

Sentence-BERT による、レコメンドへのユーザー意図の反映

青井孝史¹ 久保田崇文¹
¹株式会社カカコム

{aoi_takafumi, kubota_takafumi}@kakaku.com

概要

本論文では、株式会社カカコムが運営する求人ボックスを題材として、大規模で更新頻度が高い検索サービスにおける高性能なレコメンドシステムの設計方針を提示する。文ベクトルを生成するのに適した深層学習モデル Sentence-BERT を用いて、閲覧履歴からユーザーを表現するベクトルを作成することで、パーソナライズされた推薦を可能にする。さらに Item2vec の思想を応用したファインチューニングを行うことで、文章の意味のみならずユーザーの意図をベクトル表現に反映することが可能になる。

1 はじめに

求人ボックスは、正社員、アルバイト、パート、契約社員など多様な雇用形態の求人情報が掲載されているプラットフォームサービスであり、会員登録をしていないユーザーであっても自由に求人の検索を行うことができる。求人ボックスには 2000 万件を超える求人情報が掲載されており、新規求人の追加や古い求人の削除など、情報更新が日々行われている(※2024年9月時点)。

膨大かつ更新頻度の高い求人情報の中からユーザーが自身に適した求人を効率的に見つけられるよう、求人ボックスではレコメンドシステムを導入している。近年、従来の仕組みに加えて新たなアプローチを導入し、大幅に機能を刷新した結果、より良いレコメンドを提供できるようになった。初期のレコメンド手法(以下、旧求人レコメンドと呼ぶ)では、推薦内容をパーソナライズすることができず、求人間の類似性判断にユーザーの意図を反映することも難しかった。そこで一部の機能に新しいレコメンド手法を導入し、パーソナライズ推薦と、ユーザーの意図を反映した類似性判断とが可能になるよう設計した。その結果、レコメンド経由での求人クリック数が増加し、サービス利用の活性化が確認された。

この論文では、新しく導入したレコメンド手法(以下、新求人レコメンドと呼ぶ)の設計方針と、その実現に用いた技術的アプローチについて詳述する。

2 既存のレコメンド手法の課題

ユーザーに対して自動的にアイテムを推薦するアルゴリズムは、コンテンツベースフィルタリングと協調フィルタリング、および両者のハイブリッド方式の3種類に大別できる。旧求人レコメンドでは主に MoreLikeThis というコンテンツベースフィルタリングに類する方法を活用していた。これに対し、新求人レコメンドはコンテンツベースフィルタリングの利点は受け継ぎつつ、協調フィルタリングの考え方を活用した Item2vec という手法の利点を取り入れた手法となっている。

2.1 MoreLikeThis の利点と課題

2.1.1 利点：全ての求人が推薦候補となる

旧求人レコメンドではさまざまな手法を用いていたが、ここでは中心的な役割を果たしていた MoreLikeThis というアルゴリズムについて解説する。この方法では、ユーザーがある求人を閲覧した際、その求人文中から TF-IDF[1]を用いて特徴語(他の文ではあまり登場しないが、対象にしている文にはよく登場する単語)を抽出し、それらの特徴語がどの程度の重要性を有しているかを基準に他の求人文をスコアリングする。スコアが高い求人を閲覧中の求人と類似していると判断し、ユーザーに推薦する仕組みである。

このようにアイテムの内容だけを利用して推薦を行う手法は、コンテンツベースフィルタリングと呼ばれる。コンテンツベースフィルタリングの利点は、説明文やタグなどアイテムの内容に関する情報さえあれば、そのアイテムを推薦候補にできることである。前述のように、求人ボックスは掲載される求人

情報の更新頻度が高く、新しい求人が日々追加される。コンテンツベースフィルタリングはそのような新規求人をすぐに推薦候補にできる点が優れている。

2.1.2 課題：ユーザーの意図を反映できない

MoreLikeThis を用いたレコメンドは、個々の求人を閲覧すると、その求人に類似する(特徴語を共有する)求人が推薦結果として表示される仕組みであり、興味関心が全く異なるユーザーでも、同じ求人をクリックすれば同じ求人が推薦されていた。ユーザー個人の特性や行動文脈を考慮してパーソナライズした推薦を行えるようにはなっていなかった。パーソナライズ推薦を行えなかった理由のひとつは、閲覧中の求人だけに類似する求人を探す仕組みだったことにある。各ユーザーが閲覧した複数の求人から特徴語を抽出する仕組みに拡張すれば、行動履歴に応じてユーザーごとに推薦結果を変えること自体は可能である。しかし、MoreLikeThis において類似性の判断基準となる特徴語は TF-IDF に基づき機械的に抽出されるものであり、ユーザーの意図や関心を考慮したものではない。ユーザーが特徴語とは別の単語に注目して求人を閲覧していた場合、推薦結果はユーザーの期待から外れる可能性がある。また、ユーザーの意図は必ずしも特定の単語に紐づけられるほど明確に言語化できるとは限らない。このように MoreLikeThis によるレコメンドはユーザーの潜在的な意図を十分に捉える仕組みになっておらず、パーソナライズ推薦をするには新しい仕組みを導入する必要があった。

2.2 Item2vec の利点と課題

新求人レコメンドでは、ユーザーの意図を汲んだパーソナライズ推薦を実現するために、Item2vec という手法の考え方を取り入れている。Item2vec は、ユーザー行動データをもとにアイテム間の関係性を学習し、類似ユーザーや共通閲覧アイテムを手がかりに推薦を行う協調フィルタリングの考え方を活用した手法である[2,3]。

2.2.1 利点：ユーザーの意図を学習できる

通常、協調フィルタリングの効果を十分に発揮するには、ユーザーの長期的な行動データが必要となるが、求人ボックスでは会員登録やログインを必須としておらず、同一ユーザーの行動履歴を長期間にわたって追跡することを確実にできない。そのた

め、活用できる行動データはセッションのような短期的な履歴に限られる。Item2vec はセッション単位の短期的なデータを活用する手法であり、長期的な行動履歴を必要としない。そのため、求人ボックスにおいても、そのポテンシャルを十分に発揮できる手法である。

Item2vec は Word2vec における単語と文章の関係をアイテムと閲覧履歴に当てはめたものである。Word2vec では、文中の単語の共起情報を基に各単語のベクトル表現を学習する[4]。具体的には、共起頻度が高い単語どうしは同じ文脈で利用される可能性が高いとして、それぞれのベクトルが類似するように学習する。Item2vec ではアイテムを単語、ユーザーのアイテム閲覧履歴を文章とみなして、Word2vec と同様に共起情報を基に各アイテムのベクトル表現を学習する。同じセッションで一緒に閲覧されることが多いアイテムどうしは、類似したベクトルをもつようになる。同一セッションで閲覧されることが多いアイテム群は Item2vec により似たベクトルを持つようになり、そのことは「意味的に近い」と表現される。この表現は、そのアイテム群がユーザーにとって何らかの関連性を持っている(近い)ことを示しており、「意味」とは一連のアイテムを閲覧したユーザーの意図と言い換えることもできる。Item2vec は、閲覧履歴の背後にあるユーザーの意図をベクトルとして捉える手法だといえる。

2.2.2 課題：推薦候補が限られてしまう

Item2vec に限らず、協調フィルタリング方式の推薦には、全く閲覧されたことのないアイテムを推薦候補に含めることが原理的にできないコールドスタート問題がある[5]。コールドスタート問題は求人ボックスにおいては特に深刻となる。なぜなら求人ボックスでは日々新しい求人情報が追加され古い求人が削除されていくため、まだ誰にも閲覧されていない新しい求人の比率が非常に高いからである。したがって、Item2vec によるレコメンドアルゴリズムを導入しても、推薦候補となる求人は全体のごく一部に限られてしまう。

2.3 ハイブリッドアプローチの必要性

MoreLikeThis は全ての求人を推薦候補とできるが、ユーザーの意図を推薦に反映することが難しい。一方で、Item2vec はユーザーの意図を捉えることができるが、閲覧履歴に含まれない求人を推薦候

補とすることができない。新しいレコメンドシステムは両者の長所を併せ持つことが求められる。すなわち、未閲覧求人を含めた全求人推薦対象にでき、推薦内容にユーザーの意図を反映させられることが求められる。

3 新しいレコメンド手法の設計

新求人レコメンドの仕組みの概要を図1に示す。

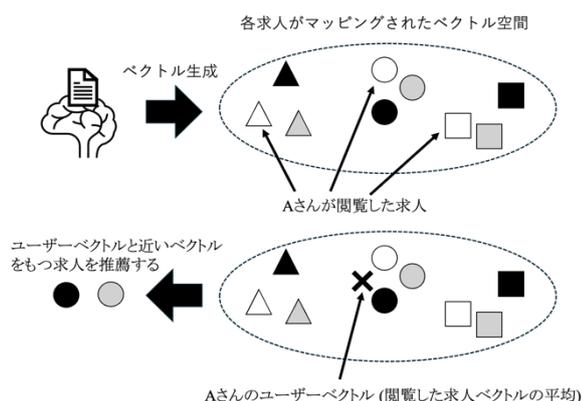


図1 新求人レコメンドで推薦する仕組み

それぞれの求人文をベクトルに変換し、各ユーザーが一定期間に閲覧した全ての求人文ベクトルの平均を、そのユーザーを表現するユーザーベクトルとする。このユーザーベクトルと類似したベクトルをもつ求人を各ユーザーに推薦する仕組みとなっている。この仕組みであれば、求人文さえあれば求人をベクトル化できるので、まったく閲覧されたことのない求人も含めて、全ての求人を推薦候補に含めることが可能である。この点についてはコンテンツベースフィルタリング方式の利点を受け継いでいる。

求人文をベクトルに変換するために用いているのは Sentence-BERT[6]である。また、ユーザーの意図をベクトルに反映させるために、ユーザーの閲覧履歴を用いて Sentence-BERT のファインチューニングを行なっている。ファインチューニングの方法には、Item2vec の考え方を取り入れている。この章では、新求人レコメンドに求められる要件について、それらが何故 Sentence-BERT を用いることで満たされるのか、詳述する。

3.1 Sentence-BERT によるベクトル生成

新求人レコメンドでは、ユーザーごとに異なる閲覧履歴からユーザーベクトルを求める。このユーザ

ーベクトルを基に類似度計算を行うことで、推薦のパーソナライズを可能にしている。ユーザーベクトルは閲覧した求人文のベクトルを平均するだけで求まるので、1 つでも求人を閲覧すればユーザーベクトルが作られる。そのため、ユーザー登録してもらって長期の閲覧履歴を追跡する必要はない。

求人文ベクトルの平均であるユーザーベクトルを基に類似度計算を行うには、足し合わせて平均をとっても各求人文ベクトルの意味が保持できるようになっていなければならない。今回、求人文をベクトル変換するのに用いられる Sentence-BERT は、自然言語処理の様々なタスクに高い性能を発揮する深層学習モデルである BERT[7,8]を、高精度な文章ベクトルを生成できるよう改良したものである。BERT は文を構成する各トークンに対応するベクトルを出力するのに最適化されたモデルであり、文全体の意味を捉えたベクトルを出力する用途には最適化されていない。それに対して Sentence-BERT は、文どうしの意味的類似性を評価するよう最適化されており、文全体を表現する単一のベクトルを生成する。Sentence-BERT で生成された文ベクトルは文の意味を反映しており、複数の文のベクトルの平均をとって文章全体の意味を捉える操作に適している。したがって各求人文ベクトルの平均からユーザーベクトルを生成する今回の用途に、Sentence-BERT は適したモデルだといえる。

Sentence-BERT で生成された求人文ベクトルは文全体の意味を反映している。そのため、旧求人レコメンドのように、特定の特徴語にのみ着目してユーザーが実際に着目していたキーワードの意味を取りこぼすことはない。ただし、事前学習しかしていない Sentence-BERT により生成されたベクトルには、求人文の文章としての意味は反映されても、文章に表れないユーザーの意図は反映されない。求人文ベクトルにユーザーの意図を反映させるには、次に述べるファインチューニングを行う必要がある。

3.2 ファインチューニング

新求人レコメンドでは、推薦内容にユーザーの意図を反映させるためにユーザーの閲覧履歴を用いて Sentence-BERT のファインチューニングを行う。図2に、ファインチューニングの効果を示す。ファインチューニングなしでは Sentence-BERT がつくるベクトル空間は求人文の文章としての意味を反映するだけである。このベクトル空間にユーザーの閲覧意

図を反映させるために、閲覧履歴を用いたファインチューニングを行う。

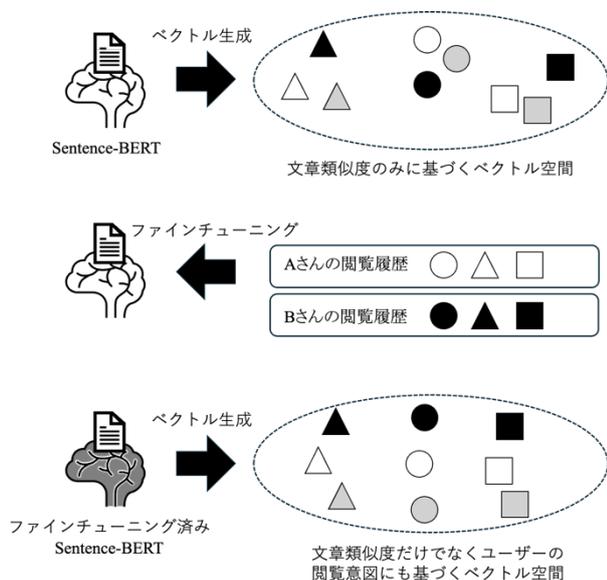


図 2 ファインチューニングの効果

新求人レコメンドでは、Siamese ネットワークを用いて Sentence-BERT のファインチューニングを行う。Siamese ネットワークでは2つの文とその2文間の類似度を入力とする。それぞれの入力文をベクトル化して計算した類似度と、入力として与えた類似度との差が小さくなるように学習する。これにより、類似した文章どうしのベクトルが近くなる。同じセッションに登場した求人ペアの類似度を 1、そうでない求人ペアの類似度を 0 とし、2つの求人文とその類似度を入力とする。2つの求人文をベクトル化してコサイン類似度を計算し、同じセッションに登場した求人ペアのコサイン類似度は 1 に近づくように、そうでない求人ペアのコサイン類似度は 0 に近づくようにモデルを学習する。こうして学習された Sentence-BERT モデルは、セッション内で共起する頻度が高い求人どうしを類似したベクトル、共起しない求人どうしを類似しないベクトルとして、それぞれ生成するようになる。

このファインチューニング方法は Item2vec の考え方を応用したものである。Item2vec ではベクトルを生成するために閲覧履歴を用いるのを、この方法ではすでに生成されたベクトルを修正するために閲覧履歴を用いる。Item2vec と同様に、同じセッションで閲覧されることが多い求人のベクトルは、閲覧時のユーザーの意図が共通しているという意味で「意

味的に近く」なる。ベクトル空間が文章の意味だけでなくユーザーの意図を反映するようになるので、ベクトルの類似度を活用したレコメンドがユーザーの意図を捉えられる可能性が高くなると期待できるのである。

たとえば、データアナリストを募集する2つの求人があり、両者とも統計学の知識やプログラミング経験を求める旨が書かれた、似たような求人文であるとする。しかし、一方は標準レベルの技術力に加えてドメイン知識があることを求める一文があり、もう一方は高度な技術力を求める反面ドメイン知識は求めないという一文がある場合、2つの求人は文章としては似ていても、求職者にとっては「意味的に近い」。一方、同じデータアナリストの募集ではあるが、データ分析の対象となる分野が異なるために文章が類似しない2つの求人があるとする。この場合、もし両者がドメイン知識を求めている場合は、データ分析のスキルを有する求職者にとっては自分が応募資格を満たす「意味的に近い」求人となるだろう。ファインチューニングにより、前者の求人ペアは類似度が下がるように、後者の求人ペアは類似度が上がるように、それぞれのベクトルが補正されることになる。この補正がなされることにより、ユーザーベクトルと似たベクトルをもつ求人を推薦したとき、推薦された求人がユーザーのニーズに合致する可能性は高まると考えられる。

4 まとめ

本論文では、求人ボックスに新しく導入したレコメンド手法の設計方針について述べてきた。新しい手法では、Sentence-BERT によって求人文からベクトルを生成する。閲覧履歴を用いたファインチューニングを行うことで、Sentence-BERT が生成するベクトルは、文章としての意味だけでなくユーザーの閲覧意図をも表現できるようになる。ユーザーの意図を反映した推薦を行える新手法を加えることにより、レコメンド経由でのクリック数が増加した。この成果は、文章類似度とユーザー意図とを反映させたベクトル空間を用いるレコメンド技術の有効性を示すものであるといえる。この新手法を導入した後も、求人ボックスではユーザーベクトル作成のリアルタイム化など、レコメンドの最適化を続けている。今後もさまざまな技術を駆使し、サービス全体の向上を図っていく予定である。

参考文献

- [1] Aizawa, Akiko. An information-theoretic perspective of tf-idf measures. **Information Processing & Management**, 2003, 39. 1: 45-65.
- [2] Barkan, Oren; Koenigstein, Noam. Item2vec: Neural item embedding for collaborative filtering. In: **2016 IEEE 26th International Workshop on Machine Learning for Signal Processing (MLSP)**. IEEE, 2016. p. 1-6.
- [3] Barkan, Oren; Caciularu, Avi; Katz, Ori; Koenigstein, Noam. Attentive item2vec: Neural attentive user representations. In: **ICASSP 2020-2020 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)**. IEEE, 2020. p. 3377-3381.
- [4] Mikolov, Tomas; Sutskever, Ilya; Chen Kai; Corrado Greg S; Dean Jeff. Distributed representations of words and phrases and their compositionality. In: **26th International Conference on Neural Information Processing Systems (NIPS 2013)**, 2013, p. 3111-3119.
- [5] Gope, Jyotirmoy; Jain, Sanjay, Kumar. A survey on solving cold start problem in recommender systems. In: **2017 International Conference on Computing, Communication and Automation (ICCCA)**. IEEE, 2017. p. 133-138.
- [6] Reimers, Nils; Gurevych, Iryna. Sentence-BERT: Sentence embeddings using siamese BERT-networks. In: **Conference on Empirical Methods in Natural Language Processing and International Joint Conference on Natural Language Processing (EMNLP-IJCNLP)**. ACL, 2019. p. 3980-3990.
- [7] Devlin, Jacob; Chang, Ming-Wei; Lee, Kenton; Toutanova, Kristina. BERT: Pre-training of deep bidirectional transformers for language understanding. In: **Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies (NAACL-HLT 2019)**. 2019. p. 4171-4186.
- [8] Choi Hyunjin; Kim, Judong; Joe, Seongho; Gwon, Youngjune. Evaluation of BERT and ALBERT sentence embedding performance on downstream NLP tasks. In: **2020 25th International conference on pattern recognition (ICPR)**. IEEE, 2021. p. 5482-5487.