



コンビニの商品売上の予測モデリング

-外的要因による売上の変化-

東京情報大学 総合情報学科 情報システム学系
知能情報システム研究室 永井ゼミ



はじめに

- 現在、コンビニには新商品や定番商品などの様々な商品が販売され、消費者のニーズに対応している。商品の充実とバランスを考えるため、過去の需要データを参考とし、売れ残りと品切れの発生を最小限に抑えるための最適発注法を考察する。

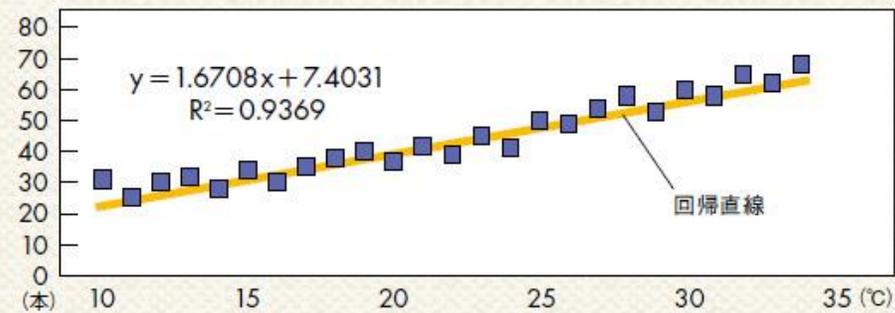
動機



● 気温で変わるビールの売り上げ (図1)

気温 (°C)	10	11	12	13	14	15	16	17	18	19	20	21	22
売上実績 (本)	31	25	30	32	28	34	30	35	38	40	37	42	39
気温 (°C)	23	24	25	26	27	28	29	30	31	32	33	34	
売上実績 (本)	45	41	50	49	54	58	53	60	58	65	62	68	

● 散布図にしてみると (図2)





「シェア」「気温」「曜日効果」「ロケーション」「群集構造」「居住人数」の6つの項目によって、どれだけ信頼できるかを計量化しています。



課題

- “気温が高いほどビールが売れる”は本当か？
- 雨の日、気温が下がる、曜日によって、、、どのぐらいの影響がある。
- 商品売上の予測モデリングを作る



研究目的

- 利益最適化を用いた結果、利益を最大にする
経済発注量を求める
- データセットを作り ディープラーニングを
用いて商品売上の予測モデリングを作成する



開發環境、言語

- Ubuntu 16.04
- Tensorflow
- sklearn
- Anaconda
- Python



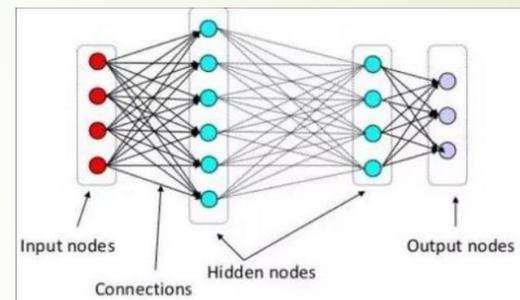
開発について

- データセット
- データを収集、加工、など
- 学習モデルを建てる
- やりたいこと

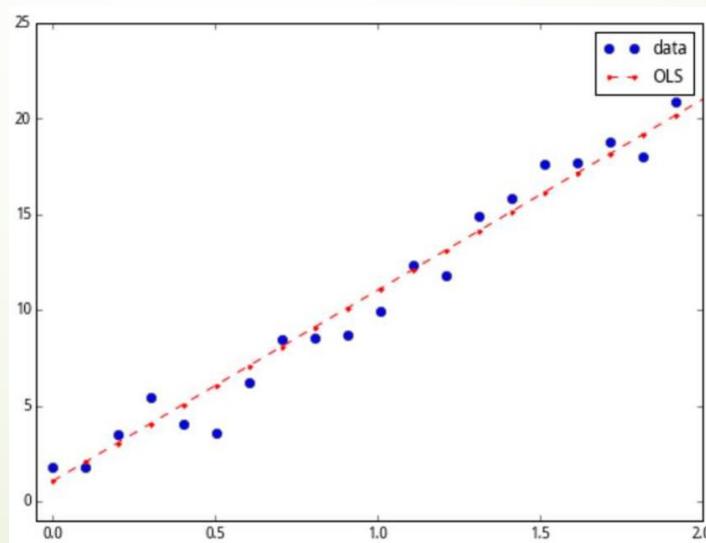
進捗

➤ ARIMAモデル (自己回帰移動平均モデル)

$$X_t = c + \sum_{i=1}^p \varphi_i X_{t-i} + \varepsilon_t.$$



➤ 訓練データ



➤ Import

➤ scikit-learn関係をインポートします

```
1. # 計算で使うかもしれないNumPy
2. import numpy as np
3. # データフレームを使うためにPandas
4. import pandas as pd
5.
6. # 可視化ライブラリ
7. import matplotlib.pyplot as plt
8. import seaborn as sns
9. sns.set()
10. # もしもJupyter notebookで使う場合は、以下のコマンドが必要
11. %matplotlib inline
```

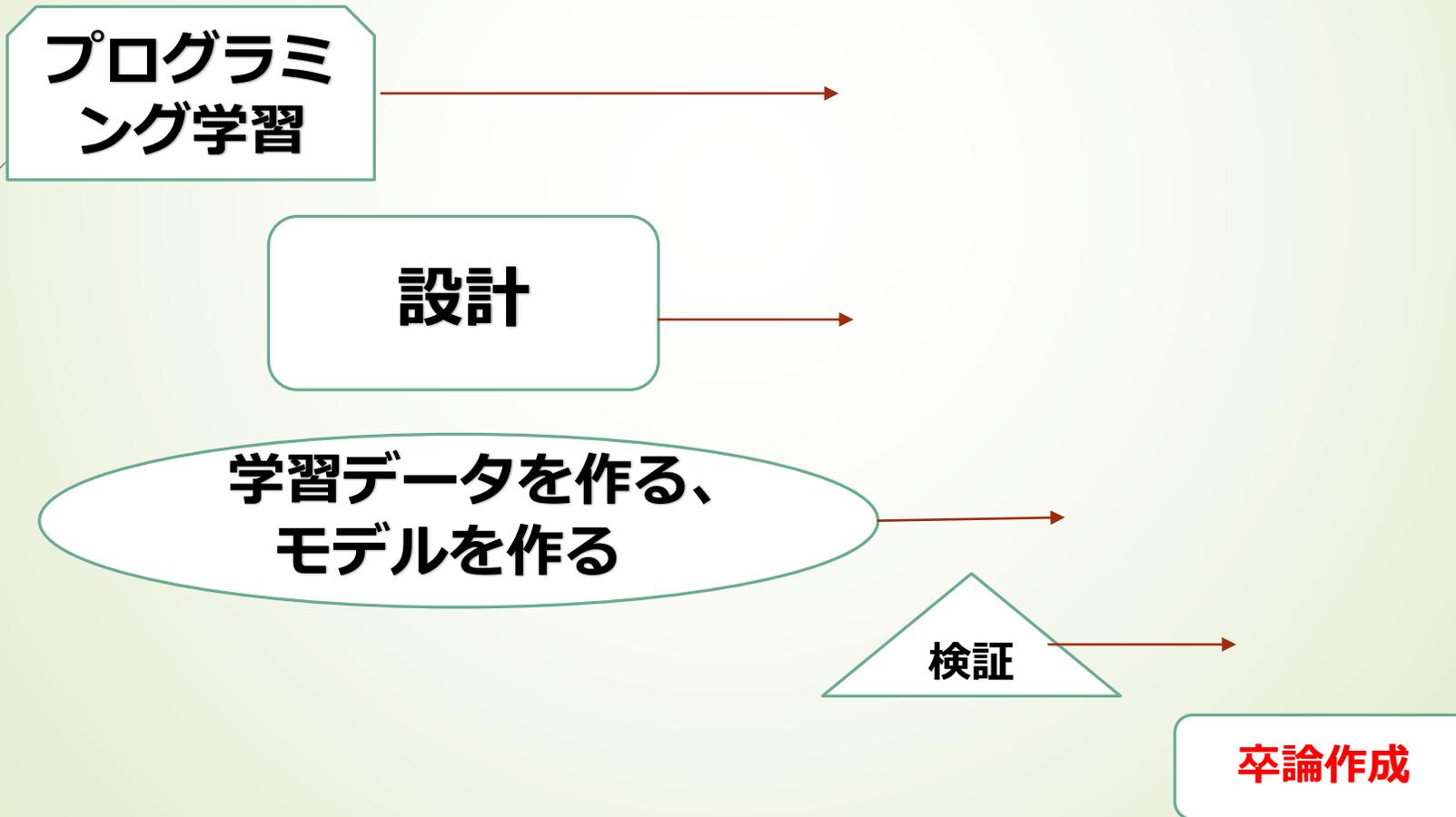
```
1. # データセット
2. from sklearn import datasets
3. # ランダムフォレスト回帰のクラスをRFRというあだ名を付けてimport
4. from sklearn.ensemble import RandomForestRegressor as RFR
5.
6. # 教師データとテストデータに分割してくれる
7. from sklearn.model_selection import train_test_split
8.
9. # 平均二乗誤差を計算する関数
10. from sklearn.metrics import mean_squared_error
11.
12. # もしパラメータ探索がしたい場合は以下もimport
13. from sklearn.model_selection import GridSearchCV
14.
15. # もしデータの標準化がしたい場合は以下もimport
16. from sklearn.preprocessing import StandardScaler
```

データの読み込み

```
1. # データフレーム化
2. train_df = pd.DataFrame(train_data, columns=boston.feature_names)
3.
4. # 上から10行を表示
5. train_df.head(10)
```

	CRIM	ZN	INDUS	CHAS	NOX	RM	AGE	DIS	RAD	TAX	PTRATIO	B	LSTAT
0	0.38214	0.0	6.20	0.0	0.504	8.040	86.5	3.2157	8.0	307.0	17.4	387.38	3.13
1	24.39380	0.0	18.10	0.0	0.700	4.652	100.0	1.4672	24.0	666.0	20.2	396.90	28.28
2	9.72418	0.0	18.10	0.0	0.740	6.406	97.2	2.0651	24.0	666.0	20.2	385.96	19.52
3	0.62356	0.0	6.20	1.0	0.507	6.879	77.7	3.2721	8.0	307.0	17.4	390.39	9.93
4	13.91340	0.0	18.10	0.0	0.713	6.208	95.0	2.2222	24.0	666.0	20.2	100.63	15.17
5	14.33370	0.0	18.10	0.0	0.614	6.229	88.0	1.9512	24.0	666.0	20.2	383.32	13.11
6	1.38799	0.0	8.14	0.0	0.538	5.950	82.0	3.9900	4.0	307.0	21.0	232.60	27.71
7	0.22188	20.0	6.96	1.0	0.464	7.691	51.8	4.3665	3.0	223.0	18.6	390.77	6.58
8	13.67810	0.0	18.10	0.0	0.740	5.935	87.9	1.8206	24.0	666.0	20.2	68.95	34.02
9	8.71675	0.0	18.10	0.0	0.693	6.471	98.8	1.7257	24.0	666.0	20.2	391.98	17.12

研究スケジュール



参考文献

- ➡ ① 小林 和宏：回帰分析,
http://homepage2.nifty.com/crop_shimane-u/index.html.
- ➡ ② A・C・ハーベイ：時系列モデル入門, 東京大学出版会