International Journal of Scientific and Technical Advancements
IJSTA
ISSN: 2454-1532

Cancer Classification Using Elitism PSO Based Lezy IBK on Gene Expression Data

Rashmi Nagpal[1], Rashmi Shrivas[2]
[1, 2]Department of Computer Science and Engineering, Mats University Aarang C.G., India
Email address: [1]rashmi.nagpal006@gmail.com

***Abstract*—**DNA micro-arrays now permit scientists to screen thousands of genes simultaneously and determine whether those genes are active, hyperactive or silent in normal or cancerous tissue. Because these new micro-array devices generate bewildering amounts of raw data, new analytical methods must be developed to sort out whether cancer tissues have distinctive signatures of gene expression over normal tissues or other types of cancer tissues.

In this paper, we address the problem of selection of a small subset of genes from broad patterns of gene expression data, recorded on DNA micro-arrays. Using available training examples from cancer and normal patients, we build a classifier suitable for genetic diagnosis, as well as drug discovery. Previous attempts to address this problem select genes with correlation techniques. We propose a new method of gene selection utilizing Elitism Particle Swarm Optimization (EPSO) based on Recursive Feature Reduction (RFR). We demonstrate experimentally that the genes selected by our techniques yield better classification performance and are biologically relevant to cancer.

***Keywords*—**MND (Most Non Dominant); EEG (Electroencephalogram); ranker algorithm; classification.

I. INTRODUCTION

Cancer is one of the dreadful diseases found in most of the living being, which is one of the challenging studies for research in the 20th century. There has been lot of proposals from various researchers on cancer classification and detailed study is still on in the domain of cancer classification [6] .In order to gain deep information for related to the cancer classification for researchers know the very closer look the cancer diseases. Researchers also searching the problem they related to the cancer diseases. Various data mining tools are available for the cancer classification with the help of gene expression data [32]. The important aim is the analysis of the collection of genes they can use in an Expressions forms for the purpose of classification of the cancer diseases. However gene classification is a very challenging because its unique problem and massive no of irrelevant attributes (gene).gene expression data can be found through the micro array data analysis. The data format is a matrix. In which row shows the one particular gene and the column shows the samples. Mainly gene expression classification is a vital role is building the robust classifier and outcomes a class label for each used samples. Accuracy plays the very important role in cancer classification with the help of micro array data analysis. Micro array data analysis is a very efficient step for cancer classification. Micro array analysis is included 3 steps 1.dimensionality reduction 2.feature selection and 3.Cancer Gene Classification. Cancer occurs as a result of mutations, or abnormal changes, in the genes responsible for regulating the growth of cells and keeping them healthy. The genes are in each cell's nucleus, which acts as the "control room" of each cell. Normally, the cells in our bodies replace themselves through an orderly process of cell growth: healthy new cells take over as old ones die out. But over time, mutations can "turn on" certain genes and "turn off" others in a cell. That changed cell gains the ability to keep dividing without control or order, producing more cells just like it and forming a tumor [BreastCancer.ORG]. In this paper studied related to which genes are more closely related to the breast cancer. Breast cancer is an uncontrolled growth of breast cells [BreastCancer.ORG]. Breast cancer refers to a malignant tumor that has developed from cells in the breast [BreastCancer.ORG]. More than 50% of women suffering from breast cancer are in the 25 to 50 years age group. Breast cancer can also develop in men but it is rare condition .Breast cancer is the most common cancer in Indian cities [WEBMD]. Sign of breast cancer is Changes in the size or shape of a breast, Dimpling or thickening of some of the skin on a part of a breast, The nipple becoming inverted, Rarely Blood discharge occurring from a nipple, rash around the nipple, which can look similar to a small patch of eczema, Pain in a breast, A lump in the breast [patient.co.uk]. There are various stages for diagnose the breast cancer [breastcancer.org]

Stage 0:-Cancer cells remain inside the breast duct, without invasion into normal adjacent breast tissue.

Stage IA:-The tumor measures up to 2 cm AND the cancer has not spread outside the breast; no lymph nodes are involved.

Stage IB:-There is a tumor in the breast that is no larger than 2 centimeters, and there are small groups of cancer cells – larger than 0.2 millimeter but not larger than 2 millimeters – in the lymph nodes.

Stage IIA:-cancer cells are found in the axillary lymph nodes (the lymph nodes under the arm).

Stage IIB:-The tumor is larger than 2 but no larger than 5 centimeters and has spread to the axillary lymph nodes.

Stage IIIA:-Cancer is found in axillary lymph nodes that are sticking together or to other structures, or cancer may be found in lymph nodes near the breastbone.

Stage IIIB:-The tumor may be any size and has spread to the chest wall and/or skin of the breast.

Rashmi Nagpal and Rashmi Shrivas, "Cancer classification using elitism PSO based lezy IBK on gene expression data," *International Journal of Scientific and Technical Advancements*, Volume 1, Issue 4, pp. 19-23, 2015.

Stage IIIC:-The cancer has spread to lymph nodes either above or below the collarbone.
Stage IV: -The cancer has spread or metastasized to other parts of the body.

Our proposed method is used to improve the classification accuracy of breast cancer. Elitism particle swarm optimization (EPSO) is a population based stochastic optimization technique developed by Dr. Eberhart and Dr. Kennedy in 1995, inspired by social behavior of bird flocking or fish schooling. EPSO is artificial intelligence technique that can be used to find optimum solution to extremely critical and impossible to maximization and minimization the numeric problems. EPSO is a very simple searching method and easily implemented and few parameter adjust mainly the velocity.

II. Data Sets Descriptions

In this paper we have applied our proposed algorithm to microarray datasets. The proposed article uses UCI () dataset and it is a one of the machine learning repository center. The dataset having nine features and one class attribute and the description of data set is given below. The nature of dataset is that it is multivariate and attribute characteristics is integer and number of instances are 106.

Car	**Carcinoma**	**21**
Fad	Fibro-adenoma	15
Mas	Mastopathy	18
Gla	Glandular	16
Con	Connective	14
Adi	Adipose	22
Total		106

9 features:

I0	**Impedivity (ohm) at zero frequency**
PA500	phase angle at 500 KHz
HFS	high-frequency slope of phase angle
DA	mpedance distance between spectral ends
AREA	area under spectrum
A/DA	area normalized by DA
MAX IP	maximum of the spectrum
DR	distance between I0 and real part of the maximum frequency point
p	length of the spectral curve

References: - J. Jossinet (1996) Variability of impedivity in normal and pathological breast tissue. Med. & Biol. Eng. & Comput, 34: 346-350.

JE Silva, JP Marques de Sá, J Jossinet (2000) Classification of Breast Tissue by Electrical Impedance Spectroscopy. Med & Bio Eng & Computing, 38:26-30.

III. Extreme Value Removal

The extreme value removal is a part of data cleaning step for data mining. The procedure for applying the extreme value theorem is to first establish that the function is continuous on the closed interval [14]. The next step is to determine the critical points in the given interval and evaluate the function at these critical points and at the end points of the interval. If the function f(x) is continuous on closed interval [a, b] then f(x) has both a maximum and a minimum on [a, b] [15]. In proposed method inter-quartile range [IQR] is used for extreme value calculations. IQR is major of variability based on dividing the dataset into

Quartiles [16] proposed article found two instances after removal of this breast cancer data set contain 104 instances for classification task [36].

IV. Feature Selection

Feature selection is a task of machine learning tool. It selecting feature in data sets that are most useful or most relevant for the problem. Feature selection method can be used to identify and discard unneeded, discursive and redundant attributes from data sets. Feature selection to improve accuracy and decrease training time.

V. Problem Definition

The cancer classification is the very critical and challenging job. Many researchers has already practice in this field so they planed many algorithms related to data mining such that decision tree methods, the linear discrimination analysis, the RBF network, Genetic algorithm etc. Most proposed cancer classification methods are related to the data mining or soft computing area. Like nearest neighbor analysis, Back propagation network analysis, Fuzzy logic analysis. Mostly of the methods are work fine on binary-class problems and not provide well result in multi-class problems. Most researchers only concerned with the accuracy of the classification. One another problem is gene classifiers proposed are quite computationally expensive they cannot afford to the all people. Exact classification of cancers based on microarray gene expressions is very crucial for doctor to select a proper treatment. Gene expression data, obtained by DNA micro arrays, has been used to investigate the biological terms of tumors and to concatenate expression patterns with clinical results for patients in various stages and different types of diseases.

VI. Proposed Methodology for EPSO

EPSO (Elitism Particle Swarm Optimization) is one of the Features selection or Attribute selection methods. Feature selection is important techniques for identifying informative genes in microarray datasets [5]. That means selection of small informative genes form huge data sets. In this reason multiple evolutionary methods are used for comparing the small samples data with large number of genes. Stack of the computational methods face difficulties to select the small subset [5]. This paper proposes a new evolutionary machine learning approach EPSO. EPSO is a population based approach. EPSO forever provides better solution for the problem. EPSO is the progressist method. This emulates the social behavior of living organisms like: bird flocking and fish schooling. Here candidate solution moves hereabouts a D-dimensional search space. In this paper use the EPSO for breast cancer data sets and experimental analysis for better

Rashmi Nagpal and Rashmi Shrivas, "Cancer classification using elitism PSO based lezy IBK on gene expression data," *International Journal of Scientific and Technical Advancements*, Volume 1, Issue 4, pp. 19-23, 2015.

accuracy outcomes from various classification and search methods. In EPSO each particle has a position and moves based on an updated velocity [base]. Each particle in a population has a fitness value computed from a fitness function. The main features of a particle in basic EPSO are position, velocity, and ability to exchange information with its neighbors, ability to memorize a previous position, and ability to use information to make a decision. Given the easy implementation of EPSO (through a few parameters adjustments), it has become a popular optimization algorithm that has been widely used in many fields to solve various problems, including gene selection [base]. Gene selection, a particle represents a potential solution that is gene subset in an n-dimensional space.

Basic algorithm of EPSO:

Step1: EPSO is initializing with group of random particles (solutions) by [“Kennedy, J. and Eberhart”].

Step2: Search by optimum updating particle.(in every iteration, each particle is updated by following two "best" values. First one is the best solution (fitness).this value is stored and it is called pbest. Another best value selected by the population of the particles with particle swarm optimizer .This value is called global best value and gbest. When a particle takes part of the population as its topological neighbors, the best value is a local best and is called lbest [33].

Step3: After find this two value the particle updates its velocity and positions (each particle Calculate fitness value .If the fitness value is better than the best fitness value (pBest) in history set current value as the new pBest).

Step4: Choose the particle with the best fitness value of all the particles as the gBest for each particle Calculate particle velocity and Update particle position. For this calculation is performed with this equation invented by "Kennedy, J. and Eberhart".

v[] = v[] + c1 * rand() * (pbest[] - present[]) + c2 * rand() * (gbest[] - present[]) (a) present[] = persent[] + v[] (b)

v [] is the particle velocity.

present [] is the current particle (solution).

rand () is a random number between (0,1).

c1, c2 are learning factors usually c1=c2=2.Flow Chart of EPSO.

CFS Feature Set Evaluation: - In this paper using the combination of EPSO and CFS feature set evaluation for feature selection. In machine learning and statistics, feature selection, also known as variable selection, attribute selection or variable subset selection, is the process of selecting a subset of relevant features. Feature selection technique is that the data contains many redundant or irrelevant features. Redundant features are those which provide no more information than the currently selected features, and irrelevant features provide no useful information. CFS is a correlation-based filter method CFS from [Hal98]. It gives high scores to subsets that include features that are highly correlated to the class attribute but have low correlation to each other Let S be an attribute subset that has k attributes, rcf models the correlation of the attributes to the class attribute, rff the intercorrelation between attributes. meritS = k rcf / sqrt(k+k(k-1) rff)

Cfs is a very fast filter method provided by weka they can evaluates sets.

VII. Methodology

The proposed methodology is useful for finding the non-redundant and noisy features from feature set. If huge number of features is used for classification of breast cancer then classification is a very critical task because there classification time and experimental cost is very high. But with the help of EPSO grow the classification accuracy and save the high experimental cost. EPSO is always population based stochastic optimization technique. EPSO responsible for classifications accuracy. Proposed work flow Steps given below and describes how to get accuracy using EPSO with attribute evaluation.

Step1: Take input as breast cancer data sets it is having 10 attributes. 9 features+1class attribute 106 instances.

Step2: Create training, testing and validation set.

Step3: Create feature selection set using different feature PSO search with classification technique.

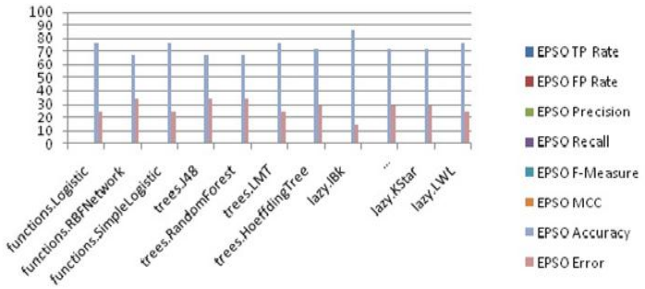
3.1 PSO + Function Logistic
3.2 PSO + Function RBF Network
3.3 PSO + Function Simple Logistic
3.4 PSO + Tree j48
3.5 PSO + Tress Random Forest
3.6 PSO + Tree LMT
3.7 PSO + Tree Hoeffding Tree
3.8 PSO + Lazy IBK
3.9 PSO + bayes.NaiveBayesUpdateable
3.10 PSO + lazy.KStar
3.11 PSO + lazy.LWL

VIII. Result and Analysis

Experiment is performed with java and machine learning tool (Weka), for multivariate data set, 46 classifier accuracy result is shown in figure [3], for measuring the performance of classifier various matrices are present and these are mapped in table [4]. In this only top 11classifier.

Classfication	TP Rate	FP Rate	Precision	Recall	F-Measure	MCC	Accuracy	Error
functions.Logistic	0.762	0.044	0.816	0.762	0.761	0.728	76.1905	23.8095
functions.RBFNetwork	0.667	0.083	0.611	0.667	0.599	0.561	66.6667	33.3333
functions.SimpleLogistic	0.762	0.049	0.825	0.762	0.754	0.726	76.1905	23.8095
trees.J48	0.667	0.087	0.492	0.667	0.559	0.524	66.6667	33.3333
trees.RandomForest	0.667	0.049	0.705	0.667	0.625	0.62	66.6667	33.3333
trees.LMT	0.762	0.049	0.825	0.762	0.754	0.726	76.1905	23.8095
trees.HoeffdingTree	0.714	0.092	0.794	0.714	0.684	0.846	71.4286	28.5714
lazy.IBk	0.857	0.029	0.912	0.857	0.863	0.848	85.7143	14.2857
bayes.NaiveBayesUpdateable	0.714	0.092	0.794	0.714	0.684	0.846	71.4286	28.5714
lazy.KStar	0.714	0.092	0.794	0.714	0.684	0.846	71.4286	28.5714
lazy.LWL	0.762	0.049	0.825	0.762	0.754	0.726	76.1905	23.8095

Rashmi Nagpal and Rashmi Shrivas, “Cancer classification using elitism PSO based lezy IBK on gene expression data,” *International Journal of Scientific and Technical Advancements*, Volume 1, Issue 4, pp. 19-23, 2015.



IX. EXPERIMENTAL ENVIRONMENT

In this paper select the one popular tool of data mining is WEKA 6.7.11. Weka is a collection of Machine learning algorithm for data analysis tasks. This algorithm is directly applied on the data's. Weka is capable for processes the big amount of data. Weka contain tool for data preprocessing attribute selection, classification regression, clustering, visualization. Weka software written in Java Code. Weka provide user graphical interface, easy access to Functionality and platform portability. Weka is a comprehensive collection of data preprocessing and modeling techniques [35]. Weka do the multiple data mining steps [35]:-

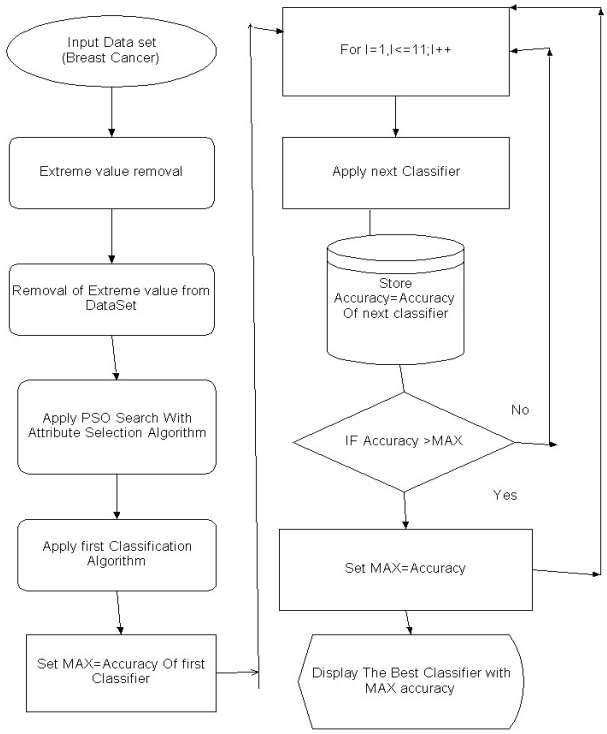
Processes	Description
Data cleaning	The removal of noise and inconsistent data.
Data integration	The combination of multiple sources of data.
Data selection	The data relevant for analysis is retrieved from the database.
Data transformation	The consolidation and transformation of data into forms appropriate for mining.
Data mining	The use of intelligent methods to extract patterns from data.
Pattern evaluation	Identification of patterns that is interesting.
Knowledge presentation	Visualization and knowledge representation techniques are used to present the extracted or mined knowledge to the end user

X. CONCLUSION

In this study we have seen that our proposed approach works well then other existing approaches. The selection of genes that is really informative of the tumor (cancer) classification concentrate in an important role play in successful gene expression data experiment. In this paper apply only EPSO LAZY IBK classifier in our proposed method other classifiers can also be applied for comparative study. Through EPSO obtains higher classification accuracy in Breast cancer data sets as well as less number of inductive features to identify deceases. This has better performance than other existing evolutionary method.

The outcomes show that the EPSO is helpful tool for gene selection and mining the high dimensional data.

ACKNOWLEDGMENT



REFERENCES

[1] T. Latkowski and S. Osowski, "Data mining for feature selection in gene expression autism data," *Expert Systems with Applications*, vol. 42, issue 2, pp. 864-872, 2015.

[2] F. Kurshumliu, L. Gashi-Luci, S. Kadare, M. Alimehmeti, and U. Gozalan, "Classification of patients with breast cancer according to nottingham prognostic index highlights significant differences in immunohistochemical marker expression," *World Journal of Surgical Oncology*, 12:243, pp. 1-5, 2014.

[3] J. Zhaohua, L. Fu, W. Yao, S. Xiaohu, X. Chong, and L. Yanchun, "Differential gene expression analysis on microarray data of breast cancer based on subgroup statistic methods," *International Conference on. IEEE Biomedical Engineering and Biotechnology (ICBEB)*, pp. 167-171, 2012.

[4] Q. Shen, W. M. Shi, and W. Kong, "Hybrid particle swarm optimization and tabu search approach for selecting genes for tumor classification using gene expression data," *Computational Biology and Chemistry*, vol. 32, issue 1, pp. 53-60, 2008.

[5] N. B. Hannan, Md. A. Mottalib, S. J. Kabeer, and A. M. Sultan, "MFS-PSO: A modified PSO method for optimizing gene selection," *International Journal of Computer Applications*, vol. 67, no. 1, pp. 38-42, 2013.

[6] P. Rajeswari and G. S. Reena, "Human liver cancer classification using microarray gene expression data," *International Journal of Computer Applications*, vol. 34, no. 6, pp. 25-37, 2011.

[7] V. P. Khobragade and A. Vinayababu, "A classification of microarray gene expression data using hybrid soft computing approach," *International Journal of Computer Science Issues (IJCSI)*, vol. 9, issue 6, no. 2, pp. 246-255, 2012.

[8] M. Vimaladevi and B. Kalaavathi, "Cancer classification using hybrid fast particle swarm optimization with backpropagation neural network," *International Journal of Advanced Research in Computer and Communication Engineering*, vol. 3, issue 11, pp. 8410-8414, 2014.

[9] A. H. Chen and M. C. Lee, "Novel approaches for the prediction of cancer classification," *International Journal of Advancements in Computing Technology (IJACT)*, vol. 3, issue 3, pp. 30-39, 2011.

[10] I. Guyon, J. Weston, S. Barnhill, and V. Vapnik, "Gene selection for cancer classification using support vector machines," *Machine Learning*, vol. 46, issue 1, pp. 389-422, 2002.

Rashmi Nagpal and Rashmi Shrivas, "Cancer classification using elitism PSO based lezy IBK on gene expression data," *International Journal of Scientific and Technical Advancements*, Volume 1, Issue 4, pp. 19-23, 2015.

[11] A. Halder and S. Misra, "Semi-supervised fuzzy K-NN for cancer classification from microarray gene expression data," *IEEE First International Conference on Automation, Control, Energy and Systems (ACES)*, pp. 1-5, 2014.
[12] S. Mahmoudi, B. S. Lahijan, and H. R. Kanan, "ANFIS-based wrapper model gene selection for cancer classification on microarray gene expression data," *IEEE 13th Iranian Conference on Fuzzy Systems (IFSC)*, 2013.
[13] S. Kanwal, I. A. Taj, and A. Farooq, "A novel classification technique for cancer diagnostics based on microarray gene expression profiling (MGEP)," *IEEE International Conference on Emerging Technologies (ICET)*, pp. 1-6, 2012.
[14] R. Zhang, G. B. Huang, N. Sundararajan, and P. Saratchandran, "Multicategory classification using an extreme learning machine for microarray gene expression cancer diagnosis," *IEEE/ACM Transactions on Computational Biology and Bioinformatics (TCBB)*, vol. 4, issue 3, pp. 485-495, 2007.
[15] M. C. P. de Souto, A. C. Lorena, N. Spolaor, and I. G. Costa, "Complexity measures of supervised classifications tasks: a case study for cancer gene expression data," *IEEE International Joint Conference on Neural Networks (IJCNN)*, pp. 1-7, 2010.
[16] K. Fukuta, T. Nagashima, and Y. Okada, "Leaf: leave-one-out forward selection method for cancer classification using gene expression data," *IEEE/ACIS 9th International Conference on Computer and Information Science (ICIS)*, pp. 31-36, 2010.
[17] W. Chen, H. Lu, M. Wang, and C. Fang, "Gene expression data classification using artificial neural network ensembles based on samples filtering," *IEEE International Conference on Artificial Intelligence and Computational Intelligence (AICI'09)*, vol. 1, pp. 626-628, 2009.
[18] W. Luo, L. Wang, and J. Sun, "Feature selection for cancer classification based on support vector machine," *IEEE WRI Global Congress on Intelligent Systems (GCIS'09)*, vol. 4, pp. 422-426, 2009.
[19] C. H. Zheng, P. Zhang, D. Zhang, X. X. Liu, and J. Han, "Gene expression data classification based on non-negative matrix factorization," *IEEE International Joint Conference on Neural Networks (IJCNN)*, pp. 3542-3547, 2009.
[20] X. Hang, "Cancer classification by sparse representation using microarray gene expression data," *IEEE International Conference on Bioinformatics and Biomeidcine Workshops (BIBMW)*, pp. 174-177, 2008.
[21] Z. Qizhong, "Gene selection and classification using non-linear kernel support vector machines based on gene expression data," *IEEE/ICME International Conference on Complex Medical Engineering*, pp. 1606-1611, 2007.
[22] F. Chu and L. Wang, "Applying RBF neural networks to cancer classification based on gene expressions," *International Joint Conference on Neural Networks (IJCNN'06)*, pp. 1930-1934, 2006.
[23] R. Hewett, A. Goksu, and S. Datta, "From cancer gene expression data to simple vital rules," *IEEE Region 5 Conference*, pp. 329-334, 2006.
[24] W. Zhao, G. Wang, h. Wang, H. Chen, H. Dong, and Z. Zhao, "A novel framework for gene selection," *International Journal of Advance Computer Technology*, vol. 3, 2011, pp. 184-191.
[25] T. Madeswaran and G. M. K. Nawaz, "Classification of micro array gene expression data using statistical analysis approach with personalized fuzzy inference system," *International Journal of Computer Applications (IJCA)*, vol. 31, no. 1, pp. 5-12, 2011.
[26] S. Li, X. Wu, and M. Tan, "Gene selection using hybrid particle swarm optimization and genetic algorithm," *Soft Computing*, vol. 12, issue 11, pp. 1039-1048, 2008.
[27] M. Vimaladevi, and B. Kalaavathi, "A microarray gene expression data classification using hybrid back propagation neural network," *Genetika*, vol. 46, issue 3, pp. 1013-1026, 2014.
[28] G. S. Reena and P. Rajeswari, "A survey of human cancer classification using micro array data," *International Journal of Computer Technology and Applications (IJCTA)*, vol. 2, issue 5, pp. 1523-1533, 2011.
[29] G. Giarratana, M. Pizzera, M. Masseroli, E. Medico, and P. L. Lanzi, "Data mining techniques for the identification of genes with expression levels related to breast cancer prognosis," *Ninth IEEE International Conference on Bioinformatics and BioEngineering*, pp. 295-300, 2009.
[30] R. Malpani, M. Lu, D. Zhang, and W. K. Sung, "Mining transcriptional association rules from breast cancer profile data," *IEEE International Conference on Information Reuse and Integration (IRI)*, pp. 154-159, 2011.
[31] M. A. Hall, "Correlation-based feature selection for machine learning," Diss. The University of Waikato, 1999.
[32] L. Y. Chuang, C. S. Yang, K. C. Wu, and C. H. Yang, "Correlation-based gene selection and classification using Taguchi-BPSO," *Methods of Information in Medicine*, vol. 49, issue 3, pp. 254-268, 2010.
[33] J. Kennedy and R. C. Eberhart, "A discrete binary version of the particle swarm algorithm," *IEEE International Conference on Systems, Man, and Cybernetics. Computational Cybernetics and Simulation*, vol. 5, 1997.
[34] R. C. Eberhart and J. Kennedy, "A new optimizer using particle swarm theory," *Proceedings of the Sixth International Symposium on Micro Machine and Human Science*, vol. 1, pp. 39-43, 1995.
[35] S. B. Jagtap and B. G. Kodge, "Census data mining and data analysis using WEKA," *International Conference in Emerging Trends in Science, Technology and Management*, Singapore, 2013.
[36] R. Nagpal, R. Shrivas, M. Sahu, and S. Shirke, "Performance evaluation of different classifier on breast cancer," *International Journal of Technical Research and Applications (IJTRA)*, vol. 3, Issue 3, pp. 130-133, 2015.

Rashmi Nagpal and Rashmi Shrivas, "Cancer classification using elitism PSO based lezy IBK on gene expression data," *International Journal of Scientific and Technical Advancements*, Volume 1, Issue 4, pp. 19-23, 2015.