Personalized Cancer Treatment by Using Naive Bayes Classifier

Bekir Karlık and Emre Öztoprak, *Member, IACSIT*

***Abstract*— Pharmacogenomics is the application of genomic technologies to drug discovery and development, as well as for the elucidation of the mechanisms of drug action on cells and organisms. DNA microarrays measure genome-wide gene expression patterns and are an important tool for pharmacogenomics applications, such as the identification of molecular targets for drugs, toxicological studies and molecular diagnostics. DNA based classification is difficult because the sequence of DNA differs from among all individuals. Genome-wide investigations generate vast amounts of data and there is a need for soft computational methods such as artificial intelligence and expert systems to manage and analyze this information. This study presents a Naive Bayes classifier to a novel approach on application of pharmacogenetics to personalized cancer treatment using data of TPMT polymorphisms.**

***Index Terms*— Personalized medicine, pharmacogenetics; TPMT polymorphism; Naïve Bayes classifier; cancer treatment.**

I. Introduction

Personalized medicine (pharmacogenetics) is an extension of traditional approaches to understanding and treating illness. Since the beginning of the study of medicine, physicians have employed evidence found through observation to make a diagnosis or to prescribe treatment. In the past, this was presumably tailored to each individual, but personalized medicine makes treatment more specific. Pharmacogenomics aims to understand pharmacological response with respect to genetic variation. Essential to the delivery of better health care is the use of pharmacogenomics knowledge to answer questions about therapeutic, pharmacological or genetic aspects [1]. As pharmacogenetics researchers gather more detailed and complex data on gene polymorphisms that effect drug metabolizing enzymes, drug target receptors and drug transporters, they will need access to advanced statistical tools to mine that data.

Recently, soft computational methods such as artificial intelligence (AI), artificial neural networks (ANN), and genetic algorithms are used for Genome-wide investigations generate vast amounts of data and analyze of this information [2]. These methods include approaches from classical biostatistics, such as logistic regression or linear discriminant analysis, and supervised learning methods from computer science, such as support vector machines and artificial neural networks [3]. Warnecke-Eberz U. et al. (2010) used TaqMan low-density arrays and analysis by ANN predicts response to neoadjuvant chemo radiation in esophageal cancer. Neoadjuvant radio chemotherapy of locally advanced esophageal cancer is only effective for patients with major histopathological response. A total of 17 genes were selected to predict histopathologic tumor response to chemo radiation (cisplatin, 5-fluorouracil, 36 Gy).For gene-expression analysis quantitative TaqMan low-density arrays was applied. Expression levels in pretreatment biopsies of 41 patients (cT2-4, Nx, M0) were compared with the degree of histopathologic regression in resected specimens applying univariate, multivariate and artificial neuronal network analyses. Multivariate analysis of the marker combination provided response prediction with 75.0% sensitivity, 81.0% specificity and 78.1% accuracy. Artificial neuronal network analysis was the best predictive model for major histopathologic response (80% sensitivity, 90.5% specificity and 85.4% accuracy), representing a clinically practical system. Low-density-array RT-PCR analyzed by artificial neuronal network predicts histopathologic response to neoadjuvant chemoradiation in their patient collective, and could be used to further individualize treatment strategies in esophageal cancer [4]. Lin E. Et al. (2006) demonstrated that a trained artificial neural network model is a promising method for providing the inference from factors such as single nucleotide polymorphisms, viral genotype, viral load, age and gender to the responsiveness of interferon [5].

ANN is an adaptable system that can learn input-output relationships through repeated presentation of data and is capable of generalizing to new, previously unseen data. It can be trained by submitting several sets of input data with their associated output, and during the training, the network learns to associate particular sets of inputs with particular outputs by adapting its free parameters [6]. Chao-Cheng Lin et al. (2008) aimed to train and validate artificial neural networks (ANN), using clinical and pharmacogenetic data, to predict clozapine response in schizophrenic patients. Five pharmacogenetic variables and five clinical variables were collated from 93 schizophrenic patients taking clozapine, including 26 responders. ANN analysis was carried out by training the network with data from 75% of cases and subsequently testing with data from 25% of unseen cases to determine the optimal ANN architecture. Then the leave-one-out method was used to examine the generalization of the models. The optimal ANN architecture was found to be a standard feed-forward, fully-connected, back-propagation multilayer perceptron. The overall

Manuscript received April 11, 2012; revised April 16, 2012.

B. Karlık is with the Department of Computer Engineering, Engineering Faulty, Mevlana University, 42003, Selçuklu-Konya, Turkey (e-mail: bkarlik@mevlana.edu.tr).

E. Öztoprak is with the Medicine Faculty, Mevlana University, 42003, Selçuklu-Konya, Turkey (e-mail: eoztoprak@mevlana.edu.tr).

accuracy rate of ANN was 83.3%, which is higher than that of logistic regression (LR) (70.8%). By using the area under the receiver operating characteristics curve as a measure of performance, the ANN outperformed the LR (0.821±0.054 versus 0.579±0.068; $p < 0.001$). The ANN with only genetic variables outperformed the ANN with only clinical variables (0.805±0.056 versus 0.647±0.066; $p = 0.046$). The gene polymorphisms should play an important role in the prediction [7].

Serretti and Smeraldi (2004) tested a neural network strategy for a combined analysis of two gene polymorphisms. A Multi Layered Perceptron (MLP) model showed the best performance and was therefore selected over the other networks. One hundred and twenty one depressed inpatients treated with fluvoxamine in the context of previously reported pharmacogenetic studies were included. The polymorphism in the transcriptional control region upstream of the 5HTT coding sequence (SERTPR) and in the Tryptophan Hydroxylase (TPH) gene was analyzed simultaneously. A MLP network composed by 1 hidden layer with 7 nodes was chosen. 77.5 % of responders and 51.2% of non-responders were correctly classified (ROC area = 0.731 – empirical p value = 0.0082). Finally, they performed a comparison with traditional techniques. A discriminant function analysis correctly classified 34.1 % of responders and 68.1 % of non-responders ($F = 8.16$ $p = 0.0005$).Overall, their findings suggest that neural networks may be a valid technique for the analysis of gene polymorphisms in pharmacogenetic studies. The complex interactions modeled through NN may be eventually applied at the clinical level for the individualized therapy [8].

Cosgun E. et al. (2011) have applied three machine learning approaches: Random Forest Regression (RFR), Boosted Regression Tree (BRT) and Support Vector Regression (SVR) to the prediction of warfarin maintenance dose in a cohort of African Americans. They have developed a multi-step approach that selects SNPs, builds prediction models with different subsets of selected SNPs along with known associated genetic and environmental variables and tests the discovered models in a cross-validation framework. Preliminary results indicate that their modeling approach gives much higher accuracy than previous models for warfarin dose prediction. A model size of 200 SNPs (in addition to the known genetic and environmental variables) gives the best accuracy. The R(2) between the predicted and actual square root of warfarin dose in this model was on average 66.4% for RFR, 57.8% for SVR and 56.9% for BRT. Thus RFR had the best accuracy, but all three techniques achieved better performance than the current published R(2) of 43% in a sample of mixed ethnicity, and 27% in an African American sample. At the end; they say that machine learning approaches for high-dimensional pharmacogenetic prediction, and for prediction of clinical continuous traits of interest, hold great promise and warrant further research [9].

E. Himes et al. (2009) sought to relate candidate gene SNP data with bronchodilator response and measure the predictive accuracy of a model constructed with genetic variants. Bayesian networks, multivariate models that are able to account for simultaneous associations and interactions among variables, were used to create a predictive model of bronchodilator response using candidate gene SNP data from 308 Childhood Asthma Management Program Caucasian subjects. The model found that 15 SNPs in 15 genes predict bronchodilator response with fair accuracy, as established by a fivefold cross-validation area under the receiver-operating characteristic curve of 0.75 (standard error: 0.03). Bayesian networks are an attractive approach to analyze large-scale pharmacogenetic SNP data because of their ability to automatically learn complex models that can be used for the prediction and discovery of novel biological hypotheses [10].

Larder B. Et al. (2008) describe that the development and application of ANN models as alternative tools for the interpretation of HIV genotypic drug resistance data. A large amount of clinical and virological data, from around 30,000 patients treated with antiretroviral drugs, has been collected by the HIV Resistance Response Database Initiative (RDI, www.hivrdi.org) in a centralized database. Treatment change episodes (TCEs) have been extracted from these data and used along with HIV drug resistance mutations as the basic input variables to train ANN models. They performed a series of analyses that have helped define the following:

- The reliability of ANN predictions for HIV patients receiving routine clinical care;
- The utility of ANN models to identify effective treatments for patients failing therapy;
- Strategies to increase the accuracy of ANN predictions; and
- Performance of ANN models in comparison to the rules-based methods currently in use [11].

Sabbagh and Darlu (2006) investigated the ability of several pattern recognition methods to identify the most informative markers in the CYP2D6 gene for the prediction of CYP2D6 metabolizer status. Four data-mining tools were explored: decision trees, random forests, artificial neural networks, and the multifactor dimensionality reduction (MDR) method. Marker selection was performed separately in eight population samples of different ethnic origin to evaluate to what extent the most informative markers differ across ethnic groups. Their results show that the number of polymorphisms required predicting CYP2D6 metabolic phenotype with a high accuracy can be dramatically reduced owing to the strong haplotype block structure observed at CYP2D6. MDR and neural networks provided nearly identical results and performed the best. Data-mining methods, such as MDR and neural networks, appear as promising tools to improve the efficiency of genotyping tests in pharmacogenetics with the ultimate goal of pre-screening patients for individual therapy selection with minimum genotyping effort [6].

Pharmacogenomics is a new field which uses genetic information to estimate drug treatment response. The person's drug-therapeutic or toxic molecular genetic basis of response to context clarification on the new drugs and genes engaged in the discovery of the target points to a branch of sciences. Currently, there are only a few pharmacogenetic diagnostic tests available, and clinical guidelines for pharmacogenetically tailored therapy are lacking [12]. In clinical pharmacology, detailed data about the complex molecular mechanisms of the interactions between drug(s) and organism become available. Most notably, the target

genes of many drugs are being discovered and the differential genes expression induced by drugs can be investigated by microarray techniques. However, genetic variation can account for as much as 95 percent of variability in drug disposition and effects [13].

One of the best-developed examples of pharmacogenetics applied to clinical practice is the enzyme thiopurine methyltransferase (TPMT). TPMT is responsible for the degradation of azathioprine and mercaptopurine, which are commonly used to treat acute leukemia, inflammatory bowel disease, rheumatoid arthritis, and transplant immune suppression. Drug metabolizing enzymes participate in the neutralizing of xenobiotic and biotransformation of drugs. Polymorphisms in the drug-metabolizing enzyme coding genes alter the activity of these enzymes for some substrates. Thiopurine S-methyltransferase (TPMT) is a cytosolic enzyme that catalyzes the S-methylation of aromatic and heterocyclic sulfhydryl compounds like 6-mercaptopurine (6MP), which is used to treat patients with acute lymphoblastic leukemia (ALL). Polymorphisms in the genes encoding cytochrome p450 (CYP) and thiopurine S-methyl transferase (TPMT) enzymes catalyze the metabolic reactions of several drugs. These polymorphisms might be responsible for adverse drug reactions. TPMT activity is related to the outcome and/or toxicity of therapy. Patients with inherited very low levels of TPMT activity are at increased risk for thiopurine-induced toxicity, when treated with standard doses of these drugs [14].

In this study, we proposed a Naïve Bayesian classifier learning for automatic detection of cancer, cancer risk, and, not-risk by using pharmacogenomics data which provides to describe data of TPMT polymorphisms. Bayes classifiers can be trained very efficiently in a supervised learning setting. We consider the relationship between supervised learning, or function approximation problems, and Bayesian reasoning. We begin by considering how to design learning algorithms based on Bayes rule. The codes of the pharmagenucleotides in the genes are used letters of alphabet such as A, C, G, T which are not numerical values. To prevent this handicap the ASCII codes of these letters can be used as input values of Naïve Bayes classifier.

II. Material and Methods

The Bayesian Classification represents a supervised learning method as well as a statistical method for classification. Assumes an underlying probabilistic model and it allows us to capture uncertainty about the model in a principled way by determining probabilities of the outcomes. It can solve diagnostic and predictive problems. Bayesian classification provides practical learning algorithms and prior knowledge and observed data can be combined. Bayesian Classification provides a useful perspective for understanding and evaluating many learning algorithms. It calculates explicit probabilities for hypothesis and it is robust to noise in input data [18]. Depending on the precise nature of the probability model, naive Bayes classifiers can be trained very efficiently in a supervised learning setting. In many practical applications, parameter estimation for naive Bayes models uses the method of maximum likelihood; in other words, one can work with the naive Bayes model without believing in Bayesian probability or using any Bayesian methods.

The Naive Bayes Classifier technique is based on the so-called Bayesian theorem and is particularly suited when the dimensionality of the inputs is high. Despite its simplicity, Naive Bayes can often outperform more sophisticated classification methods. A naive Bayes classifier is a simple probabilistic classifier based on applying Bayes' theorem with strong (naive) independence assumptions. A more descriptive term for the underlying probability model would be independent feature model. Depending on the precise nature of the probability model, naive Bayes classifiers can be trained very efficiently in a supervised learning setting [15].

The Naive Bayes classifier is designed for use when features are independent of one another within each class, but it appears to work well in practice even when that independence assumption is not valid [16]. It classifies data in two steps:

- Training step: Using the training samples, the method estimates the parameters of a probability distribution, assuming features are conditionally independent given the class.
- Prediction step: For any unseen test sample, the method computes the posterior probability of that sample belonging to each class. The method then classifies the test sample according the largest posterior probability.

The class-conditional independence assumption greatly simplifies the training step since you can estimate the one-dimensional class-conditional density for each feature individually. While the class-conditional independence between features is not true in general, research shows that this optimistic assumption works well in practice. This assumption of class independence allows the Naive Bayes classifier to better estimate the parameters required for accurate classification while using less training data than many other classifiers. This makes it particularly effective for datasets containing many predictors or features.

In spite of their naive design and apparently over-simplified assumptions, naive Bayes classifiers often work much better in many complex real-world situations than one might expect. Recently, careful analysis of the Bayesian classification problem has shown that there are some theoretical reasons for the apparently unreasonable efficacy of naive Bayes classifiers. An advantage of the Naive Bayes classifier is that it requires a small amount of training data to estimate the parameters (means and variances of the variables) necessary for classification. Because independent variables are assumed, only the variances of the variables for each class need to be determined and not the entire covariance matrix [17].

Consider a supervised learning problem in which we wish to approximate an unknown target function f: X Y, or equivalently P(Y|X). Applying Bayes rule, we can see that P(Y= yi|X) can be represented as

$$P(Y=y_i|X=x_k) = [P(X=x_k|Y=y_i)P(Y=y_i)]/[\sum_j P(X=x_k|Y=y_j)P(Y=y_j)]$$

Where y_m denotes the mth possible value for Y, x_k denotes the kth possible vector value for X, and where the summation in the denominator is over all legal values of the random variable Y. One way to learn P(Y|X) is to use the training data to estimate P(X|Y) and P(Y). We can then use these estimates, together with Bayes rule above, to determine $P(Y|X = x_k)$ for any new instance x_k [18]. The main difference between Logistic Regression and Naïve Bayes: Logistic Regression directly estimates the parameters of P(Y|X), whereas Naive Bayes directly estimates parameters for P(Y) and P(X|Y). We often call the former a discriminative classifier, and the latter a generative classifier [19-20].

In this study, used TPMT data was collected from 4 healthy persons and 8 leukemia patients by Fatih University, Department of Biology in Istanbul in Turkey. Genomics DNA was extracted from peripheral blood by using proteinase K/salting out method. Here, the codes of the nucleotides in the genes are used letters of alphabet such as A, C, G, T which are not numerical values. To prevent this handicap the ASCII codes of these letters will be used as input values. Each class of output determines Cancer (continue to drugs), Cancer (discontinue to drugs), and No cancer (no drug). All of these numbers of nucleotide is used as inputs of Bayesian classifier and all of these codes have to be converted to numerical values.

There are some kinds of computer programs about Bayes classification and RapidMiner is one of that programs. The usage of that program is very simple. RapidMiner can make classifications by using different methods. So RapidMiner was used for Bayes classification. RapidMiner (formerly YALE) is the most comprehensive open-source software for intelligent data analysis, data mining, knowledge discovery, machine learning, predictive analytics, forecasting, and analytics in business intelligence [17]. Working the program, we put our training data to application by from that Example Source option (see Fig.1).

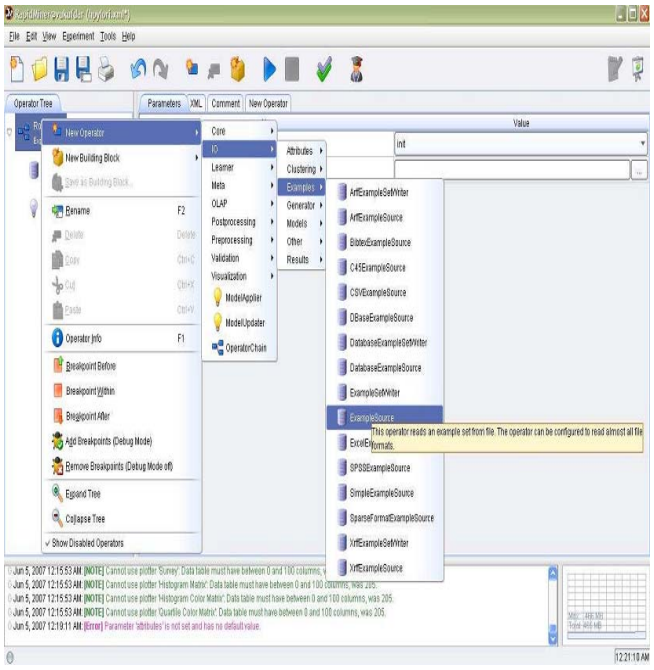


Fig. 1. Example source

Next step is applying the classification algorithm to that selected example source so; we select a learner for the program. This learner is the algorithm of the training data. On the "New Operator" panel, we select "Learner", than we choose the learning type. For our classification we choose supervised learning. We give our test data for the program and this test data is given as input-output operation. For these example sets which we selected for the program, need a Model Applier to apply training TPMT data and testing TMPT data. Testing data was recorded from 8 different persons.

On the "New Operator" panel, we select "Learner", than we choose the learning type. For our classification we choose Supervised learning. For our classification we choose Supervised learning. Than we choose the Bayes option from the Weka panel and after that W-Naive Bayes is selected as shown in Fig. 2. Example sets which is selected for the program, need a Model Applier to apply training data and test data. That Model Applier combines all of the data which we give to the program and it gives us the results about out test data.

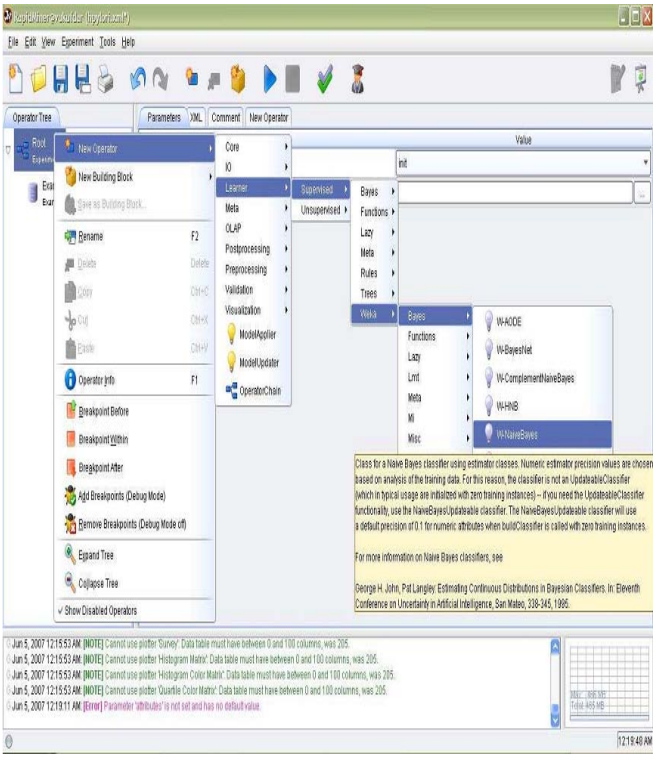


Fig. 2. Selection of W-Naive Bayes

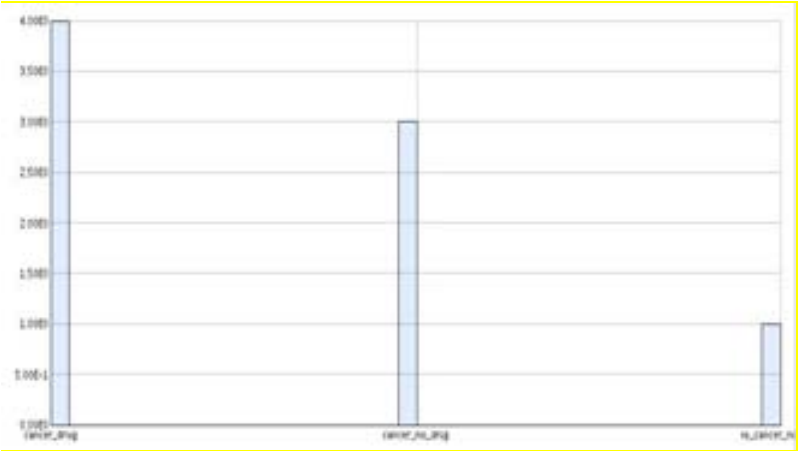
After selecting the Model Applier, program gives us the results about the test data. Fig. 3 (a) shows RapidMiner program outputs for the testing. Here each class output has been described by logical 1. So that logic 0 describe the other situation which it doesn't belongs to this class.

Example set consists of 8 examples of test data which includes 5 special attributes and 31 regular attributes. Fig. 3 (b) shows the test results of 8 different persons. According the Fig. 3 (b), we obtained 4 persons as Leukemia cancer which they need to continue to drugs such as azathioprine and mercaptopurine (see left your hand side). 3 persons have

Leukemia cancer which they don't need to continue drugs which is described middle of the bar. Only one of them has no cancer problem and no needs to take any drug which is described right side of the bar.

ro...	prediction(Cancer)	confidence(cancer_drug)	confidence(cancer_no_drug)	confidence(no_cancer_no_drug)	c...	testing.data...	testing.data...	testing.data...
1	cancer_drug	1	0	0	0	0.840	0.840	0.840
4	cancer_drug	1	0	0	0	0.840	0.840	0.840
5	cancer_drug	1	0	0	0	0.710	0.710	0.840
6	cancer_drug	1	0	0	0	0.650	0.840	0.840
2	cancer_no_drug	0	1	0	0	0.670	0.650	0.670
3	no_cancer_no_drug	0	0	1	0	0.840	0.840	0.710
7	cancer_no_drug	0	1	0	0	0.710	0.670	0.710
8	cancer_no_drug	0	1	0	0	0.670	0.650	0.670

(a)



(b)

Fig. 3. (a) The program outputs (b) The test results

III. Conclusion

One of the main factors preventing a more efficient use of new pharmacological treatments for chronic diseases like for example hypertension, cancer, Alzheimer disease or obesity is represented by the difficulty of predicting "a priori" the chance of response of the single patient to a specific drug. This study presents detection of drug therapy of leukemia cancer by using Naïve Bayes Classifier Proposed study supports use of personalized drug therapy in clinical practices. This tool can be used for treatment of variety diseases with similar characteristics. This method may provide tools for clinical association studies and help find genetic TPMT (or SNPs) involved in responses to therapeutic drugs or adverse drug reactions. Moreover, this preliminary study can be improved and applied to solve similar treatment problems.

In future work, we can compare these two methods with the other supervised or unsupervised classifiers methods such as neural networks, fuzzy classifier, Support Vector Machines (SVM), Self-Organization Future Maps (SOM) etc.

References

[1] M. Dumontier and N. Villanueva-Rosales, "Towards pharmacogenomics knowledge discovery with the semantic web," *Briefings in Bioinformatics*, vol. 10(2), pp. 153-163, March 2009.

[2] M. Ringnér, C. Peterson, J. Khan, "Analyzing array data using supervised methods," *Pharmacogenomics*, vol. 3(3), pp. 403-15, 2002.

[3] W. Shannon, R. Culverhouse, J. Duncan, "Analyzing microarray data using cluster analysis," *Pharmacogenomics*, vol. 4(1), pp. 41-52, 2003.

[4] U. Warnecke-Eberz, R. Metzger, E. Bollschweiler, S.E. Baldus, R. P. Mueller, H. P. Dienes, A. H. Hoelscher, P. M. Schneider , "TaqMan low-density arrays and analysis by artificial neuronal networks predict response to neoadjuvant chemoradiation in esophageal cancer,"*Pharmacogenomics*, vol. 11(1), pp. 55-64, 2010.

[5] E. Lin, Y. Hwang, S. C. Wang, Z.J. Gu, E. Y. Chen, "An artificial neural network approach to the drug efficacy of interferon treatments,"*Pharmacogenomics,* vol. 7(7), pp. 1017-24, 2006.

[6] A. Sabbagh and P. Darlu, "Data-Mining Methods as Useful Tools for Predicting Individual Drug Response: Application to CYP2D6 Data," *Hum Hered*. vol. 62, pp. 119–134, 2006.

[7] C.C. Lin, Y.C. Wang, J. Y. Chen, Y.J. Liou, Y. M. Bai, I. C. Lai, T.T. Chen, H. W. Chiu, Y. C. Li, "Artificial neural network prediction of clozapine response with combined pharmacogenetic and clinical data," *Computer Methods and Programs in Biomedicine*, vol.91,pp.91–99, 2008.

[8] A. Serretti and E. Smeraldi, "Neural network analysis in pharmacogenetics of mood disorders," *BMC Medical Genetics*, vol. 5(27), pp. 1-6, 2004.

[9] E. Cosgun, N. A. Limdi, C. W. Duarte, "High-dimensional pharmacogenetic prediction of a continuous trait using machine learning techniques with application to warfarin dose prediction in African Americans," *Bioinformatics*, vol. 27(10), pp. 1384-1489, May 2011.
[10] E. Himes, A. C. Wu, Q. L. Duan, B. Klanderman, A. A. Litonjua, K. Tantisira, M. F. Ramoni, and S. T. Weiss, "Predicting response to short-acting bronchodilator medication using Bayesian networks," *Pharmacogenomics*, vol. 10(9), pp. 1393–1412, September 2009.
[11] Larder, D. Wang, A. Revell, "Application of artificial neural networks for decision support in medicine," *Methods Mol Biol*., vol. 458, pp. 23-36, 2008.
[12] E. Lanfear, H. L. McLead, "Pharmacogenetics: using DNA to optimize drug therapy," *American Family Physician*, vol. 76(8), pp. 1179-1182, 2007.
[13] G. Floares, "Using computational intelligence to develop intelligent clinical decision support systems," in *Proc. 6th International Conference on Computational Intelligence Methods for Bioinformatics and Biostatistics*, Springer-Verlag Berlin, 2010, pp. 266-275.
[14] S. Zhou, "Clinical Pharmacogenomics of Thiopurine S-methyltransferase," *Current Clinical Pharmacology*, vol. 1, pp. 119-128, 2006.
[15] M. M. Ameyaw, E. S. Collie-Duguid, R. H. Powrie, et al. "Thiopurine methyltransferase alleles in British and Ghanaian populations," *Hum Mol Genet.*, vol. 8, pp. 367-370, 1999.
[16] M. A. Sayitoglu, I. Yildiz., O. Hatirnaz, U. Ozbek, "Common Cytochrome p4503A (CYP3A4 and CYP3A5) and Thiopurine S-Methyl Transferase (TPMT) Polymorphisms in Turkish Population," *Turk J Med Sci*., vol. 36, pp. 11-15, 2006.
[17] F. Jensen, An Introduction to Bayesian Networks, Springer-Verlag, 1996.
[18] B. Karlik, A. Avci, A. T. Yabanigul, "Classification of Helicobacter Pylori According to National Strains Using Bayesian Learning,"*Mathematical & Computational Applications*, vol. 14(3), pp. 241-251, 2009.
[19] T. Mitchell, *Machine Learning*, McGraw Hill, 1997.
[20] A.Y. Ng, and M. I. Jordan, "On Discriminative vs. Generative Classifiers: A comparison of Logistic Regression and Naive Bayes," in *Advances in Neural Information Processing Systems*. Cambridge, MA: MIT Press, 14, 2002, pp. 609-616.

Prof. Dr. Bekir Karlık received his BS, MS, and PhD degrees from Yıldız Technical University respectively. He has been working as a dean of the Engineering Faculty at Mevlana University in Konya, Turkey. His research interests include artificial intelligence, artificial neural networks, machine learning, pattern recognition, computer vision, and telemedicine. Prof. Karlık has published over 180 papers in many well-known international journals and proceedings of refereed conference since 1991, and many among these papers have been indexed by SCI Expanded. He has been editor-in-chief of two international journals, the editor of 7 journals, and reviewers of 25 international journals.Prof. Karlık is senior member of IACSIT (Member NO: 80343985), members of Association for Scientific Research, and International Intelligent Knowledge Systems Society. He has numerous Scientific Published Awards by TUBITAK, Fatih University, and Mevlana Universty.

Emre Öztoprak received his Doctor of Medicine degree from Ege University in Izmir in Turkey. He has been studying on PhD at Department of Pharmacology, Faculty of Medicine, Konya University. He is a lecturer in Faculty of Medicine, Mevlana University, Konya, Turkey. He is interested in pharmacogenetics, personalized medicine, pharmaceutical production.

He has awards such as Second Regional Science Olympiad, Cross-Level Examination First Rank in Siirt Province Secondary Schools, and Secondary High School Championship in Siirt.