

製造装置消耗品の突発受注予測手法に関する検討 Prediction method for burst orders of consumable items of manufacturing equipment

小山 光[†]
Koyama Hikaru

1. はじめに

製造装置のメンテナンスビジネスにおいて、顧客満足度の観点から、できるだけ短い納期で消耗品を納入することが求められている。顧客からの受注に対して間を置かず応じるには、顧客の受注量・頻度を予測し、在庫を準備しておく必要がある。一方、顧客からの受注量・頻度には、目安となる数量、周期などのパターンがある訳ではなく、製造装置メーカーでは突発的な受注への対応が課題となっている。

本報告では、消耗品の受注履歴データから機械学習(Random Forest, RNN)を用いて、突発的な受注を予測する手法を検討した。具体的には、各品目における 1, 2, 3 ヶ月先の受注を、最少受注量の 2 倍以内で予測することを目標とし、その後、各品目の上位 20%以上を越える受注数を「突発受注」として予測できるかを検証した。

2. 関連研究

適正な在庫を持つこと、顧客からの受注を予測することは、サプライチェーンマネジメント分野での一般的な課題である。成書も多くある伝統的な分野であるが、21 世紀に入ってから IT の進展・普及もあり、現在も注目されている分野でもある[1][2]。しかし、突発的な受注予測に着目すると、関連研究は非常に少なく、例えば突発的な大口受注の数量を極値分布を用いて解析した報告[3]がある。また、受注履歴データを等間隔で取得された時系列データと考えれば、機械学習による予測手法が報告[4][5]されており、特に遠い将来の予測を目的としている。

3. 分析手法およびデータ前処理

本報告で用いたデータは、微細なパターンを加工する製造装置の消耗品の受注実績データである。期間は 2015 年 1 月から 2018 年 5 月(41 ヶ月分)、品目数は 11 品目である。突発受注と呼ばれる、極端に受注数量が多い受注実績を含む品目を選んでいく。図 1 に、11 品目中 2 品目の受注推移を示す。品目 1 では、2016 年 2 月から 2017 年 2 月までの受注数が前後の期間と比較して、2 倍以上多い。また、品目 2 では、2015 年 8 月から 2016 年 8 月にかけて 3 回、2017 年 2 月から 2018 年 6 月にかけて 3 回、前後の期間と比較して 3~5 倍程度多い受注を受けている。

消耗品の受注数推移を学習・予測するために、機械学習手法である Random Forest と RNN を用いた。Random Forest は決定木系の代表的な手法として、RNN は Neural Network 系の代表的な手法として選択した。これらの機械学習手法は、入力とするデータ構造が異なるため、それぞれの機械学習手法に適したデータ前処理が必要である。Random Forest については、目的変数に、ある品目の来月(1 ヶ月先)の受注数を設定した場合、今月のすべての品目の受注数を説明変数とすることとした。2, 3 ヶ月先の受注数を予測す

る場合は、対応する説明変数の時点をずらせばよい。1 つの目的変数に、1 行の説明変数が対応するため、入力データは行列型データとなる。

一方、RNN については、Random Forest と同様に、ある品目の来月(1 ヶ月先)の受注数を目的変数とした場合、すべての品目の 24 ヶ月分の受注数を説明変数とすることとした。1 つの目的変数に、11 品目×24 ヶ月分の説明変数が対応するため、入力データはテンソル型データとなる。

両手法とも、38 ヶ月分のデータを学習に、残り 3 ヶ月分を検証に用いた。

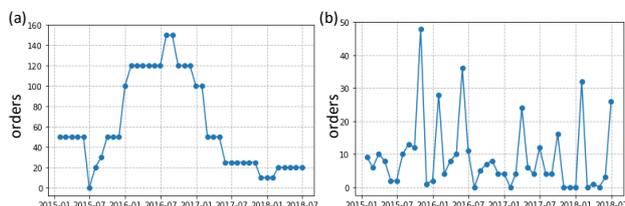


図 1: 受注数の推移 (a)品目 1, (b)品目 2

4. 1,2,3 ヶ月先の受注予測結果

図 2 に、品目 1, 2 の 1 ヶ月先の受注を、Random Forest で学習・予測した結果を示す(ページ数の都合で RNN による学習・予測結果はここでは省略し、発表当日に報告する)。なお、品目 1 は予測精度が最良だった品目、品目 2 は予測精度が最も悪かった品目である。横軸は、2015 年 2 月を 1 とした時の月数、縦軸は受注実績である。青い実線が実際の受注実績、青いプロットが学習データによる予測値、オレンジのプロットが未知のデータを Random Forest モデルに入力した場合の予測値である。プロットのエラーバーは、乱数条件を変えて 100 回取得した予測値の標準偏差を示している。

学習データによる予測は、受注実績に対してほぼ一致しており、Random Forest による学習が成功していることが判る。また、未知データに対する予測も、受注実績に対してかなり近い値を予測できており、標準絶対誤差(Mean Absolute Error:MAE)は、2.0 個であった。目標予測精度を、各品目の最少受注数(品目 1 の場合 10 個)の 2 倍以内(品目 1 の場合 20 個以内)と設定していたので、十分目標を達成している。

一方、図 2(b)に示した品目 2 の予測結果では、学習データによる予測は受注実績とおおむね一致しているが、未知データの予測(39, 41 ヶ月目の予測)が実績値から大きく外れている。MAE も 11.6 個であり、最少受注数に対して差が大きい。これは、直前の 38 ヶ月目の予測が実績値から大きく外れていることが影響している可能性がある。

[†]株式会社日立製作所 Hitachi, Ltd.

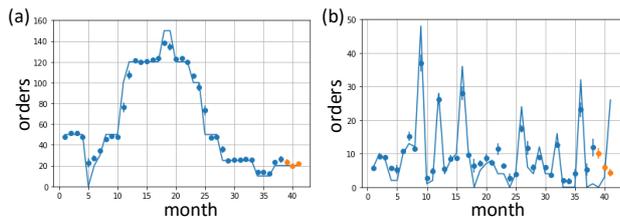


図 2:1 ヶ月先の受注数の学習・予測結果
(a) 品目 1, (b) 品目 2

図 3 に品目 1 の 2, 3 ヶ月先を予測した結果を示す。学習データに対する予測はおおむね実績値と一致しているが、11~13 ヶ月目のずれが 1 ヶ月先予測と比較して大きい。これらの月は、他の月と比較して受注数が多く、また、受注増加の立ち上がりでもある。このような急な変化の予測に対しては、予測精度が低下する可能性を示唆している。また、未知データに対する予測でも、2 ヶ月先予測では 1 点、3 ヶ月先予測では 2 点が実績値と比較して大きく異なった受注数を予測している。このことから、予測時点が将来になるに従い、予測精度が低下する可能性が示唆される。

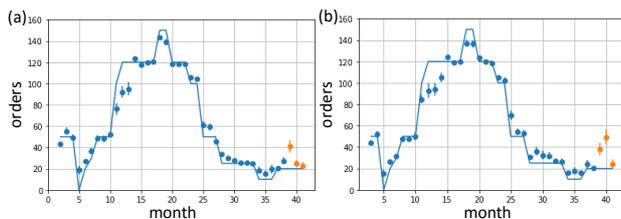


図 3:品目 1 の 2,3 ヶ月先の受注数の学習・予測結果
(a) 2 ヶ月先, (b) 3 ヶ月先

5. 考察

5.1 Random Forest と RNN の比較

本報告では、消耗品の受注数を予測するために、Random Forest と RNN を用いた。ページ数の都合で記載できなかったが、両手法の予測精度を比較すると、11 品目中 10 品目で Random Forest の方が、予測精度が高かった。

RNN は信号処理、自然言語処理など時間方向の変化、事象の順序に意味があるデータ、いわゆる時系列データの学習・予測に有効であることが知られている。本報告で対象とした消耗品の受注数も時間方向に整理できる時系列データであるため、RNN による学習・予測を試みたが、目標とした予測精度を満足した品目数は 11 品目中 6 品目であった。

RNN より Random Forest の方が、予測精度が高かった理由として、RNN と Random Forest に入力するデータ形状の違いとハイパーパラメータ設定が考えられる。本報告では、RNN への入力データは、24 ヶ月分を学習データ、直後の 1 ヶ月を教師データとした。この学習・教師データを 1 組とすると、今回の 41 ヶ月分の受注データからは、18 組のデータしか取ることができない。3 組のデータを精度検証用に充てると、15 組のデータで学習することになる。一方、Random Forest は第 3 節で述べたように受注データの対応を 1 時点分ずらせばよいだけなので、検証用に 3 組のデータ

を除いたとしても、38 組のデータを学習・予測に用いることができる。このように Random Forest の方が、学習・予測に用いるデータ数を多く取れたことが、RNN より予測精度が高かった要因と考えられる。

5.2 突発予測に関する検討

前触れなく多量の受注が届く「突発受注」が予測できているかを検証した。本報告では、41 ヶ月分の最後の 3 ヶ月分を検証データとした。この 3 ヶ月分 11 品目の受注実績に対して、突発受注を予測できているかを調べた。「突発受注」の定義は、各品目の上位 20% より多い受注とした。予測できたか否かは、予測した受注数が上位 20% を越えていれば、予測できたと判定した。表 1 に、実際に突発受注であったか否か、突発受注と予測したか否かを整理した混同行列を示す。表 1 から、正解率 0.76、適合率 0.86、再現率 0.46 であった。全体の件数に対する突発受注の件数の比が 0.39 なので、再現率の観点からは、単純な確率による推測よりやや良い程度である。一方、適合率は 0.86 と高い値であり、本手法により突発受注と予測されれば、高い確率で実際に突発受注が届くと予測できる。

表 1: 突発受注の実績・予測に関する混同行列

		actual	
		Yes	No
prediction	Yes	6	1
	No	7	19

6. おわりに

本報告では、製造装置の消耗品の突発受注を予測することを目的とし、Random Forest と RNN を用いて、1, 2, 3 ヶ月先の受注予測を試みた。突発受注を各品目の上位 20% より多い受注と定義し、11 品目、3 ヶ月分のデータを用いて検証したところ、適合率 0.86 で突発受注を予測することができた。今後の課題は、新しい受注データに対する予測精度の検証と予測に基づいた在庫計画、消耗品手配の業務設計が挙げられる。

参考文献

- [1] 2018 年版中小企業白書, 中小企業庁編 (2018).
- [2] 山口雄大, この 1 冊ですべてわかる需要予測の基本, 日本実業出版 (2018).
- [3] 中塚昭宏 他, “極値分布を用いた突発的な大口需要の予測法に関する研究”, 日本経営工学会論文誌 Vol. 69 No. 3 (2018).
- [4] S. B. Taieb *et al*, “Multi-output modeling for multi-step-ahead time series forecasting”, *Nerocomputing* 73 (2010).
- [5] L. Zhang *et al*, “Iterated time series prediction with multiple support vector regression models”, *Nerocomputing* 99 (2013).