International Journal of Pure and Applied Mathematics
Volume 118 No. 18 2018, 1775-1781
ISSN: 1311-8080 (printed version); ISSN: 1314-3395 (on-line version)
url: http://www.ijpam.eu
Special Issue

ijpam.eu

Optimal Detection of Lung cancer using Bees Algorithm

R.Abirami[1],L.Andrea[2],M.Diviya[3]
Assistant Professor[1,2,3], Rajalakshmi Engineering College[1,2,3],Chennai,
.

***Abstract*---In the world history of cancer decease, lung cancer plays a important role. It is significant to diagnose lung cancer at premature stage itself. To enhance detection radiologist use scans and X-ray images out of which CT scans seems to provide better result. So the proposed work deals with classifying data of CT scan images. Once classification is done the classified data is then optimized to improve the results with high accuracy. Several optimization algorithms like Ant Colony Optimization ,Genetic Algorithm, Bees Algorithm, Particle Swarm Optimization, Multi Swarm Optimization are compared and bees algorithm proves to give better accuracy result.**

***Keywords*--- Bees Algorithm , classification , Genetic algorithms, Cancer, optimization,X-ray.**

I.INTRODUCTION

Optimization is performed to get an accurate results. CT scan images are used to identify and to classify cancer at early stage. After classification an optimal algorithm is found for detecting cancer. To achieve this goal following objectives are highlighted : To classify CT scan images various image mining methods like preprocessing, segmentation, feature extraction is done before . Then segmentation of nodules is done using otsu's thresholding method from the CT scan image which is followed by classification which done by using classifiers based on features extracted. The last step is to evaluate the efficiency of our proposed algorithm over other optimization algorithms.

Optimization is the process or an act of finding an another way or an alternative which can be cost effective or which achieves the highest performance for the method that is used. Optimizing the results in medical field will help to predict the cancer or any other disease at early stage through various optimization algorithms. The other applications in which optimization can be used are in computer simulation or for illustrating problems related to business. For performing optimization there are various Optimization algorithms like Genetic Algorithm, Ant Colony Optimization, Bees Algorithm, Particle Swarm Optimization, Multi Swarm Optimization which can be used to get accurate results.

The paper is organized as follows: In section 2 related work is discussed . In Section 3 presents different optimization algorithms are discussed. Section 4 describes about the framework model. In Section 5 we discuss about implementation and results . Section 6 presents conclusion and future work.

II. RELATED WORK

Different methodologies were implemented to identify or predict lung cancer. Data mining methodologies are useful for lung cancer classification. One of the optimization technique Ant Colony Optimization (ACO) helps in increasing or decreasing the disease prediction value through several iterations and finds the the shortest track . As it was time consuming process to remove this difficulty Reduced Order Constrained Optimization (ROCO) was used. Which reduced the number of iterations by minimizing the trial and error method but does not give exact grouping and correctness of dataset[1].

An efficient technique based on association rule classification with optimization technique was carried out by Suman Mishra , Prateek Gupta. It establishes choice optimization methods and separation based on negative and positive supports classified by minimum support. The result shows the effectiveness of the approach used [2].

The study by C. Bhuvaneswari, P.Aruna, D. Loganathan was based on texture based segmentation and identification of diseases based on lungs from the computed tomography images. The texture based features are extracted. Different feature selection techniques which includes Information Gain, Principal Component Analysis are used. Genetic algorithm is combined with these feature selection techniques which helps in optimal initialization of clusters. Clustering techniques like fuzzy C means clustering and segmentation methods like watershed segmentation are combined with outputs of the feature extracted. For classification purpose Naive Bayes classifier is used or applied on the datasets. To identify the exact parameter the parameter selection method is used. The results proved that correlation based feature selection produced best results which is proved by measuring the performance. [3].

A Decision support system which helps doctors to analyse data in medical field. The study was conducted by K. Rajeswari P, Mahadev Shindalkar , Nikhil Thorawade , Pranay Bhandari by using the system to identify the risks related to

disease. DSS is an interactive software system which assists the doctors, stock analyzer to make approximate and faster decisions in the given context. The volume of data in fields like medicine or industry is huge and beyond capacity of human to extract useful information. Apriori algorithm along with genetic algorithm and fuzzy logic were used to obtain correlations and patterns from datasets. It eliminates human error and high accuracy is obtained [4].

For minimizing parameters of breast cancer, lung cancer and heart disease Kaberi Das , Smruti Rekha Das, Pradeepta Kumar Panigrahi, Debahuti Mishra used Principal Component Analysis (PCA) on data sets. Kernal selection methods are compared to get accuracy. Particle Swarm Optimization along with Support Vector Machine helps in achieving better performance. This paper focused on Support Vector Machine trained using linear, polynomial and radial basis function (RBF) kernels and applying Particle Swarm Optimization to each kernels for each data set to get better accuracy [5].

A. Ghanbarzadeh, D.T. Pham, S. Otri , S. Rahim , M. Zaidi, conducted study on searches done by honeybees is between the best sites. They search for the best sites with both speed and accuracy which is one of the important factors to get an optimal solution In addition to it decision or conclusion is taken among the solutions which helps in finding out the solutions of optimization problems. The bees communicate with each other through waggle dance to inform bees to find food resources and location of nest sites. In addition to it certain parameters are taken like distance and direction to find the best results.The efficiency and robustness of the new algorithm is found [6].

In 2005 researches Paola Campadelli, Elena rasiraghi and Giorgio Valentin used image processing techniques and CAD systems for detecting lung nodules. have proved to be effective for the improvement of radiologists diagnosis. A system which automatically detect lung nodules from Radiographs of chest images are demonstrated. The system extract a set of candidate regions by using it onto the radiograph three different and consecutive multi-scale schemes. The comparison of the result obtained show the efficacy of the multi-scale framework.Learning systems using as input different sets of features have been experimented using Support Vector Machines (SVMs) for candidates classification [7].

III.OPTIMIZATION ALGORITHMS

Swarm Based Optimization methods are used to perform searches to find optimal solution. The solutions are obtained by performing several iterations.The results obtained in each iteration is refined to improve the performance. Swarm Based optimization includes ACO, GA and Particle Swarm optimization.

A. Ant Colony Optimization (ACO)

ACO helps in minimizing or maximizing the disease prediction value through several iterations and finds the shortest path. This type of optimization problem is a non-greedy based search algorithm. While ants search for their food source they split out a chemical substance called pheromone. Based upon the amount of pheromone deposited on the ground the possibility that the other ants takes the same track or follows the same track as that of the previous ants is found. It is found by several trail and error methods. ACO optimization is mainly based upon trail and error method.

B. Genetic Algorithm (GA)

Genetic algorithm deals with natural selection and genetic recombination. The current population is found and a the solutions are found based on the current population. Then Genetic operators like mutation and crossover are used or applied on current population to generate new population. Genetic algorithm performs global search which increases the performance by exploiting the new search**.**

C. Particle Swarm Optimization (PSO)

The solution that are obtained from each iterations are taken as particles that change their position with time. Each solution compares the results of itself with the other neighboring solutions. It remembers the previous best visited positions or solutions hence best solution is obtained by comparing the results of the previous iterations. Particle Swarm Optimization is the combination of both local and global search.

D. Bees Algorithm (BA)

Searches done by honeybees is between the best sites which are also called as nodes. They search for the best nodes with both speed and accuracy which is one of the important factors to get an optimal solution . In addition to it decision or conclusion is taken among the solutions which helps in finding out the solutions of optimization problems. The bees communicate with each other through waggle dance to inform bees to find food resources and location of nest sites. In addition to it certain parameters are taken like distance and direction to find the best results. The algorithm is referred from reference paper [6].

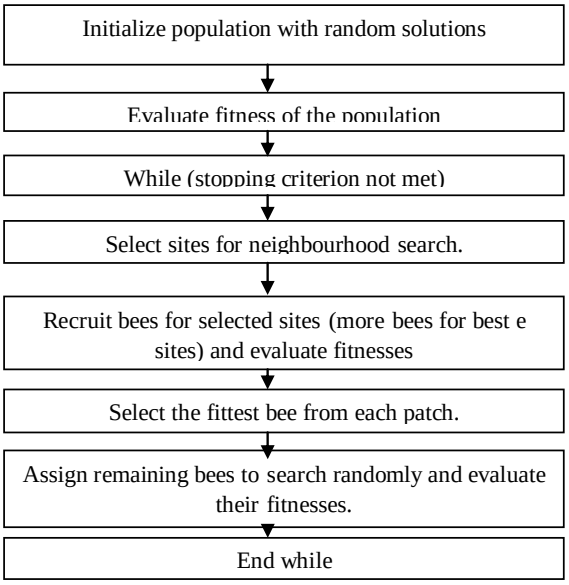

Figure 1: Procedure for algorithm

IV.PROPOSED MODEL

The proposed model combines optimization along with classification. There are two techniques to treat cancer they are conventional method and through Intensity Modulated Radio Therapy (IMRT). In proposed model treatment method IMRT is used along with optimization. Noise removal, Segmentation, Feature extraction and classification are done in the CT scan image of the lung. The training set is given to different optimal algorithms and the best algorithm is found by comparison to find which is faster to detect cancer.

The key advantages of the proposed system includes fast removal of noise and edge preservation of the image due to the use of median filter. Since Otsus method is used segmentation is fast to compute. The correctness and more desirable results are found based upon the data. Satisfactory or effective optimization algorithm that is used which improves the detection of the disease. The proposed framework is as follows which includes several steps for classification and that is followed by optimization to obtain best results.The ct scan image is given for noise removal. The filtered image is segmented using Otsus segmentation technique.The structural and textural features are extracted from the segmented image. Based upon the textural features classification is done using various classifier. The features are given to several optimization algorithms and the best algorithm is found based on comparison.

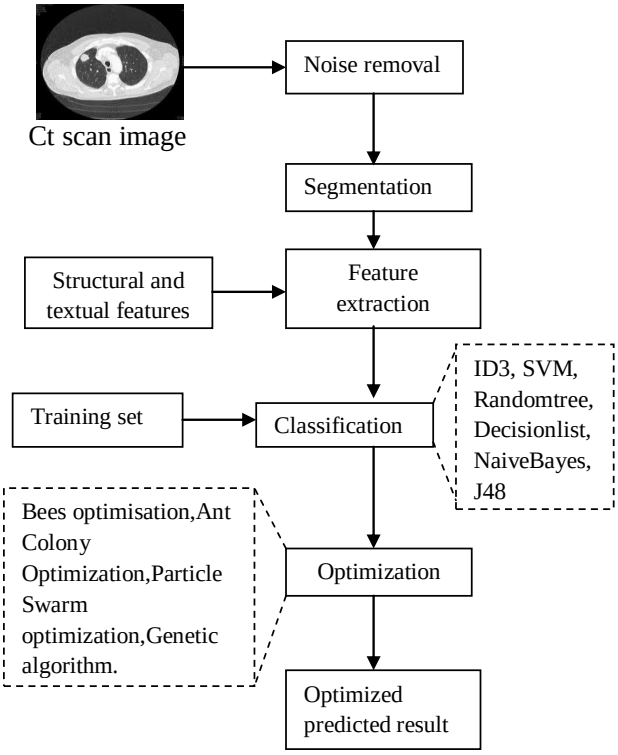

Figure 2: Framework for optimization

V.IMPLEMENTATION

Patient scans the organ. The scanned image is then segmented using certain image mining methods. Once the image is segmented and the nodule is found the features are extracted to find the depth and size of the nodule. Then classification is done based on features extracted. The data are given to several optimization algorithms and the best one if found by comparing the results which makes the detection faster and accurate. Usually IMRT technique is used for treatment of cancer which includes treatment through chemicals , and through radiations. The medical physicist gives dosages based upon the optimal result.

At first the noise has to be removed from the CT scan image of lung. Noise removal is similar to data cleansing where the unwanted informations are removed. To remove noise from the image different filters has to be used .The best way to remove salt and pepper noise is by using median filters. The purpose of using Median filter is that it is a non-linear filter that not only removes noise but also preserves edges in the image. Once the noise is removed a filtered image is obtained.

The next techniques that has to be used is segmentation. The filtered image quality has to be improved and segmented to get the location of the nodule or the affected cell in the lung. Segmentation means dividing the image on basis of certain values. Segmentation is the process of separating lung nodule from the other part of CT scan images. Segmentation is usually done by using thresholding methods. Otsus thresholding method is used. On using this thresholding method the filtered image has to be segmented into two regions using imquantize specifying the threshold levels. Otsus method first converts the grayscale image into binary image and does the segmentation based on the values specified. The purpose of using Otsus thresholding is that it can operate directly on grayscale histogram which is fast to compute. After segmentation the next step that has to be done is feature extraction. Features has to extracted to identify the characteristics of the image. The features include the characteristics of the image which is nothing but the region properties of the nodule that was segmented using thresholding method. The structural features will include area, perimeter, mean, standard deviation of the image. The textural properties will be beams, beamlets, target value, organ at risk are parameters used by radio-physicts.

For performing classification various classifiers are used and the best classifier is obtained by comparing the results of other classifiers. Training set is given to different classifiers and results are compared. The classification finds whether it belongs to lung or prostate case. Then different optimization algorithms like Ant colony optimization(ACO), Bees Algorithm(BA), Genetic Algorithm(GA), Particle Swarm Optimization(PSO) are used. Training set is given to all these optimization algorithms and there performance is to be measured and compared. Bees algorithm proves to be the best optimal algorithm.

The results found after the above implementation steps

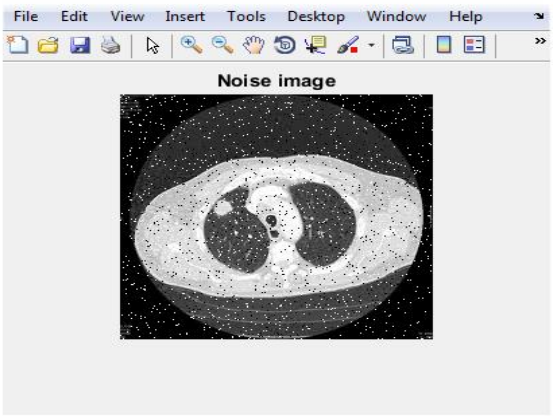

Figure 3: Noise image

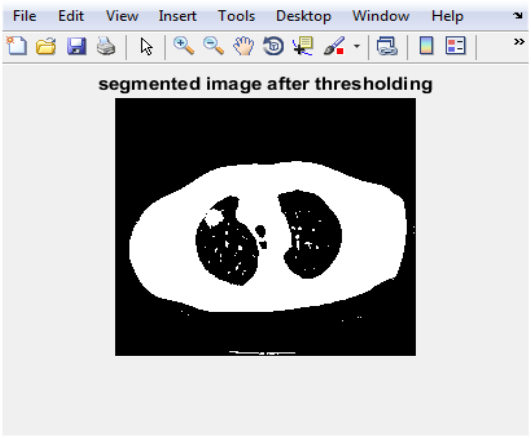

Figure 4: Segmented image

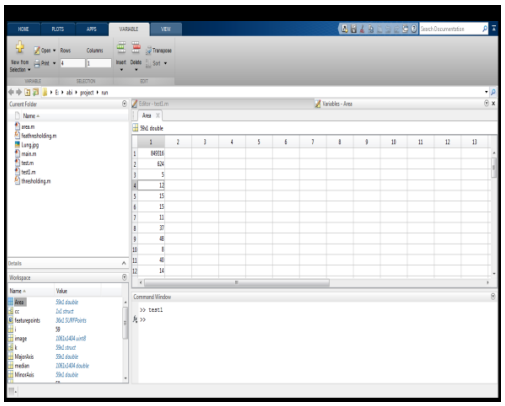

Figure 5: Area of the nodule

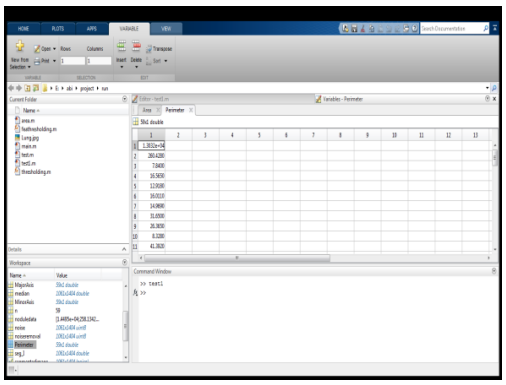

Figure 6: Perimeter of the nodule

VI.TEST RESULT OF CLASSIFIERS

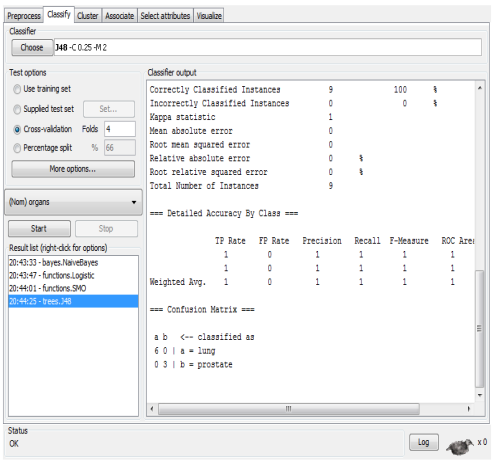

Figure 7: J48 Decision tree

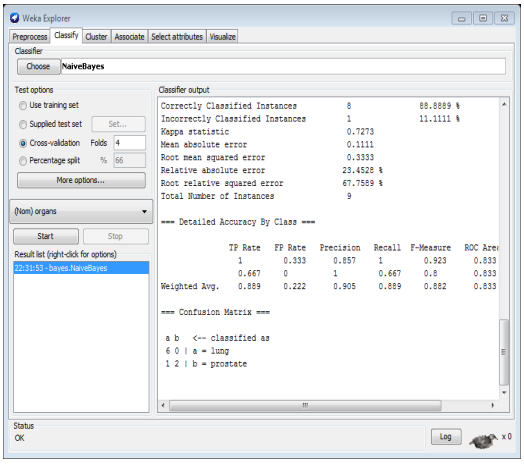

Figure 8: NavieBayes

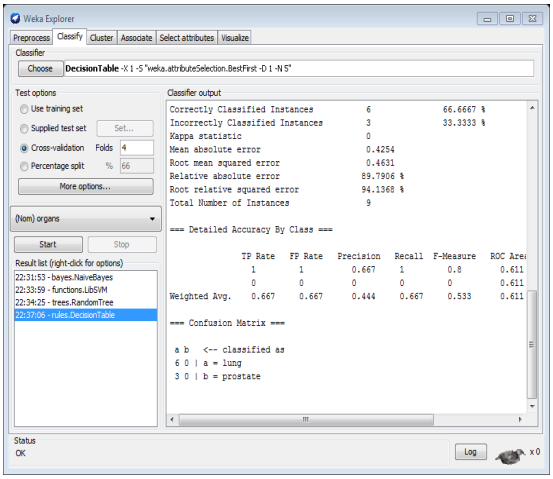

Figure 9: Decision list

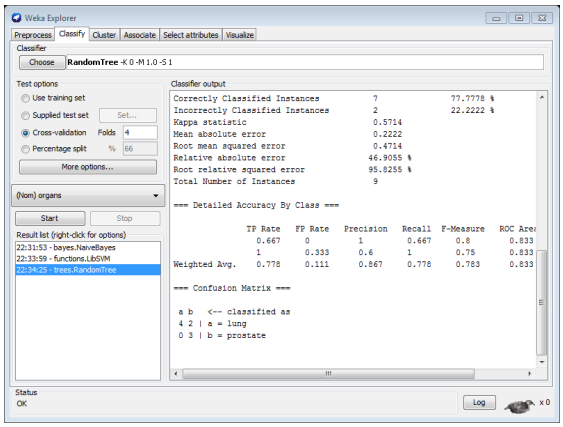

Figure 10: Random tree

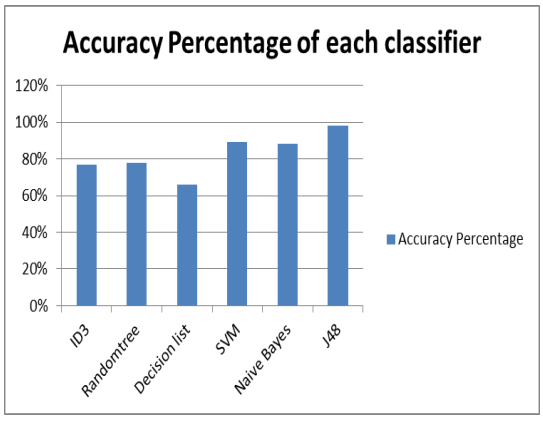

Figure 11: Performance chart

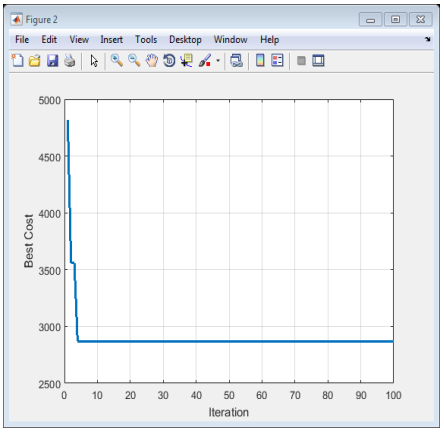

Figure 12: Ant Colony optimization

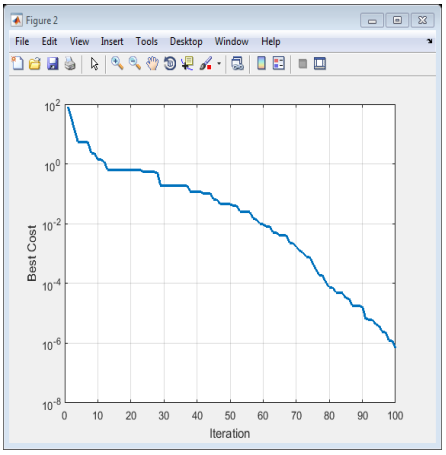

Figure 13: Particle Swarm Optimization

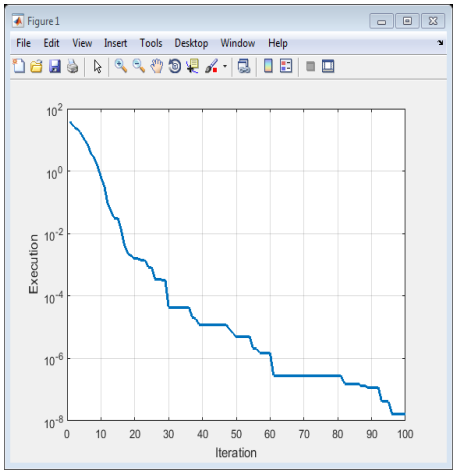

Figure 14: Bees Algorithm

Table I. Execution Result Table

Algorithm	Iterations	Execution Time
Ant Colony Optimization	100	2900
Particle Swarm Optimization	100	10^-6
Bees Algorithm	100	10^-8

VII.CONCLUSION AND FUTURE WORK

Lung cancer has to be predicted at premature level using image mining methods to reduce the number of death rate. In addition to it optimization is added to get accurate result. Since median filter is used fast removal of noise and edge preservation of the image is achieved. Computation of segmentation is done by using otsus method which is better and faster. Classification is done using weka tool to get correct accurate classified result .

Image segmentation is a primary step in image analysis used to separate the input image into meaningful regions. There are many approaches developed for the segmentation of images. To improve their efficiency optimization algorithms can be combined with segmentation approaches in future to accurately detect the tumor in the scanned image.

REFERENCE

[1] T. Sowmiya, M. Gopi, M. New Begin, L.Thomas Robinson," Optimization of Lung Cancer using Modern Data Mining Techniques," International Journal of Engineering Research,Vol .3, Issue.5, pp :309-314,May 2014.

[2]. Suman Mishra,Prateek Gupta," An Association Rrule Data Classification with Optimization,"(IJCSIT) International Journal of Computer Science and Information Technologies, Vol. 5, 2014.

[3] C. Bhuvaneswari P.Aruna,D. Loganathan,"Advanced Segmentation Techniques Using Genetic Algorithm for Recognition of Lung Diseases from CT Scans of Thorax," International Journal of Engineering Research and Applications (IJERA) Vol. 3, Issue 4, pp.2517-2524,Aug 2013.

[4] K. Rajeswari P,Mahadev Shindalkar,Nikhil Thorawade,Pranay Bhandari," DSS Using Apriori Algorithm,Genetic Algorithm And Fuzzy Logic,"International Journal of Engineering Research and Applications (IJERA),Vol. 3,Issue 4,pp.132-136,Aug 2013.

[5] Smruti Rekha Das,Pradeepta Kumar Panigrahi,Kaberi Das,Debahuti Mishra," Improving RBF Kernel Function of Support Vector Machine using Particle Swarm Optimization," International Journal of Advanced Computer Research,Vol.2,Issue-7 Dec 2012.

[6] D.T. Pham, A. Ghanbarzadeh, S. Otri , S. Rahim , M. Zaidi "The Bees Algorithm – A Novel Tool for Complex Optimisation Problems", Manufacturing Engineering Centre, Cardiff University, Cardiff CF24 3AA, UK .

[7] JunzoWatada, Keisuke Aoki, Masahiro Kawano, Muhammad SuzuriHitam, Dual Scaling Approach to Data M Journal of Advanced Computational Intelligence Intelligent Informatics (JACIII), Vol. 10, No. 4, pp. 441-447, 2006.

[8] Jiawei Han and MichelineKamber, "Data Mining Concepts and Techniques." San Francisco, CA: Elsevier Inc, 2006.

[9] Irene M. Mullins et al., "Data mining and clinical data repositories: Insights from a667,000 patient data set," Computers in Biology and Medicine, vol. 36, pp. 1351-1377, 2006.

[10] Sue S. Yom, Zhongxing Liao, H. Helen Liu, Susan L. Tucker, Chao-Su Hu, XiongWei, Xuanming Wang, Shulian Wang, Radhe Mohan, James D. Cox, and Ritsuko Komaki, "Initial evaluation of treatment-related pneumonitis in advanced-stage non-small-cell lung cancer patients treated with concurrent chemotherapy and intensity-modulated radiotherapy," International Journal ofRadiation Oncology*Biology*Physics, 68, 94 – 102 (2007).

[11] R Lu, RRadke, L Happersett, J Yang, Chui, E Yorke, and A Jackson, "Reduced-orderconstrained optimization in IMRTplanning," Physics in Medicine and Biology, 53, 6749–6766 (2008).

[12] American Cancer Society, "Cancer facts & figures 2010," http://www.cancer.org/acs/groups/content/@epidemiologysurveilance/documents/document/ acspc-026238,2010.

[13] C. C. Ling et al., A Practical Guide to Intensity-Modulated Radiation Therapy ,Medical Physics Publishing, 2004.

[14] Paola Campadelli, Elena Casiraghi, Giorgio Valentini," Lung nodules detection and classification,"Oct 2005.

[15] PreetiKhare, Hitesh Gupta, "Finding Frequent Pattern with Transaction and Occurrences based on Density Minimum Support Distribution", International Journal of Advanced Computer Research (IJACR), Volume-2 Number-3 Issue-5 September-2012.

[16] Leena A Deshpande, R.S. Prasad, "Efficient Frequent Pattern Mining Techniques of Semi Structured data: a Survey", International Journal of Advanced Computer Research (IJACR) Volume-3 Number-1 Issue-8 March-2013.

[17] Ashutosh K. Dubey and Shishir K. Shandilya," A Novel J2ME Service for Mining Incremental Patterns in Mobile Computing", Communications in Computer and Information Science, 2010,Springer LNCS.

[18] Ms. KumudbalaSaxena, Dr. C.S. Satsangi, "A NonCandidate SubsetSuperset Dynamic Minimum Support Approach for sequential pattern Mining", International Journal of Advanced Computer Research (IJACR), Volume-2 Number-4 Issue-6 December-2012.

[19] Ashutosh Kumar Dubey, Animesh Kumar Dubey, VipulAgarwal, YogeshverKhandagre, "Knowledge Discovery with a SubsetSuperset Approach for Mining Heterogeneous Data with Dynamic Support",Conseg-2012.

[20] Anshuman Singh Sadh, NitinShukla," Association Rules Optimization: A Survey", International Journal of Advanced Computer Research (IJACR), Volume-3 Number-1 Issue-9 March-2013.