iJRASET
International Journal For Research in
Applied Science and Engineering Technology

# INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

**Volume:** 12 **Issue:** V **Month of publication:** May 2024

**DOI:** https://doi.org/10.22214/ijraset.2024.62002

www.ijraset.com

Call: 08813907089 | E-mail ID: ijraset@gmail.com

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 12 Issue V May 2024- Available at www.ijraset.com

# Chronic Kidney Disease Prediction by using Naive Bayes

Dhanush M[1], Hency Raj[2], Pratik Bothra[3], Raman Zanwar[4], Dr. S.Nagraj[5]
*Student at Computer science and Business Systems, Jain (Deemed-to-be University), Bangalore - 562112, Karnataka, India.*

***Abstract: Chronic kidney disease CKD is a chronic kidney problem that affects the human kidneys and causes it to not work properly or causes complete kidney failure, leads to dialysis or causes other related diseases and reduces the quality of life symptoms of this disease cannot be identified in the preliminary stage, only very few people are aware of this disease and can predict symptoms at an early stage, an earlier CKD predictor model should be available improved with higher prediction accuracy and precision, hence the need for a decision support system that helps nephrologists in times of emergency therefore, in this research, a naive Bayesian classifier is used for classification along with Hierarchy based selection nb cb h nb classifier works efficiently with huge datasets and reduces computational complexity speed of prediction and disease severity analysis with nbare extremely higher***
***Keywords: Chronic, Kidney, Disease, naïve bayes, CKD, computational, complexity speed, naïve bayes classifiers, data, machine learning.***

## I. INTRODUCTION

Chronic kidney disease (CKD) is a prevalent medical condition that affects millions of people Early detection and accurate prediction of CKD worldwide can significantly improve the patient's condition results and guide for appropriate treatment plans in this study we propose the use of naïve a Bayesian algorithm for CKD prediction using its simplicity and efficiency categorical data processing naive bayes algorithm is a probabilistic classification method that
assumes independence between features, calculates the posterior probability of each class employing the Bayes theorem and the input features, chooses the class with the highest probability as a predicted label, this algorithm is particularly suitable for CKD prediction because it can discrete features commonly associated with this disease such as age blood pressure glucose levels and albumin levels, we used a labeled data set to implement the CKD prediction model consisting of patient records including various clinical and laboratory data set measurements was split into training and test sets, with the training set used to estimate the probabilities requires a naive Bayesian algorithm, our experimental results show promising accuracy and confidence in CKD prediction using a naive Bayesian CKD prediction model algorithm developed in this study has the potential to be integrated into clinical decision support systems assistance to healthcare professionals in the early identification and intervention of patients at risk CKD's simplicity and efficiency as a naive bayes algorithm make it a practical choice real time prediction application

.

## II. METHODOLOGY

1) Data Collection: The first step is to collect relevant patient data, including medical records, laboratory test results and demographic information. This data forms the basis for training and testing the prediction system.
2) Data preprocessing: Before deploying the system, the collected data needs to be cleaned, adjusted and put into a format that is appropriate for analysis. This ensures that the forecast the system can efficiently process and interpret data.
3) Model development: The prediction system relies on machine learning algorithms develop a predictive model. This includes training the model using the collected data, selection appropriate functions and optimization of model performance.
4) Validation and Evaluation: It is important to validate and evaluate the prediction system in advance development. This includes assessment of accuracy, precision, recall and other performance metrics to ensure system reliability and efficiency.
5) Integration with healthcare systems: The prediction system should be seamless integrated into existing healthcare systems to enable easy access and use in healthcare professionals. This integration may include collaboration with IT departments and compliance relevant regulations on personal data protection and security.
6) User interface and accessibility: A user friendly interface should be developed to enable this it allows healthcare professionals to easily interact with the prediction system. This interface should provide clear and interpretable predictions together with any supporting information or recommendation.

7) Deployment and Monitoring: Once the prediction system is ready, it can be deployed healthcare settings. Continuous monitoring is essential to ensure its continued performance,Determine any problems, then update or enhance as needed.
8) Collaboration and feedback: Collaboration with healthcare professionals is key throughout the deployment process. Their insights and feedback can help improve the system,improve predictions and address potential usability issues.
9) Ethical considerations: It is important to consider ethical aspects such as patient consent, privacy and responsible use of data. Ensures compliance with ethical guidelines and regulations that the prediction system is deployed in an accountable and transparent manner.
10) Continuous improvement: The prediction system should be continuously evaluated and improved based on actual performance and feedback from healthcare professionals. This an iterative process helps improve the accuracy and usability of the system over time.

By following these steps and considering these factors, chronic kidney disease can be predicted the system can be successfully deployed to assist healthcare professionals in obtaining information decisions and improving patient outcomes.

## III. ARCHITECTURE

The architecture design of a chronic kidney disease prediction system typically includes different components working together usually includes a function of preprocessing data collection extraction algorithms and machine learning algorithms, the first data is collected from various sources, e.g. patient medical records and demographic information into which this data is pre-processed cleaned and converted into a suitable format for analysis other relevant features are extracted from pre-processed data, these characteristics may include factors such as age, blood glucose, blood pressure levels and other medical indicators that can help predict chronic kidney disease. Once the features are extracted, machine learning algorithms are applied to train a predictive model, this model learns patterns and relationships in the data it is supposed to generate predictions about the probability of chronic kidney disease in a given patient architecture the design also includes evaluation and validation of the model to ensure its accuracy and reliability may include techniques such as cross-validation and performance metrics to assess prediction system capabilities overall CKD prediction architecture design .The goal of the system is to use data and machine learning techniques to provide accurate predictions and assist healthcare workers in the early detection and treatment of disease

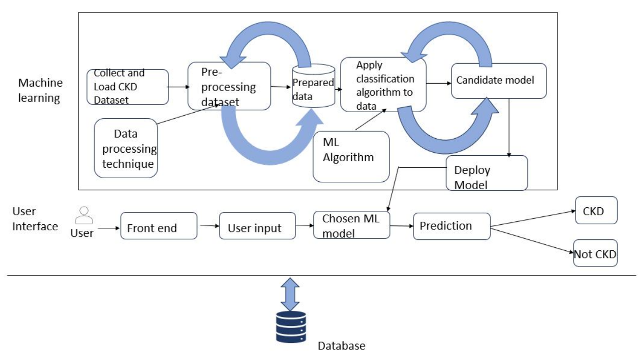



Figure. 1 Architecture design of the chronic kidney disease prediction system

## IV. PROJECT OBJECTIVES

1) In a project focused on chronic kidney disease using Naive Bayes, the objectives would revolve around developing a classifier that can accurately predict the presence or absence of the disease based on certain features. The first objective would be to build the Naive Bayes classifier itself. This involves training the model using a dataset that contains information about patients with and without chronic kidney disease, along with their corresponding features like age, blood pressure, and blood tests.
2) Once the classifier is built, the next objective would be to evaluate its performance. This can be done by using metrics such as accuracy, precision, recall, and F1 score. These metrics provide insights into how well the classifier is able to correctly classify instances of chronic kidney disease.
3) To further improve the classifier's performance, the project could involve exploring different feature selection techniques. This means identifying the most relevant features that have the greatest impact on the prediction accuracy. By selecting the right set of features, the classifier can become more efficient and accurate in identifying chronic kidney disease.

4) Another objective could be to investigate the impact of different hyperparameter settings on the classifier's performance. Hyperparameters are parameters that are set before the learning process begins, and they can significantly affect the model's performance. By experimenting with different hyperparameter values, the project aims to find the optimal configuration that maximizes the classifier's accuracy.
5) Lastly, it would be interesting to compare the performance of Naive Bayes with other classification algorithms for predicting chronic kidney disease. This comparison helps determine whether Naive Bayes is the most suitable algorithm for this particular task or if other algorithms, such as decision trees or support vector machines, might provide better results.
6) Overall, the objectives of a project on chronic kidney disease using Naive Bayes would involve building an accurate classifier, evaluating its performance, exploring feature selection techniques, investigating hyperparameter settings, and comparing its performance with other algorithms. By achieving these objectives, the project aims to contribute to the development of effective and reliable methods for diagnosing chronic kidney disease.

## V. OVERVIEW

1) Accuracy: A good prediction system strives for high accuracy in identifying probability chronic kidney disease in patients. Its goal is to minimize false positives and false negatives provide reliable predictions.
2) Data-driven: A predictive system relies on data such as patient records and medical tests results, make accurate predictions. The more complex and diverse the data, the better system performance.
3) Machine Learning: The system uses machine learning algorithms to analyse and learn from the data. These algorithms identify patterns, correlations, and features that contribute to this prediction of chronic kidney disease .
4) Personalized: Every patient is unique and a good prediction system takes that into account individual characteristics such as age, gender, medical history and other relevant factors provide personalized predictions.
5) Early detection: One of the key goals of the chronic kidney disease prediction system is identify the disease at an early stage. Early detection allows early intervention and management to potentially slow disease progression.
6) Integration with healthcare systems: The prediction system is designed to seamlessly integrate with existing healthcare systems and allow healthcare professionals to access a use predictions in your clinical decision-making process.
7) Continuous improvement: A good forecasting system is dynamic and continuous improves over time. Adapts to new data, includes feedback from healthcare professionals, and updates its algorithms to improve its predictive capabilities.

## VI. TYPES OF CHRONIC KIDEY DISEASE

Chronic kidney disease (CKD) encompasses a range of conditions that affect the kidneys and their ability to function properly over an extended period. Let's explore some of the common types:

1) Diabetic nephropathy: This type of CKD occurs as a result of long-term, uncontrolled diabetes. High blood sugar levels can damage the tiny blood arteries in the kidneys, leading to reduced kidney function. It is a significant complication of diabetes and one of the leading causes of CKD worldwide.
2) Hypertensive nephropathy: High blood pressure, when left untreated or poorly controlled, can cause damage to the blood vessels and filters in the kidneys. Chronic renal disease may result from this over time. Managing blood pressure levels is crucial in preventing or slowing down the progression of this type of CKD.
3) Glomerulonephritis: Glomerulonephritis refers to inflammation of the glomeruli, which are the tiny filters within the kidneys. Infections, autoimmune illnesses, and other underlying problems may be the cause. The inflammation impairs the kidneys' ability to filter waste products and excess fluid from the blood, resulting in CKD.
4) Polycystic kidney disease (PKD): PKD is a genetic disorder characterized by the growth of multiple cysts in the kidneys. These cysts gradually enlarge and impair kidney function over time. PKD can be inherited from one or both parents, and its progression varies from person to person.
5) Interstitial nephritis: This type of CKD involves inflammation of the kidney tubules and the surrounding tissue. It can be caused by certain medications, infections, autoimmune disorders, or exposure to toxins. Prompt identification and treatment of the underlying cause are crucial in managing interstitial nephritis.
6) Obstructive nephropathy: When there is a blockage in the urinary tract, such as a kidney stone or an enlarged prostate, it can lead to kidney damage over time. The obstruction prevents the normal flow of urine, causing pressure to build up and affect kidney function.

It's important to note that these are just a few examples of the types of chronic kidney disease. Other less common types, such as inherited conditions or certain autoimmune disorders, can also contribute to CKD. Each type of CKD may require different approaches to treatment and management, including lifestyle changes, medications, and, in some cases, dialysis or kidney transplantation.

## VII. CHALLENGES AND GAPS

1) Limited feature set: Some existing systems may have a limited feature set for the prediction of CKD. This can potentially miss important risk factors or biomarkers which could contribute to more accurate predictions.
2) Missing advanced feature selection: Some systems may not use advanced feature selection techniques to identify the most important features for CKD prediction.This can result in sub-optimal performance and potentially include irrelevant or redundant features, leading to reduced accuracy.
3) Class Imbalance: Imbalanced data sets where the number of CKD instances markedly different from non-CKD cases, may pose problems. Existing systems may not adequately address issues of class imbalance, which may lead to bias estimates and the minority class's lesser accuracy.
4) Limited dataset size: The size and diversity of the dataset used for training and testing may affect the performance of the Naive Bayes classifier. Smaller or less diverse datasets may not capture the full range of CKD patterns and characteristics, leading to reduced accuracy and generalizability.
5) Healthcare Workflow Integration: Seamless integration of the prediction system into existing healthcare workflows can be a challenge. Requires coordination with healthcare workers, adapting to their routines and ensuring that the system does it won't disrupt their workflow or add additional burden.
6) Continuous updates and maintenance: The prediction system must be regularly updated and maintained to keep pace with advances in medical knowledge and evolving patient population. This requires constant monitoring, retraining of models, and incorporating new data to ensure system accuracy and relevance.
7) Ethical considerations: Ethical challenges include ensuring the privacy of patients and data safety and informed consent. Patient confidentiality protection and compliance ethical guidelines in the use of sensitive health data are essential to maintain trust and ensure responsible use of the prediction system. Solving these challenges and filling identified gaps is key to success implementation and use of the chronic kidney disease prediction system in healthcare Settings. It requires collaboration between healthcare professionals, data scientists and policy makers to overcome these problems and ensure the effectiveness of the system and ethical use.

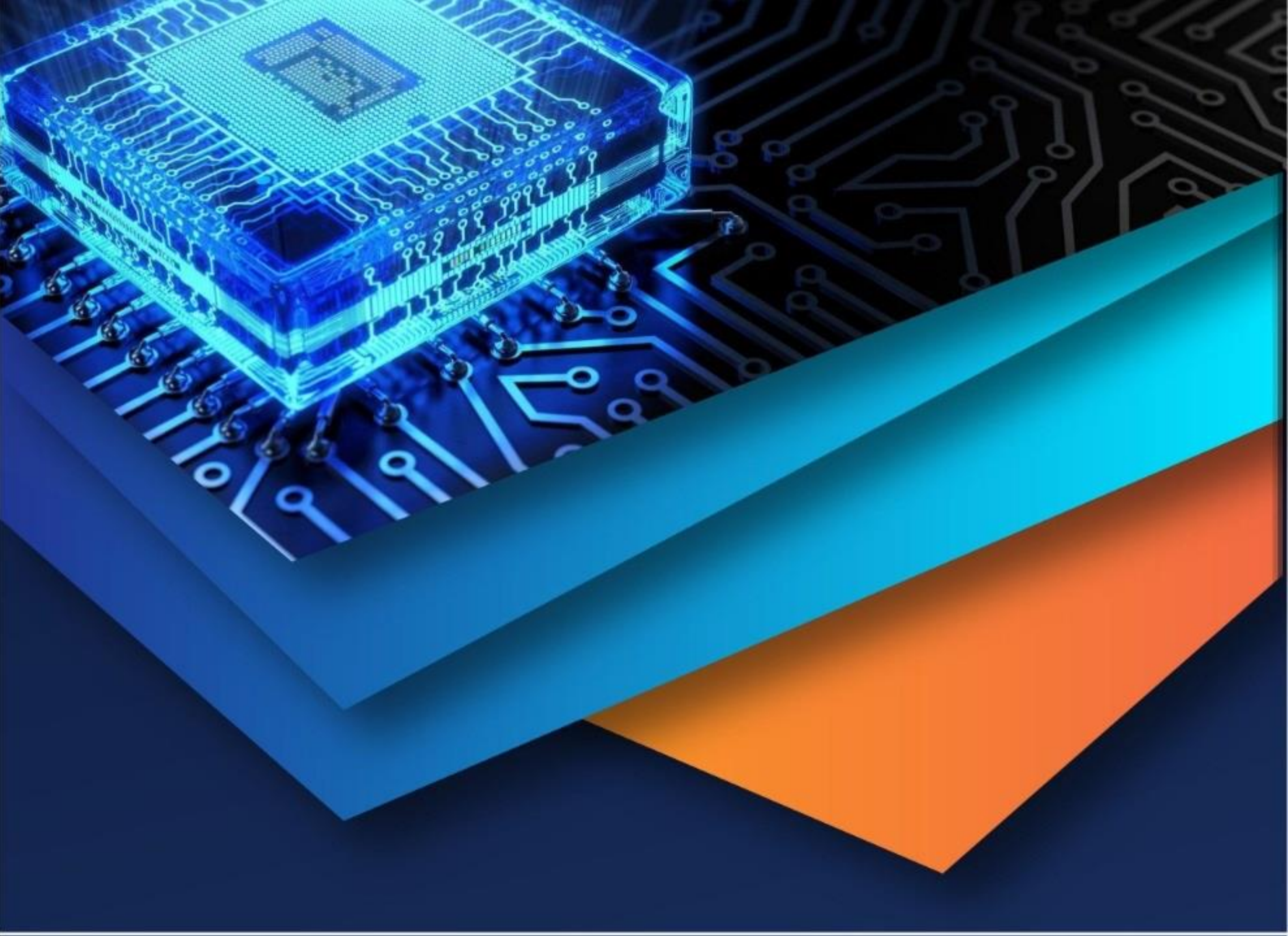
## VIII. CONCLUSION

In conclusion, the use of the Naive Bayes algorithm for chronic kidney disease (CKD) prediction shows promising results and offers several advantages in the field of healthcare. According to using a comprehensive set of clinical and laboratory measurements as functions, uses advanced feature selection techniques and potentially including ensemble classification Naive Bayes classifier can provide accurate and efficient predictions for CKD. Simplicity, efficiency and ability to process categorical data of the Naive Bayes algorithm is a suitable choice for predicting CKD. It can normally handle discrete functions efficiently associated with CKD risk factors, allowing healthcare professionals to identify at-risk patients and initiate early interventions. Additionally, the probabilistic nature of Naïve Bayes the classifier provides interpretability and allows health professionals to understand the factors contributing to the prediction of CKD. But it's crucial to understand the constraints, existing systems, such as limited feature sets, potential class imbalance issues, and the need for further validation on different datasets. Addressing these limitations through a comprehensive solution feature selection technique, dealing with class imbalance, and exploring file classification methods can increase the performance and reliability of CKD prediction systems based on Naive Bayes Algorithm.

## REFERENCES

[1] Imesh Ekanayake & Damayanthi Herath. (2020). Chronic kidney disease prediction using machine learning methods, In 2020 Moratuwa Engineering Research Conference (MERCon).
[2] Pronab Ghosh, et al. (2020). Optimization of prediction method of chronic kidney disease using machine learning algorithm, In 2020 15th International Joint Symposium on Artificial Intelligence and Natural Language Processing (iSAI-NLP)
[3] Selvi Sampath. (2021). Prediction of early symptoms for chronic kidney disease by ensembled classifiers, Cultural Studies Critical Methodologies, 12(4), 2321-2335K. Elissa, "Title of paper if known," unpublished.

[4] M. Islam, et al. (2020). Machine learning based image classification of papaya disease recognition, In 2020 4th International Conference on Electronics, Communication and Aerospace Technology (ICECA).
[5] Nikhila. (2021). Chronic kidney disease prediction using machine learning ensemble algorithm, In 2021 International Conference on Computing, Communication, and Intelligent Systems (ICCCIS)M. Young, The Technical Writer's Handbook. Mill Valley, CA: University Science, 1989.
[6] Rayan Alanazi. (2022). Identification and prediction of chronic diseases using machine learning approach, Journal of healthcare engineering
[7] Vijayalakshmi & V. Sumalatha. (2020). Survey on diagnosis of chronic kidney disease using machine learning algorithms , In 2020 3rd International Conference on Intelligent Sustainable Systems (ICISS)
[8] Akash Maurya, et al. (2019). Chronic kidney disease prediction and recommendation of suitable diet plan by using machine learning In 2019 International Conference on Nascent Technologies in Engineering (ICNTE)




10.22214/IJRASET

INDEX COPERNICUS 45.98

IMPACT FACTOR: 7.129

TOGETHER WE REACH THE GOAL

IMPACT FACTOR: 7.429


# INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY