International Journal of Applied Engineering Research
ISSN 0973-4562 Volume 10, Number 10 (2015) pp. 25287-25295

http://www.ripublication.com

Empirical Evaluation of Machine Learning Algorithms In Detecting Cardiac Arrhythmia

Dr. L.V. Nandakishore [1], Dr. S. Aruna [2]
[1] *Professor, Department of Mathematics, Dr MGR Educational and Research Institute University, Chennai-95*
[2] *Asst Professor, Department of Software Applications, A.M. Jain college, Chennai-114*
lvnandakishore@gmail.com[1],arunalellapalli@yahoo.com[2]

Abstract

In this paper comprehensive empirical comparison was carried out with various machine learning algorithms. The aim of the research was to analyze the classifier performance on real world data with respect to disease detection. The Cardiac arrhythmia dataset from UCI machine learning depositary was used for the experiment. The experiments were conducted in WEKA. The performance criterion was compared between five machine learning tools namely Naïve Bayes, Support vector machines, Radial basis neural networks and Decision trees J48 and simple CART in detecting cardiac arrhythmia.

Keywords: Cardiac arrhythmia, Machine learning Naïve Bayes, Bayes Net, Radial basis neural networks, Support vector machines, Decision trees

Introduction

Arrhythmias are abnormal heart rhythms symbolize a serious threat to the patient recovering from acute myocardial infarction, especially ventricular arrhythmias like ventricular tachycardia (VT) and ventricular fibrillation (VF) [1]. In the United States more than 850,000 people are hospitalized every year for arrhythmia [2]. Ventricular tachycardia kills an estimated 120,000 people in the UK each year [3]. The detection of arrhythmia is an important task in clinical reasons which can initiate life saving operations [4]. It is a challenging task even for experts when they are put into stress and immense work load due to the constant increase in the number of patients. Several methods for automated arrhythmia detection have been developed in the past few decades to simplify the monitoring task [5]. Electrocardiogram (ECG) records the electronic activities of the heart, and has been widely adopted for diagnosing cardiac arrhythmia [6]. The state of cardiac health is generally reflected in the shape of the

ECG waveform and heart rate and contains important pointers to the nature of the disease attacking the heart [7].

Machine learning provides various tools to build automated systems for disease detection. They make use of statistics for fundamental concepts adding more advanced AI heuristics and algorithms to achieve its goals. ML algorithms based systems are fast and effective for a single disease. Pattern recognition is essential for the diagnosis of new diseases. ML plays a vital role in recognizing patterns in the data mining process. It searches for the patterns within the patient database. Searching and recognizing patterns in the biochemical state of diseased people is very relevant to understanding of how diseases manifest or drugs act. This information can be utilized for disease prevention, disease management, drug discovery thus improving health care and health maintenance.

The predictive performance and generalization power of machine learning algorithms plays a vital role in classification of diseases. Typically high sensitivity and specificity is required to rule out other diseases. This reduces subsequent diagnostic procedures which causes additional efforts and costs for differential diagnosis of the disease. Additionally high predictive accuracy, speedy processing, results interpretation and visualization of the results are also mandatory for good screening systems. In this paper performance of five machine learning algorithms namely Support vector machines (SVM), Naïve bayes, Radial basis function networks (RBFN) and decision trees such as classification and regression trees (CART) and J48 were analyzed with respect to arrhythmia classification. The Cardiac arrhythmia dataset from the UCI machine learning depository was used for the experiment. The rest of the paper is organized as follows. Related work was given in Section 2. Section 3 gives a brief introduction to the data set and machine learning methods, framework used for the experiment. Section 4 gives the results obtained and discussion about the results were given in Section 5. Concluding remarks were given in section 6 to address further research issues.

Related Work

In a comparison of 10 learning algorithms over 11 datasets after calibration with Platt's method or isotonic regression SVM perform comparably to neural nets and nearly as well as boosted trees [8]. Gorman et al., [9] reported that back propagation outperformed nearest neighbors for classifying sonar targets. Shadmehr et al., [10] showed that the performance of Bayes algorithm is better. Kirkwood et al., [11] developed a symbolic algorithm ID3 which performed better than discriminant analysis for classifying the gait cycle of artificial limbs. Spikvoska et al [12] found that a HONN (higher order neural network) performed better than ID3. Atlas et al., [13] showed that back propagation performed better than Cart. Mitchell et al., [14] compared many algorithms on the MONK's problem.

Ripley [15] compared neural networks and decision trees on the Tsetse fly data. King et al., [16] Statlog is the first comprehensive study that analyzed different data mining algorithms on large real world data sets. LeCun et al., [17] compared several learning algorithms on a handwriting recognition problem. Cooper et al., [18]

evaluated supervised learning methods on real medical data set using accuracy. Bauer et al., [19] did an empirical analysis about different statistical methods such as bagging and boosting. Lim et al., [20] compared decision trees and other methods using accuracy as the main criterion. Perlich et al., [21] conducted comparison between decision trees and logistic regression. Provost et al., [22] examined the issue of predicting probabilities of decision trees including smooth and bagged trees. Witten et al., [23] presented the comparison of different tools and techniques of data mining.

Materials and Methods

Data Set Description

The Cardiac Arrhythmia Database used in this study was obtained from the UCI Machine Learning Repository [24]. The dataset has 452 instances of 16 classes. Class 01 is normal; classes' 02-15 different types of cardiac arrhythmias and unclassified data come under class 16. Each instance has 279 attributes, of which first four attributes namely age, sex, height and weight give the general description of the patient. Remaining attributes are extracted from the standard 12 lead ECG recordings. There are 206 linear valued and the remaining are nominal valued attributes.

Machine Learning Algorithms

Naive Bayes

Naive Bayes classifier is a statistical classifier based on the Bayes theorem with the assumption of class conditional independence i.e. a feature value's effect on the given class is independent of the values of other features in making a decision therefore only the variances of the attributes for each class were calculated instead of enumerating the entire covariance matrix. Naive Bayes utilizes a probabilistic method for classification by multiplying the individual probabilities of every attribute-value pair [25].The Naive Bayes classifier can be trained very efficiently in a supervised learning setting, working better in many complex real-world situations, especially in the computer-aided diagnosis [26].

Decision trees CART and J48

Decision trees are supervised algorithms which recursively partition the data based on its attributes, until some stopping condition is reached [27] The most important feature of decision trees are their capability to break down a complex decision making process into a collection of simpler decisions, thus providing a solution, which is often easier to interpret [28].

CART

The classification and regression trees (CART) methodology proposed by [29] uses cross-validation on a large independent test sample of data to select the best tree from the sequence of trees considered in the pruning process and the dataset is split into the

two subgroups that are the most different with respect to the outcome until some minimum subgroup size is reached.

J48

Decision tree J48 [30] implements Quinlan's C4.5 algorithm [31] for generating a pruned or an unpruned tree which is an extension of Quinlan's earlier ID3 algorithm compute normalized information gain that results from choosing an attribute for splitting the data. The J48 tree comprises of nodes at every stage that are structured with the help of training examples [32].

Radial Basis Neural Networks

Radial Basis Function (RBFN) is the artificial neural network type used extensively for application of supervised learning problem [33]. Due to the simple structure of RBFN the training of networks is relatively fast and capable of universal approximation with non-restrictive assumptions [34]. The RBFN can be implemented in any types of model whether linear on non-linear and in any kind of network whether single or multilayer [35]. A mathematical justification of this can be found in the paper by Cover [36].

Support Vector Machines

Support vector machines (SVM) are a class of learning algorithms which are based on the principle of structural risk minimization (SRM) [37]. SVMs have been successfully applied to a number of real world problems, such as handwritten character and digit recognition, face recognition, text categorization and object detection in machine vision [38], [39], [40]. SVM has a number of advanced properties, including the ability to handle large feature space, effective avoidance of over fitting, and information condensing for the given data set etc.. [41]

Framework For The Experiment

Empirical comparison of machine learning algorithms had been carried out according to the framework in figure 2. As the arrhythmia dataset has about 0.33% missing values in the first step data cleaning is done to remove the missing values instances After data cleaning 295 instances were obtained. Even though the dataset to claim to be of 16 classes, many classes have no samples or samples with a lot of missing values. Hence the data set was transformed from multiclass to binary by changing the class values of normal samples to 0 and those of arrhythmia samples to 1. After the transformation to binary dataset they are trained using the classifiers and the results are analyzed. To analyze the performance of the classifiers with respect to the attributes feature selection algorithm was used to select the attributes. Then the data set was simulated by removing the features other than the selected attributes and used for training the classifiers. The results were analyzed and compared with the previous results.

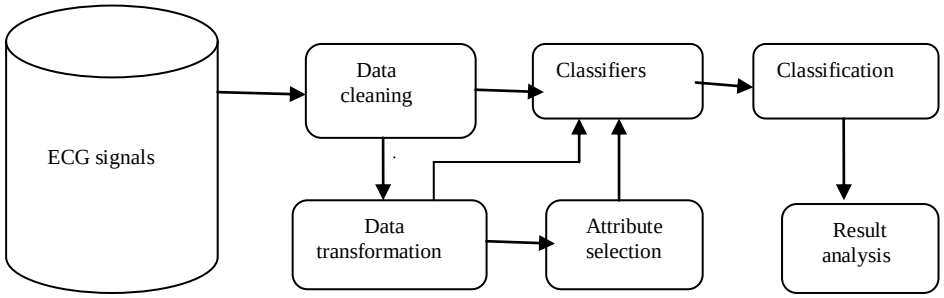

Figure 2: Experimental framework

Results

Experiments conducted in WEKA with 10 fold cross validation. Ten fold cross validation proved statistically good enough in evaluating performance of the classifier [42]. From the confusion matrix accuracy, sensitivity, specificity, mean absolute error (MAE), root mean squared error (RMSE), balanced classification rate (BCR) [43], [44] were also computed to compare the outcome of the machine learning tools.

Table 1 shows the performance comparison for the five machine learning algorithms on arrhythmia data set after data cleaning and data transformation. Correlated feature selection (CFS) was used to select the attributes for training the classifiers. Feature Subset having high, class correlation and low inter correlation are selected using best first search criterion. CFS selected 14 features (Table 2) from the 280 features. These 14 features were used to train the classifiers and the performance of the classifiers were analyzed. Table 3 shows the performance results for the ML algorithms after attribute selection.

Table 1: Performance for ML algorithms before attribute selection

Algorithms	**Accuracy (%)**	**BCR (%)**	**Sensitivity (%)**	**Specificity (%)**	**MAE**	**RMSE**
SVM	91.6	82.5	68	97	0.08	0.28
Naive bayes	87.1	82.0	75	89	0.12	0.35
RBFN	88.1	81.5	72	91	0.15	0.30
CART	92.1	81.0	68	94	0.12	0.29
J48	88.4	78.5	64	93	0.13	0.33

Table 2: Attributes selected CFS using best first search

chAVR_intrinsicRelections	chAVR_RRwaveExists
chAVR_Rwave	chv1_Qwave
Chv1_RPwave	Chv1_intrinsicRelections
Chv2_RPwave	chDII_SwaveAmp
chAVR_RPwaveAmp	chAVF_TwaveAmp
Chv1_QRSA	Chv2_SwaveAmp
Chv2_TwaveAmp	Chv1_RPwaveAmp

Table 3: Performance for ML algorithms after attribute selection

Algorithms	**Accuracy (%)**	**BCR (%)**	**Sensitivity (%)**	**Specificity (%)**	**MAE**	**RMSE**
SVM	93.9	85.0	72	98	0.06	0.23
Naive bayes	91.5	84.5	79	90	0.08	0.27
RBFN	92.8	82.0	74	90	0.10	0.24
CART	92.5	82.5	70	95	0.11	0.26
J48	91.8	82.0	70	94	0.10	0.27

Discussion

There are a number of data mining algorithms that are found useful for automatic classification of ECG signals. In [45] the authors used neural network with weighted fuzzy membership function to classify cardiac arrhythmia with an accuracy of 81.32%. In [46] Elsayad used learning vector quantization neural networks and obtained the classification accuracy of 76.92%. Zuo et al. [47] used kernel difference KNN to detect arrhythmia. They achieved an accuracy of 70.66%. In [48] authors used fuzzy weighted artificial immune recognition system for medical diagnosis. They obtained an accuracy of 80.71% for ECG arrhythmia database. Uyar et al. [49] classified arrhythmia using a serial fusion of support vector machines and logistic regression with an accuracy of 76.1%. In [50] authors used a novel pruning approach using expert knowledge for data running for diagnosing the arrhythmia with an accuracy of 68.47%. Ozcan [51] used fuzzy support vector machines for ECG arrhythmia classification. For Adaptive neuro fuzzy inference system (ANFIS) and Fuzzy support vector machines distance to class mean method they got accuracy of 79.43% and 83.33% respectively. In [52] authors proposed an effective ANN based approach for cardiac arrhythmia classification. They got classification accuracy of 82.22%, 82.35%, and 86.67% for Modular neural network model, generalized feed forward neural network model and multilayer perceptron model respectively.

Most of the algorithms produce results that are variable in nature for the same type of experiment. Thus finding the best classifier is an interesting and time consuming work. The aim of this study is to find out the classifier which can do well on the arrhythmia data set. As the ECG data set was highly imbalanced (16 classes of nil or

uneven sample distribution) and consists of lot of missing values, noisy, irrelevant attributes this study encompass the impact of the nature of the data set, selection of classifier and the preprocessing techniques used influence the predictive power of the machine learning systems. Among the five machine learning algorithms used for training SVM performed better than other learning methods in all facets except specificity. The comparison of results in table 1 and 3 also confirmed the presence the redundant, irrelevant attributes in the data set. The performance improvement of all the classifiers after attribute selection supports the claim by various researchers that not all the attributes are needed for classification and determining the best subset of variables which carries most diagnostic information is necessary to meet satisfactory classification accuracy.

Conclusions

Cardiac arrhythmias are the irregular rhythm of the heart show a serious medical problem sometimes threatening the life. Detection of arrhythmia is an important task in effective management of the disease. In this study five machine learning algorithms namely Radial basis function networks, Naïve Bayes, Support vector machines, Decision trees J48 and Simple cart were experimented using the performance metrics namely accuracy, sensitivity, specificity, mean absolute error, root mean squared error and the balanced classification rate in detecting cardiac arrhythmia. The cardiac arrhythmia dataset from the UCI machine learning depository was used for the experiment. The experiments were conducted in WEKA with 10 fold cross validation. The framework for evaluation methodology comprises of data cleaning, data transformation, classification and attribute selection. The study showed that SVM outperformed all the other classifiers with and without attribute selection in all aspects except specificity. The specificity of Naive bayes was higher than SVM in both cases.

In the present work machine learning algorithms were used for detecting the presence or absence of cardiac arrhythmias. Future work includes repeating the experiment to identify the subtypes of cardiac arrhythmia which will further explore the performance of machine learning algorithms on multiclass classification strategy. Empirical comparison between linear and ensemble classifiers, various preprocessing strategies can also be made to encompass the problem in classifying biomedical data in all facets.

References

[1] Ho Ryu, K (2008): A data mining approach and framework of intelligent diagnosis system for coronary artery disease prediction, pp. 33--34, IEICE technical report

[2] http://www.medicinnet.com

[3] http://www.heartrhythmcharity.org.uk

[4] Gyu lee, H., Ho Ryu, K (2008): A data mining approach for coronary heart disease prediction using HRV features and carotid arterial wall thickness, Bio. Eng. And informatics, 200--206

[5] Raut, R.D., Dudul, S.V (2008): Arrhythmias Classification with MLP Neural Network and Statistical Analysis, First IEEE International Conference on Emerging Trends in Engineering and Technology, pp. 553--558

[6] Zuo, W.M., Lu, W.G., Wang, K.Q., Zhang, H (2008): Diagnosis of cardiac arrhythmia using kernel difference weighted KNN classifier, Computers in Cardiology, 253--256

[7] Dale Dubin, M.: Rapid Interpretation of ECG's, USA, ISBN: 9780912912066 2001.

[8] R Caruana, and A. Niculescu-Mizil, (2006) "An empirical comparison of supervised learning algorithms" *Proceedings of the 23rd international conference on Machine learning*.

[9] R P. Gorman and T.J. Sejnowski, (1988) "Analysis of hidden units in a layered network trained to classify sonar targets", *Neural networks,* 1 (Part 1) 75-89.

[10] R. Shadmehr and Z. D'Argenio, (1990) "A comparison of a neural network based estimator and two statistical estimators in a sparse and noisy environment", In IJCNN-90 *Proceedings of the international joint conference on neural networks,* 289-292,Ann Arbor, Mi, IEEE Neural Networks Council.

[11] C. Kirkwood, B. Andrews and P. Mowforth, (1989) "Automatic detection of gait events: a case study using inductive learning techniques", *Journal of biomedical engineering,* 11(23): 511-516.

[12] L. Spikovska and M.B. Reid, (1990) "An empirical comparison of id3 and honns for distortion invariant object recognition" In Tai-90:tools for artificial intelligence: *Proceedings of the 2nd international IEEE conference,* Los Alamitos, CA, IEEE Computer Society Press.

[13] L. Atlas, J. Connor and D. Park, (1991)"An performance comparison of trained multilayer perceptrons and trained classification trees", In Systems, man and cybernetics: *Proceedings of the 1989 IEEE conference*, 915-920, Cambridge, Ma. Hyatt Regency.

[14] T. Mitchell, B. Buchanan, G. Dejon, T. Diettrich, P. Rosenbloom and A. Waibel,(1990) "Machine Learning*" Annual Review of Computer Science,* vol 4, 417-433.

[15] B. Ripley, (1993)"Statistical aspects of neural networks", *Chaos and Networks- Statistical and Probabilistic Aspects,* Chapman and Hall.

[16] R. King, C. Feng and A. Sutherland, (1995) "Statlog: Comparison of classification algorithms on large real world problems", *Applied Artificial Intelligence,* 9.

[17] Y. LeCun, L.D. Jackel, L. Bottou, A. Brunot, C. Cortes, J.S. Denker, H. Drucker, L. Guyon, U.A. Mutter, E. Sackinger, P. Simard and V. Vapnik, (1995) "Comparison of learning algorithms for handwritten digit

recognition", *International Conference on Artificial Neural Networks*, 53-60, Paris.

[18] G.F Cooper, C.F. Aliferis, R. Ambrosino, J. Aronis, B.G Buchanan, R. Cruana, M.J. Fine, C. Glymour, G. Gordon, B.H. Hanusa, J.E. Janosky, C. Meek, T. Mitchell, T. Richardson and P. Spirtes, (1997)"An evaluation of machine learning methods for predicting pneumonia mortality", *Artificial intelligence in Medicine*, 9.

[19] E. Bauer and R. Kohavi, (1999)"An empirical comparison of voting classification algorithms: Bagging, boosting and variants", *Machine Learning*, 36.

[20] T.S Lim, W.Y Loh and Y.S Shih, (2000)"A comparison of prediction accuracy, complexity and training time of thirty-three old and new classification algorithms", *Machine Learning*, 40, 203-228.

[21] C. Pertich, F. Provosi and J.S Simono, (2003)"Tree induction vs Logistic regression: a learning curve analysis", *J. Mach. Learn. Res.*, 4, 211-255.

[22] F. Provost, D. Jensen and T. Oates, (1999)"Effiecient progressive sampling", Fifth ACM SIGKDD, *International conference on knowledge Discovery and Data Mining*, San Diego, USA.

[23] I.H. Witten and E. Frank, (2000) "*Data Mining: Practical Machine learning tools and techniques with java implementations",* Morgan Kaufmann.

[24] UCI Machine Learning Depository, http://www.ics.uci.edu/~mlearn/ MLRepository.

[25] S. Belciug, (2008) "Bayesian classification vs. k-nearest neighbour classification for the non-invasive hepatic cancer detection", *Proc. 8th International conference on Artificial Intelligence and Digital Communications*.

[26] F. Gorunescu, (2006) *Data Mining: Concepts, models and techniques*, Blue Publishing House, Cluj Napoca.

[27] T. Pang-Ning, S. Michael, K. Vipin, (2008), *Introduction to Data Mining*, Low Price edn. Pearson, Education, Inc., London, ISBN 978-81-317-1472-0.

[28] L. Breiman, J. Friedman., R. Olshen, C. Stone, (1984), *Classification and Regression Trees*. Wadsworth, Belmont, CA.

[29] D.Steinberg., and P.L. Colla, (1995) "CART: Tree-Structured Nonparametric Data Analysis", *Salford Systems:* SanDiego, CA.

[30] D.Steinberg., and P.L. Colla, (1997) "CART-Classification and Regression Trees", *Salford Systems:* San Diego, CA.

[31] Ross J. Quinlan, (1992) " Learning with Continuous Classes". *5th Australian Joint Conference on Artificial Intelligence*, Singapore, 343-348.

[32] Ross Quinlan, (1993) *C4.5: Programs for Machine Learning*, Morgan Kaufmann Publishers, San Mateo, CA.

[33] M. J. L. Orr, (1996) *Radial Basis Function Networks*. Edinburgh, Scotland.

[34] J. Park, and I.W. Sandberg, (1993) "Approximation and Radial-Basis-Function Networks", *Neural Computation.* 5, 305-316.

[35] S. Haykin, (1994) *Neural Networks a Comprehensive Foundation*, New Jersey, PrenticeHall.

[36] T. M. Cover, (1965) "Geometrical and Statistical Properties of Systems of Linear with Applications in Pattern Recognition," *IEEE Transactions on Electronic Computers EC-14*, pp. 326-334.

[37] Vladimir N. Vapnik. (1998) *Statistical Learning Theory*. New York: Wiley.

[38] C. Campbell, N. Cristianini, and J.Shawe-Taylor, (1999) "Dynamically Adapting Kernels in Support Vector Machines", *Advances in Neural Information Processing Systems*,**Vol.** 11. MIT Press, 204-210.

[39] Corinna Cortes,Valdimir Vapnik, (1995) "Support-Vector Networks" *Machine Learning*, 20, 273-297.

[40] M,Pontil ,A.Verri, (1998) "Support Vector Machines for 3 D object recognition**."** *IEEE T Pattern Anal* , 20(6):637-646.

[41] You et al, (2010) "A semi-supervised learning approach to predict synthetic genetic interactions by combining functional and topological properties of functional gene network**"** *BMC Bioinformatics*, 11:343.

[42] P. Baldi, S. Brunak, Y. Chauvin, et al (2000). Assessing the accuracy of prediction algorithms for classification: and overview. *Bioinformatics*, 5(5):412–424.

[43] Sokolova, M. and Lapalme, G. (2009). A systematic analysis of performance measures for classification tasks. *Information Processing and Management* 45, 4 (Jul. 2009), 427-437.

[44] Demsar, J (2006).: Statistical comparisons of classifiers over multiple data sets. *Journal of Machine Learning Research* 7, 1-30

[45] Sang-Hong Lee, Jung-Kwon Uhm, Joon S. Lim.: Extracting Input Features and Fuzzy Rules for Detecting ECG Arrhythmia Based on NEWFM, *International Conference on Intelligent and Advanced Systems, Division of Software*, Kyungwon University, Korea

[46] Elsayad, A.M (2009): Classification of ECG arrhythmia Using Learning Vector Quantization Neural Networks, *In proceedings of Computer engineering systems* , pp. 139—144.

[47] Zuo, W. M., Lu, W. G., Wang, K. Q., Zhang, H (2008): Diagnosis of cardiac arrhythmia using kernel difference weighted KNN classifier, *Computers in Cardiology*, 253—256.

[48] Kemal Polat, Seral Şahan, Salih Güneş (2006).: A new method to medical diagnosis: Artificial immune recognition system (AIRS) with fuzzy weighted pre-processing and application to ECG arrhythmia, *Expert Systems with Applications*, 31, 264--269

[49] Uyar, Gurgen, F (2007).: Arrhythmia Classification Using Serial Fusion of Support Vector Machines and Logistic Regression, *In proceedings of Intelligent Data Acquisition and Advanced Computing Systems: Technology and Applications* (IDAACS 2007), pp. 560--565

[50] Ali Mirza Mahmood, Mrithyumjaya Rao Kuppa (2012) "A novel pruning approach using expert knowledge for data specific pruning", *Engineering with Computers*, 28, 21--30

[51] Ozlem, N. O., Gurgen, F (2010): Fuzzy Support Vector Machines for ECG Arrhythmia Detection, *In proceedings of Pattern Recognition* (ICPR 2010), pp. 2973--2976

[52] Jadhav, S.M., Nalbalwar, S.L., Ashok, A., Ghatol (2012): "Artificial Neural Network Models based Cardiac Arrhythmia Disease Diagnosis from ECG Signal Data, *International Journal of Computer Applications*, 44, 8--13