Case Retrieval Algorithm Using Similarity Measure and Fractional Brain Storm Optimization for Health Informaticians

Poonam Yadav
Department of Computer Science and Engineering, DAV College of Engineering and Technology, Maharshi Dayanand University, India

Abstract: *The management and exploitation of health Information is a demandingtask for health informaticians to provide the highest quality healthcare delivery. Storage, retrieval and interpretation of healthcare information are important stages in health informatics. Consequently, the retrieval of similar cases based on the current patient data can help doctors to identify the similar kind of patients and their methods of treatments. By taking into concern, a hybrid model is developed for retrieval of similar cases through the use of Case-based reasoning. Here, new measure called, parametric Enabled-Similarity Measure (PESM) is proposed and a new optimization algorithm called, Fractional Brain Storm Optimization (FBSO), by modifying the well known Brain Storm Optimization (BSO) algorithm with the addition of fractional calculus is proposed. For experimentation, three different patient dataset from UCI machine learning repository is used and the performance is compared with existing method using accuracy and f-measure. The average accuracy and f-measure reached by the proposed method with three different dataset is 89.6% and 88.8% respectively.*

Keywords: *Case-based reasoning, case retrieval, optimization, similarity, fractional calculus.*

Received April 1, 2015; accepted September 7, 2015

1. Introduction

Patient information retrieval is an extension of document retrieval or image retrieval [5, 14, 15]. But, it entirelydepends on the characteristics acquired by the patients and thus, it mustdeal with the issues associated with the gap that exists among the similarity measures used and the high-level semantics that a user is trying to find. Althougha lot of medical information retrieval systems are presented, it is not commonly used in real world medical applications [4, 10]. The searching of medical records hasmany challenges. If keyword based methods are used for searching the medical information, medical entities that are termed with different names cannot be accessed.

To avoid these kinds of issues resulting from keyword based methods, concept-based retrieval approaches have been proposed [16]. Thus, semantic measure and optimization-based neural network is collectively used for retrieval of patients' cases in this paper. The input for the proposed system is patient information, stored as health records which are given directly to parametric Enabled-Similarity Measure (PESM) measure along with query posted by doctor. The query is matched with stored health records to obtain the similar cases of patients. Similarly, Fractional Brain Storm Optimization (FBSO)-neural network is trained with history data with neighbour patients as output neurons. For the input query, neural network can predict its neighbour patient through its algorithmic procedure. Finally, hybrid measure combines these two results and produces the more suitable information to the doctor who can identify even well again by analysing the similar report done for those patients.

The main contributions of the paper are as follows:

- A novel measure called, PESM is proposed to match two patient cases using four different parameters with the assumption of occurrence and non-occurrence probability. Then, this measure is applied for case retrieval case.
- A novel optimization algorithm called, FBSO by changing the BSO algorithm with the addition of fractional calculus. Also, FBSO algorithm is applied to neural network trainings to identify the neighbour cases for the query patient case.
- Hybrid model is proposed newly by integrating PESM measure and FBSO-neural network for successfully retrieving patient case to easily identify the similar cases for the input query.

2. Motivating Scenario

The primary principle of Case-Based Reasoning (CBR) [1, 2, 7, 8, 11, 12] is that experience from past cases can be forced to answer new challenges. A human being's practice is called a case, and its collection is stored in a case base. The retrieval of similar cases based on the current patient data can help doctors to find similar type of patients and their treatments done along with sensitive information. Also, the record of old patients and their current health information can facilitate doctors to decide the medical prescription. These are the main motivations behind developing the proposed method. The patient information is stored in the patient case database and doctor uses medical diagnosis-support system to diagnose the similar cases to analyze their information.

Let C be the repository of patient information where each cases is represented as, $C_i\ ;1 \le i \le n$.

Here, represents the number cases stored in the database. Here, each cases is represented as attribute value pair like the representation used in case-based reasoning. The attribute-value pair for the CSR is expressed in Equation (1).

$$C_i : \{A_j, a_j\}; 0 \le j \le m \quad (1)$$

The goalhere is to do the retrieval of*k* similar cases by finding the most similar to the input query $Q_c$from '*n*' cases stored in *C*.

3. Retrieval Methods for Case-Based Reasoning Using Similarity Measure and FBSO

This section presents the proposed method for case-based reasoning by the use of new similarity measure and FBSO algorithm**.**Figure 1 shows the block diagram of the proposed retrieval method which have two major components:

1) PESM measure.
2) Designing of FBSO algorithm.

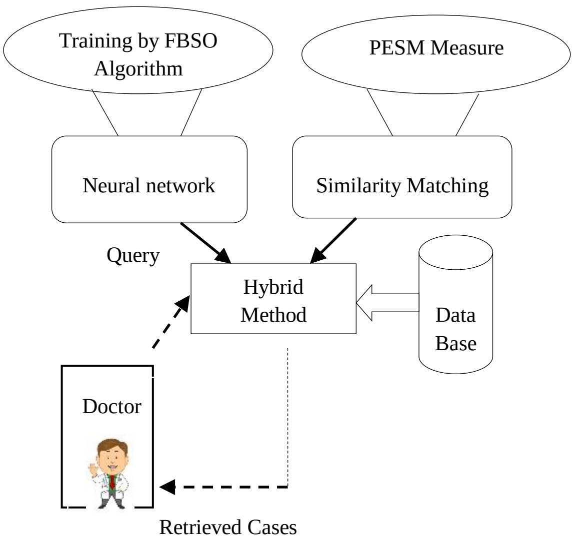


Figure 1. Block diagram of the proposed retrieval method.

3.1. PESMfor Case Retrieval

The proposed similarity measure, called PESM measure is used to retrieve the patient cases. Let us assume thatthe parameters of PESM measure x and y is initialized as, $x = 0$, $z = 1$. C_i and Cj are two cases from the data repository, represented in Equations (2) and (3).

$$C_i = \left[C_i^{(1)} \ C_i^{(2)} \ldots . C_i^{(m)} \right] \quad (2)$$

$$Cj = \left[C_j^{(1)} \ C_j^{(2)} \ldots . C_j^{(m)} \right] \quad (3)$$

Based on two cases, four parameters like, *P*1*(q)*, *P*2*(q)*, *P*3*(q)* and *P*4*(q)* are computed using the mutual occurrence of attributes in both cases. The occurrence of values in the attributes decides the similarity level of the two cases. The parameter, *P*1*(q)* provides maximum value if any one of the concern attributes' values of both cases are equivalent to the initial assignment of x. The second parameter *P*2*(q)* provides the maximum value if the concern attributes' values of both the cases are equal to x. The third parameter *P*3*(q)* can have a high degree if the values of concern attributes are not equivalent to x. The last parameter *P*4*(q)* provides the maximum similarity if the values of concern attributes should equal but not to be x. This is represented in Equations (4), (5), (6) and (7).

$$P1(q) = \begin{pmatrix} z\ ; & if\ C_i^{(q)} = x \ || \ C_j^{(q)} = x \\ x\ ; else & \end{pmatrix}; 0 \le q \le m-1 \quad (4)$$

$$P2(q) = \begin{pmatrix} z\ ; & if\ C_i^{(q)} = C_j^{(q)} = x \\ x\ ; else & \end{pmatrix}; 0 \le q \le m-1 \quad (5)$$

$$P3(q) = \begin{pmatrix} z\ ; if & C_i^{(q)} \ne C_j^{(q)} \ne x \\ x\ ; else & \end{pmatrix}; 0 \le q \le m-1 \quad (6)$$

$$P4(q) = \begin{pmatrix} z\ \ if & C_i^{(q)} = C_j^{(q)} \ne x \\ x\ ; else & \end{pmatrix}; 0 \le q \le m-1 \quad (7)$$

The parameters calculated for every values of attributes is then combined based on the following of Equations (8), (9), (10) and (11).

$$S1 = \sum_{q=1}^{m} P1(q) \quad (8)$$

$$S2 = \sum_{q=1}^{m} P2(q) \quad (9)$$

$$S3 = \sum_{q=1}^{m} P3(q) \quad (10)$$

$$S4 = \sum_{q=1}^{m} P4(q) \quad (11)$$

Based on the above values, the similarity degree, called PESM is defined as shown in Equation (12). Here, m is the number of attributes given in the case repository.

$$PESM(C_i, C_j) = \frac{1}{m}\left[\frac{z*S_1}{4} + \frac{z*_{S_2}}{2} + x*S_3 + z*S_4 \right] \quad (12)$$

For retrieval of '*k*' cases to the input case of Q_c, Q_c is matched with all the cases in the reposirtory using the PESM measure and then, cases are arranged in an descending order. Finally, top '*k*' cases are choosed to study the similar diagnosis.

3.2. FBSO for Neural Network Training

a) Fractional Brain Storm Optimization

Brain storm optimization is modified with mathematical theory called, Fractional Calculus (FC) [13], to enhance the solution searching in the predefined search space. In FBSO algorithm, ideas are represented as solution which is updated in each iteration.

- *Initialization*: Let us assume that n ideas are randomly initialized within the search space as, , where i = 1,2,....n and Dis the dimension of the solution which denotes the variable taken for optimization.
- *Grouping*: After initialization, ideas are grouped into two set of ideas based on k-means clustering algorithm where, k is number of clusters required.
- *Evaluation*: Every idea is then evaluated with fitness function.
- *Selection*: Three different probability values are used to select the clusters, selection of one or two clusters or another idea selection as per the probability values, like P5a P6b P6b3.
- *Updation*: The selected idea based on the above selection method is then updated with the following Equation. The equation used in BSO algorithm is given in Equation (13).

$$I_{t+1} = I_t + \xi N(\mu,\sigma) \tag{13}$$

The above equation can be written as,

$$I_{t+1} - I_t = \xi N(\mu,\sigma) \tag{14}$$

The left side I_{t+1}-I_t is the discrete version of the derivative of orderα=1, leading to Equation (15).

$$D^{\alpha}[I_{t+1}] = \xi N(\mu,\sigma) \tag{15}$$

The order of the velocity derivative can be generalized to a real number 0≤α≤ 1, if the FC perspective is considered, leading to a smoother variation and a longer memory effect. Therefore, the above equation can be written by considering the first r=4terms of differential derivative as shown in Equations (16) and (17).

$$I_{t+1} - \alpha I_t - \frac{1}{2}\alpha\ I_{t-1} - \frac{1}{6}\alpha(1-\alpha)I_{t-2} - \frac{1}{24}\alpha(1-\alpha)(2-\alpha)I_{t-3}6 = \xi N(\mu,\sigma) \tag{16}$$

$$I_{t+1} = \alpha I_t + \frac{1}{2}\alpha\ I_{t-1} + \frac{1}{6}\alpha(1-\alpha)I_{t-2} + \frac{1}{24}\alpha(1-\alpha)(2-\alpha)I_{t-3} + \xi N(\mu,\sigma) \tag{17}$$

Where, I_t is idea selected from the last iteration and $I_{t+1}$is to be newly generated data

$$I_t = \begin{cases} I_{ij} & ;one\ \ cluster \\ w_1\ I_{i1},j + w_2 I_{i2,j} & ;two\ \ cluster \end{cases} \tag{18}$$

$$\xi = r \log sig\left(\frac{Nc_{\max}/2_{-Nc}}{K}\right) \tag{19}$$

Where, r is a random value between 0 and 1. Nc_{max} and Nc denote the maximum number of iteration and current number of iteration respectively. K adjusts the slope of the logsig function.

- *Crossover*: After the new idea is created, a crossover between the new one and the old one is conducted. The two ideas generated by crossover, together with the old one and the created one, are evaluated using the fitness function and the old one is replaced with the best of the four.
- *Termination*: The process above repeats until m ideas are updated. Thus, one generation is finished. The iteration goes until terminal requirement is met. Then, the best idea is output as the optimal solution to the problem. It is given in Equation (20).

$$O_d = O_{H_N}^{(l)} \tag{20}$$

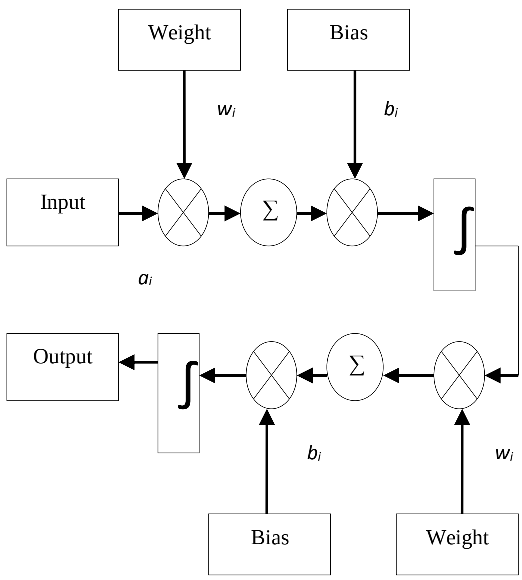


Figure 2. Block diagram of the proposed retrieval method.

b) Neural Network Training by FBSO Algorithm

Neural network training is the process of identifying the weights for neurons, suitable for the input cases taken as an input. Training is an iterative process of identifying weights by changing weights in every iteration. Here, FBSO algorithm is adapted to do neural network training by finding the optimal weights as like the output fixed as [6].

- *Idea encoding*: In FBSO algorithm, idea is encoded. The weights utilized in the neural network structure are put as vector which is an idea of the proposed FBO algorithm so the size of an idea is equivalent to the number of weights in the neural network including neuron weights and bias weights.

- *Algorithmic procedure*: At first, input weights of n vectors (ideas) are randomly initialized within the search space and k-means clustering algorithm is applied. The ideas are evaluated with fitness function. Based on three probabilities, ideas are selected and updated using the proposed updating equation and cross over are applied. The two ideas generated by crossover, together with the old one and the created one, are evaluated using the fitness function and the old one is replaced with the best of the four. This process is repeated until ideas are updated. Thus, one generation is finished. The iteration goes until terminal requirement is met.
- *Fitness function*: The fitness function is computed by giving training cases to neural network architecture and the weights in taken idea for evaluation are filled in the neural network. The output for the taken idea is computed as per the mathematical model given in the neural network architecture. So, the output can be obtained for all the input cases which is then given for the fitness function. It is shown in Equation (21).

$$fitness=\frac{1}{c}\left[\sum_{i=1}^{c} O_d - O_g\right] \tag{21}$$

Where, *O_d* is output obtained from neural network, is original neighbour index value from training cases and is number of training cases.

3.3. Hybrid Retrieval Method for Case-Based Reasoning

For a query case Q_C given by a doctor, PESM measure provides the similarity measure for every case stored in the training cases and neural network provides the neighbour index values by predicting it. The score value is then combined using the following equation and the final score value is generated for all the cases. The top-k cases are then extracted to be given for doctor to analyse the medical reports and prescriptions. It is shown in Equation (22).

$$H_{Q_c} = \frac{1}{2}\left[Q_c^{(PESM)} + Q_c^{(NN)}\right] \tag{22}$$

Here, the number of cases to be retrieved is defined under the variable, called, ‘k’. Based on the user (doctor) requirement, the number of cases to be retrieved (k) can be fixed. For example, if a doctor wants to see only one similar case for the query Q_c, then ‘k’ value can be fixed to 1 and the most similar case based on H_{Q_c} is retrieved.

4. Results and Discussion

The experimental results of the proposed method are discussed in this section and the performance of method is also discussed in detail with two different metrics.

4.1. Experimental Set up

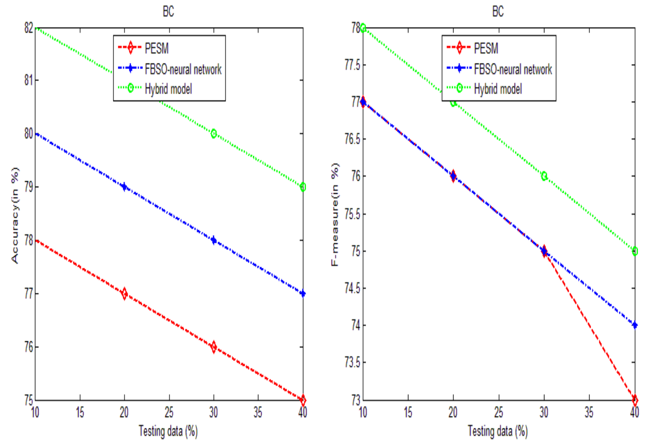
- *Platform*: The proposed retrieval method is implemented using MATLAB 8.2.0.701 (R2013b) with a system configuration of 2GB RAM Intel processor and 32 bit OS.
- *Datasets utilized*: The datasets are taken from UCI machine learning repository [3].
- *Parameter initialization*: For the reason of validating the effectiveness and usefulness of the proposed retrieval method, a set of parameters are set naturally for the process of FBSO, neural network and core method as like the values given in Table 1.

Table 1. Parameter initialization

<table>
<tr><td rowspan="2">FBSO parameters</td><td>n</td><td>P_{5a}</td><td>P_{6b}</td><td>P_{6b3}</td><td>Nc_{max}</td><td>K</td><td>μ</td><td>σ</td></tr>
<tr><td>5</td><td>0.2</td><td>0.8</td><td>0.4</td><td>2000</td><td>20</td><td>0</td><td>1</td></tr>
<tr><td rowspan="2">Neural network parameters</td><td colspan="2">N</td><td colspan="2">h</td><td colspan="2">h_1</td><td colspan="2">h_2</td></tr>
<tr><td colspan="2">3</td><td colspan="2">5</td><td colspan="2">10</td><td colspan="2">15</td></tr>
<tr><td rowspan="2">Overall parameters</td><td colspan="8">k</td></tr>
<tr><td colspan="8">5</td></tr>
</table>

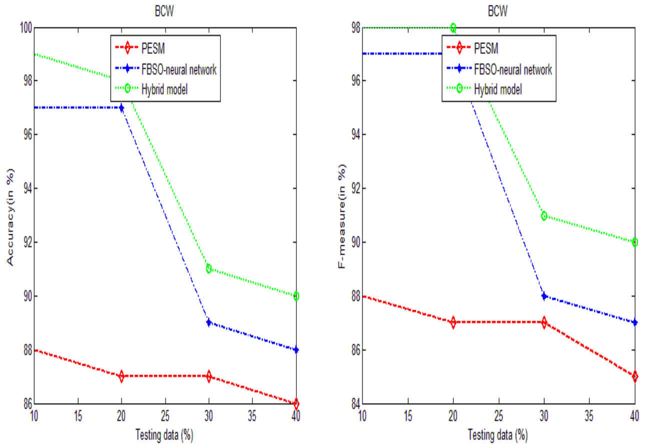
4.2. Performance Analysis

The performance of the proposed hybrid method is analyzed with two other variants of the proposed method called, PESM and FBSO-neural network. The size of the data is changed and the performance is analyzed for three different datasets. Figure 3 shows the performance graph of proposed methods in BC datasets. Here, hybrid method shows the better performance as compared with other two methods. The maximum accuracy of 82% is reached by the hybrid model when the testing data size is 10% and for the same data size, PESM method and FBSO-neural network obtains the value of 80% and 78% respectively. Similarly, hybrid model achieved 78% for the BC datasets where, PESM and FBSO-neural network obtained the same value of 77%. Figure 4-a shows the performance graph in terms of accuracy in BCW dataset and Figure 4-b shows the performance graph in terms of F-measure in BCW dataset. From the Figure 4-a, hybrid model attained about 99% accuracy by comparing with other two methods for the data size of 10%. The minimum accuracy for the hybrid model is 90% as compared with other two methods for the data size of 40%. Similarly, hybrid model attained about 98% f-measure by comparing with other two methods for the data size of 10%. The minimum accuracy for the hybrid model is 97% as compared with other two methods for the data size of 40%.



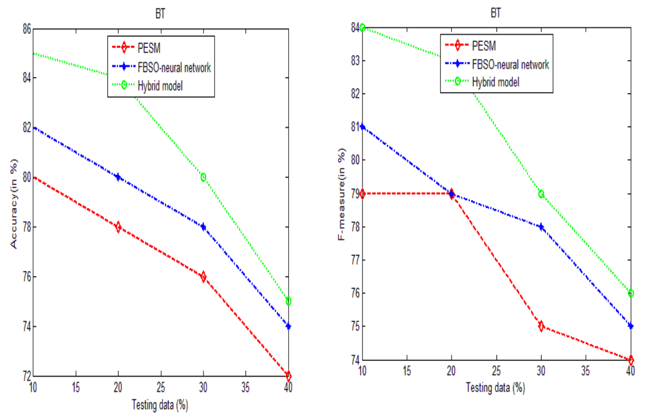
a) Accuracy. b) F-measure.

Figure 3. Performance graph in Breast Cancer (BC).



a) Accuracy. b) F-measure.

Figure 4. Performance graph in Breast Cancer Wins (BCW).



a) Accuracy. b) F-measure.

Figure 5. Performance graph in Breast Tissue (BT).

Figure 5-a shows the accuracy graph of hybrid model, PESM model and FBSO neural network in BT datasets. When the size of the testing data is 10%, hybrid model achieved 85% where, PESM and FBSO neural network obtained the value of 82% and 80% respectively. The minimum accuracy reached by the hybrid model is 75%. Figure 5-b shows the f-measure graph of hybrid model, PESM model and FBSO neural network in BT datasets. When the size of the testing data is 10%, hybrid model achieved 84% where, PESM and FBSO neural network obtained the value of 81% and 79% respectively. The minimum f-measure reached by the hybrid model is 76%.

4.3. Comparative analysis

Table 2 shows the comparative performance of the taken methods for the six different datasets using accuracy and Table 3 shows the F-measure values of different methods. From Table 2, the performance is improved when K value is increased. For the K value 3, the accuracy attained by the proposed system outperforms the accuracy achieved by the existing systems for all data sets taken for the experimentation. In BT datasets, the accuracy attained by the proposed hybrid model is 84% which is higher considering all the existing systems and the other two proposed models. Similarly for the k value 5, accuracy achieved by the proposed system is higher overcoming the accuracy of the existing systems in the NBC, BCW, BT, PID, SHD and THY datasets. The accuracy of 90% is achieved by the proposed hybrid model for the SHD which is greater compared to all the existing systems and proposed systems for the k value 5. The maximum accuracy of 99% is achieved by the proposed method in THY.From the accuracy values in the table 3, it is clear that the accuracy increases along with the value of the K. From table 3, F-measure of 77% is reached by the hybrid model in BC datasets for the K value 3. In BCW datasets, hybrid model outperformed by reaching the F-measure of 98% but other two proposed methods are less than the existing method, USIMSCAR (MV &WV) [9]. F-measure of 83% is reached by the hybrid model and shows the top performance in BT datasets. In PID data sets, the F-measure achieved by the proposed system is 88% and 88.6% for the k value 3 and 5 respectively which is far greater compared to the F- measure value of the existing methods. The proposed method PESM and FBSO-NN hasalso attained the F-measure value greater than the existing but less than the hybrid model which proves the competence of the proposed hybrid model. The maximum F-measure value of 98.5% and 98%is attained by the proposed hybrid model in the THY and BCW data sets for the K value of 5 and 3 respectively. Table 4 shows the execution time of the algorithm on different datasets.The values resembling the computation time of the optimization shows that the hybrid model requires more computation time than the PESM and FBSO-NN model since it combines the action of both of them, thereby increasing the effectiveness of the retrieval. Even though the execution time of the hybrid model is increased compared the other two proposed methods PESM and FBSO-NN, the computation time is lower compared to the existing methods USIMSCAR (MV) and USIMSCAR(WV) [9], ABC-NN and CS-NN for all the datasets taken for the experimentation shows the enhanced performance. The execution time of the hybrid model for the BC dataset is 5.4 sec which is 1.4 sec less than that of USIMSCAR (MV) [9], 0.7 sec less than that of USIMSCAR (WV) [9], 1.9 sec less than that of ABC-NN, 3 sec less than that of CS-NN. In the SHD datasets, the execution time of the proposed hybrid model is 6.2 sec whereas the existing systems require more than 7 sec for the computation. For THY datasets, the execution time elapsed by the proposed model is 4.9 sec. But the existing systems obligate more time for the computation varying over from 5 seconds in THY data sets. On comparing the execution time given in the Table 4, the computation time requirement for the hybrid model is reduced than that of existing works.

Table 2. Accuracy comparison.

DATASET	K value	BC	BCW	BT	PID	SHD	THY
USIMSCAR(MV) [9]	K=3	75.87	97.66	71.7	75.78	83.33	97.67
	K=5	76	97	71.8	76	84	97
USIMSCAR(WV) [9]	K=3	79.02	97.66	78.3	87.5	**89.63**	97.67
	K=5	79	97.1	78.5	88	**90**	96
PESM	K=3	77	87	78	86	87	88
	K=5	77.5	88	79	86.5	88	89
FBSO-NN	K=3	79	97	80	87	87	94
	K=5	79.5	97.5	81	87.5	88	92
Hybrid model	K=3	**81**	**98**	**84**	**88**	89	**98**
	K=5	**81.7**	**98.1**	**84.5**	**88.9**	90	**99**

Table 3. F-measure comparison.

DATASET	K value	BC	BCW	BT	PID	SHD	THY
USIMSCAR(MV) [9]	K=3	68.74	97.42	71.18	72.21	83.08	96.88
	K=5	69	98	72	75	84	97
USIMSCAR(WV) [9]	K=3	74.25	97.42	77.63	86.14	**89.55**	96.88
	K=5	75	97.6	78	87	**91**	97
PESM	K=3	76	87	79	85	86	89
	K=5	76	88	79.5	84.5	86.5	89.5
FBSO-NN	K=3	76	97	79	87	87	93
	K=5	77	98	76	87.6	87.5	94
Hybrid model	K=3	**77**	**98**	**83**	**88**	89	**98**
	K=5	**78.1**	**98.5**	**84**	**88.6**	90	**98.5**

Table 4. Computation time (in sec).

DATASET	BC	BCW	BT	PID	SHD	THY
USIMSCAR(MV) [9]	6.5	6.1	4.4	8.5	7.1	5.3
USIMSCAR(WV) [9]	6.1	7.3	5.1	11.1	8.1	6.6
ABC-NN	7.3	9.8	4.5	9.5	9	6.9
CS-NN	8.4	8.5	3.2	8.1	10.1	5.4
PESM	4.2	6.4	2.1	6.7	5	3.9
FBSO-NN	3.7	6.7	1.9	9.1	6.1	4.2
Hybrid model	5.4	7.4	3.2	8.1	6.2	4.9

5. Conclusions and Future Scope

We have presented a case retrieval method using similarity measure and fractional brain storm optimization for health informaticians. The proposed hybrid model is integrated with PESM measure and FBSO-neural network. At first, input data base of patient cases is given as input to PESM method and FBSO-based neural network. In PESM measure, query case is matched with historic data to identify similar cases. Also, FBSO algorithm is applied to neural network trainings to identify the neighbour cases for the query patient case. Finally, both outputs are effectively combined to obtain the final retrieval of cases to easily identify the existing diagnosis of similar patients for a doctor. The experimentation is conducted with three different patient dataset from UCI machine learning repository and performance is compared with existing method using accuracy and f-measure. The average accuracy and f-measure obtained by the proposed hybrid method over three different datasets is 89.6% and 88.8% respectively. The future work can be in the direction of including semantic and pragmatic concept in developing similarity measure for case retrieval.

References

[1] Ahn H. and Kim K., “Global Optimization of Case-based Reasoning for Breast Cytology Diagnosis,” *Expert Systems with Applications*, vol. 36, no. 1, pp. 724-734, 2009.

[2] Chuang C., “Case-based Reasoning Support for Liver Disease Diagnosis,” *Artificial Intelligence in Medicine*, vol. 53, no. 1, pp. 15-23, 2011.

[3] Datasets from UCI Machine Learning Repository, http://archive.ics.uci.edu/ml/, Last Visited, 2015.

[4] Depeursinge A., Duc S., Eggel I., and Muller H., “Mobile Medical Visual Information Retrieval,” *IEEE Transactions on Information Technology in Biomedicine*, vol. 16, no. 1, pp. 53-61, 2012.

[5] Elabd E., Alshari E., and Abdulkader H., “Semantic Boolean Arabic Information Retrieval,” *The International Arab Journal of Information Technology*, vol. 12, no. 3, pp. 311-316, 2015.

[6] Greenspan H., Müller H., and Syeda-Mahmood T., “Medical Content-Based Retrieval for Clinical Decision Support,” *in Proceedings of MICCAI International Workshop on Medical Content-Based Retrieval for Clinical Decision Support*, Nice, 2012.

[7] Guezouli L. and Kadache A., “Information Rretrieval Model Based on Neural Networks using Neighborhood,” *in Proceedings of International Conference on Information Technology and e-Services*, Sousse, pp. 1-5, 2012.

[8] Guo Y., Hu J., and Peng Y., “Research on CBR System Based on Data Mining,” *Applied Soft Computing*, vol. 11, no. 8, pp. 5006-5014, 2011.

[9] Huang M., Chen M., and Lee S., “Integrating Data Mining with Case-based Reasoning for Chronic Diseases Prognosis and Diagnosis,” *Expert Systems with Applications*, vol. 32, no. 3, pp. 856-867, 2007.

[10] Kang Y., Krishnaswamy S., and Zaslavsky A., “A Retrieval Strategy for Case-Based Reasoning

Using Similarity and Association Knowledge," *IEEE Transactions on Cybernetics*, vol. 44, no. 4, pp. 473-487, 2014.

[11] Pandey B. and Mishra R., "Case-based Reasoning and Data Mining Integrated Method for the Diagnosis of Some Neuromuscular Disease," *International Journal of Medical Engineering and Informatics*, vol. 3, no. 1, pp. 1-15, 2011.

[12] Park Y., Choi E., and Park S., "Two-step Filtering Datamining Method Integrating Case-based Reasoning and Rule Induction," *Expert Systems with Applications*, vol. 36, no. 1, pp. 861-871, 2009.

[13] Pires E., Machado J., Oliveira P., Cunha J., and Mendes L., "Particle Swarm Optimization with Fractional-order Velocity," *Nonlinear Dynamics*, vol. 61, no. 1-2, pp. 295-301, 2010.

[14] Shahid A., Afzal M., and Qadir M., "Lessons Learned: The Complexity of Accurate Identification of in-Text Citations," *The International Arab Journal of Information Technology*, vol. 12, no. 5, pp. 481-488, 2015.

[15] Taleb N., "Using Ontologies for Extracting Differences in the Dynamic Domain: Application on Cancer Disease," *The International Arab Journal of Information Technology*, vol. 13, no. 1, pp. 125-131, 2013.

[16] Zuccon G., Koopman B., Nguyen A., Vickers E., and Butt L., "Exploiting Medical Hierarchies for Concept-based Information Retrieval," *in Proceedings of the 17th Australasian Document Computing Symposium*, New Zealand, pp. 111-114, 2012.

Poonam Yadav obtained B.Tech in Computer Science & Engg. From Kurukshetra University Kurukshetra and M.Tech in Information Technology from Guru Govind Singh Indraprastha University in 2002 and 2007 respectively. She had Awarded Ph.D in Computer Science & Engg. From NIMS University, Jaipur. She is currently working as Principal in D.A.V College of Engg. & Technology, Kanina (Mohindergarh). Her research interests include Information Retrieval, Web based retrieval and Semantic Web etc. Dr. PoonamYadav is a life time member of Indian Society for Technical Education.