2. LITERATURE REVIEW

The problem of CKD had achieved great attention in the early decade of the 21st century with papers researching on its problems with few being on diagnosing patients with CKD accurately, proposing software techniques to help determine correct drug dosage for the CKD patients, study effects of specific drugs on CKD patients to name a few. In the study conducted by A. Rainey et al. [7] on the effects of anti-hypertensive drugs for CKD patients, they built a five state Markov model that calculates transition probabilities and observes the patient’s kidney function. While research initiatives like CHRONIOUS [8] have been working continuously since February 2008 to monitor COPD (Chronic Obstructive Pulmonary Disease) and CKD patients via wearable technologies for health conditions from any place.

As the applicability and knowledge of Artificial Neural Networks grew and the need for binding software and medical backgrounds together for solving major illness problems was assigned, researches started to flourish on easing the burden on doctors for performing diagnosis and medica examination of patients by fusing Machine earning and ANN techniques. As seen in the paper L. Xun, Wu Xiaoming, Li Ningshan and Lou Tanqi [9], Chinese patients with CKD were tested for Glomerular Filtration Rate (GFR) and the rate of GFR was tested using Radial Basis Function (RBF) Neural Network. And in the study proposed by T. Shen et al. [10], a Decision Based Neural Network was used to develop a low-cost method for CKD patients to evaluate their risks cardiovascular disease based on T-Wave Alternans (TWA) and Heart Rate Variability (HRV) ECG Features.

Ultrasonography images are an important diagnosis tool for medical conditions and helps in determining the internal structure of the body. These images were utilized by R. M. Pujari and V. D. Hajare [11] in their paper to determine CKD stage of a patient using Image segmentation and Feature Extraction. Another study carried out by A. Sobrinho et al. [12] on early diagnosis of CKD utilized Colored Petri Nets (CPN) and developed MultCare, a CPN model-based patient-centered tool and aids in monitoring blood pressure and glucose levels and informs patients to perform a CKD diagnosis periodically.

CKD is a slow progressing disease that is left undetected until kidney stops functioning beyond irreversible limits. Its patients are usually marked into one of the 5 stages based on GFR rates and with no cure for CKD, treatment remains as an option and if the condition falls to the bottom line, patients are given a dialysis or kidney transplant. CKD is diagnosed using blood test, urine test, kidney scan, kidney biopsy, chest X-Ray (to check for edema) and GFR test. Using these tests, various biomolecule levels are marked and accordingly treatment is given to the patient. The accurate interpretation of the combination of biomolecules in the patient’s report plays an important role in determining best medicinal dosage and similar care for the patient. Machine Learning algorithms, models and Artificial Neural Networks play an important role in this case as they have been built to find all possible relations between attributes of a dataset, find connections and determine the disease. Moreover, functionalities can be added onto these Neural Networks to recommend the best and appropriate drugs and dosages for the patients as indicated in the papers above.

The Importance of ML algorithms can be demonstrated in the paper by M. M. Luck et al. [13] in which they perform metabolic profiling using Machine Learning algorithms and Supervised Feature Extraction and prove that it provides better insight into the biological data. Moreover, in the paper by Z. Sedighi, H. Ebrahimpour-Komleh and S. J. Mousavirad [14], they perform classification of CKD using filter and wrapper methods along with ML techniques and conclude that feature selection using fewer dimensions performs better classification. U. N. Dulhare and M. Ayesha [15] in their paper propose predicting CKD using Naïve Bayes with OneR attribute selector. Using this approach, they succeeded in increasing the accuracy of prediction to 12.5% greater than the existing systems at that time. In a practical application of helping CKD patients to get monitored timely and anywhere by doctors proposed by R. Cuevas, E. L. Dominiguez and Y. H. Velazquez [16], the propose system was a telemonitoring system that required an app for the patients and a web app for the doctors to monitor the patient details and give appropriate medication or care. Applications as such make it feasible and easy for the CKD patients to get timely and routine care for their conditions.

Decision Tree based classifiers have also been put into detecting CKD of patients by M. D. Basar, P. Sari, N. Kilic and A. Akan in their paper [17] that succeeded in providing better classification than individual classification. An interesting concept of game development being combined with CKD decision making can be seen in the paper by A. C. Vis [18] which described their process of helping patients find a better way of participating along with their doctors in taking decisions for their health because in most of the cases, the patients are not given choice for selecting the treatment options for themselves due to lack of information on their behalf. This game will help them get better insight onto the various treatments available as options to them so better participation can be observed from the patient’s end and they can decide what is best for them along with their medical doctors.

Another application using MATLAB was developed for texture analysis of the ultrasound kidney images [19] and mathematical operations were performed upon those images to distinguish between and normal kidney and affected kidneys. It was concluded that root mean square over the entire kidney region and the cortex region gave best classification results among the mathematical operations performed. Support Vector Machine (SVM) algorithm for detection of CKD by Y. Amirgaliyev, S. Shamiluulu and A. Serek [20] which gave a 93% accuracy high enough to indicate that ML algorithms performed well with CKD datasets and gave acceptable results. Another research by A. Maurya, R. Wable, R. Shinde, S. John, R. Jadhav and R. Dakshayani [21] was proposed which not only predicted the CKD affected kidney but also helped in providing a healthy diet plan to the CKD patients based on potassium levels of the blood using ML algorithms. Data Science is an important field that focuses on gathering the deepest of insights from large sets of data by visualizing the dataset. This proves to be of use in the CKD dataset as Data Science will help relate all the attributes and generate patterns from it to classify a normal and affected kidney as proposed by N. V. Ganapathi Raju, K. Prasanna Lakshmi, K. G. Praharshitha and C. Likhitha [22], which achieved an accuracy of 99.29% using Random Forest and XGBoost.

Knowing he fact that CKD patients may develop anemia over a course of time which makes them more susceptible to sicknesses and decreases their immune system. It is important to have proper medications for them delivered and available to them as early as possible. J. McAllister, Z. Li, J. Liu and U. Simonsmeier [23] in their paper proposed a method to regulate hemoglobin concentration in anemia affected CKD patients using recursive ZMPC (Zone Model Predictive Control) and also regulates erythropoietin ose optimization for the patient. Another research by Arif-Ul-Islam and S. H. Ripon [24] implied the usage of boosting algorithms like AdaBoost and LogitBoost, Ant-Miner and J48 Decision Tree for comparing performance of CKD classification. In the paper proposed by R. Devika, S. V. Avilala and V. Subramaniyaswamy [25] a comparative study was carried out for CKD prediction using Naïve Bayes, Random Forest classifiers and KNN and it was concluded that Random Forest Classifier performed the best amongst the three. Another study by N. Bhaskar and S. Manikandan [26] proposed the application of Convolutional Neural Networks (CNN) fused with a SVM classifier for detection of CKD. The result was a high accuracy of 98% achieved by the CNN-SVM network. In another study conducted by Y. Wang, W. Tong and Z. Liu [27] upon the power of metabolomic biomarker model for determining its role in the prediction of CKD, an elastic-net model (ENM) was built and it was found that GFR rates alone do not help in truly identifying the CKD stage of a patient. Another group of researchers [28] proposed a diagnostic imaging tool for CKD diagnosis that evaluated the cellular level kidney imaging for Kidney Level Diagnosis which led to the finding of a new criteria for determining severity of the CKD.

Internet of Things (IoT) has grown its field of interest in minimizing human chores and making connecting lives wherever and whenever a possibility. With this in mind, A. Noor, A. Banerjee, M. F. Ahmad and M. Nazim Uddin [29] created an IoT based healthcare system for the CKD patients that has features like patient daily logs, nutrition levels of food intake, sleep pattern monitoring, water intake and other such features to help doctors in monitoring and giving proper care to the patients. In recent research carried out by S. Vashisth, I. Dhall and S. Saraswat [30], a CKD diagnosis classification was carried out using Multi-Layer Perceptron (MLP) classifier that achieved testing accuracy of 92.5%. Another study by R. Gupta, N. Koli, N. Mahor and N. Tejashri [31] focused on analyzing different ML algorithms on which algorithm performs the best among Decision Tree, Random Forest and Logistic Regression and marked the conclusion that Logistic Regression achieved highest accuracy while Decision Tree achieved highest precision.