

複数人による行先決定時における嗜好の匿名化を用いた意思表示支援

高田紗也秋[†] 北山 大輔[†]

[†] 工学院大学情報学部システム数理学科 〒163-8677 東京都新宿区西新宿 1 丁目 24-2

E-mail: ^{††} tj317167@ns.kogakuin.ac.jp, ^{††} kitayama@cc.kogakuin.ac.jp

あらまし 消極的な人は、多人数で行く飲食店などの決定時に、自分の行きたい行き先を提示しにくいということがある。そこで、自分の行きたい候補店舗に対し、それに類似する店舗、他者の候補との中間になるような候補も自動的に生成して提示することで、誰の嗜好に基づく店舗なのかをわかりにくくする手法を提案する。この事により、自分の候補店舗を、参加者で検討する候補店舗リストに加えることが容易となることが期待できる。候補リストを生成する要素の組み合わせによる比較実験を行った。結果から提案手法である候補リストの評価が最も高く、参加ユーザを部分的に組み合わせた嗜好に基づく店舗を候補リストに加えることが最も嗜好の匿名化に有効であるということが確認できた。

キーワード 飲食店、推薦

1 はじめに

近年、食べログやぐるなびのような飲食店検索サービスは数多く存在し、条件が決まっているケースでは、容易に飲食店を決定することが可能となっている。このように、飲食店を決定する際、一人で飲食店を決定することは容易だが、複数人で1つの飲食店を決定することは困難である。その理由としては、我々は「誰かの提案待ち」や、「行きたいお店はあるが、言い出しにくい」といったお互いがお互いに気を使った結果、候補店舗が提示できないという問題があると考えている。

そこで、本研究ではユーザの行きたい提案店舗が決まっている際に、「提案しづらい」という問題に着目する。この問題を解決するために、ユーザの行きたい店舗を分からなくするために、最終的に候補リストとして提示するお店は、各自の提案店舗、提案店舗に類似する店舗、複数ユーザの合成嗜好に基づく店舗を同時に提示する事によって、ユーザの提案店舗をカムフラージュできるシステムを提案する。例えば、「ペンギンカフェ」のようにユーザの行きたい店舗が特徴的で言い出しにくい場合、候補リストに他の店舗と一緒に提示されることで、誰がどの店舗を提示したのか分からなくなることを期待する。

具体的には、日常的に登録している参加者全員のお気に入り店舗を用い、その店舗の「レビュー」と「平均価格」から嗜好を抽出し、参加者の嗜好を混ぜた店舗推薦をする。このことにより、誰の嗜好に基づいた推薦店舗なのかを曖昧にすることで、「誰かの提案待ち」や「候補店舗を提示することにより目立つ」ことを解消する。本研究では、参加者がそれぞれ行きたい店舗を気を使うことなく提案することを第一の目的とし、複数人での匿名化を考慮した推薦に加え、参加者全員の嗜好を混ぜた飲食店推薦システムを提案する。

本研究の貢献は以下に示すとおりである。

- 複数人で行き先を決定するときに、候補店舗を出しにくいという問題

- 上記の問題に対する類似店舗でカムフラージュすることの効果の評価

2 関連研究

複数人で飲食店を決める際、嗜好を考慮して推薦システムを提案する研究は数多くされている。

谷川ら [1] は複数人の場合、幹事が嗜好や条件を把握する手間がかかることから、飲食店検索サービスに適用でき、参加者全員が満足する飲食店を推薦するシステムを提案している。過去に利用した飲食店を評価する事で、参加者の嗜好情報を構築している。

谷口ら [2] はグループで飲食店を決める場合に、それぞれのユーザによる飲食店選択に基づき、飲食店を推薦する手法を提案している。グループメンバーがそれぞれ選択した飲食店同士の類似度を算出し、類似度が最も高かった飲食店同士の情報から、グループとしての嗜好を抽出し、飲食店を推薦している。

奥ら [3] は web 上で提供されている膨大な情報から複数人で旅行計画に有益となる情報を得ることや、参加者それぞれの好みといった検討事項が増え困難であることから、簡単な操作のみで複数人の嗜好を反映させた観光地を推薦するシステムを提案している。複数人の嗜好の統合には、集団意思決定の応用手法である見解距離均等法を用いている。

瀬古ら [4] は互いを知っているグループに対して適したコンテンツを推薦するアルゴリズムを提案している。推薦精度向上のために、グループメンバー内の誰の嗜好が重視されるかなどのコンテンツ選択傾向に着目しており、個々人のコンテンツに対する嗜好と、グループの行動履歴を用いてグループメンバー間の力関係を推測し、その力関係に基づいて推薦スコアを算出する手法について提案している。

森田ら [5] はマイクロブログとシズルワードに着目し、嗜好とリアルタイム性を考慮した、飲食店検索システムを提案している。飲食店情報と飲食店に関するおいしさ情報を蓄積するこ

とで、各飲食店と共起するおいしさ情報を結びつけている。

山本ら [6] はユーザが過去に書いたレビューを利用し、ユーザに対して適切な店舗推薦を行うことを目的としている。クエリレビューと店舗に対して書かれたレビュー集合を一つのレビューとした場合の店舗との類似度を計算することで入力レビューがどの店舗について書かれたかを推定している。

伊藤ら [7] は嗜好情報に基づいて個々のユーザにより適した情報を提示するシステムが増加しているユーザの嗜好の変化を検出するために、EC サイトにおける時系列クラスタリングを用いている。ユーザが嗜好に合うと評価した商品にクラスタリングを適用し、そのクラスタリング結果の時系列変化を嗜好の変化として検出している。

大木ら [8] は集団意思決定の場面において、集団内の構成員が「集団の目的を共有している度合」を定量的に表現する手法を提案している。定量化された「目的の共有度合い」を用いることで、その集団の振る舞いや能力の推定を行うことができる。

これらの関連研究は、ユーザが過去に書いたレビューや、Twitter から得られる情報からユーザや複数人での嗜好を混ぜた推薦の研究を行っている。本研究では、レビューから類似度を算出している点は共通しているが、それと同時にいきたい店舗を匿名化したうえで最終的に提示させている。ユーザのいきたい店舗を匿名で提示することで本研究の目的である「意見の提案がしづらい」問題を解決している。

3 研究の概要

本研究では、ユーザのいきたい提案店舗が決まっている際に、「提案しづらい」という問題に着目した。そこで複数人で飲食店に行く際に、候補店舗を提案するプロセスを対象とする。本稿では単純化のために、参加者全員で閲覧可能な候補店舗リストがあり、そこにどのように候補店舗を追加するかというモデルについて論じる。

本研究では、候補店舗が誰の嗜好に基づくものを分からなくすることを匿名化と呼ぶ。匿名化し候補店舗を提示するだけでは、仲の良い友人や知人の場合、誰の提案した店舗であるのかが分かってしまう。それに対し、候補店舗追加時に、その店舗の類似店舗も同時に追加することで、実際に追加された店舗をわかりにくくするような匿名化が考えられる。しかしながら、類似店舗集合のレベルでは誰かの嗜好に基づくものであることはわかってしまう。そこで他の参加者の嗜好も混ぜた店舗を候補リストに追加する。そうすることで誰が提案した飲食店舗であるかが分かりづらくなり、参加者も意見の提案がしやすくなるのではないかと考えた。

また、参加者全員の嗜好を混ぜた店舗の推薦ができる事によりそれぞれの嗜好に適した未知の飲食店に対して興味を持つことができる。結果的に実際の候補店舗は行き先選ばれないかも知れないが、店舗決定のプロセス自体の満足度が向上することが期待できる。

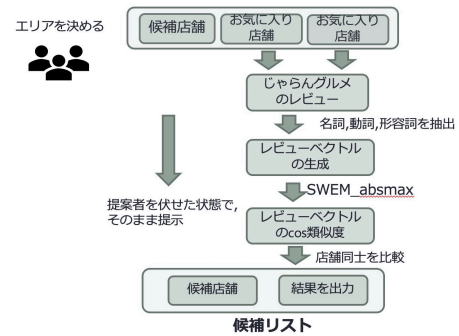


図 1 システムの流れ図

4 嗜好の匿名化を用いた意思表明支援システム

4.1 システムの概要

システムの概要図を図 1 に示す。まず、参加者は友人同士であったり、深く関わりはないが一緒にご飯を食べに行くということを前提とする。参加者全員で話し合ってもらい、エリアを決定する。参加ユーザのお気に入り店舗は、他の参加者におすすめしたい店舗や、自分が訪れてみて高評価であった飲食店を登録してあるものとする。これは、日常的に各自が登録する。例えば、食べログ等のアプリで「保存」や「お気に入り」といった機能で登録されるものを想定している。システム利用時にユーザのお気に入り店舗のうち、行き先エリアに該当する店舗があれば、行き先の候補として提案するかどうかの選択肢がそのユーザに表示される。ここで提案すると選択された店舗を、提案者が伏せられた状態で候補リストに追加する。お気に入り店舗の数に上限はなく、何個でも登録できる。登録する店舗が多いほどユーザの嗜好をより抽出できると考える。各ユーザが登録したお気に入り店舗からそれぞれの店舗のレビューを MeCab で名詞、動詞、形容詞を抽出する。そこから、各店舗のレビューベクトルを生成する。各ベクトルに類似する該当エリア内の店舗をそれぞれ k 件追加する。最後に、ユーザの組み合わせによる店舗の追加を行う。ユーザの組み合わせはユーザ数 n に対して $n!$ 存在することになるが、それでは数が多すぎるので、組み合わせからランダムに m 組み合わせを用いる。組み合わせのユーザの平均ベクトルを用いて、そのベクトルの類似店舗を l 件、候補リストに追加する。このようにして、提案したい店舗、それを隠すための類似店舗、個人の嗜好をぼかすための合成嗜好に基づく店舗によって、候補店舗リストを構築する。最終的に候補店舗リストは、行き先提案店舗数を a とすると $a + n \times k + m \times l$ 件となる。具体的な類似度の計算は 4.2 節で述べる。また、価格帯によるフィルタも行うため、それについては 4.3 節で述べる。

最終的に行き先提案として選ばれた店舗を候補リストに追加する。次に、各ユーザの類似店舗の追加を行う。行き先提案として選ばれた店舗が存在するユーザについてはその店舗のベクトルをユーザのベクトルとする。行き先提案として選ばれた店舗が存在しないユーザは、そのユーザのお気に入り店舗の平均ベクトルをユーザのベクトルとする。

また、候補リストとして提示された店舗の中から参加者全員で話し合って最終的な行き先を決定してもらう。

4.2 SWEM を用いた類似店舗の抽出

レビューの類似度を求めるために Simple Word Embedding Based Models(SWEM) という手法を用いた [9]。SWEM とは単語の各次元の最大値や平均値を文書ベクトルとして採用する手法である。レビューの名詞、動詞、形容詞からレビューベクトルを生成した。1つの店舗の飲食店ベクトルは、その店舗の全レビューベクトルの平均で生成している。また、 \cos 類似度とは、ベクトルのなす角が 0 に近づくほど値が 1 に近づく \cos の性質を利用して類似度を計算する手法である。 \cos 類似度は次の式 1 で示される。分母は A と B の大きさの積、分子は A と B の内積をとる。

$$\cos(A, B) = \frac{A \cdot B}{|A||B|} \quad (1)$$

まず形態素解析器である MeCab を用いて飲食店検索サイトのグルメレビューの形態素解析を行う。辞書は MeCab 用のシステム辞書である NEologd¹ を使用し、抽出されたレビューの名詞、動詞、形容詞に対してじゃらんのグルメレビューをコーパスに fastText で学習した日本語モデルを用いて、300次元のベクトルを得る。レビューの各キーワードベクトルの最大値を採用した SWEM_mean を文書ベクトルとする手法であり、SWEM_max は負の方向に大きいという特徴は扱われていない。このとき、高橋ら [10] の先行研究で用いられている SWEM_absmax 法を使用した。これは各次元で絶対値の最大値を用いることで、正に大きい値だけでなく、負に大きい値でもその単語の特徴とできる手法である。maxpooling とは、次元ごとの最大値を表すことのできる手法であり、minpooling とは、次元ごとの最小値を表すことのできる手法である。 $v(x)$ をあるレビューに含まれる単語ベクトルとし、 L をそのレビューの単語数とすると、 r_i^{\min} は、各ベクトルの最小値で生成するベクトルを表しており、 r_i^{\max} は、最大値で生成されるベクトルを表している。それぞれ式 2、式 3 で定義する。

$$r_i^{\min} = \text{Min} - \text{pooling}(v_1, v_2, v_3, \dots, v_L) \quad (2)$$

$$r_i^{\max} = \text{Max} - \text{pooling}(v_1, v_2, v_3, \dots, v_L) \quad (3)$$

これらを用いて、ある店舗の i 番目のレビューの SWEM_absmax を式 4 で定義しており、ここで $r_i^{\min}_1, r_i^{\max}_1$ はそれぞれのベクトルの 1 番目の要素である。 D は次元数であり、absmax は引数のうち絶対値が最大である値を返す関数である。

$$r_i = [\text{absmax}(r_i^{\min}_1, r_i^{\max}_1), \dots, \text{absmax}(r_i^{\min}_D, r_i^{\max}_D)] \quad (4)$$

4.3 平均価格によるフィルタ

候補店舗リストの店舗のうち、価格帯があまりにも異なるものは、候補店舗としてはふさわしくないため除外する。行き先

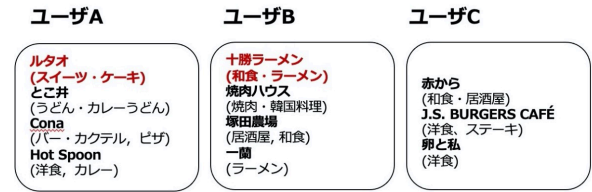


図 2 各ユーザの候補店舗とお行きに入り店舗の例

表 1 出力例

店舗名	カテゴリ	類似度
ルタオ	スイーツ・ケーキ	
十勝ラーメン	和食・ラーメン	
洋食や 純平	洋食全般	0.889
フランボワーズ	その他軽食	0.939
らーめんてつや	和食・ラーメン	0.919
Doremo LeTAO	その他軽食	0.948
ポムの樹	洋食	0.931

提案として選ばれた店舗が存在するユーザについてはその店舗に基づく類似店舗抽出時に、その店舗と 3000 円以上の差があるものを除外する。行き先提案として選ばれた店舗が存在しないユーザに基づく類似店舗抽出時には、そのユーザのお気に入り店舗の価格の平均値と 3000 円以上の差があるものを除外する。組み合わせのユーザの平均ベクトルを用いる場合では、組み合わせ元となるベクトルの店舗の価格もしくはお気に入り店舗の平均価格のいずれとも 3000 円以上の差があるものを除外する。

4.4 実行例

図 2 に各ユーザの候補店舗とお気に入り店舗の例を示す。ユーザ A, B, C の 3 人がいると仮定する。赤字で記載されている店舗は各ユーザの候補店舗で、黒字で記載されている店舗は、各ユーザの登録店舗を示しておりそれぞれの店舗のカッコ内に店舗のジャンルを示した。図 2 の例としては、ユーザ A はスイーツが食べたいが言い出しづらく、ユーザ B はラーメンが食べたいと考えており、ユーザ C は一番行きたい店舗がないという状況を表している。出力例を表 1 に示す。表 1 は、ルタオ、十勝ラーメンはユーザ A、ユーザ B の候補店舗で、洋食や 純平はユーザ C の候補店舗がないため、お気に入り店舗に基づく類似店舗を示している。また、フランボワーズとらーめんてつやは各ユーザのお気に入り店舗に基づく類似店舗を示しており、Doremo LeTAO、ポムの樹はユーザの組み合わせベクトルに基づく類似店舗を表している。

5 評価実験

提案手法による嗜好の匿名化の効果を評価するための比較実験について説明する。下記に提案手法となる候補リストを生成する要素を示す。

- 選択店舗 (X):各ユーザの候補店舗 (候補店舗がある人のみ)
- 類似店舗 (Y):各ユーザのお気に入り店舗に基づく類似店舗

¹ : <https://github.com/neologd/mecab-ipadic-neologd>

● 合成類似店舗 (Z):ユーザの組み合わせベクトルに基づく類似店舗

提案手法ではこれらの要素の和集合で候補リストが生成される。Yとしてユーザ数の3店舗, Zとして, ユーザの組み合わせより3店舗, 全ユーザによる店舗を1店舗用いる。この実験では, この組み合わせを変えることでそれぞれの要素にうおる嗜好の匿名化の効果を評価する。比較実験に用いる組み合わせを下記に示す。

比較手法1 $X \cup Y$ (7店舗)

比較手法2 $X \cup Z$ (それぞれの組み合わせを2店舗ずつ+全ユーザ由来の類似店舗を1店舗)

比較手法3 $X \cup Y$ (3店舗) $\cup Z'$ (全ユーザ由来の類似店舗)

それぞれ候補リストが9から10店舗になるようにYやZによる抽出店舗を調整した。比較手法1は合成類似店舗, 比較手法2は類似店舗, 比較手法3は参加者の組み合わせの効果を評価するために用いる。

5.1 実験方法

被験者1人と仮想ユーザ2人の3人という想定で実験を行う。10人の被験者を用いて実験を行った。また, 仮想ユーザ2人は著者が用意した候補店舗とお気に入り店舗を持つユーザである。まず, 図3に被験者に配布した説明書を示す。説明書には, 飲食店を決定する際の想定している状況と今回の実験方法を記載した。想定している状況としては, 友人とそれぞれ行きたい店舗があり, それを出し合って候補リストを作成し, その中から行き先を決定することになったことを想定している。実験方法の内容としてはじゃらんグルメの検索方法や候補店舗とお気に入り店舗の登録方法の説明が記載してある。表2にじゃらんのエリアの例を示す。じゃらんグルメのエリアは大きく4つに分かれており, 今回はarea2を使用する。被験者はarea2からエリアを決める。仮想ユーザ2人は被験者が決定したエリアに基づき設定する。また, この実験では, 実験時に被験者が候補店舗とお気に入り店舗をその場で入力した。また, 候補店舗を1店舗, お気に入り店舗を3店舗とした。その後, 提案手法と比較手法の出力結果を被験者に候補リストとして提示する。候補リストは店舗名とじゃらんグルメのURLを提示した。

被験者は4種類の候補リストを比較して, 次の4項目に回答した。評価アンケートを記載したGoogleフォームの一部を図4に示す。

- Q1:他の2人がどのようなジャンルを好んでいるか予想できるリストはどれか?
- Q2:自分の好みのジャンルが一番わかりづらくなっているリストはどれか?
- Q3:自分が好むジャンル以外の店舗で行きたい店舗が含まれたのはどのリストか?
- Q4:自分が好むジャンルの店舗を行き先の候補として他の2人に提案しやすいリストはどれか?

5.2 実験結果

評価項目の結果を表3に示す。数値は選択人数である。Q1

表2 じゃらんのエリアの例

area1	area2	area3	area4
東京	池袋・目白・板橋・赤羽	豊島区	千早
東京	八王子・立川・町田・府中・調布	多摩市	豊ヶ丘
京都	京都駅周辺	京都市山科区	安朱稲荷山町

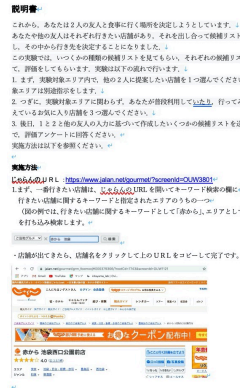


図3 説明書

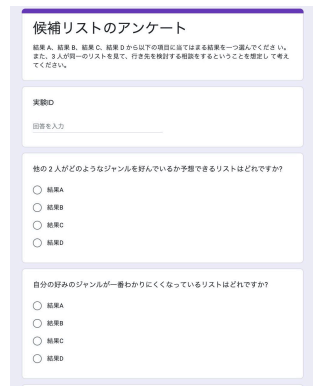


図4 評価アンケートの一部

表3 評価項目の結果

	提案手法	比較手法1 (Zの効果)	比較手法2 (Yの効果)	比較手法3 (組み合わせの効果)
Q1	1	2	3	4
Q2	5	1	3	1
Q3	4	1	5	0
Q4	6	2	1	1

では他の2人がどのようなジャンルを好んでいるか予想できるリストはどれかの結果を示している。この項目では, 他の参加者のリストが予想できなければいけないほど良いので提案手法が一番良いことが結果からわかる。Q2では自分の好みのジャンルが一番わかりづらくなっているリストはどれかの結果を示している。提案手法が最もよく, 比較手法2が次に良い結果となった。この項目では, 比較手法1と比較手法3評価が低かったことがわかる。Q3では自分が好むジャンル以外の店舗で行きたい店舗が含まれたのはどのリストかの結果を示している。評価項目4と同様に比較手法1, 比較手法3が低い結果となった。Q4では自分が好むジャンルの店舗を行き先の候補として他の2人に提案しやすいリストはどれかの結果を示している。この項目では, 比較手法の評価は全体的に低く, 提案手法のみが良い結果となった。

5.3 考察

実験の結果からQ1とQ2, およびQ4は候補リストにおける嗜好の匿名性を評価するための設問である。提案手法はQ1において低い評価となり, Q2とQ4においては高かったため, 候補リストにおいて, 嗜好の匿名性を高める効果があることが確認できる。一方, 比較手法3はQ1において高く, Q2, Q4において評価が低いため, 全ユーザの嗜好を平均化するだけでは不十分であり, 参加ユーザを部分的に組み合わせた嗜好に基づく店舗を候補リストに混ぜることが有効であることが示唆される。

Q3については, 候補リストの推薦性の高さを評価する設問である。ここにおいては, 比較手法2が高く, 次いで提案手法

が高い。この2手法に共通するのは、参加ユーザを部分的に組み合わせた嗜好に基づく店舗を候補リストに加えていることであり、この手法が推薦性も高めることを示唆している。

Q2とQ4の結果より、比較手法2は、Q2で高く、Q4で低い結果となっている。そのため、各ユーザお気に入り店舗に類似する店舗を候補リストに加えることが、匿名性を担保しつつ、店舗決定のために必要な店舗を確保するために有用であることが示唆される。

6 まとめと今後の課題

複数人で飲食店を決める際に意見の提案がしづらいという問題に着目し、匿名性を用いた店舗提示システムの提案、実行例の提示を行った。そのためにSWEMを用いた類似店舗の抽出を行い、さらに平均価格によるフィルタをし、参加者全員の類似店舗を提示した。また、参加者全員の類似店舗を提示する際に参加者の候補店舗も同時に提示することの提案を行った。

提案手法の組み合わせを変えて比較実験を行った。実験の結果から、参加ユーザを部分的に組み合わせた嗜好に基づく店舗を候補リストに加えることによって、匿名性および推薦性が高くなることが示唆された。また、各ユーザのお気に入り店舗に類似する店舗を候補リストに加えることで店舗の決定に必要な店舗を確認できることがわかった。

また、このシステムを利用するにあたり、最適な参加者の人数は2人から8人程度であると感じた。その理由として、人数が多すぎてしまうと、候補リストに提示する店舗数が増えてしまうからである。

今後の課題としては参加者の候補店舗をさらにカムフラージュできるようにレビューの類似度や価格だけでなく、客層や店内の雰囲気などの観点も含めて類似店舗を求める。また、参加者の好まない店舗を登録することによってさらに候補リストを選びやすくする。

7 謝辞

本研究の一部は、2020年度科研費基盤研究(C)(課題番号：18K11551)によるものです。ここに記して謝意を表すものとします。

文 献

- [1] 谷川佳延, 吉田博哉. 参加者の嗜好を考慮した飲食店推薦システムの開発. 第77回全国大会講演論文集, No. 1, pp. 587–588, mar 2015.
- [2] 李明昊, 佐々木淳. 複数人による選択店舗の構造類似性に基づく飲食店推薦システム. In *DIEM*, No. 1, pp. 4–6, feb 2017.
- [3] 奥蘭基, 牟田将史, 平野廣美, 益子宗, 星野准一. 複数人での旅行における嗜好分析による観光地推薦システムの提案. 情報処理学会研究報告, No. 19, mar 2015.
- [4] 瀬古俊一, 八木貴史, 茂木学, 武藤伸洋, 小林透. 行動履歴と嗜好に基づくグループ向けコンテンツ推薦手法の提案. 情報処理学会論文誌コンシューマ・デバイス&システム (CDS), Vol. 2, No. 3, pp. 56–65, dec 2012.
- [5] 山本真史, 山崎俊彦, 相澤清晴. Bag of words と skip-gram 併用によるレビュー・店舗間類似度評価とそれに基づく店舗推薦. 第78回全国大会講演論文集, 第2016巻, pp. 543–544, mar 2016.
- [6] 森田真季, 宮部真衣, 荒牧英治, 灘本明代, 吉野孝. 嗜好とリアル

- タイム性を考慮した飲食店検索システムの構築. 2018年度 情報処理学会関西支部 支部大会 講演論文集, 第2018巻, sep 2018.
- [7] 伊藤冬子, 廣安知之, 三木光範, 横内久猛. Ec サイトにおけるユーザの嗜好の変化の検出. 人工知能学会 論文集, 第23巻, mar 2009.
 - [8] 大木真, 工藤海人, 徳永弦己. 集団意思決定におけるチーム指向性の定量評価とその応用. 知能と情報, No. 4, mar 2018.
 - [9] Shen Dinghan, Wang Guoyin, Wang Wenlin, Min Martin, Renqiang, Su Qinliang, Zhang Yizhe, Li Chunyuan, Henao Ricardo, and Lawrence Carin. Baseline needs more love: On simple word-embedding-based models and associated pooling mechanisms. In *CoRR*, Vol. *abs/1805.09843*, pp. 1–13, 2018.
 - [10] 高橋輝, 北山大輔. 店舗の分散表現に対する意味演算を用いた飲食店検索手法. In *DEIM2020*, p. 5, mar 2020.