IEEE ICDM Brussels 10-13 Dec 2012

DALHOUSIE UNIVERSITY
Inspiring Minds

Rough Subspace Error Correcting Output Codes

By: **Mohammad Ali Bagheri**

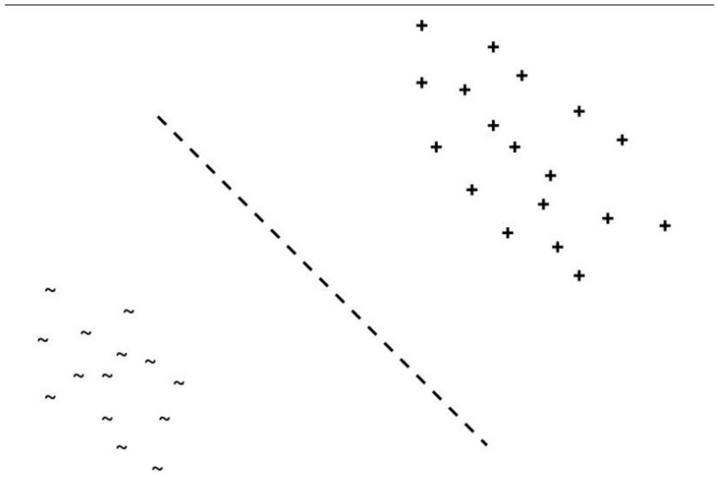
Dalhousie University
Canada

Dec 13, 2012

1. Multiclass Classification
2. Rough Subspace ECOC
3. Experiments
4. Conclusions

Introduction: Multiclass Classification

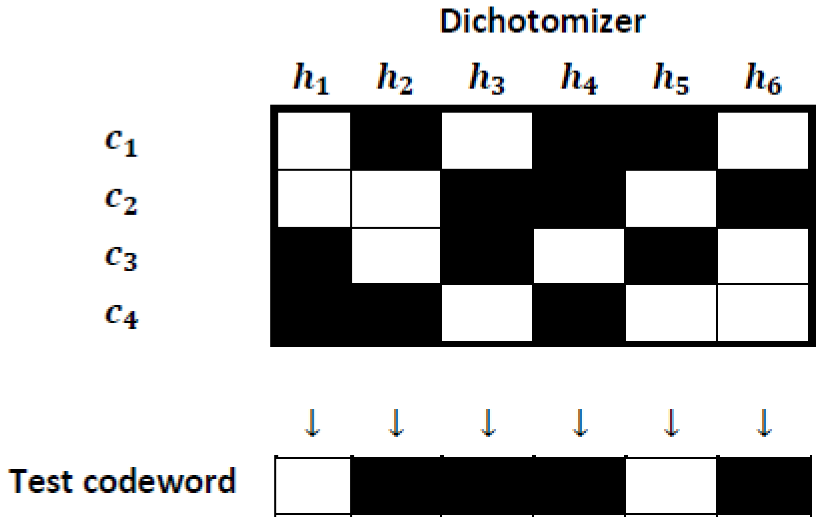
- Class binarization
- Decomposition frameworks
  - One-versus-all (OvA)
  - One-versus-one (OvO)
  - Error Correcting Output Codes (ECOC)



(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

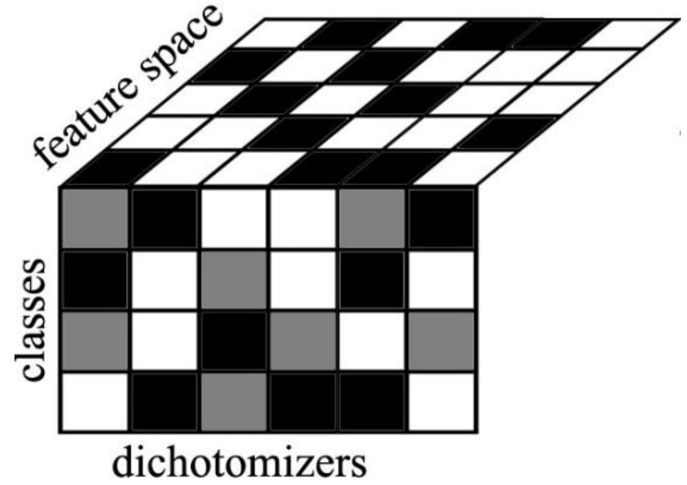
- Analysis of results of
  - (García-Pedrajas and Ortiz-Boyer, 2011)
  - (Dietterich and Bakiri, 1995, Shapire, 1997)

- Therefore, …

2. Subspace Approach to Error Correcting Output Codes

- Idea of subspace approach

- Core of the proposed method: Subspace ECOC

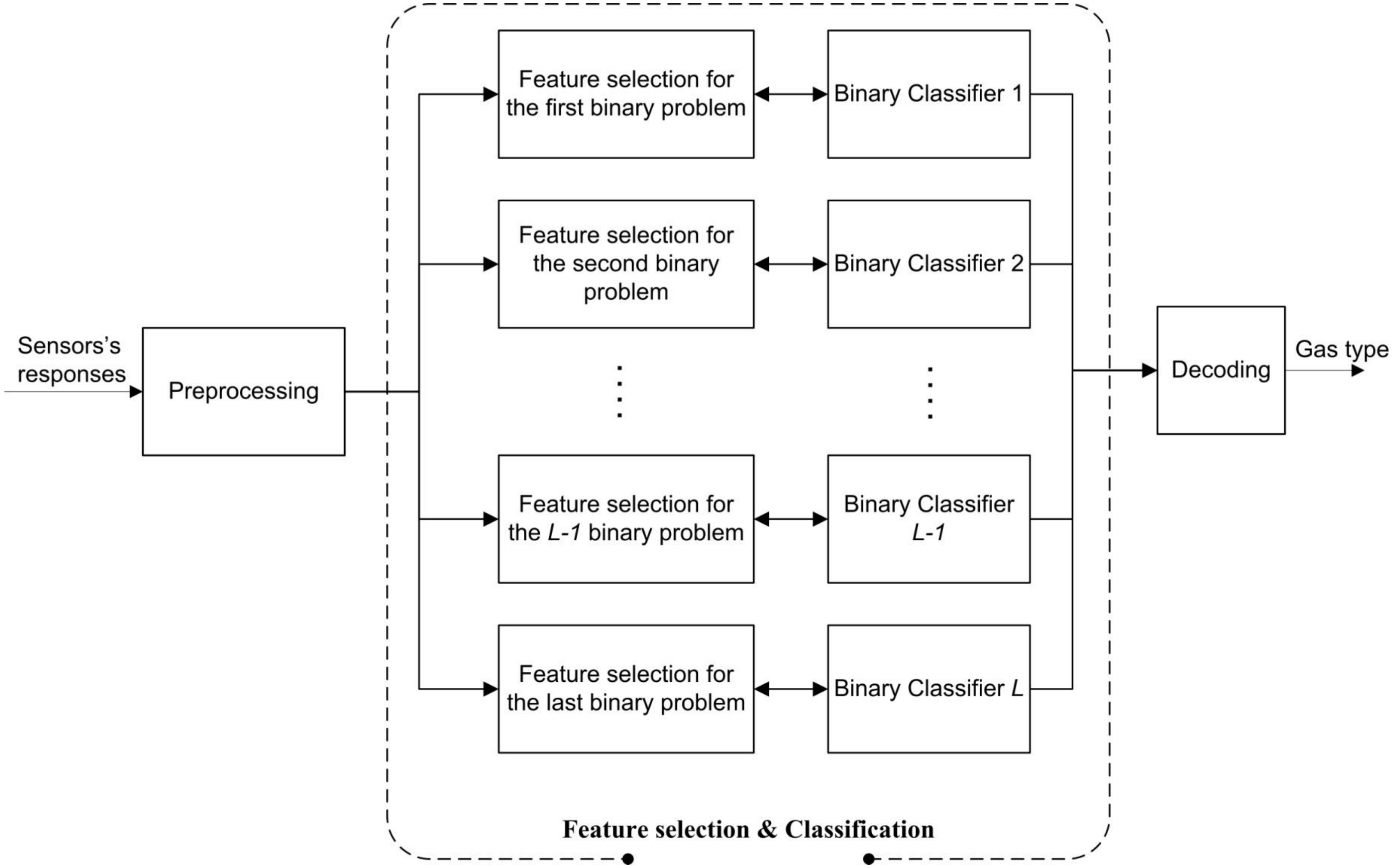


- Advantages
  - More independent classifiers
  - More accurate classifiers (provided that …)
  - Longer ECOC codes

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)


Sensors's responses
Preprocessing
Feature selection for the first binary problem
Binary Classifier 1
Feature selection for the second binary problem
Binary Classifier 2
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
Feature selection & Classification

QuickMultipleReduct algorithm

QuickMultipleReduct (C,D, N)

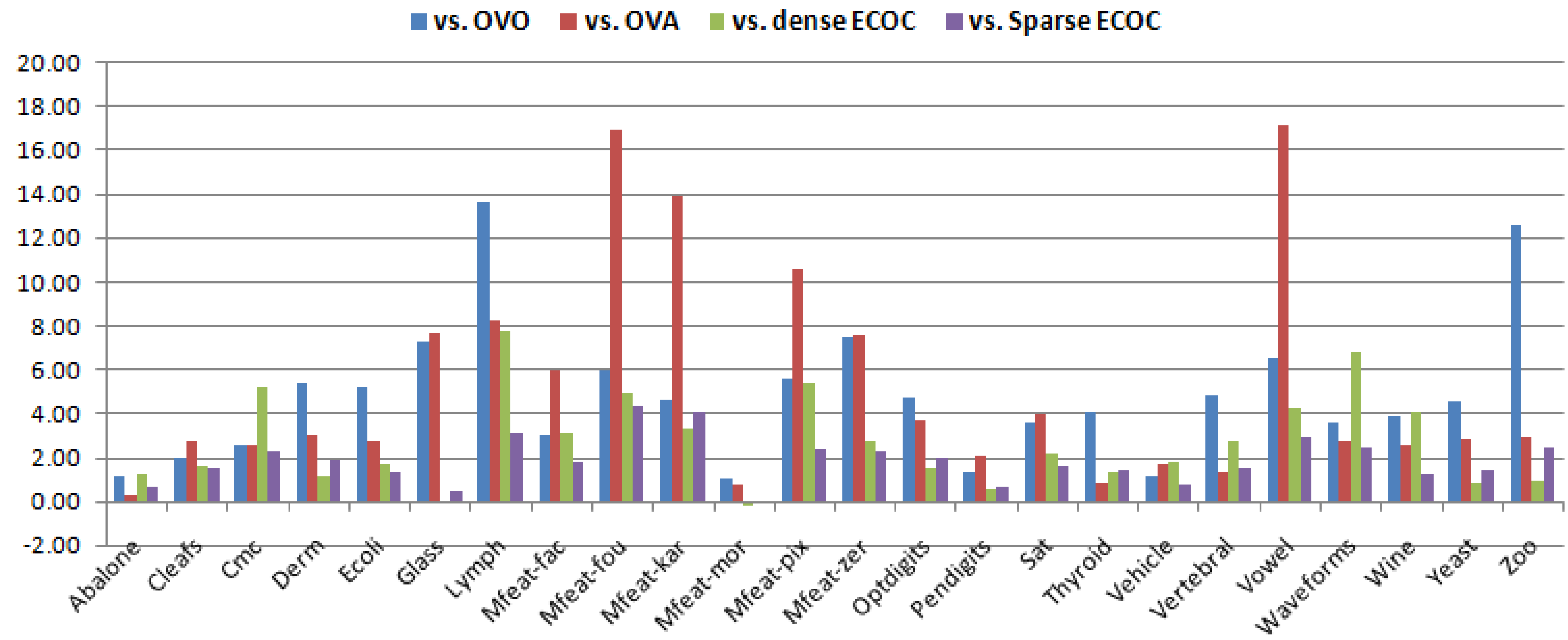
- C: the set of all conditional features.
- D: the set of decision features.
- N: number of reduct sets per dichotomizer.

1. $\forall f \in C$,
2. $\Gamma(.) = \gamma_f(D)$
3. Sort Γ on descending order
4. For $i = 1{:}\,N$
5. rand=unirand (|C|/ 2) // generate a uniform random number
6. $Init\_f = \Gamma(rand)$
7. $R_i \leftarrow \{Init\_f\}$
8. **do**
9. $T \leftarrow R_i$
10. $\forall f \in (C - R_i)$
11. if $\gamma_{R_i \cup \{f\}}(D) > \gamma_T(D)$
12. $T \leftarrow R_i \cup f$
13. $R_i \leftarrow T$
14. **until** $\gamma_{R_i}(D) = \gamma_C(D)$
15. $MultipleReducts\{i\} = R_i$
16. end
17. **Return** $MultipleReducts$

Experimental evaluation over benchmark datasets

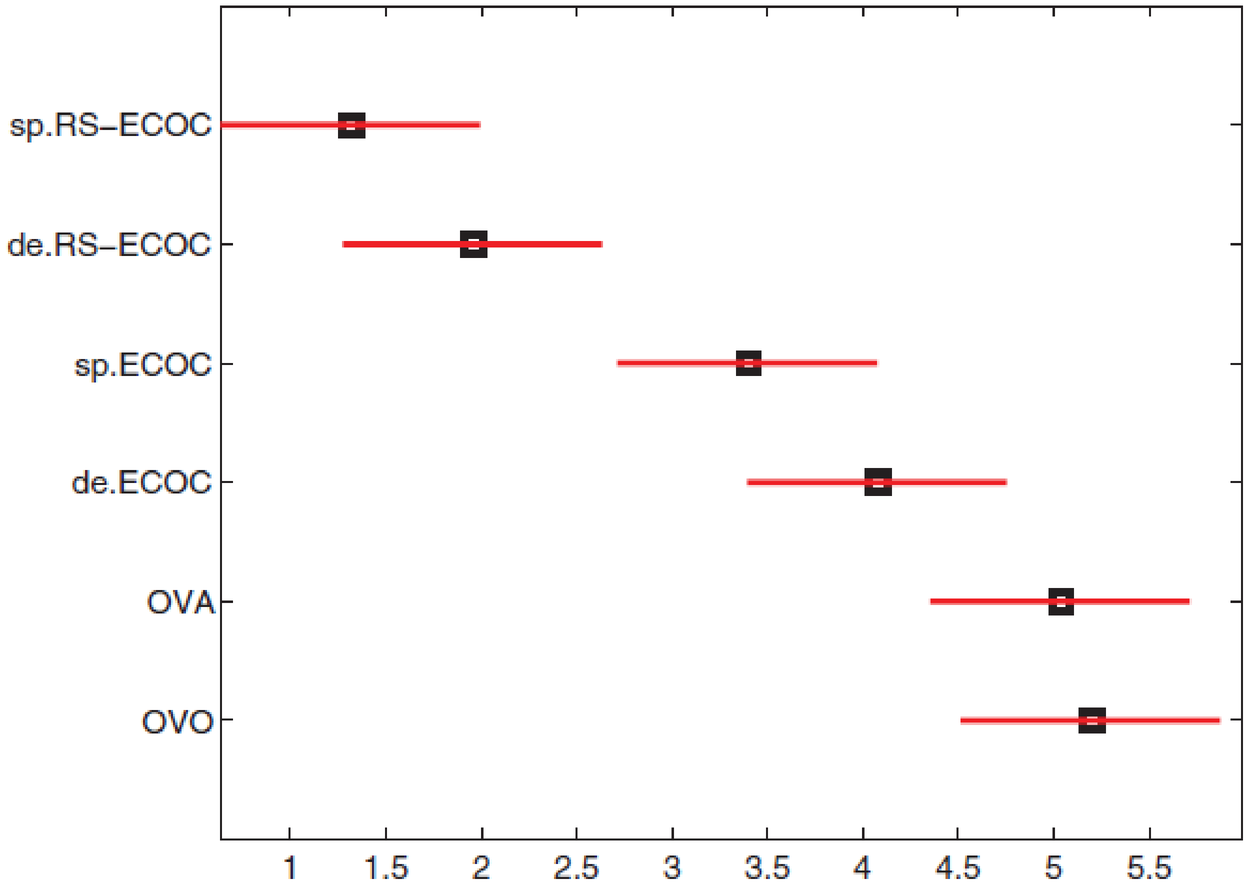
	Dataset	# instances	# features	# classes
1	Abalone	4177	8	3
2	Cleafs	4758	64	8
3	Cmc	1473	9	3
4	Derm	358	34	6
5	Ecoli	336	7	8
6	Glass	214	9	6
7	Lymph	148	18	4
8	Mfeat-fac	2000	216	10
9	Mfeat-fou	2000	76	10
10	Mfeat-kar	2000	64	10
11	Mfeat-mor	2000	6	10
12	Mfeat-pix	2000	240	10
13	Mfeat-zer	2000	47	10
14	Optdigits	5620	64	10
15	Pendigits	10992	16	10
16	Sat	6435	36	6
17	Thyroid	215	5	3
18	Vehicle	846	18	3
19	Vertebral	310	6	3
20	Vowel	528	10	11
21	Waveforms	5000	40	3
22	Wine	178	13	3
23	Yeast	1484	8	10
24	Zoo	101	16	7

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