Neurocomputing 99 (2013) 76–86

Contents lists available at SciVerse ScienceDirect

Neurocomputing

journal homepage: www.elsevier.com/locate/neucom

A nonnegative sparse representation based fuzzy similar neighbor classifier

Jie Xu [a,b], Jian Yang [a,*]

[a] School of Computer Science and Technology, Nanjing University of Science and Technology, Nanjing 210094, PR China
[b] Shaoguan University, Shaoguan 512024, China

ARTICLE INFO

Article history:
Received 13 September 2011
Received in revised form
30 March 2012
Accepted 7 June 2012
Communicated by M. Sato-Ilic
Available online 1 July 2012

Keywords:
Fuzzy set theory
Similar neighbor
Classification

ABSTRACT

This paper presents a new classification technique by combining the well-known Sparse Representation based algorithm with the theory of Fuzzy Set. The basic idea of this work is that samples with the same class-labels should be more similar to each other than those with different class-labels. Based on this similarity rule, we first impose the nonnegative coefficient constraints on the sparse representation based algorithm and obtain the desirable similar neighbors. Then by introducing the theory of Fuzzy Set into our work, we construct the fuzzy class membership matrix and then assign the decision membership of the query sample to each class. The class assigned with the dominant decision membership is wanted. The proposed approach is called the Nonnegative Sparse Representation based Fuzzy Similar Neighbor Classification (FSNC). FSNC has the following properties: (a) the neighbor parameter K is not needed to be set in advance, and K is adaptively set by the algorithm itself; (b) similar neighbors are also generated adaptively and contain much more similar properties of the query sample; (c) the degree of similarity of data is clearly recorded in the sparse nonnegative coefficient vector; (d) the fuzzy decision rule is effective and the proposed classifier FSNC is simple. Experimental results conducted on the Wine database from UCI, the AR face database, the CENPARMI handwritten numeral database, and the PolyU palmprint database show that the new proposed classification technique outperforms some other state-of-the-art classifiers.

© 2012 Elsevier B.V. All rights reserved.

1. Introduction

Data classification has been an active research topic in the machine learning field for many years. The goal of data classification is to automatically assign data samples to a set of predefined categories. For a general classification problem, provided that the prior probabilities and the state conditional densities of all classes are known, the Bayesian decision theory works well with satisfying results [1]. However, in real-world applications, the data distribution is unknown in most cases. To overcome this difficulty, many non-Bayesian techniques have been developed. The K-nearest Neighbor Classifier (KNNC) is one of them [2]. In KNNC, a query object is assigned a class-label according to the most common class among its K nearest neighbors. From this point of view, KNNC is a simple, easily interpretable technique and can achieve an acceptable accuracy rate. Despite these advantages, however, the limitation of the standard KNNC is to place equal weights on all the selected neighbors regardless of their respective distance from the query point [3]. To overcome this limitation and further improve the performance of KNNC, Keller introduced the theory of Fuzzy Set [4] into KNNC, and proposed a fuzzy version of KNNC, called the fuzzy KNN classifier (FKNNC) [5]. In FKNNC, for a query sample, the degree of membership to different classes is assigned instead of assigning the query sample to a certain class. This provides us another useful method for classification. In such a way, the accuracy of the classification can be guaranteed even when dealing with those data affected by the varying illumination, view conditions and facial expressions.

Due to the validity of FKNNC, subsequently, many improvements have been developed. For example, Bian and Mazlack [6] introduced the Rough Set theory into FKNNC and proposed the Fuzzy-Rough Nearest-Neighbor Classification approach. Chen et al. [3] optimized the parameter of FKNNC, and proposed a non-parametric classifier, called adaptive Fuzzy K-nearest Neighbor Method.

Although improvements have been made in those extended versions of FKNNC, a common problem they still encounter is using the Euclidean distance as the measure metric. Of course, under the common assumption that the samples with the same class-label are lying on the same manifold and the small patches of the manifold are smooth, it is reasonable for them to measure the differences among data in manifold with the Euclidean distance. However, in many empirical classification problems, the complex and uncertain data distributions on manifold make the measured distance unreliable. For example, when the manifolds overlap, two high-dimensional points, which are actually far away from each other in the geodesic distance,

* Corresponding author. Tel./fax: +86 25 8431 7297x307.
E-mail addresses: njxujie@yahoo.cn (J. Xu), csjyang@mail.njust.edu.cn (J. Yang).

0925-2312/$ - see front matter © 2012 Elsevier B.V. All rights reserved.
http://dx.doi.org/10.1016/j.neucom.2012.06.018

are easily misclassified by the nearest neighbors in the sense of the Euclidean distance. Under this circumstance, it seems to be meaningless to take the Euclidean distance as the measure metric to evaluate the relationships among data, since the captured discriminative information are not reliable. What is the proper measure metric for the evaluation of the relationships among data? And how can one use this reasonable measure metric to select the desirable and typical samples as the neighbors? These problems have become the key points for FKNNC and its variants to improve recognition performance. Recently, the Sparse Representation based Classification (SRC) [7] has been proposed and successfully applied to pattern recognition. Motivated by the sparse coefficient vector of SRC, we find a way to address the questions which occurred in FKNNC. Let us have a close look at the decision rule of SRC and find that the combination of the samples associated with the nonzero coefficients can be most approximate to the query sample. On the other hand, a query sample is assigned to the class associated with most nonzero coefficients. Moreover, the nonzero coefficients with different signs give us a direct explanation in sense of reconstruction. As we know, the samples associated with the positive coefficient make a positive contribution to reconstructing the query sample and vice versa. In SRC, the given query sample is described as a combination of the training samples involving both additive and subtractive interactions. It means that the query sample is a combination of some element features (properties). However, the fact that features (properties) can "cancel each other out" using subtraction is contrary to the intuitive notion of combining the parts to form a whole [8]. In other words, if the coefficients are all nonnegative, one can learn the part-based features (properties) from the query sample. The learned part-based properties can be viewed as the "similarity" between the query sample and other samples. Accordingly, the positive items of the coefficient vector can be treated as the degree of similarity between the query sample and the corresponding samples. The larger the coefficients are, the more similar the associated sample and the query sample are. Thus, the samples associated with the nonzero coefficients should have more or less common properties of the query sample and can be considered as the similar neighbors of the query sample. In contrast with the limitation suffered from the measure metric with the Euclidean distance, selecting the samples associated with the positive coefficients as the similar neighbors seems to be more reasonable. More importantly, the similar neighbors and their quantity are generated adaptively without artificially setting. As a result, for different query samples, the number of similar neighbors is different. This phenomenon accords with the actual distribution of data and can reflect the intrinsic structure of manifold.

We denote the sparse representation based algorithm with nonnegative constraints as the Nonnegative Sparse Representation (NSR) based algorithm. The desirable nonnegative coefficient vector can be obtained by the NSR algorithm. The similar neighbors are associated with the items of the positive coefficient vector. After obtaining the favorable similar neighbors, we introduce the theory of Fuzzy Set into our work and propose a simple yet effective classification technique called Nonnegative Sparse Representation based Fuzzy Similar Neighbor Classifier (FSNC). FSNC is fundamentally based on the FKNNC algorithm. However, unlike FKNNC taking the Euclidean distance as the measure metric to choose the desirable neighbors, FSNC takes the similar features into account and selects the samples which have similar properties of the query sample as the neighbors. These selected samples in most cases are not the nearest neighbors of the query sample in the sense of the Euclidean distance, while they have similar features as the query sample and are beneficial for classification. The effectiveness of the proposed FSNC is evaluated on the Wine database from UCI, the PolyU database, the CENPARMI handwritten numeral database and the AR face database.

The rest of the paper is organized as follows: Section 2 describes the Fuzzy K-nearest Neighbor Classifier (FKNNC) and the Sparse Representation based Classifier (SRC). Nonnegative Sparse Representation based Fuzzy Similar Neighbor Classifier (FSNC) is proposed in Section 3. The advantages of FSNC over the related methods are shown in Section 4. Section 5 clearly shows the role of the fuzzy class membership matrix U in connecting the query sample and the categories. The experimental results are shown in Section 6. The conclusions and our future works are described in Section 7.

2. Related works

2.1. Fuzzy K-nearest neighbor classifier (FKNNC)

FKNNC is developed based on the K-nearest neighbor (KNN) algorithm. Different from most classification approaches, FKNNC assigns the class membership to a query sample by using the theory of Fuzzy Set, instead of assigning the query sample to a particular class [5]. Specifically, let $X=[x_1,\ldots,x_N]\in R^{D\times N}$ where D is the dimension of data and N is the size of the data set. The degree of membership of sample x_j, $j=1,\ldots,N$, in each class is recorded by a element of the fuzzy class membership matrix, $\hat{U}=\{\hat{u}_{ij}\}_{c\times N}$, and satisfies the following conditions:

$$\sum_{i=1}^{c}\hat{u}_{ij}=1 \text{ and } <\sum_{j=1}^{N}\hat{u}_{ij}<N,$$

where c is the number of classes. The fuzzy class membership matrix $\hat{U}$ plays an important role in the FKNNC algorithm. Now we present the specific algorithm of FKNNC as follows:

Step 1: Give a K-nearest neighbor parameter K in advance and search the K-nearest neighbors for each training sample from the whole training set.

Step 2: Construct the fuzzy class membership matrix $\hat{U}$, whose element $\hat{u}_{ij}$ denotes the degree of membership to the ith class with respect to the jth sample. Specifically:

$$\hat{u}_{ij}=\begin{cases}0.51+0.49(n_{ij}/K) & \text{If } i \text{ is the same as the label of the } j\text{th sample}\\ 0.49(n_{ij}/K) & \text{Otherwise}\end{cases}$$

where n_{ij} stands for the number of the neighbors of the jth sample with the class-label of i.

Step 3: Assign the decision membership to the query sample y using the following formula:

$$\hat{v}_i(y)=\frac{\sum_{j=1}^{K}\hat{u}_{ij}(1/\|y-x_j\|^{2/(m-1)})}{\sum_{j=1}^{K}(1/\|y-x_j\|^{2/(m-1)})}, \tag{1}$$

where m is the fuzzy strength parameter, which is used to determine how heavily the distance is weighted when calculating each neighbor's contribution to the membership value, and its value is usually chosen as $m\in(1,+\infty)$.

Step 4: Classification. The decision rule of FKNNC is if $\hat{v}_s(y)$ is maximum among the c values of $\hat{v}_i(y)$, $i=1,\ldots,c$, the query sample y will be assigned to the sth class.

2.2. Sparse representation based classifier (SRC)

The basic idea of SRC [7] is to represent a query sample as a sparse linear combination of all training samples; the sparse nonzero representation coefficients are supposed to concentrate on the training samples with the same class-label as the query sample. Specially, given a data matrix $X=[x_1,\ldots,x_N]\in R^{D\times N}$, for a query sample y, we represent y in an over complete dictionary

Download English Version:

https://daneshyari.com/en/article/407883

Download Persian Version:

https://daneshyari.com/article/407883

Daneshyari.com