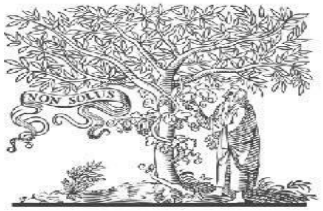
International Journal for Innovative Engineering and Management Research
PEER REVIEWED OPEN ACCESS INTERNATIONAL JOURNAL
www.ijiemr.org

COPY RIGHT



ELSEVIER SSRN

IJIEMR Transactions, online available on 10th Apr 2023. Link :http://www.ijiemr.org/downloads.php?vol=Volume-12&issue=Issue 04

10.48047/IJIEMR/V12/ISSUE 04/100

Title **CHRONIC KIDNEY DISEASE DETECTION**

Volume 12, ISSUE 04, Pages: 807-813

Paper Authors

Y. Venkata Narayana, Mohammad. Sana Minaaj, Muvva. Jhansi, Palavai. Abhinaya, Mukkapati. Sivathmika Sri

USE THIS BARCODE TO ACCESS YOUR ONLINE PAPER

To Secure Your Paper As Per UGC Guidelines We Are Providing A Electronic Bar Code

International Journal for Innovative Engineering and Management Research
PEER REVIEWED OPEN ACCESS INTERNATIONAL JOURNAL
www.ijiemr.org

Chronic Kidney Disease Detection

Y. Venkata Narayana, Assistant Professor, Department of IT,
Vasireddy Venkatadri Institute of Technology, Nambur, Guntur Dt., Andhra Pradesh.

Mohammad. Sana Minaaj, **Muvva. Jhansi**, **Palavai. Abhinaya**,
Mukkapati. Sivathmika Sri
UG Students, Department of IT,
Vasireddy Venkatadri Institute of Technology, Nambur, Guntur Dt., Andhra Pradesh.

Abstract

Chronic kidney disease (CKD) is still a significant public health issue despite improvements in surgical therapy and medication. Researchers from all around the world have lately created high-performance ways for diagnosis, treatment, and preventative therapy due to the rising prevalence of CKD. These solutions can be more effective if the users are aware of the aspects that are pertinent to the issue. In addition to clinical evaluation, medical data analysis for patients can aid healthcare professionals in the early diagnosis of diseases. Although numerous attempts have been made to improve the effectiveness of the intelligent algorithms that analyse health data to predict CKD, there is still room for improvement. This work intends to give a comprehensive categorization and prediction model for kidneyrelated illnesses. These loss functions, which are employed as a part of a modified Deep Belief Network (DBN) classification approach, are Categorical Cross-entropy activation function and SoftMax activation function, respectively. The proposed model outperforms earlier models since it has a 98.5% accuracy and an 87.5% sensitivity for predicting chronic renal illness (CKD). Modern deep learning algorithms can enhance clinical judgement and enable early prediction of CKD and related stages, according to a data analysis of the available information. This approach might aid in halting the progression of renal disease.

Keywords: Restricted Boltzmann Machine, CKD, Contrastive Divergence

Introduction

When kidney function is hampered by a problem or condition, chronic kidney disease results. The kidney damage worsens over a period of months or years.

Despite improvements in surgical care and therapy, chronic kidney disease (CKD) remains a public health issue. Scientists from all over the world are especially interested in creating highperformance approaches for the disease's detection, treatment, and prevention therapy because

to the recent increase in CKD. Learning the characteristics relevant to the issue can improve performance. Medical data analysis for the patients can help the healthcare partners in early disease prediction in addition to the clinical evaluation. Due to its high mortality rate, chronic kidney disease (CKD) has attracted a lot of attention. According to the World Health Organization, chronic diseases are becoming more prevalent in developing countries (WHO). Renal failure results from later-stage CKD, which is curable in the early stages. In 2016, 336 million men and 417 million women died from chronic renal illness worldwide, accounting for 753 million deaths overall. Kidney illness is referred described as a "chronic" condition since it worsens over time and affects how the urinary system works. Other health problems, such as high and low blood pressure, diabetes, nerve damage, and bone problems, many of which have symptoms with cardiovascular disease, manifest as a result of the buildup of waste materials in the blood. Patients with CKD are at risk for diabetes, hypertension, and heart disease (CVD). In the late stages of CKD, patients have side effects that impair the immune and neurological systems. In underdeveloped countries, patients may already be in severe stages, requiring dialysis or kidney transplants.

Literature Survey

[1] **Abdelaziz et al**. presented a hybrid CKD-based Cloud-IOT intelligent model employing Linear Regression (LR) and Neural Network (NN). While NN is used to make predictions, LR is used to identify key elements that have an impact on CKD. A hybrid cognitive framework was also established on Windows Azure as an example of a cloud technology environment to forecast CKD and aid clinicians in smart cities.

[2] **Ani et al.** Using classification approaches including probabilitybased Naive Bayes, neural networkbased Back Propagation (BPN), LDA classification, node-based decision trees, and lazy learners, a system for promoting medical decision-making has been built and developed.

[3] **Anupama et al**. developed a Grab Cut segmentation technique-based intracerebral hemorrhage (ICH) diagnosis model that combined Gabor noise removal with synergic deep learning (SDL) for feature extraction to enhance image quality.

[4] **Arora and Sharma** employed three classification techniques to diagnose CKD at the earliest possible stage, Naïve Bayes, J48 and SOM. Using WEKA, a data mining tool, experiments with these methods on these datasets were carried out. The best algorithm for a trustworthy and early diagnosis of chronic renal disease was ultimately determined to

be J48. Of the three early-stage algorithms, it had the best accuracy.

[5] **Başar et al.** used the Adaboost ensemble learning technique to provide a diagnosis for chronic renal disease. Classifiers built on decision trees were used in the diagnosis procedure. Area under the curve, kappa, mean absolute error, and root mean squared error were used to calculate the category efficiency (AUC). The results showed that the Adaboost ensemble training strategy gave classification results that were superior to those produced by human classification.

[6] **Besra and Majhi** started with preprocessing, categorise the events as CKD or NOTCKD, and calculate GFR stage %. The emergence of many classifiers, including Naive-Bayes, SOM, IB1, VFI, Multi-classifier, and Random Forest, marked the beginning of the classification approach. Finally, the rate and phases of renal activity were assessed using the GFR test method.

[7] **Boukenze et al**. used a number of learning methods to predict renal illness, including SVM, MLP, DT (C4.5), Bayesian Network, and KNN. According to the findings, MLP and C4.5 have the highest accuracy rates. Based on a variety of criteria, they compared these algorithms and chose the one that was the most effective.

[8] **Chaitanya and Rajesh et al**. Attempts were made to incorporate the K-nearest neighbor, (ANN+GA), and (ANN+GSA) approaches. In terms of accuracy, sensitivity, and specificity, GA-ANN algorithms surpassed the other algorithms in terms of classification efficacy.

[9] **Dutta and Bandyopadhyay** For the purpose of identifying CKD patients, a neural network with ten folds cross validation approach was designed in preference to decision tree, support vector machine, gradient boosting classifier, and K Nearest Neighbors. According to experimental data based on accuracy, f1 score, kappa score, and MSE, the neural network with 10 folds cross validation model showed the best performance.

[10] **Eroğlu and Palabaş** In contrast to decision trees, support vector machines, gradient boosting classifiers, and K Nearest Neighbors, a neural network technique using ten folds cross validation was created to identify the CKD patients. Using experimental data based on accuracy, f1 score, kappa score, and MSE, the neural network with 10 folds cross validation model showed the best performance.

[11] **Gharibdousti et al.** Decision Tree, Linear Regression, Super Vector Machine, Naive Bayesian, and Neural Network were used as categorization techniques.

International Journal for Innovative Engineering and Management Research
PEER REVIEWED OPEN ACCESS INTERNATIONAL JOURNAL
www.ijiemr.org

[12] **Gupta et al.** constructed a deep learning classification system to extract the most useful and important characteristics from the chronic kidney disease dataset using the stacked autoencoder model and the SoftMax classifier for prediction.

[13] **Jojoa et al.** combining the identification criteria (CKD). Two crucial aspects, including prognosis, enabled for the development of a better and more precise treatment strategy. These factors included current research techniques that allowed for a systematic study of the illnesses and gave quantifiable prognostic risk factors.

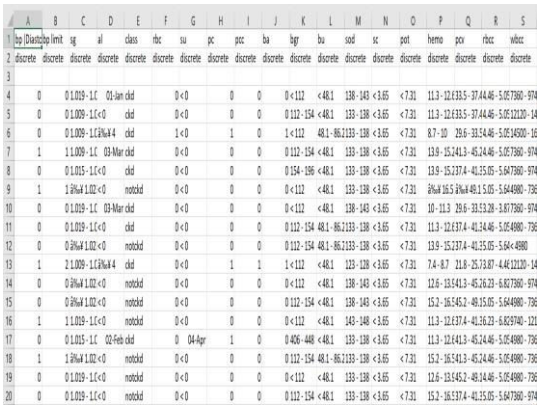
Problem Identification

"Chronic renal disease" is a word that is frequently used to describe a wide range of kidney illnesses. Chronic renal disease is an additional name for it. The disease affects five to ten percent of people worldwide. A global health issue is chronic renal disease. One of the main reasons why a higher percentage of chronic kidney disease cases occur in developing and underdeveloped countries than in developed ones, where most people receive routine check-ups and diagnoses, is that the majority of cases of chronic kidney disease go undiagnosed or are later diagnosed. If machine-based learning algorithms are used to quickly and accurately identify chronic renal disease, a clinician will be able to attend to and treat more patients in less time. This is different from the situation where a clinician must carry out the diagnosis process totally by hand.

Methodology

Data Collection

The patient's renal disease file served as the work's primary data source. This information came from UCI. It offers details on 400 individuals, 250 of whom have CKD and 150 of whom have not, and has 25 qualities—14 nominal, 11 numerical, and 1 class attribute.



	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S
1	bp (Diastolic limit	sg	al	class	rbc	su	pc	pcc	ba	bgr	bu	sod	sc	pot	hemo	pcv	rbcc	wbcc	
2	discrete	discrete	discrete	discrete	discrete	discrete	discrete	discrete	discrete	discrete	discrete	discrete	discrete	discrete	discrete	discrete	discrete	discrete	discrete
3																			
4	0	0 1.019 - 1.(	01-Jan	ckd		0 < 0		0	0	0 < 112	< 48.1	138 - 143	< 3.65	< 7.31	11.3 - 12.6	33.5 - 37.4	4.6 - 5.05	7360 - 974	
5	0	0 1.009 - 1.(	<0	ckd		0 < 0		0	0	0 112 - 154	< 48.1	133 - 138	< 3.65	< 7.31	11.3 - 12.6	33.5 - 37.4	4.6 - 5.05	12120 - 14	
6	0	0 1.009 - 1.(	3-Mar	ckd		1 < 0		1	0	1 < 112	48.1 - 86.2	133 - 138	< 3.65	< 7.31	8.7 - 10	29.6 - 33.5	4.6 - 5.05	14500 - 16	
7	1	1 1.009 - 1.(	03-Mar	ckd		0 < 0		0	0	0 112 - 154	< 48.1	133 - 138	< 3.65	< 7.31	13.9 - 15.2	41.3 - 45.2	4.6 - 5.05	7360 - 974	
8	0	0 1.015 - 1.(	<0	ckd		0 < 0		0	0	0 154 - 196	< 48.1	133 - 138	< 3.65	< 7.31	13.9 - 15.2	37.4 - 41.2	5.05 - 5.64	7860 - 974	
9	1	1 1.(	1.02	notckd		0 < 0		0	0	0 < 112	< 48.1	133 - 138	< 3.65	< 7.31	16.5 - 18.(	48.1 - 5.05	- 5.64	4980 - 736	
10	0	0 1.019 - 1.(	03-Mar	ckd		0 < 0		0	0	0 < 112	< 48.1	138 - 143	< 3.65	< 7.31	10 - 11.3	29.6 - 33.5	3.2 - 3.67	7360 - 974	
11	0	0 1.019 - 1.(	<0	ckd		0 < 0		0	0	0 112 - 154	48.1 - 86.2	133 - 138	< 3.65	< 7.31	11.3 - 12.6	37.4 - 41.2	4.6 - 5.05	4980 - 736	
12	0	0 1.(	1.02	notckd		0 < 0		0	0	0 112 - 154	48.1 - 86.2	133 - 138	< 3.65	< 7.31	13.9 - 15.2	37.4 - 41.2	5.05 - 5.64	4980	
13	1	2 1.009 - 1.(	3-Mar	ckd		0 < 0		1	1	1 < 112	< 48.1	123 - 128	< 3.65	< 7.31	7.4 - 8.7	21.8 - 25.7	3.87 - 4.4	12120 - 14	
14	0	0 1.(	1.02	notckd		0 < 0		0	0	0 < 112	< 48.1	138 - 143	< 3.65	< 7.31	12.6 - 13.5	41.3 - 45.2	6.2 - 6.82	7360 - 974	
15	0	0 1.(	1.02	notckd		0 < 0		0	0	0 112 - 154	< 48.1	138 - 143	< 3.65	< 7.31	15.2 - 16.5	45.2 - 48.1	5.05 - 5.64	4980 - 736	
16	1	1 1.019 - 1.(	<0	notckd		0 < 0		0	0	0 < 112	< 48.1	143 - 148	< 3.65	< 7.31	11.3 - 12.6	37.4 - 41.2	6.2 - 6.82	9740 - 121	
17	0	0 1.015 - 1.(	02-Feb	ckd		0	04-Apr	1	0	0 406 - 448	< 48.1	133 - 138	< 3.65	< 7.31	11.3 - 12.6	41.3 - 45.2	4.6 - 5.05	4980 - 736	
18	1	1 1.(	1.02	notckd		0 < 0		0	0	0 112 - 154	48.1 - 86.2	133 - 138	< 3.65	< 7.31	15.2 - 16.5	41.3 - 45.2	4.6 - 5.05	4980 - 736	
19	0	0 1.019 - 1.(	<0	notckd		0 < 0		0	0	0 < 112	< 48.1	133 - 138	< 3.65	< 7.31	12.6 - 13.5	45.2 - 48.1	4.6 - 5.05	4980 - 736	
20	0	0 1.019 - 1.(	<0	notckd		0 < 0		0	0	0 112 - 154	< 48.1	133 - 138	< 3.65	< 7.31	15.2 - 16.5	37.4 - 41.2	5.05 - 5.64	7360 - 974	

Data Pre-processing

Since it calls for preparing, extracting, and translating data into a machinereadable format, it is the most crucial task. Raw data makes forecasts impossible since it comprises incomplete information, imprecise information, and improper formats. Straightforward imputation, which converts the attribute's mode value and class to numerical values of "1" for ckd and "0" for non-ckd, respectively, was used to fill in some of the dataset's missing cells.

Proposed System

The suggested deep network prediction model for chronic renal disease contains six

layers. The first layer serves as the input and is made up of 24 nodes, whilst the second layer receives its input from the first layer and is made up of 13 nodes. 28 nodes make up the third layer, which receives input from the second layer. With the fourth layer, 8 nodes are employed, and these nodes take inputs from the third layer. The fourth layer serves as the fifth layer's input and has four nodes. Finally, since it uses binary classification, the output layer comprises two nodes. There are three stages in the model.

The input layer acquires the data during the initial step, sometimes referred to as the pre-training phase. The second stage, known as training, establishes the local optimal parameters after it constructs a standard RBM. The parameters are trained using a single continuous unsupervised learning approach. RBMs then appear to be stacked greedily, layer by layer, to reach the local optimal values, and the weighting factor estimate is completed in a single step. Underfitting, a problem that affects many other networks, can also be eliminated as a result of the significant training time reduction.

The second stage of fine-tuning, in which all network parameters are changed using the Back Propagation algorithm to get the global optimal, uses the up-down algorithm, a contrastive form of the wakesleep approach that is used to train the DBNs with labelled data.

The trained model with the optimal weights and biases is obtained during the third stage of classification (test), which enables accurate classification of data.

Implementation

Restricted Boltzmann Machine

RBM is a two-layer, non-partisan graphic design with symmetrical relationships defined by a weight matrix W between these two levels, a collection of visible units v, and a set of invisible units h. It describes the joint probability of the RBM of the visible units v and the hidden units h. The Contrastive Divergence (CD) training approach pre-trains the RBM using training data. As can be seen in, the stochastic steepest ascending method is used to make CDs. Before gathering training data for h0, equation initially assumes the hidden state of h0. Assume that v1 of the reconstruction is in h1 of the covert state. The conditional probabilities of the hidden state are shown for the visible state v.

Deep Belief Network

RBM that have been trained with CD are stacked to create interactive systems known as Deep Belief Networks. The greedy strategy for learning to operate with DBN training layer by layer is driven by the method that selects the best position for each layer, and the following stacked RBM layer employs the besttrained values to get the best location once more. A complicated hierarchical structure is used to generate the training data for DBN. DBN is an

International Journal for Innovative Engineering and Management Research
PEER REVIEWED OPEN ACCESS INTERNATIONAL JOURNAL
www.ijiemr.org

unsupervised learning method as opposed to neural networks for back propagation and vision. The pre-training portion of a DBN's training approach can alleviate this problem by choosing the initial weight values of the network using unsupervised learning-based training as opposed to selecting them at random.

Results and Conclusion

Actual Results of the work

Missing Value Imputation

We utilised mode imputation to estimate missing values in the original data set, maintaining data integrity and making sure no bias was added.

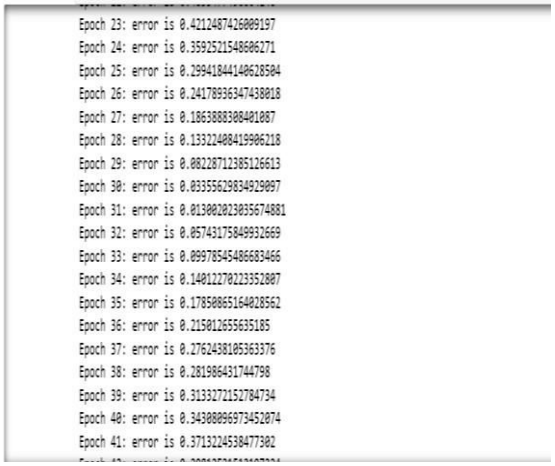
Analysis of the Results obtained

Using mode and median imputation, the outliers in the data set are eliminated and replaced for nominal and numerical properties. The top six features from this experiment are chosen using the chisquare test, which is used to choose the characteristics.

	white_blood_cell_count	blood_glucose_random	blood_urea	serum_creatinine	packed_cell_volume	albumin
0	7800.0	121.0	36.0	1.2	44.0	1.0
1	6000.0	122.0	18.0	0.8	38.0	4.0
2	7500.0	423.0	53.0	1.8	31.0	2.0
3	6700.0	117.0	56.0	3.8	32.0	4.0
4	7300.0	106.0	26.0	1.4	35.0	2.0
...	...	...	...	...	...	...
395	6700.0	140.0	49.0	0.5	47.0	0.0
396	7800.0	75.0	31.0	1.2	54.0	0.0
397	6600.0	100.0	26.0	0.6	49.0	0.0
398	7200.0	114.0	50.0	1.0	51.0	0.0
399	6800.0	131.0	18.0	1.1	53.0	0.0

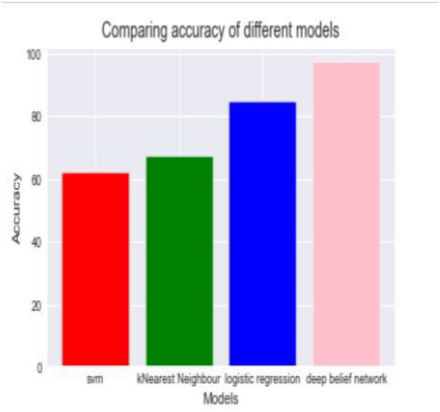
400 rows × 6 columns

Feature Selection



The accuracy obtained by using Deep Belief Network obtained is as follows

The accuracies obtained by using different algorithms is as follows.



Limitations

Methods of data processing that lack intelligence (mean/mode substitution). Averages are used to replace missing values.

examines it without considering relevant clinical information.

References

[1]https://www.ajkd.org/article/S0272-6386(13)00590-8/fulltext

[2]https://www.researchgate.net/publication/329521615_Chronic_Kidney_Disease_A_Predictive_model_using_Decision_Tree

[3]researchgate.net/profile/Sai-PrasadPotharaju/publication/325871632_Classi fication_of_nonchronic_and_chronic_kidney_disease_usin g_SVM_neural_networks/links/5b2a33d1a6fdcc72db4c86d0/Classification-of-nonchronic-and-chronic-kidney-diseaseusing-SVM-neural-networks.pdf

[4]https://ieeexplore.ieee.org/abstract/document/8819654