

ウェブ検索ログを用いたユーザの商品購買行動予測

中田 祐誠[†] 村本 直樹[†] 山本 岳洋^{††} 藤田 澄男^{†††} 大島 裕明^{†,††}

[†] 兵庫県立大学 応用情報科学研究科 〒650-0047 兵庫県神戸市中央区港島南町 7-1-28

^{††} 兵庫県立大学 社会情報科学部 〒651-2197 神戸市西区学園西町 8-2-1

^{†††} ヤフー株式会社 〒102-8282 東京都千代田区紀尾井町 1-3

E-mail: †aa19q507@ai.u-hyogo.ac.jp, aa18c508@ai.u-hyogo.ac.jp, ohshima@ai.u-hyogo.ac.jp,

††t.yamamoto@sis.u-hyogo.ac.jp, †††sufujita@yahoo-corp.jp

あらまし 本研究では、ウェブ検索ログを用いてユーザの商品購入の有無を予測する手法を提案する。EC サイトでのオンラインショッピングが一般化するにつれて、ウェブ検索を通して商品について事前に調べ、EC サイトで実際に商品を購入する行動をとるユーザが増えている。そうしたユーザのウェブ検索のクエリログを分析することで、ユーザが近い将来その商品を購入するかどうかを予測することができると考えられる。そこで本研究では、ユーザのウェブ検索クエリログから、そのユーザが商品を購入するかどうかを予測する手法を提案する。具体的には、商品を購入したユーザと購入していないユーザの過去のウェブ検索クエリログを収集し、分類器を構築する。構築された分類器を分析することで、予測の精度や、予測に必要な期間、購買を示唆する特徴的な語があるか、といったことを検証する。実験ではカメラの購入予測を題材として、Yahoo!検索が提供するウェブ検索のクエリログと、Yahoo!ショッピングにおける商品購買履歴を用いて検証を行った。

キーワード 購買予測, ウェブ検索ログ

1 はじめに

EC サイトでのオンラインショッピングが一般化するにつれて、ウェブ検索を通して商品について事前に調べ、EC サイトで実際に商品を購入する行動をとるユーザが増えている。2017年に広告および市場調査会社 Survata が実施した調査では、オンラインで商品を購入する際に、36%のユーザが最初に検索エンジンで検索を行い、特定の商品の購入を考えていない場合には、49%のユーザが検索エンジンを使用することが報告されている。[4] このように、ウェブ検索は商品購買において重要なプロセスとなっている。

これらの報告は、ユーザがウェブ検索エンジンに投入したクエリのログ、すなわちウェブ検索ログを分析することで、ユーザが近い将来その商品を購入するかどうかを予測できることを示唆している。これまでに、EC サイト上におけるショッピング行動に関する分析 [5] [6] や、ウェブ検索において商品を探るための検索をしたのか、商品の購入をするための検索をしたのかを分類する研究 [3] などがなされているが、ウェブ検索ログから将来的な商品購買行動を予測する研究は、著者らの知る限り行われていない。

そこで本論文では、ユーザのウェブ検索クエリログから、そのユーザが商品を購入するかどうかを予測する手法を提案する。具体的には、商品を購入したユーザと購入していないユーザの過去のウェブ検索クエリログを収集し、分類器を構築する。構築された分類器を分析することで、予測の精度や、予測に必要な期間、購買を示唆する特徴的な語があるか、といったことを検証する。実験ではカメラの購入予測を題材として、Yahoo!

検索が提供するウェブ検索のクエリログと、Yahoo!ショッピングにおける商品購買履歴を用いて検証する。

本研究の構成は以下の通りである。2 節では、オンラインショッピング行動やオンライン購入の予測に関する関連研究を述べる。3 節では本研究での問題定義とデータセット構築について説明する。4 節では提案した手法を説明し、5 節では提案手法に関する実験の評価を行う。最後に 6 節でまとめと今後の課題について述べる。

2 関連研究

本節では、オンラインショッピング行動やオンライン購入の予測に関する関連研究について述べる。本研究のようにウェブ検索ログを用いた購買行動予測を行う研究は散見されないが、将来の予測を行う一連の流れを持つ研究は、多数行われている。

以下、商品購買予測に関連する研究について述べるとともに、本研究について類似する性質を考える。

2.1 商品の購買行動モデル

購買決定プロセス (buyer decision process) は、消費者が商品を購入する一連のプロセスをモデル化したものである [2]。購買決定プロセスには、第 1 段階として問題やニーズの認識があり、第 2 段階に情報検索段階がある。そして第 3 段階に代替製品との比較評価、第 4 段階に購買決定が行われる。購買決定にはこのような 4 つのプロセスがあり、特に、Dewey は購買決定プロセスにおいて情報検索段階は、ユーザが問題や必要性を認識した後、解決策であると感じるものを見つけるために取るべき次のステップと定義している [1]。また、1 節で述べたよう

に、多くのユーザが商品購買に際してウェブ検索を用いることが報告されている [4]。こうした研究は、ユーザの購買行動プロセスに関する情報がウェブ検索エンジン上に表れる可能性が高いことを示していると考えられる。

2.2 EC サイトにおける検索行動分析

これまで、EC サイト上における検索行動分析に関する研究がいくつかなされている [5] [6]。たとえば、Su らは EC サイト上におけるユーザの検索意図を分類し、それぞれの意図に特徴的な行動を明らかにしている。彼らは EC サイト上における検索意図として、すでに想定している商品へのページに到達するための検索である Target Finding、購入する意思はあるものの具体的な商品が決まっておらず、購入すべき商品を決めるための検索である Decision Making、具体的な目的を持たず EC サイト上をブラウジングする Exploration という 3 つの意図を提案している。

本研究ではウェブ検索を対象としており、EC サイト上の検索行動を扱ったものではないが、こうした検索意図に関する検索はウェブ検索においてもなされていると考えられ、ウェブ検索行動から購買行動を予測する上で関連が深いと考えられる。

2.3 ウェブ検索における行動分析

ウェブ検索行動の分析に関する研究は多くなされている。本研究に特に関連の深い研究としては、Guo らの研究があげられる [3]。Guo らはクエリやクリック、マウス操作といった情報から検索タスクを推定するモデルを提案している。その研究のなかで、彼らは検索タスクを (1) 商品の購買のため、商品について調査をするための検索タスク、(2) 実際に商品を購入するための検索タスク、の 2 種類に分類する実験を行っている。

本研究は、ウェブ検索行動と商品の購買行動との関係を扱う、という点で本研究と類似している。一方、Guo らの研究では、与えられた検索タスクが、商品の購買を行うための検索なのかどうかを分類するのに対して、本研究は、検索ユーザが将来的に商品を購入するかどうか、を予測するという点で Guo らの研究とは異なる。将来的な購買行動予測という点では、Guo らの研究における商品について調査するための検索タスクに関するクエリが購買予測に重要な特徴を表すと考えられる。

3 問題定義とデータセット構築

本節ではまず、本研究で取り組む問題設定について述べる。その後、本問題で取り扱うウェブ検索ログおよび商品購買履歴情報について述べ、購買者ユーザと非購買ユーザの選定方法について述べる。

3.1 ウェブ検索ログからの商品購買行動予測

本研究で扱う問題について定義する。いま、ある時刻 T において、ユーザ u が区間 $[T - t_{\text{data}}, T]$ にウェブ検索エンジンに入力したクエリ集合 $Q_u = \{q_1, \dots, q_n\}$ が与えられるとする。このとき、このユーザ u が区間 $[T, T + t_{\text{purchase}}]$ 以内に、商品を購入するかどうかを予測することが本研究で扱う問題である。

すなわち、本研究では、購買予測を 2 値分類問題として定式化した。

本研究では、時刻の単位は日とした。たとえば、ある時点でのユーザの 3 ヶ月間のクエリログを用い、そのユーザがいまから 30 日以内に商品を購入するかどうかを分類する問題を考える。その場合、 $t_{\text{data}} = 90$ 日、 $t_{\text{purchase}} = 30$ 日とすれば、問題として定式化される。

3.2 ウェブ検索ログ

本研究で用いたウェブ検索ログについて説明する。Yahoo!検索が提供するウェブ検索において、2016 年 10 月 1 日から 2017 年 10 月 31 日までの約 1 年間のウェブ検索ログを使用した。本研究では、検索ユーザが投入するクエリに含まれる単語情報のみからユーザの購買行動をどの程度予測できるのかを検証することに注力するため、検索ユーザが投入したクエリのみに着目し、クリック等の情報は用いなかった。

3.3 商品購買履歴

本研究で用いた商品購買履歴について説明する。Yahoo!ショッピングにおける 2016 年 10 月 1 日から 2017 年 10 月 31 日までの約 1 年間の商品購買履歴を使用した。本研究では、購買行動を予測する対象の商品として、デジタル一眼カメラとコンパクトデジタルカメラに関する商品を対象とした。カメラを題材とした理由は、他社モデルとの比較検討がされやすく、また、幅広い年代のユーザに利用されていると考えられるドメインだからである。

3.4 予測対象とする検索ユーザの選定

本節では、本研究で購買行動予測の対象とする検索ユーザの選定方法について述べる。本研究では、カメラについて検索したことがある検索ユーザを対象にし、購買するかどうかを予測する。このような方法をとる理由としては 2 点ある。

1 点目は、手法の実用性を考えた際に、カメラに関心があるユーザの中で、購買するユーザと購買しないユーザを発見することが重要であると考えたからである。カメラに関心があり、将来的な購買行動が高い確率で見込まれるユーザに対しては、商品の購買サイトへ誘導するような広告を提示することで、ユーザの購買行動を促進することができると考えられる。また、カメラについて検索しながら将来的な購買行動が見込まれないユーザに対しては、購買サイトに関する広告ではなく、購買意識を高めるようなウェブサイトに関する広告を提示することで、ユーザの購買行動に繋げることができると考えられる。

2 点目は、カメラを購入する検索ユーザを正例、そうでない検索ユーザを負例としてデータセットを構築してしまうと、分類器が単にカメラに興味があるユーザかそうでないユーザのみを捉えるような分類器となってしまう、購買に関する特徴を捉えた分類器が構築できない可能性があるためである。

以降、本研究であつかう検索ユーザの選定方法について詳しく述べる。

3.4.1 フィルタリングクエリの選定

カメラに関心があり、購買行動予測の対象とする検索ユーザ

を選定するため、本研究ではカメラに関する単語を含むクエリで検索したことがあるユーザを選定した。この、カメラに関する単語を含むクエリを本研究ではフィルタリングクエリと呼ぶ。フィルタリングクエリの選定方法は以下の通りである。

フィルタリングクエリを選定するため、まずはカメラ購入者がどのようなクエリで検索することが多いのかを分析した。具体的には、3.3節で述べた商品購買履歴でカメラに関する購買記録があるユーザに対して、3.2節で述べた1年分のウェブ検索ログを用いてクエリの頻度を集計し、カメラに関する特徴的なキーワードを手で確認した。その結果、“カメラ”というキーワードや、“Canon”や“Nikon”などのメーカー名、“eos”や“pentax”などの型番に関するキーワードがみられた。そこで、「カメラ」というキーワード、メーカー名や型番に関するキーワードを計15件人手で用意し、これらのキーワードを含むクエリをフィルタリングクエリとして定義した。実際、商品購買履歴に購入情報があるユーザのうち、これらのフィルタリングクエリで検索したことがあるユーザは30.9%であった。フィルタリングクエリにより予測の対象とするユーザを絞ることで、カメラに興味を持つユーザから購入するユーザとそうでないユーザに分類することができ、より実用的な分類器が構築できると考えられる。

3.4.2 購入ユーザと非購入ユーザの抽出

今述べたフィルタリングクエリを利用して、カメラに関する商品購買履歴がありかつ、フィルタリングクエリで検索したことがあるユーザを抽出した。本研究では、このフィルタリングクエリを用いて以下の2種類のデータセットを構築した。

(a): 翌日にカメラを購入するかどうかを推定するためのデータセット:

まず、フィルタリングクエリで期間内に検索したことがあるユーザをクエリログより抽出する。その後、期間内にカメラを購入したことがあるユーザを抽出する。さらに、検索回数が少ないユーザを除去するため、1年間の総検索回数が50回以下のユーザを除去した。この操作により、12,037名の購入ユーザを抽出した。本データセットでは、この12,037名のユーザを正例として扱い、各ユーザについてカメラを購入した日の前日から3ヶ月間のクエリログを推定に用いる。また、負例についてはフィルタリングクエリで検索したことがあるユーザで、かつ今回データとして利用した1年間の期間内に商品購買履歴がない検索ユーザから正例と同数のユーザ12,037名を無作為に抽出し、非購入ユーザ、つまり負例として扱った。非購入ユーザのクエリログについては、フィルタリングクエリを含む検索が投入された日から3ヶ月間のクエリログを予測に用いる。このようにして用いたデータセットは、3.1節で述べた問題設定としては、 $t_{\text{purchase}} = 1$ 日として、推定していると捉えることができる。

(b): ある日時から1ヶ月以内に購入するかどうかを推定するためのデータセット:

(a)で述べたデータセットの集め方の問題点は、正例ユーザがカメラの購入日から遡って3ヶ月間、負例ユーザはフィルタリングクエリを検索してから3ヶ月間のクエリログを予測に用い

ており、使用するクエリログの期間が正例と負例によって大きく異なってしまう点である。

そこで、より実用的なデータセットとして、ある日時を対象とし、そこから1ヶ月以内にカメラを購入するかどうかを推定するモデルを作成するためのデータセットを構築した。具体的には、2017年9月1日を対象の日時とし、そこから3ヶ月間遡った区間、すなわち2107年6月1日から2017年9月1日までの3ヶ月間を予測に用いるユーザクエリログとする。まず、年間検索回数が50回以上のユーザで、かつ、6月1日から9月1日の3ヶ月間の中でフィルタリングクエリを使用したことがあるユーザを抽出する。その後、9月1日から10月1日までの1カ月の間に商品を購入したユーザを正例ユーザ、商品を購入していないユーザを負例ユーザとして扱う。この条件をみえず正例ユーザをクエリログから2,500人無作為に抽出するとともに、同数の負例ユーザを抽出した。このようにして用いたデータセットは、3.1節で述べた問題設定としては、 $t_{\text{purchase}} = 30$ 日として、推定していると捉えることができる。また、より実際に分類器を適用する状況に近いと考えられる。

本研究では(a)、(b)ともに3ヶ月という期間のクエリログを予測に用いる。3ヶ月という設定の妥当性については、5節で述べる実験にて検証する。

4 提案手法

本節では、ユーザのウェブ検索ログを用いて購買行動を予測する方法について述べる。本研究では、クエリに含まれる、購買行動に関係がある単語の影響を検証するため、比較的シンプルなモデルによる検証を行う。具体的には、TF-IDFで得られた特徴に基づく分類器を構築する。

4.1 ウェブ検索ログからの特徴抽出

本研究では、ユーザのウェブ検索ログからTF-IDFにより単語の重みづけを行い、ユーザの特徴を作成する。いま、データセット中の全ユーザを $U = \{u_1, \dots, u_N\}$ とする。いま、あるユーザ $u \in U$ のウェブ検索ログを $Q^u = \{q_1, q_2, \dots, q_m\}$ とする。各クエリ $q_i (1 \leq i \leq m)$ は m 個の単語集合 $q_i = \{w_1, w_2, \dots, w_m\}$ として表現される。このとき、ユーザ u のある単語 $w_j (1 \leq j \leq m)$ に対するTFIDFスコア $\text{tfidf}_{u,j}$ を以下の式で計算する。

$$\text{tfidf}_{u,j} = \text{tf}_{u,j} \cdot \text{idf}_j$$

$$\text{tf}_{u,j} = \sum_{q_i \in Q^u} \text{tf}(q_i, w_j)$$

$$\text{idf}_j = \log \frac{\sum_u |Q^u|}{df(w_j)}$$

$\text{tf}(q_i, w_j)$ はクエリ q_i 中に単語 w_j が出現する回数、 $df(w_j)$ は全ユーザ U の全クエリのうち、単語 w_j を含むクエリの数を表す。これにより、語彙のサイズを V とすると、検索ユーザの特徴は V 次元のベクトルとして表現することができる。

4.1.1 分類器

本研究では、購入予測を与えられたユーザが商品を購入する

かどうかの2クラス分類問題として扱う。本研究では以下の2つの手法により分類器を構築する。

- サポートベクタマシン (SVM)
- ロジスティック回帰

両者とも、さまざまな機械学習タスクにおいて広く用いられているモデルである。特に、ロジスティック回帰では構築されたモデル内の各説明変数に対する係数の値をみることで、その変数が予測に与える影響を容易に解釈することができる。これにより、構築した分類器がどの程度購買予測を示唆する単語を捉えられているのかを検証することができると考えられる。

4.1.2 評価指標

評価指標として、正解率、適合率、再現率、 F_1 値を用いた。本研究では購入ユーザを正例、非購入ユーザを負例として扱う。従って、True Positive (TP), True Negative (TN), False Positive (FP), False Negative (FN) は以下の様に定義される。

- TP: 商品を購入すると予測して、商品を購入した
- TN: 商品を購入しないと予測して、商品を購入しなかった
- FP: 商品を購入すると予測して、商品を購入しなかった
- FN: 商品を購入しないと予測したが、商品を購入した

これらを用いると、正解率 (Accuracy), 適合率 (Precision), 再現率 (Recall), F_1 値 (F_1) は以下の式で表される:

$$\text{Accuracy} = \frac{\text{TP} + \text{TN}}{\text{TP} + \text{FP} + \text{TN} + \text{FN}}$$

$$\text{Precision} = \frac{\text{TP}}{\text{TP} + \text{FP}}$$

$$\text{Recall} = \frac{\text{TP}}{\text{TP} + \text{FN}}$$

$$F_1 = \frac{2 \cdot \text{Recall} \cdot \text{Precision}}{\text{Recall} + \text{Precision}}$$

5 実験

本研究で行った実験について述べる。本実験の目的は、以下を明らかにするためである。

RQ1 ウェブ検索ログに基づく購買行動予測の精度はどの程度か。

RQ2 予測に用いる検索ログの期間はどの期間が良いか。

本節では、まず、RQ1 を明らかにするため、3.4.2 節で述べたデータセットに対して、予測を行った結果について示す。その後、RQ2 を明らかにするため、さまざまな期間による予測精度を検証するための実験結果について示す。

5.1 予測精度

まず、3.4.2 節で作成したデータセット (a) に対して分類器を構築し、分類精度を検証した。本データセットは、正例 (購入) ユーザと負例 (非購入) ユーザがそれぞれ 12,037 名で同数となっているデータセットである。分類器には SVM を使い、データセットの 80% を訓練データ、20% をテストデータとしてモデルの学習、精度の検証を行った。SVM のカーネルには RBF カーネルを用い、ハイパーパラメータは訓練データに対して 5 分割交差検証によるグリッドサーチで求め、 $\gamma = 1, C = 10$ を用いた。

表 1 SVM によるデータセット (a) を用いた予測結果

正解率	適合率	再現率	F_1
0.739	0.705	0.821	0.759

表 2 符号が正で係数の値の大きな単語上位 12 件 (データセット (a))

デジカメ	24 時間テレビ	ii
花火	coolpix	x100f
ドラクエ 11	j5	ランナー
d850	ダブルレンズキット	b700

表 3 符号が負で係数の値の大きな単語上位 12 件 (データセット (a))

olympus	canon	ヨドバシカメラ
nikon	ビックカメラ	pen
カメラ	カメラのキタムラ	年賀状
ライブカメラ	プリンター	nikoniko

表 1 にテストデータに対する SVM による分類結果を示す。表より、 F_1 値が 0.759 と高い値になっていることが分かる。将来的な購買予測という問題設定を踏まえると、この分類精度は比較的高いと考えられる。分類精度が高い原因として、正例ユーザと負例ユーザの検索ログの集め方の違いにあると考えられる。正例ユーザは購入日直前から遡って 3ヶ月のウェブ検索ログを使用したのに対して、負例ユーザはフィルタリングクエリで検索してから 3ヶ月のウェブ検索ログを用いている。この違いにより、正例ユーザのウェブ検索ログには、購買を示唆する単語が多く登場していると考えられる。

次に、得られた分類器がどのような単語によって予測を行っているのかを分析した。ロジスティック回帰により得られた分類器の各単語に対応する回帰係数の絶対値の大きな単語は入力データの正例・負例を予測する際に重要な単語であると判断できる。さらに、係数の符号が正である単語は入力データがその単語を含む場合は正例と予測される可能性が高い、すなわち購買を示唆する単語であり、符号が負であれば入力データは負例である可能性が高いと予測される可能性が高い、すなわち非購買を示唆する単語であると考えることができる。回帰係数の値の大きな単語を正負の符号両方で確認することで、分類に特徴的な単語を分析することができる考えた。

データセット (a) の訓練データで学習したロジスティック回帰から得られた、回帰係数の絶対値の高い単語を符号が正の単語、負の単語を表 2, 3 にそれぞれ示す。なお、ロジスティック回帰によるテストデータの分類精度は F_1 値が 0.742 であり、SVM と同程度の精度であった。表より、正例と判断されるユーザは、“デジカメ”などのカメラの名称や“d85”や“b700”などの型番やモデル名のキーワードが複数見られた。また、“ダブルレンズキット”のような商品の付属品のキーワードが含まれていることから、カメラの商品だけでなく付属品などを検索

表 4 SVM によるデータセット (b) を用いた予測結果

正解率	適合率	再現率	F_1
0.542	0.548	0.592	0.569

するユーザが、より正例と判断されたのではないかと考えられる。負例と判断されるユーザは、“olympus”などのメーカー名や“ビックカメラ”などの店舗名などのキーワードが複数見られた。データセット (a) のウェブ検索ログの抽出方法では、単にメーカー名のみを検索するユーザは負例と判断されやすいと考えられる。このことから、購入ユーザと非購入ユーザの検索クリの特徴的な違いとして、購入ユーザは非購入ユーザに比べて、カメラに関してより具体的な情報を調べているのではないかと考えられる。また、“ビックカメラ”などの店舗情報が上位に位置することから、負例ユーザには、オンラインでの購買でなく店舗の情報を調べるユーザが含まれていると考えられる。

続いて、より実用的なデータセットだと考えられる、3.4.2 節で作成したデータセット (b) による検証を行う。訓練データとテストデータの分割方法やパラメータについては先述の実験と同様に行い、 $\gamma = 1, C = 100$ を用いた。表 2 に、データセット (b) における、テストデータに対する分類結果を示す。表より、 F_1 値が 0.569 となり、より実用的な問題を考えた場合、検索行動から購買を予測することは容易ではないことがわかる。前回との分類精度の違いは、ウェブ検索ログの集め方の違いにあると考えられる。今回の対象となるユーザの選定方法は、期間中にフィルタリングクエリを検索したユーザであり、期間の序盤に検索したユーザや終盤に検索したユーザも含まれる。すなわち、購買を示唆する単語の登場する回数がユーザごとに異なり、正例と負例ユーザの差が前回に比べ、あまりない状況であると考えられる。

次に分類において、影響を与える重要な特徴を分析する。データセット (a) と同様に、ロジスティック回帰による分類器を構築し、分類に重要な単語を分析した。得られた分類器のテストデータに対する予測結果は F_1 値が 0.53 であった。データセット (b) の訓練データで学習したロジスティック回帰から得られた、回帰係数の絶対値の高い単語を符号が正の単語、負の単語を表 5, 6 にそれぞれ示す。

購買を示唆する単語は、データセット (b) を用いた場合も、“デジカメ”などのカメラの名称や、“ute”や“dc”などモデル名に関するキーワードが複数見られる結果となった。非購買を示唆する単語は、“ヨドバシカメラ”などの店舗名や、“札幌”などの地名のキーワード複数見られた。データセット (b) による分類器でも、商品を型番やモデル名まで検索しているユーザが正例に分類されやすい結果となった。負例と判断されるユーザは、データセット (a) と同様に、オンラインでの購買でなく店舗の情報を調べるユーザが多数含まれると考えられる。また、表 6 には含まれなかったが、上位 20 位の中に“通販”というキーワードが見られた。このことから、今回のデータセットで負例として用意したユーザは他の EC サイトでの購入や、比較

表 5 符号が正で係数の値の大きな単語上位 12 件 (データセット (b))

デジカメ	ニンテンドースイッチ	ute
ハイエース	セット	24
11	ブルートゥース	f1
dmc	dc	ef

表 6 符号が負で係数の値の大きな単語上位 12 件 (データセット (b))

札幌	ヨドバシカメラ	レシピ
福岡	diy	ビックカメラ
スイッチ	京都	画像
ミニ四駆	熊本	ひよっこ

検討を行うユーザも含まれていると考えられる

また、本実験では、正例と負例の割合を同数としていたが、実際には比率は大きく異なると考えられる。今後は、データセットの正例と負例のユーザの比率を考慮することで、より実用的な問題を検証していきたい。

5.2 期間の検証

本研究では 3ヶ月間のウェブ検索ログを用いて予測を行った。分類に用いる期間により、予測精度は変わると考えられる。そこで、本節では購買予測に適した検索期間を検証する。検証した期間は、半年、3カ月、1カ月、1週間の 4 種類である。

本実験のデータセットは以下の手順で用意した。まず、フィルタリングクエリで検索したことがあり、かつ年間 50 回以上検索を適用したユーザからコンパクトデジタルカメラあるいはデジタル一眼カメラを購入したことのあるユーザを 5,000 人抽出し、購入日から遡って 3ヶ月のクエリログを用いた。その後、同数の非購入ユーザを抽出し、フィルタリングクエリを含むクエリで検索した日から 3ヶ月の期間を用いた。すなわち本データセットは前節のデータセット (a) と同様に、正例ユーザは翌日にカメラを購入しているユーザである。実験条件は前節と同様に、80%のデータを訓練データ、残りをテストデータとし、分類を構築した。

表 7 に異なる検索期間における予測精度を示す。表より、1週間の期間のデータを使用した場合、他の期間に比べ分類精度が高くなっており、6ヶ月のデータを使用すると分類精度が若干下がっていることが分かる。1週間の期間で精度が最も高くなった結果として、今回の実験で用いたデータセットは、正例ユーザは翌日にカメラの購買をしているため、購買を表す特徴が直近のクエリに表れやすかったのではないかと考えられる。購買行動の予測については、購入とみなす期間をどのように設定するかによって、分類に適した期間は変わってくる。今後は、多様な問題設定で期間の検証をしていく必要があると考えられる。

6 まとめ

本研究では、ウェブ検索ログを用いてユーザの商品購買行動

表 7 異なる期間による分類精度の検証

	正解率	適合率	再現率	F_1
半年	0.684	0.688	0.685	0.686
3 カ月	0.710	0.705	0.703	0.717
1 カ月	0.702	0.757	0.585	0.659
1 週間	0.739	0.695	0.853	0.756

を予測する手法を提案した。ウェブ検索ログには購買を示唆する特徴的な語が含まれると仮定し、カメラの購入予測を題材として、ウェブ検索のクエリログと、商品購買履歴を用いて検証を行った。

今回はデータセットの構築と分類器による精度の検証や、分類に重要と判断された単語の分析を行った。分類器がどのような情報を用いて購買行動を予測しているのか、特徴的な単語が購買日から遡ってどのような時期に出現するのかといったことを分析することで、ユーザのウェブ検索上での購買プロセスをより詳細に理解することができると思われる。

今後、他の機械学習モデルによる学習、クエリ以外の検索ログを用いた推定をしていく必要がある。また、今回の実験では、正例と負例の割合を同数としていたが、実際の環境では正例と負例の割合は大きく異なると思われる。今後は、データセットの正例と負例のユーザの比率を考慮することで、より実用的な問題を検証していきたい。

謝 辞

本研究の一部は JSPS 科学研究費助成事業 JP16H02906, JP18H03494, JP17H00762, JP18H03244, JP18H03243 による助成を受けたものです。ここに記して謝意を表します。

文 献

- [1] J. Dewey. How to think. *Boston: DC Heath & Co*, 1910.
- [2] J. F. Engel, D. T. Kollat, and D. Roger. Blackwell, consumer behavior. *New York*, 1973.
- [3] Q. Guo and E. Agichtein. Ready to buy or just browsing?: detecting web searcher goals from interaction data. In *Proceedings of the 33rd international ACM SIGIR conference on Research and development in information retrieval*, pages 130–137, 2010.
- [4] G. Murga. Amazon Takes 49 Percent of Consumers' First Product Search, But Search Engines Rebound. <https://blog.survata.com/amazon-takes-49-percent-of-consumers-first-product-search-but-search-engines-rebound>.
- [5] P. Sondhi, M. Sharma, P. Kolari, and C. Zhai. A taxonomy of queries for e-commerce search. In *Proceedings of the 41st International ACM SIGIR Conference on Research & Development in Information Retrieval*, pages 1245–1248, 2018.
- [6] N. Su, J. He, Y. Liu, M. Zhang, and S. Ma. User intent, behaviour, and perceived satisfaction in product search. In *Proceedings of the Eleventh ACM International Conference on Web Search and Data Mining*, pages 547–555, 2018.