The Ciência & Engenharia - Science & Engineering Journal
ISSN: 0103-944X
Volume 11 Issue 1, 2023
pp: 276 - 281

Chronic Kidney Diseases Detection Using machine Learning

Prof. Swati Salunkhe , Divya Gaikwad , Poonam Humbe , Shubham Kale , Sakshi Lokhande

Abstract – Chronic kidney disease (CKD) and chronic renal disease (CRD) (CKD). Chronic renal disease refers to illnesses that affect your kidneys and reduce their ability to keep you healthy. Consequences include nerve damage, high blood pressure, anaemia, weak bones, and a lack of nourishment. Early detection and treatment can frequently stop chronic renal disease from getting worse. Data mining is the technique of obtaining knowledge from huge datasets. Utilizing previous data to find trends and guide decisions going forward is the aim of data mining. This task is the result of the convergence of a number of recent trends, including the declining cost of large data storage devices, the increasing simplicity of data collection over networks, the expansion of dependable and efficient machine learning algorithms, and the declining cost of computational power, which enables the use of computationally intensive techniques for decision-making. Machine learning has already produced useful applications in fields like assessing results from medical research, spotting fraud, spotting bogus users, etc.For the purpose of predicting chronic diseases, various data mining categorization methodologies and machine learning algorithms are used. The goal of this study is to develop a new decision-support system for forecasting chronic renal disease. This study compares the accuracy, precision, and execution time of Support Vector Machine (SVM) and K-Nearest Neighbor (KNN) classifiers for the prediction of CKD.

I. Introduction

As every human being is aware, the kidney is a vital organ. To put it simply, the kidney and excretion system gather and dispose of all the harmful and unneeded material produced by the body. It is a risky kidney condition that results in a slow decline in kidney function. Over a number of years, ckd causes a gradual and sporadic decline of renal function. The following symptoms may appear in a patient with ckd who is not treated for it in the early stages: Blood pressure, anaemia, weekbeans, inadequate nourishment, health problems, and nerve damage, reduced immune response because advanced stages can create severe accumulations of fluids, electrolytes, and waste products in blood and body.

Because CKD develops slowly and has symptoms that are not unique to the condition, it is crucial to identify it early. Machine learning can be useful in this situation to determine if the patient has CKD or not as some patients have no symptoms at all. In order to identify chronic renal disease and train a prediction model, machine learning uses old CKD patient data.

The Ciência & Engenharia - Science & Engineering Journal
ISSN: 0103-944X
Volume 11 Issue 1, 2023
pp: 276 - 281

A. MOTIVATION:

Acute renal failure can have an impact on virtually every organ in your body. Early diagnosis of CKD should be advantageous as it allows doctors to start treating patients effectively for moderate disease, reducing kidney function decline, and delaying or preventing renal failure.

B. PROBLEM DEFINITION

Machine learning with an SVM algorithm for the diagnosis of chronic renal illness.

C. PROJECT OBJECTIVE

- To watch SVM classifier evaluate and forecast CKD
- To create a trustworthy machine learning model that accurately predicts CKD.
- Using support vector machines to anticipate kidney diseases
- Data mining is used to draw conclusions from datasets and to uncover hidden information.

D. SCOPE OF PROJECT:

The scope projects have extensive management experience with older multi-morbidity patients at risk for renal deterioration and associated clinical consequences, such as alterations in functional status.

II. **System Architecture**:

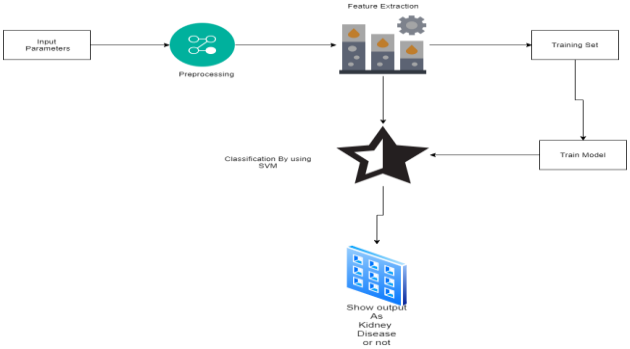


Fig. 1 System Architecture

III. **Algorithm**:

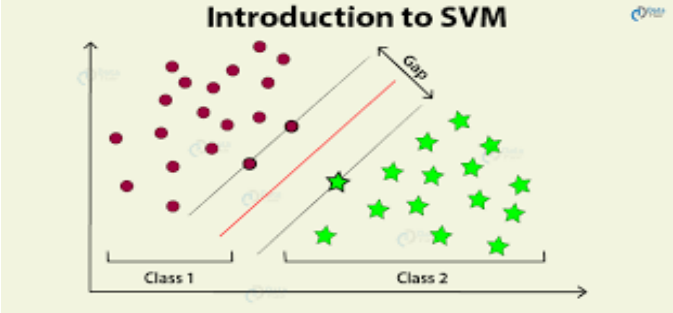
Support-vector machines (SVMs, also known as support-vector networks) are learning algorithms and supervised learning models that analyse data used for regression analysis and classification. An SVM training procedure creates a model from a series of training examples that have previously been identified as belonging to one of two categories, resulting in a non-probabilistic binary linear classifier that can categorise fresh examples into one of two groups (although methods such as Platt scaling exist to use SVM in a probabilistic classification setting). An SVM model maps the instances as points in space with as much room as feasible separating

The Ciência & Engenharia - Science & Engineering Journal
ISSN: 0103-944X
Volume 11 Issue 1, 2023
pp: 276 - 281

the examples of the different categories.Then, based on the side of the gap on which they fall, new samples are projected into that same area and predicted to belong to a category. Two different SVMs

Data that can be separated into two classes by a single straight line are used for linear SVM. The classifier used is referred to as a Linear SVM classifier, and this type of data is known as linearly separable data.

Non-linear SVM is used for non-linearly separated data, which indicates that a dataset is non-linear if it cannot be classified using a straight line, and a non-linear SVM classifier is utilised.



IV. Methodology

We pre-process the data and decide which components of our method are most important. Only those features are maintained; the rest are eliminated. Among the models used to train and test model are SVM, Random forest, and hybrid neural network models. The effectiveness of various algorithms is compared using various performance metrics.

Dataset - Provide dataset (By this, we mean that the collected data should be uniform and understandable for a machine that doesn't see data in the same way that people do.)

Preprocessing -

Real-world data typically includes noise, missing values, and may be in an undesirable format, making it impossible to build machine learning models on it directly..

Feature Extraction: It tries to decrease the amount of features in a dataset by generating new features from the ones that already exist (and then discarding the original features). The majority of the information in the original collection of features should then be summarised by this new, smaller set of features.

Feature Selection

```
df.head()
```

	sg	al	sc	hemo	pcv	htn	classification
0	1.020	1.0	1.2	15.4	28	1	0
1	1.020	4.0	0.8	11.3	22	0	0
2	1.010	2.0	1.8	9.6	15	0	0
3	1.005	4.0	3.8	11.2	16	1	0
4	1.010	2.0	1.4	11.6	19	0	0

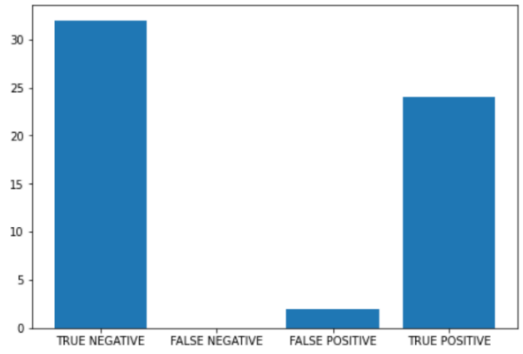
Attributes selected through feature selection

The Ciência & Engenharia - Science & Engineering Journal
ISSN: 0103-944X
Volume 11 Issue 1, 2023
pp: 276 - 281

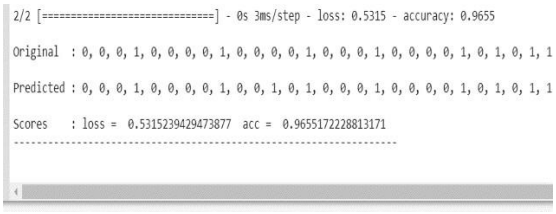
Classification - The Classification algorithm, which uses supervised learning to categorise new observations in light of training data, is used to recognise new observations. In classification, a programme makes use of the dataset or observations that are provided to learn how to categorise fresh observations into various classes or groups.

	CHRONIC KIDNEY DISEASE		
TEST RESULT		Predicted NO	Predicted YES
	Actual NO	31	0
	Actual YES	2	25

Confusion Matrix



Confusion matrix results



Comparison of original value and the predicted value with accuracy

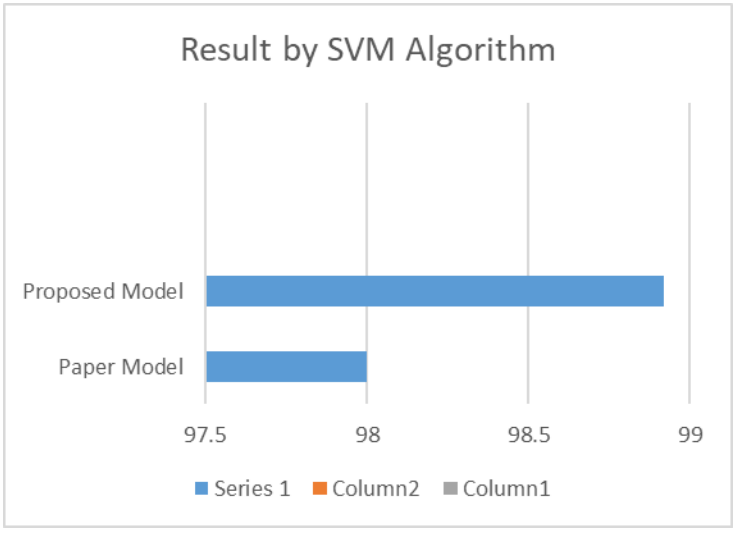
Advantages :

There are no significant symptoms of CKD prior to the disease's progression, but the patient may experience symptoms including fatigue and a loss of appetite.

The Ciência & Engenharia - Science & Engineering Journal
ISSN: 0103-944X
Volume 11 Issue 1, 2023
pp: 276 - 281

V. Expected Results

To gather the dataset age, we employ the following representation. - age bp - sg for blood pressure Al's specific gravity the albumin
sugar rbc, su cells in the red blood
pc - pus cell clumping - pus cell pcc
ba Bgr bacterium - Random blood glucose
bu - urea blood
sc - Sod serum creatinine - sodium kettle potassium hemo and PCV haemoglobin - wc for packed cell volume - wbc rc (white blood cell count) - count of red blood cells
Hypertension definition: HTN - Type 2 diabetes cad coronary artery disease
pedal edoema, hunger, and peane - class for anaemia
The information would be collected in India during a two-month period with 25 features (eg, RBC count, WBC count, etc). The classification, which can be either "ckd" or "not ckd" (ckd stands for chronic renal disease), is the target. 400 rows are present. The data needs to be cleaned because it contains NaNs and needs to have the numeric features made to float. In essence, we will specify to destroy ALL ROWS containing Nans with no threshold, which means that every row containing even one NaN will be eliminated.



VI. Conclusion and Future Work

As we have already seen, data mining and machine learning are being used in the medical industry. A SVM is utilised in this study to predict CKD. Even so, the classifiers were effective at predicting other illnesses. In this project, the performance of various SVM algorithms is compared in terms of their ability to predict Chronic Kidney Disease. Out of one classifiers, according to our analysis. SVM classifier outperformed the competition. The likelihood of CKD prediction has increased.

Reference

[1]http://www.iosrjournals.org/iosrjce/papers/Vol10

[2]S. Sadhana Rao," Sixth Sense Technology", Proceedings of the International Conference on Communication and Computational Intelligence– 2010, pp.336-339.

[3]Game P. M., Mahajan A.R,"A gestural user interface to Interact with computer system ", International Journal on Science and Technology (IJSAT) Volume II, Issue I, (Jan.- Mar.) 2011, pp.018 – 027.

[4]International Journal of Latest Trends in Engineering and Technology Vol. (7)Issue(4), pp.055-062.

[5]Dhanikonda, S. R., Sowjanya, P., Ramanaiah, M. L., Joshi, R., Krishna Mohan, B. H., Dhabliya, D., & Raja, N. K. (2022). An efficient deep learning model with interrelated tagging prototype with segmentation for telugu optical character recognition. Scientific Programming, 2022 doi:10.1155/2022/1059004

[6]Imperial Journal of Interdisciplinary Research (IJIR) Vol-3, Issue-4, 2017.

[7]Christy, A., Vaithyasubramanian, S., Mary, V.A., Naveen Renold, J. (2019)," Artificial intelligence based automatic decelerating vehicle control system to avoid misfortunes ", International Journal of Advanced Trends in Computer Science and Engineering, Vol. 8, Issue. 6, Pp. 3129-3134.

[8]G. M. Gandhi and Salvi, "Artificial Intelligence Integrated Blockchain For Training Autonomous Cars," 2019 Fifth International Conference on Science Technology Engineering and Mathematics (ICONSTEM), Chennai, India, 2019, pp. 157-161. Jesudoss A., and Subramaniam, N.P., "EAM: Architecting Efficient Authentication Model for Internet Security using Image-Based One Time Password Technique"'

[9]Indian Journal of Science and Technology, Vol. 9 Issue 7, Feb. 2016, pp. 1-6.

[10] Praveena, M.D.A., Eriki, M.K., Enjam, D.T.," Implementation of smart attendance monitoring using open-CV and python ",Journal of Computational and Theoretical Nanoscience, Vol. 16, Number 8 pp:3290-3295 · August 2019 .

[11] Jain, V., Beram, S. M., Talukdar, V., Patil, T., Dhabliya, D., & Gupta, A. (2022). Accuracy enhancement in machine learning during blockchain based transaction classification. Paper presented at the PDGC 2022 - 2022 7th International Conference on Parallel, Distributed and

[12] Grid Computing, 536-540. doi:10.1109/PDGC56933.2022.10053213

[13] M.S.Roobini, DrM.Lakshmi,(2019),"Classification of Diabetes Mellitus using Soft Computing and Machine Learning Techniques", International Journal of Innovative Technology and Exploring Engineering, ISSN: 2278-3075, Volume-8, Issue- 6S4