【土木計画学研究・講演集 No.20(1) 1997年11月】

多層パーセプトロンに基づく学習型知識獲得手法の開発

A Study on Knowledge-Aquiring Method Based on Multi-Layered Perceptrons

紀伊雅敦*,土井健司**

By Masanobu KII, Kenji DOI

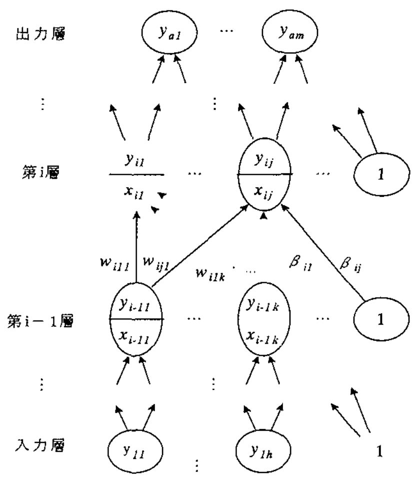
1.はじめに

ニューラルネットワークは学習といわれる漸近的なパラメータの調整により高精度な近似能力並びに自己組織的にデータの入出力特性を定めるといった特長を持つ情報処理機械である．このため入出力対のデータセットのみわかっており，それらの信号間の一般的な加工規則が不明な問題に対して有効である．ニューラルネットワークから学習により得られた情報をデータセットの内部構造として抽出，解釈できるならば，現象構造の分析上非常に有用なものとなるであろう．しかし，学習により得られた情報は重みにより表現されており，それを数式により解釈しようとすると非常に複雑となる．そのため，計画分野での多くの従来研究はその近似能力のみに着目し，ニューラルネットワークはブラックボックス的に利用されてきた．本研究では学習で得られた重みが与える入出力対の影響関係を命題論理を用いて整合的に扱うことで，その因果関係を明らかにする方法論を提示する．これにより多層パーセプトロン(MLP)といわれる階層型ニューラルネットワークの意味解釈を論理形式で行なうことが可能となる．

2.MLP の特長と問題点

本研究では，図－1のような層状に並んだ複数のユニットと層間のユニットのつながりによって構成される MLP を想定する．各層において，入力と出力の受け渡しはユニットを介して次式のように行われる．

$$x_{ij} = \sum_k w_{ijk} y_{i-1k} + \beta_{ij} \qquad (1)$$



図－1 Multi-Layered Perceptron (MLP)

$$\mathbf{y}_i = f(\mathbf{x}_i) = \frac{1}{1+\exp(-\mathbf{x}_i)} = \frac{1}{1+\exp(-\mathbf{W}_i \mathbf{y}_{i-1})} \qquad (2)$$

ただし，x_{ij}は第 i 層の j 番目ユニットの入力値であり，y_{i-1}は第 i-1 層の k 番目ユニットの出力値，w_{ijk}は第 i-1 層のユニット k と第 i 層のユニット j の結びつきの強さを表わす重み，β_{ij}は第 i-1 層の常に 1 を出力するバイアスユニットの重みである．また，$f(x)$はシグモイド関数であり，$\mathbf{x}_i$ は第 i 層の入力ベクトル(ユニット番号は省略)，$\mathbf{y}_i$, $\mathbf{y}_{i-1}$ はそれぞれ第 i 層，第 i-1 層からの出力ベクトル，$\mathbf{W}_i$は w_{ijk}とβ_{ij}を要素とする重み行列である．式(1)，(2)で表わされる出力の受け渡しと変換は，入力層から最終の出力層へと順次伝達される．MLP における学習は教師データに基づく重み行列 $\mathbf{W}$ の調整により行われる．

以上のようなMLPには，先に述べたような特長があるが，短所としては，1)汎化の問題に関わるモデル選択の問題と2)モデルの意味解釈の問題があげられる．本研究では後者の問題に着目し，それを解決する方法論を提示する．

key words:情報処理，ニューラルネットワーク

* 学生員 工修 東京工業大学土木工学専攻
** 正員 工博 東京工業大学情報環境学専攻
〒152 目黒区大岡山 2-12-1
TEL:(03)5734-2695, FAX:(03)3726-2201

3.KT アルゴリズム

重みに着目した MLP の解釈を示唆する研究はこれまでにも見られる[2)3)]. Fu[1)]は MLP の重みから知識を抽出する手法を提案し，これを KT(Knowledgetron)と名づけた．この手法は学習済みの MLP において 2 つの層ごとに上位層のユニットと下位層のユニットの関係を命題論理の含意で表わす．図－2 にそのアルゴリズムを示す．ただし，ここでは説明される上位層のユニットのもつ命題を概念 c とし，下位層のユニットのもつ命題を属性としている．ここでは，ルールに用いられる属性に対応する重みの和と，説明されるユニットの設定した閾値との大小比較により適切なルールを決定する．

KT はこれらの組み合わせを探索木により求めるが，それらを網羅的に探索すると探索空間は属性数に対し指数的に増大する．そこで，図に示す 3 つのヒューリスティックス，H1, H2, H3 を用い探索空間を減少させる．ただし，この手法では 1 つのルールに用いることのできる属性数を k としてあらかじめ定める．こうして残された組み合わせの属性をすべて論理積で結んだものが得られる含意の前件となり，上位層のユニットの持つ概念が後件となる．またそれぞれの組み合わせは別のルールを形成し，同一のルールでは表わされない．

4.MLP からの命題論理式の導出

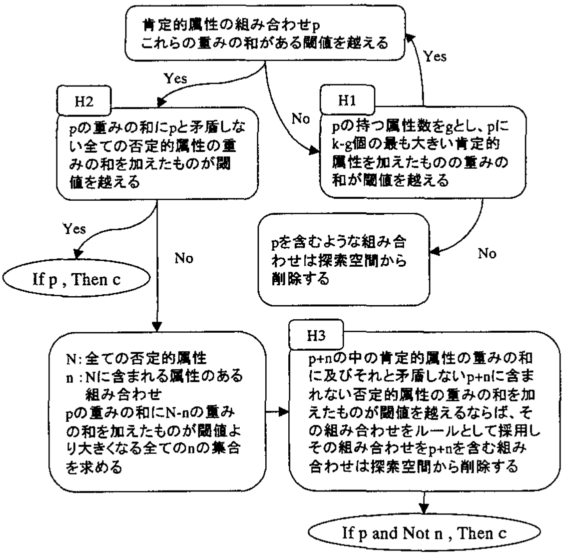
以上に示した KT は重みからプロダクションルールによる知識の導出を目的とした方法であり，得られるルールは一方向の含意にとどまり，必ずしも出力される命題の成立条件を網羅するものではない．つまり含意の前件に関する情報が網羅されない恐れがある．この問題に対し本研究では，より網羅性の高い導出手法の構築を目的とする．

以下では着目するユニットを u_i，その前階層の任意のユニットを u_j とし，u_i 及び u_j の持つ命題を p_i, p_j と表わす．

まず，バイアス項を除く重み係数 w を式(3)のように全て正となるように変換する．ただし，第 1 項の$-w_{ij}$, $1-y_j$ をそれぞれ，w_{ij}', y_j'と置き換え，第 3, 4 項をまとめて β_{ij}'とする．

$$\sum_j w_{ij} y_j = \sum_{\forall j \in (w_{ij}<0)} -w_{ij}\left(1-y_j\right) + \sum_{\forall j \notin (w_{ij}<0)} w_{ij} y_j$$

$$+\beta_{ij} + \sum_{\forall j \in (w_{ij}<0)} w_{ij} \quad (3)$$



図－2 KT アルゴリズム

ここで，変換した係数に対応するユニットの論理的な表現は元の表現の否定として扱い，意味は逆となる．これにより前階層のユニットに対する当該ユニットの単調非減少を保証する．

次にユニット u_i の出力 y_i は u_i が持つ命題 p_i の真理値とし，対象とする MLP 上では y_i は[0,1]のあらゆる値を取りうるため，次のような閾値 q を用いてその 2 値化を行なう．

$$z_i = \begin{cases} 1, & if \quad y_i = f\left(x_i\right) \geq q \\ 0, & if \quad y_i = f\left(x_i\right) < q \end{cases} \quad (4)$$

ここに，z_i は命題 p_i の 2 値化された真理値であり，y_i が閾値 q 以上の場合には p_i は十分に真であるとして z_i=1 とする．逆に y_i が q 未満のときには十分な真理値を持たないものとして z_i=0 とする．

次に，y_iと同様に出力 y_jを 2 値化された z_j で置き換え，$f(x)$の単調性より式(4)を変形すると，次式が得られる．

$$z_i = \begin{cases} 1, & if \quad \sum_j w_{ij} z_j + \beta_i \geq f^{-1}(q) \\ \\ 0, & if \quad \sum_j w_{ij} z_j + \beta_i < f^{-1}(q) \end{cases} \quad (5)$$

次に，(5)式に対応する命題論理式を考える．そのために，各ユニットに対して具体的に 2 値の真理値を当てはめ，ユニットの真理値に応じて集合 J とその補集合 $\overline{J}$ を次式のように定義する．

$$j \in J\ ,\ \forall z_j = 1 \qquad (6)$$

$$j \in \bar{J}\ ,\ \forall z_j = 0 \qquad (7)$$

式(6)は命題が真であるユニットの集合を表わし，式(7)は命題が偽であるユニットの集合を表わす．

個々の命題 p_j を 'and'で結んだものを $\wedge\ p_j$ と表わすならば，式(6)より $\bigwedge_{j\in J} p_j$ は明らかに真である．ここで，仮にユニットiに影響する前階層のユニットの集合が式(6)で表わされる集合 J であるとすれば，u_i の論理式は論理積と含意により次のように構成される．

$$\bigwedge_{j\in J} p_j \rightarrow p_i \qquad (8)$$

上式は，「$\bigwedge_{j\in J} p_j$ が真であるならば p_i も真である」という論理表現である．

次に，式(5)の条件において式(7)を満たすユニット u_j は関与しないため，これを以下のように書き直す．

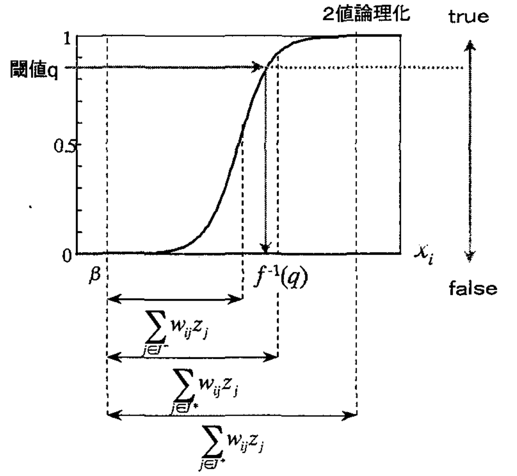
$$z_i = \begin{cases} 1, & if \quad \sum_{j\in J} w_{ij}z_j + \beta_i \geq f^{-1}(q) \qquad (9.a) \\ 0, & if \quad \sum_{j\in J} w_{ij}z_j + \beta_i < f^{-1}(q) \qquad (9.b) \end{cases}$$

ここで，$z_i=1$ となる式(9.a)の条件式は式(8)が成り立つための必要条件である．なぜなら，式(6)の下で式(9.a)が成り立つならば，必ず p_i は真となるためである．

一方，式(8)の十分条件を考えるならば，'and'の性質より，集合 J に属する任意の命題 p_j が偽の場合には p_i が偽となることが条件となる．これを示すために，式(8)を満たす集合 J の部分集合 J^- を考える．このとき，$^\forall J^-$ が式(9.b)を満たすことが式(8)の十分条件となる．この条件は次式で与えられる．

$$\sum_{j\in J^-} w_{ij}z_j + \beta_i < f^{-1}(q) \qquad (10)$$

ただし，古典命題論理では含意は前件の否定または後件の肯定により真とされるので，式(9.a)が成り立つことが式(8)が成り立つための必要十分条件となる．しかし，式(10)が成り立たないならば，$\bigwedge_{j\in J} p_j$ が偽である場合にも式(8)が成立する．これは前件が偽である含意が成立することを意味し，無関係な命題を結び付けていることになる．本研究では，そのような含意は前階層のユニットが当該ユニットの説明になっていないものとしてこれを適切な知識獲得とはみなさない．そのため，式(9.a)と式(10)の両者が同時に成り立つことが必要十分条件であるとする．このような古典論理における含意の持つ違和感を除去するための論理体系として適切さの論理などが挙げられる．なお，以下の式(11)でこの含意は同値に置き換えられるため，以降は古典命題論理として論理演算が行なわれる．



図－3　重みパラメータによる真理値の2値化

結局，集合 J が式(9.a)の条件を満たし，なおかつそのすべての部分集合 $^\forall J^- \subset J$ が式(10)を満たすならば，J は式(8)を満たすための必要十分な大きさの集合であるとし，これを J^* と表記する．

式(8)を満たす集合 J^* の群を$\{J^*_s\ ,\ s=1,\cdots,S\}$とすれば，$\{J^*_s\}$のみが式(9.a)を満たすため，第 i 層の命題 p_i が真ならば前階層の第 j 層においては$\{J^*_s\}$のうちのいずれかは真なる命題を有するユニットの集合であることが要請される．逆に，既に述べたように，ユニットの集合が J^*_s であれば p_i は真となる．したがって，以下の同値関係が導かれる．

$$p_i \leftrightarrow \bigvee_{s=1}^{S}\left(\bigwedge_{j\in J^*_s} p_j\right) \qquad (11)$$

但し，'∨'は論理和 'or'を表わす記号であり，$\bigvee_{s=1}^{S}$ は1〜Sまでの論理和を意味する．

図－3は，3つの集合 J^*, J^-, J^+ におけるそれぞれの u_i の出力 y_i を見たものである．但し，$J^* \subset J^+$ とする．このとき，J^- については $y_i<q$ より u_i において $\bigwedge_{j\in J^-} p_j$ は真とはなり得ない．次に，J^+ については $y_i \geqq q$ であるが，J^* を部分集合として含むことから必要十分な集合とは言えない．J^* は $y_i \geqq q$ であり，なおかつ他の J^*_s を含まないためその集合から得られる $\bigwedge_{j\in J^*} p_j$ は p_i を表わす論理式の一部となる．

	目的		プロセス		対応する用語	
	KT	本手法	KT	本手法	KT	本手法
共通点	学習済み MLP の重みに基づく命題論理形式による解釈を求める.		重みの和を用いてそれが閾値を越えるときにルールを採用する. 探索空間を減少させる方法として，採用されたルールを包含する組み合わせの探索は行わない.		属性 (attribute) 概念 (concept)	命題
相違点	プロダクションシステムとして用いるため前件に論理和を含まない IF－THEN 形式のルールを導出する.	MLP の解釈のみを行なうために特に得られる論理の形式は問わない．ここでは比較的意味解釈が容易であると考えられる選言標準形を導出した.	得られるルールは前件が論理積のみで結ばれた含意であるため出力の概念の完全性が保証されない.	得られる知識は前階層のユニットと当該ユニットとの同値関係で与えられるため完全性が保証される.	ルール	知識，論理式
			重みの符号により肯定的属性と否定的属性に分けて段階的にルールを導出する.	バイアス項を除く負の符号を持つ重みを変換し一括して知識を導出する.		
			探索空間を縮小するため，1ルール当たり用いることのできる属性数をkとしてあらかじめ定める.	左記のような探索空間の縮小は行わない.		
			得られる形式はプロダクションルールである.	得られる形式は選言標準形である.		

表－1　KTと本手法の比較

よって，$\bigwedge_{j\in J^*} p_j$ は u_i の持つ命題 p_i を表わす論理式とはならない．J^*は $y_i \geqq q$ であり，なおかつ他の J^*_S を含まないためその集合から得られる $\bigwedge_{j\in J^*} p_j$ は p_i を表わす一つの論理式とみなされる.

5.KT と本手法との比較

本手法と KT の共通点，相違点を表－1にまとめる．論理的な特徴として，両者とも得られる論理式の妥当性を表す健全性は保証されているが，KT では MLP に対する知識の網羅性を表す完全性を保証できない場合がある[1]．それは 1)ルールが含意で与えられているため出力層のユニットのもつ概念を表す論理式が，その概念が成立する全ての場合を表しているか否かは保証されず，2)1つのルール当たりに用いることのできる属性の数を k と定めることにより探索空間の削除を行っているが，k が入力変数の数より小さければ採用すべき全ての論理式を導出できない可能性があるためである．それに対して，本手法では知識は同値関係として与えられ，なおかつその論理式に用いることのできる命題の数に制限はないので出力ユニットのもつ命題が真である全ての場合を知識として導出する．ただし，KT において k を入力変数の数と同じ値とし，全てのルールを論理和で結ぶ解釈を行うならば，実質的に本手法と同一のルールを得る.

また，KT においては属性に対応する重みの符号によってルール導出の手続きを分けて行っているが，本手法では属性の論理上の意味を逆にすることで全ての重みの符号を正とし，一括した処理を行っている．そのためアルゴリズムは KT よりも単純化されている.

6.おわりに

本稿では KT と本研究で開発した知識獲得手法との比較を行った．両者はその目的，プロセスともほぼ同様の手法であるといえるが，幾つかの点で本手法に優れた点があることが示された.

今後の課題として，MLP を解釈するためのより緻密な論理表現方法に関する研究，ボルツマンマシン等の相互結合型ニューラルネットワークの利用可能性に関する研究が挙げられる.

[1] 論理学における健全性とは「X から A を推論可能ならば X から A を証明可能である」ことを意味し，すなわちこの推論が妥当であることを示す．完全性とは「X から A を証明可能ならば，X から A を推論可能である」ことであり，推論が網羅的であることを示す．ここでは論理式の導出方法に対応させ，得られる論理式が妥当であることを健全とし，その論理式群が網羅的であることを完全とする.

参考文献

1) L.M.Fu : Rule Generation from Neural Networks, IEEE Transactions on Systems,Man and Cybernetics, Vol. 24, No.8, pp.1114-1124, 1994.
2) 佐々木恵一，田村亨，桝谷有三，斉藤和夫：ニューラルネットワークを用いた市街化過程の基礎的分析，土木計画学研究講演集，No.18(2), pp101-104, 1995.
3) 入江文平，川人光男：多層パーセプトロンによる内部表現の獲得，電子情報通信学会論文誌，D-Ⅱ，Vol.j73-D-II, No.8, pp.1173-1178,1990.