



新しい音楽推薦システムの構築と有効性

東京学芸大学附属国際中等教育学校5年2組18番 竹内理恵



要旨

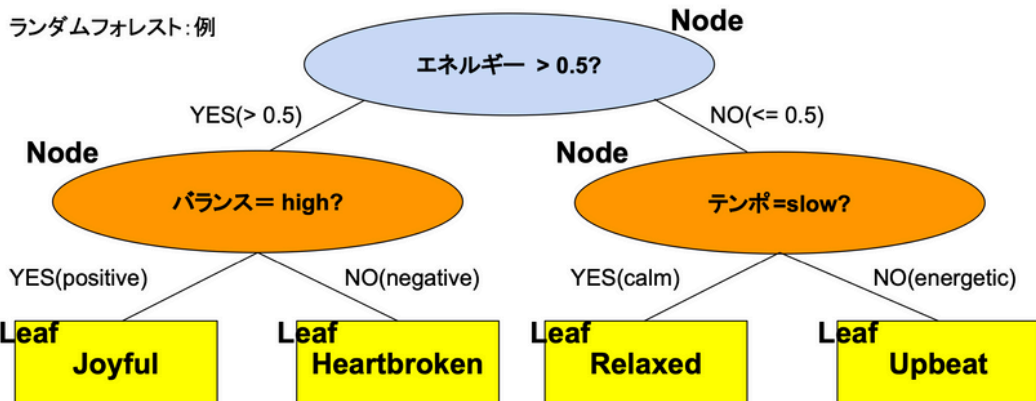
従来の音楽アプリでは過去の行動履歴や選択パターンに基づいてユーザーの好尚を学習し、推薦を行う「協調フィルタリング」が用いられている。この方法はユーザーの評価情報にも基づいて推薦アイテムを決定するためリアルタイムのユーザーの感情を感知し、それに基づいた推薦を行うことは不可能である。本研究ではSpotify APIから得た音楽的特徴量とSpotifyのプレイリストデータを用いて感情別に分類することで、ユーザーの感情に沿った推薦システムを可能にする。現時点では主観的な感情ラベル付けを行い、自身の感性による推薦が可能となったが他のユーザーのユーザーエクスペリエンスの向上の検証はまだできていない。

背景

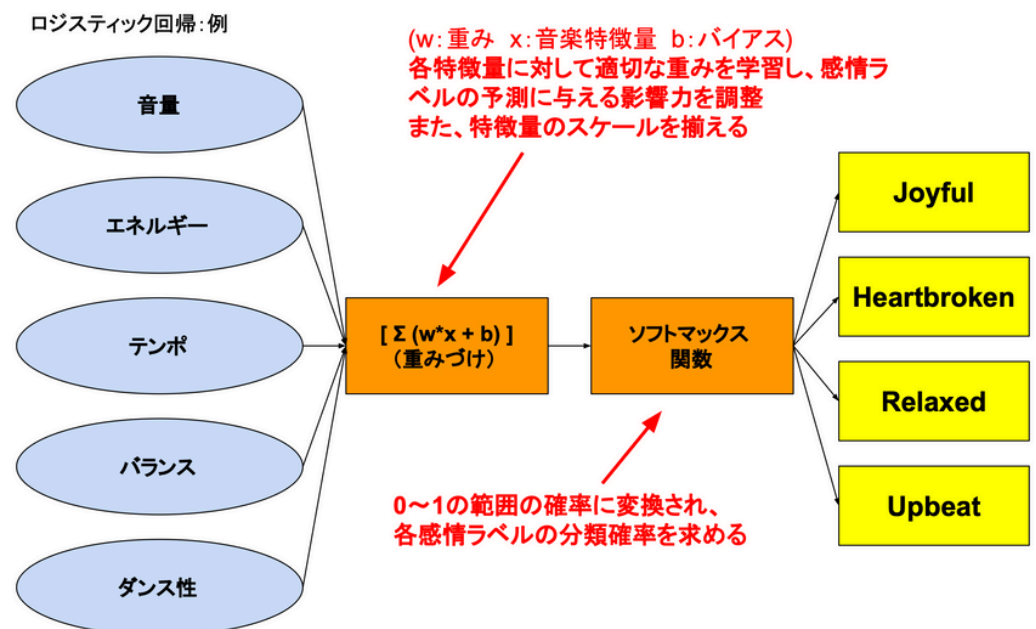
従来の音楽システムでは過去の行動履歴や選択パターンに基づいて推薦を行う「協調フィルタリング」が用いられている。この方法はユーザーの評価情報に基づいて推薦アイテムを決定するため音楽の推薦に適している。[1]しかし、ユーザーの情報によって推薦が行われるためユーザーの感情に基づいた推薦は不可能である。

研究方法

本研究ではSpotify APIを利用して、50曲と100曲を含む2つのプレイリストから楽曲の特徴量データ ('energy', 'tempo', 'danceability', 'loudness', 'valence') を収集した。各楽曲を自分で聴き、それぞれの楽曲に「喜び (Joyful)」「心が傷ついた (Heartbroken)」「落ち着いた (Relaxed)」「元気 (Upbeat)」の感情ラベルを手動で付与してサンプルを作成した。そして、ランダムフォレストとロジスティック回帰の2つのアルゴリズムで特徴量データと感情ラベルを学習させた。



上記の図ではランダムフォレストに使われる決定木の1つを図で表したものであり、本研究では決定木を100本利用する。各決定木が予測した結果をハード投票で最終予測を行った。



学習データは80%をモデル学習のための訓練用、モデルの汎化性能を評価するための20%をテスト用として分割した。各アルゴリズムでモデルを構築し、50曲と100曲のデータセットで実施し、それぞれの精度の比較を行なってモデルの評価を行った。

今後の展望

現段階の研究でサンプル数が少なく学習データの不足が見られたため、楽曲数を増やすと同時に感情ラベルの拡充を行い、感情ラベルの定義を明確化することでモデルの精度を高める。また、ユーザーエクスペリエンスの向上に向けて客観的な意見を得る必要があり、実際にモデルを使用してもらい、評価を得る。ユーザー一人ひとりが音楽を聴いて評価することでユーザー固有のプロファイルが作成され、聞けば聞くほどサンプル数が増えるのでユーザーの感覚に合った識別ができるようになる。

本研究の目的

音楽アプリのユーザーの心理的ニーズや感情状態を把握し、その情報を活用して**ユーザーの感情に基づいた、最適な音楽を提供するための推薦システムを構築**することを目標として様々なアプローチを試みた。

結果分析

ランダムフォレスト: 小規模データで高精度

<(50曲) 精度: 70.00%>

ラベル	precision	recall	F-1 score	support
relaxed	0.00	0.00	0.00	1
upbeat	0.80	0.67	0.73	6
heartbroken	1.00	1.00	1.00	2
joyful	0.33	1.00	0.50	1

<(100曲) 精度: 40.00%>

ラベル	precision	recall	F-1 score	support
relaxed	0.50	0.25	0.33	4
upbeat	0.36	0.67	0.47	6
heartbroken	1.00	0.25	0.40	4
joyful	0.33	0.33	0.33	6

ロジスティック回帰: 大規模データで高精度

<(50曲) 精度: 50.00%>

ラベル	precision	recall	F-1 score	support
relaxed	0.00	0.00	0.00	1
upbeat	0.75	0.50	0.60	6
heartbroken	0.50	0.50	0.50	2
joyful	0.25	1.00	0.40	1

<(100曲) 精度: 45.00%>

ラベル	precision	recall	F-1 score	support
relaxed	1.00	0.25	0.40	4
upbeat	0.35	1.00	0.52	6
heartbroken	0.00	0.00	0.00	4
joyful	0.50	0.17	0.25	6

50曲のプレイリストデータを用いた場合relaxedの精度、再現率が共に0.00%であり、100曲のプレイリストデータでもロジスティック回帰の結果においてheartbrokenが精度、再現率共に0.00%であるためそれぞれのクラスのサンプル数が少なく、学習するのにデータが不十分である。また、ランダムフォレストに比べ、ロジスティック回帰の精度が落ちた理由として、ロジスティック回帰で利用される線形分離は予測結果で感情ラベルが混ざり合っている複雑な場合曲線や複雑な形状になるため正確に分類することが難しい可能性がある。そのため今後は本研究に適している非曲線 (ランダムフォレストなど) を用いたアルゴリズムでの方法を利用する。

参考文献

- 野崎 海斗, 柴田 祐樹, 高間 康史(2022). 「特定アイテムへのフィードバックを重視した協調フィルタリング手法の提案」. 『第38回ファジィシステムシンポジウム 講演論文集』. 31-32
- Casper Hansen. 「Implementing logistic regression from scratch in Python」. 『IBMdeveloper』. <https://developer.ibm.com/articles/implementing-logistic-regression-from-scratch-in-python/> 閲覧日2024年10月15日
- Hawaii. 「[機械学習]ランダムフォレストを理解する」. 『Qiita』. <https://qiita.com/Hawaii/items/5831e667723b66b46fba> 閲覧日2024年10月21日