A Hybrid Intelligent Model for Enhancing Healthcare Services on Cloud Environment

Ahmed Abdelaziz\(^1\)  
Department of Information Systems  
Higher Technological Institute,  
Cairo, Egypt

Ahmed S. Salama\(^2\)  
Department of Computer and Information Systems, Sadat Academy for Management Sciences,  
Cairo, Egypt

A.M. Riad\(^3\)  
Department of Information Systems, Faculty of Computers and Information,  
Mansoura University, Egypt

Abstract—Cloud computing plays a major role in addressing the challenges of healthcare services such as diagnosis of diseases, telemedicine, maximize utilization of medical resources, etc. Early detection of chronic kidney disease on cloud environment is a big challenge that is facing healthcare providers. This paper concentrates on the using of intelligent techniques such as Decision Tree, Clustering, Linear Regression, Modular Neural Network, and Back Propagation Neural Network to address this challenge. In this paper, the researchers propose a hybrid intelligent model based on cloud computing for early revealing of chronic kidney disease. Two intelligent techniques were used: linear regression and neural network. Linear regression was used to define crucial factors that have an impact on chronic kidney disease. The proposed model for early revealing of chronic kidney disease was built using Neural Network. The accuracy of proposed model is 97.8%. This model outperforms on the other models existed in the previous works in terms of the accuracy and precision, recall and F1 score.

Keywords—Chronic kidney disease; linear regression; neural network; cloud computing

I. INTRODUCTION

Healthcare plays an important role in saving people's lives. Healthcare provides modern trends such as telemedicine, diagnosis of diseases and etc. Cloud computing plays an important role in supporting health care services. It is composed of three main services as follows:

- Medical Infrastructure.
- Medical Platform.
- Medical Software.

Cloud computing allows many patients to retrieve their medical information anywhere and anytime by using different devices such as smart phone, personal computer and etc. There are many advantages of cloud computing for healthcare services that are including the following:

- Usability
- Speed
- Accessibility
- Disaster Recovery
- Cost Savings

Fig. 1. Statistics of Chronic Kidney Disease.
Early detection of diseases is one of the most important challenges for health care services on cloud computing [7]. This paper focuses on Chronic Kidney Disease because it is one of the most serious diseases facing humanity in modern times. Many statistics indicate the growing number of Chronic Kidney patients worldwide as shown in figure 1.

This paper introduces a hybrid intelligent model for early detection of Chronic Kidney disease on cloud computing environment by using the machine learning tool on windows azure. The rest of the paper is organized as follows: section two introduces a background overview; section three presents the related work; section four introduces the proposed intelligent model; section five introduces the experimental results; and finally, section six presents the conclusion and future work.

II. BACKGROUND AND OVERVIEW

This section presents a review of Chronic Kidney disease, neural network and regression analysis as follows:

A. Overview of Chronic Kidney Disease

Chronic kidney disease is considered one of the most dangerous diseases that are facing the worldwide. There are many factors that are influencing chronic kidney disease as shown in Table 1. A data set of chronic kidney disease from machine learning repository was used. The data set contains 800 instances.

<table>
<thead>
<tr>
<th>No</th>
<th>Factor</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Anemia</td>
<td>Nominal</td>
</tr>
<tr>
<td>2</td>
<td>Age</td>
<td>Numerical</td>
</tr>
<tr>
<td>3</td>
<td>Pedal edema</td>
<td>Nominal</td>
</tr>
<tr>
<td>4</td>
<td>Blood pressure</td>
<td>Numerical</td>
</tr>
<tr>
<td>5</td>
<td>Appetite</td>
<td>Nominal</td>
</tr>
<tr>
<td>6</td>
<td>Specific gravity</td>
<td>Nominal</td>
</tr>
<tr>
<td>7</td>
<td>Coronary artery disease</td>
<td>Nominal</td>
</tr>
<tr>
<td>8</td>
<td>Albumin</td>
<td>Nominal</td>
</tr>
<tr>
<td>9</td>
<td>Diabetes mellitus</td>
<td>Nominal</td>
</tr>
<tr>
<td>10</td>
<td>Sugar</td>
<td>Nominal</td>
</tr>
<tr>
<td>11</td>
<td>Hypertension</td>
<td>Nominal</td>
</tr>
<tr>
<td>12</td>
<td>Red blood cells</td>
<td>Nominal</td>
</tr>
<tr>
<td>13</td>
<td>Red blood cell count</td>
<td>Numerical</td>
</tr>
</tbody>
</table>

B. Overview of Linear Regression

Linear regression has two types of regressions which are simple linear regression and multiple linear regressions. General equation of linear regression is formulated as follows:

\[ M = \beta_0 + \beta_1 y_1 + \beta_2 y_2 + \ldots + \beta_N y_N + \epsilon \]  

(1)

Where:

- \( M \) is the dependent variable
- \( y_1, y_2, y_N \) are the independent variables

- \( \beta_i \) represents the regression coefficient
- \( \epsilon \) represents the random error component
- \( \beta_0 \) represents the \( y \) intercept

In linear regression, there are two main criteria’s to measure the performance of the proposed model which are coefficient of determination and feature weight of independent variables as shown in figure 2.

![Fig. 2. Two Main Outputs in Linear Regression.](image)

C. Overview of Neural Network

Neural network is used to implement many intelligent applications such as diseases diagnosis and data classification and etc. There are some types of neural networks such as:

- Feed-forward network
- Back-propagation network
- Modular neural network
- Recurrent neural network

The neural network consists of input layers, hidden layers and output layers as shown in figure 3.

![Fig. 3. General Structure of Neural Network.](image)
III. RELATED WORK

This section introduces many researches that seek to detect chronic kidney disease as follows:

Lambodar Jena and et al, presented a novel model to detect chronic kidney disease by using data mining classification. This research is used WEKA tool. The accuracy of the proposed model is 95%. Naïve Bayes outperforms on decision tree. The importance of the research is to predict chronic kidney disease by naïve Bayes technique [8].

K. R. Anantha and et al, introduced a model which is decision tree technique for predicting chronic kidney disease on WEKA tool. This research shows that decision tree outperforms on clustering technique. The accuracy of the proposed model is 91%. The importance of the research is to predict chronic kidney disease by decision tree [9].

Basma B. and et al, presented a model for detecting chronic kidney disease by using support vector machine. The results show that support vector machine better than decision tree in terms of execution time and accuracy. The accuracy of the proposed model is 95%. The importance of the research is to detect chronic kidney disease by support vector machine [10].

Asif S. and et al, introduced intelligent model for revealing chronic kidney disease by using clustering technique on WEKA tool. This research shows clustering technique better than decision tree in order to accuracy. The accuracy of the clustering technique is 96%. The importance of the research is to reveal chronic kidney disease by clustering technique [11].

RUEY K. C. and et al, presented an approach for detecting of chronic kidney disease on cloud computing environment. The results show that back propagation neural network outperforms on modular neural network and feed forward neural network in order to accuracy and precision. The accuracy of the proposed system is 94.7%. The importance of the study is to detect chronic kidney disease on cloud environment [12].

Chien. Z. W. and et al, introduced fuzzy expert system for revealing of chronic kidney disease on cloud computing environment. This research shows that fuzzy expert system better than neural network in order to accuracy and mean square error. The accuracy of the proposed model is 88.4%. The importance of the study is to reveal chronic kidney disease by fuzzy expert system on cloud environment [13].

Stuti N. and et al, introduced a new model based on Naïve Bayes to predict chronic kidney disease on cloud environment (Google Application Engine). The results show that Naïve Bayes outperforms on neural network in order to accuracy and precision. The accuracy of the proposed model is 97.1%. The importance of the research is to predict chronic kidney disease on Google Application Engine [14].

Anu B. and et al, introduced survey of many researches that seek to predict chronic kidney disease by using intelligent techniques. The importance of the research is to review many researches of chronic kidney disease [15].

Table 2 Introduces Summary of Intelligent Techniques of Chronic Kidney Disease Researches.

<table>
<thead>
<tr>
<th>No</th>
<th>Factor</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>14</td>
<td>Pus cell</td>
<td>Nominal</td>
</tr>
<tr>
<td>15</td>
<td>White blood cell count</td>
<td>Numerical</td>
</tr>
<tr>
<td>16</td>
<td>Pus cell clumps</td>
<td>Nominal</td>
</tr>
<tr>
<td>17</td>
<td>Packed cell volume</td>
<td>Numerical</td>
</tr>
<tr>
<td>18</td>
<td>Bacteria</td>
<td>Nominal</td>
</tr>
<tr>
<td>19</td>
<td>Hemoglobin</td>
<td>Numerical</td>
</tr>
<tr>
<td>20</td>
<td>Blood glucose random</td>
<td>Numerical</td>
</tr>
<tr>
<td>21</td>
<td>Potassium</td>
<td>Numerical</td>
</tr>
<tr>
<td>22</td>
<td>Blood urea</td>
<td>Numerical</td>
</tr>
<tr>
<td>23</td>
<td>Sodium</td>
<td>Numerical</td>
</tr>
<tr>
<td>24</td>
<td>Serum Creatinine</td>
<td>Numerical</td>
</tr>
<tr>
<td>25</td>
<td>Class</td>
<td>Nominal</td>
</tr>
</tbody>
</table>

TABLE II. INITIAL FACTORS THAT INFLUENCE ON CHRONIC KIDNEY DISEASE CON’T

TABLE III. SUMMARY OF RELATED WORK

<table>
<thead>
<tr>
<th>No</th>
<th>Category</th>
<th>Some Approach</th>
</tr>
</thead>
</table>
| 1  | intelligent techniques applied of chronic kidney disease | ➢ quantitative approach  
➢ Naïve Bayes  
➢ mean analysis  
➢ mean time and modeling regression  
➢ time series  
➢ Clustering  
➢ Feed Forward neural network  
➢ Back Propagation neural network  
➢ Modular neural network  
➢ Decision tree  
➢ Support vector machine  
➢ Fuzzy expert system |

Through related work, intelligent applications are not accurate on cloud computing environment for revealing of chronic kidney disease. So, this paper introduces a new intelligent model for revealing chronic kidney disease on cloud environment.

IV. THE PROPOSED APPROACH

This section introduces an intelligent model to reveal Chronic Kidney disease on cloud environment. As shown in figure 4, this intelligent model has three steps:

1) Survey recent studies related to Chronic Kidney disease to develop a candidate list of factors that influence this disease.

2) Identify the critical factors form this candidate list using linear regression.

3) Develop an intelligent model for revealing chronic kidney disease based neural network.
V. THE HYBRID INTELLIGENT MODEL FOR HEALTHCARE SERVICES

This section presents the proposed hybrid model for healthcare services on cloud computing. It includes two intelligent techniques which are explained below.

A. Linear Regression

Linear regression is used to identify the critical factors of Chronic Kidney disease (CFCKD). It is composed of one dependent variable and thirteen independent variables. It is formulated as follows:

\[ M = \beta_0 + \beta_1 CFCKD_1 + \beta_2 CFCKD_2 + \ldots + \beta_{13} CFCKD_{13} + \epsilon \]  

Where:

- \( M \) is the dependent variable (degree of influence of candidate factors on Chronic Kidney disease)
- \( CFCKD_1, CFCKD_2, \ldots, CFCKD_{13} \) are the independent variables (candidate factors).

This section also introduces linear regression algorithm to identify critical factors of Chronic Kidney disease.

Algorithm 1. The Algorithm of the Linear Regression to Identify Crucial Factors that influence on Chronic Kidney Disease.

1. Input: \( \epsilon \) (degree of influence proposed factors on chronic kidney disease)
2. Output: \( \lambda \) (independent variables factors of chronic kidney disease)
3. \( \lambda \) = final list of crucial factors of chronic kidney disease
4. MLR = Multiple Linear Regression
5. FW = Feature Weight
6. SSR = Sum of Squares Regression
7. SST = Sum of Squares Total
8. \( \hat{y}_i \) = the prediction
9. \( y_i \) = the true value.
10. \( SSR = \sum (\hat{y}_i - y_i)^2 \)
11. \( SST = \sum (y_i - \bar{y})^2 \)
12. Build the MLR model based on the set of \( \epsilon \) and \( \lambda \)
13. Estimate the MLR model
14. Check coefficient of determination
15. coefficient of determination = \( \frac{SSR}{SST} \)
16. If (coefficient of determination < 0.5)
17. Change \( \epsilon \)
18. Go to step 13
19. Else
20. Accept the MLR model
21. End If
22. Check FW for each variable to define \( \lambda \)
23. If (FW < 0.05)
24. Approve \( \lambda \)
25. Else
26. Reject the other \( \epsilon \)
27. End If
28. Return \( \lambda \).

B. Neural Network

Neural network has 13 critical factors as input layer, one hidden layer and two outputs (Chronic Kidney or No Chronic Kidney).

This section introduces neural network algorithm to reveal of Chronic Kidney disease.

Algorithm 2. The Proposed Algorithm of the Neural Network to Reveal of Chronic Kidney Disease.

1. Input: \( \epsilon \) (critical factors of Renal failure), \( \lambda \) (hidden layers)
2. Output: \( \hat{y} \) (Detection of Renal failure)
3. NNM = Neural Network model
4. \( \lambda \) = Accuracy of the NNM
5. Create NNM with \( \epsilon \) and \( \lambda \)
6. Verify NNM
7. End If
8. If (\( \lambda > 0.95 \))
10. Accept NNM
11. Go to step 17
12. Else
13. Update the number of \( \varphi \)
14. Go to step 6
15. End If
16. Identify \( \lambda \) via validate NNM
17. Return \( \Psi \)

VI. EXPERIMENTAL RESULTS

This section presents the execution of the proposed hybrid model on windows azure as shown in figure 5.

A. Execution of Linear Regression on Windows Azure

This section introduces the execution of linear regression on windows azure. Linear regression was used to define critical factors of Chronic Kidney disease. Table 5 shows the critical factors of Chronic Kidney disease via the feature weights. Whenever, feature weight is less than 0.05, the factor is medically important. For example, the feature weight for Serum Creatinine (feature weight = 0.04) is less than 0.05, then this factor is medically important as shown Chronic Kidney in figure 6.

![Figure 5. Execution of Hybrid Model on Windows Azure.](image-url)
condition statement of factors of chronic kidney disease.

**TABLE V. SUMMARY OF CRITICAL FACTORS OF CHRONIC KIDNEY DISEASE**

<table>
<thead>
<tr>
<th>No</th>
<th>Factor Name</th>
<th>FW</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Anemia</td>
<td>-0.002</td>
</tr>
<tr>
<td>2</td>
<td>White Blood Cell Count</td>
<td>0.0007</td>
</tr>
<tr>
<td>3</td>
<td>Packed Cell Volume</td>
<td>-0.002</td>
</tr>
<tr>
<td>4</td>
<td>Hemoglobin</td>
<td>0.03</td>
</tr>
<tr>
<td>5</td>
<td>Potassium</td>
<td>-0.0006</td>
</tr>
<tr>
<td>6</td>
<td>Sodium</td>
<td>0.002</td>
</tr>
<tr>
<td>7</td>
<td>Serum Creatinine</td>
<td>-0.05</td>
</tr>
<tr>
<td>8</td>
<td>Blood Urea</td>
<td>0.003</td>
</tr>
<tr>
<td>9</td>
<td>Blood Glucose Random</td>
<td>0.005</td>
</tr>
<tr>
<td>10</td>
<td>Sugar</td>
<td>0.04</td>
</tr>
<tr>
<td>11</td>
<td>Specific Gravity</td>
<td>-0.002</td>
</tr>
<tr>
<td>12</td>
<td>Blood Pressure</td>
<td>-0.00005</td>
</tr>
<tr>
<td>13</td>
<td>Age</td>
<td>0.02</td>
</tr>
</tbody>
</table>

**B. Execution of Neural Network on Windows Azure**

This section introduces the execution of neural network on Windows Azure. NNM is composed of 13 neurons in input layer, one hidden layer and two outputs. The neural network type used in this paper is the two class neural network. Table 6 shows the performance of the neural network. It defines accuracy, precision, recall and f1 score. The accuracy of the proposed model is 97.8% as shown in figure 7.

**TABLE VI. PERFORMANCE ATTRIBUTES OF NNM**

<table>
<thead>
<tr>
<th>True Positive</th>
<th>False Negative</th>
<th>Accuracy</th>
<th>Precision</th>
</tr>
</thead>
<tbody>
<tr>
<td>51</td>
<td>0</td>
<td>0.978</td>
<td>0.962</td>
</tr>
<tr>
<td>False Positive</td>
<td>True Negative</td>
<td>Recall</td>
<td>F1 Score</td>
</tr>
<tr>
<td>2</td>
<td>39</td>
<td>1.000</td>
<td>0.981</td>
</tr>
</tbody>
</table>

**C. Case Study**

Three real cases of patients were conducted on Windows Azure to reveal of Chronic Kidney disease (CKD) as shown in table 7 and figure 8.

**TABLE VII. DATA OF CKD PATIENTS**

<table>
<thead>
<tr>
<th>No</th>
<th>Factor Name</th>
<th>Patient 1</th>
<th>Patient 2</th>
<th>Patient 3</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Age</td>
<td>49</td>
<td>59</td>
<td>29</td>
</tr>
<tr>
<td>2</td>
<td>Blood Pressure</td>
<td>71</td>
<td>81</td>
<td>101</td>
</tr>
<tr>
<td>3</td>
<td>Specific Gravity</td>
<td>2.007</td>
<td>1.026</td>
<td>1.2</td>
</tr>
<tr>
<td>4</td>
<td>Sugar</td>
<td>0</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>5</td>
<td>Blood Glucose Random</td>
<td>127</td>
<td>132</td>
<td>121</td>
</tr>
<tr>
<td>6</td>
<td>Blood Urea</td>
<td>57</td>
<td>19</td>
<td>21</td>
</tr>
<tr>
<td>7</td>
<td>Serum Creatinine</td>
<td>3.9</td>
<td>1.2</td>
<td>1.3</td>
</tr>
<tr>
<td>8</td>
<td>Sodium</td>
<td>113</td>
<td>142</td>
<td>131</td>
</tr>
<tr>
<td>9</td>
<td>Potassium</td>
<td>2.6</td>
<td>3.6</td>
<td>2.2</td>
</tr>
<tr>
<td>10</td>
<td>Hemoglobin</td>
<td>12.2</td>
<td>15.9</td>
<td>11</td>
</tr>
<tr>
<td>11</td>
<td>Packed Cell Volume</td>
<td>33</td>
<td>54</td>
<td>51</td>
</tr>
<tr>
<td>12</td>
<td>White Blood Cell Count</td>
<td>5700</td>
<td>6900</td>
<td>5100</td>
</tr>
<tr>
<td>13</td>
<td>Anemia</td>
<td>Yes</td>
<td>No</td>
<td>Yes</td>
</tr>
<tr>
<td>14</td>
<td>Class</td>
<td>CKD</td>
<td>NOCKD</td>
<td>NOCKD</td>
</tr>
<tr>
<td>15</td>
<td>Probability</td>
<td>0.00005</td>
<td>0.99993</td>
<td>0.99997</td>
</tr>
</tbody>
</table>

![Performance Attributes of Hybrid Model on Windows Azure.](image)
hybrid model includes two techniques: linear regression and neural network. The proposed model was implemented on windows azure. The accuracy of the proposed model is 97.8%. Three real cases of patients were conducted on windows azure to empirically validate the proposed model.

It is recommended as future work:

- To use logistic regression for determining critical factors that effect Chronic Kidney disease.
- To uses deep learning for revealing Chronic Kidney disease on cloud computing environment.

REFERENCES