

NRIデータ分析コンテスト2020

小売業の過去の売上データ、気象データをもとに、
2020年6月の特定カテゴリの販売数を予測する

株式会社野村総合研究所

2020年8月6日



コンペの説明

■ 概要

- 小売業の過去の売上データ、気象データをもとに、2020/6/1～6/30の販売数を予測してください。統計分析・機械学習など手法は問いませんが、再現可能であることが要件になりますので、再現できるプログラム、Excelなども、最終的には合わせて提出してください。

■ 課題

- 与えられたデータを使い、2020/6/1～2020/6/30の、店別商品グループ別の販売数を予測してください。

■ 評価基準

- **基本的に定量評価を行います。定量評価以外でも優れた分析アプローチなどは定性的に評価します。**
 - ・ 定量評価：評価対象となる330レコードのRMLSE（log化した指標）で評価します。
 - ・ 定性評価：データから得られた考察の深さ、アプローチ方法の有用性、ディスカッション（同報投稿）への貢献などを評価します。
- **毎月、月末締めで、それまでに提出頂いたチームのscore(LB)を翌月中間発表する予定です。public=privateです。提出チャンスが8月末、9月末、10月末の3回あります。**
- **最終評価は、最後に提出したものを採用します。**
- 販売数予測に限らず、このデータを使った自由な分析を行って、レポートを提出頂いても構いません。（例えば、コロナの影響により、2020/3以降の売上は大きく減少していますので、このデータを使って、コロナが売上に与えた影響に関する分析や、未来予測を行う、など。）
- 提供データ以外のデータを使用しても構いませんが、コンペ参加者全員に公開することを条件とします。提供データ以外のデータを使用する場合は、9月末までに宣言してください。

データの説明

データファイル名	カラムの説明
train.csv	<ul style="list-style-type: none">➤ 「target」が予測対象列です。日別店別商品グループ別の販売数です。時系列データです。 nichi：日付です。(YYYYMMDD形式) group_mise：店です。X～Zの3種類です。 group_item：商品グループです。A～Eの5種類です。なおA～Eの商品グループはいずれも食品です。 target：商品販売数です。
test.csv	<ul style="list-style-type: none">➤ 空欄になっている「target」を予測してください。
kyaku.csv	<ul style="list-style-type: none">➤ 店の客数を正規化したものです。 kyaku_param：店グループ別に0～1に正規化した客数です。
uriage.csv	<ul style="list-style-type: none">➤ 店の売上を正規化したものです。 uriage_param：店グループ別に0～1に正規化した売上です。
food_uriage.csv	<ul style="list-style-type: none">➤ 店の食品売上割合（総売上に占める食品売上の比率）です。 food_uriage_wariai：店グループ別の食品売上割合です。（例：0.7の場合、総売上に対して食品売上が70%）
tenki.csv	<ul style="list-style-type: none">➤ 天気情報です。 kion_max, kion_min, kion_ave：最高気温、最低気温、平均気温です。 kion_asa, kion_hiru, kion_yoru：朝気温、昼気温、夜気温です。 kousuiryou, kousuiryou_asa, kousuiryou_hiru, kousuiryou_yoru：降水量の、日合計、朝、昼、夜です。 kousuiryou_jikan, kousuiryou_asa_jikan, kousuiryou_hiru_jikan, kousuiryou_yoru_jikan：降水時間の日合計、朝、昼、夜です。 tenki_id2, tenki_id2_asa, tenki_id2_hiru, tenki_id2_yoru：天気コード（2桁）の日合計、朝、昼、夜です。➤ 天気コード（2桁）の意味 10：晴 20：くもり 30：雨 40：雪

The text is framed by two decorative swooshes. The top swoosh is a gradient bar transitioning from blue on the left to red on the right, arching over the text. The bottom swoosh is a solid blue bar, also arching under the text.

Share the Next Values!