Patient Record Retrieval Using Bhattacharya coefficient with Crow Search Algorithm

Dr.Poonam Yadav

Assistant Professor, D.A.V College of Engineering & Technology,

Kanina, Haryana 123027, India

Abstract: In recent years, health information management and utilization is the demanding task to health informaticians for delivering the prominence healthcare services. Extracting the similar cases from the case database can aid the doctors to recognize the same kind of patients and their treatment details. Accordingly, this paper presents the method called CS-BCF for retrieving the similar cases from the case database. Initially, the patient's case database is constructed with details of different patients and their treatment details. If the new patient comes for treatment, then the doctor collects the information about that patient and sends the query to the CS-BCF. The CS-BCF system matches the input query with the patient's case database and retrieves the similar cases. Here, the CS algorithm is used with the BCF for retrieving the most similar cases from the patient's case database. Finally, the Doctor gives treatment to the new patient based on the retrieved cases. The performance of the proposed method is analyzed with the existing methods, such as PESM, FBSO-neural network, and Hybrid model for the performance measures DCG and fall-out. The experimental results show that the proposed method attains the higher DCG of 19.324 and the minimum fall-out 0.042when compared to the existing methods.

Keywords: case retrieval, BCF, CSA, similar cases, fall-out.

1. Introduction

Now a day, the enlargement of information technology has altered the entire human life which includes medical and healthcare behaviors. In medicine and healthcare, the information technology is used in several fields, such as medical imaging systems, medical diagnosis systems, EPR (Electronic Patient Record), EMR (Electronic Medical Record), HIS (Hospital Information System), and so on. These systems have various functions, but the purpose of increasing the effectiveness and efficiency of the medical behaviors is similar in all these systems [19] [6] [35]. A medical information storage and retrieval system is a precious tool used by the healthcare professionals for investigating the medical cases. In clinical decision making and healthcare analysis, the EHR (Electronic Health Records) has become a valuable tool

[20] [21]. In applications, like clinical decision support and patient cohort identification, the IR (Information Retrieval) that retrieves the documents from the huge collection of data depending on the user query has become popular [22] [2].

MCR (Medical case retrieval) is the process of determining the descriptions of health records of patients or diseases which are similar to the query sent by the medical experts. MCR plays the significant role in clinical decision support systems [23] and employs the concept of case-based reasoning (CBR) [24]. In case-based reasoning, the most similar cases are retrieved from the case database for a given query, and the processes, such as diagnosis and treatment are applied to the patient. Anyhow, in medical research and education, MCR is the important problem because it permits to choose the exciting cases and create the datasets for medical studies visiting the case-based criteria [7]. The case-based medical information retrieval system gives power to the healthcare experts to determine the similar cases and related publications from the medical information repositories [5].

The CBR utilizes the experience from the previous cases to solve the new problems. A case is defined as the experience of the individual and the case database consists of a number of cases. Every case in the case database is represented by a problem description and the solution description. The CBR consists of four phases, namely retrieval phase, reuse phase, revise phase and retain phase from which the retrieval is the important phase because the victory of CBR systems depends on the performance of this phase [25]. The major goal of the retrieval phase is to retrieve the cases from the case database which are similar to the query. If the retrieval phase does not retrieve the similar cases for the query, then the CBR system provide the incorrect solution to the new problem [3]. Additionally, the information retrieval is facilitated by advanced compression methods [31] [32] and modern data transmission technologies [29] [30] as well as intelligence methods exploited in many applications [27] [28].

This paper introduces the case retrieval system named as CS-BCF for retrieving the similar case from the patient's case database. Initially, the patient's case database is constructed with health details of different patients and their treatment details. If the new patient comes for treatment, then the doctor collects the health information about that patient and sends the query to the CS-BCF. The CS-BCF system matches the input query with the patient's case database and retrieves the similar cases. Here, the CSA algorithm is utilized with the BCF for retrieving the most similar cases from the patient's case database. Since many heuristic methods and intelligence methods [27] [28] are utilized in many applications [33] [34]. Finally, the Doctor gives treatment to the new patient based on the retrieved cases.

The rest of the paper is organized as follows: Section 2 presents the literature review, and Section 3 presents the system model of the proposed case retrieval method. In Section 4, the proposed case retrieval method is described. Results and discussions are provided in Section 5 and Section 6 concludes the paper.

2. Literature Review

This section presents the review of existing research works on case retrieval and the major challenges in the case retrieval process.

2.1 Motivation

Here, the five existing works on case retrieval is discussed, and the advantages and disadvantages of each work are specified. Sungbin Choi *et al.* [1] have proposed a case retrieval method by integrating a semantic concept-based term-dependence feature and formal retrieval model to enhance the performance of the ranking. The advantages of this method are that it was robust and effective, enhances the primary evaluation metrics. The disadvantage is that the recall of this method was poor. Yanshan Wang *et al.* [2] have proposed two NLP-empowered IR models, POS-BoW and POS-MRF, which integrates the automatic POS-based term weighting techniques into Markov Random Field (MRF) IR models and bag-of-word (BoW). This method was efficient than the existing frequency based and heuristic based weighting techniques. This method had two drawbacks; the first one is it depends on the tagging performance of POS, the second one is this method did not consider the external data resources.

Yong-Bin Kang *et al.* [3] have proposed a Retrieval Strategy named as USIMSCAR for Case-Based Reasoning Using Similarity and Association Knowledge (AK). The advantage of this method is that the AK can be built from the case base in a simple manner. This method was not suitable for complex structures, like semantic web-based cases, hierarchical cases, and object-oriented cases. Adrien Depeursinge *et al.* [4] have proposed mobile access to peer-reviewed medical information depends on the content based and textual search visual image retrieval with optimized communication bandwidth and good execution speed. The limitation of this method was that this method did not have the accessibility to the camera. André Mourão *et al.* [5] have proposed a medical information retrieval system for multimodal medical case-based retrieval with unsupervised rank fusion. This method helps the users to retrieve the relevant information and minimizes the frustration. The results generated by this method were poorer because of the performance difference among the image runs and text.

2.2 Challenges

The medical applications present a number of challenges for the CBR researchers and drive advances in research. Significant research issues are listed below:

- Feature extraction is the problematical task in the topical medical CBR systems because of a complex data format in which the data comes from free-text format or time series or images or sensors.
- Most of the CBR systems are based on the knowledge of the experts to perform feature selection and weighting. Cases with concealed features influence the performance of retrieval.
- Most CBR systems evade the automatic adaptation techniques because of several problems, like risk analysis, reliability, a large number of features, rapid changes of medical knowledge, the complexity of medical domains, and so on [18].
- A drawback which is general to the methods [26] is that these methods describe similarity for a single term. The other drawback is that most of these methods do not deal with the medical domain, they only deal with the SNOMED-CT database [17].

3. System Model

This section presents the system model of the proposed CS-BCF case retrieval method. Here, the major aim is to extract the medical cases from the patient case database which are similar to the input query. Figure1. represents the system model of the proposed CS-BCF for retrieving similar cases from the case database for the new patient. Here, the Doctor gathers the medical information, such as blood pressure, height and weight of the patient, and so on from the new patient and sends these details in the form of a query to the medical diagnosis system which compare these details with the patient case database and extracts the similar cases. Then, the doctor provides the treatment for the new patient depending on the retrieved cases. Each case in the patient case database is represented is represented as an attribute- value pair, and for each case, the problem description and the corresponding solution description are represented.

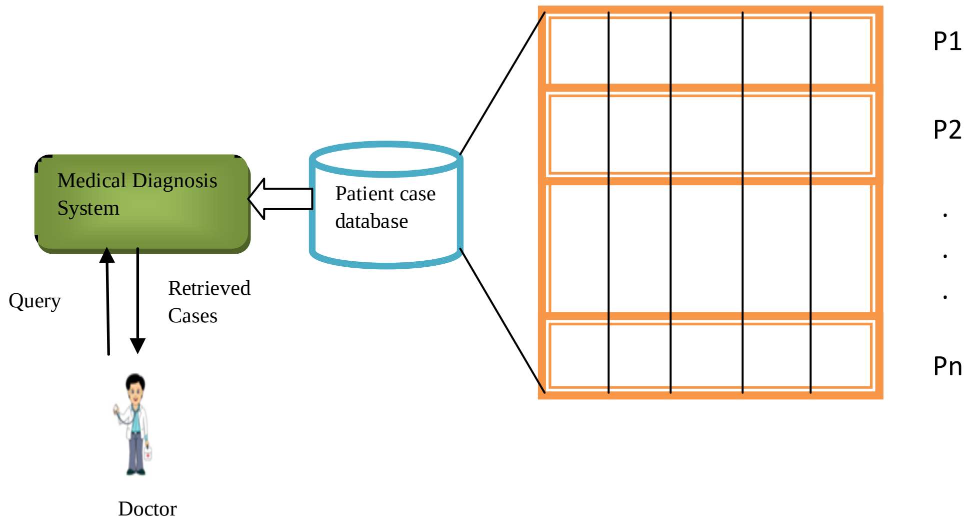


Figure1. System model of CS-BCF case retrieval method

4. Proposed CS-BCF for retrieving similar cases from the patient case database

This section presents the proposed for extracting the similar cases from the patient's case database. The proposed system utilizes the BCF function and CS algorithm for extracting the cases which are similar to the input query. Figure2. shows the block diagram of the proposed for extracting the similar cases from the patient's case database. The medical details about patients, such as height and weight of the patients, blood pressure, and so on are gathered and kept in a patient's case database, and every case is represented as an attribute value. When the new patient comes for a treatment, the doctor collects the medical details of that patient and sends the input query to the. The system matches the input query with the cases presented in the patient's case database and extracting the cases which are similar to the input query. Here, the CS algorithm is utilized with the BCF for extracting the most similar cases from the patient's case database. Finally, the Doctor provides the treatment to the new patient depending on the retrieved cases.

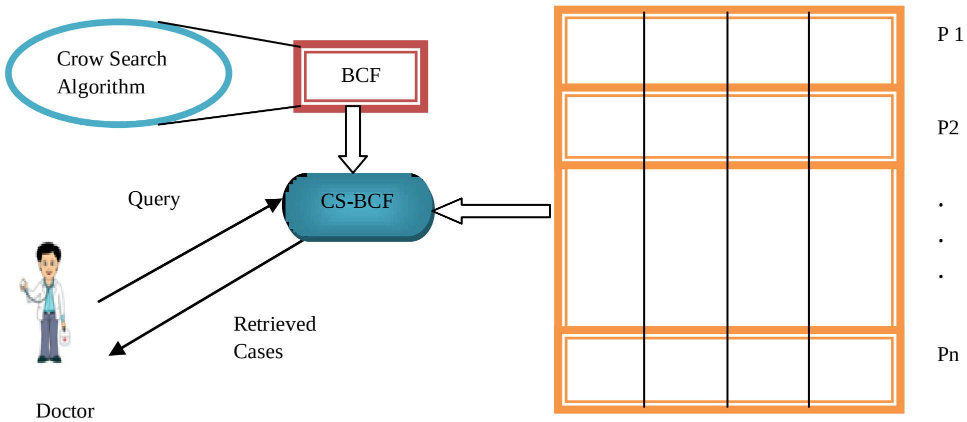


Figure2. Block Diagram of the proposed case retrieval model

Consider the case database of patients in which the medical information about the various patients is stored. The medical information includes weight and height of the patients, blood pressure rate, and so on. Also, the treatment taken for the diseases are kept in the case database. The following equation represents the patient case database,

$$CD = \{CD_i : 1 \le i \le n\} \tag{1}$$

where, CD is the patient's case database and n is the number of medical cases stored in the database. Every case in the case database is represented as an attribute. This can be represented as follows,

$$CD_i = \{a_j ; 0 \le j \le m\} \tag{2}$$

where, a_j represents the attribute and m represents the number of attributes. When the new patient comes for the treatment, the doctor gathers the details about diseases of patients, the treatments which are previously taken by those patients, and send the query to the medical diagnosis system. The query send by the doctor is represented as follows,

$$Q = \{q_p ; 0 \le p \le s\} \tag{3}$$

where, Q represents the query. q_p represents the number of queries. Then, the input query send by the doctor is matched with the patient's case database and the cases which are similar to the input query are retrieved from the case database. Here, the similarity function calculation is performed by the BCF.

4.1 BCF for calculating the similarity function

The similarity among the input query and the medical case in the patient's case database is calculated by the Bhattacharyya Coefficient in Collaborative filtering (BCF). The BCF combines the global and local similarity values to acquire the most extreme similarity value. Collaborative filtering (CF) is the triumphant recommendation system [8-10] in the modern years. Here, the recommendation of items to the user is proficient by investigative the other users or other item's rating information. The motive for utilizing the collaborative filtering is that it has a number of advantages, such as domain independent and precise. The Bhattacharyya measure is engaged in numerous areas, like pattern recognition [11, 12], signal processing, and image processing which calculates the similarity between two probability distributions [13].

The input query is matched with the medical cases existing in the patient's case database, and the cases which are similar to the input query are fetched from the database. The similarity among the input query and the medical cases which are already offered in the patient's case database is calculated by the equation (4).

$$BCF(Q, CD_i) = \alpha\left[\sum_{j=1}^{m}\sum_{p=1}^{s} BC(j,p)\, loc_{cor}(r_j, r_p)\right] + \beta\left[\sum_{j-1}^{m}\sum_{p=1}^{s} BC(j,p)\, loc_{med}(r_j, r_p)\right] \tag{4}$$

where, $BC(j,p)$ represents the global similarity information between the input query p and the case attribute j.The local similarity value grasps the major role, and it creates the local similarity information. α and β values are calculated by the Crow search algorithm. If the input query p and the case attribute j are similar to each other, then $BC(j,p)$ enhances the local similarity among p and j. If the input query p and the case attribute j are different from each other, then $BC(j,p)$ decreases the importance of local similarity between p and j. There are two functions to calculate the local similarity, namely $loc_{cor}(\ )$ and $loc_{med}(\ )$. The $loc_{cor}(\ )$ function measures the correlation amongst the case attribute and the input query which can be denoted in the following equation,

$$loc_{cor}(r_j, r_p) = \frac{(r_j - \bar{r})(r_p - \bar{r})}{\sigma_i, \sigma_j} \tag{5}$$

where, r_j represents the case attribute and r_p represents the input query. σ_i is the standard deviation of the i^{th} case and σ_j is the standard deviation of the attribute j. Then, the local similarity is calculated by the $loc_{med}(\ )$ function is denoted as follows,

$$loc_{med}\left(r_j, r_p\right) = \frac{\left(r_j - r_{med}\right)\left(r_p - r_{med}\right)}{\sqrt{\sum_{j=1}^{m}\left(r_j - r_{med}\right)^2}\sqrt{\sum_{p=1}^{s}\left(r_p - r_{med}\right)^2}} \tag{6}$$

where, r_{med} is the median of the case attributes. r_j represents the case attribute and r_p represents the input query. The global similarity information among the input query and the medical cases in the case database is calculated by the function $BC(j, p)$. The BC similarity among the case attribute j and the query p is calculated as follows,

$$BC(j, p) = \sum_{k=1}^{z}\sqrt{\left(\overset{j}{P_k}\right)\left(\overset{p}{P_k}\right)} \tag{7}$$

where, $BC(j, p)$ produces the global similarity information among the input query p and the case attribute j. The following equation calculates the value of $\overset{j}{P_k}$.

$$\overset{j}{P_k} = \frac{\#k^j}{n} \tag{8}$$

where, $\#k^j$ is the number of cases having k^{th} attribute value and n is the number of cases in the patient case database. $\overset{p}{P_k}$ can be calculated by the following equation,

$$\overset{p}{P_k} = \frac{\#k^p}{n} \tag{9}$$

where, $\#k^p$ represents the number of queries having the k^{th} attribute value and n is the number of cases in the patient case database.

k- retrieval: The input query is matched with the medical cases presented in the patient case database, and the BCF function computes the similarity between the query input and the medical cases presented in the case database. The BCF function calculates the local and global similarity among the cases and integrates these similarity values to obtain the optimal similarity. Finally, the k numbers of similar cases having the most similarity are taken as the retrieved cases.

4.2 Optimal weight coefficient using Crow search Algorithm

Crow search Algorithm (CSA) [14] is a population-based procedure that works based on the behavior of crows which kept a large amount of food in this idea that crows store their excess food in trouncing places and regain it when required. CSA has smaller number of parameters to regulate therefore it is simple to implement.

The major principles of CSA are,

- Crows are living together (flock).
- Crows remember the position of their trouncing places.
- Crows follow each other to do the stealing.
- Crows preserve their caches from being stolen.

Let assume that the D dimensional environment which has a number of crows. Here, the dimension of the solution is two which denotes the alpha and beta.

The flock size (number of crows) is represented as N and the position of crow C is represented as a vector, $Y^{C,I}(C=1,2,\ldots,N; I=1,2,\ldots,I_{\max})$ where $Y^{C,I}=\left[Y_1^{C,I},Y_2^{C,I},\ldots,Y_D^{C,I}\right]$ and $I_{\max}$ is the maximum number of iterations. Every crow has a memory where the position of its trouncing place is memorized. At iteration I, the position of the trouncing place of crow C is represented as $P^{C,I}$. Consider the crow C_2 needs to visit its trouncing place $P^{C_2,I}$ at iteration I, then the crow C_1 follows the crow C_2 to reach the trouncing place of crow C_2. In this situation, two states may happen.

State 1: Crow C_2 does not know that crow C_1 is following it. Here, crow C_1 reaches the trouncing place of crow C_2 and the new position of the crow C_1 is calculated as follows,

$$Y^{C_1,I+1}=Y^{C_1,I}+x_{C_1}\times L^{C_1,I}\times\left(P^{C_2,I}-Y^{C_1,I}\right) \tag{10}$$

where, x_{C_1} is a random number with a uniform distribution between 0 and 1. L is the flight length of crow C_1 at iteration I.

State 2: Crow C_2 knows that crow C_1 is following it. Here, the crow C_2 will fool crow C_1 by go to the other position to protect its cache from being stolen.

The state 1 and state 2 is represented as follows,

$$Y^{C_1,I+1}=\begin{cases} Y^{C_1,I}+x_{C_1}\times L^{C_1,I}\times\left(P^{C_2,I}-Y^{C_1,I}\right) & ;\, x_{C_2}\geq A^{C_2,I} \\ a\ random number & ;\, otherwise \end{cases} \tag{11}$$

where, $A^{C_2,I}$ is the awareness probability of crow C_2 at iteration I and x_{C_2} is a random number with a uniform distribution between 0 and 1.

Algorithm:

Step 1: Initialization:

The optimization problem, decision variables and constraints are defined. Then, the adjustable parameters of CSA, flock size (N), a maximum number of iterations $(I_{\max})$, flight length (L) and awareness probability (A) are valued.

Step 2: Initialize the memory and position of crows:

Here, N numbers of crows are crows are positioned in a D-dimensional search space randomly as the members of the flock. Every crow in the flock represents the feasible solution.

$$Crows = \begin{bmatrix} Y_1^1 & Y_2^1 & \dots & Y_D^1 \\ Y_1^2 & Y_2^2 & \dots & Y_D^2 \\ \vdots & & & \\ Y_1^N & Y_2^N & \dots & Y_D^N \end{bmatrix} \tag{12}$$

Initialize the memory of every crow. Here, the crows conceal their foods in their initial positions, because at the initial iteration they do not have any experience.

$$Memory = \begin{bmatrix} P_1^1 & P_2^1 & \dots & P_D^1 \\ P_1^2 & P_2^2 & \dots & P_D^2 \\ \vdots & & & \\ P_1^N & P_2^N & \dots & P_D^N \end{bmatrix} \tag{13}$$

Step 3: Fitness Calculation:

Apply the weight coefficients given in the corresponding solution to the BCF function to find the k – retrieved cases from the training data. Then the retrieved case is used to find the f-measure which is the fitness.

Step 4: Generation of new position

If one crow needs to generate its new position, then it follows any one of the crows and determines its new position. Consider the crow C_1 needs to find its new position, then it follows the crow C_2 and finds its new position. The new position of the crow C_1 is calculated by the equation (11). This process is repeated for all the crows.

Step 5: Determine the feasibility of the new position

Here, the feasibility of the new position of the crow is checked. If the new position is feasible, then it is updated otherwise the crow stay in the original place.

Step 6: Fitness Calculation of the new positions

Apply the weight coefficients given in the corresponding solution to the BCF function to find the k –retrieved cases from the training data. Then the retrieved case is used to find the f-measure which is the fitness.

Step 7: Memory updating:

The memory of the crow is updated as follows,

$$P^{C_1,I+1} = \begin{cases} Y^{C_1,I+1} & ;\ f\left(Y^{C_1,I+1}\right) is\, better\, than\, f\left(P^{C_1,I}\right) \\ P^{C_1,I} & ;\ otherwise \end{cases} \tag{14}$$

where, $f(\bullet)$ is the objective function.

Step 8: Termination:

The above steps are continued until the maximum iteration $I_{\max}$ reaches.

5. Results and Discussion

This section presents the experimental results of the proposed case retrieval method and the comparative analysis of the proposed method with the existing case retrieval methods, such as PESM [15], FBSO-neural network [15], and Hybrid model [15] for two different data sets, namely breast cancer and breast cancer wins.

5.1 Experimental Setup

Platform: The proposed case retrieval technique is experimented in a personal computer with 2GB RAM and 32-bit OS and implemented using MATLAB 8.2.0.701 (R2013b).

Datasets used: The experimentation of the proposed method is performed in two datasets such as, breast cancer and breast cancer wins which are taken from the UCI machine learning repository [16].

Evaluation metrics: Evaluation metrics considered for analyzing the performance of the proposed ***method*** are Fall-out and Discounted Cumulative Gain (DCG).

Fall-out: It is defined as the proportion of retrieved non-relevant medical cases and the total number non-relevant medical cases.

$$fall-out = \frac{\left|\{non-relevent\, medical cases\} \cap \{retrieved\, medical cases\}\right|}{\left|\{non-relevent\, medical cases\}\right|} \tag{15}$$

DCG: It is defined as a measure of ranking quality. In information retrieval, it is utilized to quantify the efficacy of web search engine algorithms or related applications.

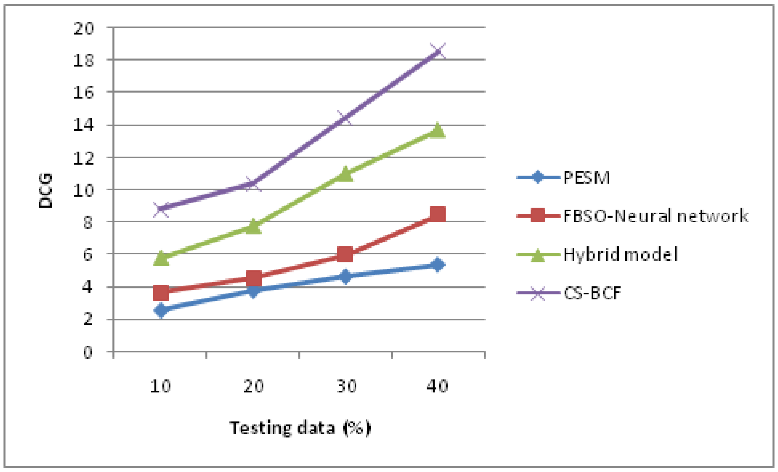
5.2 Performance Analysis

Here, the performance of the proposed method is analyzed with the existing methods, such as PESM, FBSO-neural network, and Hybrid model for the performance measures DCG and Fall-out. Here, the performance is analyzed for two different data sets.

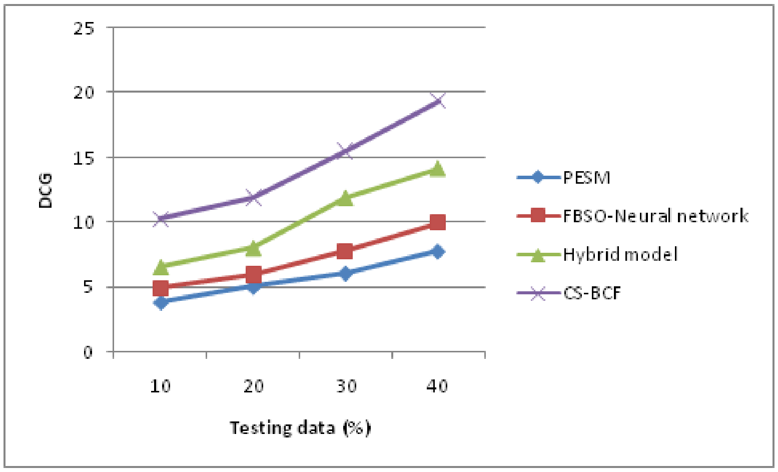
a) DCG

Figure 3 shows the DCG curve of the proposed CS-BCF case retrieval method and the existing methods, such as PESM, FBSO-neural network, and Hybrid model for different size of testing data. Figure 3 (a) shows the DCG curve of the proposed case retrieval method and the existing methods for breast cancer data set. When the size of the testing data is 10%, DCG of the proposed method is 8.785 while the DCG of the existing methods, such as PESM, FBSO-neural network, and Hybrid model is 2.578, 3.654, and 5.789 respectively. When the size of the testing data is 20%, the DCG of the existing methods, such as PESM, FBSO-neural network, and Hybrid model is 3.786, 4.546, and 7.754 respectively, on the other hand, the DCG of the proposed method CS-BCF is 10.348. Similarly, when the size of the testing data is 30% and 40%, the proposed method has the maximum DCG than the existing methods, such as PESM, FBSO-neural network, and Hybrid model.

Figure 3 (b) shows the DCG curve of the proposed case retrieval method and the existing methods for breast cancer wins data set. When the size of the testing data is 20%, DCG of the proposed method is 11.927 while the DCG of the existing methods, such as PESM, FBSO-neural network, and Hybrid model is 4.984, 5.947, and 8.026 respectively. When the size of the testing data is 30%, the DCG of the existing methods, such as PESM, FBSO-neural network, and Hybrid model is 6.045, 7.762, and 11.873 respectively, on the other hand, the DCG of the proposed method CS-BCF is 15.467. Similarly, when the size of the testing data is 10% and 40%, the proposed method has the maximum DCG than the existing methods, such as PESM, FBSO-neural network, and Hybrid model. From figure 3, it can be shown that the proposed method has the higher DCG than the existing methods.



a) Breast Cancer



b) Breast Cancer Wins

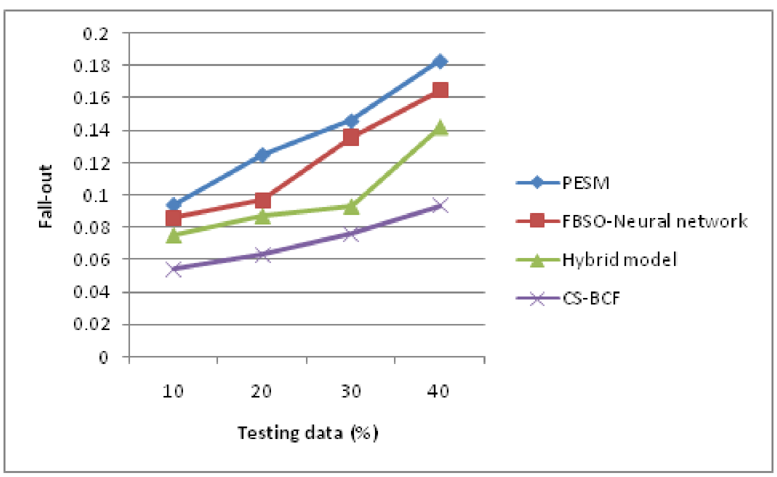
Figure 3: Illustration of DCG curve of the proposed CS-BCF and the existing methods, such as PESM, FBSO-neural network, and Hybrid model

b) Fall-out

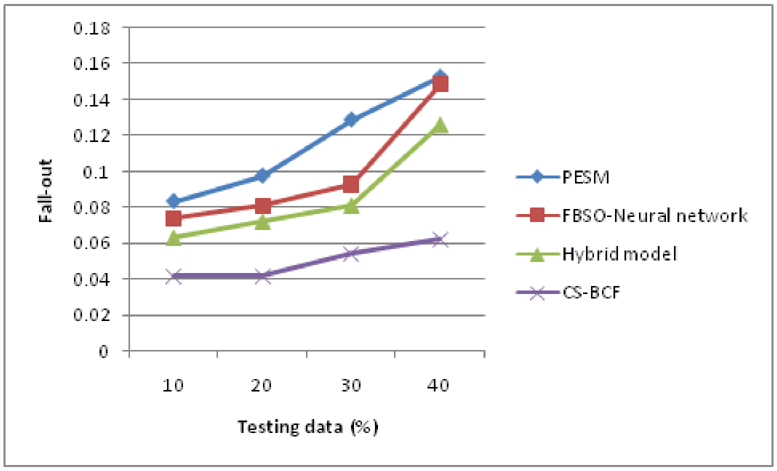
Figure 4 shows the fall-out of the proposed CS-BCF case retrieval method and the existing methods, such as PESM, FBSO-neural network, and Hybrid model for different size of testing data. Figure 4 (a) shows the fall-out of the proposed case retrieval method and the existing methods for breast cancer data set. When the size of the testing data is 10%, the fall-out of the proposed method is 0.054 while the fall-out of the existing methods, such as PESM, FBSO-neural network, and Hybrid model is 0.0941, 0.086, and 0.075 respectively. The proposed method has the fall-out of 0.063, on the other hand, the existing methods, such as PESM, FBSO-neural network, and Hybrid model have the fall-out of 0.125, 0.0965, and

0.087 respectively for 20% of the testing data. For 30% of the testing data, the existing methods, such as PESM, FBSO-neural network, and Hybrid model have the fall-out of 0.146, 0.136, and 0.093 respectively, on the other hand, the proposed method has the fall-out of 0.076. Similarly, for 40% of the testing data, the proposed method has the minimum fall-out than the existing methods.

Figure 4 (b) shows the fall-out of the proposed case retrieval method and the existing methods for breast cancer data set. When the size of the testing data is 10%, the fall-out of the proposed method is 0.042 while the fall-out of the existing methods, such as PESM, FBSO-neural network, and Hybrid model is 0.0832, 0.074, and 0.063 respectively. The proposed method has the fall-out of 0.042, on the other hand, the existing methods, such as PESM, FBSO-neural network, and Hybrid model have the fall-out of 0.0975, 0.081, and 0.072 respectively for 20% of the testing data. For 30% of the testing data, the existing methods, such as PESM, FBSO-neural network, and Hybrid model have the fall-out of 0.129, 0.093, and 0.081 respectively, on the other hand, the proposed method has the fall-out of 0.054. Similarly, for 40% of the testing data, the proposed method has the minimum fall-out than the existing methods. From figure 4, it can be shown that the proposed CS-BCF method has the minimum fall-out than the existing methods, such as PESM, FBSO-neural network, and Hybrid model.



a) Breast Cancer



b) Breast Cancer Wins

Figure 4. Illustration of Fall-out of the proposed CS-BCF and the existing methods, such as PESM, FBSO-neural network, and Hybrid model

6. Conclusion

This paper presents the method called CS-BCF for retrieving the similar cases from the case database. Initially, the patient's case database is constructed with details of different patients and their treatment details. If the new patient comes for treatment, then the doctor collects the information about that patient and sends the query to the CS-BCF. The CS-BCF system matches the input query with the patient's case database and retrieves the similar cases. Here, the CS algorithm is used with the BCF for retrieving the most similar cases from the patient's case database. Finally, the Doctor gives treatment to the new patient based on the retrieved cases. The performance of the proposed method is analyzed with the existing methods, such as PESM, FBSO-neural network, and Hybrid model for the performance measures DCG and fall-out. The experimental results show that the proposed method attains the higher DCG of 19.324 and the minimum fall-out 0.042when compared to the existing methods.

References

[1] Sungbin Choi, Jinwook Choi, Sooyoung Yoo, Heechun Kim, and Youngho Lee, "Semantic concept-enriched dependence model for medical information retrieval," Journal of Biomedical Informatics, vol. 47, pp. 18-27, February 2014.

[2] Yanshan Wang, Stephen Wu, Dingcheng Li, Saeed Mehrabi, and Hongfang Liu, "A Part-Of-Speech Term Weighting Scheme for Biomedical Information Retrieval," Journal of Biomedical Informatics, vol. 63, pp. 379-389, October 2016.

[3] Yong-Bin Kang, Shonali Krishnaswamy, and Arkady Zaslavsky, "A Retrieval Strategy for Case-Based Reasoning Using Similarity and Association Knowledge," IEEE Transactions on Cybernetics, vol. 44, no. 4, pp. 473-487, April 2014.

[4] Adrien Depeursinge, Samuel Duc, Ivan Eggel, and Henning Muller, "Mobile Medical Visual Information Retrieval," IEEE Transactions on Information Technology in Biomedicine, vol. 16, no. 1, pp. 53-61, January 2012.

[5] André Mourão, Flávio Martins, and João Magalhães, "Multimodal medical information retrieval with unsupervisedrank fusion," Computerized Medical Imaging and Graphics, vol. 39, pp. 35–45, January 2015.

[6] Shihchieh Chou, Weiping Chang, Chin-Yi Cheng, Jihn-Chang Jehng and Chenchao Chang, "An Information Retrieval System for Medical Records & Documents," In Proceedings of the IEEE conference on Engineering in Medicine and Biology Society, 2008.

[7] Mario Taschwer, "Textual Methods for Medical Case Retrieval," Technical Report No. TR/ITEC/14/2.01, Institute of Information Technology (ITEC), Alpen-Adria-University at Klagenfurt, Austria, May 2014.

[8] G. Adomavicius, A. Tuzhilin, "Toward the next generation of recommender systems: a survey of the state-of-the-art and possible extensions," IEEE Transaction on Knowledge and Data Engineering, vol. 17, no. 6, pp.734–749, 2005.

[9] J. Bobadilla, F. Ortega, A. Hernando, A. Gutirrez, "Recommender systems survey," Knowledge-Based Systems, vol. 46, pp. 109–132, 2013.

[10] X. Su, T.M. Khoshgoftaar, "A survey of collaborative filtering techniques," Journal of Advances in Artificial Intelligence, vol. 2009, 2009.

[11] T. Kailath, "The divergence and Bhattacharyya distance measures in signal selection," IEEE Transaction on Communication Technology, vol. 15, no. 1, pp. 52–60, 1967.

[12] F. Nielsen, S. Boltz, "The Burbea-Rao and Bhattacharyya centroids," IEEE Transaction on Information Theory, vol. 57, no. 8, pp. 5455–5466, 2011.

[13] Bidyut Kr. Patra, Raimo Launonen, Ville Ollikainen, Sukumar Nandi, "A new similarity measure using Bhattacharyya coefficient for collaborative filtering in sparse data," Journal of Knowledge-Based Systems archive, vol. 82, no. C, pp. 163-177, July 2015.

[14] Alireza Askarzadeh, "A novel metaheuristic method for solving constrained engineering optimization problems: Crow search algorithm," Computers & Structures, vol. 169, pp. 1-12, June 2016.

[15] Poonam Yadav, "Case Retrieval Algorithm Using Similarity Measure and Adaptive Fractional Brain Storm Optimization for Health Informaticians," Arabian Journal for Science & Engineering, vol. 41, no. 3, pp. 829-840, March 2016.

[16] Datasets from UCI Machine Learning Repository, http://archive.ics.uci.edu/ml/.

[17] A. B. Spanier, D. Cohen, and L. Joskowicz, "A new method for the automatic retrieval of medical cases based on the RadLex ontology," International Journal of Computer Assisted Radiology and Surgery, vol. 12, no. 3, pp. 471-484, November 2016.

[18] Shahina Begum, Mobyen Uddin Ahmed, Peter Funk, Ning Xiong, and Mia Folke, "Case-Based Reasoning Systems in the Health Sciences: A Survey of Recent Trends and Developments, " IEEE Transactions on Systems, Man, and Cybernetics—Part C: Applications and Reviews, vol. 41, no. 4, pp. 421-434, July 2011.

[19] H. Yan, Y. Jiang, J. Zheng, C. Peng, and Q. Li, “A multilayer perceptron-based medical decision support system for heart disease diagnosis,” Expert Systems with Applications, vol. 30, pp. 272-281,2006.

[20] H. Moen, F. Ginter, E. Marsi, L.-M. Peltonen, T. Salakoski, and S. Salantera, “Care episode retrieval: distributional semantic models for information retrieval in the clinical domain,” BMC medical informatics and decision making, vol.15, no. 2, 2015.

[21] D. Blumenthal and M. Tavenner, “The meaningful use regulation for electronic health records,” New England Journal of Medicine, vol.363, no. 6, pp. 501–504, 2010.

[22] W. R. Hersh, “Information Retrieval: A health and biomedical perspective,” 3rd Edition, Springer, 2003.

[23] K. Kawamoto, C. A. Houlihan, E. A. Balas, and D. F. Lobach, “Improving clinical practice using clinical decision support systems: a systematic review of trials to identify features critical to success,” British Medical Journal, vol.330, no. 7494, pp.765, 2005.

[24] A. Aamodt and E. Plaza, “Case-based reasoning: Foundational issues,” methodological variations, and system approaches. AI Communications, vol. 7, no. 1, pp.39-59, 1994.

[25] R. Lopez De Mantaras, D. McSherry, D. Bridge, D. Leake, B. Smyth, S. Craw, B. Faltings, M. L. Maher, M. T. Cox, K. Forbus, M. Keane, A. Aamodt, and I. Watson, “Retrieval, reuse, revision and retention in case-based reasoning,” Knowledge Engineering, vol. 20, no. 3, pp. 215–240,2005.

[26] Lee W-N, Shah N, Sundlass K, and Musen M, “Comparison of ontology-based semantic-similarity measures,” In Proceedings of the Annual Symposium on AMIA, pp.384–388, 2008.

[27] K. S. S. Rao Yarrapragada and B. Bala Krishna, "Impact of tamanu oil-diesel blend on combustion, performance and emissions of diesel engine and its prediction methodology", Journal of the Brazilian Society of Mechanical Sciences and Engineering, pp. 1-15, 2016.

[28] Rao Yerrapragada. K. S.S, S.N.Ch. Dattu .V, Dr. B. Balakrishna,"Survey of Uniformity of Pressure Profile in Wind Tunnel by Using Hot Wire Annometer Systems", International Journal of Engineering Research and Applications, vol.4 (3), pp. 290-299, 2014.

[29] Kavita Bhatnagar and S. C. Gupta, "Investigating and Modeling the Effect of Laser Intensity and Nonlinear Regime of the Fiber on the Optical Link", Journal of Optical Communications, 2016.

[30] K Bhatnagar and SC Gupta, "Extending the Neural Model to Study the Impact of Effective Area of Optical Fiber on Laser Intensity", International Journal of Intelligent Engineering and Systems, vol.10, 2017.

[31] B.S. Sunil Kumar, A.S. Manjunath, S. Christopher, "Improved entropy encoding for high efficient video coding standard", Alexandria Engineering Journal, In press, corrected proof, November 2016.

[32] BSS Kumar, AS Manjunath, S Christopher," Improvisation in HEVC Performance by Weighted Entropy Encoding Technique" Data Engineering and Intelligent Computing, 2018.

[33] S Chander, P Vijaya, P Dhyani,"Fractional lion algorithm–an optimization algorithm for data clustering",Journal of computer science, 2016.

[34] P. Vijaya, Satish Chander,"Fuzzy Integrated Extended Nearest Neighbour Classification Algorithm for Web Page Retrieval", Proceeding of the International Conclave on Innovations in Engineering and Management, 2016.

[35] P Yadav and RP Singh, "An Ontology-Based Intelligent Information Retrieval Method for Document Retrieval", International Journal of Engineering Science, 2012.