

# ホームセンターにおける売れ筋商品を主とした売上予測と在庫政策

2015SS072 竹村仁志

指導教員：鈴木敦夫

## 1 はじめに

近年、POS(Point Of Sales) システムの導入により、各商品の販売日時や売れた個数などを正確に把握することができるようになってきた。このシステムで集積された情報を用いることで顧客傾向や販売動向などの分析から販売促進に生かすことが期待されている。現在では様々な小売業でこの膨大なデータの利用法が模索されている。

今回、委託研究を受けたホームセンターでは、POS システムで集積したレシートデータをもとにオペレーションズ・リサーチ (OR) や統計的手法を用いて数年にわたり研究を行ってきた。OR とは、数学的・統計的モデル、アルゴリズムなどを利用することによって、複雑なシステムにおいて「制約条件を満たした最適解」となるよう決定する科学的手法である。(参考文献 [1])

### 1.1 使用したデータ

本研究では、ホームセンターから提供された店舗集計データを用いる。この集計データは POS システムを用いて集計されており、部門、JAN コード、商品漢字名、規格漢字名、売上日付、売上数量、在庫数量、発注単位、発注数量が含まれている。JAN コードとは、各商品に与えられた固有のコードであり、コードのみで商品の識別が可能である。本研究ではこれに加え天気データを用いて、詳しく購買傾向を確認する。

本研究では使用した各データの期間は以下のとおりである。

- 2016/02/29~2018/12/31 の売れ筋商品 400 品目の店舗集計データ
- 2016/02/29~2018/11/27 の天気データ
- 2018/12/01~2018/12/31 の天気予報データ (参考文献 [3])

## 2 深層学習による予測

今回のデータの予測には深層学習を用いるうえで効率的な Visual Mining Studio を使用した。Visual Mining Studio(以下 VMS と省略する)とは Visual Analytics Platform 上で動作し、データマイニングのため前処理やデータの分析・処理など、高機能なツール群を直感的に利用できることのできるツールのことである。その中には Deep Learning 機能 (Deep Learner) が含まれており、今回の研究ではこの機能を用いて深層学習を行う。

層数は一般に 3 層異常が推奨されており、説明変数の数なども考慮し試行錯誤した結果が 3 層が安定するため 3 層で行った。(参考文献 [2])

また、設定したパラメーターは以下のようである。

- 説明変数  
月、週、曜日、休日、最低気温、最高気温、降水量合計
- RNN 層 1  
LSTM 出力次元数 16
- RNN 層 2  
Model Optimizer (GRU, LSTM, simpleRNN) 出力次元数 16
- 全結合層  
ランプ関数 (ReLU) 出力次元数 16

## 3 予測の検証

### 3.1 変数・層の変更による予測精度の向上

中間発表後売上上位 10 % の商品の予測精度向上のためパラメーターの調整を行った。また、本論 3.2 章で日ごとに予想して発注する期間をまとめるのが最適だと立証されているため、中間発表時点との比較のため同じ期間で相関係数を求めた。これによる相関係数の変化や売れ方の考察を表をもとに行った。

まず、日ごとの売上予測から 1 週間にまとめたものが表 1 のようになる。左列から順に実際の売上数量、予測した売上数量、予測した売上数量に対する実際の売上数量の割合となっている。予測したものを 1 週間ごとにまとめたものの相関係数の変化は商品 A で 0.499 → 0.793, 商品 B で 0.788 → 0.740, 商品 C で 0.634 → 0.770, 商品 D で -0.309 → 0.690 となった。商品 A と商品 D の商品 C の相関係数が大幅に上がったが、商品 B の相関係数は僅かではあるが下がった。

10 日分にまとめたものが表 2 のようになる。同じく左列から順に実際の売上数量、予測した売上数量、予測した売上数量に対する実際の売上数量の割合となっている。予測したものを 10 日間ごとにまとめたものの相関係数の変化は商品 A で 0.952 → 0.882, 商品 B で 0.718 → 0.835, 商品 C で 0.161 → 0.762, 商品 D で -0.731 → 0.639 となった。商品 B, 商品 C, 商品 D で大幅に向上したが商品 A で僅かではあるが下がった。

以上の結果を見てわかるように全体的には大幅な進展が見られた。とくにこの 4 品目すべてで両期間ともに相関係数を 0.6 以上に変化させたのは予測を行う上でとても大きな進展である。しかし予測を行った中には僅かであるが下がっているものも存在する。全体の精度を向上させながら個々の精度も下げない品目ごとのパラメーター設定も視野に入れて改善をしていくことにより、更なる予測精度向上につながると考察される。

表1 日ごとの予測を1週間にまとめたもの

商品A			商品B		
1週間					
331	362.6499	1.095619	133	130.7387	0.982998
292	292.4658	1.001595	83	100.2697	1.208068
311	318.1184	1.022889	103	107.0882	1.039691
208	270.3063	1.29955	79	95.04391	1.203087
230	261.9173	1.138771	105	122.3438	1.165179
296	281.9595	0.952566	61	110.8236	1.81678
278	272.5887	0.980535	118	118.7001	1.005933
商品C			商品D		
1週間					
21	42.09048	2.004309	69	64.4568	0.934157
100	59.90304	0.59903	141	89.2421	0.632923
62	62.82584	1.01332	147	60.51648	0.411677
36	47.35635	1.315454	32	38.36763	1.198988
64	61.36113	0.958768	49	47.6875	0.973214
42	59.65902	1.420453	105	68.34033	0.65086
92	64.42127	0.700231	59	69.65729	1.180632

表2 日ごとの予測を10日間にまとめたもの

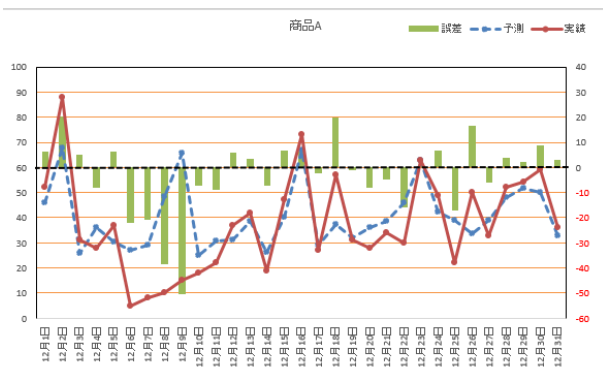
商品A			商品B		
10日間					
416	454.8751	1.09345	169	174.1058	1.030212
478	485.618	1.015937	143	150.085	1.049546
283	357.3246	1.262631	98	137.853	1.406663
360	384.5789	1.068275	136	167.0436	1.228262
444	411.4626	0.926718	144	171.1507	1.188547
商品C			商品D		
10日間					
43	65.32515	1.519189	120	103.0089	0.858408
140	88.17062	0.62979	235	105.0303	0.446937
74	78.94014	1.066759	42	58.60641	1.395391
56	84.3647	1.506513	137	80.49505	0.587555
112	92.64633	0.827199	85	100.4383	1.181627

### 3.2 売上上位商品の考察

#### ● 商品 A

競合商品である商品 F が特売日であった 12/6~12/10 の期間は競合し明らかな売上の低下がある。一方で商品 A が特売日であった 12/1~12/3 の売上は全日とも予測よりも多くの数が売れていることがわかる。大きな誤差が出ている 12/18, 26 は天気の影響が大きく外れた日であり 18 日は最低気温が 5.3 度、26 日は最高気温が 4.0 度の差がある。特売日の誤差を考慮せずに考えると、天気予報取得と予測周期の増加により、実際のデータに近いような値が予測できるのではないかと考察する。また、特売日には競合する商品なども視野に入れた何らかの説明変数の制定が必須であると分かった。

表3 商品 A 12月売上



### 3.3 12月の予測結果から検証する適切な発注頻度

売上上位 4 品目の日ごとの売上予測を 1 週間ごとにまとめたものと 10 日間にまとめたものどどちらが相関係数が高く実用的であるかといった検証では、以下のようになった。

表4 まとめ期間の相関係数比較

	1週間	10日間
商品A	0.671243	0.497352
商品E	0.335972	0.228144
商品F	-0.65377	0.234662
商品G	0.933409	0.895356

商品 F は本論 5.2 章で示したように特売日の影響をとっても大きく受けたため相関係数があまり参考にならない結果となったが、全体的には 10 日間でまとめるより 1 週間でまとめるほうが良い結果となることが示された。これより適切な発注頻度は 1 週間であるということがわかった。

## 4 おわりに

本研究では、ホームセンターから提供された売上数量過去データの分析と予測を行った。昨年度までは購買傾向の分析と売上増加のための販売促進策を考えていたが、今年度は売上という面で大きく影響する在庫コストについて最適化を行うことを目指しこの研究は行われた。在庫政策を考えるうえで最も必要な予測は売上数量に関するものである。

売上予測では一部商品群で相対誤差 2 割以内というものを出すことが出来た。また予測結果を一週間にまとめた際が最も実データに近い予測値を出しているということから売上予測を用いる際の発注頻度についても見出すことが出来た。この結果を用いて不要在庫を減らすとともに利益の向上に役立てることが出来ると予想される。

今後の展望としては、実用化できるような更なる予測精度の向上と売上予測自動化が挙げられる。

### 参考文献

- [1] 水野健斗・中島良隆・大原有紀子・杉本雄亮：ホームセンターの顧客データの分析と販売促進策の考察，2017 年度南山大学理工学部システム数理学科卒業論文，2018。
- [2] <https://www.hellocybemetics.teach>，今更聞けない Deep Learning の話【ユニット・層・正則化・ドロップアウト】，2018/10/11
- [3] <https://www.accuweather.com/ja/jp/japan-weather>，accuweather，2018/11/27