Article can be accessed online at http://www.publishingindia.com

Foetal State Determination Using Support Vector Machine and Firefly Optimisation

Subha V.*, Murugan D.**

Abstract

In recent days, it has been found that vast amount of data is available in medical field that aid the doctors in disease diagnosis. Data mining techniques have been applied to extract knowledge from these medical data so that disease prediction becomes easy. In this paper, cardiotocogram (CTG) data is classified using support vector machine (SVM). An optimised feature subset is produced using Firefly Algorithm (FFA) with SVM evaluation. The results show that the performance of classification is improved with the optimal reduced feature set than with full feature set.

Keywords: SVM Classifier, Cardiotocography, Firefly Algorithm, Feature Selection

Introduction

Cardiotocography (CTG) was introduced in late 1960s in obstetrics for assessing the foetal status (Hasan, 2013). It is used to determine the foetal state both during pregnancy and delivery. It is a combination of two signals: Foetal Heart Rate (FHR) and Uterine Contractions (UC). CTG is used to determine those babies who lack oxygen during labour. Then decision is made by the obstetricians whether to deliver the baby by caesarean section or by natural birth. Visual analysis of CTG data often leads to incorrect interpretations and hence computer assisted systems are essential in classifying the recorded CTG data for decision making by the obstetricians (Ersen *et al.*, 2013).

Many techniques have been reported in literature for analyzing cardiotocogram data. In a study presented in (Hasan, 2013), SVM classifier has been used to classify the foetal state into two classes. In addition, genetic algorithm is used to select the relevant features to improve the performance of the classifier. In another study (Ersen *et al.*, 2013), Least Squares-SVM, Particle Swarm Optimisation, and binary decision tree are used for classification of CTG data. An accuracy of 91.62% is obtained with three classes. Adaptive neuro fuzzy inference system has been used for classifying CTG data which classifies foetal state into two classes (Hasan *et al.*, 2013). In (Peterek *et al.*, 2013), CTGs are classified using random forest classifier with feature reduction technique. Foetal distress prediction using discriminant analysis, decision tree, and artificial neural network (Huang *et al.*, 2012) shows that highest accuracy is obtained with artificial neural network. (Hakan *et al.*, 2012) classified CTG recordings using artificial neural network and simple logistics and concluded that the latter overcome the performance of artificial neural networks. Classification of CTG data using Naïve Bayes Classifier with feature selection approaches is reported in (Mohamed *et al.*, 2013). Modular neural network has been applied to foetal state classification (Shivajirao *et al.*, 2011) which classifies the data into three classes. Sundar *et al.* (2013a) adopted artificial neural network based classifier that overcomes the performance of other clustering algorithms in CTG classification. Naive Bayes Classifier (Sundar *et al.*, 2013b) has been adopted for classifying CTG records which classifies the data into three classes.

* Assistant Professor, Department of Computer Science & Engineering, Manonmaniam Sundaranar University, Tirunelveli, Tamil Nadu, India. E-mail: subha_velappan@yahoo.co.in

** Associate Professor, Department of Computer Science & Engineering, Manonmaniam Sundaranar University, Tirunelveli, Tamil Nadu, India. E-mail: dhanushkodim@yahoo.com

This paper aims at classifying CTG data using SVM classifier. Experiments have been conducted in two phases. In the first phase, CTG records are classified using the complete feature set. Later, in the second phase optimal feature set has been obtained by Firefly Algorithm with SVM evaluation and the results were compared.

Materials and Methods

Material Description

The cardiotocography dataset has been downloaded from UCI Machine Learning Repository (Bachie et al. 2013). This dataset contains 2126 foetal cardiotocograms belonging to three different classes with 21 attributes and one class attribute. The CTGs are classified by three expert obstetricians and classification labels are assigned to them. The dataset consists of measurements of foetal heart rate (FHR) and uterine contractions (UC). The classification is with respect to a foetal heart rate class code (N-Normal, S-Suspect and P-Pathologic).

Attribute Information

1. LB - FHR baseline (beats per minute)
2. AC - Number of accelerations per second
3. FM - Number of foetal movements per second
4. UC - Number of uterine contractions per second
5. DL - Number of light decelerations per second
6. DS – Number of severe decelerations per second
7. DP - Number of prolonged decelerations per second
8. ASTV - Percentage of time with abnormal short term variability
9. MSTV - Mean value of short term variability
10. ALTV - Percentage of time with abnormal long term variability
11. MLTV - Mean value of long term variability
12. Width - Width of FHR histogram
13. Min - Minimum of FHR histogram
14. Max - Maximum of FHR histogram
15. Nmax - Number of histogram peaks
16. Nzeros - Number of histogram zeros
17. Mode - Histogram mode
18. Mean - Histogram mean
19. Median - Histogram median
20. Variance - Histogram variance
21. Tendency - Histogram tendency
22. CLASS- Foetal state class code (Normal=1; Suspect=2; Pathologic=3)

Methods

Support Vector Machines

One of the popular machine learning tools used for classification and regression is Support Vector Machine (SVM). Supervised learning is used to classify the points to one of two disjoint half-spaces. It uses nonlinear mapping to convert the original data into higher dimension. The objective is to construct a function which will correctly predict the class to which the new point belongs and the old points belong (Yashima *et al.*, 2012). A hyperplane is built by the SVM to separate the data points into different classes with a maximum margin.

According to Vapnik's theory (Cortes *et al.*, 1995), let $\{(a_1,b_1),....,(a_m,b_m)\}$ be assumed as the given training data sets, where $a_i \subset R^n$ represents the input space of the sample and $b_i \subset R$ for $i = 1,.....,m$ represents respective target value, where m denotes the number of elements in the training dataset. The errors are tolerable as long as they are less than ε value. Eq. (1) is solved to estimate linear regression in SVM:

$$\text{Minimise } 0.5\|w\|^2 + D\sum_{i=o}^{m}(\zeta + \zeta^*)$$

$$\text{subjected to } \begin{cases} b_i - \langle w, a_i\rangle - p \le \varepsilon_i + \zeta \\ \langle w, a_i\rangle + p - b_i \le \varepsilon_i + \zeta^* \\ \zeta_i, \zeta_i^* \ge 0, \quad i = 0,...,m \end{cases} \quad (1)$$

where, w is a normal vector, p is a scalar quantity, D represents a regularisation constant, ε is the insensitive loss function, and the slack variables, ζ, ζ^*, correspond to the size of the excess deviation for upper and lower deviations, respectively.

Lagrangian multipliers $(\alpha_i, \alpha_i^*, \eta, \eta^*)$ are used to solve Eq. (1). The obtained generic equation is written as,

$$f(a) = \sum_{i-1}^{l}(\alpha_i - \alpha_i^*)\langle a_i, a\rangle + p \quad (2)$$

This methodology can be used for solving the linear regression problems. By replacing x_i by a mapping in to the feature space $\phi(i)$ and thereby linearising the relationship between x_i and y_i, the same method can be extended for nonlinear regression problems as well. The solution of Eq. (2) will thus become;

$$f(a)=\sum_{i-1}^{l}(\alpha_i-\alpha_i{}^*)K\langle a_i,a\rangle+p \tag{3}$$

where, $K(a_i,a)=\langle\phi(a_i),\phi(a)\rangle$ and $K(a_i,a_j)$ is the kernel function.

Over the period of years, Radial Basis Function has become the choice of the researchers as kernel function for SVM because of its accurate and reliable performance (Sudheer *et al.*, 2014). Therefore, the Radial Basis Function (RBF) has been adopted in this work.

Firefly Optimisation Algorithm

Fireflies are insects producing a flashing light by a natural biochemical process which is called as bioluminescence. The male flying fireflies generally produce this flashing light to attract the flightless females. As a response, the female fireflies also produce a continuous or flashing light. The mating partners produce the flash patterns distinctly by encoding the information on identity and gender (Fister *et al.*, 2013). Hence, the flashing lights primarily serve as a communication with their mating partners, means to attract their prey and a protective warning mechanism.

Firefly algorithm uses the light intensity (H) which will decrease with the increase in the square of the distance (d^2). When the distance from the light source increases, the brightness decreases and thereby results a decrease in attractiveness. The objective function to be optimised is formed based on this phenomenon.

Following are the three idealised rules which are used in FFA;

- All fireflies are considered as unisex.
- Their attractiveness varies is directly proportional to the intensity of light.
- The light intensity of any firefly is directly influenced by the landscape of the fitness function that is considered.

FFA has been devised based two things; the variation of light intensity and the attractiveness of firefly. The light emitted by the firefly is termed as intensity. On the other hand, the intensity of a firefly seen by other fireflies (Yang, 2010) is known as attractiveness. The light intensity of a firefly which actually represents a solution in FFA, is directly proportional to the value of fitness function.

As mentioned earlier, the light intensity (H) varies with distance (d) as;

$$H(d)=\frac{H_s}{d^2}\,; \tag{4}$$

where, H_s is the intensity at the source.

With an original light intensity of source (H0) and a constant light absorption coefficient (γ), the light intensity of a firefly will vary with the distance (d) in the medium as;

$$H=H_0e^{-\gamma d} \tag{5}$$

From equations (4) and (5), it can be written in Gaussian form as;

$$H(d)=H_0e^{-\gamma d^2} \tag{6}$$

As stated above, the attractiveness β of a firefly is proportional to the light intensity seen by the other fireflies. Hence;

$$\beta=\beta_0e^{-\gamma d^2} \tag{7}$$

where, β_0 is the attractiveness at d=0.

The distance between any two fireflies X_i and X_j is expressed as the Euclidean distance;

$$d_{ij}=\left\|x_i-x_j\right\|=\sqrt{\sum_{k=1}^{n}\left(x_{i,k}-x_{j,k}\right)^2} \tag{8}$$

where, $x_{i,k}$ is the k^{th} component of the spatial coordinate x_i.

It can be written for the 2D case as;

$$d_{ij}=\sqrt{\left(x_i-x_j\right)^2+\left(y_i-y_j\right)^2} \tag{9}$$

The firefly i is attracted by another more attractive firefly j as given by,

$$x_i=x_i+\beta_0e^{-\gamma d_{ij}{}^2}\left(x_j-x_i\right)+\alpha\varepsilon_i \tag{10}$$

where, α is the randomisation parameter and ε_i is vector of random numbers generated from Gaussian distribution.

The movement of firefly has three terms: the current position of i^{th} firefly, attraction to another one and a random movement. Therefore, the FA has three parameters, the randomisation parameter (α), the attractiveness (β), and the absorption coefficient (γ) which can be adjusted to modify the performance of FFA. The pseudo code of FFA is shown in Figure 1 (Yang, 2010).

Figure 1: Firefly Algorithm Pseudo Code

Begin
Define light absorption coefficient (γ), initial attractiveness (β_0), randomization parameter (α) and maximum generation
Objective function $f(x),\ x=(x_1,x_2,......x_m)^T$
Generate initial population of fireflies
$x_i\ (i=1,2,....,n)$
Determine the light intensity H_i at x_i by $f(x_i)$
while (t < Maximum Generation)
for i=1: n all n fireflies
    for j=1: n all n fireflies
if (H_i<H_j), Move firefly i towards j;
end if

FFA Optimised Feature Subset Selection Using SVM Evaluation

Feature selection (FS) is the process of excluding the irrelevant features present in a dataset which otherwise decreases the classification performance. It is the problem of finding an optimal subset of n features from the complete set of m features. Feature selection techniques can be either wrapper based or filter based. While wrapper methods use the performance of a classifier for evaluating the feature subsets, filter methods use any feature evaluation technique. Recently, evolutionary algorithms such as Artificial Bee Colony Ant Colony Optimisation, Particle Swarm Optimisation etc., have been introduced for finding feature subsets. In this work, the Firefly optimisation algorithm has been used with SVM as evaluation function for finding the feature subset. The features in the dataset are represented as a binary string of 0's and 1's. The presence of a particular feature is denoted as 1 and its absence as 0. The classification accuracy of SVM has been used as objective function for evaluating the feature subset.

The objective function is given by,

$$\text{Accuracy} = \frac{(TP+TN}{(TP+TN+FP+FN)} \tag{11}$$

where, TP - True Positives

TN - True Negatives

FP - False Positives

FN - False Negatives

There are 50 fireflies used in each iteration and they search in a random feature space and find an optimal solution for the maximum number of iterations which is selected as 100. The light intensity of firefly is proportional to the objective function which is the classifier accuracy. The subset of features which maximizes the classification accuracy of SVM till the termination condition (i.e. maximum number of iterations) is reached is chosen as the optimal feature subset.

The parameters of Firefly Algorithm are listed in Table 1.

Table 1: Firefly Algorithm Parameters

No. of fireflies	50
No. of iterations	100
Randomisation parameter (α)	0.5
Attractiveness (β)	0.2
Light absorption coefficient (γ)	1

Results & Discussion

The experiments have been carried out using the original dataset and the optimal reduced dataset. The results are given in Table 2. It is found that the average accuracy is 86.78% with full feature set and the same is achieved as 91.95% with optimal feature set. The performance measures Sensitivity, Specificity, Positive Predictive Value and Negative Predictive Value also depicts that optimal feature set improves the classification performance of SVM. These metrics are defined as:

Table 2: Comparison of SVM Accuracy (Class-wise) with and Without Feature Selection

	Full feature set accuracy (%)	*Optimal feature set accuracy (%)*
Normal	94.44	95.64
Suspect	66.77	77.62
Pathologic	72.15	81.25
Average accuracy	88.75	91.95

$$\text{Sensitivity} = \frac{TP}{(TP+FN)}$$

$$\text{Specificity} = \frac{TN}{(TN+FP)}$$

$$\text{Positive Predictive Value } (PPV) = \frac{TP}{(TP + FP)}$$

$$\text{Negative Predictive Value } (NPV) = \frac{TN}{(TN + FN)}$$

Table 3: Performance Metrics (Average) of SVM with and Without Feature Selection

Performance Metrics (%)	*Without FS*	*With FS*
Sensitivity	77.30	84.83
Specificity	90.22	93.83
PPV	78.56	83.22
NPV	90.70	93.26

Table 3 also shows that performance of SVM improves with feature subset selection in terms of sensitivity, specificity, positive predictive value and negative predictive value.

For better illustration, Figure 2 and 3 show the results of Table 2 and 3 pictorially.

Figure 2: Comparison of SVM Accuracy (Class-Wise)

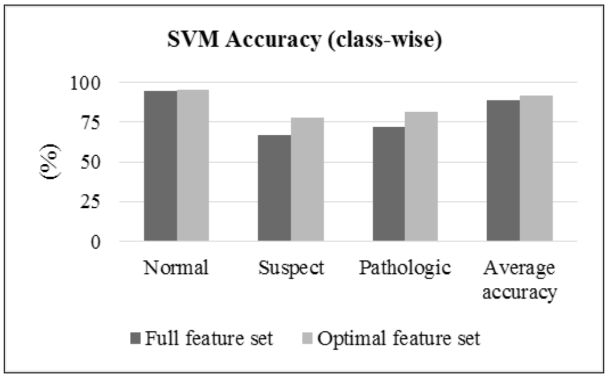
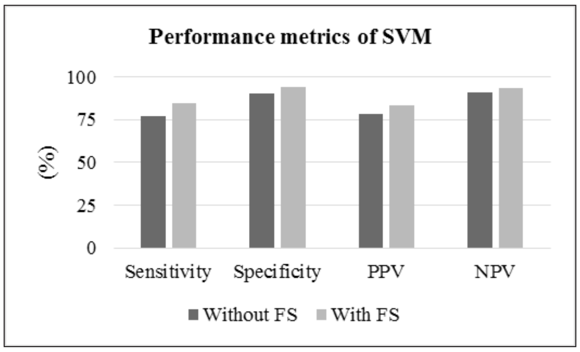


Figure 3: Performance Metrics (Average) of SVM Classifier



Conclusion

In this paper, the CTG records have been classified using Support Vector Machines. Firefly algorithm has been used to optimise the feature set. The results show that classification performance improves with optimal feature set than with full feature set. This may help the obstetricians in making accurate decisions from CTG recordings. In future, this work can be extended with hybrid FFA techniques. Other optimisation algorithms can also be applied to CTG classification to improve the accuracy. Rough sets with FFA can also be considered for feature selection so that classification performance may improve further.

References

Ahuja, Y., & Yadav, S. K. (2012). Multiclass classification and support vector machine. *Global Journal of Computer Science and Technology Interdisciplinary*, 12(11), 14-20.

Bache, K., & Lichman, M. (2013). *UCI Machine Learning Repository*. Irvine, CA: University of California, School of Information and Computer Science. Retrieved from http://archive.ics.uci.edu/ml.

Ch, S., Sohani, S. K., Kumar, D., Malik, A., Chahar, B. R., Nema, A. K., Panigrahi, B. K., & Dhiman, R. C. (2014). A support vector machine-firefly algorithm based forecasting model to determine malaria transmission. *Neurocomputing*. 129, 279-288.

Cortes, C., & Vapnik, V. (1995). Support vector networks. *Machine Learning*. 20(3), 273-297.

Fister, I., Fister, Jr., I., Yang, X. S., & Brest, J. (2013). A comprehensive review of firefly algorithms. *Swarm and Evolutionary Computation*. 13, 36-46.

Huang, M. L., & Hsu, Y. Y. (2012). Fetal distress prediction using discriminant analysis, decision tree, and artificial neural network. *Journal of Biomedical Science and Engineering*, 5(9), 526-533.

Jadhav, S., Nalbalwar, S., & Ghatol, A. (2011). *Modular Neural Network Model Based Foetal State Classification.* IEEE International Conference on Bioinformatics and Biomedicine Workshops, (pp. 915-917).

Menai, M. E. B., Mohder, F. J., & Al-mutairi, F. (2013). Influence of feature selection on naïve bayes classifier for recognizing patterns in cardiotocograms. *Journal of Medical and Bio-engineering*. 2(1), 66-70.

Ocak, H. (2013). A medical decision support system based on support vector machines and the genetic algorithm for the evaluation of fetal well-being. *Journal of Medical Systems*, 37(2), 9913.

Ocak, H., & Ertunc, H. M. (2013). Prediction of fetal state from the cardiotocogram recordings using adaptive neuro-fuzzy inference systems. *Neural Computing & Applications*, 23(6), 1583-1589.

Sahin, H., & Subasi, A. (2012). *Classification of Fetal State from the Cardiotocogram Recordings using ANN and Simple Logistic*. In 3rd International Symposium on Sustainable Development, Sarajevo, (pp.499-505).

Schiezaro, M., & Pedrini, H. (2013). Data feature selection based on Artificial Bee Colony algorithm. *EURASIP Journal on Image and Video Processing*. 47, 1-8.

Sundar, C., Chitradevi, M., & Geetharamani, G. (2013a). An overview of research challenges for classification of cardiotocogram data. *Journal of Computer Science*, 9(2), 198-206.

Sundar, C., Chitradevi, M., & Geetharamani, G. (2013b). An analysis on the performance of naïve bayes probabilistic model based classifier for cardiotocogram data classification. *International Journal on Computational Sciences & Applications*, 3(1), 17-26.

Tomas, P., Krohova, J., Dohnalek, P., & Gajdos, P. (2013). *Classification of Cardiotocography Records by Random Forest*. IEEE 36th International Conference on Telecommunications and Signal Processing, (pp. 620-623).

Uzer, M. S., Yilmaz, N., & Inan, O. (2013). Feature selection method based on artificial bee colony algorithm and support vector machines for medical datasets classification. *The Scientific World Journal*.

Yang, X. S. (2010). *Nature-inspired metaheuristic algorithms* (2nd ed). UK, Luniver Press.

Yılmaz, E., & Kılıkçıer, C. (2013). Determination of fetal state from using LS-SVM with particle swarm optimization and binary decision tree. *Computational and Mathematical Methods in Medicine*.