Rough Subspace Error Correcting Output Codes

By: Mohammad Ali Bagheri

Dalhousie University
Canada

Dec 13, 2012
1. Multiclass Classification
2. Rough Subspace ECOC
3. Experiments
4. Conclusions
Introduction: Multiclass Classification

- Class binarization

- Decomposition frameworks
  - One-versus-all (OvA)
  - One-versus-one (OvO)
  - Error Correcting Output Codes (ECOC)

(b) One-versus-One Classification
c \((c-1)/2\) classifiers, one for each pair of classes.
Here: + against ~
Error Correcting Output Codes

- Introduction

- Coding
  - Binary vs. Ternary
  - Static vs. Dynamic

- Decoding
Analysis of ECOC

- **Introduction**
  - The focus of the ECOC-based methods: maximizing row and/or column separation

- **Independence**
  - cornerstone of the ECOC framework

- **Accuracy**
Analysis of ECOC (cont.)

- Discussion on the strategies in the ECOC literature for designing independent classifiers

- Analysis of results of
  - (García-Pedrajas and Ortiz-Boyer, 2011)
  - (Dietterich and Bakiri, 1995, Shapire, 1997)

- Therefore, ...
2. Subspace Approach to Error Correcting Output Codes

- Idea of subspace approach

- Core of the proposed method: Subspace ECOC

- Advantages
  - More independent classifiers
  - More accurate classifiers (provided that ...)
  - Longer ECOC codes
Rough set subspace ECOC

Challenges

1) How to find good feature subsets?
2) How many feature subsets should be selected for each dichotomizer?

Out solutions (out of many!)

1) Rough Set Feature Selection
2) 10 feature subsets (reduct sets)
Feature selection & Classification

Sensors's responses

Preprocessing

Binary Classifier 1

Feature selection for the first binary problem

Binary Classifier 2

Feature selection for the second binary problem

... ... ...

Feature selection for the L-1 binary problem

Binary Classifier L-1

Feature selection for the last binary problem

Binary Classifier L

Decoding

Gas type
QuickMultipleReduct algorithm

QuickMultipleReduct (C, D, N)

- C: the set of all conditional features.
- D: the set of decision features.
- N: number of reduct sets per dichotomizer.

1. ∀f ∈ C,
2. Γ(.) = γ_f(D)
3. Sort Γ on descending order
4. For i = 1: N
5. rand = unirand (|C|/2) // generate a uniform random number
6. Init_f = Γ(rand)
7. R_i ← {Init_f}
8. do
9. T ← R_i
10. ∀f ∈ (C - R_i)
11. if γ_{R_i∪f}(D) > γ_T(D)
12. T ← R_i ∪ f
13. R_i ← T
14. until γ_{R_i}(D) = γ_C(D)
15. MultipleReducts{i} = R_i
16. end
17. Return MultipleReducts
Experimental evaluation over benchmark datasets

<table>
<thead>
<tr>
<th>Dataset</th>
<th># instances</th>
<th># features</th>
<th># classes</th>
</tr>
</thead>
<tbody>
<tr>
<td>Abalone</td>
<td>4177</td>
<td>8</td>
<td>3</td>
</tr>
<tr>
<td>Cleafs</td>
<td>4758</td>
<td>64</td>
<td>8</td>
</tr>
<tr>
<td>Cmc</td>
<td>1473</td>
<td>9</td>
<td>3</td>
</tr>
<tr>
<td>Derm</td>
<td>358</td>
<td>34</td>
<td>6</td>
</tr>
<tr>
<td>Ecoli</td>
<td>336</td>
<td>7</td>
<td>8</td>
</tr>
<tr>
<td>Glass</td>
<td>214</td>
<td>9</td>
<td>6</td>
</tr>
<tr>
<td>Lymph</td>
<td>148</td>
<td>18</td>
<td>4</td>
</tr>
<tr>
<td>Mfeat-fac</td>
<td>2000</td>
<td>216</td>
<td>10</td>
</tr>
<tr>
<td>Mfeat-fou</td>
<td>2000</td>
<td>76</td>
<td>10</td>
</tr>
<tr>
<td>Mfeat-kar</td>
<td>2000</td>
<td>64</td>
<td>10</td>
</tr>
<tr>
<td>Mfeat-mor</td>
<td>2000</td>
<td>6</td>
<td>10</td>
</tr>
<tr>
<td>Mfeat-pix</td>
<td>2000</td>
<td>240</td>
<td>10</td>
</tr>
<tr>
<td>Mfeat-zer</td>
<td>2000</td>
<td>47</td>
<td>10</td>
</tr>
<tr>
<td>Optdigits</td>
<td>5620</td>
<td>64</td>
<td>10</td>
</tr>
<tr>
<td>Pendigits</td>
<td>10992</td>
<td>16</td>
<td>10</td>
</tr>
<tr>
<td>Sat</td>
<td>6435</td>
<td>36</td>
<td>6</td>
</tr>
<tr>
<td>Thyroid</td>
<td>215</td>
<td>5</td>
<td>3</td>
</tr>
<tr>
<td>Vehicle</td>
<td>846</td>
<td>18</td>
<td>3</td>
</tr>
<tr>
<td>Vertebral</td>
<td>310</td>
<td>6</td>
<td>3</td>
</tr>
<tr>
<td>Vowel</td>
<td>528</td>
<td>10</td>
<td>11</td>
</tr>
<tr>
<td>Waveforms</td>
<td>5000</td>
<td>40</td>
<td>3</td>
</tr>
<tr>
<td>Wine</td>
<td>178</td>
<td>13</td>
<td>3</td>
</tr>
<tr>
<td>Yeast</td>
<td>1484</td>
<td>8</td>
<td>10</td>
</tr>
<tr>
<td>Zoo</td>
<td>101</td>
<td>16</td>
<td>7</td>
</tr>
</tbody>
</table>
Experimental results

Advanced classification accuracy of the RSS-ECOC in comparison with rival methods
Statistical analysis

(b) MLP Classifier
Conclusions

- Summary of the ECOC analysis
- Research contribution
- When the subspace approach works?
- Computational complexity
Thanks