

需要予測システム A-Forecast (自動機械学習)

- データ入力や結果出力、地図や図の操作方法、表の操作方法などは「Streamlit操作方法.pdf」を御覧ください。
- リファレンスマニュアルも合わせて御覧ください。
- 導入コンサルティングやユーザーの要望に応じたカスタマイズやAPI提供が可能です。必要な場合、お問い合わせください。

データ作成時の注意点:列の名前はサンプルと同じにする必要がある.列の省略不可.

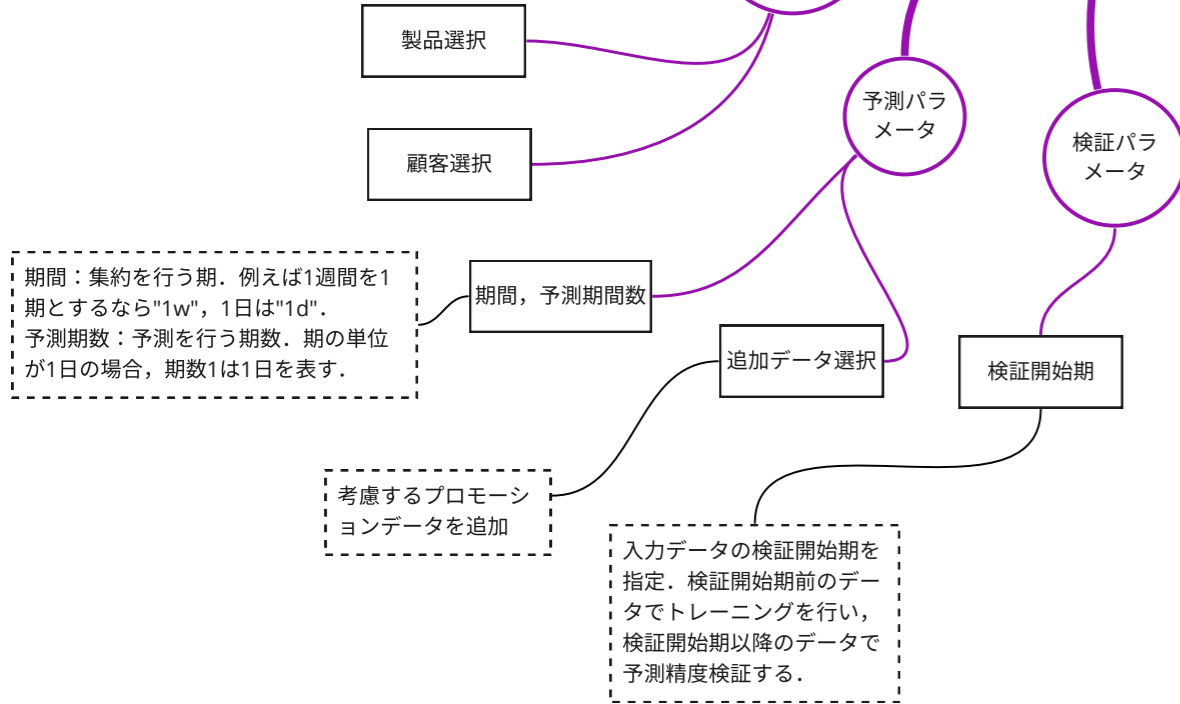
需要データ(青文字データは必須)

- date:日付
- cust:顧客名
- prod:製品名
- demand:需要量
- 予測で考慮したい要素1:任意の列名でデータ入力可能
- 予測で考慮したい要素2:任意の列名でデータ入力可能
- 予測で考慮したい要素... n: ... 考慮したい要素は何種類でも追加可能

プロモーションデータ

予測で考慮したい日付以外の要素 (プロモーションや天気などなんでもOK)

- date:日付
- 考慮したい要素 1
- 考慮したい要素 2



パラメータ設定で選択した内容で予測で使うデータを抽出する.更に, 予測用データの日付データなどを分解する.

評価尺度選択

- mean absolute error (MAE):平均絶対誤差
- mean squared error (MSE):平均二乗誤差
- root mean squared error (RMSE):平均二乗誤差の平方根
- R2:決定係数(coefficient of determination)
- root mean squared log error(RMSLE):対数平均二乗誤差
- Mean Absolute Percentage Error (MAPE):平均絶対パーセント誤差

最良モデル数:同じ数の予測結果の分析が可能

機械学習による予測

Modelはモデルの名前, その他列は評価尺度の値
「評価尺度選択」で選択された評価尺度の誤差が少ない順にソートされる.

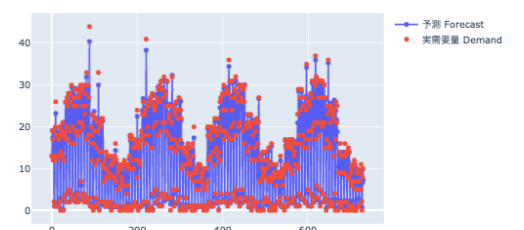
表示行数選択: 5

表で表示

| | Model | MAE | MSE | RMSE | R2 | RMSLE | MAPE | TT (Sec) |
|----------|---------------------------------|--------|--------|--------|--------|--------|--------|----------|
| rf | Random Forest Regressor | 1.1567 | 2.3772 | 1.5418 | 0.9514 | 0.3319 | 0.1753 | 0.1800 |
| lightgbm | Light Gradient Boosting Machine | 1.3149 | 3.6515 | 1.9109 | 0.9253 | 0.3159 | 0.2834 | 0.0400 |
| xgboost | Extreme Gradient Boosting | 1.3293 | 3.9298 | 1.9824 | 0.9196 | 0.2735 | 0.2326 | 0.3500 |
| gbt | Gradient Boosting Regressor | 1.5259 | 4.3517 | 2.0861 | 0.9109 | 0.3678 | 0.3130 | 0.0700 |
| dt | Decision Tree Regressor | 1.3699 | 4.3836 | 2.0937 | 0.9183 | 0.3749 | 0.2273 | 0 |

[Download csv file](#)

AutoMLによる予測 Forecasting by AutoML



予測結果

表示行数選択: 5

表で表示

| Is_month_end | Is_month_start | Is_quarter_end | Is_quarter_start | Is_year_end | Is_year_start | Elapsed | Label |
|--------------|----------------|----------------|------------------|-------------|---------------|------------|---------|
| 0 | 0 | 1 | 0 | 1 | 0 | 1546300800 | 13.0300 |
| 1 | 0 | 0 | 0 | 0 | 0 | 1546397200 | 19.1300 |
| 2 | 0 | 0 | 0 | 0 | 0 | 1546473600 | 17.4200 |
| 3 | 0 | 0 | 0 | 0 | 0 | 1546560000 | 12.1000 |
| 4 | 0 | 0 | 0 | 0 | 0 | 1546646400 | 18.8600 |

[Download csv file](#)

平均2乗誤差 2.377213698630137

表示行数選択: 5

表で表示

| promo_0 | promo_1 | Week | Year | Month | Day | Dayofweek | Dayofyear | Is_month_end | Is_month_start | Is_quarter_end | Is_quarter_start | Is_year_end | Is_year_start | Elapsed | Label |
|---------|---------|------|------|-------|-----|-----------|-----------|--------------|----------------|----------------|------------------|-------------|---------------|------------|---------|
| 0 | 0 | 53 | 2020 | 12 | 31 | 3 | 366 | 1 | 0 | 0 | 0 | 0 | 0 | 1546646400 | 18.8600 |

[Download csv file](#)

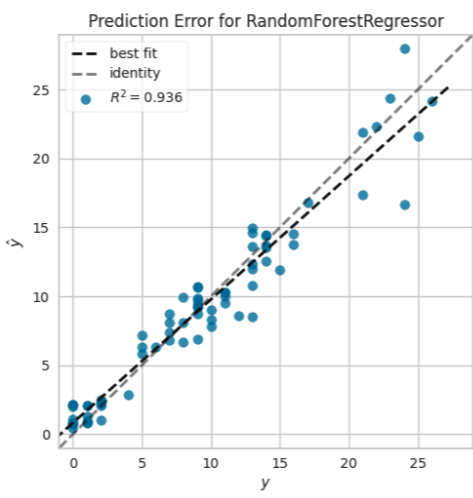
データ探索

右側が入力データ (トレーニング+検証) で左側が予測結果.

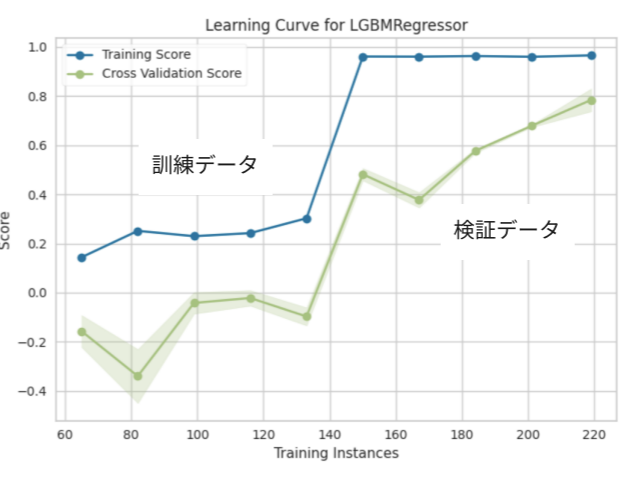
- 'Year': 年
- 'Month': 月
- 'Day': 日
- 'Dayofweek': 曜日
- 'Dayofyear': 年初からの経過日数
- 'Is_month_end': 月の最終日
- 'Is_month_start': 月の開始日
- 'Is_quarter_end': 四半期の終了日
- 'Is_quarter_start': 四半期の開始日
- 'Is_year_end': 年の最終日
- 'Is_year_start': 何の開始日
- 'Elapsed': 経過時間
- Label: 予測値 (トレーニングデータの場合,)

「モデルを選択」で選択されたモデルに対して「モデル可視化」で選択された項目の図を表示する.

ターゲットと予測値の関係



学習曲線



特徴量の重要度

