6B-02

2 値判別器を用いた多値分類方式のシステム評価（続）

平澤　茂一†　雲居　玄道†　小林　学††　後藤　正幸†　稲積　宏誠†††
早稲田大学†　早稲田大学†　湘南工科大学††　早稲田大学†　青山学院大学†††

1　はじめに

1970 年代後半，J. Pearl とA. Crolotte は質問回答システム(QA システム) の記憶容量と誤りのトレードオフについて，情報縮約論を用いて解析・評価した[4]．その結果，各種のQA (Question Answering) システムに対し，小さな誤りを許容すれば大きく記憶容量を削減できる条件を明らかにした．このような条件を「フレキシブル」(Flexible)や「エラスティック」(Elastic)と呼んでいる．筆者らはこれを拡張したトレードオフ評価モデルを提案し[2]，各種の情報システムに適用してきた．これにより，情報システムを設計するに先立って，システムの規模が大きくなるに従い相対的に効率がよい性質「効果的エラスティック」(Effectively Elastic)を持つか否かを判定することが出来る．

本研究は，先の検討結果[3]と同様，表題のテーマで多値分類システムを構成した際，どのような性質を持つか平均的な性能を評価したもので，分類誤り確率を最小化することを狙ったものではない．ここでは[3]に引き続き，対象とするデータセットを文書分類問題ではなくUCI machine learning repositor[5]のベンチマークデータを使用しており，データセットの特性によらず，評価モデルの有効性を検証する．

2 トレードオフ評価モデル

2.1 レートひずみ関数

システム評価モデルは，実在するトレードオフ関係を持つ変数間の振る舞いを評価するために，情報理論におけるレート歪関数 $R=R(D)$に基づいて構築されたている[4]．ここで，R はレート，D は歪で一般に下に凸な減少関数（トレードオフ曲線）である．

2.2 システム評価モデル

トレードオフ関数に新たにシステムの規模を示す L を導入する[4][2]．さらに $R = R(D)$ を D 軸，R 軸ともそれぞれ $D_{\max}$, $R_{\max}$ で正規化する．すなわち，$r = R/R_{\max}$, $d = D/D_{\max}$ とし，関数

$$r = r(d; L) \tag{2.1}$$

に注目する．式(2.1)の r-d 関数を用いて，トレードオフが望ましい特性を示す性質を表 2.1 に示す．

表 2.1：r-d 関数の特性評価

名称	性質
(1) フレキシブル (Flexible)[4]	ある L (> 1) に対し，システム A の r-d 関数がシステム B のそれより原点方向にあるとき，システム A はシステム B より「フレキシブル」であると言う．
(2) エラスティック (Elastic)[4]	任意の d で，r が L の減少関数のとき，「エラスティック」であると言う．
(3) 効果的エラスティック (Effective lastic)[2]	任意の d で，r が L の下に凸な関数のとき，「効果的エラスティック」であると言う．
(4) トリビアルエラスティック (Trivial Elastic)[4]	$d(0;L)$が L の減少関数のとき，「トリビアルエラスティック」であると言う．
(5) マージナルエラスティック (Marginal Elastic)[2]	任意の r において，d が L の下に凸な減少関数のとき，「マージナルエラスティック」と言う．

3 多値分類システムの構成

まず，r-d 関数を用いて多値分類システムを評価するために，表 3.1 に変数の対応表を示す．

表 3.1：多値分類システムの評価（対応表）

情報縮約論	システム評価モデル	多値分類システム
レート (R)	投資コスト(r)	2 値判別器数 (n)
歪 (D)	性能劣化(d)	分類誤り確率 (p_e)
	規模(L)	カテゴリ数 (M)

次に，カテゴリ C の数が M (≥3)である分類問題を考えると，D 個の訓練データは個々に$(\boldsymbol{x},C_i)$の形で与えられる．ここで，C_i は第 i のカテゴリを示す．（カテゴリが未知の）テストデータ $\boldsymbol{y}$ は，訓練データで学習した結果を用いて，$\boldsymbol{y}$ が属するカテゴリ C_i を推定する．図 3.1 に多値分類システムのブロック図を示す．結局，2 値分類器で多値分類を実行する重要な構成部は図中の 2 値マトリックスで示した ECOC-Matrix [1]である．ここでは，これを符号語表 (Code Word Table)と呼び，次のように行列 W でモデル化する．

$$\begin{aligned} W &= [w_{ij}] \ (w_{ij} \in \{0,1\}, i = 1,2,\cdots,M, j = 1,2,\cdots,N) \\ &= [\boldsymbol{d}_1^{\mathrm{T}}, \boldsymbol{d}_2^{\mathrm{T}}, \cdots, \boldsymbol{d}_N^{\mathrm{T}}] \\ &= [\boldsymbol{x}_1, \boldsymbol{x}_2, \cdots, \boldsymbol{x}_M]^{\mathrm{T}} \end{aligned} \tag{3.1}$$

ここで，T は転置を示す．

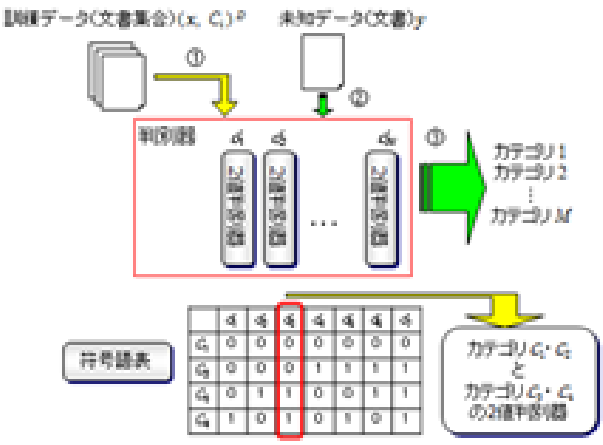


図 3.1：多値分類システムの構成

A System Evaluation for Construction Methods of Multiclass Problems using Binary Classifiers

† Shigeichi HIRASAWA, Gendo KUMOI, Masayuki GOTO : Waseda University
†† Manabu KOBAYASSHI : Shonan Institute of Technology
††† Hiroshige INAZUM : Aoyama Gakuin University

最後に，表 3.2 に今回用いた多値分類問題のベンチマークデータと実験データの仕様を示す．

表 3.2：ベンチマークデータと実験データの仕様

ベンチマークデータ（Letter Recognition Data Set, [6]）仕様	
カテゴリ(数)	A～Zの英文字(26)
文書の特徴ベクトル（次元）	整正数(16)
データ数	20,000件(A～Z, 各カテゴリ毎に734～813件)
実験データ（抽出し使用したデータ）仕様	
カテゴリ(数)	A～Mの英文字(10～13)
実験データ数	合計4,400件(A～M)
訓練データ	500件/カテゴリ, 合計4,000件
テストデータ	50件/カテゴリ, 合計400件

4 符号語表の作成と評価

4.1 Exhaustive 符号の生成

Exhaustive 符号は分類性能に寄与しない列ベクトルを除く最も網羅的（exhaustive）な符号語表で，M=5 のケースを表 4.1 に示す．この Exhaustive 号の符号長 $N_{\max}$ は次式を満たす[1].

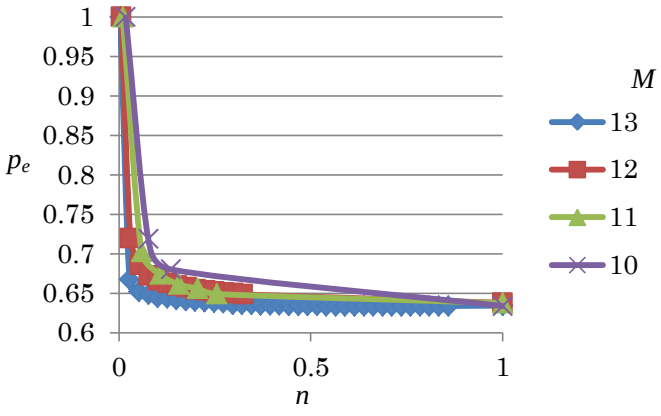
$$N_{\max} = 2^{M-1} - 1 \qquad (4.1)$$

表 4.1：Exhaustive 符号 (M=5, $N_{\max}$=15, D=7)

	d_1	d_2	d_3	d_4	d_5	d_6	d_7	d_8	d_9	d_{10}	d_{11}	d_{12}	d_{13}	d_{14}	d_{15}
C_1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
C_2	0	0	0	0	0	0	0	0	1	1	1	1	1	1	1
C_3	0	0	0	0	1	1	1	1	0	0	0	0	1	1	1
C_4	0	0	1	1	0	0	1	1	0	0	1	1	0	0	1
C_5	0	1	0	1	0	1	0	1	0	1	0	1	0	1	0

4.2 短縮 Exhaustive 符号と性能評価

ここではわずかな性能の劣化（分類誤り確率 P_e の増加）を許容して，コストを下げる（判別器数 N を減少させる）短縮 Exhaustive 符号を考える．与えられたシステムの規模 M で，コスト（判別器数，すなわち符号長）N を変え，その時の性能劣化の割合（分類誤り確率）P_e を求める．ここで，得られた結果は N に対する P_e の平均値である[1]．これを与えられた M に対し $N_{\max}$, $P_{e\max}$ で正規化し，$n = N/N_{\max}$, $p_e = P_e/P_{e\max}$ で示したものが図 4.1 である．



[1] すなわち，符号長 Nでは $N_{\max}$ 個の列ベクトルから N 個の列ベクトルの組み合わせを全て求めこれを平均している．また，符号語とカテゴリの対応をランダムとするために，M_0=8 個から M 個を選ぶ組み合わせを計算し，その平均をとっている．

図 4.1：投資コスト割合 n と性能の劣化割合 p_e のトレードオフ関係(パラメータ：システムの規模 M)

4.3 実験結果と考察

一般に多値分類器では，分類カテゴリ数 M が増加すると分類誤り確率 P_e は増加する．2 値判別器を用いた多値分類システムでも判別器数 N が一定の場合，同様である．ここでは，与えられた M で x, y 軸をそれぞれの最大値で正規化していることに注意されたい．すなわち，図 4.1 は与えられた M で，相対的な投資コストと性能の劣化割合のトレードオフ曲線を示している．その結果，両者は

(1) 下に凸なトレードオフ曲線となり，「フレキシブル」である．
(2) 一部，n→1 の近傍を除き M→大となるに従いトレードオフ曲線は原点方向に向かい，「エラスティック」である．

5 むすび

先の検討結果[3]に引き続き，「文書分類問題」とは異質の「手書き英文字認識問題」を適用した．その結果，トレードオフ関係は同様な性質を持つことが明らかとなった．

今後，符号理論の成果を導入して構成的な特定の符号を用いた場合について検討をしたい．人工データに対する多値分類システムの詳細な性能評価を行い，その振る舞いを明らかにすることも今後の課題である．

謝辞 本研究の一部は独立行政法人日本学術振興会学術研究助成基金助成金基盤研究(B) 26282090 の助成による．

参考文献

[1] T.G. Dietterich and G. Bakiri: “Solving multi-class learning problems via error-correcting out-put codes," Journal of Articial Intelligence Re-search, Vol.2, pp.263-286, 1995.

[2] S. Hirasawa, and H. Inazumi, “A system evaluation model by using information theory,” 30th Joint National Meeting, ORSA/TIMS, MB35.3, Philadelphia, PE. USA, Oct. 1990.

[3] 平澤茂一，雲居玄道，小林学，後藤正幸，稲積宏誠，「2 値判別器を用いた多値分類方式のシステム評価，」2016 年経営情報学会秋季全国研究発表大会，講演番号：B1-2，大阪，2016 年 9 月 15-16 日．

[4] J. Pearl, and A. Crolotte, “Storage space versus validity of answers in probabilistic question answering systems,” IEEE Trans. Inform. Theory, vol. IT-26, no. 6, pp.633-640, Nov. 1979.

[5] UCI machine learning repository, URL: http://www.trifields.jp/uci-machine-learning-repository-datasets-956