気分テキストによるコーヒー選択の予測

—Sentence-BERTを用いた自由記述文からの直接分類—

定方 晶　東京工科大学大学院

キーワード:感情マーケティング，レコメンドシステム，消費者行動　　日本マーケティング学会 マーケティングカンファレンス 2026

① 研究背景・目的

・コーヒー市場の多様化に伴い，消費者の選択が複雑化

・気分や体調に合わせてコーヒーを選ぶ行動が一般化

・既存レコメンドは購買履歴ベースが主流；「今の気分」に基づく推薦は少ない

⇒ 「今の気分」テキスト入力だけで最適なコーヒーを提案するシステムの実現

■ 研究目的

自由記述テキストからコーヒー種類を予測する機械学習モデルを構築し，その有効性を検証する

■ リサーチクエスチョン

RQ1:Sentence-BERTで気分テキストをベクトル化し，SVMで直接分類できるか？
RQ2:感情カテゴリを介した従来手法と比較して，直接分類はどちらが高精度か？

② 提案手法

■ 手法A(提案):Sentence-BERT + SVM による直接分類

気分テキスト（自由記述） → Sentence-BERT（384次元） → SVM（RBFカーネル） → コーヒー種類（予測）

■ 手法B(比較):感情カテゴリを介した従来的アプローチ

気分テキスト（自由記述） → LLMによる感情ラベリング → 感情カテゴリ（4種類） → SVM→コーヒー種類

・手法A(提案):1ステップで実現。中間カテゴリ不要，誤差の伝播なし。
・手法B(比較):LLMラベル付与後に分類。解釈しやすいがLLM誤差が伝播する可能性。
※ Sentence-BERT:paraphrase-multilingual-MiniLM-L12-v2(多言語対応，384次元)

③ LLMラベリング検証(手法Bの前処理)

項目	内容
対象データ	995件(Q3自由記述テキスト)
使用モデル	gpt-oss-20b＋qwen3.8-27b(Groq経由)
ラベルの種類	感情カテゴリ4種(癒し系・覚醒系・習慣系・その他)
整合性指標	Cohen’s κ = 0.638(辞書×LLM，752件)
一致率	一致率 77.0%(辞書×LLM，752件)

κ = 0.638 →「実質的一致(Substantial Agreement)」(Landis & Koch，1977)

① 感情カテゴリを入力とするSVM(手法B)として実装済み(2クラス精度:63.7%±4.8%，+13.7pt)
② BERTファインチューニング用の教師データとして活用予定
③ カテゴリ別コーヒー推薦ルールの構築(解釈可能なレコメンドシステムへの応用)

④ 分析手法の詳細

■ 特徴量エンジニアリング

特徴量	手法	次元数
テキストベクトル	Sentence-BERT(paraphrase-multilingual-MiniLM-L12-v2)	384次元
感情カテゴリ	LLMラベル(4種)のone-hot符号化(手法Bの入力)	4次元
入力の使い分け	手法A:テキストベクトルのみ/ 手法B:感情カテゴリのみ	384/4次元

■ 分類器・学習設定

設定項目	内容
分類器	SVM(RBFカーネル)/Random Forest(300木，比較用)
クラス不均衡対処	class_weight='balanced'(少数クラスを自動重み付け)
モデル評価	5分割層化交差検証(Stratified K-Fold)
評価指標	Accuracy(平均±標準偏差)，ランダムベースライン比
ベースライン	2クラス:50.0%，6クラス:16.7%(均等分類の場合)
実装環境	Python 3.11，scikit-learn，sentence-transformers

⑤ データ収集

項目	内容
調査方法	Freeasy(オンラインアンケートパネル)
調査期間	2025年～2026年
有効回答数	995件(テキスト欠損・無効回答を除外後)
Q1(選択肢)	飲みたいコーヒーのイメージ(8択)
Q3(自由記述)	今の気分や状況(モデル入力)
目的変数	コーヒー種類(2クラス/6クラス)

■ 6クラスの内訳

クラス	コーヒー種類	件数
1	ブラック系	181件
2	エスプレッソ単体	126件
3	エスプレッソ＋スチームミルク系	165件
4	エスプレッソ＋フォームミルク系	146件
5	エスプレッソ＋少量ミルク系	178件
6	ドリップ＋ミルク系	199件

⑥ 分析結果

■ 手法別の分類精度(5分割交差検証)

手法	分類	精度(mean±SD)	差
SBERT+SVM(手法A)	2クラス	**58.6% ±5.0%**	**+8.6pt**
SBERT+SVM(手法A)	6クラス	**22.0% ±2.8%**	**+5.3pt**
Random Forest(比較)	2クラス	61.4% ±3.1%	+11.4pt
感情+SVM(手法B)	2クラス	63.7% ±4.8%	+13.7pt
感情+SVM(手法B)	6クラス	17.6% ±2.6%	+0.9pt

■ 考察

・2クラスでは全手法がランダムベースライン(50.0%)を超え，気分テキストが嗜好予測に有効であることを示した
・2クラスでは感情カテゴリ経由(手法B)63.7%が直接分類(手法A)58.6%を上回ったが，差(+5.1pt)は標準偏差と同程度であり，優劣は断定できない
・6クラスでは直接分類22.0%が感情カテゴリ経由17.6%を上回った(ランダム16.7%)。4カテゴリへの要約で細かな違いが失われる可能性がある。追加データ収集・BERTファインチューニングによる改善の余地が大きい

⑦ まとめ・今後の課題

① Sentence-BERTによる気分テキストのベクトル化＋SVMにより，コーヒー種類の直接分類に成功
② 2クラス精度58.6%(±5.0%)，6クラス精度22.0%(±2.8%)を達成(ベースライン超え)
③ 感情カテゴリ経由の手法B(LLMラベリング κ=0.638)は2クラスで63.7%，6クラスで17.6%。2クラスは手法Bが高め，6クラスは手法Aが高い

■ マーケティング的含意

・感情状態テキストがコーヒー嗜好の予測に有効であることを示唆
・「気分」という主観的データのマーケティング活用可能性を提示
・カフェ等でのリアルタイムレコメンド実装の基盤技術となりうる

■ 今後の課題

① 6クラス分類精度の向上(追加データ収集・均等化)
② 日本語BERTのファインチューニングによる精度向上
③ Streamlitアプリとしてのサービス実装・ユーザー評価

■ 主要参考文献

Reimers & Gurevych (2019). Sentence-BERT: Sentence Embeddings using Siamese BERT-Networks. EMNLP.
Landis & Koch (1977). The measurement of observer agreement for categorical data. Biometrics, 33, 159-174.