

# SVMによる誤差を含むクラス分類における CMOS アニーリングマシンの性能評価

水木 直也<sup>†</sup> 福原 諒河<sup>‡</sup> 森下 誠<sup>‡</sup> 河合 直聡<sup>§</sup> 片桐 孝洋<sup>‡, §</sup> 星野 哲也<sup>§</sup> 永井 亨<sup>§</sup>

名古屋大学 情報学部 コンピュータ科学科<sup>†</sup> 名古屋大学 大学院情報学研究科<sup>‡</sup>

名古屋大学 情報基盤センター<sup>§</sup>

## 1. はじめに

本発表では CMOS アニーリングマシン[1]を利用し、サポートベクターマシン(SVM)により誤差を有するクラス分類を行う際の性能評価を行う。先行研究[2]において、CMOS アニーリングマシンでの SVM 実装が検討されたが、データが誤差を含む場合の性能評価は行われていない。そこで、本研究では先行研究を進展させ、データが誤差を含む場合の性能評価を行う。

## 2. Support Vector Machine (SVM)

SVM とは最も近い点までの距離が遠くなるように決定境界を決める分類手法である。そのアルゴリズムについて線形分離可能な場合と不可能な場合に分けて紹介する。

### 2.1. SVM のアルゴリズム(線形分離可能)

$n$ 次元空間上の集合を  $n-1$ 次元の超平面で分離できることを線形可能と定義し、全部で  $N$ 個のデータが存在する場合を考える。超平面の方程式は

$$W^T X + b = 0 \quad \dots(1)$$

である。ここで  $X, W \in \mathbb{R}^n$  である。  $i$  番目データ  $X_i \in \mathbb{R}^n$  がクラス 1 に属するとき 1, クラス 2 に属するとき -1 を返す  $t_i \in \mathbb{R}$  を導入すると、条件式は

$$t_i(W^T X_i + b) > 0 \quad (i = 1, 2, 3, \dots, N) \quad \dots(2)$$

となる。

$N$ 次元空間上の 1 点と超平面との距離  $d$  は

$$d = \frac{|w_1 x_1 + w_2 x_2 + \dots + w_n x_n + b|}{\sqrt{w_1^2 + w_2^2 + \dots + w_n^2}} = \frac{|W^T X_i + b|}{\|W\|} \quad \dots(3)$$

なので、式(1)-(3)からマージン  $M$  を最大化するという条件は

$$\max_{w, b} M \frac{t_i(W^T X_i + b)}{\|W\|} \geq M \quad (i = 1, 2, 3, \dots, N) \quad \dots(4)$$

と表せる。

また  $\|W\| = \frac{1}{M}$  となるように標準化すると

$$\max_{w, b} \frac{1}{\|W\|}, \quad t_i(W^T X_i + b) \geq 1 \quad (i = 1, 2, 3, \dots, N) \quad \dots(5)$$

となる。計算量を削減するため、次のように変形する。

$$\min_{w, b} \frac{1}{2} \|W\|^2, \quad t_i(W^T X_i + b) - 1 \geq 0 \quad (i = 1, 2, 3, \dots, N) \quad \dots(6)$$

以上が線形分離可能な問題を SVM で解く際の大きな方法である。

### 2.2. SVM のアルゴリズム(線形分離不可能)

線形分離ができない場合のアプローチは、大きく「誤分類を許容する」と「高次元の座標に変換する」の 2 種類がある。本研究では、この両方を活用する。

#### 2.2.1. 誤分類を許容する

線形分離不可能な場合、式(6)の制約条件を満たせず学習できなくなるため、式(7)で表すスラック変数  $\varepsilon_i \in \mathbb{R}$  を導入する。

$$\varepsilon_i = \max\{0, 1 - t_i(W^T X_i + b)\} \quad \dots(7)$$

このスラック変数  $\varepsilon_i$  を、制約条件に加えある程度の誤分類を許容するようにすると

$$t_i(W^T X_i + b) - 1 + \varepsilon_i \geq 0 \quad (i = 1, 2, 3, \dots, N) \quad \dots(8)$$

と変形できる。

スラック変数  $\varepsilon_i$  は以下のような性質を持つ。

- $0 \leq \varepsilon_i \leq 1$  のときマージン範囲内である。
- $1 < \varepsilon_i$  のとき誤分類が発生する。
- $\varepsilon_i$  が大きくなればなるほど誤分類の度合いが大きい。

以上より、式(6)の誤分類を許容する関数に変換すると、

$$\min_{w, b} \frac{1}{2} \|W\|^2 + C \sum_{i=1}^n \varepsilon_i,$$

$$t_i(W^T X_i + b) - 1 + \varepsilon_i \geq 0 \quad (i = 1, 2, 3, \dots, N) \quad \dots(9)$$

となる。このとき、係数  $C$  が大きいほど誤分類の最小化関数に対する影響が大きくなる。そのため、 $C$  が大きいことは過学習寄りになっていると

Performance Evaluation of CMOS Annealing Machine for SVM-based Classification with Error

<sup>†</sup>Naoya Mizuki, Information system, Computer science, School of Informatics, Nagoya University

<sup>‡</sup>Ryoga Fukuhara, Makoto Morishita, Graduate School of Informatics, Nagoya University

<sup>§</sup>Masatoshi Kawai, Takahiro Katagiri, Tetsuya Hoshino, Toru Nagai, Information Technology Center, Nagoya University

みなせる．そのため， $C$ が小さいと未学習，汎化寄りになっていることを示す．

### 2.2.2. 高次元の座標に変換する

線形分離不可能な問題を，写像 $\phi$ により線形分離可能な高次元の座標に変換する．その後，逆変換し解を求める．しかし射影後の内積計算のコストが大きいので，写像 $\phi$ は直接定義しない．

本研究では

$$K(X_i, X_j) = \exp\left(-\frac{\|X_i - X_j\|^2}{2\sigma^2}\right) = \exp(-\gamma\|X_i - X_j\|^2) \dots (10)$$

で定義される RBF カーネルを使用した． $\gamma$ は「1点の学習データが識別面に影響を与える範囲」を表すパラメータである． $\gamma$ が大きくなるほど1点の影響範囲が小さい，すなわち，曲率が大きい識別面となることを意味する．

### 3. CMOS アニーリングマシン

CMOS アニーリングマシン[3]は，イジングモデルにより最適化処理を行う．そのため，記憶素子である SRAM の構造を活用できる．本研究では，日立が開発した非ノイマン型コンピュータを利用する．

具体的には，解きたい組合せ最適化問題を制約なし 0-1 整数計画問題，かつ 2 次までの決定変数の項のみの形へ帰着させ，専用マシンで解く．対象の問題をイジングモデルで定式化し，マシン上にマッピングすることができれば，問題を解くことができる．ここでイジングモデルとは，正方格子状に配置したスピン (+1 or -1) からなる格子モデルである．このモデルにおいては，各スピン同士は磁氣的に不均一に相互作用する．そのため，系の全エネルギーは，以下のコスト関数  $E$  で与えられる．

$$E = \sum_{i=1}^M \sum_{j \neq i} J_{ij} \sigma_i \sigma_j - \sum_{i=1}^M h_i \sigma_i \dots (11)$$

ここで  $M$  は全スピン数， $\sigma_i = \pm 1$  はサイト  $i$  のスピン変数， $J_{ij}$  は異なるスピン間の相互作用強度， $h_i$  はサイト  $i$  の局所磁場である．つまり，対象の問題に対して対応したイジングモデル ( $h_i, J_{ij}$ ) を構成する．次に，その最低エネルギー状態 (スピン配列) を求めることで，組合せ最適化問題の最適解を求めることができる．

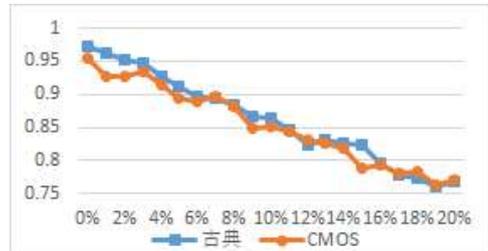
### 4. 問題設定

本研究では，2 次元上に乱数を生成し，ある関数より大きい小さいかで 2 クラスに分類する問題を取り扱う．誤差が 0%, 1%, 2%, ..., 20% の各ケースについて 1600 のデータを 5 セットずつ生成する．このデータを，各 100 の訓練データ，1000 の検証データ，および 500 のテストデータに分割し，

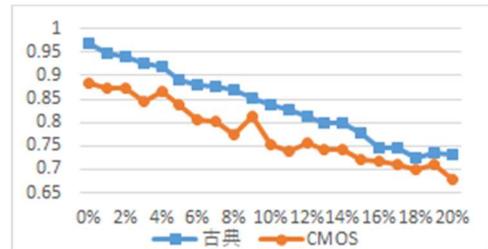
ホールドアウト法を適用した．

古典コンピュータでの SVM の実行は Windows10, Intel Core i9-13900KF 搭載のパソコンで，Ubuntu 22.04.3 (WSL) を使用した．CMOS アニーリングマシンは先行研究[4]の環境を用いた．

### 5. 実験結果とまとめ



(a) 線形分離可能の場合



(b) 線形分離不可能の場合

図 1 実験結果 (X 軸は誤差, Y 軸は正答率 [%])

図 1(a) より，線形分離可能な問題については古典の平均精度が 86.2%，CMOS の平均精度が 85.3% とわずかに古典の方が良い結果となった．また図 1(b) より，線形分離不可能な問題については古典の平均精度が 83.9%，CMOS の平均精度が 77.8% と 6 ポイントほど古典のほうが精度がよい．

先行研究[2]に対し，ハイパーパラメータチューニングを行うことで性能向上が確認できた．しかし，線形分離不可能での性能差の原因，CMOS の処理時間が古典に比べて長い点について分析する必要がある．また，自動チューニングの適用[4]も今後の課題である．

**謝辞** 本研究は JSPS 科研費 JP19H0566 の助成を受けたものです．

#### 参考文献

- [1] M. Yamaoka, "A 20k-spin Ising chip for combinatorial optimization problems with CMOS annealing, IEEE International Solid-State Circuits Conference, 2015.
- [2] 福原ほか, "CMOS アニーリングにおけるサポートベクターマシンの性能評価", 情処研報告 2023-HPC-190, pp. 1-6, 2023.
- [3] CMOS アニーリングマシン - Annealing Cloud Web <https://annealing-cloud.com>
- [4] 森下ほか, "量子コンピューティングへの自動チューニングの適用と評価", 情処研報告 2023-HPC-188, pp. 1-7, 2023.