International Journal of Scientific Research in Engineering and Management (IJSREM)
Volume: 08 Issue: 04 | April - 2024 SJIF Rating: 8.448 ISSN: 2582-3930

# CHRONIC KIDNEY DISEASE PREDICTION USING MACHINE LEARNING

NANDAM VENKATA KUSUMA SAI[1], KOTHURI GNANA SARASWATHI[2], KUNCHALA REVANTH KUMAR[3], POTHALA RAJESH[4], Dr. N. GOPALA KRISHNA[5]

[1,2,3,4] *Student, Department of Computer Science and Engineering, Tirumala Engineering College*
[5] *Professor, Department of Computer Science and Engineering, Tirumala Engineering College*

---

**Abstract -** The term –chronic kidney disease‖ means lasting damage to the kidneys that can get worse over time. If the damage is very bad, your kidneys may stop working. This is called kidney failure, or end-stage Kidney disease (ESRD). Kidney disease patients have the potential to get into the chronic phase and chronic kidney disease (CKD) is a decrease in kidney function gradually. So, doctors can diagnose kidney disease patients. So, predicting whether patients with Kidney disease have entered a phase of chronic kidney disease or not by showing the best accuracy result of comparing supervised classification machine learning algorithms. The aim is to investigate machine learning- based techniques for CKD forecasting by predicting results in the best accuracy. The analysis of the dataset by supervised machine learning technique (SMLT) to capture several information like, variable identification, univariate analysis, bi-variate and multi-variate analysis, missing value treatments and analyze the data validation, data cleaning/preparation, and data visualization will be done on the entire given dataset.

Additionally, to compare and discuss the performance of various machine learning algorithms from the given hospital dataset with an evaluation classification report, identify the confusion matrix and to categorizing data from priority and the result shows that the effectiveness of the proposed machine learning algorithm technique can be compared with the best accuracy with precision, Recall and F1 Score.

**Key Words:** Chronic kidney, Damage, Univariate, bi-variate, Multi-variate.

## 1.INTRODUCTION

Diabetes and high blood pressure are the two main causes of chronic kidney disease. Diabetes is characterized by high blood sugar levels, causing damage to the kidneys and heart, blood vessels and eyes. Moreover, poor control of high blood pressure can be a major cause of heart attack, stroke, and chronic kidney disease. Other conditions that affect the kidneys are glomerulonephritis, hereditary diseases, dysplasia, kidney stones, tumors, recurrent urinary tract infections, metabolic diseases, obesity,and age.

The most important and effective parameter for the evaluation of renal function is the glomerular filtration rate (GFR), which practically evaluates the ability of the kidney to filter blood. The glomerular filtration rate is the best measure of renal functionand is usually assessed (eGFR) by the results of a creatinine blood test

## 2. LITERATURE REVIEW

A literature review is a body of text that aims to review the critical points of current knowledge on and/or methodological approaches to a particular topic. It is secondary sources and discuss published information in a particular subject area and sometimes information in a particular subject area within a certain time period. Its goal is to bring the reader up to date with current literature on a topic and forms the basis for another goal, such as future research that may be needed in the area and precedes a research proposal and may be just a simple summary of sources. Usually, it has an organizational pattern and combines both summary and synthesis.

A summary is a recap of important information about the source, but a synthesis is a re-organization, reshuffling of information. It might give a new interpretation ofold material or combine new with old interpretations or it might trace the intellectual progression of the field, including major debates. Depending on the situation, the literature review may evaluate the sources and advise the reader on the most pertinentor relevant of them.

## 3.EXISTING SYSTEM

An accurate estimation of glomerular filtration rate (GFR) is clinically crucial for kidney disease diagnosis and predicting the prognosis of chronic kidney disease (CKD). Machine learning methodologies such as deep neural networks provide a potential avenue for increasing accuracy in GFR estimation.

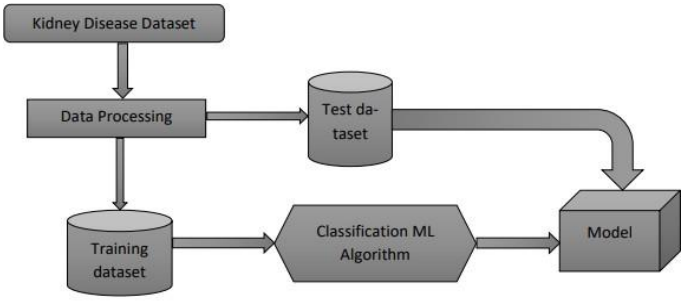
We developed a novel deep learning architecture, a deep and shallow neural network, to estimate GFR (dlGFR for short) and examined its comparative performance with estimated GFR from Modification of Diet in Renal Disease (MDRD) and Chronic Kidney Disease Epidemiology Collaboration (CKD-EPI) equations. The GFR model jointly trains a shallow learning model and a deep neural network to enable both linear transformation from input features to a log GFR target, and non-linear feature embedding for a stage of kidney function classification. We validate the proposed methods on the data from multiple studies obtained from the NIDDK Central Database Repository.

## 4. PROPOSED SYSTEM

The proposed method is to build a machine learning model for the classification of kidney disease. The process carries from data collection where past data related to kidney disease are collected. Data mining is a commonly used technique for processing enormous data in the healthcare domain. The kidney disease if found before proper treatment can save lives. Machine learning is now applied and mostly used in health care where it reduces the manual effort and a better model makes error less which leadsin saving the life. The data analysis is done on the dataset proper variable identification is done that is both the dependent variables and independent variables are found. Then proper machine learning algorithms are applied to the dataset where the pattern of data is learnt. After 15 applying different algorithms, a better algorithm is used for the prediction of the outcome. Architecture of Proposed model.

### Advantages for proposed system

We are using Machine Learning Technique to predict Kidney Disease.




Architecture of Proposed model

**Fig 3.2. Architecture of proposed model**

Algorithms are compared and the best model is evaluated for better prediction.

Performance metrics of different algorithms are compared and a better prediction is done.

**Early Detection and Prevention:** Implementing population-based screening programs and risk assessment tools can facilitate early detection of CKD and identify individuals at risk for progression. Emphasizing preventive strategies, lifestyle modifications, and risk factor management can delay the onset and slowthe progression of CKD.

**Patient Care:** Adopting a patient-centered approach that emphasizes shared decision-making, individualized care plans and providing education, resources, and support for self-management, treatment adherence, and lifestyle modifications can improve patient satisfaction, health literacy, and treatment outcomes.

**Quality Improvement:** Establishing quality improvement for CKD management can standardize care practices, promote evidence-based interventions, and monitor healthcare outcomes. Regular audits, feedback mechanisms, and benchmarking can drive continuous quality improvement and ensure adherence to best practices.

**Health Information Technology (HIT) Solutions:** Implementing electronichealth records (EHRs), clinical decision support systems, and interoperable data exchange platforms can streamline documentation, facilitate information sharing, and improve care coordination among healthcare providers.

## 5. SOFTWARE IMPLEMENTATION

### 5.1 GENERAL REQUIREMENTS

Requirements are the basic constrains that are required to develop a system. Requirements are collected while designing the system. The following are the requirements that are to be discussed.

1. Functional requirements
2. Non-Functional requirements
3. Environment requirements

A. Hardware requirements

B. software requirements

**Functional requirements**

The software requirements specification is a technical specification of requirements for the software product. It is the first step in the requirements analysis process. It lists requirements of a particular software system. The following details to follow the special libraries like sk-learn, pandas, numpy, matplotlib and seaborn.

**Non-Functional Requirements**

Process of functional steps,

1. Problem defines
2. Preparing data
3. Evaluating algorithms
4. Improving results
5. Prediction the result.

**REQUIREMENT SPECIFICATIONS**

**Software Requirements:**

Operating System: Windows

Tool: Anaconda with Jupyter Notebook

**Hardware requirements:**

Processor: Intel i5 , i7

Hard disk: Minimum 80 GB

RAM: Minimum 4 GB

## 5.2. SOFTWARE DESCRIPTION

Anaconda is a free and open-source distribution of the Python and R programming languages for scientific computing (data science, machine learning applications, large-scale data processing, predictive analytics, etc.), that aims to simplify package management and deployment. Package versions are managed by the package management system –Conda.‖ The Anaconda distribution is used by over 12 million users and includes more than 1400 popular data-science packages suitable for Windows, Linux, and MacOS. So, Anaconda distribution comes with more than 1,400 packages as well as the Conda package and virtual environment manager called Anaconda Navigator and it eliminates the need to learn to install each library independently. The open-sourcepackages can be individually installed from the Anaconda repository with the conda install command or using the pip install command that is installed with Anaconda. Pip packages provide many of the features of conda packages and in most cases they can work together. Custom packages can be made using the conda build command, and can be shared with others by uploading them to Anaconda Cloud, PyPI or other repositories. The default installation of Anaconda2 includes Python 2.7 and Anaconda3 29 includes Python 3.7. However, you can create new environments that include any version of Python packaged with conda.

## 6. SYSTEM TESTING

The purpose of testing is to discover errors. Testing is the process of trying to discover every conceivable fault or weakness in a work product. It provides a way to check the functionality of components, sub assemblies, assemblies and/or a finished product It is the process of exercising software with the intent of ensuring that the Software system meets its requirements and user expectations and does not fail in an unacceptable manner. There are various types of test. Each test type addresses a specific testing requirement.

**TYPES OF TESTINGS**

### 6.1. Unit Testing

Unit testing involves the design of test cases that validate that the internal program logic is functioning properly, and that program inputs produce valid outputs. All decision branches and internal code flow should be validated. It is the testing of individual software units of the application .it is done after the completion of an individual unit before integration. This is a structural testing, that relies on knowledgeof its construction and is invasive. Unit tests perform basic tests at component level and test a specific business process, application, and/or system configuration. Unit tests ensure that each unique path of a business process performs accurately to the documented specifications and contains clearly defined inputs and expected results.

### 6.2. Integration Testing

Integration tests are designed to test integrated software components to determine if they actually run as one program. Testing is event driven and is more concerned with the basic outcome of screens or fields. Integration tests demonstrate that although the components were individually satisfaction, as shown by successfully unit testing, the combination of components is correct and consistent.

### 6.3. Functional Testing

Functional tests provide systematic demonstrations that functions tested are available as specified by the business and technical requirements, system documentation, and user manuals.

Functional testing is centered on the following items:

Valid Input : identified classes of valid input must be accepted.

Invalid Input : identified classes of invalid input must be rejected.

Functions : identified functions must be exercised.

Output : identified classes of application outputs must be exercised.

Systems/Procedures: interfacingsystems or procedures must be invoked.
Organization and preparation of functional tests is focused on requirements, key functions, or special test cases. In addition, systematic coverage pertaining to identify Business process flows; data fields, predefined processes, and successive processes must be considered for testing. Beforefunctional testing is complete, additional tests are identified and the effective value of current tests is determined.

### 6.4. System Test
System testing ensures that the entire integrated software system meets requirements. It tests a configuration to ensure known and predictable results. Anexample of system testing is the configuration oriented system integration test. System testing is based on process descriptions and flows, emphasizing pre-driven process linksand integration points.

### 6.5. Whitebox Testing
White Box Testing is a testing in which in which the software tester has knowledge of the inner workings, structure and language of the software, or at least its purpose. It is purpose. It isused to test areas that cannot be reached from a black box level.

### 6.6. BlackBox Testing
Black Box Testing is testing the software without any knowledge of the inner workings, structure or language of the module being tested. Black box tests, as most other kinds of tests, must be written from a definitive source document, such as specification or requirements document, such as specification or requirements document. It is a testing in which the software under test is treated, as a black box .you cannot –see‖ into it. The test provides inputs and responds to outputs without considering how the software works.

### Test strategy and approach

Field testing will be performed manually and functional tests will be written in detail.

### 6.7. TEST OBJECTIVES
- All field entries must work properly.
- Pages must be activated from the identified link.
- The entry screen, messages and responses must not be delayed.

### Features to be tested
- Verify that the entries are of the correct format
- No duplicate entries should be allowed
- All links should take the user to the correct page.

### INTEGRATIONTESTING
Software integration testing is the incremental integration testing of two or more integrated software components on a single platform to produce failures caused by interface defects.

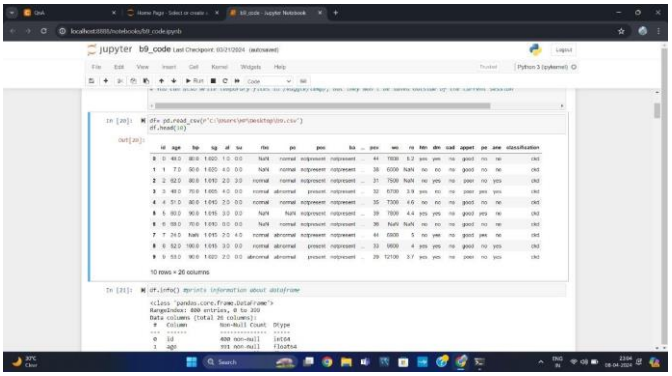
The task of the integration test is to check that components or software applications, e.g. components in a software system or – one step up – software applications at the company level –interact without error.

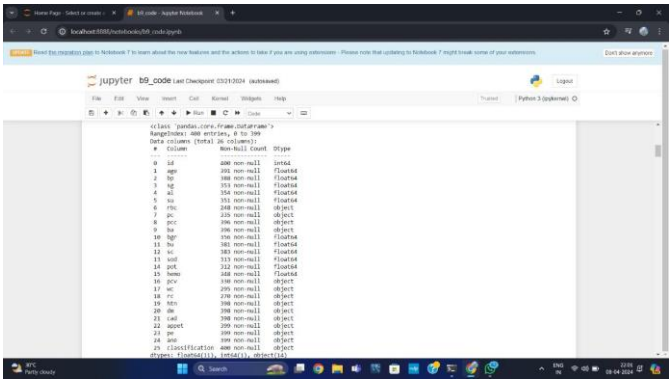
### Test Results
All the test cases mentioned above passed successfully. No defects encountered.

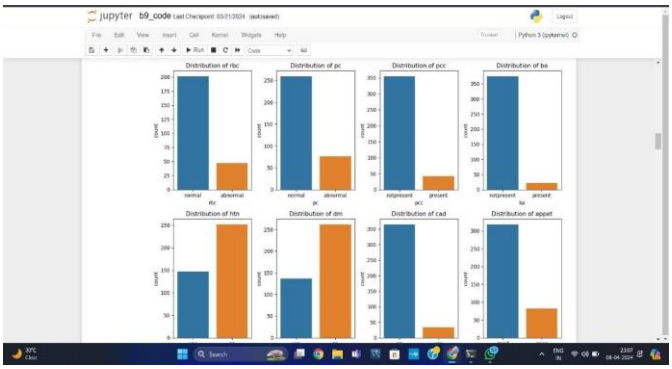
### Acceptance Testing
User Acceptance Testing is a critical phase of any project and requires significant participation by the end user. It also ensures that the system meets the functional requirements.

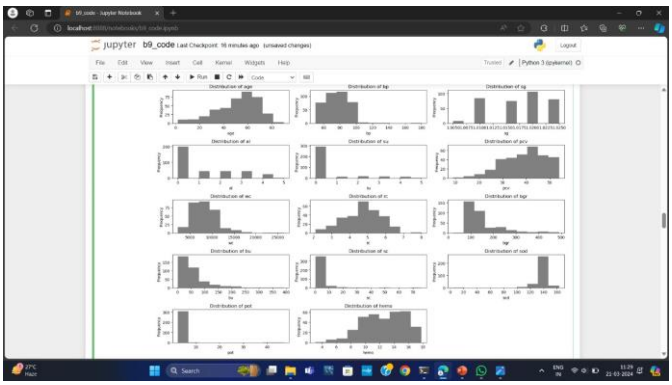
## 7. OUTPUT SCREENS

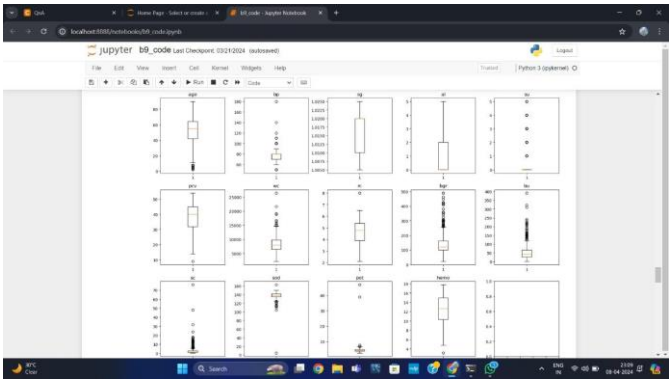
In this module the trained deep learning model is converted into hierarchical data format file which is then deployed in our project for providing better user interface and predicting the output whether the given image is CKD / Not CKD.

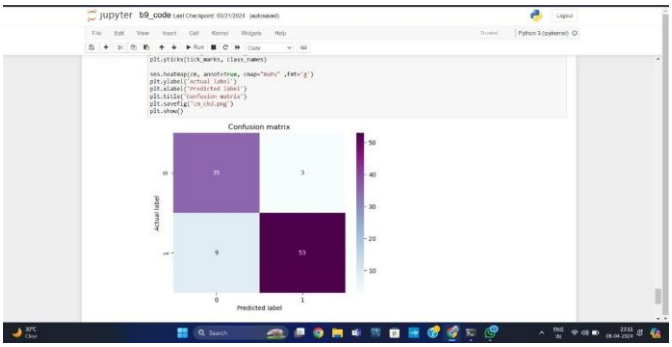













## 8. CONCLUSION

The analytical process started from data cleaning and processing, missing value, exploratory analysis and finally model building and evaluation.In this the models tha are constructed using CKD patients are then trained and validated using the input parameters that were discussed earlier.

When applying a filter feature selection approach to the reemaining attributes, it was discovered that heamoglobin, albumin, and specific gravity had the biggest impact when it comes to predicting CKD.This was the case after the method was used.The best accuracy on public test set is higher accuracy score is will be found out.This application can help to find the Prediction of Chronic Renal Disease.

## 9. REFERENCES


- H. Wang and P. Avillach, ‐Diagnostic classification and prognostic prediction using common genetic variants in autism spectrum disorder: Genotype-based deep learning,‖ JMIR medical informatics, vol. 9, no. 4, p. e24754, 2021.
- Mazzini SS, Arani RH, Hashem Inia M, Tohidi M, Azizi F, Hadeeth F. High incidence of chronic kidney disease among Iranian diabetic adults: using CKDEPI and MDRD equations for estimated glomerular filtration rate. Korean Diabetes J. 2021;45(5):684-97
- Zeba Z, Fatema K, Sumit AF, Zinnat R, Ali L. Early screening of chronic kidney disease patients among the asymptomatic adult population in Bangladesh. J Prev Epidemiol. 2020;5(1):e10–e.
- Harer MW, Charlton JR, Tipple TE, Reidy KJ. Preterm birth and neonatal acute kidney injury: implications on adolescent and adult outcomes. J Perinatol. 2020;40(9):1286–95. 39. Al Salmi I, Hannawi S. Birth weight and susceptibility to chronic kidney disease. Saudi J Kidney Dis Transplant. 2020;31(4):717
- Mahajan C, Tiwari V, Divyaveer SS, Patil MR, Banerjee A, Bagur V, et al. Spectrum of renal biopsies; a three-year data from a tertiary care Centre of eastern India. J Nephropharmacol. 2020;9(2):e20–e.
- E.-H. A. Rady and A. S. Anwar, ‐Prediction of kidney disease stages using data mining algorithms,‖ Informatics in Medicine Unlocked, p. 100178, 2019.
- Chaitanya SMK, Rajesh Kumar P. Detection of chronic kidney disease by using artificial neural networks and gravitational search algorithm. In: Saini H, Singh R, Patel V, Sanithi K, Ranganayakulu S, editors. Innovations in electronics and communication engineering. Springer. Singapore, 2019. pp. 441–8.
- Kalaiselvi K. Ant colony optimization (ACO) based feature selection and extreme learning machine (ELM) for chronic kidney disease detection. IJASSR. 2019;4(1). 71