Comparative Review of Chronic Kidney Disease Using Machine Learning Approaches

Mr. Dammati Pavan Kumar[1], T. Geethika[2], S. Sai Bhavani Kartheek[3], T. Divya[4], M. Rajarajeswari[5]
[1]*Asst. Prof, Dept of CSE, Tirumala Engineering College, Narasaraopet, Guntur, A.P., India*
[2,3,4,5]*B. Tech Students, Dept of CSE, Tirumala Engineering College, Narasaraopet, Guntur, A.P., India*

ABSTRACT - Kidney failure is one of the chronic diseases that is becoming a major public health concern around the world. It is a condition in which the kidneys have been weakened and are unable to clean the blood as effectively as they can. Excess fluid and waste can lead to more health problems in our bodies, which can take a long time to diagnose. Furthermore, since there are no early signs, the disease is only discovered at a late critical end point. As this Chronic Disease becomes a problem in today's world, large-scale research is being conducted to predict the prevalence of this Chronic Disease using machine learning. With the use of various techniques such as Support Vector Machine, Decision Tree, Nave Bayes, Random Forest, and others, machine learning is playing a huge role in the healthcare system, such as identifying diseases and diagnosis, drug development and manufacturing, and smart health records. Based on evidence from numerous research papers, this paper compares the accuracy of pre-existing techniques for predicting chronic kidney disease. Furthermore, this analysis takes into account the various attributes from either an existing database or a real-world database by using a variety of machine learning techniques. It is concluded that using machine learning on real-world datasets with all possible attributes yields an accurate estimate of the occurrence of chronic kidney disease.

Keywords: Chronic Kidney, Machine learning, SVM

I. INTRODUCTION

Kidney is a very important part of our body; in fact, we cannot live without it because it filters waste and excess fluid from our bodies. However, kidney failure is now a very popular health problem around the world. It is a chronic condition, which means it lasts a long time or takes a long time to diagnose [1]. Chronic Kidney Disease is a condition in which the kidneys are weakened and are unable to clean the blood as well as they can. Excess fluid and waste can cause additional health problems in our bodies, and this problem threatens people all over the world [2]. As kidney disease progresses and causes permanent damage to our bodies, it can eventually lead to kidney failure, necessitating dialysis or a kidney transplant for survival. Chronic kidney disease is estimated to affect one out of every ten people in India. Early detection and treatment can contribute to the prevention of kidney failure. The best way to assess kidney function or predict kidney disease stages is to monitor the Glomerular Filtration Rate (GFR) on a regular basis [3]. GFR is determined by taking into account age, gender, race, and blood creatinine level.

On the basis of GFR, there are five phases of chronic kidney disease:

Kidney disease is usually asymptomatic at first and can go undiagnosed before it progresses. It is sometimes referred to as the "silent illness" [4].

Stage 1: When the measured GFR is greater than 90 ml (GFR>90ml), the kidney function is considered normal.

Stage 2: If the approximate GFR is between 60 and 80 mL, mild kidney function loss begins.

Step 3: If the measured GFR is between 30-59 ml, the stage is considered minor to negative.

Stage 4: If the approximate GFR is between 15 and 29 ml, we can predict that the significant loss of kidney function has begun.

Step 5: If the measured GFR is less than 15 ml (GFR15 ml), this is the final stage of kidney failure that requires dialysis or a kidney transplant to survive.

Dialysis: When 90 percent of your kidney function fails, you need an artificial procedure to clear your blood from extra fluid; we have two options for this. Hemodialysis is the process of removing fluid from your blood using an artificial kidney[5]. However, when a tube pipe is fitted into a pistula in your arm or belly to purify your blood, this is referred to as peritoneal. Both procedures are carried out under the supervision of a doctor.

End-Stage Renal Disease (ESRD): A condition in which transplantation or dialysis are the only options for treating complete and stable kidney failure.

Glomerular Filtration Rate (GFR): The rate at which the kidneys remove waste and excess fluid, measured in millilitres per minute.

Proteinuria occurs when urine produces a high concentration of an unusual protein known as albumin.

Though kidney disease is asymptomatic in the early stages, if we can detect it early and take appropriate medical intervention, we can avoid chronic kidney disease[6]. If a specialist system detects kidney failure by examining the patient's symptoms, such as blood and urine tests [7]. Then it gives doctors enough time to treat or cure the patient's illness. Machine learning algorithms are now playing a significant and accurate role in the healthcare field.

II. RELATED WORK

Machine Learning is outstanding within the Artificial Intelligence sector, as it focuses on developing systems that work intelligently, and it has also covered the most real-life effects for business [8]. The primary purpose of ML is to allow machines to operate in self-learning mode without being specifically programmed. With the aid of its algorithm implementation, it makes programmes capable of learning, growing, and changing on their own when exposed to new data. If the observable output of a task assigned by a computer programme improves when it gains more and more knowledge, this is referred to as machine learning. Similarly, when a machine makes decisions and makes predictions based on data, this is also referred to as machine learning.

Machine Learning is primarily used to solve classification, regression, and clustering problems [9]. The methodology we will use is primarily determined by the type and category of data we will provide to our model. To implement machine learning algorithms, the techniques available are Supervised, Unsupervised, Semi-Supervised, and Reinforcement.

The technological world's latest advances are allowing us to create a technology that will assist physicians in detecting chronic diseases at an early stage and curing them. Machine learning is performing admirably in the healthcare system. There are numerous applications of machine learning in the healthcare field, including disease identification and diagnosis, drug development and production, medical imaging diagnoses, smart health records, and improved radiotherapy [8]. There are many chronic illnesses in the world, such as cancer, lung disease, heart disease, kidney failure, and so on. With no early signs and people just being aware of the problem at the end point, machine learning is actually offering a plethora of solutions to early detect and cure your problem.

The health protection system in machine learning is dependent on the computer's broad computational power and the logic competence of the doctor [13]. Both doctors and machines look for patterns, but doctors cannot estimate each patient's pulse like an efficient computer can. As a result, the computer will complete all of these tasks and show the results to the doctor for approval. In healthcare, a decision-support system is used. Decision support systems assist doctors in early disease exploration by maintaining understanding of health issues and background information of particular patients [12]. It keeps track of the insurance policy, the refund details, the fee, the account receivable, and the account payable. This system also aids in the identification of the patient's situation, and the doctor will then advise you about when and how to use the medication to treat the illness.

III. PROPOSED ARCHITECTURE

To strengthen the study and help fill the research gap found by the systematic review, an experiment based on CKD diagnoses in developed countries was carried out. Based on a previous analysis, the experiment used a dataset of subjects with or without CKD from the University Hospital Prof. Alberto Antunes at the Federal University of Alagoas (UFAL) in Brazil. To identify the CKD risk, a subset of the techniques defined when answering RQ1.1, namely, RF, NB, J48 decision tree, KNN, SVM, and MLP neural network, were used. This subset of machine learning techniques was also chosen in response to RQ1.3. Figure 1 depicts the planned architecture.

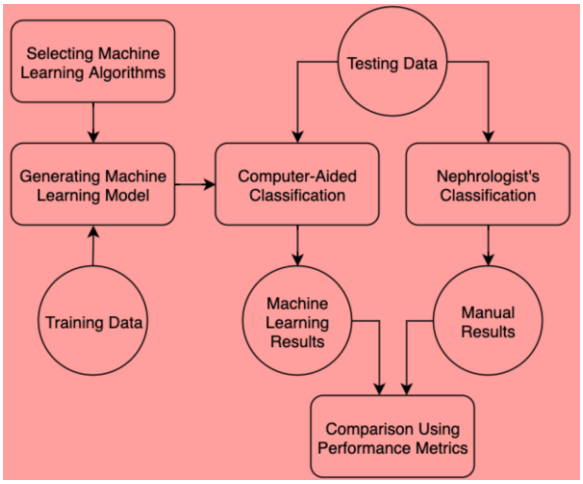

Figure 1. Proposed Architecture

IV. RESULTS AND OBSERVATION

The following research compares the accuracy of various techniques for predicting the existence of chronic kidney disease in a person, using different attributes from either a pre-existing database or a real-life dataset. Figure 2 depicts a comparison of the accuracy of SVM and DT algorithms with 14 attributes from 400 patient records downloaded from the UCI Repository[18]. SVM produces the highest results, with 96.7 percent accuracy, but the data strength is not higher due to size and missing aspects.

Figure 3 depicts a comparison of the accuracy of the LOGR, SVM, MLP, and RPART algorithms with 24 attributes of a dataset downloaded from the UCI Repository. MLP provides the best results with 99.5 percent accuracy, but it

only finds the association between input parameters to get the desired result, for example, creatinine and urea have a close relationship, and the value of most of the attributes is ignored to obtain better accuracy.

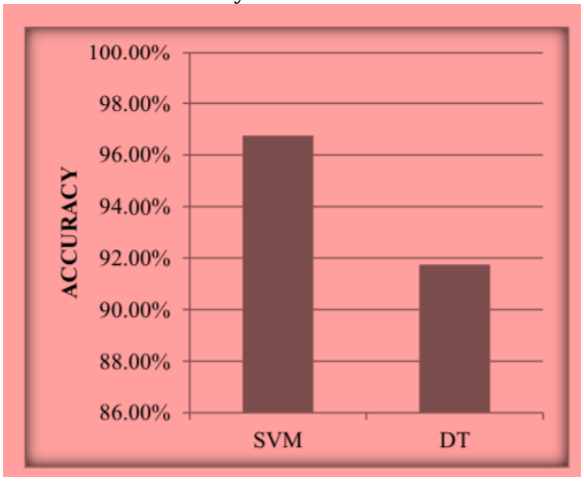

Figure 2. Ccomparison of SVM and DT techniques

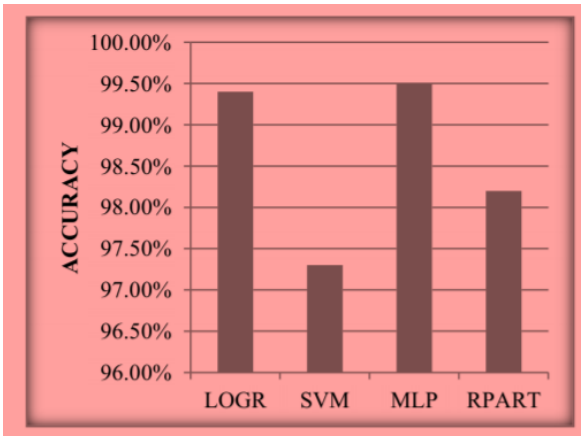

Figure 3. Ccomparison of Accuracy

V. CONCLUSION

This paper presents a Comparative Analysis of Machine Learning Techniques focused on a chronic kidney disease dataset of various attributes from either an existing database or a real-world database. Support Vector Machine (SVM), Decision Tree (DT), Nave Bayes (NB), Random Forest (RF), Recurrent Neural Network (RNF), and other machine learning algorithms are used. The study looks at how datasets influence an algorithm's accuracy level. For example, with just 14 attributes, SVM achieves an accuracy of 96.7 percent, while MLP achieves an accuracy of 99.5 percent with 24 attributes. However, in order to achieve greater precision, the importance of the majority of the attributes is ignored[8]. In such a scenario, certain characteristics that are influenced by factors such as prescription medications influencing the patient's blood pressure may have a high chance of being incorrectly predicted. As a result, working with real-world data with all possible attributes for chronic kidney disease prediction is more reliable, as shown in figure 6, with all metabolite features on the basis of blood samples from randomised hospitals, RNN provides the best result with 95.6 percent precision.

VI. REFERENCES

[1] A. C. Webster, E. V. Nagler, R. L. Morton, and P. Masson, "Chronic Kidney Disease," The Lancet, vol. 389, no. 10075, pp. 1238–1252, March 2017, 10.1016/S0140-6736(16)32064-5.

[2] E. J. Lamb, A. S. Levey, and P. E. Stevens, "The kidney disease improving global outcomes (kdigo) guideline update for chronic kidney disease: Evolution not revolution," Clinical Chemistry, vol. 59, no. 3, pp. 462–465, March 2013, 10.1373/clinchem.2012.184259.

[3] National Institute for Health and Care Excellence, "Chronic kidney disease in adults: assessment and management," Clinical guideline CG182, Jan. 2015.

[4] L. A. Inker, B. C. Astor, C. H. Fox, T. Isakova, J. P. Lash, C. A. Peralta, M. K. Tamura, and H. I. Feldman, "Kdoqi us commentary on the 2012 kdigo clinical practice guideline for the evaluation and management of ckd," American Journal of Kidney Diseases, vol. 63, no. 5, pp. 713–735, March 2014, 10.1053/j.ajkd.2014.01.416.

[5] M. Kinaan, H. Yau, S. Martinez, and P. Kar, "Concepts in diabetic nephropathy: From pathophysiology to treatment," Journal of Renal and Hepatic Disorders, vol. 1, no. 2, pp. 10–24, June 2017, 10.15586/jrenhep. 2017.17.

[6] R. C. C. Sesso, A. A. Lopes, F. S. Thome, J. R. Lugon, and E. A. Burdmann, "Brazilian Dialysis Census, 2009," Brazilian Journal of Nephrology, vol. 32, no. 4, pp. 380–384, Dec. 2010, 10.1590/S0101- 28002010000400007.

[7] M. Eichelberg, T. Aden, J. Riesmeir, A. Dogac, and G. B. Laleci, "2010 report of the Brazilian dialysis census," Brazilian Journal of Nephrology, vol. 33, no. 4, pp. 442–447, Dec. 2011, 10.1590/S0101-28002011000400009.

[8] R. C. C. Sesso, A. A. Lopes, F. S. Thome, J. R. Lugon, and C. T. Martins, "Brazilian Chronic Dialysis Survey 2016," Brazilian Journal of Nephrology, vol. 39, no. 3, pp. 380–384, Sept. 2017, 10.5935/0101-2800.20170049.

[9] F. S. Thome, R. C. Sesso, A. A. Lopes, J. R. Lugon, and C. T. Martins, "Brazilian chronic dialysis survey 2017," Brazilian Journal of Nephrology, vol. 41, no. 2, pp. 1–7, March 2019, 10.1590/2175-8239-jbn-2018-0178.

[10] Brazilian Ministry of Health. (2016). Expended amounts. [Online]. Available: http://portalarquivos.saude.gov.br/images/pdf/2016/abril/04/Valoresgastos. pdf

[11] G. Abraham, S. Varughese, T. Thandavan, A. Iyengar, E. Fernando, S. A. Naqvi, R. Sheriff, H. Ur-Rashid, N. Gopalakrishnan, and R. K. Kafle, "Chronic kidney disease

hotspots in developing countries in South Asia," Clinical Kidney Journal, vol. 9, no. 1, pp. 135–141, Feb. 2016, 10.1093/ckj/sfv109.

[12] F. S. Meirelles. (2017). 28th annual survey of administration and use of information technology in companies. [Online]. Available: eaesp.fgvsp.br/ensinoeconhecimento/centros/cia/pesquisa

[13] E. T. Lwoga, and R. Z. Sangeda, "ICTs and development in developing countries: A systematicreview of reviews," The Electronic Journal of Information Systems in Developing Countries, vol. 85, no. 1, pp. 1–17, Jan. 2019, https://doi.org/10.1002/isd2.1206.

[14] D. W. Cockcroft, and M. H. Gault, "Prediction of creatinine clearance from serum creatinine," Nephron, vol. 16, no. 1, pp. 31–41, 1976, 10.1159/000180580.