MACHINE LEARNING FOR CHRONIC KIDNEY DISEASE: ADVANCES, CHALLENGES, AND FUTURE RESEARCH DIRECTIONS

Hassan Iftikhar[1], Sawera Kanwal[*2], Salwa Muhammad Akhtar[3], Rida Imtiaz[4], Aisha Habib[5]

[1]*Department of Computer Science, Baba Guru Nanak University, Nankana Sahib. Pakistan*
[*2]*Department of Computer Science, University of South Asia, Lahore. Pakistan*
[3]*Department of Information Systems, University of management and technology, Lahore, Pakistan*
[4]*Department of Zoology, University of Lahore. Pakistan*
[5]*Department of Computer Science, University of South Asia, Lahore. Pakistan*

[1]hassan.iftikhar@bgnu.edu.pk, [*2]Sawera.kanwal@usa.edu.pk, [3]salwa.akhtar@umt.edu.pk, [4]ridaimtiaz159@gmail.com, [5]ayesha.habib@usa.edu.pk

Abstract
Millions of individuals worldwide are affected by chronic kidney disease (CKD), a serious public health issue. To stop CKD from progressing to end-stage renal disease (ESRD), early identification and effective care are essential. However, conventional approaches to CKD diagnosis and prognosis are frequently insufficient, resulting in postponed therapy and subpar results. In the field of medicine, especially the diagnosis and prediction of CKD, machine learning (ML) has emerged as a potential approach. Large volumes of data can be analyzed by ML algorithms, which can also identify patterns that may elude human specialists. This may lead to more accurate CKD diagnosis and prognosis calculations, as well as the creation of novel, strong therapeutic approaches. In this article, we examine recent advances in machine learning (ML) in the field of chronic kidney disease (CKD), including research that employed ML to predict treatment outcomes and make diagnoses and prognoses. We also go over the difficulties and restrictions associated with using ML in the study of CKD, as well as potential future lines of inquiry. Our analysis highlights the potential of ML to improve CKD diagnosis, prognosis, and treatment, while also underscoring the need for further research to realize this potential fully.

INTRODUCTION

It is estimated that 10 percent of the world population is chronically affected by renal disease. It has a significant impact on the lives of individuals and society as a whole, and is a severe societal health issue. Early detection and management are necessary to prevent the development of the disease and improve the quality of life. Chronic kidney disease (CKD) is very common in the world, with a prevalence of 10-15 percent in the United States and 10.8 percent in China [1]. In Mexico, the most prevalent type of the disease is 14.7 percent among adults [1]. Chronic kidney disease (CKD) is a condition in which the kidneys fail to function correctly, resulting in the inability to filter blood and remove metabolic waste products. Consequently, trash accumulates, which can be detrimental to your health [2].

One such emerging global health issue is chronic kidney disease, which is targeted mainly by the elderly. It consists of diverse renal dysfunction and may lead to kidney failure. The United States is experiencing increasing kidney failure rates, which result in poor outcomes and high costs. However, there are people with chronic renal disease who are stable, particularly those who are above 65 years. [5]

The two primary risk factors of chronic renal disease are high blood pressure and diabetes. Uncontrolled high blood pressure may lead to heart attacks, strokes, and renal disease, unlike diabetes, which has the potential to damage the kidneys, heart, blood vessels, and eyes. Others are glomerulonephritis, hereditary diseases, dysplasia, kidney stones, malignancies, metabolic diseases, obesity, and old age. Advanced chronic kidney disease symptoms are fatigue, insomnia, loss of appetite, leg and ankle oedema, peripety, swelling under the eyes, skin urticaria, and frequent night urine discharge [3]. A greater emphasis is being placed on the early identification of chronic renal disease due to the expanding global population. This highlights the importance of this issue in the medical community, as effective illness management and therapy depend significantly on early detection [4]. Patients with chronic kidney disease progress through several phases that eventually lead to end-stage renal failure, which necessitates expensive kidney transplants or dialysis to preserve life. However, many people in underdeveloped nations cannot afford these procedures [6].

The timely identification and diagnosis of the disease determines the success of CKD management and the prevention of its outcomes. Machine learning (ML) has provided an opportunity to detect and diagnose a large number of diseases, including CKD. It is increasingly finding applications in the medical sector due to the surge in technology. Machine learning algorithms can be applied to healthcare management, overcoming the shortcomings of traditional methods by creating more accurate and efficient predictions. Through these predictive technologies, the healthcare industry can make informed decisions and improve operational management by analyzing massive amounts of data to identify trends and generate predictions [6]. Machine learning can be used to predict chronic kidney disease (CKD). This condition has been successfully identified and understood in the medical community through the analysis of data from previous CKD patients to train a predictive algorithm. In this manner, disease prediction can be more accurate and efficient, which is why machine learning can help improve the healthcare sector [7].

Millions of individuals worldwide are impacted by CKD, a serious public health issue that is linked to a high risk of mortality and morbidity. Many hospitals maintain databases of patients with chronic renal disease, and it is feasible to find functional patterns that can guide decision-making and help with disease prediction by analyzing this data using data mining techniques. Given the enormous number of persons affected by chronic kidney disease, the knowledge that might be discovered through such analysis could have profound effects [8]. The conventional clinical risk prediction models for CKD may have shortcomings, including reliance on limited patient data and an inability to consider the complexity of the disease. By analyzing large amounts of patient data, including both clinical and genetic data, machine learning can potentially enhance the accuracy of CKD risk prediction by revealing the relevant risk factors that traditional models may otherwise obscure. The aim of applying machine learning methods to CKD is to develop more precise and exclusive models for predicting disease progression. This may lead to prompt interventions and more customized treatment that would improve the outcomes and quality of life of patients [9]. The primary purpose of this manuscript is to provide an in-depth review of the different machine learning-based chronic kidney disease prediction models, in a bid to encompass all the pertinent issues of the study on CKD prediction models, such as the use of different data sources to gather data, encoding schemes to convert the categorical data into a machine learning algorithm's understandable format, contemporary

challenges, and prospects, this systematic literature review. Considering study behavior, exposure, and clinical value, we searched for the most recent developments in the sphere of illness diagnostic performance based on medical imaging pre-trained algorithms in relation to clinical professionals. This study was conducted to determine the diagnostic performance of transfer learning algorithms when applied in conjunction with medical professionals [10].

The primary goal of this study is to conduct a thorough analysis of the literature on pre-trained algorithms for disease diagnosis, with a focus on comparing the accuracy of these algorithms with that of human diagnosticians. To achieve this, we developed a taxonomy for describing diseases based on medical imaging, and we utilised this taxonomy to evaluate the efficacy of transfer learning algorithms. To better understand their relative performance, we also compared the outcomes of these algorithms with the diagnostic precision of health professionals. Overall, our work illustrates the capabilities and limitations of pre-trained algorithms for illness diagnosis in comparison to human diagnosticians, offering a contemporary overview of their potential.

The rest of the paper is organized as follows: Section II identifies the relevant work. Section III of this review outlines the methodology, detailing the study's objectives, search strategy, article selection criteria, abstract-based keywording, and quality assessment standards. Section IV contains an analysis and findings of the publications under evaluation. The entire discussion, suggested taxonomy, and overall layout of a prediction model for chronic kidney disease are presented in the summary of the research findings in Section V. Section VI addresses unresolved issues and challenges, discussing possible avenues for future research. The review concludes in Section VII.

II.RELATED WORK

In recent years, scholars have demonstrated that machine learning algorithms can perform complex calculations and data analysis to make decisions that benefit human life. Machine learning algorithms have consistently delivered promising results in experiments across various fields, including mathematics, science, business, medicine, and everyday life. [13]. The development of devices and approaches to monitor and predict multiple diseases, with a focus on the most prevalent ones in human life, is becoming increasingly significant in the eyes of scholars and isctors nowadays. This article predicts the risk of chronic kidney disease using machine learning methods and methods for addressing small datasets. [12]. Several scholars have applied various machine learning algorithms to chronic kidney disease cases in the past, utilising a range of tools to assess their validity.[14]. Based on variables such as socioeconomic position, age, education, gender, and the outcomes of a mini-mental state examination, a machine learning algorithm was developed to predict chronic kidney disease. The model's prediction accuracy is high at 91.3%.[15]. We attempted to evaluate the capability of different machine learning algorithms to provide CKD early detection in this research [17].

Although there have been extensive studies in this area, we focused on the use of predictive modelling as our strategy enhancement tool. With machine learning and predictive analytics, we can develop predictive models by researching the connection between information variables and the target group's characteristics. The predictive modeling enabled the incorporation of other variables or the identification of the most significant causes of CKD because it made it possible to measure the attributes more accurately. This strategy enabled us to develop a reliable prediction model and gain a deeper understanding of CKD. With this data, we aimed to maximize accuracy while minimizing input characteristics using principal component analysis (PCA). We also attempted to develop a machine learning-based model with the highest accuracy [7].

The authors of this article propose a hybrid model to forecast Chronic Kidney Disease (CKD) that involves the Logistic Regression (LR) and Random Forest (RF) algorithms. They contrasted their proposed model with six other

machine learning (ML) models: LR, RF, Support Vector Machine (SVM), K-Nearest Neighbor (kNN), Naive Bayes (NB), and Feedforward Neural Network (FNN). The highest percentage, 99.83, was achieved by their proposed hybrid model in predicting CKD. The effectiveness of the NB, K-Star, SVM, and J48 classifiers was also evaluated by the scientists using WEKA software. They also found that the J48 algorithm performed the best in comparison with the other algorithms, achieving an accuracy rate of 99%. In general, the study suggests the potential applicability of the proposed hybrid model and J48 algorithm in clinical practice for accurate diagnosis of CKD.[16].

Previous research on CKD has employed data analysis using ML techniques; however, no big data platforms have been deployed. This motivated us to explore the application of big data platforms, specifically, Apache Spark, to study and analyze CKD data. We are also examining the effectiveness of hybrid strategies that combine feature selection techniques with ML classification algorithms and those that combine feature selection techniques with ensemble algorithms. Spark enables us to handle large quantities of CKD data more efficiently, and the hybrid methodologies we are using may help improve the accuracy of CKD diagnosis and prediction [18]. The future ways in which machine learning may predict chronic kidney disease include Incorporating More types of data. The majority of research to date has utilised clinical and demographic data only. To enhance prediction accuracy, future researchers may consider incorporating other types of data, such as genetic or imaging data. Creation of individualized models: The existing models of predicting CKD are usually formulated using population-based information, which may not be representative of the risk of an individual. This could be improved by having individualized models that would consider individual characteristics of the person. Using longitudinal data: Because CKD is a degenerative condition, using data collected over time may increase prediction precision. The identification of patients at risk of CKD development and the

guidance of therapy choices could both benefit from longitudinal data—electronic health record chronic integration. EHRs are a treasure trove of patient information that could be leveraged to enhance CKD prediction algorithms. Additionally, integration with EHRs can make it easier to apply predictive models in clinical settings. Validation in a range of patient populations: The majority of research to date has concentrated on specific groups; therefore, it is crucial to verify predictive models across various patient demographics to ensure their applicability and efficacy in different settings.

Using machine learning, Mohammed and Beshah created a self-learning knowledge-based system to identify and treat the first three phases of chronic kidney disease. They developed a prototype that enables patients to ask questions of the knowledge-based system and receive answers as a result of their research, which used a modest dataset. They generated the rules using decision trees and reported an overall accuracy of 91% for the prototype. In a study on the prediction of chronic renal disease, Priyanka et al. used the Naive Bayes algorithm. They evaluated its performance in comparison to that of KNN, SVM, Decision trees, and ANN. In contrast to other algorithms, their investigation revealed that the Naive Bayes algorithm had the highest accuracy, at 94.6%.

The study by Almasoud and Ward aimed to measure the performance of machine learning algorithms using a subset of features to predict chronic kidney disease. They selected the predictive attributes based on Pearson correlation, ANOVA, and Cramér's V test, and then evaluated the models using LR, SVM, RF, and GB machine learning models. Their analysis had the highest accuracy in Gradient Boosting with an F-measure of 99.1. Through the study of the data of patients with CKD, Yashfi suggested using the machine learning algorithms to predict the likelihood of chronic kidney disease (CKD) development. 20 out of 25 features were retrieved with the help of the Random Forest (RF) and Artificial Neural Network (ANN) algorithms to predict the risk of developing CKD. Their data showed that RF was the most

accurate, with 97.12% accuracy in predicting the possibility of developing CKD.

Rady and Anwar utilised their research to compare the effectiveness of various machine learning algorithms, including Probabilistic Neural Networks (PNN), Multilayer Perceptron (MLP), Support Vector Machine (SVM), and Radial Basis Function (RBF) in predicting phases of kidney illness. The investigation was conducted using a small dataset and a limited number of features. The researchers found that PNN was the most accurate in the overall classification, with a rate of 96.7%, which implies that it can be considered an effective means of predicting the phases of renal disease. On a publicly available dataset of healthy patients and patients with kidney disease, Poonia et al. constructed and evaluated prediction models for kidney disease. They achieved this through the application of several machine learning algorithms, including KNN, ANN, SVM, NB, and logistic regression, as well as feature selection algorithms such as RFE and Chi-Square. The researchers concluded that the logistic regression-based prediction model, ha"orable features as selected by the Chi-Square method, had the highest accuracy of 98.75. In an attempt to predict BCD, Vinod examined the performance of seven supervised machine learning models: K-Nearest Neighbor, Decision Tree, Support Vector Machine, Random Forest, Neural Network, Naive Bayes, and Logistic Regression. The researchers found that k-NN was the best model to use in predicting the BCD, and it had the highest accuracy of 97 percent.

Various studies have been conducted to identify the use of machine learning techniques to predict chronic kidney disease. To increase the performance of the model, the amount of data, its quality, and the time of collection are significant. It is challenging to recommend a course of treatment based on the disease stage, as past studies have primarily focused on predicting the two groups.

III. RESEARCH METHODOLOGY

The systematic literature review (SLR) is an extensive process that helps identify, assess, and select the most significant research on a specific issue from all that is already known. In this work, the recommendations of the SLR, initially suggested in 2004, have been followed to ensure the collection and presentation of the analysed and extracted results as objective data. The purpose of this strategy is to establish a rigorous and transparent method for conducting literature reviews, ensuring that conclusions are based on a solid and reliable body of data. Generally, the SLR process is intended to reduce bias and enhance the accuracy and reliability of the results [11]. The research methodology that will be employed in this SLR is depicted in Figure 1. This process is appropriate and reflective with six levels, namely: 1) Specifying the goals of the research, 2) An explanation of research inquiries, 3) creation of a search plan, 4) paper screening and selection, 5) using keywords to group papers, 6) information synthesis and extraction.

A. RESEARCH OBJECTIVES (RO)

The primary purposes of this study are the following:

- RO1: The most crucial aspect is to be aware of the high-tech research on the topic of predicting chronic kidney disease in a high-tech manner.

- RO2: Descriptive visualization of the existing data sources, feature representation techniques, and machine learning algorithms used within the particular area of interest.

- RO3: A hierarchical classification of the different methods of feature encoding, and a suggested taxonomy will be proposed that will provide the most effective schemes that have found extensive application.

- RO4: One of the hypotheses is that the proposed generic framework will guide researchers in further developing the field.

- RO5: Identifying the main challenges and unsolved problems to uncover potential opportunities for future studies.

-

B. RESEARCH QUESTIONS (RQ)

This SLR has been designed to efficiently address the main research questions. An

extensive search strategy has been developed to identify and extract the most significant articles for the review. Table I outlines the research questions addressed in the current study, along with their primary drivers.

TABLE I: RQ and primary motivations

|	Research Question	Major Motivation
RQ1	What is the most effective machine learning algorithm for predicting chronic kidney disease risk and progression?	The effectiveness of a machine learning algorithm in predicting chronic kidney disease risk and progression depends on various factors and can vary across different datasets.
RQ2	How can machine learning be used to optimize treatment and management strategies for patients with chronic kidney disease?	To improve patient outcomes by providing more accurate and personalized treatment recommendations, reducing the risk of disease progression and complications, and discovering new and more effective treatments.
RQ3	How can healthcare systems and providers better coordinate and communicate with patients to improve CKD management and outcomes?	By using technology to facilitate communication and information-sharing between healthcare providers and patients.
RQ4	What are the gaps and challenges of using machine learning for chronic kidney disease?	
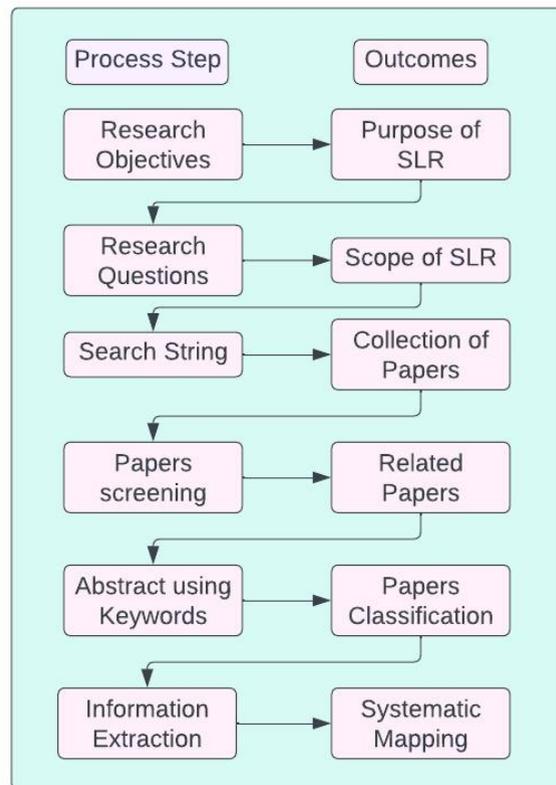
FIGURE 1: An SLR process model

C. SEARCH SCHEME

The creation of a search strategy that ensures the identification and collection of potentially essential articles on the topic of interest is the most crucial step in a systematic literature review (SLR). To find the most pertinent articles from the collection, this procedure entails designing an inclusion/exclusion plan, selecting the right literary resources to search, and developing a search string. Following the collection of the articles, various parts of them are evaluated qualitatively and experimentally to reflect distinct viewpoints connected with the research. This strategy aims to make sure that the literature evaluation is methodical, rigorous, and thorough, and that the chosen articles are of a high caliber and pertinent to the research topic being addressed.

- **SEARCH STRING**: To conduct an efficient and objective study, a keyword search and compilation of the existing studies in the field were performed. The primary concepts were analyzed with reference to the research questions to find the relevant keywords and terms applied in the area of the study, which served to guarantee the validity and applicability of the search string. To formalize a search string to find the most relevant articles, in Table II, the keywords with the synonyms were compiled as finalized keywords. The sign + means that a particular keyword appears, and the sign - means that a specific keyword does not appear. Such a process helps ensure the integrity of the research process by verifying that all relevant papers have been collected.

TABLE II. Search terms and keywords

Terms (Keywords)	Synonyms / Keywords
+ Computer-Based (CB)	Machine-based, Technological, Machine learning (ML)
+ Chronic Kidney Disease (CKD)	Chronic Renal Disease (CRD), Chronic renal failure (CRF), Chronic kidney insufficiency (CKI)
- Mass Spectrometry (MS)	Spectrometry (SY), Spectrometric (SC), Spectrum (SM)
- End-Stage Renal Disease (ESRD)	kidney failure.

To create a search string from the finalized keywords and alternative phrases, the logical operators "AND" and "OR" were employed. The wild character "*" was then used to represent zero or more characters as necessary. The "OR" operator was used to include more search alternatives, and the "AND" operator was used to connect the terms and define the search criteria to get precise search results. This made it possible to focus the search and only return pertinent results. Three distinct parts comprise the search string. To focus on results that are computer-based or computational, the first section is used. The forecast of chronic kidney disease is covered in the following section, and the results of non-computer-based investigations, including lab-based trials, are limited in the last section. The mathematical formulation of the search string is represented in equation 1.

$$R = \forall \left[(CB \lor MB \lor ML) \land (CRD \lor CRF \lor CKI) \neq (MS \lor SY \lor SC \lor SM \lor ESRD) \right] \quad (1)$$

The mathematical foundation of the search string is shown in Equation (1). R stands for the search outcomes found using the search string. The word "for all" is represented by the symbol "," while "OR" and "AND" are represented by the letters "" and ", respectively. When combined with the search terms specified in Table II, these symbols formalize the entire search string for the chosen repository. Equation (1) may be used to express a generalized form of the search term as follows: ((computer-based OR Machine-based OR machine learning) AND ("Chronic Renal Disease" OR "Chronic renal failure" OR "Chronic kidney insufficiency" NOT (Spectrometry OR Kidney Disease)

- **LITERATURE RESOURCES** To carry out the literature search in the online repositories focused on the research publication and collection, the most well-known and topic-specific journals have been selected. The details of the chosen repositories, search terms, and the results are presented in Table III.

TABLE III. Publisher-wise search strings

Repository	Search Strings
PLOS	"Computer-based" or computer-based or machine learning) and (chronic renal disease or chronic renal failure or prediction of chronic kidney disease), not mass spectrometry
PMC	("Computer based"[all fields] or computer-based [all fields] or machine learning [all fields]) and ("chronical renal disease "[all fields] or " chronic renal failure "[all fields] or " chronic kidney insufficiency "[all fields] or "prediction of chronic kidney disease"[all fields]) not (spectrometry or spectrometric [all fields] or spectrum [all fields])

MDPI	Computer-based or machine learning and chronic renal disease or chronic renal failure or chronic kidney insufficiency, and not mass spectrometry, not kidney failure)
Springer Link	Note: a full stop is included at the end of the first and second lines to indicate the start and the end of the search query, respectively.
Science Direct	Cor(computer-based or machine-based or machine learning) and (chronic renal disease or chronic renal failure or chronic kidney insufficiency) and not (spectrometry or kidney failure).
IEEE Explore	((((("All metadata": "computer-based") or "all metadata": machine-based) or "all metadata": machine learning) and "all metadata": chronic renal disease) or "all metadata": chronic renal failure) not "all metadata": spectrometry)

INCLUSION AND EXCLUSION CRITERIA:

Parameters defined for inclusion criteria (IC) and All articles were subject to exclusion criteria (EC), which aimed to eliminate studies that involved lab-based experimental identification or prediction of chronic kidney disease for their effects, or were involved in other irrelevant procedures, like

IC 1) Include studies that were primarily conducted for computational prediction of chronic kidney disease.
IC 2) Presence of underlying conditions, such as diabetes or high blood pressure, that increase the risk of kidney damage and progressive kidney function decline.
EC1) Studies that do not include human subjects.
EC2) Studies that focus on acute kidney injury rather than chronic kidney disease.
EC3) Studies that have a high risk of bias or poor methodological quality.

- **SELECTION OF RELEVANT PAPERS** The period from January 2010 to December 2023 was selected for the publication search due to its relevance. When the primary search method is employed, numerous research articles can be found; however, not all of them are highly relevant to the formulated research questions, and duplication exists. To retrieve essential papers, the searched papers should be re-assessed and filtered. The process of identifying the relevance of the articles, as well as screening them, was carried out using the described method in [28]. The first step involves filtering the research by title and removing duplicates. In this study, several publications were not relevant to the chosen field. The initial approach was to filter them out based on their titles. Abstracts were reviewed meticulously in the second phase of selection, and papers that reported computational aspects in the selected field were incorporated. Moreover, in addition to examining the articles that met the inclusion criteria, an exclusion criterion was applied to filter out articles that discussed the prediction of therapeutic peptides but made no significant contribution to the study or were clinical studies. The selected articles were then included in the subsequent level of assessment after the clarification of the method.

- **ABSTRACT-BASED KEYWORDING** The screening and categorizing of the articles were also done using the two-stage abstract-based keywording

procedure as outlined in [22] to identify the most relevant articles. The initial stage of analysing the abstract involved identifying the most topical keywords, understanding the article's basic idea, and assessing its contribution to the relevant sphere. Next, a profound grasp of the contribution of the research to the field has been determined with the help of the keywords obtained after analyzing a vast number of articles. Ultimately, these keywords were used to categorize the articles to be reviewed in the mapping process. The selected publications will be classified into four major categories: AIP, AACP, and AMP, based on the significance of the treatments. These subjects were chosen because they are directly related to peptide-based treatment-dependent topics. Other areas, such as "Chronic Disease (CD)" and Cell-Chronic Disease (CCD), are also excluded from the scope of this study. Lastly, to clarify, AMP is considered and classified here as AMP, despite exhibiting other known medicinal properties, such as antiviral, antibacterial, and antifungal effects [29, 30].

F. QUALITY ASSESSMENT CRITERIA:
A systematic review has an essential element, which is a systematic analysis of the articles that were included in the review. The designs of such studies differ, which is why the sequential evaluation scheme presented in [31] served as a reference point for the quality assessment. The process entails several steps, including quality evaluation, consideration of all essential research-related variables, such as the notion, study plan, mode of data collection, mode of data analysis, discussion, and data results [32]. Based on the above principles, the questionnaire was developed in cooperation with other authors to facilitate the quality of the research (see Table IV). The quality of the assessment has also been improved using internal and external quality criteria as proposed by [33]. The internal standards assess the quality of an article, while external standards evaluate the reliability and uniformity of the publication source. The evaluation/ scoring of the external quality has been based on the application of the Computer Science Conference rankings (CORE) and Journal Citation Reports (JCR) [34]. The score is the total of the constituent scores for each criterion. The result can range from 0 to 10, and it can be classified as high, average, or low based on whether the consequence exceeds 8, falls between 6 and 8, or is less than 6.

TABLE IV: Quality Assessment Questionnaire.

Sr.	Assessment Questions	Expected Answers	Score
Internal Scoring			
1	Was the abstract well described?	a. Yes	a. 0.5
		b. Intermediate	b. 1
		c. No	c. 1
2	Was the background/literature review described in detail?	a. Yes	a. 0.5
		b. Intermediate	b. 0.5
		c. No	c. 1
3	Was the result assessment well described, valid, and reliable?	a. Yes	a. 0.5
		b. Intermediate	b. 1
		c. No	c. 1
4	Was the conclusion relevant and effectively based on results?	a. Yes	a. 0.5
		b. Intermediate	b. 1
		c. No	c. 1.5
External scoring (based on publication source)			
5		a. CORE rank A	a. 1.5

	A study published in CORE ranked conference, proceedings, and symposium.	b. CORE rank B	b. 0.5
		c. CORE rank C	c. 0.5
		d. No CORE ranking	d. 1
6	A study published in JCR is a ranked journal	a. JCR rank Q1	a. 1
		b. JCR rank Q2	b. 2
		c. JCR rank Q3/Q4	c. 0.5
		d. No JCR ranking	d. 0

III. DATA ANALYSIS

provides a precise evaluation of all the articles selected, which were chosen after compiling the results. To successfully respond to the research questions, the selected publications were examined. The search results produced by the specified search string are discussed in the first section of this paragraph. The evaluation score is followed by its description, and the final section provides an in-depth discussion that addresses the research questions

The primary search process retrieved a total of 2,269 articles from various online sources. These articles serve as the foundation for developing benchmark datasets, feature extraction techniques, and computational models for predicting therapeutic peptides. The selection procedure, as detailed in the previous section, was applied to this collection. The step-by-step selection workflow is illustrated in Figure 2, while the results of each phase are summarized in Table V.

A. SEARCH RESULTS

TABLE V. Publisher-Based Stage-Wise Selection Process

Database/Repository	Primary Search	P-I	P-II	p-III
PLOS	5	2	2	2
Springer Link	28	20	18	18
IEEE Xplore	17	17	15	12
Mdpi	21	21	15	15
Total	71	60	50	47

Phase I (P-I) selection of 60 papers was conducted by the two authors, based on titles. The extraneous articles on the topic of study were further filtered after the elimination of the duplicate articles in phase (P-II) in accordance with the inclusion and exclusion criteria discussed in the preceding section. In the analysis, a phase III/P-III full-text-based analysis was conducted on 50 articles, and 47 articles were deemed most relevant and incorporated into this SLR for data extraction and analysis. Using the example of the search process, papers were also found during the search process, such as those on chronic kidney disease due to their binding properties, but not in Spectrometry (SY), Spectrometric (SC), or Spectrum (SM), thus indicating individuals with no relation whatsoever.

As indicated in the search strategy in Table III, the highly developed digital libraries (DL) that publish research studies from various journals, conferences, and workshops were employed to select the studies used in this systematic literature review. In Figure 3, the DL-wise distribution ratio of some of the selected articles is presented. The former is SPRINGER (38.29% share), followed by MDPI (31.91%), PLOS ONE (4.25%), and IEE Xplore (25.53%). Table V above already presents the distribution ratio and the status of publisher-based stage-wise selection of the chosen studies.
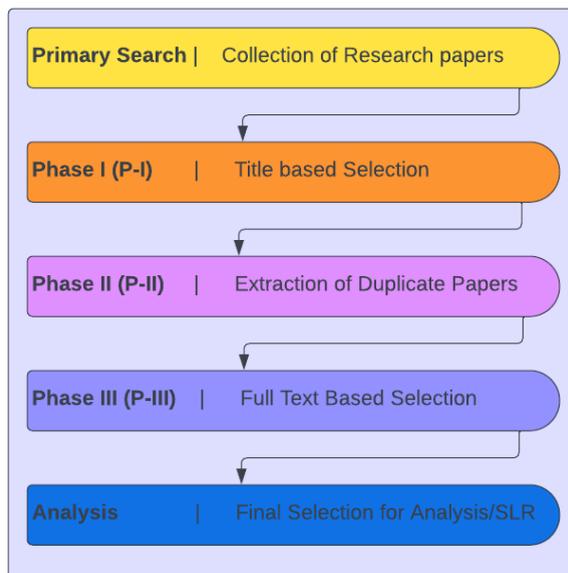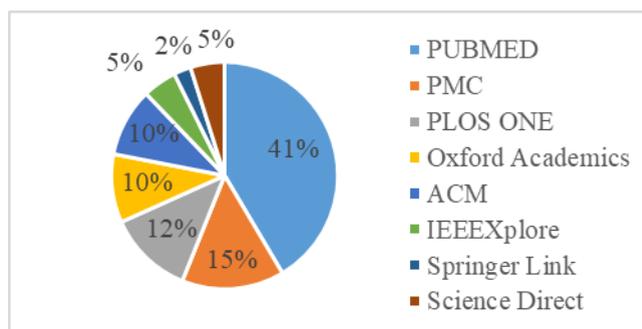
FIGURE 2. Selection procedure

FIGURE 3. DL-wise ratio of selected studies

B. QUALITY ASSESSMENT SCORE

According to the scoring scheme presented above, scores were assigned to each of the selected studies, which were then evaluated using both internal and external criteria, as shown in Table VI. By looking at the results of the scores based on internal and external criteria in terms of I-Score and E-Score respectively and publication type respectively, it can be seen that out of all the calculated articles 32 percent of the articles scored the highest mark of over 8 and was regarded as being in the high-ranking category, 58 percent of the articles in the results scored the average rank and only 4 articles in the results were in the low-ranking category based on the criteria. These facts are illustrated in Figure 4, which shows that people are generally confident about the quality of the selected articles. Quite the contrary, there was no quality-based elimination.
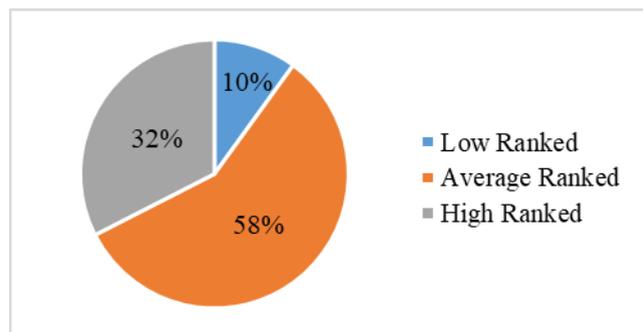
FIGURE 4. Percentage of score-based articles ranking

TABLE VI. Quality assessment score and classification

Ref No.	Publication		Classification		Empirical Validation	Internal Scoring			External Scoring		Total Score
	Channel	Year	Machine learning algorithm	Dominating Factors		1	2	3	4	5	
[1]	Journal	2020	None	Position monitoring	Validated	1	1	1	1	0.5	4.5
[2]	Journal	2020	None		Not validated	0.5	0.5	0	0.5	0	1.5
[3]	Journal	2018	None	Not applicable	Not validated	0.5	0.5	0	0.5	0	1.5
[4]	Journal	2021	None		Not validated	1	1	1	1	0	4
[5]	Conference	2018	None	No feature	Not validated	1	0.5	0.5	1	0	3
[6]	Journal	2019	None		Validated	1	0.5	1	1	1.5	4.5
[7]	Journal	2019	None	Location tracking shockwave generator.	Validated	1	1	1	0.5	0.5	4
[8]	Conference	2016	Decision Tree		Validated	1	0.5	1	1	0	3.5
[9]	Journal	2020	None	Image scanning and placement.	Validated	0.5	0	0.5	0.5	0.5	2
[10]	Conference	2019	None		Not validated	1	0.5	0.5	0	0	2

[11]	Journal	2021	Logistic regression	Shockwave, audio recording, and tracking.	Validated	1	1	1	1	0.5	4.5
[12]	Conference	2018	Logistic regression		Validated	1	1	0.5	1	0	3.5
[13]	Journal	2018	None	Shock-generating GPS tracker.	Not validated	0	1	0.5	0	0.5	2
[14]	Conference	2021	None		Validated	0.5	1	1	1	0	3.5
[15]	Conference	2018	Hidden Markov Model	Position tracking	Validated	1	1	1	1	0	4
[16]	Journal	2022	Logistic regression		Validated	1	0.5	0.5	1	0.5	3.5
[17]	Conference	2021	None	Position monitoring	Not validated	0.5	0	0	1	0.5	2
[18]	Journal	2019	None		Not validated	0.5	0.5	0.5	1	2	4.5
[19]	Journal	2021	None	Location detection	Not validated	0	0.5	0.5	0	0.5	1.5
[20]	Conference	2020	Outlier detection		Validated	1	1	1	1	0	4
[21]	Conference	2017	Decision tree	GPS tracking system	Not validated	1	1	1	1	0	4
[22]	Journal	2021	Logistic regression		Validated	1	1	1	1	0	4
[23]	Conference	2022	None	Location tracking alarm.	Validated	1	1	1	0	0	3
[24]	Conference	2016	None		Validated	1	1	1	0	0	3
[25]	Conference	2021	None	Position tracking	Not validated	1	0.5	0.5	0.5	0	2.5
[26]	Conference	2021	None		Validated	1	0.5	1	1	1.5	5
[27]	Journal	2019	Logistic Regression	Location service	Validated	0	0.5	0.5	0.5	0.5	2
[28]	Conference	2019	None		Validated	1	1	0.5	1	0	3.5
[29]	Conference	2022	None	GPS tracking function	Validated	1	0.5	0.5	0.5	0	2.5

[30]	Conference	2021	None		Not validated	1	0.5	0.5	0.5	0	2.5
[31]	Journal	2020	None	Position monitoring	Not validated	1	1	0.5	0.5	0.5	3.5
[32]	Journal	2021	Hidden Markov Model		Not validated	0.5	0.5	0.5	0.5	0.5	2.5
[33]	Journal	2020	Logistic regression	Location detection, call image capture, and alarm.	Validated	0.5	1	1	0.5	0	4
[34]	Conference	2020	None		Validated	1	0.5	0.5	1	0	3

C. ASSESSMENT AND DISCUSSION OF RESEARCH QUESTIONS

The selected studies were analyzed based on the research questions outlined in Table I, and the extracted information was discussed using a piecemeal evaluation approach. A total of 47 articles were analysed, and the resulting facts and findings were synthesized and summarized to provide a comprehensive understanding of the application of machine learning in chronic kidney disease.

1) Assessment of question 1: what is the most effective machine learning algorithm for predicting

chronic kidney disease risk and progression?
The best machine learning algorithm for predicting the risks and progression of chronic kidney disease (CKD) largely depends on the characteristics of the data and the research question being addressed. Since each of the two algorithms has its advantages and disadvantages, it is often necessary to critically analyze and compare the different approaches to arrive at the most relevant one. The most widely used algorithms in predicting CKD are logistic regression, decision trees, random forests, support vector machines (SVMs), and artificial neural networks (ANNs). Logistic Regression is a popular statistical model used for classifying problems involving binary classification. It is used in CKD prediction, with the assumption that the presence of the disease (CKD vs. non-CKD) is predicted by the use of predictor variables that encompass demographics, medical history, and laboratory outcomes.

- **Decision Trees:** This type of algorithm divides the data into subsets recursively until a final decision is reached. They provide justifiable models for identifying CKD risks and manifestations of the disease.
- **Random Forests:** It is a collection of various decision trees that are applied to increase the quality and power of prediction and minimise overfitting. Random forests yield promising results in the analysis of multidimensional CKD data.
- **Support Vector Machines (SVMs):** The SVMs are implemented to categorize the data as the optimal hyperplane to divide the patients and those without CKD. They can be applied to high-dimensional data and offer good performance in predicting CKD.
- **Artificial Neural Networks (ANNs):** ANNs are modelled after humans, and each node in the network is connected to another node. And the input in the network is processed to achieve predictions. ANNs can lead to non-linear and difficult-to-comprehend relationships between patient risk and progression in CKD, provided sufficient patient data is available.

Ref. No	Model	Best Model	Paper Category	Accuracy
[35]	Logistic Regression +Decision Tree +Random Forest +Neural Network +Naive Bayes + SVM	Random Forest	IEEE Conference	87.50%
[36]	DNN	DNN	IEEE	96%
[37]	Regression + Clustering	Clustering	IEEE	0.824
[38]	Linear regression + +Random Forest + Gradient Boosting	Random Forest	International Conference(dl.acm.org)	80%
[39]	Decision tree +Random Forest +Neural Network +Naive Bayes +SVM	Random Forest	International Conference(researchgate.net)	89.10%
[40]	Random Forest	Random Forest	dl.acm.org	88%
[41]	XGBoost	XGBoost	arxiv.org	80%
[42]	Linear +logistic regression +decision trees +random forest +gradient boosting	NA	Springer	
[43]	Artificial Neural Network	Artificial Neural Network	Springer	85%
[44]	DNN	DNN	Elsevier	90%
[45]	logistic regression+ decision tree+ random forest+(SVM) + (KNN)	Random Forest	IEEE International Conference	83%
[46]	Random Forest +Gradient Boosting+ econometrics (OLS)	Random Forest	NA	80%
[47]	decision tree, random forest+ (SVM)	NA	IEEE Symposium	92%
[48]	DNN	DNN	Springer	90%
[49]	decision tree +random forest+ SVM+LSTM	LSTM	pubsonline.informs.org	90%

[50]	linear regression, decision tree, random forest, and support vector regression	NA	Springer	80%
[51]	Linear regression +decision tree +random forest + SVM	LSTM	aisel.aisnet.org	90%
[52]	multiple linear regression +decision tree +random forest + SVM+LSTM	For Future use, deep learning techniques	Springer Survey	
[53]	multiple linear +decision tree +random forest +SVM +LSTM	LSTM	Emerald	95%
[54]	decision tree, random forest, and support vector machine (SVM)	NA	Elsevier	92%
[55]	multiple linear regression, decision tree, random forest, and support vector +LSTM	LSTM	Elsevier	95%
[56]	multiple linear regression, decision tree, random forest, and support vector regression	NA	Elsevier	90%
[57]	multiple linear regression, decision tree, random forest, and support vector + LSTM	NA	arxiv.org	92%
[58]	multiple linear regression, decision tree, random forest, and support vector + tacking fusion model	stacking fusion model	Springer	92%
[59]	Decision Tree, Random Forest, and Neural Network	NA	upcommons.upc.edu	72%
[60]	Random Forest and Gradient Boosting	NA	IEEE International Conference	85%
[61]	Decision Tree, Random Forest, and Neural Network	NA	dl.acm International Conference	72%
[62]	XGBoost	NA	Springer	80%
[65]	Random Forest, Neural Network, and Support Vector Machine	NA	inderscienceonline.com	85%

2)

Assessment of question 2: How can machine learning be used to optimize treatment and management strategies for patients with chronic kidney disease?

Machine learning brings effective means to improve the treatment and management of chronic kidney disease (CKD).

• **Anticipating disease development**: ML models can predict disease progression in

specific patients, allowing clinicians to identify those at high risk and develop strategies to delay or avoid their development.

• **Determining treatment response:** ML can assess the most effective treatment in a particular patient by examining laboratory findings, medical history, and other information about the patient, which helps in making improved treatment decisions.

• **Individualization of treatment plans**: ML is used to design individualized treatment plans using patient-specific information, so that the care is personalized to the needs of each specific patient.

• **Resource allocation optimization**: ML can identify patients at risk of hospitalization or readmission, enabling healthcare providers to utilize resources more effectively and eliminate unnecessary hospitalizations.

• **Earlier detection of complications**: ML models are capable of identifying complications earlier, allowing them to act and prevent serious consequences, e.g., anemia or bone disease.

To conclude, machine learning can enhance CKD management by forecasting disease progression, identifying effective treatments, personalizing care, optimizing resources, and facilitating early detection of complications. This information-based strategy enables clinicians to make sound judgments, which ultimately lead to improved patient outcomes and quality of life.

Ref. No	Data Source	Data Source Link	Accessible
[66][68][70][71]	Chronic Kidney Disease	Kaggle	Yes
[69][80][79][67]	CKD	Kaggle	Yes
[72][73]	Kidney Disease Dataset	Kaggle	Yes
[74][75][76][77]	CT KIDNEY	Kaggle	Yes
[80][82][84][85][81]	chronic kidney disease	Kaggle	Yes
[86][83]	Causes of Deaths around the World	Kaggle	Yes
[65][60]	Diabetes Health Indicators	Kaggle	Yes
[57][54][80]	MovieLens	Kaggle	Yes
[79]	IFOA-GRNN	Kaggle	Yes
[81][30]	Malaysia Box Office	NA	NA

3) **Assessment of question 3: How can healthcare systems and providers better coordinate and communicate with patients to improve CKD management and outcomes?**

Healthcare systems and providers can better coordinate and communicate with patients to improve CKD management and outcomes by implementing the following strategies:

- **Patient education:** Providing patients with educational resources about CKD, including lifestyle modifications and treatment options, can improve patient engagement and understanding of their condition. This can lead to better adherence to treatment plans and improved outcomes.

- **Shared decision-making:** Involving patients in the decision-making process about their care can improve patient satisfaction and engagement, leading to better outcomes. This can be achieved by providing patients with information about their treatment options and involving them in the decision-making process.

- **Care coordination:** Coordination between healthcare providers, including primary care physicians, nephrologists, and other specialists, can improve communication and collaboration, leading to better CKD management. This can be achieved through the use of shared electronic health records and regular communication between healthcare providers.

- **Patient-centred care:** Focusing on the patient's needs, preferences, and goals can improve patient satisfaction and engagement, leading to better outcomes. This can be achieved through regular communication with patients, listening to their concerns, and involving them in the decision-making process.

- **Use of technology:** Implementing technology, such as patient portals and mobile apps, can improve communication between patients and healthcare providers, leading to better engagement and adherence to treatment plans.

Healthcare systems and providers can enhance CKD management and outcomes by coordinating and communicating more effectively with patients. This can be achieved through patient education, patient involvement in shared decision-making, enhanced care coordination, a focus on patient-centred care, and the use of technology to improve communication. By implementing these strategies, healthcare providers can improve patient outcomes and quality of life.

4) Assessment of question 4: What are the gaps and challenges of chronic kidney disease using machine learning?

Machine learning has the potential to enhance the diagnosis, treatment, and management of chronic kidney disease (CKD), but several challenges must be addressed. These challenges include issues related to data quality and availability, model interpretability, generalizability, ethics and privacy, as well as a lack of standardization in the management of CKD. Machine learning in CKD requires collaboration between clinicians, data scientists, and other stakeholders to address these challenges and develop high-quality datasets, interpretable models, ethical and secure data sharing practices, and standardized protocols. Data quality and availability: The quality and availability of data continue to be significant challenges in CKD management. Many datasets are incomplete, inconsistent, or contain missing values, which can negatively impact the performance of machine learning models. Additionally, data collection can be time-consuming and costly, particularly for rare or chronic conditions like CKD.

- **Interpretability:** Machine learning models can be complex and challenging to interpret, especially for clinicians who lack a background in data science. This can make it difficult to understand how the model is making predictions and identify potential biases or limitations.

- **Generalizability:** Machine learning models trained on one dataset may not perform well on another, particularly if the data distribution or patient population differs significantly. This can make it challenging to apply machine learning models to diverse patient populations or in different healthcare settings.

- **Ethics and privacy:** Machine learning models that utilize patient data must be developed and deployed with consideration for both ethics and privacy. Ensuring that patient data is secure and used appropriately is critical for building trust and ensuring patient safety.

• **Lack of standardization: There is a lack of standardization** in CKD management, particularly in terms of diagnostic criteria, treatment protocols, and outcome measures. This can make it challenging to develop machine learning models that are broadly applicable across different healthcare settings and patient populations.

To address the gaps and challenges of using machine learning in CKD, collaboration between clinicians, data scientists, and other stakeholders in healthcare is necessary. Working together, they can develop high-quality datasets, interpretable machine learning models, and ethical and secure data sharing practices to improve CKD management and patient outcomes.

A. TAXONOMY OF FEATURE ENCODINGS

The machine learning models the taxonomy of CKD, which typically does not involve the diagnosis or treatment of CKD. However, predictive models based on machine learning can be used to predict patients who are likely to develop CKD and predict outcomes like disease progression or response to therapy. Some of the standard models used in CKD research include decision trees, random forests, support vector machines, neural networks, and logistic regression. These models are typically trained on electronic health records, lab findings, and medical imaging data to identify patterns and risk factors associated with CKD development or progression.

Although these methods have demonstrated potential in enhancing CKD diagnosis and management accuracy, they remain in their early stages of development and require further substantiation before being considered standard clinical practice. It is also important to emphasize that machine learning should not replace clinical experience and proven diagnostic techniques, but rather complement them to provide accurate diagnoses and effective treatments. Each of the encoding strategies is addressed in some detail in Section III, and their hierarchical structure is illustrated in Figure 7.
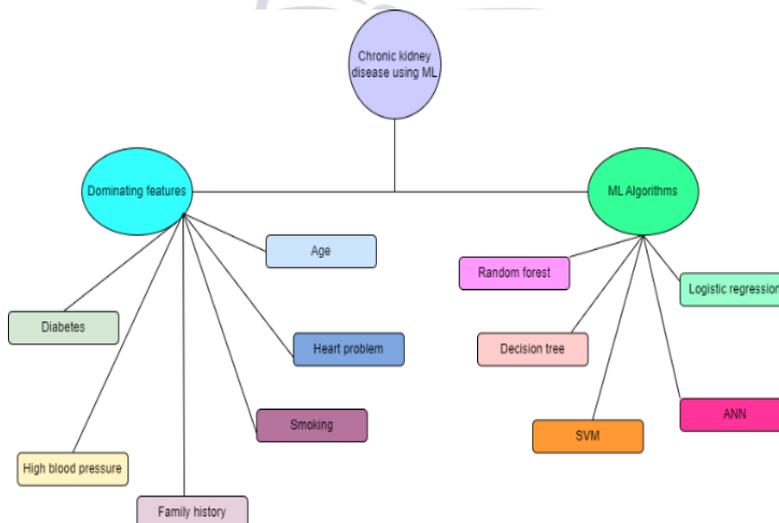

FIGURE 7. Feature encoding hierarchy

B. FRAMEWORK FOR PEPTIDE PREDICTION MODEL

Specific machine learning (ML) models for diagnosing or treating chronic kidney disease (CKD) are typically not included in the taxonomy of this disease. Nevertheless, one can use ML algorithms to develop predictive models that identify patients at risk of developing CKD and predict their future outcomes, which may include disease progression or response to

treatment. Decision trees, random forests, support vector machines, neural networks, and logistic regression are widely used models in research on CKD. These methods are typically trained on electronic health record information, laboratory test results, and medical imaging to identify patterns and risk factors associated with the occurrence and progression of CKD. Despite the promising opportunities offered by these models, they are still in their nascent stage and require verification before they can be widely used to supplement diagnostic accuracy and treatment strategies in clinical practice. Notably, ML must not be used as an alternative to clinical knowledge and existing diagnostic devices to make correct decisions and treat patients appropriately.

CKD is a long-term disorder where the kidney malfunction is caused by impaired damage affecting its functionality, and this ranks as a high global burden to society, with high morbidity and mortality rates. The ability to accurately predict the course of the disease is a significant challenge in CKD management, as it is vital for designing an effective treatment plan and achieving better patient outcomes. ML has demonstrated good potential in this regard, particularly in identifying high-risk populations, predicting disease progression, and supporting the development of individualized treatment plans. Nevertheless, several obstacles are curbing the application of ML in CKD care. Moreover, issues related to ethics (including patient privacy protection and reducing algorithm bias) should be addressed carefully. To create strong and clinically useful AI-based models, collaboration among clinicians, data scientists, and ML experts is essential. This type of collaboration will not only result in technically sound models but also in models that are proven in the real-world clinical setting without sacrificing ethical considerations. To sum up, although there are still serious flaws, ML has tremendous opportunities to enhance the diagnosis, treatment, and management of CKD. Interdisciplinary partnerships and technological development can be leveraged to advance research on CKD and enhance these tools to achieve better patient outcomes and quality of care. Developing a machine learning model for CKD is a complex task that requires collaboration between different experts to ensure that the model is accurate, relevant, and ethical. Here are some reasons why a multidisciplinary approach is essential:

- **Clinical context:** Machine learning models for CKD need to be developed in close collaboration with clinical experts, who can provide insights into the clinical context and the specific challenges faced by patients with CKD. This includes understanding the disease progression, the risk factors, and the impact of CKD on patient outcomes. Clinical experts can also help to identify relevant outcome measures and clinical endpoints that the machine learning model should predict.

- **Data quality:** Developing an accurate machine learning model requires high-quality data that is representative of the population of interest. Data scientists can collaborate with clinical experts to curate and preprocess data, ensuring it is relevant, accurate, and complete. They can also utilize techniques such as feature engineering to identify the most crucial features for the model to accurately predict CKD outcomes.

- **Ethical implications**: Machine learning models for CKD can have significant moral implications, particularly in the areas of privacy, bias, and fairness. Data scientists and machine learning specialists need to work with clinical experts to ensure that the model is developed in a way that respects patient privacy, minimizes bias, and promotes fairness in the prediction and diagnosis of CKD. Model validation and monitoring: Once the machine learning model is developed, it must be validated and monitored to ensure its accuracy and relevance over time. This requires collaboration between clinical experts and machine learning specialists to identify appropriate validation metrics, to test the model on new data, and to monitor its performance in real-world settings. In summary, developing a machine learning model for CKD requires collaboration between different experts to ensure that the model is accurate, relevant,

and ethical. By working together, clinical experts, data scientists, and machine learning specialists can maximize the impact of the model on patient outcomes and improve the diagnosis and treatment of CKD.

V. ISSUES AND CHALLENGES

CKD is a complex disease that poses numerous challenges and issues for patients, healthcare providers, and researchers. Here are some of the key challenges and issues related to CKD:

- **Prevalence and burden**: CKD is a significant public health problem, affecting millions of people worldwide. The prevalence of CKD is increasing due to the ageing population, rising rates of diabetes and hypertension, and other risk factors. CKD is associated with a significant burden of morbidity and mortality, as well as high healthcare costs.

- **Diagnosis and screening:** CKD is often asymptomatic at its initial phases and, therefore, difficult to diagnose and intervene in time. Suitable screening is a challenging process because it often requires multiple laboratory tests and imaging interventions. Moreover, the most effective screening strategies and standardized diagnostic criteria are not clearly agreed upon, which makes it difficult to detect it at an earlier stage.

- **Treatment and management:** The management of CKD is based on the cause of the condition as well as its progression stage. The available treatment modalities include lifestyle changes and pharmacotherapy, dialysis, or kidney transplants, depending on the severity of the case. Regardless of these availabilities, some types of CKD lack effective treatment, and access to correct treatment is still limited in most areas.

- **Comorbidities and complications:** CKD is usually associated with such comorbid conditions as anaemia, bone disorders, cardiovascular disease, and cognitive decline. Such complications further complicate the management of diseases and necessitate a multidisciplinary approach to patient care.

- **Data quality and standardization:** One of the most significant gaps in CKD research is that data collection and reporting methods are not standardized. This lack of consistency prevents comparison of results between studies and populations. The quality of data and the encouragement of standardization are the steps to better research and better clinical outcomes.

- **Health differences:** CKD is disproportionately found in vulnerable population groups such as racial and ethnic minorities, low-income, and rural populations. These disparities require a comprehensive approach that addresses the social determinants of health, promotes health equity, and improves access to efficient and timely care.

VI. CONCLUSION

In this paper, a systematic literature review has been done, which gives a detailed discussion of qualitatively chosen research papers found within the field of computational prediction of CKD. This article presents a comprehensive discussion of qualitatively selected research papers, now accessible in the field of computational prediction of CKD, to offer a review of the literature. To research the purpose of choosing the fourth published articles, a clear and systematic methodology was used. This paper has provided an analysis of the different ML classifiers, feature extraction, and data sources. A model has been provided as a roadmap for future developments in the field, incorporating best practices learned from the selected publications. The feature encodings have been grouped and presented accordingly, following the taxonomic order. Moreover, the problems, challenges, and opportunities have been identified that may guide scholars of the field. Moreover, instead of publishing all the experimental models, as was the case in the past, it has been recommended that workers in the area provide the most reliable prediction models so that they can be compared fairly in the future. Although multiple predictors of renal diseases are currently available in the field, the constantly growing range of datasets emphasizes the need to adopt new approaches and qualitative ML techniques for exploring information that can assist in chronic kidney disease.

REFERENCES

[1] Gurubani Gulati, Dr. Satinder Singh, "Modern Era and Security of Women: An Intellectual Device," International Research Journal of Engineering and Technology (IRJET), 2020.

[2] Ms. K. Opika, Ms. C. Shilpa Rao, "An Evolution of Women's Safety System: A literature review," An International Bilingual Peer-Reviewed Research Journal, 2020.

[3] B.Sindhu Bala, M.Swetha, M. Tamilasari, D.Vinodha, "Survey on women safety using IoT," International Journal of Computer Engineering in Research Trends, 2018.

[4] Agrima Agrawal, Ankita Maurya, Prof. Amruta Patil, "Voice-controlled tool for anytime safety of women," Journal of Emerging Technology and Innovative Tools, 2021.

[5] Navya R Sogi, Priya Chatarjee, UNethra, V Suma, "SMARISA: A Raspberry Pi-based Smart Ring for Women's Safety Using IoT".

[6] Wasim Akram, Mohit JainC. Sweetlin Hemalatha, "Design of a Smart Safety Device for Women using IoT," Procedia of Computer Science, 2019.

[7] B. Sathyasri, U. Jaishree Vidhya, G. V. K. Jothi Sree, T. Pratheeba, K. Ragapriya, "Design and Implementation of Women Safety based on IoT technology," International Journal of Recent Technology and Engineering (IJRTE, 2019.

[8] Raja, S. Pravinth; Rachel, S. Sheeba; R, Sapna, "Women's Safety with a Smart Foot Device," in 2021 4th International Conference on Computing and Communications Technologies (ICCCT), 2021.

[9] T. Sowmya, D. Treevani, D. Keerthana, A Vasantha Laxmi, "Women's safety system using IoT," International Research Journal of Engineering and Technology, 2020.

[10] Aishwarya, "IoT-Based Smart Security Gadget for Women's Safety," in 2019 1st International Conference on Advances in Information Technology (ICAIT), 2019.

[11] Sumanth Pagadala, Lakshmi Prasanna, Anusha Reddy, "A Novel ML- Supported IoT Device for Women Security," International Research Journal Of Engineering and Technology(IRJET), 2021.

[12] Muskan, Teena; Khandelwal, Manisha; Pandey, Purnendu Shekhar, "Women Safety Device Designed Using IoT and Machine Learning," in 2018 IEEE SmartWorld, Ubiquitous Intelligence & Computing, Advanced & Trusted Computing, Scalable Computing & Communications, Cloud & Big Data Computing, Internet of People and Smart City Innovation, 2018.

[13] A.Jesudoss, "Smart Solution for Women Safety Using IOT," International Journal of Pure and Applied Mathematics, 2019.

[14] Prottasha Ghosh, Emran Hasan, "Smart Security Device for Women Based on IoT Using Raspberry Pi," in 2nd International Conference on Robotics, Electrical and Signal Processing Techniques (ICREST), DHAKA, Bangladesh, 2021.

[15] Seth, Debojyoti; Chowdhury, Ahana; Ghosh, Shreya, "A Hidden Markov Model and Internet of Things Hybrid Based Smart Women Safety Device," in 2nd International Conference on Power, Energy and Environment: Towards Smart Technology (ICEPE), 2018.

[16] S. Srinivasan, P. Muthu Kannan & R. Kumar, "A Machine Learning Approach to Design and Develop a BEACON Device for Women's Safety," 2022.

[17] Manikumar, M. Murugan, "Guardian device for women - a survey and comparison study," in Second International Conference on Robotics, Intelligent Automation and Control Technologies (RIACT 2021), Chennai, India, 2021.

[18] Lauren F. Cardoso, Sara Landers, "Recent and emerging technologies: Implications for women's safety," Technology in Society, 2019.

[19] Sagarika Das, Baswakiran, Shweta Dasar, Jagadeesha, Shrinivas Rao, "women's security system," International Journal of Engineering Research & Technology (IJERT), 2021.

[20] Hyndavi, Rakesh, "Smart Wearable Device for Women's Safety Using IoT," in 5th International Conference on Communication and Electronics Systems (ICCES), 2020.

[21] Sohini Roy, "MoveFree: A ubiquitous system to provide women safety," in Third International Symposium on Women in Computing and Informatics, WCI 15At Kochi, India.

[22] Sri Raksha S, Yerrapadu Raghavendra Reddy, Ediga Indu Meghana, Kakanuru Manikanteswara Reddy, Prabhat Kumar Panda, "Design of a smart women safety band using IOT and machine learning," International Journal of Contemporary Architecture, 2021.

[23] Daniyal, M., Qureshi, M., Marzo, R. R., Aljuaid, M., & Shahid, D. (2024). Exploring clinical specialists' perspectives on the future role of AI: evaluating replacement perceptions, benefits, and drawbacks. BMC Health Services Research, 24(1), 587.

[24] Harikiran, G C; Menasinkai, Karthik; Shirol, Suhas, "Smart security solution for women based on Internet of Things(IOT)," in 2016 International Conference on Electrical, Electronics, and Optimization Techniques (ICEEOT), 2016.

[25] Saranya N, "A Smart friendly IoT device for women safety with GSM and GPS location tracking," in Proceedings of the Fifth International Conference on Electronics, Communication and Aerospace Technology, 2021.

[26] B.Ramesh Reddy, T.SowjanyaN.Bala, Subrahmanayam G. Mahantesh, Prudhvi, "IoT-based smart protective equipment for women," 2021.

[27] A. Bhate, Parveen Sultana H, "Smart wrist band for women's security using logistic regression technique," Blue Eyes Intelligence Engineering and Sciences Publication, 2019.

[28] Ruman, Saha, Saikat, "Safety Assistant And Harassment Prevention For Women," 2019.

[29] V. Kanchana Devi, Alluru Kavya, D. Shruthi, IoT-SDWD: Internet of Things-Based Security Device for Women in Danger, Smart Innovation, Systems and Technologies book series (SIST, volume 248), 2021.

[30] K. Venkatesh, S. Parthiban, P. S. Kumar, and C. V. Kumar, "IoT-based Unified approach for Women safety alert using GSM," in Third International Conference on Intelligent Communication Technologies and Virtual Mobile Networks (ICICV), 2021.

[31] Mona Chaware, Mrudula Nimbarte, "Smart Safety Gadgets for Women: A Survey," Journal of the University of Shanghai for Science and Technology, 2020.

[32] Nikhil Balu Nanekar, Gangaram Kauthale, "Woman Safety Device," International Advanced Research Journal in Science, Engineering and Technology, 2021.

[33] Dr. K. M. Anand Kumar, M Prakash R, "Smart Garb - A Wearable Safety Device for women," International Journal for Research in Applied Science & Engineering Technology (IJRASET, 2020.

[34] Sunehra, Dhiraj; Sreshta, V. Sai; Shashank, V.; Goud, B. Uday Kumar, "Raspberry Pi Based Smart Wearable Device for Women Safety using GPS and GSM Technology," in 2020 IEEE International Conference for Innovation in Technology (INOCON), 2020.

[35] Khan, A., Qureshi, M., Daniyal, M., & Tawiah, K. (2023). A novel study on machine learning algorithm-based cardiovascular disease prediction. Health & Social Care in the Community, 2023(1), 1406060.

[36] J. C. L. L. Y. L. C. F. C. &. C. B. Qin, "A machine learning methodology for diagnosing chronic kidney disease," 2019.

[37] Y. S. S. &. S. A. Amirgaliyev, "Analysis of chronic kidney disease dataset by applying machine learning methods," IEEE Access, 2018.

[38] Qureshi, M., Khan, S., Bantan, R. A., Daniyal, M., Elgarhy, M., Marzo, R. R., & Lin, Y. (2022). Modeling and forecasting monkeypox cases using stochastic models. Journal of Clinical Medicine, 11(21), 6555.

[39] P. C. Chittora, P., Chaurasia, S., Chakrabarti, P., Kuma, "Prediction of chronic kidney disease-a machine learning perspective," Vols. 17312-17334, no. 21, p. 9, 2021.

[40] K. A. & P. G. Padmanaban, "Applying machine learning techniques for predicting the risk of chronic kidney disease," no. IEEE Access, pp. 1-6, 2016.

[41] Iftikhar, H., Daniyal, M., Qureshi, M., Tawiah, K., Ansah, R. K., & Afriyie, J. K. (2023). A hybrid forecasting technique for infection and death from the mpox virus. Digital Health, 9, 20552076231204748.

[42] P. S. F. J. M. S. S. A. S. A. A. A. &. K. A. A. Ghosh, "Optimization of prediction method of chronic kidney disease using machine learning algorithm," in 15th International Joint Symposium on Artificial Intelligence and Natural Language Processing (iSAI-NLP), 2020.

[43] B. M. S. M. S. F. T. F. T. A. K. &. K. S. Gudeti, " A novel approach to predict chronic kidney disease using machine learning algorithms," in 4th International Conference on Electronics, Communication and Aerospace Technology (ICECA), 202.

[44] Hussain, I., Qureshi, M., Ismail, M., Iftikhar, H., Zywiołek, J., & López-Gonzales, J. L. (2024). Optimal features selection in the high dimensional data based on robust technique: Application to different health database. Heliyon, 10(17).

[45] N. &. M. S. Bhaskar, "A deep-learning-based system for automated sensing of chronic kidney disease," pp. 1-4, 2019.

[46] W. H. S. D. P. K. D. M. &. K. K. A. D. C. P. Gunarathne, "Performance evaluation on machine learning classification techniques for disease classification and forecasting through data analytics for chronic kidney disease (CKD)," in 17th International Conference on Bioinformatics and bioengineering (BIBE).

[47] Qureshi, M., Ishaq, K., Daniyal, M., Iftikhar, H., Rehman, M. Z., & Salar, S. A. (2025). Forecasting cardiovascular disease mortality using artificial neural networks in Sindh, Pakistan. BMC Public Health, 25(1), 34.

[48] I. U. &. H. D. Ekanayake, "Chronic kidney disease prediction using machine learning methods," in Moratuwa Engineering Research, 2020.

[49] Iftikhar, H., Hashem, A. F., Qureshi, M., & Rodrigues, P. C. (2025). Clinical application of machine learning models for early-stage chronic kidney disease detection. Diagnostics, 15(20), 2610.

[50] R. A. A. K. A. A. N. Y. A. N. S. A. M. F. O. S. O. .. &. A. N. Alassaf, "Preemptive diagnosis of chronic kidney disease using machine learning techniques," in International conference on innovations in information technology (IIT), 2018.

[51] Q. S. C. T. W. &. L. Y. Bai, "Machine learning to predict end-stage kidney disease in chronic kidney disease," Springer, 2022.

[52] K. M. & Z. H. M. Zubair Hasan, "Performance evaluation of ensemble-based machine learning techniques for

prediction of chronic kidney disease. In Emerging Research in Computing, Information, Communication and Applications, in ERCICA, 2018.

[53] F. F. C. G. D. G. D. H.-B. T. &. C. A. Sanmarchi, "Predict, diagnose, and treat chronic kidney disease with machine learning: a systematic literature review," pp. 1-17, 2023.

[54] A. S. A. S. R. A. M. &. M. A. N. Abdelaziz, "A machine learning model for predicting chronic kidney disease based on Internet of Things and cloud computing in smart cities. Security in smart cities: models, applications, and challenges.," pp. 93-114, 2019.

[55] M. A.-J. D. H. A. L. P. & A. A. J. Alloghani, "Performance-based prediction of chronic kidney disease using machine learning for high-risk cardiovascular disease patients. Nature-inspired computation in data mining and machine learning," pp. 187-206, 2020.

[56] A. A. E. & K. J. Subasi, "Diagnosis of chronic kidney disease by using random forest. In CMBEBIH, in the International Conference on Medical and Biological Engineering, 2017.

[57] J. K. M. K. R. M. L. J. E. L. V. B. .. &. T. G. E. Weaver, "Deep learning imaging features derived from kidney ultrasounds predict chronic kidney disease progression in children with posterior urethral valves. Pediatric Nephrology, pp. 1-8, 2022.

[58] A. S. P. K. &. P. M. Motwani, "Novel machine learning model with wrapper-based dimensionality reduction for predicting chronic kidney disease risk," in Soft Computing and Signal Processing: Proceedings of 3rd ICSCSP, Singapore, 2021.

[59] D. A. &. S. T. M. Debal, "Chronic kidney disease prediction using machine learning techniques," vol. 9(I), pp. 1-19, 2022.

[60] L. P. B. N. S. M. S. &. T. S. Jena, "Risk prediction of kidney disease using machine learning strategies," in Intelligent and Cloud Computing: Proceedings of ICICC 2019, Singapore, 2021.

[61] A. S. G. P. B. T. S. G. D. &. K. A. Khamparia, "Chronic kidney disease classification with multimedia data learning using a deep stacked autoencoder network. Multimedia Tools and Applications," p. 79, 2020.

[62] S. K. M. B. K. L. S. E. L. I. H. K. Y. A. .. &. P. J. T. Lee, "Machine Learning–Aided Chronic Kidney Disease Diagnosis Based on Ultrasound Imaging Integrated with Computer-Extracted Measurable Features," Digital Imaging, vol. 35(5), 2022.

[63] H. P. B. &. R. S. Kriplani, "Prediction of chronic kidney diseases using deep artificial neural network technique. In Computer-aided Intervention and diagnostics in clinical and medical images, Springer International Publishing, 2019.

[64] L. N. S. & S. R. Jena, "Chronic disease risk (CDR) prediction in biomedical data using machine learning approach," in Computing and Communication: Proceedings of ICAC 2019, Singapore, 2020.

[65] H. D. M. H. & C. A. Polat, "Diagnosis of chronic kidney disease based on support vector machine by feature selection methods," vol. 41, pp. 1-11, 2017.

[66] B. H. A. & M. H. Boukenze, "Predicting chronic kidney failure disease using data mining techniques," in Ubiquitous Networking 2: Proceedings of the UNet'16 2, Singapore, 2017.

[67] S. Pal, "Chronic Kidney Disease Prediction Using Machine Learning Techniques. Biomedical Materials & Devices, pp. 1-7, 2022.

[68] V. &. I. M. Shanmugarajeshwari, " IoT-Based Prediction of Chronic Kidney Disease Using Python and R Based on Machine and Deep Learning Algorithms,"

in Next Generation of Internet of Things: Proceedings of ICNGIoT, Singapore, 2022.

[69] S. A. B. & F. Y. Sossi Alaoui, " Statistical and predictive analytics of chronic kidney disease," in Advanced Intelligent Systems for Sustainable Development (AI2SD'2018) Vol 4: Advanced Intelligent Systems Applied to Health, 2019.

[70] M. R. A. M. A. A. N. &. B. T. Ahmed, " Computational intelligence approaches for prediction of chronic kidney disease," in In Advances in Distributed Computing and Machine Learning: Proceedings of ICADCML 2021, Singapore, 2022.

[71] F. P. A. V. W. A. D. I. &. D. N. T. Schena, "Prediction of chronic kidney disease and its progression by artificial intelligence algorithms," vol. 35(8), 2022.

[72] S. &. C. B. De, "Development of chronic kidney disease prediction system (CKDPS) using machine learning technique. In Intelligent Data Communication Technologies and Internet of Things, in ICICI 2019, 2020.

[73] A. K. A. S. S. &. P. C. Sarkar, "Detection and Evaluation of Chronic Kidney Disease Using Different Regression and Classification Algorithms in Machine Learning. In Advances in Electronics, Communication and Computing." In Proceedings of ETAEERE 2020, Singapore, 2021.

[74] Z. Z. X. & Z. Z. Chen, "Clinical risk assessment of patients with chronic kidney disease by using clinical data and multivariate models," International Urology and Nephrology, 2016.

[75] J. R. & P. E. Lambert, " Oppositional firefly optimization-based optimal feature selection in chronic kidney disease classification using a deep neural network," vol. 13(4), 2022.

[76] I. G. M. K. & H. G. M. Saha, "Classification System for Prediction of Chronic Kidney Disease Using Data Mining Techniques," in Advances in Data and Information Sciences: Proceedings of ICDIS 2021, Singapore, 2022.

[77] V. &. S. P. Sharma, "Machine Learning-Based Algorithms for Prediction of Chronic Kidney Disease: A Review. Advances in Data and Information Sciences.".

[78] M. S. L. F. N. S. F. P. C. P. Granal, "Prediction Tool to Estimate Potassium Diet in Chronic Kidney Disease Patients Developed Using a Machine Learning Tool: The UniverSel Study. Nutrients," vol. 12(12), 2022.

[79] Y. T. L. C. H. C. K. S. C. C. P. &. P. T. W. Liao, "Data augmentation based on generative adversarial networks to improve stage classification of chronic kidney disease," vol. 12(1), p. 352, 2021.

[80] D. M. U. B. A. P. H. P. K. M. D. Swain, " A Robust Chronic Kidney Disease Classifier Using Machine Learning," vol. 12(1), p. 212, 2023.

[81] H. Y. C. C. C. C. C. F. C. Y. C. C. C. &. T. Y. L. Kao, "Associations between Sex and Risk Factors for Predicting Chronic Kidney Disease," vol. 19(3), 2022.

[82] M. H. L. H. H. H. Y. C. L. T. S. C. M. &. F. Y. W. Tsai, "Hazardous Effect of Low-Dose Aspirin in Patients with Predialysis Advanced Chronic Kidney Disease Assessed by Machine Learning Method Feature Selection." vol. 9, p. 1484, 2021.

[83] C. S. F. J. M. H. M. R. A. S. G. P. T. Z. .. &. I. S. M. Mondol, "Early Prediction of Chronic Kidney Disease: A Comprehensive Performance Analysis of Deep Learning Models." vol. 9, p. 308, 2022.

[84] A. W. B. A. N. B. K. H. C. Y. &. C. K. C. Jallow, "Machine Learning Approach for

Chronic Kidney Disease Risk Prediction Combining Conventional Risk Factors and Novel Metabolic Indices. Applied Sciences," vol. 12, p. 12001, 2022.

[85] C. T. C. Y. P. K. Y. T. &. L. C. M. Su, "Machine Learning Models for the Prediction of Renal Failure in Chronic Kidney Disease: A Retrospective Cohort Study. Diagnostics," vol. 12, p. 2454, 2022.