

グラフ畳み込みニューラルネットワークによる 日本の都市発展の予測に関する研究

岡室 智貴¹・Jan-Dirk Schmöcker²・福田 幸二³・
中尾 聡史⁴・孫 文哲⁵・山田 忠史⁶

¹ 非会員 京都大学大学院工学研究科 (〒615-8540 京都府京都市西京区京都大学桂)
E-mail: okamuro.tomoki.n19@kyoto-u.jp

² 正会員 京都大学大学院工学研究科 (〒615-8540 京都府京都市西京区京都大学桂)
E-mail: schmoecker@trans.kuciv.kyoto-u.ac.jp

³ 正会員 株式会社日立製作所研究開発グループ基礎研究センター
(〒350-0395 埼玉県比企郡鳩山町赤沼 2520 番地)
E-mail: koji.fukuda.jf@hitachi.com

⁴ 正会員 京都大学大学院工学研究科 (〒615-8540 京都府京都市西京区京都大学桂)
E-mail: nakao@trans.kuciv.kyoto-u.ac.jp

⁵ 正会員 京都大学大学院工学研究科 (〒615-8540 京都府京都市西京区京都大学桂)
E-mail: wz.sun@trans.kuciv.kyoto-u.ac.jp

⁶ 正会員 京都大学経営管理大学院 (工学研究科併任) (〒606-8317 京都府京都市左京区吉田本町)
E-mail: yamada.tadashi.2x@kyoto-u.ac.jp

都市の発展を予測することは、国土計画や都市計画において極めて重要である。そこで、本研究では、交通ネットワークをグラフ構造で表現し、グラフ畳み込みニューラルネットワーク (GCN) を用いた都市発展の予測を行った。GCN を用いて 10 年後の人口や経済指標を予測するモデルを構築したところ、決定係数は約 0.9 を示し、その精度が高いことが確認された。また、構築したモデルを用いて、1976 年において山陽新幹線が整備されていないケースにおける予測を行い、整備されていたケースと比較したところ、山陽新幹線の整備が 1986 年の人口や経済に正の影響を与えていることが確認された。

Key Words: urban growth, graph convolutional neural network, transportation network, prediction model

1. はじめに

都市は、人口増加や経済成長を促しながら、発展してきた。世界銀行の融資分野別比率を見ると、交通インフラ整備への融資比率が大きく¹⁾、交通インフラ整備は都市発展の重要な要因の一つと考えられてきた。実際に、高速道路や鉄道の整備が人口増加や経済成長をもたらすことを指摘した研究は多く見られる²⁾³⁾⁴⁾⁵⁾。また、都市の発展は移動需要の増加をもたらし、それによって交通インフラ整備の需要も高まることから、都市の発展と交通インフラには相互関係があることも認められている⁶⁾。

交通インフラ整備には、多大な時間と費用を要するため、国土計画や都市計画において、その効果を正確に評価し、予測していくことが求められている。我が国では、

国土交通省が、交通インフラ整備が都市の人口や経済に与える影響を可視化しようと試みてきた⁷⁾。そして、帰着先が分かる空間的応用一般均衡 (Spatial Computable General Equilibrium: SCGE) モデルが用いられ、交通インフラ整備がもたらす経済効果の評価がなされている。しかし、SCGE は、パラメータの設定等により推定結果に幅が生じるため、確定的な解釈が困難であることが指摘されている⁸⁾。

また、既往研究を概観すると、都市成長モデルに関するものは数多く存在するものの、交通ネットワークや都市間の相互関係を十分に考慮できていない。人口増加や経済成長といった都市発展のプロセスは、非常に複雑であることから、その要因を特定することは困難であり、都市発展を予測する分析手法は一般化されていないのが

現状である。

そこで、本研究では、深層学習を用いた都市発展の予測を試みる。具体的には、交通ネットワークをグラフ構造で表現し、ニューラルネットワーク (Neural Network: NN) を用いて、人口や事業所数、農業生産額、工業製品出荷額などを予測するモデルを構築する。さらに、予測結果の解釈を行い、交通インフラ整備が都市の発展にもたらす影響について考察する。

2. 既往研究の整理と本研究の位置づけ

(1) 既往研究

様々な分野において、都市発展について研究がなされてきた。その一つが経済学で発展した都市経済学という分野である⁹⁾。Von Thünen¹⁰⁾の付け値地代理論を Alonso¹¹⁾が都市成長モデルに応用したことから都市経済学の歴史が始まったと言われる¹²⁾。これは、後に Mills¹³⁾と Muth¹⁴⁾によって発展され、Alonso-Muth-Mills モデルとしても知られている。この Alonso-Muth-Mills モデルにより、交通費用の低下によって、郊外に比べて中心部付近の人口密度が低下することが実証されている。この結果は、交通インフラ整備が都市の人口を決定する重要な要因であることを示唆している。そして、Hansen¹⁵⁾によって、アクセシビリティが土地利用に与える影響が明らかになり、交通と土地利用の間の相互関係が考慮されるようになる。Aljoufie ら⁶⁾は、サウジアラビアの Jeddah 市を対象に、リモートセンシング技術と地理情報システムを用いた時空間的な分析を行い、交通インフラの拡大は、人口増加、空間的拡大、土地利用の変化と強い相関関係があることを示した。このように、土地利用と交通が相互に影響し合うという知見が蓄積されるようになるが¹⁶⁾、既存のモデルでは十分に土地と交通の相互関係を捉えきれていないという課題も指摘されている¹⁷⁾。

特に、都市は長い時間をかけて形成されるものであることから¹⁸⁾¹⁹⁾、都市発展の予測においては長期的な視点が重要になる。Berger ら²⁰⁾は、スウェーデンの 150 年間のデータを対象に、交通インフラの短期的効果と長期的効果の双方を分析したところ、鉄道建設は長期にわたって、都市発展に影響を与えた可能性があることを示唆している。Banerjee ら²¹⁾もまた、都市発展を捉えるには、長期及びマクロの視点で分析することが重要であると説いている。しかし、都市の成長は不可逆的であり²²⁾、その時間的・空間的複雑さから、短期的にしか確実な予測ができないことも指摘されており²³⁾、統計的手法による長期での都市発展の予測は困難であると考えられる²⁴⁾。そこで、近年では、時系列データなどを利用したデータ駆動型のアプローチが注目を集めている。特に、柔軟な

非線形モデルである NN は、都市経済学において有用な方法であると考えられている¹⁷⁾。なぜなら、統計的モデルでは、都市発展の動的で非線形なプロセスを捉えることが困難だが、NN はその複雑で不定な非線形性の挙動を扱うことができるからである。実際に、土地利用の変化の予測において、NN を用いた研究が存在している¹⁷⁾²⁵⁾。

また、都市発展を予測するにあたって、交通ネットワークの影響も考慮していく必要がある。Berger ら²⁰⁾は、鉄道網の拡大でネットワークにアクセスした町の人口が短期間で相対的に大きく増加したことを確認している。増加の主な要因は、ネットワークに接続されていない近隣の町からの経済活動の移転であることを明らかにしている。都市の発展は、その周辺の地方都市との間の相互作用や、さらにその周辺にある大都市との相互作用についても考慮していく必要がある²⁶⁾。しかし、交通ネットワークの成長の評価や、交通アクセシビリティが都市発展に与える影響についての研究は不足している²⁷⁾。

(2) 本研究の位置づけ

このように都市・地域の発展を表すモデルは様々あるが、その多くは長期的な発展には焦点を当てておらず、ネットワークの影響を十分に考慮できていない。以上を踏まえ、本研究では、グラフ畳み込みニューラルネットワーク (Graph Convolutional Network: GCN) を用いた都市発展の予測を試みる。

GCN は、その名の通り、グラフ構造に適合した NN である²⁸⁾。NN 手法は、長期の時系列データの予測において、高い精度を出すことが期待されている²⁹⁾。また、GCN では、他ノードとの接続性を考慮して、ノード情報を更新できることから、ネットワークを考慮したモデルの構築が可能である²⁸⁾。さらに、GCN は、変数に正規性や定常性などの仮定を必要としないため、統計的手法と比べて多くの指標を、高い精度で予測できるものと期待できる。

3. 分析方法

(1) 入力データ

本研究では、1976 年から 2016 年の日本全土を対象とした分析を行った。ゾーニングは、全国幹線旅客純流動調査で設定されている 207 生活圏とする。207 生活圏は国土交通省が全国の旅客流動の実態を定量的かつ網羅的に把握する際に定められたものである。

a) ノード

ノード特徴量として、人口 (1 指標)・経済 (7 指標)・交通 (2 指標) に関する値の計 10 種類を扱う。表

表-1 ノード特徴量

特徴量名	調査名	調査主体	集計年
人口	国勢調査	総務省	1975, 1980, 1985, 1990, 1995, 2000, 2005, 2010, 2015, 2020
事業所数 従業者数	事業所・企業統計調査経済センサス	総務省	1975, 1981, 1986, 1991, 1996, 2001, 2006, 2009, 2014, 2019
農業産出額	生産農業所得統計	農林水産省	1975, 1980 から毎年
製造品出荷額等	工業統計調査	経済産業省	1971 から毎年
卸売販売額 小売販売額	商業統計	経済産業省	1974, 1976, 1979, 1982, 1985, 1988, 1991, 1994, 1997, 1999, 2002, 2004, 2007, 2012, 2014, 2016
課税対象所得	市町村税課税状況等の調査	総務省	1975 から毎年
新幹線駅数	国土数値情報鉄道時系列	国土交通省	
高速道路 IC 数	国土数値情報高速道路時系列	国土交通省	

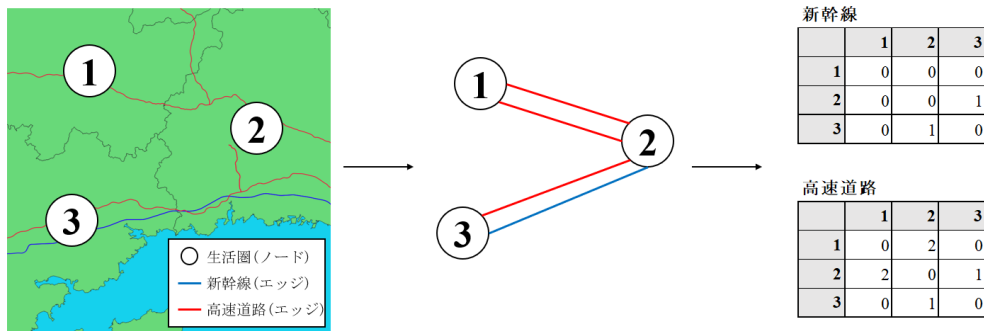


図-1 接続リンク数の表現方法

-1 にそれぞれの調査機関及び調査名，集計年についてまとめた。なお，データの欠損値については線形補間を行った。

b) エッジ

本研究では，新幹線と高速道路のネットワークをエッジとみなした。エッジ情報は国土数値情報にある時系列データより入手した。なお，新幹線についてはフル規格とミニ規格の種類があるが，本研究ではミニ規格のものも新幹線として扱った。

エッジの特徴量には，接続リンク数（新幹線と高速道路の2種類），近隣生活圏数，最小所要時間（新幹線と高速道路の2種類）の5種類を用いた。接続リンク数は，生活圏間を繋ぐ新幹線や高速道路の本数を表す（図-2）。同一路線が複数回生活圏の境界をまたいでも，それらは1本のリンク数として数えた。近隣生活圏数は，生活圏の代表都市から半径 50km 内にある他代表都市が属する生活圏を近隣とした。所要時間は，新幹線や高速道路を利用した際の最小所要時間を表し，国土交通省の総合交通システム NITAS を用いて計算した。新幹線による移動時間は JR の乗車ダイヤから計算されたものであり，高速道路による移動時間は車の平均旅行速度を用いて，ダイクストラ法で算出されたものである。なお，駅での待ち時間は考慮しなかった。

(2) モデル

a) グラフ畳み込み層

GCN は，ディープラーニングをグラフデータに適用する手法である。本研究では生活圏をノード，生活圏間の関係をエッジと定義した。ノード数は 207 で，人口，事業所数といったノード特徴やエッジは時間と共に変化する。各時間帯について， N 個のノード $v_i \in \{1, 2, \dots, N\} \in \mathcal{V}$ と R 種類のエッジ $(v_i, v_j)^{(r)} |_{r \in \{1, 2, \dots, R\}} \in \mathcal{E}$ で無向グラフ $G = (\mathcal{V}, \mathcal{E})$ として表すことができる。各ノードは F 個の特徴を持ち，ノード特徴行列は $X \in \mathbb{R}^{N \times F}$ と表記される。つまり，本研究では $N = 207$ ， $F = 10$ ， $R = 5$ である。エッジは R 個の重み付き隣接行列 $A^{(r)} |_{r \in \{1, 2, \dots, R\}} \in \mathbb{R}^{N \times N}$ として書くことができる。本研究では，スペクトラルアプローチで畳み込み演算を一般化し，グラフに適用した。グラフ畳み込みネットワークの第 $(l+1)$ 層における出力は次式(1)で表される²⁸⁾。

$$H^{(l+1)} = \sigma \left(\tilde{D}^{-\frac{1}{2}} \tilde{A} \tilde{D}^{-\frac{1}{2}} H^{(l)} \theta^{(l)} \right) \quad (1)$$

$H_t^{(l)} \in \mathbb{R}^{N \times F}$ は l 層における出力であり， $l = 0$ のときは入力層になる。つまり， $H^{(0)} = X_t \in \mathbb{R}^{N \times F}$ となる。 σ は活性化関数である。 $\tilde{D} = [d_{ij}]$ は $\tilde{A} = [a_{ij}]$ の次数行列であり， $\tilde{d}_{ii} = \sum_j \tilde{a}_{ij}$ より，算出される。 $\tilde{A} = A + I$ より，隣のノードだけでなく，自身の特徴量も考慮する。 $\theta^{(l)}$ は学習対象となる重み行列である。

本研究では複数のエッジを考慮するために，Schlichtkrull ら³⁷⁾の R-GCN 手法を用いて，式(1)を下記のように変換した。

$$H^{(l+1)} = \sigma \left(H^{(l)} \theta^{(l)} + \sum_{r=1}^R \tilde{D}_r^{-\frac{1}{2}} \tilde{A}_r \tilde{D}_r^{-\frac{1}{2}} H^{(l)} \theta^{(l)} \right) \quad (2)$$

式(2)に従い、ネットワークを考慮してノード情報を更新する。

b) Graph U-nets 層と Jumping Knowledge 層

本研究では、伝搬精度を高めるため、U-Net の手法を用いた。U-Net はエンコーダとデコーダから構成されている。U-Net のエンコーダは、入力した画像を何度か畳み込みし、特徴を抽出する。一方、デコーダは、エンコーダによって抽出された特徴を受け取り、逆畳込み処理を行い、アップサンプリングする。アップサンプリングとは、小さな特徴マップを大きな特徴マップにする処理だが、単純にアップサンプリングすると、物体の位置情報をうまく捉えられない。そこで、U-Net は、各階層において、エンコーダの特徴マップをデコーダの特徴マップに連結させている。Gao ら³⁰⁾は、GCN で U-Net を扱える Graph U-Nets を提案している。Graph U-Nets は、graph Pooling と graph Unpool で構成されている。本研究では、このプーリングとアンプーリング作業を 3 回繰り返した。さらに本研究では、Xu ら³¹⁾によって提案された Jumping Knowledge 層を取り入れた。これによって、適切な範囲で周辺ノードの特徴を取り込むことができる。

(3) SHAP

統計的手法と異なり、解釈を目的としないニューラルネットワークは回帰係数などを得ることができず、その挙動は、ブラックボックスと言われる²⁷⁾。しかし、実際に政策などを考える際は、「そのモデルがなぜその予測を行ったのか」や「どの様な特徴が効いているのか」を知ることが重要である。そこで本研究では、特徴量のインパクトを見るために、Lundberg ら³²⁾によって開発された SHAP (SHapley Additive exPlanations) を導入した。SHAP は、協力ゲーム理論のシャープレイ値を機械学習に応用したものである。シャープレイ値は、協力ゲーム理論において、複数プレイヤーの協力によって得られた利得を各プレイヤーに公正に分配するための手段の一つとして活用されている。プレイヤー j のシャープレイ値 ϕ_j は次式で表される。

$$\phi_j = \sum_{S \subseteq P \setminus \{j\}} \frac{|S|! (|P| - |S| - 1)!}{|P|!} \times [f_{S \cup \{j\}}(x_{S \cup \{j\}}) - f_S(x_S)] \quad (3)$$

P は全プレイヤーの集合であり、 S はその部分集合である。 x_S は集合 S のプレイヤーの特徴になる。 f は特性関数であり、 $f_{S \cup \{j\}}$ と f_S は、プレイヤー j が集合 S に属する場合と属さない場合の期待配当の合計を表す。つまり、

プレイヤー j のシャープレイ値 ϕ_j は、平均的な結果から特定の結果に寄与した貢献度を表す³³⁾。機械学習の予測モデルにこの概念を適応すると、各ノード特徴量がプレイヤーに該当し、報酬は予測モデルの出力結果になる。すなわち、SHAP では、ひとつひとつの特徴量がモデル予測値に与える貢献度をシャープレイ値で表される。

SHAP は、シャープレイ値の説明を線形モデルである Additive Feature Attribution Methods で表現している。具体的に、 $\mathbf{X} = (X_1, \dots, X_M)^T$ を特徴量とした機械学習モデル $f(\mathbf{X})$ について考える。モデルの平均的な予測結果 $E[f(\mathbf{X})]$ と、あるインスタンスに対する予測 $f(\mathbf{X})$ の差分を各特徴量の貢献度 ϕ_j に分解する。 $f(\mathbf{x}) - E[f(\mathbf{X})] = \sum \phi_j$ より、 $E[f(\mathbf{X})]$ を ϕ_0 に置き換え、特徴量の存在有無を考慮すると、バイナリ値 $z \in \{0,1\}^M$ を用いて、次式で SHAP を定式化できる。

$$f(z) = \phi_0 + \sum_{j=1}^M \phi_j z_j \quad (4)$$

なお、実際には正確な SHAP 値を求めるのは困難である⁴⁰⁾。そこで、モンテカルロサンプリングを取り入れ、以下のように近似する。

$$f(z) = E[f([z_S, z'_S]) | z'_S \in \bar{S}] \quad (5)$$

S は z の非負インデックス集合、 \bar{S} は S の補集合、 z_S は \bar{S} にある特徴量の欠損値である。特徴量の独立性とモデルの線形性を仮定すると、式(5)の右辺は、 $f([z_S, E[z'_S]])$ と近似できる。

本研究で使用する説明モデルは、DeepExplainer³²⁾に基づいている。しかし、DeepExplainer は、グラフ構造に適用していないため、本研究ではグラフ構造に対応した DeepExplainer の改良版、GraphExplainer³⁴⁾を用いた。

4. 結果と考察

本研究では、ノード特徴量の内、新幹線駅数と高速道路 IC 数を除く 8 つの指標に対して、10 年後の値を予測するモデルを構築し、その精度を検証した。本章では、人口を予測したモデルの結果について考察を行う。

(1) モデル精度

表-2 の通り、10 年間隔の組み合わせを 7 組作成し、両端を除いた 5 組の訓練データ(青)を GCN で学習させ、端の 2 組をテストデータ(赤)として、モデルの精度を検証

表-2 time step の組み合わせ

Time step	0	1	2	3	4	5	6
入力層	1976	1981	1986	1991	1996	2001	2006
出力層	1986	1991	1996	2001	2006	2011	2016

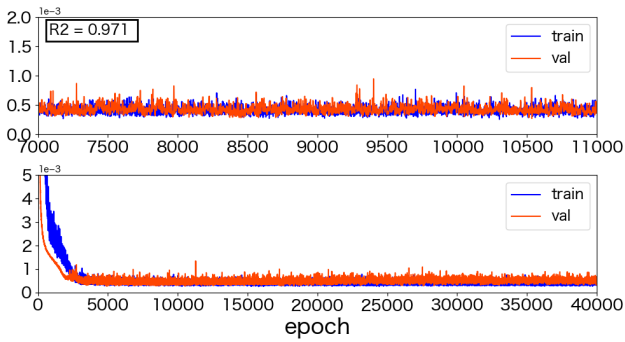


図-2 人口モデルの損失曲線

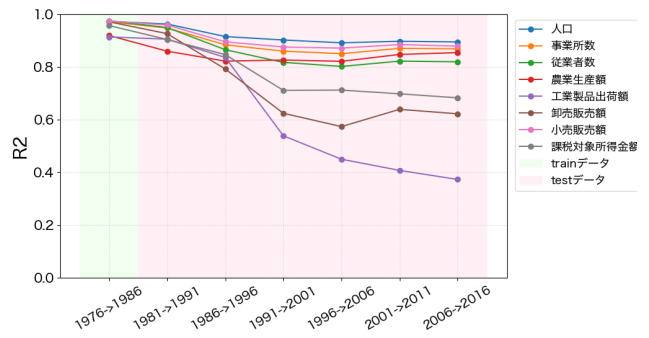


図-5 決定係数の推移 (time step 0 を学習させたケース)

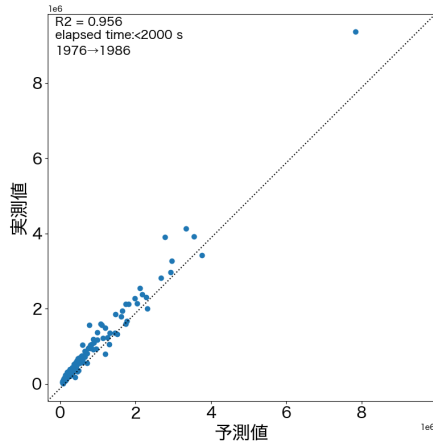


図-3 time step 0における予測値と実測値の散布図

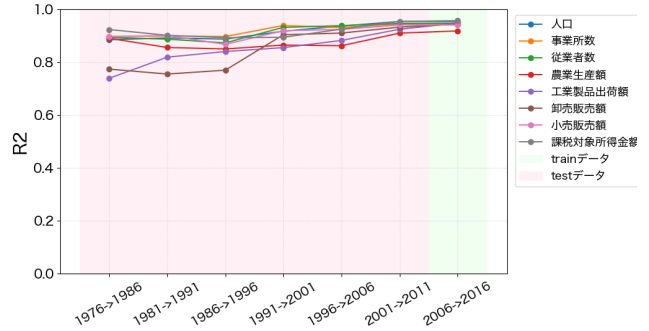


図-6 決定係数の推移 (time step 6 を学習させたケース)

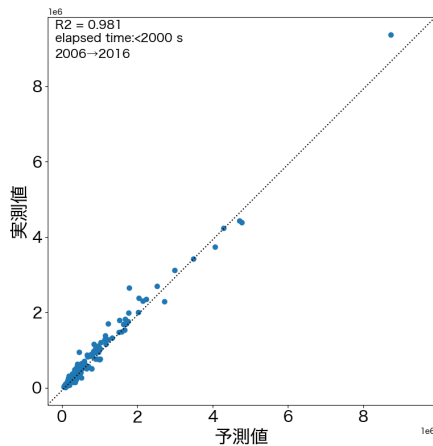


図-4 time step 6における予測値と実測値の散布図

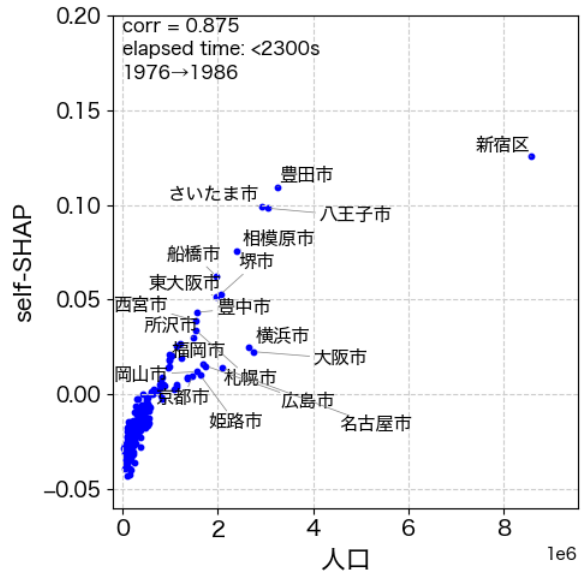


図-7 人口と人口予測の self-SHAP 値の関係

した。モデルは、損失関数（平均二乗誤差）を用いて、最適なパラメータを決めたのち、その精度を決定係数で評価した。図-2 は、人口を予測するモデルの損失の変化を表したものである。

学習率 0.0001, エポック数 9000 のパラメータを採用し、テストデータの予測における決定係数を算出すると、平均で 0.971 を示した。これは、時系列分析による精度より高かった。なお、予測値と実測値（真値）の散布図を図-3 と図-4 に示した。

訓練データが十分であることを検証するため、7 組の time step の内、time step 0 もしくは time step 6 の 1 組だけを訓練データ、残りの 6 組をテストケースにしたケースを扱った。図-6 と図-7 は人口と経済指標を予測したモデルの精度の変化を表したものである。多くの経済指標は、訓練データの時点から離れるほど、その精度が低下した。しかし、人口、事業所数、従業者数、農業生産額、小売販売額の決定係数はいずれも 0.8 以上で安定した。time step 6 を学習したモデルより、time step 0 を学習したモデ

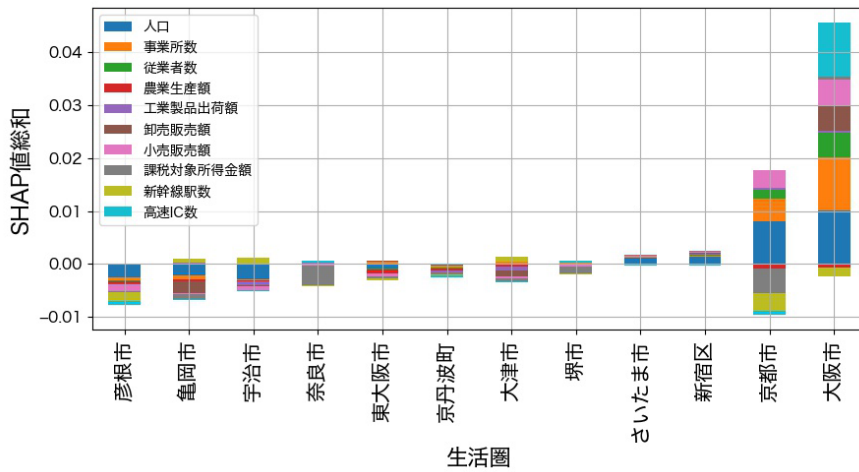


図-8 京都の人口に対する周辺 SHAP (time step 6 のケース, 絶対値 0.001 以上)

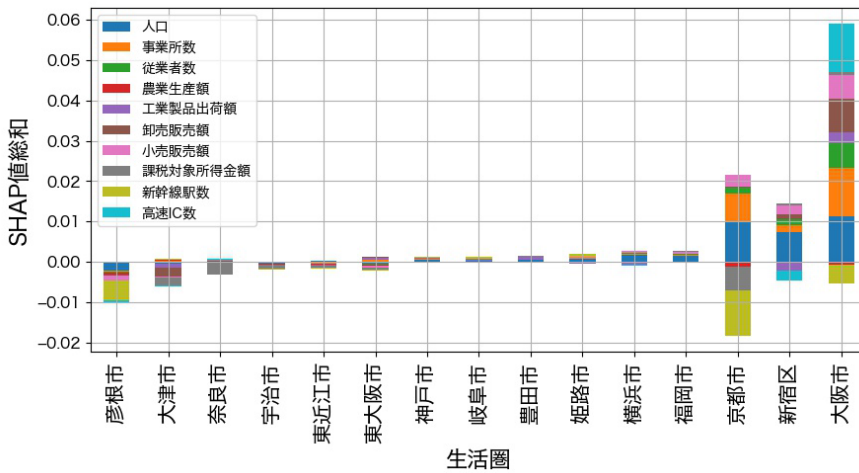


図-9 京都の人口に対する周辺 SHAP (time step 0 のケース, 絶対値 0.001 以上)

ルの精度が全体的に低いのは、1976年は交通網が小さく、交通ネットワークに関する入力情報が少ないことが原因と考えられる。

(2) self-SHAP

SHAP 値は、予測値への各ノード特徴量の貢献度を表す。なお、SHAP 値の信頼性は予測モデルの精度に依存するが、本モデルの決定係数は 0.9 以上であることから、SHAP 値を使うに値すると判断した。time step 1~5 を学習したモデルを使用して、time step 0 の 2 種類の SHAP 値を算出した。予測対象のノード自身の特徴量が予測結果に与える影響を self-SHAP と定義する。

図-7 は、人口とその人口の self-SHAP の関係を表す。生活圏の人口が多いほど、10 年後の人口予測に対して正の影響が大きいという相関関係が見られた。また、傾きの異なる 2 つの回帰直線が引けるように見える。東京、名古屋、大阪、神戸、広島、福岡など、東海道新幹線及び山陽新幹線上に位置する生活圏は、下の回帰直線付

近に位置する。これは、新幹線接続の影響による社会増加の効果が表れ、自然増加の効果が相対的に弱まったからだと考えられる。

(3) 周辺 SHAP

GCN では、ネットワークの繋がりを考慮しているため、各ノードに与える周辺ノードの影響が異なる。そこで、周辺の生活圏による影響を評価するときは、任意の生活圏に着目する必要がある。本研究では、京都府の生活圏の一つである京都と和歌山県の生活圏の一つである新宮に着目した。

図-8 は、Time step 6 における京都の将来人口に影響を与えた周辺の特徴量 (周辺 SHAP) を表す。縦軸は周辺 SHAP 値を表しており、横軸には、生活圏を記載している。なお、生活圏の代表都市を生活圏名として記載している。また、右に行くほど、SHAP 値の総和が大きくなるように記載している。京都は、大阪からの周辺 SHAP が最大であり、彦根や亀岡からは負の影響を受けてい

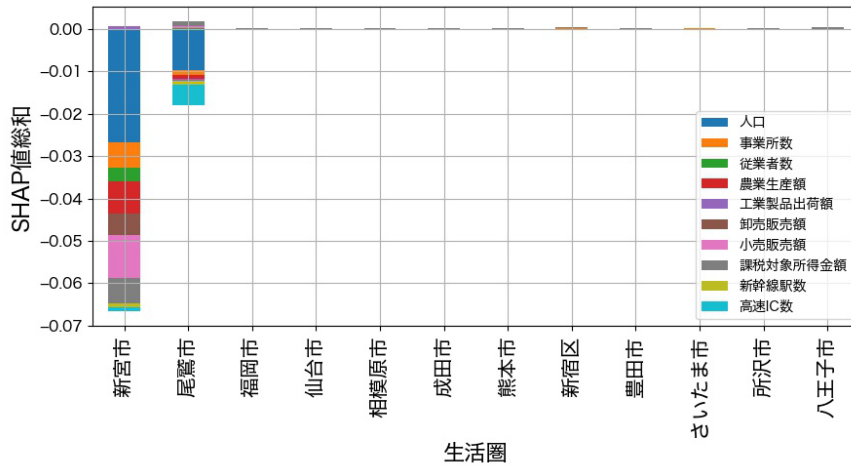


図-10 新宮の人口に対する周辺 SHAP (time step 6 のケース, 絶対値 0.0001 以上)

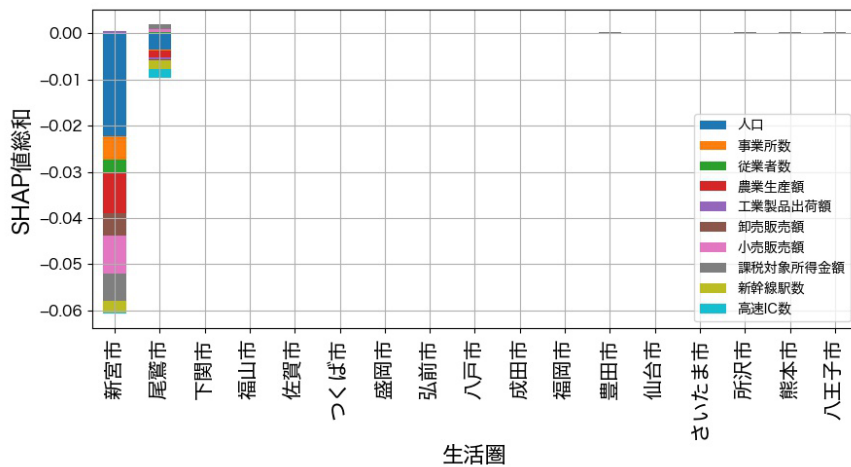


図-11 新宮の人口に対する周辺 SHAP (time step 0 のケース, 絶対値 0.0001 以上)

る。Time step 0 での予測を表した図-9 では、横浜、福岡、豊田など関西圏外の生活圏も多く見られるが、time step 6 での予測を表した図-8 ではそれらが減っている。この理由として、関西圏のネットワークが強化され、遠方の大都市より、付近の都市の影響のほうが大きくなったことが考えられる。

一方で、和歌山県の南部に位置する新宮生活圏の周辺 SHAP は、関東の大都市からのものが大きく、周辺都市からの値は小さいことが確認された (図-10, 図-11)。新宮は、京都と比較して、周辺生活圏との交通の接続性が弱いため、このような結果になったと考えられる。

(4) ケーススタディ

GraphExplainer⁴²⁾により算出した SHAP 値は、ノード特徴量による貢献度しか示さない。しかし、交通インフラを整備した場合、交通ネットワークが強化されることから、新幹線駅数や高速道路駅数の特徴量の貢献度だけでなく、エッジの貢献度を知ることも重要である。しかし、本モデルでは、エッジの貢献度を算出することができない

表-3 山陽新幹線と東北自動車道の整備効果

ノード特徴量	山陽新幹線	東北自動車道
人口	約 156 万人	約 10 万人
事業所数	約 6 万所	約 0.08 万所
従業者数	約 63 万人	約 0.4 万人
農業生産額	約 1470 億円	約 364 億円
工業製品出荷額	約 4.39 兆円	約 2.13 兆円
卸売販売額	約 2.36 兆円	約 0.066 兆円
小売販売額	約 1.50 兆円	約 0.036 兆円
課税対象所得金額	約 4500 万円	約 45 万円

ことから、現存している新幹線路線や高速道路を省いたケースを仮定し、予測結果を比較することで整備効果を確認した。

本研究では、1976年に山陽新幹線や東北自動車道が無かったケースを想定し、1986年の日本全体の人口や経済の予測を行った。山陽新幹線は、1972年に大阪-岡山間、1975年に岡山-福岡間が開通した。東北自動車道は1972年に開通し、1975年にさいたま市と仙台市を繋ぐ高速道路となっている。1986年において上記のインフラが整備

されているケースの予測値から、整備されていない予測値を引いた値を表-3 に示す。つまり、この値が山陽新幹線の東北自動車道のインフラの整備効果を表している。

山陽新幹線の全ての項目において、正の値が産出された。特に、工業製品出荷額、卸売販売額、小売販売額はいずれも 1 兆円以上の差が見られる。しかし、東北自動車道のケースでは、いくつかの項目で負の値も見られた。

5. おわりに

本研究では、交通ネットワークをグラフ構造で表現し、GCN を用いた都市発展の予測を行った。都市発展を表す指標として、人口、事業所数、従業者数、農業生産額、工業製品出荷額、卸売販売額、小売販売額、課税対象所得の 8 つを採用し、それぞれの指標に関して、10 年後の値を予測するモデルを構築した。

それぞれのモデルの精度を決定係数により確認したところ、いずれも 0.9 に近い値を示した。時系列分析により構築した AR モデルとの比較を行ったところ、GCN を用いて構築したモデルの方が、予測精度が高いことが確認された。また、構築したモデルを用いて、現存している新幹線路線や高速道路を省いたケースにおける人口や経済指標の予測を行った。その結果、1976 年において山陽新幹線が整備されていた場合と整備されいなかった場合を比較したところ、整備されていた場合において、人口や経済が大きいたことが確認された。一方で、東北自動車道の整備は、人口や経済に対して負の影響を与える結果となった。

本研究では、10 年後の人口や経済を予測するモデルを構築したが、インフラの整備の効果はより長期的に表れることが考えられることから、入力層と出力層の時間の間隔についても考慮していく必要がある。また、本研究では、10 年間隔の数組のデータを作成し、それらを独立した形で学習させてモデルを作成したが、これらを連続して考えることが望まれる。今後は、LSTM などの再帰型ニューラルネットワーク (Recurrent Neural Network: RNN) を検討していく必要がある。

参考文献

- 1) The World Bank: Annual Report, 2016
- 2) Baum-Snow, N., Did Highways Cause Suburbanization?, *The Quarterly Journal of Economics*, Vol.122 (2), pp.775-805, 2007.
- 3) Duranton, D., Turner, M.A., Urban growth and transportation, *The Review of Economic Studies*, Vol.79 (4), pp.1407-1440, 2012.
- 4) Atack, J., Bateman, F., Haines, M., & Margo, R. A. Did railroads induce or follow economic growth?: Urbanization and population growth in the American Midwest, 1850-1860. *Social Science History*, Vol.34 (2), pp.171-197, 2010.
- 5) Donaldson, D. Railroads of the Raj: Estimating the Impact of Transportation Infrastructure. *American Economic Review*, Vol.108 (4-5), pp.899-934, 2018.
- 6) Aljoufie, M., Zuidgeest, M., Brussel, M., & van Maarseveen, M. Urban growth and transport: Understanding the spatial and temporal relationship. *WIT Transactions on the Built Environment*, Vol.18 (1), pp.315-328, 2011.
- 7) 国土交通省. 社会資本整備審議会・交通政策審議会交通体系分科会計画部会専門小委員会「ストック効果の最大化に向けて：その具体的戦略の提言」, 2009. <https://www.mlit.go.jp/policy/shin-gikai/sogo08_sg_000220.html>(2022 年 3 月現在)
- 8) 佐藤啓輔, 小池淳司. SCGE モデルを活用した道路整備のストック効果最大化検討. *土木学会論文集 D3 (土木計画学)*, Vol.76 (2), pp.114-127, 2020.
- 9) Fujita M. *Urban Economic Theory*. Cambridge University Press, Cambridge, 1989.
- 10) Von Thunen, J. H. *Von Thunen's Isolated State*. Oxford, Pergamon Press, 1826.
- 11) Alonso, W. *Location and Land Use*. Cambridge, MA: Harvard University Press 7, 1964.
- 12) 野田順康. 都市化・都市成長の世界的潮流. *都市政策研究*, Vol. 18, pp.1-8, 2017.
- 13) Mills, E. S. An aggregative model of resource allocation in a metropolitan area. *American Economic Review Papers and Proceedings*, Vol.57 (2), pp.197-210, 1967.
- 14) Muth, R. F. *Cities and Housing: The Spatial Pattern of Urban Residential Land Use*. Chicago, University of Chicago Press, 1969.
- 15) Hansen, W. G. How accessibility shapes land use. *Journal of the American Institute of Planners*, Vol.25 (2), pp.73-76, 1959.
- 16) Acheampong, R. A., and Silva, E. Land use-transport interaction modeling: A review of the literature and future research directions. *Journal of Transport and Land Use*, Vol.8, pp.1-28, 2015.
- 17) Berling-Wolff, S., and Wu, J. Modeling Urban Landscape Dynamics: A Case Study in Phoenix, USA. *Urban Ecosystems*, Vol.7 (3), pp.215-240, 2004.
- 18) Rietveld, P., and Bruinsma, F., *Is Transport Infrastructure Effective?* Springer: Verlag Berlin Heidelberg, 1998.
- 19) Hanlon, W. W., and Hebllich, S. *History and urban economics*. *Regional Science and Urban Economics*, 2021.
- 20) Berger, T., and Enflo, K. Locomotives of local growth: The short-and long-term impact of railroads in Sweden. *Journal of Urban Economics*, Vol.98, pp.124-138, 2017.
- 21) Banerjee, A., Duflo, E., and Qian, N. On the Road: Transportation Infrastructure and Economic Growth in China. *Journal of Development Economics*, Vol.145, 2020.
- 22) Engelen, G. The Theory of Self-Organization and Modeling Complex Urban Systems. *European Journal of Operational Research*, Vol.37 (1), pp.42-57, 1988.
- 23) Cheng, J. *Modelling Spatial & Temporal Urban Growth*. International Institute for Geo-Information Sciences and Earth Observation, Enschede, 2003

- 24) Karlaftis, M.G., and Vlahogianni, E.I. Statistical methods versus neural networks in transportation research: differences, similarities and some insights. *Transportation Research, Part C. Emerging Technologies*, Vol.19 (3), pp.387–399, 2011.
- 25) DeTienne, K.B., Detienne, D.H., and Joshi, S.A. Neural networks as statistical tools for business researchers. *Organizational Research Methods*, Vol.6(2), pp.236–265, 2003.
- 26) Mills, E.S. A thematic history of urban economic analysis. *Brookings-Wharton Papers on Urban Affairs*, pp.1–52, 2000.
- 27) Chen, S., Claramunt, C., and Ray, C. A spatio-temporal modelling approach for the study of the connectivity and accessibility of the Guangzhou metropolitan network. *Journal of Transport Geography*, Vol.36, pp.12–23, 2014.
- 28) Kipf, T. N., and Welling, M. Semi-Supervised Classification with Graph Convolutional Networks. *ICLR*, 2017.
- 29) Laptev, N., Yosinski, J., Li, L., and Smyl, S. Time-series extreme event forecasting with neural networks at uber. *International Conference on Machine Learning*, Vol.34, pp.1–5, 2017.
- 30) Gao, H., Ji, S. Graph u-nets. *International conference on machine learning*. PMLR, p.2083–2092, 2019.
- 31) Xu, K., Li, C., Tian, Y., Sonobe, T., Kawarabayashi, K., and Jegelka, S. Representation learning on graphs with jumping knowledge networks, 2018.
- 32) Lundberg, S., and Lee, S. A Unified Approach to Interpreting Model Predictions. *Advances in Neural Information Processing Systems 30*, Curran Associates, Inc., 2017.
- 33) Molnar, C. 9.5 Shapley Values. *Interpretable Machine Learning: A Guide for Making Black Box Models Explainable*, 2022. (<https://christophm.github.io/interpretable-ml-book/shapley.html>) (2022 年 3 月現在)
- 34) Fukuda, K. GraphExplainer. Hitachi Internal Report, 2020. Unpublished.

(Received July 1, 2022)

(Accepted November 1, 2022)

FORECASTING URBAN GROWTH FROM TIME SERIES TRANSPORTATION NETWORK DATA USING GRAPH CONVOLUTIONAL NEURAL NETWORK

Tomoki OKAMURO, Jan-Dirk Schmöcker, Koji FUKUDA,
Satoshi NAKAO, Wenzhe SUN and Tadashi YAMADA