

# 2022.4.21

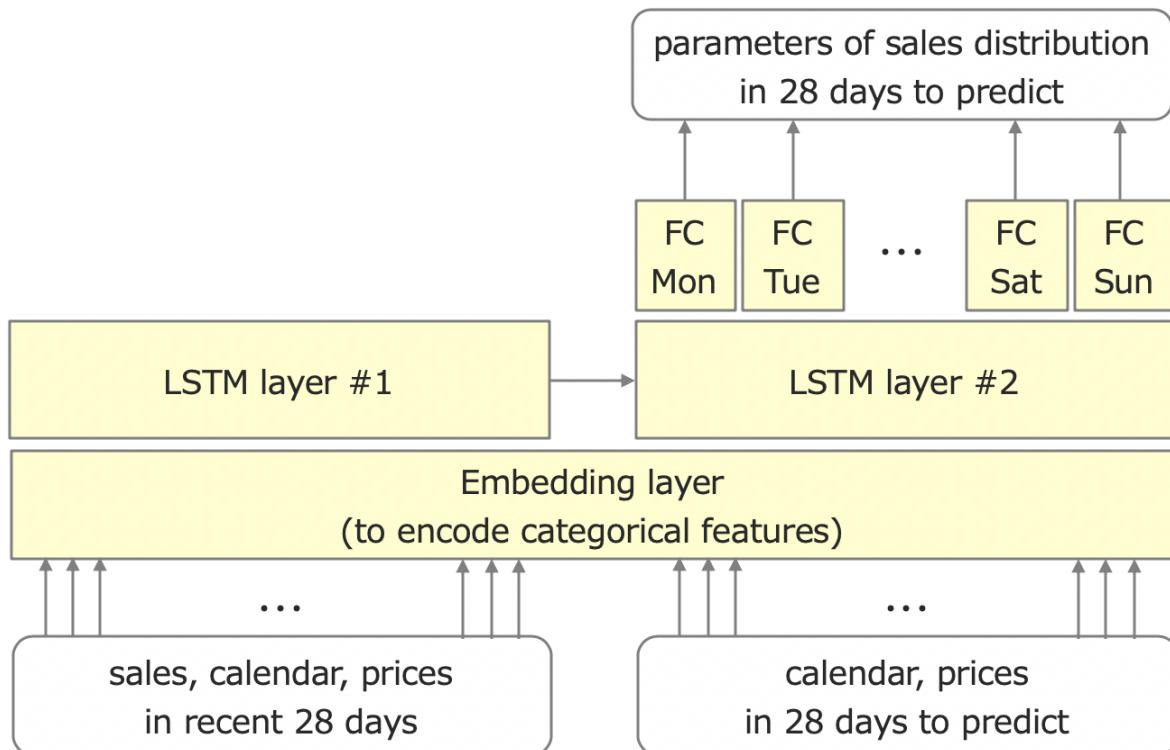
## 基于LSTM的基准预测尝试

主要参考 **Robust recurrent network model for intermittent time-series forecasting** (M5点预测第三) 以及 M5 不确定性预测第四 ([Back to \(predict\) the future - Interactive M5 EDA](#)) [【日本語】Back to \(predict\) the future - M5 EDA 6/1版](#) 的设计,但出于学习考虑还是尝试自己搭建网络, 并未直接调用其代码文件。目前主要聚焦特征与网络结构。

包含特征如下:

- 价格特征: 原始价格、相对同cat内的最大值的比例、相对同dept内的最大值的比例;
- 日历特征: 日期在一年中的比例、星期几 (可表示成在一周的比例) 、距开售第一天的距离
- 销售特征: 原始销量 (预测期首期前1-28期滞后值) 、7期移动平均、28期移动平均 (预测期首期前1-28期)
- 节假日 (event name 31类, 嵌入后6类; event type5类, 嵌入后2类)
- SNAP
- 与间断相关的特征: 上两个正销售额间的距离、距上一个正销售的距离
- id: store (10-4) 、state (3-2) 、cat (3-2) 、dept (7-3) 、item (3049--30) 的id

网络结构: 暂且先部分仿照不确定性预测的网络 (M5不确定性第四名), 用嵌入层处理分类变量, 用LSTM输入28期特征数据预测28期 (这两个论文都是28期输入, 而非用全长数据 (1900左右) 预测); 原文对一周七天各设置一个模型, 自写网络暂不考虑。损失函数也学该模型, 使用负二项的负似然作为目标。下图是M5第四名的网络结构:



自己尝试设计的网络结构如下:

**(嵌入层——Dropout——) LSTM层——Dropout层——线性层**; 输入数据时长是28期, 网络输出是负二项分布随时间变的两个参数 (28 \* 30490 \* 2)。也尝试了不加入类别特征的网络, 此时网络去掉前两层。

目前预测出现一些问题: **最大的问题是损失不下降不收敛**

由于只输入了最近28期的数据，日期变量、假日变量意义不明显；以类别特征作为输入之一时，由于类别再输入中维度比较大，故多数预测结果不准且趋同；去掉类别特征后训练效果也不佳，表现为损失没有收敛的下降趋势。

去掉类别变量后在refer set 的结果如下：可以看到，在小分位数的结果较好，大分位数结果差，说明预测结果偏小。（观察预测结果也如此，预测结果普遍小）

分位数	0.01	0.025	0.165	0.25	0.5	0.75	0.835	0.975	0.99
LSTM	0.0139	0.0347	0.2287	0.3337	0.5761	0.6360	0.5643	0.2414	0.1502

今后的改进：

- 考虑更多的历史数据作为输入（使用28期的窗口作为输入段，因时间序列的真实长度不同）
- 此外，本节目的是使用机器学习/深度学习的全局基准预测，也不一定要用LSTM/DeepAR作为基准模型。

## 基于启发式优化的组合权重计算

这里尝试了两种策略进行组合优化的权重计算：**模拟退火与粒子群优化**。设计的组合如下：

$$\sum_{i=1}^M w_i f_i + b$$

这里  $w_i$  指模型权重，不随时间、样本变； $b$  是截距项。在优化时，设定所有待求变量的范围为 [-1,1]，不设其它约束限制。优化的预测结果如下，结果显示粒子群优化的效果不佳，但模拟退火能够在一些分位数上改进组合结果。不过，这个改进并不是很明显。

分位数	0.01	0.025	0.165	0.25	0.5	0.75	0.835	0.975	0.99
模拟退火	0.0144	0.0359	0.2177	<b>0.3090</b>	<b>0.4910</b>	<b>0.4894</b>	<b>0.4214</b>	<b>0.1351</b>	0.0703
粒子群	0.0149	0.0368	0.2200	0.3179	0.5259	0.5149	0.4441	0.1370	0.0738
固定效应组合	<b>0.0143</b>	<b>0.0356</b>	<b>0.2168</b>	0.3091	0.4915	0.4897	0.4217	<b>0.1351</b>	<b>0.0702</b>
简单平均	0.0177	0.0397	0.2217	0.3161	0.5091	0.5071	0.4327	0.1362	0.0710

模拟退火与粒子群优化的系数如下所示。可以看到，模拟退火学习的权重基本与基准方法的结果好坏类似（结合基准方法结果看），而粒子群算法得到的权重绝对值偏大且有较大负值，可能存在陷入局部最优的现象。

分位数 (模拟退火)	<b>0.01</b>	<b>0.025</b>	<b>0.165</b>	<b>0.25</b>	<b>0.5</b>	<b>0.75</b>	<b>0.835</b>	<b>0.975</b>	<b>0.99</b>
gam	0.0761	0.1491	0.0829	0.0814	0.1166	0.0817	0.0877	0.1737	0.1716
vz	0.0642	0.1240	0.1850	0.0347	0.0072	0.0452	0.1227	0.2578	0.1783
wss	0.1245	0.0908	0.1628	0.1562	-0.0270	0.0389	0.0037	0.0557	0.0763
poisson_damped	-0.0023	-0.0033	-0.0963	-0.0515	0.0013	0.0041	-0.0661	-0.1707	-0.1966
poisson_undamped	-0.0034	-0.0506	0.1010	0.1007	0.1928	0.3745	0.2508	0.1642	0.2416
poisson_static	-0.0086	0.0256	0.2340	0.2614	0.4221	0.3907	0.5576	0.3411	0.2476
nb_static	0.1877	0.1018	0.0751	0.0547	-0.0370	-0.0751	-0.0079	0.0285	0.1486
nb_damped	0.1233	0.1836	0.0605	0.0715	0.0267	-0.0520	0.0110	0.0458	0.0251
nb_undamped	0.0335	0.0640	0.1192	0.2504	0.2040	0.1550	0.0476	0.0777	0.0820
截距	-0.0003	0.0000	0.0000	0.0000	0.0002	0.0000	0.0002	0.1808	0.2141

分位数 (粒子群)	<b>0.01</b>	<b>0.025</b>	<b>0.165</b>	<b>0.25</b>	<b>0.5</b>	<b>0.75</b>	<b>0.835</b>	<b>0.975</b>	<b>0.99</b>
gam	-0.0016	-0.0645	0.1843	0.0423	-0.2002	0.0655	0.1025	0.2374	0.1608
vz	-0.1273	-0.0562	-0.1875	-0.0941	-0.5233	-0.2083	-0.2718	-0.1188	0.0730
wss	-0.2087	-0.0381	0.0734	-0.2159	-0.3644	-0.4110	-0.4646	-0.2798	-0.4782
poisson_damped	-0.2388	-0.0736	-0.2097	-0.1615	0.9611	0.2405	0.2607	0.2073	0.3886
poisson_undamped	-0.0817	-0.0679	0.1494	0.5869	0.1468	0.0331	0.2009	0.1986	0.1325
poisson_static	0.1230	0.1199	0.2116	0.1710	0.1988	-0.0441	-0.1428	0.1146	-0.3195
nb_static	0.5064	0.5974	0.5278	0.4914	-0.2543	0.7130	0.7479	0.6020	0.5826
nb_damped	-0.0007	-0.1823	0.0755	-0.0491	0.5064	0.1532	0.2749	-0.0026	0.1366
nb_undamped	0.0200	-0.1655	0.3118	-0.0244	0.2894	0.5931	0.4296	0.1739	0.2690
截距	0.0000	-0.0006	-0.0022	0.0010	0.0000	0.2614	0.2664	0.1082	0.6199

尝试以上算法也是为了考虑今后使用基于特征的组合预测时，设置非约束权重组合的可行性。进一步，可以尝试组合权重随样本变、随时间变的优化（对于启发式算法，可能是每条数据分别训练权重）。

## 之后计划

---

- 继续尝试使用深度学习/机器学习的全局基准预测
- 随样本、时间变化的权重（启发式优化尝试）