Comprising Feature Selection and Classifier Methods with SMOTE for Prediction of Male Infertility

Bekir Karlık, Abdulkerim M. Yibre, and Barış Koçer

***Abstract*—**The aim of this study is on prediction of male fertility problems by using different classifier algorithms and applying feature selection methods. Used data in this study is gathered from UCI data repository. The ratio of between normal and abnormal (alter) samples is more than 7. This indicates the dataset is originally imbalanced in order to balance the examples of the dataset SMOTE technique which is applied for a better accuracy and representative result of classifiers. Feature selection and classification methods are comprised for prediction of male fertility. In this study, MLP, Naïve Bayes, Random Forest, KNN, and SVM classifiers were used. Comparison results show that Naive Bayes classifier has better classification accuracy as 90.65% than the others.

***Keywords*—**Classifier, feature selection, fertility, prediction

I. Introduction

According to World Health Organization (WHO), infertility is explained as the impotency of a sexually active, non-contracepting couple to achieve pregnancy in a year [1]. It is reported that a male contribution in infertility is 45-50%. Studies show that there are different causes of male infertility. Male fertility problems are related to the quality of the semen produced [2]. A research undertaken in Nigeria by involving sixty male volunteers indicated that cadimium as a cause of infertility [3]. Exposure to chemicals such as pesticides and solvents is highly related to low threshold values of sperm [4]. In the study the researchers concluded that environmental factors contribute highly for infertility. It is reported that 10-15% of sever male infertility is due to genetic reasons including chromosomal aberrations and single gene mutations [5]. The standardized outline published by WHO is followed in clinics for investigation of male comprising of full medical history and physical examination. The investigations includes semen analysis, hormonal investigation, microbiologic assessment, Ultrasonography and testicular biopsy.

Machine learning approaches in the medical areas has been used in the examination and diagnosis of diseases such as arrthymia [6-8], asthma and bronchitis [9-10], breast cancer [11], retinopathy [12], epileptic seizure [13-15], diabetes [16-17], hepatitis [18], lumbar disc hernia [19-21], otitis media [22], pediatric nutritional requirements [23] etc.

B. Karlık, Selçuk University, The Engineering Faculty, Department of Computer Engineering, 42075, Selçuklu, Konya, TURKEY, (corresponding author: +90-332-2231981; fax: +90-332-2231981; e-mail: bkarlik@selcuk.edu.tr).

A.M. Yibre, Selçuk University, The Engineering Faculty, Department of Computer Engineering, 42075, Selçuklu, Konya, TURKEY, (e-mail: abdukerimm@gmail.com).

B. Koçer, Selçuk University, The Engineering Faculty, Department of Computer Engineering, 42075, Selçuklu, Konya, TURKEY, (e-mail: bariskocer@selcuk.edu.tr).

Previous study on fertility has been by Gil et al [24] by using three types of machine learning algorithms such as MLP, Decision Tree and SVM. They have used UCI data repository. The data is composed of 88 normal and 12 altered examples (instances) which shows existence of imbalance between normal and altered examples. They have found the highest accuracy as 86% using SVM method.

According to [25] and [26] data imbalance has impact on the representativeness of classification accuracy. Taking this into account, in this study we have employed SMOTE technique to resample the original data. Moreover, performances of different machine learning algorithms (or classifiers) such as Multilayer Perceptron (MLP), Naïve Bayes, Random Forest, K-Nearest Neighbor (K-NN), and Support Vector machines (SVM) applied on fertility data, were compared in this study. Aiming at boosting the accuracy of classifiers, feature selection techniques namely; Information Gain ratio attribute selector, Principal Component Analysis (PCA) and SVM based attribute selectors, which are all available in WEKA, were used.

The continuing subsections of the paper is organized as follows. In the next section, we discussed classification methods employed in the study including Multilayer Perceptron Network, Naïve Bayes, Random Forest, K-Nearest Neighbor, and Support Vector Machines. Section 3 describes the material and methods used in this study. Under this section, the data source for the study and tools used is discussed. In section 4, experimental data processing with methods, and result comparisons of used classifiers are presented in detail. Finally, some concluding remarks are provided in Section V.

II. USED CLASSIFIERS

A. Naïve Bayes Classifier

Bayes classification method is based on Bayes' theorem. It is a statistical classifier that can be predicted class (or label) membership probabilities such as the probability a given example belongs to each class [27].

$$P(X|C_i) = \prod_{k=1}^{n} P(X_k|C_i) = P(X_1|C_i)P(X_2|C_i) \dots P(X_n|C_i) \quad (1)$$

where X_n describes the nth possible value for X, X_i describes the kth possible vector value for X, and where the summation in the denominator is over all legal values of the random variable C [18].

B. K-Nearest Neighbor (K-NN)

K-NN is one of lazy classifiers. Lazy classifiers the learners wait till the last cycle before creating a model to classify a given set of examples. Nearest-neighbor classifiers are based on learning by comparison of a given test dataset with training dataset which are similar to it where the samples of training data set are defined by n attributes. Each example is indicates a point in n-dimensional space. K-NN searches pattern space for the *k* training data samples that are nearest to the unknown example. Nearness is measured by using a distance metric like Euclidean distance whic is measured between two points or data samples, for example X_1= (x_{11}, x_{12}... x_{1n}) and X_2 =(x_{21}, x_{22}... x_{2n}) is measured by

$$dist(X_1, X_2) = \sqrt{\sum_{i=1}^{n}(X_{1i} - X_{2i})^2} \qquad (2)$$

K-NN can be computationally expensive when making estimation or classification. They also require efficient storage techniques. However, K-NN assists incremental learning, and can be modeled with complicated decision spaces.

C. Artificial Neural Network (ANN)

An ANN is a set of parallel processing elements connected input/desired-output in which each connection has a weight associated with it. Learning in neural network is accomplished by updating the weights beween two layers so as to be able to estimate class of the input. Error back-propagation training algorithm performs learning on multi-layer feed-forward neural network structure by iteratively learning a set of weights for estimation of the class [27]. A multi layered perceptron (MLP) generally consists of three layers fully connected to the next layer. These layers are; Input layer gets inputs, a hidden layer, and the output layer which is generated result of classification. Layers with the exception of the input layer are processing element with a nonlinear transfer function. When input dataset are applied to the input layer, the nodes (or artificial neurons) operates calculations in the sequential order layers until an output is generated at the output layer [24]. The inputs are provided at the same time into the input layer. The outputs of the first hidden layer can be an input for the other hidden layer if necessary more than one hidden layer.

D. Support Vector Machines (SVM)

SVM is a useful classifier capable of classifying both linear and nonlinear dataset. SVM is used a nonlinear mapping to convert the training data into a higher dimension. It investigates for linear optimal within the separating hyperplane, a decision boundary separating the tuples of one class from the other. SVM is discovered hyperplane using support vectors and margins. SVM classifiers are very true and much less inclined to over fitting than the other classifiers. In addition the support vectors provide a compact description of the learned vector. However, the training time for SVM classifier can be slow than the other methods.

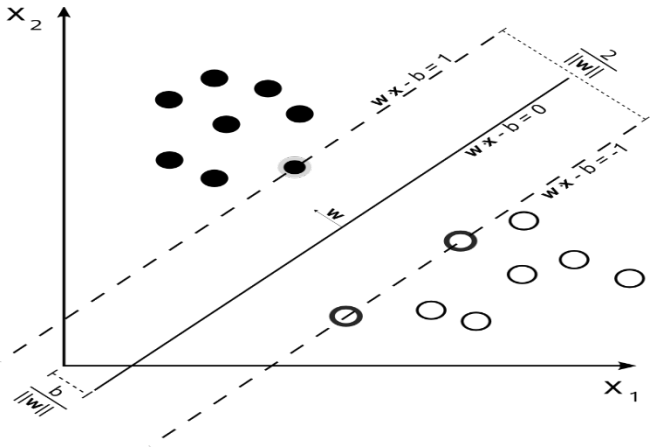

Fig. 1 Structure of linearly separable classes in Support Vector Machines

E. Random Forest Classifier

Random forest classifier (RFC) is among ensemble learning techniques proven to be popular and effective methods in pattern recognition [29]. RFC can be seen as one classifier containing several classification algorithms. Given a learning set L=((M1,N1),…, (Mn, Nn) made of n vectors, M element of X, where X is a set of numerical or symbolic observation and N element of Y, where Y is a set of class for classification problems, a classifier mapping X into Y. Classification results are produced by each tree. In RFC, the principle is building binary sub-tress applying training bootstrap samples coming from the learning sample and selecting randomly at each node. Then classification having the most votes over all the tress in the forest is selected by the decision forest [29].

For each decision tree random forest training is accomplished where each classifier's training data set is generated by randomly drawing N examples, with N the size of the original training data set. So as to form the final classifier the learning system produces a classifier from data sample and collects all the classifiers generated from different test. Every classifier registers a vote for the class to which a given instance belongs to and the instance is labeled as member of the class with the most votes. A class to the instance is randomly selected in case more than one class jointly receives the maximum number of votes. RFC has disadvantage of visualizing decision tree forest, rather it is more of a black box.

III. MATERIALS AND METHODS

A. Data Normalization

The data for the study is collected from UCI biomedical data repository. It is gathered from 100 young healthy volunteer among the University of Alicane students. The volunteers provide semen samples and also had fulfilled a questionnaire their healthy status and living habits.

A total of ten attributes (Nine attributes and one label or class attribute) are available from UCI dataset repository (see Table 1).The data is normalized based on rules set by the researchers [24]. The values of the attributes are normalized between -1 and 1. The same normalization method is adopted for this study. As to our knowledge there are no many researches done using this data set.

TABLE I. ATTRIBUTES AND VALUES

No	Attribute Name	Values	Normalization
1	Season	1) winter 2) spring 3) summer 4) fall	(-1, -0.33, 0.33, 1)
2	Age	18-36	(0, 1)
3	Childish disease	1) Yes 2) No	(0, 1)
4	Accident or series trauma	1) Yes 2) No	(0, 1)
5	Surgical intervention	1) Yes 2) No	(0, 1)
6	High fever within 1 year	1) less than 3 months, 2) greater than 3 months, 3) no	(-1, 0, 1)
7	Alcohol	1) several times, 2) everyday, 3) several times a week, 4) once a week, 5) No	(0, 1)
8	Smoking	1) never, 2) occasional, 3) daily	(-1, 0, 1)
9	Hours spent sitting per day	0-24	(0, 1)
10	class	Normal or Altered	(N, A)

The data is composed of 88 normal and 12 altered examples (instances) in which normal examples are more than 7 times bigger than altered ones. This indicates the existence of imbalance between examples which can have impact on the performance of classification algorithms. In order to address this problem, SMOTE data re-sampling technique for minority examples (altered) is adopted. SMOTE allows to synthetically increase the number of the minority examples based on percentage to increase and the ratio of the imbalance between the minority and majority examples [25-28]. Aiming at achieving a better classification accuracy, in this study the minority examples are increased by 100%, 200%, 300%, 400%, 500% and 600%. Among the six, 600% SMOTE has registered best average performance on the selected classifier algorithms. Hence the minority examples are increased by 600%, i.e the examples are increased from 12 to 84 making the total examples 172. WEKA 3.6 classifier an Orange 2.7.8 are used for date preprocessing and classification task.

B. SMOTE- Synthetic Minority Over-sampling TEchnique

Predictive accuracy is typically used to measure performance of classifiers [25-26]. However this approach is not appropriate when data is imbalance for three reasons [26]. Firstly, if searching is guided by accuracy ratio, it benefits the majority examples. Secondly, classification rules that estimate positive class are highly specialized thus ignoring negative classes. Thirdly, it is difficult to identify between minority class and noise examples. Due to this, studies indicate re-sampling of data is important to alleviate the aforementioned problem. SMOTE is one of the techniques used to resample examples in imbalanced dataset [25-26]. SMOTE is a technique to over-sample the minority class by taking each minority class sample and introducing synthetic examples along the line segments joining any or all of the *k* minority class nearest neighbors [25-26]. SMOTE technique gains the benefits of avoiding the over-fitting problem of the minority class by interpolating new minority class instances rather than duplicating the existing instances [25].

Synthetic samples are generated as follows:

• Compute the difference between the feature vectors considered and its nearest neighbor.

• Multiply this difference by a random number between 0 and 1,

• Add the feature vectors considered. This give rise to the selection of a random point along the line segment between two special properties. This approach effectively pushes the decision zone of the minority class to make more general. For this study SMOTE from 100% to 600% has been experimented and among them the SMOTE 600% has performed best in terms ROC (see Table-II; *due to shortage of space we did not included the results for SMOTE 100&200*%). Therefore, the classification task is undertaken after resampling the minority class samples (altered) 600% rising the number of samples from 12 to 84.

TABLE II COMPARISON OF SMOTE VALUES ACCORDING TO REGION UNDER CURVE (ROC)

	SMOTE	300%	400%	500%	600%
Sample Size	Altered	48	60	72	84
	Normal	88	88	88	88
Accuracy (%)	NB	93.7	94	93	97.7
	MPN	88	86	86	90
	KNN	93.6	89.1	87	95.3
	SVM	81.1	80	80	82
	RF	92.1	92	93.7	95.4
	AVG	89.7	88.22	87.94	92.08

C. Attribute Selection

Studies indicate that dimensionality reduction of dataset boost performance of classification [29]. Three attribute selectors namely; Information gain ratio attribute evaluator, Principal Component Analysis and SVM attribute selector, from WEKA are employed on the re-sampled dataset.

The ratio of information gain evaluates the worth of an attribute by measuring the information gain with is considered to the class. By setting 4 attributes as a threshold, age, surgical intervention, accident and season attributes are selected.

Principal Component Analysis (PCA) performs a principal components and conversion of the dataset which is used in adequate with a grade search. Dimensionality reduction is performed by selecting adequate eigenvectors to account for some percentage of the variance in the original dataset default 0.95%. Attribute noise is able to filter by converting to the PC space, eliminating some of the worst eigen-vectors, and then converting back to the original space again.

SVM attribute selector interprets the worth of an attribute with an SVM classifier. Attributes are ordered by the square of the weight assigned by the SVM. Attribute selection of multi-class problems is processed by ranking attributes for each class one by one using a one-vs-all method and then "dealing" from the top of each set to grant the last ranking. Thus, four attribute including alcohol, age, accident and high fever are generated.

The three different attribute selectors are compared through intensivee experiments employing the selected

classification algorithms. The attribute selectors reduced the number of attributes from nine to four. For this study, an attribute selector is selected based on its contribution on classification accuracy. Among the three SVM attribute selector has contributed the best for classification accuracy. Therefore, the experimentation ,which is discussed under experimentation and results section, is undertaken after applying SVM attribute selector on the original dataset.

IV. Experimentation and Results

The performance of the classifiers selected for the study are evaluated using classification accuracy, precision, recall, sensitivity, specificity, F-measure and AUC.

$$\text{Accuracy} = (TP+TN)/(TP+FP+FN+TN) \times 100\% \quad (3)$$

$$\text{Sensitivity} = TP/(TP+TN) \times 100\% \quad (4)$$

$$\text{Specificity} = TN/(FP+TN) \times 100\% \quad (5)$$

In the equations, TP defines the number of true positives; FN describes the number of false negatives; TN defines the number of negatives; finally, FP defines the number of false positives.

Receiver operating characteristic (ROC) curves are also used to measure performance of the classifiers. It plans the true positive rate as a function of the false positive rate.

Support Vector Machines (SVM) classifier comprising SVM attribute selector. Applying SVM attribute selector on the dataset attributes: alcohol, high fever, age and accident are selected. Radial Basis Functions (RBF) Kernel function it was adopted which is a useful selection to solve complex problems. The RBF kernel nonlinearly maps are sampled into a higher dimensional space. Thus, in contrast to the linear kernel, it may process a situation when the relation between classes and attributes is non-linear [6].

K-Nearest Neighbor (K-NN) classifier comprising with SVM attribute selector parameters: metrics is Euclidean distance and weighting is based on distance. Experiments involving different values of k (k=3,k=5,k=7 and k=10) has been done. But as depicted in Table III. kNN has shown the best classification accuracy of 87.22% when k=7.

TABLE III: COMPARISON OF CLASSIFIERS PERFORMANCE AFTER APPLYING SVM ATTRİBUTE SELECTOR

Classifier	CA	Sens	Spec	ROC	F1	Prec	Recall
SVM	78.4	79.5	77.4	85.7	79.1	78.6	79.5
NB	90.7	92.2	89.3	94.0	91.0	90.0	92.0
RF	89.5	93.1	85.7	95.3	90.1	87.2	93.2
kNN	87.2	85.2	89.3	94.3	87.2	89.3	85.2
MPN	83.7	84.0	83.3	92.8	84.0	84.0	84.0

The other classifier is Random Forest (RF). For this classifier experiment has been done involving different number of trees =10, 15 and 20. Better perfomance is acheived for number of trees =20. When we used to MLP architecture of neural network, it has performed best accuracy=85.46%, when the numbers of hidden layers is 15, learning rate is 0.1 for maximum iteration as 1000.

However, when the number of iteration is 500 its classification accuracy is 83.73%, which has only 1.73% difference. Since making the iteration 1000 for this narrow difference consumes computation time and memory, the latter classification accuracy (83.73%) preferable.

Among the four classifiers, NB has scored the highest classification accuracy as 90.7% and ROC 94%.

Using the same parameter setting as the above experiments, results of classification algorithms comprising of principal component analysis (PCA) an Information gain attribute selectors are shown in Table VI and Table V.

TABLE IV. CLASSIFIERS PERFORMANCE COMPRISING PC ATTRIBUTE SELECTOR

Classifier	**CA**	**Sens**	**Spec**	**ROC**	**F1**	**Prec**	**Recall**
kNN	77.9	75.0	80.9	87.5	77.7	80.5	75.0
SVM	70.9	71.6	70.2	75.5	71.6	71.6	71.6
Neural Network	85.5	82.9	88.1	89.9	85.4	87.9	82.3
Naive Bayes	69.0	72.7	65.5	76.6	70.7	68.8	72.7
Random Forest	73.7	71.6	76.2	84.5	73.7	75.9	71.6

TABLE V. CLASSIFIERS PERFORMANCE COMPRISING INFORMATION GAIN RATIO

Cassifier	**CA**	**Sens**	**Spec**	**ROC**	**F1**	**Prec**	**Recall**
kNN	85.4	86.4	84.5	95.3	85.9	85.4	86.4
SVM	77.8	78.4	77.4	87.7	78.4	78.41	78.4
Neural Network	80.1	80.6	79.8	87.9	80.7	80.7	80.7
Naive Bayes	86.5	87.5	85.7	92.1	87.0	86.5	87.5
Random Forest	88.3	92.0	84.5	95.9	89.0	86.2	92.0

The SVM attribute selector has contributed the highest in achieving highest classification accuracy. Hence, the classification is undertaken on the dataset containing four attributes (alcohol, age, accident and high fever are generated). 10-fold cross validation is followed for all the classifiers. In addition, 70% of the re-sampled dataset is used to train classifiers and the remaining for testing. The experimentation has done involving K-nearest neighbor, Naive Bayes, Support Vector Machine, Multilayer Perceptron Network and Random Forest classifiers. The performance of the classifiers is evaluated by taking in to account classification accuracy, precision, recall, sensitivity, specificity and f-measure.

As shown in Table III, Naive Bayes achieved classification accuracy of 90.7%, sensitivity 92.2% and specificity 89.3%. Next to Naive Bayes is Random Forest classifier scoring 89.5%% classification accuracy. However, as mentioned in reference [26] in the presence of imbalanced dataset, it is appropriate to use the ROC curve as performance metric.

In this regard, Naive Bayes and Random forest classifiers has registered 94.03% and 95% ROC percentage respectively as compared to the other algorithms. Previous study by [24] using the same dataset (as used in this study) using SVM and MLP classifiers has obtained good accuracies as 86% and 86% respectively as it can be seen in Table VI.

TABLE VI. PREVIOUS RESULTS BY [24] USING THREE CLASSIFIERS

	MLP	SVM	Decision Tree
Classification accuracy %	86	86	84
Sensitivity %	94.1	97.7	96.5
Specificity %	40	20	13.3
Positive predictive value %	89.9	87.4	86.3
Negative predictive value %	54.5	60	40

In addition, specificity 40% for MLP and 20% for SVM are registered. This study has registered a better result in terms of accuracy (90.65%, and specificity (89.3%) as compared to [24]. This is due to re-sampling done using SMOTE technique and feature selection technique employed on the dataset.

V. CONCLUSION

In this paper, different supervised learning classification algorithms such as MLP, k-NN, Naïve Bayes, Random Forest, and SVM are proposed with SMOTE technique employed to re-sample the dataset. Feature selection method (SVM attribute selector) has reduced the dimensionality of the dataset by reducing the number of attributes to four. By employing five different classifiers on the re-sampled and dimensionally were reduced dataset, Naive Bayes classifier has shown the best classification accuracy as 90.7%, and ROC percentage 94% which has obtained better accuracy than previous study given Table VI. Thus, applying SMOTE to re-sample minority class and feature selection techniques enable predict male fertility conditions as nor mal normal or altered.

In the future work, this proposed methods can be applied to big data sets to verify the observed relations.

REFERENCES

[1] World Health Organization, "WHO manual for the standardised investigation and diagnosis of infertile couple.," Campridge University Press, Campridge , 2000.

[2] M. Ramalingam, S. Kini, T. Mahmood, "Male fertility and infertility," Obsterics, gynaecology and reproductive medicine," pp. 326-332, 2014.

[3] A. Oluyemi, A. O. Arowojolu, B. O. Shittu, J. I. Anetor, "Cadmium toxicity: a possible cause of male infertility in Nigeria," Reproductive Bilogy, vol. 6, no. No.1, pp. 17-30, 2006.

[4] A. Oliva, A. Spira, L. Multigner "Contribution of environmental factors to the risk of male infertility," Human Reproduction , vol. 16, no. 8, pp. 1768-1776, 2001.

[5] A. Ferlin, B. Arredi, C. Foresta, "Genetic causes of male infertility," Reproductive toxicology , vol. 22, pp. 133-141, 2006.

[6] Y. Ozbay, B. Karlik, "A Recognition of ECG arrhythmia using artificial neural networks," Proc. of the 23rd Annual International Conference IEEE Engineering in Medicine and Biology Society, 25-28 October, 2001, Istanbul, Turkey

[7] Y. Ozbay, R. Ceylan, B. Karlik, "Integration of type-2 fuzzy clustering and wavelet transform in a neural network based ECG classifier", Expert Systems with Applications, vol. 38, pp. 1004-1010, 2011.

[8] R. Ceylan, Y. Ozbay, B. Karlik, "Comparison of type-2 fuzzy clustering based cascade classifier models for ECG arrhythmias", Biomedical Engineering: Applications, Basis and Communications (BME), vol. 26, no. 6, 2014, 1450075

[9] G. Guclu, B. Karlik, H.R. Oz, "Respiratory sound classification fuzzy c-means, princepal component analysis, and artificial neural networks", 1st Inter. Symposium on Computing in Science & Engineering Bioengineering Congress, 3-5 June, 2010, Kuşadası, Turkey

[10] F. Z. Gogus, B. Karlik, G. Harman, "Classification of Asthmatic Breath Sounds by Using Wavelet Transforms and Neural Networks", International Journal of Signal Processing Systems, vol. 3, no. 2, pp. 106-111, December 2015.

[11] B. Karlik, S. Unlu, "Practical artificial neural networks software for breast cancer diagnosis", Ukrainian Journal of Telemedicine and Medical Telematics, v. 6, 2, pp. 63-68, 2008.

[12] F. Demirezen, B. Karlik, A. Okatan, "Automatic recognition of retinopathy diseases by using wavelet based neural network", The First International Conference on the Applications of Digital Information and Web Technologies, vol. 1-2, pp. 461-464, 2-4 August, 2008, Ostrava, Czech Republic

[13] A. Alkan, Y.G. Sahin, B. Karlik, "A novel mobil epilepsy warning System", Lecture Notes in Artificial Intelligence, 4304, pp. 922–928, Springer-Verlag, 2006

[14] B. Karlik, T.G. Apari, "Development of effective telemedicine software for epileptic seizure recognition", International Journal of Computing & Information Technology, 1(2): pp. 91-99, 2009

[15] B. Karlik and S. Hayta, "Comparison machine learning algorithms for recognition of epileptic seizures in EEG", IWBBIO 2014 (International Work-Conference on Bioinformatics and Biomedical Engineering), 7-9 April 2014, Granada, Spain

[16] B. Karlik, S.S. Cemel, "Diagnosing Diabetes from breath odor using artificial neural networks", Turkiye Klinikleri J Med Sci. vol. 32 (2), pp. 331-336, 2012.

[17] O. Karan, C. Bayraktar, H. Gumuskaya, B. Karlik, "Diagnosing diabetes using neural networks on small mobile devices", Expert Systems with Applications, vol. 39, pp. 54-60, 2012

[18] B. Karlik, "Hepatitis disease diagnosis using backpropagation and the naive bayes classifiers", BURCH Journal of Science and Technology, vol. 1(1), pp. 49-62, 2011

[19] B. Karlik, S. Kul, M. Ozates, "Diagnosis of lumbar disc hernia from images using artificial neural network", 2nd International Conference on Electrical Engineering and Technologies, Hammamet, Tunisia, 3-5 November, 2008

[20] B. Karlik, S. Kul, "Diagnosis of lumbar disc hernia using wavelet transform and neural networks", Ukrainian Journal of Telemedicine and Medical Telematics, vol. 7, No.1, pp. 10-15, 2009

[21] B. Karlik, S. Kul, "Determination lumbar disc hernia MRI with artificial neural networks", the 2nd Workshop on Soft Computing in Image Processing and Computer Vision (SCIPCV), July, 12-15, 2010, Las Vegas, USA

[22] C. Bayraktar, B. Karlik, F. Demirezen, "Automatic diagnosis of otitis media diseases using wavelet based artificial neural networks", 1st Inter. Symposium on Computing in Science & Engineering Bioengineering Congress, 3-5 June, 2010, Kuşadası, Turkey

[23] B. Karlik, A. Ece, "Pediatric Nutritional Requirements: Determination with Neural Networks", Mathematical & Computational Applications, 3, (3), pp. 169-175, 1998

[24] D. Gil, J. L. Girela, J. De Juan, M. J. Gomez-Torres, M. Johnsson, "Predicting seminal quality with artificial intelligence methods," Expert Systems with Applications, no. 39, pp. 12564–12573, 2012.

[25] J. Luengo, A. Fernández, S. García, F. Herrera, "Addressing data complexity for imbalanced data sets: analysis of SMOTE-based oversampling and evoutionaly undersampingl," Soft Computing , no. 15, pp. 1909-1936, 2011.

[26] N. V. Chawla, K. W. Bowyer, L. O. Hall, W. P. Kegelmeyer, "SMOTE-synthetic minority over-sampling technique," Journal of Artificial Intelligence Research , no. 16, pp. 321-357, 2002.

[27] J. Han, M. Kamber, J. Pei, Data Mining Concepts and techniques, 3rd ed., Walthan, MA: Morgan Kaufmann Publishers, 2003.

[28] P. Jeatrakul, K. W. Wong, C. C. Fung, "Classification of imbalanced data by combining the complementary neural network and SMOTE algorithm," ICONIP, vol. 2, pp. 152-159, 2010.

[29] A. T. Azar, H. I. Elshazly, A. E. Hassanien, and A. M. Elkorany, "A random forest classifier for lymph diseases," Computer Methods and Programs in Biomedicine , pp. 465-473, 2014.

Bekir Karlık was born in Karaman, Turkey, in 1964. He received his BS, MS, and PhD degrees from Yildiz Technical University, Istanbul, in 1988, 1991, and 1994 respectively. He has been working as a distinguished professor at Department of Computer Engineering, Engineering Faculty, Selcuk University in Konya, Turkey. His research work focuses on soft computing, machine learning, pattern recognition, biometric recognition, computer vision, biomedical data processing, and telemedicine.

Prof. Karlık authored and co-authored approximately 200 research papers in national or international academic journals and conferences, as well as 2 books and 3 edited books. He has been editor-in-chief and of advisory board of 2 journals, the editorial board member of 9, and reviewers of 41 journals. He has supervised 13 graduate students, and 3 PhD students, the program for those 9 of who are on-going now.

He is recipient of several awards, and referee or reviewer for more than hundreds journals and numerous conferences. He also served on dean, head of departments, and several faculty and institute committees. He served Honorary Chair, Co-chair, and Session Chair, as well as TPC member or Keynote Speaker for a number of conferences. He is senior member of IACSIT and member of Association for Scientific Research.

Abulkerim M. Yibre was born in Tissabay, Ethiopia, in 1984. He received his BS and MS degrees from Addis Ababa University,Ethiopia in 2006 and 2013 respectively. He worked as a Lecturer at Department of Computer Engineering, Bahir Dar Institute of Technology(BiT) , Bahir Dar University in Bahir Dar,Ethiopia . His research work focuses on machine learning,pattern recognition and Knowledge Based systems. He is currently a PhD student in Selcuk University in Konya, Turkey.

Barış Koçer was born in Ordu, Turkey, in 1981. He received his BS, MS, and PhD degrees from Selçuk University, Konya, in 2003, 2006, and 2012 respectively. He has been working as a assistan professor at Department of Computer Engineering, Engineering Faculty, Selcuk University in Konya, Turkey. His is interested in machine learning, transfer learning, optimization.

Dr. Koçer authored and co-authored approximately 30 research papers in national or international academic journals and conferences. He has been reviewers of journal of knowledge based systems.