Information Technology Journal, 2012
ISSN 1812-5638 / DOI: 10.3923/itj.2012.

Medical Data Classification with Naïve Bayes Approach

K.M. Al-Aidaroos, A.A. Bakar and Z. Othman
Center for Artificial Intelligence Technology, Faculty of Information and Science Technology,
Universiti Kebangsaan Malaysia, Selangor, Malaysia

Abstract: Medical area produces increasingly voluminous amounts of electronic data which are becoming more complicated. The produced medical data have certain characteristics that make their analysis very challenging and attractive. In this study we present an overview of medical data mining from different perspectives; including characteristics of medical data, requirements of systems dealing with such data and the different techniques used for medical data mining. Among the different approaches we emphasize on the use of Naïve Bayes (NB) which is one of the most effective and efficient classification algorithms and has been successfully applied to many medical problems. To support our argument, empirical comparison of NB versus five popular classifiers on 15 medical data sets, shows that NB is well suited for medical application and has high performance in most of the examined medical problems.

Key words: Data mining, classification, medical data, naïve bayes

INTRODUCTION

Nowadays modern hospitals are well equipped with monitoring and other data collection devices resulting in enormous data which are collected continuously through health examination and medical treatment. All this led to the fact that medical area produces increasingly voluminous amounts of electronic data which are becoming more complicated.

In the past, various statistical methods have been used for modeling in the area of disease diagnosis. These methods require prior assumptions and are less capable of dealing with massive and complicated nonlinear and dependent data (Lin, 2009). However, data mining has proven to be more powerful and effective and it provides processes for discovering useful patterns from large data sets (Thongkam *et al*., 2008). These data mining techniques are generally classified into supervised and unsupervised models. Clustering techniques which are unsupervised learning, have emerged as popular techniques for pattern recognition and image processing (Abed and Zaoui, 2011; Velmurugan and Santhanam, 2011) and have also been applied to problems with medical data (Kittaneh *et al*., 2012). However, in this paper, we are concerned with predictive (i.e., supervised) methods which require the data to include a special response attribute, known as the class attribute and therefore known as classification models.

The importance of Medical Data Mining (MDM) is to assist the physician to make the final decision without hesitation, minimizing diagnostic errors (especially from inexperienced physicians), improving diagnostic speed and increasing the quality of medical treatment (Maria, 2002; Bai and Srivatsa, 2006; Lin, 2009; Temurtas *et al*., 2009).

In this study we review MDM from different perspectives. We start with highlighting the special characteristics of medical data and discussing the requirements of data mining systems to cope with medical data problems and difficulties. We present a review of some of those proposed methods in the medical domain to show what are the different techniques and methods which have been applied to medical data.

Among the different techniques used in MDM we concentrate on Naïve Bayes approach. We discuss its features and justify why it is suited for application in the medical domain, supporting our discussion with successful medical applications. Based on the evidence we found from the literature, we conducted an empirical comparison of NB against five popular classifiers which represent different learning approaches and their results were analyzed.

BACKGROUND OF MEDICAL DATA MINING

Characteristics of medical data: The data gathered in medicine is generally collected as a result of patient-care activity to benefit the individual patient and research is only a secondary consideration. As a result, medical data contain many features that create problems for the data

Corresponding Author: K.M. Al-Aidaroos, Data Mining and Optimization Research Group,
Center for Artificial Intelligence Tehnology, 43600 UKM Bangi, Selangor, Malaysia

mining techniques and they might be in a format which is not suitable for the direct application of those techniques (Delen *et al*., 2005; Thongkam *et al*., 2009).

In general, medical collections, diagnoses and treatments are subject to error rates, imprecision and uncertainty (Pattaraintakorn *et al*., 2005). As with any large databases and due to the collection method, medical databases may contain missing values and can introduce noisy, redundant, incomplete or inconsistent data (Delen *et al*., 2005).

In a detailed discussion of the main differences of data mining in medicine from that in other fields, Cios and Moore (2002) discussed four major points about the uniqueness of medical data. First point is the heterogeneity and complexity which is a result of medical data being collected from various images, interviews with the patient, laboratory data and the physician's observations and interpretations. Second point is about the special ethical, legal and social constraints which relate to privacy and security considerations, fear of lawsuits or possible injury to the patient. Statistical philosophy is the third point and this is because of poor physical formulae or equations for characterizing medical data and the violation of statistical assumptions in medical data. Finally is the special status of medicine itself, because outcomes of medical care are life-or-death and they apply to everybody.

Requirements for systems dealing with medical data: For a data mining system to be useful in solving medical problems, the following features are desired:

- **Handling missing values and noisy data:** In real medical data sets, missing values are frequently present and most patients' records lack certain data. This can be a result of certain tests not performed or certain questions that were not asked (Kononenko, 2001; Cios and Moore, 2002; Bellazzi and Zupan, 2008). Therefore, medical mining systems have to be able to appropriately deal with such incompleteness of the data. Some data mining approaches are robust to missing values while other approaches deal with this requirement through preprocessing of the data. In addition to missing values, medical data are characterized by their incorrectness, inconsistency, redundancy, sparseness and inexactness. For this reason, in most cases, a robust data preprocessing system is required in order to draw any kind of knowledge from even medium-sized medical data sets (Lavrac, 1999; Sumathi and Sivanandam, 2006)
- **High performance and efficiency of the produced model:** For a medical diagnostic system to be accepted by the user, its accuracy must be as high as possible. In most cases several approaches are tested on the available data and the one with best performance is considered. However, for small differences in predictive performance it might be necessary to take into account other features for selecting the appropriate method (Kononenko, 2001; Ohno-Mochado, 2001). Efficiency of the data mining method used is also important, because the final application is a user interactive and for many optimal solutions they are usually time consuming (Li *et al*., 2005)
- **Transparency of the model:** Data mining techniques differ in their degree of transparency, i.e., the users' ability to analyze and understand how the patterns were generated. For some techniques which are considered as "black boxes", their results may not be accepted by the end user, especially when producing unexpected solution (Lavrac, 1999). In medical applications the user should be able to use the model's logic to explain how the conclusion was reached which may significantly increase a physician's confidence in the model (Bellazzi and Zupan, 2008)
- **Interpretability and understandability of results:** Interpretability and acceptability by the medical community intervene in favour of a method that may not have the highest predictive performance (Ohno-Mochado, 2001). In general, users do not care how sophisticated a data mining method is but they do care how understandable its results are (Li *et al*., 2005). It is crucial for a medical diagnosis system to be able to explain and justify its decisions when diagnosing a new patient (Lavrac, 1999)
- **Reduction of the number of tests and generalization:** Since the collection of medical data is sometimes expensive and harmful for the patients, it is desirable to have a system that is able to reliably diagnose with a small amount of data (Kononenko, 2001). However this should not result in overfitting situations and the produced model must be able to perform well with unseen cases (Bellazzi and Zupan, 2008)
- **Protecting the privacy of data:** When dealing with medical data it is important to protect the privacy and sensitive information from disclosure and to identify possible ways to have secure channels for transfering medical data (Chen *et al*., 2010; Nabi *et al*., 2010)

Techniques and methods used in medical data mining: The increasing availability of various data mining methods and tools requires medical informatics researchers and practitioners to systematically select the

most appropriate strategy to cope with clinical prediction problems. In particular, the mechanisms that make them better suited for the analysis of medical databases based on the discussed characteristics and requirements of medical data mining.

By reviewing the literature on medical data mining, we can find various techniques applied to a variety of medical problems with varying degrees of success. Some of these applications are particular and involve individual learning technique while others integrate/hybridize two or more techniques to enhance the resulting model. In the following we discuss these techniques and their applications.

Soft computing methods have been widely used for medical data mining and proven to be well suited to cope with the special characteristics of medical data such as imprecision and uncertainty. For example: Fuzzy Logic (John and Innocent, 2005; Blessia *et al*., 2011), Rough Sets (Tsumoto, 1999; Hassanien *et al*., 2008), Genetic Algorithms (Laurikkala *et al*., 1999; Goletsis *et al*., 2004) and Neural Networks (Yao and Liu, 1999; Lisboa, 2002).

Statistical methods have been considered, by many researchers, less capable of dealing with massive, non-linear and dependent data (such as the health care data). However some predictive statistical approaches such as the proposed model by Cong and Tsokos (2010), the k-Nearest Neighbour (k-NN) (Dzeroski and Lavrac, 1996), Logistic Regression (LR) (Kuhnert *et al*., 2000) and Bayesian Classifiers (Kononenko *et al*.,1998; Tsymbal and Puuronen, 2002), have been successfully applied to medical data; especially the naïve bayes method which we will discuss in more detail in the following section.

Decision Tree (DT) algorithm is one of the most popular classification algorithms used for data mining. It has been applied to medical data providing competitive performance as compared to other approaches as discussed by Delen *et al*. (2005) and Kuo *et al*. (2001).

Agent-based systems and artificial immune systems (AISs) have been also applied to medical problems. Examples of their use for medical applications are given by Lanzola *et al*. (1999), Hudson and Cohen (2002), Polat *et al*. (2005) and Latifoglu *et al*. (2008).

Recently, the need of a hybrid data mining approach is widely recognized by the data mining community and much current work in data mining tends to hybridize diverse methods (Hassan and Verma, 2007). In Medical domain there are a lot of hybrid models which have been proposed, such as Evolutionary decision tree (Podgorelec and Kokol, 2001), Polynomial Fuzzy DT (Mugambi *et al*., 2004), ANN with MARS (Multivariate Adaptive Regression Splines) (Chou *et al*., 2004) and Fuzzy AIS with k-NN (Sahan *et al*., 2007). Most of the above mentioned methods combine two or three methods, while some researchers have proposed combining more models, such as Hassan and Verma (2007) which combines self-organizing map (SOM), k-means and naïve Bayes with a neural network based classifier.

Apart from improving (or hybridizing) existing data mining techniques, other attempts to enhance the final predicted output are based on improving the quality of the data itself. Approaches which fall under this category aim to study the medical data itself and apply different techniques to the data such as Decomposition using structured rule-feature matrix (Kusiak, 2001), discretization (Abraham *et al*., 2006), filtering outliers (Podgorelec *et al*., 2005) and filtering with over-sampling (Thongkam *et al*., 2009).

However, among the different approaches and techniques used for medical applications, in this paper we are concerned with the use of Naïve Bayes (NB) for medical classification. In the following we discuss its basic features and how it suits for this domain.

NAÏVE BAYES

Naïve Bayesian classifier, or simply naïve bayes (NB), is one of the most effective and efficient classification algorithms. It is a simple probabilistic classifier based on applying Bayes' theorem with strong (naïve) independence assumptions.

Given a set of training instances with class labels and a test case E represented by n attribute values (a_1, a_2... a_n), Bayesian classifiers use the following equation to classify E:

$$C_{NB}(E)=\arg_{c\in C}\max P(c)\prod_{i=1}^{n}P(a_i\mid c) \quad (1)$$

where, c_{NB} (E) denotes the classification given by NB on test case E.

Although independence is generally a poor assumption, in practice NB often competes well with much more sophisticated techniques. In a large-scale comparison of naïve Bayes classifier with state-of-the-art algorithms for decision tree induction, instance-based learning and rule induction, conducted by (Domingos and Pazzani, 1997) on standard benchmark datasets; the authors found NB to be sometimes superior to the other learning schemes, even on datasets with substantial feature dependencies.

A variety of adaptations to NB in the literature have been studied in order to improve upon its good performance while maintaining its efficiency and simplicity. For more details on features of NB and an

overview of variants of NB classifiers, please refer to (Al-Aidaroos *et al*., 2010).

NB has proven its effective application, often reported as "surprisingly" accurate, in text classification, medical diagnosis and systems performance management (Rish, 2001). However, as mentioned previously in this paper we are concerned on its application to medical data and how it handles the different problems in this domain. In the following, based on the discussed requirements of medical data mining systems, we see how this approach is applicable for mining medical data.

Medical data mining with NB: Kononenko (2001) considered NB as a benchmark algorithm that in any medical domain has to be tried before any other advanced method. While Abraham *et al*. (2006) argue, based on their study, that simple methods are better in medical data mining and this makes NB performs well for such data. Compared to other classifiers, NB is simple, computationally efficient, requires relatively little data for training, do not have lot of parameters and is naturally robust to missing and noise data (Al-Aidaroos *et al*., 2010).

One of the main advantages of NB approach which is appealing to physicians, is that all the available information is used to explain the decision. This explanation seems to be "natural" for medical diagnosis and prognosis i.e. is close to the way how physicians diagnose patients (Zelic *et al*., 1997).

When dealing with medical data, naïve bayes classifier takes into account evidence from many attributes to make the final prediction and provides transparent explanations of its decisions and therefore it is considered as one of the most useful classifiers to support physicians' decisions.

Successful applications of NB to medical data have been reported by many researchers in the literature (Kononenko *et al*., 1998; Demsar *et al*., 2001; Abraham *et al*., 2006). Kononenko *et al*. (1998) compared NB with six algorithms (Assistant-R, Assistant-I, LFC, backpropagation, k-NN and semi-NB). The result was that NB classifier outperformed all the algorithms on five out of eight medical diagnostic problems. However, even with small data sets, naïve bayes have shown that it can construct reasonably accurate prognostic models as proved by Demsar *et al*. (2001), who used naïve bayes classifier with a data set which includes only 68 patients. In a comparative study of discretization methods for medical data mining, conducted by Abraham *et al*. (2006), it suggests that on an average the NB classifier with MDL discretization seems to be the best performer compared to popular variants of NB and non-NB classifiers (such as DT, k-NN and LR).

EMPIRICAL COMPARISON

Here we present an empirical comparison of Naïve Bayes algorithm with five popular algorithms on 15 medical data sets. The selected algorithms are: Logistic Regression (LR), KStar (K*), Decision Tree (DT), Neural Network (NN) and a simple rule-based algorithm (ZeroR). These algorithms were chosen because they represent quite different approaches to learning and they have been used in medical data mining applications as discussed earlier. We used K* which retrieves the nearest stored example using an entropic measure instead of Euclidean distance (Cleary and Trigg, 1995), to represent instance-based learning instead of the k-Nearest Neighbor because it produced better results for the selected data sets. We also included ZeroR in this comparison which simply predicts the majority class in the training data, because it is usually used by many comparison studies as a baseline for other methods.

Motivation: Before we go to the experiment details, in this subsection is our motivation for conducting this comparison although there have been other comparative studies which include NB in their experiments.

Firstly, to support our conclusion, based on the reviewed literature, that naïve Bayes suits classification problems in the medical domain as it satisfies most of the MDM requirements.

Secondly, our experiment is different because most of the comparative studies use UCI data sets from a wide range of domains, while in this comparison analysis we focus on problems which are all from medical domain. Although Abraham *et al*. (2006) conducted a comparative analysis on NB with respect to medical data (only 6 data sets included); their objective was to study the effect of discretization methods in improving NB's accuracy rather than comparing the overall performance of NB approach with other approaches.

Finally, to provide explanation of the bad performance reported by some of the comparative studies in the literature. Jamain and Hand (2005) conducted a large meta-analysis of comparative studies of supervised classification methods and they spotted some anomalous results relating to the naïve bayes method. Their conclusion regarding the two famous comparative studies, Statlog project and Zarndt study, is that the former mislabeled another method as naïve Bayes and the later used the right method but possibly incorrectly

Table 1: Medical data sets used for the experiment

Medical problem	No. of Instances	No. of attributes		No. of Classes
		Numeric	Nominal	
Breast cancer wisc	699	0	9	2
Breast cancer	286	0	9	2
Dermatology	366	1	33	6
Echocardiogram	132	8	2	2
Liver disorders	345	6	0	2
Pima diabetes (Indians)	768	8	0	2
Haberman	306	2	1	2
Heart-c (Cleveland)	303	6	7	2
Heart-statlog	270	5	8	2
Heart-h (Hungarian)	294	5	7	2
Hepatitis	155	6	13	2
Lung cancer	32	0	56	3
Lymphography	148	0	18	4
Postoperative patient	90	0	8	3
Primary tumor	339	0	17	21

reported its provenance. Based on our literature review on naïve Bayes with MDM we agree with the explanation given by Jamain and Hand (2005), that the bad performance of naïve Bayes in some studies is a result of the choice of data sets that different studies have made and in this section we empirically support our argument.

Experimental setup: In this paper, WEKA version 3.7.0 (Hall *et al*., 2009) was used for the evaluation of the performance of NB against the five selected schemes. The six algorithms (NB: Naïve Bayes, LR: Logistic, K*: KStar, DT: J48, NN: Multilayer Perceptron and ZeroR) were employed based on their default parameters of the WEKA application. We used 10-fold cross-validation to minimize the bias associated with the random sampling of the training and holdout data samples (Kohavi, 1995).

Selected data sets: In our experiment, 15 real-world medical problems from the UCI machine learning repository (Asuncion and Newman, 2007) were selected for evaluating the performance of all algorithms. Table I gives a summary of the characteristics of these data sets. Some of these data sets contain attributes which are all numeric or all nominal while other data sets contain mixed type of attributes. The number of class values for most of the data sets is two, indicating the presence or absence of a disease, but there are some data sets with more than two classes, as shown in Table 1.

Evaluation measures: Two established measures were used to evaluate the performance of NB as compared to the other five algorithms (schemes); namely prediction accuracy and area under ROC curve.

All accuracy estimates were obtained by averaging the results from 10 runs of 10-fold cross-validation. In other words, each scheme was applied 100 times to generate an estimate for a particular data set. In the following, two results for a data set are considered to be "significantly different" if the difference is statistically significant at the 5% level according to the corrected resampled t-test (Witten and Frank, 2005).

An ROC (Receiver Operating Characteristics) graph is a technique for selecting classifiers based on their performance and it is commonly used in medical decision making. However, an ROC curve is a two-dimensional depiction of classifier performance and to compare classifiers a common method is to calculate the area under the ROC curve (AUC) (Fawcett, 2006). In our experiment, we computed the area under the ROC (AUC) for the 10 two-class data sets.

RESULTS AND DISCUSSION

Based on the conducted experiments, the predictive accuracy results for NB versus LR, K*, DT, NN and ZeroR are presented in Table 2; while the computed AUC values are shown in Table 3. For both tables, the entries which are marked as bold are those algorithms achieving the highest value for each data set.

The wins given at the bottom of Table 2 shows that NB outperforms the other algorithms in 8 out of 15 data sets. LR is the second best algorithm followed by DT, based on the wins of each algorithm. In other words, depending on the number of wins we can see that NB on the average gives better performance than the other methods on the 15 data sets.

However, to compare NB's performance against each method, in the following we analyze the results by considering if any has a statistically significant improvement or degradation; denoted as •, ◦, respectively; over NB. From Table 2, we can see that LR is significantly worse than NB on four data sets and significantly better on one. K* is significantly better than NB on one data set and significantly worse on eight. DT is also significantly better on only one data set but significantly worse on five data sets. NN is significantly better on one data set and worse on three. Finally, as expected ZeroR, is significantly worse on 9 data sets and has no significant improvement over NB on any of the data sets. It is important to notice here that all the significant improvements over NB are limited to the Liver disorders data set. In other words, NB is significantly worse than other approaches in only one data set among the 15 data sets.

Table 3 shows the AUC results on the two class data sets. AUC's value is always between 0 and 1.0 but no realistic classifier should have an AUC less than 0.5 (Fawcett, 2006). It is clear from the displayed results that ZeroR is the worse in terms of AUC and its value is 0.5 for all data sets.

Table 2: Comparative analysis based on predictive accuracy

Medical problem	NB	LR	K*	DT	NN	ZeroR
Breast cancer wisc	**97.30**	92.98◦	95.72◦	94.57◦	95.57◦	65.52◦
Breast cancer	72.70	67.77◦	73.73	**74.28**	66.95	70.30
Dermatology	**97.43**	96.89	94.51◦	94.10◦	96.45	30.60◦
Echocardiogram	95.77	94.59	89.38◦	**96.41**	93.64	67.86◦
Liver disorders	54.89	**68.72•**	66.82•	65.84•	**68.73•**	57.98
Pima diabetes (Indians)	75.75	**77.47**	70.19◦	74.49	74.75	65.11◦
Haberman	**75.36**	74.41	73.73	72.16	70.32◦	73.53
Heart-c (Cleveland)	83.34	**83.70**	75.18◦	77.13◦	80.99	54.45◦
Heart-statlog	**84.85**	84.04	73.89◦	75.59◦	81.78	55.56◦
Heart-h (Hungarian)	83.95	**84.23**	77.83◦	80.22	80.07	63.95◦
Hepatitis	**83.81**	**83.89**	80.17	79.22	80.78	79.38
Lung cancer	**53.25**	47.25	41.67	40.83	44.08	40.00
Lymphography	**84.97**	78.45	83.18	78.21	81.81	54.76◦
Postoperative patient	68.11	61.11◦	61.67	69.78	58.44	**71.11**
Primary tumor	**49.71**	41.62◦	38.02◦	41.39◦	40.38◦	24.78◦
Wins	8/15	5/15	0/15	2/15	1/15	1/15

•,◦Statistically significant improvement or degradation over NB

Table 3: Comparative analysis based on Area Under ROC curve (AUC)

Medical problem	NB	LR	K*	DT	NN	ZeroR
Breast cancer wisc	**0.99**	0.95◦	**0.99**	0.97◦	**0.99**	0.50◦
Breast cancer	**0.70**	0.63◦	0.67	0.61◦	0.65	0.50◦
Echocardiogram	0.69	0.68	0.70	**0.71**	**0.71**	0.50◦
Liver disorders	0.64	0.72•	0.70	0.65	**0.73•**	0.50◦
Pima diabetes (Indians)	0.82	**0.83**	0.73◦	0.75◦	0.80	0.50◦
Haberman	0.65	0.65	**0.70**	0.57	0.66	0.50◦
Heart-c (Cleveland)	**0.91**	**0.91**	0.83◦	0.77◦	0.88	0.50◦
Heart-statlog	**0.91**	0.90	0.84◦	0.77◦	0.88	0.50◦
Heart-h (Hungarian)	**0.91**	**0.91**	0.82◦	0.78◦	0.88	0.50◦
Hepatitis	**0.86**	0.84	0.81	0.67◦	0.81	0.50◦
Wins	6/10	3/10	2/10	1/10	3/10	0/10

•,◦Statistically significant improvement or degradation over NB

In terms of the number of wins given at the bottom of Table 3; we can see that NB is the best algorithm outperforming the other algorithms in 6 out of 10 data sets. LR and NN are the second best algorithms with 3 wins, followed by K* which wins in 2 out of the 10 data sets.

It is interesting to note that, in terms of AUC, NN is better than LR as compared to NB. Although both approaches have the same number of wins but when compared to NB and considering significant tests, we can see the difference. Although both algorithms, NN and LR, are significantly better than NB in one data set but NN has no significant degradation as compared to NB while LR is significantly worse than NB on two data sets. Both DT and K* algorithms have no significant improvement over NB, however, K* seems to be better than DT since it is significantly worse on only 4 data sets while DT is significantly worse on 7 data sets.

To conclude about the AUC results, the number of wins shows that NB is better than the other methods but if we consider significant differences then we can see that NN is significantly better in one case with no significant degradation over NB. However, as with the prediction accuracy, the NB is outperformed on the Liver disorders data set; this might indicate that the attribute dependencies are present in this data set.

Based on the experimental results, it is important to notice the following. Although NN significantly outperformed NB in terms of AUC measure on one data set with no degradation, but the "black box" nature of NN gives poor explanation capabilities to the results produced by NN; and as discussed previously in the requirements of MDM this is an important feature which can intervene in favor of a method that may not have the highest performance. So, in our case for the tested medical problems NB is considered to be the best choice because of its simplicity, explanation capability and efficiency as compared to NN.

CONCLUSION AND FUTURE WORK

This study reviewed the current state of medical data mining from different perspectives. Naïve Bayes classification approach has been discussed and its main features are highlighted based on the medical mining requirements. Based on the experimental results we prove empirically its suitability to the medical domain problems as compared to other approaches. The experimental

results show that NB is better than the compared approaches on most of the used medical data sets.

However, since NB has been widely criticized due to its unrealistic independence assumption and as hybridization is widely used to overcome problems of different individual techniques; our main direction for future work is to investigate the hybridization of NB with other approaches which have dependency detection capability to improve the performance of NB and propose new algorithm for medical mining applications.

REFERENCES

Abed, H. and L. Zaoui, 2011. Partitioning an image database by K_means algorithm. J. Applied Sci., 11: 16-25.

Abraham, R., J.B. Simha and S.S. Iyengar, 2006. A comparative analysis of discretization methods for medical data mining with Naive Bayesian classifier. Proceeding of the 9th International Conference on Information Technology (ICIT`06), Dcember 18-21, 2006, Bhubaneswar, pp: 235-236.

Al-Aidaroos, K.M., A.A. Bakar and Z. Othman, 2010. Naive Bayes variants in classification learning. Proceeding of the International conference on Information Retrieval and Knowledge Management (CAMP 2010), March 17-18, 2010, Shah Alam, Selangor, pp: 276-281.

Asuncion, A. and D. Newman, 2007. UCI machine learning repository of machine learning databases. University of California, School of Information and Computer Science, Irvine, CA. http://www.ics.uci. edu/~mlearn/ MLRepository.html

Bai, V.T. and S.K. Srivatsa, 2006. Wireless tele care system for intensive care unit of hospitals using bluetooth and embedded technology. Inform. Technol. J., 5: 1106-1112.

Bellazzi, R. and B. Zupan, 2008. Predictive data mining in clinical medicine: Current issues and guidelines. Int. J. Med. Inform., 77: 81-97.

Blessia, T.F., S. Singh, A. Kumar and J.J. Vennila, 2011. Application of knowledge based system for diagnosis of osteoarthritis. J. Artif. Intel., 4: 269-278.

Chen, T.S., J. Chen and Y.H. Kao, 2010. A novel hybrid protection technique of privacy-preserving data mining and anti-data mining. Inform. Technol. J., 9: 500-505.

Chou, S.M., T.S. Lee, Y.E. Shao and I.F. Chen, 2004. Mining the breast cancer using artificial neural networks and multivariate adaptive regression splines. Expert Syst. Applica., 27: 133-142.

Cios, K.J. and G.W. Moore, 2002. Uniqueness of medical data mining. Artif. Intell. Med., 26: 1-24.

Cleary, J.G. and L.E. Trigg, 1995. An Instance-based learner using an entropic distance measure. Proc. Int. Conf. Mach. Learn., 5: 108-114.

Cong, C. and C.P. Tsokos, 2010. Statistical modeling of breast cancer relapse time with different treatments. J. Applied Sci., 10: 37-44.

Delen, D., G. Waller and A. Kadam, 2005. Predicting breast cancer survivability: A comparison of three data mining methods. Artificial Intell. Med., 34: 113-127.

Demsar, J., B. Zupan, N. Aoki, M.J. Wall, T.H. Granchi and J.R. Beck, 2001. Feature mining and predictive model construction from severe trauma patient`s data. Int. J. Med. Inform., 63: 41-50.

Domingos, P. and M. Pazzani, 1997. On the optimality of the simple Bayesian classifier under zero-one loss. Machine Learning, 29: 103-130.

Dzeroski, S. and N. Lavrac, 1996. Rule induction and instance-based learning applied in medical diagnosis. Technol. Health Care, 4: 203-221.

Fawcett, T., 2006. An introduction to ROC analysis. Pattern Recog. Lett., 27: 861-874.

Goletsis, Y., C. Papaloukas, D.I. Fotiadis, A. Likas, and L.K. Michalis, 2004. Automated ischemic beat classification using genetic algorithms and multicriteria decision analysis. IEEE Trans. Biomed. Eng., 51: 1717-1725.

Hall, M., E. Frank, G. Holmes, B. Pfahringer, P. Reutemann and I.H. Witten, 2009. The WEKA data mining software: An update. SIGKDD Explorations Newslett., 11: 10-18.

Hassan, S.Z. and B. Verma, 2007. A hybrid data mining approach for knowledge extraction and classification in medical databases. Proceeding of the IEEE Seventh International Conference on Intelligent Systems Design and Applications, Octuber 20-24, 2007, Rio de Janeiro, pp: 503-508.

Hassanien, A.-E., M.E. Abdelhafez and H.S. Own, 2008. Rough sets data analysis in knowledge discovery: A case of Kuwaiti diabetic children patients. Adv. fuzzy syst., Vol. 2008. 10.1155/2008/528461.

Hudson, D.L. and M.E. Cohen, 2002. Use of intelligent agents in the diagnosis of cardiac disorders. Disorders. Compu. Cardiol., 29: 633-636.

Jamain, A. and D.J. Hand, 2005. The naive bayes mystery: A classification detective story. Pattern Recogni. Lett., 26: 1752-1760.

John, R.I. and P.R. Innocent, 2005. Modeling uncertainty in clinical diagnosis using fuzzy logic. IEEE Trans. Syst. Man Cybern., 35: 1340-1350.

Kittaneh, R., S. Abdullah and A. Abuhamdah, 2012. Iterative simulated annealing for medical clustering problems. Trends Applied Sci. Res., (In Press).

Kohavi, R., 1995. A study of cross-validation and bootstrap for accuracy estimation and model selection. Proceeding of 14th International Joint Conference on Artificial Intelligence, (IJCAI`95), Standfard, pp: 1137-1143.

Kononenko, I., 2001. Machine learning for medical diagnosis: history, state of the art and perspective. Artif. Intell. Med., 23: 89-109.

Kononenko, I., I. Bratko and M. Kukar, 1998. Application of Machine Learning to Medical Diagnosis. In: Machine Learning and Data Mining: Methods and Applications, Michalski, R.S., R.S. Michalski, I. Bratko and M. Kubat (Eds.). J. Wiley, New York.

Kuhnert, P.M., K.-A. Do and R. McClure, 2000. Combining non-parametric models with logistic regression: An application to motor vehicle injury data. Comput. Statist. Data Analys., 34: 371-386.

Kuo, W.J., R.F. Chang, D.R. Chen and C.C. Lee, 2001. Data mining with decision trees for diagnosis of breast tumor in medical ultrasonic images. Breast Cancer Res. Treat., 66: 51-57.

Kusiak, A., 2001. Decomposition in data mining: A medical case study. Proceedings of the SPIE Conference on Data Mining and Knowledge Discovery: Theory, March 27, 2001, Tools and Technology III, pp: 267-278.

Lanzola, G., L. Gatti, S. Falasconi and M. Stefanelli, 1999. A framework for building cooperative software agents in medical applications. Artif. Intell. Med., 16: 223-249.

Latifoglu, F., K. Polat, S. Kara and S. Gunes, 2008. Medical diagnosis of atherosclerosis from carotid artery Doppler signals using principal component analysis (PCA), k-NN based weighting pre-processing and Artificial Immune Recognition System (AIRS). J. Biomed. Inform., 41: 15-23.

Laurikkala, J., M. Juhola, S. Lammi, and K. Viikki, 1999. Comparison of genetic algorithms and other classification methods in the diagnosis of female urinary incontinence. Method. Inform. Med., 38: 125-131.

Lavrac, N., 1999. Selected techniques for data mining in medicine. Artificial Intell. Med., 16: 3-23.

Li, J., A. Wai-Chee Fu, H. He, J. Chen and H. Jin, 2005. Mining risk patterns in medical data. Proceedings of the 11th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, Aug. 21-24, Chicago, Illinois, USA., pp: 770-775.

Lin, R.-H., 2009. An intelligent model for liver disease diagnosis. Artif. Intell. Med., 47: 53-62.

Lisboa, P.J.G., 2002. A review of evidence of health benefit from artificial neural networks in medical intervention. Neural Netw., 15: 11-39.

Maria, P.R.F., 2002. An intelligent, interactive, web-based platform for effective clinical education through VR and distance learning modules. Inform. Technol. J., 1: 126-131.

Mugambi, E.M., A. Hunter, G. Oatley and L. Kennedy, 2004. Polynomial-fuzzy decision tree structures for classifying medical data. Knowledge-Based Syst., 17: 81-87.

Nabi, M.S.A., M.L.M. Kiah, B.B. Zaidan, A.A. Zaidan and G.M. Alam, 2010. Suitability of using SOAP protocol to secure electronic medical record database transmission. Int. J. Pharmacol., 6: 959-964.

Ohno-Mochado, L., 2001. Modeling medical prognosis: survival analysis techniques. J. Biomed. Inform., 34: 428-439.

Pattaraintakorn, P., N. Cercone and K. Naruedomkul, 2005. Hybrid intelligent systems: Selecting attributes for soft-computing analysis. Proceeding of the 29th Annual International Computer Software and Applications Conference, July 26-28, 2005, Department of Math., Mahidol University, Thailand, pp: 319-325.

Podgorelec, V. and P. Kokol, 2001. Towards more optimal medical diagnosing with evolutionary algorithms. J. Med. Syst., 25: 195-219.

Podgorelec, V., M. Hericko and I. Rozman, 2005. Improving mining of medical data by outliers prediction. Proceedings of 18th IEEE Symposium on Computer-Based Medical Systems, June 23-24, 2005, Dublin, Ireland, pp: 91-96.

Polat, K., S. Sahan, H. Kodaz and S. Gunes, 2005. A new classification method for breast cancer diagnosis: Feature selection artificial immune recognition system (FS-AIRS). Lect. Notes Comput. Sci., 3611: 830-838.

Rish, I., 2001. An empirical study of the naive bayes classifier. Proceedings of the IJCAI-01 Workshop on Empirical Methods in Artificial Intelligence, (WEMAI`01), Watson Research Center, Hawthorne, pp: 41-46.

Sahan, S., K. Polat, H. Kodaz and S. Gunes, 2007. A new hybrid method based on fuzzy-artificial immune system and k-NN algorithm for breast cancer diagnosis. Comput. Biol. Med., 37: 415-423.

Sumathi, S. and S.N. Sivanandam, 2006. Data mining in biomedicine and science. Stud. Comput. Intell., 29: 499-543.

Temurtas, H., N. Yumusak and F. Temurtas, 2009. A comparative study on diabetes disease diagnosis using neural networks. Expert Syst. Appl., 36: 8610-8615.

Thongkam, J., G. Xu, Y. Zhang and F. Huang, 2008. Breast cancer survivability via AdaBoost algorithms. Proc. 2nd Aust. Workshop on Health Data Knowledge Manage., 80: 55-64.

Thongkam, J., G. Xu, Y. Zhang, and F. Huang, 2009. Toward breast cancer survivability prediction models through improving training space. Expert Syst. Appl., 36: 12200-12209.

Tsumoto, S., 1999. Knowledge discovery in medical multi-databases: A rough set approach. Principles Data Min. Knowl. Dis., 1704: 147-155.

Velmurugan, T. and T. Santhanam, 2011. A survey of partition based clustering algorithms in data mining: An experimental approach. Inform. Technol. J., 10: 478-484.

Witten, I.H. and E. Frank, 2005. Data Mining Practical Machine Learning Tools and Techniques. 2nd Edn., Morgan Kaufman, San Francisco, CA.

Yao, X. and Y. Liu, 1999. Neural networks for breast cancer diagnosis. Proceedings of the 1999 Congress on Evolutionary Computation, July 06-09, 1999, Washington, DC, USA.

Zelic, I., I. Kononenko, N. Lavrac and V. Vuga, 1997. Induction of decision trees and bayesian classification applied to diagnosis of sport injuries. J. Med. Syst., 21: 429-444.