**主要背景**

Rossmann在7个欧洲国家中运营超过3000家药店。现在，Rossmann商店经理被要求预测接下来六周每日的销量。商店销量受到很多因素的影响，包括促销活动，竞争对手行为，学校与法定节假日，季节性以及商店的地点。成千上万个药店经理的预测都是基于他们各自的情况，这导致预测结果的准确性存在巨大的差异。

**问题描述**

**数据集和输入**

数据集包含如下字段：

* Id – 药店的id,代表了药店和对应日期
* Store – 每个药店的独有id
* Sales – 在某一日的销量
* Customers - 在某一日的顾客数量
* Open - 标记药店是否开张: 0 = 不开张, 1 = 开张
* StateHoliday - 代表法定节假日. Normally all stores, with few exceptions, are closed on state holidays. Note that all schools are closed on public holidays and weekends. a = public holiday, b = Easter holiday, c = Christmas, 0 = None
* SchoolHoliday - indicates if the (Store, Date) was affected by the closure of public schools
* StoreType - differentiates between 4 different store models: a, b, c, d
* Assortment - describes an assortment level: a = basic, b = extra, c = extended
* CompetitionDistance - distance in meters to the nearest competitor store
* CompetitionOpenSince[Month/Year] - gives the approximate year and month of the time the nearest competitor was opened
* Promo - indicates whether a store is running a promo on that day
* Promo2 - Promo2 is a continuing and consecutive promotion for some stores: 0 = store is not participating, 1 = store is participating
* Promo2Since[Year/Week] - describes the year and calendar week when the store started participating in Promo2
* PromoInterval - describes the consecutive intervals Promo2 is started, naming the months the promotion is started anew. E.g. "Feb,May,Aug,Nov" means each round starts in February, May, August, November of any given year for that store

**解决方案**

**基准模型**

**评估指标**

评估指标基于均方根误差(RMSPE)

其中代表一个商店在一天的销量， 代表对应的预测。在评分中，销量为0的商店会被忽略。

**项目设计**