

埋め込み空間上のトピック分布を考慮したアイテム推薦モデルに関する研究

1X16C039-9 小野雄生
指導教員 後藤正幸

1. 研究背景・目的

情報技術の発展に伴い、EC サイトや動画配信サイト等において、ユーザーの購買や閲覧に関する行動履歴がデータとして蓄積されるようになった。ユーザーの嗜好が反映された行動履歴データに基づき、選択する可能性の高いアイテムを推薦することは、顧客満足度の向上を図る上で重要な課題となる。一般にアイテムの選択履歴は、多くのユーザーが共通して選ぶ人気アイテムと、個人の嗜好によって選択されるアイテムによって構成されている。ユーザーの人気アイテムに対する嗜好度合いは、アイテムの性質やイベント等の季節性によって異なる。したがって、EC サイト等の運営者は、このような嗜好度合いの変動に合わせて、個人が嗜好するアイテムと人気アイテムのどちらを重視して推薦するかの戦略を変更する必要がある。しかし、協調フィルタリング(以下、CF)や内容ベースフィルタリング等の従来の推薦手法では、人気アイテムとその他のアイテムを区別しない推薦を行ってしまうため、状況に応じて適切な戦略を選択することができない。以上の従来手法の課題を解決するため、本研究では、マルチトピックモデルの一種である Embedded Topic Modeling(以下、ETM)[1]に着目する。この手法は、自然言語処理の分野に用いられている手法であり、文書ごとに異なる特徴的な単語を持つトピックと、多くの文書に共通して出てくる単語を持つトピックを同時に抽出する手法として知られている。この手法をユーザー行動履歴データに適用した場合、多くのユーザーが選択する人気アイテムが頻出するトピックと個人の嗜好性が強く表れるトピックが抽出されることが期待できる。そこで、本研究では、上述の ETM の特性を利用して、ユーザーの嗜好性が強く反映されるアイテム及び人気アイテムの比重を自由に変更することのできる推薦モデルを提案する。提案手法により、事前の前処理を行わなくても、推薦時にこれらのアイテムの比率を変更することができ、EC サイトが戦略に応じて推薦方針を自由に変えることが可能となる。さらに、実データに提案モデルを適用し、推薦されたアイテムの分析、推薦精度の算出を行うことで、提案モデルの有用性を検証する。

2. 準備

Embedded Topic Modeling(以下、ETM)[1]とは、Latent Dirichlet Allocation(以下、LDA)[2]等に代表される、マルチトピックモデルの一種である。マルチトピックモデルとは、文書内の単語が生成される過程をモデリングした言語モデルであり、トピックが持つ単語分布と文書ごとのトピック分布を算出することによって、文書の持つ潜在的な意味構造を把握することができる。ETMでは、同一の特徴空間上にトピックベクトルと単語ベクトルを埋め込むことでトピックの持つ単語分布を算出している。これによって、ETMは大規模な語彙に対する頑健性を獲得しているだけでなく、多

くの文書に共通して出現している単語(以下、stop words)を埋め込み空間における近傍領域に集めることを可能にしている。図1で示すように、ETMは大きく分けて上部のトピック分布構成部分と下部の単語分布構成部分から成っている。トピック分布構成部分ではニューラルネットワーク(NN)から得られたパラメータ δ_d からトピック分布 θ_d を算出しており、単語分布構成部分では単語の埋め込み表現 ρ とトピックの埋め込み表現 α の内積を計算し、トピックと単語の類似度を算出している。こうして得られたトピック分布と単語分布をかけ合わせ、それを合計することで単語の出現確率を計算している。

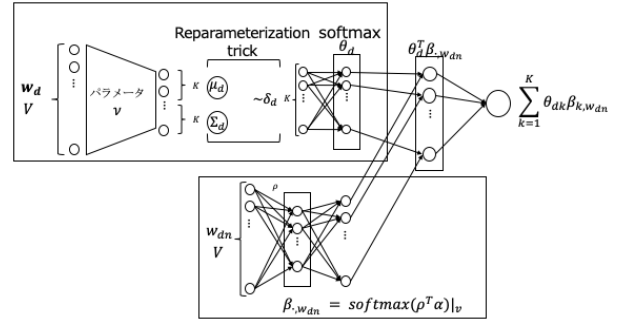


図1: ETMのモデル概要

また、ETMのパラメータの推定は以下の目的関数を最大化させるように、誤差逆伝播法を用いて行う。なお、ここで $q(\delta_d; w_d, \nu)$ は、 w_d を入力、 ν をパラメータ、 δ_d を出力するNNモデルであり、 $KL(\cdot)$ はカルバック・ライブラー擬距離を表す。

$$\mathcal{L}(\alpha, \rho, \nu) = \sum_{d=1}^D \sum_{n=1}^{N_d} \mathbb{E}_q[\log p(w_{dn} | \delta_d, \alpha, \rho)] - \sum_{d=1}^D KL(q(\delta_d; w_d, \nu) || p(\delta_d)) \quad (1)$$

3. 提案手法

文書の潜在的な意味構造であるトピックは、その文書を構成する単語から決定される。これと同様に、ユーザーの選択における嗜好は、それを構成するアイテムから決定づけられる。トピックモデルであるETMを行動履歴に利用すると、多くのユーザーが共通して選択する人気アイテム(自然言語処理におけるstop wordsに対応)が集まったトピックを抽出すると同時に、ユーザー個人の嗜好がよく反映されているトピックを抽出することが期待できる。本研究ではこの点に着目し、多くのユーザーが共通して選択するアイテムと、個人の嗜好によって選択されるアイテムの割合を自由に変更することのできる推薦モデルを提案する。具体的には、ETMの適用を通して得られたトピックごとのアイテム分布の解釈性によって、前者のトピック集合を Z_{pop} 、後者を Z_{pref}

とし、これらのトピック分布に対して重み付けを行うことによって、アイテムの視聴確率を計算することを考える。この重みを $\lambda (0 \leq \lambda \leq 1)$ としたとき、ユーザー u がアイテム i を選択する確率 $\tilde{p}(i|u)$ を以下のように定義する。

$$\begin{aligned} \tilde{p}(i|u) := & \lambda \sum_{z \in Z_{pop}} p(z|u)p(i|z) \\ & + (1 - \lambda) \sum_{z \in Z_{pref}} p(z|u)p(i|z) \end{aligned} \quad (2)$$

提案手法では、式 (2) のパラメータ λ の制御により、推薦における人気アイテムと個人の嗜好によるアイテムの重要度を調節することができる。 λ を 0 に近づければ定番アイテムを重視した推薦を行い、1 に近づけるほど個人の嗜好を重視した推薦を行う。また、 $\lambda=0.5$ の場合、式 (2) は、重み付けを行わないアイテム選択確率 $p(i|u)$ と等価である。 k 個のアイテムを推薦する場合、全アイテムの中から、 $\tilde{p}(i|u)$ を最大にする上位 k 個のアイテムを推薦リストとする。

4. 実データ分析

提案手法の有効性を示すため、提案手法を映画視聴履歴データに対して適用し、分析を行う。

4.1. 分析条件・方法

ユーザー数は 6,000 人、アイテム数は 21,329 個、レコード数は 890,423 件である。このデータを 5,400 人の学習データと 600 人のテストデータに分割し、テストデータの一部を視聴データとし、それ以外をマスクする。本実験は、テストデータ内のユーザーにおいて、一部の視聴データのみから推薦リストを作成し、リスト内に占めるマスク化された実際の視聴履歴の割合を求めることで、その予測精度を評価した。また、トピックごとのアイテム分布における上位 10 アイテムのうち、総視聴頻度の上位 20 アイテムが 8 割以上を占めているトピックを Z_{pop} とし、その他のトピックを Z_{pref} とした。なお学習時における埋め込み空間の次元数は 300、トピック数は 20 とした。

4.2. 分析結果と考察

4.2.1. 提案手法の基本的な推薦精度

まず、提案手法の基本的な性能を確認するために、テストデータ内のユーザーの視聴アイテム数を変化させたときの予測精度の推移を図 2 に示す。なお、比較手法として、代表的なトピックモデルである LDA と、ユーザーベース協調フィルタリングを用い、これらの手法と同じ設定で比較するため、ここでは $\lambda = 0.5$ とした。図 2 から、提案手法は比較手法に対し、全体的に優れた精度を示していることがわかる。また、ユーザーの視聴済みアイテム数の変化に対する精度の変化についても、比較手法である LDA と CF ではそれぞれ 38%、39% 精度が低下しているのに対し、提案手法は 18% の低下に留まっていることが確認できる。

4.2.2. λ による推薦内容の変化

次に λ の変化による推薦内容の変化を観察するため、予め、アイテムごとのトピック分布を基に、全アイテムを c_{pop} と、 c_{pref} の 2 クラスに分け、式 (2) の λ を変化させたときの c_{pop} と、 c_{pref} の予測精度を図 3 に示す。なお、視聴

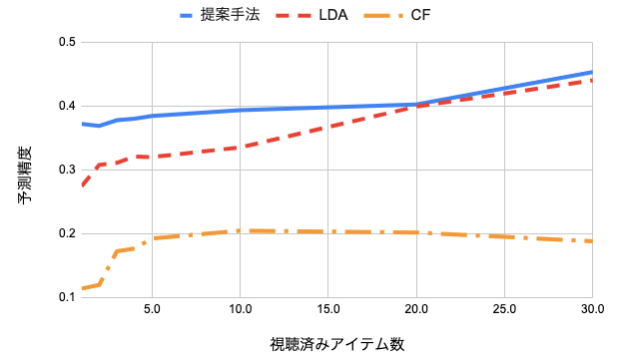


図 2: 視聴アイテム数ごとの推薦リストの予測精度

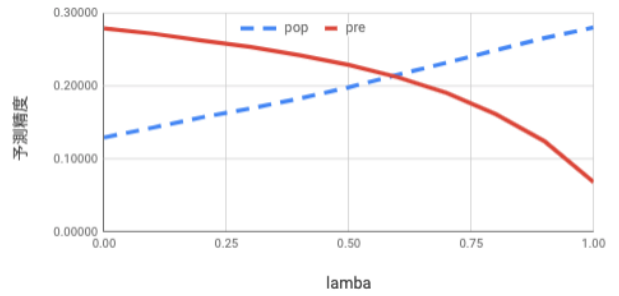


図 3: pop と $pref$ の予測精度推移

済みアイテム数は 30 とし、式 (3) を用いて、全てのアイテム $i = \{1, \dots, I\}$ に対してその所属クラス c_i を定める。

$$c_i = \begin{cases} c_{pop} & \text{if } \arg \max_z p(i|z) \in Z_{pop} \\ c_{pref} & \text{if } \arg \max_z p(i|z) \in Z_{pref} \end{cases} \quad (3)$$

図 3 より、 λ の値が 1 に近づくほど、全ユーザーによる視聴頻度が上位の人気アイテムの予測精度が高くなり、一方で λ の値が 0 に近づくほど、個人の嗜好が反映されたトピックに所属するアイテムの予測精度が上昇していることが確認できる。また、 pop と $pref$ の予測精度の合計が最大になる λ を選択することで、ユーザーが将来選択する確率が最も高くなるようなアイテムの推薦を行うことも可能となる。

5. まとめと今後の課題

本研究ではユーザーの行動履歴に対して、ETM で抽出されるトピックの性質を利用し、トピックに対して重み付けを行うことで、嗜好性が強いアイテムと人気アイテムの比重を自由に変更することのできる推薦モデルを提案した。また、実際の映画視聴履歴データを用いて提案手法の有用性を確認した。今後の課題としては、最適な λ の決定方法や、トピックの適切な分割方法の検討などが考えられる。

参考文献

- [1] Adji B.Dieng, Francisco J.R.Ruiz, David M.Blei, "Topic Modeling in Embedding Spaces" *Transactions of the Association for Computational Linguistics*, vol.8, pp.439-453, 2020
- [2] David M.Blei, Andrew Y.Ng, Michael I.Jordan, "Latent Dirichlet Allocation," *Journal of Machine Learning Research*, vol.3, pp.993-1022, 2003