOps in healthcare prediction.
- Articles of real-world application of ML in clinical settings for prediction models.
- Studies of different ML models in healthcare prediction.

Papers selected based on the selection criteria then underwent inclusion and exclusion criteria to narrow down the findings.

Inclusion and exclusion criteria. The inclusion and exclusion criteria were then defined to decide whether or not a particular study should be included in this systematic review [14].

Inclusion criteria

1. Studies on healthcare prediction
2. Studies written in English
3. Studies published in the last five years (2020–2024)

Exclusion criteria

1. Studies that did not provide clear methodologies or results related to ML applications in the healthcare sector
2. Irrelevant titles, abstracts and keywords
3. Review papers

In the view of [15], extraction flow is very helpful in selecting papers. According to PRISMA guidelines [11], the first step is to identify papers equivalent to our research objectives. In this study, papers related to healthcare prediction using ML were searched. As stated in Table 1, 52 papers were retrieved from IEEEXplore, 61 papers from ScienceDirect, 63 papers from SpringerLink and 41 papers from the SAGE publication database, making a total of 217 papers to be used in this review study. Seventy-two papers were excluded as they did not meet the established inclusion criteria (review papers).

In the second step, all papers were screened based on their title, abstracts, and full content. In this phase, 58 papers were removed as they had irrelevant titles and abstracts. Then, another 26 papers were removed due to not having enough explanation of ML algorithms, methods, and features used in their studies. From 217 papers, 61 papers were selected. Figure 1 shows the PRISMA flowchart of this selection process.

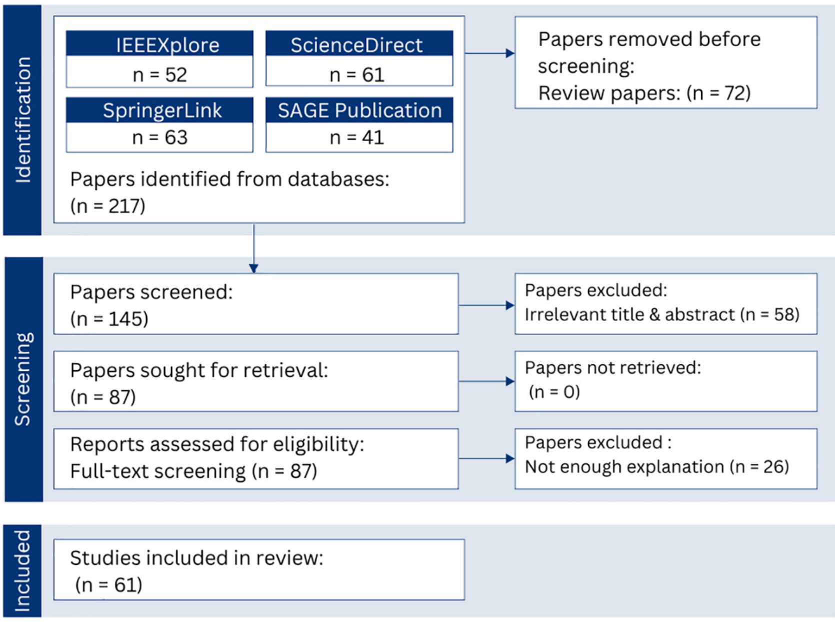

Fig. 1. PRISMA flowchart of papers identification for review

2.3 Reporting the review

This section provides tabulated data of all papers extracted from the 2.2 phase. Table 2 displays the paper ID and author/references for each of the included papers.

Table 2. Paper ID and author list

Paper Id	Author/Reference	Paper Id	Author/Reference	Paper Id	Author/Reference
P1	[20]	P22	[41]	P43	[62]
P2	[21]	P23	[42]	P44	[63]
P3	[22]	P24	[43]	P45	[64]
P4	[23]	P25	[44]	P46	[65]
P5	[24]	P26	[45]	P47	[66]
P6	[25]	P27	[46]	P48	[67]
P7	[26]	P28	[47]	P49	[68]
P8	[27]	P29	[48]	P50	[69]
P9	[28]	P30	[49]	P51	[70]
P10	[29]	P31	[50]	P52	[71]
P11	[30]	P32	[51]	P53	[72]
P12	[31]	P33	[52]	P54	[73]

(Continued)

Table 2. Paper ID and author list *(Continued)*

Paper Id	Author/Reference	Paper Id	Author/Reference	Paper Id	Author/Reference
P13	[32]	P34	[53]	P55	[74]
P14	[33]	P35	[54]	P56	[75]
P15	[34]	P36	[55]	P57	[76]
P16	[35]	P37	[56]	P58	[77]
P17	[36]	P38	[57]	P59	[78]
P18	[37]	P39	[58]	P60	[79]
P19	[38]	P40	[59]	P61	[80]
P20	[39]	P41	[60]		
P21	[40]	P42	[61]		

3 RESULTS

In this section, a comprehensive analysis of features, methods and algorithms of ML and MLOps in healthcare prediction is provided. Each subsection answers a specific research question in detail.

3.1 RQ1: What features are most commonly utilised by researchers in the healthcare sector for ML applications?

Many researchers agree that the selection and quality of features within a dataset play a vital role in enhancing the accuracy and reliability of ML models [81]. Proper feature selection boosts model performance and improves interpretability and generalisation across various datasets and applications.

Different researchers have mentioned different features that are best for their studies. From 61 papers reviewed, nine feature categories were detected to be commonly used by researchers. They were demographic information, clinical data, laboratory test results, symptom descriptions, previous and current treatment data, behavioural factors, electronic health records (EHR) and image data.

Table 3 below shows the general description of each feature category, highlighting the types of attributes included in each feature category. Some attributes expressed the same meaning but were labelled differently. Not all nine features were addressed by every paper; some papers only mentioned one, two, three, four, or five of the features.

Table 3. Feature categories description

Feature Id	Feature	Description
F1	Demographic Information	Age, gender, ethnicity and socioeconomic status of patients
F2	Clinical Data	Medical history, family history of diseases, current medications and vital signs
F3	Laboratory Test Results	Blood test results, imaging results (e.g., X-rays, MRIs) and genetic markers
F4	Symptom Descriptions	Patient-reported symptoms and their severity

(Continued)

Table 3. Feature categories description *(Continued)*

Feature Id	Feature	Description
F5	Treatment Data	Previous treatment outcomes and response to medications
F6	Behavioural Factors	Lifestyle choices such as smoking, alcohol consumption and physical activity levels
F7	Environmental Factors	Exposure to toxins or pollutants and living conditions
F8	Electronic Health Records (EHR)	Comprehensive patient records including notes from healthcare providers
F9	Imaging Features	Features extracted from medical imaging data using techniques like radiomics

According to Figure 2, clinical data emerged as a primary feature utilised by many researchers to enhance the performance accuracy of their ML models. [41] found that their system is showing better performance when integrated with clinical data. Table 4 shows a result tabulated by [45] showing the clinical dataset achieves the highest accuracy compared to other datasets. Additionally, a study by [82] emphasised the importance of various clinical features in predicting how patients will recover after surgery. Clinical data has the highest recurrence percentage at 68.85%, followed closely by demographic information at 85.25%.

The laboratory test results showed a significant impact with only 39.34%. Treatment and imaging features each accounted for 11.48% and 13.11%, respectively, while EHR appeared in just 8.20% of the reviewed studies. Symptom descriptions, behavioural factors, and environmental factors have the lowest recurrence rates, each at only 3.28% in this systematic literature review.

Despite the focus of all reviewed papers on the healthcare sector, many do not recognise symptom descriptions, behavioural factors and environmental factors as critical features to include in their datasets. This oversight may stem from the fact that many patients experience symptoms that healthcare providers fail to recognise or document. Furthermore, behavioural and environmental factors are often intertwined with numerous variables, including biological, psychological and social influences. This complexity can make it challenging for researchers to effectively isolate and analyse these factors within their studies.

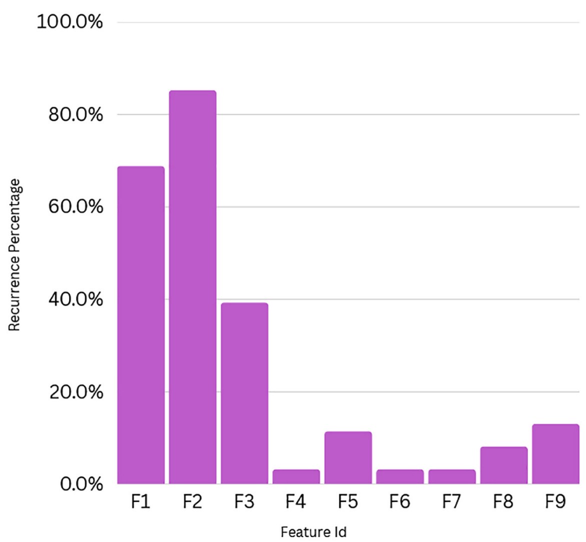

Fig. 2. Recurrence percentage of commonly used feature categories in healthcare prediction

Table 4. Example of dataset evaluation

Detecting Heart Diseases - Dataset Evaluation.

Dataset	Accuracy (%)
Proposed regular medical monitoring and Electronic Clinical Data (ECD)	99.4
Framingham Heart Study Dataset	89.7
Cleveland Heart Disease Dataset	91.6
MIMIC-III Dataset	93.2

[83] emphasised the importance of feature combinations in enhancing the performance of prediction in ML algorithms. Specifically, they highlighted that "the multiple features reveal underlying patterns in the data, which is essential for accurate predictions."

Figure 3 illustrates that over 18% of studies opted to combine multiple features in their datasets. Specifically, 39.3% of the studies utilised two features, 31.1% employed three features, 9.8% incorporated four features and 1.6% of the studies combined five features in their datasets.

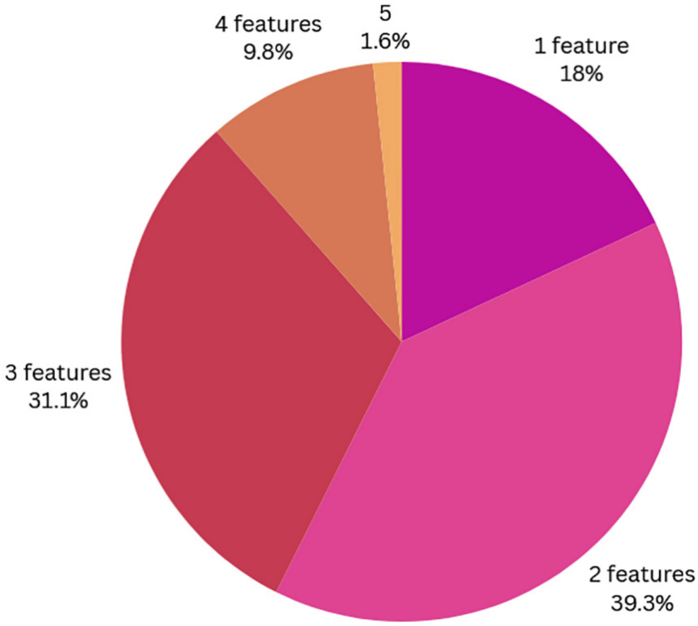

Fig. 3. Percentage of studies that utilised one or more features

Table 5 below records paper IDs for each feature category. 52 papers used clinical data in their study. Data like medical history, family history of diseases, current medications, and vital signs are certainly influential in shaping the ML model. Next, 42 papers indicate that demographic information is vital data for training the ML model. A study highlighted that many ML models failed to report sufficient demographic data, which led to biased outcomes and limited the applicability of the models across different populations [83]. Out of 61 studies, 24 studies utilised laboratory test results to help elevate the predictive models. Table 4 also indicates that treatment data and imaging feature data were prevalent in the studies reviewed. Seven studies utilised treatment datasets, while another eight focused on imaging features. Treatment data allows for personalised models that predict how patients respond to different treatments. This is especially crucial in fields like psychiatry, where individual responses to medications can vary [81]. On the other hand, [84] found that imaging features provide a more detailed understanding of medical conditions, like tumour characteristics, improving prediction accuracy. Additionally, EHR datasets, which contain comprehensive patient information including notes from healthcare providers, were identified as a crucial predictor element in healthcare prediction models. Five studies reviewed highlighted EHR data as an essential feature in this context. Although there were only several studies that concentrated on the symptom descriptions dataset, the behavioural factors dataset, and the environmental factors dataset, each of these datasets was utilised by two papers, leading to a total of six distinct papers involved.

Table 5. Paper ID used by each feature category

Feature Id	Paper Id
F1	P2, P5, P8, P9, P11, P13, P14, P15, P19, P20, P21, P22, P23, P25, P27, P31, P33, P34, P35, P36, P37, P38, P39, P40, P41, P42, P43, P44, P45, P46, P47, P48, P49, P50, P51, P52, P53, P55, P56, P57, P58, P61
F2	P2, P3, P4, P5, P6, P7, P8, P9, P10, P11, P12, P13, P14, P15, P18, P19, P20, P21, P22, P23, P25, P26, P27, P30, P31, P32, P33, P34, P35, P36, P37, P38, P39, P40, P41, P42, P43, P44, P45, P46, P47, P48, P49, P50, P51, P52, P54, P55, P56, P57, P58, P60
F3	P3, P15, P16, P18, P20, P21, P22, P24, P25, P27, P28, P31, P33, P35, P36, P39, P47, P48, P49, P51, P55, P56, P57, P58
F4	P33, P55
F5	P27, P30, P39, P41, P44, P45, P57
F6	P53, P58
F7	P34, P42
F8	P12, P23, P26, P29, P57
F9	P1, P4, P7, P14, P17, P22, P54, P59

3.2 RQ2: How do ML methods used by researchers in healthcare prediction influence the evaluation of model accuracy and effectiveness?

There are two widely used ML methods in predictive modelling: classification and regression [85]. Table 6 tabulated data on methodologies used by each study. Both approaches are essential for making predictions based on labelled datasets but serve different purposes and are applied to different types of problems.

Classification is used when the output is categorical, indicating that it represents distinct classes or labels [86]. For instance, it can categorise data into "Yes" or "No," or into multiple classes like blood type A, B, AB, or O. Classification is selected when the goal is to predict a category based on input features, making it suitable for tasks like image recognition, medical diagnosis and sentiment analysis.

On the other hand, regression is commonly applied when the output involves continuous data that can take any real value. It focuses on predicting numerical outcomes rather than categories. Common applications of regression include forecasting values such as house prices, temperature prediction, or sales figures. Regression is particularly beneficial when estimating a value based on input variables, making it ideal for tasks like financial forecasting and demand prediction [86].

Figure 4 shows the percentage of each methodology employed in this study. As shown, classification was utilised in 85.2% of the studies, followed by regression tasks at 8.2%. Additionally, 6.6% of the studies employed both classification and regression tasks. Most studies have used the classification method as they intend to get the result in categorical data, such as classifying patients as high risk or low risk in cancer [66], yes or no classification of relapse status in breast cancer patients [49], and temperature classification of the eye region [26]. To ensure the effectiveness of an ML model, it is essential to first consider the choice of ML methods to achieve the best and most accurate results. While most papers reviewed in this SLR have selected classification as their method, this does not imply that classification is the best choice for all prediction models.

Table 6. Methodologies used by each study

Method Used	Paper Id
Classification	P1, P2, P3, P4, P5, P7, P8, P9, P10, P11, P12, P14, P15, P16, P17, P18, P19, P20, P21, P22, P23, P24, P25, P26, P28, P29, P30, P31, P32, P33, P34, P35, P36, P37, P38, P39, P40, P41, P42, P43, P44, P45, P47, P49, P50, P51, P53, P54, P55, P56, P58, P59
Regression	P6, P13, P48, P52, P61
Both	P27, P46, P57, P60

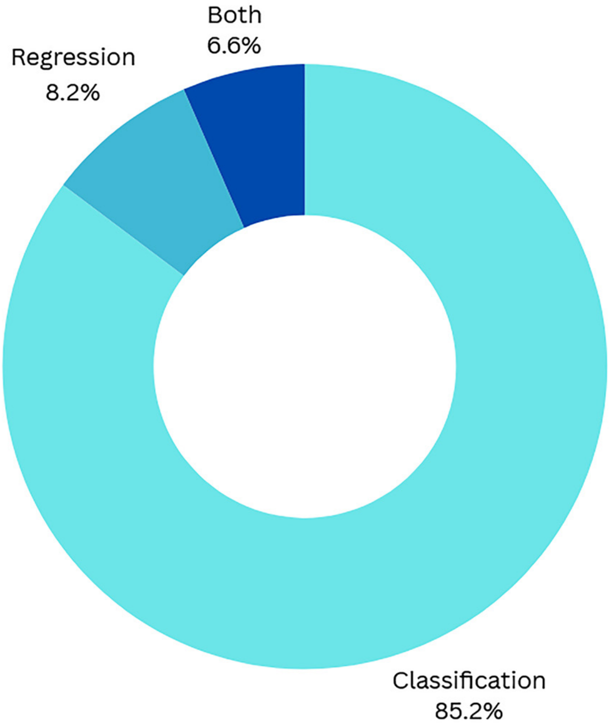

Fig. 4. Percentage of methodologies commonly used in healthcare prediction

3.3 RQ3: Which algorithms are considered the most effective in ML models within the healthcare sector?

The choice of algorithm in ML plays a critical role in determining the performance and accuracy of prediction models. Different algorithms are designed to handle various types of data and relationships. For instance, algorithms that excel in processing linear relationships may struggle with non-linear data, potentially compromising prediction accuracy [87]. Different algorithms also depend on different tasks.

Table 7 provides an overview of algorithm performance within this study. The data demonstrates a strong preference among researchers for utilising hybrid algorithms in their models. The table presents performance metrics (precision, recall, AUC, F1 score and accuracy) for various ML algorithms evaluated in this study. However, due to different metrics reported across studies, there is a considerable amount of missing data, and it limits the ability to make direct comparisons.

Algorithms included in this study are hybrid algorithms, random forest (RF), gradient boosting (GB), neural network (NN), decision tree (DT), support vector machine (SVM), naïve bayes (NB), long short-term memory (LSTM), K-nearest neighbour (KNN), ada boost (AB) and logistic regression. Extreme gradient boosting (XGB), light gradient boosting (LGB) and hist gradient boosting (HGB) are categorized under the GBM category. Artificial neural networks (ANN) and convolutional neural networks (CNN) are categorised under the neural network category.

Table 7. Algorithm performance details

Bil.	Algorithm	Paper Id	Precision	Recall	AUC	F1-Score	Accuracy
1	Hybrid	P3	82	85	90	83	83.1
2		P5	–	–	98.5	–	96.9
3		P6	94.29	91.7	–	93	95.6
4		P14	98	–	–	98	97
5		P16	85.65	85.7	–	–	87.4
6		P20	–	–	91	–	–
7		P22	95.31	85.7	–	90	94
8		P23	99	98.9	–	99.3	99.4
9		P25	–	71.3	80.7	48.6	74.4
10		P26	95.9	96	–	96	95.2
11		P27	–	–	85.2	–	–
12		P28	–	–	98.9	98.4	98.9
13		P37	–	–	–	–	97.2
14		P49	99.23	99.1	–	–	99.1
15		P55	–	–	81	–	71
16		P59	81.7	78.3	86.1	78.4	79.3
17		P60	–	–	81.7	–	–

(Continued)

Table 7. Algorithm performance details *(Continued)*

Bil.	Algorithm		Paper Id	Precision	Recall	AUC	F1-Score	Accuracy
18	RF		P1	96.49	–	99.2	–	97.4
19			P2	–	–	87.2	44.4	95.4
20			P8	97	97.6	–	97	97.9
21			P11	–	–	92	83	81
22			P12	–	–	90.2	–	91.3
23			P15	–	–	92	–	88
24			P18	–	–	79.2	–	93.7
25			P31	–	–	72.2	–	97.9
26			P33	–	–	96.2	–	87.3
27			P38	98.61	97.9	99.9	98.5	98.9
28			P41	–	–	88.5	–	86.8
29			P45	–	–	–	89.2	94.8
30			P56	–	–	–	–	91.5
31			P57	96.6	96.4	–	–	96.4
32	GB	GBM	P21	–	–	–	–	95.1
33			P50	–	–	90.2	–	–
34			P61	98.65	100	–	99.3	99.1
35			P51	81.6	95.2	95.2	87.9	93.2
36		HGB	P42	99.9	–	–	99.8	99.8
37		LGB	P10	–	–	85.5	79.5	80
38		XGB	P29	94.13	91	–	92.4	96.9
39			P40	–	–	97	–	87.8
40			P58	–	–	–	–	82.5
41			P61	–	–	–	–	89.9
42	NN	ANN	P13	–	–	88.5	–	98.2
43			P43	88	90	94	89	89
44			P47	100	100	–	100	100
45			P52	76.2	91.2	88.6	73.3	89.5
46		CNN	P9	83.47	–	–	82.3	82.3
47		GNN	P30	97	99	98	98	98
48		NN	P24	94	94	–	90	94
49	DT		P7	94.44	94.4	–	94.5	96.1
50			P19	97.5	–	98	97.5	98
51			P44	92.31	–	97.5	96	95.7
52			P46	93.53	91.5	–	92.4	99

(Continued)

Table 7. Algorithm performance details *(Continued)*

Bil.	Algorithm	Paper Id	Precision	Recall	AUC	F1-Score	Accuracy
53	SVM	P17	95.27	94.6	–	94.9	94.7
54		P36	–	–	–	–	81.7
55		P54	–	–	–	60.9	82.5
56	NB	P35	98	98	–	98	98.1
57		P53	14.8	14.8	86.9	11.4	83.4
58	LSTM	P4	–	–	98.7	92.3	97.1
59	KNN	P32	–	–	81.5	–	72.8
60	AB	P39	88.26	88.5	–	88.4	94.4
61	LR	P48	–	–	–	–	77.3

This study has concluded that RF algorithms are the most effective single approach for making predictions in the healthcare sector. [20] agreed that RF is very powerful due to its robust feature selection process, effective handling of complex datasets, and superior performance compared to other algorithms. This statement is supported by [21], who stated that RF is a highly effective algorithm within their clinical decision support system, achieving high accuracy in disease prediction tasks. Fourteen studies concluded RF performs with the highest accuracy compared to other algorithms in their study.

Decision tree is structured like a flowchart, where each internal node represents a feature (or attribute), each branch signifies a decision rule, and each leaf node indicates the outcome. [26] mentioned that the flowchart-like structure allows users to easily understand how decisions are made, which is particularly valuable in fields like healthcare and finance where transparency is crucial. According to Table 6, most papers that acknowledged DT as their best performance algorithm used clinical data as their feature selection dataset.

Neural network mimics the way biological neurone's function using weighted connections between nodes. Each node receives inputs, processes them using an activation function and produces an output to be sent to the next layer. This architecture allows neural networks to learn complex patterns in data [43]. [43] also highlighted various applications of neural networks across different fields, demonstrating their versatility in tasks, such as image recognition, natural language processing and predictive analytics.

Gradient boosting is an ensemble machine-learning technique that builds models in a stage-wise fashion. It combines the predictions of several weak learners, typically DTs, to create a strong predictive model [53]. XGBoost was recognised as the most popular algorithm chosen by researchers within the gradient-boosting category.

Support vector machine is a powerful supervised ML algorithm primarily used for classification and regression tasks. It excels in binary classification problems, where the goal is to separate data points into two distinct classes. [55] highlighted that SVM is particularly effective for classification tasks, especially in scenarios with high-dimensional data. This capability makes it suitable for medical applications where numerous features may be present [55].

Naïve bayes is based on Bayes' theorem, which calculates the posterior probability of a class given the features. It assumes that the features are conditionally independent given the class label, which simplifies computation [72]. [72] also pointed out that NB performs competitively against other algorithms, particularly in scenarios where the independence assumption holds or when dealing with high-dimensional data.

Hybrid algorithms have been reported to achieve the highest accuracy in 17 studies. A hybrid algorithm is defined as one that combines different methods to effectively tackle challenging problems. This approach aims to leverage the strengths of various algorithms, improving solution quality, convergence speed, and robustness [45]. [74] highlighted various applications of the hybrid algorithm, showcasing its adaptability in solving real-world optimisation problems across different fields. P28 shows the real application of a hybrid algorithm in their cancer diagnosis model called En-MinWhale that combines SVM, DT, multilayer perceptron, and RF as the base learners to make initial predictions [47]. Figure 5 shows the workflow of En-inWhale.

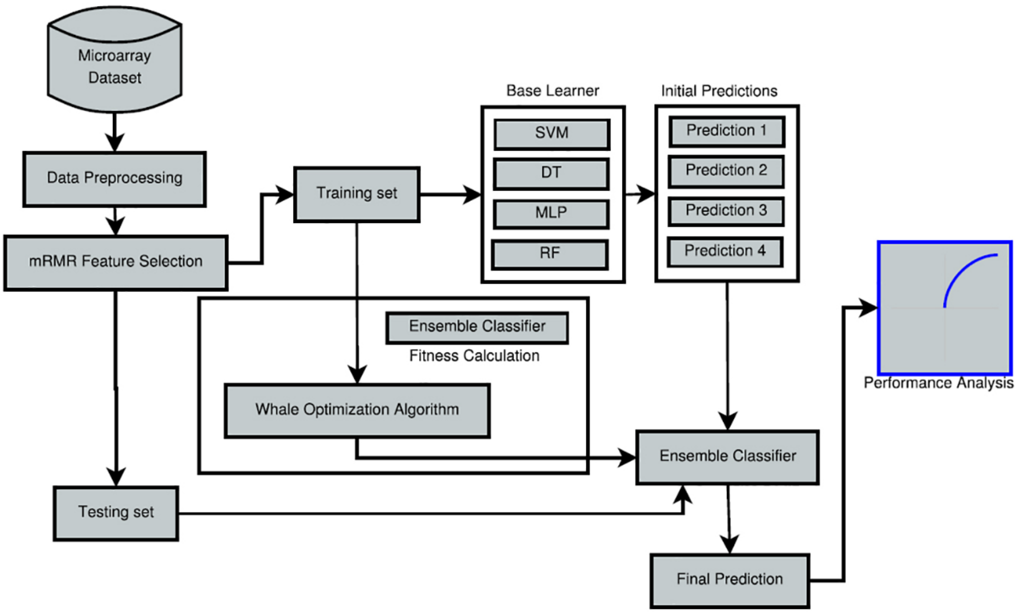

Fig. 5. Example of real-world application of hybrid algorithm

Table 8 summarises the performance value of ML algorithms across different metrics. The averages recorded are significantly affected by missing values. RF exhibits high precision (97.18%) and recall (97.29%), indicating a strong balance between minimising false positives and false negatives. However, its AUC is lower than the other categories, and its F1 score is substantially lower than its precision and recall values. DT demonstrates high accuracy (97.2%), but its low AUC significantly affects the results. NN can also be categorised as a high-performing algorithm based on its high recall and accuracy averages. GB and hybrid algorithms are concluded as potentially strong performers. Despite limited data, it displays generally high average scores across precision, recall, AUC, and F1score. Conversely, NB and SVM fall under low-performing categories. It gives low precision and recall values compared to other algorithms. LSTM, KNN, AdaBoost and LR each provide only a single data point and are not representative of their performance. It is important to highlight that all the values are highly dependent on the specific dataset and experimental setup used in each study.

Table 8. Algorithm performance summary (average values)

Algorithm	Precision (Avg)	Recall (Avg)	AUC (Avg)	F1-Score (Avg)	Accuracy (Avg)
Hybrid	92.34	87.96	88.12	87.18	90.61
RF	**97.18**	**97.29**	89.66	82.42	92.73
GB	93.57	95.4	91.9825	91.78	91.59
NN	89.78	94.84	92.27	88.76	92.99
DT	94.445	92.945	65.18	**95.1**	**97.2**
SVM	95.27	94.63	0	77.9	86.29
NB	56.4	56.4	86.85	54.7	90.735
LSTM	–	–	**98.7**	92.3	97.1
KNN	–	–	81.5	–	72.8
AdaBoost	88.26	88.49	–	88.39	94.38
LR	–	–	–	–	77.33

4 DISCUSSION

Machine learning has fundamentally improved the landscape of healthcare prediction by enabling more accurate assessments and clinical decisions. Conventionally, prediction models relied on statistical methods to predict outcomes. However, these techniques often struggle to capture the complex interactions of vast clinical data. This paper showcases a five-year systematic review of the application of ML in predictive models within the healthcare sector. The analysis of 61 papers included in this review emphasises the most popular and effective features and algorithms that influence the performance of these predictive models.

In healthcare predictive models, nine features were commonly utilised, including demographic information, clinical data, laboratory test results, symptom descriptions, treatment data, behavioural factors, environmental factors, EHR, and imaging features. Clinical data stood out as the most commonly used feature. Furthermore, many researchers were found to combine multiple features in their datasets to improve the effectiveness of their models. The use of suitable datasets and optimal feature selection can also significantly enhance the performance of prediction models.

The evaluation of common methods employed by researchers in healthcare prediction revealed classification as a more effective approach for running predictive models compared to the regression approach. Nevertheless, some researchers opted to combine classification and regression methods in their models.

The findings of this study identified the most effective algorithms for healthcare prediction models. Hybrid algorithms, RF, GB, NN and DT emerged as popular choices among researchers. These algorithms showed the highest accuracy in their respective study. Specifically, hybrid algorithms have been reported to achieve the highest accuracy in 17 studies. This studies has also included the value of performance metrics (precision, recall, AUC, F1 score, accuracy) of various ML algorithms. However, the variability in reported metrics across different studies resulted in a considerable amount of missing data, thereby limiting the ability to make direct comparisons. Furthermore, it is important to highlight that all the values are highly dependent on the specific dataset and experimental setup used by each study. Performance is highly dependent on the specific datasets, pre-processing techniques, and experimental setups.

Although ML has significantly improved healthcare prediction models, its complexity requires MLOps integration. Nevertheless, there are also limitations regarding appropriate timing, interoperability challenges and the need for skilled personnel.

In conclusion, the implementation of ML into healthcare prediction has significantly enhanced the accuracy and effectiveness of clinical decision-making. By leveraging a combination of features and advanced algorithms, researchers can develop robust predictive models that better capture the complexities of clinical data. The findings from this review underscore the importance of using appropriate methodologies, particularly classification techniques and hybrid algorithms, to optimise predictive performance in healthcare settings. As ML continues to evolve, its application in healthcare will likely lead to even greater advancements in patient care and outcomes. Nonetheless, the integration of these technologies also presents challenges, such as ensuring data privacy, maintaining regulatory compliance and addressing the ethical implications of AI-driven decision-making. Therefore, future research in ML for healthcare should focus on carefully managing model deployment and exploring approaches such as blockchain technology to enhance security.

5 LIMITATIONS

This study has certain limitations that should be acknowledged. Firstly, most of the reviewed papers raised concerns regarding imbalanced datasets, highlighting the need for more appropriate datasets and noting how both excessively large and insufficient datasets could bias the results. Secondly, conducting a comprehensive comparison between models is challenging, as many papers do not provide essential information, such as pre-processing steps or the attributes of dataset features. Lastly, there were only a limited number of papers available on MLOps studies. Therefore, this study cannot focus solely on MLOps; it must also incorporate ML to expand the dataset and increase the number of studies included.

6 ACKNOWLEDGEMENT

This study was funded by the Ministry of Higher Education Malaysia (MOHE) through the Fundamental Research Grant Scheme (FRGS) with project reference code: FRGS/1/2023/ICT06/UNISZA/02/1. A special thanks to the Centre for Research Excellence & Incubation Management (CREIM) of Universiti Sultan Zainal Abidin (UniSZA).

7 CONFLICT OF INTEREST

The authors declare no conflict of interest.

8 REFERENCES

[1] R. Batra and S. Batra, "Review paper on technological innovation management," *Palarch's Journal of Archaeology of Egypt/Egyptology*, vol. 17, no. 9, 2020. https://archives.palarch.nl/index.php/jae/article/download/4656/4604/8916

[2] C. Hong, L. Sun, G. Liu, B. Guan, C. Li, and Y. Luo, "Response of global health towards the challenges presented by population aging," *China CDC Weekly*, vol. 5, no. 39, pp. 884–887, 2023. https://doi.org/10.46234/ccdcw2023.168

[3] C. Boeri *et al.*, "Machine learning techniques in breast cancer prognosis prediction: A primary evaluation," *Cancer Medicine,* vol. 9, no. 9, pp. 3234–3243, 2020. https://doi.org/10.1002/cam4.2811

[4] N. Mehta, A. Pandit, and S. Shukla, "Transforming healthcare with big data analytics and artificial intelligence: A systematic mapping study," *Journal of Biomedical Informatics*, vol. 100, p. 103311, 2019. https://doi.org/10.1016/j.jbi.2019.103311

[5] Rishabh Software, "Machine learning use cases in healthcare and key benefits," 2022. Retrieved from https://www.rishabhsoft.com/blog/machine-learning-in-healthcare-use-cases-benefits

[6] M. Razzak, S. Naz, and A. Zaib, "Deep learning for medical image processing: Overview, challenges and future," *arXiv preprint arXiv.1704.06825*, 2017. https://doi.org/10.48550/arXiv.1704.06825

[7] M. Javaid, A. Haleem, R. P. Singh, R. Suman, and S. Rab, "Significance of machine learning in healthcare: Features, pillars and applications," *International Journal of Intelligent Networks*, vol. 3, pp. 58–73, 2022. https://doi.org/10.1016/j.ijin.2022.05.002

[8] F. S. Alreshidi *et al.*, "Fed-CL- an atrial fibrillation prediction system using ECG signals employing federated learning mechanism," *Scientific Reports,* vol. 14, 2024. https://doi.org/10.1038/s41598-024-71366-7

[9] N. Sreeja Vidya Sai Venkata, "AI model deployment in healthcare: MLOps innovations and challenges," *International Journal of Computer Science and Information Technology Research*, vol. 6, pp. 2025–2031, 2025.

[10] H. Rafik, A. Maizate, and A. Ettaoufik, "Data security mechanisms, approaches, and challenges for e-Health smart systems," *International Journal of Online and Biomedical Engineering (iJOE),* vol. 19, no. 2, pp. 42–66, 2023. https://doi.org/10.3991/ijoe.v19i02.37069

[11] A. Basile, F. Calefato, F. Lanubile, G. Mallardi, and L. Quaranta, "An MLOps solution framework for transitioning machine learning models into eHealth systems," in *Proceedings of the 4th National Conference on Artificial Intelligence (Ital-IA 2024)*, 2024. https://ital-ia2024.it/submission/524/paper

[12] C. Janiesch, P. Zschech, and K. Heinrich, "Machine learning and deep learning," *arXiv preprint arXiv.2104.05314*, 2021. https://doi.org/10.48550/arXiv.2104.05314

[13] E. Bisong, "Building machine learning and deep learning models on Google Cloud Platform," *A Comprehensive Guide for Beginners*, 2019. https://doi.org/10.1007/978-1-4842-4470-8

[14] A. Bodor, M. Hnida, and N. Daoudi, "Machine learning models monitoring in MLOps context: Metrics and tools," *International Journal of Interactive Mobile Technologies (iJIM),* vol. 17, no. 23, pp. 125–139, 2023. https://doi.org/10.3991/ijim.v17i23.43479

[15] B. Kitchenham, O. P. Brereton, D. Budgen, M. Turner, J. Bailey, and S. Linkman, "Systematic literature reviews in software engineering: A systematic literature review," *Information and Software Technology,* vol. 51, no. 1, pp. 7–15, 2009. https://doi.org/10.1016/j.infsof.2008.09.009

[16] D. Moher, D. G. Altman, A. Liberati, and J. Tetzlaff, "PRISMA statement," *Epidemiology,* vol. 22, no. 1, p. 128, 2011. https://doi.org/10.1097/EDE.0b013e3181fe7825

[17] M. L. Rethlefsen *et al.,* "PRISMA-S: An extension to the PRISMA statement for reporting literature searches in systematic reviews," *Systematic Reviews*, vol. 10, 2021. https://doi.org/10.1186/s13643-020-01542-z

[18] H. Nawang, M. Makhtar, and W. M. A. F. Wan Hamzah, "A systematic literature review on student performance predictions," *International Journal of Advanced Technology and Engineering Exploration,* vol. 8, no. 84, 2021. https://doi.org/10.19101/IJATEE.2021.874521

[19] B. Hutton, F. Catala-Lopez, and D. Moher, "The PRISMA statement extension for systematic reviews incorporating network meta-analysis: PRISMA-NMA La extensión de la declaración PRISMA para revisiones sistemáticas que incorporan metaanálisis en red: PRISMA-NMA," *Med. Clin. (English Edition),* vol. 147, no. 6, pp. 262–266, 2016. https://doi.org/10.1016/j.medcle.2016.10.003
[20] Z. Huang and D. Chen, "A breast cancer diagnosis method based on VIM feature selection and hierarchical clustering random forest algorithm," *IEEE Access*, vol. 10, pp. 3284–3293, 2021. https://doi.org/10.1109/ACCESS.2021.3139595
[21] K. N. Singh and J. K. Mantri, "A clinical decision support system using rough set theory and machine learning for disease prediction," *Intelligent Medicine*, vol. 4, no. 3, pp. 200–208, 2024. https://doi.org/10.1016/j.imed.2023.08.002
[22] V. V. Khanna *et al.*, "A decision support system for osteoporosis risk prediction using machine learning and explainable artificial intelligence," *Heliyon,* vol. 9, no. 12, p. e22456, 2023. https://doi.org/10.1016/j.heliyon.2023.e22456
[23] M. Fadel, N. Elseddeq, R. Arnous, Z. Ali, and A. Eldesouky, "A fast accurate deep learning framework for prediction of all cancer types," *IEEE Access*, vol. 10, pp. 122586–122600, 2022. https://doi.org/10.1109/ACCESS.2022.3222365
[24] R. Ghorbani, R. Ghousi, A. Makui, and A. Atashi, "A new hybrid predictive model to predict the early mortality risk in intensive care units on a highly imbalanced dataset," *IEEE Access*, vol. 8, pp. 141066–141079, 2020. https://doi.org/10.1109/ACCESS.2020.3013320
[25] N. O. El-Ganainy, I. Balasingham, P. S. Halvorsen, and L. A. Rosseland, "A new real-time clinical decision support system using machine learning for critical care units," *IEEE Access,* vol. 8, pp. 185676–185687, 2020. https://doi.org/10.1109/ACCESS.2020.3030031
[26] M. A. Jaddoa *et al.*, "An approach-based machine learning and automated thermal images to predict the dark-cutting incidence in cattle management of healthcare supply chain," *Engineering Applications of Artificial Intelligence,* vol. 135, p. 108804, 2024. https://doi.org/10.1016/j.engappai.2024.108804
[27] V. Chang, M. A. Ganatra, K. Hall, L. Golightly, and Q. A. Xu, "An assessment of machine learning models and algorithms for early prediction and diagnosis of diabetes using health indicators," *Healthcare Analytics*, vol. 2, p. 100118, 2022. https://doi.org/10.1016/j.health.2022.100118
[28] K. K. Patro *et al.*, "An effective correlation-based data modeling framework for automatic diabetes prediction using machine and deep learning techniques," *BMC Bioinformatics*, vol. 24, 2023. https://doi.org/10.1186/s12859-023-05488-6
[29] D. Dash and M. Kumar, "An ensemble-based stage-prediction machine learning approach for classifying fetal disease," *Healthcare Analytics*, vol. 5, p. 100322, 2024. https://doi.org/10.1016/j.health.2024.100322
[30] F. Mesquita and G. Marques, "An explainable machine learning approach for automated medical decision support of heart disease," *Data and Knowledge Engineering*, vol. 153, p. 102339, 2024. https://doi.org/10.1016/j.datak.2024.102339
[31] I. Nissar *et al.*, "An intelligent healthcare system for automated diabetes diagnosis and prediction using machine learning," *Procedia Computer Science*, vol. 235, pp. 2476–2485, 2024. https://doi.org/10.1016/j.procs.2024.04.233
[32] T. Cui *et al.*, "Applying machine learning techniques in nomogram prediction and analysis for SMILE treatment," *American Journal of Ophthalmology*, vol. 210, pp. 71–77, 2020. https://doi.org/10.1016/j.ajo.2019.10.015
[33] Y. Wang *et al.,* "Automated multimodal machine learning for esophageal variceal bleeding prediction based on endoscopy and structured data," *Journal of Digital Imaging*, vol. 36, pp. 326–338, 2023. https://doi.org/10.1007/s10278-022-00724-6

[34] Q. Li, A. Campan, A. Ren, and W. E. Eid, "Automating and improving cardiovascular disease prediction using machine learning and EMR data features from a regional healthcare system," *International Journal of Medical Informatics*, vol. 163, 2022. https://doi.org/10.1016/j.ijmedinf.2022.104786
[35] S. Zhou, C. Hu, S. Wei, and X. Yan, "Breast cancer prediction based on multiple machine learning algorithms," *Technology in Cancer Research and Treatment*, vol. 23, 2024. https://doi.org/10.1177/15330338241234791
[36] R. Martínez-Licort *et al.*, "Breast carcinoma prediction through integration of machine learning models," *IEEE Access*, vol. 12, pp. 134635–134650, 2024. https://doi.org/10.1109/ACCESS.2024.3431998
[37] D. A. Debal and T. M. Sitote, "Chronic kidney disease prediction using machine learning techniques," *Journal of Big Data*, vol. 9, 2022. https://doi.org/10.1186/s40537-022-00657-5
[38] R. O. Alabi, A. A. Mäkitie, M. Pirinen, M. Elmusrati, I. Leivo, and A. Almangush, "Comparison of nomogram with machine learning techniques for prediction of overall survival in patients with tongue cancer," *International Journal of Medical Informatics*, vol. 145, 2021. https://doi.org/10.1016/j.ijmedinf.2020.104313
[39] F. Liu, J. Yao, C. Liu, and S. Shou, "Construction and validation of machine learning models for sepsis prediction in patients with acute pancreatitis," *BMC Surgery*, vol. 23, 2023. https://doi.org/10.1186/s12893-023-02151-y
[40] W. Wei, S. Tongping, and W. Jiaming, "Construction of a clinical prediction model for complicated appendicitis based on machine learning techniques," *Scientific Reports*, vol. 14, 2024. https://doi.org/10.1038/s41598-024-67453-4
[41] E. K. Jadoon, F. G. Khan, S. Shah, A. Khan, and M. Elaffendi, "Deep learning-based multi-modal ensemble classification approach for human breast cancer prognosis," *IEEE Access*, vol. 11, pp. 85760–85769, 2023. https://doi.org/10.1109/ACCESS.2023.3304242
[42] A. Sheik Abdullah, V. Naga Pranava Shashank, and D. Altrin Lloyd Hudson, "Disseminating the risk factors with enhancement in precision medicine using comparative machine learning models for healthcare data," *IEEE Access*, vol. 12, pp. 72794–72812, 2024. https://doi.org/10.1109/ACCESS.2024.3400023
[43] I. Zenbout, A. Bouramoul, S. Meshoul, and M. Amrane, "Efficient bioinspired feature selection and machine learning-based framework using omics data and biological knowledge databases in cancer clinical endpoint prediction," *IEEE Access*, vol. 11, pp. 2674–2699, 2023. https://doi.org/10.1109/ACCESS.2023.3234294
[44] M. M. Rahman *et al.*, "Empowering early detection: A web-based machine learning approach for PCOS prediction," *Informatics in Medicine Unlocked*, vol. 47, p. 101500, 2024. https://doi.org/10.1016/j.imu.2024.101500
[45] S. A. Alzakari *et al.*, "Enhanced heart disease prediction in remote healthcare monitoring using IoT-enabled cloud-based XGBoost and Bi-LSTM," *Alexandria Engineering Journal*, vol. 105, pp. 280–291, 2024. https://doi.org/10.1016/j.aej.2024.06.036
[46] X. Liu, H. Niu, and J. Peng, "Enhancing predictions with a stacking ensemble model for ICU mortality risk in patients with sepsis-associated encephalopathy," *Journal of International Medical Research*, vol. 52, no. 3, 2024. https://doi.org/10.1177/03000605241239013
[47] A. Panigrahi *et al.*, "En-MinWhale: An ensemble approach based on MRMR and whale optimization for cancer diagnosis," *IEEE Access*, vol. 11, pp. 113526–113542, 2023. https://doi.org/10.1109/ACCESS.2023.3318261
[48] C. Felix *et al.*, "Explainable machine learning for predicting conversion to neurological disease: Results from 52,939 medical records," *Digital Health*, vol. 10, 2024. https://doi.org/10.1177/20552076241249286

[49] A. Basaad, S. Basurra, E. Vakaj, M. A. Aleskandarany, and M. M. Abdelsamea, "GraphX-Net: A graph neural network-based Shapley values for predicting breast cancer occurrence," *IEEE Access*, vol. 12, pp. 93993–94007, 2024. https://doi.org/10.1109/ACCESS.2024.3424526

[50] A. Hennebelle, H. Materwala, and L. Ismail, "HealthEdge: A machine learning-based smart healthcare framework for prediction of type 2 diabetes in an integrated IoT, edge, and cloud computing system," *Procedia Computer Science,* 2023, vol. 220, pp. 331–338. https://doi.org/10.1016/j.procs.2023.03.043

[51] N. Divyashree and K. S. Nandini Prasad, "Improved clinical diagnosis using predictive analytics," *IEEE Access,* vol. 10, pp. 75158–75175, 2022. https://doi.org/10.1109/ACCESS.2022.3190416

[52] M. Cox *et al*., "Interpretable machine learning for the prediction of amputation risk following lower extremity infrainguinal endovascular interventions for peripheral arterial disease," *Cardiovascular and Interventional Radiology,* vol. 45, pp. 633–640, 2022. https://doi.org/10.1007/s00270-022-03111-4

[53] Z. B. Tadese *et al*., "Interpretable prediction of acute respiratory infection disease among under-five children in Ethiopia using ensemble machine learning and Shapley additive explanations (SHAP)," *Digital Health,* vol. 10, 2024. https://doi.org/10.1177/20552076241272739

[54] Q. Zhang *et al*., "Machine learning in risk prediction of continuous renal replacement therapy after coronary artery bypass grafting surgery in patients," *Clinical and Experimental Nephrology,* vol. 28, pp. 811–821, 2024. https://doi.org/10.1007/s10157-024-02472-z

[55] Y. Su *et al*., "Machine learning model for prediction of permanent stoma after anterior resection of rectal cancer: A multicenter study," *European Journal of Surgical Oncology,* vol. 50, no. 7, 2024. https://doi.org/10.1016/j.ejso.2024.108386

[56] Y. Qu, Z. Lin, Z. Yang, H. Lin, X. Huang, and L. Gu, "Machine learning models for prognosis prediction in endodontic microsurgery," *Journal of Dentistry*, vol. 118, 2022. https://doi.org/10.1016/j.jdent.2022.103947

[57] R. M. Eickhoff *et al*., "Machine learning prediction model for postoperative outcome after perforated appendicitis," *Langenbeck's Archives of Surgery*, vol. 407, pp. 789–795, 2022. https://doi.org/10.1007/s00423-022-02456-1

[58] D. Zuo, L. Yang, Y. Jin, H. Qi, Y. Liu, and L. Ren, "Machine learning-based models for the prediction of breast cancer recurrence risk," *BMC Medical Informatics and Decision Making*, vol. 23, 2023. https://doi.org/10.1186/s12911-023-02377-z

[59] D. Y. Lee *et al*., "Machine learning-based prediction model for postoperative delirium in non-cardiac surgery," *BMC Psychiatry*, vol. 23, 2023. https://doi.org/10.1186/s12888-023-04768-y

[60] A. Erdogan Yildirim and M. Canayaz, "Machine learning-based prediction of length of stay (LoS) in the neonatal intensive care unit using ensemble methods," *Neural Computing and Applications*, vol. 36, pp. 14433–14448, 2024. https://doi.org/10.1007/s00521-024-09831-7

[61] N. H. Kim *et al*., "Machine-learning model for predicting depression in second-hand smokers in cross-sectional data using the Korea national health and nutrition examination survey," *Digital Health*, vol. 10, 2024. https://doi.org/10.1177/20552076241257046

[62] R. Ali, J. Hussain, and S. W. Lee, "Multilayer perceptron-based self-care early prediction of children with disabilities," *Digital Health*, vol. 9, 2023. https://doi.org/10.1177/20552076231184054

[63] I. Gomez-Rios, E. Egea-Lopez, and A. J. Ortiz Ruiz, "ORIENTATE: Automated machine learning classifiers for oral health prediction and research," *BMC Oral Health*, vol. 23, 2023. https://doi.org/10.1186/s12903-023-03112-w

[64] T. Yang *et al.*, "Postoperative delirium prediction after cardiac surgery using machine learning models," *Computers in Biology and Medicine,* vol. 169, 2024. https://doi.org/10.1016/j.compbiomed.2023.107818

[65] S. Bani Hani and M. Ahmad, "Predicting mortality amongst Jordanian men with heart attacks using the chi-square automatic interaction detection model," *Health Informatics Journal,* vol. 30, no. 3, 2024. https://doi.org/10.1177/14604582241270830

[66] H. Qi *et al.*, "Prediction methods of common cancers in China using PCA-ANN and DBN-ELM-BP," *IEEE Access,* vol. 10, pp. 113397–113409, 2022. https://doi.org/10.1109/ACCESS.2022.3215706

[67] A. S. Abdalrada, J. Abawajy, T. Al-Quraishi, and S. M. S. Islam, "Prediction of cardiac autonomic neuropathy using a machine learning model in patients with diabetes," *Therapeutic Advances in Endocrinology and Metabolism,* vol. 13, 2022. https://doi.org/10.1177/20420188221086693

[68] K. Abnoosian, R. Farnoosh, and M. H. Behzadi, "Prediction of diabetes disease using an ensemble of machine learning multi-classifier models," *BMC Bioinformatics,* vol. 24, 2023. https://doi.org/10.1186/s12859-023-05465-z

[69] J. Wan and Y. Zeng, "Prediction of hepatic metastasis in esophageal cancer based on machine learning," *Scientific Reports,* vol. 14, 2024. https://doi.org/10.1038/s41598-024-63213-6

[70] M. Rahimi, M. R. Afrash, S. Shadnia, B. Mostafazadeh, P. E. T. Evini, M. S. Bardsiri, and M. Ramezani, "Prediction of the prognosis of poisoned patients undergoing hemodialysis using machine learning algorithms," *BMC Medical Informatics and Decision Making*, vol. 24, 2024. https://doi.org/10.1186/s12911-024-02443-0

[71] M. Kumar, S. Kumar Singh, and S. Kim, "Predictive analytics for mortality: FSRNCA-FLANN modeling using public health inventory records," *IEEE Access*, vol. 12, pp. 81252–81264, 2024. https://doi.org/10.1109/ACCESS.2024.3411162

[72] M. Islam, M. Usman, A. Mahmood, A. A. Abbasi, and O. Y. Song, "Predictive analytics framework for accurate estimation of child mortality rates for Internet of Things enabled smart healthcare systems," *International Journal of Distributed Sensor Networks*, vol. 16, no. 5, 2020. https://doi.org/10.1177/1550147720928897

[73] L. Tian *et al.*, "Radiomics-based machine-learning method for prediction of distant metastasis from soft-tissue sarcomas," *Clinical Radiology*, vol. 76, no. 2, pp. 158.e19–158.e25, 2021. https://doi.org/10.1016/j.crad.2020.08.038

[74] S. V. Babu, P. Ramya, and J. Gracewell, "Revolutionizing heart disease prediction with quantum-enhanced machine learning," *Scientific Reports*, vol. 14, 2024. https://doi.org/10.1038/s41598-024-55991-w

[75] A. Hennebelle, L. Ismail, H. Materwala, J. Al Kaabi, P. Ranjan, and R. Janardhanan, "Secure and privacy-preserving automated machine learning operations into end-to-end integrated IoT-edge-artificial intelligence-blockchain monitoring system for diabetes mellitus prediction," *Computational and Structural Biotechnology Journal*, vol. 23, pp. 212–233, 2023. https://doi.org/10.1016/j.csbj.2023.11.038

[76] A. Farrag, Z. M. Fadlullah, M. M. Fouda, and N. S. Almalki, "Survival-based treatment planning using stage-specific machine learning models," *IEEE Access*, vol. 11, pp. 134404–134420, 2023. https://doi.org/10.1109/ACCESS.2023.3337117

[77] H. A. Al-Jamimi, "Synergistic feature engineering and ensemble learning for early chronic disease prediction," *IEEE Access*, vol. 12, pp. 62215–62233, 2024. https://doi.org/10.1109/ACCESS.2024.3395512

[78] D. J. Houserman *et al.*, "The viability of an artificial intelligence/machine learning prediction model to determine candidates for knee arthroplasty," *Journal of Arthroplasty*, vol. 38, no. 10, pp. 2075–2080, 2023. https://doi.org/10.1016/j.arth.2022.04.003

[79] P. Liu and S. Fei, "Two-stage prediction of comorbid cancer patient survivability based on improved infinite feature selection," *IEEE Access,* vol. 8, pp. 169559–169567, 2020. https://doi.org/10.1109/ACCESS.2020.3016998

[80] K. N. Kunze, E. M. Polce, A. Patel, P. M. Courtney, and B. R. Levine, "Validation and performance of a machine-learning derived prediction guide for total knee arthroplasty component sizing," *Archives of Orthopaedic and Trauma Surgery,* vol. 141, pp. 2235–2244, 2021. https://doi.org/10.1007/s00402-021-04041-5

[81] A. M. Chekroud *et al*., "The promise of machine learning in predicting treatment outcomes in psychiatry," *World Psychiatry,* vol. 20, no. 2, pp. 154–170, 2021. https://doi.org/10.1002/wps.20882

[82] S. Simeone *et al*., "Delirium in ICU patients following cardiac surgery: An observational study," *Journal of Clinical Nursing,* vol. 27, nos. 9–10, pp. 1994–2002, 2018. https://doi.org/10.1111/jocn.14324

[83] H. N. Huang *et al*., "Employing feature engineering strategies to improve the performance of machine learning algorithms on echocardiogram dataset," *Digital Health,* vol. 9, 2023. https://doi.org/10.1177/20552076231207589

[84] C. Severn, K. Suresh, C. Görg, Y. S. Choi, R. Jain, and D. Ghosh, "A pipeline for the implementation and visualization of explainable machine learning for medical imaging using radiomics features," *Sensors,* vol. 22, no. 14, p. 5205, 2022. https://doi.org/10.3390/s22145205

[85] De Gruyter, "Classification and regression algorithms," *Machine Learning Algorithms for Predicting or Categorizing Data,* 2023. https://www.degruyter.com/document/doi/10.1515/9783111323749-003/html

[86] P. Strecht, L. Cruz, C. Soares, J. Mendes-Moreira, and R. Abreu, "A comparative study of classification and regression algorithms for modelling students' academic performance," in *Proceedings of the 8th International Conference on Educational Data Mining*, 2015. https://www.educationaldatamining.org/EDM2015/proceedings/short392-395.pdf

[87] M. Badawy, N. Ramadan, and H. A. Hefny, "Healthcare predictive analytics using machine learning and deep learning techniques: A survey," *Journal of Electrical Systems and Information Technology,* vol. 10, 2023. https://doi.org/10.1186/s43067-023-00108-y

9 AUTHORS

Nur Farah Afifah Ahmad Sukri earned her BSc in Computer Science (Software Development) and MSc in Computer Science from Universiti Sultan Zainal Abidin in 2016 and 2019, respectively. She is currently pursuing a PhD in the Department of Information Technology at the Faculty of Informatics and Computing, Universiti Sultan Zainal Abidin in Terengganu, Malaysia. Previously, she worked as a lecturer and laboratory assistant at Universiti Malaysia Terengganu before starting her PhD. Her research interests include natural language processing, machine learning, and deep learning (E-mail: sl4722@putra.unisza.edu.my).

Wan Mohd Amir Fazamin Wan Hamzah received a B.I.T. degree in Software Engineering, M.Sc. and Ph.D. degrees in Computer Science from the Universiti Malaysia Terengganu in 2003, 2010 and 2016, respectively. He is currently a senior lecturer with the Faculty of Informatics and Computing, Universiti Sultan Zainal Abidin (UniSZA), Terengganu, Malaysia. His research interests include learning analytics, machine learning, gamification, and e-learning (E-mail: amirfazamin@unisza.edu.my).

Mohd Kamir Yusof received Ph.D. degree from Universiti Malaysia Terengganu. He is currently a senior lecturer in the Department of Computer Science, Faculty of

Informatics and Computing. University Sultan Zainal Abidin (UniSZA), Terengganu, Malaysia. His research interests include data integration, and mobile and web application development (E-mail: mohdkamir@unisza.edu.my).

Ismahafezi Ismail is a Senior Lecturer at the Faculty of Informatics and Computing, Universiti Sultan Zainal Abidin (UniSZA), Terengganu, Malaysia. His research focuses on virtual reality, augmented reality, computer games, and animation (E-mail: ismahafezi@unisza.edu.my).

Harmy Mohamed Yusoff is the Dean of the Faculty of Medicine, UniSZA, with extensive experience and expertise in the field of Family Medicine. With an excellent academic background, he obtained his Master of Medicine (Family Medicine) from Universiti Sains Malaysia in 2002. Prior to that, he obtained his Doctor of Medicine (MD) in 1995 and Bachelor of Science in Medicine (BSc) in 1993, both from Universiti Kebangsaan Malaysia.

Azliza Yacob is a senior lecturer and a researcher at University College TATI (UCTATI). She received her Bachelor of Science (Computer) and Master of Science (Information Technology – Manufacturing) at University Teknologi Malaysia (UTM) in 2003 and 2005. She also completed her PhD in Computer Science, from University Malaysia Terengganu in 2018. Her research interest includes e-learning, quality control, decision-support systems, the computer industry and artificial intelligence (E-mail: azliza@uctati.edu.my).