

# クラス所属確率を用いた事例ごとの分類器選択

高橋 和子  
敬愛大学 国際学部  
takak@u-keiai.ac.jp

## 1 はじめに

本稿では、サポートベクターマシン (SVM) による文書分類の精度を高めるために、素性の異なる複数の分類器を構築し、クラス所属確率を用いて各事例ごとに最適な分類器を選択するアンサンブル学習を提案する。

機械学習においては、複数の分類器を組み合わせ、それらの結果を統合することで個々の分類器よりも予測精度を上げるアンサンブル学習が有効な場合が多く (Sebastiani, 2002; 麻生他, 2003; 元田他, 2006), 代表的な方法としてバギングやブースティングがある。バギングは、リサンプリングにより元のデータセットと同じサイズのデータセットを複数個作成して、各データセットに同じアルゴリズムを適用してバリエーションの異なる複数の分類器を構築し、個々の分類器による予測結果に対して、カテゴリ型の場合には多数決により、連続値である回帰問題の場合には平均値や中央値により最終決定を行う方法である (Breiman, 1996; 麻生他, 2003; 元田他, 2006)。また、ブースティングは、逐次的に事例の重みを変化させながら分類器を構築していき、個々の分類器による予測結果に異なる重み付けをして最終決定を行う方法であり (麻生他, 2003; 元田他, 2006; Wu et al, 2008), 代表的なアルゴリズムに AdaBoost がある。

しかし、文書分類において分類精度の高さが評価されている SVM (Joachims, 1998) にバギングやブースティングを適用する場合には、以下の点が問題となる。まず、バギングについては、バイアス - バリエーションの理論 (Breiman, 1996) により、誤差をバイアス (予測に用いたモデルに由来する誤差)、バリエーション (学習に用いた訓練データのサンプリングの揺らぎに由来する誤差)、基本的に減らせない誤差の 3 つの部分に分解したとき、SVM のような高バイアスのモデルはもともとバリエーションの占める要素が少ないために、低バイアスのモデルほどにはリサンプリングによる効果が期待できない (Torii and Liu, 2007; 神島他, 2008)。また、ブースティングについては、SVM はブースティングに必要な重みを直接的に反映させることができないため、それに代わる対策が必要で (工藤他, 2002; 松田他, 2007; Li et al, 2008), うまく代替できない場合には有効ではない。

複数の分類器があるとき、全クラスについての分類精度 (分類器が正解した事例数を全事例で割った値) の平均が最も高い分類器は、どの事例においてもつねに他の分類器を上回って正解するわけではなく、この分類器が不正解の事例に対して、より低い分類精度の分類器が正解する場合も観察される。したがって、各事例ごとに正解した分類器を選択することができれば、全体として正解事例の数が増え分類精度の向上が期待できよう。

本稿ではこのような考えに基づき、SVM において多様な結果が得られるような分類器を複数個構築し、事例ごとに正解の可能性が高いと考えられる分類器を選択する方法を提案する。構築される分類器はできる限り多様なクラスを予測することが望ましいが、SVM においてはサポートベクターが分類に大きく関与するため、リサンプリングにより訓練事例を変化させるより学習のための素性を変化させる方が効果的であると考えられる。したがって、本稿では素性の選択をさまざまに変えることにより分類器の構築を行う。この点で提案手法はバギングと異なるが、同時並行的に複数の分類器を構築する点では類似する。提案手法においては、複数の分

類器の中から正解する可能性が高い分類器をいかにうまく選択できるかが重要である。今回は、多数決による方法、分類器の出力するスコア (分類スコア) を用いる方法、分類スコアにより推定したクラス所属確率を用いる方法の 3 つについて検討し、クラス所属確率を用いる方法の有効性を実験的に示す。

以下、次節で関連研究について述べた後、3 節で提案手法について説明する。4 節で実験と考察を行い、最後にまとめと今後の課題について述べる。

## 2 関連研究

ここでは、関連研究として、神島らの研究 (神島他, 2008) および Torii and Liu の研究 (Torii and Liu, 2007) について述べる。

神島らの研究は、バギングの考え方を利用し、より多様な事例が多数含まれると考えられる野生データ (整合性のある概念に基づいてラベル付けされた事例事例とそうではない事例が混在する) に注目したリサンプリングにより、多様な分類器の構築を提案する。Torii and Liu の研究は、SVM においてはバギングが有効ではないとして、bag-of-words に対して情報利得の計算により上位からランキングを行って、素性として利用する順位を変化させることで多様な分類器の構築を提案し、多数決により最終決定を行う。いずれも、有効なアンサンブル学習のために多様な分類器を構築しようとする点で本稿と共通するが、神島らの研究は事例選択である点が異なる。Torii and Liu の研究は、素性選択である点も本稿と同様であるが、最終決定を多数決により行う点で異なる。

## 3 提案手法

提案手法の概略は、次の通りである。

- STEP1 素性の選択方法を変化させて複数の分類器を構築する
- STEP2 各事例に対して個々の分類器ごとにクラスを予測する
- STEP3 各事例ごとに適切な分類器を選択し、その分類器が予測したクラスを最終決定とする

STEP3 における適切な分類器の選択方法については 3.1 節で述べる。

### 3.1 分類器の選択方法

ここでは、複数の分類器の中から適切な分類器を選択する方法として、今回は次の 3 つを検討する。

- 多数決による方法 (以下、「多数決法」と略す)
- 分類器の出力するスコア (分類スコア) を用いる方法 (以下、「分類スコア法」と略す)
- 分類スコアにより推定したクラス所属確率を用いる方法 (以下、「クラス所属確率法」と略す)

「多数決法」はバギングにおいてしばしば用いられる方法で、最も頻度の多いクラスを予測した分類器を選択する。「分類スコア法」は、各事例に対する結果に付随して出力される分類スコアの中で最も大きな値をもつクラスを出力した分類器を選択する。SVM において

は、分類スコアは分離平面からの距離である。今回、2 値分類である SVM を one-versus-rest 法 (kressel, 1999) により多値分類器として拡張したために、予測されたクラスがランク付けされて、各クラスごとの分類スコア間に次の関係：第 1 位に予測されたクラスの分類スコア > 第 2 位に予測されたクラスの分類スコア > ... があるが、今回は簡単のため、第 1 位に予測されたクラスの分類スコアのみを対象とする。提案手法である「クラス所属確率法」は、分類器が予測したクラスがどの程度確からしいかを示すクラス所属確率 (Platt, 1999; Zadrozny and Elkan, 2002; Niculescu-Mizil and Caruana, 2005; Langford and Zadrozny, 2005; 高橋他, 2008; Takahashi et al, 2008) を推定し、最も大きなクラス所属確率の推定値をもつクラスを出力した分類器を選択する。

### 3.2 クラス所属確率の推定方法

クラス所属確率の推定は、第 1 位に予測されたクラスに対して実験的に有効性が示された次の 2 つのいずれかにより行う (高橋他, 2008; Takahashi et al, 2008)。

- ロジスティック回帰式を用いるパラメトリックな方法 (特に第 1 位から第 3 位までに予測されたクラスの分類スコアを用いる場合が有効)
- 「正解率表」を作成して用いるノンパラメトリックな方法 (特に第 1 位および第 2 位に予測されたクラスの分類スコアを用いる場合が有効)

ロジスティック回帰式を用いる方法は、第 1 位から第 3 位に予測されたクラスの分類スコア ( $f_1, f_2, f_3$ ) を、次のロジスティック回帰式

$$P_{Log}(f_1, f_2, f_3) = \frac{1}{1 + \exp(\sum_{i=1}^3 A_i f_i + B)} \quad (1)$$

に代入して直接計算する。ただし、(1) 式におけるパラメタ (4 個) は最尤法により推定しておく必要があるため、訓練データをさらに訓練データと評価データに分けて学習した結果を用いる。簡単のため、分類スコアが 1 個の場合におけるパラメタの推定方法を以下に示す。与えられた事例の分類スコアを  $f^i$  とすると、正解 ( $Y^i = 1$ ) である確率は  $P_{Log}(f^i; A, B)$ 、不正解 ( $Y^i = 0$ ) である確率は  $1 - P_{Log}(f^i; A, B)$  であるため、 $Y^1, \dots, Y^n$  を得る同時確率を  $A, B$  の関数と考えれば、次の尤度関数が得られる (東大教養学部統計学教室 (編), 1992)。

$$L(A, B) = \prod_{Y^i=1} P_{Log}(f^i; A, B) \times \prod_{Y^i=0} [1 - P_{Log}(f^i; A, B)] \quad (2)$$

「正解率表」を用いる方法においても、あらかじめ訓練データを訓練データと評価データに分割して学習を行い、評価データの正誤状況を調査する。この方法では、第 1 位および第 2 位に予測されたクラスの分類スコアのそれぞれを区間 (例えば 0.1) で分け、各区間 (セル) ごとに正解率 (各セル内の正解事例数 / 各セル内の全事例数) を算出した正解率表を作成しておく。評価事例の分類スコアから正解率表内の該当セルを探し、そのセルの正解率を用いて間接的にクラス所属確率法の推定値とする<sup>1</sup>。

なお、クラス所属確率を事後確率と考えるためには、すべてのクラスに対してそれぞれのクラス所属確率を

<sup>1</sup> 「正解率表」を用いる方法は、分類スコアの区間設定が適切であればロジスティック回帰を用いる方法より良好な結果が得られたが、安定性の問題が存在した (高橋他, 2008; Takahashi et al, 2008)。

求めて和が 1 になるように正規化する必要があるが、本稿では注目するクラスに関してそのクラスに所属するかどうかに関心があるため、正規化までは行っていない<sup>2</sup>。

## 4 実験と考察

提案手法の有効性を調査するために、2005 年 SSM 調査 (社会階層と社会移動に関する全国調査) により収集された職業データを 390 個の国際標準職業分類 (ISCO) コード (Bureau of Statistics, 2001) に分類するタスク (「ISCO 職業コーディング」 (2005 年 SSM 調査研究会, 2006; 2005 年 SSM 調査研究会, 2007) ) に適用し、単独の分類器や他の分類器選択手法との比較を行った。

### 4.1 実験設定

#### データセット

用いたデータセットは、2005 年 SSM 調査データのうち有職者 (本人現職, 本人初職, 配偶者職) 計 16,089 サンプルである。訓練データと評価データの分割は、10 分割交差検定によった。すなわち、14,480 サンプルを訓練データ、1,609 サンプルを評価データとし、訓練データと評価データを変化させて 10 回の実験を行った。ロジスティック回帰式におけるパラメタ推定および正解率表作成のためには、各訓練データごとに 10 分割交差検定を行ってこれを訓練データと評価データに分割し、この評価データにおける正解 / 不正解の状況 (2 値) を用いた。また、正解率表作成のための区間幅は 0.1 とした。

2005 年 SSM 調査における職業データは、すでに調査終了後に行われた職業コーディングにより、SSM コードと ISCO コードの 2 種類の職業コード (各 1 個) が付与されている。本稿においては、この ISCO コードを正解として扱った。

#### 素性の選択

今回用いたデータセットは調査データで、職業データ以外にも多様なデータを利用することができたため、情報利得を用いずに以下のような素性の選択を行った。

まず、2005 年 SSM 調査データにおける職業データは、「仕事の内容」 (自由回答)、「従業先事業の種類」 (自由回答)、「従業上の地位と役職」 (13 個の選択回答)、「従業先事業の規模」 (13 個の選択回答) の 4 種類で構成されるが、高橋ら (高橋他, 2005) にしたがって、今回も「従業先事業の規模」を除く 3 種類を基本素性とした。次に、職業データ以外の種類の異なる素性を組み合わせながら基本素性に追加した。今回追加した素性は、「学歴」 (6 個の選択回答)、「性別」 (2 個の選択回答)、「職業コーディングによりすでに付与された「SSM コード」 (約 200 個) の 3 つである<sup>3</sup>。ここで、SSM コードは、調査データである学歴や性別とは性質が異なる素性であるといえる。以上により、今回は次のような 8 種類の分類器を構築した。

- [ 分類器 A ] 基本素性のみ
- [ 分類器 B ] 基本素性, 学歴
- [ 分類器 C ] 基本素性, 正解 SSM コード
- [ 分類器 D ] 基本素性, 学歴, 正解 SSM コード

<sup>2</sup> 高橋ら (高橋他, 2008) の実験においては、正規化した値としないう値の差がほとんどなかったことが報告されている。

<sup>3</sup> 学歴は、ISCO コードは分類に際して「スキル」という概念が用いられるが、直接これを測るデータが収集されていないため、最も近いものとして代用した。SSM コードは、高橋ら (高橋他, 2005) により、SSM 職業コーディングにおける有効な機械学習手法として、すでに存在するルールベース手法による自動分類システムの出力結果を機械学習の素性として用いる方法が提案されており、これを参考にした。今回用いられるクラスは SSM コードとは異なる ISCO コードであり、さらに追加されるコードも対象とするクラスとは別の体系のコードであるが、職業コードであるという共通点を考慮した。

表 2: 単独の分類器 (E ~ H) における分類精度 (平均)

| 分類器  | E     | F     | G     | H            |
|------|-------|-------|-------|--------------|
| 分類精度 | 0.692 | 0.696 | 0.735 | <b>0.739</b> |

[ 分類器 E ] 基本素性, 性別

[ 分類器 F ] 基本素性, 学歴, 性別

[ 分類器 G ] 基本素性, 性別, 正解 SSM コード

[ 分類器 H ] 基本素性, 学歴, 性別, 正解 SSM コード

#### 分類器と評価尺度

SVM は本来は 2 値分類器であるため, いずれも one-versus-rest 法を用いて多値分類器に拡張し (kressel, 1999), カーネル関数は高橋ら (2005) にしたがって線型カーネルを用いた. また評価尺度は分類精度を用いた.

#### 4.2 予備実験

提案手法についての実験を行う前に, 4 種類の分類器 (A から D) により, 分類精度および正解の出現状況について調査した (表 1 参照). 表 1 において太字は分類器の中で最も高い値を示す. 表 1 より, 単独の分類器における分類精度は SSM コードを素性としてもつ分類器 C と D の値が高く, 分類器 D > 分類器 C > 分類器 B > 分類器 A の順であった. 分類器の正解状況は, 1 個の分類器だけが正解の場合 1.9%, 以下 2 個正解の場合 12.7%, 3 個正解の場合 (1 個だけ不正解) 2.5%, すべて正解の場合 62.6% であった. このうち特に 1 個の分類器だけ正解した場合に注目すると, 正解事例数は分類精度の高さと関係なくどの分類器においてもほぼ等しかった. すなわち, 事例ごとにみると必ずしも分類精度の高い分類器だけが正解しているわけではなく, より分類精度の低い他の分類器が正解している場合もあることが確認できた. したがって, 各事例ごとに正解した分類器があればその予測クラスを最終決定とすることが効果的であると考えられ, 実際, 最適な分類器を選択できた場合の分類精度は 79.7% となり, 単独の分類器の中で最も高い分類器 D の値 73.7% を 6.0% 上回った (高橋, 2008b).

分類器 E から H の分類精度については表 2 に示す. 8 種類すべての分類器の中で分類精度が最も高かったのは分類器 H の 73.9% であったため, 以下の実験ではこの値をベースラインとした. また, 分類器が 8 種類の場合につねに最適な分類器を選択できたときの分類精度は, 単独の分類器の中で最も高い値を 6.6% 上回る 80.5% であったため, この値を目標値とした.

#### 4.3 実験結果と考察

分類器の選択方法別の分類精度は表 3 に示すような結果であった. 表 3 において, 多数決法で同数の分類器が正解した場合 (例えば正解と不正解の分類器数が 4 個ずつの場合) は, 分類器 H が含まれるグループを選択した. 太字はすべての選択方法の中で最も高い値を示す.

表 3 より, 分類器の選択方法としてクラス所属確率法の分類精度は 74.6% (正解率表) および 74.1% (ロジスティック回帰) で, 単独で最も高かった分類器 H をそれぞれ 0.7%, 0.2% 上回ったが, 他の方法は 0.1% (多数決法), 0.8% (分類スコア法) 下回った. これより, 複数の分類器から事例ごとに適切な分類器を選択する方法はつねに有効であるわけではなく, クラス所属確率を用いる場合のみ有効なことがわかった. しかし, 最もよかったクラス所属確率法 (正解率表) でも, 分割したデータセット 10 個のすべてで分類器 H を上回ったわけではなく, 2 個は下回っており (0.9%, 0.3%), また, 目標値 80.5% との差も大きい (5.9%), 今後, より一

表 4: 分類器の多様性による分類精度 (平均) の違い (クラス所属確率 (正解率表) の場合)

| 用いた分類器    | A,B,C,E | A,B,E,F |
|-----------|---------|---------|
| 分類精度      | 0.739   | 0.696   |
| 単独の分類器    | 0.734   | 0.696   |
| 単独の分類器との差 | 0.005   | 0.000   |

般的なタスクに対する実験を行って, 有効性の確認を行う必要がある<sup>4</sup>.

次に, クラス所属確率の再現率 (発見できた正解事例数 / 全正解事例数) を正解した分類器数が 1 個の場合, 2 個の場合, ..., 8 個すべての場合に分けて調査すると, 順に 2.9% (0.7%), 10.7% (14.2%), 28.3% (26.9%), 74.6% (70.5%), 92.9% (92.3%), 96.8% (96.8%), 100.0% (100.0%), 100.0% (100.0%) であった. ( ) のない値は正解率表, ( ) 内の値はロジスティック回帰を用いた場合の再現率を示す. 正解した分類器が 3 個以下の場合の再現率はいずれの選択方法も非常に低かったが, その中では分類スコア法 (22.8%) が最も高く, クラス所属確率法は 13.4% (14.2%) という結果であった. クラス所属確率の精度の向上がより適切な分類器の選択につながると考えると, クラス所属確率の推定精度をさらに高める研究が必要である.

最後に, クラス所属確率 (正解率表) において, 分類器の多様性の違いによる分類精度の変化を調査するために, 分類器の多様性が高いグループ (分類器 A, B, C, E) と低いグループ (分類器 A, B, E, F) の比較を行った (表 4 参照). ここで, 分類器 A, B, C, E のグループは他の素性とは性質が異なる素性 (SSM コード) をもつ分類器 C を含むため, これを含まないグループより多様性が大きいとした. 表 4 において単独の分類器とは, 単独の分類器の中で分類精度が最も高い分類器の分類精度を示す. 表 4 より, 分類器の多様性が大きい方が有効性が高かった. 表 4 は分類器が 4 種類の場合であるが, さらに多様性が増すと考えられる 8 種類の場合には, 表 3 に示したように有効性はさらに高かった. 以上より, 提案手法は分類器が類似している場合には有効ではなく, 多様性が増すほど効果的であると考えられるため, できる限り性質の異なる素性を用いて多様な分類器を構築する必要がある.

#### 5 おわりに

本稿では, SVM における分類精度を高めるためのアンサンブル学習として, 素性を変化させて多様な分類器を構築し, 各事例ごとにクラス所属確率が最も高い分類器を選択する方法を提案し, その有効性を実験的に示した. 今後の課題としては, 提案手法の有効性を確認するために, 適用するデータセットをより一般的なものに変えて実験を行うことおよび, 分類精度だけでなく AUC (Area Under the ROC Curve) による評価も行うことが必要である.

謝辞 2005 年 SSM 調査データの利用に関して, 2005SSM 研究会の許可を得た.

#### References

- 2005 年社会階層と社会移動調査研究会. 2007. 2005 年 SSM 日本調査 コード・ブック.
- 2005 年社会階層と社会移動調査研究会. 2006. 2005 年 SSM 調査 日本・韓国・台湾調査票.

<sup>4</sup> 今回の実験に用いたタスクは, 文書分類タスクとしては, 事例に含まれる素性の数が平均で 15 個程度で非常に少ない上にクラスの数 が 390 個と非常に多い点で, やや特殊であったと考えられる.

表 1: 単独の分類器 (A ~ D) における分類精度 (平均) と正解の出現状況

| 分類器         | A            | B            | C            | D            | 分類器を適切に<br>選択できた場合 |
|-------------|--------------|--------------|--------------|--------------|--------------------|
| 分類精度        | 0.689        | 0.693        | 0.734        | <b>0.737</b> | <b>0.797</b>       |
| 1 個の分類器だけ正解 | 0.004        | <b>0.005</b> | 0.004        | <b>0.005</b> | 0.019              |
| 2 個の分類器が正解  | 0.045        | 0.044        | <b>0.084</b> | 0.083        | 0.127              |
| 3 個の分類器が正解  | 0.014        | 0.018        | 0.020        | <b>0.023</b> | 0.023              |
| すべての分類器が正解  | <b>0.626</b> | <b>0.626</b> | <b>0.626</b> | <b>0.626</b> | 0.626              |

表 3: 分類器の選択方法における分類精度 (平均) (8 種類の分類器を用いた場合)

| 分類器の選択方法  | 多数決法   | 分類スコア法 | クラス所属確率法<br>(ロジスティック回帰) | クラス所属確率法<br>(正解率表) |
|-----------|--------|--------|-------------------------|--------------------|
| 分類精度      | 0.738  | 0.731  | 0.741                   | <b>0.746</b>       |
| 分類器 H との差 | -0.001 | -0.008 | 0.002                   | <b>0.007</b>       |

麻生英樹, 津田宏治, 村田昇. 2003. パターン認識と学習の統計学 新しい概念と手法. 岩波書店.

Bureau of Statistics; International Labour Office. 2001. Coding Occupation and Industry. Bureau of Statistics; International Labour Office.

L. Brieman. 1996. Bagging predictors. In *Machine Learning* 24(2), pp. 123–140.

Y-S. Dong and K-S. Han. 2004. A comparison of several ensemble methods for text categorization. In *Proceedings of IEEE 2004 International Conference on Services Computing (SCC 2004)*, pp. 419–422.

T. Joachims. 1998. Text Categorization with Support Vector Machines: Learning with Many Relevant Features. In *Proceedings of the European Conference on Machine Learning*, pp. 137–142.

神高敏弘, 濱崎雅弘, 赤穂昭太郎. 2008. 飼い慣らし - 飼育・野生混在データからの学習. 第 22 回人工知能学会, pp. 1–4.

U. Kressel. 1999. Pairwise classification and support vector machines. In *Advances in Kernel Methods Support Vector Learning*, pp. 255–268. MIT Press.

工藤拓, 松本裕治. 2002. Support Vector Machine を用いた Chunk 同定. 自然言語処理 Vol.19 No.5, pp.3–22.

J. Langford and B. Zadrozny. 2005. Estimating class Membership Probabilities using Classifier Learners. In *Proceedings of the Eighth International Conference on Knowledge Discovery and Data Mining (AI-STAT'05)*, pp.694–699.

X. Li, L. Wang, and E. Sung. 2008. AdaBoost with SVM-based component classifiers. In *Engineering Applications of Artificial Intelligence* 21(5) pp.785–795.

松田博義, 滝口哲也, 有木康雄. 2007. 弱識別器に SVM を用いた AdaBoost の検討. 信学技報 Vol.107 No.405, pp.109–114.

元田浩, 津本周作, 山口高平, 沼尾正行. 2006. データマイニングの基礎. オーム社.

A. Niculescu-Mizil and R. Caruana. 2005. Predicting Good Probabilities With Supervised Learning. In *Proceedings of the 22nd International Conference on Machine Learning (ICML'05)*, pp. 625–6323.

J. C. Platt. 1999. Probabilistic Outputs for Support vector Machines and Comparisons to Regularized

Likelihood Methods. In *Advances in Large Margin Classifiers*, pp. 1–11. MIT Press.

F. Sebastiani. 2008. Machine Learning Automated Text Categorization. In *ACM Computing Surveys* 34(1), pp.1–47.

高橋和子, 高村大也, 奥村学. 2005. 機械学習とルールベース手法の組み合わせによる自動職業コーディング. 自然言語処理 Vol.12 No.2, pp. 3–24.

高橋和子. 2008a. 機械学習による ISCO 自動コーディング. 2005 年 SSM 調査シリーズ 12 社会調査における測定と分析をめぐる諸問題, pp.47–68.

高橋和子, 高村大也, 奥村学. 2008. 複数の分類スコアを用いたクラス所属確率の推定. 自然言語処理 Vol.15 No.2, pp. 3–38.

高橋和子. 2008b. 素性選択によるアンサンブル学習に関する一考察. 情報処理学会研究報告 2008-NL-188, pp. 23–28.

K. Takahashi, H. Takamura, and M. Okumura. 2008. Direct estimation of class membership probabilities for multiclass classification using multiple scores. In *Knowl Inf Syst* (doi:10.1007/s10115-008-0165-z). Springer London.

D. Tao, X. Tang, X. Li, and X. Wu. 2006. Asymmetric Bagging and Random Subspace for Support Vector Machines-Based Relevance Feedback in Image Retrieval. In *The IEEE Transactions on Pattern analysis and machine intelligence (TPAMI)* 28(7), pp.1088–1099.

M. Torii and H. Liu. 2007. Classifier ensemble for biomedical document retrieval. In *Proceedings of the Second International Symposium on Languages in Biology and Medicine (LBM 2007)*.

東京大学教養学部統計学教室 (編). 1992. 基礎統計学 自然科学の統計学. 東京大学出版会.

X. Wu, V. Kumar, J. R. Quinlan, J. Ghosh, Q. Yang, H. Motoda, G. J. McLachlan, A. Ng and B. Liu, P. S. Yu, Z-H. Zhou, M. Steinbach, D. J. Hand, and D. Steinberg. 2008. Top 10 algorithms in data mining. In *Knowl Inf Syst* 14, pp.1–37. Springer London.

B. Zadrozny and C. Elkan. 2002. Transformation Classifier Scores into Accurate Multiclass Probability Estimates. In *Proceedings of the Eighth International Conference on Knowledge Discovery and Data Mining (KDD'02)*, pp. 694–699.