J. Basic. Appl. Sci. Res., 3(10)277-280, 2013

ISSN 2090-4304
Journal of Basic and Applied Scientific Research
www.textroad.com

© 2013, TextRoad Publication

Moment Based Thresholding in Binary Classification

A.R. Tayfe Ayremlou*, Ali Farzan, Mina Khoshrangbaf

Computer Dept., Shabestar branch, Islamic Azad University, Shabestar, IRAN

Received: June 20 2013
Accepted: July 17 2013

ABSTRACT

Binary classification is the process of assigning one of two different labels or classes to each of data patterns. One of simplest methods for classifying data patterns is thresholding and specifying best threshold value is a major issue in these algorithms. This paper proposes a thresholding algorithm in which threshold value is determined by using third order momentum values. Performance of the algorithm is evaluated based on the classification accuracy and compared with well known MiniMax algorithm. Results reveal that the proposed algorithm achieves higher accuracy with respect to the MiniMax.

KEY WORDS: Threshold Value, Momentum, Binary Classification, Skewness, Probability Density Function, Expected Value

I. INTRODUCTION

Let L be a finite and non-empty set of labels or classes. In a classification task involving L, each input instance or data pattern has only one single label or class belonging to L. This kind of classification arises in different fields. In many databases, text documents, videos, music or movies are tagged with labels. In a multi-dimensional feature spaces, classification of the data patterns to the various classes is based on the whole or major part of the features [1-3]. But in some cases, just one simple feature carries most of variances between data samples and so, can be simply used for classifying data patterns. Classification based on simply one feature can be done by different methods [4-6]. Among these methods and one of the most common forms of them, is thresholding. During the thresholding process, individual patterns are marked as "Class-A" if their value of the appropriate feature is greater than some threshold value and as "Class-B" otherwise [7-8]. This type of classification is assumed as the basic and simplest form of classification. Moreover, thresholding can be used as an intermediate step in a large framework which is developed for complex classification applications. Results of the classification are strongly related to the estimated threshold value and exact computation of it is a major issue in the context. There are some methods in specifying the threshold value. In between, receiver operating characteristic (ROC) curve is a powerful and popular method [9]. Despite of its predominance in estimating the threshold value, it never considers the moment parameters in probability density function (PDF) of features in specifying the perfect threshold. This paper propose a new method based on the moment values in computing the threshold.

Usually, a threshold-based classifier corresponds to a unique feature and the classifier is determined by thresholding the feature according to a certain criterion. The criterion is often related to the training error. Threshold-based classifiers have been widely used for object detection and recognition. For example, Viola's face detection algorithm [10] uses Haar-like rectangle features and the difference value between two rectangles, together with a proper threshold and a sign to form a base classifier. If the difference is larger than the leaned threshold, the pattern in question is classified as non-face, and otherwise as face.

Yang et al. [11] proposed a method to employ the difference of Gabor features and compute the optimal threshold to recognize faces. In addition to Gabor features, local binary pattern (LBP) [12] features have also been widely used in threshold-based object recognition [13-15]. Another method is proposed by Rodriguez and Marcel to adopt pixel-based classifiers for face verification [16]. The features in their method are simply raw pixel values. If the pixel value exceeds a learned threshold, a decision is made to claim that the face is either client or imposter.

One of the major issues in thresholding algorithms is to specify the best threshold value. Previous researches use various parameters and methods such as Entropy [17-18], Covariance [19-20] or DBT [21] for estimating the threshold values. Here in this paper, momentum values are used for calculating a perfect threshold value. That is, to achieve a threshold value by which we can design more accurate classifiers. The proposed method outperforms well known miniMax thresholding method by exploiting the central moment values. As we know, the moments of any random variable or feature characterize the properties of probability density function. These properties convey lots of information about the distribution of feature values and are so helpful in making decision on the best value of

***Corresponding Author**: Alireza Tayfe Ayremlou, Computer Dept., Shabestar branch, Islamic Azad University, Shabestar, IRAN. Tel.: +98 (914) 3636148; Email: t_ayremlou@yahoo.com

threshold value. The main goal of the proposed method is to involve these properties in specifying the perfect threshold value by which we can make most accurate decision of classifying the feature vectors.

Results of evaluating the proposed method and comparing it with the commonly used popular MiniMax method reveals that if the properties of the probability density function are taking into account, we can compute the threshold value exactly.

II. METHOD

Proposed method is developed based on the moment values. In order to define the moments, we have to define the expected values first. The expected value of a function $g(x)$ of a continuous random variable is defined as [22]

$$E[g(x)] = \int_{-\infty}^{\infty} g(x)p(x)dx$$

If the random variable is discrete the definition becomes

$$E[g(x)] = \sum_{i=1}^{n} g(x_i)P(x_i)$$

According to the above definitions, we can define the nth central moment of random variable, x, as [3]

$$\mu_n = E[(x-m)^n]$$

Where m is the mean value of random variable, x.

In general, moments are used to characterize the probability density function of a random variable. For example, the second, third, and fourth central moments are intimately related to the shape of the probability density function of a random variable. The second central moment (the centralized variance) is a measure of spread of values of a random variable about its mean value, the third central moment is a measure of skewness (bias to the left or right) of the values of x about the mean value [23], and the fourth moment is a relative measure of flatness.

Third central moments along with the mean values of selected feature in two classes are used to estimate the threshold value. Detail description of the algorithm is so that if the third central moments in both classes have the same signs (either negative or positive), the skewness of probability density function for both classes are toward the same directions. In this case, the threshold value is set according to the difference between mean of the classes. If this difference is greater than 0.05 the threshold is set to the average value of the mean of two classes. But if it is smaller or equal to 0.05, the threshold is set to the mode of the feature values through the both classes. In the other hand if the signs are opposite, the skewness of two classes are in opposite direction and so the threshold value is set to the smallest mean value of the feature along the two classes.

The algorithm can be formulated according to the following equations

$$\gamma = \{ \begin{array}{ll} 0 & equal \quad signs \\ 1 & opposite\ signs \end{array}$$

$$\delta = \{ \begin{array}{ll} 0 & |m_1 - m_2| < 0.05 \\ 1 & |m_1 - m_2| > 0.05 \end{array}$$

And the threshold value is

$$\left(\gamma \times \min_i(m_i)\right) \quad + \quad \left((1-\gamma) \times \delta \times \frac{m_1+m_2}{2}\right) + \left((1-\gamma) \times (1-\delta) \times {\scriptstyle Mod(x_i)}\right)$$

III. RESULTS

Performance of the algorithm is evaluated by using two different popular datasets. One of them is a heart disease dataset (Adapted from http://archive.ics.uci.edu/ml/datasets/ Statlog+%28Heart%29 (Accessed 5/6/2012)) and the other one is a vertebral dataset (Adapted from http://archive.ics.uci .edu/ml/datasets/Vertebral+Column# (Accessed 5/6/2012)). The former one includes 13 features and the later one has 6 features. Threshold values for all of the features are estimated using the proposed method and also well known MiniMax algorithm, and accuracies of classification are estimated in all the cases. Figure. 1 and Figure. 2 compare the accuracy values of MiniMax method and the proposed algorithm. It can be revealed that almost in all the cases, our proposed method outperforms the popular MiniMax method and achieves higher classification accuracy. In spite of the fact that maximum achieved accuracy is not so prominent, but it is necessary to mention that these are the results of classification just by using one of the features and so, are acceptable.

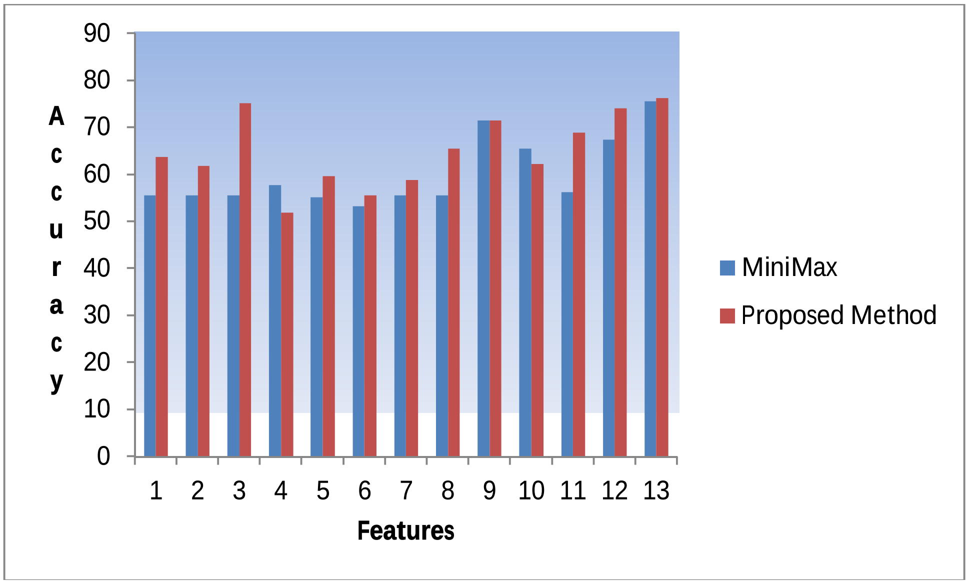

Figure. 1. Classification accuracy of proposed method in comparison to the MiniMax algorithm using Statlog dataset

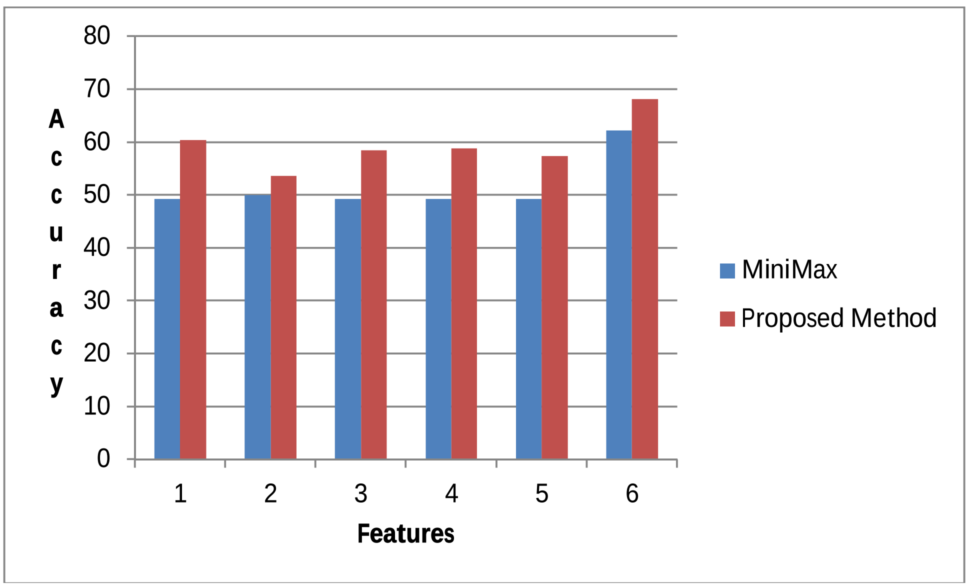

Figure. 2. Classification accuracy of proposed method in comparison to the MiniMax algorithm using Vertebral dataset

Acknowledgment

The authors declare that they have no conflicts of interest in the research.

IV. REFERENCES

[1] Ramón Quevedo J, Luaces O, Bahamonde A. Multilabel classifiers with a probabilistic thresholding strategy. Pattern Recognition. 2011.

[2] Ramaki AA, Atani RE, Abadi RKI, Tavaghoe M. Enhancement Intrusion Detection using Alert Correlation in Co-operative Intrusion Detection Systems. 2013.

[3] Mousavi B. A Novel 2D Texture Classifier For Gray Level Images. 2012.

[4] Pang Y, Deng J, Yuan Y. Incremental threshold learning for classifier selection. Neurocomputing. 2012.

[5] Mortazavi D, Mashohor S, Mahmud R, Jantan AB, editors. Comparison of 3S multi-thresolding with fuzzy C-means method2009: IEEE.

[6] Mirak HT. A new Energy Efficient Connected Target Coverage for WSNs. 2012.

[7] Poletti E, Zappelli F, Ruggeri A, Grisan E. A review of thresholding strategies applied to human chromosome segmentation. Computer Methods and Programs in Biomedicine. 2012.

[8] Negahban F, Shafieian MA, Rahmanian M. Various Novel Wavelet–Based Image Compression Algorithms Using a Neural Network as a Predictor. 2013.

[9] Drakhshanfar H, Rafsanjani MS, Shojaee M, Reza H. Accuracy of Base Deficit in Diagnosis of Intra-Abdominal Injury in Pediatrics with Blunt Abdominal Trauma. Blood transfusion. 2013;8:6.3.

[10] Viola P, Jones MJ. Robust real-time face detection. International journal of computer vision. 2004;57(2):137-54.

[11] Peng Y, Shiguang S, Wen G, Li SZ, Dong Z, editors. Face recognition using Ada-Boosted Gabor features. Automatic Face and Gesture Recognition, 2004. Proceedings. Sixth IEEE International Conference on; 2004 17-19 May 2004.

[12] Ojala T, Pietikainen M, Maenpaa T. Multiresolution gray-scale and rotation invariant texture classification with local binary patterns. Pattern Analysis and Machine Intelligence, IEEE Transactions on. 2002;24(7):971-87.

[13] Ahonen T, Hadid A, Pietikainen M. Face Description with Local Binary Patterns: Application to Face Recognition. Pattern Analysis and Machine Intelligence, IEEE Transactions on. 2006;28(12):2037-41.

[14] Zhang G, Huang X, Li S, Wang Y, Wu X. Boosting local binary pattern (LBP)-based face recognition. Advances in biometric person authentication. 2005:179-86.

[15] Shan C, Gritti T, editors. Learning discriminative lbp-histogram bins for facial expression recognition2008.

[16] Rodriguez Y, Marcel S. Boosting Pixel-BASED Classifiers FOR Face Verification. 2004.

[17] Kapur J, Sahoo PK, Wong A. A new method for gray-level picture thresholding using the entropy of the histogram. Computer vision, graphics, and image processing. 1985;29(3):273-85.

[18] Farhan M, Martin PT. Benefits of Route Guidance System in a Combined Modeling Framework with Variance in Intervals and Equipped Demand. 2012.

[19] Otsu N. A threshold selection method from gray-level histograms. Automatica. 1975;11:285-96.

[20] Rehman O, Manzoor B, Khan R, Ilahi M, Khan Z, Qasim U, et al. A survey on Indoor Localization Techniques in Wireless Body Area Sensor Networks. 2013.

[21] Chowdhury MH, Little WD, editors. Image thresholding techniques1995: IEEE.

[22] Rabbani S, Kamal N, Salim M. Testing the Weak-Form Efficiency of the Stock Market: Pakistan as an Emerging Economy. 2013.

[23] Kazemi H, Abad EFN. Studying and Identifying Operational Strategies to Improve Compliance Tax in Value-Add Tax System. 2013.