

日本のガス使用量における決定要因に関する研究 —多層パーセプトロンモデルによる分析—

A Study on Determining Factors of Amount of Gas Used in Japan
— Analysis using multi-layer perceptron model —

許 慎碩¹ 村上 英治^{1,2} 高橋 大志¹

Myeongseok HEO¹ Eiji MURAKAMI^{1,2} and Hiroshi TAKAHASHI¹

¹慶應義塾大学大学院経営管理研究科 ²アズビル金門(株)

¹Graduate School of Business Administration, Keio University ²Azbil Kimmon Corporation

Abstract: Gas and electric power are one of the essential social infrastructure for social and economic activities. The purpose of this research is to construct a gas' consumption model suitable for the distribution and sales market of LP gas in Japan by utilizing temperature. As a method of predicting demand, we would like to first create a model of power consumption and then apply it to LP gas. Analysis of customers' gas consumption can infer from what standpoint they were consuming public goods. Through this, we would like to offer users' reasonable price and usage to expand their choices and contribute to the efficient use of social resources. Furthermore, analysis of LP gas usage through deep learning is possible.

1. はじめに

社会インフラエネルギーについての研究は数多く行われている。しかし、LP ガスの使用量についての研究は、使用率が低い等の理由から、相対的に十分な研究が行われてこなかった。

都市ガスが普及していない地域においては LP ガスの利用率が高く、小さなベンダーが各地域で競争している傾向にある。地域別の LP ガスの使用量についての研究を進めることで、今後ベンダーと消費者の双方にプラスのメリットをもたらすことが期待される。このような背景から、本研究では、ガス使用量の決定要因に関する研究を進める¹。

2. 関連研究

電力消費量に焦点を当てた研究は数多く報告されている。例えば、UCI(University of California Irvine) Machine Learning Repository が提供している Electricity Load Diagrams Dataset[1]を活用した電力消費量についての研究[4]では、季節ごとの電力の消費量の変化に着目している。本研究では、電力市場などにおける先行研究の成果を取り込み、日本のガス

市場に関する研究を行う[2-7]。

3. データおよび分析結果

本分析では、LP ガス消費量の季節ごとの消費量を分析した。分析対象期間は 2019 年-2020 年であり、国内における複数の都市を対象として分析を行った。

3.1 気温と使用量の関係性

Fig. 1 は、静岡市における気温と LP ガスの使用量の関連性を示したグラフである。

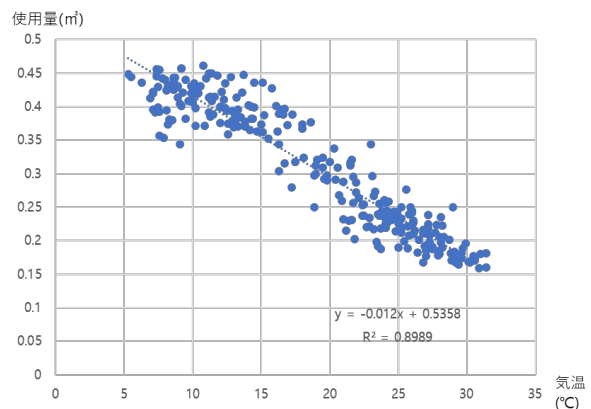


Fig.1 気温、LP ガス使用量の関連性 (静岡市)

¹ 本研究を実施するにあたりアズビル金門(株)から研究の支援を受けた。ここに記して感謝する。

3.2 各気温におけるLP ガス使用量の推定

Table. 1 は、各温度における各利用者のガス使用量の推計値(概略)を示したものである。推計においては、最も広く知られた教師有学習モデルの一つである K 最近傍法 (KNN : K-Nearest Neighbor Algorithm) を用い、データセットを構築した² [8-10]。

Table.1. 気温とガス消費量の関連性 (概略)

気温(°C)	ID1(m ³)	ID2(m ³)	...
:	:	:	...
5.6	0.19	:	...
5.7	0.18	:	...
5.8	0.15	:	...
5.9	0.13	:	...
:	:	:	...

3.3 利用者の分類 : K-means によるクラスタリング

構築したデータセットを対象としてクラスタリングを実施した。本稿では、クラスター数を 4 として分析を実施した。

分析の結果、(1) 使用量の変動が大きいクラスター、(2) 気温の影響が小さいクラスター、(3) 気温の上昇とともに使用量が大きく減少するクラスター、(4) 気温の上昇とともに使用量が緩やかに減少するクラスターなどを特定することができた。

Fig. 2 は、気温の上昇とともに使用量が大きく減少するクラスターに含まれるサンプルを示したものである。

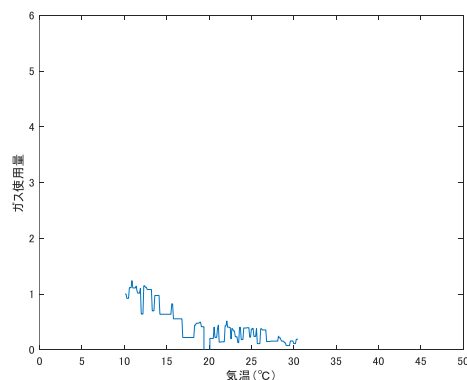


Fig.2 気温と LP ガス使用量 (サンプル)

3.4 多層パーセプトロンモデルの詳細

静岡の 2020 年 1 月から 2020 年 9 月のデータに限定し、以下の四半期別に集計した。

- (1) 2020_Q1 : 2020 年 1 月~3 月 (気温低い)
- (2) 2020_Q2 : 2020 年 4 月~6 月 (気温中間)
- (3) 2020_Q3 : 2020 年 7 月~9 月 (気温高い)

次は以下のデータ間の差分を取り、

- (1) 差分 Q2-Q1 : (2020_Q2-2020_Q1)
- (2) 差分 Q3-Q2 : (2020_Q3-2020_Q2)

差分 Q2-Q1 と差分 Q3-Q2 の符号を求める。(+, 0, - の 3 種類) 正常なパターンならば差分 Q2-Q1 と差分 Q3-Q2 が全部「-」になり、それ以外は全部「正常ではない」となる。

本論文では、静岡のガスデータ (差分 Q2-Q1、差分 Q3-Q2、パターン (正常、正常ではない)) における多層パーセプトロンモデルによる分析を行う[11]。分析においては、入力層 (2 ノード)、隠れ層 (3 ノード)、出力層 (2 ノード) とした。また、隠れ層と出力層の活性化関数は二値分類のため、シグモイド関数を採用した。深層学習時に使用された教師データとテストデータは、それぞれ既存データの 50% を使用した。学習回数(Epoch)を 200 回、バッチサイズは 4 に設定した。学習結果として loss は 3-4% 程度、accuracy は 98% 程度となった。

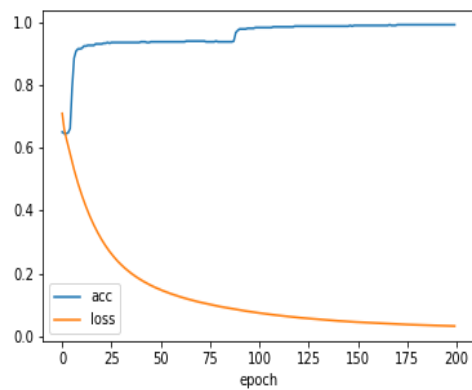


Fig.2.多層パーセプトロンモデルの深層学習グラフ

² 本稿では K=5 とした場合の結果を示す。

Table 2 は、分析結果を示したものである。

Table.2. 実際、予測結果の分析結果

予測 \ 実際	Normal	Abnormal
Normal	279	4
Abnormal	2	157

4. まとめおよび今後の課題

本分析では、日本の LP ガス使用量と気温の関連性について分析を行った。本研究では、静岡市のデータに焦点を当て分析を行い、LP ガス使用料データから利用者をいくつかのタイプに分類できることを見出した。今後の課題としては、分析対象地域の拡大や、より詳細なユーザ属性を考慮した分析などが挙げられる。

参考文献

- [1] <https://archive.ics.uci.edu/ml/datasets/ElectricityLoadDiagrams20112014>
- [2] 江田健二: かんたん解説!! 1時間でわかるガス自由化入門, <https://pps-net.org/taivavi-interview> (2017年6月)
- [3] 村上英治: 価値創出を指向するメーターデータプラットフォームガスミエール™, *Azbil Technical Review*, pp. 20-23, (2020年4月)
- [4] Yasirli Amri, Amanda Lailatul Fadhilah, Fatmawati, Novi Setiani, Septia Rani: Analysis Clustering of Electricity Usage Profile Using K-Means Algorithm, IOP Publishing, (2016)
- [5] <https://archive.ics.uci.edu/ml/datasets/ElectricityLoadDiagrams20112014>
- [6] Sebastián de la Torre, José M. Arroyo, Antonio J. Conejo, and Javier Contreras: Price Maker Self-Scheduling in a Pool-Based Electricity Market: A Mixed-Integer LP Approach, *IEEE TRANSACTIONS ON POWER SYSTEMS*, Vol. 17, No. 4, pp. 1037-1042, (NOVEMBER 2002)
- [7] Alireza Rahimi Vahed, Teodor Gabriel Crainic, Michel Gendreau, Walter Rei: A Path Relinking Algorithm for a Multi-Depot Periodic Vehicle Routing Problem, *CIRRELT*, pp.1-33, (2013)
- [8] Altman, N. S. (1992). "An introduction to kernel and nearest-neighbor nonparametric regression". 《The American Statistician》46 (3): 175-185.
- [9] 元田浩, 津本周作, 山口高平, 沼尾正行: データマイニングの基礎, オーム社 (2006)
- [10] Trevor Hastie, Robert Tibshirani, Jerome Friedman: *The Elements of Statistical Learning Data Mining, Inference, and Prediction*, Springer New York (2009) (杉山将, 井手剛, 神尾敏弘, 栗田多喜夫, 前田英作(監訳) 井尻善久, 井出剛, 岩田具治, 金森敬文 他 (訳) 統計的学習の基礎 —データマイニング・推論・予測—, 共立出版 2014)
- [11] EEG signals classification using the K-means clustering and a multilayer perceptron neural network model Umut Orhan a , Mahmut Hekim a , Mahmut Ozer, *Expert Systems with Applications* 38(2011)