

選択的不感化ニューラルネットによる特徴量の有効性の分析

染野 翔一[†] 堀江 和正^{††} 森田 昌彦^{†††a)}

Analysis of Feature Effectiveness Using a Selective Desensitization Neural Network

Shoichi SOMENO[†], Kazumasa HORIE^{††}, and Masahiko MORITA^{†††a)}

あらまし 機械学習によって関数近似を行う場合、入力変数（特徴量）の中に出力と無関係のものがあると近似精度が低下するケースが多いが、完全に無関係かどうかを判断するのは容易でない。本論文では、選択的不感化ニューラルネット（SDNN）を用いて、各入力変数が出力に関係するか、すなわち特徴量が有効かどうかを判別する方法を提案する。人工的に作成した関数近似課題を用いて学習実験を行ったところ、SDNNは無効変数が多く含まれていても、比較的少ないサンプルで十分に高い近似精度が得られること、無効変数のみに関係するSDNNの結合荷重はそれ以外のものに比べて分散が小さくなることがわかった。また、このことを利用して、出力と相関がないが関係する変数、限られた場合にのみ関係する変数や、単独では効果が薄い相乗効果を発揮する変数のペアも無効変数と区別できるという結果が得られた。このことは、SDNNの計算コストの大幅な削減に利用できるとともに、従来よりも効果的な特徴選択や非線形データ解析の手法の開発につながると考えられる。

キーワード 関数近似、選択的不感化ニューラルネット、特徴選択、無効変数

1. ま え が き

機械学習を用いて特徴量に基づくパターン分類や関数近似を行う場合、入力する特徴量の中に出力値の決定に貢献せず、出力との関係性がないものが含まれている場合がある。このような特徴量のことを無効な特徴量、これに対応する入力変数を無効変数と呼ぶ。こうした無効変数は、入力空間に対するサンプル密度の相対的な低下につながり、多くの関数近似器において近似精度の低下を引き起こす [1], [2] ことから、高精度な近似を行うためには除去することが望ましい。

しかしながら、特徴量が有効か否かを判別するのはそれほど容易ではない [3]~[5]。最も簡単な有効性の判別方法は入出力間の相関を求めることだが、これは

入出力関係が非線形の場合適用できない。また、他の特徴量と組み合わせると初めて有効となるもの（組み合わせ有効な特徴量） [6] や、ある特定の状況下でのみ出力に影響する特徴量（部分的有効な特徴量） [7] について、その有効性を正しく判別する現実的な方法は知られていない。

本研究では、入出力関係を機械学習を用いてモデル化し、そのモデルを解析することによってどの入力変数（特徴量）が無効であるか判別することを考える。このような場合、モデル化を行う関数近似器は、以下の特徴を有する必要があると考えられる。

(1) 無効特徴量や部分的有効な特徴量が多数含まれていても高精度に関数を近似できること。無効特徴量によって近似精度が大きく低下するのでは、学習後の解析結果が信頼できない。また、部分的有効な特徴量の出力への寄与を正しく捉えられる必要がある。

(2) 各入力出力にどのように寄与したのか、解析が容易であること。特徴量の除外や追加を繰り返して有効性を評価する方法よりも手間がかかるのではあまり意味がない。

しかし、一般的な関数近似器ではこれらの条件を満たしていない。例えば、放射状基底関数ネットワークは、無効特徴に弱い。サポートベクター回帰 (SVR)

[†] 筑波大学大学院システム情報工学研究科，つくば市
Graduate School of Systems and Information Engineering,
University of Tsukuba, 1-1-1 Tennodai, Tsukuba-shi, 305-8573 Japan

^{††} 筑波大学計算科学研究センター，つくば市
Center for Computational Sciences, University of Tsukuba,
1-1-1 Tennodai, Tsukuba-shi, 305-8577 Japan

^{†††} 筑波大学システム情報系，つくば市
Faculty of Engineering, Information and Systems, University
of Tsukuba, 1-1-1 Tennodai, Tsukuba-shi, 305-8573 Japan

a) E-mail: mor@bcl.esys.tsukuba.ac.jp
DOI:10.14923/transinfj.2018JDP7075

は非線形カーネルを用いないとモデル化を十分に行えないが、カーネルを介するため元の特徴量の有効性を解析しづらい。また、深層ニューラルネットは、入力時間の演算が非常に多いため解析が難しい。

3層の多層パーセプトロンを用いて特徴量の有効性を判定した研究 [8] はあるが、無効特徴のみを対象としている。実問題では部分的有効特徴が含まれる場合が多いと思われるが、多層パーセプトロンは部分的有効特徴を苦手としており [7]、この研究でも扱われていない。

そこで本研究では、近年開発された選択的不感化ニューラルネット (Selective Desensitization Neural Network: SDNN) [9] に着目した。SDNN は表現能力が非常に高いだけでなく、無効変数や部分的有効変数があっても高精度なモデル化が可能である [1], [7]。また、一般的な SDNN では学習時に結合荷重の最適化が行われる層が一つしかないうえ、各荷重には特定の二つの入力変数しか関わらないため、荷重の解析によって特徴量または特徴量ペアの寄与を他のニューラルネットに比べて容易に把握できると考えられる。

しかしながら、SDNN を用いて特徴量の有効性を調べる具体的な方法は不明であり、部分的有効特徴や組み合わせで有効性が高まる特徴が含まれる場合に無効特徴と区別できるかも明らかでない。また、無効変数が含まれる場合の SDNN の関数近似能力、特にサンプル数と近似精度との関係は調べられていない。以上をふまえて本研究では、真値が明らかな人工関数の近似課題を用いて、SDNN の無効変数に対するロバスト性を確認するとともに、学習後の結合荷重から無効変数や無効変数のペアを判別する方法を提案する。

2. 選択的不感化ニューラルネット

SDNN は、分散表現された入力の情報統合操作である「選択的不感化」を導入した順伝搬型ニューラルネットである [9]。本研究では一般的な関数近似問題で用いられる、パターンコーディング [10], [11]、選択的不感化、並列パーセプトロンからなる SDNN を使用する (図 1)。本章では SDNN の各要素の具体的な操作について述べる。

2.1 パターンコーディング

SDNN へ入力されたアナログ値は、まずパターンコーディングにより多次元の 2 値パターン (コードパターン) に変換される。具体的には、入力を取りうる範囲を q 個に分割し、それぞれにコードパターン

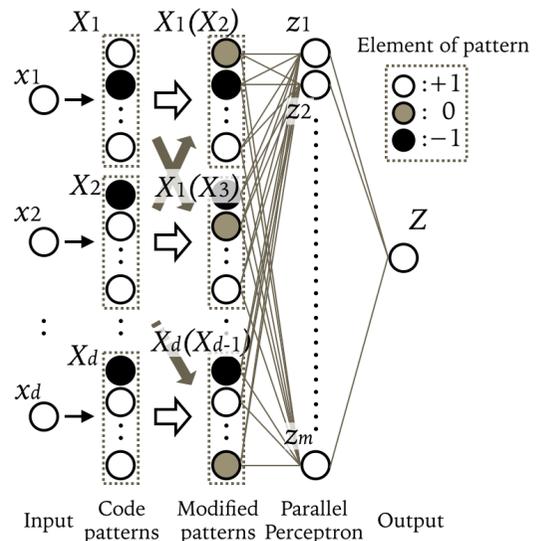


図 1 関数近似を行う SDNN の構造

Fig. 1 Structure of an SDNN for function approximation.

P^1, P^2, \dots, P^q を対応づけたルックアップテーブルを作成することで、入力をコードパターンに変換する。

コードパターン P^1, \dots, P^q には、同数の ± 1 からなる n 次元 2 値パターン (n は偶数) を用いる。また、2 つのコードパターンは表す値が近いほどパターン間の相関が大きく、離れるほど 0 に近づくことが望ましい。そのようなコードパターンの作成方法は幾つか存在するが [1], [12]、本研究では以下の手順を用いる。

まず、+1 と -1 が半分ずつになるパターンをランダムに作成して P^1 とする。 P^1 の成分の中から +1 及び -1 をとるものを r 個ずつランダムに選び、それらの符号を反転して得られるパターンを P^2 とする。同様に、 P^2 の成分のうち $2r$ 個の符号を反転して P^3 とする。以下同様に、 P^{k-1} の $2r$ 個の成分を反転することによって順次 P^k を作成する。

コードパターンの次元数 n 、入力範囲の分割数 q 、隣接パターンの反転数 r は SDNN の表現能力・汎化能力に影響を与える。本研究では、予備実験を基にパラメータの組み合わせを選択、全ての実験において同一のパラメータを使用した。

2.2 選択的不感化

入力を分散表現した複数のパターンがあるとき、あるパターンの一部を他のパターンに合わせて不感化 (中立値に変換) し、これらがもつ情報を統合することを選択的不感化と呼ぶ (図 2)。本研究では不感化す

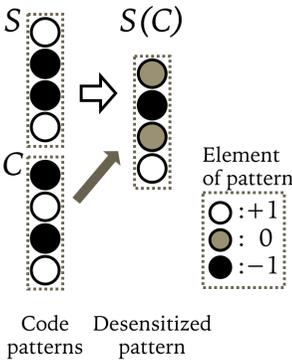


図2 選択的不感化の例
Fig.2 Example of selective desensitization.

るパターンを修飾パターン，不感化されるパターンを被修飾パターンと呼ぶことにする。

不感化の手順の説明のため，具体的にパターン $S = (s_1, s_2, \dots, s_n)$ をパターン $C = (c_1, c_2, \dots, c_n)$ で不感化する場合について考える。また， S と C の各要素にはあらかじめランダムに決められた1対1の対応付けがなされているとする。

s_i と c_j に対応があるとき，不感化後の素子 s'_i は，

$$s'_i = \frac{1 + c_j}{2} s_i \quad (1)$$

と表される。

$c_j = +1$ のとき， s_i の値はそのまま保持され， $c_j = -1$ のときには中立値 (= 0) に変換される。今， C の要素のうち半数は -1 であるため，この操作により半数の要素は 0，残り半数の要素は ± 1 の値をもつ3値パターン $(s'_1, s'_2, \dots, s'_n)$ が得られる。このパターンを「 C によって修飾された S 」と呼び， $S(C)$ と表す。 s_i と c_i に対応付けがなされている場合の例を図2に示す。

u 個の入力変数とそれに対応したパターンがあるとき，基本的にはこれらのパターンの全組み合わせで不感化操作を行い， $u(u-1)$ 個の3値パターンを作成する。

2.3 並列パーセプトロン

選択的不感化の結果得られた全パターンは，並列パーセプトロン[13]に入力される。並列パーセプトロンは $\{0, 1\}$ のいずれかの値を出力する単純パーセプトロンを並列にならべたもので，各単純パーセプトロンの出力値の合計に合わせて全体の出力を決定する。

並列パーセプトロンの学習には誤り訂正学習が用いられる[13]。例えば，ある入力に対し， m 個の単純

パーセプトロンのうち l 個が1を出すべきなのに，実際には $k (< l)$ 個が1を出力したとする。このとき，0を出した $m - k$ 個のうち，内部電位がしきい値に近いものから順に $l - k$ 個選び，これらが誤っているとみなして学習を行う。 $k (> l)$ の場合は1を出した単純パーセプトロンを対象に，内部電位がしきい値に近いものから $k - l$ 個学習を行う。

3. 実験 1

SDNN の無効変数に対するロバスト性を定量的に確認するとともに，学習後の結合荷重から無効変数を判別する方法を探るために数値実験を行った。ここでいうロバスト性とは，無効変数が含まれている場合でもできるだけ少ないサンプル数で誤差の増加を抑えられることを意味する。先行研究で使用された2変数関数[12]に無効変数を加えた6変数関数を人工的に作成し，これをSDNNとSVRを用いて近似したときの精度を比較した。またSDNNの結合荷重の分布と入力変数の有効性との関係を調べた。

3.1 標的関数と学習サンプル

本実験では，以下の6変数関数 $f(\mathbf{x})$ を使用した。

$$\mathbf{x} \in \{(x_1, x_2, \dots, x_6) | \forall x_i \in [0, 1]\}$$

$$f(\mathbf{x}) = \begin{cases} 1 & ((x_1 - 0.5)^2 + (x_2 - 0.5)^2 \leq 0.04) \\ \frac{1 + x_1}{2} \sin^2(6\pi\sqrt{x_1 x_2}^2) & (\text{otherwise}) \end{cases} \quad (2)$$

x_1, x_2 は $f(\mathbf{x})$ の出力に寄与する有効変数， x_3, \dots, x_6 が無効変数である。上記関数の定義域(6次元の入力空間)中に設定した0.001間隔の格子点(1001⁶個)からランダムに100000点を選び，これらの一部または全部を学習サンプルとして学習と評価を行った。近似誤差の評価には，ランダムに選んだ学習サンプルとは別の10000点における平均2乗誤差平方根(Root mean squared error: RMSE)を用いた。

3.2 関数近似器とその設定

SDNN に用いたパラメータを表1に示す。なお並列パーセプトロンの素子のうち k 個が1を出した場合，最終的な出力値は $Z = 0.005k - 0.2$ とした。

このSDNNに対して各サンプルを500回学習させた。ただし，学習の最中に学習サンプルに対する平均絶対値誤差が0.01を下回った場合，その時点で学習を打ち切った。

SVR はガウシアンカーネルを用い， $c =$

表 1 実験で使用した SDNN の各パラメータ

Table 1 Parameters of the SDNN used in the experiment.

パラメータ	設定値
入力範囲の分割数 q	101
隣接パターンの反転数 r	2
パターンの要素数 n	200
単純パーセプトロン数 m	280

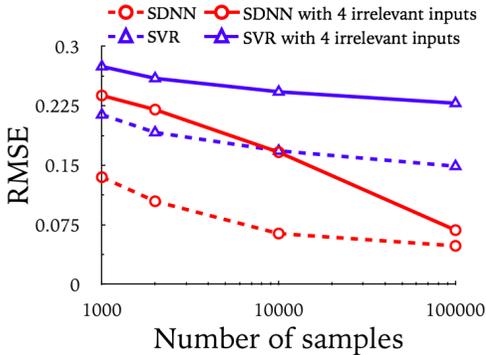


図 3 サンプル数に対する平均 2 乗誤差平方根 (点線は有効変数 (2 変数) のみ, 実線は有効変数 (2 変数) と無効変数 (4 変数) の両方を用いて近似した)

Fig. 3 Root mean squared error (RMSE) against the number of training samples.

$\{2^{-5}, 2^{-4}, \dots, 2^5\}$, $\gamma = \{2^{-5}, 2^{-4}, \dots, 2^5\}$, $p = \{2^{-8}, 2^{-7}, \dots, 2^{-1}\}$ の範囲でグリッドサーチを行った. SVR の実装には libSVM [14] を用いた.

3.3 結 果

SDNN と SVR のサンプル数に対する近似誤差の変化を比較した (図 3). SVR の場合, 無効変数による誤差の増加量は, サンプル数にほとんど依らないのに対し, SDNN ではサンプルが増えるにつれて誤差の増加量が減り, サンプル数が 100000 のときには無効変数の影響はごくわずかになっている. また, サンプル数が 10000 以上のとき, 無効変数がある場合の SDNN の方が, 無効変数がない場合の SVR よりも近似精度が高い.

以上のように, SDNN の方がより少ないサンプルで無効変数の影響を抑えることができている. 無効変数に対してロバストだといえる. なお, 標的関数を変えて同じ実験を行ったところ, 一様でならぬ関数の場合には SVR の方が誤差が少ないことがあった. しかし, 急激な変化を含む複雑なものになるほど SDNN の優位性が高まる傾向が見られ, またどの場合も SDNN の方が少ないサンプルで無効変数の影響を抑えられる

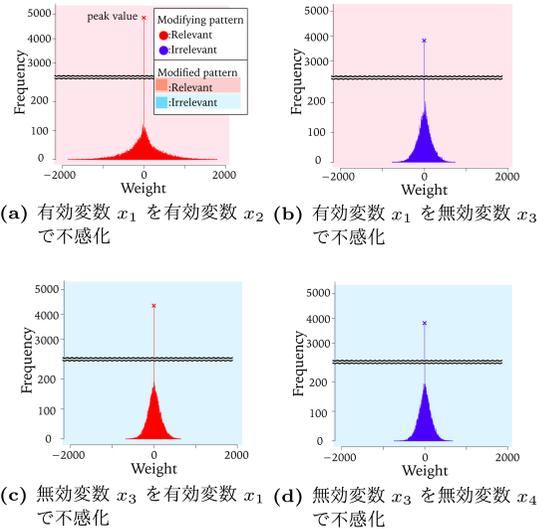


図 4 SDNN の結合荷重のヒストグラム

Fig. 4 Histogram of the synaptic weights of the SDNN.

という点は一貫していた. 無効変数に対する SDNN のロバスト性を示した先行研究 [1], [16] も合わせると, この結果はかなり一般性をもつと考えられる.

図 4 は, 10000 点のサンプルを学習した時点での各素子群の n 個の素子から m 個の単純パーセプトロンへの mn 個の結合荷重の分布を示したものである. 入力変数 x_1, \dots, x_6 は, パターンコーディング及び選択的不感化によって $X_1(X_2), \dots, X_6(X_5)$ の 30 個のパターンに変換される. これらのパターンは, 修飾及び被修飾パターンが有効変数か無効変数かによって 4 通りに分類される. パターン $X_\mu(X_\nu)$ を表現する素子群を $G_{\mu,\nu}$ と表したとき, これらも 4 種類に分けられる. そこで, この 4 種類の素子群に関する結合荷重の分布を比較した. ヒストグラムの色は被修飾変数の有効性 (赤: 有効, 青: 無効) を, 背景は修飾変数の有効性を表している.

この図より, 有効性の有無の組み合わせによって, 結合荷重の分布が異なることがわかる. 特に有効変数同士の場合, 荷重が 0 の結合が多数ある一方で絶対値の大きなものも見られるのに対し, 無効変数同士の場合には比較的分布の偏りが小さい. この分布の違いを反映した数値的な指標が得られれば, それを基に入力変数の有効性を判断できる可能性がある.

そのような指標として, 分布の分散, 歪度, 尖度及び絶対値の平均を候補とし, 各素子群, すなわち被修

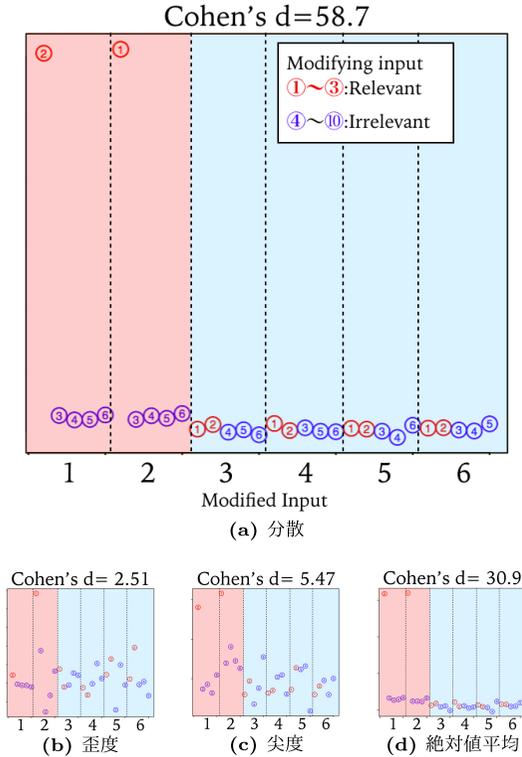


図 5 様々な指標で集約した入力の変数ペアに関する結合荷重
 Fig. 5 Variance, skewness, kurtosis and mean of the synaptic weights for each variable pair.

飾変数と修飾変数の組み合わせ（以下、変数ペアと呼ぶ）に関する結合荷重についてそれぞれ計算した結果を図 5 に示す。図の横軸は変数ペア、縦軸は各指標値を表す。図 5a～図 5d を比べると、有効変数同士のペアとその他の違いが大きく、入力の有効性を判断する上で最も効果的な指標は分散だと考えられる。またこれらの指標ごとに Cohen's d [17], [18] を計算したところ、分散の場合が最も大きかった。この値は 2 つのクラスの平均値の差をそれぞれの標準偏差を用いて正規化したもので、これらが大きいほどクラスを分割するのに有効であることから、分散が指標として最も優れているといえる。

4. 実験 2

実験 1 で用いた関数は、個々の変数が有効か無効か明確に別れていた。しかし、実際には部分的に有効な変数や、単独では有効でない変数が混じっていることも想定される。このような場合にも、SDNN の結合荷

表 2 各入力変数の有効性
 Table 2 Effectiveness of each input.

入力	有効性
x_1	有効変数
x_2, x_3	組み合わせ有効変数
x_4, x_5	部分的有効変数
x_6, \dots, x_{10}	無効変数

重の分散に基づいて有効性を判断できるかどうか検討した。標的関数とサンプル数 (10000 点)、格子点ではなく連続値をサンプルとした以外の条件は実験 1 と同じである。

4.1 標的関数

本実験では、標的関数 $f(x)$ として以下の式で表されるものを使用した。

$$\begin{aligned}
 & \mathbf{x} \in \{(x_1, x_2, \dots, x_{10}) | \forall x_i \in [0, 1]\} \\
 & k(\mathbf{x}) = x_1^2 + x_1(x_2 + x_3 - 0.5) \\
 & \quad - 2(x_2 + x_3 - 0.5)^2 \\
 & \quad - x_1 + 2(x_2 + x_3 - 0.5) + 0.1 \\
 & g(\mathbf{x}) = \begin{cases} k(\mathbf{x}) + x_4 - 0.5 & (x_1 < 0.1) \wedge (x_2 < 0.1) \\ k(\mathbf{x}) + x_5 - 0.8 & (x_5 > 0.8) \\ k(\mathbf{x}) & (\text{otherwise}) \end{cases} \\
 & f(\mathbf{x}) = \begin{cases} 0 & (g(\mathbf{x}) < 0) \\ 1 & (g(\mathbf{x}) > 1) \\ g(\mathbf{x}) & (\text{otherwise}) \end{cases} \quad (3)
 \end{aligned}$$

本関数では、 x_1, x_2, x_3 が有効変数である。特に x_2 と x_3 は常に和の形で出力に影響を与えることから単独での有効性は低いが、組み合わせると非常に有効であると考えられる。 x_4, x_5 は、特定の条件下でのみ出力に影響する部分的有効変数であり、特に x_4 は出力との相関がほぼ 0 であるため相関分析では有効性を判断できない。 x_6, \dots, x_{10} は無効変数とした (表 2)。

4.2 結果

実験 1 と同様に SDNN を用いて標的関数の近似を行ったところ、真値との誤差 (RMSE) は約 0.035 であり、高精度での近似ができた。変数ペアごとの結合荷重の分散を図 6 に示す。まず、有効変数、部分的有効変数、無効変数と比較すると、被修飾変数の有効性順に分散値が大きくなるが見て取れる。組み合わせ有効変数については、これらのペアの組み合わせに対する分散が x_2, x_3 の片方のみ含まれるペアに比べ

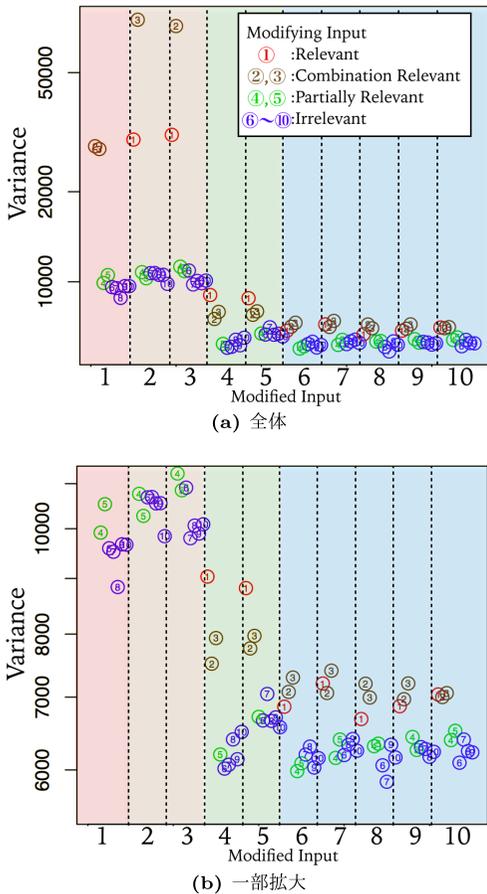


図 6 各不感化パターンの結合荷重についての分散
Fig. 6 Variance of the synaptic weights for each variable pair.

非常に大きかった。以上のことから、組み合わせ有効変数、部分的有効変数についても有効変数と同様に、分散値を用いることによって無効変数と区別することが可能であった。

この結果を利用して、計算コストをどの程度削減できるか検討した。SDNN で関数近似を行う場合、特別な事前知識がない限り、全ての組み合わせで不感化を行うのが一般的であるが、そうすると計算コストが入力次元数の 2 乗に比例して増大してしまう。結合荷重の分散値が大きい変数ペアのみを用いれば、その後の計算コストを減らすことができると考えられる。そこで、90 個の変数ペアを分散値の大きい順に並べ、最初の k ペアのみを用いて最初から学習をやり直して近似誤差を求めた。変数ペアの数 k に対する近似誤差の変化を図 7 に示す。結合荷重の分散値が大きい 9 個

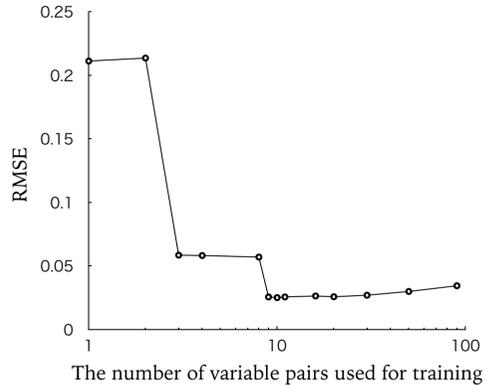


図 7 推定に使用した変数ペア数に対する近似誤差 (変数ペアは荷重の分散が大きいものから使用)
Fig. 7 RMSE against the number of variable pairs used to train the SDNN; the variable pairs were used in order from the ones with the largest variance of the synaptic weights.

のペア $X_2(X_3)$, $X_3(X_2)$, $X_3(X_1)$, $X_2(X_1)$, $X_1(X_2)$, $X_1(X_3)$, $X_3(X_4)$, $X_3(X_6)$, $X_3(X_5)$ を用いて学習したとき、全 90 ペアを用いたときよりも誤差が小さくなった。この 9 ペアに含まれる変数は、一部無効なものも含まれるが、大部分が有効な変数 (組み合わせ有効変数と部分的有効変数を含む) であり、無効変数同士のパアは含まれていない。また 50 番目までの変数ペアには無効変数同士のパアは含まれていないが、 $k = 50$ のときの誤差は $k = 9$ のときとほとんど同じであった。

9 ペアのみ用いた場合、全ペアを使用した場合に比べ計算コストは約 10 分の 1 に削減できる。それ以上有効変数を含むペアを追加しても誤差は減少しないこと、逆に多少無効変数を含むペアが含まれていても誤差が増えることはないということから、使用するペアを多めに選ぶことによって、必要な変数を誤って取り除いてしまう危険性を避けられると考えられる。このように、結合荷重の分散を指標に使用する変数ペアを決めることによって、近似誤差の最小化と計算コストの大幅な削減が可能であることがわかった。

5. むすび

無効変数を含む関数近似課題について数値実験を行い、選択的不感化ニューラルネットワーク (SDNN) が出力に無関係な無効変数に対して十分にロバストであること、無効変数を含んだまま学習した後の結合荷重の分布から変数の有効性を分析できる可能性があることを

示した．特に，入力変数のペアごとの結合荷重の分散の大きさから，無効変数のペアがほぼ特定でき，これによって無効な特徴量を判別できることを明らかにした．更に，有効だが出力と相関がない，特定の場合にのみ有効，2つの特徴量を組み合わせないと有効性が低い，といった特徴量を含む場合でも，それらを残して無効である可能性の高いものだけを削除することによって，近似精度を保ちつつ計算コストを大幅に削減できることを示した．

本論文で扱った関数近似課題は2種類のみであるが，これまでに多数の課題で同様の実験を行い，同様の結果を得ている．このことから本研究の結果はある程度一般性があると考えられるが，理論的な裏付けがあるわけではない．今後，更にさまざまな条件下で実験を繰り返すとともに，有効な入力変数に関して分散が大きくなる理由について解析を進めることが課題である．

本研究の結果は，新たな特徴選択の手法を与えると考えられる．特に部分的に有効な特徴量や2つの組み合わせによってはじめて有効になる特徴量を誤って取り除くことがないため，統計的指標を用いる従来の特徴選択手法よりも幅広い問題への適用が期待できる．また，特徴量のさまざまな組み合わせを試す方法に比べて，計算コストが圧倒的に少なく済む．ただし，本論文の方法のままでは有効だが冗長な特徴量を除外できない（例えば同じ特徴量が2重に入力されていた場合，どちらも必要と判断される）ため，まだ改良の余地がある．従来の特徴選択手法と定量的に比較することも今後の課題である．

また，特定の状況のみで意味をもつ変数や相乗効果をもつ変数を含む非線形データを解析する手法としても有望である．ただし，実験2の結果が示すように，結合荷重の分散が大きいほど重要な変数とは限らないため，例えばサンプル数や特徴量の組み合わせを変えることによって，各変数がどのように出力に貢献しているのか具体的に解析する手法を開発する必要がある．

謝辞 本研究は，JSPS 科研費 JP18H03304 の助成を受けたものである．

文 献

- [1] 新保智之, 山根 健, 田中文英, 森田昌彦, “選択的不感化ニューラルネットを用いた強化学習の価値関数近似,” 信学論 (D), vol. J93-D, no.6, pp.837–847, June 2010.
- [2] M.L. Bermingham, R. Pong-Wong, A. Spiliopoulou, C. Hayward, I. Rudan, H. Campbell, A.F. Wright, J.F. Wilson, F. Agakov, P. Navarro, and C.S. Haley, “Application of high-dimensional feature selection: evaluation for genomic prediction in man,” *Scientific Reports*, vol.5, article number 10312, 2015. DOI:10.1038/srep10312
- [3] Y. Saeys, I. Inza, and P. Larranaga, “A review of feature selection techniques in bioinformatics,” *Bioinformatics*, vol.23, no.19, pp.2507–2517, 2007.
- [4] I. Guyon and A. Elisseeff, “An introduction to variable and feature selection,” *J. Machine Learning Research*, vol.3, pp.1157–1182, 2003.
- [5] P. Ravisankar, V. Ravi, G.R. Rao, and I. Bose, “Detection of financial statement fraud and feature selection using data mining techniques,” *Decision Support Systems*, vol.50, no.2, pp.491–500, 2011. DOI:10.1016/j.dss.2010.11.006
- [6] O. Abedinia, N. Amjadi, and H. Zareipour, “A new feature selection technique for load and price forecast of electrical power systems,” *IEEE Trans. Power Systems*, vol.32, no.1, pp.62–74, 2016. DOI:10.1109/TPWRS.2016.2556620
- [7] T. Tanno, K. Horie, J. Izawa, and M. Morita, “Robustness of selective desensitization perceptron against irrelevant and partially relevant features in pattern classification,” *Proc. 24th International Conference on Neural Information Processing*, pp.520–529, 2017.
- [8] 棚橋裕輔, 多層パーセプトロンによる可読性と汎化性に優れた回帰分析に関する研究, 名古屋工業大学博士学位論文, 2009.
- [9] 森田昌彦, 村田和彦, 諸上茂光, 末光厚夫, “選択的不感化法を適用した層状ニューラルネットの情報統合能力,” 信学論 (D-II), vol. J87-D-II, no.12, pp.2242–2252, Dec. 2004.
- [10] 堀江和正, 末光厚夫, 丹野智博, 森田昌彦, “選択的不感化ニューラルネットによる表面筋電位からの手首関節角速度推定,” 信学論 (D), vol. J99-D, no.6, pp.617–629, June 2016.
- [11] T. Tanno, K. Horie, T. Kobayashi, and M. Morita, “Effect of patten coding on pattern classification neural networks,” *Int. J. Machine Learning and Computing*, vol.5, no.4, pp.339–343, 2015.
- [12] 野中和明, 田中文英, 森田昌彦, “階層型ニューラルネットの2変数関数近似能力の比較,” 信学論 (D), vol. J94-D, no.12, pp.2114–2125, Dec. 2011.
- [13] P. Auer, H. Burgsteiner, and W. Maass, “A learning rule for very simple universal approximators consisting of a single layer of perceptrons,” *Neural Netw.*, vol.21, no.5, pp.786–795, 2008.
- [14] C.C. Chang and C.J. Lin, “libSVM: A library for support vector machines,” *ACM Trans. Intelligent Systems and Technology*, vol.2, no.3, pp.1–27, 2011. DOI:10.1145/1961189.1961199
- [15] X. Glorot and Y. Bengio, “Understanding the difficulty of training deep feedforward neural networks,” *Proc. Thirteenth International Conference on Artificial Intelligence and Statistics*, pp.115–123, 2010.

cial Intelligence and Statistics, pp.249–256, 2010.

- [16] T. Kobayashi, T. Shibuya, and M. Morita, “Q-learning in continuous state-action space with noisy and redundant inputs by using a selective desensitization neural network,” J. Advanced Computational Intelligence and Intelligent Informatics, vol.19, no.6, pp.825–832, 2015. DOI:10.20965/jaciii.2015.p0825
- [17] K.D. Vohs, N.L. Mead, and R.M. Goode, “The psychological consequences of money,” Science, vol.314, no.5802, pp.1154–1156, 2006.
- [18] A.D. Kramer, J.E. Guillory, and J.T. Hancock, “Experimental evidence of massive-scale emotional contagion through social networks,” Proc. Natl. Acad. Sci., vol.111, no.24, pp.8788–8790, 2014.

(2018年11月29日受付, 2019年4月22日早期公開)



染野 翔一

平 29 筑波大学・工学システム学類卒. 平 30 現在筑波大学システム情報工学研究科博士前期課程在学中. 神経回路網による情報処理の研究に従事.



堀江 和正 (正員)

平 24 筑波大学・工学システム学類卒. 平 29 同大大学院・博士後期課程修了. 同年, 同大研究員を経て平 30 より同大計算科学研究センター助教. 機械学習による生体信号の認識, 処理に関する研究に従事.



森田 昌彦 (正員)

昭 61 東大・工・計数卒. 平 3 同大大学院博士課程修了. 日本学術振興会特別研究員, 東京大学工学部助手を経て, 平 4 筑波大学電子・情報工学系講師. 同大機能工学系助教授などを経て, 平 19 同大大学院システム情報工学研究科教授. 現在同大システム情報系教授. 脳の情報処理機構及び神経回路網による情報処理の研究に従事. 平 5 日本神経回路学会研究賞, 平 6 同学会論文賞, 平 11 日本心理学会研究奨励賞受賞.