Perceptual Relativity-based Local Mean Center Classifier

Xianfa Cai, Guihua Wenb, Jia Weib, Zhiwen Yub
and Yongming Caia

ABSTRACT

Despite \( k \) -nearest neighbors and its variants perform very well in many applications, they usually suffer from such limitations as sensitivity to noisy, sparse and imbalance data which results to a dramatic performance degradation. However human cognition has its unique ability to deal with these issues from the different perspective. Motivated by this character, this paper proposes a feasible strategy called Perceptual Relativity-based Local Mean Center Classifier by using the relative transformation to local mean center classifier (RLMC). Firstly, relative transformation will be performed over the training samples to build the relative space and find \( k \) nearest neighbors in the relative space. The advantage of relative transformation is that it improves the distinguishing ability among data points and diminishes the impact of noise on classification. Experimental results on both real and simulated data suggest that the proposed approach often gives the better results in classification and robustness.

INTRODUCTION

In machine learning, computer vision or pattern recognition, classification is a very common but important task which identifies to which of a set of categories a new sample belongs. Series classification algorithms have been proposed for decades, among which \( k \) -nearest neighbors (KNN) \([1, 2]\) is one of the well-known local classifier. KNN not only does not need to make any assumptions on the overall distribution of the training data but just need only one free parameter such that it’s very simple but suitable for problems. Despite KNN algorithm often performs very well in many applications, one of the main disadvantages of KNN is that each sample is equally important in classification. So it will be very difficult to classify when the distance information among its nearest neighbors is not always negligible and becomes very large outside the region of high density. Therefore, researchers have worked out their way to apply complex distance functions trick to

\(^{a}\)School of Medical Information Engineering, Guangdong Pharmaceutical University, Guangzhou, 510006, China

\(^{b}\)School of Computer Science and Engineering, South China University of Technology, Guangzhou 510006, China
select more suitable nearest neighbors, such as locally adaptive distance[3, 4], the kernel trick[5], neighbor counting[6], and data gravitation[7]. Besides distance trick, researchers also establish some efficient decision rules. Keller et al [8] proposed FKNN which applied the fuzzy set to KNN classifier. T. Denoeux [9, 10, 11] used dempster’s rule to improve the classifier. However, they are still confronted with many problems[12]. Another disadvantage of the classical KNN classifier is that it can be frequently influenced by the imbalance problem, where the data in one class heavily outnumbers the data in another class and the class boundary can be skewed towards the class with few data samples [13]. In this case, the performance of KNN easily decreases. Y. Mitani et al [14, 15] designed a LMC classifier which uses the categorical k nearest neighbors of the query sample to calculate the local mean centers per class and to classify the unseen query sample in terms of the minimum distance between the query and these centers. Beyond the imbalance problem, the method can also deal with noises and small sample size problem well. Boyu Li et al [16] present another classifier based on local probability centers (LPC) which considers the situation that class conditional distributions are nonseparable and overlapping or the local hyper plane [17,18]. However in real world life, there are abundant of the sparse, noisy and imbalance data which will influence the performance of these classifiers. Considering that they are sensitive to the selection of neighborhood, a novel algorithm of perceptual relativity-based local mean center classifier (RLMC) is proposed from a different perspective in this paper, to overcome this problem by utilizing the perceptual relativity in terms of cognitive psychology to reduce its negative influences.

There are two main contributions of RLMC in this paper: (1) The perceptual relativity has been applied to improve the performance of classification on the sparse, noisy or imbalance data, indicating the possibility of other perceptual laws in cognitive psychology being considered for classification. (2) LMC use k nearest neighbors of the query sample from the same class to compute the local center and then to classify the query sample in terms of the distance between the query sample and each center. Applying the idea of relative transformation to LMC, we design a new classifier called RLMC that applies relative transformation to select k nearest neighbors from each class and then perform the classification. In RLMC, firstly, relative transformation will be performed over the training samples to build the relative space and find k nearest neighbors in this space. The advantage of relative transformation is that it improves the distinguishing ability among data points and diminishes the impact of noise on classification.

The rest of this paper is organized as follows: Section 2 presents some basic concepts. A novel method is designed in Section 3. The proposed method is evaluated through experiments in Section 4. The paper is concluded with a summary and discussion of possible future work in Section 5.

**ELEMENTARY CONCEPTS**

In this section, we present the concept that our approach is based on and which will serve as building block.
Relative Transformation

Compared with machine classification, human being has natural ability in classification on the sparse, noisy and imbalance data, from which we can get some inspiration. To nicely perform the classification on this kind of data, it’s natural for machine to learn from human being. The existing classification approaches, such as those to recognize faces, gene expression data, require hundreds if not thousands of samples for training, while human visual recognition just need train through very few samples[19]. This is because humans routinely classify objects according to both their individual attributes and membership in higher order groups, where individual attributes may be influenced and regulated by their group [20]. This can be illustrated from Fig.1.

![Figure 1. Human visual perception is relative.](image1.png)

It can be observed from Fig. 1 that the circle $x$ looks bigger than its original size as it is surrounded by smaller circles while the circle $y$ appears smaller than its original size as it is surrounded by bigger circles. Consequently, when we observe $x$ and $y$ simultaneously, $x$ is perceived to be bigger than $y$, although they are of the same sizes in fact [21]. This cognitive characteristic is very important for us to distinguish an object from its surrounding objects and can be then formalized using geometry model to process the data more efficiently. One way is to define a transformation on the original space to build a new space whose dimensions are composed of all points in the original space. The newly created space is called the relative space and can be generated through relative transformation. Relative transformation can make the data more distinguishable [22]. Some data can be distinguishable in the relative space while they cannot be identified in the original space. The relative transformation is also simple and efficient in dealing with noisy data or outliers which can be illustrated in Fig.2.

![Figure 2. Relative transformation on noisy data where (a) Original space (b) Relative space.](image2.png)
The Fig.2 demonstrates that the point $x_i$ may be regarded as a noisy point or an outlier in the original space since it is far away from the other three points which is consistent with human perception. However, $d(x_i,x_j)$ is equal to $d(x_j,x_i)$ in the original space, which means that the point $x_j$ has the same possibility with the point $x_i$ to be taken as a nearest neighbor of the point $x_i$. This is inconsistent with human perception. In contrast, in the relative space, the outlier or noisy point becomes further away from the normal points, so it can be recognized easily. Furthermore, it may also make points which originally lie on the same surface of the manifold closer to each other and points from the different surfaces further away from each other, which is especially useful to the sparse data. Finally, this approach has a simple mathematical basis and it allows a compact mathematical description of arbitrarily shaped neighborhood in the original space. The relative transformation can be formulated as following: $f^*: X \rightarrow Y \subseteq R^n, y_i = f^*(x_i) = (d_{ij},...,d_{ik}) \in Y, d_j = \|x_i - x_j\|$

where $n$ is the number of elements in dataset $X$, the point $x_i$ in the original space is mapped to the point $y_i \in R^n$ in the relative space.

**OUR PROPOSED METHOD**

**RLMC**

This section presents a new classifier that applies the relative transformation to LMC, called RLMC. It finds $k$ nearest neighbors for the query sample from each class and then performs the relative transformation over all these nearest neighbors to build the relative space. Subsequently, each local mean center is constructed in the relative space. The class label of the query is assigned according to the distance between the query and the local mean center of each class.

Algorithm RLMC $(x,X,k)$ /* $x$ be the test sample, $X$ be training samples, $k$ be the selected nearest training samples from each class */

Step 1: Find $k$ nearest neighbors for the test sample $q$ from each class $\omega$, denoted as $\omega(\ldots)$

Step 2: Build the relative space by

$(\ldots) = ((\ldots) \cup \{\})$

Step 3: Compute the local mean vector for each class $\omega$ by using $k$ nearest neighbors as follows:

$\bar{x}_j = \frac{1}{k} \sum_{m=1}^{k} x_m, x_m \in x'(q,k)$

Step 4: Compute the distance between test sample $x$ and the local mean vectors $\bar{x}_j$ of the class $w_j$
\[ d_j = \| x - \bar{x}_j \| \]

Step 5: According to the distance \( d_j \) of each class \( w_j \), i.e. \( \{(w_j, d_j)\} \), test sample will be assigned to the class \( w_j \) with minimum distance \( d_{\text{min}} \).

EXPERIMENTS

Experimental Setup

To validate the proposed RLMC on the performance, we compare it with the baseline approaches through experiments on benchmark data sets. These baseline approaches are KNN, FKNN, EKNN, LMC, HKNN and LPC. In experiment, the error rate is taken as the measure of performance. The parameter \( k \) takes the value from \( \{3,6,9, \ldots, 30\} \) while the kernel parameter for LPC and RLMC takes the values from \( \{0.1,0.2, \ldots, 0.9\} \). The parameter \( C \) for HKNN takes the values from \( \{0.1,0.2, \ldots, 0.9\} \). Euclidean distance is taken in all compared classifiers. When classifying, we performed 5 times five-fold cross validations on each data. On each partition, the parameters are determined for each compared classifier through five-fold cross validations on the training samples, and then applied to perform the classification over the testing samples. Finally the average error rate of 5 times is reported.

A. On artificial data sets

To validate RLMC’s robustness to noise, we apply artificial data ring norm data set [23] which is appended with different random Gaussian noises according to the experimental purpose. The ring norm data is a 20-dimensional 2 class classification example where each class is drawn from a multivariate normal distribution. Class 1 has mean zero and covariance 4 times the identity. Class 2 has mean \((2/\sqrt{20},2/\sqrt{20}, \ldots, 2/\sqrt{20})\) and unit covariance.

It can be observed from Fig.3 that RLMC performs best at many noise cases, followed by EKNN. FKNN and KNN become the worst, illustrating they are much sensitive to the noise. The experimental results mean that RLMC is robust to noise disturbance.

To validate RLMC with the better ability to deal with the curse of the dimensionality problem, we do experiments on ring norm data[23] and p-dimensional norm data[16]. It can be observed that RLMC performs best at most dimension on ring norm data, followed by EKNN. As the dimension increases, the average classification error of RLMC decreases to nearly 0 while that of KNN and FKNN go up quickly to almost 80%. The same goes with the p-dimensional norm data.
B. On real data sets

It is highly likely that simulated data are beneficial to classification, but not correspond to real situations that are likely to occur in practice. Thus, in this section, to validate the performance of the competing classification methods, we examine it
through experiments on benchmark data sets illustrated as Table 1. They are Wine, Ionosogere, Glass, Scale, Monks and SPECT. It can be observed from Table 2 that RLMC outperforms most compared classifiers of the real datasets, indicating that its performance is most stable. These results do indicate the significant value of the proposed idea and the classifier.

**Table 1. Data sets used in experiments.**

<table>
<thead>
<tr>
<th>No.</th>
<th>data set</th>
<th>attributes</th>
<th>classes</th>
<th>size</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Wine</td>
<td>13</td>
<td>178</td>
<td>3</td>
</tr>
<tr>
<td>2</td>
<td>Ionosphere</td>
<td>34</td>
<td>351</td>
<td>2</td>
</tr>
<tr>
<td>3</td>
<td>Glass</td>
<td>19</td>
<td>210</td>
<td>7</td>
</tr>
<tr>
<td>4</td>
<td>Vowel scale</td>
<td>4</td>
<td>150</td>
<td>3</td>
</tr>
<tr>
<td>5</td>
<td>Monks</td>
<td>30</td>
<td>569</td>
<td>2</td>
</tr>
<tr>
<td>6</td>
<td>SPECT</td>
<td>18</td>
<td>846</td>
<td>4</td>
</tr>
</tbody>
</table>

**Table 2. AVERAGE CLASSIFICATION ERRORS (%) FOR REAL DATA.**

<table>
<thead>
<tr>
<th>Data</th>
<th>KNN</th>
<th>EKNN</th>
<th>FKNN</th>
<th>HKNN</th>
<th>LMC</th>
<th>LPC</th>
<th>RLMC</th>
</tr>
</thead>
<tbody>
<tr>
<td>Wine</td>
<td>2.80±0.68</td>
<td>2.57±0.64</td>
<td>2.79±0.69</td>
<td>2.12±0.90</td>
<td>2.57±0.63</td>
<td>2.36±0.62</td>
<td>2.13±0.72</td>
</tr>
<tr>
<td>Ionosphere</td>
<td>15.43±0.51</td>
<td>11.85±0.95</td>
<td>15.15±0.50</td>
<td>9.29±1.20</td>
<td>10.48±0.89</td>
<td>10.65±1.14</td>
<td>6.4±0.88</td>
</tr>
<tr>
<td>Glass</td>
<td>32.88±4.16</td>
<td>32.10±0.97</td>
<td>32.18±2.10</td>
<td>38.69±1.17</td>
<td>31.36±2.72</td>
<td>31.76±2.72</td>
<td>31.52±1.5</td>
</tr>
<tr>
<td>Scale</td>
<td>4.32±0.36</td>
<td>4.00±0.38</td>
<td>2.70±0.37</td>
<td>2.18±0.39</td>
<td>2.42±0.41</td>
<td>2.56±0.42</td>
<td>1.90±0.40</td>
</tr>
<tr>
<td>Monks</td>
<td>16.29±0.73</td>
<td>18.61±2.58</td>
<td>20.78±1.12</td>
<td>14.30±1.72</td>
<td>18.66±0.74</td>
<td>20.23±1.58</td>
<td>12.68±1.27</td>
</tr>
<tr>
<td>SPECT</td>
<td>17.75±1.15</td>
<td>17.52±0.70</td>
<td>18.36±1.93</td>
<td>16.33±1.24</td>
<td>17.15±0.47</td>
<td>16.85±0.87</td>
<td>16.40±0.63</td>
</tr>
</tbody>
</table>

**CONCLUSIONS AND FUTURE WORK**

Because of its simplicity, generality and distinguished ability in dealing with sparse and noisy data, the advantage of RLMC achieves amazing result but also is robust to noise. The most importance of RLMC seems that it opens a new direction as a fundamental methodology to develop new local classifiers which simulate all kinds of cognitive laws in terms of cognitive psychology. We will also develop more variants of techniques in terms of more cognitive laws and apply them to local classifiers.

**ACKNOWLEDGEMENTS**

The authors thank anonymous reviewers and editors for their valuable suggestions and comments on improving this paper. The research leading to these results has received the support of the Natural Science Foundation of China under Grant No. 61501128, the PhD Start-up Fund of Natural Science Foundation of Guangdong Province(2015A030310267, 2016A030310300) and Guangdong Province Youth Innovation Talent Project(2014KQNCX139).

**REFERENCES**