

# 画像分類におけるアニーリングの活用

理学専攻・情報科学コース 1940645 坂倉 佑季

## 1 はじめに

「組合せ最適化問題」とは、様々な制約のもとで複数の選択肢の中から、ある指標（価値）を最もよくする変数の値（組合せ）を求める問題のことである。インターネットが広く普及し、IoT（モノのインターネット化）や AI が発展してきた現代において、膨大な量のデータの入手が手軽に収集可能になり、大量の選択肢の中から最適なものを選択することが重要になってきている。しかし、このような組合せ最適化問題は、データ量が多くなるにつれ組合せの数が爆発的に増加し、従来のコンピュータでは解くのに非常に時間がかかる。そんな最適化問題を高速かつ高精度に解いてくれると期待されているのが、アニーリングマシンである。本研究では、組合せ最適化問題の中でも様々な分野で応用されている画像分類に着目し、アニーリングを活用して解析を行うことを考えた。手法として、アニーリングマシン用の Boosting アルゴリズムである QBoost[1] と、ソフトマックス関数を用いた量子機械学習 [2] を用いて画像分類を実行し結果を考察した。

## 2 QBoost を用いた解析

### 2.1 モデルと方法

従来の Boosting アルゴリズムは弱識別器に対する重み  $w_i$  が連続値であり、アニーリングマシンにそのまま載せることができないので、アニーリングマシン用の Boosting アルゴリズムである QBoost を用いることにする [1]。

QBoost では弱識別器の重みは 0 か 1 の値を取り、 $i$  番目の識別器を使用するかどうかを表している。最適な重み  $w_i$  の組み合わせを求めることで、どの識別器を使えば出来るだけ少ない数の弱識別器でより良い性能を得ることができるかを明らかにすることができる。最適な重み  $w_i$  の組み合わせを決定するために、アニーリングマシンを用いる。問題を QUBO の形式 (1) にするために識別器の出力と正解との差の二乗を最小化する問題とする。 $d$  番目の画像ベクトル  $\mathbf{x}^{(d)}$  に対する教師データを  $y^{(d)}$  とすると、コスト関数は式 (2) で与えられる。

$$H = \sum_i h_i w_i + \sum_{i < j} J_{ij} w_i w_j \quad (1)$$

$$\begin{aligned} H(\mathbf{w}) &= \sum_{d=1}^D \left( \frac{1}{N} \sum_{i=1}^N w_i c_i(\mathbf{x}^{(d)}) - y^{(d)} \right)^2 + \lambda \sum_{i=1}^N w_i \\ &= \sum_{i=1}^N \left( \lambda - \frac{2}{N} \sum_{d=1}^D c_i(\mathbf{x}^{(d)}) y^{(d)} \right) w_i \\ &\quad + \frac{1}{N^2} \sum_{i,j=1}^N \left( \sum_{d=1}^D c_i(\mathbf{x}^{(d)}) c_j(\mathbf{x}^{(d)}) \right) w_i w_j \\ &\quad + \text{const.} \end{aligned} \quad (2)$$

ここで、 $c_i(\mathbf{x}) = \pm 1$  は弱識別器  $i$  の出力、 $D$  は画像の

枚数、 $N$  は弱識別器の数、 $\lambda$  は過学習を防ぐための正則化パラメータである。

### 2.2 手書き数字画像の分類

手書き数字データセット MNIST [3] の多値分類を行う。QBoost は二値分類を行うアルゴリズムであるため、0 とそれ以外、1 とそれ以外、 $\dots$ 、9 とそれ以外を分類する強識別器を 10 個用意し、多値分類を行う。その際スコア  $f(\mathbf{x})$  を次式のように定義した。

$$f(\mathbf{x}) = \sum_{i=1}^N w_i c_i(\mathbf{x}) \quad (3)$$

同じ画像に対して判定が複数出た場合、このスコアの値が大きい方の判定を採用する。正則化項のパラメータ  $\lambda = 3$ 、弱識別器の数は 60 個、木の深さは 5、訓練データとテストデータの割合は 8:2 と設定した。この時、訓練データ数を増やした際の正答率を表したものが図 1 である。

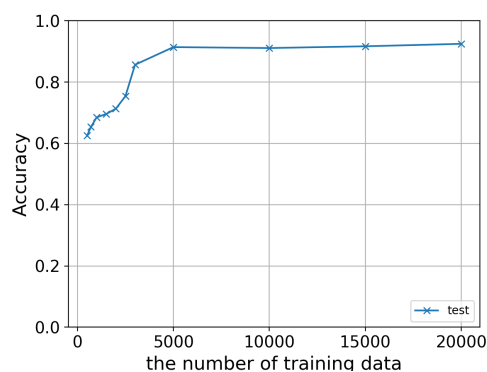


図 1: 手書き数字画像の多値分類結果

訓練データを増やすにつれ、正答率が上昇し、訓練データが 5000 枚あれば 9 割近い正答率を得ることがわかった。また、多値分類を行う際のスコアが機能していることが確認でき、定義が妥当であることが検証できた。平均正答率は平均 0.91 であり、数字 0 ~ 9 のどれにも分類されなかった画像が 6.1%、間違えて判定されたものが 2.9% であることがわかった。

### 2.3 考察

QBoost を用いる利点は、スパース性によりデータ量を削減することができることと、アニーリングマシンを使用する際に必要となるビット数が画像の大きさに依存しないことである。欠点は、QBoost を用いて多値分類を行った際、どれにも分類されない画像が出てくること、何箇所かのピクセル値だけを用いて判定を行なっているため、カラー画像や人物画像などの画像は判定が難しいということが挙げられる。ただ、どれにも分類されない画像が出てくることを利用し、学習データの分布に含まれていなかった入力データ (Out-of-Distribution) の検出に応用できると考えられる。

### 3 量子機械学習を用いた解析

#### 3.1 モデルと方法

アニーリングマシンを用いてトレーニングを行うことにより、データセット内のクラスが与えられる確率を最大化する重みを見つける。つまり、負の対数尤度を QUBO の形式に変換し、アニーリングマシンを用いて最小化することにより重みを求める。そして、求めた重みとソフトマックス関数を用いて、ある画像に対する各クラスの確率を求めることができる。確率が最も高いクラスに基づいて予測ラベルを割り当て、画像の判定を行う [2]。\$i\$ 番目の \$M\$ 次元特徴ベクトル \$\mathbf{x}\_i\$ に対応するクラスを \$y\_i\$ とすると、コスト関数は次式のようになる。

$$\mathcal{L} \approx \sum_{k=1}^{K-1} \mathbf{w}_k^T (\mathbf{b}_k + \mathbf{h}') + \sum_{k=1}^{K-1} \mathbf{w}_k^T \mathbf{J}' \mathbf{w}_k - \sum_{k=1}^{K-1} \sum_{j \neq k} \mathbf{w}_j^T \mathbf{J}'' \mathbf{w}_k \quad (4)$$

$$\mathbf{b}_k = \sum_{i: y_i = k} -\mathbf{x}_i \quad (5)$$

$$\mathbf{h} = \frac{1}{K} \sum_i \mathbf{x}_i \quad (6)$$

$$\mathbf{J}' = \frac{K-1}{2K^2} \sum_i \mathbf{x}_i \mathbf{x}_i^T \quad (7)$$

$$\mathbf{J}'' = \frac{1}{2K^2} \sum_i \mathbf{x}_i \mathbf{x}_i^T \quad (8)$$

式 (4) をアニーリングマシンで最小化することにより、クラスに対応する最適重み \$\mathbf{w}\$ が求まる。求めた重みとソフトマックス関数を使用し、画像を分類する。

#### 3.2 CIFAR-10 画像の分類

飛行機、自動車、鳥、猫、鹿、犬、カエル、馬、船、トラックの 10 種類の画像が入った CIFAR-10 [4] を用いて画像分類を行なった。分類結果を可視化したものが図 2 である。左側は判定した画像、その下に書かれているものは判定結果とその確率、括弧の中が正解ラベルとなっている。右側のグラフは、それぞれのクラスに対する確率を棒グラフにしたものである。ここで、赤は間違えて判定、青は正しく判定されたものを表している。正答率としては、0.63 となった。

間違えた画像に着目すると、間違えてはいるものの 2 番目か 3 番目に高い確率が得られていることが確認できた。また、10 種類の画像としてではなく、動物か乗り物かで見るとだいたい判定できているのではないかと考えられる。そこで次に、動物と乗り物の画像で二値分類を実行した。ここでは、動物 3 種類と乗り物 3 種類の画像を用いて分類を行う。結果を表 1 に示す。8 割近い正答率を得ることができ、動物か乗り物かの二値分類に成功した。

#### 3.3 考察

ソフトマックス関数を用いた手法における画像分類ではカラー画像などの複雑な画像に対しても良い正答率が得られることが確認できた。また、どのラベルに対してどれだけの確率となっているのかを可視化でき

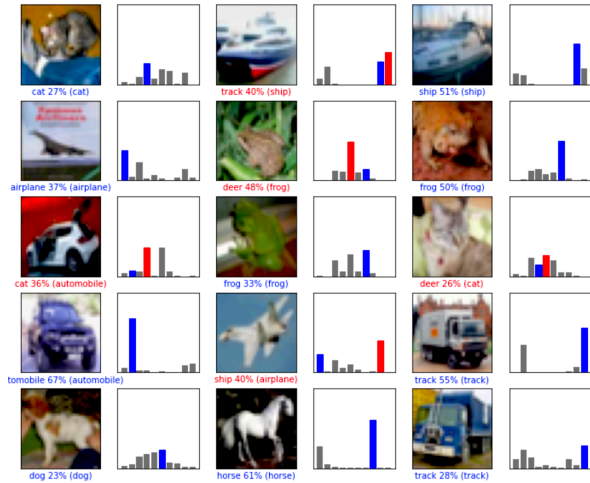


図 2: 分類結果

動物 or 乗り物	ソフトマックス関数を用いた手法
(猫・犬・鹿) or (飛行機・車・船)	0.81
(鳥・犬・蛙) or (飛行機・車・船)	0.73
(鳥・蛙・馬) or (車・船・トラック)	0.76
(鹿・犬・馬) or (車・船・トラック)	0.79

表 1: 動物と乗り物の分類結果

るため、「間違えていても 2 番目に高い確率が与えられていた」などの情報を得られることは利点と考えられる。

### 4 まとめと今後の課題

二値分類アルゴリズムである QBoost において、スコアを定義することにより多値分類に応用することに成功し、手書き数字の多値分類で約 9 割の正答率を得ることができた。しかし、カラー画像などの複雑な画像に対しては精度があまり良くないため、もう一つの手法としてソフトマックス関数を用いた画像分類を行った。こちらの手法ではカラー画像に対して 6 ~ 7 割近い正答率を得た。従来の手法と比較し、アニーリングマシンを用いた画像分類の精度向上や実行時間短縮の方法を検討するのが今後の課題である。

#### 参考文献

- [1] H. Neven, et al., “Training a Large Scale Classifier with the Quantum Adiabatic Algorithm”, arXiv:0912.0779. (2009)
- [2] Richard Y. Li, et al., “Unconventional machine learning of genome-wide human cancer data”, arXiv:1909.06206. (2019)
- [3] 手書き数字データセット (mnist), <https://www.kaggle.com/c/digit-recognizer/data> (2018 年 10 月 23 日アクセス)
- [4] CIFAR-10 dataset, <http://www.cs.toronto.edu/kriz/cifar.html> (2019 年 10 月 20 日アクセス)