ISSN 2229-5518

Diagnosis of sleep Apnea disease with ECG, SPO2 through using of support vector machine (SVM)

Mohamad Amin alikhani,Behzad Ebrahimnezhad sani,Javad Haddadnia

Abstract—The aim of this article is to present new method on the basis of support vector machine (SVM) in order to diagnose obstructive sleep Apnea through features of ECG SPO signals. For achieving our aim, we consider two parameters signal namely, the rate of oxygen concentration in blood and relative pressure of blood circulation. Then, we designed support vector machine (SVM) in Matlab environment and applied patients' information to network. This designed network is diagnosed on the basis of inputs and instructional data. Moreover, it will simulate the prediction of sleep Apnea with high accuracy.

Index Terms—Sleep apnea- support vector machine (SVM)-PCA-ECG-SPO2

1 INTRODUCTION

Respiration and sleep are necessary for our survival. We spend about one third of our lives asleep. During sleeping, some changes occur in our body which causes to create vulnerability to some kinds of sleep disorders. For instance, our respiration may be natural in awakening hours. However, because of the changes in our body during sleeping, we may be afflicted with disorders. Apnea during sleeping is to stop airway in nose and mouth intermittently during sleeping. In fact, lasting at least 10 seconds in the case of Apnea is of importance. Sleep Apnea syndrome is a kind of clinical disorders which is result from repeated Apnea during sleeping which includes signs such as extremely loud heavy snoring, outbreak of respiratory interruption, abruptly and repeated awakening, following respiration interruption, morning headache, fatigue, sleepiness. Apnea during sleeping can be CSA, central (OSA) or mixed. Obstructive Apnea is defined as interruption measurable airflow without stopping respiration efforts simultaneously. However, central Apnea occurs when there are not airflow, the stomach and ribcage respiratory efforts. Mix Apnea is has both of central and obstructive Apnea. Obstructive sleep Apnea is the most common form of Apnea. It appears that 5_20 0/0 of matures suffers from obstructive respiratory disruption. Patients who suffer from respiratory interruption of obstructive sleep are susceptible to sleepiness. Furthermore, because of hypoxias and hyperkapin attacks, they are prone to cardiovascular disease heart failure. Polisonografy is recognized as a standard method of sleep Apnea diagnosis. The definition of polisonografy of syndrome obstructive sleep Apnea during sleeping is on the basis of the numbers of abnormal respiratory event which occur every hour of sleeping. In fact, it is called and exapnehypopeneh. Usually, it is believed that 5 events of …apnehyponie is considered normal in an hour. However, it is abnormal for a society group of children in every hour. This respiratory disorder during sleeping can cause to reduce the quality of sleeping. Hence, it has destructive effect on quality of awakening hours. Excessive daytime sleepiness is one of the main signs in patients suffering from this syndrome. Studies show that 2_3 % of workers who work in industries afflicted with sleepiness because of syndrome of obstructive sleep Apnea.

Therefore, this syndrome can be considered as a general healthy problem. Because it causes to increase outbreak of working event and road disasters.

Currently, analysis of the patient's polysomnography (PSG) is considered as the effective diagnosis of the sleep apnea [5]. PSG requires overnight recordings of several electrophysiological signals during a night's Sleep such as electrocardiogram, respiratory effort, airflow, etc. in sleep laboratories using specific systems and participating personnel [6]. To derive respiration from electrocardiography (ECG) this is a simple, low cost and non-invasive recording is an alternative way. The correctness of such an idea using the comparison example recordings of the ECG derived respiration (EDR) with common respiration measurements are showed in Moody et al. work [7]. Thus, different methods have been proposed to derive respiratory signal from the ECG [8-10]. Numerous studies show that EDR methods using the band-pass filter application can be accepted as the most successful methods in the field of apnea detec-

1-Master student of electronic Engineering, Hakim Sabzevari University
2- Master student of electronic Engineering, Hakim Sabzevari University
3-Associate Professor of Biomedical Engineering, Hakim Sabzevari University

ISSN 2229-5518

tion[11,12].Numerous methods are available in the literature to detect sleep apnea based on the evaluation of PSG [13-16]. These methods are mostly based on analysis of the frequency and amplitude. These algorithms can detect sleep apnea at 80% - 90% precision [17]. There are several algorithms for the detection of sleep apnea based on the evaluation of the EDR signals [18-20]

2 PATHOPHYSIOLOGY

Mechanism OSA of general society is similar to mechanism OSA of patients afflicted with heart failure. In fact, heart failure may help to upper airway instability through effecting on neck venous congestion. It appears that when patients are in a supine position during sleeping, increasing pressure helps to increase collapse of airway. The role of this mechanism is shown in OSA outbreak as an agent among many of pathogenesis namely, obesity, senility in heart failure.
Increasing the pressure of filling upon on the rate of the heart helps to occur OSA. It is probability that this change is responsible for significant common between obstructive and central events in patients suffering from heart failure and change between OSA and CSA in repetitive experiment.

3 ELEMENTS ANALYSIS

Advances and development in collecting data and storing capabilities in recent decades to have a bulk of information in many sciences. Researchers carried out studies in different subjects including engineering, biology, astronomy, and economics with more and more observation.
Traditional statistics methods lost their efficiency because of two reasons.
The first reason: increasing observation numbers
The second reason: it is of high importance. Increasing the number of variables related to one observation. The number of variables which measured for every observation called dimension. Variable is used more in statistics .whereas, it is used as a feature or attribute in computer science and learning machine. Bed data have many dimensions. In spite of the fact that they create opportunity and calculation challenges. One of the data problems with big dimensions is that all of data features found in hidden data are not necessary. That's reason why among fields is to decrease data dimensions. In fact, it is one of significant discussions.

4 THE METHODS OF DIMENSIONS REDUCTION:

It is divided into two groups:
1-The methods based upon feature extraction:
These methods reduce multi dimensions space into one space with fewer dimensions. So, there is less number of feathers.
Indeed, theses features have available information in primary features. These methods are divided into two groups namely, linear and non linear.
2-The methods based upon selecting feature:
The methods try to decrease data dimensions with selecting subset from primary features. This chapter deals with mathematics poof. But, a little work on it was done. More concepts and applications of methods were done in this chapter. Indeed, we present clear examples in order to better comprehension.

5 METHODS BASED UPON FEATURE EXTRACTION

As mentioned above, the methods based upon feature extraction reduce multi dimension to one space or fewer dimension .These methods are divided into two groups namely linear and nonlinear. The former is simple and easy to understand. They are to find global flat subspace. While nonlinear is hard to understand and its analysis is difficult. They are to find locally flat subspace. Linear method includes FA, PCA, DWT, DFT. Nonlinear method is as follows:

- **Principal Curves**
- **Setf Oeganizing Maps**
- **Vestor Quantization**
- **Genetic and Evolutionary Algorithms**
- **Regression**

The issue of data dimension reduction can be mentioned as mathematics:
One random variable –p dimension x=(x1, x2…) we want –k variable –k (dimension) S= (s1, s2,….sk). Find first K ≤ P.., secondary s", the content exists in x. these concepts have special criteria. Linear methods try to obtain k component from linear component of p, primary component.

$$s_{\mathrm{i}} = \mathrm{w}_{\mathrm{i},1} + w_{i,2}x_2 + w_{i,p}x_p (1)$$

WKP is linear rate matrix. Techniques PCA is the best method for describing data dimension as linear. In fact, reduction information is led to less input information than other methods.

$$s = WX \Rightarrow (2)$$
$$i = 1,2, \ldots K$$

6 PRINCIPAL COMPONENT ANALYSIS (PCA)

In this method, new characteristic axial is defined for data and data is mentioned on the basis of these new characteristic axial. The first axial should be in a direction which data points have high variance.high**variance** indicates that the **data** points are very **spread** out around the mean and from each other. That is, there is the **mostvariance** in the x **direction**.
The second axial should be vertical on the first axial. It should be in a direction which data points have high variance. In fact, with maximum variance, the data is

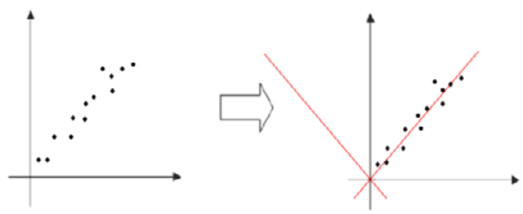
ISSN 2229-5518

most spread. Figure 1_5 is shown for data.
Method PCA is known by other names such as :

- **KARHUNEN Loeve Transform (KLT)**
- **Hotelling Transform**
- **Empirical Orthogonal Function (EDF)**

Before dealing with method in details, we mentioned math and statistical concepts related to this method. This concept includes standard deviation, covariance, special vector and special values.

.



. figure 1: new axial with regard to special vector in a direction of the most density point

7 PRIMARILY CONCEPTS IN PCA

PCA have detail mathematic basics. As general insights from PCA mathematic discussion, it is proved that space N dimension can be defined N vertical vector. Furthermore, each point of this space is shown through coefficients multiplication in vertical vector (**vertical** bars on either side of the **vector)**.
As an example, in a space of two dimensions, we have two axial vertical. All of the points of two dimension space are shown by these two axial. Furthermore, three dimension space…….to N dimension.
PCA obtained special vectors or vertical vectors and special values i.e. coefficient of these vectors.
Statistical concepts:
Let x is string of values. The mean of this value is obtained in the following equation

$$\bar{X} = \frac{1}{n}\sum_{i=1}^{n} X_i \quad (3)$$

Variance is calculated from the following equation:

$$V(X) = \frac{1}{n-1}\sum_{i=1}^{n}(X_i - \bar{X})^2 \quad (4)$$

The reason why we use n_1 instead of n, is that it is assumed that all of values are not available but a number of these values are selected and put in x. that means x is sample collection not total data. So, n_1 is near to real variance.
The disadvantage of variance is that it is not comparable with attribute. If we assume that attribute is according to meter$X_i - \bar{X}$, it is according to meter. But $(X_i - \bar{X})$ is according to squere meter.
So, it is not comparable. This disadvantage of variance is standard deviation foe elimination of square root.

$$\sigma = \sqrt{V(X)} \ (5)$$

The criteria mentioned above, it is only information related to one dimension. In fact, it does not give any knowledge about relation among different dimensions, we can find the relations data different dimension with using of covariance. Let us have the other strain from number which is shown with Y. covariance between x.y:

$$Cov(X,Y) = \frac{1}{n-1}\sum_{i}(X_i - \bar{X})(Y_i - \bar{Y}) \quad (6)$$

: According to whether positive or negative, cov(x,y) shows that x,y change in one direction or in two opposite direction.
If large values x,values larger than average, are related to large values y or small values x with large value y, it will be negative covariance. Or put it in another way,
If cov(x,y)…, x,y change similarly
If cov(x,y)…, x,y change similarly
If cov(x,y)= 0, it is not calculated that x,y are independent(in case of x,y are normal)
Covariance among all of dimension can be calculated in pairs and saved in one matrix. This matrix is called matrix covariance one symmetrical square matrix.
For example, if we have three dimension namely, x, y, z, matrix covariance is:

$$C = \begin{pmatrix} cov(x,x) & cov(x,y) & cov(x,z) \\ cov(y,x) & cov(y,y) & cov(y,z) \\ cov(z,x) & cov(z,y) & cov(z,z) \end{pmatrix} \quad (7)$$

The reason for being symmetrical matrix covariance is that:

$$Cov(X,Y) = \frac{1}{n-1}\sum_{i}(X_i - \bar{X})(Y_i - \bar{Y})cov(y,x) \ (8)$$

Principal component analysis (PCA) is foundered on the basis of Pearson. Generally, PCA' aim is to produce new variables (well arranged on the basis of importance rank). In fact, it is as compound of primary variables and without correlation. Geometrically, PCA can be the periods of primary characteristic axial to new vertical axial.
In case of being arranged on the basis of primary variables variance, one of the main reasons for carrying out PCA is that we can find smaller groups from basic variables which can describe data. For doing this study, the first main multi element can form many scattering from primary data. Reduction data dimension is of importance.
However, data display in smaller dimension is not necessarily to get to interpretable data.

ISSN 2229-5518

8 Classification by Using Support Vector Machine (SVM)

Support vector machine is a binary classifier that separates two classes with a linear boundary. One of the advantages of SVM is classes' separations due to their dispersion. This technique as a separator shows an efficient optimal operation in data classification. In the training phase, for a number of feature vectors with known class, the optimal boundary between two classes is determined. In SVM, by using all training vectors and an optimization algorithm, a number of training examples are obtained that make up the class boundaries. That these training vectors are support vectors. Training examples are considered as support vectors. And they have the minimum distance to the boundary of class boundary and by using them; an optimal linear separation boundary is obtained to separate two classes. [21] for investigating SVM, we assume that data are formed of two separate classes and the number of training vectors L are xi and two classes are labeled with yi=1 and yi=-1. Two linear separation boundaries between two classes are chosen so all samples of first class with label yi=1 are places on one side of boundary and all samples of second class with label yi=-1 are placed on the other side of boundary. Also the closest distance between training vectors of two classes should be maximum in the perpendicular direction on separation boundary. A linear separation boundary can be obtained from the equation w.x+b=0, generally that x Is a point on the separation boundary, w is the perpendicular vector on separation boundary and w.x is the inner product of two vectors. So, if xi is a support vector yi (w.xi+b) =1 and if xi is not support vector, we have yi(w.xi+b)>1. In figure 5-1, optimal linear boundary is shown between two separate classes.

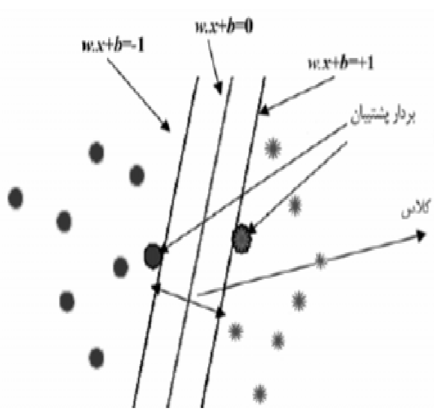


Figure 2: optimal linear boundary for two separate classes

Due to choosing b and w values, there are many separation boundaries that separate two classes with zero error. To obtaining an optimal separation boundary, first we obtain two nearest training samples between two classes and then the distances between two samples are calculated in the perpendicular direction on two boundaries that separates two classes. A boundary which has the maximum margin between two classes is the optimal separation boundary and it is obtained from the following equation:

$$\min_{w,b} \frac{1}{2}\left|\mathbf{w}\right|^2 \quad (,y_i\ (\mathbf{w.x_i} + b) \geq 1 \qquad i=1,\dots,L(9)$$

To solve the above optimization problem, the method of indefinitely Lagrange multipliers can be used. And in this method, the above optimization is as follows that iα is the Lagrange coefficient.

$$\max_{\alpha_1,\dots,\alpha_L}\left[-\frac{1}{2}\sum_{i=1}^{L}\sum_{j=1}^{L}\alpha_i y_i (\mathbf{x_i}.\mathbf{x_j}) y_j \alpha_j + \sum_{i=1}^{L}\alpha_i\right]$$

(10)

α0 i = 1,…..,L_i ≥

$$\sum_{i=1}^{L}\alpha_i y_i = 0$$

(11)

By solving above equation, the w value is equal to $w = \sum_{i=1}^{L}\alpha_i y_i \mathbf{x}_i$. And iα for support vectors is higher than zero and for non-zero vectors equal to zero. Then, after finding w by using the following equation, the b valie is calculated for each support vector and final b is obtained by averaging all obtained b.

$$[y_i\ (\mathbf{w.x_i} + b) - 1] = 0 \qquad i=1,\dots\dots,L \qquad (12)$$

α i

The final separation equation is obtained from the following equation:

$$\boldsymbol{f}(\mathbf{x}) = \text{sgn}\ (\mathbf{w.x} + b) \qquad (13)$$

By using above method, we can obtain the linear boundary between two separate classes but if two classes are UN -separable, separating classes with linear separation boundary has error. In this case, first, we transfer the feature vectors of different classes through a non-linear Φ to a space with higher dimensions and then by using linear function or hyper plane, this process is done in new class separation space. In new space and by using previous equations and replacing xi with xi (Φ) and considering an error value for each vector, the optimal separation boundary is calculated. Optimizing problem for finding optimal separation boundary, reach to this equation:

$$\max_{\alpha_1,\dots,\alpha_L}\left[-\frac{1}{2}\sum_{i=1}^{L}\sum_{j=1}^{L}\alpha_i y_i (\Phi(\mathbf{x_i}).\Phi(\mathbf{x_j})) y_j \alpha_j + \sum_{i=1}^{L}\alpha_i\right]$$

0 i = 1,…..,L_i ≥αC ≥(14)

ISSN 2229-5518

$$\sum_{i=1}^{L} \alpha_i y_i = 0 \quad (13)$$

In the above, C is a constant number and specifies the error value and if a larger number is selected, a greater value of error is obtained. In the above equation, instead of using Φ, we use a core function, as follows:

$$k(\mathbf{x}_i, \mathbf{x}_j) = \Phi(\mathbf{x}_i)\,\Phi(\mathbf{x}_j) \quad (15)$$

Sp, instead of) xj (Φ) xi (Φ) the core function k (xi, xj) is considered and optimization problem solving is done. Core functions are symmetric positive definite functions.
Three important functions are used as core function:
1-a polynomial function K(x,y) =(x.y+1)d that d is the degree of polynomial.

2. GRBF Function: $K(x, y) = \exp\left(-\frac{\|\mathbf{x}-\mathbf{y}\|^2}{2\sigma^2}\right)$ that σ is the width of Gaussian function.

3. ring function K(x, y) = tanh (k(x.y) - μ) that K and μ are scale and offset parameters.

Simulation

After receiving data ECG, Spo2 from site physionet, we reduced dimensions with PCA method for 14 members. Then, we selected 6 elements as a fuzzy neural network (input element), then, the feature vector of every person is trained through target related to fuzzy neural network. The numbers of attachment function for every result is reported as follows:

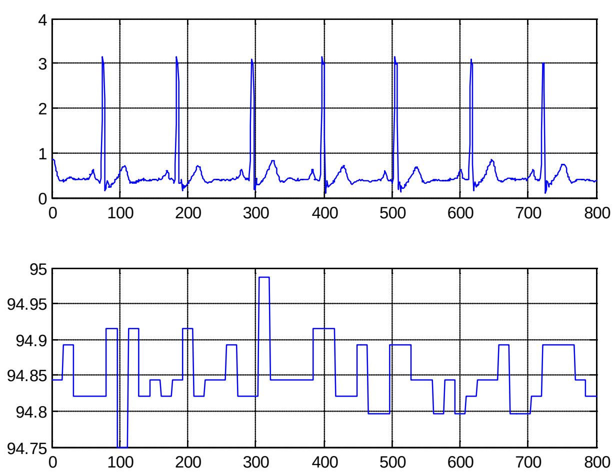


Figure3.Graph 1 is related to ECG and the second graph to SPO2

Target	Classification Target	Classification Output
25	2	2
25	2	2
91	3	3
16	2	2
13	1	1
12	1	1
39	3	3
5	1	1
2	1	1
12	1	1
7	1	1
34	3	3

Conclusion

In this article, we presented new method on the basis support vector machine (SVM) in order to diagnose obstructive sleep Apnea through ECG Spo2 signal features. For achieving our aim, we consider two parameters signs including the rate of oxygen concentration in blood and relative pressure of blood circulation. Then, support vector machine (SVM) was designed in Matlab environment. Moreover, patients' information was applied to network. The result of simulation shows that this designed network diagnose input and instructional data. Furthermore, it predicts sleep Apnea with high accuracy.

References

[1]C. Guilleminault, J. van den Hoed and M. Mitler, "Clinical Overview of the Sleep Apnea Syndromes," In: A. R. Liss, Ed., Sleep Apnea Syndromes, New York, 1978, pp.1-11.
[2] T. Young, M. Palta, J. Dempsey, et al., "The Occurrence of Sleep Disordered Breathing among Middle-Aged Adults," The New England Journal of Medicine, Vol. 328, 1993, pp. 1230-1235.
[3] American Academy of Sleep Medicine, "Sleep Related Breathing Disorders in Adults: Recommendations for Syndrome Definition and Measurement Techniques in Clinical Research," Sleep, Vol. 22, 1999, pp. 667-689.
[4] T. Young, P. E. Peppard and D. G. Gottlier, "Epidemiology of Obstructive Sleep Apnea, a Population Health Perspective," American Journal of Respiratory and Critical Care Medicine, Vol. 165, 2002, pp. 1217-1239.
[5] M. Cabrero-Canosa, E. Hernandez-Pereira and V. Moret-Bonillo, "Intelligent Diagnosis of Sleep Apnea Syndrome," Engineering in Medicine and Biology Magazine,Vol. 23, No. 2, 2004, pp. 72-81.

ISSN 2229-5518
[6] A. H. Khandoker, J. Gubbi and M. Palaniswami, “Automated Scoring of Obstructive Sleep Apnea and Hypopnea Events Using Short-Term Electrocardiogram Recordings,”IEEE Transactions on Information Technology in Biomedicine, Vol. 13, No. 6, 2009, pp. 1057-1067.

[7] G. B. Moody, R. G. Mark, A. Zoccola and S. Mantero,“Derivation of Respiration Signal from Multi-Lead ECGs,”Computers in Cardiology, 1985, pp. 113-116.

[8] C. O’Brien and C. Heneghan, “A comparison of Algorithms for Estimation of a Respiratory Signal from the Surface Electrocardiogram,” Computers in Biology and Medicine, Vol. 37, 2007, pp. 305-314.

[9] J. E. Mietus, C. K. Peng, P. Ch. Ivanov and A. L. Goldberger, “Detection of Obstructive Sleep Apnea from Cardiac Interbeat Interval Time Series,” IEEE Computers in Cardiology, Vol. 27, 2000, pp. 753-756.

[10] M. O. Mendez, D. D. Ruini, O. P. Villantieri, M. Matteucci, T. Penzel, S. Cerutti and A. M. Bianchi, “Detection of sleep Apnea From Surface ECG Based on Features Extracted by an Autoregressive Model,” Annual International Conference of the IEEE Engineering in Medicine and Biology Society—EMBC, Lyon, 2007, pp. 6105-6108.

[11] D. Cysarz, R. Zerm, H. Betterman, M. Frühwirth, M. Moser and M. Kröz, “Comparison of Respiratory Rates Derived from Heart Rate Variability, ECG Amplitude,
and Nasal/Oral Airflow Ambulatory Single-Lead ECG,”Annals of Biomedical Engineering, Vol. 36, No. 12, 2008, pp. 2085-2094.

[12] C. Avcı, S. Beşli and A. Akbaş, “Performance of the EDR Methods; Evaluations Using the Mean and Instantaneous Respiratory Rates,” The 5th International Conference on Bioinformatics and Biomedical Engineering, Wuhan,10-12 May 2011.

[13] L. J. Epstein, D. Kristo and P. J. Strollo, “Clinical Guideline for the Evaluation, Management and Long-Term Care of Obstructive Sleep Apnea in Adults,” Journal of Clinical Sleep Medicine, Vol. 5, No. 3, 2009, pp. 263-276.

[14] F. C. Yen, K. Behbehani, E. A. Lucas, J. R. Burk and J. R. Axe, “A Noninvasive Technique for Detecting Obstructive and Central Sleep Apnea,” IEEE Transaction on Biomedical Engineering, Vol. 44, No. 12, 1997, pp. 1262-1268.

[15] American Sleep Disorders Association Task Force, “The Chicago Criteria for Measurements, Definitions, and Severity of Sleep Related Breathing Disorders in Adults,”Associated Professional Sleep Societies Conference, New Orleans, 1998.

[16] M. E. Tagluk and N. Sezgin, “Classification of Sleep Apnea through Sub-band Energy of Abdominal Effort Signal Using Wavelets + Neural Networks,” Journal of Medical Systems, Vol. 34, No. 6, 2010.

[17] J. R. Stradling, “New Approaches to Monitoring SleepRelated Breathing Disorders,” Sleep, Vol. 18, No. 3, 1996,pp. 77-84.

[18] P. Várady, T. Micsik, S. Benedek and Z. Benyó, “A Novel Method for the Detection of Apnea and Hipopnea Events in Respiration Signals” Transactions on Biomedical Engineering, Vol. 49, No. 9, 2002, pp. 936-942.

[19] J. Y. Tian and J. Q. Liu “Apnea Detection Based on Time Delay Neural Network,” The 27th Annual International Conference of the IEEE Engineering in Medicine and Biology Society, Shanghai, September 2005, pp. 2571-2574.

[20] L. S. Correa, E. Laciar, V. Mut, A. Torres and R. Jane,“Sleep Apnea Detection Based on Spectral Analysis of Three ECG—Derived Respiratory Signals,” Annual International Conference of the IEEE Engineering in Medicine and Biology Society, 3-6 September 2009, pp. 4723-4726.

[21]Chang, Chih-Chung, and Chih-Jen Lin. "LIBSVM: a library for support vector machines." ACM Transactions on Intelligent Systems and Technology (TIST) 2.3 (2011): 27.