

Previsão do número de visitantes em restaurantes

Kaggle Competition

Cleilton C. D. da Silva
cleiltoncarlos@copin.ufcg.edu.br

Por que a previsão?



Por que a previsão?

- Compra de ingredientes
- Alocação de funcionários

Como prever número de clientes em restaurantes?

- O clima
- Feriado
- Final de semana
- Reservas em sites de compras
- Promoção

Experimento

- Validação Cruzada 2-fold
- Regressão Ridge e Lasso (sklearn)
- `scoring="neg_mean_squared_log_error"`
- Dados
 - Hot Pepper Gourmet (hpg)
 - AirREGI / Restaurant Board (air)

Experimento

ホーム > 各種データ・資料 > 過去の気象データ・ダウンロード

データ検索
関連ページ

過去の気象データ

過去の地域平均気象データ

取り纏めてみませんか?
気候リスク管理

過去の気象データ・ダウンロード

!! 重要なお知らせ

このページでできること

検索条件の設定方法

気象データの表記等

ダウンロードファイルの形式

検索条件

選択済みのデータ量 0% 100% (上限)

地点を選ぶ

項目を選ぶ

期間を選ぶ

表示オプションを選ぶ

画面に表示 ▶

CSVファイルをダウンロード ▶

すべての選択済みの地点をクリア

一回のリクエストで表示・ダウンロードできるデータ量は上限があります(右ナベグラフ参照)。また、このページへのアクセスが集中したり、リクエストのデータ量が多い場合は、表示・ダウンロードまで時間がかかる場合があります。

まず、都道府県を選んでください

選択地点・項目をクリア

選択された地点 観測項目

← 地点を選択してください

選択された項目

日平均気温
日最高気温
日最低気温
降水量の日合計

削除
削除
削除
削除

Modelo 1 (features)

reserve_visitors

holiday_flg

weekend

(Utilizado somente dados de Air database)

Modelo 2 (features)

reserve_visitors

holiday_flg

weekend

from_air_database

from_hpg_database

(Utilizado dados de Air e HPG database)

Modelo 3 (features)

reserve_visitors

holiday_flg

weekend

(Inclusão de dados em que não houve reserva)

Modelo 4 (features)

Fatores:

- General information (3)
- Weather information (4)
- Restaurant information (6)

Interações

- Weather * Date (1)
- Reserve * Weather (2)
- Date * Restaurant (13)
- Reserve * Restaurant (4)

(Preenchimento de NaN baseado na média do dia da região)

Modelo 5 (features)

General information:

- reserve_visitors
- holiday_flg
- weekend

Weather information:

- t_media
- daylight_hours
- ave_cloud_cover

(Preenchimento de NaN baseado apenas na média do distrito)

Modelo 6 (features)

General information:

- reserve_visitors
- holiday_flg
- weekend

Weather information:

- t_media
- daylight_hours
- ave_cloud_cover

(Preenchimento de NaN baseado na média do distrito, considerando o ano e o mês.)

Modelo 7 (features)

Fatores

- General information (3)
- Restaurant information (11)

Interações

- Date * Restaurant (10)
- Reserve * Restaurant (5)

Resultados

Model	Training RMLSE (Ridge)	Training RMLSE (Lasso)	Test RMLSE (Ridge)
1	0.591	0.591	-
2	0.656	0.656	-
3	0.829	0.829	-
4	0.810	0.812	0.830
5	0.831	0.832	0.852
6	0.829	0.830	0.847
7	0.813	0.815	0.824

Resultados

Model	Training RMLSE (Ridge)	Training RMLSE (Lasso)	Test RMLSE (Ridge)
1	0.591	0.591	-
2	0.656	0.656	-
3	0.829	0.829	-
4	0.810	0.812	0.830
5	0.831	0.832	0.852
6	0.829	0.830	0.847
7	0.813	0.815	0.824

Resultados

Model	Training RMLSE (Ridge)	Training RMLSE (Lasso)	Test RMLSE (Ridge)
1	0.591	0.591	-
2	0.656	0.656	-
3	0.829	0.829	-
4	0.810	0.812	0.830
5	0.831	0.832	0.852
6	0.829	0.830	0.847
7	0.813	0.815	0.824

Resultados

Model	Training RMLSE (Ridge)	Training RMLSE (Lasso)	Test RMLSE (Ridge)
1	0.591	0.591	-
2	0.656	0.656	-
3	0.829	0.829	-
4	0.810	0.812	0.830
5	0.831	0.832	0.852
6	0.829	0.830	0.847
7	0.813	0.815	0.824

Conclusões

- Tratamento de NaNs
- Seleção de features
- Features com mais influência:
 - weekend
 - reserve_visitors
 - holiday_flg
 - is_dining_bar
 - visitors * italian_french