Cosponsored by

Paper id: SMART 0064

Title of paper: Analysis of Chronic Kidney Disease using Machine Learning Techniques

Name of authors

1.T Pavani

2. D Murali Krishna

3. K. Ajita Lakshmi

4.G.N.S.Lakshmi

5.G.Pooja Hitha

6.G.Akshitha

7. Ch.Tejaswini

Shri Vishnu Engineering College, Bhimavaram

Introduction:

- Chronic Kidney Disease (CKD) is an important cause of morbidity and morality worldwide. It stands in 12th place in India for its mortality rate and in 3rd place worldwide.
- CKD treatment can be expensive, especially in advanced stages. Machine learning can help in cost-effective screening and management by reducing unnecessary tests and hospitalizations.
- Timely diagnosis and intervention can slow down the progression of CKD and reduce mortality rates. Machine learning plays a pivotal role in identifying high-risk patients.
- Machine learning models can analyze patient data to detect CKD at an early stage when symptoms may be subtle. This early intervention can significantly improve the quality of life for patients.

Literature:

Authors	Title	Findings
Dibaba Adeba Debal and Tilahun Melak Sitote	Chronic Kidney Disease Prediction Using Machine Learning Techniques	XGBoost - 82.6% Decision Tree - 77.5% Random Forest - 78.3% SVM - 78.78%
K.R.Anantha Padmanaban and G.Parthiban	Applying Machine Learning Techniques For Predicting The Risk Of Chronic Kidney Disease	Naïve Bayes - 86% Decision Tree - 91%
Saurabh Pal	Chronic Kidney Disease Prediction Using Machine Learning Techniques	Decision Tree - 95.92%
Virginia A. Dines and Vesna D. Garovic	Menopause and chronic kidney disease	CKD mostly found in women/girls compared to men/boys.

Aim:

- To preprocess the selected dataset.
- To design, develop, and train machine learning models using CKD datasets for the detection of Chronic Kidney Disease.
- To conduct a comparative analysis of the performance metrics (such as accuracy, sensitivity, selectivity, precision, and F1-score) for machine learning models in the detection of Chronic Kidney Disease.
- To analyze the importance of feature engineering and the use of feature ranking algorithms in identifying the effective features for detection of CKD.

Abstract:

- Chronic kidney disease (CKD)is a serious global health concern that has a high rate of morbidity and mortality and can cause other disorders.
- The early stages of CKD are characterized by a lack of evident symptoms, which might make patients unaware of the disease.
- Patients with early-stage chronic kidney disease (CKD) can benefit from early therapy, which can also slow the disease's progression.
- Therefore, having an effective model is crucial for CKD early diagnosis. Because machine learning models execute identification tasks quickly and accurately, they can help therapists accomplish this goal.
- Using machine learning techniques such as Support Vector Machines, Decision Trees, Neural Networks, and KNN for the detection of chronic kidney disease is the goal of this proposal.
- Feature Engineering algorithms such as ANOVA, MRMR, and CHI2 algorithms are used to determine the features for predicting CKD.

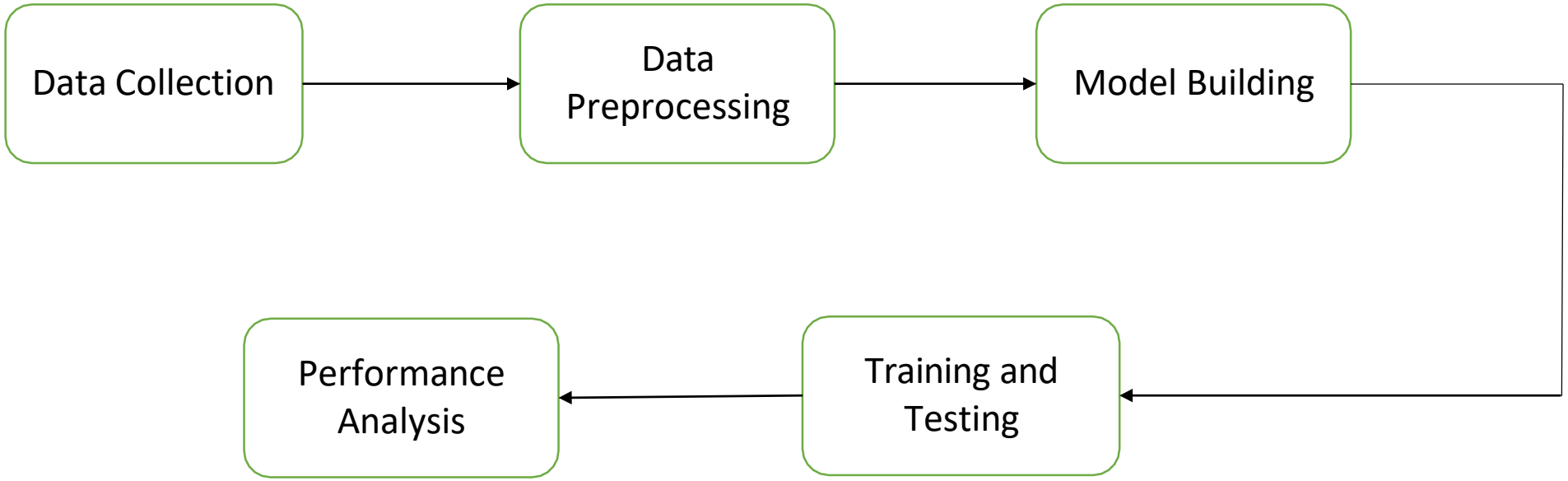
Existing models/design:

- Dibaba Adeba Debal and Tilahun Melak Sitote used RF, DT, and SVM algorithms to predict CKD. Out of all three algorithms, Random Forest has given the highest accuracy with 99.7% followed by Decision Tree with 98.5% accuracy and Support Vector Machine with 96.9% accuracy.
- K. R. Anantha Padmanaban and G. Parthiban has predicted CKD using machine-learning methods like DT and NB. NN and data clustering have also been utilized to improve prediction accuracy. With decision tree classification, they achieved the highest accuracy up to 91%, and with the Naïve Bayes algorithm, 86%. They used the Clinic Foundation Heart Disease Dataset for classifying the risk of CKD.

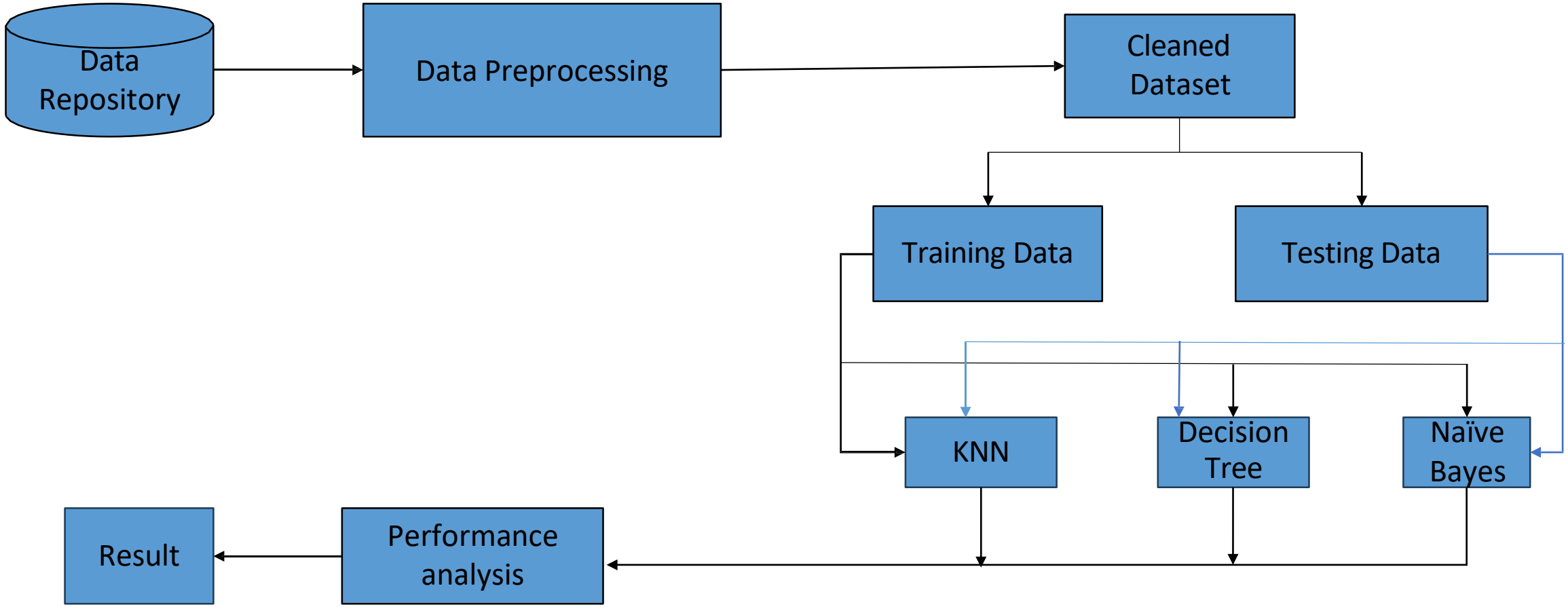
Proposed model/design:

- The aim of this project is to develop a machine learning model for symptoms-based disease prediction. The goal is to create an efficient and accurate system that can predict diseases based on a set of symptoms reported by patients.
- Traditionally, diagnosing diseases has heavily relied on the expertise and experience of medical professionals. However, this process can be time-consuming and may lead to errors or delays in diagnosis. By leveraging machine learning techniques, we can automate the disease prediction process and potentially improve the accuracy and efficiency of diagnoses.
- Development of chronic kidney analyzer using machine learning techniques helps to solve the above problem.

Flow chart/algorithm:



Block diagram/Design flow:



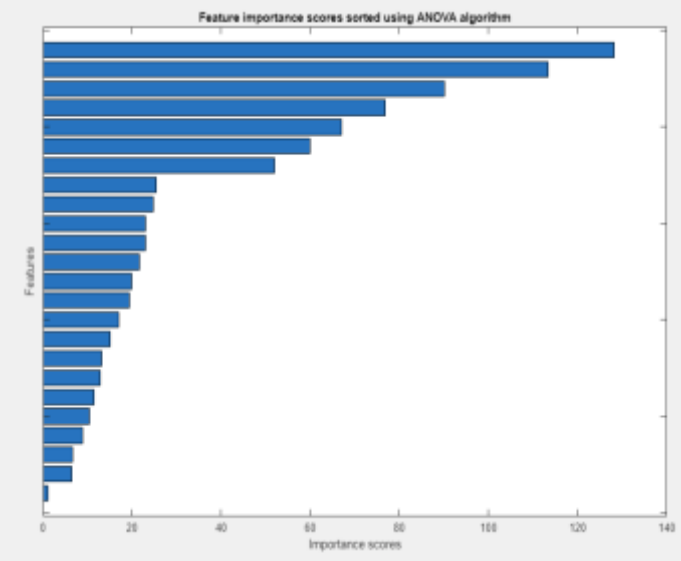
Novelty of Research :

- This research uniquely combines KNN, Decision Trees, SVM, and Neural Networks to optimize CKD classification, enhancing predictive accuracy and robustness
- Feature selection methods such as Chi2, MRMR, and ANOVA algorithms are employed to improve model interpretability and performance, providing deeper insights into key predictors of CKD.
- The research explores and applies state-of-the-art machine learning algorithms and architectures, potentially leading to novel insights and advancements in CKD classifications.

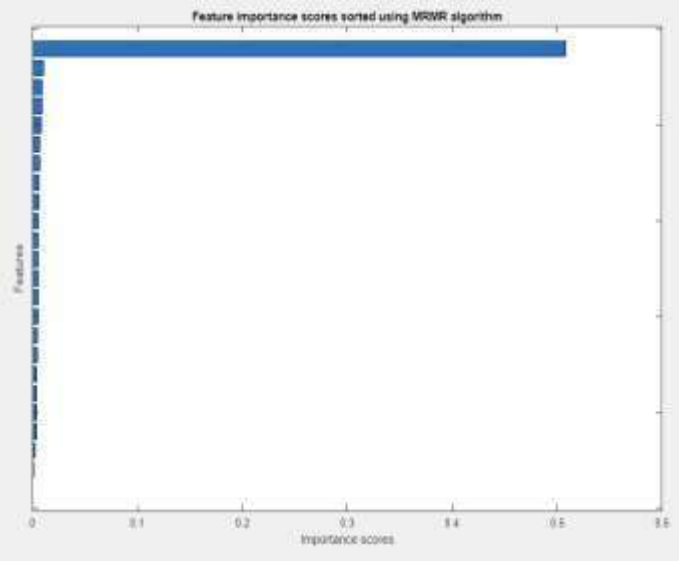
Results(1):

Algorithm	Training Accuracy	Testing Accuracy	Sensitivity	Selectivity	Precision	F1-Score
Fine KNN	98.9	99.2	97.8	100	100	98.8
Medium KNN	95	97.5	93.75	100	100	96.77
Coarse KNN	78.2	80	65.2	100	100	78.9
Fine Tree	98.6	99.2	100	98.68	97.77	98.87
Medium Tree	98.6	99.2	100	98.68	97.77	98.87
Linear SVM	98.6	100	100	100	100	100
Quadratic SVM	98.2	100	100	100	100	100
Cubic SVM	98.6	100	100	100	100	100
Linear NN	98.6	100	100	100	100	100
Cubic NN	98.6	100	100	100	100	100

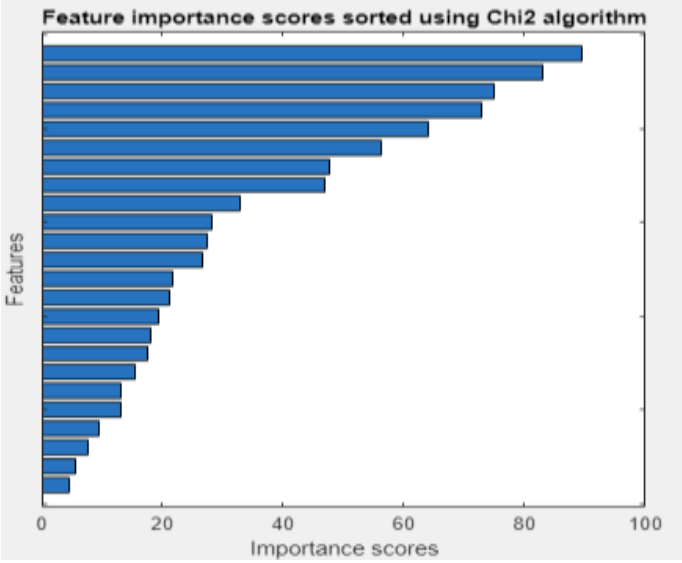
Results(2):



Feature importance using ANOVA algoritm



Feature importance using MRMR algoritm



Feature importance using CHI2 algoritm

Applications:

- Early diagnosis and risk stratification
- Personalized treatment plans
- Clinical decision support systems
- Population health management
- Predictive monitoring
- Research and drug development

Advantages:

- Enhanced Diagnostic Accuracy
- Comprehensive Feature Analysis
- Potential for Real-World Application

Limitations:

- Data Quality and Availability
- Computational Resources
- Model Interpretability

Conclusion:

We had done analysis of our model by using 3 different training testing ratios and with 4 different algorithms. By comparing different ratios, we got better accuracy for 70:30 ratios. We got better accuracy for Neural Networks algorithm compared to KNN, Decision Tree, and Support Vector Machine algorithms.3 feature selection algorithms such as CHI2, MRMR, and ANOVA are used to decrease the number of features and they are analyzed using the 4 algorithms with 3 different training testing ratios. Up to 18 features, the accuracy for the models is similar while using CHI2 feature selection algorithm. Upto 20 features, the accuracy for the models is similar while using MRMR feature selection algorithm. Up to 22 features, the accuracy for the models is similar while using ANOVA feature selection algorithm.

References

1. K. R. Anantha Padmanaban and G. Parthiban, “Applying Machine Learning Techniques for Predicting the Risk of Chronic Kidney Disease” Indian Journal of Science and Technology, Volume 09, Issue 29, doi: 10.17485/ijst/2016/v9i29/93880, August 2016 ISSN (Print): 0974-6846 ISSN (Online): 0974-5645.
2. Dibaba Adeba Debal and Tilahun Melak Sitote, “Chronic Kidney Disease Prediction Using Machine Learning Techniques”, Debal and Sitote Journal of Big Data-Springer, Debal and Sitote Journal of Big Data, doi: https://doi.org/10.1186/s40537-022-00657-5.
3. Sathiya Priya S and Suresh Kumar M, “Chronic Kidney Disease Prediction Using Machine Learning”, International Journal of Computer Science and Information Security (IJCSIS), Volume 16, Issue 04, April 2018.
4. Dr. Vijayaprabakaran. K, Pratheek Reddy. P, Puthin Kumar Reddy.T, Munnaf. K, Reddi Prasad. G, “Chronic Kidney Disease Diagnosis Using Machine Learning”, International Research Journal of Engineering and Technology (IRJET), Volume: 08, Issue: 06 | June 2021, e-ISSN: 2395-0056, p-ISSN: 2395-0072.
5. Virginia A. Dines & Vesna D. Garovic, “Menopause and chronic kidney disease”, doi: https://doi.org/10.1038/s41581-023-00717-w.
6. Gazi Mohammed Ifraz, Muhammad Hasnath Rashid, “Comparative Analysis for Prediction of Kidney Disease Using Intelligent Machine Learning Methods”, Hindawi,Computational and Mathematical Methods in Medicine, Volume 2021, Article ID 6141470, 10 pages, doi :https://doi.org/10.1155/2021/6141470.
7. Asif Salekin, John Stankovic, “Detection of Chronic Kidney Disease and Selecting Important Predictive Attributes”, International Conference on Healthcare Informatics (IEEE), October 2016, doi :10.1109/ICHI.2016.36.
8. Anant Agrawal, Harshit Agrawal, “Disease Prediction Using Machine Learning”, International Conference on Internet of Things and Connected Technologies (ICIoTCT),2018, ISSN: 1556-5068, HTTPS://WWW.SSRN.COM/LINK/3RD-ICIOTCT-2018.HTML.
9. Saurabh Pal, “Chronic Kidney Disease Prediction Using Machine Learning Techniques”, doi: https://doi.org/10.1007/s44174-022-00027-y.
10. Rameshchandra, Kothapalli, “Enhancing Kidney Stone Diagnosis: A Fusion Approach of FCM and CNN for Precise Detection”,doi:10.1109/ICDCECE60827.2024.10549262


Thank You!