

自動抽出された顧客欲求で重み付けられた飲食店の順序付け Ranking Restaurants Weighted by Customer Desires Automatically Extracted

須藤 晃平†

佐々木 優太‡

島川 博光†

Kouhei Sudou

Yuta Sasaki

Hiromitsu Shimakawa

1. はじめに

外食する際、顧客はさまざまな基準で飲食店を選ぶ。選ぶ基準としては、メニューや雰囲気、価格、場面などがある。しかし、顧客は必ずしも顕在的な基準で飲食店を選んでくるとは限らない。顧客は自身で表現できない潜在的な基準で飲食店を選んでくることがある。そのため、従来の飲食店の検索システムや推薦システムは顧客の潜在的な基準で飲食店を提示できていない。例えば、従来の飲食店の検索システムはジャンルやキーワードを指定して飲食店を検索しなければならない。しかし、飲食店を選ぶのが潜在的な基準である場合、顧客は自身で表現することができないので、検索結果に目標の飲食店を見つけることができない。従来の推薦システムでは、顧客の来店履歴の類似性などで飲食店を推薦している [6]。しかし、来店履歴は飲食店の特徴の一部しか表していない。飲食店の特徴の一部が顧客の潜在的な基準を満たしているか分からない。

本論文では顧客を満足させることができる飲食店は顧客の潜在的な嗜好を満たしている仮定する。そこで本論文では飲食店のテキストデータをトピック分析することで顧客の潜在的な基準を見つけ、潜在的な基準の類似度を基に飲食店を推薦して提示する手法を提案する。

2. 現在の情報推薦システム

2.1 既存の推薦システム

現在商用で利用されている多くの推薦システムは構造化データを用いて商品やサービスを推薦している。構造化データとは、顧客データや購買履歴データに表形式で扱えるものを指す。例えば、宅配サービスのアマゾンの推薦システム [2] はアイテムベースの協調フィルタリングを用いている [4]。協調フィルタリングは購買データを元に顧客と顧客嗜好の類似性を定義し、自分に似た人が購入して自分が購入していない商品を推薦する。ただ、協調フィルタリングは顧客データが少ないときには使用することができない。また、購買データがない商品は推薦することができないという課題がある。協調フィルタリングは商品の機能の関連性や商品閲覧の共起性を考慮していない点でも、ユーザの満足度を満たせない課題をもつ。

2.2 テキストからの嗜好抽出

金山 [5] は自然言語処理技術を用いてユーザ作成コンテンツからの評価や嗜好の抽出の必要性を述べている。テキストなどの非構造化データを用いた推薦はユーザ作成コンテンツから顧客の嗜好を抽出し、顧客の嗜好の類似性を基に推薦する。しかし、ユーザ作成コンテンツから顧客の嗜好を抽出する場合、コンテンツを作成する顧客数に影響される [7]。そのため、店舗の発信情報から顧客の嗜好を抽出する必要がある。

3. 顧客欲求での飲食店の順序付け

3.1 顧客欲求に適合した推薦モデル

本論文は顧客の日常店舗をもとに顧客の嗜好に適した非日常店舗を推薦するモデルを提案する。

本論文では、日常店舗は顧客が二回以上訪れたことのある店舗とする。また、非日常店舗は顧客が一回も訪れたことのない店舗とする。日常好みベクトルとは日常店舗から顧客の嗜好を反映させたベクトルである。非日常好みベクトルとは非日常店舗における顧客の嗜好を表すベクトルである。店舗ベクトルとは非日常店舗の一店舗の料理や雰囲気などの特徴を表したベクトルである。

本手法の概要図を図 1 に示す。

まず、本手法は日常店舗の本や雑誌のテキストをトピック分析することで顧客の好みの傾向を表すトピックを抽出する。また、非日常店舗のウェブサイトなどの発信情報をトピック分析することで店舗の特徴を表すトピックは抽出される。

抽出されたトピックを基にベクトルを作成する。日常店舗のテキストデータから抽出されたトピックを表す単語の出現確率を基に日常好みベクトルは算出される。非日常好みベクトルは一非日常店舗のテキストデータから抽出されたトピックを基に作成されたアンケートを顧客が回答することで算出する。店舗ベクトルは、一非日常店舗のテキストデータから抽出されたトピックを表す単語の出現確率を基に算出される。

日常好みベクトルと非日常好みベクトルの各軸の相関を調べる。日常好みベクトルと非日常好みベクトルは同一ユーザ群から算出されている。そのため、日常好みベクトルと非日常好みベクトルの軸で相関があれば、相関のある軸は軸変換をする。座標変換された軸は日常好みベクトルと店舗ベクトルを同軸上で表現することができる。

本手法では、同軸上で表現した日常好みベクトルと店舗ベクトルの \cos 類似度を計算する。日常好みベクトルと店舗ベクトルの \cos 類似度の大きさにより、顧客は web

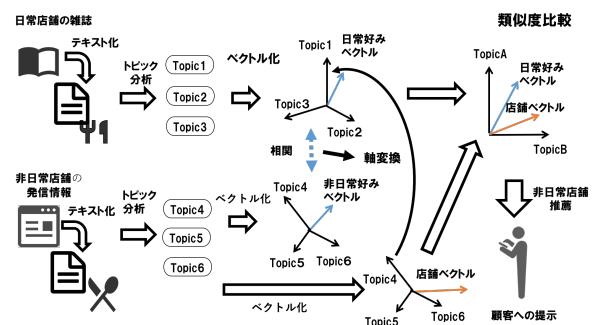


図 1: 顧客欲求に適合した推薦モデル

†立命館大学情報理工学部

‡立命館大学大学院理工学研究科

またはアプリによって非日常店舗が推薦される。

3.2 日常店舗, 非日常店舗のトピック抽出

店舗の特徴は, ホームページや SNS, 雑誌などの媒体から得ることができる. 本手法は, 日常店舗のトピックを本や雑誌という媒体のテキストデータから抽出する. また, 非日常店舗のトピックは発信情報や口コミ, レビューなどのテキストデータから抽出する.

トピック分析はテキストデータから店舗がどのような潜在的なトピックを持っているかを示す. 本手法では, トピックモデルとして LDA (Latent Dirichlet Allocation) [1] モデルを採用する. LDA のトピック数やトピックを構成する要素数は LDA の評価指標で決定する. 評価指標にはトピックが人間にとって解釈しやすいかを表す "coherence" や, モデル M のもとで正解を選ぶ難しさを表す "perplexity" がある [3]. 日常店舗のトピックは, 評価指標を元に m 個抽出される. また, 非日常店舗のトピックも同様に評価指標を基に n 個抽出される.

3.3 トピックを軸とするベクトルの生成

本手法では日常店舗, 非日常店舗のトピック分析によって得られたトピックを用いて日常好みベクトル, 非日常好みベクトル, 店舗ベクトルを生成する.

日常好みベクトルの各要素は一店舗の発信情報に出現する単語のうち, トピックを構成する単語が出現する条件付き確率から算出する. 日常好みベクトルの軸は, 日常店舗のテキストデータから抽出されたトピックである.

非日常好みベクトルは, 顧客が回答した非日常店舗に対するアンケート結果から算出する. 顧客はアンケートを5段階のリッカート尺度で回答する. 本手法では, アンケートの質問項目及び非日常ベクトルの軸を非日常店舗の発信情報テキストデータから抽出されたトピックとする.

店舗ベクトルの各要素は, 全店舗の発信情報に出現する単語のうち, 一店舗の発信情報を構成する単語が出現する条件付き確率から算出する. 日常好みベクトルの軸は, 日常店舗のテキストデータから抽出されたトピックである.

3.4 好みベクトルの相関に基づく軸変換

日常好みベクトルと非日常好みベクトルは, 同一ユーザ群から算出されている. そのため, 日常好みベクトルと非日常好みベクトルが表現されている軸は相関をとることができる. 相関のある軸は, 非日常好みベクトルが表現される軸から日常好みベクトルが表現される軸へ変換される. 軸変換は相関のある軸を同軸上で表現するために行う. 相関のない軸は関連性の低い軸であると考え, 次元削除する.

3.5 類似度比較による推薦

日常好みベクトルと店舗ベクトルは同軸上で表現され, 類似度を比較する. 本論文では類似度の計算手法として \cos 類似度を用いる. \cos 類似度は2本のベクトルの角度を用いて類似度を評価する指標である. \cos 類似度は-1から1の値をとる. \cos 類似度が-1の時, 日常好みベクトルと店舗ベクトルのなす角は180度となり, 類似度が最も低い. \cos 類似度が1の時, 日常好みベクトルと店舗ベクトルのなす角は0度となり, 類似度が最も高い. 日常好みベクトルと店舗ベクトルの \cos 類似度が1に近

い順にウェブまたはアプリ上で非日常店舗が顧客へ自動的に推薦される.

4. おわりに

顧客が飲食店を潜在的な基準で選ぶ際, 顧客は潜在的な基準を自身で表現することができない. そのため, 特定の単語で検索することで目標の飲食店を見つけることができない.

本論文では, 飲食店の発信情報をトピック分析すること顧客の潜在的な基準を見つけ, 潜在的な基準の類似度を基に飲食店を推薦する手法を提案した. この手法を用いることにより, 構造化された顧客の評価やユーザ作成コンテンツからではなく飲食店の発信情報から顧客の嗜好を抽出することができる. そのため, 飲食店を評価した顧客数が少なくても顧客の嗜好に適合した訪れたことのない店舗を推薦することができる.

参考文献

- [1] David M Blei, Andrew Y Ng, and Michael I Jordan. Latent dirichlet allocation. *Journal of machine Learning research*, Vol. 3, No. Jan, pp. 993–1022, 2003.
- [2] Greg Linden, Brent Smith, and Jeremy York. Amazon.com recommendations: Item-to-item collaborative filtering. *IEEE Internet computing*, No. 1, pp. 76–80, 2003.
- [3] David Newman, Jey Han Lau, Karl Grieser, and Timothy Baldwin. Automatic evaluation of topic coherence. In *Human Language Technologies: The 2010 Annual Conference of the North American Chapter of the Association for Computational Linguistics*, pp. 100–108. Association for Computational Linguistics, 2010.
- [4] Badrul Munir Sarwar, George Karypis, Joseph A Konstan, John Riedl, et al. Item-based collaborative filtering recommendation algorithms. *Www*, Vol. 1, pp. 285–295, 2001.
- [5] 金山博ほか. 利用者の好みをとらえ活かす-嗜好抽出技術の最前線-: 4. ユーザ作成のコンテンツに着目した嗜好・評判抽出 2) テキストを用いた評判と嗜好の分析. *情報処理*, Vol. 48, No. 9, pp. 1001–1007, 2007.
- [6] 神寫敏弘ほか. 推薦システムのアプローチ (1). *人工知能学会誌*, Vol. 22, No. 6, pp. 826–837, 2007.
- [7] 土方嘉徳. 嗜好抽出と情報推薦技術. *情報処理学会論文誌*, Vol. 47, No. 4, p. 2015, 2006.