

検索行動に基づく購買満足度の関係分析

柳田 雄輝[†] 加藤 誠^{††} 河田 友香^{†††} 山本 岳洋^{†††,††††} 大島 裕明^{†††,††††}
藤田 澄男^{††††}

[†] 筑波大学大学院 情報学学位プログラム 〒 305-8550 茨城県つくば市春日 1-2

^{††} 筑波大学 図書館情報メディア系 / JST さきがけ 〒 305-8550 茨城県つくば市春日 1-2

^{†††} 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

^{††††} 兵庫県立大学 大学院情報科学研究科 〒 650-0047 兵庫県神戸市中央区港島南町 7-1-28

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

E-mail: [†]ynagi2@klis.tsukuba.ac.jp, ^{††}mpkato@acm.org, ^{†††}{fa19c021,t.yamamoto}@sis.u-hyogo.ac.jp,

^{††††}ohshima@ai.u-hyogo.ac.jp, ^{†††††}sufujita@yahoo-corp.jp

あらまし 本研究では、EC サイトにおける商品購入の前後に行われる Web 検索行動と購買満足度の関係分析を行った。具体的には、EC サイトで商品を購入したユーザの Web 検索ログをその検索意図に基づいて分類し、その時間変化が購買満足度によって異なるかを調査した。その結果、購買に満足した場合とそうでない場合において検索意図の時間変化が異なること、また、購入者や商品の性質によってその差異が異なることを明らかにした。さらに、分析で得た知見をもとに購買満足度の推定にも取り組み、商品購入の前後に行われる Web 検索行動によってある程度の正解率で満足度が推測できることを示した。

キーワード 検索行動, 検索意図, 購買満足度, Web 検索, ユーザ行動分析

1 はじめに

近年オンラインショッピングにて商品を購入する人々は増加しており、その利用者を満足させることは重要である。PwC の調査によると、週に 1 回以上スマートフォンなどのモバイル端末から商品を購入している人の割合は、2019 年は 24%であったが、2021 年 12 月には 41%とその割合が増加している [1]。今後もオンラインショッピングにて商品を購入する人々の割合は増加することが見込まれているため、その利用者を満足させることには大きな効果が期待される。

このような背景から購買満足度に関する研究は広く行われているが、オンラインショッピングにおける購入者の情報探索行動と購買満足度の関係についてはまだ十分な理解が得られていない。例えば、Richins と Bloch は車を例として、耐久消費財購入後の満足度とその時間的な変化について分析し、商品に対する関心の高さによって満足度の時間的な変化が異なることを明らかにしている [2]。Tsotsos は知覚された商品の質が、Söderlund は購買経験が購買満足度や商品の再購入に影響を及ぼすことを明らかにしている [3], [4]。また、情報検索の分野においてはユーザ検索行動の分析が広く行われている。中でも Su らは EC サイトにおける検索行動について、各検索セッションに検索意図を割り当て、ユーザの検索行動や割り当てた検索意図と、検索結果に対する満足度の関係性を分析している [5]。Zhang らはサーチエンジンの広告について、広告に表示された商品を購入する前後でユーザのその商品に対する広告消費行動が異なることを明らかにしている [6]。このように、商品に対す

る関心の高さ、商品の質、顧客の特性として購買経験が購買満足度に影響を及ぼすこと、そしてユーザの検索行動と商品に関する検索結果についての関係性は明らかにされている。しかし、購入者の情報探索行動も購買満足度に影響を及ぼすと考えられるが、その関係性についてはまだ十分な理解が得られていない。

そこで本研究では、EC サイトにおける商品購入の前後に行われる情報探索行動と購買満足度の関係を分析する。オンラインショッピングにおける情報探索行動の記録として Web 検索ログを用い、購買満足度を表す評価として EC サイトにおける商品のレーティングを用いる。Web 検索行動と購買満足度の関係分析として、例えば、検索回数の多いユーザは満足しやすいといった傾向が見られるか分析する。本研究のリサーチクエスションは次の通りである：

- RQ1.** 商品を購入し満足したユーザとそうでないユーザの商品に関する検索行動は異なるのか？
- RQ2.** 商品の特性が変われば、商品を購入し満足したユーザとそうでないユーザの商品に関する検索行動は異なるのか？
- RQ3.** ユーザの特性が変われば、商品を購入し満足したユーザとそうでないユーザの商品に関する検索行動は異なるのか？
- RQ4.** 購買前、または、購買後の検索行動から満足度は予測できるか？

RQ1 を検証するために、EC サイトで商品を購入したユーザの Web 検索ログをその検索意図に基づいて分類し、その時間変化が購買満足度によって異なるかを調査する。RQ2 と RQ3 については、RQ1 の分析について商品やユーザの特性でそれぞれ二分し、特性ごとに購買に満足したユーザとそうでないユーザの購買に関する検索行動の違いがあるか検証する。RQ4 に

については、ユーザの時区間ごとの検索回数を用いてユーザが満足しているかどうかを予測する2値分類問題に取り組むことで検証する。

データセットには Yahoo!検索の Web 検索ログと Yahoo!ショッピングの購買ログそれぞれ13ヶ月分およびレビューデータを用いた。分析の結果、購買に満足した場合とそうでない場合において検索意図の時間変化が異なること、また、購入者や商品の性質によってその差異が異なることが明らかになった。さらに、分析で得た知見をもとに購買満足度の推定にも取り組み、商品購入の前後に行われる Web 検索行動によってある程度の正解率で満足度が推測できることを示した。

本研究の貢献は次の3点である：

- 商品購入の前後に行われる情報探索行動と購買満足度の関係を明らかにする新たな問題に取り組んだ。
- 購買に満足した場合とそうでない場合において検索意図の時間変化が異なること、また、購入者や商品の性質によってその差異が異なることを明らかにした。
- 購買前後の Web 検索行動からユーザの満足度を予測するタスクに取り組み、Web 検索行動によってある程度の正解率で満足度が推測できることを示した。これにより、レビューをしていないユーザに対してもある程度満足度が推測できることを示した。

本論文の構成は次の通りである。2節では商品の購買満足度やユーザの検索行動に関する関連研究について述べる。3節では分析の準備として、分析に用いるデータの詳細や検索意図について述べる。4節では購買に満足したユーザとそうでないユーザについて、検索行動に違いが見られるか分析した結果について報告する。5節では4節で得られた分析結果をもとに、購買満足度を予測するタスクに取り組む。6節では本論文の結論を述べる。

2 関連研究

本節では顧客満足度とユーザ検索行動の2トピックそれぞれの関連研究について述べる。

2.1 顧客満足度

顧客満足度とは、知覚された商品のパフォーマンスが顧客の期待に対しどの程度であったかを評価したものである。したがって、パフォーマンスが期待ほどでなければ顧客は不満を持ち、パフォーマンスが期待以上であれば顧客は満足する。そして、顧客満足度は売り手の収益やサービスの利用継続などに関係することがわかっている [7], [8]。例えば Hallowell は、顧客の満足度はロイヤリティに結びつき、ロイヤリティは収益に関係することを明らかにしている [7]。

また、満足度に影響を及ぼす要因についても研究されている [9], [10]。Richins と Bloch は車を例として、耐久消費財購入後の満足度とその時間的な変化について分析している [2]。その結果、製品に対する関心の高い消費者は、関心の低い消費者と比べ、商品を所有している期間中の満足度がわずかに高いこ

とを明らかにしている。その一方で、購入後2ヶ月間では、製品に対する関心の高い消費者の満足度は低下しており、関心の低い消費者の満足度は上昇していたことも明らかにしている。Tsiotsou は知覚された商品の品質と満足度が購入に与える影響を調査し、知覚された商品の質が購買満足度や商品の再購入に影響を及ぼすことを明らかにしている [3]。Söderlund は顧客の親近感と満足度の関係を分析している [4]。顧客がこれまでにあるレストランを利用した回数を親近感としたとき、親近感が平均よりも高い群は平均よりも低い群と比べ、顧客満足度や再利用意図がより極端になることを示している。また、サービスのパフォーマンスが高い場合には、親近感が平均よりも高い群の方が高い満足度を得ていることを明らかにしている。

このように、商品に対する関心の高さ、商品の質、顧客の特性として購買経験が購買満足度に影響を及ぼすことが明らかにされているが、情報探索行動と結びつけた分析はされていない。本研究では、EC サイトにおけるレビューのレーティングを顧客満足度、検索回数を商品に対する期待や関心の大きさとみなす。そして、既存研究で明らかにされている知見がオンラインショッピング及び Web 検索行動にもあてはまるか、また Web 検索行動がどのような影響を及ぼしているか検証する。

以上のように、本研究では RQ1 に加え、RQ2, RQ3 の通り、商品やユーザの特性が変わった場合に、購買に満足したユーザとそうでないユーザで商品に関する検索行動が異なるかどうかを分析する。

2.2 ユーザ検索行動

ユーザ検索行動、特に購買に関する検索行動の分析を行った研究もなされている [11], [12], [13], [14], [15]。中でも Su らは EC サイトにおける検索行動について、検索の動機、商品の絞り込み具合、戦略を観点として各検索セッションに Target finding (TF), Decision making (DM), Exploration (EP) の3種類の検索意図を割り当てている。その後、ユーザの検索行動や割り当てた検索意図と、検索結果に対する満足度の関係性を分析している [5]。Su らの分析はユーザの検索行動や検索意図と満足度の関係を分析している点で本研究と類似しているが、扱う満足度が異なる。Su らの分析における満足度とは、被験者実験によって得た検索体験に対する満足度を指す。検索体験に対する満足度と購買行動の関係については、満足が必ずしも購買につながるとはいえないとし、購買行動を検索体験に対する満足度とすることは適切でないと述べている。対して本研究では、ユーザが商品に付与したレーティングを用いることで、ユーザの商品購買後も含めた検索行動と商品購入後に得る購買満足度の関係を分析する。加えて、本研究ではユーザのこれまでの購買経験や商品の特性を加味した分析まで行う。Zhang らはサーチエンジンの広告について、広告に表示された商品を購入する前後でユーザのその商品に対する広告消費行動が異なることを明らかにしている [6]。Zhang らの分析は商品の購買前後の検索行動を分析している点で本研究と類似している。しかし、Zhang らは広告消費行動の分析をしていることに対し、本研究ではユーザが購買後に付与する満足度の分析を行う点で異

なる。

検索行動に理解にあたり、検索意図の分類法を提案する研究もされている。Web 検索意図の分類として、Broder は Web 検索を目的によって Navigational, Informational, Transactional の 3 つに分類し、Jansen らはこの分類に基づいたクエリの分類を行っている [16], [17]。Rao らは Broder の分類に加え、Web 検索における商品に関連する意図として Comparison, Support を加えた分類法を提案している [18]。Ashkan と Clarke は広告を一定回数以上クリックしているクエリについて、商用と非商用、Navigational と Informational にそれぞれ分類している [19]。EC サイトにおける意図の分類法として、Su らが商品検索意図を提案している他、Moe は EC サイト内のユーザー行動意図をクラスタリングすることで分析している [5], [20]。Sondhi らは EC サイトのクエリについて、CTR やトークンの長さなどを特徴量としてクエリを 5 カテゴリにクラスタリングしている [21]。これらの研究は検索行動を理解するために検索意図の分類や予測などを行っている。対して本研究では、Web 検索行動について検索意図を割り当てたのち、EC サイトでの購買と紐付けて分析を行う。本研究では購買における情報探索行動の分析として、どの程度情報を絞り込んで検索しているかに注目するため、Su らが提案した検索意図を Web 検索ログに適用し分析する。

以上より、本研究では RQ1 の通り、購買満足度と購買前後の検索行動の関係の分析に取り組む。そして、RQ4 を検証するため、購買前後の検索行動に基づく購買満足度の推定にも取り組む。

3 対象データ

本節では、本研究で用いるデータセットについて述べた後、用いるデータに対して検索意図を割り当てる方法について示す。

3.1 データセット

本研究では商品購買前後の Web 検索行動と満足度の関係を分析するため、Web サーチエンジンである Yahoo! 検索の Web 検索ログと、EC サイトである Yahoo! ショッピングの購買ログおよびレビューデータを用いる。Web 検索ログと購買ログは 2016 年 10 月から 2017 年 10 月を対象期間とし、レビューデータについては、レビューは購買後に行われる性質上、2016 年 10 月から 2018 年 3 月を対象期間とした。購買ログおよびレビューデータと Web 検索ログを紐付けることで、対象期間に商品を購入したユーザーの Web 検索行動の分析を可能とする。

購買ログについては、ある程度価格に幅があるカテゴリを選択することで、価格に応じて Web 検索行動の違い、例えば高価な商品を買う際には入念に商品の検討が行われているといった傾向があるか比較できるようになる。そこで本研究ではカメラカテゴリを対象とし、対象期間にカメラカテゴリの商品を購入したユーザーについて分析する。

Web 検索ログについては、日常的に Yahoo! 検索にて Web 検索を行っているユーザーを対象とするため、対象期間において毎

月 10 日以上検索しているユーザーに絞って分析する。

レビューデータには 5 段階で評価されるレーティングの他にレビュー文が付与されている。レーティングの値は必ずしも購買に対する満足度を表しているとはいえず、その質も多様である。レビューの文字数と記載内容について目視したところ、100 文字以上のレビューにおいては商品について詳細に述べているものが多いことがわかった。一方で、100 文字以上のレビューの中には同じ文章を使いまわして掲載している文や、購買前に付与されている文が見られた。以上を踏まえ、本研究では 100 文字以上のレビュー文を持ち、かつ文字列が完全一致するレビューが過去に行われていない、商品購買後に付与されたレビューのレーティングに限定しレビューを抽出する。また、5 段階のレビューレートについてもレートとその記載内容について目視したところ、レートが 3 以下のレビューについては主に不満足であることを示す内容が記されている傾向が見られた。この傾向から、本研究ではレートが 4 以上であれば満足、レートが 3 以下であれば不満足とする。

3.2 意図の割り当て

本研究では各 Web 検索クエリに意図を割り当て、その頻度や時間変化を捉えることで検索行動を分析する。Web 検索クエリには商品の購買に関係あるものとそうでないものが存在する。加えて、同じカテゴリの商品について検索している場合でもその絞り込み度合いは異なることがある。例えば、クエリ“カメラ おすすめ”で検索したユーザーはカメラを購入したいことしか決まっていないと考えられる。一方、クエリ“EOS R5 価格”で検索したユーザーは EOS R5 を購入しようか検討している段階であると考えられる。そのため、各クエリに商品の絞り込み度合いを表すラベルを割り当てることによって検索行動を捉える。本研究では購買における情報探索行動の分析として、どの程度情報を絞り込んで検索しているかに注目するため、Su らが提案した検索意図を用いて分析する。

Su らはアノテーションに基づいて EC サイトにおける検索の観点を集約し検索意図を提案している。しかし、Web 検索クエリは EC サイトにおけるクエリログと比べ多様な観点を含むうえ、その量は膨大であるため、アノテーションに基づき検索意図を割り当てることは困難である。したがって、本研究では自動で検索意図の割り当てを行う。自動での検索意図の割り当てを行うにあたり、ルールベースの分類では条件を満たすクエリを確実に抽出できる利点がある。しかし、商品カテゴリごとに、そのカテゴリに関する語の一覧を作成しルールベースで分類するのは困難である。例えばカメラカテゴリにおいてレンズという語はカメラに関係する語だと考えられるが、このようなカメラに関係するがカメラと明示されていない語を網羅することは難しい。そこで本研究では、ルールベースでは捉えきれなかった語を捉えるため、ルールベースの分類を行った後に弱教師あり学習による分類も行うことで、ルールでは捉えきれなかった傾向を捉えた意図の割り当てを行う。以上より、本研究では Su らによる TF/DM/EP の 3 つによる意図の分類法を参考しつつ、ルールベースの分類と弱教師あり学習による分類の

両方を用いてクエリに意図を割り当てる。

3.2.1 ルールベースの分類による意図の割り当て

ルールベースの分類では、TF/DM/Other の3つの意図を用い、次のルールにてクエリに意図を割り当てる。

a) TF (Target finding)

ユーザは商品のカテゴリやブランドを知っており、商品を購入する、あるいは的を絞ったキーワードで検索している段階を指す。ユーザが自身の購入したい商品のカテゴリやブランドを既に知っているのであれば、メーカー名やブランド名で検索して商品を吟味したり、その商品の商品名や型番で検索しその商品を購入したりすると考えられる。したがって、クエリ内に購入商品のカテゴリに関するメーカー名、型番、商品名のいずれかを含むクエリにこの意図を割り当てる。メーカー名、型番、商品名については商品比較サイトの価格.com¹をクロールし、メーカー名、型番、商品名の一覧を作成した。

b) DM (Decision making)

商品カテゴリは知っているが、何をかうべきかについては不明瞭な段階を指す。ユーザが自身の購入したい商品のカテゴリは明らかになっているが、他の要素についてはまだ不明瞭である場合、その商品のカテゴリ名をクエリに含んだ検索を行うことが考えられる。例えば、カメラを購入したい場合はまず、クエリ“カメラ おすすめ”で検索するといったことが考えられる。したがって、TFの条件を満たさないが、その商品のカテゴリ名を含むクエリにこの意図を割り当てる。

c) Other

Suらはブラウジングや商品を絞り込んでいない段階の意図としてEP (Exploration)を示している。ECサイトの検索ログを用いているSuらに対し、本研究ではWeb検索ログを用いている。そのため、TFおよびDMのいずれの条件も満たさないクエリが商品の探索そのものに関係する可能性は低いと考えられる。したがって、商品には関係ない意図として新たにOtherという意図を割り当てる。

3.2.2 弱教師あり学習の分類による意図の割り当て

弱教師あり学習の分類による意図の割り当てではルールベースでは捉えきれなかった語を捉えるため、ルールベースの分類で意図がOtherに分類されたクエリのみを用いる。弱教師あり学習による意図の割り当てでは、クエリを空白で分割して得たBag of Wordsを入力、そのクエリに割り当てる意図を出力として、クエリに含まれている語に基づく意図の分類を実現する。加えて、ユーザは商品について調べる際、商品の絞り込み度合いを変えながら検索すると考えられるため、意図を割り当てるにはクエリの時系列も考慮する必要がある。そのため、Web検索ログをセッション分割し、各セッションに注目して弱教師あり学習の正例と負例に相応しいクエリを抽出する。セッションは連続したクエリの間30分以上の時間が空いている場合はタイムアウトしていると仮定し作成した。ここで、セッション集合は $\mathcal{N}(s)$ と表す。任意のセッション $s_i \in \mathcal{N}(s)$ はクエリ列であるため、 s_i の k 番目のクエリを $q_{i,k}$ 、 n を0以上の整数と

して $s_i = \{q_{i,0}, \dots, q_{i,n}\}$ と表し、正例と負例を次のように定める。

a) 正例

Web検索を用いて商品について調べる場合、ユーザはクエリを何度か再構成しながら情報を探索すると考えられる。そのため、商品を探しているクエリは時系列的に連続して出現しやすいと推測できる。したがって、 m を整数として $0 < m < n$ を満たす任意の $q_{i,m}$ について、 $q_{i,m-1}, q_{i,m+1}$ の両方が型番、商品名、カテゴリ名のいずれかを含む場合、 $q_{i,m}$ はOtherに分類されていても商品の探索に関係している可能性が高い。なお、ルールベースによる分類ではメーカー名を含むクエリもTFと分類している。しかし、メーカー名のみ含まれているクエリは他のカテゴリの商品を探している可能性も考えられるため、正例とみなす条件には含めないこととする。このとき $q_{i,m}$ はメーカー名、型番、商品名の一覧にある語を含まないことも踏まえ、意図をDMであるとみなし正例とする。加えて、前述の条件のみでは対象とする商品カテゴリ以外でも出現しやすい語が含まれる可能性がある。そのため、クエリ内での語の共起に注目し、PMI (Pointwise Mutual Information) が負の語および出現頻度10回未満の語を除く。この条件により、カテゴリ名と共起しにくい語や、他のカテゴリでよく共起する語を含まないようにする。以上の条件を満たしたクエリについて目視したところ、カメラカテゴリにおいてカメラ専門店のECサイトや、レンズのような付属品を表す語を多く含んでいたことが確認できた。そのため、これらのクエリを正例として用いる。

b) 負例

Web検索クエリは多様であり、それらの大半はその商品の購買に関係のないものであると考えられる。したがって、任意の $q_{i,m}$ に対してOtherが割り当てられているとき、 s_i は購買に関係しないセッションであると推測できるため、任意の $q_{i,m}$ を負例とする。

c) 評価

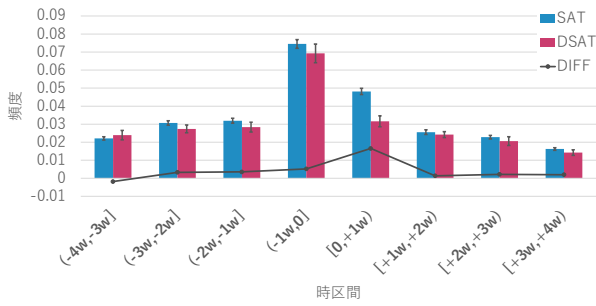
以上の条件を満たした正例および負例のクエリを正例と負例の比が1:1になるようにサンプリングし学習に用いる。学習にはサポートベクターマシン、ランダムフォレスト、ロジスティック回帰を用いたが、いずれのモデルでも予測結果はほぼ変わらなかった。そのため、より早く学習を行えるロジスティック回帰を用い、正例または負例のラベルが付与されていないクエリについてラベルを割り当てる。本モデルによる予測の結果について評価するため、予測結果のうち、無作為に抽出したクエリと割り当てたラベル25組について、そのクエリがカテゴリに関係するクエリであるかどうかを著者が見て判定した。その結果、正解率0.84、F1値0.80で分類できていることが確認された。

以上で得た、ルールベースの分類と機械学習による分類の両方によってクエリに割り当てた意図を用いて検索行動を分析する。

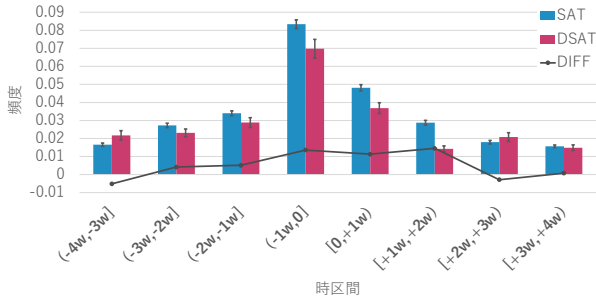
4 購買満足度の分析

本節では、3節で述べたデータに割り当てた検索意図の時間

1: 価格.com: <https://kakaku.com/>

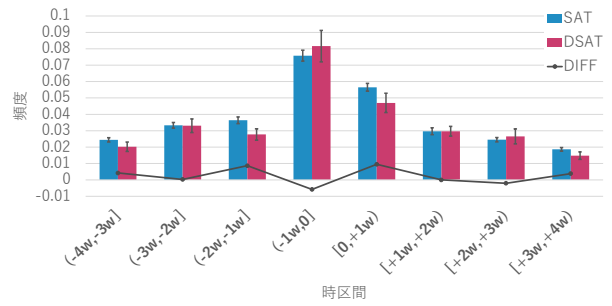


(a) TF クエリの時間変化

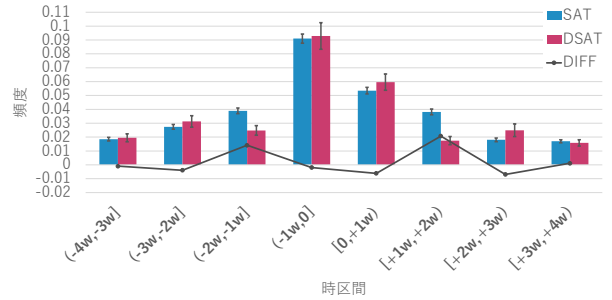


(b) DM クエリの時間変化

図 1 検索意図の時間変化



(a) TF クエリの時間変化



(b) DM クエリの時間変化

図 2 価格が中央値よりも高い場合における検索意図の時間変化

変化を分析することで RQ1, RQ2, RQ3 を検証する。

分析にあたっては、検索ログを購買前後それぞれについて分割し、各時区間におけるクエリに割り当てられた検索意図の頻度を求めることで検索意図の時間変化を分析した。購買の直前・直後だけでなく長期的な購買に関する Web 検索行動を捉えるため、購買前後それぞれについて、購買から 1 週間以内、1 週間から 2 週間以内、2 週間から 3 週間以内、3 週間から 4 週間以内、4 週間より後の 5 時区間ずつ、合計 10 時区間に分割した。検索意図の時間変化を分析するに際し、各ユーザが各時区間においてどの程度検索しているかを明らかにするため、購買に満足したユーザとそうでないユーザそれぞれについて各クエリに割り当てられた検索意図の出現回数をカウントした。分析にあたっては、ユーザごとに総検索回数は異なるが、極端に多く検索しているユーザの影響を強く受けたくない方が望ましいため、カウントした検索意図について正規化を行い算出した。具体的には、式 (1) で得られる $n_{t,u}$ を全ユーザについて平均し、時区間 t における検索意図の頻度として各時区間における頻度を算出した。

$$n_{t,u} = \frac{\text{時区間 } t \text{ におけるユーザ } u \text{ の検索回数}}{\text{ユーザ } u \text{ の全期間の検索回数}} \quad (1)$$

4.1 商品を購入し満足したユーザとそうでないユーザの商品に関する検索行動は異なるのか？

本小節では RQ1 について検証する。図 1a に TF が割り当てられたクエリ (以下、TF クエリ) について、図 1b に DM が割り当てられたクエリ (以下、DM クエリ) について、購買に満足したユーザとそうでないユーザの検索意図の時間変化を示す。これらのグラフの横軸は購買から検索までの時区間を表しており、例えば、(-1w, 0] は購買前 1 週間以内、[+1w, +2w] は購買後 1 週間以上 2 週間以内の時区間を表している。購買前後 4

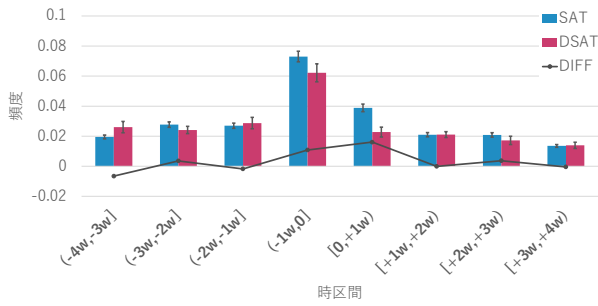
週間より後の時区間については、その頻度の割合が他の区間よりも非常に大きくなるため、グラフではそれら以外の 8 時区間の頻度を示している。縦軸は正規化して得た検索意図の頻度を表している。棒グラフは満足したユーザとそうでないユーザそれぞれについて各時区間における検索意図の頻度を表している。例えば図 1a では、各時区間における TF クエリについて、満足したユーザの頻度は SAT の棒グラフで、そうでないユーザの頻度は DSAT の棒グラフで示している。折れ線グラフは各時区間について、SAT の値から DSAT の値を引いた値を示している。

図 1a と図 1b から、満足したユーザはそうでないユーザと比べ、TF クエリと DM クエリの両方において購買前後 2 週間以内の頻度が高く、特に TF クエリにおける購買後 1 週間以内はその傾向が顕著にみられる。この傾向は検索の頻度を製品に対する関心とみなしたとき、Richins と Bloch の製品に対する関心の高い消費者は、関心の低い消費者と比べ満足度が高いという知見を反映していると考えられる。購買後 1 週間以内における TF クエリの頻度の差については、商品に満足している場合はそうでない場合と比べ、その商品の使い方や詳細などについて調べるためにこのような差がみられたと推測できる。

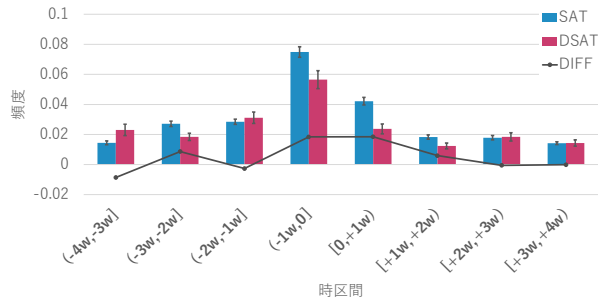
以上より、商品を購入し満足したユーザとそうでないユーザの商品に関する検索行動は異なる傾向にあることが推測された。

4.2 商品の特性が変われば、商品を購入し満足したユーザとそうでないユーザの商品に関する検索行動は異なるのか？

本小節では RQ2 について検証する。RQ2 の検証にあたっては、商品の特性の中でも価格に注目することで、高価な商品を購入する場合とそうでない場合で満足する Web 検索行動が異なるかを明らかにする。本分析では商品を価格の中央値で二分



(a) TF クエリの時間変化



(b) DM クエリの時間変化

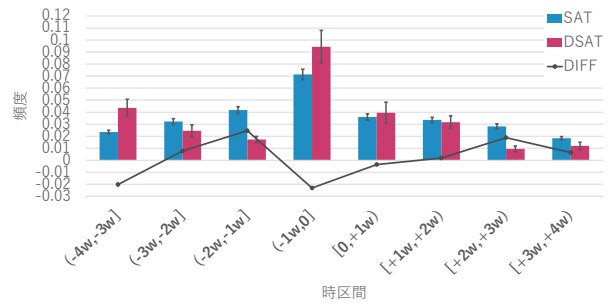
図3 価格が中央値以下の場合における検索意図の時間変化

し、購入した商品の価格が中央値よりも高い場合とそうでない場合で購入に満足したユーザーとそうでないユーザーの商品に関する Web 検索行動が異なるか検証する。

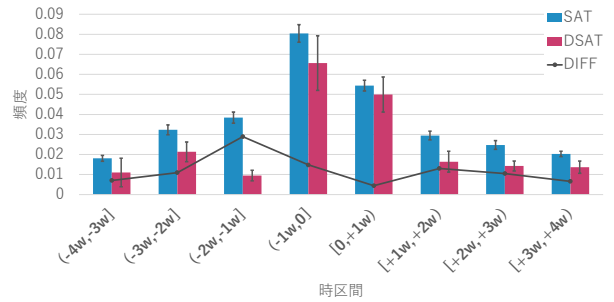
価格が中央値よりも高い場合の結果として、図 2a に TF クエリについて、図 2b に DM クエリについて、購入に満足したユーザーとそうでないユーザーの検索意図の時間変化を示す。価格が中央値以下の場合の結果として、図 3a に TF クエリについて、図 3b に DM クエリについて、購入に満足したユーザーとそうでないユーザーの検索意図の時間変化を示す。図 1a や図 1b と同様に、これらのグラフの横軸は購入から検索までの時区間を、縦軸は正規化して得た検索意図の頻度を表している。棒グラフは満足したユーザーとそうでないユーザーそれぞれの各時区間における検索意図の頻度を表しており、折れ線グラフは各時区間について、満足したユーザーの検索意図の頻度からそうでないユーザーの検索意図の頻度を引いた値を示している。

図 2a と図 2b から、価格が中央値よりも高い場合はいずれの検索意図においても、満足していないユーザーの方が購入前 1 週間以上 2 週間以内の頻度は低く、購入前 1 週間以内の頻度が高いことがわかる。この傾向は、満足しているユーザーはそうでないユーザーと比べ購入前 1 週間以内以外にも商品カテゴリに関する検索をしていることを意味する。したがって、価格の高い商品を購入する場合においては購入直前だけでなく、商品について長期的に調べている方が満足度が高くなる可能性があるかと推測できる。図 3a と図 3b から、価格が中央値以下の場合については、価格で二分しなかった場合と同様の傾向が見られる。したがって、価格が低い場合においては、長期的に調べることが満足につながるとはいえないと推測できる。

以上より、商品の特性が変われば、商品を購入し満足したユーザーとそうでないユーザーの商品に関する検索行動は異なる傾

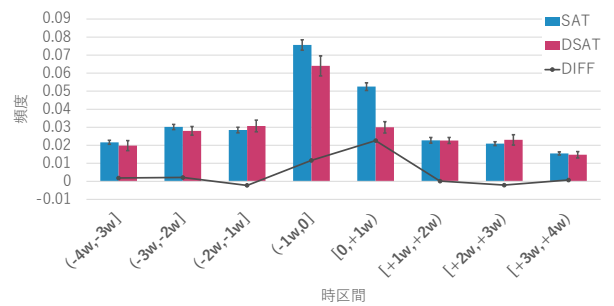


(a) TF クエリの時間変化

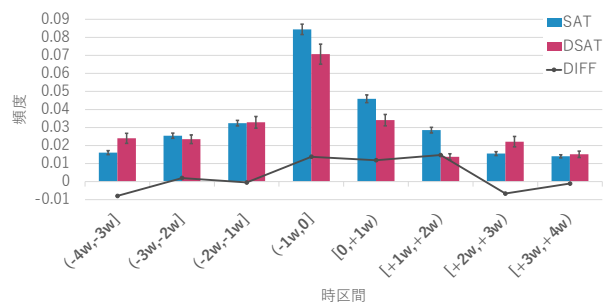


(b) DM クエリの時間変化

図4 購買経験が複数回あるユーザーにおける検索意図の時間変化



(a) TF クエリの時間変化



(b) DM クエリの時間変化

図5 購買経験が1回のみユーザーにおける検索意図の時間変化

向にあることが推測された。

4.3 ユーザの特性が変われば、商品を購入し満足したユーザーとそうでないユーザーの商品に関する検索行動は異なるのか？

本小節では RQ3 について検証する。RQ3 の検証にあたっては、分析期間内でカメラカテゴリの商品を複数回購入しているユーザーはカメラについて高い興味を持っていると考えられるため、ユーザーの特性の中でも分析期間内の購買回数に注目する。

この観点により、商品カテゴリに対し高い興味を持っている場合とそうでない場合で満足する Web 検索行動が異なるかを明らかにする。本分析では分析期間内でカメラカテゴリの商品を複数回購入しているユーザであるかどうかで二分することで、ユーザの特性によって購入に満足したユーザとそうでないユーザの商品に関する Web 検索行動が異なるか検証する。

分析期間内の購買回数が複数回であるユーザの結果として、図 4a に TF クエリについて、図 4b に DM クエリについて、購入に満足したユーザとそうでないユーザの検索意図の時間変化を示す。分析期間内の購買回数が 1 回であるユーザの結果として、図 5a に TF クエリについて、図 5b に DM クエリについて、購入に満足したユーザとそうでないユーザの検索意図の時間変化を示す。これらのグラフの横軸と縦軸は図 1a や図 1b と同様である。

図 4a と図 4b より、購買経験が複数回あるユーザは満足した場合、TF クエリと DM クエリの両方で购买前 1 週間以内だけでなく、购买前 3 週間以上 4 週間前から徐々に頻度が高くなっていく傾向がみられる。一方、購買経験が複数回あるユーザは満足しない場合、TF クエリと DM クエリの両方で购买前 3 週間以上 4 週間前から購入に近づくにしたがい頻度が低くなる傾向がある。そして、购买前 1 週間以内になると急激に頻度が高くなっており、特に TF クエリにおいては満足しているユーザ以上にその頻度が高くなっている。これらの傾向から、商品に対し高い興味を持っているユーザにおいては、徐々に検索の量を増やすような探索ではなく、検索の量が減少していった後、購買直前に集中して調べる場合に不満足になると推測できる。図 5a と図 5b のように、購買経験が複数回あるユーザはその割合が少ないため、購買経験が 1 回のユーザについては特性で二分しない場合と似た傾向がみられた。

以上より、ユーザの特性が変われば、商品を購入し満足したユーザとそうでないユーザの商品に関する検索行動は異なる傾向にあることが推測された。

5 購買満足度の予測

本節では、ユーザの各時区間の検索回数を用いてユーザが満足しているかどうかを予測する 2 値分類問題に取り組むことで RQ4 を検証する。満足度予測にあたっては、4 節にて商品を購入し満足したユーザとそうでないユーザの商品に関する検索行動は異なることが明らかになったため、各検索意図における各時区間の頻度を用いた予測を行う。加えて、検索意図だけでなく、購買前の Web 検索クエリを用いた予測にも取り組むことで、購買満足度とクエリ語の関係を明らかにする。

予測に用いるデータとして、本研究で用いるレビューデータのうち 83.0% はレートが 4 以上であり、本研究において満足とみなされる。したがって、レビューデータは満足度が満足に偏った不均衡データとなっているといえる。そのため、レビューデータは満足と不満足が同数になるようにアンダーサンプリングした上で学習に用いることで、学習における予測ラベルの偏りが生じないようにした。

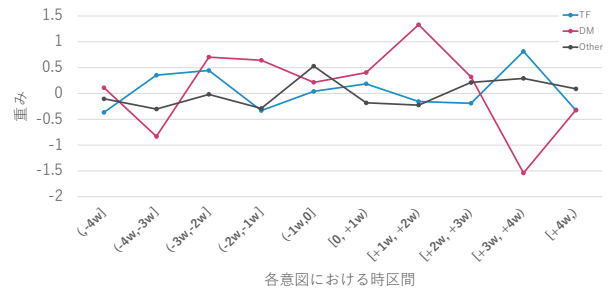


図 6 検索意図の頻度における重み

予測にあたっては、どの特徴量が購買満足度に影響しているかを明らかにするため、ランダムフォレストとロジスティック回帰を用いた。ランダムフォレストは木の本数を 1 本から 20 本、木の深さを 1 から 5、ロジスティック回帰は正則化の強さを表す係数の候補を 10^{-2} から 10 倍刻みで 10^5 までとして、グリッドサーチを用いて最適なパラメータを探索し決定した。本学習ではアンダーサンプリングを行っているため、学習に十分な量のデータを用いることができていない可能性がある。そのため、Nested Cross Validation によって学習と評価を行い、検証用データに関する Cross Validation とテスト用データに関する Cross Validation のいずれについても 3-Fold で分割した。

5.1 検索意図の頻度による予測

検索意図の頻度による予測では、合計 3 種類の各検索意図について、合計 10 区間の各時区間ごとの割合を並べた 30 次元ベクトルを用いて満足度の予測を行った。その結果、正解率はランダムフォレストで 0.47、ロジスティック回帰で 0.49 となり、今回用いたデータにおいては、検索意図の割合から満足度を予測することは困難であることが示された。ロジスティック回帰による予測で得た各特徴量の重みを図 6 に示す。

図 6 では、横軸を各検索意図における時区間を示しており、縦軸は各意図における各時区間の特徴の重みを示している。図 6 では、DM クエリの重みの絶対値が大きい傾向がある。しかし、購買直前・直後の重みの絶対値が相対的に小さいなど、4 節で得た購買満足に関する特徴を捉えられていないと考えられる。購買に関する特徴を捉えられていない原因として、アンダーサンプリングをした結果、十分な学習が行えず満足するユーザに共通する特徴を見つけ出せなかったことが推測される。

5.2 クエリ語による予測

クエリ語による予測では、購買前のクエリを空白で分割して得た Bag of Words を入力として満足度の予測を行った。この際、Web 検索クエリには購買に関係のない語も多く含まれるため、10 人以上のユーザが用いていた語に限定して Bag of Words を作成した。予測の正解率はランダムフォレストで 0.45、ロジスティック回帰で 0.53 となり、Web 検索クエリから満足度を予測できる可能性が示された。満足と不満足それぞれについて予測に効いている語を明らかにするため、より正解率の高いロジスティック回帰による予測における、重みの上位 5 語と下位 5 語を図 7 に示す。

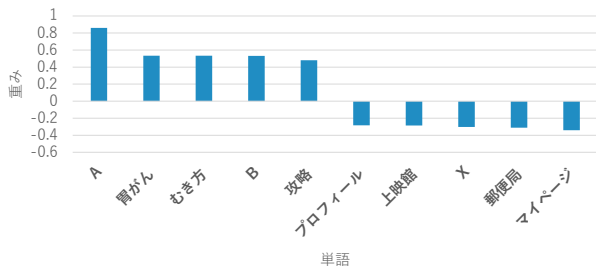


図 7 クエリ語による予測における重み

図 7 の縦軸は重みを、横軸は重みの上位 5 語と下位 5 語を示している。なお、図中の A はある総合 EC サイトを、B はあるカメラ専門店の EC サイトを、X はセンシティブワードを示している。図 7 より、重みの下位 5 語ではカメラの購買に関連する語が現れていないが、上位 5 語の中には総合 EC サイトやあるカメラ専門店の EC サイトがみられた。この結果から、満足したユーザは Yahoo!ショッピング以外の EC サイトにて商品と比較後、Yahoo!ショッピングで購入していると考えられる。

また、購買 1 週間前に 10 回以上用いられていたクエリ語にも着目した。その結果、メーカー名やブランド名は満足度に関わらず共通して見られたが、“価格”や“比較”などの語が購買に満足した場合はそうでない場合と比べよりも高い割合でみられた。この結果からも、購買前に EC サイトや商品について吟味することがユーザの満足につながると考えられる。

6 ま と め

本研究では、EC サイトにおける商品購入の前後に行われる Web 検索行動と購買満足度の関係分析を行った。具体的には、EC サイトで商品を購入したユーザの Web 検索ログをその検索意図に基づいて分類し、その時間変化が購買満足度によって異なるか調査した。その結果、購買に満足したユーザはそうでないユーザと比べ、購買前後 2 週間以内の頻度が高いことがわかった。また、価格の高い商品を購入する場合においては購買直前だけでなく、商品について長期的に調べているユーザの方が満足度が高い傾向がみられた。商品に対し高い興味を持っているユーザは、徐々に検索量を増やすような探索ではなく、検索量が減少していった後、購買直前に集中して調べ購買した場合に不満足になる傾向がみられた。分析で得た知見をもとに購買満足度の推定にも取り組み、購買前後に行われる Web 検索行動によって満足度が推測できる可能性があることを示した。

今後の課題として、カメラカテゴリ以外についての分析や、購買満足度の予測正解率の向上などが挙げられる。

謝辞 本研究は JSPS 科学研究費助成事業 JP18H03244, JP18H03494, JP21H03774, JP21H03775 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] PwC. December 2021 global consumer insights pulse survey. <https://www.pwc.com/gx/en/industries/consumer-markets/consumer-insights-survey.html>, (accessed 2021-12-21).

[2] Marsha L. Richins and Peter H. Bloch. Post-purchase product satisfaction: Incorporating the effects of involvement and time. *Journal of Business Research*, 23(2):145–158, 1991.

[3] Rodoula Tsotsou. The role of perceived product quality and overall satisfaction on purchase intentions. *International journal of consumer studies*, 30(2):207–217, 2006.

[4] Magnus Söderlund. Customer familiarity and its effects on satisfaction and behavioral intentions. *Psychology & Marketing*, 19(10):861–879, 2002.

[5] Ning Su, Jiyin He, Yiqun Liu, Min Zhang, and Shaoping Ma. User intent, behaviour, and perceived satisfaction in product search. In *WSDM*, pages 547–555, 2018.

[6] Qianyun Zhang, Shawndra Hill, and David Rothschild. Post purchase search engine marketing. In *WWW*, pages 663–670, 2018.

[7] Roger Hallowell. The relationships of customer satisfaction, customer loyalty, and profitability: an empirical study. *International journal of service industry management*, 7(4):27–42, 1996.

[8] Anders Gustafsson, Michael D. Johnson, and Inger Roos. The effects of customer satisfaction, relationship commitment dimensions, and triggers on customer retention. *Journal of marketing*, 69(4):210–218, 2005.

[9] Ruth Mugge, Hendrik N. J. Schifferstein, and Jan P. L. Schoormans. Product attachment and satisfaction: understanding consumers’ post-purchase behavior. *Journal of consumer Marketing*, 27(3):271–282, 2010.

[10] Todd A. Mooradian and James M. Olver. “i can’t get no satisfaction:” the impact of personality and emotion on post-purchase processes. *Psychology & Marketing*, 14(4):379–393, 1997.

[11] Sharon Hirsch, Ido Guy, Alexander Nus, Arnon Dagan, and Oren Kurland. Query reformulation in e-commerce search. In *SIGIR*, pages 1319–1328, 2020.

[12] Rohan Kumar, Mohit Kumar, Neil Shah, and Christos Faloutsos. Did we get it right? predicting query performance in e-commerce search. In *eCom@SIGIR*, 2018.

[13] Yangyang Guo, Zhiyong Cheng, Liqiang Nie, Xin-Shun Xu, and Mohan Kankanhalli. Multi-modal preference modeling for product search. In *MM*, pages 1865–1873, 2018.

[14] Bernard J. Jansen and Simone Schuster. Bidding on the buying funnel for sponsored search and keyword advertising. *Journal of Electronic Commerce Research*, 12(1):1–18, 2011.

[15] Beibei Li, Anindya Ghose, and Panagiotis G. Ipeirotis. Towards a theory model for product search. In *WWW*, pages 327–336, 2011.

[16] Andrei Broder. A taxonomy of web search. In *ACM SIGIR forum*, volume 36, pages 3–10, 2002.

[17] Bernard J. Jansen, Danielle L. Booth, and Amanda Spink. Determining the informational, navigational, and transactional intent of web queries. *Information Processing & Management*, 44(3):1251–1266, 2008.

[18] Nikitha Rao, Chetan Bansal, Subhabrata Mukherjee, and Chandra Maddila. Product insights: Analyzing product intents in web search. In *CIKM*, pages 2189–2192, 2020.

[19] Azin Ashkan and Charles L. A. Clarke. Characterizing commercial intent. In *CIKM*, pages 67–76, 2009.

[20] Wendy W. Moe. Buying, searching, or browsing: Differentiating between online shoppers using in-store navigational clickstream. *Journal of consumer psychology*, 13(1-2):29–39, 2003.

[21] Parikshit Sondhi, Mohit Sharma, Pranam Kolari, and ChengXiang Zhai. A taxonomy of queries for e-commerce search. In *SIGIR*, pages 1245–1248, 2018.