

ニューラルネットワークを用いた 知識獲得手法の開発と土地利用分析への適用

土井 健司¹・紀伊 雅教²

¹正会員 工博 東京工業大学助教授 大学院情報理工学研究科(〒152 東京都目黒区大岡山 2-12-1)

²学生員 工修 東京工業大学 大学院理工学研究科土木工学専攻博士課程

人口や社会経済指標に示される成熟化傾向とは裏腹に、都市部での土地利用や立地現象は多様化および複雑化の一途をたどっており、その分析のための方法論研究が近年盛んに行なわれている。それらの手法の 1 つであるニューラルネットワークは多様な入出力関係を表現でき、かつ高度な非線形現象をも記述しうる強力な手法として認識されているが、その構造上の複雑さのため論理的な裏付けなしにブラックボックス的に適用されることがほとんどである。本研究ではニューラルネットワーク内部の入出力を命題論理と対応づけることで内部状態を論理式により表現し、モデルとして利用する際に問題とされていたニューラルネットワークの解釈、評価の支援方法を開発し、それを土地利用分析へと適用する。

Key Words: neural network, knowledge-acquiring, propositional calculus, land use analysis

1. はじめに

都市部における土地や不動産はその立地条件に関する個別性が強く、同じ地区や街区にあっても同一の需給圏や市場という概念を設定することは必ずしも妥当ではない。したがって、立地や土地利用に関する精緻な分析を実施する上では、従来の経済学的な接近法の範囲にとどまらず、より多様な要因を網羅できかつオペレーショナルな分析モデルを指向する工学的接近法の重要性が増している。

こうした背景の下に、近年、パターン認識的アプローチの代表的な手法の一つであるニューラルネットワークが注目を集めており、特に線形現象のみならず非線形現象に対する高精度な近似能力は従来の分析手法の枠組みを大きく変えうるものとして受けとめられている。しかしながら、従来研究では多くの場合ニューラルネットワークの識別能力や近似能力の高さのみが注目され、モデルの特徴や意味解釈に関しては十分な議論がなされていないのが実状である。

ニューラルネットワークを知識獲得手法として利用する試みは、1990 年前後からファジィ制御、知識ベースシステムの分野で見られる。これらは主として言語的な知識や命題論理を用いた記号的な知識、あるいはルールを学習結果から導くことを意図したものである。ファジィ制御においてはその代表的なものはネットワークの中間層ユニットにヒューリスティックなファジィ推論ルールを対応させ、感度分析を通じて因果

性の強いルールを導出したり、学習過程で逐次ルールを追加しながら適合性の高いルール群を選択しようとするものである^{1,2}。但し、ここでは個々の中間層ユニットを推論ルールに対応させる必要性から比較的簡単なネットワーク構造を前提としなければならない。また、相互結合型のニューラルネットワークを用いて連想記憶をモデル化しようという試みも見られる³。知識ベースシステムにおけるニューラルネットワークの利用として Fu⁴による KT(Knowledgetron) アルゴリズム、Towell and Shavlik⁵による M-of-N アルゴリズムなどがあげられる。これらは、学習済みニューラルネットワークの持つ重みパラメータを情報として、プロダクションルールを導出するものである。ただし、前者ではアルゴリズム中に多数のヒューリスティックスを抱え、計算量が膨大なものとなることが指摘されており、また得られる知識は、出力に対する入力十分性のみを保証する IF-THEN ルールにとどまり、必要十分なルールを獲得するためには別の処理機構を必要とする。後者は重みパラメータの統合によりネットワーク構造を単純化し、KTに見られた計算負荷の問題を軽減したものである。ただし、この方法は統合化の妥当性や根拠は明らかとされておらず、また、得られるルールの形式が Boolean タイプの特殊な IF-THEN 形式にとどまっている。

以上のような既往の研究に対し、本研究では入力変数やネットワークの形態等に制約を置くことのない知識獲得の考え方を示し、その方法として Fu や Towell and Shavlik らによ

る IF-THEN ルールの導出方法を命題論理式の導出方法へとより一般化し、入出力間の論理的意味付けにおいて必要十分性を満たす知識の獲得方法の開発を目的とする。

次に、この知識獲得方法を土地利用分析へと適用し、ミクロレベルでの立地を規定する諸要因の抽出を行なうとともに、要因の寄与関係を論理的に構造化する。さらに、得られた知識を用いてニューラルネットワークモデルの論理的整合性に関する事後評価を行なう。

以下では、まず既存のニューラルネットワークモデルの適用上の課題を整理し、それに対する改善方策としての本研究のアプローチを示す。

2. ニューラルネットワークモデルの適用上の課題と本研究の位置づけ

本研究では、図-1のように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階層型ネットワークモデルを想定する。ここで、入力層は入力原データを0~1に基準化し次の層へと出力するものであり、出力層は目的変数の値を出力する最終層である。これらの入力層と出力層で挟まれている層は中間層といわれる。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のように行われる。まず、当該層 i のユニットの入力は前階層 $i-1$ の出力を用いて次のように与えられる。

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

ここで、 x_{ij} は第 i 層の j 番目ユニットの入力値、 $y_{i-1,k}$ 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、また β_{ij} は第 $i-1$ 層の常に 1 を出力するユニットの重みである。ユニットへの入力値は次のようなシグモイド関数により変換され出力値を与える。

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

ここで、 \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、 $i-1$ 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、ニューラルネットワークにおける学習は教師データに基づく重み行列 \mathbf{W} の調整によって行われる。

以上のような階層型ニューラルネットワークには、高精度な近似能力に加え、現象構造の自己組織化能力といった特長があり、入出力データ間の構造が不明な問題に対しても有効な解析手法を与えることが可能である⁹⁾。一方短所とし

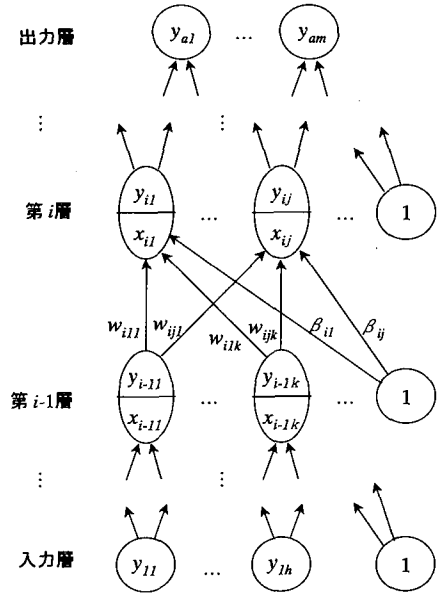


図-1 階層型ニューラルネットワーク

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。前者はモデルのサイズに関する、学習データに対する近似能力と、汎化能力といわれる未学習データに対する推定能力とのトレードオフ問題であり、離散値問題については汎化を数学的に保証する定理が示されているもの⁷⁾⁹⁾、連続値問題については適切な方法論が得られていない。すなわち、連続値をとるニューラルネットワークのパラメータ推定においてはローカルミニマムや大きい条件数などの困難のために通常是最尤推定量が得られず、モデル選択の目安となる情報量基準が得られない。このような問題に対して近年、和田、川人¹⁰⁾は3層ネットワークを対象として、最尤推定量が得られない場合にも Cross Validation を用いて情報量規準を求める方法を提案している。

一方、後者の意味解釈の問題とは、ニューラルネットワークがブラックボックス的に用いられており、学習により得られた情報に意味付けを行なうことが難しいことを指す。この問題に対して、知識ベースシステムに関わる Fu や Towell and Shavlik の研究に加え、佐々木、田村ら¹¹⁾は重みパラメータの絶対値を直接的な判断基準として因果関係の把握を試みている。この方法には明確な根拠が示されていないが、ニューラルネットワークをそれ単体で帰納法的知識獲得のツールとして活用しようとする点で新たな試みと言える。

また、入江、川人¹²⁾は、ニューラルネットワークモデルにおけるユニット数の設定がモデル構造に一定の拘束条件を与えることに対応し、ネットワーク中にデータの特徴づけるパラ

メータが得られることを実験的に示している。しかしそこでは、中間ユニット数が入力層よりも少なく、出力が入力データ自身への写像であるような砂時計モデルが用いられており、特殊な構造設定における考察にとどまっている。

本研究では以上の2種の問題のうち、まず意味解釈の問題に着目し、1) 学習モデルが獲得した知識を論理的に表現する方法論を開発することを主目的とする。次に、2) 得られた論理表現を用い、論理的整合性の観点から、ネットワーク構造の事後評価、更新を行なう方法についても提案し、モデル選択の問題についても対処を試みる。(図-2 参照)

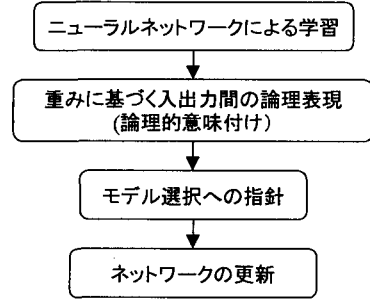


図-2 本研究のフロー

3. ニューラルネットワークを用いた知識獲得手法

(1) 関連する既往研究

ニューラルネットワークを論理モデルに利用しようという試みは、初期のニューラルネットワーク研究において既に見られる。McCulloch and Pitts¹³⁾は0, 1の値を取る出力値を命題論理における真理値と関係づけ、図-3に示すようなニューラルネットワークに基づく論理演算の表現を試みた。

$$\text{[入力]} \quad x_i = \sum_j w_{ij} y_j \quad (3)$$

$$\text{[出力]} \quad y_i = \begin{cases} 1, & x_i \geq \beta_i \\ 0, & x_i < \beta_i \end{cases} \quad (4)$$

このモデルにおいて、入力 x_i は前階層からの出力 y_j の重み w_{ij} による加重和で表わされる。ユニットの出力 y_i はこの入力 x_i と閾値 β_i の大小比較に基づき、0または1の値として決定される構造となっている。

図-3において、 y_1 と y_2 は前階層のユニットからの出力、 x_3 は当該ユニットにおける入力、 y_3 は出力を表わし、上図においては入力が閾値 $\beta = 3/2$ を超えるときには1が出力され、そうでないときには0が出力される構造となっている。同様に下図においては閾値を $\beta = 1/2$ とおき、入力 x_3 と閾値 β の大小関係から1あるいは0が出力される構造となっている。この時、出力値 y_1 、 y_2 及び y_3 を命題真理値とみなし、0を「偽」、1を「真」と判断するならば上図は論理演算における「and」、下図は「or」を表わすことになる。これらのケースでは重み w と閾値 β の値を適切に設定することにより、基本的な論理演算が表現されることを示したものであるが、より高度な演算についてもそれらを組み合わせることにより表現することができる。

(2) ニューラルネットワークからの命題論理式の導出

以上に示したMcCulloch and Pittsのモデルは所与の論理演算及び論理式を満たすようなネットワーク構造を規範的に与えているものであり、学習によってネットワーク内部にどのような表現が得られたかを知るためには推定されたパラメー

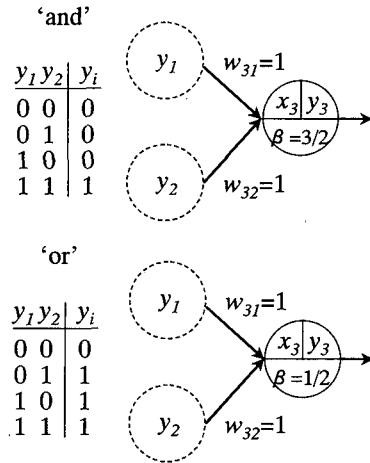


図-3 McCulloch and Pitts のモデルによる論理演算の表現

タから各ユニットの持つ命題論理式を導出することが必要となる。本研究ではこの論理式の導出をもって知識の獲得とする。以下では特に断りのない限り、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、McCulloch and Pittsのモデルと同様、ユニット u_i の出力 y_i は u_i が持つ命題 p_i の真理値であるとみなす。但し対象とする学習型ニューラルネットワーク上では 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2値化を行なう。

$$z_i = \begin{cases} 1, & \text{if } y_i = f(x_i) \geq q \\ 0, & \text{if } y_i = f(x_i) < q \end{cases} \quad (5)$$

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 を超える場合には 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 を超えない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、式(5)におけるシグモイド関数 $f(x)$ は単調増加関数であることから^[1]、式(1)を用いて次の関係が得られる。

$$z_i = \begin{cases} 1, & \text{if } \sum_j w_{ij} y_j + \beta_i \geq f^{-1}(q) \\ 0, & \text{if } \sum_j w_{ij} y_j + \beta_i < f^{-1}(q) \end{cases} \quad (6)$$

また、上式を McCulloch and Pitts のモデルに対応させるために出力 y_j を 2 値化された z_j で置き換える。

$$z_i = \begin{cases} 1, & \text{if } \sum_j w_{ij} z_j + \beta_i \geq f^{-1}(q) \\ 0, & \text{if } \sum_j w_{ij} z_j + \beta_i < f^{-1}(q) \end{cases} \quad (7)$$

次に、上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 \bar{J} を次式のように定義する。

$$\forall j \in J, z_j = 1 \quad (8)$$

$$\forall j \in \bar{J}, z_j = 0 \quad (9)$$

式(8)は命題が真であるユニットの集合を表わし、式(9)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’ で結んだ論理式(連言, conjunction) を $\bigwedge_{j \in J} p_j$ と表わすならば、式(8)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 に影響を及ぼす前階層のユニットの集合が式(8)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って次のように構成される^[2]。

$$\bigwedge_{j \in J} p_j \rightarrow p_i \quad (10)$$

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(7)の条件において式(9)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。

$$z_i = \begin{cases} 1, & \text{if } \sum_{j \in J} w_{ij} z_j + \beta_i \geq f^{-1}(q) \quad (11a) \\ 0, & \text{if } \sum_{j \in \bar{J}} w_{ij} z_j + \beta_i < f^{-1}(q) \quad (11b) \end{cases}$$

このとき、 $z_i=1$ が成り立つこと、すなわち式(11.a)の条件式は式(10)が成り立つための必要条件である。なぜなら、式(8)の下で式(11.a)が成り立つ時には、必ず p_i が真となるためである。

一方、式(10)の十分条件を考えるならば、‘and’の性質より、集合 J における任意の命題 p_j が偽の場合には p_i が偽となることが条件となる。この十分条件を示すために、式(10)を満

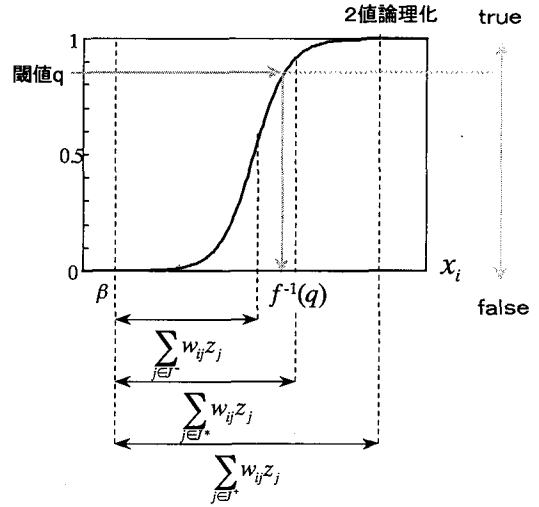


図-4 重みパラメータによる真理値の 2 値化

たす集合 J の部分集合 J' を考える。このとき、 $\forall J'$ に対して式(11)において $z_j=0$ が成り立つこと、すなわち式(11.b)の条件式は式(10)が成り立つための十分条件となり^[3]、以下のように示される。

$$\sum_{j \in \bar{J}'} w_{ij} z_j + \beta_i < f^{-1}(q) \quad (12)$$

結局、集合 J が式(11.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(12)を満たすならば、 J は式(10)を満たすための必要十分な大きさの集合であると、これを J^* と表記する。また、式(10)の含意(\rightarrow)は隣接する i, j の 2 層間の命題の一方の関係を示すものであるが、同値関係が導かれるならば任意の層のユニットの命題を入力層の命題を用いて表わすことができる。そこで、式(10)を満たす集合 J^* の群を $\{J_s^*, s=1, \dots, S\}$ とすれば、 $\{J_s^*\}$ のみが式(11.a)を満たすものであるため、第 i 層の命題 p_i が真であるならば前階層の第 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。

$$\bigvee_{s=1}^S \left(\bigwedge_{j \in J_s^*} p_j \right) \leftrightarrow p_i \quad (13)$$

但し、 \bigvee 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-4 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^+ は J^* を部分集合として含む任意の集合であり、 $J^* \subset J^+$ とする。このとき、 J^- については $y_i < q$ であることから u_i において $\bigwedge_{j \in J^-} p_j$ は真とはなり

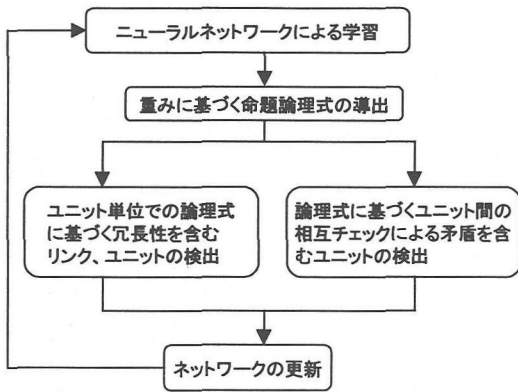


図-5 論理テストの流れ図

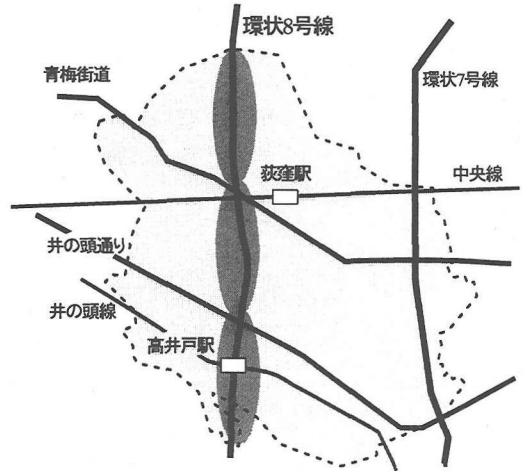


図-6 適用の対象地域

得ない。次に、 J^+ については $y_i \geq q$ 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。よって、 $j \in J^+, P_j$ は u_i の持つ命題 p_i を表わす論理式とはならない。

(3) 論理テストに基づくモデルの更新方法

和田、川人¹⁹⁾らが指摘するように、適切なモデル選択はモデルの適合性および汎化に直結する問題であり、モデルの選択は中間層のユニット数ならびにそれらを結ぶリンクの配置を意味する。ここでは、それらを定める判断基準として論理的整合性という概念を導入する。図-5は、(2)で導出した命題論理式に基づきネットワーク内に存在する冗長性及び矛盾性を検出し、ネットワーク構造を更新する手順を示したものである。ここで言う冗長とは、不必要なユニットあるいはリンクを含む状態であり、常に真または常に偽と判断される命題を持つユニット、及び論理式において影響を与えないリンクはそれぞれ冗長と見なされる。また、矛盾とはネットワーク内にあるユニットが相互に逆の影響を与えることである。それらのユニットは結果として影響を打ち消しあうため不必要なものであり、これは同一層内の異なるユニットの論理式の相互比較により検出される。このようなユニットあるいはリンクを削除することにより冗長、矛盾を減少させる。こうして更新されたネットワークを再び学習させる。

4. ニューラルネットワークの土地利用分析への適用

(1) 土地利用分析への適用の意図

局域レベルでの土地利用現象を扱う上では、土地は他の市場財に比べ個別性が高く、代替性が低いなどの特殊性を有する。また、土地は単に供給されたり消費されるだけの対象ではなく、同一の土地をめぐっても多様な利用動機や制約が存在する。したがって現象のモデル化に際しては、背後にある構造をアプリオリにあるいは規範的に仮定するより

も、まずその構造自身を抽出することが必要である。このような観点を踏まえたとき、以上に開発したニューラルネットワーク型の知識獲得システムは、土地利用分析への適用に際して次のような特長を有する。

- i) 土地利用現象を要因データ群から土地利用データへの写像であるとみなし、その関数を構造的自由度の高いニューラルネットワークにより推定させる。ニューラルネットワークはデータを記述する任意の関数の近似が可能であり、また並列処理により複数のメカニズムを同時に推定することが可能である。その際、中間層ユニットの内部関数を非線形シグモイド関数に設定することにより、要因データの組み合わせが多様な意味を持ちうることになる。したがって、先述の特殊性や多様性を有する土地利用メカニズムの抽出に有効であると考えられる。
- ii) ただし、ニューラルネットワークにより表現される構造は直接的な解釈が容易ではなく、なおかつその適切さを判断することが困難である。そこで抽出された構造を命題論理式を用いて記述することにより、論理上の適切性を有し、また、真理値という指標に基づき、信頼性の評価が可能なるモデルを構築しうる。

以下の土地利用分析への適用においては、こうした特長の検証を意図し、土地利用が安定的な一般市街地ではなく、多様な立地主体が混在し、かつ商業的な立地需要が大きく土地利用変化が顕著に見られる幹線道路沿道地域を対象とする。

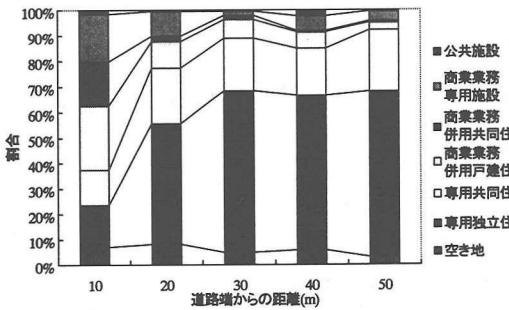
(2) 対象地域の概要

対象地域としては図-6に示す環状8号線沿道地域を取り上げ街区単位での土地利用状況をデータ化した。土地利用の用途については土地利用現況図によって把握し、建替

表-1 調査データ

対象地域	杉並区内の環状8号線沿道地域	
抽出サンプル	116街区のうち、環8に接する街区は57)	
使用データ	住宅地図	画地面積、道路端からの距離、建物階数(3階以上)
	土地利用現況図	用途、構造、建物階数(3階未満)
	都市計画図	用途指定、指定容積率
データ収集年次 : 1986, 1991, 1996		

A: 道路端からの距離帯と建物用途構成



B: 道路端からの距離帯と建物階数構成

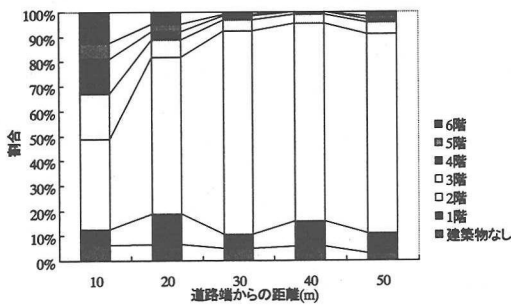


図-7 対象地域の土地利用構成

わりや土地利用変化については住宅地図から抽出した。沿道地域の範囲は、道路端から約50mとした。以下、表-1に調査データの概要をまとめる。このデータより、道路端からの距離帯別に土地利用状況を集計したものが図-7である。これより、道路端から20m帯までは、用途および高度利用の状況が距離とともに大きく変化するが、それ以遠においては目立った変化が見られないことがわかる。また、10m帯においては住居、商業、住商併用、空き地等の多様な土地利用形態が混在している。次に、幹線道路からの距離帯別の土地利用変化率を示したものが図-8である。ここでは用途の変更がなされた場合および建替えがなされた場合を土地利

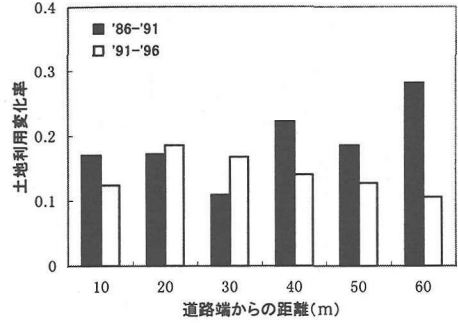


図-8 道路端からの距離と土地利用変化率

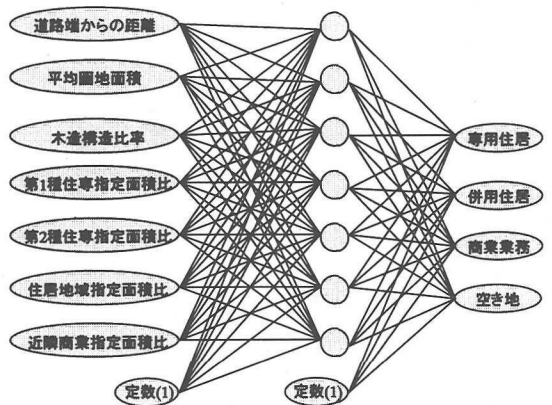


図-9 分析に使用したネットワーク

用変化と定義し、1986-91年および1991-96年の2つの期間での道路端からの距離と土地利用変化率との関係を示している。これより、1986年から91年にかけては概ね道路端から離れるほど変化率が高く、1991年から96年にかけては10m帯を除き道路端から離れるにしたがって変化率が低下するという傾向が示されている。

(3)ニューラルネットワークに基づく土地利用比率の表現

まず、ニューラルネットワークを街区単位での土地利用形態別の面積比率の算定に適用した結果を示す。分析に使用した入出力変数及びネットワーク構造は図-9に示す通りである。中間ユニット数の設定については、まず入力ユニット数と同数とし、十分な再現性が得られない場合増やしていくものとした。一般的に中間ユニット数が多いほど現況再現性は高くなるが、パラメータ数は階乗的に増加する。ただし、このようなモデルの初期構造決定に関する問題についてはここでは対象としていない。入力データとしては、画地の平均規模を表わすものとして街区平均画地面積、幹線道路への近接性を表わすものとして道路端からの距離、建物構造を表

わすものとして木造構造比率、用途指定の状況を表わすものとして用途指定面積比率(第1種住専, 第2種住専, 住居地域, 近隣商業地域の地域別)を用い、教師データは各々の街区における 1)専用住居, 2)住商併用施設, 3)商業業務施設, 及び 4)空き地のそれぞれの土地利用形態毎の面積比率を同時に用いて学習させた。サンプル数は街区数で116である。また、中間ユニット数は入力数と同じ7とした。なお、学習回数はすべての計算で30万回とした。

図-10は、1996年のデータを用いて得られたモデルの現況再現性を示している。また、図中には、同じ環8沿道地域から抽出した20の未学習データ(学習させていない街区データ)に対する空間移転性を併記している。未学習データについては、対象地域の全サンプル街区データをランダムに並べ、その中から5つおきに抽出した。本モデルは4種類の土地利用形態(用途)の面積比率を同時に学習させたにも関わらず、それぞれの用途別に見ても良好な現況再現性が得られており、専用住居については実測値と推定値の相関係数で0.933と高い適合性が得られている。また、未学習データに対する移転性についてみると、住居系の用途については比較的高い相関係数が得られているのに対して、商業・業務施設および空き地に関しては十分な相関係数は得られていない。過渡的な土地利用形態である空き地についてはそもそも十分な移転性は期待できないが、商業・業務施設については、現況比率が0~0.2の範囲に偏りサンプルが十分なバラツキをもたないことに起因して、推定されたモデルの汎化能力が制約されたものと考えられる。

(4)ニューラルネットワークに基づく土地利用変化の表現

次に、街区単位の土地利用変化の状況をニューラルネットワークモデルを用いて推定したものを図-11に示す。ここでは、用途の変更に加え、同じ用途であっても建て替わりが生じた場合を土地利用変化と定義する。モデルは1991-96年の土地利用変化率を、平均画地面積、道路端からの距離、平均容積率、木造構造比率、空き地比率、住居系用途比率を入力として学習させた。図中には学習データに対するモデルの再現性を示し、併せて未学習データに対する空間移転性を示している。これより、モデルの学習データに対する再現性は高いものの、未学習データに対する移転性は極めて低いことがわかる。このことから、土地利用の変化率を対象とした当該モデルの汎化能力は非常に低いと言わざるを得ない。

なお、時点間の土地利用変化の因果関係を考慮していないことが当該モデルの汎化能力の低さの原因となっている可能性が考えられる。既往のニューラルネットワーク研究においては、前時点の結果をフィードバック学習に取り込む方法論¹⁹⁾も提案されているが、本研究で扱った2時点間の建替え率の相関は-0.163と低かったことから、こうした動学的学習方法については試みていない。

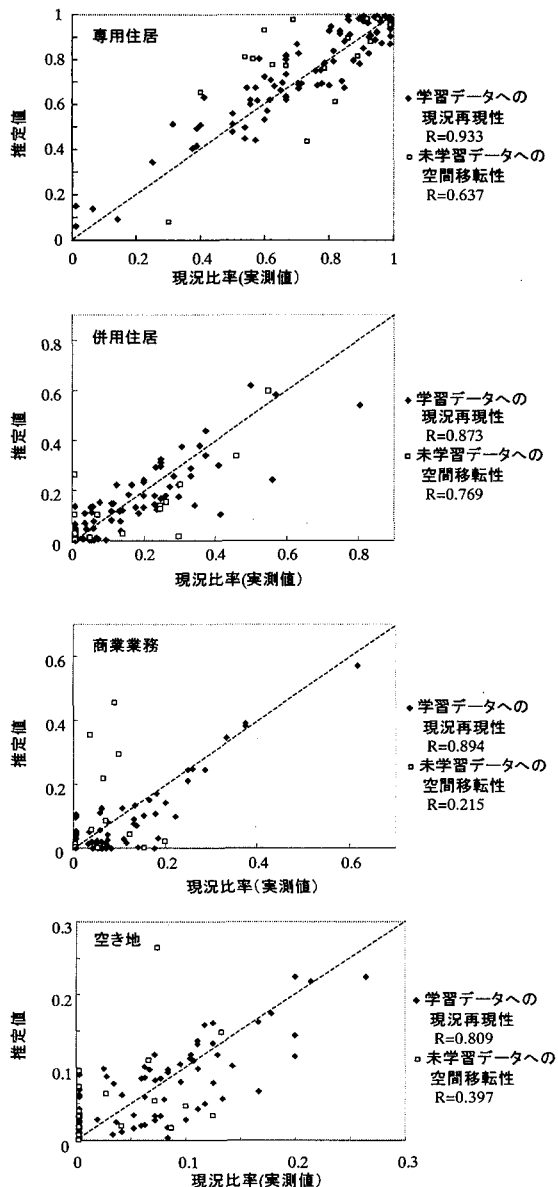


図-10 土地利用比率モデルの現況再現性と空間移転性

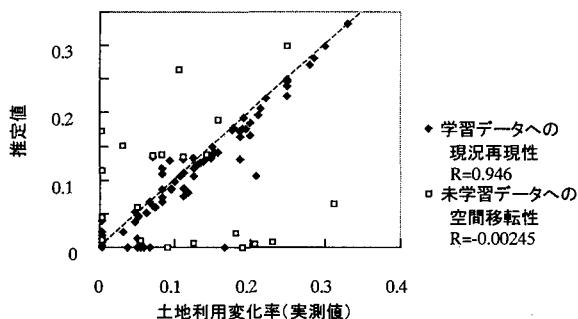


図-11 土地利用変化モデルの現況再現性と空間移転性

5. 論理表現に基づく知識の獲得とモデルの更新

表-2 土地利用比率への影響要因の寄与

(1) 獲得された知識—モデルの解釈—

土地利用比率モデルにおける入力変数の影響関係をまとめたものが表-2である。ここでは、3.(2)に示した知識獲得方法により式(13)を導出し、その右辺を構成する各々の連言 $\bigwedge_{j \in J} P_j$ を表中の説明要因の作用和として表現している。有効セットとして示した表の各列は、 $j \in J_j^*$ に対応する。また、各要因の立地への作用の方向については正負の符号により区別しており、その要因が各用途の立地に、正の影響を与える場合には「+」、負の影響を与える場合には「-」、設定した閾値水準では影響を与えない場合には「 \square 」としている。なお、ここで用いる有効セットという名称は、全ての J のうち式(11. a)の条件を満たし、なおかつ $\forall J \subset J$ が式(12)の条件を満たす、必要十分な集合 J^* を意味する。有効セット数の多さは現象メカニズムの多様性を示しており、専用住居や併用住居の立地は商業業務施設の立地や空き地としての利用に比べて、より多くの構造的特徴を持つことを意味している。住居系の立地は比率的にも大きく、かつ沿道から後背地までの広い範囲に分布しているという状況(図-7 参照)を見れば、その立地メカニズムがより多様であることは十分納得されよう。

なお、閾値の設定は論理表現の導出結果に影響を及ぼす。本研究では、数値シミュレーションにおいて安定した挙動を示す範囲の閾値レベルを採用した(次節参照)。

表-2 の結果から当該地域における土地利用のメカニズムとして以下のような知見が得られている。

- 1) 専用住居については、各セットを横並びで見ると、概ね道路端から離れているか、商業的用途の指定比率が低い街区ほど、その現況比率が高いことがわかる。ただし、有効セット No. 1~4 のように道路端からの距離が大きく、商業的用途指定比率が低いことが専用住居比率の高さに寄与しているケースと、No. 5, 6 のように道路端からの距離の影響は受けず、平均画地面積が小さく、住居地域指定比率が高いことが、寄与しているケースも存在することが分かる。後者のケースは、住居地域指定のなされた沿道直近地域における専用住居のメカニズムを表わしているものと考えられる。また、No. 7 のように明らかにこれらとは異なるケースも存在する。
- 2) 併用住居については幹線道路に近く、画地面積が小さく、さらに1種住専、2種住専地域の指定比率が低い(有効セット No. 1, 2)か、もしくは近隣商業指定比率が高い(有効セット No. 3, 4)ことが、併用住居比率の高さに寄与していることが読み取れる。No. 5 のように、これらと異なり、道路端からの距離が寄与しないケースも存在する。
- 3) 商業業務施設は道路端に近く、画地面積が小さく、近隣商業指定が低いほど、その比率が高いことが読み取れる。ただし、これ以外の要因については、有効セット No. 1, 2 の

専用住居 閾値レベル0.8	有効セット						
	1	2	3	4	5	6	7
道路端からの距離	+	+	+	+	\square	\square	\square
平均画地面積	+	\square	\square	\square	-	-	+
木造構造比率	\square	+	\square	\square	+	\square	\square
第1種住専指定比率	\square	\square	+	\square	\square	+	\square
第2種住専指定比率	\square	\square	\square	\square	-	-	+
住居地域指定比率	\square	\square	\square	-	+	+	\square
近隣商業指定比率	-	-	-	-	-	-	-

併用住居 閾値レベル0.8	有効セット				
	1	2	3	4	5
道路端からの距離	-	-	-	-	\square
平均画地面積	-	-	-	-	-
木造構造比率	+	-	\square	\square	-
第1種住専指定比率	-	-	\square	-	-
第2種住専指定比率	-	-	\square	+	\square
住居地域指定比率	-	\square	-	\square	+
近隣商業指定比率	-	\square	+	+	+

商業業務 閾値レベル0.8	有効セット		
	1	2	3
道路端からの距離	-	-	-
平均画地面積	-	-	-
木造構造比率	+	+	-
第1種住専指定比率	\square	\square	-
第2種住専指定比率	+	+	\square
住居地域指定比率	\square	-	-
近隣商業指定比率	-	-	-

空き地 閾値レベル0.8	有効セット
	1
道路端からの距離	+
平均画地面積	-
木造構造比率	-
第1種住専指定比率	-
第2種住専指定比率	-
住居地域指定比率	+
近隣商業指定比率	-

- のように1種住専指定比率の影響は受けず、木造構造比率が高いことが、商業業務施設の現況比率の高さに寄与しているケースと、No. 3 のように木造構造比率が低く、1種住専指定比率が低いことが寄与しているケースが存在する。
- 4) 空き地については、単一の有効セットのみが得られており、道路端からの距離が離れており、平均画地面積が小さく、また、木造構造比率、1種住専、2種住専指定比率、近隣商業指定比率のいずれも低く、なおかつ住居地域の指定比率が高いほど、空き地としての土地利用比率が高いことが示されている。

表-3 土地利用変化率への影響要因の寄与

1986年-91年		有効セット		
閾値レベル0.8		1	2	3
道路端からの距離	+	+	■	■
平均画地面積	-	-	-	-
平均容積率	+	■	■	■
木造構造比率		-	+	
空き地比率	+	+	+	
住居系用途比率		-	-	

1991年-96年		有効セット					
閾値レベル0.8		1	2	3	4	5	6
道路端からの距離	+	-	-	-	-	■	■
平均画地面積	+	+	+	-	-	+	+
平均容積率	+	-	-	+	+	■	■
木造構造比率	+	+	+	+	■	■	■
空き地比率		-	+	+	+	+	+
住居系用途比率	+	+	■	+	-	+	+

次に、土地利用変化モデルにおける入力変数の影響関係を示したものが表-3である。ここでは、1986-91年と1991-96年の2つの期間ごとの学習結果を示している。1986-91年では、平均画地面積が小さく、空き地比率の高い街区において土地利用変化率が大きいことがわかる。また、これ以外の要因についても概ね道路端からの距離が大きく、住居系用途比率が低いほど変化率が大きいことが読み取れる。一方、1991-96年においては、各要因の影響関係は多様化し明確な傾向は読みとれない。ただし1986-91年との対比で見れば、道路端からの距離が小さいほど、平均画地面積が大きいほど、また住居系用途比率が高いほど土地利用変化率が高いケースが目立ち、土地利用変化のメカニズムが時点間で大きく変容していることが読み取れる。

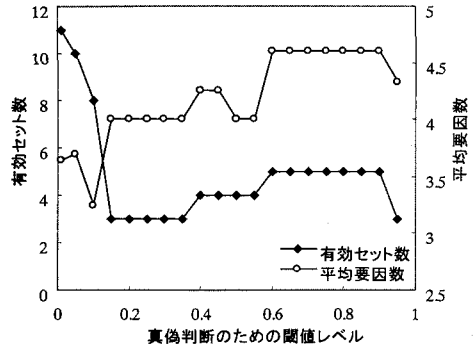
また、平均画地面積と平均指定容積率との間には相反する影響が見られ、平均画地面積が小さい場合には平均指定容積率が高い場合に変化率が大きいことがわかる。これらの結果のうち道路端からの距離の影響に関しては、図-8に示した現状ともほぼ一致している。

以上ではニューラルネットワークにおける各要因個別の影響関係を示したに過ぎないが、意味解釈上それら要因相互の影響関係についても情報を得ることが望ましい。本稿では、これについて十分な検討を行なってはいないが、入力変数間に強い代替関係が存在する場合とそうでない場合とでは、導かれる論理表現の結果が異なるなどの知見を得ている(付録参照)。

(2) 閾値設定の影響

前節に示した論理式の導出において、閾値の設定が及ぼす影響をみるために、閾値レベルに応じて導かれる有効セット数と1セット当たりの平均影響要因数を算定した。図-12

1986-1991年



1991-1996年

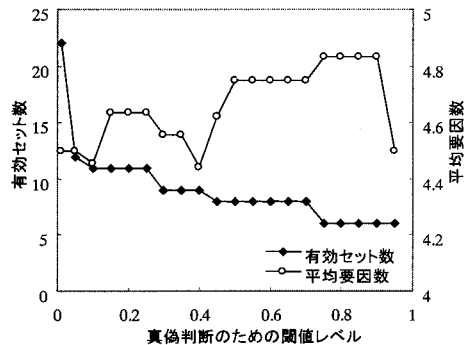


図-12 土地利用変化モデルにおける閾値設定の影響

は土地利用変化モデルを対象として0.05間隔の閾値レベルに変化に対するセット数および平均要因数の変化を示したものである。この図より、1986-91年においては閾値レベルが0.6~0.9において両者はともに安定した値を得ている。一方、1991-96年においては、0.5~0.7の範囲と0.75~0.9の範囲において両者は比較的安定している。閾値は、ユニットの出力がそれを越えるときに十分真であると判断する指標である。設定する閾値レベルが低いと影響の小さな要因の組み合わせも真であるとみなされ、有効なセットであると判断される可能性がある。また、逆にこれが高すぎると、十分真とみなされるためには影響の小さい要因をも用いる必要があり、そうした要因の混入のため閾値の変化に対するセット数、平均要因数の変化が不安定になる。よってここでは、より高いレベルでの安定域の閾値を採用することとし、前節の土地利用変化モデルにおける命題論理表現の導出に際しては、1986-91年および91-96年の両期とも閾値レベルを0.8と設定した。

(3) 論理的整合性の検討とモデルの更新

ここでは、まず、4.(3)において汎化能力の低さが指摘された土地利用変化モデルを対象として、設定したネットワーク構造の論理的整合性の再検討を行う。1991-96年を対象

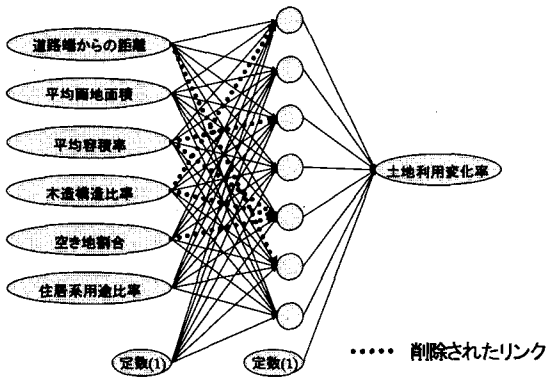


図-13 ネットワークの更新

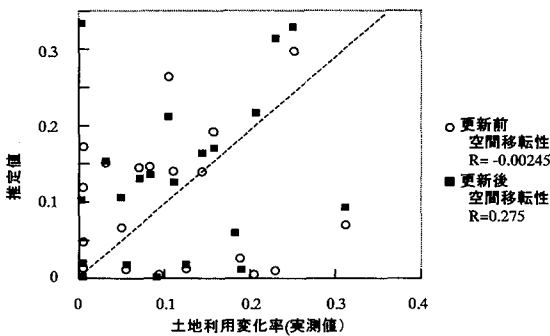


図-14 更新されたネットワークの再現性と空間移転性

とした土地利用モデルにおいて図-5 の手順でテストを実施した結果、リンクに関しては図-13 に示すように 49 リンク中 6 つが冗長であると判断された。ただしユニットに関する冗長性と矛盾性は共に検出されていない。冗長性及び矛盾性を持つユニットが検出されなかった理由としては、ネットワークの初期設定において中間層ユニット数を入力値数と同等数に限定していたためと考えられる。

次に、不整合リンクが削除された更新後ネットワークに対して、元のネットワークと同じ教師データに基づく学習を行わせたモデルの未学習データに対する移転性(相関係数)を、元のネットワークのそれと共に図-14 に示す。なお、学習データに対する再現能力は、更新前のネットワークが相関係数で 0.946、更新後のネットワークが 0.942 であり両者の間にほとんど差が見られない。一方、未学習データに対する空間移転性については、更新前のモデルが相関係数で -0.00245 と全く移転性がないのに対し、更新後には 0.275 へと若干向上していることがわかる。

モデル構造を単純化したにも関わらず、現況再現性が維持され、かつ未学習データへの移転性を若干ながら向上させたことは、以上に示した論理的整合性の検討方法の有効性を示す一つの証左と言えよう。

6. おわりに

従来、ニューラルネットワークモデルでは、内部の入出力関係を代数的に記述すると非常に複雑となるため、数式の上で、個々のパラメータやモデル全体のもつ意味を、現象と対応づけて捉えることが困難とされてきた。これに対しファジィ制御や知識ベースの分野では、学習によりルールを得るための、内部関数やネットワークの設定方法や、出力に対するパラメータの相対的な寄与関係をルールで表現する方法論等が研究されてきた。ただし、前者ではネットワークの構造を制約する必要があり、後者では集約された知識を得るには別の処理機構が必要とされる。一方、手法の汎用性の観点から言えば、意味解釈の際にモデル構造には手を加えず、また、得られる知識は適切な形にまとめられていることが望ましい。このため本研究ではネットワーク構造に制約を置くことなく、入出力の同値関係を表わす命題論理式という形で知識を導出する方法論を提示した。

また、本手法の土地利用分析への適用を通して、次のような知見が得られた。

- ①沿道地域における土地利用は、個々の用途においても複数の構造的特徴を有しており、単純なメカニズムによっては表現できないことを確認し、本研究ではそれらを互いに独立な有効セット群として得た。これはニューラルネットワーク自身の並列処理機能にもよるが、本研究ではさらに各有効セットに対して構造解釈上の意味付けを加えた。
- ②有効セット数の多さに示されるように、専用住居や併用住居の立地は商業業務施設の立地や空き地としての利用に比べて、より多様な構造的特徴を持つことが捉えられた。
- ③土地利用変化のメカニズムについて、1986年～91年というバブル期と91年～96年の沈静期では差違が見られ、特に道路端からの距離および画地面積の影響に大きな変化が見られる。

ただし、方法論の構築を主目的とする本稿においては、土地利用分析への適用に際しては、比較的容易に入手しうる土地および建物属性のみのデータセットでの分析にとどまっており、それゆえ抽出される知識も限定的である。現実の土地利用を記述する上では、立地者や地権者の意思決定に関する考察が重要となることは言うまでもない。したがって、これら主体の個人属性や、相隣関係、地域属性等を考慮した上での知識の抽出が必要となる。さらに方法論上の課題としては、論理式導出における真偽判断のための閾値レベルの設定についての理論的考察、無限値論理などを用いたより緻密な論理表現方法の開発が必要とされる。また、本研究で示した論理的整合性という判断基準に基づくモデル選択とモデルの汎化との関連性については、まだ知見が不足しており今後の課題としたい。加えて、本研究で示した論理的整合性に基づくネットワークの更新方法については、その手

続きは手作業に依存したものであり、手続きの自動化による、より効率的な更新方法が必要とされる。

付録：入力変数間の影響関係

入力変数間の相互関係が、獲得される論理式に及ぼす影響を捉えるために、ここでは次式のように 2 つの変数からなるCES関数を取りあげ、その弾力性パラメータを変化させた場合の論理表現の違いを考察する。但し $a=0.5$ とした。

$$f = \left(ax_1^{-\rho} + (1-a)x_2^{-\rho} \right)^{-\frac{1}{\rho}} \quad (14)$$

表-4 は式(14)における弾力性パラメータを $\rho = -1, 0, 1$ とした 3 つの場合に、ニューラルネットワークの学習結果から導出されるそれぞれの論理表現の結果を示している。 $\rho = -1$ のときは式(14)は線形関数となり、入力変数が相互に完全代替することを表わす。 $\rho = 0$ のときは代替弾力性が 1 の場合、 $\rho = 1$ のときは代替弾力性が 0.5 である場合に相当する。分析に用いるニューラルネットワークは図-15 に示すように中間層に 3 つのユニットを持つ階層構造である。2 つの入力変数を一様乱数により与え、それらを式(14)に代入して得られる f の値を教師データとして用いた。学習により得られたニューラルネットワークの学習データに対する適合性は相関係数で 0.995 である。

表-4 より $\rho = -1$ のときには閾値の広い設定範囲において $p(x_1)$ or $p(x_2)$ が導かれているのに対し、他の 2 ケースではそうした論理和の関係は得られていない。このことは一定水準の出力値を得るためには入力変数 x_1, x_2 のいずれかが高い値をとれば良いことを意味しており、 x_1 と x_2 の強い代替的關係を反映した結果と解釈される。 $\rho = 0$ 及び $\rho = 1$ のケースでは、閾値の設定が 0.1~0.5 程度までは $p(x_1)$ and $p(x_2)$ or $(\neg p(x_1) \text{ and } \neg p(x_2))$ という結果が導かれているが、より高い閾値レベルでは $p(x_1)$ and $p(x_2)$ という関係が導かれている。これらのことは一定水準の出力値を得る上では x_1 と x_2 が共にある一定以上の値を取る必要があることを意味しており、2 つの命題間の論理積の関係を表わすと解釈される。なお、いずれの ρ のケースにおいても閾値の高い水準では 'false' が得られており、高すぎる閾値の設定では適切な論理式が得られないことを意味している。

以上のように、適切な閾値レベルを設定することにより、入力変数間の代替・補完関係は命題論理式に反映されることが示される。

補注

[1] 重み係数 w を式(6)に従って、すべて正となるように変換する。ただし、第 1 項目の $-w_{ij}$ 、 $1-y_j$ をそれぞれ、 w_{ij}' 、 y_j' と置き換え、第 3 項目、第 4 項目をまとめて β_{ij}' とする。

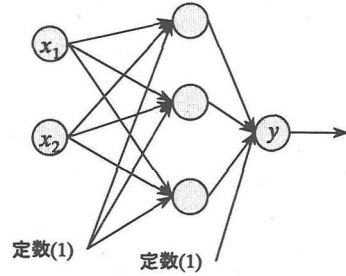


図-15 使用ネットワーク

表-4 CES関数の論理表現

	$\rho = -1$	$\rho = 0$	$\rho = 1$
0.1	true	false	true
0.2	$p(x_1)$ or $p(x_2)$	$p(x_1)$ and $p(x_2)$	$p(x_1)$ and $p(x_2)$
0.3		or	or
0.4		$(\neg p(x_1) \text{ and } \neg p(x_2))$	$(\neg p(x_1) \text{ and } \neg p(x_2))$
0.5			$p(x_1)$ and $p(x_2)$
0.6		$p(x_1)$ and $p(x_2)$	
0.7	$p(x_1)$ and $p(x_2)$		
0.8			

* $p(x_1), p(x_2)$ はそれぞれ x_1, x_2 の値をとるユニットの命題を表わす。

** '¬' は命題の否定を表わす論理記号

$$\sum_j w_{ij} y_j = \sum_{j \in \{w_{ij} < 0\}} \{-w_{ij}(1-y_j)\} + \sum_{j \in \{w_{ij} > 0\}} w_{ij} y_j + \beta_{ij} + \sum_{j \in \{w_{ij} < 0\}} w_{ij} \quad (1')$$

これにより入力に対する出力の単調非減少を保証する。ただし、その際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。そうすることにより前階層のユニットに対する当該ユニットの単調増加を保証する。

[2] KTにおけるIF-THENルールは式(10)の形で与えられ、出力命題に対する入力命題の十分性のみを示すものである。これに対し本研究で導出する式(13)は入出力間の必要十分性を示すものである。

[3]古典命題論理において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。したがって、式(11.a)が成り立つか、あるいは式(12)が成り立つことが式(10)が成り立つための必要十分条件となる。しかし、ここでは前件が偽となるときに後件が真となることは無関係な命題を結び付けることになるため、それらの因果関係に違和感が存在するということが指摘されてきた¹⁾。本研究の場合、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(11.a)と式(12)の両者が同時に成り立つことが必要十分条件であるとしている。

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が、本研究ではそのような論理体系も視野に入れている。なお、式(13)においてこの含意は同値に置き換えられているため、以降は古典命題論理として論理演算が行なわれる。

参考文献

- 1) 増田達也, 太田俊彦: ファジィ・ニューラルネットワークによるメンバーシップ関数および推論ルールの獲得手法, 電気学会論文誌 C, vol.114, No.11, pp.1185-1193, 1994.
- 2) 赤羽旗一: 記号処理型のニューラルネットワークモデルと集合で表わした知識表現, 人工知能学会誌, vol.11 No.4, pp.566-573, 1996.
- 3) 豊田秀樹: 非線形多変量解析, 朝倉書店, 1996.
- 4) Li Min Fu: Rule Generation from Neural Networks, *IEEE Transactions on Systems, Man, and Cybernetics*, Vol.24, No.8, pp.1114-1124, 1994.

- 5) Towell, G. and Shavlik, J.: The extraction of refined rules from knowledge based neural networks, *Machine Learning*, Vol.131, pp.71-101, 1993.
- 6) 清水英範, 宮城俊彦, 森光正: ニューラルネットワークの空間相互作用モデルへの適用可能性, 土木計画学研究講演集, No.16(1), pp.343-348, 1993.
- 7) Baum, E. B. and Haussler, D.: What size net gives valid generalization?, *Neural Computation*, 1, pp. 151-160, 1989.
- 8) Yaser S. Abu-Mostafa: The Vapnik-Chervonenkis Dimension: Information versus Complexity in Learning, *Neural Computation*, 1, pp.312-317, 1989.
- 9) 新井政彦: 多層パーセプトロンの中間素子数について, 人工知能学会誌, vol.8, No.2, pp.230-234, 1993.
- 10) 和田安弘, 川人光男: 新しい情報量規準と Cross Validation による汎化能力の推定, 電子情報通信学会論文誌, (D-2) J74-D-2, No.7, pp.955-965, 1991.
- 11) 佐々木恵一, 田村亨, 榎谷有三, 斉藤和夫: ニューラルネットワークを用いた市街化課程の基礎的解析, 土木計画学研究講演集, No.18(2), pp.101-104, 1995.
- 12) 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.
- 13) McCulloch, W.S. and Pitts, W.: A logical calculus of the ideas immanent in nervous activity, *Bullet. Math. Biophysics*, 5, pp.115-133, 1943.
- 14) 鈴木宏典, 川口雄一郎, 福田敦: 土地区画整理地区における一般化市街化曲線, 土木計画学研究講演集, No.18(2), pp.97-100, 1995.
- 15) 吉浦紀晃: 知識ベースのための適切さの論理に関する研究, 東京工業大学博士論文, 1996.

(1997.5.29 受付)

A STUDY ON KNOWLEDGE-AQUIRING METHOD BASED ON NEURAL NETWORK AND ITS APPLICATION TO LAND USE ANALYSIS

Kenji DOI and Masanobu KII

Roadside land use is one of the major factors which influence a long-term urban form and environmental quality. This study aims to develop a practical method for analyzing roadside land use pattern and its change based on a neural network model. Neural network model is regarded as a powerful method to describe complicated phenomena with a high non-linearity, however, it is often utilized as a black box, without logical interpretation. This study tries to overcome the crucial shortcoming existing in neural network by adding a detecting function of controversial and redundant description and to enlarge the applicability of neural network to land use analysis.