

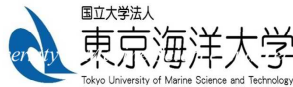


目次

- ・ 研究背景・目的
- ・ 研究対象と方法
- ・ 需要予測手法の紹介
- ・ データ分析と結果
- ・ まとめ

ニューラルネットワークを用いた需要予測モデル における入力変数の選定に関する研究

東京海洋大学大学院
海洋科学技術研究科 海運ロジスティクス専攻
1855015 廖 静妍
黒川研究室



2

発表の流れ

- ・ 研究背景・目的
- ・ 研究対象と方法
- ・ 需要予測手法の紹介
- ・ データ分析と結果
- ・ まとめ



研究背景：経営環境の変化

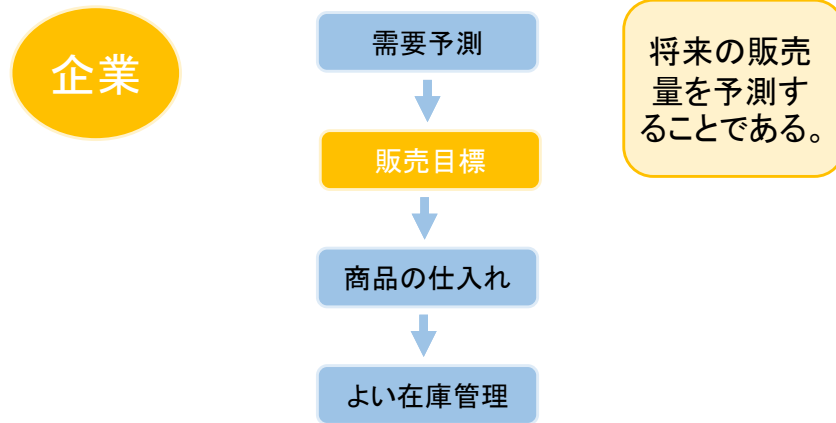
- ・ 企業を取り巻く諸要因には、経済的、社会的、自然的、文化的要因など多くの要因があり、近年、激しく変化している。なかでも、消費者の嗜好がバブル崩壊後に著しく変化している。そして、企業はこれらの激しい変化に対応し、存続、発展していかななくてはならない。



3

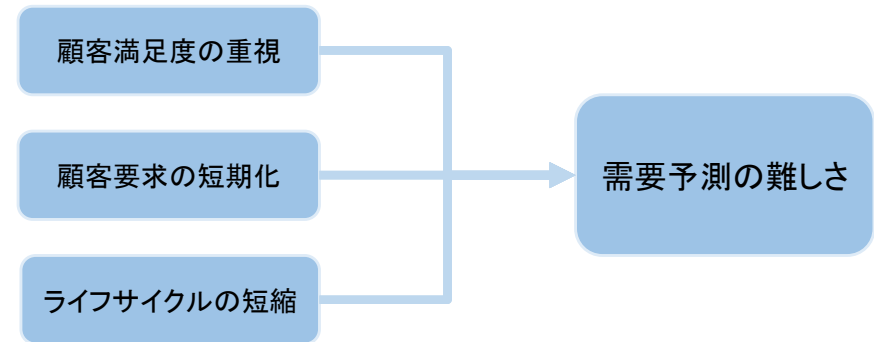
4

研究背景：需要予測の重要性



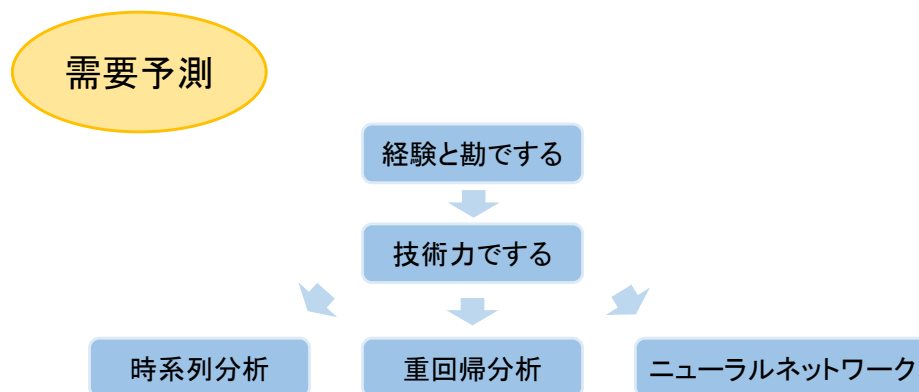
5

研究背景：需要予測の難しさ



6

研究背景：需要予測の手法



7

研究目的



- ・本研究では、時系列モデル、重回帰モデルとニューラルネットワークモデルを使って需要予測を行い、予測結果によってニューラルネットワークモデルの精度を検証する。
- ・また、入力変数が異なるニューラルネットワークを用いたモデルを使って需要予測を行い、予測結果によって入力変数の選定が需要予測精度に与える影響を検討することを目的とする。

8

発表の流れ

- ・ 研究背景・目的
- ・ 研究対象と方法
- ・ 需要予測手法の紹介
- ・ データ分析と結果
- ・ まとめ



研究対象

- ・ 対象ショップ：ネット通販ショップ
- ・ 商品：化粧品（450種）
- ・ 期間：2014年8月から～2018年7月



9

研究対象：統計データ

シルクミネラルファンデーション
(詰め替え用)

2014		2015		2016		2017	
商品番号	販売量	商品番号	販売量	商品番号	販売量	商品番号	販売量
f_mspf_001	3236	v_msf_c_r-1	2925	v_msf_f-1	5524	v_msf_f-1	4736
v_mpf_f-2	2983	v_msf_f-1	2882	v_msf_c_r-1	4249	v_msf_c_r-1	4520
v_msf_f-1	2912	v_mcc_f-1	2161	v_puff1_cs	3207	v_mcc-1	3152
v_uvfm_f-1	2338	v_puff1_cs	2124	v_mmb2_f-1	2506	v_mmb2-1	3084
v_mpf_f-1	2045	v_uvfm_c_r-1	2001	v_msf_f-2	2475	v_msf_f-2	2376
v_msf_f-2	1769	v_msf_ss-1-2	1719	v_uvfm_c_r-1	2443	v-bm-tr	2326
v_msf_ss-1-2	1719	v_msf_c_r-5	1693	v_mpf_f-2	2045	v_uvfm_c_r-1	2218
v_msf_f-5	1700	v_uvfm_f-1	1689	v_msf_c_r-5	1980	v_uvfm_f-1	1959
v_msf_ss-5-1	1596	v_puff1	1620	v_uvfm_f-1	1976	v_msf_c_r-2	1860
v_mmb_f-1	1546	v_mpf_f-2	1612	v_msf_c_r-2	1827	v_puff1	1801
v_uvfm_ss-1	1489	v_msf_f-2	1449	v_puff1	1732	v-uvcc	1799
msf_s_set	1366	v_msf_ss-5-1	1428	v_msf_f-5	1635	v_msf_c_r-5	1785
v_puff1	1089	v_msf_c_r-2	1392	v_uvmp_f	1481	v-msmb	1640
v_msf_c_r-1	833	v_msf_f-5	1356	v_mpf_c_r-2	1395	v_msf_f-5	1630
v_mif_f-2	729	v_mmb_f-1	1247	v_msf_ss-1-2	1327	v_msf_r-1	1500



研究方法

- ・ ソフト
EXCEL、SPSS
- ・ モデル
時系列モデル（ARIMAモデル）
重回帰モデル
ニューラルネットワークモデル（隠れ層は1層）



11

12

発表の流れ

- ・ 研究背景・目的
- ・ 研究対象と方法
- ・ 需要予測手法の紹介
- ・ データ分析と分析結果
- ・ まとめ



需要予測手法：時系列分析

- ・ 時間の経過とともに不規則に変動する現象の記録が時系列である。
- ・ ARIMAモデルは、自己回帰和分移動平均モデルと言われており、時系列分析の一種である。
- ・ 定義式：ARIMA (p,d,q)



13

需要予測手法：重回帰分析

- ・ 重回帰分析は、複数の説明変数を用い、目的変数を表す回帰式を算出し、次の事柄を明らかにする手法である。
- ①関係式に用いた説明変数の従属変数に対する貢献度（影響度、重要度）
- ②関係式に用いた説明変数の重要度ランキング
- ③予測
- ④潜在能力



ニューラルネットワークを用いた需要予測

- ・ 我々の脳は、ニューロンと呼ばれる神経細胞がシナプスと呼ばれる結合部位を介して多数結合してできているネットワークである。
- ・ ニューラルネットワークは、大量な比較的単純な情報処理要素（ユニット）を相互に結合したネットワーク上で、信号をやり取りすることにより人間の脳を真似た情報処理しようとする試みである。

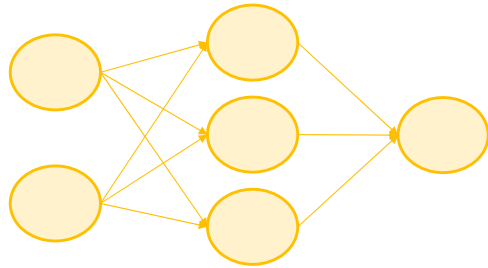


15

16



多層パーセプトロン

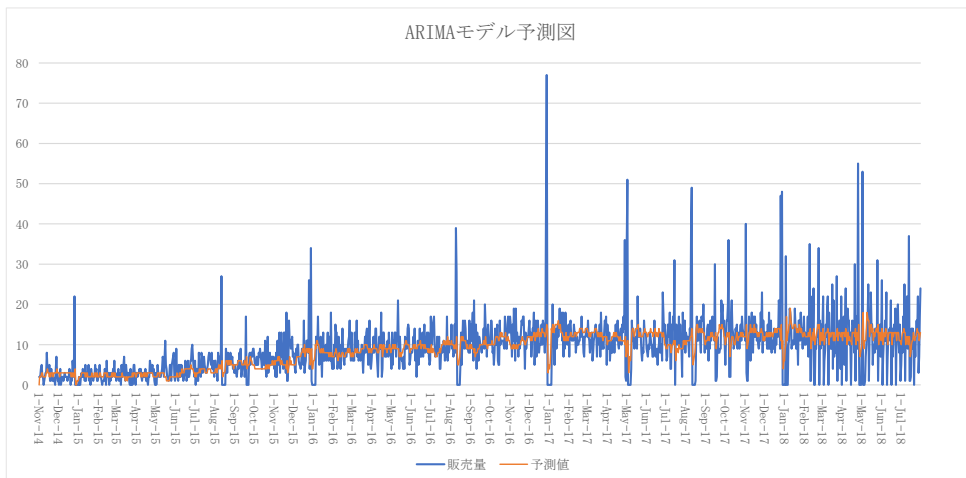


- ・ 研究背景・目的
- ・ 研究対象と方法
- ・ 需要予測手法の紹介
- ・ データ分析と結果：時系列分析とニューラルネットワーク分析
- ・ まとめ

時系列分析：ARIMAモデル



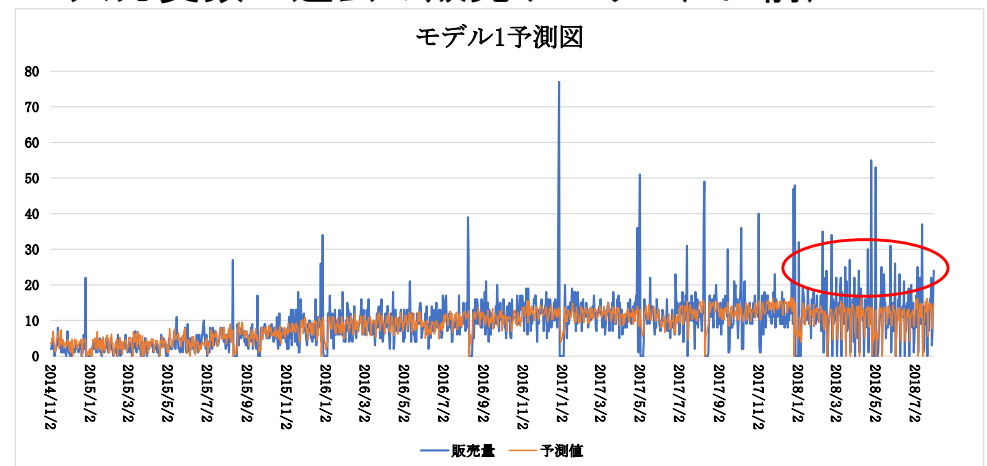
- ・ 目的変数：販売量
- ・ モデル式：ARIMA(1,0,13)



ニューラルネットワークモデル1



- ・ 出力変数：販売量
- ・ 入力変数：過去の販売データ（1日前）



分析結果の比較：誤差確認1

- ・ 指標：偏差平方和


モデル	ARIMA	ニューラルネットワーク
偏差平方和	48106	44036

同様なデータに対して需要予測をし、時系列モデルより、ニューラルネットワークモデルの予測精度が高い。



21

需要予測手法：重回帰分析

- ・ 目的変数：販売量
- ・ 説明変数： 



23

発表の流れ

- ・ 研究背景・目的
- ・ 研究対象と方法
- ・ 需要予測手法の紹介
- ・ データ分析と結果：重回帰分析とニューラルネットワーク分析
- ・ まとめ



22

データの処理：相関分析

- ・ 相関分析は二つの変数間の関係を分析する手法である。
- ・ 変数 x の値が大きいほど、変数 y の値も大きい場合を正の相関関係という。変数 x の値が大きいほど、変数 y の値が小さい場合を負の相関関係という。変数 x の値と、変数 y の値の間に直線関係が成立しない場合を無相関という。



24



グループ1		グループ2		グループ3	
日数	相関係数 (高い)	日数	相関係数 (やや高い)	日数	相関係数 (低い)
26	0.26173	13	0.21705	4	0.17673
20	0.26498	9	0.22452	3	0.18091
23	0.26578	12	0.22526	6	0.18198
29	0.27236	18	0.22631	5	0.18610
30	0.30057	19	0.23233	11	0.19264
7	0.31281	15	0.23492	24	0.19998
1	0.31861	22	0.24319	10	0.20364
28	0.31980	16	0.25065	2	0.20666
21	0.34507	27	0.25778	25	0.20769
14	0.36211	17	0.26114	8	0.21375



- ・ 目的変数：販売量
- ・ 説明変数：
 - ①グループ1（相関性が高い：0.26以上）：1、7、14、20、21、23、26、28、30日前の販売量
 - ②グループ2（相関性がやや高い：0.21～0.26）：9、12、15、16、17、18、19、22、27日前の販売量
 - ③グループ3（相関性が低い：0.2以下）：2、5、6、8、10、11、24、日前の販売量

重回帰分析の結果（グループ1）

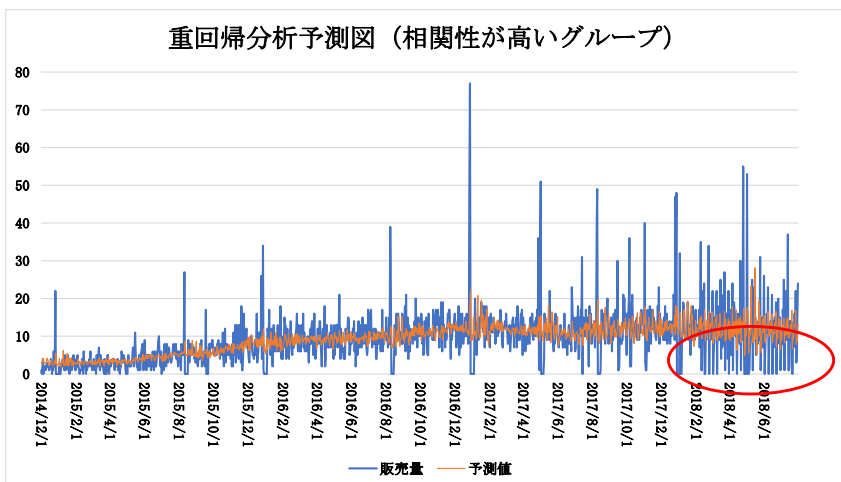


重回帰分析の結果（グループ2）

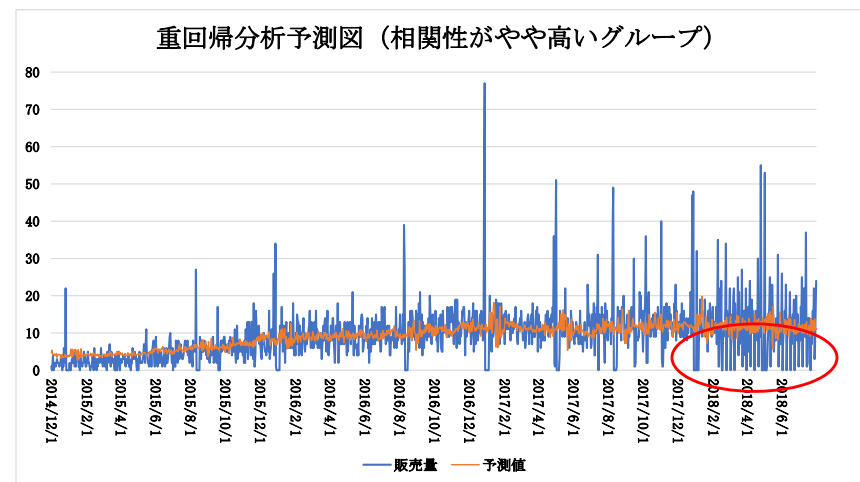


重回帰分析予測図（相関性が高いグループ）

重回帰分析予測図（相関性がやや高いグループ）



回帰統計	
重相関 R	0.534729
重決定 R2	0.285935
補正 R2	0.2811
標準誤差	5.972272
観測数	1339

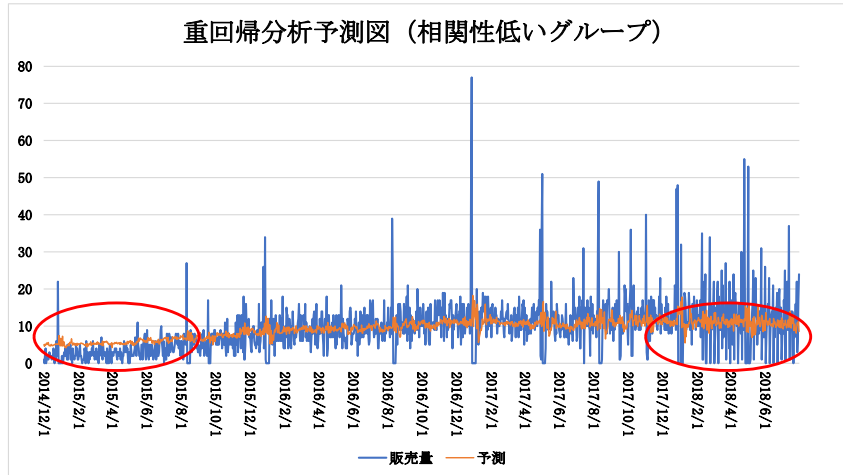


回帰統計	
重相関 R	0.428558
重決定 R2	0.183662
補正 R2	0.178134
標準誤差	6.385659
観測数	1339

重回帰分析の結果（グループ3）



重回帰分析予測図（相関性低いグループ）



回帰統計	
重相関 R	0.344296
重決定 R2	0.11854
補正 R2	0.113904
標準誤差	6.63049
観測数	1339

ニューラルネットワークモデルの設定



- 出力変数：販売量
- 入力変数：
 - モデル②過去の販売データ（相関係数が高いグループ）
 - モデル③過去の販売データ（相関係数やや高いグループ）
 - モデル④過去の販売データ（相関係数が低いグループ）

29

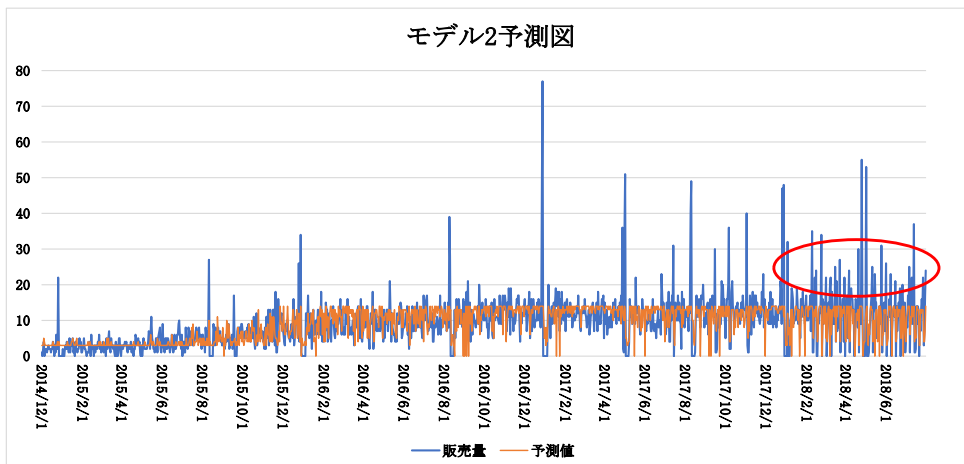
30

ニューラルネットワークモデル②



- 出力変数：販売量
- 入力変数：過去の販売データ（相関係数が高いグループ）

モデル2予測図

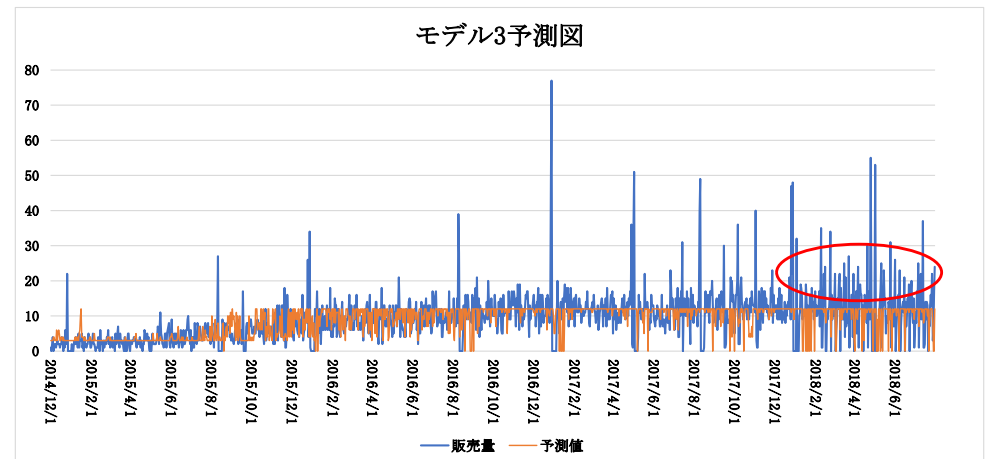


ニューラルネットワークモデル③



- 出力変数：販売量
- 入力変数：過去の販売データ（相関係数やや高いグループ）

モデル3予測図

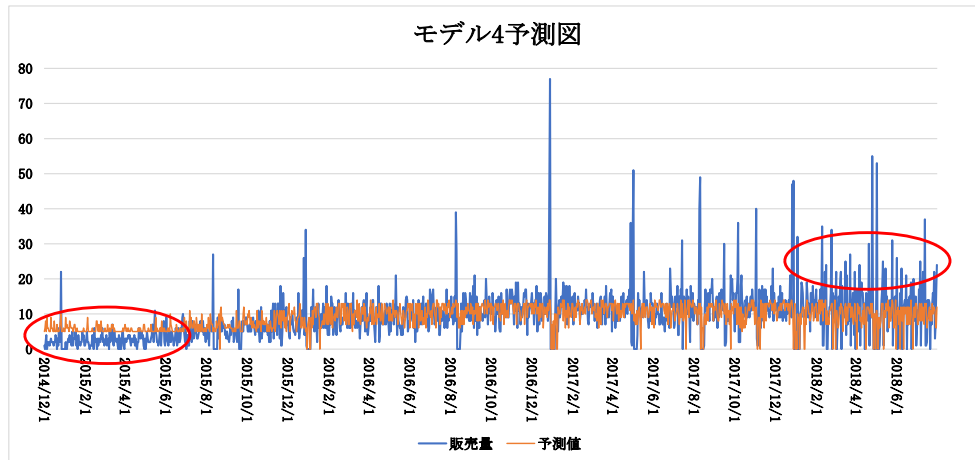


31

32

ニューラルネットワークモデル④

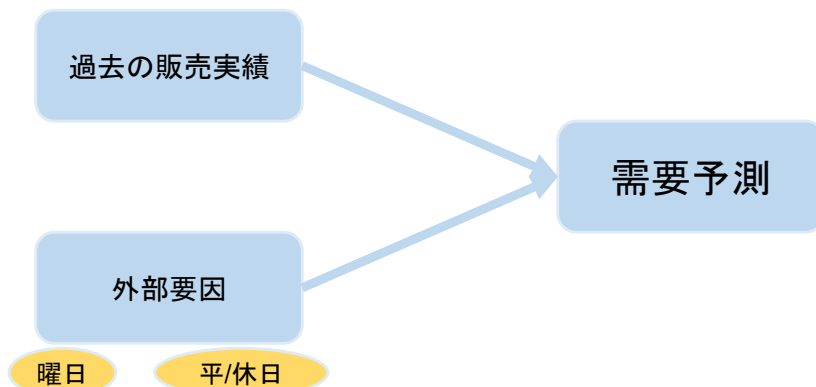
- 出力変数：販売量
- 入力変数：過去の販売データ（相関係数が低いグループ）



33

今まで分析した結果のまとめ

- トレンドを予測できない



35

分析結果の比較：誤差確認1

- 指標：偏差平方和

説明変数	重回帰モデル	ニューラルネットワーク
グループ1	47403	47271
グループ2	54192	54767
グループ3	58515	55535

ニューラルネットワークモデルに与える入力変数とし、出力変数との相関が高い変数を選択すると、モデルの精度が高くなる。

34

分散分析

- 分散分析：観測データにおける変動を誤差変動と各要因およびそれらの交互作用による変動に分解することにより、要因および交互作用の効果を判定し、統計的仮説検定の一手法である。

36

分散分析

①ケース幅：曜日

変動要因	変動	自由度	分散	観測された分散比	P-値	F 境界値
グループ間	2690.673	6	448.4455	9.384007	4.12E-10	2.105227
グループ内	65087.62	1362	47.78827			
合計	67778.30	1368				

②ケース幅：平日休日

変動要因	変動	自由度	分散	観測された分散比	P-値	F 境界値
グループ間	234.555	1	234.5549	7.167091538	0.008061	3.889839
グループ内	6348.971	194	32.72665			
合計	6583.526	195				

曜日/平日休日に変化/交替すると販売量に変化するという。
曜日/平日休日の変化が販売量の多少に影響を与える。



ニューラルネットワークモデル

出力変数：販売量

入力変数：

モデル⑤：過去の販売データ（相関係数が高いグループ）、曜日

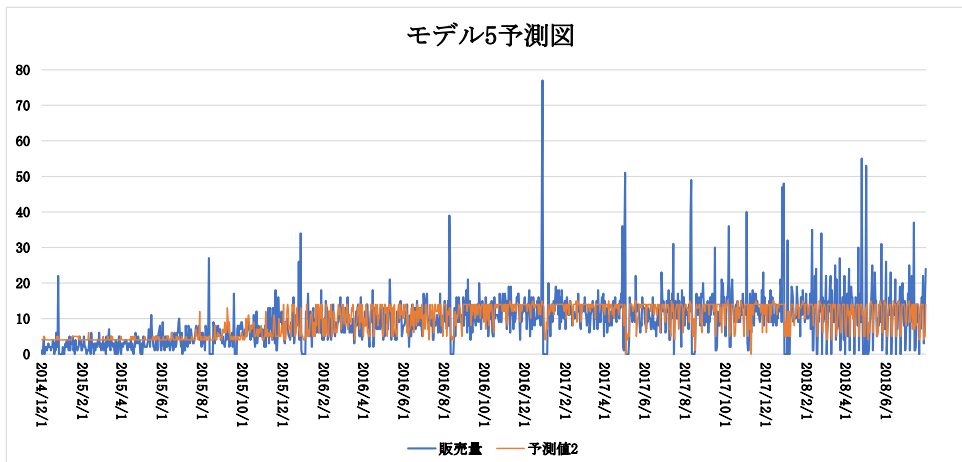
モデル⑥：過去の販売データ（相関係数が高いグループ）、平日休日



ニューラルネットワークモデル⑤

出力変数：販売量

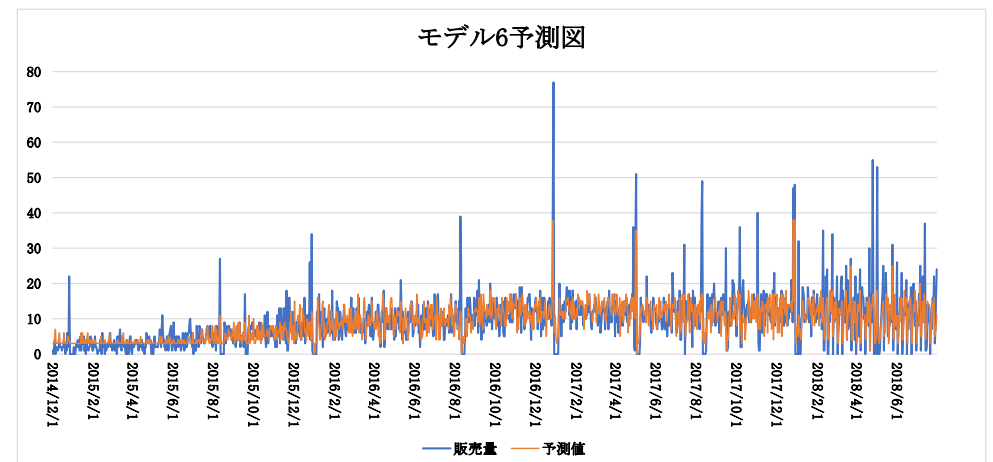
入力変数：過去の販売データと曜日情報



ニューラルネットワークモデル⑥

出力変数：販売量

入力変数：過去の販売データと平日休日情報





モデルの入力変数	偏差平方和
相関係数が高いグループ、曜日	43789
相関係数が高いグループ、平日休日	35207
相関係数が高いグループ	47271

ニューラルネットワークモデルに与える入力変数として、販売量に影響を与える因子を選定すると、モデルの精度が高くなる。

41

発表の流れ

- ・ 研究背景・目的
- ・ 研究対象と方法
- ・ 需要予測手法の紹介
- ・ データ分析
- ・ 分析結果
- ・ まとめ



43



- ・ 商品番号：v_msf_ss-5-1

モデル	ARIMA	ニューラルネットワーク
偏差平方和	5784	5268

説明変数	重回帰モデル	ニューラルネットワーク
相関係数が高いグループ	5645	5571
相関係数がやや高いグループ	6307	6372
相関係数が低いグループ	6703	6779

モデルの入力変数	偏差平方和
相関係数が高いグループ、曜日	5080
相関係数が高いグループ、平日休日	5044
相関係数が高いグループ	5571

42

まとめ

- ・ 商品の需要予測をする際に、機械学習によるニューラルネットワークを用いたモデルは、需要予測の精度向上に効果があるといえる。
- ・ また、ニューラルネットワークモデルに与える入力変数として、出力変数との相関が高い変数や曜日など、販売量に影響を与える因子を選定することにより、モデルの精度が高くなることが分かった。



44

今後の課題



- ・ ①予測する際に、用いるデータの期間の検討
- ・ ②まったく新しい商品の予測モデルの検討

ご清聴ありがとうございました。