

A I による不動産分析

— A I の説明可能性問題に向けての視点 —

小長谷 一 之
今 井 隆 志
清 水 千 弘

『都市経営研究』第1巻 創刊号 2021年3月
大阪市立大学 都市経営研究科
都市経営研究会

大阪市立大『都市経営研究』第1巻 創刊号（通巻1号） 2021年3月

■ 研究ノート ■

73頁～84頁

AIによる不動産分析

— AIの説明可能性問題に向けての視点 —

小長谷一之（大阪市立大学大学院都市経営研究科教授）・今井隆志（大阪市立大学大学院都市経営研究科非常勤講師、エナリスみらい研究所プレジデント）・清水千弘（東京大学空間情報科学研究センター特任教授）

The Analysis of Artificial Intelligence Method on Real Estate Data: Toward the 'Explainability Problem'

Kazuyuki KONAGAYA (Professor, Graduate School of Urban Management, Osaka City University), Takashi IMAI (Adjunct Lecturer, Graduate School of Urban Management, Osaka City University, President of ENERES Future Institute) and Chihiro SHIMIZU (Specially Appointed Professor, CSIS, The University of Tokyo)

【要旨】

不動産情報、特に不動産価格形成は、複雑なメカニズムであると想定されている。本小論では、通常の不動産価格のデータについてパラメトリックな方法とディープラーニングの手法の比較、そして、今のAIの最大の課題の説明可能性問題についての作業をのべることを目的とする。データは、できるだけ性質の良い典型的な住宅地的地域として、神奈川県横浜市の隣接する住宅地域である、港北区、神奈川区、保土ヶ谷区の3区をとりあげ、住宅系の代表点の公示地価を説明するモデルを構築した。AI手法はブラックボックス問題といい、要因の根拠が直接明確でないという大きな課題があり結果の見える化をおこなった。結果として、中間層において劇的にデータの要約をおこなっていること、横浜市の住居系の不動産の公示地価では、説明変数の影響力は、第1変数（距離変数）の説明力が大きく、これに前面道路方位がサブの要因で考慮されており、第3変数の用途地域指定はほとんど影響がないこと、社会科学等の性質の良いデータでは、概略的には「ReLU関数をつかったニューラルネットワークの判断＝（発火パターン関数）＋（線形関数の貼り合わせ）と把握できる」場合が多いと予想できること、等がわかった。

【キーワード】

不動産情報、AI（人工知能）、予測可能性、説明可能性、発火パターン関数

【Abstract】

Real estate information especially pricing process, are thought to be complex mechanism. Nonparametric methods in general compared to ordinary parametric (written down in some sort of formula) methods has been remarkably developed using AI (artificial intelligence) methods based on deep learning. Official land prices in housing areas with good environment in Kohoku ward, Kanagawa ward and Hodogaya ward of Yokohama City are analyzed. this study try to improve AI models applied to Japanese real estate data, and make some insight to black box problem for the establishment of explanatory AI systems. The results are as follows: 1) Information are drastically summarized in

hidden layers, 2) Distance to nearest-neighbor station variables are most important factor to influence official land prices with secondary factor of front road orientation, 3) The results to apply Neural network with ReLU function to social science data are supposed to be interpreted as 'Nerve firing pattern + linear relations'.

【Keywords】

Real estate information, AI (artificial intelligence), Prediction Performance, Explanatory Performance, Nerve Firing Pattern Function

I. 問題の設定

本小論では、通常の不動産価格のデータについてパラメトリックな方法とディープラーニングの手法の比較、そして、今のAIの最大の課題の説明可能問題についての作業をのべることを目的とする。

不動産情報、特に不動産価格形成は、複雑なメカニズムであると想定されている。通常の計量経済的分析で行われてきた多重回帰式近似（パラメトリック推計）と異なり、単純な式であらわすことのできない複雑な関係を広義の意味でノンパラメトリック推計と総称するが、近年、深層学習（ディープラーニング）手法の発展によりAIが普及すると関連のミドルウェアが使えることがわかった。筆者の一人がBoston Housing Dataで、村上他編著（2020）でおこなわれた結果は、通常手法に比べ劇的に予測能力が改善しており、今後、複雑な関係を正確にとらえる能力をもったAI手法の不動産研究への応用が期待される。本研究ではCSISの日本での不動産データをもちい、AI手法の改良を試みる。特に、AI手法はブラックボックス問題といい、要因の根拠が直接明確でないという大きな課題があり、この点の改善も試みる。

II. データ

1. 良好な環境をもつ住宅地系の地価

データとしては、バブル直後の地価の状況を見るために、「国土交通省地価調査（首都圏H13）」を用いる。このデータは、基準地番号、調査年、所在、施設又は隣接名称、地価（ $\text{¥}/\text{m}^2$ ）（ $\text{¥}/10\text{a林}$ ）、地積（ m^2 ）、形状、利用区分構造、前面道路（側面）、給排水（ガス 水道 下水）、最寄駅、距離（m）、法規制、建ぺい率_容積率（%）、区域区分、利用現況、X座標、Y座標、KCODEなどからなっている。

できるだけ性質の良い典型的な住宅地的地域として、神奈川県横浜市の隣接する住宅地域である、港北区、神奈川区、保土ヶ谷区の3区をとりあげ、住宅系の代表点46ケースをとる。



図1：区の所在、横浜市ホームページによる

2. 変数の要約

- 1) 目的変数としては、「公示地価」。
- 2) 説明変数は、「最寄り駅までの距離」、以下のように集約した「前面道路方位」、「用途地域指定」をインデックス化したものとする（良好な住宅地として有利と思われるものに対し得点）。

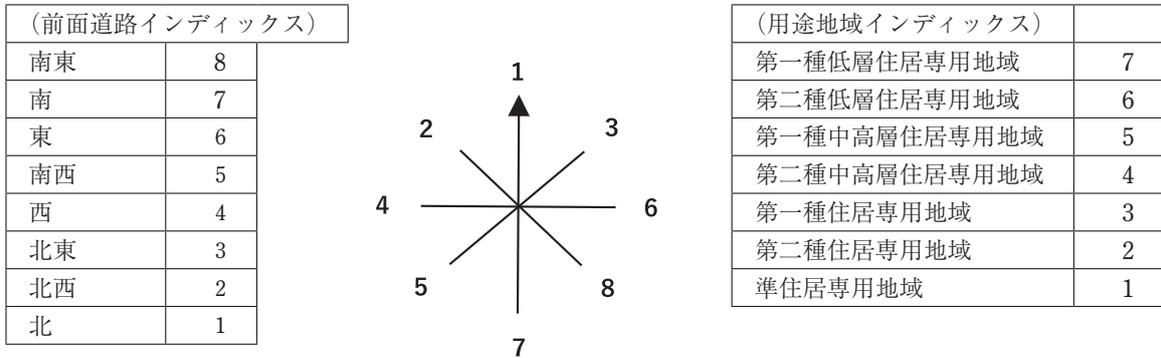


図2：変数

Ⅲ. AIの表現力と説明可能性問題

1. ニューラルネットワークの構造

(1) 重みと学習をもつネットワーク

小長谷 (2020a) にあるように、今のAIは、人間等の神経回路網をまねたものなので、神経細胞に対応するノードに入力データが入り、ある閾値を超えれば、活性化し、次のノードにつたえるという構造をしている。ラメルハートの「バックプロパゲーション (逆伝播法) (ラメルハート1989) により、最終的な判断結果を現実の判断結果と比較し、ニューラルネットの判断が正しければその接続を強化し、悪ければ弱めるという風に、もともどって、結果をフィードバックし、もとの接続 (の強さを表す重みW) を調整することであり、これによりコンピュータが「学習」ができることになった。自分で自分を改良することができるようになったのである。

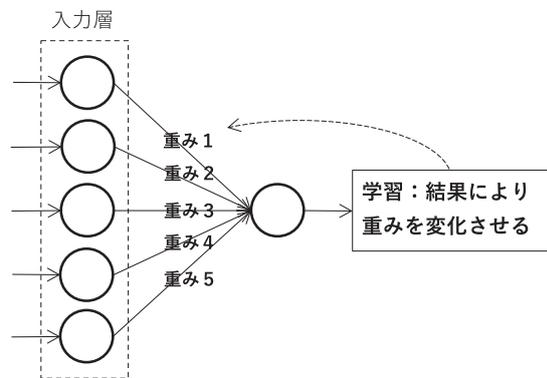


図3：小長谷 (2020a) の図11-2再掲

(2) 生物の細胞の判断機能を表現する活性化関数 (activation function)

ニューラルネットワークは、上記のように「ある閾値をこえ」と、「決断」し、次に判断をつたえるという機能が一番重要で、自明でない非線形の効果が入っている。この判断の瞬間を表現するものを「活性化関数 (activation function)」という。活性化関数は、ニューラルネットワーク神経回路網を模した、各ノードに対応する神経細胞の活性化の様相を表現する重要な関数であり、自明でない、非線形化の源泉である。2つの代表的な関数が使われる。

(a) シグモイド関数 = $1 / (1 + \exp(-ax))$

(b) ReLU関数 (Rectified Linear Unit、「レルー」と読む) = $\max(0, x) \dots (1)$

(a) は、tanhなどと同じで、-無限大から少しずつ増加し、ゼロ点付近で急激に立ち上がり、+無限大にむかって、漸近的に1に近づく関数である。

(b)は、さらに急激に、ゼロ点までゼロを返し、ゼロ点で傾き1の入力と同じ出力を返す関数である。

当初は、微分可能性を重視してシグモイド関数が使われることも多かったが、ディープラーニングの研究で有名なヒントンなどの主張もあり、現在は、ReLU関数が優位性をもっているという認識である。

(ReLU関数が有利な点1) AIはなんらかの最適化問題をとくために、上記の「バックプロパゲーション」で重みを調整する。これは最大最小問題だから関数の傾き(勾配)を評価する。ところがニューラルネットワークの層が深くなると、この勾配が敏感に変化しなくなり、調整ができない勾配消失問題(たとえば小長谷aの180頁)がおこる。とくにシグモイド関数等がこの問題が深刻なので、ReLU関数を用いる。

(ReLU関数が有利な点2) ニューラルネットワークのもともとの発想が、生物の神経細胞をシミュレートしたもので、神経が興奮して「発火」するかそうでないかをより忠実に表現するので、好ましい。

このため、現在はAIではReLU関数が主流となっており、本稿でもReLU関数を用いる。

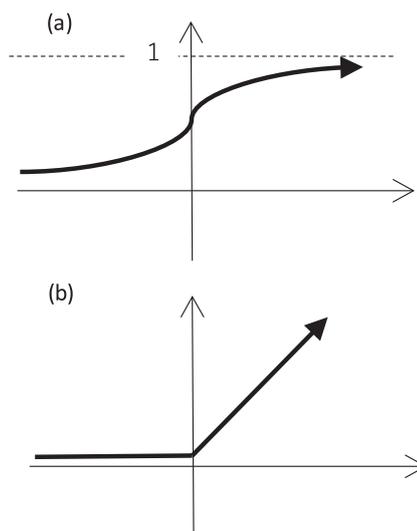


図4：代表的な活性化関数

(3) 一般的な関係

ある層 x_i から次の層 y_j への関数は、大きく言って、

$$y_j = \phi(\sum_i w_{ij} \times x_i + b_j) \cdots (2)$$

ϕ = 活性化関数

b_j は発火点を調整する次層ごとの調整位置を表す閾値の負符号、バイアス。

2. AIの表現可能性

これについては、3つの大きな発見が軸となっている。

(1) ミンスキーの定理 (中間層のない2層の限界)

隠れ層(中間層)のない、入力と出力だけのニューラルネットワーク(単純パーセプトロン(Simple perceptron))は、ミンスキーが

「中間層をもたない2層では、単純な線で引く単純判断(データを線で分ける分類判断)しかできない、線形非分離な問題を解けない」

ことを示してしまった(「ミンスキーの悪魔」といわれるミンスキーの不可能性定理)ため、一気に下火になってしまった(ミンスキー1969)。

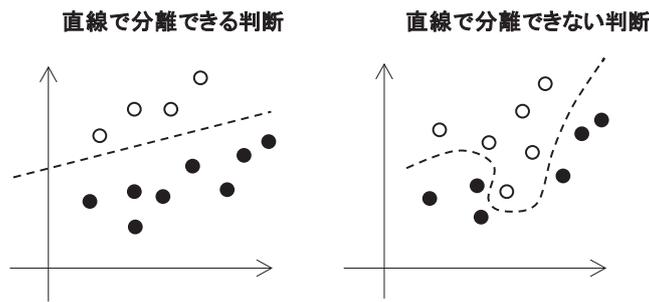


図5：小長谷 a (2020) の図11-4再掲、単純なニューラルネットワーク〈中間層なし〉の限界

(2) 表現可能性定理 (中間層をもつ場合あらゆる関数を近似できる可能性)

このようなことから、実際の脳と同じように、神経細胞をつなげ複雑化する、隠れ層 (中間層) を設けるアイデアがでてきた。このアイデアが思いのほか強力であることは、以下の定理からもわかる。

「ニューラルネットワークは、隠れ層が1層でもあると、(可微分で連続であるものは全て) あらゆる関数 (関係性) を近似できるという万能性をもつことが証明されている。」

これは、1変数のときは以下のような直感的理解で説明できる。以下、高木 (2018) を参考に説明する。式 (1) (2) から、隠れ層 (中間層) を作り、3層にするだけで、以下の図6の上図のようなパラメータをとれば任意の高さの「山」を表現できる。ここでは、活性化関数として、階段関数に近い形のきついシグモイド関数を想定しているが、ReLU関数だと三角の山になるだけで結論はほぼ同じである。さらに、同じ3層でも以下の図6の下図のように、ノードを追加すれば任意の高さの「谷」も表現できる。基本的には、任意の微分可能な関数は山谷からなっているので、このように細かく細分化することにより、任意の関数関係を、ノード数や層数を追加することにより、いくらでも表現できることになる。

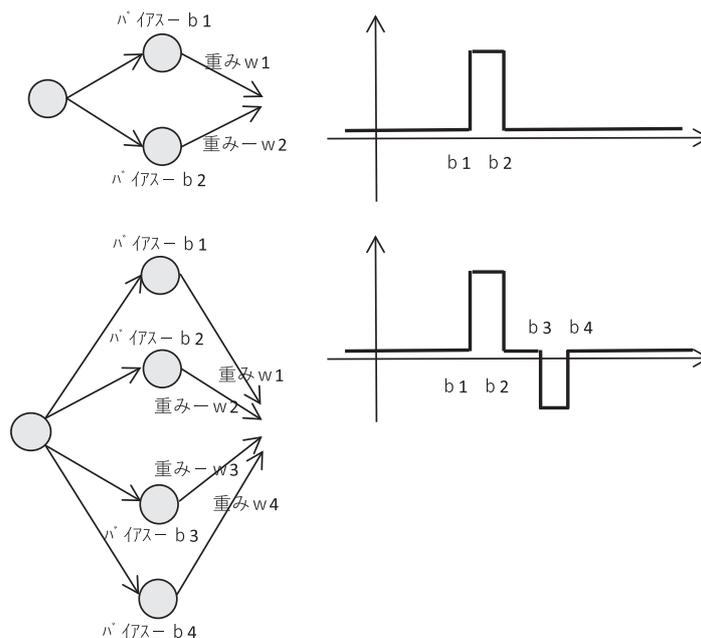


図6：AIの表現可能性を拡大する中間層の役割。高木 (2018) 等を参考に筆者作成。

(3) 接神経核理論 (NTK : Neural Tangent Kernel) (ノードとパラメータを増やせば、学習結果として、必ず問題解決する可能性)

わずか2年程前に得られた驚くべき結果がある(Jacot2018,甘利2020,甘利ら2020)。ランダム結合のニューラルネットワークを学習の初期値として使うと強力的に課題解決できる、すなわち、

「初期値としてランダムに結合したニューラルネットワーク(神経回路網)を用いれば、ノード数(神経素子数)、すなわち、パラメータの次元が十分に大きければ(データ数より十分大きければ)、学習の結果としてあらわれる正解は、初期値のごく近くにある。」

これをNTK理論という。Nは「神経」のニューラル、Tは「接する」意味のタンジェントを意味する。これは、ニューラルネットワークは、ノード数、層数を十分多くすれば、すなわちパラメータを多くとればどんな問題でもある意味最適化の解に近づくことができるという「AIの万能性」を示す強力な結果なのである。

これは、これまでのニューラルネットワークの研究者を相当驚かせた。というのは、ニューラルネットワークは、学習メカニズムモデルであり、それは、上述のように、最適化問題であり、実際には最大最小問題である。ところが微分などの局所的方法で解けるのは極値であり、大局的な最大最小値の必要条件であり十分条件でない。一般に最適化問題は、勾配降下法など、勾配の方向に沿って微分的に少しずつ動かして極値にいたるようにして極値解に達する。以下のように局所解である極値Bが間違っただ底にとらわれた場合、大局的な正解のAにはいけない。その異なる場合を判別し真の解に到達しなければならない。このことから、最適化問題では、確率的にジャンプさせる遺伝的アルゴリズムや量子的なアニーリングが開発されてきた。

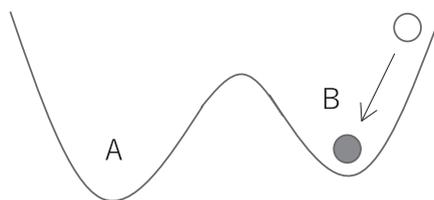


図7：最大最小問題と極値問題の違い。

ところが、このNTKの結果は、そのような心配をあまりしなくてもニューラルネットワークは「十分にパラメータを多くすれば」意外にも正解にたどり着けるといのである。ニューラルネットワークの万能性を示す驚くべき結果であるが、パラメータを多くすれば自由度が多くなり、高次元空間になるので、正解にたどり着く障害になる制約が少なくなるのではないと思われる。

3. これまでと全く違うモデル概念—生物はどのようにデータのモデルをつくるのか？

(1) 膨大な自由度で無限に発想し課題解決する。

ところで、上記のNTKの結果は、ある意味、これまでの数理モデルの考え方に革命をもたらすような、非常に重大な意味をもっている。

これまで通常の数理的なモデルは、ある意味で方程式的なものを、入力データの量、入力変数と、制約条件などをほぼ同じにして、いわば、制約をかけて、解を同定してきた。そこでは、データ、説明変数とパラメータは同程度であるから解が決まってくるという発想があった。ある意味、制約条件を大きく、パラメータは小さくする「決定論的モデル」の世界である。

ところが、ニューラルネットワークによるAIはそうではなく、パラメータ(ノードやレイヤー)を非常に

大きくすることにより、どのような解でも一応到達できるということなのである。いわば、自由度が多い、パラメータを多くとって確率的に正解にいたる「確率的解決モデル」の世界なのである。

理由はわからなくても、解決はできるのだということ、解決してしまえばよいということである。

じつは、この非常にネットワーク、パラメータが多いというのは、われわれ人間が最終的に動物から進化して発達させてきた脳の特徴である。

人間の脳の構造と機能

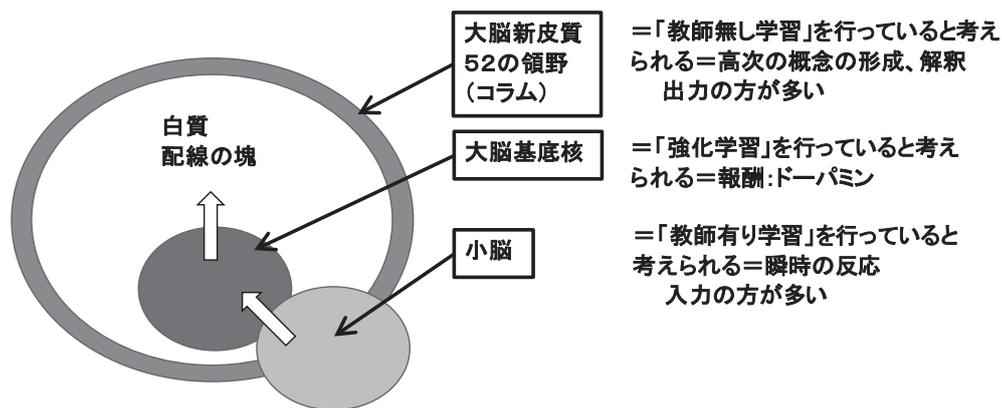


図8：小長谷 b (2020) の図12-4 再掲、脳神経系の階層性。

大脳は新皮質の細胞よりも、白質と呼ばれる配線の束の方が数が大きいほど、パラメータが多く、その中で、教師無し学習をおこなっている。このパラメータが多いということが、われわれ人間が、無限に近いほど自由な発想をおこない、なんとか課題解決することの一つの理由と考えられる。

甘利らの情報幾何学では、本来、ニューラルネットワークの制約条件の視点から、学習の軌道とMilnorアトラクターによって研究を進めてきたが、上記のようなNTKの驚くべき結果は、あまり制約が問題とならず結局はなんとか正解にたどりつく可能性を示唆しており驚かされている。「深層学習では腕力でパラメータ数 p を十分に大きくしてやってみると、結構うまくいく (甘利ほか2020)」ことが知られており、「一般の非線形系では局所解が無数に現れるため、勾配法は役に立たないというのが常識・・・(であり、しかし)・・・深層学習はこの常識を破り、やってみればうまくいくという話で押し通した (甘利2020)」 「・・・深層学習はこれを神話と化した。やってみてうまくいけばよいという腕力主義の勝利である (甘利2020)」とのべている。

すなわち、地球上の環境の中で生命が生きていくために発達させてきた、多数のパラメータをもつニューラルネットワークシステムは、通常の設定論的数理モデルとは考え方が異なり、理由はわからなくても、正解にたどりつく可能性のたかい驚異的なメカニズムであるということが出来る。

(2) 説明はあとからーニューラルネットワークのブラックボックス問題

このようにして、多数のパラメータをもつニューラルネットワークは、理由はわからなくても、課題解決だけは「直感で? 経験で?」できるという、まさにわれわれ人間の脳と似た特性をもっている。実際この強力な能力で、いまや人間の顔などの特徴を「カンで」人間より低いエラー率で予測し、自動運転も低速やガイドウェイのある条件ならなんなくこなすようになってきている。しかし、そのAIの判断が、説明が難しい、根拠を示せないというブラックボックス性があり、「AIの説明可能性」問題がでてきた。

小長谷 (2020b) では、以下のような「AI応用3法則」のうち、(A) がクリアされれば、その応用分野では商用、実用化段階になってくるということ述べた。

「(応用原則A) (エラー減少成績の存在) エラー率が人間を下回った場合。A Iのエラーを許容できる場合 (絶対間違いが許されないのではなく、エラー率が少なければA Iの利用が許容できる場合)」

「(応用原則B) (時間的余裕の存在) 判断を人間が確認できる時間的余裕がある場合。人間が最終判断したという形になる場合」

「(応用原則C) (フォローの存在) 結果に大きな法的責任が生じない場合。A Iがエラーした場合の人間のフォローを用意できる場合」

ところが、原則 (B) (C)、特に (C) がクリアされない場合は、法的責任がでてくる。実はこの典型例の分野が自動運転であり、どれだけ精密につくっても、たとえ天文学的数字の小さな確率でも、一旦事故を起こした場合、その事故の責任問題がでてくる。そのときに、原因は、信号無視なのか歩行者の飛び出しなのか理由、根拠があった方がよい。すなわち、今後のA Iの発展のためにも、問題をとくA Iの説明変数の要因を「見える化」する「A Iの説明可能性」を推進した方がよいのである。そこで本稿では、不動産データをもちい、A Iが解を得たとして、どのようなことがおこっているのかその構造をさぐる、A Iモデルの解の構造をできるだけ「見える化」することを試みる。

IV. 関数構造

1. 前提

(1) 一般的な関係

ある層 x_i から次の層 y_j への関数は、大きく言って、

$$y_j = \phi (\sum_i w_{ij} \times x_i + b_j)$$

ϕ = 活性化関数

b_j は発火点を調整する次層ごとの調整位置を表す閾値の負符号、バイアス。
である。

(2) 入力処理

最初の入力層から隠れ層第1層のところだけ、値が大きく違っていると、最初の非線形関数にかけることにより、スケールフリーの結果とならないので、値の大きさをそろえるスケール変換をする。すなわち、最初の入力層から隠れ層第1層のところだけ、ボックス型正規化

最大値、最小値を1、0とする正規化
をおこなっている。

(3) 出力処理

最後の中間層から全体出力へは、活性化関数はかけない。

2. 全体を書き下す。

隠れ3層、中間層は16ノードとする。

(1) 第0層 (入力層) : x_i (3変数 : $i = 1 \sim 3$ 、正規化されている)

(2) 第1層 : $y^1_j = \phi (\sum_i w^{01}_{ij} \times x_i + b^1_j)$ (16変数 : $j = 1 \sim 16$)

(3) 第2層 : $y^2_k = \phi (\sum_j w^{12}_{jk} \times y^1_j + b^2_k)$ (16変数 : $k = 1 \sim 16$)
 $= \phi (\sum_j w^{12}_{jk} \times \phi (\sum_i w^{01}_{ij} \times x_i + b^1_j) + b^2_k)$

$$\begin{aligned}
 (4) \text{ 第3層: } y^{3l} &= \phi(\sum k W^{23kl} \times y^{2k} + b^{3l}) \quad (16 \text{変数: } l=1 \sim 16) \\
 &= \phi(\sum k (W^{23kl} \times \phi(\sum j W^{12jk} \times y^{1j} + b^{2k}) + b^{3l})) \\
 &= \phi(\sum k (W^{23kl} \times \\
 &\quad \phi(\sum j W^{12jk} \times \phi(\sum i W^{01ij} \times x_i + b^{1j}) + b^{2k}) + b^{3l}))
 \end{aligned}$$

$$\begin{aligned}
 (5) \text{ 第4層 (出力層): } y^{\text{out}} &= \sum l W^{34l} \times y^{3l} + b^4 \quad (1 \text{変数}) \\
 &= \sum l W^{34l} \times \\
 &\quad \phi(\sum k (W^{23kl} \times \\
 &\quad \phi(\sum j W^{12jk} \times \phi(\sum i W^{01ij} \times x_i + b^{1j}) + b^{2k}) + b^{3l})) \\
 &\quad + b^4
 \end{aligned}$$

V. フィッティング結果

Googleが提供しているTensorFlowのVer.2を使用する。

隠れ3層、中間層は16ノードとした。

回帰分析のmac = 30416

AIのmac = 29958

となり、回帰分析より良い状況である。

VI. 結果の解釈

ところで、現在のAIをめぐる最大の問題にブラックボックス問題がある。

そこで、最終的に得られた解である、モデルのパラメータ群（重み行列W、バイアスペクトルB）をすべて可視化してみる。

1. パラメータ群（重み行列W、バイアスペクトルB）

最終的なパラメータの傾向は以下のとおりである。

(1) 入力層（第0層）から第1層までの重み行列 W^{01} をみると、絶対値で0.5以上は、第1変数から第1層の第1、3～4、6～8、11～13、15～16番目のノードに向かうリンクと、第2変数から第1層の第1、3番目のノードに向かうリンクに大きな値がある。

(2) 第1層のノードで事実上「発火」しているもの、すなわち、出力が事実上0でないのは、第1、3～4、6～8、11～13、15～16番目のノードである。

(3) 第1層から第2層までの重み行列 W^{12} をみると、絶対値で事実上0でないのは、第1層の第1、3～4、6～8、11～13、15～16番目のノードから、第2層の第1、3、5～7、9～14、16番目のノードに向かうリンクに大きな値がある。

(4) 第2層のノードで事実上「発火」しているもの、すなわち、出力が事実上0でないのは、第1、3、5～7、9～14、16番目のノードである。

(5) 第2層から第3層までの重み行列 W^{23} をみると、絶対値で事実上0でないのは、第2層の第1、3、5～7、9～14、16番目のノードから、第3層の第9、11、13番目のノードに向かうリンクに大きな値がある。

(6) 第3層のノードで事実上「発火」しているもの、すなわち、出力が事実上0でないのは、第9、11、13番目のノードである。

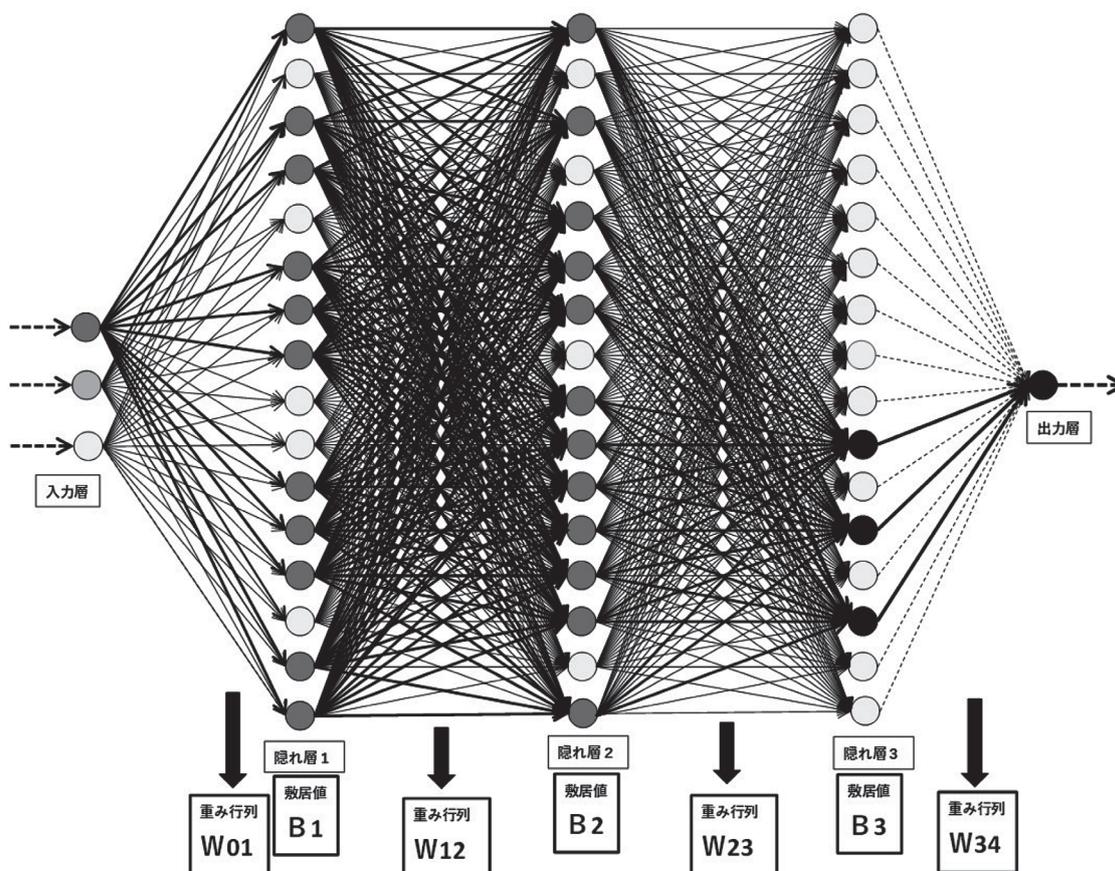


図9：不動産データの解=パラメータ群（重み行列W、バイアスペクトルB） 筆者作成（重み行列Wが0より有意に大きいリンクを太線、バイアスペクトルBが有意に大きいノードを黒丸で表現している。）

2. パラメータ群の変動からみた特徴

以下の3つの結果がえられる。

(1) 概念の要約

この図9では、重み行列Wが0より有意に大きいリンクを太線、バイアスペクトルBが有意に大きいノードを黒丸で表現している。

これをみると、特に第2層から第3層に移るところで、劇的に変数、パラメータの縮約がおこっている。第2層から第3層までの重み行列 W^{23} をみると、絶対値で事実上0でないものとしては、第2層は12個のノードがあるのに、第3層はたった3つのノードに要約されている。第3層のノードで事実上「発火」しているもの、すなわち、出力が事実上0でないのは、第9、11、13番目のノードである。

このことはAIがここで劇的にデータの要約をおこなっていることがわかる。

(2) 説明変数

横浜市の住居系の不動産の公示地価では、説明変数の影響力は、第1変数（距離変数）の説明力が大きく、これに前面道路方位がサブの要因で考慮されている。第3変数の用途地域指定はほとんど影響がない。

3. 特にReLU関数の場合の傾向

(1) ニューラルネットワークでは、非線形が強調されてきた。

人工神経の活性化関数は、ネットワークを強化または単純化するような特性を持つものが選ばれる。実際、線形伝達関数を使った多層パーセプトロンには、全く等価な単層ネットワークが必ず存在する。したがって、多層ネットワークの利点を生かすには非線形関数が必須であると考えられてきた。

(2) しかし、社会科学等のこのような「素直な」データでは、発火パターンと線形関数の重ね合わせになる場合が多いと考えられる。すなわち、標語的には、以下のような傾向が強いといえる。

(仮説)

社会科学等のデータでは、概略的には、

ReLU関数をつかったニューラルネットワークの判断 = (発火パターン関数) + (線形関数の貼り合わせ) と把握できる場合が多い。

といえる。

このことは、計量経済学分野でも似たような傾向を指摘できる。最近それらの分野ででてきた部分識別法(奥村2018など)なども似た問題意識がみとめられるので、それゆえ、社会科学・社会工学におけるAI分析の推進は、それらの方法を補完するような役割も示唆しているといえる。

(謝辞) 本研究は、東京大学空間情報科学研究センターの以下のプロジェクトによるデータを使用した。また同プロジェクトの推進にあたりセンターの柴崎亮介先生はじめ皆様にはお世話になりました。ここにお礼申し上げます。新規共同研究名称「AI(人工知能)手法を用いた不動産データの予測可能性の向上に関する研究」(A Study on the Fitting and Explanatory Performance in the Real Estate Prediction Models Using AI methods Compered to Ordinary Methods)

【参考文献】

- 甘利俊一 (2016) 『脳・心・人工知能－数理で脳を解き明かす』講談社。
- 甘利俊一 (2020) 「深層神経回路網の幾何－高次元推測」『数理科学』第689号 (2020年11月号 特集：情報幾何学の探究－基礎と応用、現状と展望に迫る)。
- 甘利俊一・鈴木大慈・穴井宏和 (2020) 「AIの常識と未解決問題」『数学文化』第34号、日本評論社。
- 五木田和也 (2016) 『コンピュータで「脳」がつかれるか』技術評論社。
- 伊庭斉志 (2015) 『進化計算と深層学習 (ディープラーニング)』オーム社。
- 今井隆志 (2020) 「第4章 AIソフト：テンソルフローの使い方」村上他編『AIと社会・経済・ビジネスのデザイン』日本評論社。
- 大関真之 (2016) 『機械学習入門－ボルツマン機械学習から深層学習まで』オーム社。
- 奥村綱雄 (2018) 『部分識別入門－計量経済学の完結的アプローチ』日本評論社。
- 唐木田亮 (2020) 「深層神経回路網の幾何－統計神経学とのつながり」『数理科学』第689号 (2020年11月号 特集：情報幾何学の探究－基礎と応用、現状と展望に迫る)。
- 国土交通省 (2019) 「日本が主導してきた自動運転技術に関する国際ルールが国連で合意！～衝突被害軽減ブレーキの国際基準の成立～」令和元年6月28日告示。
- 小長谷一之 (2015) 「第8章：インターネットと都市」『角川インターネット講座10－第三の産業革命経済と労働の変化』KADOKAWA。
- 小長谷一之他編 (1999) 『マルチメディア都市の戦略－シリコンアレーとマルチメディアガルチ』東洋経済新報社。
- 小長谷一之 (2014a) 「都市構造の変容」近畿都市学会編『都市構造と都市政策』古今書院。
- 小長谷一之 (2014b) 「特集：高度化した都市の構築を－単なる縮小から「高付加価値都市(スマートシティ)へ」『日刊建設産業新聞』。
- 小長谷一之 (2020a) 「第11章 AIの発展と課題総括－AIそのもののデザインのもつ意味と課題」村上他編『AIと

- 社会・経済・ビジネスのデザイン』日本評論社。
- 小長谷一之(2020b)「第12章 AIと社会・経済のデザイン」村上他編『AIと社会・経済・ビジネスのデザイン』日本評論社。
- 清水千弘(2020)『不動産市場分析のデータ資源』『日本不動産学会誌』Vol33, No4。
- 清水千弘編著(2020)『不動産テック: Real Tech』。
- 高木聡(2018)『TensorFlowを使った「ニューラルネットワーク」の構築法』工学社。
- 田中宏和著(2016)『計算論的神経科学』森北出版。
- 松尾豊(2015)『人工知能は人間を超えるのか-ディープラーニングの先にあるもの』KADOKAWA。
- 村上憲郎・服部桂・近勝彦・小長谷一之編著(2020)『AIと社会・経済・ビジネスのデザイン』日本評論社。
- 村田昇(2020)「機械学習における確率的推測」『数理科学』第689号(2020年11月号 特集:情報幾何学の探究-基礎と応用、現状と展望に迫る)。
- Motor Magazine編集部(2020/01/08)「2019年12月の改正道路交通法で、自動運転レベル3実用化へ法整備が整った」。
- Motor Fan illustrated編集部(2019/07/01)「国土交通省:日本が主導してきた自動運転技術に関する国際ルールが国連で合意」。
- 涌井良幸・涌井貞美(2017)『ディープラーニングがわかる数学入門』技術評論社。
- 渡辺正峰(2017)『脳の意識 機械の意識-脳神経科学の挑戦』中央公論新社。
- A.Jacot, F.Gabriel and C.Hongler(2018) 'Neural tangent kernel: Convergence and generalization in neural networks' '*Neurips*'.
- Minsky, Marvin, and Papert S. (1969) "*Perceptron*", Cambridge, MA: MIT Press (ミンスキー、S.パパート(1971)『パーセプトロン-パターン認識理論への道』、斎藤正男訳、東京大学出版会)。
- Geoffrey E. Hinton & R. R. Salakhutdinov (2006) 'Reducing the Dimensionality of Data with Neural Networks', "*Science*", 313 (5786), pp.504-507.
- Krizhevsky, Alex, Ilya Sutskever, and Geoffrey E. Hinton. (2012) 'Imagenet classification with deep convolutional neural networks' "*Advances in neural information processing systems*".
- Rumelhart, David E., Hinton, Geoffrey E. and Williams, Ronald J. (1986) 'Learning representations by back-propagating errors' "*Nature*" 323 (6088): pp.533-536.
- Rosenblatt, Frank(1958) 'The Perceptron: A Probabilistic Model for Information Storage and Organization in the Brain' "*Psychological Review*" 65 (6): pp.386-408.
- (ホームページ)
- 横浜市ホームページ「18区の地図」<https://www.edu.city.yokohama.jp/study/town/18ku01.htm>
- 第14回全脳アーキテクチャ勉強会「深層学習を越える新皮質計算モデル」報告レポート(2016/08/15)松田卓也(2016)「大脳新皮質のマスターアルゴリズムの候補としてのHierarchical Temporal Memory (HTM) 理論」発表資料 <http://wba-initiative.org/1653/>
- 第4回全脳アーキテクチャシンポジウム, 教育事業 - event, 全脳アーキテクチャシンポジウム(2019年11月23日)松尾豊(2019)「理解するとは何か? - 高次元科学と記号処理」発表資料 <https://wba-initiative.org/3701/>