

時系列予測モデル構築における データプロファイリング・前処理支援手法の開発

株式会社日立製作所 工藤文也

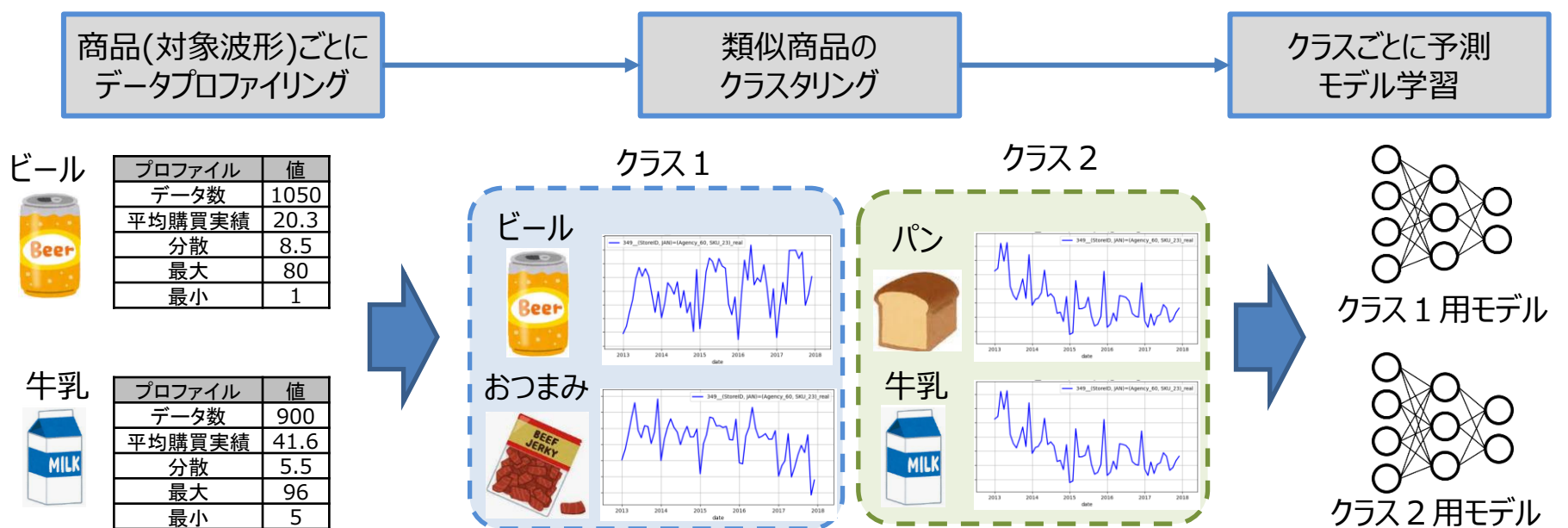
背景・課題

- ・小売りでは、数万商品の需要予測が求められる
- ・ニューラルネットワークなど大規模モデルが用いられるが、**傾向の異なる波形が混在すると予測精度が低下**してしまう
- ・分析者の負担を増やさず予測精度向上したい

アプローチ

- ・予測したい商品の統計情報を用いて波形をクラスタリングし、類似する波形に分類
- ・分類されたクラスごとに予測モデルを構築することで、一つのモデルで予測する場合と比べて予測精度が向上するか検証

データプロファイリング×クラスタリングによる波形の自動分類



実験結果

| 評価指標 | (1)従来TFT モデル(1クラス) | (2)提案TFTモデル (前処理 + TFT) | | | |
|-------|-----------------------|-------------------------|------|------|-------|
| | | 2クラス | 4クラス | 6クラス | 10クラス |
| RMSE | 707 | 522 | 498 | 443 | 453 |
| sMAPE | 0.61 | 0.6 | 0.59 | 0.55 | 0.56 |
| MASE | 1.6 | 1.64 | 1.48 | 1.54 | 1.59 |

モデル : TFT (Temporal Fusion Transformer) [1]
データ : Volume Forecasting (Stallion & Co.) [2]
評価指標 : RMSE, sMAPE, MASE ※小さいほど予測精度高

[1] <https://arxiv.org/abs/1912.09363>

[2] <https://www.kaggle.com/datasets/utathya/future-volume-prediction>

まとめ・今後の課題

■まとめ

- ・購買傾向の類似する波形同士がクラスタリングできている様子を確認
- ・クラスタリングにより一つのモデルで予測する従来と比べて、予測精度が向上することを確認

■今後の課題

- ・予測モデルを変えた場合やクラス数に応じたモデルパラメータのチューニング