International Journal of Scientific Research in Engineering and Management (IJSREM)
Volume: 07 Issue: 12 | December - 2023 SJIF Rating: 8.176 ISSN: 2582-3930

PREDICTION OF CHRONIC KIDNEY DISEASE

1.Alekhya.K, 2.Amulya..G, 3.Amulya.G, 4.Amulya.J, 5.Ananaya.S, 6.Aneesh.P, 7.Senthil kumar.K

1,2,3,4,5,6 Student, 7 Assistant Professor
Artificial Intelligence & Machine Learning, Department Of Computer Science And Engineering,
Malla Reddy University, Hyderabad, Telangana, India

ABSTRACT

Chronic Kidney Disease is a prevalent and weakening medical condition with a substantial impact on public health. Early detection and prediction reduces its adverse effects and improve patient outcomes. In this project, wepropose a comprehensive predictive analysis framework utilizing advanced machine learning algorithms to forecast the onset and progression of the disease. The dataset comprises a diverse range of clinical, population, and laboratory features collected from a group of chronic kidney disease patients over a defined period. Preprocessing techniques are employed to handle missing values and normalize data. Subsequently, various machine learning algorithms, including Random Forest, Support Vector Machine, and Gradient Boosting , are trained on the processed dataset to learn complex patterns and relationships within the data. Additionally, feature importance analysis is undertaken to identify the most influential variables contributing to prediction accuracy. The machine learning algorithms can effectively aid clinicians in predicting chronic kidney diseases, enabling timely interventions and personalized treatment plans. This research contributes to the growing knowledge on the application of machine learning in healthcare, particularly in chronic disease prediction.

.

INTRODUCTION

The code performs sentiment analysis on a dataset using machine learning techniques. Sentiment analysis involves determining the sentiment or emotion expressed in textual data, often categorized as positive, negative, or neutral. In this context, the code utilizes natural language processing (NLP) and machine learning algorithms to analyze a dataset containing textual sentences and associated sentiments. The primary steps include data loading, preprocessing, visualization, model training, and evaluation. Notable techniques employed in the code encompass text tokenization, stopwords removal, lemmatization, and the utilization of a Support Vector Machine (SVM) with a linear kernel for sentiment prediction. The evaluation metrics include accuracy, a classification report, a confusion matrix, and visualizations such asROC and precision-recall curves. The ultimate goal is to build and assess a sentiment analysis model capable of accurately categorizing the sentiment conveyed in the provided dataset.

LITERATURE REVIEW

1.1. Learning theories

Machine Learning Models: Utilizing machine learning algorithms, such as decision trees, random forests, or support vector machines, can help predict CKD based on various factors like age, blood pressure, diabetes status, and more.
Statistical Models: Regression analysis and survival analysis can be employed to identify significant predictors of CKD, considering variables like genetic predisposition, lifestyle factors, and medical history.

1.2. Self-determination theory (SDT)

Self-determination theory (SDT) explains people's inherent motivational tendencies for learning, growing, and connecting with others.
However, these tendencies are not automatic and they can be supported.

PROBLEM STATEMENT

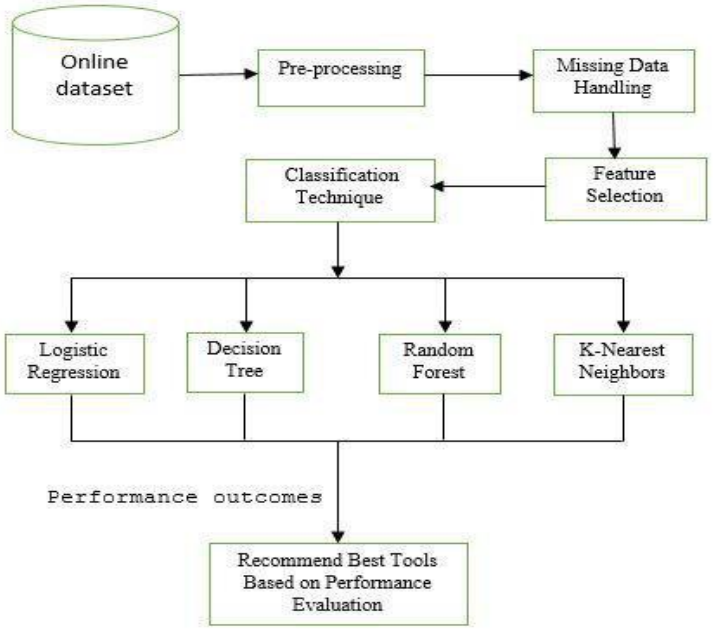
The prediction of kidney disease plays a pivotal role.
Chronic Kidney Disease (CKD) is a prevalent health issue with significant implications for patient well-being. Early detection and accurate prediction of CKD progression are crucial for timely intervention and improved patient outcomes.
Traditional methods for assessing CKD progression rely on static clinical markers, often leading to delayed identification of deteriorating renal function.
This research aims to address the limitations of current approaches by leveraging machine learning algorithms to develop a predictive model for CKD progression.
The primary objectives are:

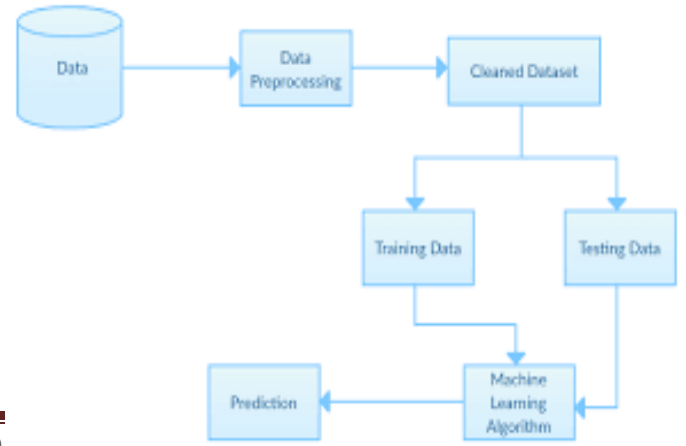
- Data integration
- Algorithm selection and optimization
- Real-time monitoring
- Feature selection

METHODOLOGY

Predictive modelling is used to analyse the data and thepredict the outcome. Predictive modeling used to predict the unknown event which may occur in the future. In this process, we are going to create, test and validate the model. There are different methods in predictive modelling. They are learning, artificial intelligence and statistics. Once we create a model, wea
can use many times, to determine the probability of outcomes. So, predict model is reusable. Historical datails used to train an algorithm. The predictive modelling process is an iterative the processand ofteninvolves training the model, using multiple models on the same dataset.

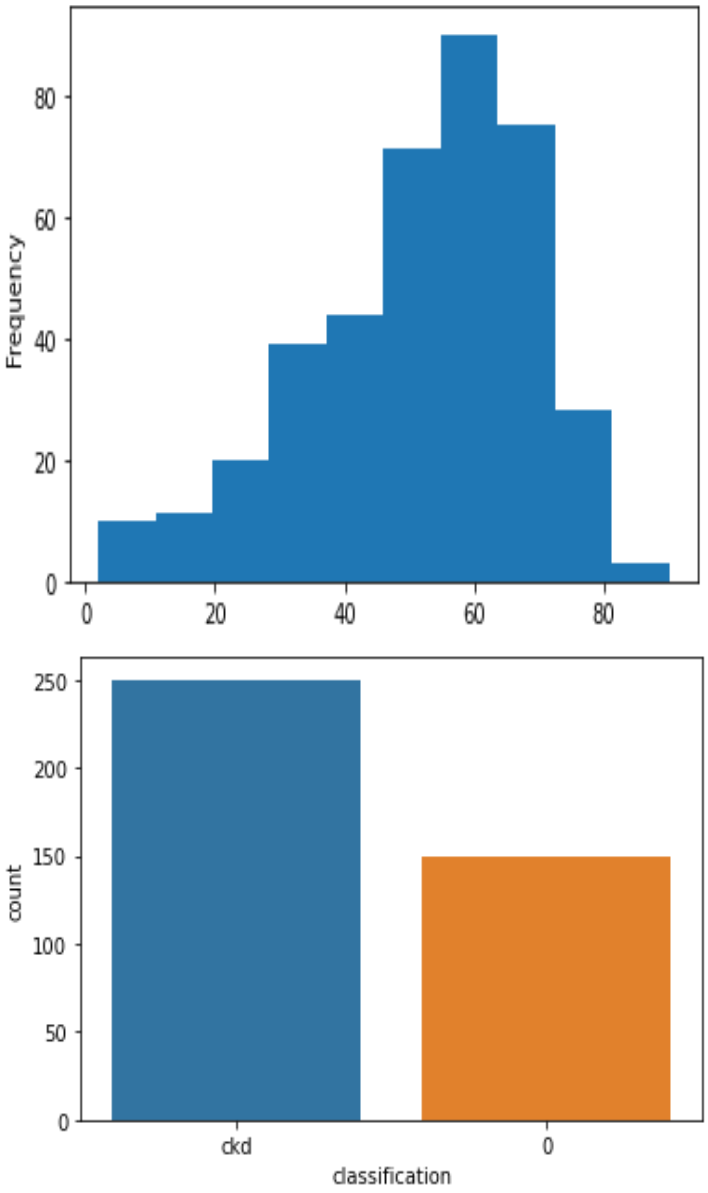
ARCHITECTURE





EXPERIMENTAL RESULTS

OUTPUT:



DFD/ER/UML DIAGRAM:

```
from sklearn.ensemble import RandomForestClassifier
from sklearn.datasets import make_classification
X, y = make_classification(n_samples=1000, n_features=4,
                           n_informative=2, n_redundant=0,
                           random_state=0, shuffle=False)
clf = RandomForestClassifier(max_depth=2, random_state=0)
clf.fit(X, y)
print(clf.predict([[0, 0, 0, 0]]))
```

[1]

CONCLUSION

This paper is aimed to observe and analyze the results obtained by applying different machine learning algorithms in the medical field in order to predict chronic kidney failure. Here we presented a prediction algorithm to predict CKD at an early stage. The dataset shows input parameters collected from the CKD patients and the models are trained and validated for the given input parameters. Decision tree, Random Forest and Support Vector Machine and Gradient Boosting and KNN learning models are constructed to carry out the diagnosis of CKD. The performance of the models is evaluated based on the accuracy of prediction. The results of the research showed that Gradient Boosting model better predicts CKD in comparison to Decision trees and Support Vector machines. The comparision can also be done based on the time of execution, feature set selection as the improvisation of this research..

FUTURE WORK

Integration of multi-model data:incorperating the Diverse data sources such as genomic,protomic and clical data to enhance prediction accuracy.
Real time monitoring:Creating systems for continuous monitoring of patient data to detect early science of CKD progessions or complications.

REFERENCES

[Yu et al., 2002] H. Yu, J. Han, & K. Chang. PEBL: Positive
example based learning for Web pageclassification using an
SVM. KDD-02, 2002.

[Liu et al., 2002] B. Liu, W. Lee, P. Yu, & X. Li. Partially su
pervised classification of text documents. ICML-2002.

[Muslea et al., 2002] I. Muslea, S. Minton & C. Knoblock. Ac
tive + semi-supervised learning =robust multi-view learning. I
CML-2002, 2002.