

Doc2Vecを用いた 飲食店向け不動産店舗賃料推定モデルの構築

Construction of Rent Estimation Model for Restaurants Using Doc2Vec

鶴山優季子¹ 諏訪博彦¹ 小川祐樹² 荒川豊¹ 安本慶一¹

Yukiko Tsuruyama¹ Hirohiko Suwa¹ Yuki Ogawa² Yutaka Arakawa¹ Keiichi Yasumoto¹

¹ 奈良先端科学技術大学院大学
¹ Nara Institute of Science and Technology
² 立命館大学
² Ritsumeikan University

概要

飲食店向け不動産物件の賃料は、不動産会社のベテラン営業職員が培ってきた経験や勘といった暗黙知に基づいて決定されている。賃料の決定要因としては、物件固有の情報である静的情報、物件周辺の情報である動的情報、物件の特徴などを含む潜在的情報が挙げられている。潜在的情報は、ベテラン営業職員による指標化が難しいとされる情報である。本研究では潜在的情報として、物件に付与されているキャッチコピーを用いる手法を提案する。キャッチコピーはDoc2Vecによりベクトル表現に変換し、ノイズを除去するため、品詞の選別を行なった。その結果、品詞の選別を行うことで推定精度が向上し、重回帰分析を用いた場合に、決定係数が0.611と最も高い値が得られた。

キーワード：機械学習，データマイニング，賃料推定，自然言語処理

Abstract

The rent for restaurants is determined based on tacit knowledge such as experience and intuition cultivated by veteran sales man of real estate companies. Determinants of the rent include static information that is specific to the property, dynamic information that is around the property, potential information that include features of the property. Potential information is difficult to index by veteran sales man. In this research, we propose a method using catch phrase given to property as potential information. We use the catch phrase that was vectorized by Doc2Vec, and in order to remove noise, we selected part of speech. As a result, the estimation accuracy was higher to select part of speech, and when multiple regression analysis was used, the highest coefficient of determination 0.611 was obtained.

Keywords: Machine Learning, Data Mining, Rent Estimation, Natural Language Processing

1 はじめに

機械学習の発展に伴い、不動産の分野においても物件の価格推定 [1][2][3] などの営業支援が活発に行われている。一方で、飲食店を対象とした不動産に着目すると、不動産会社のベテラン営業職員が長年培ってきた、経験や勘といった暗黙知に基づいた価格の決定がされている。この手法だと、賃料を決定している要因が明確ではなく、人によって賃料の価格が異なるといった問題や、新人営業職員への知識継承が困難であるといった問題が生じる。飲食店向け物件の選定には、

ガスや排気設備、視認性や通行量など固有の属性が存在するため、一般的な住宅の価格推定モデルの手法をそのまま取り入れることはできない。そのため、飲食店向け不動産物件の賃料決定は、ベテラン営業職員の暗黙知に基づいて行われている現状がある。

荒川ら [4] の研究では、野中ら [5] のSECIモデルの考え方をベースとし、飲食店向け不動産物件用の、ベテラン営業職員の暗黙知に基づいた賃料推定モデルを提案している。ここで荒川ら [4] は、飲食店向け不動産物件の賃料決定には、3つの要因があることを明らかにしている。その3つの要因を、静的情報、動的情

報, 潜在的情報と呼んでおり, 静的情報は物件固有の情報, 動的情報は物件周辺の情報, 潜在的情報は物件の特徴を含む情報として, 指標化している. このうち, 静的情報と動的情報は, ベテラン営業職員による指標化が容易な要素である. 具体的には, 静的情報は物件の坪数や階数など, 動的情報は物件の視認性や物件周辺の通行量などが挙げられている. 一方で潜在的情報は, ベテラン営業職員による指標化が難しい情報であり, 具体的な指標は述べられていない. そこで荒川らは, 潜在的情報として物件の付与されているキャッチコピーを用いる手法を提案している. キャッチコピーを形態素解析し, 名詞と形容詞を抽出することで, これらの単語が賃料に与える影響を調査している. しかしこの手法では, 文脈が考慮されない, また同一の単語でなければ検出されず, 曖昧さが無いといった課題がある.

もう一つの先行研究 [6] では, 荒川ら [4] の課題であるキャッチコピーの文脈や曖昧さを考慮するため, キャッチコピーに Doc2Vec[7] を用いる手法を提案している. Doc2Vec[7] は, 機械学習を用いた, 任意の長さの文書を任意の次元数にベクトル化する技術である. この技術を用い, キャッチコピー全文をベクトル化したものをモデルに組み込んでいる. その結果, ベクトル化したキャッチコピーを取り入れた場合の方が, 取り入れない場合よりも推定精度が良くなることが得られている. しかしその推定精度は, 品詞を選別している荒川らの推定精度よりも低くなっており, キャッチコピーの全文を組み込んだことによる多数のノイズが含まれている可能性がある.

本研究では, 飲食店向け不動産物件の賃料推定システムを構築することを目的とする. また賃料推定モデルの構築は, 先行研究 [4][6] をベースとして行う. これらの先行研究には, キャッチコピーを賃料推定モデルに組み込む際の課題がある. これらを, 課題 1 として文脈や曖昧さの考慮, 課題 2 としてノイズの除去とする. 本研究では, 課題 1 を解決するために Doc2Vec 技術を, 課題 2 を解決するために品詞の選別を行う手法を提案し, その評価を行う. 賃料推定モデルの構築には, 重回帰分析とランダムフォレストを用い, 交差検証により評価を行う. データセットは, 契約が成立した 184 件の物件を対象としている. その結果, キャッチコピーの品詞を選別した方が, 全文を組み込んだ場合よりも決定係数が高く, 平均二乗誤差が低くなること得られた. また, キャッチコピー中の名詞, 動詞, 形容詞を抽出し, 重回帰分析によりモデルを構築した場合に決定係数が最も高くなり, その値は 0.611 であった.

本稿の構成について述べる. 2 章では, 賃料推定に関する研究について述べる. 3 章では先行研究と, それに基づいて提案する本研究の手法について述べる. 4 章では提案した手法の評価を行い, 5 章では考察, 6 章でまとめとする.

2 関連研究

不動産物件の価格推定に関する研究について述べる. Wu ら [1] は台湾において, 住宅選定に影響があるとされている風水を組み込んだ, 住宅物件の価格推定モデルを提案している. モデルの構築には, バック

プロバゲーションニューラルネットワーク, ファジーニューラルネットワーク, また彼らが独自に開発したハイブリッド遺伝子ベースのサポートベクター回帰からなる複数のアルゴリズムを用いて比較を行なっている. その結果, いずれの手法においても風水を考慮した方がより良い推定精度が得られている.

三浦ら [2] は, インターネット上にある膨大な情報から地域ごとの評判に関する変数を抽出し, 不動産側が所有しているデータベース上の情報と組み合わせる方法を提案している. 評判に関する変数は, インターネット上で地名を検索し, 形態素解析により抽出されている. モデルの構築にはヘドニック分析法が用いられ, 不動産側のデータベースのみで推定した場合は, 決定係数が 0.528, インターネット上の情報を組み合わせた場合は, 決定係数が 0.560 となる結果が得られている.

Chiarazzo ら [3] は, 交通システムと地域ごとの環境の質が不動産の価格に影響を与えていると考え, 人口ニューラルネットワークを用いて検証している. その結果, 全 42 要素ある属性のうち, 8 番目に環境汚染に関する属性が, 15 番目付近に交通に関する属性があげられることを明らかにしている.

このように, 賃料推定に関する研究は多数存在するが, 対象とする不動産は一般住宅であり, 選定の基準が異なる飲食店用の不動産には同じ手法を適用することはできない. 本研究では, これらの関連研究を参考にしながら, 飲食店向け不動産用の賃料推定モデルを構築する.

3 提案手法

本研究では, 2 つの先行研究 [4][6] に基づき, 飲食店向けの賃料推定モデルを構築する. 本章では先行研究のモデル構築方法と課題, またその課題を解決するための本研究の手法について述べる.

3.1 先行研究

本節では, 本研究の基盤となる先行研究について述べる. 荒川ら [4] の研究では, 野中ら [5] の SECI モデルに基づき, 飲食店向け不動産物件の賃料推定モデルを構築している. SECI モデルとは暗黙知を伝承するための知識創造モデルであり, 言語化できない知識を暗黙知, 言語化できる知識を形式知と呼んでいる. 野中ら [5] によると, 知識創造は (1) 共同化→(2) 表出化→(3) 連結化→(4) 内面化→(1) 共同化といったサイクルを繰り返すことで可能になるとしている. SECI モデルは知識を表出化して形式知にし, それらを連結化することで, 概念として継承が可能であることを示している. このモデルに基づく, ベテラン営業職員の暗黙知は, 表出化して形式知にし, それらを連結化することで, 知識として継承が可能となる (図 1).

荒川らは SECI モデルに基づき, 暗黙知の表出化を行うため, ベテラン営業職員に対し詳細なインタビューを行った. その際に得られた特徴量の取捨選択を行なった結果, 賃料の決定要因として, 物件の固有の情報である静的情報, 物件周辺の情報である動的情報, またベテラン営業職員による指標化が困難である

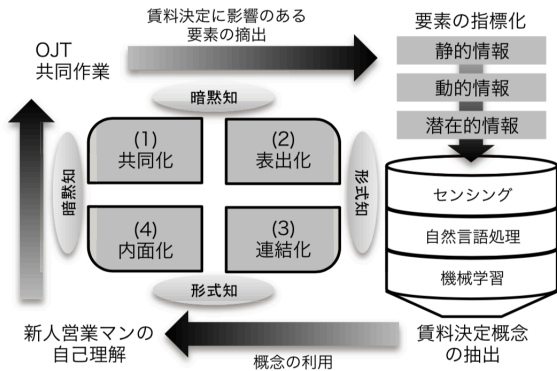


図 1: 先行研究のアプローチ [4]

潜在的情報の3つの要因が得られている。次に、これら3つの要因の具体的な特徴量について述べる。

3.1.1 静的情報

静的情報とは、年数が経っても基本的には変化しない、物件固有の情報である。ベテラン営業職員への詳細なインタビューにより、静的情報には、物件の坪数、駅までの時間、階数、居抜きが挙げられている。居抜きとは、物件に付帯しているガス、テーブル、カウンターなどの設備を表しており、居抜きであれば1、そうでなければ0として指標化している。これらの特徴量は、不動産会社が運営しているウェブページ上などで公開されているため、改めて指標化を行う必要はない。

3.1.2 動的情報

動的情報とは、物件周辺の日々変化する情報である。インタビューにより、動的情報には、視認性、通行量、地域ポテンシャルが挙げられている。視認性と通行量は、ベテラン営業職員2名による5段階評価の平均値をとり、視認性×通行量を物件の見つけやすさとして指標化している。積をとる理由は、例えば視認性は高いが通行が全くないといった場合に、見つけやすさを低く評価したいためである。地域ポテンシャルは、対象物件の最寄駅の平均坪単価に物件の坪数をかけた値としている。

3.1.3 潜在的情報

潜在的情報とは、ベテラン営業職員による指標化が困難な情報である。そのため、インタビューにより具体的な指標は得られていない。しかし、先行研究では潜在的情報として、物件に付与されているキャッチコピーに着目している。キャッチコピーには、具体例として、「大通り交差点すぐそば!」、「焼肉店居抜き店舗」、「視認性良好、角地店舗です」などがある。

次に、2つの先行研究によるキャッチコピーの解析方法について述べる。

(1) 単語ごとの正負判定

1つ目の先行研究 [4] では、キャッチコピーに含まれる名詞、形容詞を形態素解析により求め、それらの

単語が賃料に対して正の影響を与えるか、負の影響を与えるかを算出している。上記で挙げた例では、賃料に対して正の影響を与える単語として「大通り」、賃料に対して負の影響を与える単語として「焼肉」、「角地」が得られている。

しかしこの手法では、キャッチコピーの文脈が考慮されていないといった問題や、同一の単語でなければ検出できないといった問題がある。従って、キャッチコピーの文脈や曖昧さを考慮できるモデルが求められる。これを課題1とする。

(2) 全文のベクトル化

一方、2つ目の先行研究 [6] では、課題1を解決するため、Doc2Vec [7] を用いた手法を提案している。Doc2Vec [7] とは、機械学習を用いた任意の長さの文書をベクトル化することができる技術である。類似の手法である Word2Vec は、具体例として、「王様」 - 「男」 + 「女」 = 「女王様」といった、概念の足し引きをすることができる。またベクトル表現であることから、文書の類似度を算出することができる。Doc2Vec は、Word2Vec 技術を文書レベルに拡張した技術であり、単語の意味的表現を文書レベルで学習することができる。この Doc2Vec の技術を用い、1件のキャッチコピーを1文書と見なし、ベクトル化を行なっている。

24000件以上のキャッチコピーを用い、全文をベクトル化したものをモデルに組み込んだ結果、1つ目の先行研究 [4] よりも、精度が下がるという結果が得られている。これは、キャッチコピー全文をベクトル化したため、名詞と形容詞のみを抽出していた先行研究 [4] よりも、ノイズが多数含まれた可能性が考えられる。従って、それらのノイズの除去を行う必要がある。これを課題2とする。

3.2 本研究

ここでは、先行研究の課題1と課題2を解決するため、本研究の手法について述べる。課題1を解決するため、引き続き Doc2Vec 技術を用いる。また課題2を解決するため、キャッチコピー中の品詞の抽出と取捨選択を行う。

3.2.1 課題1: Doc2Vec

先行研究 [6] に引き続き、本研究ではキャッチコピーの解析に Doc2Vec 技術 [7] を用いる。Doc2Vec [7] は、機械学習を用いた、任意の長さの文書を任意の次元数にベクトル化することができる技術である。ベクトル化することで、類似度の高い単語や文書を検索することができ、曖昧さを持たせることができる。

24000件以上の文書を用いて、Doc2Vec のモデルを構築する。モデルを構築する際に主に使用されるパラメータとして、ベクトルの次元数、ウィンドサイズ、最小カウント数がある。ウィンドサイズは、現在の単語と予測する単語の最大の距離、最小カウント数は、出現回数が低いものを無視する閾値である。これらの値は、予備解析を行い、ベクトルの次元数は1、ウィンドサイズは3、最小カウント数は10と決定した。

また、このモデルを構築する際に、キャッチコピーの全文を組み込むとノイズが多く含まれる可能性が生じている。そのため、品詞の抽出と取捨選択を行い、モデルを構築する。

表 1: キャッチコピーから抽出された品詞

品詞	出現回数	単語の例
名詞	193325	店舗 雰囲気 常連客 ビル
記号	70496	、。」「
助詞	30208	の で に を
動詞	11207	落ち着き 面する 有り
助動詞	10071	です ます な
接頭詞	7807	現半前貸
形容詞	2240	暖かい 白い 高い 多い
副詞	1554	こじんまり 絶えず すぐ
連体詞	149	大きな この そんな
フィラー	10	(誤解析)
感動詞	1	(誤解析)

3.2.2 課題 2: 品詞の選別

本研究では、キャッチコピーに含まれている品詞と、それらの品詞が賃料に対しどのような影響を与えているかを調べる。

24000 件以上の物件のキャッチコピーから品詞を抽出した結果、表 1 が得られた。このうち、フィラーと感動詞は形態素解析の際の誤りであることを確認している。また名詞は、全てのキャッチコピーに最低でも 1 回は含まれていること、また名詞以外は、含まれていないキャッチコピーがあることも確認した。

表 1 の単語の例を確認すると、文書中で意味を持つ品詞として、名詞、動詞、形容詞、副詞、連体詞があげられる。これらのことから、キャッチコピーから抽出する品詞の組み合わせを、表 2 のように決定した。名詞以外を含まないキャッチコピーを確認したため、品詞を抽出する際は、必ず名詞を含むこととした。Doc2Vec によるベクトル化の際に、文書が空であることを防ぐためである。

3.2.3 提案モデルの構築

指標化された構成要素と賃料の関係を連結化するため、機械学習法を用いて賃料推定モデルを構築する。本研究では、機械学習法として重回帰分析 (Linear Regression, LR) とランダムフォレスト (Random Forest, RF) を用いた。ランダムフォレストの木の数は、予備解析により 20 と決定した。モデル構築には、契約が成立した物件のうち、賃料が 60 万円以下の 184 件の物件を対象としている。高額物件は特殊なものが多く、外れ値になりうるため、推定の対象外とした。モデルを構築する際の要素には、静的情報、動的情報、潜在的な情報を用いた。具体的には、静的情報として坪数、駅徒歩時間、居抜き、階数、動的情報として、地域ポテンシャル、通行量×視認性を用いた。また潜在的な情報として、品詞を選別し、Doc2Vec によりベクトル化したキャッチコピーを用いた。

推定モデルの精度の評価をするため、決定係数 R^2 と平均二乗誤差 $RMSE$ を用いる。決定係数の求め方を式 (1) に、平均二乗誤差の求め方を式 (2) に示す。式中の y_i は真値、 \hat{y}_i は推定値、 \bar{y} は真値の平均値を表す。また n は物件数であり、今回は $n = 184$ である。

表 2: 抽出する品詞の組み合わせ

番号	品詞の組み合わせ
1	名詞
2	名詞 動詞
3	名詞 形容詞
4	名詞 副詞
5	名詞 連体詞
6	名詞 動詞 形容詞
7	名詞 動詞 副詞
8	名詞 動詞 連体詞
9	名詞 形容詞 副詞
10	名詞 形容詞 連体詞
11	名詞 副詞 連体詞
12	名詞 動詞 形容詞 副詞
13	名詞 動詞 形容詞 連体詞
14	名詞 動詞 副詞 連体詞
15	名詞 形容詞 副詞 連体詞
16	名詞 動詞 形容詞 副詞 連体詞

決定係数は、モデルの当てはまりの良さを表し、最も良いスコアは 1 である。一方で平均二乗誤差は、モデルの当てはまりの悪さを表し、最も良いスコアは 0 である。

$$R^2 = 1 - \frac{\sum_{i=0}^n (y_i - \hat{y}_i)^2}{\sum_{i=0}^n (y_i - \bar{y})^2} \quad (1)$$

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=0}^n (y_i - \hat{y}_i)^2} \quad (2)$$

4 賃料推定モデルの評価

機械学習法により、構築した賃料推定モデルの評価を行うため、品詞の選別と先行研究との比較をそれぞれ行う。

4.1 品詞別の比較

表 3 に、品詞の組み合わせ別の推定結果を示す。また比較のため、キャッチコピーの品詞を選別せず、全文をモデルに組み込んだ場合の推定結果も示す。LR の場合は、キャッチコピー全文を組み込んだ場合よりも、品詞を選別した方が決定係数が高くなり、平均二乗誤差が小さくなるという結果が得られた。一方で RF の場合は、決定係数、平均二乗誤差の両方において、品詞を選別したことによる差は得られなかった。組み込んだ品詞別に比較すると、決定係数が最も高くなったのは、データ番号 6 の名詞、動詞、形容詞の組み合わせであった。しかし、その他の品詞の組み合わせと比較すると、大差は得られなかった。

表 3: 品詞の組み合わせ別推定結果

データ番号	決定係数		平均二乗誤差	
	LR	RF	LR	RF
全文	0.548	0.559	63,643	69,882
1	0.607	0.561	63,554	68,091
2	0.607	0.552	63,525	68,995
3	0.608	0.542	63,464	68,566
4	0.607	0.549	63,556	70,740
5	0.605	0.532	63,827	67,148
6	0.611	0.557	63,357	71,152
7	0.606	0.583	63,801	69,062
8	0.608	0.592	63,698	68,058
9	0.605	0.551	63,845	71,566
10	0.606	0.557	63,357	71,152
11	0.608	0.584	63,655	65,047
12	0.607	0.568	63,512	69,820
13	0.609	0.568	63,513	69,178
14	0.607	0.573	63,775	66,164
15	0.607	0.555	63,692	69,421
16	0.609	0.536	63,502	69,263

表 4: 名詞の細分類とその例

細分類	出現回数	単語の例
一般	83521	徒歩 居抜き 階
固有名詞	49164	新宿 靖国通り 牛スジ
接尾	21964	沿い 階 丁目
サ変接続	17832	営業 出店 相談
数	10927	1 5 2 0
形容動詞語幹	6429	シンプル 綺麗 減多
副詞可能	2118	夜 週末 今
非自立	917	方 ため 以上
代名詞	418	何 あなた そこ
ナイ形容詞語幹	18	限り 間違い 問題
接続詞的	8	兼
動詞非自立的	6	ご覧
特殊	3	そう

4.2 先行研究との比較

本手法において、最も決定係数が高かったデータ番号 6 (表 3) の推定結果を、図 2 に示す。また、先行研究 [4] の推定結果も図 3 に示す。横軸が実際の賃料、縦軸が推定した賃料を表している。中央の実線が理論線を表しており、その上下にある実線は理論線の $\pm 30\%$ を表している。

本手法では、重回帰分析によりモデルを構築しており、抽出している品詞は名詞、動詞、形容詞、決定係数は 0.611、平均二乗誤差は 63,357 である。先行研究では、ランダムフォレストによりモデルを構築しており、抽出している品詞は名詞、形容詞、決定係数は 0.738、平均二乗誤差は 52,494 である。

先行研究では、賃料が低い場合に高く見積もってしまう傾向がある一方で、本手法では、全体的にばらつきが確認された。

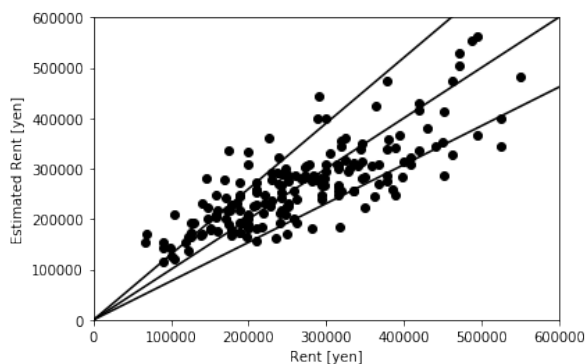


図 2: 本手法の推定結果

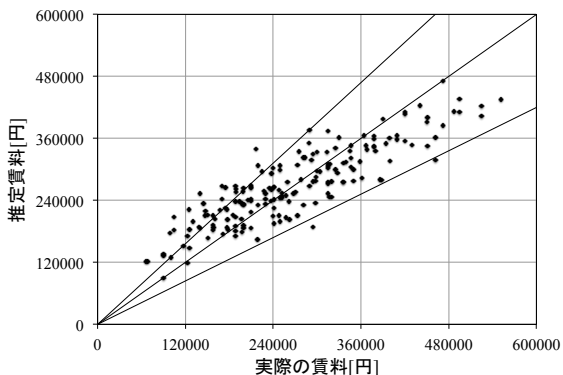


図 3: 先行研究の推定結果 [4]

5 品詞別の比較に関する考察

ノイズを除去するため、キャッチコピーに含まれている品詞を抽出し、選別を行なった。表 1 により、キャッチコピーには 10 種類程度の品詞が含まれていることが確認されたが、そのうち、名詞、動詞、形容詞、副詞、連体詞の 5 種類の品詞のみを抽出し、解析を行なった。しかし、決定係数や平均二乗誤差において、品詞の組み合わせによる差はほとんど得られなかった (表 3)。

今回、Doc2Vec によるモデル構築の際に用いたキャッチコピーには、必ず名詞が含まれており、その他の品詞は含まれていないものもあった。そのため、Doc2Vec によりベクトル化する際には、必ず名詞を含むこととした。表 1 を確認すると、24000 件以上のキャッチコピーには、20 万回近くの名詞が出現していることがわかる。一方でその他の品詞は、動詞で 1 万回程度、また連体詞で 150 回程度であり、名詞の出現回数名詞と比較するとかなり少ないことがわかる。このことから、表 3 の推定精度に変化がなかったのは、名詞の出現頻度の多さにより、その他の品詞の影響が小さく

なってしまった可能性が高いと考えられる。

また20万回近く出現した名詞の細分類を確認したところ、表4に示すような結果が得られた。出現頻度が多いものから順に見ると、一般や固有名詞、接尾が挙げられており、これらの物件の場所や雰囲気などを多く含んでいることがわかる。名詞の一般に含まれている単語のうち、更に出現頻度の高い単語を見ると、上位5以内に「店舗(10709回)」、「徒歩(8910回)」、「居抜き(6447回)」、「階(3852回)」が挙げられている。これらは、既に静的情報としてモデルに組み込まれている要素であり、潜在的情報として組み込む場合は、こういった情報を除く必要があると考える。また先行研究と比較すると、本手法はばらつきが大きく、全体的に低く見積もっていることが確認された。これも、潜在的情報として抽出すべき情報に、静的情報が含まれていたことが原因である可能性が高い。

6 おわりに

本研究では、飲食店向け不動産物件の賃料推定システムを構築することを目的とし、先行研究をベースとして賃料推定モデルを構築した。賃料の決定要因としては、物件固有の情報である静的情報、物件周辺の情報である動的情報、また不動産会社のベテラン営業職員による指標化が難しい潜在的情報の3つが挙げられた。指標が挙げられなかった潜在的情報については、物件に付与されているキャッチコピーを用いた。キャッチコピーの文脈や曖昧さを考慮するため、Doc2Vecによりベクトル化を行い、またノイズを除去するため、ベクトル化する際の品詞を選別した。モデルの構築には、重回帰分析とランダムフォレストの2つの機械学習法を用いた。その結果、重回帰分析でモデルを構築した場合は、キャッチコピーの全文を組み込むよりも、品詞を抽出した方が推定結果が良くなることが得られた。また、名詞、動詞、形容詞を抽出した際に最も決定係数が高く、平均二乗誤差が低くなるという結果が得られた。

しかし、その他の組み合わせと比較すると、決定係数に大きな差は得られなかった。今回は、品詞を抽出する際は必ず名詞を含むこととしているが、その名詞の出現頻度が、他の品詞よりもはるかに多かったため、他の品詞の影響が小さくなったと考えられる。

また、先行研究[4]の結果と比較すると、本手法ではばらつきが大きく、実際の賃料よりも低く見積もっている傾向があった。キャッチコピー中に含まれている名詞を確認すると、「徒歩」や「居抜き」、「階」といった、静的情報としてモデルに組み込まれている情報が含まれていることが確認された。潜在的情報として組み込む場合は、これらの情報は除く必要があると考えられる。

今後の課題としては、除ききれていないノイズの除去が挙げられる。また、より精度の高い賃料推定モデルの構築のため、築年数や前借主の営業年数など、新たな指標、変数の取り入れなどが考えられる。

参考文献

- [1] Wu, C. H., Li, C. H., Fang, I. C., Hsu, C. C., Lin, W. T., and Wu, C. H.: Hybrid genetic-based support vector regression with shui theory for appraising real price, *First Asian Conference on Intelligent Information and Database Systems*, pp.295–300 (2009)
- [2] Takahumi Miura, and Yasuhi Asami: Hedonic Analysis for Estimation of Condominium Rent Utilizing WEB Information, *Procedia Social and Behavioral Science* 21, pp.147–156 (2011)
- [3] Chiarazzo, V., Caggiani, L., Marinelli, M., and Ottimanelli, M.: A Neural Network based model for real estate price estimation considering environmental quality of property location, *Transportation Research Procedia*, Vol.3, pp.810–817 (2014)
- [4] 荒川周造, 諏訪博彦, 小川祐樹, 荒川豊, 安本慶一, 太田敏澄: 暗黙知センシングに基づいた飲食店向き不動産店舗の賃料推定, 人工知能学会全国大会論文集 2017年度人工知能学会全国大会(第31回)論文集, 一般社団法人人工知能学会(2017)
- [5] Nonaka, I., and Takeuchi, H.: The knowledge creation company: how Japanese companies create the dynamics of innovation (1995)
- [6] 鶴山優季子, 諏訪博彦, 小川祐樹, 荒川豊, 安本慶一: キャッチコピーを用いた飲食店用不動産賃料推定モデルの改善, 社会情報システム学シンポジウム(2019)
- [7] https://deepage.net/machine_learning2017/01/08/doc2vec.html: (最終閲覧日 2019-02-22)