

消費者行動とシステムデザインの結びつき

The Linkage Between Consumer Behavior and System Design

九大（非）*チャップマン アンドリユー

Andrew Chapman*

*Kyushu University

1. 概要

本研究は、個人の行動や嗜好が、環境・経済・社会の持続可能性に関する意識をどの程度推定できるかを探るものである。従来は、統計解析を用いた調査研究が主流であったが、本研究では統計的アプローチに併せて、機械学習手法の両方を用いて比較・検証を行った。特に、非特定個人情報に基づき、日常行動から人々の価値観を推定するモデルの有効性に焦点を当てている。最終目的はエネルギーシステムデザインであって、人々の嗜好、選考にあった将来システムの構築とともに、カーボンニュートラル達成に貢献する。

2. 背景

持続可能な社会実現に向けて、個人の行動変容が重要であるとされる中、温室効果ガス排出の 6 割以上が家庭消費など個人由来であるという報告もある (Ivanova et al., 2016)。しかし、環境配慮行動の背景にある意識や価値観を把握するには、従来のアンケートによる手法には限界があり、特に回答者への負担の削減、回答の質の向上が何よりも重要であることがこれまでの研究で明らかとなった (Chapman, 2023)。近年、アンケート調査の分析において機械学習の活用が広まりつつあり、計算能力の高さに加え、品質管理やリアルタイムでのデータ監視、傾向の把握といった点でも利点があるとされている (Piroddi, 2022)。特に、ニューラルネットワークは従来の統計的回帰分析では捉えにくい非線形な関係性のモデル化に優れており、意思決定に関わる変数の分類には決定木などの手法が高い精度を示すことが報告されている (Chan et al., 2020)。また、階層構造のあるデータや協調フィルタリング、教師あり・半教師あり・教師なし学習など、さまざまなアプローチが応答予測に応用されている (Gharibshah and Zhu, 2021)。さらに、エネルギー関連分野においても、機械学習はエコ効率の評価や省エネルギー化に貢献しており、スマートメーターや IoT センサーと組み合わせることで、建物のエネルギー効率向上や熱的快適性、室内空気質の改善にも寄与している (Himeur et al., 2021)。本研究では人々にとって重要である持続可能性の指標を測るアンケートの結果を用いて、人口動態は、日常及び消費行動が重要性に及ぼす影響を予測とエネルギーシステムデザインへの応用を目的とする。

3. 研究手法

本研究では、消費者の購買行動および持続可能性に対する重要度認識に影響を与える要因を明らかにするため、先進的な機械学習技術を用いた包括的な計算フレームワークを構築した。分析対象は、製品カテゴリごとの消費者の回答データであり、属性情報としては、人口統計 (demographics)、行動パターン (behaviours)、および生活経験 (lived experience) を含む。

分析にはアンケート調査データを用い、指定された特徴量に対する欠損値は除去した。加えて、都道府県コードを日本の地域別にマッピングすることで、地理的要因による行動の違いを文脈的に補足した。

各製品カテゴリにおいて設定された複数の目的変数 (例：環境、安全性、価格など) に対し、XGBoost 分類器による特徴量選択を実施した。この際、F1 スコア (weighted) を性能指標として、グリッドサーチ付き交差検証を行い最適なモデル構成を探索した。選択された主要特徴量を用いて、Random Forest、Multi-Layer Perceptron (ニューラルネットワーク)、およびこれらのアンサンブル手法による複数分類モデルの学習と予測を行った (図 1)。

各モデルは、訓練データ (全体の 80%) およびテストデータ (20%) に分割して適用し、分類精度、macro-F1 スコア、weighted F1 スコア、クラス別 precision-recall スコアなど複数の評価指標で性能を比較した。また、1~5 のリッカート尺度を用いた三値分類 (重要でない、中立、重要) に変換し、混同行列により分類結果の可視化も行った。分類閾値はカテゴリごとの分布に応じて適応的に調整した (例：重要でない=1-2、中立=3、重要=4-5 など)。さらに、各モデルによる予測の説明性を高めるため、SHAP (SHapley Additive exPlanations) 解析を実施し、特徴量の寄与度を可視化した。これには、SHAP summary plot、特徴量重要度の棒グラフ、個別予測のウォーターフォールプロット、特徴量の相互作用を示す部分依存プロット、およびクラス別の説明統計等を含む。各目的変数に対して重要とされた特徴量一覧とそれに対応する予測性能 (F1 スコア、

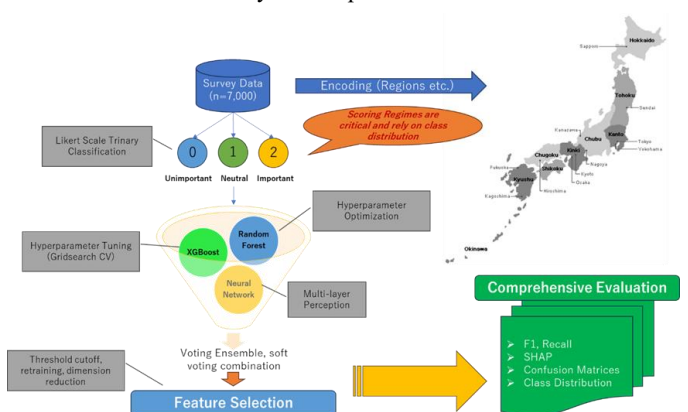


図 1. 解析モデルの構成

リコールなど)を比較し、どの人口統計・生活経験要因が予測精度に大きく寄与したかを明示した。このように、本研究は、データ前処理、モデル構築、性能評価、説明可能性分析を統合的に行うことで、消費者意識に影響を与える要因を定量的かつ解釈可能な形で抽出・分析することを目的としている。

4. 結果

機械学習モデルによる予測結果は、①三値分類におけるクラス分布、②人口統計情報のみを用いた場合と生活経験情報を含めた場合の比較、③モデル間の予測精度および重要特徴量の比較、の3点に基づいて報告する。

図2は、食品に関する重要度の例として、価格、環境影響、社会的側面に関するクラス分布(重要でない、中立、重要)を示しており、それぞれの項目における回答傾向の違いが確認できる。たとえば、価格は「重要」とする回答が多く、環境は比較的バランスが取れており、社会的側面は「中立」回答が多い傾向がみられた。これ

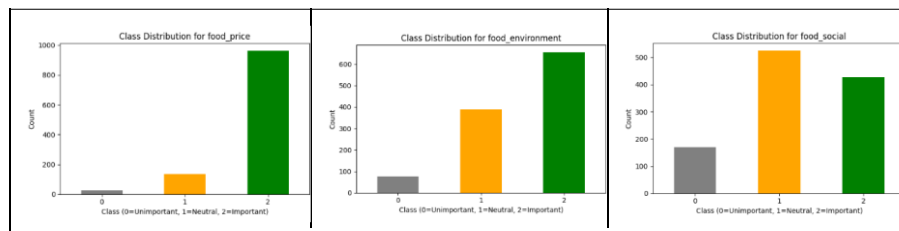


図 2. 食品に対する嗜好のクラス分布

らの分布に応じた三値分類(1=重要でない、2=中立、3=重要)を採用することで、モデル精度の向上を図った。この分類法に基づく予測精度は、価格に対して最も高く、74.7~87.6%を記録し、特にニューラルネットワーク(MLP)によるモデルがわずかに優れた成績を示した。維持費、安全性、品質についても、それぞれ最大で82.7%、81.8%、82.7%と高い精度が得られた。一方、社会的側面や環境要素などについては予測精度が相対的に低かった。

図3では、予測精度が高かった上位10項目(a)と、それらにおける主要な特徴量(b)を示している。

年齢と世帯収入が最も影響力のある人口統計要因であり、これに続いて生活経験要因、教育、性別などが重要であった。地域別では、関東地域に居住していることが一部の項目(車の価格、安全性、維持費など)の予測に影響を与えていた。また、生活経験要因を除外した場合、モデルの予測精度はむしろ向上した。これは、人口統計情報のほうが欠損が少なく、モデルがより多くのサンプルを学習・評価に利用できたためである。

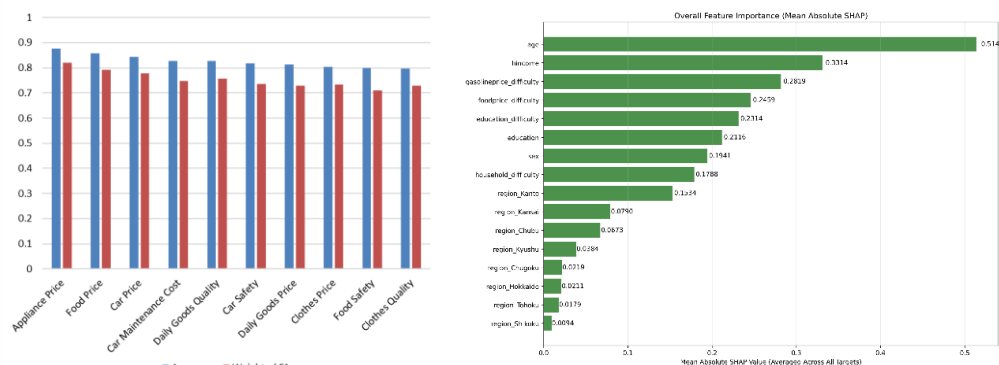


図 3. (a)モデル予測精度、(b)要素の影響度

5. 結論

本研究は、個人の人口統計情報や日常行動・生活経験に基づいて、持続可能性に関する価値観や意識をどの程度予測可能かを、機械学習手法を用いて検証した。結果として、価格や安全性、維持費などの項目は予測精度が高く、特にニューラルネットワークモデルが最も優れた性能を示した。一方、社会的側面や環境意識の予測は相対的に困難であり、モデル性能もやや低下する傾向が確認された。特徴量の重要性分析においては、年齢や世帯収入などの人口統計情報が最も影響力のある変数であり、加えて教育、性別、地域差も予測に寄与していることが明らかとなった。

これらの知見は、人々の意識や価値観を日常行動や非特定個人情報から推定するモデルの有効性と限界を示すとともに、持続可能なエネルギーシステム設計において人々の嗜好を中心としたアプローチを取り入れる際の実践的な手がかりを提供する。将来的には、より精緻な因果推論手法の導入などにより、本研究で示された手法のさらなる高度化と応用展開が期待される。

文献

- 1) Ivanova, D et al., 2016. Environmental impact assessment of household consumption. J. Ind. Ecol. 20, 526–536.
- 2) Chapman A. Enhancing Survey Efficiency and Predictive Ability in Energy System Design through Machine Learning: A Workflow-Based Approach for Improved Outcomes. Energies 2023;16.
- 3) Piroddi, L. Special Topics in Information Technology; SpringerBriefs in Applied Sciences and Technology, Switzerland, 2022.
- 4) Chan, K. Y et al., Affective Design Using Machine Learning: A Survey and Its Prospect of Conjoining Big Data. Int. J. Comput. Integr. Manuf. 2020, 33, 645–669
- 5) Gharibshah, Z.; Zhu, X. User Response Prediction in Online Advertising. ACM Comput. Surv. 2021, 54, 1–43.
- 6) Himeur, Y et al., A Survey of Recommender Systems for Energy Efficiency in Buildings: Principles, Challenges and Prospects. Inf. Fusion 2021, 72, 1–21.