

小売サービスにおける大規模データ融合による顧客行動予測システム

Customer Behavior Prediction System by
Large Scale Data Fusion in a Retail Service

山崎隼大

u320073@st.pu-toyama.ac.jp

富山県立大学情報システム工学科 3 年

December 12, 2025

背景と目的

背景

サービス産業において生産性を向上させるためには、顧客の価値観やライフスタイルの多様化に合わせ、個々の要求に過不足なく対応する「個別対応」が不可欠である。

本研究の目的

- ID-POS データ（購買履歴）と顧客アンケート（意識調査）という異種データを融合し、顧客と商品の適切なカテゴリを発見する手法を開発する。
- 発見されたカテゴリに基づき、顧客の購買行動を予測するシステムを構築する。

潜在クラスモデルによるデータ融合

異種データ（ID-POS とアンケート）を統合するため、PLSI を拡張した確率的生成モデルを提案する。

モデルの確率構造

顧客 x_i 、商品 y_j 、潜在顧客カテゴリ u_k 、潜在商品カテゴリ v_l の同時確率は以下のように定義される。

$$P(x_i, y_j, u_k, v_l) = P(u_k)P(x_i|u_k)P(v_l|u_k)P(y_j|v_l) \quad (1)$$

- $P(u_k)$: 潜在顧客カテゴリの生起確率
- $P(x_i|u_k)$: 顧客カテゴリ u_k に顧客 x_i が属する確率
- $P(v_l|u_k)$: 顧客カテゴリ u_k が商品カテゴリ v_l を選好する確率
- $P(y_j|v_l)$: 商品カテゴリ v_l に商品 y_j が属する確率

パラメータ推定

観測データ N_{ij} （顧客 x_i が商品 y_j を購入した回数）に対する対数尤度を最大化する。

対数尤度関数

$$\begin{aligned} L &= \sum_i^X \sum_j^Y N_{ij} \log P(x_i, y_j) \\ &= \sum_i^X \sum_j^Y N_{ij} \log \left\{ \sum_k^U \sum_l^V P(u_k) P(x_i | u_k) P(v_l | u_k) P(y_j | v_l) \right\} \quad (2) \end{aligned}$$

- EM アルゴリズムで最大化
- AIC によりカテゴリ数 U, V を決定

パラメータ推定

はじめに
提案手法
分析結果
予測モデル
結論

潜在顧客カテゴリ u_k 潜在商品カテゴリ v_l

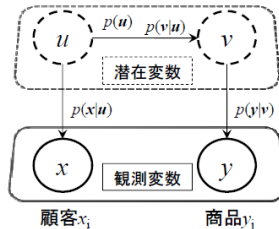


図2 提案モデル

対数尤度関数

$$AIC = -2\ell + 2[\text{Para}] \quad (3)$$

- AIC によりカテゴリ数 U, V を決定

アンケートに基づく制約の導入

ライフスタイルカテゴリのスコアを以下のように定義する。

ライフスタイルスコア k の定義

$$L_k^{(i)} = \frac{\sum_{Q_q^{(i)} \in \mathcal{L}_k^{(i)}} Q_q^{(i)}}{|\mathcal{L}_k|} \quad (4)$$

確率の定義

$$p(u_k | x_i) = \frac{L_k^{(i)}}{\sum_k L_k^{(i)}}, \quad (5)$$

$$p(u_k) = \frac{\sum_i L_k^{(i)}}{\sum_i \sum_k L_k^{(i)}}, \quad (6)$$

$$p(x_i) = \frac{\sum_j N_{ij}}{\sum_i \sum_j N_{ij}}. \quad (7)$$

カテゴリマイニングの結果

抽出された顧客カテゴリ

- こだわり消費派
- 家庭生活充実派
- 節約消費派
- アクティブ消費派
- 堅実消費派
- パパッと消費派

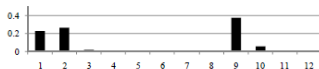
抽出された商品カテゴリ

- 健康的
- 料理用
- 高品質
- 低価格

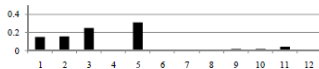
※ 提案モデルは、k-means や従来 PLSI より高い性能

カテゴリマイニングの結果

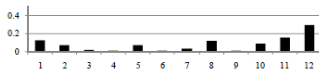
はじめに
提案手法
分析結果
予測モデル
結論



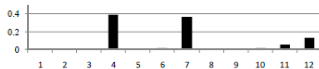
(a) $p(v|u_1)$ の推定結果 (こたわり消費派)



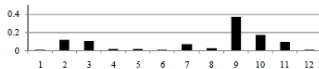
(b) $p(v|u_2)$ の推定結果 (家庭生活充実派)



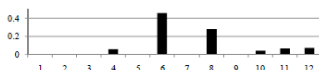
(c) $p(v|u_3)$ の推定結果 (アクティブ消費派)



(d) $p(v|u_4)$ の推定結果 (節約消費派)



(e) $p(v|u_5)$ の推定結果 (堅実消費派)



(f) $p(v|u_6)$ の推定結果 (それほど消費派)

表 2 潜在商品カテゴリへ分類された商品の傾向

カテゴリ番号	特徴的な商品
1	高品質 PB 商品
2	生野菜・生鮮品
3	日配品
4	低価格帯商品
5	鮮魚・肉類
6	肉類・パン・飲料
7	セールへの頻出商品
8	多種混合
9	高価格帯野菜
10	小サイズの野菜・日配品
11	飲料
12	惣菜・飲料

カテゴリベースの顧客行動予測

ベイジアンネットワークの構築

- ノード：顧客カテゴリ、商品カテゴリ、状況変数（時間帯・季節・安売り等）
- 条件下における購買確率を推論可能

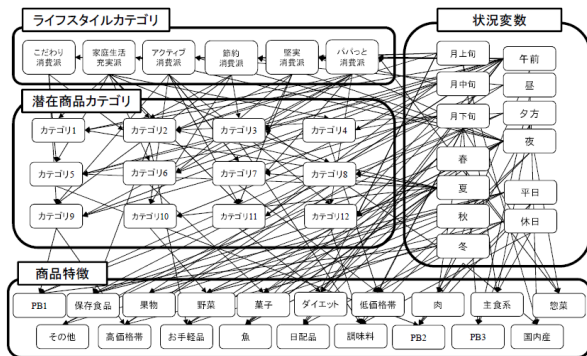


図 6 構築したベイジアンネットワークモデル

まとめと今後の展望

まとめ

- ID-POS とアンケートを融合したカテゴリ発見法を確立
- カテゴリに基づく状況依存型の購買行動予測システムを開発

応用と展望

- 小売サービス：顧客満足度向上・在庫最適化
- EC・通販サービスにも展開可能
- 予測に基づくプロモーション設計、店舗支援の実証が今後の課題