The Document is an individual work
Amateur research paper
Finished: 09/08/2020

Prediction of Chronic Kidney Disease Using Deep Learning

Yusif Ibrahimov[1]

[1]**Computer Science, French-Azerbaijani University, Baku, Azerbaijan**

ABSTRACT

During the research, 4-layer Neural Network was prepared based on the values of blood pressure(Bp), specific gravity(Sg), albumin level(Al), sugar level(Su), red blood cell level(Rbc), blood urea level(Bu), serum creatinine level(Sc), sodium level(Sod), potassium level(Pot), hemoglobin level(Hemo), white blood bell count level(Wbcc), red blood bell count level(Rbcc), and hypertension level(Htn) that taken from the results of 400 blood test reports, which can predict whether the patient has chronic kidney disease or not. Although the accuracy of the model is very high, the aim is to detect whether the tissue test(biopsy) or further tests are needed or not.

Keywords: Kidney Failure,Neural Networks , Diagnosis, Medicine

1 INTRODUCTION

Chronic Kidney Disease(CKD) is a serious health problem which is the gradual loss of kidney function. Annually, 2.4 million deaths are caused by CKD globally.
Although the exact diagnosis of CKD is blood and urine test, a significantly high number of patients are referring to the renal biopsy test to identify kidney disease. The renal biopsy is used to differentiate acute kidney disease from CKD. Since it requires the anesthesia and it is a costly and stressful procedure, to identify the loss of the kidney function, a blood test should have been done beforehand and only in the case of the confusion between the acute and chronic disease, a biopsy can be applied.
4-layer Deep Neural Network was developed to predict the existence of CKD by taking the blood test results as inputs. The example dataset received by the DNN:

	Bp	Sg	Al	Su	Rbc	Bu	Sc	Sod	Pot	Hemo	Wbcc	Rbcc	Htn
0	80.0	1.020	1.0	0.0	1.0	36.0	1.2	137.53	4.63	15.4	7800.0	5.20	1.0
1	50.0	1.020	4.0	0.0	1.0	18.0	0.8	137.53	4.63	11.3	6000.0	4.71	0.0
2	80.0	1.010	2.0	3.0	1.0	53.0	1.8	137.53	4.63	9.6	7500.0	4.71	0.0
3	70.0	1.005	4.0	0.0	1.0	56.0	3.8	111.00	2.50	11.2	6700.0	3.90	1.0
4	80.0	1.010	2.0	0.0	1.0	26.0	1.4	137.53	4.63	11.6	7300.0	4.60	0.0

The blood pressure can be measured with stethoscope by doctor or nurse.
The specific gravity or relative density were calculated gravimetrically from 25 healthy volunteers' blood and plasma. For whole blood it was determined as 1.0621 (95% confidence interval: 1.0652-1.0590) at 4^0C and 1.0506 (95% confidence interval: 1.0537-1.0475) at 37^0C.
Level of the blood sugar($C_{256}H_{387}N_{65}O_{79}S_6$) can be determined as a result of FBS (fast blood sugar) test with the help of the lancet.
Albumin ($C_{2936}H_{4624}N_{786}O_{110}S_{41}$) concentration determination refers to the decrease in the absorbance of a neutral buffered solution of bromcresol green when albumin is combined with the indicator, serum is diluted with a bromcresol green solution of sufficient concentration to allow an essentially linear shift in the absorbance of albumin.
The concentration of serum / plasma urea represents the equilibrium between the development of urea in the liver and the expulsion of urea by the kidneys in urine; therefore, elevated concentration of plasma / serum urea may be induced by increased output of urea, reduced removal of urea or a mixture of both. A blood urea nitrogen (BUN) check is used to figure out how well the kidneys function. The check is done by calculating the volume of nitrogen from urea in the body. Urea nitrogen is a waste product created in the liver when proteins are broken down by the body. The kidneys normally flush this waste away and urinate it off the body.
The rates of serum creatinine are determined by the amount of creatinine filtration in the kidneys, sex , age, muscle mass

and the analytical procedure used for measuring. Creatinine is determined in the blood and urine using a colorimetric process discovered in 1886 by Max Jaffé (1841-1911) based on Jaffé reaction in medicinal chemistry.
A blood check for sodium(Na) tests the volume of sodium found in the body. Sodium is a kind of an electrolyte. Electrolytes are minerals charged electrically that help to maintain fluid levels and the chemical balance in your body , called acids and bases. Sodium will also help the nerves and muscles work properly.
A blood check for potassium(K) tests the concentration of potassium that is in the body. Potassium is a type of an electrolyte as well. Electrolytes are substances that are electrically activated in the body , helping to regulate muscle and nerve development, sustain fluid rates and conduct certain essential functions. Your body requires potassium to help work correctly in your heart and muscles.
Hemoglobin determinations are usually performed by an automated cell counter filled to a predetermined level from a tube of well-mixed EDTA-anticoagulated blood. Both types of hemoglobins are transformed to the colored protein cyanomethemoglobin in this experiment, and determined by a colorimeter.
A WBC count will identify latent pathogens in the body and warn doctors about undiagnosed medical problems, such as autoimmune diseases, immune defects and blood disorders. Wbcc and Rbcc can be measured by the laboratory assistant using microscopic counting process.
High blood pressure (HBP or hypertension) occurs when the blood pressure remains too high due to the weight of the blood pressing through the walls of your blood vessels. If blood pressure reliably reaches >130 mmHg systolic and >80 mmHg diastolic, hypertension is identified. Blood pressure is measured using a blood pressure monitor, a non-invasive tool capable of measuring strain inside the arteries, transmitting numerical values using a sphygmomanometer or an electronic system.
The aim of Deep Learning is imitating human intelligence with the help of Artificial Neural Networks(ANN). ANN contains three distinct types of layers including the Input Layer, the Hidden Layer(s), and the Output Layer. Deep neural networks (DNN) are special types of ANNs that have multiple hidden layers. The DNN identifies the correct mathematical relation between the input and output.
There are multiple connections between units within and between layers. These connections have strengths or "weights" that are "learned" by the network. Information in the network is stored in these interconnection weights.
DNNs can be applied to many different problems like computer vision,speech recognition, natural language processing, audio recognition, social network filtering and etc.
Lots of kinds of medical tasks , including analysis of EKG patterns, decision-making in pathology, texture analysis in ultrasound, lesion detection in SPECT images, differential diagnosis from chest radiograph, prediction of pulmonary embolism from ventilation/perfusion scan, breast cancer analysis and decision-making in mammography has been performed using DNNs.

2 MATERIALS AND METHODS

2.1 Deep Neural Networks Architecture

The DNN for failure prediction was developed as 4-layer architecture with 3 hidden layers and one output layer. Input feature values are the blood test results assigned by the lab assistants. The hidden layers perform non-linear transformations of the inputs to get the probability of chronic failure in the output layer.
The network is subjected to three processes, namely training, validation, and test, to develop a network model. In the case of training, the network is educated for an output prediction depending on the input data. The validation process is used for tuning the hyperparameters of a model. The test process is used for performance evaluation.
During the training process, the data set is passing through the whole network, and it is stopped when the error has reached the desired value.
Back Propagation (BP) process is the main part of the neural network training. BP is the way to tune the parameters (weights and biases) of neural network based on the error calculated at the end of the network. Proper tuning of the parameters causes error reduction and making more reliable and generalized model. As mentioned above, the training process stops when the the error of the network converges. The error is calculating by the formula:

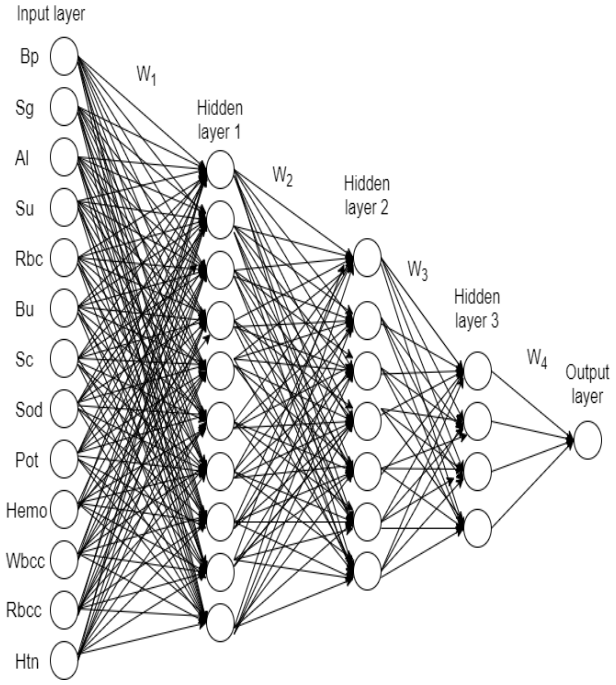
$$H_p(q) = -\frac{1}{N}\sum_{i=1}^{N} y_i log(p(y_i)) + (1-y_i) log(1-p(y_i))$$

The formula given above is called "Binary Cross Entropy Formula" and it is applied to the Binary Classification problem. y_i is the real label of i^{th} example which is verified by the doctors. $p(y_i)$ is the probability or prediction of the i^{th} example. It can also be defined as the extended form of the Shannon Entropy Formula. For this network, it is converted as:

$$C(\omega, b) = -\frac{1}{N}\sum_{i=1}^{N} y_i log(o_i) + (1-y_i) log(1-o_i)$$

where ω (weight) and b (bias) are the parameters of the network. y_i is the real label of i^{th} example which is verified by the doctors. o_i is the output of i^{th} example calculated by the network.
As mentioned before, the DNN architecture developed for the chronic disease prediction has 1 input layer, 3 hidden layers and 1 output layer. Input layer contains the the features lead to detect the kidney anomality. Since the number of features is 13, the number of neurons in input layer is 13. Number of neurons in first hidden layer is 10, for the second hidden layer is 7 and for the last hidden layer is 4. The DNN architecture performs binary classification, therefore, output layer has just one neuron.

Generated DNN structure for the prediction is shown below:



2.2 Training and Testing the Neural Network

After the DNN was built, the input data was normalized (standardized) in order to improve the training characteristics of the data set. Applied Formula:

$x_{stand} = \frac{x-\mu}{\sigma}$, where $\mu = \frac{1}{N}\sum_{i=1}^{N} x_i$ and $\sigma = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(x_i-\mu)^2}$

The network trained using 324 examples, validation has been done on 36 cases and the evaluation process of the DNN was done on 40 samples. The whole training dataset can be combined inside of X matrix and all baled can be combined inside of the Y matrix. Training dataset can be divided into some mini batches. After the multiplication of X matrix with the W_1 matrix (+ bias vector), $X \times W_1 + b$ has been activated with relu function in the first hidden layer. Characteristic of relu (Rectified Linear Unit) function: $relu(z) = max(0,z)$.

The main advantage of using the ReLU function over other activation functions is that it does not activate all the neurons at the same time. This means that the neurons will only be deactivated if the output of the linear transformation is less than 0. Therefore, the output of the first hidden layer will be: $A_1 = relu(X \times W_1 + b_1)$. The same process will be done in the second hidden layer and the answer will be $A_2 = relu(A_1 \times W_2 + b_2)$, then $A_3 = relu(A_2 \times W_2 + b_2)$. In the turn of the output layer activation, the sigmoid function will be used instead of relu, since the answer should be calculated with probabilistic approach. : $sigmoid(x) = \frac{1}{1+e^{-x}}$. The answer is between is 0 and 1. The process is starting from the input layer and going till the output layer is called forward propagation process. During the forward propagation process Dropout regularization is also applied to eliminate the high variance and overfitting.

Since it's the final computed values, it should be compared with the actual labels and the error ought to be calculated with binary cross entropy formula. Using the optimization algorithms, gradients will be calculated and the weights will be updated starting from the last layer till the first layer. This process is back propagation process. This forward and back propagation process will be repeated till the convergence of the error. The number of the repetition of the whole dataset through the all layers is number of epochs. Update formulas:

In each iteration t:
$dW = \frac{\partial C(\omega,b)}{\partial W}$ and $db = \frac{\partial C(\omega,b)}{\partial b}$
$V_{dW} = \beta_1 V_{dW} + (1-\beta_1)dW$ and $V_{db} = \beta_1 V_{db} + (1-\beta_1)db$
V_{dW}, V_{db} are momentum elements
$S_{dW} = \beta_2 S_{dW} + (1-\beta_2)dW^2$ and $S_{db} = \beta_2 S_{db} + (1-\beta_1)db^2$
S_{dW}, S_{db} are RMSprop elements
To minimize the bias:
$V_{dW}^{corr} = \frac{V_{dW}}{1-\beta_1^t}$ and $V_{db}^{corr} = \frac{V_{db}}{1-\beta_1^t}$
$S_{dW}^{corr} = \frac{S_{dW}}{1-\beta_2^t}$ and $S_{db}^{corr} = \frac{S_{db}}{1-\beta_2^t}$
Update:
$W = W - \eta * \frac{V_{dW}^{corr}}{\sqrt{S_{dW}+\varepsilon}}$ and $b = b - \eta * \frac{V_{db}^{corr}}{\sqrt{S_{db}+\varepsilon}}$

After obtaining the error convergence, since the answers of the last layers are probabilities (between 0 and 1), there is a need for convergence. If the answer is greater than threshold (for this case 0.5) convert it to 1, else 0.

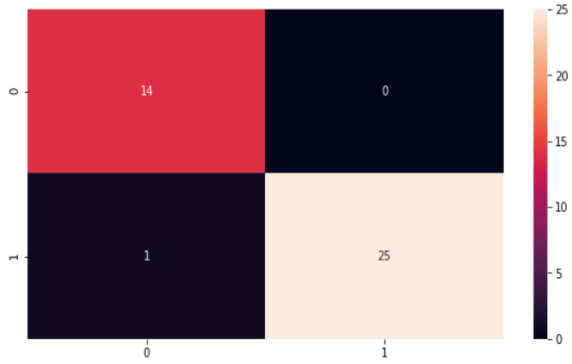
It is the end of the training process. During the training, the weights and biases were updated. During the validation process, the correct values of the hyper parameters were tuned. Correct values of hyperparameters

Hyperparameter	**Value**
Number of hidden layers	3
Number of neurons in hidden layer 1	10
Number of neurons in hidden layer 2	7
Number of neurons in hidden layer 2	4
Number of epochs	100
Mini batch size	64
Activation functions	relu and sigmoid
β_1	0.9
β_2	0.99
ε	10^{-8}
learning rate (η)	with decay
Dropout rate	0.25

The purpose of the testing procedure is evaluation of the model with the test data set. The accuracy of the model is 97.5%

3 RESULTS

Confusion matrix analysis
Sensitivity is the ability expressed with the rate at which the system can determine real patients among those that were definitely diagnosed to be sick.
The confusion matrix of the test is:



The accuracy of the model is $\frac{TP+TN}{TP+TN+FP+FN} = 97.5\%$

Precision of the healthy cases is 93%, for sick cases is 100%
Recall of the healthy cases is 100%, for sick cases is 96%

Precision of healthy cases is what proportion of healthy identifications was actually correct.
Precision of sick cases is what proportion of sick identifications was actually correct.

Recall (sensitivity) of the healthy cases: What proportion of actual healthy cases was identified correctly
Recall (sensitivity) of the sick cases: What proportion of actual sick cases was identified correctly

Referring to the confusion matrix analysis, if the network decides the kidney failure it's 100% true, there is no need for further analysis. However, if the system decides that the patient is healthy, further test may be required since the precision of the healthy examples is 93%. Means that, 7 % of the healthy predictions was wrong.

REFERENCES

[1] Ismail Saritas
Prediction of Breast Cancer Using ANN.
Springer Science+Business Media, LLC 2011

[2] National Kidney Foundation
Kidney Disease Can be Treated.
30 East 33rd Street New York
http://www.kidney.org

[3] National Kidney Foundation
Chronic Kidney Disease (CKD) Symptoms and causes.
http://www.kidney.org

[4] Health Line
Blood Pressure Readings Explained.
https://www.healthline.com

[5] Raymond J Trudnowski, Rodolfo C Rico
Specific Gravity of Blood and Plasma at 4 and 37 °C.
Clinical Chemistry, Volume 20, Issue 5, 1 May 1974, Pages 615–616,

[6] Health Line
Serum Albumin Test.
https://www.healthline.com

[7] Peltarion
Binary crossentropy.
https://peltarion.com/

[8] Health Line
High Potassium.
https://www.healthline.com

[9] Medine Plus
Sodium Blood Test.
https://medlineplus.gov

[10] Very Well Mind
Serum Blood Levels and Medication.
https://www.verywellmind.com

[11] Very Well Health
How Hypertension Is Diagnosed.
https://www.verywellhealth.com

[12] Joscilin Mathew; Parvathy Sankar; Matthew Varacallo.
Physiology, Blood Plasma.
https://www.ncbi.nlm.nih.gov

[13] Science Direct *Confusion Matrix.*
https://www.sciencedirect.com/