

ポートフォリオ

物件の賃料予測と 条件変更の影響推定

目次

1. 概要
2. 背景・目的
3. 賃料予測モデル
4. 不動産投資判断のための因果推論
5. ビジネス導入に向けて

概要

機械学習モデルの作成と因果効果の推定

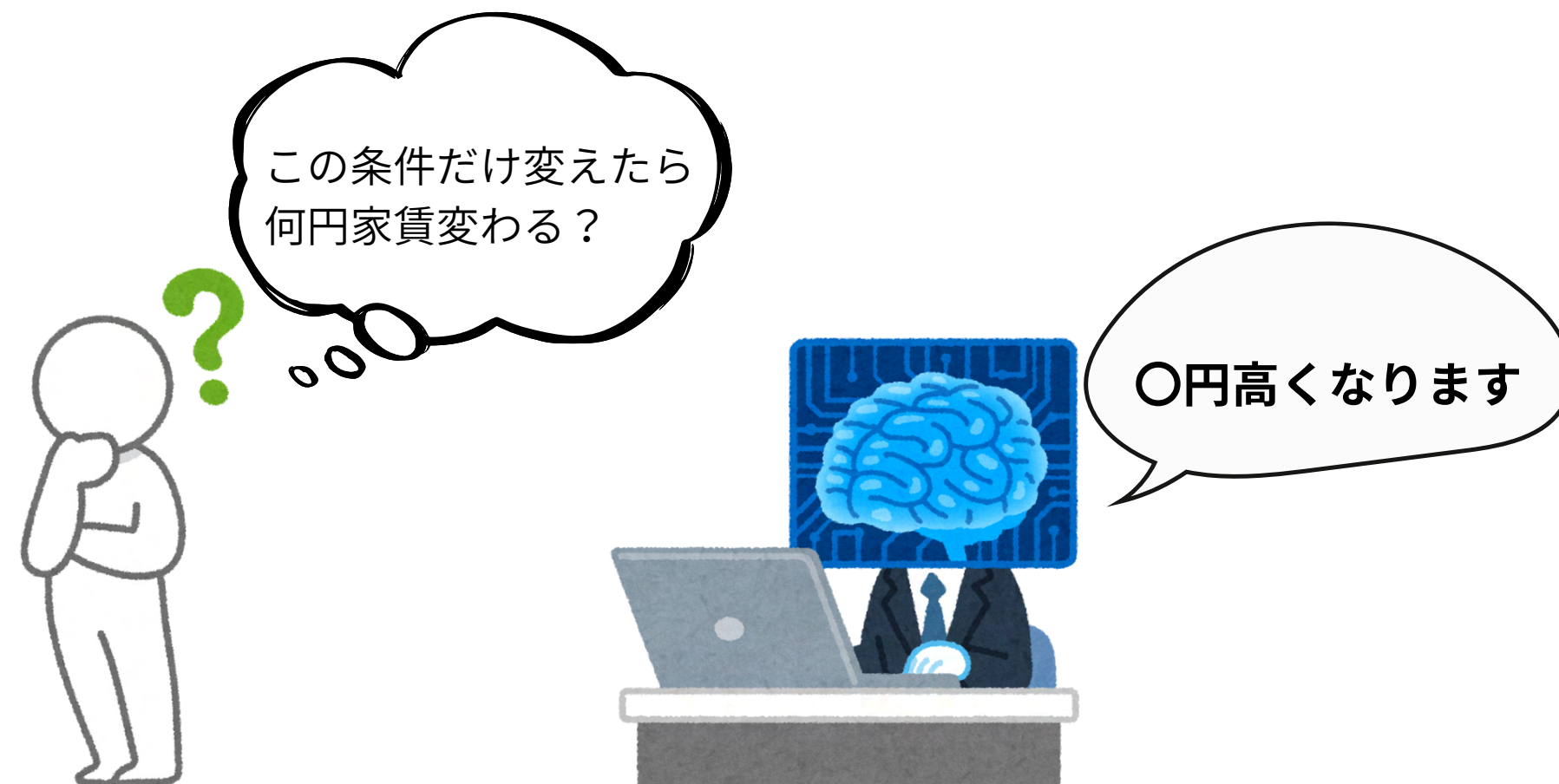
①賃料を予測する機械学習モデルの作成

【物件情報】

- ・ 渋谷駅5分
- ・ 新築
- ・ 10階
- ・ 1LDK
- ・ オートロック



②不動産投資のための因果推論

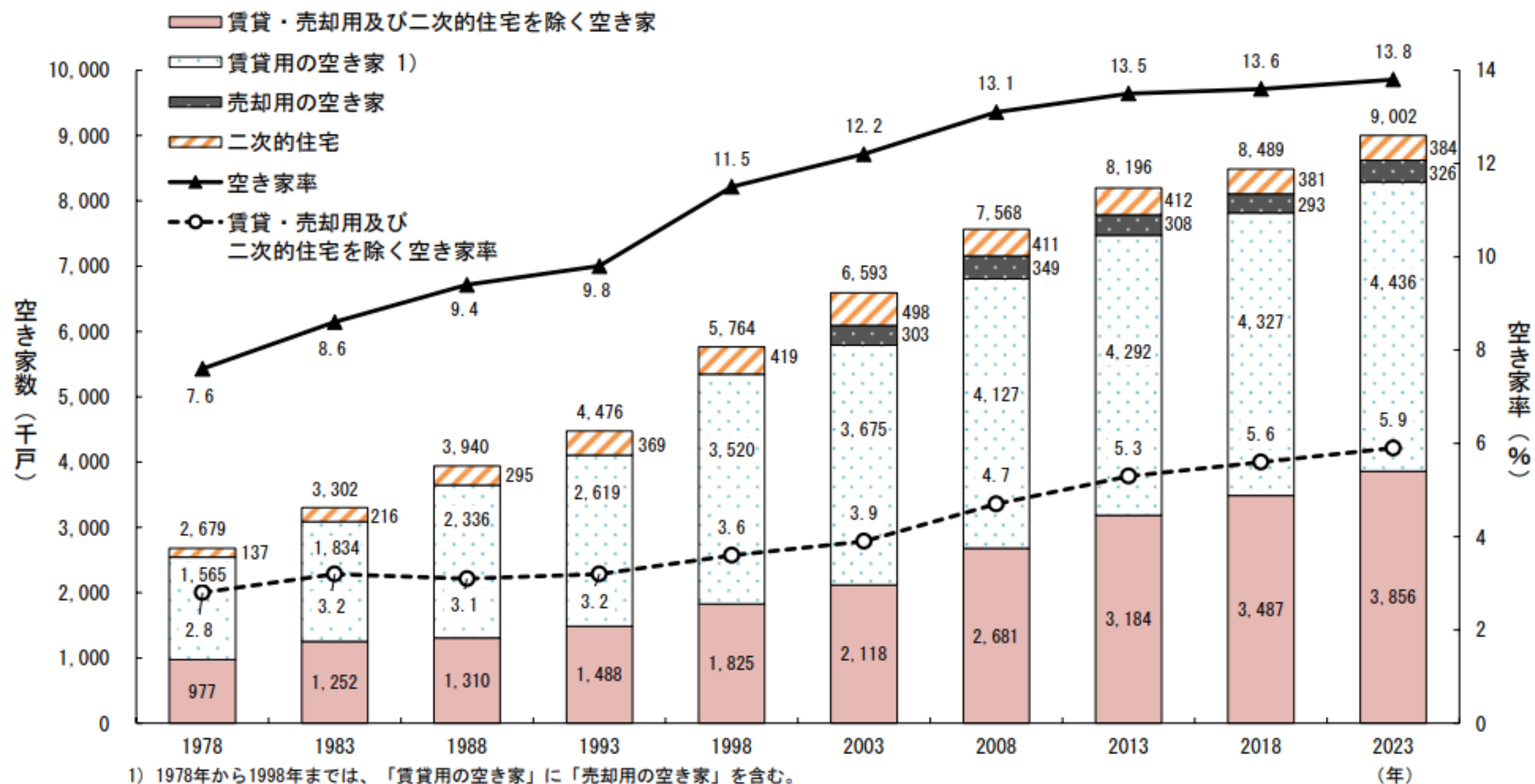


背景 · 目的

背景

空家は増加傾向が続く

図 2-1 空き家数及び空き家率の推移ー全国（1978年～2023年）



出典：「令和5年住宅・土地統計調査」（総務省統計局）

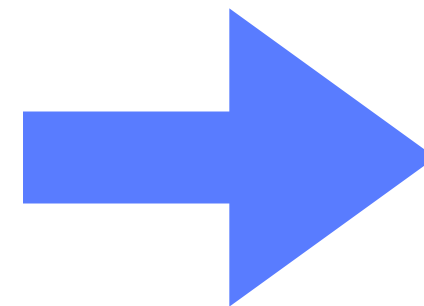
背景

空室リスクと賃料査定の必要性

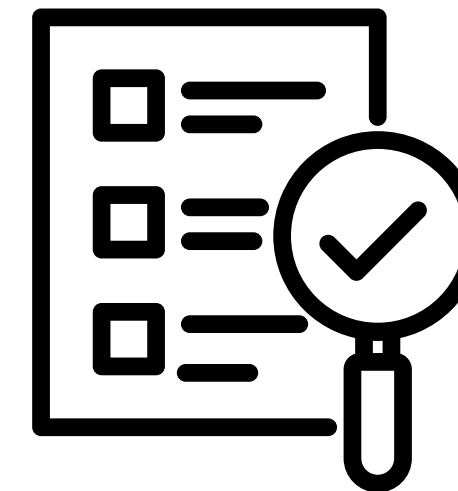
空室リスク増・収益減



人口減少や競合物件の増加により、
空室リスクはさらに高まる
空室率の増加は大きな収益減を招く



「最適な賃料予測」が必要

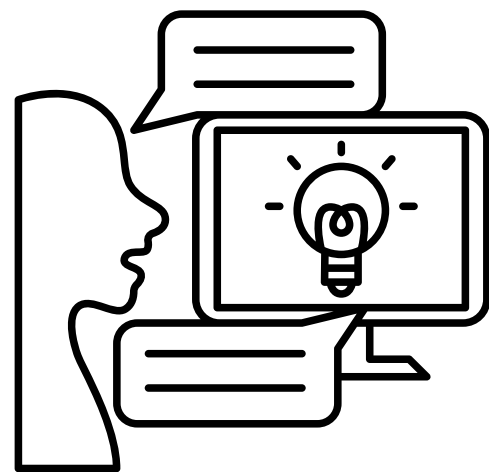


空室リスクを抑えながら収益を最大化できる
適切な賃料を査定する必要がある

今回の目的

機械学習モデル・因果推論の活用目的

査定業務の効率化・自動化



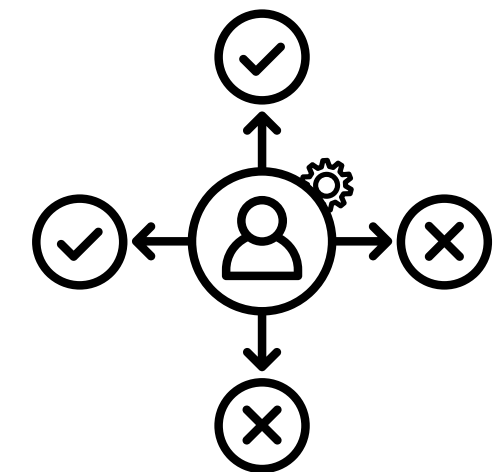
複雑な条件を加味した
賃料査定は時間が掛かるが、
自動かつ短時間で査定可能に

収益の最大化



設備や立地条件など多種多様なデータを
モデルに反映することで、市場に即した
賃料を算出することが可能に

不動産投資の判断材料の提供



不動産投資における
客観的な判断材料を提供することで
より正確な収益シミュレーションや
費用対効果の算出が可能に

データについて

データ・コンペ詳細

SIGNATE 「第1回 国土交通省 地理空間情報データチャレンジ ～国土数値情報編～」
物件情報サービス「LIFULL HOME`S」に掲載された
全国の賃貸マンション・アパートの賃料と物件情報

開催期間	2024/10/15～2024/12/13
評価指標	RMSE(二乗平均平方根誤差)
学習データ	58.4万件(2019年1月～2021年7月掲載分) / 賃料1万円~350万円の物件
テストデータ	38.4万件(2022年1月～2023年7月掲載分)
特徴量	物件情報・立地情報を中心に300以上
欠損値・外れ値	実データそのままのため、大量に欠損値・外れ値が存在
その他	外部データとして国土交通省整備のオープンデータ「国土数値情報」の利用が必須 座標情報をもとにした効率的なデータフレームの結合が必要



賃料予測モデル について

アプローチ

予測対象(目的変数)の変更

課題

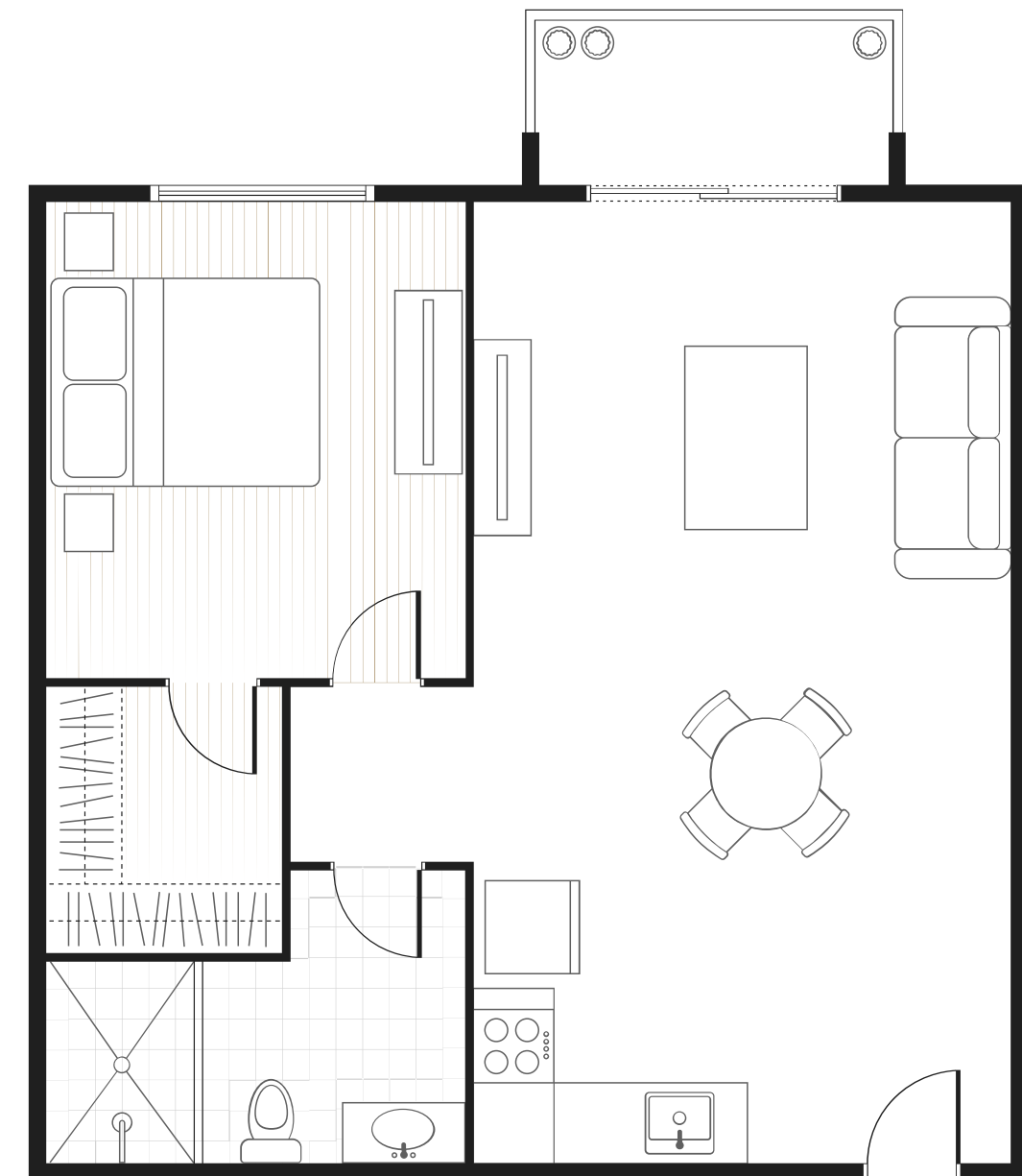
賃料予測に「専有面積」が非常に強く影響
他の特徴量(築年数や駅からの距離)が学習されにくい問題が発生

対処

予測対象を「単位面積当たり賃料」へ変更
 $\text{予測値(単位面積当たり賃料)} \times \text{専有面積} = \text{賃料(元の予測対象)}$

効果

面積の影響を軽減し、他の特徴量をより適切に評価できるように
モデルの予測性能が向上



アプローチ

地域特性の反映

背景・問題意識

都心の一等地と郊外・地方では、同じ物件条件でも賃料は大きく異なる
地域特性をうまくモデルに学習させることが予測精度向上のカギ

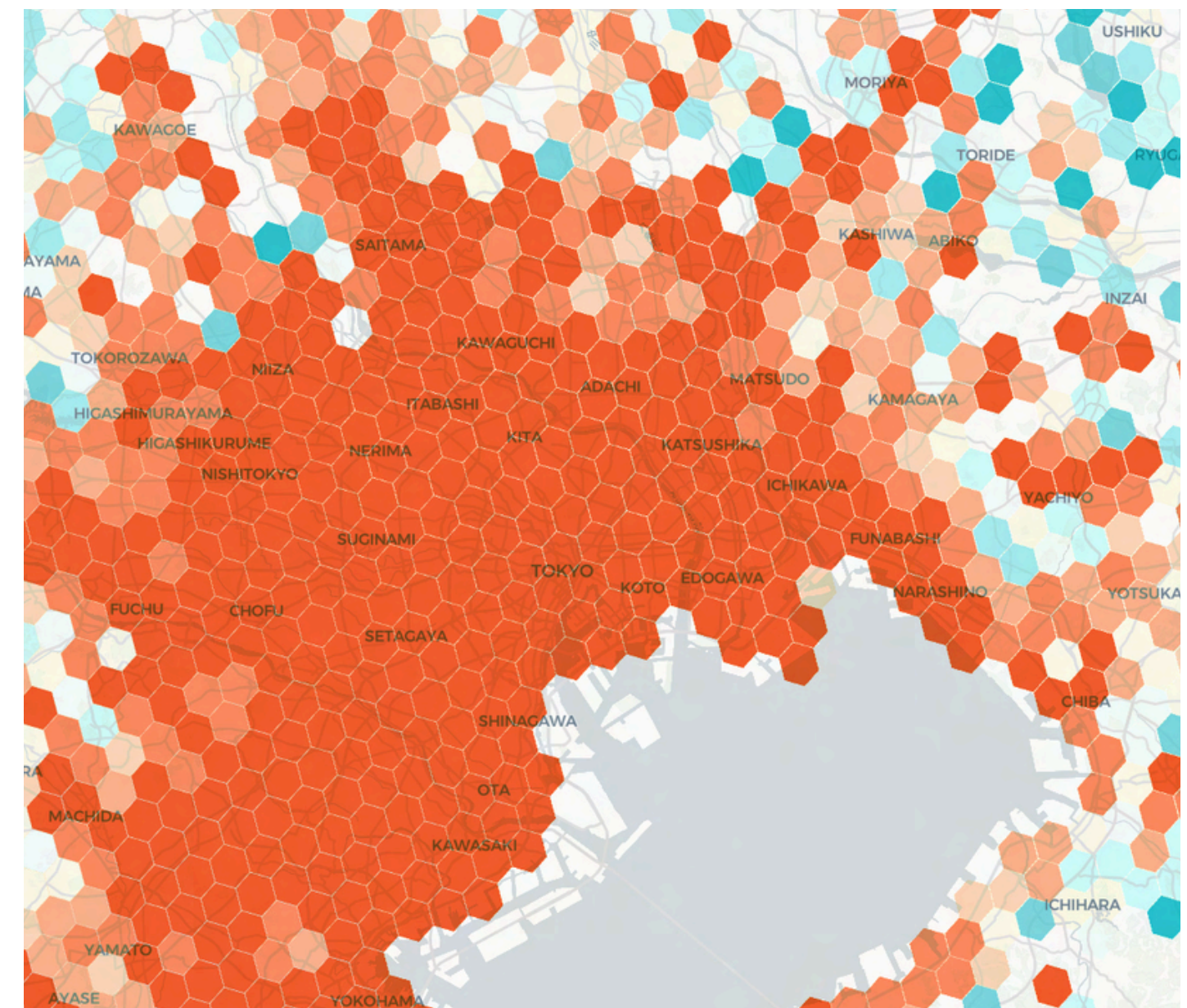
方針

- ・ H3インデックス（グリッドシステム/右図）
 - ・ 郵便番号単位
 - ・ 最寄り駅単位
- など複数スケールで地域を分割

実装

単位面積当たり賃料の平均*を
複数の地域区分で算出しモデルに学習

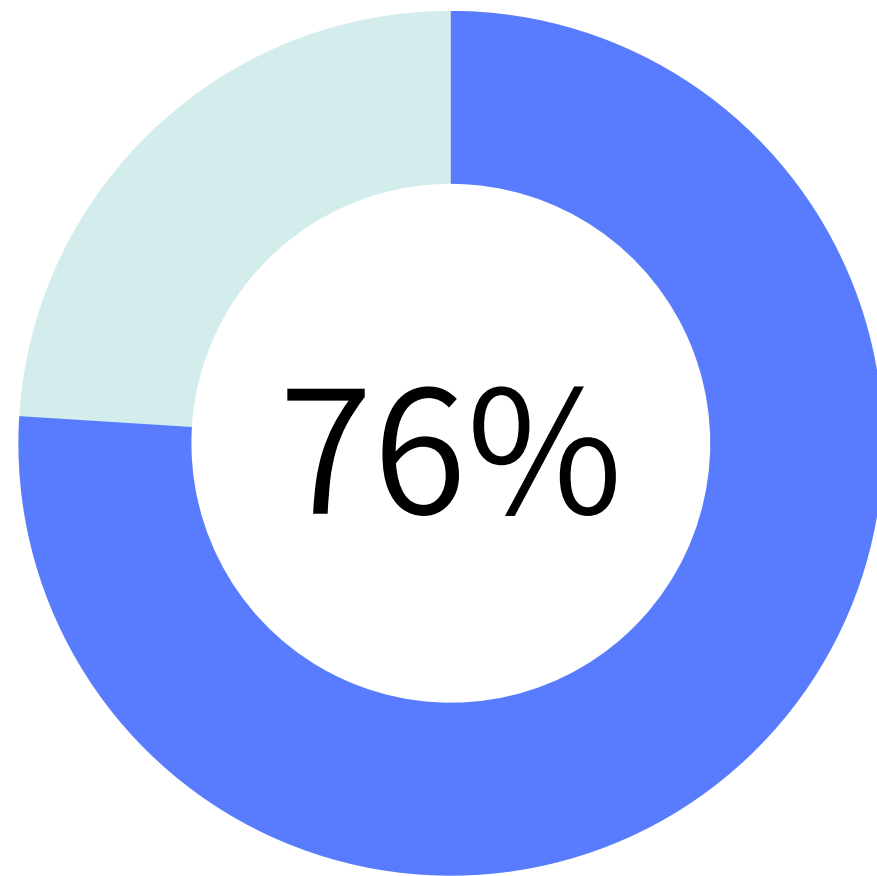
*正確にはターゲットエンコーディングした値



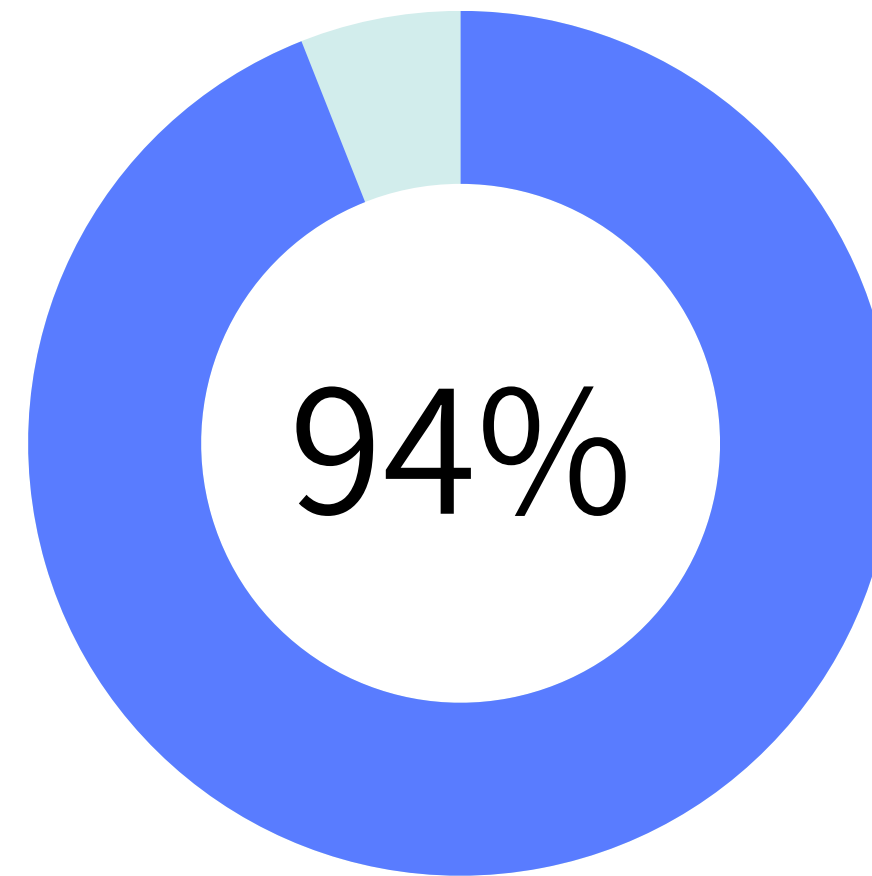
モデルの精度

誤差(価格)

誤差5千円以内の物件割合



誤差1万円以内の物件割合

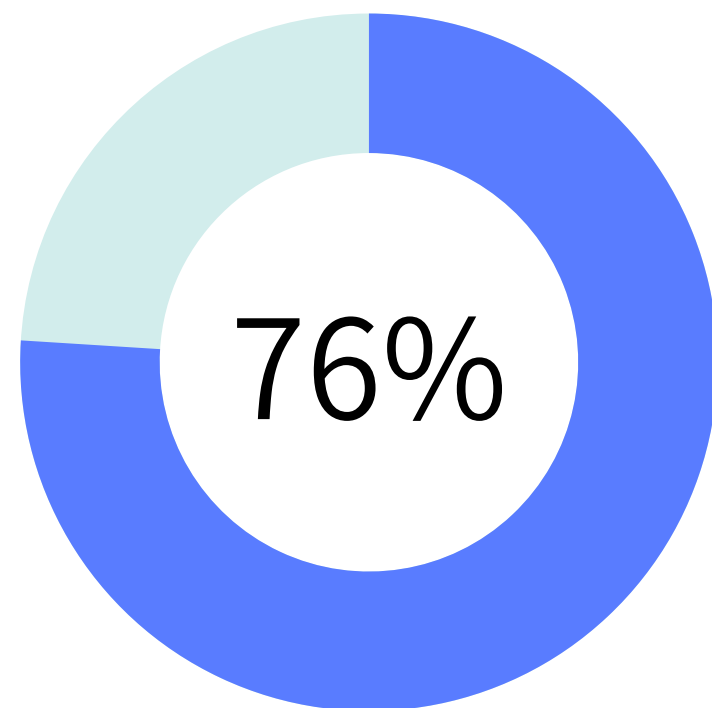


※学習データでの交差検証の精度であり、未知のデータに対する精度ではありません。

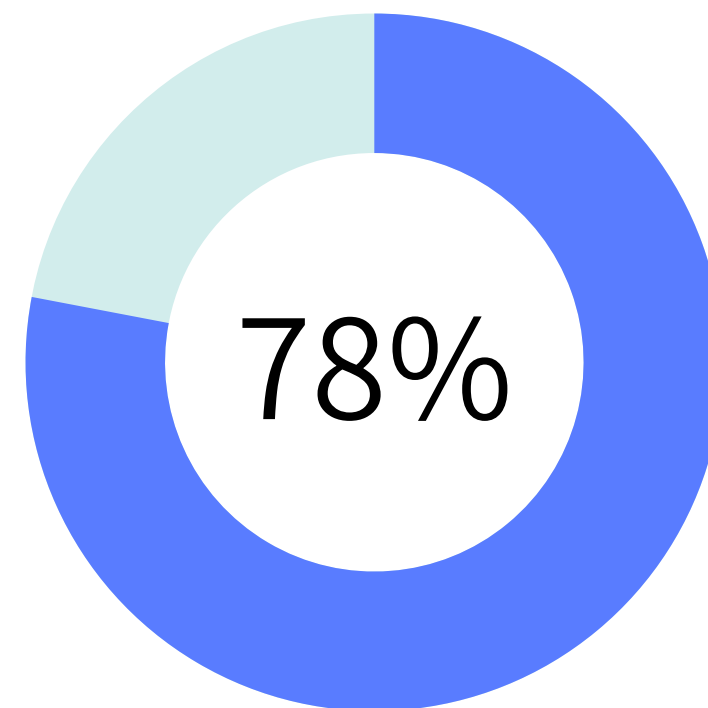
モデルの精度

誤差率

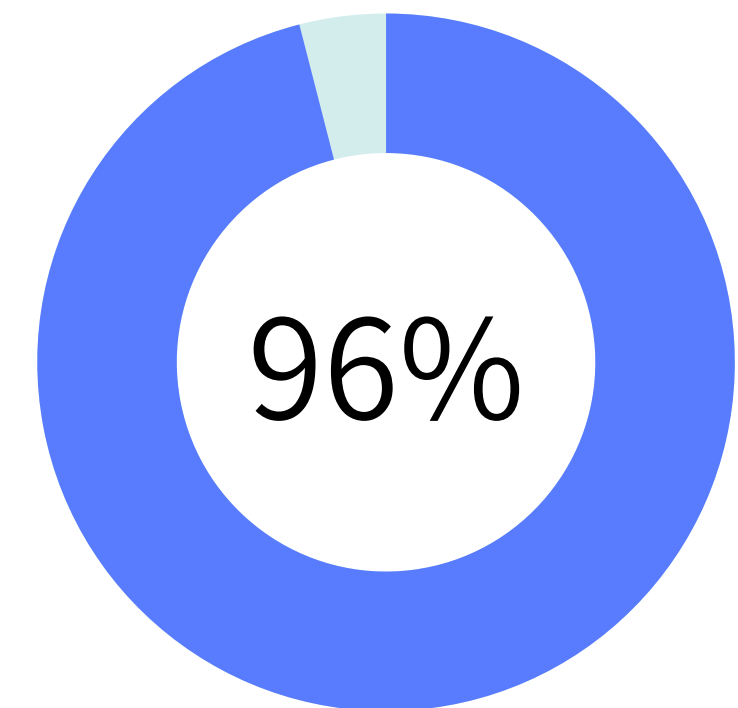
誤差5%の物件割合



誤差10%以内の物件割合



誤差20%以内の物件割合



※学習データでの交差検証の精度であり、未知のデータに対する精度ではありません。

結果

コンペ成績

銀メダル獲得

参加者数528人中13位(チーム参加・アンサンブルweight0.5)
個人のモデルでは35位相当(銀メダル圏内)

タイプ	コンペティション名	開催期間	賞金総額	投稿数	順位
田	第1回 国土交通省 地理空間情報データチャレンジ ～国土 数値情報編～ モデリング部門	2024/10/15 ~ 2024/12/13	1位 10万円、2位 5万円、3位 3万 円、4～10位 賞状	36	● 13 位 / 528人投稿

⇒テストデータに対しても一定の精度を確保できていた

企業賞(LIFULL賞)受賞

評価基準：予測が難しい物件（外れ値など）に対して斬新的な発想で精度向上に取り組んだ分析レポート
【レポート内容】

- ① 学習データの目的変数の外れ値の検出と補正
- ② 専有面積のカラムの外れ値の補正
- ③ 空間結合を用いた効率的な最寄り駅情報の結合

予測値の根拠を解釈する

ビジネス導入には「なぜその予測値になったのか」が必要



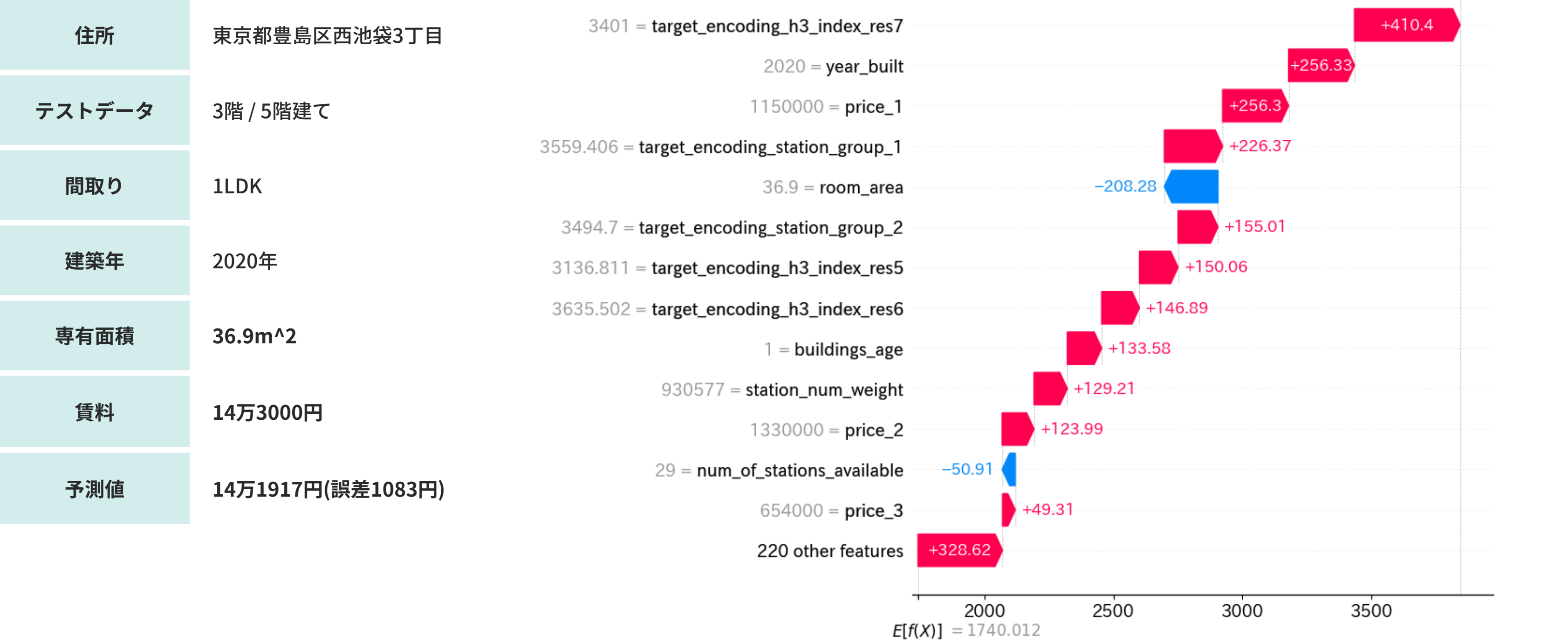
この物件は15万円
この物件は120万円
この物件は1万円...



適当に予測してない？

予測値の根拠を解釈する

SHAPを用いた個別の物件に対する予測の解釈



不動産投資判断のための 因果推論

不動産投資判断のための因果推論

なぜ因果推論を用いるのか

背景

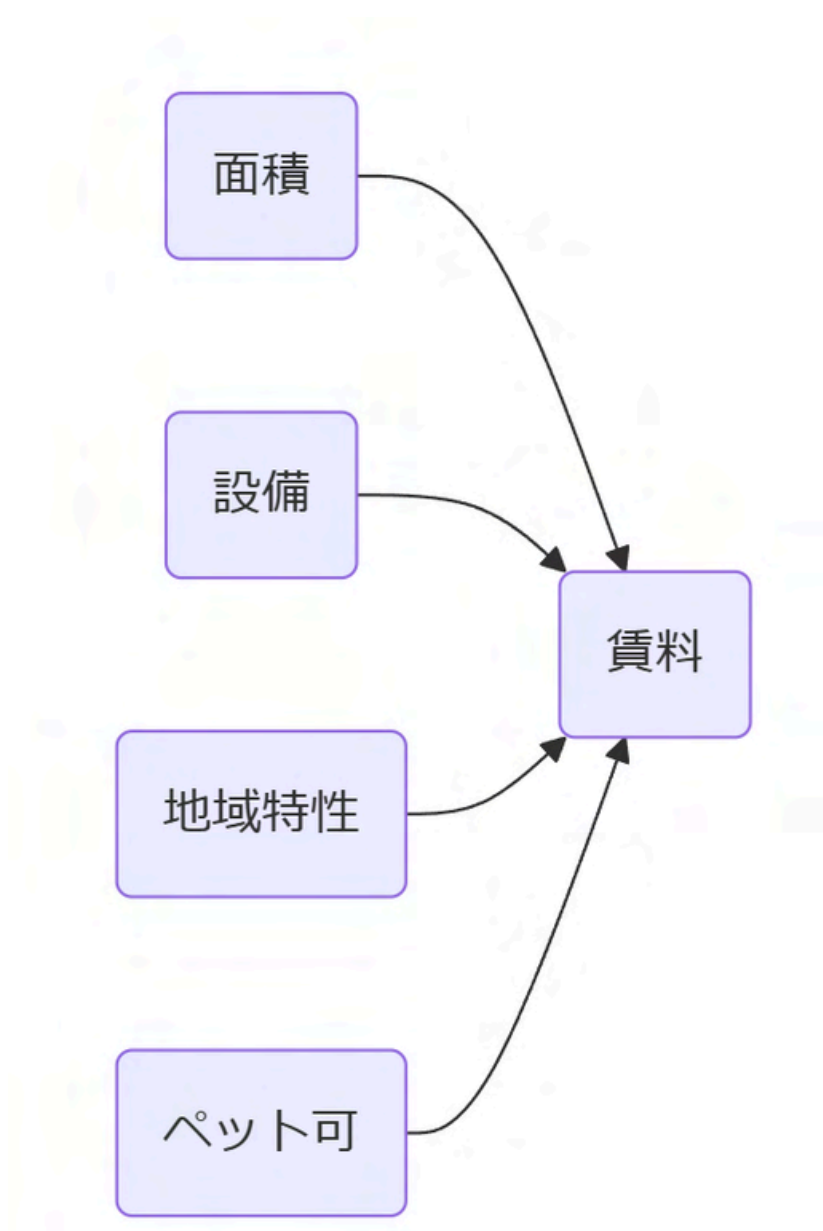
不動産投資では、賃料設定以外にも物件の条件変更を検討する場面が多い
例) ペット条件、リフォーム・設備投資など

想定事例

既存の物件の収益向上のために
「ペット可への切り替えによりどの程度賃料を上げられるか？」を知りたい

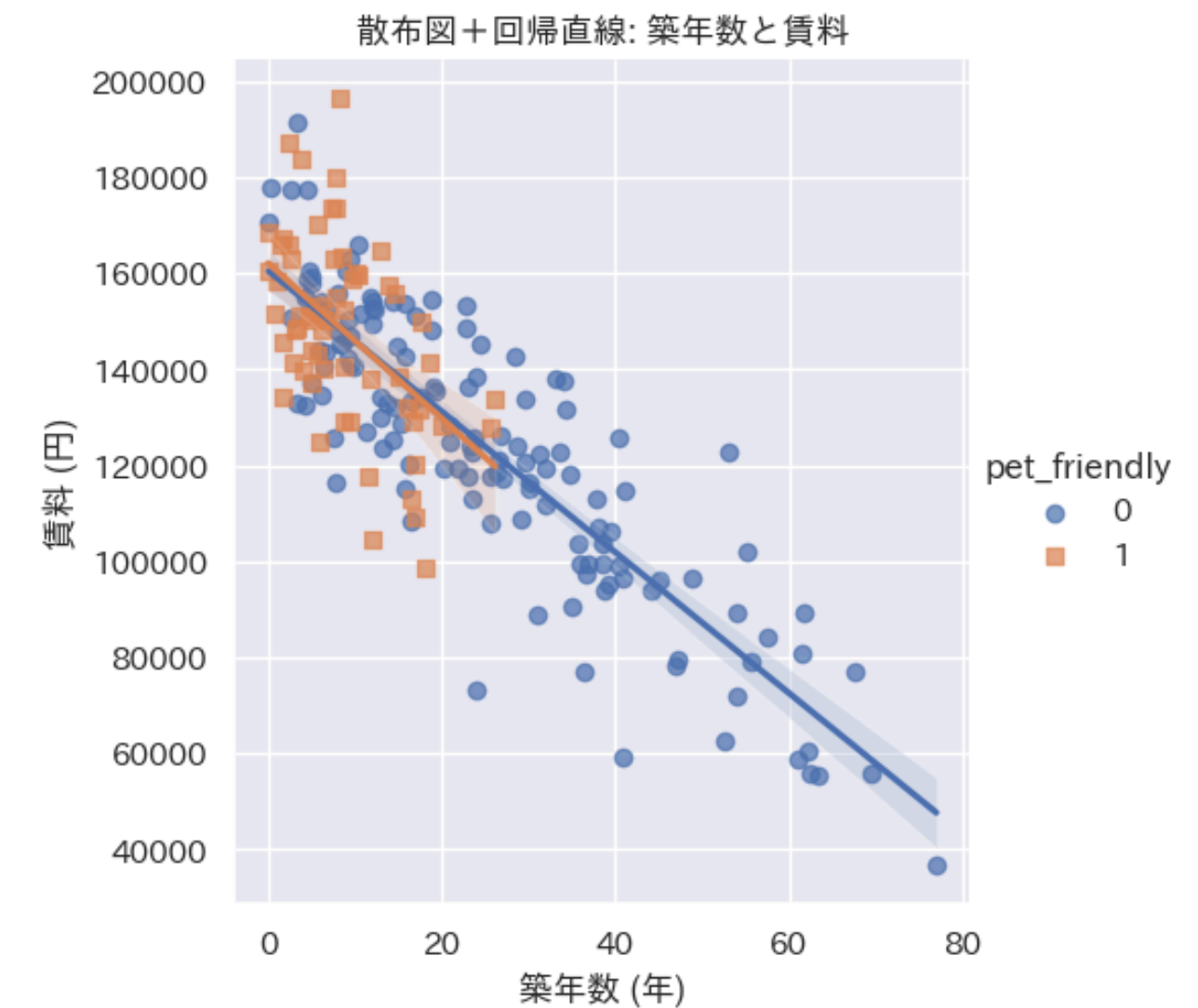
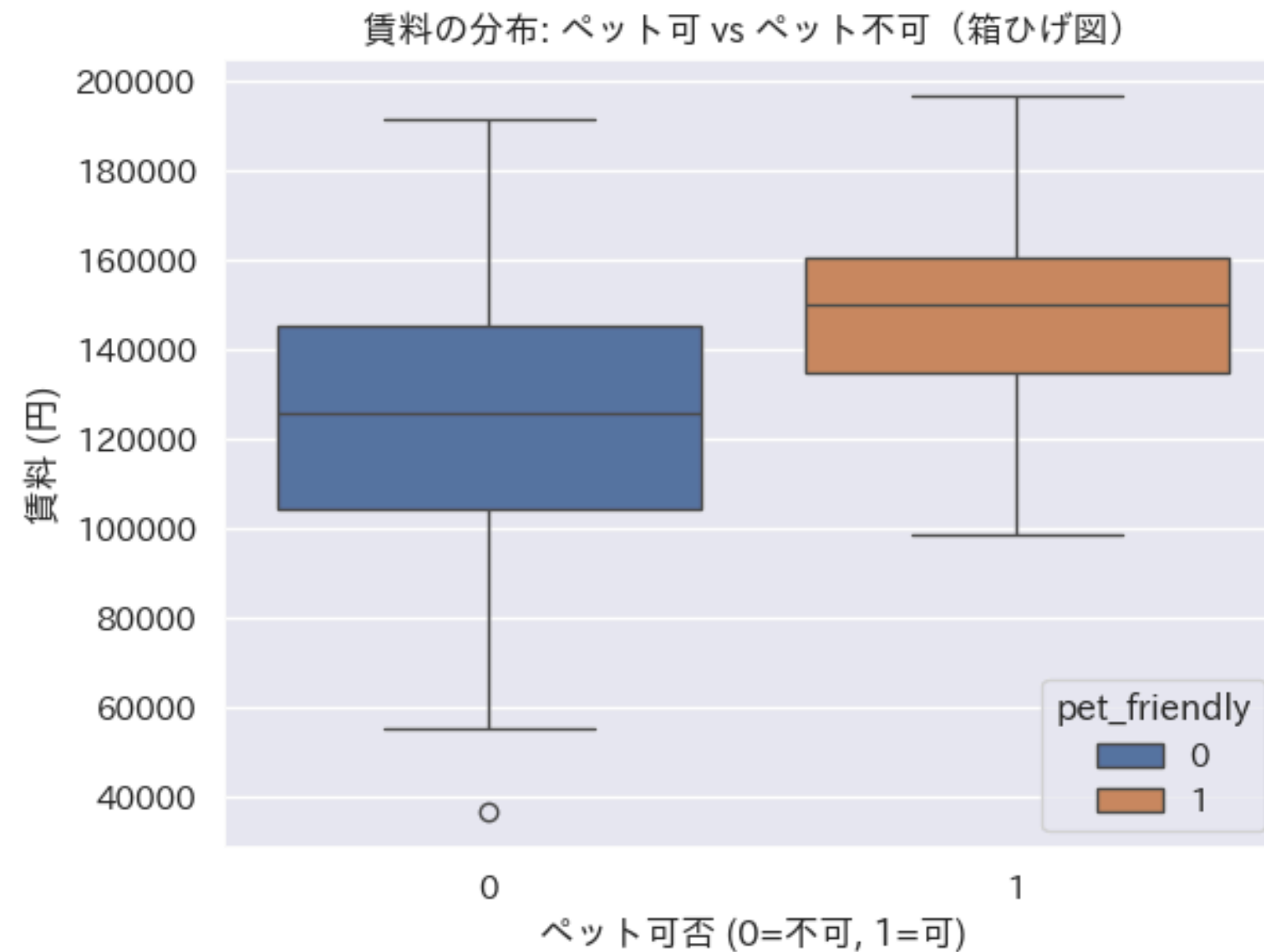
機械学習の問題

機械学習モデルに単純に「条件をペット可に変更した」データを入力すると、
既存の相関関係をそのまま引き継ぐためバイアスのかかった結果になる可能性がある。



不動産投資判断のための因果推論

疑似相関



$$\text{賃料} = \text{切片} - (\text{築年数} \times 1500) + \text{ノイズ}$$

⇒条件を変えただけでモデルの予測賃料が変わったとしても、
その条件の因果効果を説明しているわけではない

不動産投資判断のための因果推論

傾向スコアマッチングを用いた因果効果の推定

因果推論の必要性

条件変更による「本当の効果」を推定するためには、
変数同士の相関や既存データの偏りを排除する必要がある

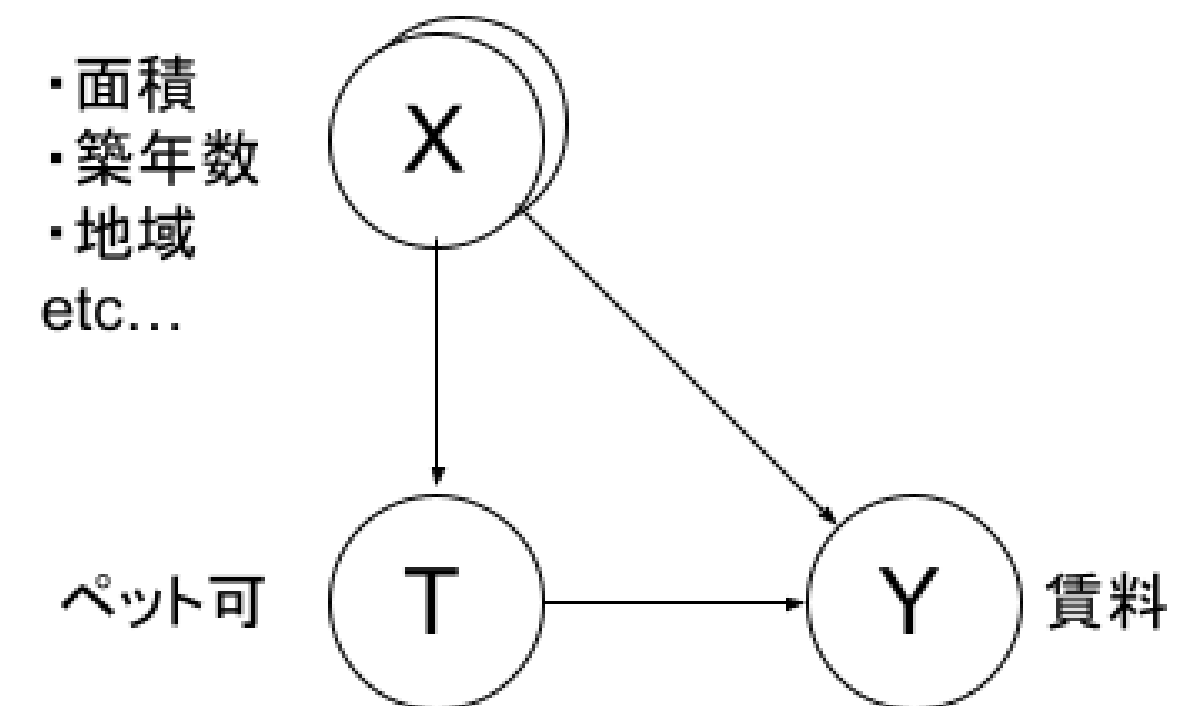
アプローチ

傾向スコアマッチングによって「ペット可物件」と「ペット不可物件」を
できるだけ条件が近いサンプル同士で比較することでバイアスを排除する

結果

ペット可にすることで1m²あたり25円の賃料の増加が期待される。

※条件変更による「空室期間の短縮」「入居期間の長期化」など
賃料以外のメリットも考慮する必要がある。





ビジネス実装に向けて

ビジネス実装に向けて

改善策

「空室期間」のデータ取得

この機械学習モデルは既存の物件が全て「**適正価格**」である前提

すぐに入居が決まる物件



適正より安い？

なかなか入居が決まらない物件



適正より高い？

⇒予測値にバイアスが含まれている可能性

ご清聴ありがとうございました