BREAST CANCER DIAGNOSIS VIA DATA MINING: PERFORMANCE ANALYSIS OF SEVEN DIFFERENT ALGORITHMS

Zehra Karapinar Senturk [1] and Resul Kara[1]

[1]Department of Computer Engineering, Duzce University, Duzce, Turkey

ABSTRACT

According to World Health Organization (WHO), breast cancer is the top cancer in women both in the developed and the developing world. Increased life expectancy, urbanization and adoption of western lifestyles trigger the occurrence of breast cancer in the developing world. Most cancer events are diagnosed in the late phases of the illness and so, early detection in order to improve breast cancer outcome and survival is very crucial.

In this study, it is intended to contribute to the early diagnosis of breast cancer. An analysis on breast cancer diagnoses for the patients is given. For the purpose, first of all, data about the patients whose cancers' have already been diagnosed is gathered and they are arranged, and then whether the other patients are in trouble with breast cancer is tried to be predicted under cover of those data. Predictions of the other patients are realized through seven different algorithms and the accuracies of those have been given. The data about the patients have been taken from UCI Machine Learning Repository thanks to Dr. William H. Wolberg from the University of Wisconsin Hospitals, Madison. During the prediction process, RapidMiner 5.0 data mining tool is used to apply data mining with the desired algorithms.

KEYWORDS

Data mining, breast cancer diagnoses, RapidMiner

1. INTRODUCTION

There is a great increase in data owned as we come from past to present, and so, to control and manage those rapidly increasing data get harder evenly. When the calculations that are kept on papers did not suffice to store data and also when to find a data got harder, the need for easy manageable and relatively big systems appeared. It is started to keep rapidly increasing data in computer hard discs through the proliferation of computer usage. Although the usage of only computer hard discs seems to be solution at first glance, difficulties in some operations like accessing data that takes up large spaces in memories and making changes in some data directed people to the idea of database management systems. The facility to make the operations on stored data easily is provided by those systems. The operations that normally take a lot of time to achieve are realized in a short period of time with error rate minimization thanks to database management systems. However, current database management systems become insufficient when the needs require obtaining more information from data. The need of gathering more and more information from data is felt in about all areas of life and the methods are considered to satisfy

these needs. Realistic predictions for the future are made by analyzing the data on hand with the developed methods. The process of obtaining information from data is then called as "data mining".

Data mining can be defined as analyzing data from different perspectives and summarizing it to obtain useful information. Information here may be used for the purposes like increasing income or decreasing costs. Technically, data mining is the process of finding certain relationships or models among dozens of area in very big relational databases.

The purpose of this study is to make analysis to be used for diagnoses of breast cancer illness with data mining. Thus, the leeway problem which is vital in cancer illnesses will vanish and the acquired time may then be used for the treatment of the illness.

In the literature, there are many studies done on cancer detection and/or data mining. [7] used data mining for the diagnosis of ovarian cancer. For the analysis, serum proteomics that distinguish the serum ovarian cancer cases from non-cancer ones are used. An SVM (Support Vector Machine) based method is applied and statistical testing and GA (Genetic Algorithms) based methods are used for feature selection. [6] aimed to propose a new 3-D microwave approach based on SVM classifier whose output is transformed to a posteriori probability of tumor presence. Gene expression data sets for ovarian, prostate and lung cancers are analyzed in another paper [13]. An integrated gene search algorithm (preprocessing: GA and correlation based heuristics, making predictions/ data mining: decision tree and SVM algorithms) for genetic expression data analysis is proposed. In [11] the clinical and imaging diagnostic rules of peripheral lung cancer by data mining techniques that are Association Rules (AR) of knowledge discovery process and Rough Set (RS) reduction algorithm and Genetic Algorithm (GA) of generic data analysis tool (ROSETTA) are extracted [11]. [14] deals with complementary learning fuzzy neural network (CLFNN) for the diagnosis of ovarian cancer. CLFNN-micro-array, CLFNN-blood test, CLFNN-proteomics demonstrates good sensitivity and specificity. So, it is shown that CLFNN outperforms most of the conventional methods in ovarian cancer diagnosis. [15] applies the classification technology to construct an optimum cerebrevascular disease predictive model. Classification algorithms used are decision tree, Bayesian classifier, and back propagation neural network.

The objective of [3] is to develop an original method to extract sets of relevant molecular biomarkers (gene sequences) that can be used for class prediction and as a prognostic and predictive tool. With the help of the analysis of DNA microarrays, molecular biomarkers are generated and this analysis is based on a specific data mining technique: Sequential Pattern Discovery.

The performance of data classification by integrating artificial neural networks with multivariate adaptive regression splines (MARS) approach is explored for mining breast cancer pattern [1]. This approach is based on firstly to use MARS in modeling the classification problem, then obtained significant variables are used as input variables of designed neural networks model. A comparison of three data mining techniques artificial neural networks, decision trees, and logistic regression is realized in a study to predict the survivability of breast cancer [2]. Accuracy rates are found as 93.6%, 91.2%, and 89.2% respectively. Many aspects of possible relationships among DNA viruses and breast tumors are considered [9]. Feasible clusters in DNA virus combinations that depend on the observed probability of breast cancer, fibro adenoma and normal mammary tissue are created in this study and viral prerequisites for breast carcinogenesis and the protectives are determined. Obtaining bioinformatics about breast tumor and DNA viruses, and building an

accurate diagnosis model for breast cancer and fibro adenoma are aimed [4]. A hybrid SVM-based strategy with feature selection to render a diagnosis between the breast cancer and fibro adenoma and to find important risk factor for breast cancer is constructed. DNA viruses, HSV-1, EBV, CMV, HPV and HHV-8 are evaluated. There is also another study related to breast cancer. Breast cancer pattern is mined using discrete particle swarm optimization and statistical method [16]. Besides, to detect breast cancer, association rules (AR) and neural network (NN) are used this time [5]. AR is used to reduce the dimension of the database and NN is used for intelligent classification. In Menendeza et al (2010), a Self-Organizing Map (SOM) based clustering algorithm for preprocessing of samples from a breast cancer screening program is introduced. Prediction of the recurrence of breast cancer is investigated [7]. The accuracy of Cox Regression and SVM algorithms are compared and it is shown that a parallelism of adequate treatment and follow-up by recurrence prediction prevent the recurrence of breast cancer.

In this study, different from the studies stated above, breast cancer is tried to be predicted whether as a benign or malignant case through seven different algorithms which have not been tried for breast cancer yet in the literature and a performance analysis is aimed to be performed.

2. MATERIALS AND METHODS

In this study, data mining is applied to the health sector. Possible cancer diagnoses for new patients whose other data (laboratory results) exists in hospital databases, but diagnoses have not been determined yet are to be predicted using the data of the patients whose breast cancer have been diagnosed before. Different algorithms have been used for the operation of predicting and the one with the high confidence can then be preferred.

The required data about breast cancer patients have been taken from UCI Machine Learning Repository thanks to Dr. William H. Wolberg from the University of Wisconsin Hospitals, Madison. This data includes 699 samples with 10+1 attributes (1 for class). These attributes are as follows:

```
    #  Attribute                      Domain
-- ---------------------------------------------------------
    1. Sample code number             id number
    2. Clump Thickness                 1 - 10
    3. Uniformity of Cell Size         1 - 10
    4. Uniformity of Cell Shape        1 - 10
    5. Marginal Adhesion              1 - 10
    6. Single Epithelial Cell Size    1 - 10
    7. Bare Nuclei                    1 - 10
    8. Bland Chromatin               1 - 10
    9. Normal Nucleoli                1 - 10
   10. Mitoses                       1 - 10
   11. Class:              (2 for benign, 4 for malignant)
```

In this data set we have 458 benign and 241 malignant cases. There were some attributes having "?" value and those are removed from the set in the data preprocessing phase that is before mining.

After the data is obtained and cleared, they are divided into two sets as training and testing. Some of them are used in training phase and the rest are used for testing the algorithms. Then, data is transferred to RapidMiner data mining tool and breast cancer diagnosis for each sample in the test set is predicted with seven different algorithms which are Discriminant Analysis, Artificial Neural Networks, Decision Trees, Logistic Regression, Support Vector Machines, Naïve Bayes, and K-NN. Last but not least, the performance analysis including these algorithms is realized and the best one for breast cancer is determined.

Prediction mechanism in RapidMiner can be summarized as shown in the figure below. Model box here stands for the selected algorithm. In our case, this structure will be established and run 7 times for our 7 algorithms.
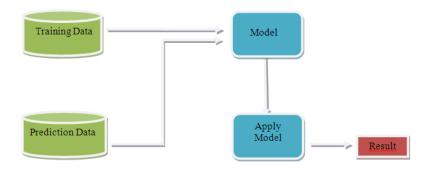
Figure1. Modeling of Breast Cancer Diagnoses System

For the performance analysis is schematically shown below in Figure 2. In this process 10-fold cross validation is used.
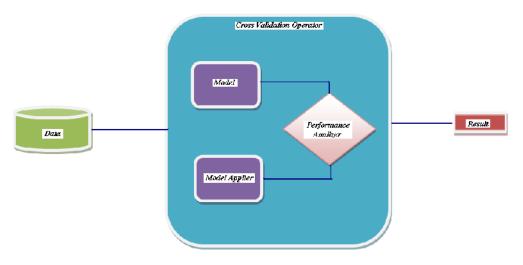
Figure2. Performance analysis mechanism

The algorithms used in RapidMiner for the diagnosis of breast cancer are given below with the explanations in RapidMiner 5.0 Help.

2.1. Discriminant Analysis

Discriminant analysis in RapidMiner is applied with nominal labels and numerical attributes. It is used to determine which variables discriminate between two or more naturally occurring groups, it may have a descriptive or a predictive objective. Discriminant analysis is performed in three ways as linear, quadratic, and regularized in RapidMiner. In linear case, a linear combination of features which best separates two or more classes of examples is tried to be found. Then, the resultant combination is used as a linear classifier. Linear Discriminant analysis is somewhat like the variance analysis and regression analysis with some difference. In these two methods, the dependent variable is a numerical value while it is a categorical value in LDA (Linear Discriminant Analysis). LDA is also related to principle component analysis (PCA) and factor analysis (both look for linear combinations of variables which best explain the data), but PCA and other methods does not consider the difference in classes while LDA attempts to model the difference between the classes of data.

Quadratic Discriminant Analysis (QDA) is closely related to linear discriminant analysis (LDA), where it is assumed that the measurements are normally distributed. Unlike LDA however, in QDA there is no assumption that the covariance of each of the classes is identical.

The regularized discriminant analysis (RDA) is a generalization of the LDA and QDA. Both algorithms are special cases of this algorithm. If the alpha parameter is set to 1, RDA operator performs LDA. Similarly if the alpha parameter is set to 0, RDA operator performs QDA.

In our problem we applied linear form of discriminant analysis.

2.2. Artificial Neural Networks (Multi Layer Perceptron)

Multi Layer Perceptron is a classifier that uses back propagation to classify instances. This network can be built by hand, created by an algorithm or both. The network can also be monitored and modified during training time. The nodes in this network are all sigmoid (except for when the class is numeric in which case the output nodes become unthresholded linear units). Parameters of this algorithm are:

L: Learning Rate for the backpropagation algorithm. (Value should be between 0 - 1, Default = 0.3). *Range:* real; -?-+?
M: Momentum Rate for the backpropagation algorithm. (Value should be between 0 - 1, Default = 0.2). *Range:* real; -?-+?
N: Number of epochs to train through. (Default = 500). *Range:* real; -?-+?
V: Percentage size of validation set to use to terminate training (if this is non zero it can pre-empt num of epochs. (Value should be between 0 - 100, Default = 0). *Range:* real; -?-+?
S: The value used to seed the random number generator (Value should be >= 0 and and a long, Default = 0). *Range:* real; -?-+?
E: The consequetive number of errors allowed for validation testing before the netwrok terminates. (Value should be > 0, Default = 20). *Range:* real; -?-+?
G: GUI will be opened. (Use this to bring up a GUI). *Range:* boolean; default: false
A: Autocreation of the network connections will NOT be done. (This will be ignored if -G is NOT set) *Range:* boolean; default: false
B: A NominalToBinary filter will NOT automatically be used. (Set this to not use a NominalToBinary filter). *Range:* boolean; default: false

H: The hidden layers to be created for the network. (Value should be a list of comma separated Natural numbers or the letters 'a' = (attribs + classes) / 2, 'i' = attribs, 'o' = classes, 't' = attribs .+ classes) for wildcard values, Default = a). *Range:* string; default: 'a'
C: Normalizing a numeric class will NOT be done. (Set this to not normalize the class if it's numeric). *Range:* boolean; default: false
I: Normalizing the attributes will NOT be done. (Set this to not normalize the attributes). *Range:* boolean; default: false
R: Resetting the network will NOT be allowed. (Set this to not allow the network to reset). *Range:* boolean; default: false
D: Learning rate decay will occur. (Set this to cause the learning rate to decay). *Range:* boolean; default: false[12]

2.3. Decision Trees

RapidMiner generates a Decision Tree for classification of both nominal and numerical data. A decision tree is a tree-like graph or model. It is more like an inverted tree because it has its root at the top and it grows downwards. This representation of the data has the advantage compared with other approaches of being meaningful and easy to interpret. The goal is to create a classification model that predicts the value of a *target attribute* (often called *class* or *label*) based on several input attributes of the Example Set. In RapidMiner an attribute with *label* role is predicted by the Decision Tree operator. Each interior node of tree corresponds to one of the input attributes. The number of edges of a nominal interior node is equal to the number of possible values of the corresponding input attribute. Outgoing edges of numerical attributes are labeled with disjoint ranges. Each leaf node represents a value of the *label* attribute given the values of the input attributes represented by the path from the root to the leaf.

Decision Trees are generated by recursive partitioning. Recursive partitioning means repeatedly splitting on the values of attributes. In every recursion the algorithm follows the following steps:

- An attribute A is selected to split on. Making a good choice of attributes to split on each stage is crucial to generation of a useful tree. The attribute is selected depending upon a selection criterion which can be selected by the criterion parameter.
- Examples in the Example Set are sorted into subsets, one for each value of the attribute A in case of a nominal attribute. In case of numerical attributes, subsets are formed for disjoint ranges of attribute values.
- A tree is returned with one edge or branch for each subset. Each branch has a descendant subtree or a label value produced by applying the same algorithm recursively.

In general, the recursion stops when all the examples or instances have the same *label* value, i.e. the subset is pure. Or recursion may stop if most of the examples are of the same *label* value. This is a generalization of the first approach; with some error threshold. However there are other halting conditions such as:

- There are less than a certain number of instances or examples in the current subtree. This can be adjusted by using the minimal size for split parameter.
- No attribute reaches a certain threshold. This can be adjusted by using the minimum gain parameter.
- The maximal depth is reached. This can be adjusted by using the maximal depth parameter.

Pruning is a technique in which leaf nodes that do not add to the discriminative power of the decision tree are removed. This is done to convert an over-specific or over-fitted tree to a more general form in order to enhance its predictive power on unseen datasets. Pre-pruning is a type of pruning performed parallel to the tree creation process. Post-pruning, on the other hand, is done after the tree creation process is complete [12].

2.4. Logistic Regression

Logistic Regression operator is a Logistic Regression Learner. It is based on the internal Java implementation of the *myKLR* by Stefan Rueping. *myKLR* is a tool for large scale kernel logistic regression based on the algorithm of Keerthi etal (2003) and the code of *mySVM*. For compatibility reasons, the model of *myKLR* differs slightly from that of Keerthi et al (2003). As *myKLR* is based on the code of *mySVM*; the format of example files, parameter files and kernel definition are identical. Detailed information is given in the next section(Support Vector Machines). This learning method can be used for both regression and classification and provides a fast algorithm and good results for many learning tasks. *mySVM* works with linear or quadratic and even asymmetric loss functions [12].

2.5. Support Vector Machines

A basic description of the SVM is that the standard SVM takes a set of input data and predicts, for each given input, which of the two possible classes comprises the input, making the SVM a non-probabilistic binary linear classifier. Given a set of training examples, each marked as belonging to one of two categories, an SVM training algorithm builds a model that assigns new examples into one category or the other. An SVM model is a representation of the examples as points in space, mapped so that the examples of the separate categories are divided by a clear gap that is as wide as possible. New examples are then mapped into that same space and predicted to belong to a category based on which side of the gap they fall on.

More formally, a support vector machine constructs a hyperplane or set of hyperplanes in a high- or infinite- dimensional space, which can be used for classification, regression, or other tasks. Intuitively, a good separation is achieved by the hyperplane that has the largest distance to the nearest training data points of any class (so-called functional margin), since in general the larger the margin the lower the generalization error of the classifier. Whereas the original problem may be stated in a finite dimensional space, it often happens that the sets to discriminate are not linearly separable in that space. For this reason, it was proposed that the original finite-dimensional space be mapped into a much higher-dimensional space, presumably making the separation easier in that space. To keep the computational load reasonable, the mapping used by the SVM schemes are designed to ensure that dot products may be computed easily in terms of the variables in the original space, by defining them in terms of a kernel function $K(x,y)$ selected to suit the problem. The hyperplanes in the higher dimensional space are defined as the set of points whose inner product with a vector in that space is constant [12].

2.6. Naïve Bayes

A Naive Bayes classifier is a simple probabilistic classifier based on applying Bayes' theorem (from Bayesian statistics) with strong (naive) independence assumptions. A more descriptive term for the underlying probability model would be 'independent feature model'. In simple terms, a Naive Bayes classifier assumes that the presence (or absence) of a particular feature of a class

(i.e. attribute) is unrelated to the presence (or absence) of any other feature. For example, a fruit may be considered to be an apple if it is red, round, and about 4 inches in diameter. Even if these features depend on each other or upon the existence of the other features, a Naive Bayes classifier considers all of these properties to independently contribute to the probability that this fruit is an apple.

The advantage of the Naive Bayes classifier is that it only requires a small amount of training data to estimate the means and variances of the variables necessary for classification. Because independent variables are assumed, only the variances of the variables for each *label* need to be determined and not the entire covariance matrix [12].

2.7. K-Nearest Neighborhood

The k-Nearest Neighbor algorithm is based on learning by analogy, that is, by comparing a given test example with training examples that are similar to it. The training examples are described by n attributes. Each example represents a point in an n-dimensional space. In this way, all of the training examples are stored in an n-dimensional pattern space. When given an unknown example, a k-nearest neighbor algorithm searches the pattern space for the k training examples that are closest to the unknown example. These k training examples are the k "nearest neighbors" of the unknown example. "Closeness" is defined in terms of a distance metric, such as the Euclidean distance.

The k-nearest neighbor algorithm is amongst the simplest of all machine learning algorithms: an example is classified by a majority vote of its neighbors, with the example being assigned to the class most common amongst its k nearest neighbors (k is a positive integer, typically small). If k = 1, then the example is simply assigned to the class of its nearest neighbor. The same method can be used for regression, by simply assigning the label value for the example to be the average of the values of its k nearest neighbors. It can be useful to weight the contributions of the neighbors, so that the nearer neighbors contribute more to the average than the more distant ones.

The neighbors are taken from a set of examples for which the correct classification (or, in the case of regression, the value of the label) is known. This can be thought of as the training set for the algorithm, though no explicit training step is required.

The basic k-Nearest Neighbor algorithm is composed of two steps: Find the k training examples that are closest to the unseen example. Take the most commonly occurring classification for these k examples (or, in the case of regression, take the average of these k label values). [12]

3. RESULTS

It is easily achieved to access to the golden data from raw data by data mining techniques in recent years at which competition and time concept gain importance gradually. Data mining techniques are started to be used frequently especially in the area of medicine since the data amount accumulates more in that area when compared to the other areas and since the analysis of data is vitally important.

A study on early diagnoses of breast cancer illnesses that mostly ends with deaths is presented in this paper. Patient data gathered from UCI is used for the purpose. According to the selected prediction model, some diagnostics are predicted with some confidence values for the patients

whose diagnoses are unknown by using the cancer events whose diagnostics have been determined before.

The patients whose diagnoses are known are also included in the model as if their diagnoses are unknown to test how much the system is coherent with the reality. Which algorithm is more effective for the situation is shown by using different algorithms during the observation.

Accuracy values for each algorithm are shown separately in tables. In short, we can say that Support Vector Machines algorithm with previously stated attributes is the best for breast cancer prediction. Accuracy rates of each algorithm are shown below and in the last figure, a comparison is given and the best one is emphasized.

Table 1: Discriminant Analysis with RapidMiner

	True 0	True 1
Prediction 0	368	29
Prediction 1	6	196
Accuracy	98.40%	87.11%

Table 2: Multi Layer Perceptron (an Artificial Neural Network) with RapidMiner

	True 0	True 1
Prediction 0	357	6
Prediction 1	17	219
Accuracy	95.45%	97.33%

Table 3: Decision Trees with RapidMiner

	True 0	True 1
Prediction 0	359	16
Prediction 1	15	209
Accuracy	95.99%	92.89%

Table 4: Logistic Regression with RapidMiner

	True 0	True 1
Prediction 0	364	14
Prediction 1	10	211
Accuracy	97.33%	93.78%

Table 5: Support Vector Machines with RapidMiner

	True 0	True 1
Prediction 0	362	9
Prediction 1	12	216
Accuracy	96.79%	96.00%

Table 6: Naïve Bayes with RapidMiner

	True 0	True 1
Prediction 0	356	5
Prediction 1	18	220
Accuracy	95.19%	97.78%

Table 7: K-NN with RapidMiner

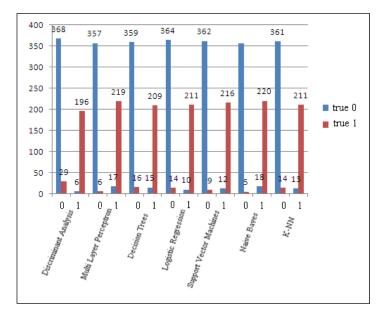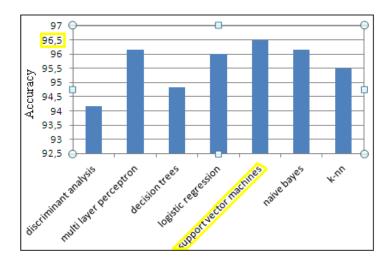|	True 0	True 1
Prediction 0	361	14
Prediction 1	13	211
Accuracy	96.52%	93.78%

Figure 3.True 0 and True 1 values for prediction of 0 and 1

Zeros and ones in the horizontal axis of the above figure stand for the prediction of zero and one respectively.

Figure 4.Accuracy percentages of algorithms

4. DISCUSSION

Data mining can be applied to other types of cancers too. In the literature, some of them have already been used in some cancer types, but it can be expanded. The same prediction mechanism (breast cancer diagnosis) can be performed with different attributes (not with the 10 attributes stated above) and the effectiveness of the other factors can be analyzed.

REFERENCES

[1] Choua, S.M., T.S. Leeb, Y. E. Shaoc, and I.F. Chen (2004) "Mining breast cancer pattern using artificial neural networks and multivariate adaptive regression splines", Expert Systems with Applications,No. 27,pp133-142

[2] Delen, D., G. Walker, and A. Kadam (2005) "Predicting breast cancer survivability: a comparison of three data mining methods", Artificial Intelligence in Medicine,No.34,pp 113-127

[3] Fabregue, M., S. Bringay, P. Poncelet, M. Teisseire, and B. Orsetti (2011) "Mining microarray data to predict the histological grade of a breast cancer", Journal of Biomedical Informatics. No.44,pp12-16

[4] Huang, C., H.C. Liao, and M.C. Chen (2008) "Prediction model building and future selection with support vector machines in breast cancer diagnosis". Expert Systems with Applications, No. 34,pp578-587

[5] Karabatak, M., and M.C. Ince(2009) "An expert system for detection of breast cancer based on association rules and neural network", Expert Systems with Applications, No.36,pp3465-3469

[6] Kerheta, A., M. Raffettob, A. Bonia, and A. Massa (2006) "An SVM-based approach to microwave breast cancer detection", Engineering Applications of Artificial Intelligence, No.19,pp807-818

[7] Kim K.S., W. Kim, K.Y. Na, J.M. Park, J.Y. Kim, K.Y. Lee, J.E. Lee, S.W. Kim, R.W. Park, and Y.S. Jung(2010) "New recurrence prediction model for breast cancer by data mining",pp136

[8] Lia L., H. Tanga, Z. Wua, J. Gonga, M. Gruidlb, J. Zoub, M. Tockmanb, and R. A. Clark (2004), "Data mining techniques for cancer detection using serum proteomic profiling", Artificial Intelligence in Medicine, No.32,pp71-83

[9] Liao H.C.,and J.H. Tsai (2007), "Data mining for DNA viruses with breast cancer, fibroadenoma, and normal mammary tissue", Applied Mathematics and Computation, No.188, pp989-1000

[10] Menéndeza, L. Á., F.J. de Cos Juez, F. Sánchez Lasheras, and J.A. Álvarez Riesgo (2010) "Artificial neural networks applied to cancer detection in a breast screening programme", Mathematical and Computer Modelling. No.52,pp983-991

[11] Qiang, Y., Y. Guo, X. Li, Q. Wanga, H. Chenc, and D. Cuicc (2007) "The diagnostic rules of peripheral lung cancer preliminary study on data mining techniques", Journal of Nanjing Medical University. No.21(3),pp190-195

[12] RapidMiner 5.0 Help (2013)

[13] Shah, S., and A. Kusiak (2007) "Cancer gene search with data mining and genetic algorithms", Computers in Biology and Medicine. No.37,pp251-261

[14] Tan, T. Z., C. Queka, G. S. Ng, and K. Razvi (2008) "Ovarian cancer diagnosis with complementary learning fuzzy neural network", Artificial Intelligence in Medicine. No.43, pp207-222,

[15] Yeh, D.Y., C.H. Cheng, and Y.W. Chen (2011) "A predictive model for cerebrovascular disease using data mining", Expert Systems with Applications. No.38, pp8970–8977

[16] Yeh, W.C., W.W. Chang, and Y. Y. Chung (2009) "A new hybrid approach for mining breast cancer pattern using discrete partical swarm optimization and statistical method", Expert Systems with Applications. No.36, pp8204-8211

Authors:

Zehra Karapinar Senturk graduated from Computer Engineering department of Atilim University, Ankara, Turkey in 2009. Currently she is a Ph.D. student at Computer Engineering department of Gazi University, Ankara, Turkey and research assistant at Duzce University, Duzce, Turkey. Her main interests are data mining, image processing, human-computer interaction, and biomedical engineering.

Resul Kara graduated from Electrics-Electronics Engineering department of Gazi University, Ankara, Turkey in 1996. He received his Ph.D. degree from Sakarya University, Turkey in 2009. Currently he is the head of department of Computer Engineering at Duzce University. His main interests are communication, computer networks, and optimization.