

不動産情報のデジタル化と可視化は、収益化の鍵になるか？

原題: Digitalize = Visualize → Monetize?

日本不動産研究所 REA-Tech 研究開発グループ

南川しのぶ

はじめに

不動産関連情報のデジタル化に期待される効用は絶大である。当研究所では4年前に、所内データの整備・先進技術活用による所内業務の効率化・データ活用による新しい知見の発見と新規ビジネスへの寄与を期待して、専門のグループを設置した。本稿はこのグループの活動報告を趣旨としている。

ビッグデータは、整備しなければ、単に容量の大きな情報の塊に過ぎない。本稿では、これに対して、第一章: デジタル化すること自体が新規ビジネスの要素になる例と、第二章: デジタル化した上で、利害関係者の立場に応じた情報の選択と開示の方法を工夫すること、すなわち可視化(visualize)が新規ビジネスの要素になる例、さらに、第三章: 高度な統計解析やAIを使った解析による新しい知見の発見が、新規ビジネスの端緒になる例を報告する。

私達を取り巻く情報は日々肥大し、既に個人が手に取って理解できる量を凌駕した。不動産については各社のいかんともしがたい事情により個人に割り当てられる業務量・対象範囲は拡大し、全国化・アジア全域化・世界化してかつ多層化する。事情の異なる地域現物たる不動産について日々判断を求められ、そのピッチは加速する。これが不動産を取り巻くニューノーマルだとすれば、このニューノーマルにおける新しい物差しと算盤を作ることが当グループの使命だと思いたい。

既往研究

不動産関連情報のデジタル化を出発点とするデジタルトランスフォーメーション(以下「DX化」)に注目が集まる理由は、不動産ライフサイクル全般における効率化・安全確保・将来予測の精度向上などと、データ活用によるビジネスチャンス獲得に大別される。Brennenらは、混同されがちな digitization と digitalization について整理した[1]。前者を「写真・動画・テキストなどのアナログデータをデジタル形式に変換するデジタル化のアクションまたはプロセスのこと」とし、「個々の情報のアナログな流れをデジタルビットに変換する物的プロセス」と定義している。また後者を「社会生活の多くのドメインがデジタル通信とメディアインフラを中心に再構築される方法」と定義している。不動産データのデジタル化による効果測定については、学会[2]やカンファレンス[3]においてその進展を確認することができる。また、ジャーナルも盛んに公開されており、デジタル化に特化したものも開催されている[4]。

立地選定の研究は日本においても多い。例えば百合本らの研究[5]はアメリカを対象としてAHP(Analytic Hierarchy Process)法を利用して生産拠点の立地選定のための意思決定支援システムを構築した。またGISを活用した立地選定システムについての亀田らの研究[6]は、北九州地域を対象として、風力、用地の勾配、土地利用と社会的状況等の諸条件についてGISを用いて総合的に評価し、風力発電所に最適な候補地を選定している。

候補地の選定には企業の出退店担当者の暗黙知を表出することで精度向上が期待できる。荒川らはベテラン営業マンの暗黙知の表出を試み、賃料推定の精度向上に寄与したことを報告している[7]。

本研究ではうどん店の候補地選択には店舗出店担当者の暗黙知を機械で実装するツール開発を目指した。また、データセンター候補地選択には選定条件につい

て GIS を用いて総合的に評価する試みとして行った。

統計や機械学習を用いた価格・賃料の推定の研究は盛んに行われている。[Harrison 1978]は、ボストンの住宅価格と大気汚染との関係を分析し、この論文が以降の研究に多く引用されている。日本を対象とした価格や賃料の推定、またはその構成要因に関する研究は多い。これら論文は住宅を対象とするものが多く、これはデータ入手容易性と、各物件の比較容易性が寄与していると思われる。

一方商業用不動産の価格・賃料の推計についても近年論文が出てきており、取引事例が秘匿事項とされ公開されていない日本においても、研究用公開データを使った研究や、前出の荒川らの研究[7]等、独自データを使った研究がある。

なお、日本においても、アメリカの Zillow, Zestimate, Redfin のように、各情報プロバイダから価格・賃料推定情報が公開されている。情報提供サイトである「飲食店.com」[8]は、同サイトに直近1年間で掲載された物件を集計し、駅や市区町村ごとの賃料(坪単価)について、平均値・最高値・最低値と賃料分の布図、平均相場の推移(4年分)を公開している。

本研究において、デジタル化の章では既往研究をもとに不動産実務領域で試行する場合の論点を共有する。可視化の章では、既往研究を踏襲しつつ、対象企業の高次の意思決定支援と営業マンの説明支援ツールとしての一挙両得の活用可能性を確認する。

商業用不動産の賃料検証の章では、既往の統計と機械学習による価格や賃料推定について、既往研究とは異なるデータセットを用いて頑健性を確認する。また、その精度検証を実務での運用可能性と関連させて論じる。

第一章 デジタル化

不動産テックに限らず、数十年前の IT 黎明期から何度も議論されてきたことではあるが、社内に蓄積された紙、スキャンデータ、ファイル等の情報は、情報整備の目的・範囲・許容精度を明確にし、構造化することで、初めて情報資産となる。よってこの目的の明確化と構造化するデータの特定、そしてデータ項目の定義が大切な出発点となる。と同時に、このコンサルテーションがビジネスチャンスとなる。

不動産情報の DX (digital transformation) は不動産テックに関して言えば、目的の明確化と、その目的を達するための基準+必要情報の定義、については、まださほど議論されていない。そこで、ここではいくつかの基準と収集データのセットを考えたい。

まず、鑑定評価自体は確立された議論がある。

目的: 対象不動産の時価を算出する

基準: 不動産鑑定評価基準

情報: 価格形成要因

DB: 土地/建物/権利関係の DB と、査定項目の DB
例えば、不動産管理事業に係るデジタル化ならば、

目的: 管理事業全体の収益最大化

基準: 会計基準と鑑定評価基準と物件管理基準

情報: 物件特定、会計、評価、管理に関する情報

DB: 現行の時価と賃料情報・検証のための情報一式、立地情報と、管理拠点に関する情報一式、管理状況・修繕履歴等の一覧等

ここで2つの大きなコスト局面がある。

まず、このデジタル化は、目的・範囲・精度の基準と、必要なデータを考えることが、企業の事業計画に関わる本質的な議論を呼び起こす。それは不動産の開発事業や建築の設計段階においてコンセプトの共有と設計までの作業に大きな労力と決断を要するのと似ている。

そして、散逸した状態で長らく蓄積されたデータは、その構造化に甚大な時間とコストがかかる。

この2つの局面で支援する業務を、BPO 業者・システム業者・ERP 業者等がビジネスチャンスと捉えて、参入する動きがある。おそらく企画フェーズはコンサルフィーとして数百万円、デジタル化フェーズは数千万円～数億円の投資規模となるだろう。参考として、研究所が過去 100 年分の紙資料をデータベース化するとしたら、約 20 億円かかる。また人工知能化のために鑑定 COPAS と呼ばれる学習用辞書を作ろうとすれば、要件定義だけで 2,000 万円ほどかかる。

第二章 可視化

I うどん屋出店候補地

フランチャイズの飲食チェーン企業が次の出店候補地を選ぶ際に、不動産鑑定士はどのような支援ができるだろうか。募集店舗情報をかき集めることだろうか。エリア選定から始めるだろうか。経験と勘、だろうか。この例では、うどん屋の出店候補地を、東京都内にある延床面積 20,000 m²以上のビル一覧を地図上に可視化することから始めた。なお本件は想定設定であり、受託案件ではない。

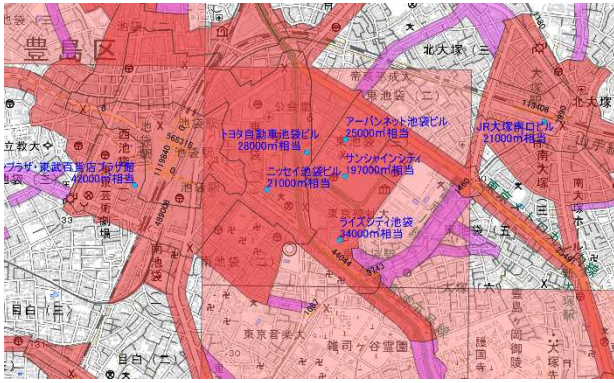
目的は最適候補地の選出である。調査対象エリアは東京都 23 区。手法は、GIS(地理空間情報システム)上に以下の要件を満たすエリアまたは地点をプロットし、全要件を満たすビルを選出すること。本件の適用に際して、空間解析や統計解析は行わない。データは、①登記簿から研究所が収集した延床面積 20,000 m²以上のビル一覧と各ビルの甲区情報(住所・竣工年・情報取得年・階層・構造・床面積・事務所相当床面積)、②国土数値情報[9]より道路密度/道路延長データ 1km メッシュ、③用途地域である。①の延床面積 20,000 m²とは、同ビル内に十分な量の就業者が見込まれる観点の選考基準である。②の道路密度/道路延長を採用した理由は、容易に取得できるオープンデータであり、視認性が高い店舗用地の確保が可能な地域を抽出するためのヒント

となる点である。このデータについて、視認性が認められること、店舗の繁華性が生まれやすいこと、当該道路が地域を分断しないこと、の観点から、幅員 13-19mの道路延長に着目し、東京全域を 5 ランクに区分した。その結果、上位から 3 ランクに該当する地域を選出した。③用途地域は、高度な容積率、または、道路沿道等の店舗繁華性が見込める商業地域・第二種中高層専用地域・近隣商業地域・準住居地域の 4 地域を選出した。

この地図に、事務所相当面積が 10,000 m²以上のビルだけを残した。そして、④出店済みのうどん店舗を載せ、⑤競合のうどん店舗を載せた。次図は顧客に提出する資料の一部である。用途地域、⑥駅の乗降客数・道路視認性ランク・ビルが一覧できる。なお、ビルは登記簿取得時の情報のため、その後のオーナーチェンジには対応していない。旧ビル名の場合がある。

このプロジェクトは、顧客企業のデータ管理議論を待たず、自社で保有するデータと、オープンデータだけでどこまで分かるか試した点が特徴的である。結果は、805 件のデータから出発し、要件に最も合致したビルを 20 件程度抽出した。依頼者は対象を絞った各ビルに対して空き室がないかを随時監視できる。また FAX 等で知らされる新規空き店舗情報と、有望地域を常に地図上で照合できる。また競合店の状況を載せてユーザーの暗黙知を含む判断材料にした。

出店担当者のもとに FAX 等で常時送られてくる空き店舗情報は全国に及ぶことがある。地元精通者でなくても一定の品質と基準での選別が容易になる点が本件の効果である。



図表: うどん出店候補地 ~池袋~

II 鑑定情報一覧サイト「すぐれあ」

社内の鑑定評価に役立つ閲覧サイトを作れるだろうか。効率化・品質向上に資する目的で、社内に GIS 閲覧サイト「すぐれあ」を作った。「SUGU」とは、すぐに鑑定評価できるようにとの思い、「REA」は、不動産鑑定評価の略字である。目的は、より多くの参考情報の共有と、一枚の地図上で混乱することなく見て、比較できる仕組みを作り、鑑定評価の品質維持・作業の効率化に役立っていることにある。対象地域は日本全土 37.8 万km²。手法は、GIS 上にエリアまたは地点をプロットすること。本件の適用に際して、空間解析や統計解析は行わない。また、現段階の仕様では、ユーザーに検索や解析を行わせる。データは、主に研究所が調査・公表している定期調査、並びに、研究所が収集した以下の 11 点を掲載している。

- 先例
- 利回りベンチマークポリゴン(住宅オフィス商業)
- 利回りベンチマーク地点(レジオフィス商業ホテル物流)
- TO-REIT
- 投資家調査(賃料水準・利回り(標準商業物流ホテル賃貸)) [10]
- ビルリスト
- 全国賃料統計[11]

- 市街地(住商工・全用途平均・ポイントデータ)[12]
- 店舗賃料トレンド
- 地価公示[13]
- 都道府県地価調査



本件で期待される効果を述べる。従来の紙資料では、一般的に用途別、都道府県別に整理されてきた情報についても、すぐれあをスクロールして比較できるようになった。以下の例では、例えば物流倉庫の利回り水準が、下図中央の川を挟んで東側が高いことが分かる。この川は、県境である。各県を横断しての同用途物件の水準比較、同一地域における用途ごとの水準比較が可能である。



すぐれあについての課題が明らかになってきた。

現状は、ユーザー数が増えず、ログが伸びなかった。それは、現状の研究所の評価フローやプロセスが、すぐれあを見なくても成立するからである。全ての鑑定士は必ずしもより多くの参考情報を求めない。それよりも、依頼者のメ切に応えつつ最短で確実な回答と、より迷いのない仕事を自らに課す。それは過去と不確実な将来に対する責任感とも、当事者意識とも言える。

このすぐれあプロジェクトを社内 GIS の次世代化に向けた一試行として整理するか、公開データを中心に

構成された公開版の普及を試みるかは、議論の途上である。すぐれあに関して先に興味を示したのは、鑑定業以外の会計監査法人や金融機関、IT システム会社内の不動産テック部署であった。不動産専門ではない企業または部署にとって、公開版であっても、可視化されて分かりやすくなった不動産ツールは魅力だと分かった。また、資金調達を裏付けに早期の急成長を課すスタートアップ企業は、このデータの上で解析エンジンを機能強化して、新しいサービスを共同提案したいと思っている。顧客企業に公開されることで、当該情報を持つ顧客に対して鑑定評価の価値を問うことになるだろう。先ほどの県境の例で言えば、鑑定評価理論を知らなくても、すぐれあを見れば、隣の県より高いのはなぜか、隣のビルより低いのはなぜかという議論はありえる。情報化が進展する中、相互の前提知識が増えれば、鑑定評価書に綴られた共通言語としての数値がより雄弁なものとなる。また、鑑定士が依頼者と市場の代弁者として査定した利回りや格差の意味を共有する助けとなるのではないか。

法律に基づく自律的な専門職業家によるチームが、遵法性の範囲で、情報の非対称性の減少につながるサービスを模索している。

Ⅲ データセンター用地の抽出

アンドリュー＝バウム博士は不動産テックを「不動産の情報、取引、管理に関する革新的な技術」と定義した[14]。革新的技術のひとつに GIS がある。GIS の活用により、これまででない効率で成果を上げることができのではないか。先述の通り、立地選定の研究は風力発電所やメガソーラーなどがある。ここでは例としてデータセンターを取り上げ、以下の 10 要件に合致した候補地を抽出した。

目的は、考え得る“全ての”適地を抽出すること。

対象地域は、東京駅から半径 100 キロ以内の全ての土

地(31,400 km²)。手法は、顧客が提示する 10 要件を可視化すること。要件ごとの複合的な計算・解釈・重み付けは行わないが、地価公示水準のメッシュ化に際しては内挿を、候補地抽出に際しては空間結合を、距離の測定に際しては空間解析を行った。出力に際して候補地のオススメ度・重要度等についての解釈は行わない。

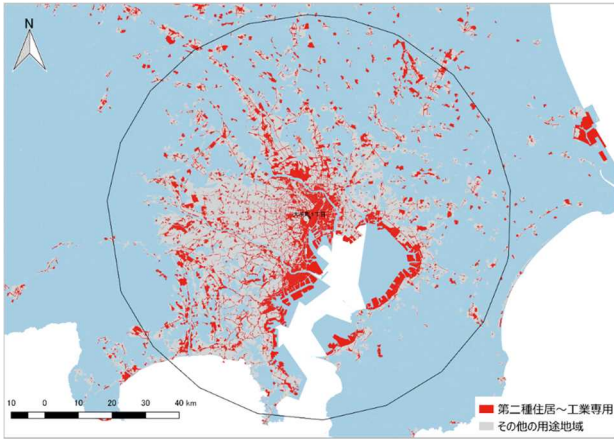
使用したデータは、

・国土数値情報[9]より駅、用途地域、地価公示、衛星画像、道路のデータ、河川の水害ハザードマップ。都道府県 HP より、津波浸水想定マップ（閲覧）である。本件もオープンデータを使って検討を行った。

本件の要件は以下の通りである。

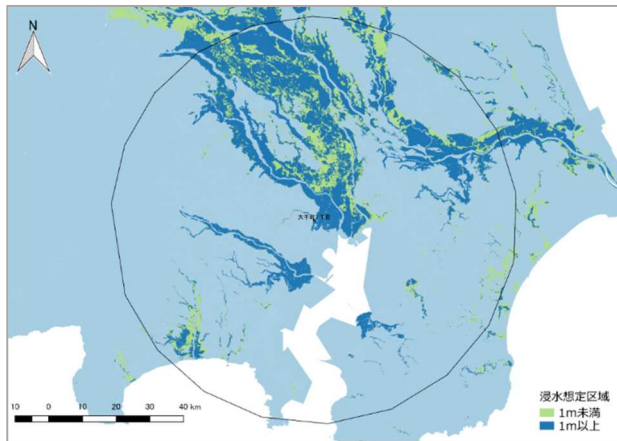
- ① 東京駅から一定距離以内
- ② 特定の用途地域
- ③ 河川からの浸水想定高が 1m 未満
- ④ 土地価格が十分に安い
- ⑤ 道路等の分断がない 10,000 m²以上の一団の土地
- ⑥ 30,000KW 以上の電力供給エリアから 1.5km 以内
- ⑦ 液状化の可能性がない
- ⑧ 津波浸水深 1m 未満
- ⑨ ガスタンク等爆発物施設から遠い
- ⑩ 空き地であること

なお上記の要件は、想定設定した要件である。受託案件ではない。また本稿では一部要件の表現をあいまいにしている。



要件①+②の実装イメージ

国土交通省「国土数値情報」をもとに編集・加工



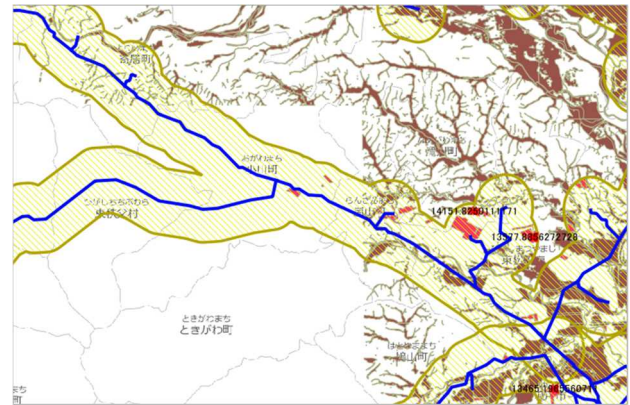
要件①+③の実装イメージ

国土交通省「国土数値情報」をもとに編集・加工

要件⑤の一团地を GIS 上に描画するために、国土数値情報より、道路データと建物データを使用し、これ以外の場所をベクタデータとして整備した。なおこの手法では用途地域内の現況山林等が含まれるが、要求精度の観点から許容した。

要件⑥の電力供給エリア関連については、次図表における黄色の帯で表現した。また、要件⑦の液状化リスクについては、約 100 年前、明治時代に作られた「迅速測図」[15]による選別を提案し、沼地・泥湿地・海等の一定の地目を条件外とした。次図は要件①～⑦電力供給エリア及び液状化リスクの可視化イメージである。なお、迅速測図の整備エリア外であることから茶色の

地図が右下で途切れている。



要件⑧の津波浸水深については、作業時点で地図データ化された情報を入手できなかったため、海に面した4県[16]について、抽出後の画地について目でチェックした。

最後に、要件⑨のガスタンク及び⑩空き地であることについては、当初航空写真に基づく機械学習を試みた。しかし、満足な結果が得られなかったため不採用とした。その主な理由は以下の通りである。

- ・駐車場なのかビルの屋上なのかを判別できない
(平地の駐車場=○、屋上P=×、ビルの屋上=×)
- ・空き地なのか収穫後の農地なのかを判別できない
- ・ガスタンク(円型)を判別できない
- ・商用利用な航空写真が最大5年古い

よって、この2要件についても、8要件に合致した候補地について、目でチェックした。

以上の作業を経て作成した地図の一例を以下に示す。
(部分と全体)



要件①～⑦の実装イメージ

国土交通省「国土数値情報」をもとに編集・加工



要件①～⑩の実装イメージ

国土交通省「国土数値情報」をもとに編集・加工

本プロジェクトの結果、31,400 km²から、207 件の候補地を得た。案件化すれば、この 207 件を台帳にすることで納品できる。また技術的にはこの 207 件に対して、標高差や日照等の自然的条件や、雇用者の通勤可能性等の条件を追加し、候補地としての精度を最上位から 10 件程度に絞ることが可能である。

本プロジェクトで得られた課題について述べる。顧客要件は変動することを想定し、可視化ツールも機動力/柔軟性が必要という教訓を得た。例えば土地価格要件が変動すれば、全 31,400 km²に対して再度の解析が必要となる。画地抽出には一晩かかる作業もあり、所要時間とマシン性能を比較した機動性の確保が必要となる。特に、本件のように、一部のデータがデジタル化されていない場合、後半の作業が目視になるので、前半の要件が変化すれば、目視のやり直しを免れない。

また、実務においては、依頼者が当該 10 要件の設定が正しいという確信を持ち、候補地確定に至るまでの支援ができるために、事業に関する知見も併せて必要と感じた。

これら課題を認識しつつも、本件の手法は他の用途にも活用できる手法と期待している。可視化とは、載せるだけだが、何の目的で、何を載せるか、読み手の解釈と判断が容易になるような情報の取捨選択が重要であ

る。また、多くの場合、価格・賃料情報を求められる。よって、不動産精通者が率先して立地選定に関わる意義があると考えている。

なお、可視化ツールの作成には先述の企画と選定に係る行程に加え、データ更新を求められる。データ更新は「間違いなくできて当たり前」であり、一般には IT の素養を要し、やってみるとハードルが高かった。このように介入価値がある一方で情報維持更新体制を要することから、ビジネス化には企画と維持の両面に配慮した体制作りを要するだろう。

第三章 解析

デジタル化と可視化ができたとすれば、整備されたデータについて解析を行い、隠れた示唆を明らかにすることができるのではないかと。以下に、現在開発中のプロジェクトを示す。

IV 24 時間営業のドラッグストア

フランチャイズチェーン本部にとって、店舗の賃料はとても大きな固定費になる。また、過去に出店営業担当が新規出店によって臨時ボーナスを得る慣行があったなど、長期間オーナーに有利な（高い）賃料が据え置かれている例がある。これらの賃料を適正化することで当社全体に大きな体質改善が見込まれる。

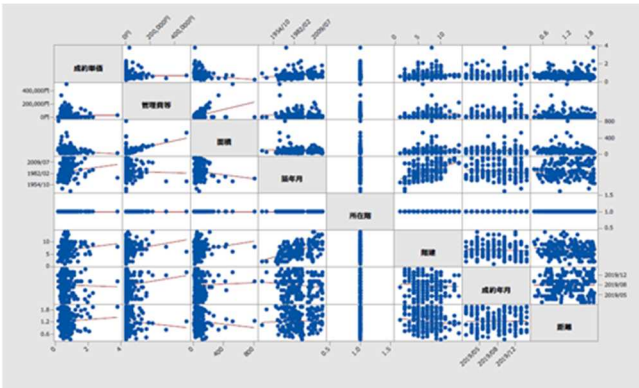
本件の目的は、依頼者から受領した物件一覧について、市場賃料との乖離度を提示することを通じて、顧客が賃料減額交渉の準備をすることにある。対象サンプルとして、東京都 23 区に存する 50 件のドラッグストアを想定設定した。実際の受託案件ではない。

手法は、物件及び物件周辺の店舗に係る取引事例を取得し、比較アプローチ及び統計的アプローチによる賃料検証を行って、結果を行列散布図及び GIS に出力する。データは、各対象不動産及び周辺の事例 50 件～200 件に係る、住所、交通、面積、坪面積、賃料、単価、

坪単価、用途、管理費、礼金、敷金、敷引、保証金、償却、築年月、所在階、構造、階建、成約年月、駅からの直線距離である。

手法の詳細を以下に示す。

まず、対象物件から一定距離に存在する 50 件~200 件の事例を取得した。そして、これら 50~200 の事例 +1 の対象不動産の、計 51~201 件の物件について、縦軸を賃料に、横軸を各変数として、行列散布図を出力し、賃料が高い割に駅から遠い、などの情報を可視化した。



対象物件を含む 201 件についての行列散布図

独自収集データに基づく編集・加工

次に、平均単価との乖離率順位表（相場感の提示と対象物件の位置づけの可視化）、標準化補正後単価の順位表（定量化できる理由なく高い物件を抽出：非線形分布を想定）、回帰補正後単価の順位表（定量化できる理由なく高い物件を抽出：線形回帰を想定）を作成した。成果物のイメージ図を以下に示す。

住所	築年	成約日	所有者	成約単価	成約単価(2)	成約単価(3)	成約単価(4)	成約単価(5)	成約単価(6)	成約単価(7)	成約単価(8)	成約単価(9)	成約単価(10)	成約単価(11)	成約単価(12)	成約単価(13)	成約単価(14)	成約単価(15)	成約単価(16)	成約単価(17)	成約単価(18)	成約単価(19)	成約単価(20)
東京都品川	2018/11	4	住友不動産	1,800,000	60,421	1	3	4	2	10	1	1	1	1	1	1	1	1	1	1	1	1	1
東京都品川	2018/12	6	新築	960,000	42,526	2	3	2	2	9	1	1	1	1	1	1	1	1	1	1	1	1	1
東京都品川	2018/12	0	住友	745,200	30,263	3	4	34	10	488	13	1	1	1	1	1	1	1	1	1	1	1	1
東京都品川	2018/11	3	住友	745,200	30,263	4	26	34	10	118	14	1	1	1	1	1	1	1	1	1	1	1	1
東京都品川	2018/12	1	新築	250,000	26,014	5	30	14	5	36	2	1	1	1	1	1	1	1	1	1	1	1	1
東京都品川	2018/12	0	住友	440,800	24,953	6	81	29	20	148	30	1	1	1	1	1	1	1	1	1	1	1	1
東京都品川	2018/12	0	新築	248,400	23,397	7	80	7	35	127	14	1	1	1	1	1	1	1	1	1	1	1	1
東京都品川	2018/12	0	新築	248,400	23,397	8	6	3	64	72	26	1	1	1	1	1	1	1	1	1	1	1	1
東京都品川	2018/11	2	住友	231,400	21,648	9	64	13	1	77	5	1	1	1	1	1	1	1	1	1	1	1	1
東京都品川	2018/12	0	新築	234,000	12,388	10	23	108	20	134	21	1	1	1	1	1	1	1	1	1	1	1	1
東京都千代田	2018/11	9	新築	690,000	16,612	11	87	53	36	178	50	1	1	1	1	1	1	1	1	1	1	1	1
東京都品川	2018/11	2	新築	216,000	15,618	12	51	14	31	65	1	1	1	1	1	1	1	1	1	1	1	1	1
東京都品川	2018/11	7	新築	324,000	12,388	13	88	20	20	134	21	1	1	1	1	1	1	1	1	1	1	1	1
東京都品川	2018/12	3	新築	220,000	11,735	14	57	31	74	162	68	1	1	1	1	1	1	1	1	1	1	1	1
東京都品川	2018/11	0	大塚	216,000	11,531	15	28	31	24	83	34	1	1	1	1	1	1	1	1	1	1	1	1
東京都品川	2018/11	1	住友	140,000	10,752	16	64	8	21	69	20	1	1	1	1	1	1	1	1	1	1	1	1
東京都品川	2018/12	6	大塚	108,000	10,499	17	34	25	21	80	20	1	1	1	1	1	1	1	1	1	1	1	1
東京都品川	2018/12	6	新築	118,800	10,090	18	75	6	43	124	31	1	1	1	1	1	1	1	1	1	1	1	1
東京都千代田	2018/11	0	新築	216,000	8,678	19	63	64	38	130	68	1	1	1	1	1	1	1	1	1	1	1	1
東京都品川	2018/11	30	大塚	150,000	8,089	20	87	9	43	147	32	1	1	1	1	1	1	1	1	1	1	1	1

※画像はサンプルデータによるイメージ

行列散布図を見ると、相関係数の高いものも低いものもある。この行列散布図は、対象不動産の競争力を多

角的に視覚化するために採用しており、相関係数が低いことが必ずしもモデルの説明性を左右しない。

結果として、対象となる 50 件を地域賃料水準に比して割高と思われる順にランキングをつけて納品することになる。リストの他に、各 50 店舗について周辺状況をレポートし、定性的な条件も示唆できる仕組みとした。

本件から得られた考察と課題を述べる。本件においては、高度な解析を行わず、可視化に留めた。その理由は、納品済みのレポートが、提案者による解釈や説明を省略しても雄弁な資料になっている点である。また、コンフリクトの観点から、賃料評価を行わずに、資料の適切性を示唆する試みであった。コンフリクトとは、同一不動産における異なる利害関係者から鑑定評価の依頼を受けることを厳に慎むルールである。

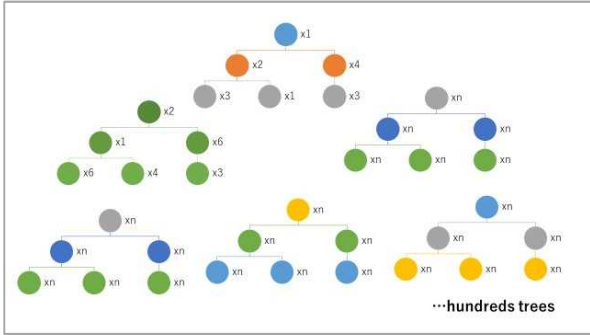
日本は、アメリカ等と違い、取引事例は秘匿情報である。典型的な個人が見られるのは、募集広告である。そして募集広告は、日々更新され、膨大で、重複も多く、価格や賃料が適正なのかは、分からない。(なお重複情報の問題、いわゆる不動産同定問題については清田らにより 2021 年 AI 学会などで定義、議論されている。)

冒頭でも述べた通り、私達を取り巻く情報は日々肥大し、既に個人が手に取り理解できる量を凌駕した。その中において、借り手による賃貸物件選定は申込の先手必勝となる。ここに、賃料査定のエッセンスを含めた可視化と解釈の大量化・自動化に備える必要性和意義がある。

V 24 時間営業のドラッグストア(2)

この案件においては、IVのモデルを拡張し、統計的機械学習による賃料を予測し、実際賃料からの乖離率を求めた。対象物件、エリア、手法は前節と同じとし、解析の一部に機械学習による推定賃料を加えて順位表を更新した。機械学習手法としては、step 関数を用いた

重回帰、決定木、XGboost 法による回帰木、ランダムフォレストによる予測結果の RMSE(二乗平均平方根誤差)を比較した結果、本件ではランダムフォレストを採用した。



図表: ランダムフォレストのイメージ図 (筆者作成)

成果物は前節同様に 50 件分のレポートである。

現状で見えてきた課題を以下に述べる。

まず、機械学習の実装と課題として、序列を制御することの難しさがある。第二章Ⅱすぐれあでも述べたが、「隣より高いのはおかしい、隣より低いのはおかしい」という序列の違和感は、不動産や統計に精通していなくても気づかれやすく、多くの場合正しい。機械学習で価格の RMSE (二乗平均平方根誤差) を改善しても、筆者はまだ序列を制御できていない。

また、機械学習を高度化しても、結局は対象不動産の目的変数をより良く説明する 1 変数を発見できれば、多くの精度面の問題が改善する。そしてそのような変数によるシンプルな説明ができれば、機械学習よりも実際の説明性は向上する可能性がある。

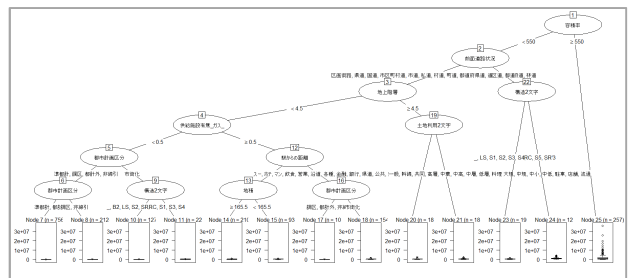
発見されたこれらの課題を、今後の検討に役立てていきたい。

第四章 課題と結論

5 件の事例を含め、多くの試行錯誤を行ってきた。その中でいくつかの課題を取り上げる。

1. 複数用途に有効な価格モデルを作ることの難しさ

以下は価格モデルを作成する際に行った決定木による価格形成要因分析の結果である。第一位の分岐、すなわち、価格の決定に最も影響が大きいと思われる分岐が、容積率であり、550%以上の土地については価格差が大きいことが分かった。都心をはじめとする高価格の地域は、鑑定評価の需要が非常に多く、また誤評価を行った場合のインパクトも大きい。別の価格モデルを組むことが示唆されるが、容積率ごとに採用する価格モデルを変えることの整合性・説明性は保てるのだろうか。ひとつの提案として、精度向上のためには、用途・容積率・価格水準ごとに価格モデルを換えることは有効と思われる。その際、分岐の定義とモデルは慎重に更新されなければならない。価格モデル設計及び精度向上に関しては、先行研究を参考にしながら、引き続き検討していきたい。



図表: 決定木による価格形成要因分析

日本不動産研究所作成

2. リーク現象

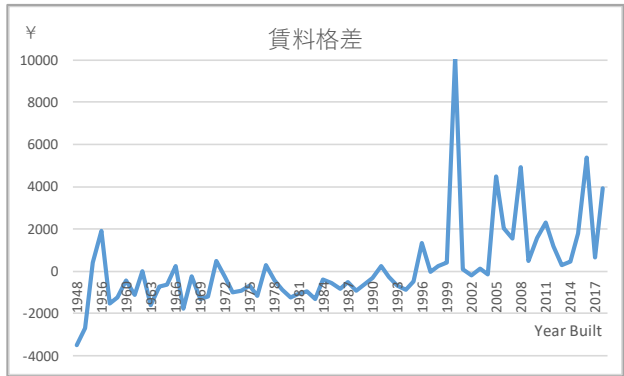
以下は証券化対象不動産の価格事例に対する単回帰の結果である。目的変数は土地価格である。目的変数に対して最も相関係数が高いのは、建物の公租公課という結果になった。このようなリーク現象が起きている価格モデルは実務では使えない。予測モデルの実運用に際しては、各変数が実務フローの中で取得できるタイミングを誤らない検討を要する。

REIT名称	1994	1995	1996	1997	1998	1999	2000	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015	2016	2017	2018	2019		
東武不動産	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	
東武不動産(株)	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
東武不動産(有)	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000

図表: REIT 公表物件の単回帰マトリクス
REIT 公開データに基づき日本不動産研究所作成

3. 変数の定義

IVとVに関して、多変量解析は、変数を連続変数とみなし、線形回帰を前提としてモデル化している。しかし、次表のように、例えば連続変数である築年数と賃料との関係だと相関係数が0.3と低いが、離散変数である建築年次と賃料との関係を散布図にした場合、変動は認められるが一定の傾向が認められる。



もちろん採用に際しては、各年次における事例の数やばらつきを検討して慎重に採否を検討すべきである。この例のように、これまで連続変数とみなして統計的アプローチにより採否を決定してきた各変数についても、統計的機械学習や空間データ解析の手法が身近になったことにより、再度の検討を行っていききたい。

今後

不動産は、人の全ての活動基盤であることから、今や重要なデータ生成装置としても注目されている。デジ

タル経済においては、情報のデジタル化自体が商材となりうることを予感している。第一章においてはデジタル化の目的や基準作りの上で参考になる既存の基準を例示した。データは客観的で公平な「共通言語」たりえることから、情報そのものが潤滑油としての機能を強め、企業活動の意思決定の高速化につながっている。

第二章においては可視化の事例、第三章においては解析の事例を紹介した。いずれも精度面の課題は残っており、引き続き各プロジェクトの顧客と検討を進めていきたい。

不動産関連データの整備力と連携力は、各企業の属する業界におけるプレゼンス向上の促進剤となる。今後、不動産関連情報は、人・モノをはじめとする他分野のデータと結合されることで、より多くの知見を見出す端緒となることが期待される。

なお、本稿において採用した情報及び技術に、新規性はない。しかし、先進的な技術は、ビジネス判断の材料として採用する場合に、その先進性ゆえに避けられる場合もある。よって、決して目新しいものではなく、安定した技術と、先進技術とを、常に検討し、ニーズに応じて両者を機動的に用いてビジネス課題を解決することが重要だと実感している。

参考文献

[1] Brennen, S., & Kreiss, D. (2014). *Digitalization and Digitization* より拙訳引用 Available: <https://culturedigitally.org/2014/09/digitalization-and-digitization/>

[2]一例として、人工知能学会全国大会 2019 オーガナイズドセッション「不動産とAI」 Available: <https://confit.atlas.jp/guide/event/jsai2018/session/2C10-14/detail?lang=ja>

https://www.jstage.jst.go.jp/article/jjsai/32/4/32_529/pdf/-char/ja (Japanese)

[3] 一例として、NYC Real Estate Tech Week 2020 Available: <http://www.realestatetechweek.nyc/>

[4] 3th Conference of interdisciplinary Research on Real Estate. *Digitalization and Real Estate* pp. 28 - 33, Groningen, September 20-21, 2018.

Available: <http://www.cirre.eu/wp-content/uploads/2016/01/Book-of-abstracts-version-DB-08.08.2018-definitief.pdf>

[5] 百合本茂 & 増井忠幸. (1991). *工場立地選定のための意思決定支援システム*, 日本経営工学会誌, 第42巻1号, pp. 37 - 44, 1991

※百合本らは1995年にファジイ線形計画モデルにより立地選定モデルを改良している。

[6] 亀田伸裕, 森信之, 田中邦博, 中山伸介. (2003). *GISを用いた風力発電所立地選定に関する研究* GIS-理論と応用, 第11巻2号, pp. 173 - 178, 2003. "Decision Support System for Selection of Plant Location"

[7] 荒川周造, 諏訪博彦, 小川祐樹, 荒川豊, 安本慶一, 太田敏澄. (2017). *暗黙知に基づく飲食店用不動産の賃料推定システム*. 情報処理学会論文誌 59(1), 33 - 42, 2018-01-15 "Rental fee estimation system of real estate properties for restaurants based on tacit knowledge"

[8] 「飲食店.com」 Available: <https://www.inshokuten.com/bukken/kanto/market/rent/>

[9] 国土交通省 国土数値情報 <https://nlftp.mlit.go.jp/ksj/>

[10] JREI Real Estate Investor Survey: JREI Real Estate Market Report. (2020 First Half) Available:

<https://www.reinet.or.jp/en/form/index.php?fm=en-contact>

[11] National Office and Apartment Rent Index: JREI Real Estate Market Report (2020 First Half)

[12] Urban Land Price Index: JREI Real Estate Market Report (2020 First Half)

[13] The National Land Agency (MLIT)'s annual survey of "listed land prices" Available: https://nlftp.mlit.go.jp/ksj/gml/datalist/KsjTmplt-L01-v2_3.html

[14] Baum, A. (2017). *PropTech 3.0: the future of real estate*. University of Oxford. UK.

[15] 歴史的農業環境閲覧システム「迅速測図」. 国立研究開発法人農業環境技術研究所

Available: <https://habs.rad.naro.go.jp/>

[16] ハザード情報: 各都道府県 HP 千葉県

<https://www.pref.chiba.lg.jp/bousaik/higaisoutei/documents/p06p07.pdf>

<https://www.pref.chiba.lg.jp/bousaik/higaisoutei/documents/p08p09.pdf>

東京都

http://www.bousai.metro.tokyo.jp/res/projects/default_project/page/001/000/401/assumption_h24_outline.pdf (P16-19)

神奈川県及び茨城県

https://disaportal.gsi.go.jp/maps/?ll=35.319327,139.439163&z=12&base=pale&ls=y2k7%7Ctsunamishinsui_raster%2C0.8&disp=01&vs=c1j010u0