

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

廖 静妍

指導教員 黒川 久幸 教授

1. はじめに

1.1 研究背景

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。そのため企業にとって、経営戦略やマーケティング戦略を立てる上で考慮しなければならないことは数多くあるが、中でも需要とされるのは販売目標、あるいは計画であろう。よい計画を立てるために、それらの目標あるいは計画の基礎として、精確な需要予測が必要となる。

しかし、商品の需要を精確に予測することが困難だと言える。「顧客要求の短期化」と「プロダクトライフサイクルの短縮」などの原因により、経営環境が激しく変化しており、今まで売っていたものが急に売れなくなることも少なくない。この問題に対し、過去は「需要予測は経験と勘である」と一部考えられていたが、近年は情報技術の発展により「需要予測は技術である」という考え方に転換した。

この需要予測でよく用いられている予測モデルとして、時系列モデルや重回帰モデルがあるが、近年ではニューラルネットワークを用いた AI を需要予測に応用した予測モデルが注目されている。機械学習によってデータ間の非線形な関係を含めた複雑な関係をネットワークで表現でき、従来の需要予測モデルより、精度が高い予測モデルを構築できると期待されている。

1.2 研究目的

そこで本研究では、時系列モデル、重回帰モデル、そしてニューラルネットワークモデルを使って商品の需要予測を行い、予測結果からニューラルネットワークモデルの精度を検証する。そして、ニューラルネットワークモデルの精度を向上させ

るために、与える入力変数の影響について検討するために、入出力データ間の相関係数に着目して入力変数をグループ分けし、相関の強さが予測精度に与える影響について検討を行うことを目的とする。

2. 研究対象と研究方法

2.1 研究対象

研究対象とする商品は、化粧品のネット通販ショップで取り扱われている商品で、売れ行きの良い主力商品とした。需要予測に用いるデータは、2014年8月から2018年7月まで、4年間の販売実績である。

2.2 研究方法

本研究では、次の3つの予測モデルを用いて、需要予測の結果を比較する。

- ・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)

そして、需要予測で用いる説明変数(入力変数)として、目的変数(出力変数)との相関係数に着目し、相関が「強い」、「やや強い」、「弱い」の3つのグループに分け、グループごとに10個の変数とするデータを作成した。

具体的には、過去1日前から30日前の販売量をもとに相関分析を行い、得られた3つのグループ・データから、予測に用いる説明変数(入力変数)の選択について検討する。

3. 従来の需要予測手法

3.1 時系列分析

時間の経過とともに変動する現象の記録データが、時系列データである。時系列データを解析する目的は、過去の情報を利用して将来の予測や意思決定に役立たせることである。

本研究で用いた時系列モデルは、ARIMA モデル

で、自己回帰和分移動平均モデルと言われている。図1に需要予測の結果を示す。また、用いたデータの期間等は、下記のとおりである。

データ分析期間：2014年11月1日～2017年12月31日

データ予測期間：2018年1月1日～2018年7月31日

ARIMAモデル：ARIMA(1, 0, 6)

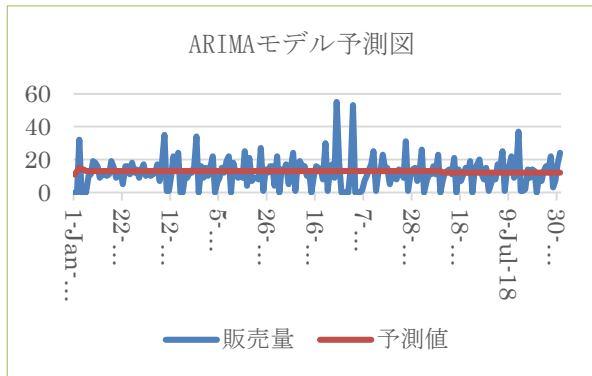


図1 ARIMAモデル予測図

3.2 重回帰分析

重回帰分析は、目的変数と複数の説明変数との関係を表したモデルで、目的変数への影響度についても分析できるモデルである。用いたデータの期間は、ARIMAモデルと同様である。

説明変数として、目的変数との相関係数に着目し、下記の3つのグループに分けた。

- (1) グループ1 (相関が強い：0.32以上)：1、14、17、20、22、27、29、30日前の販売量
- (2) グループ2 (相関がやや強い：0.28～0.31)：8、10、11、13、15、16、18、23、25日前の販売量
- (3) グループ3 (相関が弱い：0.28以下)：2、4、7、9、12、19日前の販売量

各グループの需要予測の結果を図2から図4に示す。

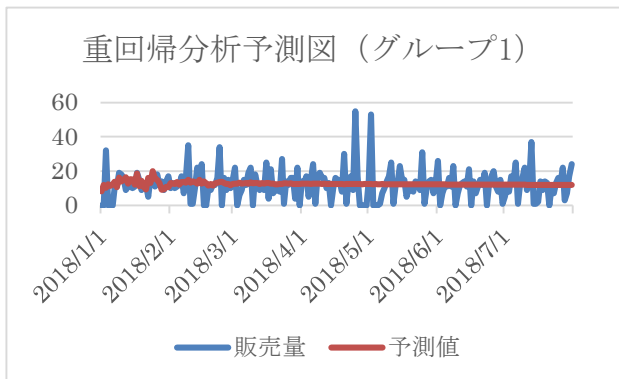


図2 グループ1予測図

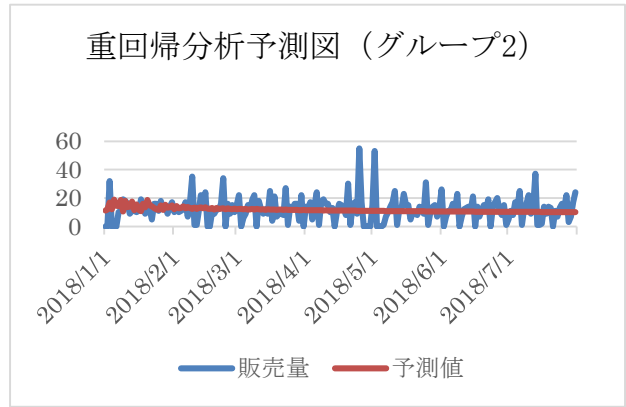


図3 グループ2予測図

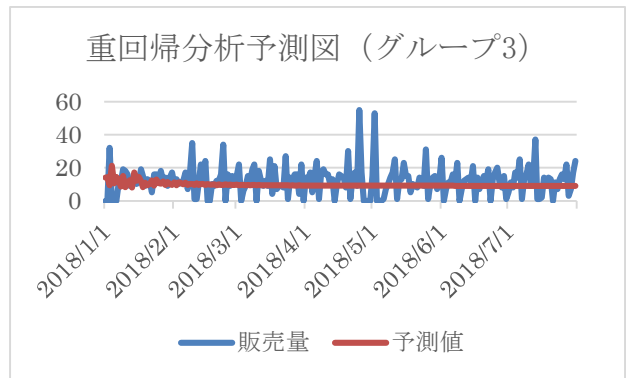


図4 グループ3予測図

4. ニューラルネットワークモデルを用いた需要予測

4.1 ニューラルネットワークの特徴と構造

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、この脳のネットワークを真似た物で、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。

本研究で用いたニューラルネットワークの構造は、図5のような階層的なネットワークである。

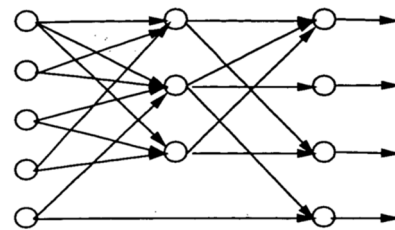


図5 階層型ネットワーク

階層型ネットワークでは、ユニットが複数の階層をなすようになり、左の層(入力層)から中間の層(隠れ層)、そして右の層(出力層)へ向か

う方向の結合のみが許される。現在、もっともよく使われている多層パーセプトロンは、このタイプのネットワークの代表例である。多層パーセプトロンは、入力変数の値に基づき、一つ以上の目的変数の予測モデルを生成できる。また、目的変数と入力変数の間の特定の関係を事前に仮定することなく、幅広い統計モデルに近似させることができる特徴を持っている。

なお、本研究では、隠れ層を1つ持つ階層型ネットワークを用いた。

4.2 ニューラルネットワークモデルの構築と分析結果

ニューラルネットワークモデルを用いた需要予測の結果を図6から図9に示す。図6は、1日前のでたのみを用いた場合で、図7から図9は重回帰分析で説明した3つの相関グループの結果である。

(1) 過去の販売データ (1 日前)

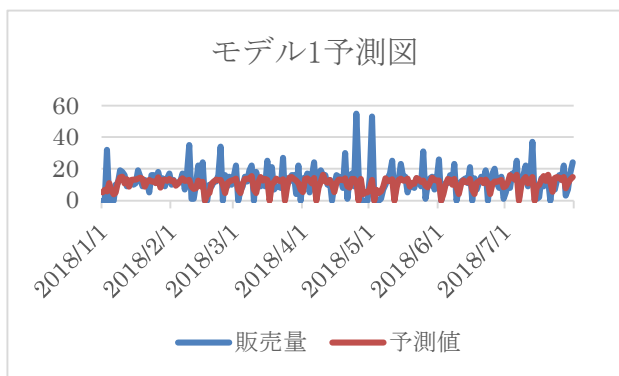


図6 (1) 予測図

(2) 過去の販売データ (相関係数が高いグループ)

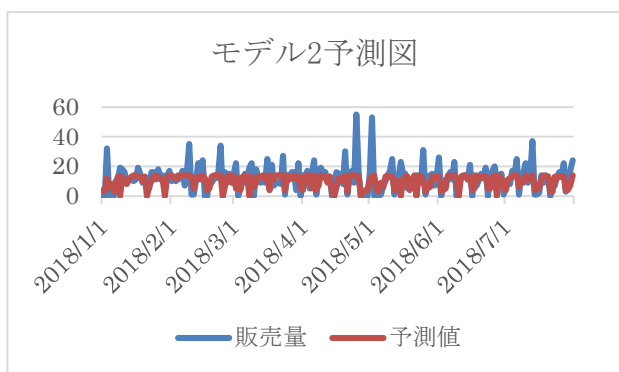


図7 (2) 予測図

(3) 過去の販売データ (相関係数がやや高いグループ)

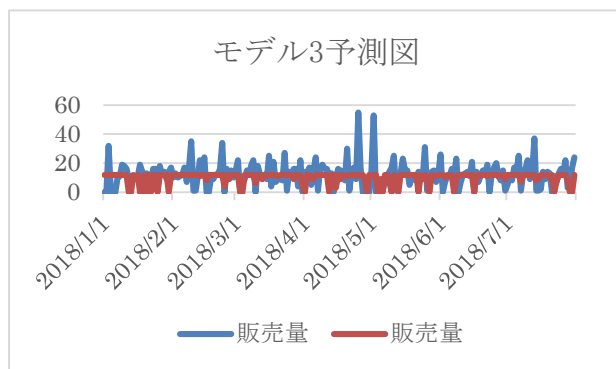


図8 (3) 予測図

(4) 過去の販売データ (相関係数が低いグループ)

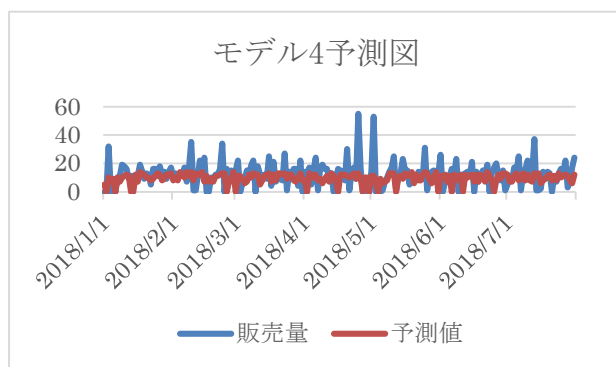


図9 (4) 予測図

次に、販売量に影響を与えている要因として考えられる曜日と平日休日の影響を考慮した予測結果を図10と図11に示す。

(5) 過去の販売データ (相関係数が高いグループ) と曜日情報

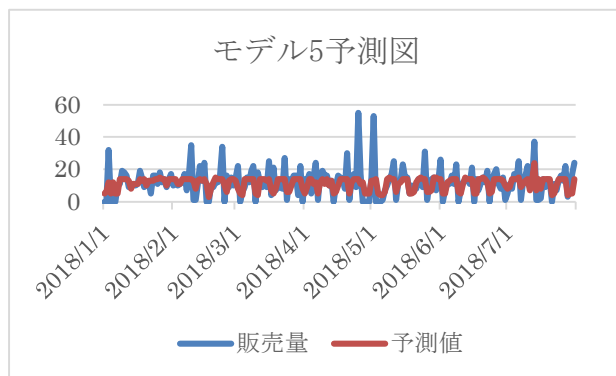


図10 (5) 予測図

(6) 過去の販売データ (相関係数が高いグループ) と平日休日情報

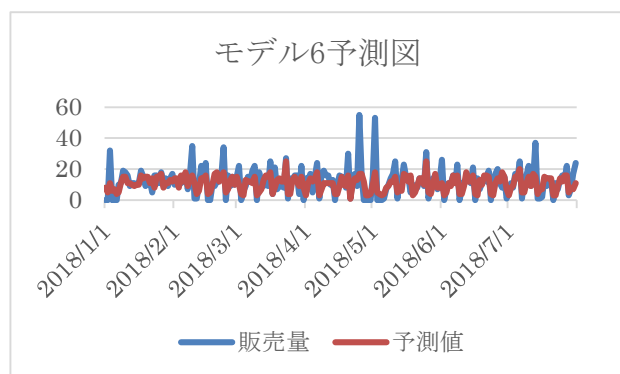


図 11 (6) 予測図

5. 分析結果の比較

5.1 従来の需要予測モデルとニューラルネットワークモデルの結果比較

3つの予測モデルの精度を比較するための指標として、偏差平方和を用いる。この偏差平方和の値が小さければ小さいほど、予測値と実際値の差が小さく、モデルの精度がよいことを表す。

表 1 ARIMA モデルとニューラルネットワークモデルの誤差

モデル	ARIMA	ニューラルネットワーク
偏差平方和	15592	12151

表 2 重回帰モデルとニューラルネットワークモデルの誤差

説明変数	重回帰	ニューラルネットワーク
グループ 1	15653	12568
グループ 2	16123	15941
グループ 3	16728	15050

表 1 と表 2 より、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く、重回帰モデルとニューラルネットワークモデルの予測精度がほぼ変わらないことが明らかとなった。

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、相関が異なる 3つのグループの結果から、販売量との相関性が高いデータを入力変数とした場合が最も精度が高いと分かった。

すなわち、出力変数と無関係な入力変数を用いては意味がなく、どのように事前に入力変数に影響を与える入力変数を選定しておくかが重要であることが分かった。

5.2 入力変数が異なるニューラルネットワークモデルの結果比較

さらに、予測精度を向上させるために、販売量に影響すると考えられる曜日と平日休日を加えた場合の予測精度についてまとめた結果を表 3 に示す。

表 3 より、曜日など販売量に影響を与える因子を入力変数としてモデルに入れることにより、さらに予測精度の向上をはかれることが分かった。

表 3 入力変数が異なるニューラルネットワークモデルの誤差

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	11450
相関係数が高いグループ、平日休日	9345

6. おわりに

本研究では、時系列モデル、重回帰モデル、そしてニューラルネットワークモデルを使って商品の需要予測を行った。その結果、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ことが分かった。

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精度が高くなることが分かった。

これらの結果から、出力変数と無関係な入力変数を用いては意味がなく、どのように事前に入力変数に影響を与える入力変数を選定しておくかが重要であることが分かった。

参考文献

- (1) 奥本勝彦：販売予測とマーケティング・モデルの選択, 多賀出版, 1997.
- (2) 浅田克暢・岩崎哲也・青山行宏：在庫管理のための需要予測入門, 東洋経済新報社, 2006.
- (3) Robert G. Brown: 在庫管理のための需要予測, 紀伊国書店, 1961.
- (4) IBM SPSS Forecasting 22
- (5) IBM SPSS Neural Networks25
- (6) 北川源四郎: 時系列解析入門, 岩波書店, 2005.
- (7) 栗田多喜夫・本村陽一: 階層型ニューラルネットワークとその周辺, 応用統計学, Vol. 22, No. 3, pp. 99-115, 1993.