

## 計測ノイズ除去に向けた PU 分類学習原理の検討

## Feasibility Study on PUC for Measurement Noise Reduction

吉田 剛  
Takeshi Yoshida鷲尾 隆  
Takashi Washio大城 敬人  
Takahito Ooshiro谷口 正輝  
Masateru Taniguchi大阪大学 産業科学研究所  
The Institute of Scientific and Industrial Research Osaka University

The needs to employ machine learning is increasing for accurate estimation and noise reduction in recent advanced measurement where its output data is enormous, complex and noisy. Particularly, the recently emerging Positive and Unlabeled Classification (PUC) can be used to classify target objects and contaminants in the measurement. However, the existing standard machine learning is based on Bayesian estimation which assumes invariance of the target population distributions, whereas they are very different depending on the objects in the measurement. In this study, we investigated the PUC to overcome this issue. We applied the method to an actual measurement problem and confirmed its significant noise reduction.

## 1. はじめに

これまで半教師あり学習に関する研究が多数行われている [Chapelle 10]. この中でも特に、多数のラベルなし事例が簡単に取得できる一方で非常に限られた個数の正例しか得られないという我々がよく直面する問題について、正例とラベルなし事例から学習し正負例を分類する問題 (PU Classification, PUC) [Elkan 08, Ward 09] が、近年とりわけ盛んに研究されてきている。

一方、IoT センシングや自然科学における観測など多くの計測技術において、対象の状態  $Y$  を観測結果  $X$  から正確に推定する技術が非常に重要になってきている。このような計測の場面においては、各観測対象の事前確率密度関数  $p(Y)$  は不明であることがほとんどである。しかし、例えば、温度計は空気や熱水などの計測対象の分布がどうであろうとも、その対象の正しい温度を計測する必要がある。最先端の分子計測装置では、計測のたび様々に混合比率が変化する目的分子と夾雑物を分類し目的分子の計測を行う必要がある [Tsutsui 10]. これを達成するために、ほとんどの場合においてセンサーのような計測装置は、 $p(Y)$  の変化に対して  $Y$  のロバストな推定ができるように、不変な  $p(X|Y)$  を安定的に実現するよう設計されている。

これに対して、標準的な機械学習におけるベイズ推定  $p(Y|X) \propto p(X|Y)p(Y) = p(X, Y)$  は、学習用事例集合とテスト用事例集合の同時確率分布  $p(X, Y)$  が同一であることを仮定し、通常はこれら学習用/テスト用事例集合のクラス事前確率  $p(Y)$  も変わらないと仮定している。

これまで提案された PUC 手法はすべてこの設定で開発されている [Elkan 08, Ward 09].  $p(X, Y)$  の全体または  $p(Y)$  と  $p(X|Y)$  のうち一方が変化する状況をコンセプトドリフト (concept drift) と呼ぶが、通常の教師あり学習において徐々にコンセプトドリフトが起こる場合に対して多くの適応的学習法が提案されている [Gama 14]. より大きなコンセプトドリフトについては、テスト用事例集合のための新しい推定器を学習用事例集合から得る多くの手法が、トランスダクティブ転移学習 [Pan 10] および共変量シフト適応 [Pfeffermann 98] において研究されている。しかし、これらの手法はすべて、新しく得ら

れるテスト用事例集合の分布に関する情報を必要とし、状態  $Y$  が不明な単一の観測値  $X$  が新たに得られるというような計測タスクには適用できない。

このような特徴をもつ計測タスクでは、上記のようなベイズ推定における困難を回避するため、正負ラベル付き学習用事例集合から学習した  $p(X|Y)$  の不変性に基づく  $X$  による  $Y$  の最尤推定が広く採用されている [Washio 17]. しかし我々が知る限り、正例集合とラベルなし事例集合から  $p(X|Y)$  を正確に推定する原理は知られていない。このような原理に基づく最尤 PUC が利用可能になれば、後述するノイズ除去技術を含んだ、計測を指向した様々な新しい情報処理技術の開発が実現可能となる。以上の背景から、本研究では以下の成果をもたらす検討を行った。

- 我々は、計測タスクを指向した最尤推定 PUC の新しい原理を検討した。これは、2 標本問題においてクラス事前確率を必要としない初めての PUC である。
- 検討 PUC 手法は、ラベルなし事例が多数得られる一方でラベル付き正例が少数しか得られないようなインバランスな学習用事例集合について正確に正負例分類ができ、現実の多くの問題に適用できる。
- 検討手法は、最先端分子計測技術におけるノイズ除去という、現実の応用問題において非常に有効に機能することが確認された。

## 2. 問題設定および関連先行研究

$X \in \mathcal{X} \subseteq \mathbb{R}^d$  ( $d \in \mathbb{N}$ ),  $Y \in \{P, N\}$  をそれぞれ特徴ベクトルおよび正負例ラベルとする。学習用に与えられた正例集合とラベルなし事例集合をそれぞれ  $D_{LP}$ ,  $D_{LU}$  とし、計測のたびに取得される事例集合  $D_{TU}$  をテスト用ラベルなし事例集合とする。 $D_{LP}$  の事例は正例周辺分布  $p_{LP}(X|Y=P)$  から i.i.d. 標本抽出され、 $D_{LU}$ ,  $D_{TU}$  の事例はそれぞれ周辺分布  $p_{LU}(X)$ ,  $p_{TU}(X)$  から i.i.d. 標本抽出される。ラベル付き負例は存在しない。

ここで、 $p(X|Y)$  に関して以下のように仮定する。

**Assumption 1.**  $p_{LP}(X|Y=P)$ ,  $p_{LU}(X)$ ,  $p_{TU}(X)$  は同一の不変な分布  $p(X|Y)$  から構成される。 □

これは特殊な仮定ではなく、過去のすべての PUC 手法においても全事例集合にわたって共通の  $p(X|Y)$  を仮定している [Elkan 08, Ward 09]. この仮定により,  $p_{LP}(X|Y = P) = p(X|Y = P)$  が成り立つので,  $Y = P, N$  についての共通の  $p(X|Y)$  と正負例のクラス事前確率  $\pi_L = p_{LU}(Y = P)$ ,  $\pi_T = p_{TU}(Y = P)$  を用いて,  $p_{LU}(X)$  と  $p_{TU}(X)$  を以下のように表せる.

$$p_{LU}(X) = \pi_L p(X|Y = P) + (1 - \pi_L) p(X|Y = N), \quad (1)$$

$$p_{TU}(X) = \pi_T p(X|Y = P) + (1 - \pi_T) p(X|Y = N). \quad (2)$$

さらに  $p(X|Y)$  を正で有界かつ  $\mathcal{X}$  において 2 回微分可能であるとし,  $\pi_L, \pi_T \in (0, 1)$  は不明だが独立に与えられるとする.

我々の問題は,  $D_{LP}$  と  $D_{LU}$  のみから,  $D_{TU}$  から抽出された未知の事例についての正確なノンパラメトリック分類器を学習することである.  $D_{LP}$  と  $D_{LU}$  は独立に与えられ, かつ相互に異なる未知のクラス事前確率に従うので, これは PUC における 2 標本問題である [Ward 09].

この問題に取り組む上で, 計測タスクに特有の問題が 2 つある. 1 つは,  $D_{TU}$  に関する情報はその事例を分類する以前には何一つ得られないということである. 前述のようにこれまでに提案された PUC では  $p(X, Y)$  の不変性を仮定するベイズ推定を用いているために, このような条件下で適用できない. 伝統的な正負例の 2 値分類問題において, コンセプトドリフト [Gama 14] や学習/テスト事例集合間の分布が大きく違う場合 [Pan 10, Pfeiffermann 98] に, 分類器を適応させる手法がこれまでに多数研究されてきている. PUC においては, Li らがデータストリームに対するマイクロクラスターのラベルが徐々に変化する場合における PUC を提案した [Li 09]. ラベル付き負例も学習に用いる通常の半教師あり学習において, Saerens らはクラス事前確率が異なるラベル付き学習事例集合とラベルなしテスト用事例集合に対して共変量シフト適応を提案した [Saerens 02]. より最近の研究としては, トランスダクティブ転移学習を用いてクラス事前確率を推定するものもある [Plessis 14b]. しかしこれらはすべて, 未知のテスト用事例を分類する以前に  $D_{TU}$  を必要とし, 我々の問題設定において適用することができない.

第 2 の問題点は,  $\pi_L$  と  $\pi_T$  は未知であるということである. したがって, 考案すべき PUC は, そもそもクラス事前確率を用いないものであるか, またはクラス事前確率の推定も同時に行うようなものでなければならない. Elkan らの PUC [Elkan 08] はベイズ推定に基づいており我々の問題設定には使えないが, 彼らの手法はクラス事前確率を用いないため前者の PUC である. この方法では,  $D_{LU}$  の分布に  $D_{LP}$  の分布を部分的に一致させることによりクラス事前確率を暗に推定している. そのために 1 標本問題を仮定しており,  $p_{LU}(X)$  から抽出される正例はある一定確率でラベル付けされ  $D_{LP}$  をなし, 残りの正負例を  $D_{LU}$  としている [Plessis 15]. しかし  $D_{LP}$  と  $D_{LU}$  は通常, 個別に生成されるので, この 1 標本問題は計測タスクにおいて常に適用できるとは限らない.

我々の問題には不適切であるが, 後者のクラス事前確率の推定を同時に行う PUC が,  $D_{LP}$  と  $D_{LU}$  が独立に提供される 2 標本問題において提案されている [Ward 09, Plessis 14a, Niu 16]. さらに,  $D_{LU}$  のクラス事前確率の推定手法も多数, 提案されている [Blanchard 10, Plessis 14a, Scott 15, Ramaswamy 16]. これらの手法はすべて  $p(X|Y)$  の非還元性, すなわち  $p(X|Y = N)$  と他の  $X$  の確率密度関数のいかなる線形結合によっても  $p(X|Y = P)$  は実現できない, という仮定に基づいている. し

かし計測タスクにおいて, この  $p(X|Y)$  の非還元性が保証されるとは限らない. さらにこの手法では, 未知の事例を分類する以前に  $\pi_T$  の推定のため  $D_{TU}$  が必要であり, 我々の問題設定には適用できない.

このような問題を解決すべく次節において,  $D_{TU}$  を事前に必要とせず,  $\pi_L, \pi_T$  の推定も必要としない, 最尤推定原理に基づく新しい PUC を検討する.

### 3. 検討手法

#### 3.1 原理

$\pi_L$  と  $\pi_T$  の違いにロバストな分類器を構成するため, クラス事前確率の影響を受けない最尤推定原理を用いる分類基準を採用する. あるラベルなしテスト事例  $x \in D_{TU}$  の最尤な  $Y$  は, 仮定 1 より以下で与えられる.

$$y = \begin{cases} P & \text{if } p(x|Y = P) \geq p(x|Y = N), \\ N & \text{if } p(x|Y = P) < p(x|Y = N). \end{cases} \quad (3)$$

この基準に関連して, 以下の補題が成立する.

**Lemma 1.** 任意の  $\pi \in (0, 1)$  を有する

$$p_\pi(X) = \pi p(X|Y = P) + (1 - \pi) p(X|Y = N)$$

について, 以下の 2 つの不等式は等価である.

$$p(X|Y = P) \geq p_\pi(X) \Leftrightarrow p(X|Y = P) \geq p(X|Y = N). \square$$

**Proof.**  $p(X|Y = P) \geq p_\pi(X)$  に上記の  $p_\pi(X)$  の定義を代入し, 以下を導く.

$$\begin{aligned} p(X|Y = P) &\geq \pi p(X|Y = P) + (1 - \pi) p(X|Y = N) \\ &\Leftrightarrow (1 - \pi) p(X|Y = P) \geq (1 - \pi) p(X|Y = N) \\ &\Leftrightarrow p(X|Y = P) \geq p(X|Y = N) \text{ s.t. } \pi \in (0, 1). \quad \blacksquare \end{aligned}$$

この補題と仮定 1 から, 直ちに我々の検討手法の中核原理である以下の定理を得る.

**Theorem 1.** 任意の  $\pi_L \in (0, 1)$  の下で与えられる以下の式は,  $\pi_L$  と独立に与えられる任意の  $\pi_T \in (0, 1)$  をもつ  $p_{TU}(X)$  に従う事例  $x \in D_{TU}$  の最尤分類基準である.

$$y = \begin{cases} P & \text{if } p(x|Y = P) \geq p_{LU}(x), \\ N & \text{if } p(x|Y = P) < p_{LU}(x). \end{cases} \quad (4)$$

$\square$

**Proof.** Eq.(1), (2) と補題 1 より, 以下が成立する.

$$\begin{aligned} p(X|Y = P) \geq p_{LU}(X) &\Leftrightarrow p(X|Y = P) \geq p_{TU}(X) \\ &\Leftrightarrow p(X|Y = P) \geq p(X|Y = N). \end{aligned}$$

仮定 1 より最後の不等式は  $\pi_L, \pi_T$  に依存しないので, 任意の  $\pi_L, \pi_T \in (0, 1)$  について, これら 3 つの不等式は等価でありかつ不変な条件を表す. したがって, Eq.(3) より, Eq.(4) は  $p_{TU}(X)$  に従う  $x \in D_{TU}$  の最尤分類基準である.  $\blacksquare$

$D_{LP}, D_{LU}$  からそれぞれ  $\hat{p}(x|Y = P) := \hat{p}_{LP}(x|Y = P)$  と  $\hat{p}_{LU}(x)$  をノンパラメトリック推定し, Eq.(4) を用いて  $x \in D_{TU}$  のラベル  $y$  を最尤推定する PUC を構成できる.

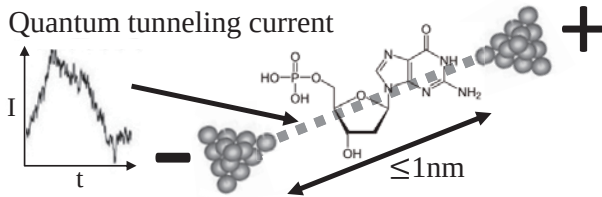


図 1: ナノギャップ電極の模式図

| 手法     | $\pi_L \simeq \pi_T$ | $\pi_L < \pi_T$ | $\pi_L \ll \pi_T$ |
|--------|----------------------|-----------------|-------------------|
| NB-E&N | 0.572                | 0.582           | 0.592             |
| KD-E&N | 0.318                | 0.322           | 0.305             |
| 検討 PUC | 0.715                | 0.671           | 0.654             |

表 1: 各手法におけるノイズ除去効果の比較

### 3.2 パラメータ最適化と性能指標

$\hat{p}(x|Y=P)$ ,  $\hat{p}_{LU}(x)$  のノンパラメトリック推定にはカーネル幅  $h$  のガウスカーネル密度を用い、この  $h$  の最適化は以下のように行う。  $D_{LP}$  と  $D_{LU}$  からランダム抽出したうち 90% をそれぞれ学習用  $D_{LP}^L$ ,  $D_{LU}^L$  とし、残りの 10% をテスト用  $D_{LP}^T$ ,  $D_{LU}^T$  として、10CV 法を用いる。このように最適化した  $h$  を用いて、すべての  $D_{LP}$ ,  $D_{LU}$  により  $\hat{p}(x|Y=P)$ ,  $\hat{p}_{LU}(x)$  の推定をする。

ところで我々の問題設定ではラベル付き負例が得られないため、上述の 10CV 法においてラベル付きの負例を用いない分類性能指標が必要となる。このような性能指標としては、分類器のパラメータ選択ではなく分類器のクラスの評価法として、AUC と同様な性能指標が多数提案されているが [Hajizadeh 14, Menon 15, Jain 17], パラメータ選択のための指標はほんの数例しか提案されていない。Lee らは 1 標本問題において、precision と recall の幾何平均と同等な指標を提案した [Lee 03]。Calvo らは、2 標本問題において、真の  $\pi_L$  がわかっている場合には通常の F-measure に一致するような、疑似 F-measure を提案した [Calvo 10]。2 標本問題に適用できるので、我々は後者の指標を採用した。疑似 F-measure は以下で定義される。

$$\tilde{F} = \frac{2\pi|\hat{D}_{LP}^T|/|D_{LP}^T|}{|\hat{D}_{LU}^T|/|D_{LU}^T| + \pi} \quad (5)$$

ここで、 $D_{LU}^T$  は  $D_{LU}$  中の真の正例集合であり、 $\hat{D}_{LP}^T$  と  $\hat{D}_{LU}^T$  はそれぞれ  $D_{LP}^T$ ,  $D_{LU}^T$  中で正例と分類された事例集合である。この疑似 F-measure は、 $\hat{D}_{LU}^T = D_{LU}^T$  のとき、つまりラベルなし事例中の正例を正しく分類できた場合に最大となることが示せる。この指標は、 $h$  の最適化だけではなく、後に述べるノイズ除去効果の性能指標としても用いる。

## 4. 実データによるノイズ除去

我々は、現実の一分子計測装置 [Tsutsui 10] におけるノイズ除去のため、検討 PUC 手法を適用し性能評価を行った。図 1 はこの計測装置の模式図である。2 つのナノ電極ギャップの間を物体が通り抜けると、量子トンネル効果により通過物体に応じた電流パルスが計測される。我々の目的は、溶液中の目的分子と夾雑物などのノイズとを、それらのパルス概形に基づいて分類することである。ここでパルス概形は、パルス開始時点から終了時点までを 10 分割した各計測電流値の平均値を成分とする 10 次元特徴ベクトルにより表現する。計測初期の短時間におけるバックグラウンド値の計測により、ノイズのみによるパルス概形の集合を、ラベル付き正例集合として得ることが

きる。引き続き一定時間による計測により、ノイズと目的分子が混在したパルス概形集合を、ラベルなし事例集合として得られる。この PU 分類問題を解くことによって未知のパルス概形を、ノイズまたは目的分子に分類することができる。

性能比較対象として、Elkan らの PUC [Elkan 08] を用い、 $\hat{p}_{LP}(X|Y=P)$  と  $\hat{p}_{LU}(X)$  の推定にはガウスアンナイブベイズ (NB-E&N) と、ガウスカーネル密度を用いたベイズ推定 (KD-E&N) の 2 通りを採用した。

真の正負例ラベルは得られないので、3.2 節で導入した疑似 F-measure をその性能指標として用いる。すなわち、テスト用の正例集合  $D_{TP}$ , ラベルなし事例集合  $D_{TU}$  のうちで正例に分類された事例集合をそれぞれ  $\hat{D}_{TP}$ ,  $\hat{D}_{TU}$  とするとき、 $\tilde{F} = 2\pi|\hat{D}_{TP}|/|D_{TP}|/(|\hat{D}_{TU}|/|D_{TU}| + \pi)$  とする。

各事例集合について、 $|D_{LP}| = 20$ ,  $|D_{LU}| = 800$  を学習用に、 $|D_{TP}| = 20$ ,  $|D_{TU}| = 100$  をテスト用に取得した。表 1 に、各 PUC 手法の疑似 F-measure の値を示す。  $D_{TU}$  として、計測の初期段階 ( $\pi_L \simeq \pi_T$ )、しばらく時間が経って夾雑物が多くなった頃 ( $\pi_L < \pi_T$ )、非常に夾雑物が増えた頃 ( $\pi_L \ll \pi_T$ ) の 3 通りについて調べた。疑似 F-measure は通常のものとは異なり  $[0, 1]$  に規格化されないが、この値が大きいほど分類性能が高いことを示す。結果、検討手法は、学習用正例数が少ないにも関わらず、 $\pi_T$  の値によらず既存手法よりはるかに高性能を示すことがわかった。

## 5. 結論

本研究において、対象の母集団分布に依存しない、計測系を指向した新しい PU 分類手法を検討した。検討手法は既存手法に比べ、ラベル付き正例とラベルなし事例の数がインバランス、かつクラス事前確率が大きく異なる場合において、特に高精度・高ロバスト性を示すことがわかった。このような、ラベル付き正例が少数しか得られない一方で、様々なクラス事前確率をもつ多数のラベル付き事例集合がたやすく得られるという、現実起こりやすいケースに対して、検討手法は非常に有効であることがわかった。

## 6. 謝辞

本研究は、JST CREST JPMJCR1666 「計測技術と高度情報処理の融合によるインテリジェント計測・解析手法の開発と応用 (研究課題名: 機械学習と最先端計測技術の融合深化による新たな計測・解析手法の展開)」の支援を受けたものである。

## 参考文献

- [Blanchard 10] Blanchard, G., Lee, G., and Scott, C.: Semi-Supervised Novelty Detection, *J. Machine Learning Research*, Vol. 11, pp. 2973–3009 (2010)
- [Calvo 10] Calvo, B., Inza, I., Larrañaga, P., and Lozano, J. A.: Wrapper Positive Bayesian Network Classifiers, *Knowledge and Information Systems*, Vol. 33, No. 3, pp. 631–654 (2010)
- [Chapelle 10] Chapelle, O., Schölkopf, B., and Zien, A.: *Semi-supervised learning*, The MIT Press (2010)
- [Elkan 08] Elkan, C. and Noto, K.: Learning classifiers from only positive and unlabeled data, in *Proc. KDD08: the 14th ACM SIGKDD Int. Conf. on Knowledge Discovery and Data Mining*, pp. 213–220 (2008)



- [Gama 14] Gama, J., Žliobaitė, I., Bifet, A., Pechenizkiy, M., and Bouchachia, A.: A survey on concept drift adaptation, *ACM Computing Surveys (CSUR)*, Vol. 46, No. 4, pp. 44:1–44:37 (2014)
- [Hajizadeh 14] Hajizadeh, S., Li, Z., Dollevoet, R. P., M.J. D., and Tax, : Evaluating Classification Performance with only Positive and Unlabeled Samples, in *Proc. S+SSPR14: Structural, Syntactic, and Statistical Pattern Recognition*, Vol. LNCS 8621, pp. 233–242 (2014)
- [Jain 17] Jain, S., White, M., and Radivojac, P.: Recovering True Classifier Performance in Positive-Unlabeled Learning, in *Proc. AAAI17: the 31st AAAI Conf. on Artificial Intelligence*, p. 3060 (2017)
- [Lee 03] Lee, W. S. and Liu, B.: Learning with Positive and Unlabeled Examples Using Weighted Logistic Regression, in *Proc. ICML03: the 20th Int. Conf. on Machine Learning* (2003)
- [Li 09] Li, X.-L., Yu, P. S., Liu, B., and Ng, S.-K.: Positive Unlabeled Learning for Data Stream Classification, in *Proc. SDM09: the 2009 SIAM Int. Conf. on Data Mining*, pp. 259–270 (2009)
- [Menon 15] Menon, A., Rooyen, B. V., Ong, C. S., and Williamson, B.: Learning from Corrupted Binary Labels via Class-Probability Estimation, in *Proc. ICML15: the 32nd Int. Conf. on Machine Learning*, Vol. 37, pp. 125–134 (2015)
- [Niu 16] Niu, G., Plessis, du M. C., Sakai, T., Ma, Y., and Sugiyama, M.: Theoretical Comparisons of Positive-Unlabeled Learning against Positive-Negative Learning, in *Proc. NIPS16: Advances in Neural Information Processing Systems*, Vol. 29, pp. 1199–1207 (2016)
- [Pan 10] Pan, S. J. and Yang, Q.: A survey on transfer learning, *IEEE Trans. on Knowledge and Data Engineering*, Vol. 22, pp. 1345–1359 (2010)
- [Pfeffermann 98] Pfeffermann, D., Skinner, C., Holmes, D. J., Goldstein, H., and Rasbash, J.: Weighting for unequal selection probabilities in multilevel models, *J. the Royal Statistical Society. Series B (Statistical Methodology)*, Vol. 60, No. 1, pp. 23–40 (1998)
- [Plessis 14a] Plessis, du M. C., Niu, G., and Sugiyama, M.: Analysis of Learning from Positive and Unlabeled Data, in *Proc. NIPS14: Advances in Neural Information Processing Systems*, Vol. 27, pp. 703–711 (2014)
- [Plessis 14b] Plessis, du M. C. and Sugiyama, M.: Semi-Supervised Learning of Class Balance under Class-Prior Change by Distribution Matching, *Neural Networks*, Vol. 50, pp. 110–119 (2014)
- [Plessis 15] Plessis, du M. C., Niu, G., and Sugiyama, M.: Class-prior Estimation for Learning from Positive and Unlabeled Data, in *Proc. ACML15: the 7th Asian Conf. on Machine Learning*, Vol. 45, pp. 221–236 (2015)
- [Ramaswamy 16] Ramaswamy, H. G., Scott, C., and Tewari, A.: Mixture Proportion Estimation via Kernel Embedding of Distributions, in *Proc. ICML16: the 33rd Int. Conf. on Machine Learning*, Vol. 5, pp. 2996–3004 (2016)
- [Saerens 02] Saerens, M., Latinne, P., and Decaestecker, C.: Adjusting the Outputs of a Classifier to New a Priori Probabilities: A Simple Procedure, *Neural Computation*, Vol. 14, No. 1, pp. 21–41 (2002)
- [Scott 15] Scott, C.: A Rate of Convergence for Mixture Proportion Estimation, with Application to Learning from Noisy Labels, in *Proc. AISTATS15: the 18th Int. Conf. on Artificial Intelligence and Statistics*, Vol. 38, pp. 838–846 (2015)
- [Tsutsui 10] Tsutsui, M., Taniguchi, M., Yokota, K., and Kawai, T.: Identifying Single Nucleotides by Tunneling Current, *Nature Nanotechnology*, Vol. 5, pp. 286–290 (2010)
- [Ward 09] Ward, G., Hastie, T., Barry, S., Elith, J., and Leathwick, J. R.: Presence-only data and the em algorithm, *Biometrics*, Vol. 65, No. 2, pp. 554–563 (2009)
- [Washio 17] Washio, T., Imamura, G., and Yoshikawa, G.: Machine learning independent of population distributions for measurement, in *Proc. DSAA17: the 4th IEEE Int. Conf. on Data Science and Advanced Analytics*, pp. 212–221 (2017)