

地域性とリアルタイムの周辺状況に応じたカラオケ楽曲推薦

田中 茂樹^{1,a)} 深澤 佑介¹

受付日 2020年8月6日, 採録日 2021年2月2日

概要: 近年機械学習を用いたレコメンド技術の発展にともない, オンラインショッピングサイトをはじめとした様々なサービスは商品のレコメンド機能を提供し, 購買の活性化などに活用している. しかし, 実店舗では Web サイトと異なりログインなどによる個人特定は難しく, コールドスタート問題によりレコメンドが困難である場合が多い. そこで, 店舗周辺の地域性とリアルタイムの人口分布を用いて客層やニーズを理解することで, 個人情報によらないレコメンド手法を提案する. 本研究では NTT ドコモの提供する国内人口分布統計から地域性とリアルタイム情報を抽出する技術を開発し, カラオケ店舗における 1 曲目のレコメンドへ活用した. その結果, MAP@10 において 12.35%の精度となり, 人気楽曲をレコメンドする手法に対し約 1.5 倍の精度であることから, カラオケドメインにおける有効性を示した. また, 定性評価により地域性はオフィス街, 学生街, 観光地といった地域性を表現していることを確認し, 他の業態におけるレコメンドでも広く活用できる可能性を確認した.

キーワード: レコメンド, コールドスタート問題, 国内人口分布統計

Song Recommendation in Karaoke Shops Considering Locality and Current Surrounding Situation

SHIGEKI TANAKA^{1,a)} YUSUKE FUKAZAWA¹

Received: August 6, 2020, Accepted: February 2, 2021

Abstract: More and more Web services provide recommend functions to maximize their business thanks to the evolution of recommendation methods using machine learning. However, most of the methods have shortcomings in recommendation in real shops as they struggle with cold-start problem. In this paper, we propose a method to recommend not based on user information or history, but locality and real-time population distribution around shops. We extracted the information from the domestic population distribution statistics provided by NTT DOCOMO and used for the first song recommendation in real Karaoke shops. The method achieved 12.35% in MAP@10 which outperformed popularity method by nearly 50%. In addition, our examination illustrates that the features extracted from the statistics possess rich locality information which could be applied in any other real shops.

Keywords: recommendation, cold-start problem, domestic population distribution statistics

1. はじめに

機械学習の進歩により, レコメンド技術は EC サイトや動画配信サイトなどの様々なサービスに導入され, ユーザ体験の向上やビジネスの活性化に貢献している. 一方で, レコメンドで広く活用されている協調フィルタリング手法

などは, 一般に個人の商品購入履歴や Web 閲覧履歴などから趣味・嗜好を理解し, 似たユーザが興味を持つ商品を関連商品として推薦するため, 履歴のないユーザへはレコメンドが困難である. 本問題はコールドスタート問題と呼ばれ, 特に実店舗ではオンラインサービスと異なりログインなどによる事前のユーザ特定が困難であることから, レコメンド導入への非常に大きな障壁となっている. Amazon では 35%の商品がレコメンド経由で購入され, Netflix では 75%の動画がレコメンド経由で視聴されていると報告され

¹ 株式会社 NTT ドコモ
NTT DOCOMO, INC., Chiyoda, Tokyo 100-0014, Japan
^{a)} shigeki.tanaka.da@nttdocomo.com

ている [7] ことからレコメンドの導入効果は大きく、実店舗においてもコールドスタート問題を解消しレコメンドを活用できればビジネス活性化につながると考えられる。

本報告ではカラオケ店舗での 1 曲目に歌唱する楽曲のレコメンドにおいて、国内人口分布統計を用いたコールドスタート問題の解決手法を提案する。カラオケにおいて特に 1 曲目の選曲を難しいと考えるユーザは多く、他人に譲ると回答する人が 3 割以上を占める^{*1}一方で、カラオケへは会社の同僚・上司・部下と行く場合が 24%^{*2}、かつカラオケを敬遠する理由として「歌うことを強要されると感じる」が 1 位で 31% にのぼることから、必ずしも順番を選べないケースがあると考えられる。カラオケでは一般に、グループのメンバーの多数が知っていて、かつ盛り上がる楽曲が好ましいとされるなど、場の空気を読んだ選曲が肝要となるが、1 番手は他者の歌唱傾向を見て合わせるができないため特に選曲の難易度が高い。そのため、歌い慣れたユーザであれば状況に応じた楽曲を 1 曲目として選択可能であっても、慣れていないユーザにとっては億劫であり、前述の 1 曲目の敬遠、ひいてはカラオケを敬遠する理由にもつながりうることから、1 曲目の円滑な選曲を助けるレコメンドの実現はユーザにも運営企業にも大きな価値を提供しうる。

歌唱履歴がなく、ログインなどによるユーザ情報も得られないユーザへのレコメンドは、一般には人気ランキングのレコメンドが考えられる。しかし、性年代などによって好まれる楽曲は異なるため、店舗が学生街にあるか、オフィス街にあるかといった地域性によって選ばれる楽曲は大きく変化する。また、同じ地域であっても時間やイベント、天候など様々な要因で店舗周辺の人口分布は刻々と変化し、楽曲の人気ランキングもリアルタイムに変わりうる。これらの地域性とリアルタイム状況に応じて 1 曲目の推薦を出し分けることができれば、ユーザにとってはスムーズな選曲につながり、運営企業にとってはカラオケ利用の活性化につながるというメリットがある。

地域性やリアルタイム状況を考慮したレコメンド手法について、単純には都道府県や時間別に人気ランキングを生成する手法が考えられるが、分割が不十分である場合や、分割過多により疎になってしまう場合があり、十分ではない。そこで、国内人口分布統計を用いた地域性特徴量の生成手法と、これを用いたレコメンドモデルを提案する。国内人口分布統計からはどこに、どの年代の方が、どれくらいいたかを 500m 四方単位で得られるが、地方であれば町丁目単位よりも細かな粒度であることから、店舗周辺の統計情報は店舗の環境を十分表していると考えられる。店舗周辺の数十日の統計からベクトルを生成し、AutoEncoder

による次元圧縮を行うことで地域性特徴量を生成する。また、レコメンドを実施する直近の国内人口分布統計からリアルタイム状況を取得し、リアルタイム周辺人口特徴量を生成する。これらの特徴量を入力とし、各楽曲の歌唱尤度を推定する深層学習モデルを学習させることで、ユーザの情報や履歴を必要としない、地域性やリアルタイム状況を考慮したレコメンドを実現する。実際にカラオケ店舗の 1 カ月間の歌唱履歴データにおいて提案手法の検証を行い、MAP@10 において、人気楽曲をレコメンドする場合に比べ約 1.5 倍となる 12.35% の精度を確認した。本報告の貢献は次のとおりである。

- (1) 静的な地域性特徴量およびリアルタイムに変化するリアルタイム周辺人口特徴量を入力とし、コールドスタートとなる 1 曲目を推薦する深層学習モデルを提案した。
- (2) 提案手法をカラオケの実データで検証し、MAP@10 で人気ランキングに比べて 1.5 倍の精度を達成した。
- (3) 地域ごとにオフィス街、学生街、観光地といった傾向がみられ、静的な地域性特徴量を考慮することの有効性を示した。
- (4) 同じ地域であっても、時間帯によって学生が多いときや観光客が多いときの傾向がみられ、リアルタイム周辺人口特徴量を考慮することの有効性を示した。

2. 関連研究

2.1 国内人口分布統計活用

NTT ドコモの国内人口分布統計は携帯電話ネットワークの運用で得られるデータを活用して生成される統計情報で、最小 250m 四方別に時間、性別、年代別の人口統計情報が得られる [18]。国内人口分布統計の活用事例は様々であり、清家ら [20] は国内人口分布統計を用いて柏市および横浜市の地域評価を行い、オフィス街・大学などの地理的特性をもつ主要メッシュにおいて共通性や自治体固有の特徴が得られることを確認し、都市計画における活用可能性を示した。NTT ドコモの提供する AI タクシーでは、リアルタイムの人口変動や天気予報をもとに、各エリアの 30 分後のタクシー需要を予想する深層学習を構築し、93%～95% の精度を達成した [21]。

2.2 機械学習を用いたレコメンド手法

レコメンド手法は一般に協調フィルタリングベースとコンテンツベースがあり、特に協調フィルタリングベースのレコメンド手法は履歴のみを用いて高い精度が得られることから、様々なサービスで活用されている [9], [13]。協調フィルタリングにはその中でもアイテムベース、ユーザベースなどいくつかの手法があるが、いずれも、アイテム間やユーザ間でクリック傾向や購買傾向の類似度テーブルを事前に生成し、ユーザの履歴から類似度の高いアイテムをレ

^{*1} <https://research-panel.jp/rpdr/view.php?eid=419375&aid=1754529>

^{*2} https://chosa.nifty.com/hobby/chosa_report_A20150109/

コメンドする。協調フィルタリングは履歴のみを用いるためコンテンツ情報が必要なく精度も高い一方、協調フィルタリング単体では時系列情報やメタ情報を加味することが困難であり、新たなレコメンド手法が近年多数提案されている。Hidasi ら [5] はユーザの一連のクリック履歴をセッションと見なし、セッション内でクリックされたアイテムの順番を深層学習の RNN [14] を用いて学習するレコメンド手法を提案した。Hidasi ら [6] はメタ情報ごとの RNN モデルを構築し、それらのスコアと組み合わせることでメタ情報も考慮する手法を提案した。ただし、これらの手法はいずれも履歴がないユーザにはレコメンドが行えない。

協調フィルタリングに対し、コンテンツベースのレコメンド手法は一般にユーザ属性やアイテムのメタ情報から類似度を計算しレコメンドを行うため履歴への依存度が低く、コールドスタート問題へのアプローチとして様々な手法が提案されている。Vozalis ら [16] は新規ユーザと既存ユーザの類似度を計算する際、ユーザ属性をベクトルとし、k 平均法を用いてユーザをクラスタする手法を提案した。また、ユーザ属性だけでなく、交友関係をグラフ表現し、ユーザの類似度を表す特徴量を抽出する手法も提案されている [10], [17]。これらの手法を用いることで履歴の乏しい新規ユーザを既存ユーザと関連付けてレコメンドが可能である一方、未ログインのようなユーザ属性も不明のユーザに対しては適用が困難である。

ユーザ情報やアイテム情報に加え、時間や天気などのコンテキスト情報を加味するレコメンド手法も提案されており、より動的なレコメンドの提供や、コールドスタート問題の緩和を行う手法が提案されている [2], [4]。Bahramian ら [1] はツアーのレコメンドにおいてユーザの履歴が乏しいことによるコールドスタート問題に対し、ユーザ周辺の POI や路面情報、移動距離などのコンテキスト情報を加味することで精度向上を行った。須藤ら [22] はウェアラブル端末から得られた情報をもとにユーザが移動中であるか、屋内にいるか、といった状況を段階的に推定し、事前に得られているユーザ嗜好と現在の状況を考慮したサービスのレコメンドを提案した。村上ら [19] はユーザの性年代と、モバイル端末から得られた現在地や日時、天気などを用いた飲食店レコメンドを提案し、特に位置情報は性年代よりも精度に寄与する結果となった。

3. 提案手法

3.1 手法の概要

本研究では、ユーザがカラオケ店舗に入店時、1 曲目に選ばれる楽曲を推定するレコメンドモデルを検討する。推定時に活用可能な情報を表 1、国内人口分布統計の例を表 2 に示す。日本全国の住所は 500 m 四方ごとにメッシュ番号が付与されており、店舗住所にもメッシュ番号があることから、表 1、表 2 はメッシュによって結合可能である。レ

表 1 レコメンド時に得られる情報

Table 1 Information available for recommendation.

属性	概要
端末 ID	操作端末に与えられたユニーク ID
店舗住所	操作端末が設置されている店舗住所
メッシュ	店舗住所の 4 次メッシュコード (500 m 粒度)
日時	レコメンド時の日時

表 2 国内人口分布統計から取得可能なデータ例。本研究では 500 m 四方ごとに 10 分おきの年齢別、性別の人口統計データを用いる

Table 2 Example of mobile spatial statistics. The data illustrates the demographics in each 500 m mesh every 10 minutes.

メッシュ	日時	15–20 才 男性	...	70–75 才 女性
XX	07/01 08:00	30 人		1 人
XX	07/01 08:10	26 人		0 人
XX	07/01 08:20	22 人		2 人
YY	07/01 08:00	10 人		12 人
YY	07/01 08:10	13 人		11 人
YY	07/01 08:20	21 人		13 人

コメンドモデルはこれらの情報を用いて、レコメンド候補の楽曲から N 個の楽曲をレコメンドする。

3.2 提案モデル

本研究の提案手法は主に以下の工程で実現される。

- (1) 学習期間の国内人口分布統計を用いた店舗周辺の地域性特徴量生成
- (2) レコメンド日時付近の国内人口分布統計を用いた店舗周辺のリアルタイム周辺人口特徴量生成
- (3) 地域性特徴量とリアルタイム周辺人口特徴量を入力とし、各楽曲が歌唱される尤度を推定する深層学習モデルの学習と推論

3.2.1 地域性特徴量生成

店舗メッシュの地域性特徴量生成は、学習期間における国内人口分布統計をメッシュごとにメッシュベクトルとして整形し、AutoEncoder を用いることで実現した [11]。表 2 で示したとおり、空間統計からはある時間帯における 15–20 才～70–75 才までの 13 の年代を性別で分け、計 26 次元の特徴量が得られる。これを 10 分ごとに得られることから、1 時間あたり 156 次元、1 日あたりでは 3,744 次元の特徴量となり、本研究では学習期間を 24 日としているため、計 89,856 次元の特徴量となる。この特徴量は通年の情報ではないにしても 3 週間強の統計情報であることから、直近の状況を表しており、将来数日から数週間のレコメンドなどには十分有益と考えられる。この特徴量が類似しているメッシュは訪問者の傾向が類似していることを意味し、たとえばオフィス街、学生街、観光地といった地域性が類似していると推察される。年代別に人口の偏りがあ

り、絶対数の少ない年代が評価されない場合を防ぐため、統計値を性年代ごとに最大値を 1、最小値を 0 とする正規化を行いメッシュベクトルとした、正規化の式を式 (1) へ示す。

$$X_{norm} = \frac{X - X_{min}}{X_{max} - X_{min}} \quad (1)$$

X は正規化前のデータ、 X_{min} および X_{max} は X における最小値と最大値、 X_{norm} は正規化後のデータを示す。提案手法では次元数の発散を防ぐため、約 9 万次元のメッシュベクトルをそのまま用いるのではなく、AutoEncoder を用いて次元圧縮を行い地域性特徴量とした。AutoEncoder モデルのアーキテクチャを図 1 に示す。AutoEncoder はメッシュベクトルを入力し、Dense 層を経由して最小 32 次元まで圧縮したのち、さらに Dense 層を経由してメッシュベクトルと同じ次元数まで拡張して出力する。各 Dense 層は入力値に重みをかけ、バイアスを加算することで N 次元の値を出力し、活性化関数として最終層のみ Sigmoid 関数、その他の層は ReLU 関数 [3] を用いて非線形へ変換する。出力されたベクトルは入力ベクトルと比較され、交差エントロピーを用いて損失が算出される。交差エントロピーの式を式 (2) へ示す。

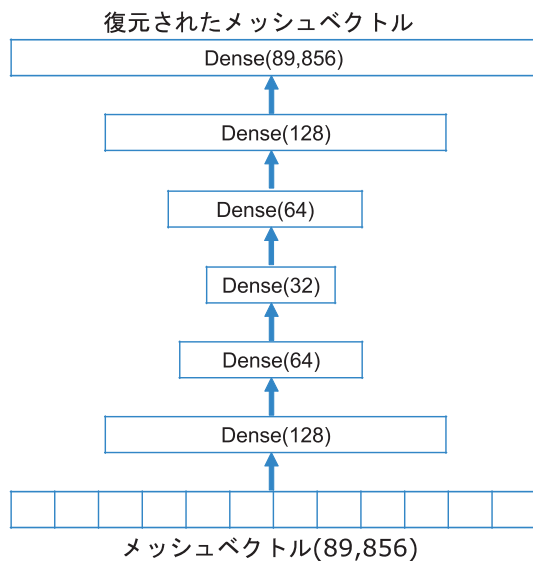


図 1 AutoEncoder モデルの概要。各層の括弧内は次元数を表す。店舗付近の 24 日間の国内人口分布統計をメッシュベクトルとして入力し、ボトルネックを経由して 32 次元まで圧縮したのち、元々の入力と同次元のベクトルを復元し出力する。モデルは入力ベクトルと復元されたベクトルの誤差が最小になるよう学習する過程で、ボトルネック部でもとの入力情報により低次元で損失の少ない表現方法を獲得する

Fig. 1 Graphical representation of the AutoEncoder model. The number of dimensions of each layer is shown in parentheses. The model takes mobile spatial statistics of a location as input, and learns better representation of the information in smaller dimensions through training.

$$H(p, q) = - \sum_x p(x) \log q(x) \quad (2)$$

モデルは損失を最小化するため、Adam 最適化 [8] を用いた誤差逆伝播法 [12] によって各層の重みを更新する。学習された AutoEncoder モデルはボトルネック部で情報の損失がなるべく少ない次元圧縮方法を獲得できているため、本研究ではモデルで 32 次元まで圧縮されたベクトルを地域性特徴量として抽出した。

3.2.2 リアルタイム周辺人口特徴量生成

店舗メッシュのリアルタイム周辺人口特徴量生成は、レコメンドを行う際、直近 30 分間の店舗周辺の国内人口分布統計を用いることで実現した。前節で示したとおり、国内人口分布統計から、10 分ごとに 26 次元の性年代別人口情報が得られるので、30 分間で計 78 次元の情報が得られる。この特徴量は店舗周辺に現在どういった性年代の方が多いかを示しており、入店したユーザの性年代もある程度相関があると考えられる。前節と同様に、性年代ごとに最大値を 1、最小値を 0 とする正規化を行いリアルタイム周辺人口特徴量とした。

3.2.3 レコメンドモデル

提案モデルのアーキテクチャを図 2 へ示す。本モデルは AutoEncoder によって店舗へ事前に付与された地域性特徴量と、店舗周辺のリアルタイム情報から得られたリアルタイム周辺人口特徴量を入力する。これらの特徴量は結合され、Dense 層と Dropout 層を 5 層ずつ経由する。Dense 層は AutoEncoder で用いられたように入力を変換して N 次

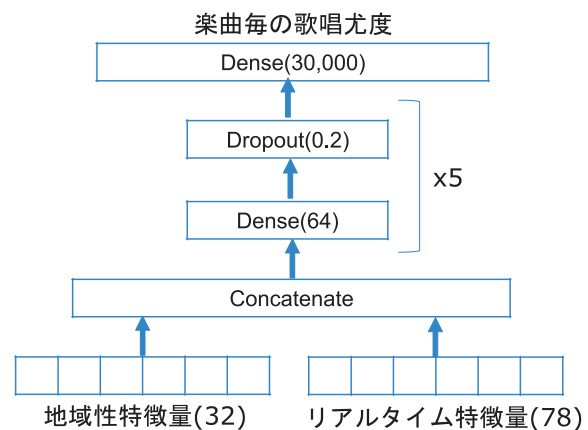


図 2 レコメンドモデルの概要。Dense 層の括弧内は次元数を表し、Dropout 層の括弧内は dropout 率を表す。店舗周辺の地域性特徴量と、レコメンド時のリアルタイム周辺人口特徴量を入力して結合し、Dense 層と Dropout 層を交互に 5 層経由したのち、レコメンド対象曲それぞれが歌唱される尤度を出力する

Fig. 2 Graphical representation of the proposed model. The number of dimensions and dropout ratio of each layer is shown in parentheses. The model inputs local information and real time information around karaoke shop acquired from mobile spatial statistics, and outputs the likelihood for each song to be played.

元の値を出力し、活性化関数を用いて非線形変換を行う。Dropout 層は任意の確率で入力値をマスクすることでモデル内に擬似的にサブモデルを形成し、アンサンブルによる汎化性能向上と類似した効果が上げられる [15]。複数層に積まれたこれらの層を経由し、Dense 層と Softmax 活性化関数によって楽曲ごとの歌唱尤度を出力する。層数や各層の次元数はハイパーパラメータであり、データやドメインの性質によって最適な値は異なるが、本研究では様々な組合せで実験を行った結果、浅く太い（層数：少，次元数：多）層構成よりも深く狭い（層数：多，次元数：少）層構成の方が過学習が起りづらく汎化性能が高かったことから、提案の構成とした。

学習時、モデルは学習データを入力して推定結果を出力し、正解データと比較して式 (2) の交差エントロピーを損失として算出する。モデルの重みは誤差逆伝播法によって損失を最小化するために更新され、最適化には Adam 最適化を用いる。

4. 性能評価

本章では、オフラインでの提案手法の精度評価結果と得られた知見について述べる。

4.1 データセット

オフライン評価のため、第一興商が提供する DAM カラオケ端末を利用してカラオケ店舗で歌唱された 2018 年 7 月の履歴を用いてデータセットを生成した。歌唱履歴の例を表 3 に示す。

歌唱履歴にはユーザの入店、退店の明示的な情報は含まれないため、同端末で 15 分以上歌唱されていない場合にユーザの入れ替わりが発生したと見なし、セッションを分割した。今回はコールドスタート問題に対する精度を評価するため、各セッションの 1 曲目に歌唱された楽曲を端末 ID や日時から推定する問題とした。また、カラオケでは一部の楽曲に人気が偏る傾向があるため、今回は最も歌唱された 3 万曲のみを対象とし、それ以外の楽曲の履歴は除外した。精度評価にはデータセットのうち最後の 1 週間を用い、それ以前を学習データとして用いた。

4.2 評価モデルとパラメータチューニング

本研究では以下のベースラインと提案手法の精度を評価

し、比較を行った。

- (1) 全国人気ランキング：全国で歌唱された回数による人気ランキングを作成し、TOP N をレコメンドするモデル
- (2) 都道府県別人気ランキング：都道府県別の人気ランキングを作成し、TOP N をレコメンドするモデル
- (3) 時間別人気ランキング：時間別（0～23 時）の人気ランキングを作成し、TOP N をレコメンドするモデル
- (4) 都道府県 × 時間別人気ランキング：都道府県別、かつ時間別の 1,128 (47 都道府県 × 24 時間) グループごとで人気ランキングを作成し、TOP N をレコメンドするモデル
- (5) 店舗別人気ランキング：各店舗で歌唱された回数による人気ランキングを作成し、TOP N をレコメンドするモデル
- (6) 提案手法：国内人口分布統計から得られた地域性特徴量とリアルタイム周辺人口特徴量をもとに各楽曲の歌唱尤度を推定し、尤度の高い TOP N をレコメンドする深層学習モデル
- (7) 提案手法（地域特徴量のみ）：国内人口分布統計から得られた地域性特徴量をもとに各楽曲の歌唱尤度を推定し、尤度の高い TOP N をレコメンドする深層学習モデル
- (8) 提案手法（リアルタイム特徴量のみ）：国内人口分布統計から得られたリアルタイム周辺人口特徴量をもとに各楽曲の歌唱尤度を推定し、尤度の高い TOP N をレコメンドする深層学習モデル

4.3 定量結果

本実験では、各モデルの推定結果が評価データの各セッションにおいて最初に歌唱される楽曲をどれだけ当てられているか、Mean Average Precision (MAP) @10 を用いて評価した。数式を以下に示す。

$$MAP@k = \frac{\sum_{q=1}^Q AP@k(q)}{Q} \quad (3)$$

$$AP@k(q) = \frac{\sum_{k=1}^N \frac{Precision@k \cdot y_k}{\sum_{i=1}^k y_i}}{N} \quad (4)$$

$$y_k = \begin{cases} 1 & (k\text{th item is correct}) \\ 0 & (\text{otherwise}) \end{cases} \quad (5)$$

Q は全セッションの 1 曲目の歌唱履歴、 q は 1 セッションの 1 曲目の歌唱履歴、 k は上位何番目までをレコメンドするかを表し、今回は 10 とした。評価結果を表 4 に示す。

この結果において、ベースラインどうしの精度を比較すると、都道府県別人気ランキングが全国人気ランキングの精度を上回っていることが分かる。また、時間別人気ランキングもわずかではあるが全国人気ランキングを上回る結果となった。この結果から、全国一律のレコメンド結果を表示するよりも、レコメンド対象となるユーザの層によってレコメンドを変化させることが有効であると推察され

表 3 歌唱履歴例

Table 3 Example of singing history.

端末 ID	歌唱楽曲 ID	歌唱開始日時	歌唱終了日時
A001	S001	07/01 18:02:13	07/01 18:06:40
A001	S002	07/01 18:07:03	07/01 18:12:10
A001	S003	07/01 19:40:50	07/01 19:44:40
A001	S004	07/01 19:46:23	07/01 19:51:05

表 4 各モデルの精度評価結果. 対全国比は全国人気ランキングに対する比で, 最も高い精度は太字で示す

Table 4 Results of Karaoke recommendation task. The best result is in bold.

Model	MAP@10	対全国比
全国人気ランキング	0.0828	—
都道府県別人気ランキング	0.0843	+1.81%
時間別人気ランキング	0.0834	+0.72%
都道府県×時間別人気ランキング	0.0796	−3.86%
店舗別人気ランキング	0.0309	−62.68%
提案手法 (地域特徴量のみ)	0.1127	+36.11%
提案手法 (リアルタイム特徴量のみ)	0.1034	+24.88%
提案手法 (両特徴量併用)	0.1235	+49.15%

る. 一方で, 都道府県×時間別人気ランキングにおいてはむしろ全国人気ランキングを下回る結果となった. 理由として, 都道府県別, 時間別ともに有意なグループであっても, これらを掛け合わせることでグループ数が膨大になり, 各グループにおける履歴数が減少してしまったためと考えられる. 同様の傾向は店舗別人気ランキングでもみられ, 適切なグループ分けが重要であることが分かる.

提案手法はいずれもすべてのベースラインを大きく上回る精度を達成し, 特に地域特徴量およびリアルタイム特徴量を併用した場合に, 従来手法である全国人気ランキングに対しては約 1.5 倍の精度となった. 地域特徴量のみ用いた場合は約 1.36 倍, リアルタイム特徴量のみの場合は約 1.25 倍の精度となったことから, それぞれの特徴量がレコメンドに寄与することが分かり, 組み合わせた場合に最高精度を達成したことから, それぞれが異なる特徴を表現していることが分かる. この実験は候補である 3 万曲から 10 曲を選択し正解を見つける問題であり, 従来手法の 8.28%はけっして低い精度ではないと考えられる. 提案手法がレコメンド対象ユーザの明確な情報をいっさい用いず, 従来手法を大きく上回る 12.35%の精度を達成した意義は大きく, 国内人口分布統計は実店舗でのレコメンドにおけるコールドスタート問題の解決に十分寄与できることを示している. 国内人口分布統計から得られた地域性特徴量とリアルタイム周辺人口特徴量については定性評価で更に検証する.

4.4 定性評価

前節の定量評価では, 国内人口分布統計から得られた店舗周辺の地域性特徴量とリアルタイム周辺人口特徴量は, レコメンドにおけるコールドスタート問題の解決に寄与することが分かった. 本節では, 特に深層学習を用いて得られた地域性特徴量が表現している情報の解釈と, 2 つの特徴量の変化にともなうレコメンド内容について評価する.

4.4.1 地域性特徴量の解釈

AutoEncoder による次元削減で得られた 32 次元の地域性特徴量がどのような地域性を表すか評価するため, いく

つかの住所と, 地域性特徴量が類似した住所を比較した. いくつかの店舗住所と, その店舗メッシュの国内人口分布統計の特徴, および地域性特徴量が類似した住所を表 5 に示す. 1 例目の東京都中央区八重洲は東京駅周辺であり, 国内人口分布統計の特徴からも平日にビジネスで利用されていることが示されている. 地域性特徴量は赤坂や新横浜, 広島市中区などと類似しており, オフィス街としての特徴が地域性特徴量で表現されていることが分かる. 2 例目の神奈川県鎌倉市は多くの文化遺産を有し, 日本屈指の観光地として知られており, 平日・休日問わず特に 75 歳前後の女性に人気であることが国内人口分布統計から分かる. 地域性特徴量は同じく歴史的建造物で有名な地域と類似している. 3 例目の東京都渋谷区は若者に人気であり, 特に週末 15 歳前後の男女が多く訪れることが国内人口分布統計に示されている. 新宿や歌舞伎町, 西池袋といった地域と地域性特徴量が類似しており, 若者に好まれる地域の傾向をとらえられていることが分かる. 4 例目の東京都新宿区高田馬場には多くの大学が存在し, 屈指の学生街として知られており, 15 歳前後の男性・女性が平日に多く滞在していることが国内人口分布統計からも読み取れる. 地域性特徴量は同じく大学が存在する中野などと類似しており, 学生街としての地域性特徴量が得られていることが分かる. これらの結果から, 国内人口分布統計から得られた地域性特徴量は一般にいわれるオフィス街, 学生街, 観光地といった情報を表現していることが分かり, 地域性を考慮したレコメンドを行う際にはカラオケに限らず, 様々な業態において非常に有用な情報であると考えられる. たとえば飲食チェーンでは店舗ごとに地域性に応じて, 学生街ではボリュームのあるメニュー, オフィス街では調理の早いメニュー, 観光地では写真写りの良いメニューの写真を店頭に掲示することで, より集客につながると推察される. また, 服飾店においても地域性を考慮した商品をより多く展示するなど, 様々な業態において汎用的に活用可能な特徴量であると考えられる.

4.4.2 2 つの特徴量とレコメンド内容の変化

2 つ目の定性評価として, 地域性特徴量とリアルタイム周辺人口特徴量の組合せによって, どのようにレコメンド楽曲が変化するかを評価する. 2 つの店舗において, リアルタイム周辺人口特徴量が変化した場合のレコメンド例を表 6 に示す. 1 例目の八重洲は地域性特徴量の定性評価でも言及されたとおり, 平日にビジネスで利用されることが地域性特徴量に示されている. 平日 21 時台でもその傾向がリアルタイム周辺人口特徴量に示されており, レコメンド例としては TOKIO の「LOVE YOU ONLY」や B'z の「ultra soul」といった, 40 歳前後の男性に好まれる楽曲が多い. 一方で, 週末には人口の傾向が変わり, 25 歳前後の男性・女性が多くみられる. レコメンド例もそれに応じて変化し, 人気ランキング上位である米津玄師の「Lemon」

表 5 店舗住所の例と、店舗メッシュの国内人口分布統計の特徴、および地域性特徴量が類似した住所の例。地域性特徴量が類似した住所は住所間の地域性特徴量のコサイン類似度が高い住所例で、括弧内はコサイン類似度を示す

Table 5 Example of shop address and features of mobile spatial statistics, and the addresses that location information are similar. The cosine similarity of location information are shown in parentheses.

店舗住所	店舗メッシュの国内人口分布統計の特徴	地域性特徴量が類似した住所（類似度）
東京都 中央区八重洲 1 丁目	平日 14 時台に 45 歳前後の男性が多い 平日 16 時台に 40 歳前後の男性が多い 金曜日 15 時台に 50 歳前後の男性が多い	東京都 港区赤坂 3 丁目 (0.998) 神奈川県 横浜市港北区新横浜 3 丁目 (0.990) 広島県 広島市中区立町 3 丁目 (0.968)
神奈川県 鎌倉市小町 1 丁目	平日 11 時～15 時に 75 歳前後の女性が多い 金曜日 11 時台に 75 歳前後の女性が多い 週末 12 時台に 75 歳前後の男性が多い	福岡県 久留米市東町 (0.999) 奈良県 奈良市東向町 (0.996) 愛媛県 松山市大街道 (0.986)
東京都 渋谷区宇田川町	週末 20 時台に 15 歳前後の女性が多い 週末 19 時台に 15 歳前後の男性が多い 金曜日 20 時台に 35 歳前後の女性が多い	東京都 新宿区歌舞伎町 1 丁目 (0.999) 東京都 新宿区新宿 3 丁目 (0.999) 東京都 豊島区西池袋 1 丁目 (0.999)
東京都 新宿区高田馬場 1 丁目	金曜日 11 時台に 15 歳前後の女性が多い 平日 10 時～12 時に 15 歳前後の男性・女性が多い 平日 14 時台に 15 歳前後の女性が多い	東京都 中野区中野 5 丁目 (0.998) 千葉県 船橋市本町 4 丁目 (0.996) 広島県 広島市中区三川町 (0.986)

表 6 地域性特徴量とリアルタイム周辺人口特徴量の組合せにおけるレコメンド楽曲の例

Table 6 Example of shop address and real time information, and top 5 recommendation.

店舗住所	時間帯	リアルタイム周辺人口特徴量の傾向	レコメンド例 (TOP5)
東京都 中央区八重洲 1 丁目	平日 21 時台	40 歳前後の男性が多い 30 歳前後の男性が多い 30 歳前後の女性が多い 35 歳前後の女性が多い 65 歳前後の女性が少ない	LOVE YOU ONLY / TOKIO ultra soul / B'z 小さな恋のうた / MONGOL800 恋 / 星野源 さくらんぼ / 大塚愛
	週末 19 時台	20 歳前後の女性が多い 25 歳前後の女性が多い 25 歳前後の男性が多い 35 歳前後の女性が多い 65 歳前後の女性が少ない	Lemon / 米津玄師 奏 (かなで) / スキマスイッチ 小さな恋のうた / MONGOL800 さよならエレジー / 菅田将暉 糸 / 中島みゆき
神奈川県 鎌倉市小町 1 丁目	平日 14 時台	65 歳前後の女性が多い 70 歳前後の女性が多い 70 歳前後の男性が多い 40 歳前後の男性が少ない 35 歳前後の男性が少ない	Lemon / 米津玄師 北の旅人 / 石原裕次郎 糸 / 中島みゆき 母なる海よ / 大月みやこ 美唄の風 / 真木柚布子
	週末 15 時台	20 歳前後の女性が多い 40 歳前後の女性が多い 70 歳前後の女性が多い 45 歳前後の男性が多い 20 歳前後の男性が多い	Lemon / 米津玄師 信濃の国—長野県歌— / 長野県歌 糸 / 中島みゆき さよならエレジー / 菅田将暉 北の旅人 / 石原裕次郎

や、スキマスイッチの「奏 (かなで)」などが表示される。2 例目の神奈川県鎌倉市は地域性特徴量の定性評価において、歴史的建造物で有名な地域としての特徴があることが分かっており、リアルタイム周辺人口特徴量では、平日 14 時頃には 70 歳前後の男性・女性が多く訪れることが示されている。この際のレコメンド例としては人気の高い「Lemon」に加え、石原裕次郎の「北の旅人」や、中島みゆきの「糸」などの楽曲があがる。対して、週末 15 時台

には対照的に観光地として幅広い年代の方が訪れ、レコメンド楽曲でも前述の楽曲に加え、「長野県歌」や菅田将暉の「さよならエレジー」など、幅広い楽曲が選ばれる。この定性評価から、同じ店舗であっても地域性特徴量とリアルタイム周辺人口特徴量の組合せによって予想されるレコメンド対象ユーザの人物像は変化し、それに合わせたレコメンドが生成されていることが分かる。

5. まとめと今後の課題

本報告では、実店舗でのレコメンドにおいて特に障害となりうるコールドスタート問題に対し、国内人口分布統計を活用した解決手法を提案した。提案手法では、国内人口分布統計から得られる店舗周辺の情報を AutoEncoder を用いて次元圧縮し、地域性特徴量を抽出した。また、レコメンドを行う直近の国内人口分布統計からリアルタイム周辺人口特徴量を生成し、これらの特徴量を用いて入店直後に歌唱される楽曲を推定する深層学習モデルを構築し、既存手法である全国一律の人気ランキングに対し約 1.5 倍の精度向上を達成した。また、定性評価を通じ、地域性特徴量にはオフィス街、学生街、観光地といった情報が表現されていることや、深層学習レコメンドモデルが地域性特徴量とリアルタイム周辺人口特徴量の組合せによって柔軟にレコメンド結果を変更し、状況に合ったレコメンドを行っていることが分かった。本手法はその状況において選択されやすい楽曲の推定をもとにレコメンドしており、本システムがない場合でも確率的に選ばれやすい楽曲を表示することとなるが、はじめに示したとおり不慣れなユーザの補助に大きな効果があると考えられ、慣れたユーザにとっても検索簡略化による UX 向上が期待される。本報告の定量・定性評価から、国内人口分布統計はカラオケ店舗におけるレコメンドを行う際のコールドスタート問題解消に寄与することが示され、また、その汎用性から業態が異なる店舗であっても活用可能であると考えられる。

本報告ではオフラインでの検証結果の報告となったが、今後実際に実店舗でのレコメンドに活用し、有効性をさらに検証していきたい。

参考文献

- [1] Bahramian, Z., Ali Abbaspour, R. and Claramunt, C.: A cold start context-aware recommender system for tour planning using artificial neural network and case based reasoning, *Mobile Information Systems 2017* (2017).
- [2] Dey, A.K.: Understanding and using context, *Personal and Ubiquitous Computing*, Vol.5, No.1, pp.4–7 (2001).
- [3] Glorot, X., Bordes, A. and Bengio, Y.: Deep sparse rectifier neural networks, *Proc. 14th International Conference on Artificial Intelligence and Statistics*, pp.315–323 (2011).
- [4] Haruna, K., Akmar Ismail, M., Suhendroyono, S., Damiasih, D., Pierewan, A.C., Chiroma, H. and Herawan, T.: Context-aware recommender system: A review of recent developmental process and future research direction, *Applied Sciences*, Vol.7, No.12, 1211 (2017).
- [5] Hidasi, B., Karatzoglou, A., Baltrunas, L. and Tikk, D.: Session-based recommendations with recurrent neural networks, arXiv preprint arXiv:1511.06939 (2015).
- [6] Hidasi, B., Quadrana, M., Karatzoglou, A. and Tikk, D.: Parallel recurrent neural network architectures for feature-rich session-based recommendations, *Proc. 10th ACM Conference on Recommender Systems*, pp.241–248 (2016).
- [7] Ian MacKenzie, C.M. and Noble, S.: How retailers can keep up with consumers, available from (<https://www.mckinsey.com/industries/retail/our-insights/how-retailers-can-keep-up-with-consumers#>).
- [8] Kingma, D.P. and Ba, J.: Adam: A method for stochastic optimization, arXiv preprint arXiv:1412.6980 (2014).
- [9] Linden, G., Smith, B. and York, J.: Amazon.com recommendations: Item-to-item collaborative filtering, *IEEE Internet Computing*, Vol.7, No.1, pp.76–80 (2003).
- [10] Meng, C., Cheng, Y., Jiechao, C. and Peng, Y.: A method to solve cold-start problem in recommendation system based on social network sub-community and ontology decision model, *3rd International Conference on Multimedia Technology (ICMT-13)*, Atlantis Press (2013).
- [11] Rumelhart, D.E., Hinton, G.E. and Williams, R.J.: Learning internal representations by error propagation, Technical Report, California University (1985).
- [12] Rumelhart, D.E., Hinton, G.E. and Williams, R.J.: Learning internal representations by error propagation, Technical Report, California University (1985).
- [13] Sarwar, B., Karypis, G., Konstan, J. and Riedl, J.: Item-based collaborative filtering recommendation algorithms, *Proc. 10th International Conference on World Wide Web*, pp.285–295 (2001).
- [14] Simon Haykin, K.M.: *Neural networks: A comprehensive foundation*, Vol.13, Cambridge University Press (1999).
- [15] Srivastava, N., Hinton, G., Krizhevsky, A., Sutskever, I. and Salakhutdinov, R.: Dropout: A simple way to prevent neural networks from overfitting, *The Journal of Machine Learning Research*, Vol.15, No.1, pp.1929–1958 (2014).
- [16] Vozalis, M. and Margaritis, K.G.: Collaborative filtering enhanced by demographic correlation, *AIAI Symposium on Professional Practice in AI, 18th World Computer Congress* (2004).
- [17] Zhang, Z.-K., Liu, C., Zhang, Y.-C. and Zhou, T.: Solving the cold-start problem in recommender systems with social tags, *EPL (Europhysics Letters)*, Vol.92, No.2, 28002 (2010).
- [18] 小田原 亨, 永田智大: 社会動態推定技術: モバイル空間統計の推計技術と応用, 電子情報通信学会誌, Vol.97, No.9, pp.806–811 (2014).
- [19] 村上 千央, 白木 孝, 桐越 孝之, 西村 健士, 河又 恒久, 白井 隆行, 小林 功ほか: 「マルチモードレコメンド基盤」におけるコンテキストウェア推薦の有効性評価, 全国大会講演論文集, データベースとメディア, pp.577–578 (2010).
- [20] 清家 剛, 三牧 浩也, 森田 祥子: 柏市および横浜市を対象としたモバイル空間統計による地域評価モデルに関する研究, 日本建築学会技術報告集, Vol.21, No.48, pp.821–826 (2015).
- [21] 石黒 慎, 菊池 悠, 深澤 佑介: リアルタイム人口を用いた stacked denoising autoencoders によるタクシー将来需要予測, 情報処理学会論文誌, Vol.60, No.1, pp.118–128 (2019).
- [22] 須藤 健太ほか: ウェアラブルセンサーを用いた利用者の状況の階層的なモデル化と情報提供サービスへの応用, 中央大学大学院研究年報理工学研究科編, Vol.39 (2009).



田中 茂樹 (正会員)

1985年生. 2016年 Birkbeck, University of London, B.Sc. Computing 学部卒業. 2017年 NTT ドコモ入社.



深澤 佑介 (正会員)

2004年東京大学大学院工学系研究科修士課程修了. 同年株式会社NTT ドコモ入社. 2011年東京大学大学院工学系研究科博士後期課程修了. 2011年より東京大学人工物工学研究センターにて協力研究員, 2016年より客員研究員を兼任. 2019年より早稲田大学イノベーション研究所招聘研究員を兼任, 現在に至る. Web マイニング, パーソナライゼーション, 確率モデルに関する研究開発を行っている. IEEE, 人工知能学会各会員. 博士 (工学).