

実空間行動範囲の店舗属性分析と潜在的興味推定に基づく広告推薦方式

大村 貴信[†] 鈴木 健太^{††} Panote Siriaritya^{†††} 栗 達^{††††} 河合由起子^{††††}

中島 伸介^{††††}

[†] 京都産業大学 コンピュータ理工学部 〒603-8555 京都府京都市北区上賀茂本山

^{††} 京都産業大学大学院 先端情報学研究科 〒603-8555 京都府京都市北区上賀茂本山

^{†††} 京都工芸繊維大学 情報工学・人間科学系 〒606-8585 京都市左京区松ヶ崎橋上町

^{††††} 京都産業大学 情報理工学部 〒603-8555 京都府京都市北区上賀茂本山

E-mail: †g1744249@cc.kyoto-su.ac.jp, ††i2086060@cc.kyoto-su.ac.jp, †††spanote@kit.ac.jp,

††††{lida,kawai,nakajima}@cc.kyoto-su.ac.jp,

あらまし スマートフォン向けの広告サービスは年々増加傾向にある。しかしながら、従来のWeb広告推薦手法の多くは、検索キーワードや閲覧履歴に基づいたキーワードマッチング等、ユーザの明示的行動の分析をベースとする手法であり、潜在的な購買者層に対して効果的に広告を推薦するには限界がある。そこで我々は先行研究においてWeb閲覧履歴を利用したユーザの潜在的興味分析に取り組み成果をあげてきた。一方、ユーザの潜在的興味はWeb空間での行動履歴だけでなく実空間の行動履歴にも現れると考え、本稿ではWeb空間だけでなく実空間行動範囲の店舗属性分析と潜在的興味推定に取り組んだので報告する。具体的には、ジオタグ付きツイートとユーザ行動ログデータからユーザの行動範囲を抽出し、OpenStreetMap(OSM)データを用いてこの行動範囲の属性分析を行うことで、実空間行動履歴の特徴抽出を行う。さらに、ある特定のスポットに訪れるユーザモデルを構築することで、対象ユーザへの広告推薦への可能性を議論する。

キーワード Web広告、ユーザプロファイリング、行動分析、潜在的興味分析

1. はじめに

スマートフォン向けの広告サービスは年々増加傾向にある。しかしながら、従来のWeb広告推薦手法の多くは、検索キーワードや閲覧履歴に基づいたキーワードマッチング等、ユーザの明示的行動の分析をベースとする手法であり、潜在的な購買者層に対して効果的に広告を推薦するには限界がある。そこで我々は先行研究においてWeb閲覧履歴を利用したユーザの潜在的興味分析に取り組み成果をあげてきた[1]。一方、ユーザの潜在的興味はWeb空間での行動履歴だけでなく実空間の行動履歴にも現れると考え、本稿ではWeb空間だけでなく実空間行動範囲の店舗属性分析と潜在的興味推定に取り組んだので報告する。具体的には、ジオタグ付きツイートとユーザ行動ログデータからユーザの行動範囲を抽出し、OSMデータを用いてこの行動範囲の属性分析を行うことで、実空間行動履歴の特徴抽出を行う。さらに、ある特定のスポットに訪れるユーザモデルを構築することで、対象ユーザへの広告推薦への可能性を議論する。

本研究は一般にジオターゲティングと言われる手法の一つに位置付けられると考える。ジオターゲティングとはユーザの位置情報を利用したマーケティング手法である。ユーザの現在地や居住地区に合わせた広告推薦が可能であり、Web広告を通じて物理空間に存在する実店舗への来店に繋げるという魅力がある。また、ユーザの行動ログを取得することでユーザの興味・関心が高い対象を推定することもできる。ただし、従来のジオ

ターゲティングの多くは、基本的に物理的な位置情報を利用するものであるのに対して、提案手法ではユーザの実空間での行動に対する意味的な分析を併せて行うものであり、独自性・新規性は高いと考えている。

本稿の構成は以下の通りである。2章では関連研究を紹介する。3章では提案手法について詳細を説明する。4章では実験の条件、結果、考察について述べる。最後に5章でまとめを記述する。

2. 関連研究

以下に、Web広告および行動予測に関連した研究について述べ、本研究との差異を示す。

内野らはユーザが次に見たい情報を予測し、それに関する広告を配信するWeb広告配信システムをkMERおよびマルコフモデルを応用した研究を行なっている[2]。本研究ではユーザが次に見たい情報を予測するのではなく、ユーザの実空間での行動分析に基づき、広告を推薦するか否か決めることを目的としている。

小河らは消費者の購買行動時における情報探索行動を考慮し、消費者が必要とする商品情報とデザインおよびメッセージを個人に合わせた、インターネット広告の構成手法を提案している[3]。本研究でデザインやメッセージをユーザに合わせるのではなく、ユーザの普段の行動から訪れるであろうPOIを予測し、広告を推薦することを目的としている。

倉島らはユーザの過去の行動履歴から次の行動を予測し、推

薦するためのジオトピックモデルを提案している[4]. この研究ではユーザに向けて興味のある場所の推薦を行なっているが, 本研究では興味のある場所の広告を推薦する. 広告を推薦することでその場所で行われているイベントやキャンペーン情報も同時に推薦することができると考える.

久松らは閲覧行動パターンを考慮した購買予兆を発見するモデルをロジット・モデルを基に作成している[5]. この研究ではユーザの閲覧行動から購買予兆を発見し, そのタイミングで広告を表示する研究を行なっているが, 本研究ではWeb空間での行動分析だけでなく, 実空間での行動分析に基づき, 広告を推薦するか否か決める目的としている.

Kan らはユーザモデリングのためにユーザごとに記憶されたシーケンシャルパターンによる生涯シーケンシャルモデリングのためのHierarchical Periodic Memory Networkを提案している[6]. 本研究ではユーザモデリングを行うだけでなく, 作成したユーザモデルを元に別のユーザが潜在的に興味のある対象を推定し, 広告を推薦することを目的としている.

Stamatina らはWeb広告キャンペーンを半自動で開発するというコンテキストでマルチキーワード推薦システムの提案を行っている[7]. この研究では広告キャンペーンのためのキーワードを推薦するシステムの提案を行っているが, 本研究では潜在的に興味のある対象の広告を推薦するシステムの提案を行う.

Xueqin Sui らはソーシャルメディア上で都市レベルの位置情報を検出し, ユーザの移動軌跡を識別する新しい手法を提案している[8]. 位置情報の検出には2種類の方法がある. 投稿内容に明示的/暗示的に位置情報が含まれていればその情報を利用する. 投稿内容に含まれていなければ, 友人たちと同一の投稿を行う傾向を利用し, 友人たちの投稿の位置情報を検出, 利用することでユーザの位置情報を推測する. この位置情報と投稿の文脈, 位置情報間の移動時間の両方に応じて平滑化することでユーザの移動軌跡を構築している. 本研究では都市レベルではなく, より小さいエリアでユーザの行動を分析し, 広告を推薦するか否か決めることを目的としている.

Chuang Song らは従来のPOI推薦の欠点である「ユーザベース強調フィルタリングではユーザの好みが十分に考慮されない」「地理的な影響力をモデル化する場合, 地理的特徴が深く検討されていない」という2つの問題を解決するための新しいPOI推薦アプローチを提案している[9]. 正規化されたチェックイン頻度を用いた行列の構築とユーザのチェックインPOIをいくつかの地域に分けて地域特徴と統合してPOI推薦を行っている. 本研究ではPOIを推薦するのではなくユーザが訪れる場所を予測し, 広告を推薦することを目的としている.

Qiang Gao らは, POI予測のためのユーザチェックイン行動の連続パターンをキャプチャする, VANext (Variation Attention based Next)を提案している[10]. 提案されたモデルは, 個々の動きのパターンをより効率的にキャプチャできるという点で, 広く使用されているRNNよりも優れている. 本研究ではPOIを予測するだけではなく, 広告を推薦することを目的としている.

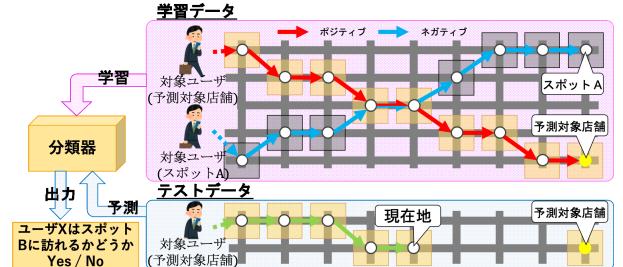


図1 提案手法のシステム概要



図2 提案システムの推薦の例

Xin Li らはロケーションベースソーシャルネットワークで収集されたチェックイン記録が疎であることからPOI推薦が困難であるという問題を解決するために文脈的特徴(時間帯, 曜日, 場所のカテゴリなど)から学習したパーソナライズされた潜在行動パターンを活用し, レコメンデーションの効果を向上させる2種類のモデルを提案している[11]. 全ユーザに対して固定的なパターン分布を学習するGPDMと, ユーザごとにパーソナライズされたパターン分布を学習するPPDMの2種類のモデルを使用することでチェックイン記録の時間の短さや距離に関係なく推薦を行うことができる. 本研究ではPOIを推薦するのではなく, 広告を推薦することを目的としている.

3. 実空間行動範囲の店舗属性分析と潜在的興味推定に基づく広告推薦方式

本章では, 実空間行動範囲の店舗属性分析と潜在的興味推定に基づく広告推薦方式の概要を解説し, 特徴抽出, 学習方法, 評価方法について説明する.

図1に提案システムの概要を示す. 本研究では, 各ユーザが予測対象となる店舗(予測対象店舗)に訪れるか否かを学習し, 分類器として作成する, 分類器に未知のユーザXの行動ログを入力したとき, ユーザXが予測対象店舗に訪れるかどうか判定する. ユーザXが予測対象店舗に訪れたユーザと類似したエリア内を行動しており, 予測対象店舗に訪れると分類器に判定された場合, 広告の推薦を行うシステムの開発を目的としている.

図2に提案システムの推薦の例を示す. ユーザが, 日常的に「カフェ」や「猫がいるペットショップ」を訪問しているような場合, このユーザは潜在的には「猫カフェ」にも興味を持つであろうと推定し, 近くの「猫カフェ」の広告を推薦することを目指している.

従来の広告推薦では, 頻繁に利用する店舗やアイテムを推薦

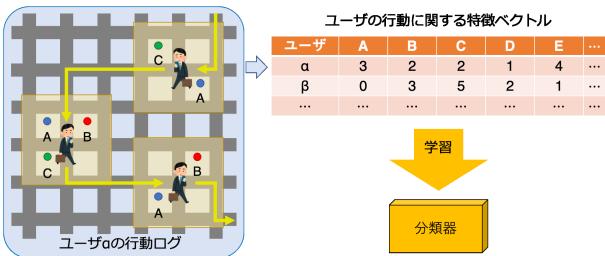


図3 ユーザ行動特徴ベクトルの抽出

したり、性別や年齢に応じて該当しそうな店舗やアイテムを推薦したりといった比較的単純な手法が採用されているが、広告主が購買層を広げるという意味ではその効果が十分とはいえない。一方、提案手法では潜在的な興味分析を行うことで、これまで広告を推薦できなかったようなユーザにも効果的な広告推薦を行える可能性があると考えている。

3.1 データ収集とポジティブ・ネガティブ分類

本研究では2種類のデータを使用する。一つはジオタグ付きツイートデータであり、2016年6月から2020年6月までの4年間に収集された約530万件のデータを採用している。

もう一つは本研究において予測の対象とする店舗(予測対象店舗)の広告配信対象者のユーザ行動ログデータと、予測対象店舗の広告表示ログデータであり、匿名化されたデータを採用している。ユーザ行動ログデータは、2019年のある1ヶ月間に収集された、約2億件のデータであり、広告表示ログデータは、同様に2019年のある1ヶ月間に収集された約20万件のデータである。

これらのデータからポジティブデータの候補とネガティブデータの候補となるデータを抽出する。ジオタグ付きツイートデータでは、あるユーザが予測対象店舗に訪れている場合、ユーザが予測対象店舗に訪れるまでの一定期間のデータを取得しポジティブデータの候補する。また、ユーザが予測対象店舗に訪れていない場合は、ランダムに選択したツイートを呴くまでの一定期間のデータを取得しネガティブデータの候補とした。このときユーザが予測対象店舗に訪れたか否かは、ツイートに含まれている “I'm at starbucks” といった内容から判断した。

また、ユーザ行動ログデータではデータに予測対象店舗に滞在しているか否かの項目、広告表示ログデータには広告をクリックしたか否かの項目がある。そのため少なくとも1回以上、予測対象店舗に滞在しているユーザデータをポジティブデータの候補とし、逆に広告をクリックしておらず、予測対象店舗に滞在していないユーザデータをネガティブデータの候補とする。

3.2 ユーザ行動の特徴抽出

本節では、ユーザの行動の特徴抽出について説明する。実空間でのユーザ行動の特徴を抽出する上で、本研究では OpenStreetMap^(注1)を利用する。OpenStreetMap とは誰でも自由に編集・利用できるオープンな地理情報データである。

図3に、ユーザ行動特徴ベクトルの抽出手法を示す。本研究では、まずジオタグ付きツイート投稿場所やユーザ行動ログ



図4 ユーザの潜在的興味として青のスポットを検出

データの位置情報検出地点(記録地点)の周辺スポット情報をOpenStreetMapから取得する。次に周辺スポット情報から記録地点周辺のスポットの数をユーザ毎、カテゴリ毎にカウントする。このカウントした情報を基に特徴ベクトルの作成を行う。すなわち、ツイートデータやユーザ行動ログデータに含まれる記録地点の特徴を、周辺に存在するスポットの数やそのバランスによって表現しようとするものである。このように記録地点の特徴を周辺に存在するスポットの数やバランスで表現することでユーザがどのようなエリアを移動したのか推定することができると考える。移動したエリアを推定することで別の地域のユーザに向けた広告推薦にも使用できる可能性があると考えている。

3.3 予測対象店舗への潜在的興味推定手法

本節では、3.2 節にて説明した学習データとなる特徴ベクトルを用いた、予測対象店舗への潜在的興味推定手法について説明する。

まず 3.2 節にて抽出した学習データとなる特徴ベクトルと各種クラス分類手法を用いて学習を行い、予測対象店舗を訪れるユーザモデルを分類器として構築する。この分類器に未知のユーザの行動ログをテストデータとして与えた時、予測対象店舗を訪れるかどうかを正しく判定するか検証を行う。この分類器に未知のユーザの行動ログを入力することで、そのユーザが予測対象店舗を訪れるかどうかを推定することが可能になるとを考えている。

4. 潜在的興味推定のためのアルゴリズム別比較実験

本章では提案手法に基づく対象店舗訪問予測精度の評価と非時系列処理アルゴリズムと時系列処理アルゴリズムの比較を目的として様々な学習アルゴリズム、2種類のデータで検証を行ったので報告する。時系列のあるデータを使用しているが、時系列を考慮する場合としない場合でどちらが高い対象店舗訪問予測精度を示すのか検証を行う。また特徴ベクトル作成時のエリアの比較と分析に使用する期間の比較についても同時に検証を行ったので報告する。

4.1 非時系列処理アルゴリズム

4.1.1 エリアサイズの比較

特徴ベクトルを作成するときに記録地点のスポット情報だけでなく、周辺のスポット情報を用いて作成することで、ユーザがどのようなエリアに移動したかを表現することができると考える。そのため、ユーザの現在位置を中心とした正方形のエリア内のスポット情報を検出している。図4にスポット情報検出

(注1) : <https://www.openstreetmap.org/>

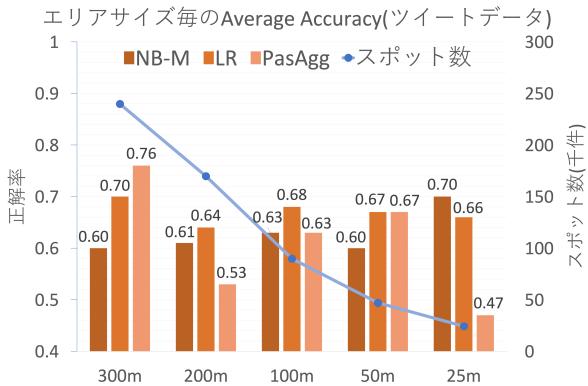


図 5 ジオタグ付きツイートデータのエリアサイズごとの正解率

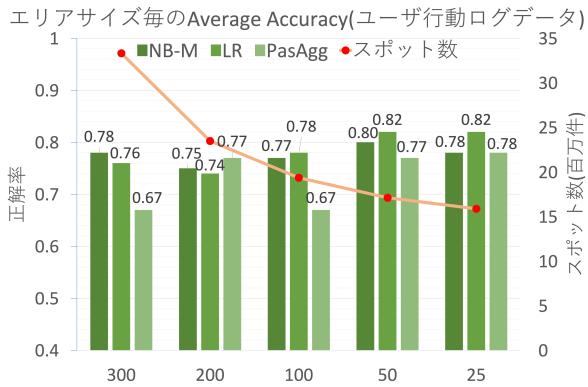


図 6 ユーザ行動ログデータのエリアサイズごとの正解率

の例を示す。このエリアが広すぎるとユーザが実際には行動していないエリアを多く含むこととなり、狭すぎるとユーザの興味を捉るために十分なデータが確保できない可能性があるためエリアの比較を行う。

正方形のサイズは 300m 四方, 200m 四方, 100m 四方, 50m 四方, 25m 四方の 5 種類のデータを用意した。ジオタグ付きツイートデータではポジティブユーザ・ネガティブユーザ共に 149 人、合計 298 ユーザの特徴ベクトルを使用した。これらのユーザは、1 か月に 20~100 件のツイートを投稿している。特徴ベクトルは、一意のユーザ ID とタグ値の 2 つの列で構成されている。

ユーザ行動ログデータではポジティブユーザ・ネガティブユーザ共に 300 人、合計 600 ユーザの特徴ベクトルを使用した。特徴ベクトルは、データを一意に識別できる番号とタグ値の 2 つの列で構成されている。

タグ値は、店舗、学校、建物などの様々なタイプの場所を参照し、ジオタグ付きツイートデータ、ユーザ行動ログデータ共に 3:2 の比率で学習データとテストデータに分割した。前処理後、Naïve Bayes-Multinomial(NB-M), Logistic Regression(LR), Passive Aggressive (PasAgg) など非時系列処理の様々な機械学習アルゴリズムを実装し、正解率で評価を行った。図 5(ツイートデータ)と図 6(ユーザ行動ログデータ)がその結果である。

図 5 では Passive Aggressive の 300m 四方が最も結果が良く、また図 6 ではロジスティック回帰モデルの 50m 四方と 25m

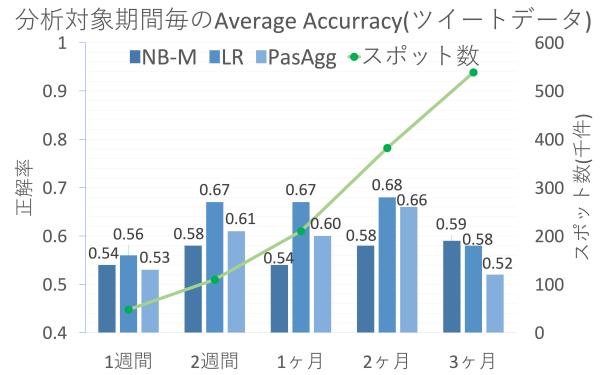


図 7 ジオタグ付きツイートデータの分析対象期間ごとの正解率

四方が最も結果が良い。またユーザ行動ログデータの方が全体的に結果が良い。特徴ベクトルに使用したスポット数がジオタグ付きツイートでは 2 万件から 23 万件であるのに対し、ユーザ行動ログデータでは 1500 万件から 3300 万件であることからジオタグ付きツイートデータはユーザ行動ログデータに比べデータ量が少ないと言える。このことからより多くの情報を考慮できる 300m 四方が最も良いとなったと考えができる。それに対してユーザ行動ログデータではデータ量が十分に確保できているため、50m 四方や 25m 四方の小さいエリアの方がユーザの興味を捉るために適していたと考えられる。これらの結果からデータ量が十分に確保できていれば小さいエリアの方が良いが、データ量が不十分であればより多くの情報を考慮するために広いエリアのほうが良いと考える。またロジスティック回帰モデルや Passive Aggressive の結果が良かったことから 2 種類のデータは線形分離可能であったと考えられる。

表 1 各期間ごとのユーザのツイート数の範囲

	ツイート数の範囲
1 週間	5-25
2 週間	10-50
1ヶ月	20-100
2ヶ月	40-200
3ヶ月	60-300

4.1.2 期間の比較

また、ジオタグ付きツイートデータにおいては分析に使用する期間の比較も行う。分析期間が短いと特徴を捉るために十分な情報が取得できない恐れがあり、長いとユーザの短期的な興味を分析することが困難になると考えられるからである。なおユーザ行動ログデータは 1 ヶ月間のデータのため期間の比較は行っていない。

期間は 1 週間、2 週間、1 ヶ月、2 ヶ月、3 ヶ月の 5 種類のデータを用意した。ポジティブユーザ・ネガティブユーザ共に 120 人、合計 240 ユーザの特徴ベクトルを使用した。正方形の面積は 300m × 300m とし、各期間の間にユーザは、表 1 の件数、ツイートを投稿している。特徴ベクトルは、一意のユーザ ID とタグ値の 2 つの列で構成されている。タグ値は、店舗、学校、建物などの様々なタイプの場所を参照し、データは 3:2 の比率

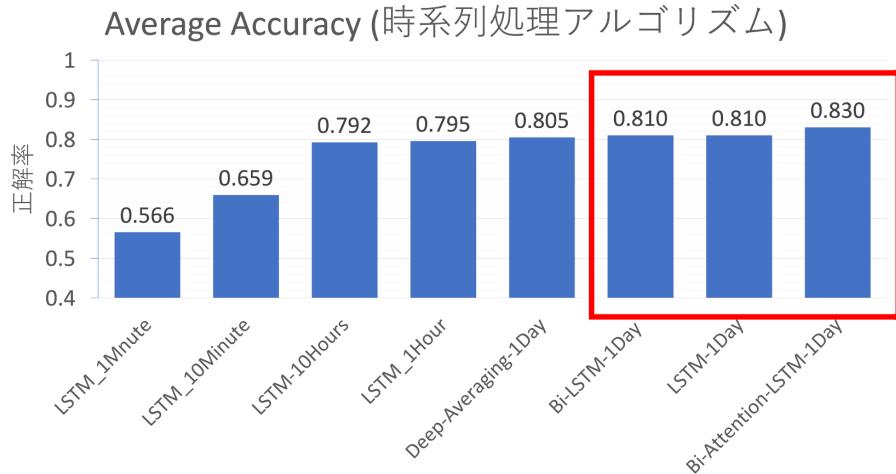


図 8 時系列処理アルゴリズムの正解率

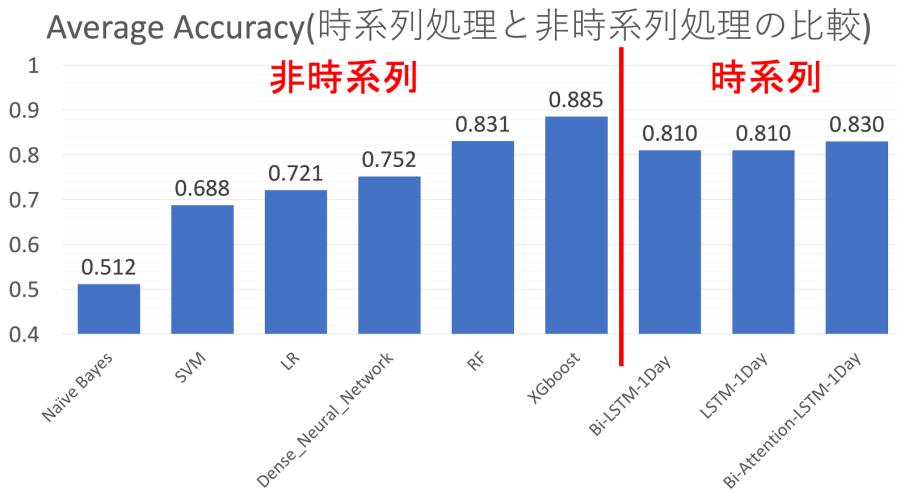


図 9 時系列処理と非時系列処理の比較

で学習データとテストデータに分割した。前処理後、エリアの比較と同様に非時系列処理の様々な機械学習アルゴリズムを実装し、正解率で評価を行った。図 7 がその結果である。

ロジスティック回帰モデルの 2ヶ月の結果が最も結果が良く、また 1 週間、3ヶ月の結果より 2 週間、1ヶ月、2ヶ月の結果が良い。これは 1 週間のデータ量ではユーザの興味を捉えるには不十分であり、3ヶ月のデータではユーザの短期的な興味を捉えるには期間が長すぎたのでは無いかと考える。この結果からジオタグ付きツイートデータにおいてユーザの興味を捉えるには 2 週間から 2ヶ月のデータが必要であると考える。

4.2 時系列処理アルゴリズム

我々の提案手法ではユーザが予測対象店舗に訪れるか否かを予測するため、ユーザの行動を分析している。ユーザの行動を分析する際に時系列の考慮を入れることで結果が良くなるのではないかと考え、実験を行った。

データはユーザ行動ログデータを使用し、エリアは 300m 四方である。一定時間に存在する OSM のタグ値から生成したベクトルを平均化し、時系列データを作成する。時系列データの Window サイズを 1 分、10 分、1 時間、10 時間、1 日とした 5

種類のデータを用意した。ポジティブユーザ・ネガティブユーザ合わせ 847 ユーザの特徴ベクトルを使用した。データは 4:1 の比率で学習データとテストデータに分割し、いくつかの時系列処理アルゴリズムを実装し、正解率で評価を行った。図 8 がその結果である。また図 9 に時系列処理アルゴリズムと非時系列アルゴリズムの比較結果を示す。

図 8 から時系列処理アルゴリズムの中で、僅差はあるが時系列データの Window サイズは 1 日が最も結果が良い。図 9 から非時系列処理アルゴリズムと時系列処理アルゴリズムを比較した場合、非時系列処理アルゴリズムである XGboost の結果が最も良い。しかし、時系列処理アルゴリズムも単一学習器と比べると良い結果を示している。このことから時系列を考慮すべき場合もあることが確認できた。時系列の順番に意味がある場合と意味がない場合があることやユーザ毎・店舗毎において分類手法の得て不得手があることが考えられるため今後も様々な条件で調査を行いたいと考えている。

5. おわりに

本稿では実空間行動範囲の店舗属性分析と潜在的興味推定に

基づく広告推薦方式について提案し、2種類のデータ、様々な学習アルゴリズムで比較検討を行ったので報告した。

特徴ベクトル作成時のエリアに関してデータ量が十分に確保できていれば小さいエリアの方が良いが、データ量が不十分であればより多くの情報を考慮するために広いエリアのほうが良いという結果が得られた。またジオタグ付きツイートデータにおいてユーザの潜在的興味を推定するために必要な期間は2週間から2ヶ月程度であるという結果に至った。

非時系列処理アルゴリズムと時系列処理アルゴリズムを比較した結果、本実験条件においては非時系列処理アルゴリズムであるXGboostが最も良い結果を示したが、対象店舗やユーザによっては他のアルゴリズムが良い結果を示す可能性があるため今後の検討課題とする。

今後についてはポジティブデータとネガティブデータの比率を変更し、AUCで評価を行うことやより広いエリア・より長い期間での比較実験、時系列処理アルゴリズムでのエリアサイズの比較、ジオタグ付きツイートデータでの時系列処理アルゴリズムも検討したいと考えている。また学習データとテストデータで別の地域のユーザデータを使用することで地域特性による差異に関して検証を行いたいと考えている。

謝 辞

本研究の一部は、2020年度国立情報学研究所公募型共同研究（20FP02）の助成を受けたものである。ここに記して謝意を表す。

文 献

- [1] 山口由莉子, Panote Siriaraya, 森下民平, 稲垣陽一, 中本レン, 張建偉, 青井順一, 中島伸介, Web広告推薦のための長期的・短期的興味を考慮したユーザの潜在的興味分析方式, 第10回データ工学と情報マネジメントに関するフォーラム(DEIM Forum 2018) B2-3, 2018.
- [2] 内野英治, 森田博彦, 下野雅芳, Web広告動的配信システムへのマルコフモデルとkMERの応用, 22nd Fuzzy System Symposium(Sapporo, Sept. 6-8, 2006) 6B1-1, 2006.
- [3] 小河真久, 原田史子, 島川博光, 消費者の情報探索行動に着目した広告の内容と表示の個別化, 情報処理学会研究報告 Vol.2010-DBS-150 No.17 Vol.2010-IFAT-99 No.17, 2010.
- [4] 倉島健, 岩田具治, 星出高秀, 高屋典子, 藤村考, 行動範囲と興味の同時推定モデルによる地域推薦, 情報処理学会論文誌データベース Vol.6 No.2 30-41, 2013.
- [5] 久松俊道, 外川隆, 朝日真弓, 生田目崇, ECサイトにおける購買予兆発見モデルの提案, オペレーションズ・リサーチ: 経営の科学, 2013.
- [6] Ren, Kan and Qin, Jiarui and Fang, Yuchen and Zhang, Weinan and Zheng, Lei and Bian, Weijie and Zhou, Guorui and Xu, Jian and Yu, Yong and Zhu, Xiaoqiang and Gai, Kun. Lifelong Sequential Modeling with Personalized Memorization for User Response Prediction. Proceedings of the 42nd International ACM SIGIR Conference on Research and Development in Information Retrieval. 2019.
- [7] Stamatina Thomaidou, Michalis Vazirgiannis, Multiword Keyword Recommendation System for Online Advertising, 2011 International Conference on Advances in Social Networks Analysis and Mining, 2011.
- [8] Sui, Xueqin and Chen, Zhumin and Guo, Lei and Wu, Kai and Ma, Jun and Wang, Guanghui. Social media as sensor in real world: movement trajectory detection with microblog.

Soft Comput. 2017.

- [9] Song, Chuang and Wen, Junhao and Li, Shun. Personalized POI Recommendation Based on Check-in Data and Geographical-Regional Influence. Proceedings of the 3rd International Conference on Machine Learning and Soft Computing. 2019.
- [10] Gao, Qiang and Zhou, Fan and Trajcevski, Goce and Zhang, Kunpeng and Zhong, Ting and Zhang, Fengli. Predicting Human Mobility via Variational Attention. The World Wide Web Conference. 2019.
- [11] Li, Xin and Han, Dongcheng and He, Jing and Liao, Lejian and Wang, Mingzhong. Next and Next New POI Recommendation via Latent Behavior Pattern Inference. ACM Trans. Inf. Syst. 2019.