

機械学習を用いて見積価格の予測 を行い、作成コストの削減を狙う

インテグレーションステップ

2020/07/05 西久保 祥二

Qiitaに掲載している内容について

今回のプレゼンテーションの概要をQiitaに掲載しました。
(そちらの方がボリュームを抑えてありますので、見易いかと思います。)

限定共有となっておりますので、下記URLよりアクセスいただきますよう
よろしくお願いいたします。

https://qiita.com/sushio_nishin/private/b400b67a035eb2f16a90

なお、こちらの資料もPPTのみで内容を理解していただけるように努めて作成しています。そのため文字情報が多い内容となっております。ご了承ください。

目次

分析の目標・背景知識

- ・ 予測・分析のターゲット
- ・ 概算見積とは？
- ・ テーマ選定の背景
- ・ 業界の構造
- ・ 概算見積作成のフロー
- ・ 概算見積の内容
- ・ 該当部署の抱えるビジネス課題
- ・ 企業の抱えるビジネス課題
- ・ ビジネス課題、解決後の世界

機械学習

- ・ A 使用するデータ
- ・ B データの準備
- ・ C データの前処理
- ・ D 予測モデルの構築・実行
- ・ E モデルの性能評価・考察
- ・ F 機械学習モデルの応用
- ・ G 概算見積の作成で実装するにあたって
- ・ H ビジネスインパクト

結び

- ・ 今後さらに実施してみたい事
- ・ 課題に取り組むにあたって
- ・ 最後に

予測・分析のターゲット

- ・ 予測する見積価格

自動火災報知設備（以下自火報）の概算見積金額の予測を目標とする。

- ・ 自火報とは

読んで字のごとく、自動で火災を報知してくれる設備。

ショッピングセンターや学校など、人の出入りする建築物に設置されている。

煙や熱などで火災を感知し、実際に火災が発生すると非常ベルや音声装置によって、建物内の人に火災を知らせる。

消防法という法律で設置が義務付けられており、法令遵守で設置を行い、点検に合格しないと建物の使用許可が行政から出されない。



概算見積とは？

- ・ 見積は大きく分けて 2 種類

概算見積を顧客に提出した後、実際に建設計画が行われ、通常見積を作成する流れとなる。
今回のターゲットは、精度の求められない概算見積である。作成コストの削減を目指す。

	内容	顧客が求める事
概算見積	建設予定の建築物を、実際に建設する場合にかかる費用を概算で把握するための見積。 どのくらいの費用が必要か、顧客が判断するための指標となる。	<ul style="list-style-type: none">・ 大まかな金額の把握・ 依頼から提出までの納期の短さ・ 修正、変更に対しての対応のしやすさ (実際に機器の数量がいくつ必要か、どのように設置するかなど、具体的な内容の把握は現時点では必要ない)
通常見積	実際に設計図の通り建築物を建設した場合に必要な機器の種類、数量を把握し、費用の詳細を把握するための見積。	<ul style="list-style-type: none">・ 正確な金額・ 正確な機器の詳細・ 正確な数量 (具体的な内容の把握が必要とされる。)

テーマ選定の背景

- ・ 以前従事していた仕事内容

自身が以前に従事していた仕事内容である。
見積の作成や設備図面の設計を行う中で、現状の非効率な業務の改善策を考える事が多く、また考えることに楽しさを感じていた。

- ・ 概算見積の作成は、生産性の乏しい業務である

概算見積の作成には、日常業務のうち1/5程度の時間を費やしていた。時間を費やしている割に、物件の受注に直接繋がること、自社の売上への貢献度は非常に乏しい。
その他受注が絡んでいる物件の対応で、設計・積算業務は十分に逼迫していた。
利益に繋がらない顧客の概算見積の依頼を拒否することができれば、業務効率の向上を図れる。
しかし現実的ではなく「いかに作成コストを減らせるか」その点に機械学習を応用できると仮定し、今回の課題を設定した。

- ・ 具体的な業務に機械学習を適用したい

機械学習を用いて業務効率の改善に貢献することを目標とし、
具体的な業務に落とし込み、実装までの過程を経験してみたいと考えた。

業界の構造

・建築業界は、典型的なピラミッド構造

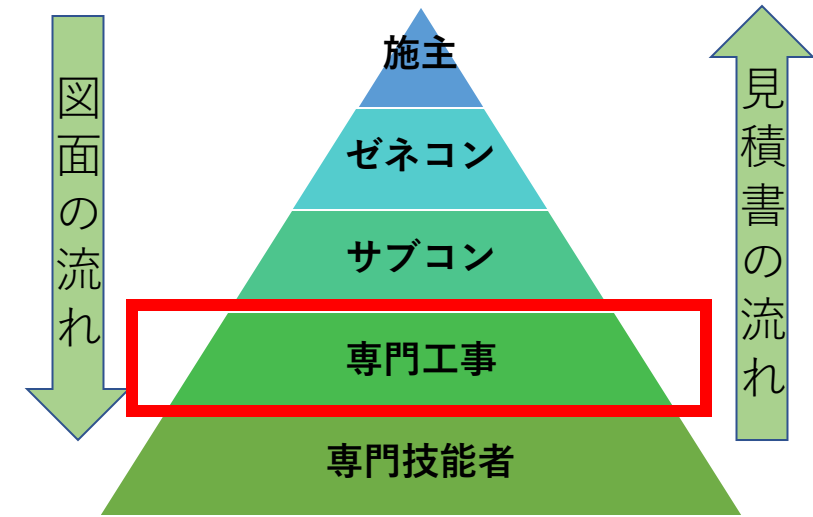
施主 : 建設工事の発注者

ゼネコン : 施主から全てを請け負う元請事業者

サブコン : ゼネコンから工事の一部を請け負う建設事業者（電気設備など）

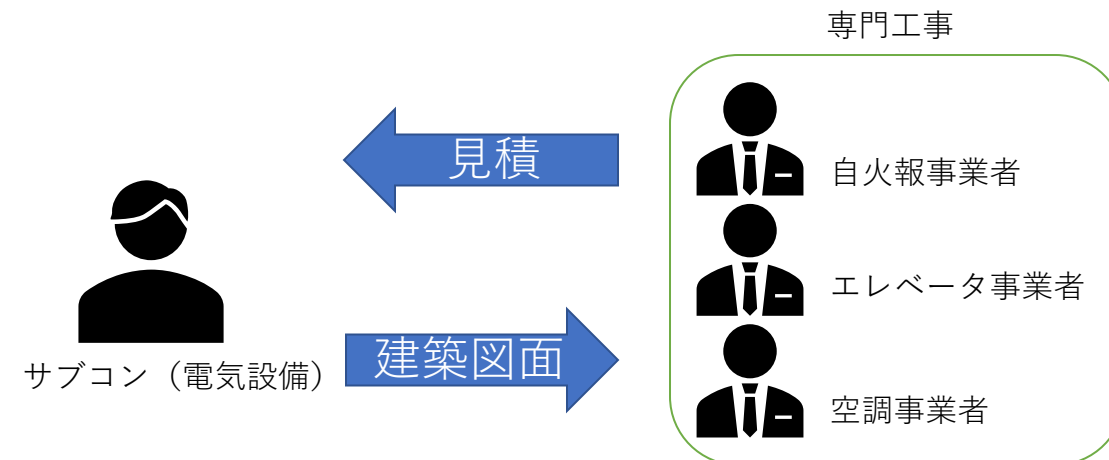
専門工事 : 自火報設備やエレベータ等、専門工事事業者 **自身の立ち位置**

専門技能者 : 実際に現場で工事を行う技能者



・各事業者間の関わり方

施主・ゼネコン・サブコンから建築図面を受領し、自社の設備費用を算出。
見積書として顧客に提出。
(直接施主から依頼を受ける場合もあるが、基本の流れは同様)



概算見積作成のフロー

顧客 : 見積の依頼、建築図面の提出



営業 : 図面の受領、設計への概算設計の依頼



設計 : 図面を元に自火報設備を手書きでプロット、概算設備図面の作成

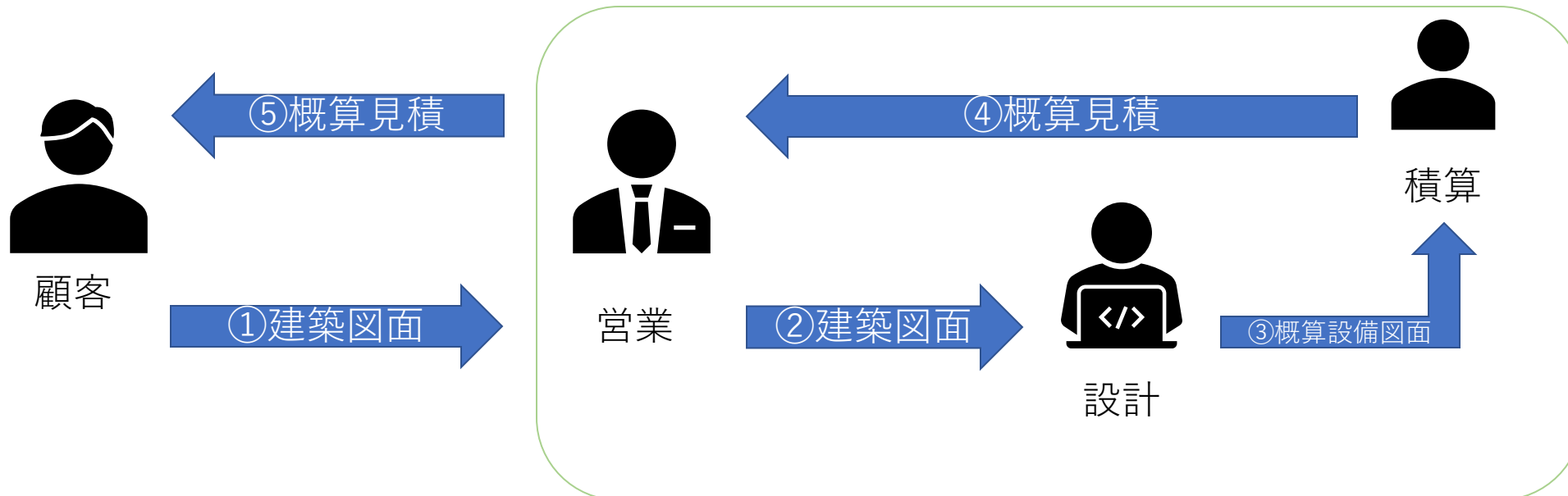


積算 : 概算設備図面を元に数量の拾い出し、概算見積の作成、営業への提出



営業 : 見積金額の調整、顧客へ提出

作成コストの高い部分



概算見積作成のフロー

営業 : 図面の受領、設計への概算設計の依頼



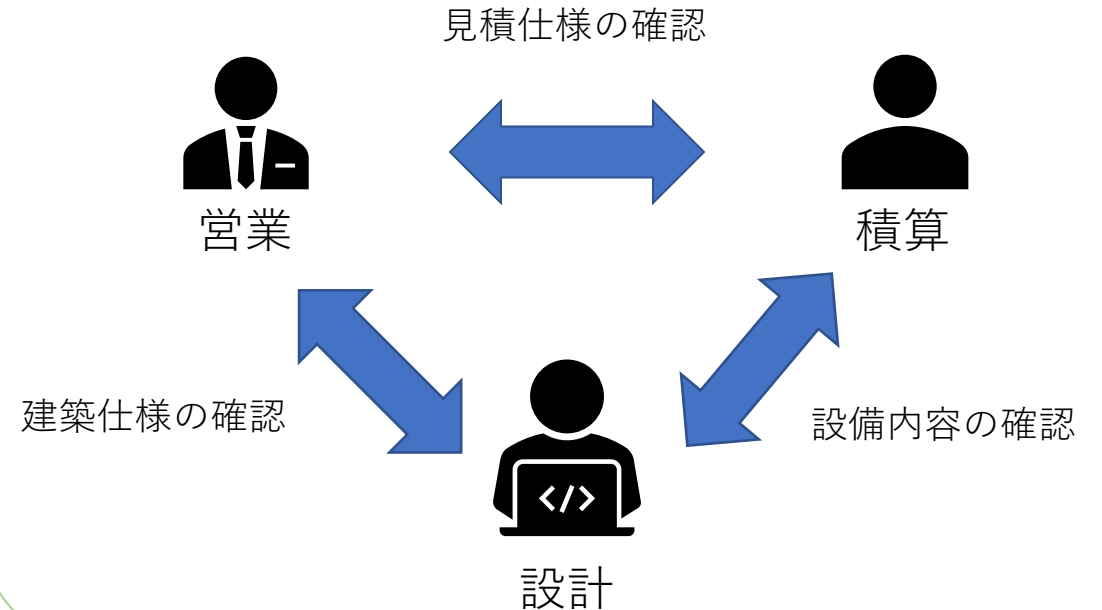
設計 : 図面を元に自火報設備を手書きでプロット、概算設備図面の作成



積算 : 概算設備図面を元に数量の拾い出し、概算見積の作成、営業への提出

・作成コストの高い業務

図面の作成や、見積の作成に時間がかかる。
各担当者間で行う意思疎通のコストが非常に高い。
建築の仕様、設備の仕様、見積の仕様に関して、
情報共有を行うことに時間を費やす。



概算見積の内容

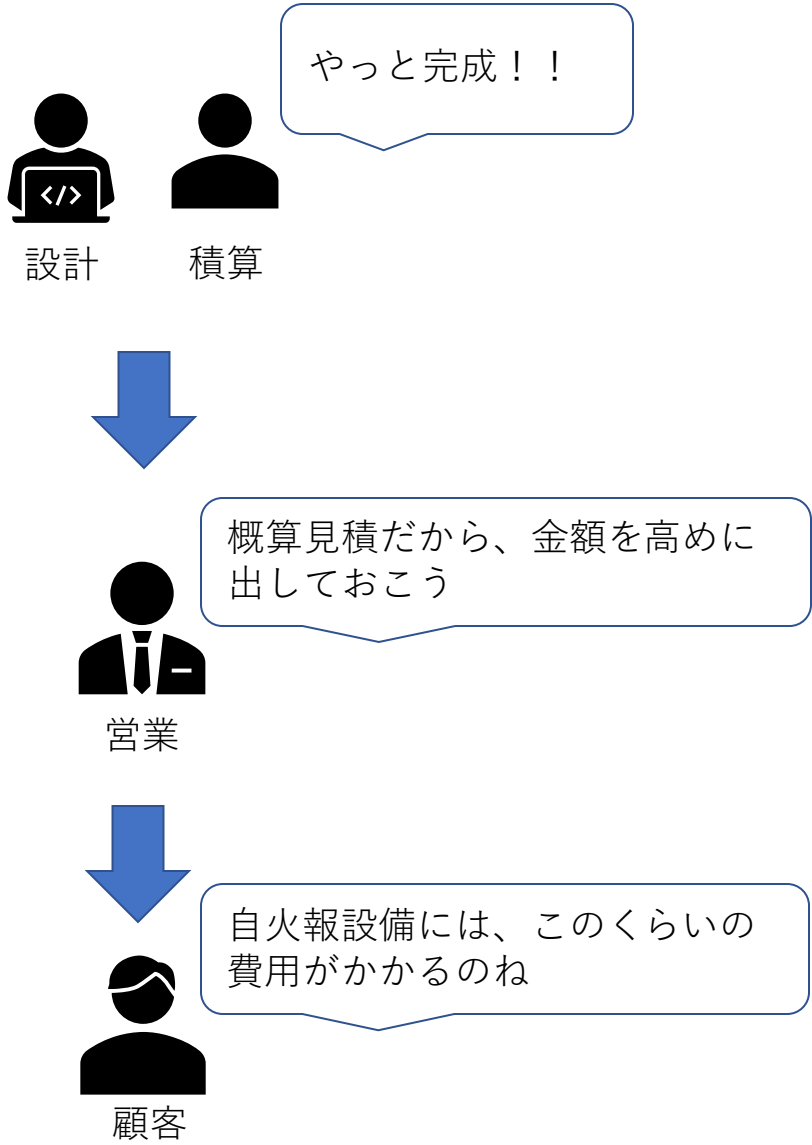
・ やっと完成。概算見積の内容

件名	〇〇マンション建築工事		
トータル金額	¥〇,〇〇〇,〇〇〇円		
機器の数量	受信機	×1	
	煙感知器	×20	
機器費計	¥〇〇,〇〇〇円		
施工費・経費・雑費	¥〇〇,〇〇〇円		

・ 営業による見積金額の調整（なんとなく、経験則）

・ 顧客は概算金額を把握
（顧客はトータル金額のみの把握が目的。概算では機器の詳細は重要視しない）

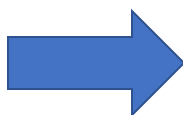
件名	〇〇マンション建築工事		
トータル金額	¥〇,〇〇〇,〇〇〇円		



該当部署の抱えるビジネス課題

- ・ 顧客 : 円滑な概算金額の把握が目標
突発的な見積依頼の慣行。
精度は重要視せず、トータル金額の把握をしたい。
依頼から見積の受領まで、レスポンスの早い対応を求める。
- ・ 営業 : 見積を早急に顧客へ提出したい
トータル金額は適宜調整するため、設計・積算の算出工程は
見積金額の一つの指標にすぎない。
- ・ 設計 & 積算 : 概算見積に費やす時間的余裕がない。
受注に関わる物件の対応で、業務は十分逼迫している。
突発的な概算依頼によって、作業工程の見直し、修正が頻発している。
- ・ 企業 : 設計・積算業務は利益を直接もたらさない
設計や積算にかかる設備費、人件費は、受注物件の工事費から充てている。

概算物件・受注物件にか
かわらず、設計費・見積
費は一切顧客に費用を請
求していない。



設計・積算自体が非プロ
フィット部門であり、会
社としては生産性を落と
すことなく削減したいコス
トである。

企業の抱えるビジネス課題

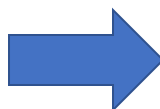
自火報メーカーは自社の利益率向上のために、非プロフィット部門の経費削減を目指している。



自火報の設備設計・積算に関しては、受注した物件の工事費から人件費等の経費を捻出している。そのため、受注物件の数を減らすことなく、業務コストの削減を検討している。



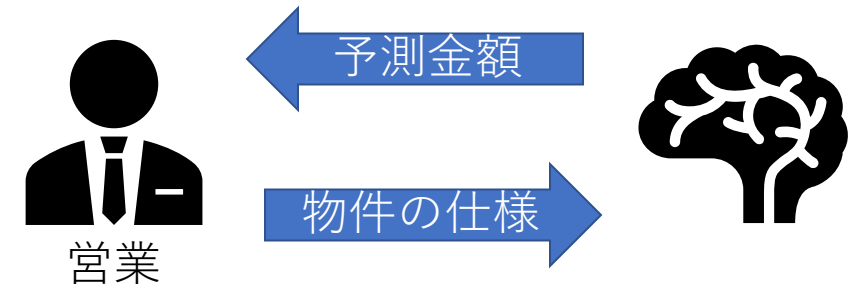
受注確率の低い物件や、精度を必要としない物件に関しては、少ない業務コストで顧客の要望に応える事で経費の削減に繋がる。



誰が	:	営業担当者が
いつ	:	顧客から概算見積依頼を受けたとき
何を	:	見積金額を
どうすれば	:	機械学習を用いて、物件の仕様から金額予測を行う事ができれば
何ができる	:	設計と積算の業務コストの削減に寄与し、会社の経費削減に貢献できるのではないだろうか。

ビジネス課題、解決後の世界

- ・ 顧客：概算見積を受け取るまでのレスポンスが早い。
- ・ 営業：機械学習を用いた概算見積作成プログラムのおかげで、客先からの急な要求に時間をかける事なく対応できる。
- ・ 設計：設備図面の必要となる物件の対応に時間をかける事ができる。（クオリティ向上）
- ・ 積算：見積依頼が削減し、時間的余裕が生まれる。
見積の見直し時間の確保、ミスの低減に繋がる。
- ・ 企業：人件費、残業の削減により、利益率が向上する。



機械学習を用いて、ビジネス課題の解決を目指す！

ビジネス課題の解決に向けて

いよいよ、機械学習の出番！

A 使用するデータ

- ・今回使用するデータ、及び収集方法

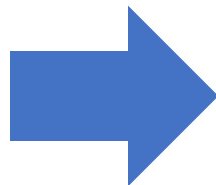
物件掲載サイト「SUUMO」にて、東京23区の賃貸物件のデータをスクレイピング

- ・賃貸物件データを用いる理由

実際の社内データの使用ができないため。

「間取り・専有面積・階層」などの建物の情報が、目的変数となる「賃料」を説明していることが予想されるため、見積作成モデルへの応用が可能と仮定した。

SUUMOの賃料予測	
賃料合計（目的変数）	¥150,000
立地	新宿区
専有面積	35m ²
建物の高さ	11階
間取り	2LDK



概算見積の金額予測	
見積合計（目的変数）	¥5,000,000
建物用途	マンション
専有面積	3500m ²
建物の高さ	5階
住戸の数	20

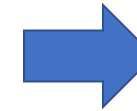
類似した変数を用いることができるのではないか？

B データの準備

- スクレイピングで収集した情報（特徴量）

The screenshot displays a real estate listing for 'FREESIA TERRACE' in Tokyo. The top section includes a photo of the building, its name, and location details. Below this is a table of apartment units. The first unit is highlighted with a red box. To the right of the table, there are buttons for '詳細を見る' (View Details) and 'お問い合わせ (無料)' (Contact Us (Free)).

階	賃料/管理費	敷金/礼金	間取り/専有面積	お気に入り
8階	15.9万円 6000円	敷 15.9万円 礼 31.8万円	1LDK 40.38m2	<input type="checkbox"/> 追加
8階	15.9万円 8000円	敷 15.9万円 礼 31.8万円	1LDK 40.38m2	<input type="checkbox"/> パノラマ 追加



マンション名
所在地
最寄駅情報
築年数
建物の高さ
物件の存する階
賃料・管理費
敷金・礼金
間取り
専有面積
詳細URL

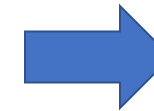
カテゴリ変数（文字情報）
量的変数

- 収集方法：requests, Beautiful Soup
- 収集時期：2020年6月上旬
- データ総数：175,032件

B データの準備

・データフレームの作成

マンション名	住所	立地1	立地2	立地3	築年数	建物の高さ	階層	管理費	敷/礼/保証/敷引、償却	間取り	専有面積	詳細URL	賃料
5 コンフォート神楽坂B棟	東京都新宿区赤城下町	東京メトロ東有線/神楽坂駅 歩4分	東京メトロ有線/江戸川橋駅 歩6分	都営大江戸線/牛込神楽坂駅 歩8分	築29年	3階建	3階	3000円	5.2万円/5.2万円	ワンルーム	9.88m2	https://suumo.jp/chintai/jnc_000043930213/?bc=...	5.2万円
6 ジュネス哲学堂	東京都新宿区西落合1	都営大江戸線/落合南長崎駅 歩6分	西武池袋線/東長崎駅 歩13分	西武新宿線/新井薬師前駅 歩17分	築26年	2階建	2階	-	5.5万円/-	1K	19.42m2	https://suumo.jp/chintai/jnc_000058044412/?bc=...	5.5万円
7 ヴィヴレコート	東京都新宿区中井1	都営大江戸線/中井駅 歩5分	西武新宿線/中井駅 歩5分	東京メトロ東有線/落合駅 歩9分	築26年	3階建	2階	5000円	5.4万円/5.4万円	ワンルーム	15.68m2	https://suumo.jp/chintai/jnc_000058304186/?bc=...	5.4万円



数値データにも文字や単位が混在した状態。

住所、立地、間取りのデータは細分化し、カテゴリ変数やダミー変数へと変換できる形にする必要がある。

・取得したデータの理解

今回は、取得した物件データの特徴理解がビジネス課題の解決に寄与しないため、データ理解は割愛する。
実務で行う上では、データ理解が必要な場合がある。

C データの前処理

- 間取りにダミー変数を適用

(ワンルーム = room_number : 1)

	room_number	DK	K	L	S
5LDK	5	1	0	1	0
1LDK	1	1	0	1	0

- 変数の細分化

東京都千代田区九段南 3



区	市町村
千代田	九段南 3

東京メトロ南北線/市ヶ谷駅 歩6分
東京メトロ半蔵門線/九段下駅 歩9分
J R 中央線/飯田橋駅 歩17分



路線 1	駅 1	徒歩 1	路線 2	駅 2	徒歩 2	路線 3	駅 3	徒歩 3
東京メトロ南北線	市ヶ谷駅	6	東京メトロ半蔵門線	九段下駅	9	JR中央線	飯田橋駅	17

- Ordinal Encoder の適用

	路線1	路線2	路線3	駅1	駅2	駅3	徒歩2	区	市町村
0	西武新宿線	東京メトロ東西線	NaN	下落合駅	落合駅	NaN	NaN	新宿	上落合 1
1	西武新宿線	東京メトロ東西線	NaN	下落合駅	落合駅	NaN	NaN	新宿	上落合 1
2	西武新宿線	東京メトロ東西線	NaN	下落合駅	落合駅	NaN	NaN	新宿	上落合 1
3	西武新宿線	東京メトロ東西線	J R 中央線	中井駅	落合駅	東中野駅	NaN	新宿	上落合 3
4	東京メトロ東西線	東京メトロ有楽町線	都営大江戸線	神楽坂駅	江戸川橋駅	牛込神楽坂駅	NaN	新宿	赤城下町



	路線1	路線2	路線3	駅1	駅2	駅3	区	市町村
0	8.0	8.0	0.0	0.0	43.0	0.0	0.0	1.0
1	8.0	8.0	0.0	0.0	43.0	0.0	0.0	1.0
2	8.0	8.0	0.0	0.0	43.0	0.0	0.0	1.0
3	8.0	8.0	19.0	1.0	43.0	35.0	0.0	3.0
4	6.0	7.0	16.0	31.0	34.0	43.0	0.0	136.0

C データの前処理

・データフレームの再構築

```
1 df = df[["マンション名", '賃料+管理費', "区", '市町村',  
2         '築年数', '建物の高さ', '階1', '専有面積',  
3         '路線1', '駅1', '徒歩1',  
4         '路線2', '駅2', '徒歩2',  
5         '路線3', '駅3', '徒歩3',  
6         '間取り', '間取りDK', '間取りK', '間取りL', '間取りS']]  
7  
8 df.columns = ['name', 'real_rent', "city", 'adress', 'age', 'hight', 'level', 'area',  
9              'route_1', 'station_1', 'distance_1',  
10             'route_2', 'station_2', 'distance_2',  
11             'route_3', 'station_3', 'distance_3',  
12             'room_number', 'DK', 'K', 'L', 'S']
```

In [5]: 1 df.iloc[0]

Out[5]:

name	クレール北綾瀬
real_rent	36000
city	22
adress	2114
age	33
hight	2
level	1
area	10.7
route_1	25
station_1	199
distance_1	18
route_2	1
station_2	382
distance_2	23
route_3	2
station_3	0
distance_3	0
room_number	1
DK	0
K	0
L	0
S	0

・外れ値対策

家賃総額が30万円未満の物件に限定

(総数170,251件)

```
In [57]: 1 #データの家賃金額の制限  
2 #ハズレ値対策  
3  
4 df = df[df['real_rent'] < 300000]  
5 df.shape
```

Out[57]: (170251, 22)

D 予測モデルの構築・実行

- 使用する特徴量は5つ

区

築年数

物件の存する階

専有面積

間取り

概算見積作成の際には、建物の立地によって金額の変動は起こらないため、立地に関する特徴量の使用は最小限に留める。

その他feature importanceの確認を経て、5つに限定した。

- 学習データとテストデータへ分割

`train_test_split (test_size=0.33)`

train = 114,068件 / test = 56,183件

- 使用する機械学習アルゴリズム

Light_GBM

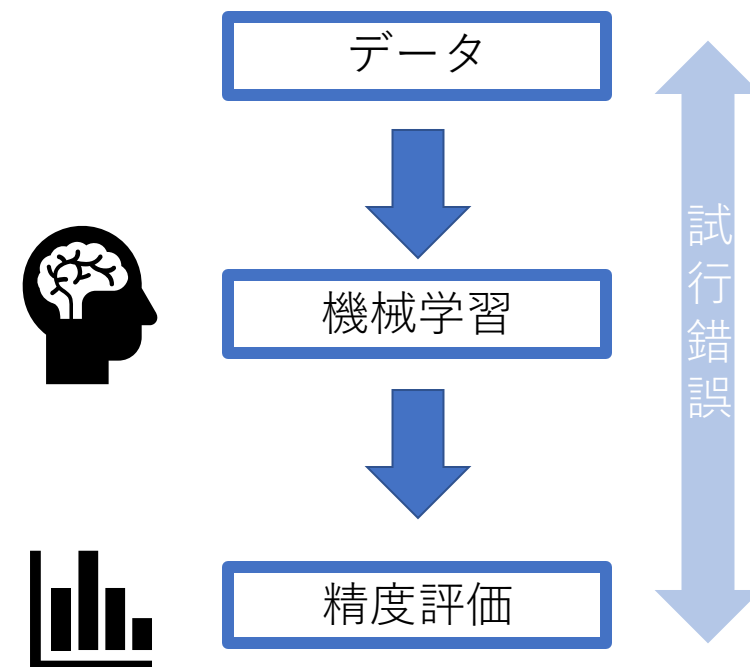
ランダムフォレスト

- 精度評価の指標

r2_score (当てはまりの良さ)

RMSE (誤差の確認)

	city	age	level	area	room_number	DK	K	L	S
0	22.0	33	1.0	10.70	1	0	0	0	0
1	22.0	45	3.0	19.80	1	1	0	0	0
2	22.0	27	1.0	14.40	1	0	0	0	0
3	22.0	27	1.0	14.40	1	0	0	0	0
4	22.0	3	2.0	9.61	1	0	0	0	0



E モデルの性能評価

- ・テストデータに対するスコア

	r2_score	RMSE
light_GBM	0.9113	14553.651
ランダムフォレスト	0.9415	11818.094

- ・今回において秀でた学習モデルは、ランダムフォレストである。
- ・light_GBMは学習コストが低く、早い。（精度も悪くない）
特徴量の修正後など、何度も学習する際に優位。
重要な特徴量の確認にも有利可能。
- ・ランダムフォレストは木の数を増やして予測を行うことで、精度の向上がみられた。
（グリッドサーチにて、木の数600、深さ30）

E モデルの性能評価

- 実際にテストデータ1件に当てはめて、感覚を掴む
精度面で優位なランダムフォレストで検証

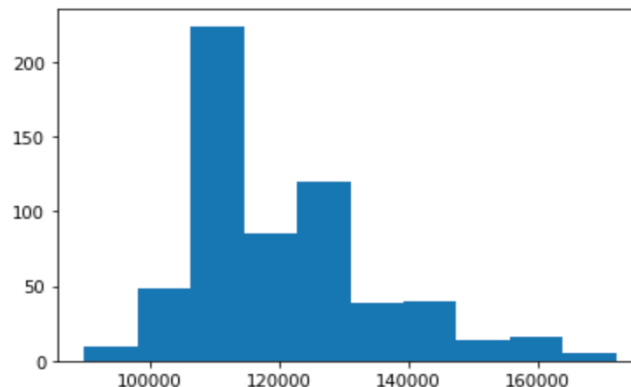
実際の価格	¥128,500円
予測価格	¥120,839円
誤差	¥7,661円

- 予測結果群の可視化

アンサンブル学習に用いる個々を

可視化することで、全体の予測傾向や
分散の大小が見てとれる。

Out[23]: (array([9., 48., 224., 85., 120., 39., 40., 14., 16.,
array([90000., 98200., 106400., 114600., 122800.,
147400., 155600., 163800., 172000.]),
<a list of 10 Patch objects>)



```
In [16]: 1 X_test.iloc[10]
```

```
Out[16]: city      9.0  
age       18.0  
level     5.0  
area      33.6  
room_number 1.0  
DK        0.0  
K         1.0  
L         0.0  
S         0.0  
Name: 14313, dtype: float64
```

```
In [17]: 1 y_test.iloc[10]
```

```
Out[17]: 128500.0
```

```
In [20]: 1 y_pred1 = rf.predict([X_test.iloc[10]])  
2  
3  
4 print("予測価格 : ",y_pred1)  
5 print("実際の価格 : ",y_test.iloc[10])
```

```
予測価格 : [120838.97882395]  
実際の価格 : 128500.0
```

E モデルの性能評価・考察

・予測結果

実際の価格	¥128,500円
予測価格	¥120,839円
誤差	¥7,661円

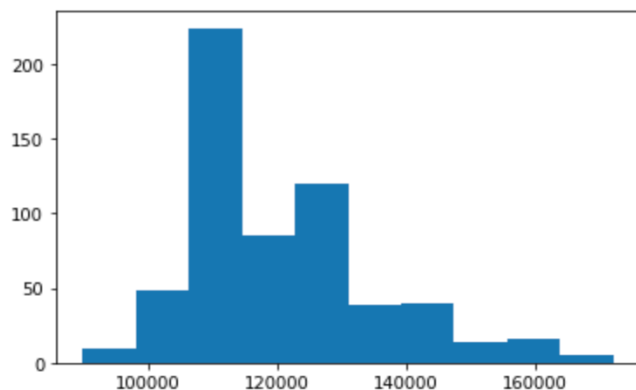
言語化

- ・ 本当の価格が¥128,500円のところ、機械学習モデルは¥120,839円と予測。
- ・ RMSEの値から、平均して \pm ¥11,818円の誤差が発生している。

「概算見積作成の際にも、同様の水準で精度が担保できるのであれば、試してみる価値があり、採用に値すると判断できる。」

・予測結果群の可視化

Out[23]: (array([9., 48., 224., 85., 120., 39., 40., 14., 16.,
array([90000., 98200., 106400., 114600., 122800.,
147400., 155600., 163800., 172000.]),
<a list of 10 Patch objects>)



言語化

「機械学習モデルは12万円と予測したが、予測数の多い値は11万円。」

「13万あたりも2番目に予測数が多く、さらに高い値にも予測結果が分散している。」

「概算見積の際には、高めに価格を設定する傾向があるため、顧客への提出は14万円としておこう。」

機械学習から導かれた予測結果が、人間の判断材料となり、ビジネス課題に対する打ち手の選択を助ける。

F 機械学習モデルの応用

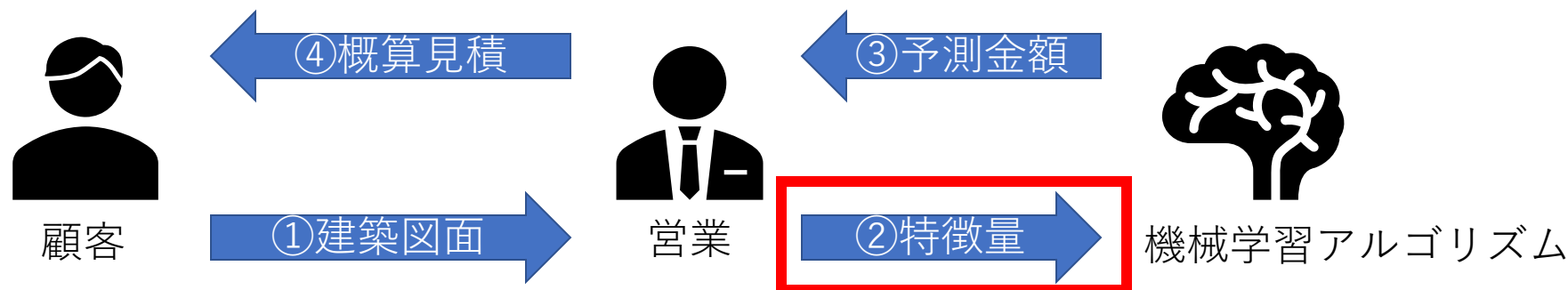
- ・機械学習モデルの構築・評価が完了！

しかし、むしろここからが本題。

- ・概算見積の金額予測に、機械学習モデルを応用するにあたって
顧客から受領した建築図面や仕様を元に、自火報設備の概算見積金額を予測することが目標である。
そのため、新しい数値データ（特徴量）を入力できるプラットフォームが必要となる。



今回は機械学習を行ったjupyter notebook上で、input関数を用いて入力値をデータフレームへと変換することとする。



任意の値を入力できる必要がある

F 機械学習モデルの応用

- input関数による特徴量の入力

物件の該当する区名を入力してください。例:「新宿」 :

物件の築年数を入力してください。例:「5」 :

物件の存する階層を入力してください。例:「2」 :

物件の専有面積を入力してください。例:「15」 :

物件の間取りを入力してください。例:「ワンルーム」,「2LDK」 :

- DataFrameに変換
- 機械学習モデルの実装
- 予測結果・分布の可視化

新しいデータに対して予測結果のアウトプットができるようになった。

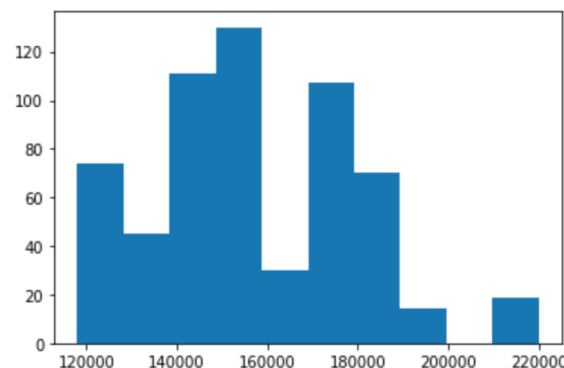
Out[15]:

	city	age	level	area	room_number	DK	K	L	S
0	0	10	4	45.0	2	1	0	1	0

In [16]:

```
1 #未知データの予測
2 y_pred = rf.predict(target_X)
3 print("予測価格 : ",y_pred)
```

予測価格 : [157619.75138889]



G 概算見積の作成で実装するにあたって（ハード面）

- データの収集

社内の物件管理システムのデータを用いる。

- 特徴量の選定

建築用途ごとに特徴量が変わる可能性があり、個別に学習データを作成する必要があると考えられる。

- 機械学習モデルの選択・実装・精度評価

実際のデータで試してみないことには、わからない。

feature importanceの確認を行う事で、特徴量の取捨選択に貢献できる。

- プラットフォームの作成

顧客から依頼を受けた新しい物件に対して、予測を行う際に必要となる。

Pythonの経験が無い人にも、使えるように工夫が必要。

- 学習モデルの更新

消防法の改正や、データの取得量に応じて更新。

賃貸物件の特徴量
区
築年数
建物の存する階
専有面積
間取り



概算見積の特徴量	
建築用途	共同住宅
防火対象物	5項 口
建築構造	耐火
受信機回線数	30回線
延べ床面積	4000㎡
地上階	5
地下階	1
消化設備	無
防排煙設備	有
住戸数	20

マンションなどの共同住宅の場合、住戸数から導入機器の数を予測しやすいため、比較的高い精度が期待できる。

G 概算見積の作成で実装するにあたって（ソフト面）

- ・ 誰が使いこなせると、メリットが大きいのか

営業担当者が自身で概算金額を算出できれば、設計・積算の業務コストの削減に寄与する。

しかし、見積内容を熟知している積算が最初に利用し、予測モデルへの心理的ハードルを下げることを第一ステップとする。

その後、徐々に営業へと研修を行いながら広めていく事が、円滑な実装に繋がると考えられる。

- ・ いきなり機械学習の1本化は現実的ではない。

まずは設計・積算を介した見積金額の算出と同時並行で進めていく必要がある。

積算・営業が予測の精度や実用性を肌感覚で感じる必要性がある。

- ・ 客先への金額算出方法の説明

納期の早さで、顧客の満足度向上が見込めるが、信憑性の説明にコストがかかる。

最初は従来の見積作成方法と同時並行していき、具体的な設計見積に移行した際に、

従来の見積と、予測モデルの見積との誤差を共に確認・評価していく必要がある。

（LIME・SHAPなどを用いて予測結果の解釈性を高める方法もあるが、より納得感を

与えるためには、従来の見積との比較が有効だと考える。）

予測モデルに初めて触れる人へ
「納得感」を与える事が重要課題となる。

（一番苦勞しそう・・・）

H ビジネスインパクト

- ・ 設計・積算業務の20%削減

設計・積算が概算見積作成に費やしている時間は、日常業務量の1/5程度である。

1日8時間労働、1ヶ月の勤務日数を20日で換算すると、1ヶ月あたり32時間。

1人あたり1年間に384時間費やしている。

営業への完全アウトソーシングを達成した場合、1人あたり384時間の削減効果が見込める。

扱う物件の件数、クオリティを据置とする場合、時給2000円換算とすると1人あたり

768,000円のコスト削減が可能となる。（残業の削減によるインパクトが大きい）

- ・ 設計・積算業務のクオリティ向上・件数の向上

概算の見積にかかる時間を削減する事で、受注に関わる物件に集中できる。

その結果、成果物のクオリティ向上、クレームやミスの削減に寄与する事が予想される。

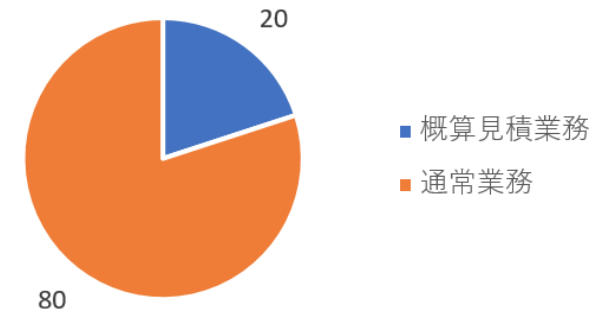
結果として、顧客満足度の向上を図ることができる。

業務の効率化により、対応できる物件の件数が向上する。

- ・ 経費の削減が可能となる。

設計・積算担当者が20人在籍していた場合、1年間で15,360,000円の経費削減に繋がる。

概算見積の占める割合



今後さらに実施してみたい事

- ・機械学習モデルを複数用意すると、営業の判断に柔軟性が加わるか？

r2_score/RMSEの水準が同レベルの予測モデルを複数用意できれば、各々の予測結果を勘案し、提出金額に対して納得感が高まるのではないだろうか。

(予測精度の面から見ると、必ずしも優位とは言えない)



- ・実際の物件データを用いた機械学習の実装（上司への提案）
 - ・金額予測を行うプラットフォームの作成
 - ・期間を設定する事で、自動で学習済モデルを更新するシステムの構築
- 1ヶ月に一回、新たなデータを取り込むなど。

課題に取り組むにあたって

・工夫した事

機械学習モデルは、あくまでビジネス課題の解決に向けたツールの一つとして位置付け

「何ができるのか」「どうビジネスの場で有効活用できるのか」この点を考えることに時間を費やした。

ビジネス課題においてどう機械学習を活かせるのか、分析方針の策定。

機械学習モデルを実装した場合に必要な、運用化までの工程の把握、構築。

(PPTの内容は、本資料のみで内容の把握が可能となるように作成した。そのため文字が多く、見づらい面もあるが、ご容赦いただきたい。)

・難しかった事

ビジネス課題の理解を深める事。

目的のために必要な情報を考え、データを収集する一連の流れの把握。

実際に現場で実装する場合に起こりうる問題点、課題点の把握。

機械学習の知識を持たない人から市民権の獲得を目指す説得方法の策定。

定量的に期待されるビジネスインパクトを説明する事。

最後に

- ・機械学習はあくまで食材であり、料理するのは人間である

人間にはよくわからない事も機械学習は導き出せる。（人間に理解できるかは別として）

小難しいことは機械学習に任せて、導き出された結果を受けて「ではどうしようか」「何ができるのか」

この点に注力して取り組むことが可能となる。これこそが機械学習の醍醐味だと感じた。

そのためには、機械学習のできること、できないことを理解している必要がある。

その結果、ビジネス課題の解決に多様な打ち手を提案できる。

そのためにも、これからも実際に手を動かして学びを深めていきたいと考えている。

実際に学ぶ事で、自身の築きたいキャリアの方向性を改めて理解できた。