

通行量センシングと機械学習に基づく飲食店用不動産賃料推定

荒川周造¹ 諏訪博彦¹ 小川祐樹² 荒川豊¹ 安本慶一¹ 太田敏澄³

概要：飲食店用不動産の賃料決定は、ベテラン営業マンが培った勘や経験からなる暗黙知を元になされており、他の営業マンに継承するための形式知化（表出化）が課題となっている。そこで根拠に基づく賃料推定システムを構築するために、立地条件などの静的な情報に加え、店舗周辺の通行量などの動的情報を考慮したシステムを構築する。本発表では、機械学習を用いて賃料推定を行い、推定精度向上のための特徴量の検討を実施したので報告する。

The real estate rent estimation for restaurants based on human traffic sensing and machine learning

SHUZO ARAKAWA¹ HIROHIKO SUWA¹ YUKI OGAWA²
YUTAKA ARAKAWA¹ KEIICHI YASUMOTO¹ TOSHIZUMI OHTA³

1. はじめに

飲食店用不動産物件における賃料決定は、従来、豊富な経験を持つベテラン営業マンが培ってきた勘や経験からなる暗黙知に基いてなされてきた。しかしながら、経験の浅い新人営業マンにおいてもベテラン営業マンと同様に賃料決定できるようになることが望まれており、他の営業マンへの暗黙知の継承が課題となっている。野中らが提案している SECI モデル[1]によれば、この課題は、暗黙知を表出化して形式知にし、さらに連結化することで、概念として伝承が可能となる。そこで、本研究ではこれまで具体的な根拠が無い暗黙知を形式知化することで、根拠に基づく賃料推定システムを構築することを目的とする。

不動産賃料推定に関する多くの先行研究では、物件の立地条件や周囲の住環境といった静的な情報を入力として、価格推定を実施しているが、それらの多くは個人向けマンションなどの住宅物件を対象とするものである。本研究で取り扱う物件は、飲食店用不動産向けに限定されており、その特殊性から、店舗前の通行量や店舗の雰囲気といった、これまで考慮されてこなかった動的情報も重要なパラメータである。筆者らが過去に取り組んできた重回帰分析による賃料推定[2]では、通行量のパラメータは営業マンの感覚を元に大雑把な値として入力されている。こうした暗黙知を根拠あるものにするため、通行量センサを用いたセンシングシステムを提案する。通行量を感覚値ではなく定量値として入力することによって推定精度の向上が期待される。また、機械学習を用いることによる推定精度向上を図る。

本研究における具体的な課題は、①通行量などの動的情報の取得、②取得データに基づく賃料推定精度の向上である。課題①を解決するために、通行量センシングシステムを提案する。提案する通行量センシングシステムは、実際の物件に設置することを念頭に、目につきにくい小型なもので、かつ低消費電力であることが求められている。従来のセンサの多くは、電源を必要としており、設置場所が制限されてきた。また、ソーラーパネル付きセンサもあるが、十分な電力量を得るには大型になるなどの問題があった。そこで、市販のエナジーハーベスティング無線モジュールをベースに要求に合ったセンサを新たに開発することで、小型・低消費電力なシステムの実現を可能にしている。また、設置の自由度を高くするため、センサ本体とデータ収集端末を分離して実装した。さらには、機械学習に用いる学習データを容易に作成可能にするには、デバイスを物理的に接続せずにデータ収集が可能であることが望ましい。そこで、収集したデータをクラウド上にアップロードすることによって、遠隔地でもデータ収集を容易に実現可能にしている。

課題②を解決するために、賃料推定手法として機械学習を用いる。機械学習手法としては、ランダムフォレスト[3]の回帰分析などを用いる。また、特徴量としては、地域ポテンシャルとしての駅推定賃料、店舗の視認性、店舗周辺の通行量、居抜き（机や椅子、カウンター、冷蔵庫などの什器が付帯した状態）の有無、階数や坪数などを用いる。特にこれまでの取り組みから、通行量を用いることによって推定精度が向上することが確認できている。これらの手法を比較検討することにより、より精度の高い賃料推定モデルを構築する。

本稿では、主に課題②の解決に着目し、賃料推定にランダムフォレストを用いることで、従来の重回帰分析のもの

¹ 奈良先端科学技術大学院大学
Nara Institute of Science and Technology

² 立命館大学
Ritsumeikan university

³ 行政情報システム研究所
The Institute of Administrative Information Systems

と比較し、内挿同士での精度向上を確認している。また、3-fold 交差検証においても同等の精度で推定可能なことを確認している。加えて、既存の特徴量の影響度合いを、各特徴量を省いて推定した際の平均二乗誤差 (RMSE) に基づき評価している。その結果、精度低下を招いていた「坪数」の項目を抽出でき、この特徴量を削減することで精度の向上が確認できている。同時に、影響度が比較的低かった「駅徒歩時間」と「階数」の項目について、積を取ることで推定誤差を減少させている。さらには新たな特徴量として店舗の「導入力」を追加し、精度向上が確認された。

以下に本稿の構成について述べる。まず、第2章では本研究に関連する研究や取り組みについて説明し、本研究の位置づけを明確にする。第3章では、店舗の動的情報を取得するため試みとして開発した通行量センシングシステムについて説明する。第4章では、機械学習に基づく賃料推定について述べ、営業マンの感覚データを用いて推定精度向上に向けた検討を行う。第5章はまとめとする。

2. 関連研究

本研究に関連する研究として、不動産賃料 (価格) 推定、通行量センシングへの取り組みが挙げられる。また、概念的な課題である暗黙知の形式知化に関する取り組みが挙げられる。以下にそれぞれについて説明し、本研究の位置づけを明確にする。

2.1 不動産賃料 (価格) 推定への取り組み

データに基づく賃料推定に関する取り組みとして、Victor[4]らは決定木とニューラルネットワークに基づく不動産価格推定を実施している。この研究では、米国キング郡を対象とし、推定の外れ値を取り除くために取引価格の上位・下位 5%の物件を除外して推定を行い、不動産価格帯ごとに分割して推定誤差を評価している。その結果、ニューラルネットワークを用いた方が良好な結果が得られることを明らかにしている。

また、Chih-Hung[5]らは、台湾での住宅選定に関わりの深い風水に着目し、風水を考慮した不動産価格推定に取り組んでいる。多くの不動産価格推定に用いられている建物そのものの属性に加え、風水におけるタブーを学習データセットに設定している。彼らはバックプロパゲーションニューラルネットワーク (BPN) とファジーニューラルネットワーク (FNN)、そして独自に開発したハイブリッド遺伝ベースのサポートベクター回帰 (HGA-SVR) からなる複数のアルゴリズムで推定した結果、いずれの手法においても風水のタブーを考慮した方がより良好な推定結果となることを示している。また、それぞれのアルゴリズムの性能を評価し、HGA-SVR、FNN、BPNの順で優れていることを明らかにしており、最も優れた HGA-SVR の平均絶対パーセント誤差 (MAPE) は 4.79%を示している。さらに、

BPN を用いて最も価格推定に影響のある属性を土地の大きさであることをつきとめた。また、風水におけるタブーでは、トイレ、窓、家の外観、ドアの順に住宅価格に影響していることを明らかにしている。本研究で取り扱う飲食店用不動産物件においても、風水を気にされる契約者の方が存在してもおかしくないため、こうしたアプローチを考慮することも選択肢の一つとして考えられる。

さらに、Vincenza[6]らは、ターラント市 (イタリア) における不動産価格に、交通システムと地域ごとの環境の質が深く関係していると考え、人工ニューラルネットワーク (ANN) を用いて不動産価格推定を行っている。データセットには、立地条件や建物の構造に加え、駅やバス停までの距離などの交通に関する属性や、SO₂、NO_x、NO、NO₂、CO、PM₁₀のそれぞれの値と最大値といった環境汚染に関する属性が含まれている。この研究でもデータセットの各属性の感度分析を行った結果、不動産価格に最も影響している要因として、ビーチへの近さや、ガレージおよびテラスの有無が求められた。環境汚染に関する項目で最も重要なものとしては、SO₂の最大値が挙げられ、全 42 ある属性の内の 8 番目であった。また、交通に関する属性も 15 番目付近に位置しており、影響があることを示している。さらには、環境汚染に関する属性を除いて学習したところ、大きな精度低下は見られなかったことから、不動産物件から工業地帯への距離を用いることで、環境汚染の属性を代用できることを明らかにしている。

以上 3 件の関連研究では、それぞれの地域に特化した手法を採用している点で共通している。地域ごとに賃料決定に影響のある属性に違いが見られるため、どの属性が重要であるかを吟味する必要がある。これらの研究は全て海外不動産を対象とした研究であったが、近年では国内においても不動産賃料推定をサービスとして提供している事例も存在している。例えば、株式会社リブセンスでは、未来型不動産サービス「IESHIL (イエシル)」[7]において、ビッグデータを活用した不動産物件のリアルタイム査定を実施しており、売買時の不動産価格や価格推移、推定賃料を提供している。また、ソニー不動産株式会社も、ソニー株式会社 R&D の持つディープラーニング (深層学習) 技術と、ソニー不動産が有する不動産査定ノウハウや不動産取引特有の知識を組合せた「不動産価格推定エンジン」[8]を開発し、不動産の成約価格を推定している。しかしながら、前述したいずれの先行研究においても、一戸建て、マンション、アパートといった住宅用不動産物件を取り扱っており、飲食店用不動産物件を扱う本研究とは対象が異なっている。

2.2 通行量センシングへの取り組み

次に通行量センシングに関連する研究について述べる。多くの先行研究では、通行量を測定する手法として、カメラ

ラを用いた取り組みがなされている。特に近年では、システムの複雑化が進んでおり、具体的な例として Calro[9]らは、ステレオカメラを用いた人数カウントシステムを構築している。また、Zeng[10]らは、複数のカメラを用いてシステムを提案している。しかしながら、カメラを用いたこれらのシステムの場合、設置やプライバシー面でのハードルが高く、飲食店用不動産物件には適していない。設置面では消費電力が高いカメラの場合は電源線が必要となる。また、カメラを設置すること自体を嫌うオーナーの方も多く、この手法を本研究で導入することは難しいといえる。

そこで、センサを用いたアプローチが注目されている。Wahi[11]らは、受動型赤外線 (PIR: Passive Infrared Ray) センサを用いてオフィス内の人数を推定する手法について提案し、シミュレーションに基づいてアルゴリズムの性能を評価している。PIR センサを用いるメリットとして、センサの単価がカメラと比較して安価であり、プライバシー面の問題にも配慮できていることが挙げられる。この研究においても、プロトタイプとして、エナジーハーベストセンサを用いたセンサノードを作成し、センシングシステムの検証を実施しており、本研究との関連性は高いといえる。また、Fang[12]らも複数の PIR センサからの連続したパルス出力を元に、サポートベクターマシン (SVM) を用いた人数カウント手法を提案している。

これらを踏まえ、本研究においては、PIR センサを用いたエナジーハーベストセンシングセンサを新たに開発することで、店舗への設置に適したシステム構築を実施する。カメラと異なり個人の識別までは対応できないところが、逆にプライバシー面においてメリットとなる他、低消費電力であるためセンサを小型にできる特長がある。一方で、先行研究であるような人数カウント機能までは搭載しない。その根拠として、本研究で必要な通行量の変数としては、絶対的な人数ではなく、物件同士の相对比较が行えるオーダーでよいことが挙げられる。具体的には、センサの一定時間内の反応回数を用いることで通行量とする。

2.3 暗黙知の表出化への取り組み

最後に暗黙知を形式知に変換する試みについて述べる。暗黙知と形式知を相互に変換するプロセスとしては、野中らが SECI モデルを提案している。SECI モデルの概要を図 1 に示す。SECI モデルとは、①共同化 (暗黙知→暗黙知) →②表出化 (暗黙知→形式知) →③連結化 (形式知→形式知) →④内面化 (形式知→暗黙知) →①共同化といったサイクルを繰り返すことによって、知識創造をしていくプロセスモデルである。本研究で取り組むのは②表出化であり、センシングとデータマイニングによって営業マンが持つ暗黙知を形式知に変換し、営業パッケージ化を実施する。また、具体的な例として、河村[13]らは申込顧客推薦システムを提案している。この研究は、飲食店向け不動産

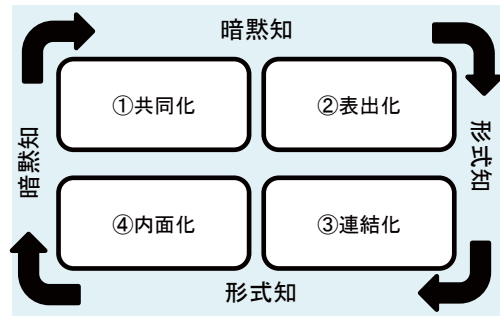


図 1 SECI モデル

営業支援のために行われており、顧客が不動産の検索から内見に移ったり、内見から申込に移ることを促すための電話営業を対象としている。これまで営業マンの勘と経験によって、数ある顧客リストの中から成約に至りやすい顧客を抽出し、電話をかけていたのに対し、経験の浅い新人営業マンではどの顧客に電話をかければ効果的なのかわからないといった問題があった。そこで、アンケート結果、物件情報、顧客情報の 3つのクラスを説明変数に機械学習を行うことで、49.2%の精度で顧客推薦を行うシステムを構築している。

本研究における暗黙知としては、賃料決定プロセスにおける営業マンの感覚が挙げられ、その中でも定量化されていない物件周辺の通行量や視認性の評価などが中心となる。本稿では通行量を定量的に取り扱えるようにすることで、暗黙知の抽出を図る。先行研究と異なる点としては、機械学習に使用する説明変数の特性が挙げられる。物件情報については共通性がある一方で、日々変化する動的情報である通行量を、センシングデバイスを利用して取得する点で異なっており、それが本研究の特徴の一つであるといえる。また、今回は取り組んでいないが、物件の視認性に関する暗黙知の抽出も必要になることが示唆されており、表出化手法に関する検討が必要である。

3. 通行量センシングシステム

第 1 章で述べた課題①を解決するために、店舗に関する動的情報である通行量などを取得するためのセンシングシステムを構築する。本システムは主にエナジーハーベスト無線センサと、ゲートウェイの役割を持つ無線受信機によって構成される。エナジーハーベスト無線センサを活用する利点として、半永久的なセンシングを可能としており、なおかつ配線が不要であることから設置の自由度が高いといった特長が挙げられる。

また、通行量に限らず、店舗内の雰囲気などのあらゆる情報をセンシング可能にするため、他の種類の EnOcean 機器にも対応可能な汎用性のあるシステムとしている。また、センサとゲートウェイを分離することによって、センサの設置の自由度を高めている。以下に、システムの構成

要素ごとの詳細について説明する。

3.1 通行量センサ

開発した通行量センサの外観を図2に示す。センサは設置しても目立ちにくいように小型のパッケージにしており、ケース込の外形(全長×全幅×全高)は81×21×45mmである。ベースのエナジーハーベスティング無線モジュールはEnOcean社の温度センサSTM431Jであり、独自に開発した拡張ボードに受動型赤外線(PIR)センサを搭載したものと組み合わせることで通行量センシングを実現可能にしている。

この温度センサには、ソーラーパネルと蓄電用キャパシタ、無線モジュールが搭載されており、ファームウェアを書き換えることによって、より汎用的な使用が可能となっている。EnOcean規格で使用している無線周波数帯域は928.35MHzであり、一般に広く使用されている無線LANやBluetooth、ZigBeeなどが使用している2.4GHz帯と比較して、回折しやすく干渉しにくいことから、実環境に強い規格であるといえる。また、通信可能距離は廊下見通し30m、鉄筋コンクリート壁越し10mであり、飲食店用不動産物件に設置するにあたり十分な距離を確保している。

さらに、夜間の照度が不十分な環境下でも永続的なセンシングを実現可能にするため、バックアップ用コイン電池を搭載している。使用しているPIRセンサはPanasonic社のEKMB1101112であり、低消費電流駆動(最小1μA)を特長としており、エナジーハーベスティングの実現に貢献している。検出可能距離は5mであり、検出範囲は縦横方向それぞれ±45度程度である。今回は unnecessaryな動体検出を防ぐことを目的に、センサ周辺をアクリル製のパイプ(図2:◆記号)で覆うことで検出範囲を制限している。

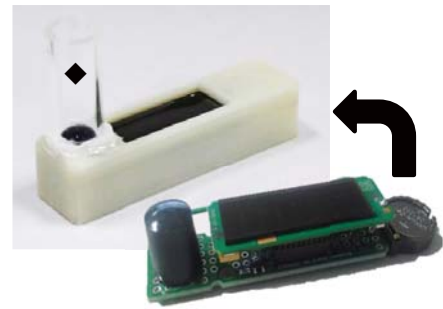


図2 通行量センサの外観

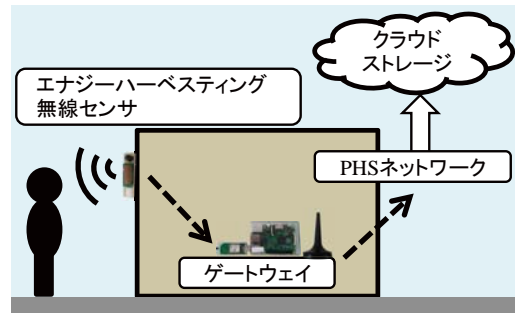


図3 データ送信の仕組み

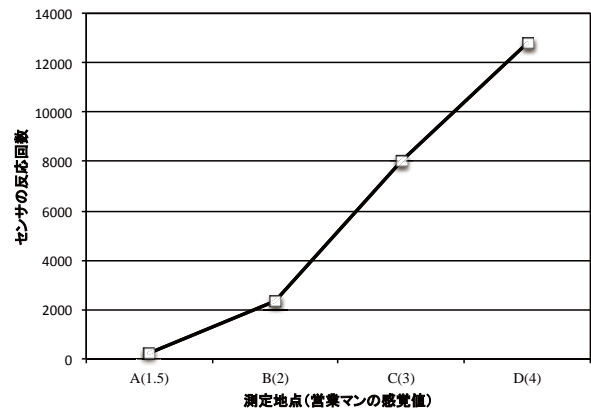


図4 営業マンの感覚データとセンサデータとの比較

3.2 ゲートウェイ

図3にゲートウェイによるデータ送信の概要を示す。ゲートウェイの役割を持つ無線受信機は、小型コンピュータのRaspberryPi2 modelBに、EnOcean受信モジュールUSB400JとABiT社のPHSモジュールを接続することで構成している。センサから受信したデータは内蔵マイクロSDカードにCSV形式で記録される。また、外部からのデータ参照を容易にできるようにするため、クラウドストレージであるDropboxに毎朝午前3時に自動的にアップロードするように構成している。RaspberryPiは、通常のPCとは異なり、自動的なクラウドストレージの同期には対応していないため、通行量が少ない深夜にセンシングを停止し、前日分のデータ送信を行っている。

また、ゲートウェイへの電源供給はUSB給電によって行うため、店舗にある商用電源の他に、短時間ならばモバイルバッテリーを用いた運用も可能である。

3.3 システム評価

構築した通行量センシングシステムの有用性を確認するために、営業マンの感覚に基づく通行量と、実際に通行量センサで取得したデータの比較を行う。本研究で対象としているのは、株式会社ABC店舗が取り扱う東京都内の飲食店用不動産物件であり、この中から営業マンの感覚データが異なる4つの地点を抽出し、センシング実験を実施している。図4に2016年2月8日(月)の夕方から翌日昼までの期間で実施した実験結果を示す。グラフの横軸に示しているA(1.5)、B(2)、C(3)、D(4)は4つの物件を示し、括弧内の数字は2名の営業マンが5段階評価した通行量の平均値(1~5:0.5刻み)である。また、縦軸はセンサの反応回数を示している。結果より、人の感覚に比例する形で反応回数が増加していることが読み取れる。このことから、構築した通行量センシングシステムでは、通行量のボリューム感を取得できるポテンシャルがあるといえる。また同時に、営業マンの感覚値であった通行量を根拠のあるデー

タとして形式化化する目処がたったことから、本システムは有用であるといえる。一方で、測定箇所が4地点に留まることからサンプル数を増やして検討をする必要があるといえる。

4. 機械学習に基づく賃料推定

課題②を解決するために、賃料推定手法として機械学習を用いる。本研究において、賃料推定で取り扱うのは東京都内の飲食店用不動産物件である。第3章で説明した通行量センシングシステムは、現在配備を進めている最中であるため、ここでは2名の営業マンから提供を受けた通行量の感覚データを用いて推定を実施する。データは、1~5の5段階であり、2名分の平均値を通行量として使用している。図5では、Open Street Map[14]に営業マンによる通行量の感覚データ（平均値）をプロットしている。

4.1 従来手法との比較

これまで、賃料推定は重回帰分析を用いて実施してきた。賃料推定に用いる指標としては、駅推定賃料（駅平均坪単価×物件の坪数）、視認性×通行量、居抜きフラグ、最寄り駅徒歩時間、階層、坪数が挙げられる。ここで、視認性×通行量は、5段階で評価した物件の視認性と、同じく5段階評価の通行量の積であり、通行量を入れると精度が上がることを確認されている。居抜きは、机や椅子、カウンター、ガスレンジ、冷蔵庫などの什器が付帯した物件を示しており、階数は物件が建物の何階にあるかを示している。これらから、推定賃料は次式によって求められる[2]。

$$\begin{aligned} \text{推定賃料} &= 137440.248 \text{ (定数)} \\ &+ 2.080 \times \text{駅推定賃料} \\ &+ 8504.796 \times (\text{視認性} \times \text{通行量}) \\ &- 23445.154 \times \text{居抜きフラグ} \\ &- 6177.402 \times \text{駅徒歩時間} \\ &- 47927.537 \times \text{階数} \\ &- 150.755 \times \text{坪数} \end{aligned} \quad (1)$$

また、(1)式を用いて求めた推定結果を図6に示す。グラフは横軸が実賃料、縦軸が推定賃料であり、グラフ内の線は±30%の範囲を示している。本手法では実際の賃料の倍近く差がついているものが多く見られることから、実運用に向けてはさらなる精度向上が必要であるといえる。また、定量的な精度評価を行うために推定値の平均二乗誤差(RMSE: Root Mean Squared Error)を(2)式によって求める。

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

ここで、Nは全予測対象数であり、 y_i は実績値、 \hat{y}_i は推定値を示している。RMSEは推定値が実績値からどれ程乖離

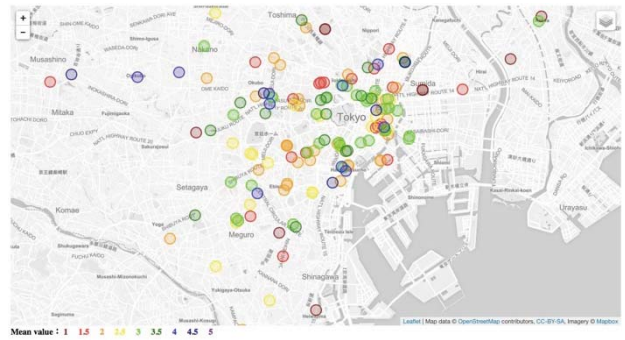


図5 Open Street Map[14]上にプロットした通行量

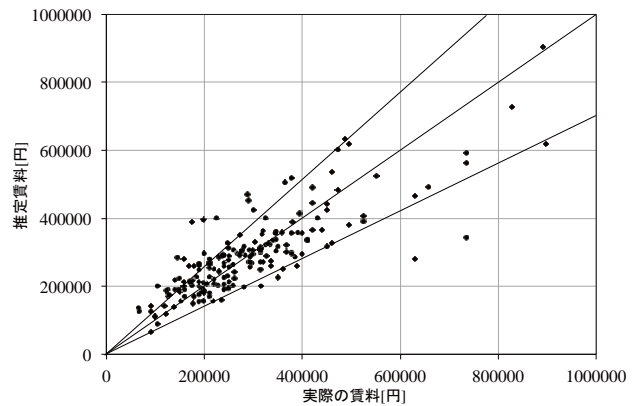


図6 重回帰分析による賃料推定結果

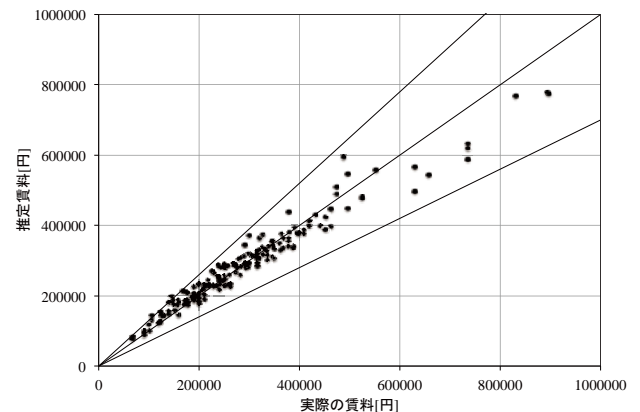


図7 ランダムフォレスト(内挿)による賃料推定結果

しているかを示しており、モデルの推定精度の悪さを評価する指標である。よって、0に近い程優れているといえる。以上に基づいて求められた重回帰分析を用いた際のRMSE値は80609.23であった。

次に、本稿では機械学習に基づいた推定システムを構築することで、精度の向上を図る。具体的には、機械学習手法の一つであるランダムフォレストの回帰分析を用いる。ランダムフォレストは、Breiman(2001)によって提案された手法であり、回帰木とアンサンブル学習の組合せである。回帰木はIf-Thenのような単純なルールで分岐できる弱学習器であり、非線形回帰分析の一つとして用いられている。また、アンサンブル学習は回帰木のような複数の学習器を融合し、汎化能力を高める手法である。

最初に、内挿である重回帰分析との単純比較を行うために、全物件データを学習させて推定を実施する。モデル構築に用いる木の数は 4000、用いる特徴量の最大数は 3 に設定している。このときの結果を図 7 に示す。重回帰分析の結果と比較すると、高い精度で推定できていることが読み取れる。また、このときの RMSE 値は 35235.51 となり、重回帰分析の半分以下にまで誤差が減少している。

次に機械学習によって作成されるモデルの汎化性能を評価するために、K-fold 交差検証を行う。これは、データセットを K 個に分割したものの内、K-1 個を学習データとして使用し、残り 1 個をテストデータとして検証することを K 回繰り返して正答率を得る手法である。ここで、扱う物件数は全部で 193 あり、分割数 K=3 としている。このときの結果を図 8 に示す。また、RMSE 値は 93785.39 となり、重回帰分析のときと同程度の結果となった。

4.2 精度向上に向けた特徴量の工夫

学習モデルのさらなる精度向上に向けて、現在使用している特徴量に新たな指標を加えたり、減らしたりすることで検討を行う。表 1 に、各特徴量を 1 つずつ減らした際の RMSE 値を示す。表中では、各特徴量を以下の様にアルファベットで表しており、その特徴量を使用する場合は○、使用しない場合は×で記載している。

- A: 駅推定賃料 (駅平均坪単価×坪数)
- B: 視認性×通行量 (1~25)
- C: 居抜きフラグ (0 or 1)
- D: 駅徒歩時間 (1~21min.)
- E: 階数 (1~5F: 地下 1F は地上 2F と同等扱い)
- F: 坪数 (4.51~73.02 坪)

一般的にこれらの特徴量の内、A,B,F については値が大きければ賃料が高くなるのが予想でき、逆に C,D,E はマイナスに働くことが予想できる。表 1 の最右列では、基準となる全部入りとの RMSE 値との差分を示している。この値が大きければ、その特徴量の影響が大きいといえる。また、正の数であれば、精度を向上させる方向に働きがあり、負の数であれば精度低下を招いているといえる。したがって、表中の delta の値に着目すると、駅推定賃料 (A) が最も推定精度に影響しており、視認性×通行量 (B) と居抜き (C) も高い影響を持っていることが確認できる。一方で、駅徒歩時間 (D) と階数 (E) の影響度はそれほど高くなく、坪数 (F) に関しては入れることによって精度低下を招いていることが確認できる。坪数 (F) に関しては駅推定賃料 (A) に含まれていることもあり、省いても構わないといえる。

次に、さらなる精度向上に向けて、坪数 (F) を除いた 5 つの特徴量をベースに新たな特徴量の追加・変更を行う。まず、影響の度合いが比較的小さかった駅徒歩時間 (D) と階数 (E) を積とすることで、特徴量の整理を行う。こ

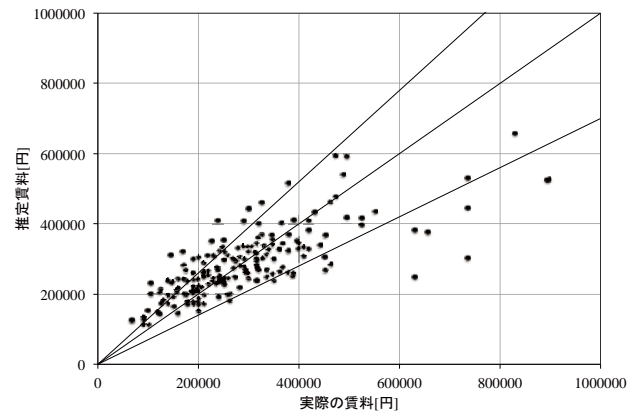


図 8 ランダムフォレスト (3-fold 交差検証) を用いた際の賃料推定結果

表 1 各特徴量の影響度合い

A	B	C	D	E	F	RMSE	delta
○	○	○	○	○	○	93785.39	0
×	○	○	○	○	○	112733.81	18948.42
○	×	○	○	○	○	97188.82	3403.43
○	○	×	○	○	○	95829.18	2043.79
○	○	○	×	○	○	93871.73	86.34
○	○	○	○	×	○	94580.55	795.16
○	○	○	○	○	×	93617.76	-167.63

表 2 特徴量を整理した際の精度評価

A	B	C	D	E	RMSE	delta
○	○	○	○	○	93617.76	0
○	○	○	D×E		91835.25	-1782.51
○	×	○	sqrt(D)×E		91987.58	-1630.18

れらを組み合わせる根拠として、この 2 つの特徴量は、顧客が店舗まで到達するための労力を示しており、共通性があるためである。ここでは、単純に掛けあわせたものと、値が大きい D の平方根を取ったものに E を掛けたものの 2 つについて検証を行う。結果を表 2 に示す。このとき、RMSE の差分である delta の値は双方とも負の数となり、元の RMSE 値よりも誤差が減ることから精度が向上することが読み取れる。また、平方根の有無は精度にほとんど影響がないといえる。したがって、駅徒歩時間 (D) ×階数 (E) を用いることによって推定精度の向上と学習に使用する特徴量の削減を実現することができたといえる。さらに、新たな特徴量として、店舗の導入力を追加する。導入力とは物件が持つ人を呼び込む程度を示しており、例として大通りの傍であることや、人の動線に面していることなどを目安として営業マンが感覚で評価したものである。この値も 1~5 の 5 段階評価であり、これを特徴量として追加することによって、RMSE 値は 91475.50 となり、わ

ずかではあるが精度の向上が確認できた。一方で、この変数を使うことはモデルに内包される暗黙知を増やすことにも繋がるため、それを表出化させる方法が必要になる課題もある。このことから、より精度向上に効果的な特徴量が見つかった際には、整理の対象となる可能性もある。

5. おわりに

本稿では、飲食店用不動産向けの賃料推定システムを実現するにあたり、動的情報を収集するための通行量センシングシステムについて説明し、機械学習に基づく賃料推定システムを提案した。これまで取り組んできた重回帰分析に基づく手法とランダムフォレストを用いた場合の推定結果を比較し、内挿同士の比較では推定精度の向上が確認できた。また、推定に使用する特徴量を比較検討した結果、精度低下を招いていた特徴量を排除でき、既存の特徴量の一部においても積をとることで整理を行い、元の状態よりも推定誤差を減少させることができた。

今後の課題として、まずニューラルネットワークやSVRなどの異なる種類の機械学習手法を適用することが挙げられる。今回使用したランダムフォレスト単体では、3-fold交差検証を行った際に元の重回帰分析の精度を上回ることができなかったことから、学習器のモデルとの相性などを考慮していく必要があると考えられる。また、使用する特徴量の検討を重ねていくことで、さらなる精度向上に取り組むことがあげられる。さらに、長期的な予定として、通行量センシングシステムを用いて収集したデータを活用し、本稿で提案した推定システムと組み合わせることで、根拠に基づいた賃料推定システムを完成させることが挙げられる。現在、東京都内でのセンサの設置を進めており、営業マンの5段階の感覚データを定量的なセンサデータに置き換えることで、暗黙知を形式化していく。また、現時点では視認性と導入力に関する評価も営業マンの感覚に依存しているため、これらに関しても暗黙知を抽出させる手法を検討していく必要があるといえる。

謝辞

本研究の一部は科学研究費補助金 挑戦的萌芽研究 15K12161 の助成を受けたものである。また、研究フィールドの提供にご協力いただいた株式会社 ABC 店舗の皆様へ感謝いたします。

参考文献

- [1] Ikujiro Nonaka, Hirotaka Takeuchi: The knowledge-creating company: How Japanese companies create the dynamics of innovation. Oxford university press (1995)
- [2] 荒川周造, 諏訪博彦, 小川祐樹, 荒川豊, 安本慶一: 通行量センサを用いた飲食店用不動産賃料推定システムの提案, 人工知能学会「社会における AI」研究会 第25回研究会
- [3] Leo Breiman: Random Forests, Machine Learning Volu

- me 45 Issue 1, pp.5-32 (2001)
- [4] Victor Gan, Vaishali Agarwal, Ben Kim: DATA MINING ANALYSIS AND PREDICTIONS OF REAL ESTATE PRICES, Issues in Information System, Volume 16, Issue IV, pp.30-36 (2015)
- [5] Chih-Hung Wu, Chi-Hua Li, I-Ching Fang, Chin-Chia Hsu, Wei-Ting Lin, Chia-Hsiang Wu: HYBRID GENETIC-BASED SUPPORT VECTOR REGRESSION WITH FUNCTIONAL THEORY FOR APPRAISING REAL ESTATE PRICE, 2009 First Asian Conference on Intelligent Information and Database Systems, pp.295-300 (2009)
- [6] Vincenza Chiarazzo, Leonardo Caggiani, Mario Marinelli, Michele Ottomanelli: A Neural Network based model for real estate price estimation considering environmental quality of property location, Transportation Research Procedia 3, pp.810-817(2014)
- [7] 株式会社リブセンス ニュースリリース: リブセンス、新たに中古不動産売買領域へ参入 未来型不動産サービス「IESHIL (イエシル)」スタート ービッグデータ活用により中古不動産売買試乗の”可視化”と”活性化”を実現ー (2015年8月27日) http://www.livesense.co.jp/news/n/2015/0827_01/attachment
- [8] ソニー株式会社 ニュースリリース: 業界最高水準の精度を実現した「不動産価格推定エンジン」を開発〜ソニー不動産が Real Estate Tech (不動産支援テクノロジー) 領域で消費者意思決定支援のための新サービスを展開〜 (2015年10月8日) <http://www.sony.co.jp/SonyInfo/News/Press/201510/15-1008/>
- [9] Carlo Pane, Marco Gasparini, Andrea Prati, Giovanni Galdi, Rita Cucchiara: A People Counting System for Business Analytics, 2013 10th IEEE International Conference on Advanced Video and Signal Based Surveillance, pp.135-140 (2013)
- [10] MA Hudong, Zeng Chengbin, Ling Charles X.: A Reliable People Counting System via Multiple Cameras, ACM Transactions on Intelligent Systems and Technology (TIST) Volume:3 Issue: 2, pp.31:1-31:22 (2012)
- [11] Florian Wahl, Marija Milenkovic, Oliver Amft: A distributed PIR-based approach for estimating people count in office environments, 2012 IEEE 15th International Conference on Computational Science and Engineering, pp.640-647 (2012)
- [12] Fang Zhu, Xinwei Yang, Junhua Gu, Ruixia Yang: A New Method for People-Counting Based on Support Vector Machine, 2009 Second International Conference on Intelligent Networks and Intelligent Systems, pp.342-345 (2009)
- [13] 河村一輝, 諏訪博彦, 荒川豊, 安本慶一: 飲食店向け不動産営業を支援する申し込み顧客推薦手法の提案, 情報処理学会研究報告, DICOMO2015, pp.588-593 (2015)
- [14] Leaflet | Map data © OpenStreetMap contributors, CC-BY-SA, Imagery © Mapbox