Intelligent Decision Support System for Depression Diagnosis Based on Neuro-fuzzy-CBR Hybrid

Victor E. Ekong¹, Udoinyang G. Inyang¹ & Emmanuel A. Onibere²

¹ Department of Computer Science, Faculty of Science, University of Uyo, Uyo, Nigeria
² Department of Computer Science, Faculty of Physical Sciences, University of Benin, Benin City, Nigeria

Correspondence: Victor E. Ekong, Department of Computer Science, Faculty of Science, University of Uyo, Akwa Ibom State, Nigeria. Tel: 234-80-5604-3359. E-mail: victor_eshiet_ekong@yahoo.co.uk

Received: May 2, 2012     Accepted: June 19, 2012     Online Published: June 25, 2012
doi:10.5539/mas.v6n7p79          URL: http://dx.doi.org/10.5539/mas.v6n7p79

Abstract
Depression disorder is common in primary care, but its diagnosis is complex and controversial due to the conflicting, overlapping and confusing nature of the multitude of symptoms, hence the need to retain cases in a case base and reuse effective previous solutions for current cases. This paper proposes a neuro-fuzzy-Case Base Reasoning (CBR) driven decision support system that utilizes solutions to previous cases in assisting physicians in the diagnosis of depression disorder. The system represents depression disorder with 25 symptoms grouped into five categories. Fuzzy logic provided a means for handling imprecise symptoms. Local similarity between the input cases and retrieved cases was achieved using the absolute deviation as the distance metric, while adaptive neuro-fuzzy inference system handled fuzzy rules whose antecedents are the mapped local similarities of each category of symptoms for global similarity measurement, upon which the retrieved cases are ranked. The 5 best matched cases are subjected to the emotional filter of the system for diagnostic decision making. This approach derives strengths from the hybridization since the tools are complementary to one another.

Keywords: depression disorder, neuro-fuzzy, decision support, similarity measure, depression diagnosis

1. Introduction
Depression is a common mental disorder in primary health care, but its diagnosis and effective treatment is controversial because of variations in clinical practice, lack of evidence on the best treatments and access to them, and because of its multifaceted nature and contested meaning of symptoms (Dowrick, 2004). Depressive symptoms range along a continuum from every day sadness, loss of interest to suicidal depression.

Although many other symptoms occur in varying combinations, the illness is a causal factor in many chronic conditions such as diabetes, cardiovascular diseases, HIV/AIDS resulting in higher costs to the healthcare system (Maja, Meifania, & Tharam, 2008; World Health Organization (WHO), 2009; Kessler, 2002). The recognition and treatment of depression is a challenging area of clinical practice, especially in primary care where there are many patients with various presentations and a multitude of causes for distress (Lester & Howe, 2008). In depression diagnosis, physicians utilize a number of cognitive behavioural therapy (CBT) assessment tools such as Becks depression inventory (BDI) and Hamilton’s Rating Scale for Depression (HRSD) or the Montgomery-Asberg Depression Rating Scale (MADRS) to establish severity levels of the disease in order to determine therapy (Ariyanti, Kusumadewi, & Paputungan, 2010; Nameroff, 2006). Classification of the disease is therefore based on the patients’ subjective description of symptoms and the physician’s judgment as to whether they meet the criteria of the diagnostic and statistical manual of mental disorders, version 4 (DSM-IV) (American Psychiatric Association (APA), 1994). Diagnosis of the disorder by primary care physicians is difficult due to the complexity and confusing nature of the disease (Mondimore, 2006; Mila, Kielan, & Michalak, 2009).

Medical diagnosis has undergone different phases of research from mathematical and statistical approaches which are mostly engaged to enhance the quality of medical data estimations, to Artificial Intelligent (AI) approaches. The inadequacy of statistical estimation techniques is that quality cannot be guaranteed when dealing with incomplete, noisy and non-linear data (Antoni, Jorge, & Paulo, 2008). AI approaches provides reasoning capability, which consists of inference from facts and rules using heuristics, pattern matching or other
search approaches. Recent developments in medical diagnostics has embraced the AI approaches such as genetic algorithms (GA), neural networks (NN), Fuzzy logic (FL), rule based systems (RBS) and Case Based Reasoning (CBR) to develop tools for diagnosis and for predicting treatment responses (Wan, Wan, & Fadzilah, 2006).

In this work, AI techniques of neural networks, fuzzy logic and CBR are combined to model a DSS for the diagnosis of depression disorders. NNs are constructed to imitate the intelligent human biological processes of learning, self-modification and adaptation. Although NN is good at handling non-linear, noisy or incomplete data it has very weak explanation mechanism which is highly desirable in medical decision support systems (MDSS) (Uzoka, Obot, & Baker, 2009). Fuzzy logic provides a means for dealing with imprecision, vagueness and uncertainties in medical data (Zadeh, 1965). Neuro-fuzzy inference systems provide self-learning intelligent systems that are capable of handling uncertainties in a diagnosis process (Jang, Sun, & Mitzutani, 1997). CBR entails the use of a set of concrete past situations, called cases, stored in a knowledge base referred to as a case base to solve a new problem (Aamodt, 1994). Combinations of CBR with other intelligent methods have been explored for more effective knowledge representation and problem solving. Hybrid CBR systems are reported in (Obot, Akin yokun, & Udoh, 2008; Prentzas, & Harzilygeroudis, 2009; Lopez-Fernandez, FdezRiverola, Rboiro-Jato, Glez-Pena, & Mendez, 2011; Begum & Mema, 2011). Among the hybrid CBR systems, a neuro-fuzzy CBR is not widely reported; the closest are combinations of FL with CBR or NN with CBR in the implementation of medical diagnosis systems. Hence, the contribution to knowledge of this work stems not only from the hybridization of NN, FL and CBR but also the fusion of neuro-fuzzy inference system and fuzzy similarity matching in the implementation of a MDSS.

2. Related Works

In Marks et al., (2004), a computer-aided CBT (CCBT) is described for delivering selfhelp for sufferers with phobic, panic, anxiety, obsessive-compulsive and depressive disorders. Although the approach greatly reduced the demand on therapist’s time, it lacked the patient - doctor feedback mechanism which is crucial for depression therapy. In Razzouk, Mari, Wainer, & Sigulem (2006), a pilot study for modeling the diagnoses of schizophrenia and other psychotic disorders is reported. In Mila et al. (2009), a fuzzy semiotic framework for modeling imprecision in the assessment of depression is presented. In Begum, Ahmed, Funk, Xiong & Scheele (2009), a case based DSS is developed for individual stress diagnosis using finger temperature profiles. The approach enabled the reuse of experience from previous cases with analyzed temperature and stress profiles. In Suhasini, Palanivel, & Ramalingam (2010), an ANN approach is described for diagnosing depression using radial basis function (RBF) and back propagation neural networks. The approach showed great promise in accurately identifying the psychiatric problem among patients, but it lacked an explanation mechanism. In Ariyanti et al. (2010), a BDI version 2 test assessments is reported that utilizes a fuzzy inference system to represent all the factors of BDI-II’s 21 questions.

3. Method

In our methodology, we propose an intelligent system that incorporates ANN, FL and CBR in the diagnosis of depression. The procedure of hybrid platforms design described in (Shanthi, Sahoo, & Saravanan, 2009; Ajith, 2005; Medsker, 1999; Fuller, 1995) are studied and modified to suit the design. The block diagram for the proposed neural network, fuzzy logic and CBR hybrid is presented in Figure 1.

![Figure 1. Block diagram of depression diagnosis IDSS hybrid platform](image-url)
3.1 Fuzzification Subsystem

As shown in Figure 1, the neural network receives fuzzified symptoms of depression as inputs. The fuzzy set of depression symptoms (Sij) and depression diagnosis (Ri) are {low, medium, high, very high} and {Near Absent, Mild, Moderate, Severe} respectively. Sij is the jth symptom of the ith patient, Ri represents the result of depression diagnosis of the ith patient. The 21 BDI-II set of parameters for depression disorder provided in Beck's (1996) and Physiological parameters described in Hildrum, Mykletun, Holman and Dahl (2008) are adopted for this work. The 21 BDI-II symptoms are grouped based on Wan, Hu, Moore and Ashford (2008) categorization of Beck's cognitive model of depression. A summary of depression disorder attributes and their categories is presented in Table 1.

### Table 1. Attributes of depression disorder

<table>
<thead>
<tr>
<th>SN</th>
<th>Category</th>
<th>Patients' attributes</th>
<th>Code</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Sadness</td>
<td>SAD</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>Crying</td>
<td>CRY</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>Irritability</td>
<td>IRR</td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>Loss of interest</td>
<td>LOI</td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>Agitation</td>
<td>AGI</td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>Self criticalness</td>
<td>SCR</td>
<td></td>
</tr>
<tr>
<td>7</td>
<td>Indecision</td>
<td>IND</td>
<td></td>
</tr>
<tr>
<td>8</td>
<td>Self dislike</td>
<td>SDL</td>
<td></td>
</tr>
<tr>
<td>9</td>
<td>Worthlessness</td>
<td>WOT</td>
<td></td>
</tr>
<tr>
<td>10</td>
<td>Cognitive factors (B)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>11</td>
<td>Pessimism</td>
<td>PES</td>
<td></td>
</tr>
<tr>
<td>12</td>
<td>Guilty feelings</td>
<td>GUI</td>
<td></td>
</tr>
<tr>
<td>13</td>
<td>Concentration difficulty</td>
<td>CON</td>
<td></td>
</tr>
<tr>
<td>14</td>
<td>Past failure</td>
<td>FAL</td>
<td></td>
</tr>
<tr>
<td>15</td>
<td>Tiredness</td>
<td>TRD</td>
<td></td>
</tr>
<tr>
<td>16</td>
<td>Changing sleep patterns</td>
<td>CSL</td>
<td></td>
</tr>
<tr>
<td>17</td>
<td>Change in appetite</td>
<td>CAP</td>
<td></td>
</tr>
<tr>
<td>18</td>
<td>Loss of interest in sex</td>
<td>LSX</td>
<td></td>
</tr>
<tr>
<td>19</td>
<td>Loss of pleasure</td>
<td>LOP</td>
<td></td>
</tr>
<tr>
<td>20</td>
<td>Motivational factors (D)</td>
<td>Loss of energy</td>
<td>LOE</td>
</tr>
<tr>
<td>21</td>
<td>Suicidal thoughts</td>
<td>SUI</td>
<td></td>
</tr>
<tr>
<td>22</td>
<td>Age</td>
<td>AGE</td>
<td></td>
</tr>
<tr>
<td>23</td>
<td>Diastolic blood pressure</td>
<td>DBP</td>
<td></td>
</tr>
<tr>
<td>24</td>
<td>Systolic blood pressure</td>
<td>SBP</td>
<td></td>
</tr>
<tr>
<td>25</td>
<td>Body mass index</td>
<td>BMI</td>
<td></td>
</tr>
</tbody>
</table>

A fuzzy set is represented by expressing it as a function and then map the elements of the set to their degree of membership. Some of the membership functions that exist are Gaussian, Triangular, Trapezoidal, S-function and L-function (Vijaya, Nehemiah, Kannan, & Bhuveswari, 2010). Triangular and Trapezoidal membership functions have been used extensively due to their computational efficiency (Jang et al., 1997). The work of Dubois & Prade (1996) guided the choice of the triangular membership function for this work and gives the general form of the triangular membership function in Equation (1)

\[
\mu(x) = \begin{cases} 
1, & \text{if } x = b \\
\frac{x-a}{b-a}, & \text{if } a \leq x < b \\
\frac{c-x}{c-b}, & \text{if } b \leq x < c \\
0, & \text{if } c \leq x 
\end{cases}
\]
where a and c are the parameters governing triangular membership functions; b represents the value for which \( \mu(x) = 1 \) and is defined as \( b = \frac{a + c}{2} \). The actual membership functions of each element in the fuzzy set are derived as described in (Jang et al., 1997; Akinyokun, Obot, & Uzoka, 2009). Examples include Low(A) = f(A,[0.10,0.25,0.40]) and Severe(R)=f(R,[0.47, 0.74,1.0]).

3.2 Training Subsystem

The fuzzified symptoms serve as inputs to the neural network models. Each Neural network is a two layered feed forward architecture which interacts with each other as shown in Figure 2.

![Figure 2. Block diagram of NN for depression DSS](image)

The algorithm used to train the networks is the back-propagation algorithm with sigmoid function for hidden and output layer neurons’ transformation. The neuro-fuzzy architecture is presented in Figure 3. In Figure 3, the NN trained by the subsystem consists of 25 nodes in the symptoms layer; each node represents a linguistic value in the fuzzy set of depression symptoms. The hidden layer has 5 nodes, each node correspond to a category of depression attributes considered in this work. The output layer consists of 4 nodes, each representing the level of severity of depression disorder.

![Figure 3. Neuro-Fuzzy Architecture of IDSS depression disorder diagnosis](image)
Let p, q and t represent the number of neurons in the input, hidden and output layers respectively. \( W_{sh} \) represents the weight vector between symptoms (input) layer nodes to hidden layer node. \( W_{hr} \) represents the weight vector between hidden layer nodes to output layer node. \( A^r \) represents the expected output for a given patient data. For each patient’s data \( S_{ij}(s_1, s_2, \ldots s_p) \), in the training set, respective weights are computed using Equations (2) to (6). The output of the input layer neurons is derived from Equation (2), the input to the hidden layer neurons are derived from Equation (3) while outputs of the hidden layer neurons are computed from Equation (4). Input to the output layer neurons and output of the output layer neurons are represented in Equations (5) and (6) respectively.

\[
O_j^i = x_j, \quad j=1, 2, \ldots, p
\]

\[
S_k^h = \sum_{j=1}^{p} O_j^i W_{jk}^{sh}, \quad k = 1, 2, \ldots, q
\]

\[
O_k^h = \frac{1}{1 + e^{-S_k^h}}, \quad k = 1, 2, \ldots, q
\]

\[
S_m^o = \sum_{k=1}^{q} O_k^h W_{km}^{ho}, \quad m = 1, 2, \ldots, t
\]

\[
O_l^o = \frac{1}{1 + e^{-S_m^o}}, \quad m = 1, 2, \ldots, t
\]

### 3.3 Inference Subsystem

In the inference subsystem, the mapping from an input (new case) to the retrieved case (target) is measured using fuzzy similarity matching and adaptive neuro-fuzzy inference system (ANFIS). The results from the fuzzy similarity matching are the parameters driving the ANFIS which provides results from which decisions are made. We denote cases in the case base as \( Z \times W \) matrices, where \( W \) represents the categories of depression symptoms and \( z \) the number of patients cases in the case base. Depression diagnostic cases are in a column vector. Let \( A \) and \( B \) be patients’ symptom and diagnostic cases respectively.

\[
A = \begin{bmatrix}
\alpha_{11} & \alpha_{12} & \alpha_{13} & \ldots & \alpha_{1w} \\
\alpha_{21} & \alpha_{22} & \alpha_{23} & \ldots & \alpha_{2w} \\
\vdots & \vdots & \vdots & \ddots & \vdots \\
\alpha_{u1} & \alpha_{u2} & \alpha_{u3} & \ldots & \alpha_{uw}
\end{bmatrix}
\]

In this work, we set \( w=5 \); the \( ith \) row vector, \( A_{i,u} = [a_{i1}, a_{i2}, \ldots a_{i5}] \) represents the \( ith \) patient’s case, where \( u=1, 2, 3, 4, 5 \); \( a_{iu} \) is the result of the actual mapping of the linguistic value to the crisp value within the degree of membership. We therefore have the row vector as in Equation (8).

\[
A = \begin{bmatrix}
a_{jw}
\end{bmatrix}_{j=1}^{y}
\]

Let \( E \) and \( T \) represent the transpose of the row vectors of new cases (input) and retrieved cases (target) respectively. Let \( L \) be the distance metric matrix; \( l_{ij} = \lvert e_{ij} - t_{ij} \rvert \) be the distance between the \( ith \) input element and \( jth \) target element. There are \( y \times y \) possible distance metrics; where \( y \) is the total number of symptoms in a category, in this work, we adopt the absolute deviation in Equation (9) as the distance metric. The general form of Matrix \( L \) is as shown in Equation (10).

\[
l(e_{ij}, t_{ij}) = \lvert e_{ij} - t_{ij} \rvert
\]

We denote the various distance metrics matrices as \( L_A, L_B, L_C, L_D, \) and \( L_E \). Each column of \( L \) represents the absolute deviation of a symptom from other of symptoms in a particular category. Let \( a_{ij} \) be the elements of \( L \), \( \bar{\delta}_{i,j} \) \( (i=1,2,\ldots,y) \) denotes the average absolute deviation of symptom \( j \) from \( i \) other symptoms in a particular category; it is given in Equation (11)
The normalized absolute deviation, $\delta_{ij}$, of the $k$th category is the local similarity measure;

$$\delta_{ij} = \frac{\sum_{i=1}^{y} \alpha_j}{y}$$

The fuzzy set \{very close, close, slightly close\} is used to represent the local similarities between the input case and target case and its membership function is defined as follows;

$$S_{L(k)} = \begin{cases} 
  "Very Close" & \text{if } \left\| \delta_{i} \right\| < 0.35 \\
  "Close" & \text{if } 0.35 \leq \left\| \delta_{i} \right\| \leq 0.65 \\
  "Slightly Close" & \text{if } \left\| \delta_{i} \right\| > 0.65 
\end{cases}$$

The global similarity measure ($S_g$) is derived from the fuzzy rule of the local similarity measure. Example of such rules is:

If ($S_{L(1)}$ is 'Very Close') and ($S_{L(2)}$ is 'Close') and ... ($S_{L(y)}$ is 'Slightly Close') then $S_g$ is 'Very Similar'.

From the rule set, the training pattern {$(\|\delta_{i}\|, S_{L(1)}), (\|\delta_{i}\|, S_{L(2)}), \ldots, (\|\delta_{i}\|, S_{L(y)})$} is the general form of the antecedent part while an element in the Fuzzy set, $S_g = \{Very Similar, Similar, Slightly Similar\}$ is the fuzzy target output. The training patterns are evaluated by the ANFIS. The fact that Mamdani’s Fuzzy Inference is intuitive, has widespread acceptance and well suited for human cognition, informed its choice as the inference mechanism (Chai, Jia, & Zhang, 2009). The architecture of the ANFIS is presented in Figure 4. It is a Multiple Input, Single Output (MISO) architecture with five layers as in (Mendel, 2001; Inyang & Inyang, 2011). Out of the five layers, the first and fourth layers consist of adaptive nodes while the second, third, and fifth layers consist of fixed nodes. The input layer (Layer 0) has 5 nodes, each corresponding to a category of depression symptoms; they pass external crisp value to layer 1. Layer 1 consists of 15 fuzzification nodes; the outputs of this layer are the fuzzy membership grade defined by:

$$O_{ij} = \mu_A(x_i), \text{ for } i=1,2,3;$$
The triangular membership function in Equation (1) such that a < x < b is adopted. These outputs serve as the inputs of Layer 2. Layer 2 is the antecedent rules layer, it provides the firing strength of the rules with and operator as the T-norm. The output is given as follows:

\[ O_{2,i} = \omega_i = \mu_A(x_i) \times \mu_B(x_2) \times \mu_C(x_3) \times \mu_D(x_4) \times \mu_E(x_5), \quad i = 1, 2, 3, 4, 5 \]  

(20)

\[ \omega_i = \text{Min} \{ \mu_A(x_i) \times \mu_B(x_2) \times \mu_C(x_3) \times \mu_D(x_4) \times \mu_E(x_5) \} \]  

(21)

The output of the normalization layer, (Layer 3) is given in Equation (22)

\[ \bar{\sigma}_i = \frac{\omega_i}{\omega_1 + \omega_2 + \omega_3 + \omega_4 + \omega_5}, \quad i = 1, 2, 3, 4, 5 \]  

(22)

\[ \bar{\sigma}_i = \frac{\omega_i}{\sum_{i=1}^{5} \omega_i} \]  

(23)

The product of each rule’s normalized firing level (\( \bar{\sigma}_i \)) and the rule’s output (\( S^g_i \)) is represented in layer 4. For the \( i \)th node

\[ R_i = \bar{\sigma}_i S^g_i(\omega_i) \quad \text{where} \quad i = 1, 2, 3, \ldots, 243 \]  

(24)

The aggregation (sum) neuron is layer 5, it computes the overall system output as the sum of all incoming signals as given in Equation (25), which represent the global similarity match of the input case and a retrieved case.

\[ S = \sum_{i=1}^{243} (\bar{\sigma}_i S^g_i(\omega_i)) \]  

(25)
The result of the ANFIS for all the cases in the case base are ranked and evaluated by the decision support engine. The best 5 matches are selected by the cognitive filter and presented for analysis by the emotional filter. Some factors that may be considered by the emotional filter may include treatment resistance, suicide risk, hospitalization status and so on. The diagnostic decisions for the best matched case is used by the physician for the current case thereafter updated and retained as a new solved case in the case base.

4. Conclusion and Future Work

We have presented a neuro-fuzzy CBR model as a decision support system for the diagnosis of depression based on overall severity of symptoms. The proposed hybrid system framework introduces a similarity matching driven neuro-fuzzy architecture that provides flexibility for physicians measuring the severity levels of symptoms and symptoms’ category. The fuzzified local similarities of symptoms’ categories form the decision variables evaluated by the Mamdani-type ANFIS whose crisp output represent the global similarity between the query case and target case. In this work we have also proposed CBR and neuro-fuzzy hybrid framework for solving real life problems. In our future work, we intend to implement the model in an environment characterized by Windows 7 operating system, Microsoft Access database management system, MatLab and Java programming languages and to evaluate the performance of the system based on standard performance metrics. The use of a hybrid intelligent system undoubtedly offers superior benefit in terms of performance and usability over solitary intelligent systems. This work therefore, proposes an approach for diagnosing depression disorder which will also be extended to quantify the severity levels of other depression related dysfunctions like diabetes, cardiac, liver, lungs and cancer diseases.

References


