

オンラインレビューに関する自己教師あり学習を利用した 説明性を有する POI 推薦手法

片山 一† 牛尼 剛聡††

†九州大学大学院 芸術工学府 〒815-8540 福岡県福岡市南区塩原4丁目9-1

††九州大学大学院 芸術工学研究院 〒815-8540 福岡県福岡市南区塩原4丁目9-1

E-mail: †katayama.hajime.498@s.kyushu-u.ac.jp, ††ushiana@design.kyushu-u.ac.jp

あらまし 近年、ユーザが Point of Interest (POI) を探す際にオンラインレビュー上の情報を参考にすることが広く行われている。しかし、全てのレビューを読むことはユーザにとって負担が大きく非現実的である。一方、一部のレビューのみを参考にすると有用な情報の見落とししてしまう可能性がある。本研究ではユーザの目的に関連する POI を推薦し、それぞれの POI に関して目的に対応した特徴を提示するシステムを提案する。提案手法では BERT を用いて、「じゃらん net」のレビューをもとに自己教師あり学習を行ったモデルを生成し実験を行った。

キーワード 地理データ活用, 観光, POI 推薦, 説明可能性, BERT

1 はじめに

近年、インターネットの普及によりユーザが Point of Interest (POI) を探す際に「じゃらん net」¹などのオンラインレビュー上の情報を参考にすることが広く行われている。レビューには投稿者が実際に POI に訪れた際に体験したことや感想などが記述されており、それぞれの POI でどのような体験が行うことができるのか詳細に知ることができる情報源の一つとなっている。しかし、POI 毎に投稿されているレビューの数は異なり、中には数千件を超える場合も数多く存在する。対象とする POI の全てのレビューを読むことは困難であり、結果的に一部のレビューのみを参考にしてしまうことが多い。一方、一部のレビューのみを参考にする場合では、ユーザが求めていた情報や POI が持つ魅力などを見落とししてしまう可能性がある。

このようなユーザの負担や情報の見落としの可能性を減らすために、本研究ではユーザの目的に関連した POI 推薦と目的に対応付けた各 POI の説明を提示するシステムの開発を目的とする。提案システムの全体像を図 1 に示す。システムではユーザが与える目的によって地図上に表示される内容が異なる。例えばユーザがクエリとして「デートにおすすめ」と入力した場合、対象地域のデートスポットと各スポットの特徴が提示される。また、別のクエリとしてユーザが「家族連れにおすすめ」と入力した場合も同様に、クエリの内容に関する POI と説明文が表示される。この際に、二つのクエリで同様の POI が推薦された場合においても、提示される説明文の内容が異なり、ユーザが欲しい情報を適切に表示することを目指す。これによりユーザが全てのレビューを読むことなく、効率的に欲しい情報を得ることができる。

本論文の貢献は以下の通りである。

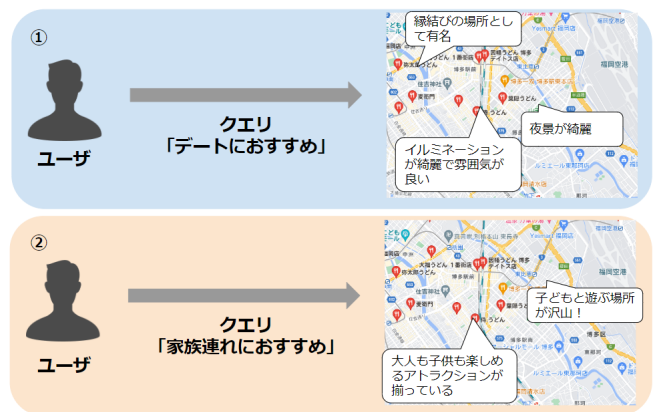


図 1 システムのイメージ図

- クエリに応じて POI 推薦と説明文自動生成を行い、ユーザがスポットを探す際の情報収集を支援するシステムを提案する。
- 日本語事前学習済みモデルの BERT にレビューの内容からスポットを予測するタスクを学習させることで、クエリから POI 予測モデルを生成する手法を提案する。

本論文の構成は以下のとおりである。2 章では、関連研究について述べる。3 章では本論文で利用するモデルについて述べ、4 章では提案手法について述べる。5 章では利用データセットについて述べる。6 章では評価実験の方法について述べ、7 章では実験結果について述べる。最後に 8 章でまとめと今後の課題について述べる。

2 関連研究

2.1 POI 推薦

近年、POI 推薦に関する研究においては、対象ユーザの過去

1: <https://www.jalan.net/>

の行動履歴や POI の特徴, 他ユーザとの関係など様々な情報をデータセットとすることで, 深層学習によって対象ユーザに対し POI を推薦することを可能にしている [1]. Hu ら [2] はグラフ埋め込み法による POI 推薦システムを提案している. 特にユーザ, POI, POI の特徴間にある高次の相互関係をペアワイズによる相互作用関係の集合に分解することで, 効率的なモデル化に取り組んだ. Yang ら [3] は, 協調フィルタリングと半教師あり学習を組み合わせて POI に対するユーザの嗜好をモデル化を行った. Ma ら [4] は, 地理的情報, カテゴリー情報, 他ユーザに関する情報, スポットの人気度という複数の情報を組み込んだ POI 推薦フレームワークを提案している. Zhao ら [5] は時間間隔, 距離感覚, 近傍の POI を POI に関するコンテキストとし, コンテキストとユーザの逐次的な訪問パターンを同時にモデル化を行い, 次にユーザがどこに訪れるべきかを推薦する手法を提案している.

2.2 説明可能性

Zhang ら [6] は e コマースでのアイテム推薦の際, 各アイテムの特徴に基づいて推薦理由の説明を用いることでユーザの購買行動に大きな差が出るとしている. このため情報推薦における説明可能性に関する研究は, 推薦システムに対する信頼性やユーザの満足度を高めるうえで重要なトピックの一つとなっている.

Li ら [7] はユーザ ID とアイテム ID に基づいて推薦と説明生成を行うモデルを提案し, Transformer での個別化された自然言語生成を行っている. Geng ら [8] は個別化されたプロンプトのテンプレート集を用いて学習を行うことで, 説明可能推薦など様々な推薦タスクを一つのフレームワークに統合することに成功している. また Ge ら [9] は推薦システムにおける公平性という問題において説明可能な公平性という考え方を導入し, Li ら [10] は説明可能な推薦システムによって生成された説明文を標準的なランキング手法で定量的に評価を行うベンチマークデータセットを提供している. このように, ユーザに対してなぜそのアイテムを推薦するのかという理由を理解し, より高度な推薦システムの開発を行うためにも説明可能な推薦に関する研究は広く行われている.

3 レビューを対象とした自然言語処理

本研究では自然言語処理モデルの BERT [11] を用いる. BERT は事前学習として以下の 2 つのタスクを行うことで入力として用いた大規模な文章コーパスから汎用的な言語のパターンを学習する.

- Masked Language Modeling(MLM)

入力の中からランダムに選ばれた 15% のトークンを [MASK] トークンで隠し, 周りの単語から [MASK] されたトークンを予測する.

- Next Sentence Prediction(NSP)

事前学習時に BERT には常に二つの文章のペアが入力される. 入力された文章が連続したものであるか, そうでないかを

当てるタスク.

BERT は一般的な文章 (Wikipedia など) をコーパスとして事前学習を行った後, 特定の領域に関連するコーパスを用いて追加で事前学習を行うことで, その領域での BERT の予測精度が向上すると考えられている. 例えば, Lee ら [12] は生物医学に関するコーパスを用いて追加で事前学習を行うことで, 一般的なコーパスのみで事前学習させた BERT に比べて, 生物医学に関する様々なテキストマイニングタスクでの精度の向上を報告している. また, 仁木ら [13] らは金融領域に関する自然言語処理のタスクにおいて, 日本語の金融コーパスを用いた追加事前学習を行うことの有効性を示している.

本論文では二つの事前学習済みモデルを用いる. 一つ目は東北大学の乾研究室が公開している日本語事前学習済み BERT モデル²である. このモデルは Wikipedia の日本語記事を用いて学習されており, 以降このモデルを「BERT-base」と呼ぶ.

二つ目は BERT-base に「じゃらん net」に記載されているレビューをコーパスとして MLM と NSP を追加学習させたモデルである. 以降, このモデルを「BERT-review」と呼ぶ. このタスクを行う意図として, 日本語モデルの中でもよりオンラインレビューの内容に特化したモデルを作成することにある. 例として表 1 を挙げる. 表は「ここは [MASK] におすすめてす。」という文章に対して, 各モデルが [MASK] に入る単語を予測し, 予測スコアの高かった順に上位 10 件を示している. BERT-review は「観光」「カップル」「ドライブ」などのように観光サイトのレビューの内容に記載されているであろうと考えられる単語が上位に予測されていることが分かる.

BERT-review がレビューに特化したモデルであり提案手法ではレビューをデータセットとして用いるため, BERT-base と比較して高精度の結果が期待できると仮定している. そこで本論文では提案手法の有効性を確認するとともに二つのモデルの比較実験も行う.

	BERT-base	BERT-review
1	主	観光
2	大人	カップル
3	あなた	ドライブ
4	非常	デート
5	週末	ハイキング
6	リスナー	非常
7	最高	散歩
8	子供	ファミリー
9	特別	女性
10	すぐ	絶対

表 1 [MASK] に入る単語を予測した際に, 各モデルでの予測スコアが上位 10 件の結果

2: <https://github.com/cl-tohoku/bert-japanese>

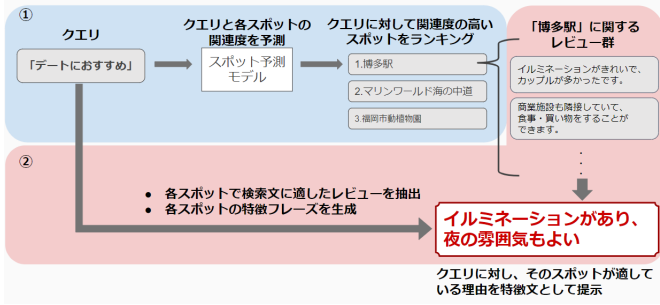


図2 提案手法

4 提案手法

4.1 提案手法の概要

本研究ではユーザが入力したクエリから関連する POI を推薦し、各 POI に関する説明文を抽出する手法を提案する。提案手法の流れを図 2 に示す。本手法では①クエリをスポット予測モデルに入力し、出力された分類スコアの高い順にスポットをランキングする。②前の処理でランキングが上位のスポットをクエリに対応する POI とし、各 POI に関するレビュー群からクエリの内容に関係するレビューを抽出する。以上の 2 つの処理を順に行う。各処理でクエリ情報を用いているため POI 推薦と説明文生成、それぞれの処理の出力がクエリと対応付いた内容となる。

4.2 スポット予測モデル

提案手法ではユーザが与えたクエリから推薦する POI を決定するために機械学習モデルを利用する。図 3 に機械学習モデル構築とシステムでの利用方法の概要を示す。

提案手法を形式的に定義する。POI 集合 $PT = \{pt_1, pt_2, \dots, pt_{|PT|}\}$ とレビュー集合 $RV = \{rv_1, rv_2, \dots, rv_{|RV|}\}$ が与えられているものとする。レビュー rv は、それが対象とする POI_{pt} とレビューテキスト T のペア $rev = (pt, T)$ として定義される。ここで、 $pt(rv)$ はレビュー rv の対象とする POI を表し、 $text(rv)$ はレビュー rv のレビューテキストを表す。

レビューテキスト T は、文 s_i の集合として以下のように定義される。

$$T = \{s_1, s_2, \dots, s_T\}$$

レビュー文 S が与えられたときに、そのレビュー文が POI_{sp} について述べられたと考えられる確率を以下の式で表す。

$$p(pt|S)$$

本研究では、ユーザのクエリテキスト q が与えられたとき、確率 $p(pt|q)$ をクエリテキスト q に対する pt の適合度ととらえる。本研究では $p(pt|q)$ を計算するために機械学習モデルを利用する。本研究では以下の式で定義されるモデルを考える。

$$\mathbb{Y} = f(\mathbf{x}; \theta)$$

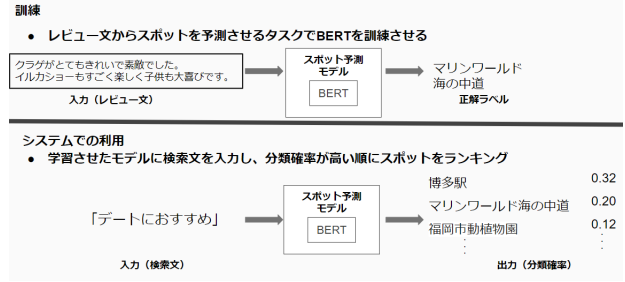


図3 スポット予測モデル

ここで、 \mathbf{x} は、対象とするテキストの埋め込みベクトルである。また、 θ はパラメータであり、 $\mathbb{Y} = (y_1, y_2, \dots, y_{|PT|})$ は予測値である。ここで、 y_i はスポット pt_i に対する確率を表し、 $\sum_j y_j = 1$ とする。

上記の機会学習モデルを学習させるために、以下の教師データ D を利用する。

$$D = \{(\mathbf{x}, \mathbf{y}) | \mathbf{x} = emb(s), \mathbf{y} = one_hot(pt), s \in text(rv), pt = pt(rv), rv \in RV\} \quad (1)$$

ここで、 $emb(S)$ は文 S の埋め込み表現であり、 $one_hot(pt_i)$ はスポット i 番目の要素のみが 1 であり、ほかの要素が全て 0 となるワンホットベクトルである。

訓練時には、「じゃらん net」に記載されているレビューを入力とし、入力のレビューが参照しているスポットを正解ラベルとしたデータセットを用いてレビューからそのレビューが書かれたスポットを予測するタスクを学習で行う。日本語事前学習済みモデル BERT に対し、レビューの内容から正解ラベルのスポットを予測させるタスクで fine-tuning を行い、スポット予測モデルを学習させる。

このようにして学習させたモデルは、入力されたテキストの内容に適したスポットほど分類確率が大きく出力すると考えられる。そこで本手法では、スポット予測モデルにクエリを入力し、出力の分類確率が高いスポット順にランキングを行い、上位のスポットを POI としてユーザに推薦する。これにより、クエリに合わせた POI を推薦することが可能であると考えられる。

また本論文では BERT-base と BERT-review のモデルを比較を行い追加事前学習を行うことによる本手法の有効性を検討する。

4.3 関連レビュー抽出

4.2 節で述べた方法により推薦する POI を決定した後、クエリに対応する説明文生成のために関連するレビューを抽出する。以下では二つの提案手法を述べる

4.3.1 cos 類似度を利用する手法

ここでは cos 類似度を利用した手法について述べる。手法の流れを図 4 に示す。

この手法ではまず、スポット予測モデルにクエリを入力し、POI として推薦された各スポットのレビューを参照する。参照

したレビューをセンテンスごとに区切り、各センテンスとクエリとの \cos 類似度を計算する。類似度が高く出たセンテンスはクエリとの内容が一致している、つまりユーザが目的としている内容が記載されたセンテンスだと考えられる。例えば図4の場合、クエリを「デートにおすすめ」とし、推薦されたPOIの中で「博多駅」のレビューのセンテンス群の中から最も類似度が高く計算されると考えられるものは、クエリと内容が似た「カップルが多かったです」と予想できる。類似度が高く計算されたセンテンスの前後には、理由付けとなる内容が記載されたセンテンスがあると仮定し（図4の場合、前の文章の「イルミネーションがきれいでした。」が対象となる）、その内容こそクエリに対してPOIが推薦された理由であると考え、特徴文として抽出を行う。

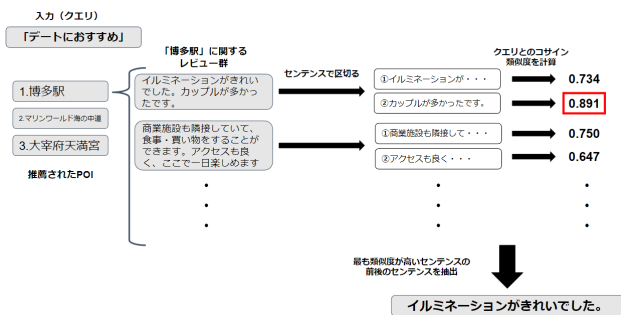


図4 \cos 類似度を用いたレビュー抽出

4.3.2 NSP タスクを利用する手法

本論文ではクエリに関連するレビュー内のセンテンスを抽出するもう一つの方法として、NSP タスクを利用した手法を提案する。手法の流れを図5に示す。

4.3.1 で提案した手法と同様にPOIとして推薦されたスポットのレビューをセンテンスごとに区切る。クエリと各センテンスをペアとしてBERTに入力し、NSPタスクにより入力された2文が連続した文章であるかどうかを予測する。全てのセンテンスの中でクエリと連続する可能性が最も高いと予測されたセンテンスが、意味的もしくは論理的に整合性がとれている、つまりクエリに対する理由となるセンテンスが抽出できると考えられる。本手法ではペアのセンテンスを入力する際、クエリの語尾に「です。」を付けた(例. クエリ: デートにおすすめ ⇒ BERT への入力: デートにおすすめです。)。これは体言止めなどのようなセンテンスを入力した際、NSPタスクでの精度が大きく落ちると考えられるためである。

また本論文では4.2と同様にBERT-baseとBERT-reviewのモデルを比較を行い追加事前学習を行うことによる本手法の有効性を検討する。

5 利用データセット

本実験では「じゃらん net」に記載されているレビューをデータセットとして用いた。BERT-review モデルを作成する際に用いたデータセットは、各都道府県からレビュー数が100件以上存在するスポットを任意に50スポットずつ選び、選ばれた

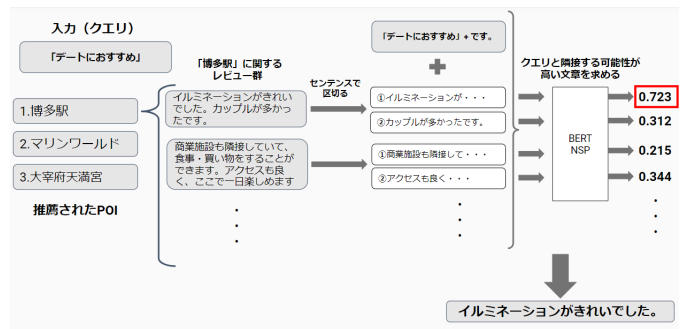


図5 NSP を用いたレビュー抽出

各スポットから最新のレビュー100件を利用した。また、一部の都道府県では100件以上レビューが存在するスポットが50スポット未満の場合もあり、その際は100件以上のレビュー数が存在する全てのスポットをデータセットとして利用した。全てのレビューをセンテンスごとに区切り、合計で747318件のセンテンスをBERT-review作成のコーパスとして用いる。

提案手法でスポット分類モデルを構成するために用いたデータセットはBERT-reviewモデルを作成する際に用いたデータセットの中から、福岡県のレビューを用いる。具体的には50件のスポットに対し各スポット100件のレビュー、合計で5000件のレビューを利用した。スポット予測モデルに関しては、5000件のレビューの内、6割を学習データ、2割を検証データ、2割をスポット予測モデルの精度を確認するためのテストデータに用いた。

6 評価実験

6.1 スポット予測モデル

5.1で述べたデータセットを用いてBERT-base, BERT-reviewそれぞれにレビューからスポットを予測させるタスクでfine-tuningを行った。学習させた二つのモデルにテストデータを用いて精度を確認すると、BERT-baseが74%, BERT-reviewが76%という結果であった。いずれのモデルも高い精度でスポットを予測することができていると考えられる。ユーザのクエリを入力した際に適切なスポットを出力できると考えられるため今回はこの2つのモデルを含めた以下の3つの手法で比較を行う。

- tf 値

レビュー内での単語の出現頻度を算出した値。本研究では各スポットに関する全てのレビューに対して形態素解析を行い名詞のみを抽出する。「デート」や「ドライブ」などユーザがスポットを探す際に目的となる名詞の出現頻度が高いスポットは、対象の目的に適したスポットであると考えられる。そこで目的となる名詞を入力し、tf値の高いスポット順にランキングを行う。tf値は以下の式で求める。

$$tf \text{ 値} = \frac{\text{スポット } s_i \text{ の全レビュー内の名詞 } n_j \text{ の出現回数}}{\text{スポット } s_i \text{ の全レビュー内の全ての名詞の出現回数の和}} \quad (2)$$

- BERT-base

4.2で提案した手法をBERT-baseを用いて説明を行ったもの。

- BERT-review (提案手法)

4.2で提案した手法をBERT-reviewを用いて説明を行ったもの。

今回データとして用いた福岡県の50件のスポットに対して、「デート」「ドライブ」「家族で出かける」「写真撮影」の4つの目的にどれほど適したスポットであるかを{1: 適切である(5点), 2: どちらかといえば適切である(4点), 3: どちらでもない(3点), 4: どちらかといえば適切でない(2点), 5: 適切でない(1点)}の5段階で回答してもらった。また回答者が該当のスポットを知らない場合は{6: 該当のスポットを知らない}と回答してもらった。回答者の総数は25名である。

本論文では各手法の整合性を図る手法としてスピアマンの順位相関係数を用いる。まずアンケートの結果から、それぞれの目的に対する各スポットの適切度を計算する。適切度の求め方は、目的に対するスポットの適切性を回答したスコアを合計し、選択肢(1)~選択肢(5)のいずれかで答えた人数で割ることでスコアの平均を計算している。適切度の値は1~5となる。適切度の高い順にランキングを行い、各スポットの順位を得る。次に各手法に対して目的を入力し、出力されたランキングの順位を得る。

各手法での順位とアンケート結果による順位の相関係数を計算することで、各手法の整合性を確認する。スピアマンの順位相関係数の値ほど手法の整合性が高いと判断する。

6.2 関連レビュー抽出

推薦されたPOIに関してクエリに対応するレビューが抽出できているか、4.3.1と4.3.2、それぞれで提案した手法で実験を行い結果を比較する。本論文では4.3.2での結果を用いて、各POIに関するレビューが正しく抽出できているかを著者が主観的に判断する。

7 実験結果

7.1 スポット予測モデル

各手法とアンケート結果によるランキングの順位相関係数の結果を図6から図9に示す。グラフでは各手法に対して大括弧内の文字を入力とした際の結果を示している。

まず本論文での提案手法であるBERT-baseとBERT-reviewの手法を比較する。いずれのクエリにおいても、BERT-reviewを利用した手法がBERT-baseと比べて強い正の相関があることが分かる。特に図8ではクエリで「ドライブにおすすめ」と入力した際、BERT-baseでは若干負の相関が見られるのに対して、BERT-reviewでは0.5817と強い正の相関が見られる。これらの結果から、スポット予測モデルを学習させる際に、レビューをデータセットとして用いた追加の事前学習を行うことが有効であると考えられる。

次にtf値による手法とBERT-reviewによる手法を比較する。tf値を用いた手法では全てのスポットに対してランキングを行うことができない場合がある。図9のように「写真撮影」

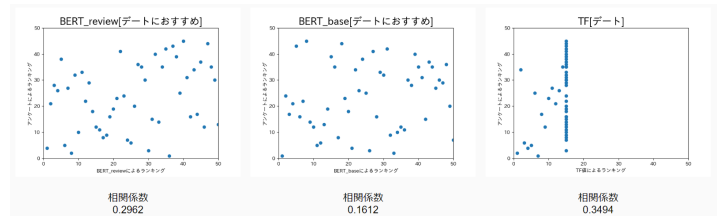


図6 結果例1

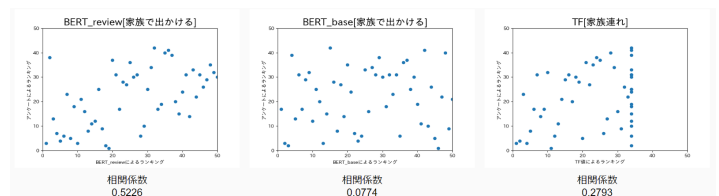


図7 結果例2

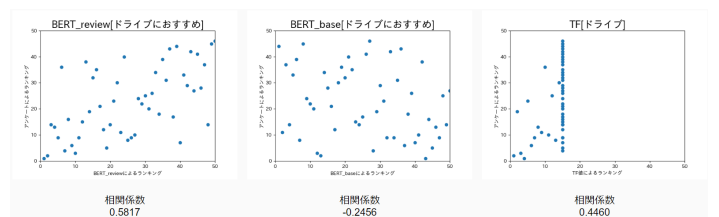


図8 結果例3

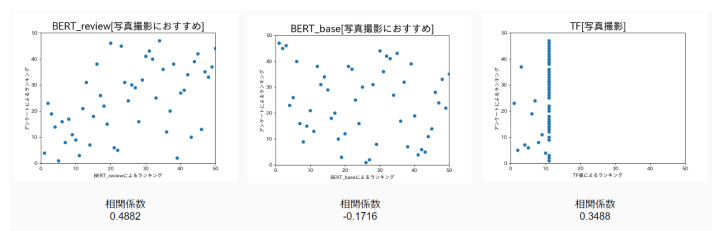


図9 結果例4

という単語のtf値を求める場合、レビュー内に入力の単語が含まれているスポットは10件しか存在せず、残りのスポットはtf値の値が0となり同じ順位としてランキングされる。しかし、BERT-reviewを用いた提案手法では全てのスポットがランキングで順位付けされ、アンケート結果との正の相関も見られるため、ユーザのクエリに対するPOIとして推薦を行う際に有効な手法であると考えられる。

7.2 関連レビュー抽出

本節ではBERT-reviewにスポット予測タスクを学習させたモデルにクエリを入力し、上位にランキングされたスポットに関するレビューの抽出を行った結果を述べる。

7.2.1 cos類似度を利用する手法

表2表3にcos類似度を用いたレビュー抽出の例を示す。表では1列目にスポット名を、2列目に抽出したレビューを、3列目にクエリとの類似度が最も高い値となったセンテンスを記載している。

表2では「皿倉山」のように「デートにおすすめ」という内容に類似したセンテンスがcos類似度の最も高い値となってお

り、抽出したレビューも「北九州エリアの絶景が見られるところといえばこの場所です。」と POI の推薦理由として妥当性があるセンテンスを抽出できている。一方で3か所とも cos 類似度が最も高く出たセンテンス内に「おすすめ」という単語が記載されており、「筑紫野 天拝の郷」と「鷺尾愛宕神社」はいずれもクエリの内容に適さないセンテンスが抽出されている。これは cos 類似度を計算する際にクエリの内容ではなく、「おすすめ」といった単語の一致度に大きく影響を受けていると考えられる。

また表3では「海の中道海浜公園」のように「何回も行ってありますが、子供が小さいうちは何回でも行けます。」といった POI 推薦の理由にならない抽象的な内容や、「チャンネルシティ博多」の「歩いてすぐなので行きました」といったスポットとは関係のない内容のセンテンスが抽出されている。

これらの結果からクエリと関連するレビュー抽出を行う際に cos 類似度は適切でないと考えられる。

7.2.2 NSP タスクを利用する方法

表4表5に cos 類似度を用いたレビュー抽出の例を示す。表4の「皿倉山」「鷺尾愛宕神社」や表5の「キリンビール福岡工場」のようにクエリに関連した内容のレビューが抽出されていることが分かる。また他のスポットに関するレビューもクエリと直接的に関連がなくとも、対象のスポットの特徴を表す内容のレビューが抽出されている。これらの結果から、クエリに関するレビューを抽出する手法として NSP タスクを用いた手法が有効であることが分かった。この手法を踏まえて今後は、クエリに適した内容かつスポットの特徴を示したレビューを抽出する制度を高めるとともに、抽出したレビューをもとに特徴文生成に取り組む。

8 ま と め

本論文では膨大に存在するオンラインレビュー情報からユーザのクエリに応じて POI を推薦し、その推薦理由となるレビューを抽出する手法を提案した。POI 推薦のためのスポット予測モデルでは日本語事前学習済みモデルに対しレビューをコーパスとした MLM を追加学習させたモデルを fine-tuning させることで高い精度を得られることが示された。レビューの抽出に関しては NSP を利用することでクエリに適したレビューを抽出することが有効な手法の一つであることが分かった。今後はレビュー抽出の手法に関しても被験者実験を行い精度を確認するとともに、より具体的に特徴文を示すための手法を考案に取り組む。

9 謝 辞

本研究は JSPS 科研費 19H04219 の助成を受けたものです。

文 献

- [1] Md.Ashraful Islam, Mir Mahathir Mohammad, Sarkar Snigdha Sarathi Das, Mohammed Eunus Ali, “A survey on deep learning based Point-of-Interest (POI) recommendations” *Neurocomputing*, 472, 306-325, 2022
- [2] Xiaojiao Hu, Jiajie Xu, Weiqing Wang, Zhixu Li, An Liu

- “A graph embedding based model for fine-grained POI recommendation” *Neurocomputing*, 428, 376-384, 2021
- [3] Carl Yang, Lanxiao Bai, Chao Zhang, Quan Yuan, Jiawei Han “Bridging Collaborative Filtering and Semi-Supervised Learning:A Neural Approach for POI Recommendation” *KDD*, 1245-1254, 2017
- [4] Yaxue Ma, Jin Mao, Zhichao Ba, Gang Li, “Location recommendation by combining geographical, categorical, and social preferences with location popularity” *Information Processing and Management* 57, 2020
- [5] Pengpeng Zhao, Anjing Luo, Yanchi Liu, Jiajie Xu, Zhixu Li, Fuzhen Zhuang, Victor S. Sheng, Senior Member, IEEE, and Xiaofang Zhou, Fellow, IEEE “Where to Go Next: A Spatio-Temporal Gated Network for Next POI Recommendation” *IEEE TRANSACTIONS ON KNOWLEDGE AND DATA ENGINEERING*, VOL 34, NO 5, MAY 2022
- [6] Yongfeng Zhang, Guokun Lai, Min Zhang, Yi Zhang, Yiqun Liu, Shaoping Ma, “Explicit Factor Models for Explainable Recommendation based on Phrase-level Sentiment Analysis” In *Proceeding of the 37th international ACM SIGIR conference on Research & development in information retrieval*, 83-92, 2014
- [7] Lei Li, Yongfeng Zhang, Li Chen, https “Personalized Transformer for Explainable Recommendation” In *Proceedings of the 59th Annual Meeting of the Association for Computational Linguistics and the 11th International Joint Conference on Natural Language Processing (Volume 1: Long Papers)*, 4947-4957, 2021
- [8] Shijie Geng, Shuchang Liu, Zuohui Fu, Yingqiang Ge, Yongfeng Zhang, “Recommendation as Language Processing (RLP): A Unified Pretrain, Personalized Prompt & Predict Paradigm (P5)” *arXiv preprint arXiv:2203.13366*, 2022
- [9] Yingqiang Ge, Juntao Tan, Yan Zhu, Yinglong Xia, Jiebo Luo, Shuchang Liu, Zuohui Fu, Shijie Geng, Zelong Li, Yongfeng Zhang, “Explainable Fairness in Recommendation” *arXiv preprint arXiv:2204.11159*, 2022
- [10] Lei Li, Yongfeng Zhang, Li Chen “EXTRA: Explanation Ranking Datasets for Explainable Recommendation” *SIGIR 2021 - Proceedings of the 44th International ACM SIGIR Conference on Research and Development in Information Retrieval*, Association for Computing Machinery, Inc, 2463-2469, 2021
- [11] Jacob Devlin, Ming-Wei Chang, Kenton Lee, Kristina Toutanova “BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding” In *Proceedings of the Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies*, pp.4171-4186, 2019
- [12] Jinhyuk Lee, Wonjin Yoon, Sungdong Kim, Donghyeon Kim, Sunkyu Kim, Chan Ho So, and Jaewoo Kang “BioBERT: a pre-trained biomedical language representation model for biomedical text mining” *Bioinformatics*, Volume 36, No.4, pp.1234-1240, 2020
- [13] 仁木裕太, 坂地泰紀, 和泉潔, 松島裕康, “再事前学習した BERT を用いた金融文書中の因果関係知識有無の判別”, *人工知能学会全国大会論文集*, 2020.

クエリ：「デートにおすすめ」

	抽出したレビュー (cos 類似度が高く出た文章に隣接した文章)	クエリとの cos 類似度が最も高く出た文章
皿倉山	北九州エリアの絶景が見られるところといえばこの場所です。	特に夏場でも涼しいのでデートにもお勧めです。
筑紫野 天拝の郷	子供たちも大喜びでした!	家族連れにおすすめです!
鷲尾愛宕神社	愛宕山に登ることでこの神社にお参りを果たせます。	山自体は普通に登ることができて難所ではないのでおすすめです。

表 2

クエリ：「家族で出かける」

	抽出したレビュー (cos 類似度が高く出た文章に隣接した文章)	クエリとの cos 類似度が最も高く出た文章
海の中道海浜公園	何回も行ってますが、子供が小さいうちは何回でもいけます。	また、家族みんなで行く計画をたてます。
麒麟ビール福岡工場	春はポピー、秋はコスモスの時期になると必ず行っています!	毎年家族で見に行くのが楽しみです。
キャナルシティ博多	歩いてすぐなので行きました。	近くのホテルに宿泊。

表 3

クエリ：「デートにおすすめ」

	抽出したレビュー
皿倉山	ロープウェイで上の方まで行くので、夜景が凄く綺麗でカップルにおすすめです。
筑紫野 天拝の郷	温泉施設も良いのですが、バイキングがおすすめです。
鷲尾愛宕神社	デートスポットで人気なのでカップルがたくさんいました。

表 4

クエリ：「家族で出かける」

	抽出したレビュー
海の中道海浜公園	散歩に行くといい運動にもなるし、季節によって色々なお花が咲いてとても良いです。
麒麟ビール福岡工場	ビール工場見学なので大人ばかりかと思っていたら子ども連れの家族も多くて驚きましたが、試飲でジュースもあるし親はビールも飲めるので見ていて納得。
キャナルシティ博多	時間はわかりませんが、噴水ショーなども行っており、音楽もかかるとも迫力があります。

表 5