

商品選択のための可視化における商品情報の次元削減法の評価 Evaluation on dimensional reduction methods of product information in visualization for product

滝口 魁人[†] 梅澤 猛[†] 大澤 範高[†]
Kaito Takiguchi Takeshi Umezawa Noritaka Osawa

1. はじめに

現在、多くの EC (electronic commerce) サイトでは、テキスト入力による商品名検索を行うことで表示する商品群の絞り込みを行うインターフェースが採用されている。表示内容の絞り込みには、商品名の他に価格やジャンルなどの属性情報も利用されている。このような属性情報を用いた絞り込み方は、利用者は自分の要求に合致する商品を効率的に表示させることができるという利点が存在している。しかし、複数のショップが出品する同じ商品が重複して表示される、不適切なキーワード設定で登録された無関係な商品が表示される、絞り込みに用いた要素以外の要素や好みについて見逃してしまうこともあるなどといった問題点も同時に存在している。 unnecessary 商品表示が増えると、手間と時間が増大し、検索内容に適合した商品群の全体像を把握することが困難になる。そこで、本研究では 2 次元ディスプレイ上に商品の類似性に応じた配置を行うことで、商品間の関係を確認することが容易なインターフェースを検討、評価する。

2. アプローチ

本研究では、検索結果として得られる商品表示における、商品同士の属性情報に基づく関係を視覚的に捉えやすいインターフェースを構築し、評価する。似た属性を持つ商品同士を近接させ、そうでない商品は離して表示するように、商品説明文を分散表現ベクトルに変換し、さらにこの分散表現を二次元へと次元削減する。また、検索結果の表示画面において、利用者は拡大・縮小操作および範囲指定によってインタラクティブに表示内容を調整することで、商品を容易に絞り込むことができるようにする。

3. 次元削減による商品情報の可視化

楽天商品検索 API[1]を使って、商品ジャンルが「スイーツ」である商品データを収集し、その中からランダムに 3,000 件を選んだ。次に、各商品の説明文をインターネット上で公開されている Doc2Vec の PV-DBOW モデル[2]を用いて 300 次元のベクトル表現へ変換した。

得られたベクトル表現に基づいて類似した商品同士を近接表示するにあたって、平面である画面に表示するためには高次元ベクトルを 2 次元へと変換する必要がある。このときに用いる次元削減法の違いが商品選択時間に与える影響について調べるために、次の 3 種類の次元削減法に基づいた可視化が可能な実験用システムを構築した。

- 主成分分析 (Principal Component Analysis - PCA)
- t-SNE (t-distributed Stochastic Neighbor Embedding)
- UMAP (Uniform Manifold Approximation and Projection)

システムの構築には、Python による Web 向けインタラクティブ可視化ライブラリである bokeh を用い、表示の拡大、

縮小、範囲指定など利用者の操作に応じて表示を随時更新する機能を実装した。作成したシステムの表示画面例を図 1 に示す。左右の大きな表示領域は、左が商品点を点として、右が商品画像として表示する領域である。図 1 では、データ点表示領域の四角形で囲んだ部分を商品画像表示領域に表示している。左上部には、検索に用いるテキスト入力欄と、表示に用いる次元削減法を選択するラジオボタンを設けた。

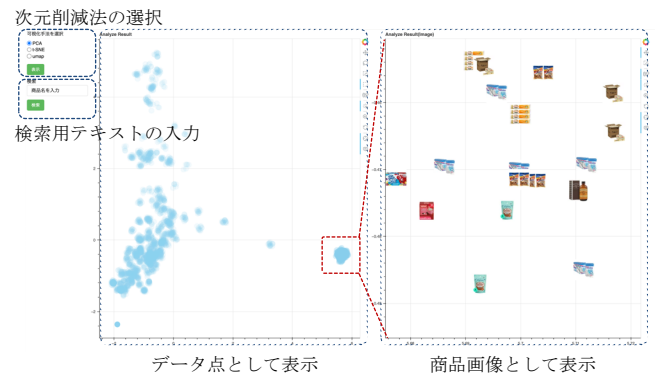


図 1 実験用システムの表示例 (PCA)

4. 評価実験

構築したシステムを使って利用者が実際に商品検索を行った際に、用いる次元削減法の違いが所要時間に与える影響を調べるために被験者実験を行った。比較のため、既存 EC サイト (楽天市場) も対象に加えた。本実験の被験者は 4 名であった。

4.1 実験手順

実験用システムの機能について説明を行った後、操作練習として実際にシステムを用いて検索と表示結果の絞り込みを行う時間を設けた。

実験用システムに用いた 3 種類の次元削減法に基づく表示法と、既存 EC サイトを使用した表示法の、計 4 種類の表示法を用いた。この 4 種類の表示法と、次に示した 4 つのタスクの全ての組み合わせについて、それぞれを完了するのに掛かる時間を計測した。

タスク 1: 「いちごの形状をしたアイスとカップアイスで、それぞれ 10 個、6 個含まれているセット商品」を検索し、選択する

タスク 2: 「ゆずショコラ風味とホワイトベリー風味のレアチーズケーキが 2 個入りの 10 セットで計 20 個含まれているギフト商品」を検索し、選択する

[†] 千葉大学 Chiba University

タスク3: 「ひとロサイズのパンにチョコレートが入っており、ブルボンから発売されている商品の 120 個セットで売られているもの」を検索し、選択する
 タスク4: 「小麦からできている和菓子でサンドされたアイスクリームが 12 個入っている商品」を検索し、選択する

1 つのタスクについて、4 通りの表示法を使って順に検索・選択を行い、これをタスク 1 からタスク 4 まで 4 回繰り返した。計測順が各表示法を用いた際の結果に影響する可能性を考え、それぞれのタスクにおける表示法の利用順をラテン方格に基づき表 1 の順序とした。

表 1 タスクにおける各表示法の計測順

順序	1	2	3	4
タスク 1	楽天	UMAP	t-SNE	PCA
タスク 2	PCA	楽天	UMAP	t-SNE
タスク 3	t-SNE	PCA	楽天	UMAP
タスク 4	UMAP	t-SNE	PCA	楽天

最後に、質問紙調査により主観評価データ収集を行った。調査には、SUS (System Usability Scale) [3]を用いた。

4.2 実験結果

各表示法における被験者のタスク完了時間は図 2 のグラフに示す結果となった。グラフには、タスク 1~タスク 4 を 4 回繰り返した際の、同一表示法を使ったときの所要時間を合計して示した。したがって 1 本のグラフは、1 人の被験者が同一表示法を使ってタスク 1~タスク 4 すべてを行ったときの合計所要時間を示しており、計測順序でみると 1 番目から 4 番目までを 1 回ずつ含んでいることになる。また、各表示法について、被験者 4 人の平均所要時間の合計も示した。

4 人の平均値に注目すると、「実験用システムを用いた場合の所要時間は、楽天市場を用いた場合よりも 210~290 秒程度短く、3 つの次元削減法の中では t-SNE、UMAP を用いた場合、PCA よりも 70~80 秒程度短い。

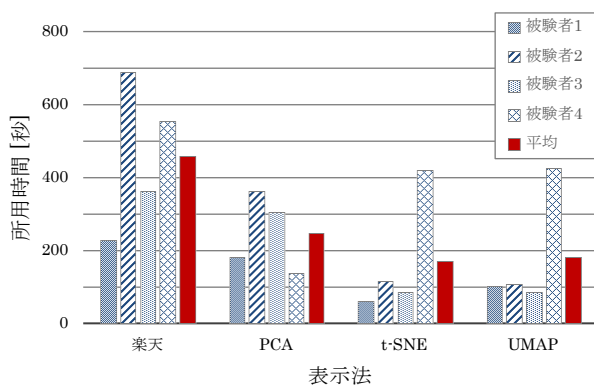


図 2 全タスク所要時間の合計

図 3 は次元削減に UMAP を用いた際の画面表示の一例である。データ点表示領域内の四角形で囲んだ小領域を選択することで、商品画像表示領域へ拡大表示することができる。データ点表示領域の拡大倍率が小さいとき、同一の商品を近傍へと表示することができているため、異なる商品

との関係性を把握しやすくなっている。また、出店元の違う同一商品や、登録されている名前の異なる同一商品を何度も確認する手間が省ける。

PCA がタスク実行に比較的時間を要した原因は、表示特性にあると考えられる。PCA による表示 (図 1) では、UMAP による表示 (図 3) に比べ、1 画面に多種多様な商品が表示される傾向がある。実際に、PCA による表示では、被験者が目的に合致しない商品を選択する事象が確認された。PCA は、t-SNE や UMAP と比べて表示に商品の類似性を強く反映できていないと考えられる。

楽天市場を使った場合、被験者は複数ある検索結果一覧ページを行き来したり、出店元や商品名だけが異なる実質的には同じ商品を重複して確認したりといった、本来不要な操作が多く、所用時間が長い原因と考えられる。

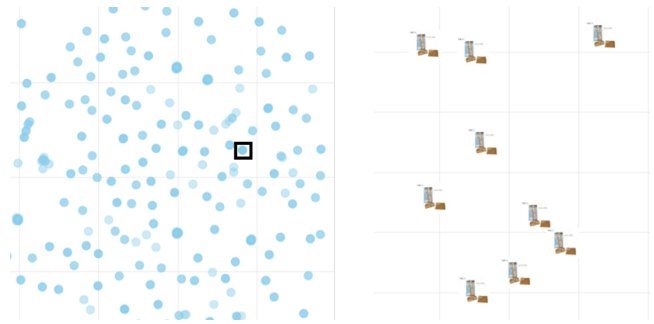


図 3 同一商品の近傍表示の例 (UMAP)

質問紙調査は概ね良好な回答を得ており、「システムの単純さ」と「システムの一貫性」で 9.4 点、「システムに対する自信」で 10.0 点と高スコアであった。被験者からは、本研究の特徴である同一商品を近傍へ表示させる機能について、「商品群の関係を捉えやすい」と好評を得られた。

5. おわりに

本研究では、商品検索における結果表示において、商品同士の属性情報に基づく関係を視覚的に捉えやすいインタフェースを構築し、評価した。被験者実験より、商品の類似性に応じた配置による表示の方が、EC サイト (楽天市場) に比べて効率的に検索を行うことができ、所用時間の短縮に寄与していることを示した。特に、3 種類の次元削減法の中でも t-SNE や UMAP を用いた表示法においては、利用者が商品間を確認することを助け、同一商品の確認を繰り返すような手間を省くことができた。

分散表現変換モデルによる埋め込みベクトルの生成や、次元削減法は商品表示に大きな影響を与える。そのため、今回用いたもの以外の方法で同様の実験用システムを構築し、利用者の目的とする商品の検索に適したモデル、次元削減法を検証する予定である。

参考文献

- [1] 楽天市場商品検索 API (<https://webservice.rakuten.co.jp/documentation/ichiba-item-search/>)
- [2] 奥田 祐樹, “日本語 WIKIPEDIA で学習した DOC2VEC モデル”, (https://yag-ays.github.io/project/pretrained_doc2vec_wikipedia/)
- [3] J.Brooke, "SUS: A quick and dirty usability scale," Usability evaluation in industry, vol. 189, no. 194, pp. 4-7, 1995.