# マルチモーダル SNS データを用いた イベント情報推薦システムにおける GPT 活用

大本 詩織 (指導教員:小口 正人)

### 1 はじめに

現在、ソーシャルネットワーキングサービス(以下、 SNS とする) 上には、その手軽さから、固定的なメディ アに掲載されていないようなローカルイベントや開催 中のイベントをはじめ、大小様々な規模のイベントに 関する情報が投稿されている。加えて、SNS では投稿 からそのユーザの情報も得ることができる. 近年, デー タの収集や分析技術は飛躍的な進歩を遂げており、様々 な情報源から膨大なデータを処理および分析すること が可能になっている. そこで、SNS 上の膨大なイベン ト情報をサーバ側に収集した上で、ユーザの SNS の投 稿や検索履歴などから趣味趣向を推測し、適切なイベ ントを推薦するシステムを提案する. 本研究では、SNS の一種である X (旧 Twitter) からイベント情報を取 得した上で、近年飛躍的な進歩を遂げている大規模言 語モデルの内の一つである ChatGPT を使用すること で、マルチモーダルな SNS データから趣味趣向の推測 を行い、加えてイベント推薦や回答精度の向上にも活 用することで、効率的なシステムの構築を行うことが できるか検証する.

### 2 提案システムの概要

本研究で提案するイベント推薦システムの概要を図1に示す.スマホの検索履歴や SNS の投稿など,本人のデバイスから得られる情報を ChatGPT に与えて,ユーザの趣味趣向を推測させる.次に,工藤ら [1] が提案する手法を参考にし, X から取ってきたイベントデータのリストを与えて,その中から取得した趣味趣向をもとにイベントを選択させ,回答精度の向上も試みた上で,ユーザに適切だと考えられるイベントを推薦結果として回答させる.



図 1: 提案システムの概要

### 3 ユーザの趣味趣向の取得

### 3.1 情報源によるユーザの趣味趣向の比較

様々な情報源から取ってきたデータを与え, それぞれから取得できるユーザの趣味趣向にどのような違いが現れるか検証した.

まず、Xから無作為に何人かのユーザを選び、プロンプトに十分な数のポストと趣味趣向を推測するよう指示を与えた. その結果、例えば、野球の試合や甲子園について投稿しているユーザの場合、「スポーツ:特に野球に関心がある」といった回答が得られ、ポストを

投稿しているユーザの趣味趣向として適切だと思われる回答を得ることができた.一方で、ポストで少し触れただけの内容も趣味趣向として挙げられてしまうといった問題も見られた.例えば、「スマホ買い替え時かな.iPadも欲しいな.」という投稿から、「スマートフォンやiPadについて触れており、デジタルデバイスに興味があることがわかります」という回答が得られるなど、実際の本人の趣味趣向とずれてしまう可能性があることが分かった.

次に、Instagram の場合の検証を行った。Instagram の場合も、無作為にユーザを選び、投稿画像と投稿文を与えて趣味趣向を推測させた。結果、適切に趣味趣向を取得することができたが、Instagram を趣味趣向の取得に使用する問題点として、ほとんどのユーザが旅行や外出などのイベントや人物、食べ物に関する投稿を行なっており、そもそもの投稿内容が偏る傾向にあるため、そこから得られる趣味趣向にも偏りが見られた。

さらに、スマホの検索履歴を与えた場合の検証も行った。statcounterによる統計 [2]によると、日本のスマホのwebブラウザのシェアは、safariの割合が一番高くなっているため、ここではsafariの履歴を与えることを考えた。SQLコマンドを使用してSafariの履歴を~/Library/Safari/History.dbからCSV形式で取得し、プロンプトに与えてユーザの趣味趣向を推測させた。その結果、適切な回答が得られることもある一方で、図2のように、与えた検索履歴によってはユーザの趣味趣向として不適切な回答が得られる場合もあり、有用な検索履歴を与える必要があるということが分かった。

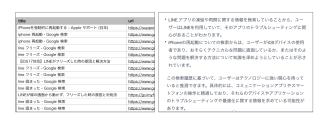


図 2: 与えた検索履歴の一部 (左) と、それを元に趣味 趣向を推測させた結果 (右)

このように、いずれの情報源を選択した場合でも、それぞれ問題点があることが分かった.

### 3.2 同一のユーザに対する趣味趣向の比較

次に、同一のユーザの X, Instagram、検索履歴の情報を与えてそれぞれ趣味趣向を取得したところ、同じユーザであるにもかかわらず、得られる趣味趣向に違いが見られた。一方で、一部共通して見られる趣味趣向もあった。そこで、重複している趣味趣向はより信憑性が高く、適切な趣味趣向である可能性が高いと考え、図3aのように優先度を高く回答するよう指示を与えたところ、図3bのような出力結果が得られ、実際のユーザ

の趣味趣向と比較すると、より高い精度でユーザの趣 味趣向を取得することができていることが分かった.

X、 Instagram、検索履歴のいずれか l つのみで見られる趣味趣向の優先度は1で す。XとInstagram、検索履歴とXなど、どれか2つで重複している趣味趣向の優 先度は2です。3つ全てに共通する趣味趣向の優先度は3です。優先度を考慮し て、このユーザの趣味趣向を推測してください。

### (a) 与えたプロンプト

#### ChatGPT

- . 優先度3(全て共通)
- ・ 歴史:X, Instagram, 検索履歴の全てで歴史への関心が見られます。
- 2. 優先度2 (2つで共通)
- ・ 音楽:XとInstagramで音楽やライブイベントへの関心が見られます。
- 3. 優先度1 (1つでのみ見られる)
  - 書道:Xでのみ見られます。
  - スポーツ:Xでのみ見られます。
  - テーマパーク・アトラクション: Instagramでのみ見られます。
  - 旅行・探検:検索履歴でのみ見られます。

以上のことから、歴史への興味が最も高く、音楽も重要な趣味の一つです。また、書道 スポーツ、テーマパーク、旅行にも興味があると考えられます。

#### (b) 出力結果

図 3: 共通する趣味趣向の優先度を考慮した場合

### 趣味趣向に基づいたイベントの推薦

趣味趣向をもとにしたイベント推薦の検証を行った. ここではあらかじめ X からイベント名をとってきて リストとしてまとめたデータを ChatGPT に渡し、そ の中からユーザの趣味趣向に合うイベントを推薦する よう指示した. イベントデータは, 2021 年の 11/20 か ら 11/24 において, 原宿, 渋谷, 新宿, 赤坂, 代々木, 池 袋、六本木の7種のキーワードで検索して得られたイ ベントのリストを使用した. 結果の一例を図4に示す。 図4を見ると、「音楽が好き」だという趣味趣向を与 えた場合には音楽に関連するイベントが推薦されてお り、適切にイベント推薦を行うことができるというこ とが分かる.



図 4: 与えたイベントデータの一部 (左) と, 趣味趣向 を与えてイベントを推薦させた結果(右)

#### 回答精度の向上 5

しかしながら、4を何回か試す中で、ハルシネーショ ンが発生するという問題が生じた. ハルシネーション とは、LLM の出力が現実の事実やユーザ入力と矛盾す る現象のことを言う. ChatGPT が、イベント名から イベントの内容を推測して推薦しているため、イベン ト内容によっては、実際とは全く異なる内容をでっち 上げて推薦する場合があった. 加えて, 実際には存在 しないイベントが推薦される場合もあった。そこで、 ChatGPT にあらかじめ備わっている web ブラウジン グ機能を使って回答精度の向上を試みた. web ブラウ ジング機能とは、ChatGPT がリアルタイムでウェブ検 索を行い、最新の情報や特定のトピックに関するデー タを取得できる機能である. 回答精度の向上を試みた 際のプロンプトと、それに対する回答を図5に示す.虚 偽の内容を含む元々の推薦結果をプロンプトに与え, その上で、以下のような指示を加えた.

「これらのそれぞれについて、以下を実行してくだ さい.

step1: 実際にイベントが存在するか調べる.

step2: イベント推薦理由を見て、 そのイベントの実 際の内容とあっているか調べる」

その結果、図5の右図のように結果を改善することが できた.



図 5: 回答精度の向上を試みた際のプロンプト (左) と それに対する回答(右)

## まとめと今後の課題

本研究では、SNS データや検索履歴などのマルチモー ダルなデータからユーザの趣味趣向を取得し、イベン トを推薦するシステムにおける、ChatGPT の活用の 可能性について検討した. ユーザの趣味趣向を取得す る際、それぞれの情報源に問題点が見られたが、それら の情報源を組み合わせることでより適切に趣味趣向を 取得することができることが分かった. また, web ブ ラウジング機能を使ってイベント推薦の結果を改善す ることができた. 今後は、ChatGPT が趣味趣向とは いえないものまで趣味趣向として回答してしまうなど といった問題が見られたため、より推薦の精度を高め ていきたい.

また、本研究ではユーザの趣味趣向の取得、イベント 推薦,回答精度の向上の3パートに分けて、それぞれの 内容について手動でプロンプトに指示を与えることで 検証を行ったが、今後はこれら一連の流れのシステム 化も行なっていきたい.

### 参考文献

- [1] 工藤瑠璃子, 榎美紀, 中尾彰宏, 山本周, 山口実靖, 小口正人: SNS データを用いた場所と時間を考慮 するイベント検索手法の提案と評価. マルチメディ ア, 分散, 協調とモバイル (DICOMO2018) シン ポジウム, pp. 1658-1664, 2018.
- [2] Mobile Browser Market Share Japan(statcounter). https://qs.statcounter.com/browser-marketshare/mobile/japan/, Dec 2021 - Dec 2022.