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.2024.61697

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

Deep Learning Techniques for Chronic Kidney Disease Risk Prediction

Muthu Janaki[1], Jashveen Kumar[2], Jeevani Dontham[3], Joel Sastry[4], Nama Kartheek[5], Prof. K. Rajeshwar Rao[6]
[1, 2, 3, 4, 5]*Department of AI-ML Malla Reddy University*
[6]*Department of AI-ML, Malla Reddy University, Maisammaguda, Hyd*

Abstract: ***Chronic kidney disease (CKD) and its symptoms are mild and gradual, often go unnoticed for years only to be realized lately. Bade, a Local Government of Yobe state in Nigeria has been a center of attention by medical practitioners due to the prevalence of CKD. Unfortunately, a technical approach In culminating the disease is yet to be attained. We obtained a record of 400 patients with 10 attributes as our dataset from Bade General Hospital. We used DNN model to predict the absence or presence of CKD in the patients. The model produced an accuracy of 98%. Furthermore, we identified and highlighted the Features importance to provide the ranking of the features used in the prediction of the CKD. The outcome revealed that two attributes; Creatinine and Bicarbonate have the highest influence on the CKD prediction.***

I. INTRODUCTION

The kidneys are a two of a kind of organs placed towards the lower back of the abdomen. Kidneys job is to strain the blood by moving out the toxic substance from the body using the bladder through urination. Kidney failure can cause death if the kidneys do not remove waste which is affected toxins. Difficulties of Kidney can be classified as acute or chronic. Chronic kidney diseases include circumstances that harm kidneys and reduce their capability to keep us fit. If kidney problem gets worse, the waste can build up at zenith levels in our blood and can cause difficulties such as hypertension, weak bones, anemia, bad nutrition, and nerve injure. Furthermore, kidney disease raises the possibility of heart and vascular diseases.

CKD can be caused by diabetes, hypertension, coronary heart disease, lupus, anemia, bacteria and albumin in the urine, complications of some drugs, sodium and potassium deficiency in the blood and a family history and many others. Early disclosure and medical care can usually avoid the worsening of chronic kidney disease. If it can get worse it can result in to kidney failure requiring dialysis. It can also result for kidney transplantation to sustain living. Chronic kidney disease (CKD) is a universal public wellbeing difficulty. 10% of the world's population is affected by CKD.

However, there is little direct evidence on how to diagnose CKD systematically and automatically. Worldwide, CKD is a serious reason of demise and disability. Chronic kidney disease (CKD) is a universal public wellbeing difficulty. 10% of the world's population is affected by CKD. However, there is little direct evidence on how to systematically and automatically diagnose CKD. Worldwide, CKD is a serious reason of demise and disability. It was the 27th focal reason in 1990 and became 18th focal reason in 2010 [1]. Near about 1 million people lose heir life in 2013 [2]. In spite of that, people of developing countries are being affected by CKD. In 2015 an orderly planned analysis, conducted and proclaimed that in high- income countries 109.9 million people had CKD in which 48.3 million and 61.7 million people were men and women respectively while the load was 387.5 million in countries with low average incomes where men were 177.4 million and women were 210.1 million [3]. According to the Kidney Foundation, out of 18 million people about 35,000 to 40,000 patients suffer from with chronic renal failure in Bangladesh each year. CKD can be caused by diabetes, hypertension, coronary heart disease, lupus, anemia, bacteria and albumin in the urine, complications of some drugs, sodium and potassium deficiency in the and a family history and many others. Early disclosure and medical care can usually avoid the worsening of chronic kidney disease. If it can get worse it can result in to kidney failure requiring dialysis. It can also result for kidney transplantation to sustain living. The motive of our research is to scrutinize the CKD data and portend the risk of CKD using Random Forest Algorithm and Artificial Neural Network.

II. LITERATURE SURVEY

Many researchers across the globe worked on various chronic kidney disease prediction models. "Predicting chronic kidney disease using a machine learning approach based on electronic health records" by Xie et al. (2021). In this study, the authors developed a machine learning model to predict CKD using electronic health records (EHR) data. The model achieved an AUC of 0.87 in predicting.

©IJRASET: All Rights are Reserved

"Development and validation of a machine learning model for prediction of incident CKD" by Xu et al. (2021). This study developed and validated a machine learning model to predict incident CKD using data from the National Health and Nutrition Examination Survey. The model achieved an AUC of 0.83 in predicting CKD. "Machine learning-basedprediction of chronic kidney disease progression: a systematic review" by Park et al. (2020). This review summarizes the current state of machine learning based CKD prediction models and their performance in predicting CKD progression. The authors concluded that machine learning models can improve the accuracy of CKD progressionprediction. "A machine learning model forpredicting progression of chronic kidney diseaseusing clinical data.And the model achieved an AUCof 0.85 in predicting CKD Progression. "Machine learning for prediction of CKD progression: a retrospective observational study" by Katayama et al. (2019). This study developed a machine learningmodel to predict CKD progression using data from a retrospective observational study. The model achieved an AUC of 0.84 in predicting CKD progression. S.Gopika, et Siddheshwar Tekale, et al.

[7] has analyzed 14 dissimilar characteristics correlated to CKD sick person and anticipated accuracy for several ML procedures alike Decision tree and Support Vector Machine. It is detected fromthe results study that decision tree algorithms gives the accuracy of 91.75% and SVM renders accuracy of 96.75%. The prediction process is less time consuming. Asif Salekin, et al.al. has performed a method of CKD Prediction with Clustering Method.The main objective is to determine the kidney function failure by using clustering algorithm. The experimental outcome revealed that the Fuzzy C means algorithm renders superior results and its accuracy 89%. Similarly Tabassum S, et al. Similarly, Tabassum S, et al. [5] has explored big data in healthcare which has been developed byconducting data mining techniques. EM is used to cluster parallel type of individual into one set. ANNand C4.5 are classification method which is used for prediction of the disease. EM got the accuracy resultof 70% which is a type of clustering algorithm. ANNand C4.5 is classification algorithm in which ANN got the accuracy result of 75% and C4.5 algorithm got the accuracy result of 96.75%. Sujata Dral, et al.Machine learning and different classification algorithms which is focused on predicting CKD status of a sick person with high accuracy. In this research they have used 5 CKD attributes out of 25 and two classification algorithms KNN and Naïve Bayes for portending CKD of a patient. KNN classifier predicted CKD with an accuracy of 100%, whereas Naïve Bayes Classifier predicted with an accuracy of 96.25%. They evaluated their method on a dataset of 400 entities, wherever 250 were CKD and they achieved a prediction accuracy of 0.993. They additionally performed feature assortment to choose the utmost appropriate features for recognizing CKD and flourishing them consistent with their certainty. Faisal Aqlan, et al. [9] has showed a DST that assist in the analysis of CKD. Data mining and analytics methods can be designedfor detecting CKD by conducting chronological patient's report and analysis proceedings. In this research DT, LR, NB and ANN were intended for detecting CKD. Houssainy A. Rady, et al. [10] haveexperienced that Effective data mining strategies is appeared to uncover besides concentrate concealed data The outcomes of relating PNN, MLP, SVM andRBF procedures have been associated. PNN process gives the maximum accuracy of 99.7%, as related with all other algorithm consequences. Sahil Sharma, et al.The outcomes of relating PNN, MLP, SVM and RBF procedures have been associated. PNN process gives the maximum accuracy of 99.7%, as related with all other algorithm consequences. Sahil Sharma, et al. [11] has applied various machine learning algorithms for CKD. 400 cases and 24 features are used for the research. The outcomes demonstrate that DTdidthe best resultswith the accuracy of 98.6%, sensitivity of 0.9720, precision of 1 and specificity of 1. Once the diseaseis detected, the website redirects the page into the page containing appropriate pesticides and chemicals with their usage and MRPs as well. It is one of the great work which identified so far as every researcher looking to detect and classify the diseases whereas this proposed methodology not only detecting the disease but also suggesting the appropriate pesticides suitable for it. In this study, we worked with 55 real life dataset and used chi- square test which is a feature selection algorithm. As mentioned earlier, we used ANN and Random Forest algorithm. The uniqueness of our study is, we merged the result of the two algorithms using logic mentioned below and predicted the risk.

III. PROBLEM STATEMENT

Chronic Kidney Disease (CKD) is a prevalent and debilitating condition affecting millions worldwide,characterized by the gradual loss of kidney functionover time. Early detection and accurate risk prediction are paramount for timely intervention andmanagement to mitigate adverse outcomes. DeeChronic Kidney Disease (CKD) is a prevalent and debilitating condition affecting millions worldwide, characterized by the gradual loss of kidney function over time. Early detection and accurate risk prediction are paramount for timely intervention and management to mitigate adverse outcomes. Deep learning techniques offer promisingavenues in this realm by leveraging complex patterns within vast datasets comprising clinical parameters, genetic markers, and lifestyle factors topredict CKD risk.

The problem statement revolves around developing robust deep learning models capable of accurately stratifying individuals' risk levels for CKD onset or progression, thereby enabling proactive healthcare interventions tailored to individual risk profiles and ultimately improving patient outcomes and healthcare resource utilization. p learning techniques offer promising avenues in this realm by leveraging complex patterns within vast datasets comprising clinical parameters, genetic markers, and lifestyle factors to predict CKD risk. The problem statement revolves around developing robust deep learning models capable of accurately stratifying individuals' risk levels for CKD onset or progression, thereby enabling proactive healthcare interventions tailored to individual risk profiles and ultimately improving patient outcomes and healthcare resource utilization.

IV. EXISTING METHODOLOGY

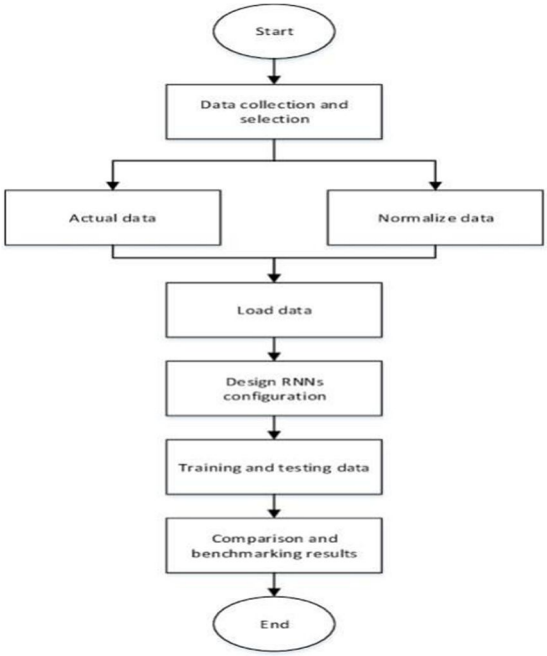
Existing methodologies for CKD risk prediction using deep learning techniques typically involve the utilization of large-scale datasets encompassing diverse clinical features, biomarkers, and patient demographics. These datasets are preprocessed to handle missing values, normalize features, and potentially augment data for improved model generalization. Deep learning architectures such as convolutional neural networks (CNNs), recurrentneural networks (RNNs), or hybrid models are then employed to extract intricate patterns and relationships from the data. Transfer learning may also be utilized to leverage pre-trained models on related tasks or datasets to enhance predictive performance, particularly in scenarios with limited labeled data. Model training often involves optimization techniques like stochastic gradient descent (SGD) or adaptive algorithms such as Adam to minimize prediction errors and fine-tune model parameters. Additionally, techniques like cross- validation and hyperparameter tuning are commonly employed to ensure model robustness and generalizability. Overall, these methodologies strive to develop accurate and clinically relevant models capable of effectively stratifying CKD risk for improved patient care and management strategiesExisting methodologies for CKD risk prediction using deep learning techniques typically involve the utilization of large-scale datasets encompassing diverse clinical features, biomarkers, and patient demographics. These datasets are preprocessed to handle missing values, normalize features, and potentially augment data for improved model generalization. Deep learning architectures such as convolutional neural networks (CNNs), recurrent neural networks (RNNs), or hybrid models are then employed to extract intricate patterns and relationships from the data. Transfer learning may also be utilized to leverage pre-trained models on related tasks or datasets to enhance predictive performance, particularly in scenarios with limited labeled data. Model training often involves optimization techniques like stochastic gradient descent (SGD) or adaptive algorithms such as Adam to minimize prediction errors and fine-tune model parameters. Additionally, techniques like cross- validation and hyperparameter tuning are commonly employed to ensure model robustness and generalizability. Overall, these methodologies strive to develop accurate and clinically relevant models capable of effectively stratifying CKD risk for improved patient care and management strategies.

V. PROPOSED METHODOLOGY

The proposed methodology for CKD risk prediction using deep learning techniques entails several steps to optimize model performance and clinical utility. Initially, a comprehensive dataset comprising clinical records, biomarkers, and demographic information is collected and preprocessed to handlemissing values, normalize features, and potentially augment data for improved model generalization. Deep learning architectures such as convolutional neural networks (CNNs), recurrent neural networks (RNNs), or attention mechanisms are then tailored to the task, with careful consideration given to model complexity and interpretability. Transfer learning may be employed to leverage pre-trained models on related tasks or datasets, especially in scenarios with limited labeled data. Model training involves optimization techniques like stochastic gradient descent (SGD) or adaptive algorithms suchas Adam, with hyperparameter tuning and cross- validation used to ensure robustness and generalizability. Additionally, attention is paid to model explainability and interpretability, facilitating clinical adoption and decision-making. Evaluation metrics such as accuracy, sensitivity, specificity, andarea under the receiver operating characteristic curve (AUC-ROC) are utilized to assess model performance, with extensive validation conducted on diverse patient cohorts to demonstrate real-world applicability and clinical efficacy. Ultimately, the proposed methodology aims to develop highly accurate and clinically relevant models for personalized CKD risk prediction, thereby enabling early intervention and improved patient outcomes. The proposed methodology for CKD risk prediction using deep learning techniques entails several steps to optimize model performance and clinical utility. Initially, a comprehensive dataset comprising clinical records, biomarkers, and demographic information is collected and preprocessed to handle missing values, normalize features, and potentially augment data for improved model generalization. Deep learning architectures such as convolutional neural networks (CNNs), recurrent neural networks (RNNs), or attention mechanisms are then tailored to the task, with careful consideration given to model complexity and interpretability.

Transfer learning may be employed to leverage pre-trained models on related tasks or datasets, especially in scenarios with limited labeled data. Model training involves optimization techniques like stochastic gradient descent (SGD) or adaptive algorithms suchas Adam, with hyperparameter tuning and cross- validation used to ensure robustness and generalizability. Additionally, attention is paid to model explainability and interpretability, facilitating clinical adoption and decision-making. Evaluation metrics such as accuracy, sensitivity, specificity, andarea under the receiver operating characteristic curve (AUC-ROC) are utilized to assess model performance, with extensive validation conducted on diverse patient cohorts to demonstrate real-world applicability and clinical efficacy. Ultimately, the proposed methodology aims to develop highly accurate and clinically relevant models for personalized CKD risk prediction, thereby enablingearly intervention and improved patient outcomes.

VI. ARCHITECTURE DIAGRAM



VII. DATA PRE-PROCESSING TECHNIQUES

Data preprocessing techniques play a crucial role inpreparing the dataset for CKD risk prediction using deep learning techniques. Initially, missing values are handled through methods such as mean or median imputation, or more advanced techniques like k-nearest neighbors (KNN) imputation, ensuring completeness of the dataset. Featurenormalization is then applied to scale numerical features to a standard range, preventing any single feature from dominating the model training process. Categorical variables are often encoded usingtechniques like one-hot encoding or label encoding to transform them into numerical representationssuitable for deep learning models. Additionally, feature engineering may involve creating new features or transformations to capture complex relationships within the data. Data augmentation techniques such as random rotation, cropping, or adding noise may also be employed to increase the diversity of the training dataset, enhancing model generalization. Overall, these preprocessing techniques aim to ensure the dataset is clean, standardized, and well-suited for training deep learning models to predict CKD risk effectively.
Data preprocessing techniques play a crucial role inpreparing the dataset for CKD risk prediction using deep learning techniques. Initially, missing values are handled through methods such as mean or median imputation, or more advanced techniques like k-nearest neighbors (KNN) imputation, ensuring completeness of the dataset. Featurenormalization is then applied to scale numerical features to a standard range, preventing any single feature from dominating the model training process. Categorical variables are often encoded usingtechniques like one-hot encoding or label encoding to transform them into numerical representationssuitable for deep learning models. Additionally, feature engineering may involve creating new features or transformations to capture complex relationships within the data. Data augmentation techniques such as random rotation, cropping, or adding noise may also be employed to increase the diversity of the training dataset, enhancing model generalization. Overall, these preprocessing techniques aim to ensure the dataset is clean, standardized, and well-suited for training deep learning models to predict CKD risk effectively.

VIII. MODULES

Deep learning techniques for Chronic Kidney Disease (CKD) risk prediction typically involve the utilization of various Python libraries and modules for model development and evaluation. Commonly used modules include TensorFlow and Keras for buildDeep learning techniques for Chronic Kidney Disease (CKD) risk prediction typically involve the utilization of various Python libraries and modules for model development and evaluation. Commonly used modules include TensorFlow and Keras for building and training deep neural networks, providing high-level APIs for constructing complexneural architectures and handling large-scale datasets efficiently.

Additionally, libraries such as scikit-learn offer tools for preprocessing data, performing feature selection, and evaluating model performance using metrics like accuracy, sensitivity, specificity, and area under the receiver operating characteristic curve (AUC-ROC). For handling medical imaging data, specialized libraries like PyRadiomics or SimpleITK may be employed for feature extraction and analysis from kidney ultrasound images. Furthermore, libraries such as matplotlib and seaborn facilitate data visualization, aiding in the exploration and interpretation of modelresults.

By leveraging these modules and libraries, researchers and practitioners can develop robust deep learning models for CKD risk prediction, contributing to early diagnosis and intervention strategies for improved patient outcomes.ing andtraining deep neural networks, providing high-level APIs for constructing complex neural architectures and handling large-scale datasets efficiently. Additionally, libraries such as scikit-learn offer tools for preprocessing data, performing feature selection, and evaluating model performance using metrics like accuracy, sensitivity, specificity, and area under the receiver operating characteristic curve (AUC-ROC).

For handling medical imaging data, specialized libraries like PyRadiomics or SimpleITK may be employed for feature extraction and analysis from kidney ultrasound images. Furthermore, libraries such as matplotlib and seaborn facilitate data visualization, aiding in the exploration and interpretation of model results. By leveraging these modules and libraries, researchers and practitioners can develop robust deep learning models for CKD risk prediction, contributing to early diagnosis and intervention strategies for improved patient outcomes.

IX. ALGORITHMS

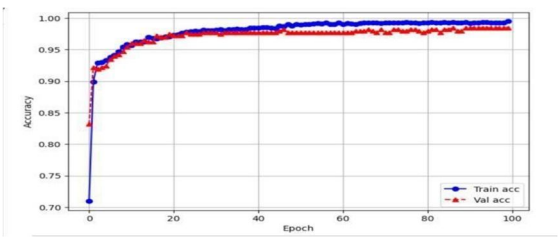
Deep learning techniques for Chronic Kidney Disease (CKD) risk prediction encompass various algorithms tailored to handle different types of data and tasks. Convolutional Neural Networks (CNNs) are commonly utilized for analyzing medical imaging data, such as kidney ultrasound scans, by extracting spatial features relevant to CKD risk factors from the images. Recurrent Neural Networks (RNNs) and their variants, such as Long Short-Term Memory (LSTM) networks, are effective for processing sequential data, such as patient health records or time series data, capturing temporal dependencies and patterns indicative of CKDprogression. Additionally, hybrid architectures combining CNNs and RNNs, known as Convolutional Recurrent Neural Networks (CRNNs), are employed to leverage both spatial and temporal information fromDeep learning techniques for Chronic Kidney Disease (CKD) risk prediction encompass various algorithms tailored to handle different types of data and tasks. Convolutional Neural Networks (CNNs) are commonly utilized for analyzing medical imaging data, such as kidney ultrasound scans, by extracting spatial features relevant to CKD risk factors from the images.

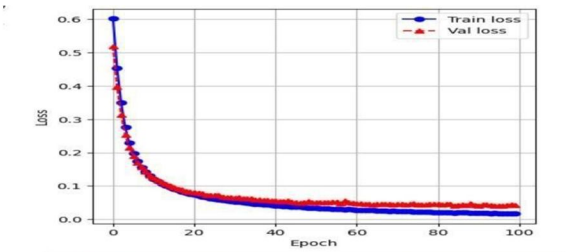
Recurrent Neural Networks (RNNs) and their variants, such as Long Short-Term Memory (LSTM) networks, are effective for processing sequential data, such as patient health records or time series data, capturing temporal dependencies and patterns indicative of CKD progression. Additionally, hybrid architectures combining CNNs and RNNs, known as Convolutional Recurrent Neural Networks (CRNNs), are employed to leverage both spatial and temporal information from heterogeneous data sources. Generative Adversarial Networks (GANs) can be utilized for data augmentation or generating synthetic medical images to augment limited datasets.

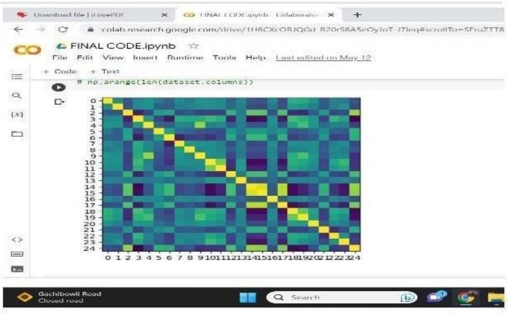
Moreover, attention mechanisms, such as Transformer models, enhance the interpretability and performance of deep learning models by focusing on relevant features or regions within the input data. By leveraging these diverse deep learning algorithms, researchers and clinicians can develop accurate and interpretable models for early CKD risk prediction, facilitating timely intervention and personalized patient care heterogeneous data sources. Generative Adversarial Networks (GANs) can be utilized for data augmentation or generating synthetic medical images to augment limited datasets. Moreover, attention mechanisms, such as Transformer models, enhance the interpretability and performance of deep learning models by focusing on relevant features or regions within the input data. By leveraging these diverse deep learning algorithms, researchers and clinicians can develop accurate and interpretable models for early CKD risk prediction, facilitating timely intervention and personalized patient care.

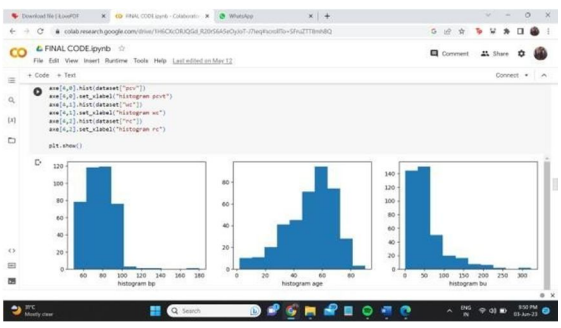
X. OUTPUT SCREENS











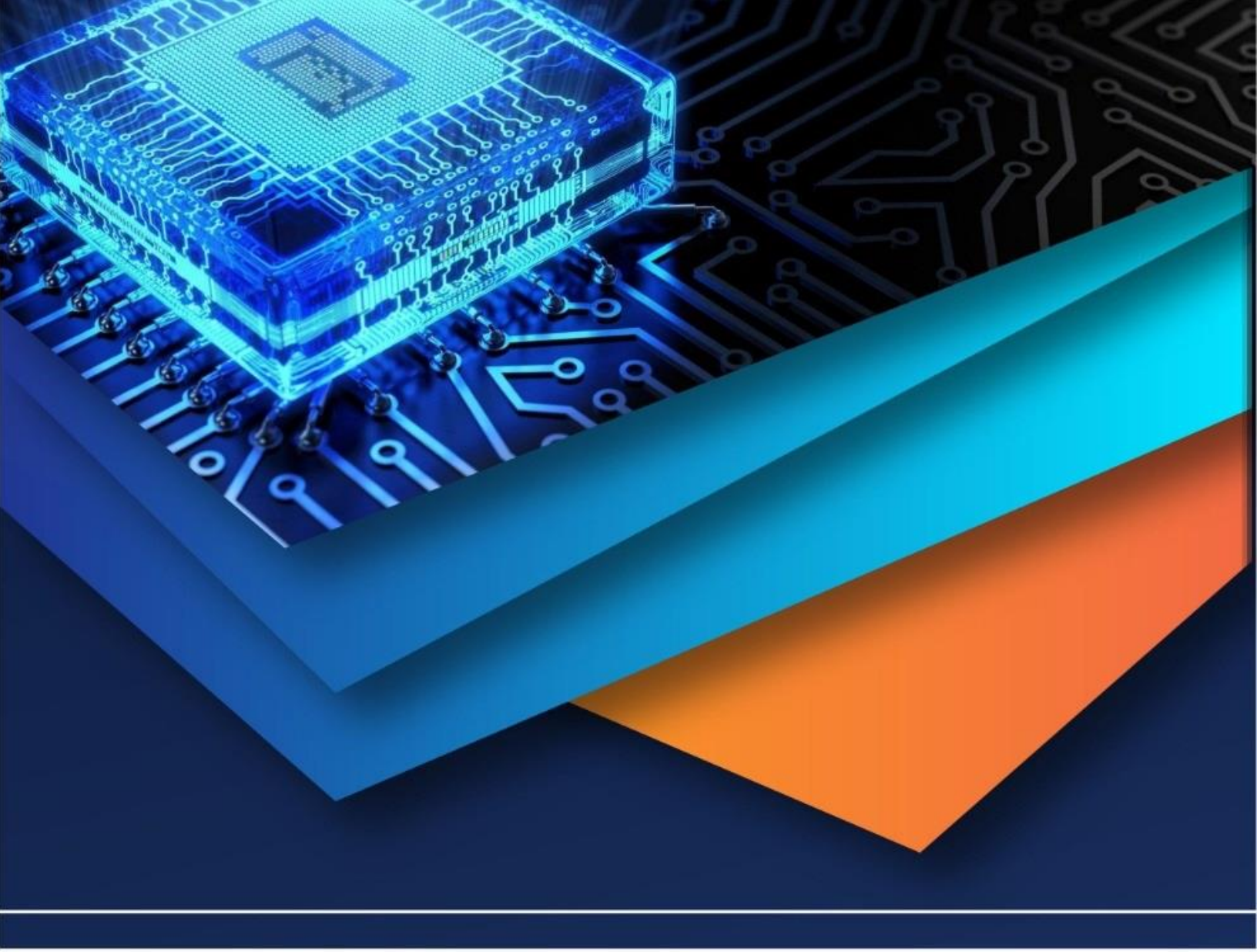
XI. FUTURE ENHANCEMENT

Further in future, our research would we extended to detection and classification of various other diseases on a single leaf and to develop an IoT based real- time monitoring system. To develop a website and launch a mobile application.

REFERENCES

[1] S. Gopika, and Dr. M. Vanitha, "Machine learning Approach of Chronic Kidney Disease Prediction using Clustering Technique", International Journal of Innovative Research in Science, Engineering and Technology, Vol. 6, no. 7,pp.14488-14496, 2017.
[2] Tabassum, Mamatha Bai B G, Jharna Majumdar,"Analysis and Prediction of Chronic Kidney Diseaseusing Data Mining Techniques", International Journal of Engineering Research in Computer Science and Engineering (IJERCSE),vol. 4, no. 9, pp.25- 31, 2017.
[3] Sujata Drall, Gurdeep Singh Drall, Sugandha Singh, Bharat Bhushan Naib, "Chronic Kidney Disease Prediction Using Machine Learning: A NewApproach", International Journal of Management, Technology And Engineering, vol. 8, no. 5, pp.278-287,2018.
[4] SiddheshwarTekale,PranjalShingavi,Sukanya Wandhekar, "Prediction of Chronic Kidney DiseaseUsing Machine Learning Algorithm",International Journal of Advanced Research in Computer and Communication Engineering, vol. 7, no. 10, pp.92-96, 2018.
[5] Asif Salekin, Jhon Stankovic, "Detection of Chronic Kidney Disease and Selecting Important Predictive Attributes", in 2016 IEEE International Conference on Healthcare Informatics 11 (ICHI), Chicago, USA, 2016, pp. 262-270.
[6] Faisal Aqlan, Ryan Markle, Abdulrahman Shamsan,"Data mining for chronic kidney disease prediction", 67th
[7] RNNual Conference and Expo of the Institute of IndustrialEngineers, United States, 2017, pp. 11789-1794.
[8] Sahil Sharma, Vinod Sharma, and Atul Sharma, "Performance Based Evaluation of Various Machine Learning Classification Techniques for Chronic Kidney Disease Diagnosis", International Journal of Modern Computer Science (IJMCS), vol. 4, no. 3, pp.11-15, 2016.

©IJRASET: All Rights are Reserved | SJ Impact Factor 7.538 | ISRA Journal Impact Factor 7.894 |



10.22214/IJRASET

45.98

IMPACT FACTOR: 7.129

IMPACT FACTOR: 7.429

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Call : 08813907089 (24*7 Support on Whatsapp)