

3 R-7

ニューラルネットによる概念の属性の重要度評価の試み

有田英一、小船隆一、豊浦潤
三菱電機(株)

1. はじめに

パターン分類をはじめとする事例の汎化能力を持つニューラルネットワーク技術を知識情報処理へ応用することを目的として、分類問題を解く階層型ニューラルネットワークを利用して、概念の分類問題における属性の重要度の評価を行なう1つの試みについて述べる。ネットワークの入力層のノードを概念の属性に、出力層のノードを分類する概念に対応させて学習を行ない、学習が終了した後、各入力ノードが各出力ノードの活性度に貢献する度合いを計算し、それを重要度の尺度とする。

2. ノードの活性度とリンクの重み

学習の終了した3層のバックプロパゲーション型ニューラルネットワークについて、学習に利用した学習パターンを入力した時の隠れユニットと隠れユニットから学習パターンに対応する出力ユニットとのリンクの重みを隠れユニットの活性度の大きさの順にソートすると、活性度の大きい隠れユニットに対応するリンクの重みは正の値をとり、活性度の小さい隠れユニットに対応するリンクの重みは負の値になっていることがわかった。図1にその例を示す。図は入力ユニット35個、隠れユニット16個、出力ユニット4個の階層型ネットワークで、入力ボタン1と出力ユニット0、入力ボタン2と出力ユニット1、入力ボタン3と出力ユニット2、入力ボタン4と出力ユニット3がそれぞれ対応するように学習させたものである。図で $o2[n]$ は隠れユニット番号 n の出力値、 $w32[i,j]$ は第 i 番目の出力ユニットと第 j 番目の隠れユニットとの間のリンクの重みを表わす。図は入力ボタン1 (Pattern 1) をネットワークに入力した時の隠れユニットの活性度と隠れ層と出力層とのリンクの重みを表わしたもので、その隠れユニットの活性度の大きさでソートすると出力ユニット0 (Output Unit No: 0) に入るリンクの重みが正のもの

負のものに分れている。入力ボタン2、3、4の場合も同様の結果であった。

3. 入力ノードの貢献度

2節の結果は、よく学習されたネットワークは効率良くリンクの重みが学習されていることを示している。このことを利用すると、ネットワークの重みの静的な解析でネットワークの挙動がある程度予測できる。即ち出力ユニット $O(i)$ が活性化されるのは $O(i)$ と正のリンクを持つ隠れユニット $H(j)$ が活性化されている場合であり、 $H(j)$ が活性化されるのは $H(j)$ と正のリンクを持つ入力ユニット $I(k)$ が活性化される場合である。そこで入力ユニット $I(k)$ の出力ユニット $O(i)$ の活性度への貢献度 (Contribution) を次のように定義する。

$$\text{Contribution}(O(i), I(k)) = \sum_{h=0}^{hun-1} \{W32(i,h) \times W21(h,k)\}$$

但し、

hun : 隠れユニットの数

$W32(i,h)$: 隠れユニット h から出力ユニット i へのリンクの重み

$W21(h,k)$: 入力ユニット k から隠れユニット h へのリンクの重み

このような入力ノードの出力ノードへの貢献度は [Saito 88][梶原90] などでも考察されている。

4. 概念の属性の重要度評価の試み

出力ノードを概念に対応させ、入力ノードを概念の属性に対応させると貢献度は分類問題における概念の属性の重要度を表わしていると考えられる。貢献度が0に近い属性は、その属性があってもなくても概念を活性化させる度合いが小さいので、概念の分類問題においてその属性の重要度は小さいと言える。貢献度が正の大きい値を持つ属性は、その属性が概念を活性化させる度合いが大きいことを表わし、逆に貢献度が負の大きい値を持つ属性は、その属性が概念を活性化させない度合いが大きいことを表わすので、貢献度の絶対値の大きい属性は概念の分類問題において、その重要度は大きいと言える。

[Winston84]の動物当ての例題を用いて簡単な実験を行なった。入力層21ノード、隠れ層16ノード、出力層7ノードの階層型ネットワークで図2に示すボタンをバックプロパゲーション法[Rumelhart86]で学習させた。図2で各行が入力ノードに対応し、各列が出力ノードに対応する。図3は貢献度を10段階に量子化したものである。図3で、右端の(MAX-MIN)の列は各行の最大値と最小値の差であり、その値が大きい属性ほど分類において重要な属性である。この例ではHas-Black-Stripe, Has-Long-Legs, Has-A-Long-Neck, Has-Dark-Spotsなどの属性が7つの動物の分類において重要な属性である。分類においては、ただ1つだけ他と違う値を持つ属性(例: 図2のSwims, Flies, Flies-Well)が重要な属性と考えられるが、図3ではHas-Dark-Spots, Has-Black-StripesのほうがSwimsなどより重要度が大きい。これはTigerとCheetahの入力ボタンが非常に似ているため、その差異を表わす属性Has-Dark-Spots, Has-Black-Stripesの重要度が高くなったと考えられる。重要度の評価においてニューラルネットワークが持つボタン全体の統計的性質を汎化する能力が生かされていると考えられる。

Pattern : 1 Output Unit No : 0 Output Unit No : 1 Output Unit No : 2 Output Unit No : 3

```

(o2i 3 i) 0.055 1. w32i 0 3 i = -1.877) w32i 1 3 i = 0.077) w32i 2 3 i = 0.168) w32i 3 3 i = 0.844)
(o2i 9 i) 0.076 1. w32i 0 9 i = -0.814) w32i 1 9 i = 1.353) w32i 2 9 i = -0.804) w32i 3 9 i = 0.067)
(o2i 14 i) 0.077 1. w32i 0 14 i = -1.161) w32i 1 14 i = 0.236) w32i 2 14 i = 1.701) w32i 3 14 i = -1.259)
(o2i 10 i) 0.081 1. w32i 0 10 i = -1.313) w32i 1 10 i = -0.639) w32i 2 10 i = 0.767) w32i 3 10 i = 0.178)
(o2i 4 i) 0.100 1. w32i 0 4 i = -0.832) w32i 1 4 i = -0.028) w32i 2 4 i = -0.775) w32i 3 4 i = 1.193)
(o2i 2 i) 0.192 1. w32i 0 2 i = -0.659) w32i 1 2 i = -0.750) w32i 2 2 i = 0.250) w32i 3 2 i = 0.670)
(o2i 13 i) 0.417 1. w32i 0 13 i = -0.373) w32i 1 13 i = -0.984) w32i 2 13 i = -0.310) w32i 3 13 i = 0.738)
(o2i 8 i) 0.791 1. w32i 0 8 i = -0.097) w32i 1 8 i = 0.596) w32i 2 8 i = 0.241) w32i 3 8 i = -1.694)
(o2i 15 i) 0.817 1. w32i 0 15 i = -0.010) w32i 1 15 i = 0.391) w32i 2 15 i = -0.945) w32i 3 15 i = -0.759)
(o2i 6 i) 0.817 1. w32i 0 6 i = 0.272) w32i 1 6 i = 0.658) w32i 2 6 i = -0.764) w32i 3 6 i = -1.057)
(o2i 1 i) 0.907 1. w32i 0 1 i = 0.430) w32i 1 1 i = -0.104) w32i 2 1 i = 0.712) w32i 3 1 i = -1.703)
(o2i 7 i) 0.924 1. w32i 0 7 i = 0.386) w32i 1 7 i = 0.341) w32i 2 7 i = -1.767) w32i 3 7 i = 0.244)
(o2i 6 i) 0.929 1. w32i 0 6 i = 0.694) w32i 1 6 i = 0.665) w32i 2 6 i = -1.498) w32i 3 6 i = -0.202)
(o2i 11 i) 0.931 1. w32i 0 11 i = 0.904) w32i 1 11 i = -1.470) w32i 2 11 i = 1.209) w32i 3 11 i = -1.178)
(o2i 0 i) 0.941 1. w32i 0 0 i = 1.366) w32i 1 0 i = -1.095) w32i 2 0 i = -0.515) w32i 3 0 i = -0.349)
(o2i 12 i) 0.960 1. w32i 0 12 i = 0.832) w32i 1 12 i = -2.496) w32i 2 12 i = -0.080) w32i 3 12 i = 0.923)
    
```

図1

	Tiger	Giraffe	Ostrich	Albatross
	Cheetah	Zebra	Penguin	
Has-Hair	○	○	○	○
Has-Feather	×	×	×	×
Gives-Milk	○	○	○	○
Flies	×	×	×	×
Does-Not-Fly	×	×	×	×
Flies-Well	×	×	×	×
Lays-Eggs	×	×	×	×
Eats-Meat	○	○	○	○
Has-Pointed-Teeth	○	○	○	○
Has-Claws	○	○	○	○
Has-Forward-Eyes	○	○	○	○
Has-Hoots	×	×	×	×
Chews-Cud	×	×	×	×
Has-Tanny-Color	○	○	○	○
Has-White-Color	×	×	×	×
Is-Black-And-White	×	×	×	×
Has-Dark-Spots	×	○	×	×
Has-Black-Stripes	○	×	○	×
Has-Long-Legs	×	×	×	×
Has-A-Long-Neck	×	×	×	×
Swims	×	×	×	×

図2

5. おわりに

バックプロパゲーション型ニューラルネットワークについて、各入力ノードの出力ノードへの貢献度を定義し、入力層のノードを概念の属性に対応させ出力層のノードを概念に対応させることにより、貢献度を分類問題における属性の重要度の尺度とする提案を行なった。

貢献度はネットワークからのルールの抽出[Saito 88]や、概念マップ[堀86][田村91]における概念間の距離の初期値の設定などに応用することができると考えられる。

参考文献

[田村91] 田村：概念マップ生成システムSPRINGS、情報処理学会第43回全国大会、P2-159 (1991)
 [堀86] 堀：単語の意味の学習について、コンピュータソフトウェア、Vol. 13, No.4, pp65-72 (1986)
 [Saito 88] K.Saito and R.Nakano: Medical Diagnostic Expert System Based on PD P Model, Proceedings of the IEEE International Conference on Neural Networks, pp1-255-1-262 (1988)
 [堀原90] 堀原：ニューラルネットワーク推論システム、情報処理学会論文誌、Vol.31, No.3, pp500-510 (1990)
 [Winston84] P.Winston and B.Horn: LISP(second edition), Addison-Wesley (1984)
 [Rumelhart86] D.Rumelhart and J.McClelland, Parallel Distributed Processing, MIT press (1986)

	Tiger	Giraffe	Ostrich	Albatross	MAX
	Cheetah	Zebra	Penguin		-MIN
Has-Hair	0	0	1	1	-3
Has-Feather	-1	-1	-2	-1	3
Gives-Milk	0	0	1	1	2
Flies	0	0	-1	0	3
Does-Not-Fly	-1	-1	-1	-1	3
Flies-Well	0	0	0	0	3
Lays-Eggs	-1	0	-1	0	2
Eats-Meat	2	1	-1	-1	3
Has-Pointed-Teeth	1	1	-1	0	2
Has-Claws	1	1	-1	0	2
Has-Forward-Eyes	1	1	-1	-1	2
Has-Hoots	-1	-1	2	2	3
Chews-Cud	-2	-1	2	3	3
Has-Tanny-Color	0	1	1	-1	2
Has-White-Color	0	-1	0	-1	4
Is-Black-And-White	0	-1	-1	-1	3
Has-Dark-Spots	-5	5	4	-2	10
Has-Black-Stripes	-4	-4	-3	-2	8
Has-Long-Legs	-1	-1	4	-2	7
Has-A-Long-Neck	-1	-1	3	-2	7
Swims	0	0	0	-2	5

図3