Detecting Chronic Kidney Disease from Blood Samples using Neural Networks

Mohamed Ghaisan Latheef*, Rajasvaran Logeswaran and Nurul Haniza Mohtar

Asia Pacific University of Technology & Innovation, Bukit Jalil, 57000 Kuala Lumpur, Malaysia

*Corresponding Author e-mail: TP047492@mail.apu.edu.my,

Abstract. This paper proposes an artificial neural network approach to automatically detecting Chronic Kidney Disease through fluid samples taken from patients. The rationale for developing such a system is given, as well as possible benefits to the patients and medical industry. Similar systems proposed in the industry and for diagnosing chronic kidney disease through other approaches such as classification algorithms are explored. A dataset to train the neural network on is collected and features analysed, as well as methodology and tools to be used in the development of the neural network.

Index Terms. Artificial Neural Networks, Medical Diagnosis, Data Analysis, Chronic Kidney Disease (CKD).

1. Introduction

It is estimated that about 1 in every 10 adults suffer from some form of kidney damage, with millions dying annually with complications related to Chronic Kidney Disease (CKD) [1]. The Global Burden of Disease study had CKD penned at rank 27th in the list of causes of total number of deaths worldwide in 1990, but rose to 18th in 2010 – a degree of list ranking upward movement second only to that of HIV and AIDS. CKD is extremely harmful and incurable, but if caught early its progress can be halted through medication and a proper diet regiment.

However, it is difficult to catch the disease early, due to the lack of warning symptoms that is outwardly visible in the early stages. Therefore, unless someone has other conditions that would make them more susceptible to kidney damage, if the person is very cautious and does regular full body check-ups, or if they took the blood or urine test for another reason and coincidentally also found CKD, normally CKD is only found in mid to late stages.

CKD is determined as when for more than three months, the kidneys are observed to be unable to effectively perform its intended functions, such as to clean blood of waste and excess products and help control blood pressure [2]. This can lead to a buildup of waste products in the patient’s body, leading to swelling and bloating of the ankles, insomnia, shortness of breath, and weakness. By the time even these symptoms are observed, it may already be too late. Early CKD shows no outward signs or symptoms, sometimes even up until the point the person had already lost 90% of their kidney functions [3].

The current methods of diagnosing CKD include measuring results of Serum Creatinine in the blood, blood glucose levels to see if the patient is diabetic (as diabetic patients have a much high prevalence for being afflicted with CKD), as well as measuring for albumin, or the presence of protein in the urine, which would be filtered out in healthy kidneys.
However, these generally require three months of monitoring of creatinine levels as well as other criteria such as hematuria, congenital malformations, etc. According to the US National Institute of Diabetes and Digestive and Kidney Diseases (NIDDK), latest research suggests that a staging/classification system to identify the risk of CKD from analyzing patient data with greater accuracy in its predictions, requires multiple factors and not simply eGFR, creatinine or albumin levels, but all three, age, and diabetes status [4]. Therefore, this project aims to use a similar set of patient data with multiple features in order to train a neural network to be able to identify patterns which indicate the presence of CKD and do binary classification.

By doing so, this project aims to ease the process of testing for CKD and decrease the several months wait time to determine if the issue is chronic, by simply using the results of one blood test and other observations and samples made/taken, with greater confidence on the diagnosis based on multiple factors instead of a single factor, which would have higher chances of being an outlier due to external factors. This is especially vital since any progression of CKD is irreversible and therefore it is in the patients’ best interests to have it quickly and accurately diagnosed in order to start treatment and take the appropriate measures to slow the damage immediately.

2. Literature Review

2.1 Domain Research

Medical diagnosis and decision-making is a high-priority high-risk field, where making accurate diagnoses are absolutely critical to the patients’ health. Artificial Neural Networks (ANNs) and other machine learning techniques are being increasingly used for cases of diseases diagnosis, with promising levels of accuracy.

Al-Shayea presented two cases of ANNs being used for diagnosis of: (1) acute nephritis disease, which occurs when the kidneys suddenly become inflamed and lead to kidney failure if left untreated, as well as (2) heart disease [5]. Feed-forward backpropagation was used, with the model for acute nephritis reported 99% and heart disease 95% accuracy rate in diagnosis. The model for the former had 6 inputs, 20 hidden neurons, 2 layers, and used Levenberg-Marquardt backpropagation as a training algorithm, with training based on improvement of mean-square error (MSE). The latter network used 2 layers, 22 inputs, 20 hidden neurons with other factors being the same as the previous.

Šter and Dobnikar tested five different medical databases (Coronary Artery Disease, Breast Cancer, Hepatitis, Pima Indians Diabetes, and Heart Disease) with ANN and simpler linear discriminant methods [6]. The performance was relatively the same, implying, they claimed, that the data was simple with the values being independent attributes. As such, they claim that complex classification systems or ANN are unnecessary in medical diagnosis, as results are high for linear discriminant and Naïve Bayes methods as well. However, it must be noted that they tested the methods with default configurations and no fine-tuning, and as they themselves accept that even a gain of a few percent in accuracy in diagnosis of patients is vital as quantitatively that could be as much as hundreds or thousands of people. This was an earlier study and ANNs have grown significantly thereafter, especially with deep learning. However, a persistent issue with ANNs is the lack of transparency – it is unable to be seen how the ANN made the decision it produced and what values the decision hinged upon. The patient would not be convinced if the doctor did not know why they are giving the patient the specific medicine and treatment by simply following the prescription of an ANN. The lack of explanation also means that human doctors are left in the dark and cannot learn from the decisions made by the ANN.

Liu et al. undertook the task of testing the accuracy of healthcare professionals as compared to deep learning techniques such as ANN [7]. In order to get as large a dataset as possible, they scoured existing articles and analyses in order to perform a meta-analysis, on studies from 1 January 2012 to 6 June 2019, for studies comparing the diagnostic performance for any disease of deep learning models as compared to healthcare professionals. Comparison of performance for 14 of those studies showed 87% pooled sensitivity and 92.5% pooled specificity for the deep learning models, with 86.4% and 90.5%, respectively, for healthcare professionals. Although the results are similar, deep learning models overall performed better even if marginally, and any minor increase in percentage is still a reflection of countless lives saved or improved by extrapolating that figure onto the quantitative value of those afflicted by such health conditions.
2.2 Similar Systems

There are many articles detailing the diagnosis of CKD using ANN or other methods, such as utilizing Support Vector Machines (SVM) and Naïve Bayes. Kriplani et al. used Deep Neural Network in order to predict CKD using 18 parameters in the input layer, although the number of hidden layer neurons, layers and network architecture were not mentioned in the paper [8]. The Deep Neural Network had a true positive rate of 95.2%, and a true negative rate of 100%, which was also compared with other classification methods, such as Logistic, Random Forest, AdaBoost, SVM and Naïve Bayes. The paper claimed that their Deep Neural Network had the highest accuracy among all the methods highlighted above. However, on further analysis of the performance results given in the paper, although the same true negative rate was obtained by all other classification methods, Naïve Bayes and Random Forest, relatively more simple algorithms, performed as well as the ANN with 95.2% true positive accuracy. Further, both AdaBoost and SVM achieved a true positive accuracy rate of 96.2%, which is higher than the Deep Neural Network. In another table provided, Naïve Bayes and Deep Neural Network were shown with an accuracy of 97.7679%, whereas AdaBoost and SVM both had 98.2143% accuracy and Random Forest had 99.1071% accuracy. Thus, their conclusion that out of all the models compared, the Deep Neural Network was the best appears erroneous according to the figures provided in their paper, as not only is it slower and higher in computational resources consumed, Naïve Bayes performed equally, with AdaBoost, SVM and Random Forest obtaining a higher accuracy.

Ahmad et al. used SVM in order to propose an auxiliary decision support tool for emergency situations, using 5 inputs: blood pressure, serum creatinine, packed cell volume, hypertension factor and anemia factor [9]. The SVM was coded in R programming language, with an accuracy of 98.34%. They also concluded in their paper that the importance of the attributes were linked to the Mean Decrease in Gini, or the mean of the variable’s total decrease in node impurity, with higher the value greater the role of the attribute. Using this, they found that packed cell volume was the most important attribute in diagnosing CKD from the data. As solely the serum creatinine levels is one of the ways CKD is diagnosed now, this creates some uncertainty as to whether it might be more accurate to rely on Packed Cell Volume values instead.

Ravindra et al. also used SVM in order to perform classification and identify CKD patients [10]. The input values were distributed into four different cases based on close association between them, with:

Case 1 - blood pressure, specific gravity, and serum creatinine;
Case 2 - albumin, sugar, blood glucose random, and hemoglobin;
Case 3 - packed cell volume, white blood cell count, and red blood cell count;
Case 4 - albumin, sugar, blood glucose random, serum creatinine, sodium, potassium, and hemoglobin.

Each of the four cases were separately tested with the SVM classifier without the other factors, and it was found that Case 2 had the highest accuracy of 93.75%. However, the accuracy is lower than other articles previously, which used a higher number of inputs. Therefore, while albumin, sugar, blood glucose random and hemoglobin seem to produce good results on their own, other inputs have demonstrated an increase in accuracy when included alongside.

Salekin and Stankovic used Random Forest and achieved 99.3% accuracy and 0.1084 Root Mean Square Error (RMSE) using 24 inputs [11]. However, they then used a variety of methods such as a machine learning wrapper method and LASSO regularization, to select 5 features from the original 24:

- specific gravity,
- albumin,
- diabetes mellitus,
- hypertension, and
- hemoglobin.

The paper highlighted a cost effective automatic detection method, and therefore reducing the number of factors needed while keeping the accuracy as level as possible would increase cost effectiveness due to less tests needed and time taken to get the results of the tests. They estimate that with the 2016 US inflation rate and medical test pricing, it would take no more than USD 45.05 for a patient test. With the
5 attributes, they achieved 98% accuracy with 0.11 RMSE, a decrease of only 1.3% accuracy after discarding 19 of the original 24 features.

Vijayarani and Dhayanand used MATLAB in order to compare the performance of SVM and an ANN (feedforward with backpropagation) on a dataset of Kidney Disease patients [12]. The ANN had an accuracy rating of 87.70%, with the SVM having an accuracy of 76.32%. However, SVM took 3.22 seconds whereas the ANN took almost double that time at 7.26 seconds. The accuracy was markedly below those of the previously mentioned systems. However, this system also had more than binary classification of CKD or Healthy Kidneys, instead it classified into 5 different cases: Normal, Acute Nephritic Syndrome, CKD, Acute Renal Failure, or Chronic Glomerulonephritis, which made the task more complex.

2.3 Review Summary
The paper by Liu et al. shows the viability of ANN and other machine learning substitutes to replace the trained eye of healthcare professionals, by simply being more accurate. However as Šter and Dobnikar highlights, there is an issue with the transparency of ANN decisions, as there is no way to identify which factor played into that decision, and how heavily. The factors can be identified outside of the ANN through various data mining and analysis methods similar to the approach by Salekin and Stankovic on the CKD features to see which play a larger role. However, there are contradictions when different methods are used to identify feature priority, such as that set of 5 only sharing 1 feature with the case of highest accuracy in Ravindra et al. ‘s findings, albumin. In addition, Ahmad et al.’s Mean Decrease in Gini approach found packed cell volume as being the most significant factor, however none of the previous papers had any significance attributed to that factor in the final features they chose for their systems.

There are issues with how ANNs make decisions, and in a way, those decisions have some element of human bias, as the developers decide which inputs to discard as unnecessary and which to keep. Salekin and Stankovic does however have a good point; we have the means and it is time to develop a more accurate cost-effective solution to identifying CKD patients faster, such that people are avoided of undue pain, distress and death.

3.0 Materials and Methods
3.1 Programming the ANN
MATLAB was used for fast prototyping as it allows native creation of ANNs through its Deep Learning Toolbox without any additional coding. Figure 1 provides a snapshot of the ANN training tool of the Deep Learning Toolbox. Following Kanban, all the tasks for the project are placed on a virtual board, such as Trello, with differing columns for tasks that have already been completed (such as deciding on system functionality, data gathering, planning schedule, finishing preliminary report, filling out ethics form, etc.), tasks in the process of being completed (finishing investigatory report) and tasks to complete in the future (clean data, train AI with data, document results, etc.). The basic principle is to keep a work-in-progress limit and stick to it, being as constant as possible without exceeding the limit and taking too many tasks at once, and continue until the project is complete.

3.2 Obtaining the Training and Testing Data
For the training and testing of the model, there were two options considered. One was to contact local hospitals for anonymous patient information for purely academic use, and the other was to find a dataset available for use online. Online was the preference, and a dataset was found with 400 patient records, 250 CKD patients, and 150 non-CKD patients, collected from Apollo Hospitals in Tamilnadu, India, over the course of 2 months. There were 24 different features, as shown in Table 1. However, the records were incomplete, with data missing or corrupted in many instances, as is often the case. If all the records with an empty value were removed, the usable records would only be slightly over 200. Thus, alternative methods of cleaning the data and retaining the incomplete records had to be looked into, such as replacing the empty values by the mode, or the most commonly seen value in the feature, but with additional consideration for the most common value seen in other records for that feature, amongst records that are otherwise most similar to the record with the missing feature.
Figure 1. Neural Network Training tool, a feature of the Deep Learning Toolbox

Table 1. Features List

| Feature               | Datatype          |
|-----------------------|-------------------|
| Age (age)             | Years             |
| Blood Pressure (bp)   | mm/Hg             |
| Specific Gravity (sg) | Nominal           |
| Albumin (al)          | Nominal (1-5)     |
| Sugar (su)            |                   |
| Red Blood Cells (rbc) | Normal/Abnormal   |
| Pus Cells (pc)        |                   |
| Pus Cell Clumps (pcc) | Present/NotPresent|
| Bacteria (ba)         |                   |
| Blood Glucose Random (bgr) | mgs/dl         |
| Blood Urea (bu)       |                   |
| Serum Creatinine (sc) |                   |
| Sodium (sod)          | mEq/L             |
| Potassium (pot)       |                   |
| Hemoglobin (hemo)     | gms               |
10 of the values were also text representing Boolean values, which also had to be cleaned to 0 and 1 values as only numeric values are accepted to train the ANN. Furthermore, it is difficult for an artificial intelligence (AI) with no context of language to make the association that “yes” is the opposite of “no” as opposed to “1” and “0”. Depending on the results of cleaning the data and the records retained, it is possible that other sources of data have to be relied on, such as contacting hospitals for anonymous patient data.

4. Results and Discussion

On average, the typical age of CKD patients were found to be around 55 to 75, whereas the age of non-CKD patients were more evenly distributed. This shows that CKD is more prevalent in the elderly, which is reasonable considering lack of genetic issues, as it takes time and sustained damage to the kidneys to develop CKD.

Diabetes Mellitus, also commonly known as Diabetes, is a risk factor which increases the chance of getting CKD. In fact, according to the American National Kidney Foundation [13], about a 30% of Type 1 and 40% of Type 2 diabetics get CKD later on in life. If diabetic patients do not regulate their blood glucose levels through medication, insulin injections and diet, the high glucose levels clog the tiny blood capillaries in the kidneys, damaging the kidneys. All of the diabetic patients in the dataset were also CKD patients, with more than half of the CKD patients being diabetic.

The Red Blood Cell count on average in CKD patients are lower than that of Non-CKD patients, ranging from 2 to 6 million instead of the healthy 4 to 7 million red blood cells per microliter of blood. A low red blood cell count is indicative on anaemia. The kidneys produces a hormone called Erythropoietin (EPO), which gets the body to produce blood cells, enough of which is not produced when Kidney function decreases due to damage from CKD [14]. As such, EPO levels drop, causing red blood cell count to drop in turn, therefore causing anemia. This leads to less oxygen carried throughout to body, general fatigue, shortness of breath, cold extremities (hands and feet), and in the worst cases, death.

Sodium and Potassium are essential mineral salts needed for the human body. However, it is possible to exceed the daily limit intake, as with water. Normally, the kidneys clear the excess salts and fluids, but an issue arises when they no longer function as they are supposed to. This leads to a fluid build-up in tissues and bloodstream, causing high blood pressure as well as nausea, weakness and abnormal heart rhythms [15]. Potassium is in 3.3 to 5.25 range for non-CKD patients with a high variance, and sodium

|                          | Nominal                        |
|--------------------------|--------------------------------|
| Packed Cell Volume (pcv) |                                |
| White Blood Cell Count (wc) | cells/cmm                     |
| Red Blood Cell Count (rc) |                                |
| Hypertension (htn)       |                                |
| Diabetes Mellitus (dm)   | Yes/No                         |
| Coronary Artery Disease (cad) |                        |
| Appetite (appet)         | Good/Poor                      |
| Pedal Edema (pe)         | Yes/No                         |
| Anaemia (ane)            |                                |
at 135 to 151. However, potassium values of CKD patients can peak at up to 48.4, about 10 times higher than the normal amount, with sodium peaking at 170.4.

There are many values in the input data that have similar linked attributes, such as hemoglobin, red blood cell count and anaemia. Naturally, low red blood cell count and low hemoglobin correlates with the development of anaemia. Similarly, there is a correlation between potassium and sodium levels with high blood pressure, as well as blood glucose levels, sugar and being diabetic. Pus cells and pus cell clumps, as well as bacteria values also correlate as the presence of bacteria causes pus, which precedes pus cell clumps. This can be used to reduce the number of inputs, or be taken into consideration when cleaning the data and adding the incomplete values. The weights that the ANN places on each individual input will be through training and not be set manually, so whether the analysis matches up with the neural network’s own decision-making process is yet to be seen.

5. Conclusion
The data analysed of the patients show a clear pattern through manual data analysis even to the human eye. An ANN may be capable of seeing far more subtle patterns in the numerical data, and can therefore present a high level of accuracy in terms of diagnosing CKD in patients through observing the values of features present in their bodily fluids. In fact, reviewing papers on a similar subject shows that the number of features can be reduced and the tests and fluids collected from a patient be reduced. Therefore, with a reduction in cost and effort, it is possible to effect a minimal decrease in the result accuracy due to the consistently high correlation shown between some values.

References
[1] National Kidney Foundation, “Global Facts: About Kidney Disease | National Kidney Foundation,” National Kidney Foundation, 2015. [Online]. Available: https://www.kidney.org/kidneydisease/global-facts-about-kidney-disease. [Accessed 29 January 2020].
[2] WebMD, “Kidney Disease (Chronic & Acute) - Causes, Risk Factors, Stages,” WebMD, 2020. [Online]. Available: https://www.webmd.com/a-to-z-guides/understanding-kidney-disease-basic-information#1. [Accessed 29 January 2020].
[3] World Kidney Day, “Chronic Kidney Disease - World Kidney Day,” World Kidney Day, 2019. [Online]. Available: https://www.worldkidneyday.org/facts/chronic-kidney-disease/. [Accessed 28 January 2020].
[4] National Institute of Diabetes and Digestive and Kidney Disease, “Identify & Evaluate Patients with Chronic Kidney Disease | NIDDK,” 2020. [Online]. Available: https://www.niddk.nih.gov/health-information/communication-programs/nkdep/identify-manage-patients/evaluate-ckd. [Accessed 28 January 2020].
[5] Q. K. Al-Shayea, “Artificial Neural Networks in Medical Diagnosis,” IJCSI International Journal of Computer Science Issues, vol. 8, no. 2, pp. 150-154, 2011.
[6] B. Šter and A. Dobnikar, “Neural Networks in medical diagnosis: Comparison with other methods,” Proceedings of the International Conference on Engineering Applications of Neural Networks, 1996.
[7] X. Liu, L. Faes, A. U. Kale, S. K. Wagner, D. J. Fu, A. Bruynseels, T. Mahendiran, G. Moraes, M. Shamdas, C. Kern, J. R. Ledsam, M. K. Schmid, K. Balaskas, E. J. Topol, L. M. Bachmann, P. A. Keane and A. K. Dennist, “A comparison of deep learning performance against health-care professionals in detecting diseases from medical imaging: a systematic review and meta-analysis,” Lancet Digital Health, vol. 1, no. 6, pp. e271-e297, 2019.
[8] H. Kriplani, B. Patel and S. Roy, “Prediction of Chronic Kidney Diseases Using Deep Artificial Neural Network Technique,” Lecture Notes in Computational Vision and Biomechanics, vol. 1, no. 1, pp. 179-187, 2019.
[9] M. Ahmad, V. Tundjungsari, D. Widianti, P. Amalia and U. Rachmawati, “Diagnostic decision support system of chronic kidney disease using support vector machine,” Second International Conference on Informatics and Computing (ICIC), vol. 1, no. 1, pp. 1-4, 2017.
[10] B. Ravindra, N. Sriraam and M. Geetha, “Classification of non-chronic and chronic kidney disease using SVM neural networks,” International Journal of Engineering and Technology, vol. 7, no. 1, pp. 191-194, 2018.
[11] International Conference on Healthcare Informatics (ICHI), vol. 1, no. 1, pp. 262-270, 2016.
[12] S. Vijayaranis and S. Dhayanan, “Kidney Disease Prediction Using SVM and ANN Algorithms,” International Journal of Computing and Business Research (IJCBR), vol. 6, no. 2, 2015.
[13] National Kidney Foundation, “Diabetes - A Major Risk Factor for Kidney Disease | National Kidney Foundation,” 2015. [Online]. Available: https://www.kidney.org/atoz/content/diabetes. [Accessed 10 February 2020].
[14] National Institute of Diabetes and Digestive and Kidney Diseases (NIDDK), “Anemia in Chronic Kidney Disease | NIDDK,” 2014. [Online]. Available: https://www.niddk.nih.gov/health-information/kidney-disease/anemia. [Accessed 10 February 2020].
[15] Bel Marra Health, “Sodium-potassium balance in diet critical in chronic kidney disease prevention,” 2015. [Online]. Available: https://www.belmarrahealth.com/sodium-potassium-balance-in-diet-critical-in-chronic-kidney-disease-prevention/. [Accessed 10 February 2020].
[16] Oddappz, “Oddappzdemands,” 2018. [Online]. Available: https://oddappzdemands.wixsite.com/foodpanda-clone/single-post/20180918The-Growing-Need-for-On-Demand-Food-Delivery-Application. [Accessed 1 December 2019].
[17] M. Nauge, “Intertek,” 2019. [Online]. Available: https://www.intertek.com/food/webinar-series/e-commerce-challenges-regulation-trends/. [Accessed 1 December 2019].
[18] L. Wyndham, “How E-Commerce Is Transforming The Food Industry,” Total Food Service, 30 May 2018. [Online]. Available: https://totalfood.com/how-e-commerce-transforming-food-industry/. [Accessed 28 April 2020].