

目的指向推薦システム実現に向けた 商品特徴とユーザー目的の関連づけ

Sopheaktra YONG[†] 浅野泰仁[†] 吉川正俊[†]

[†] 京都大学大学院情報学研究科 〒606-8501 京都市左京区吉田本町 36-1

E-mail: [†]{yong@db.soc., asano@, yoshikawa@}i.kyoto-u.ac.jp

あらまし 近年、消費者の購買行動のかなりの部分が実店舗からオンラインショッピングサイトに移行しつつある。そのようなサイトでは、売り手側は商品の特徴や機能を表示しているが、良い点は記載しても悪い点は減多に記載されないのが、消費者は自分の目的に合っている商品(例えば、鳥を撮るのに向いたデジタルカメラはどれかなど)を見極めるのに時間を費やすことが多い。そのため、消費者はしばしば他のユーザーによるレビューを閲覧することになるが、複数の商品のレビューは何百もあることが多く、非常に時間がかかってしまう。この問題を解決するために、本研究ではまず、「目的指向推薦システム」のアイデアを提案する。これは、レビューを解析してユーザーの目的に最も合致する商品を、その理由とともに推薦するものである。このシステムの実現には、商品の属性とユーザーの目的を関連づける知識が必要になる。例えば、デジタルカメラの連写機能は鳥のように高速で移動する物体を撮影するのに有用であるという知識があれば、鳥を撮る目的のユーザーに連射の速いカメラを推薦できるからである。この知識抽出のためのレビュー分析手法を提案し、Amazon.com のデジタルカメラのレビューデータに適用した結果について考察する。

キーワード 推薦システム, テキストマイニング, レビュー分析

1. はじめに

ウェブが誕生して以来、オンラインショッピングは急激に普及し、現在ではごく日常的なものとなっている。従来の実店舗型のショッピングと比較して、オンラインショッピングは物理的な移動が不要なため労力や時間などの面で有利であることに加えて、豊富な商品のデータを検索・比較できることが便利である。さらに、オンラインショッピングサイトでは、購買決定を支援するための仕組みを複数用意していることが多い。例えば、Amazon では、売り手側から提供される商品情報、消費者の購買履歴に基づいて商品を推薦するシステム、商品の購買者によるオンラインレビュー等があり、消費者はこれも参考にすることができるようになっている。

推薦システムに使われている技術としては、消費者が注目している商品に似ている商品を推薦する手法や、注目している商品を買った他の消費者が買った別の商品や、購買履歴が似ている他の消費者が買った商品を推薦する手法などが一般的である。しかし、同じような商品で機能が微妙に異なるものも世の中には数多く存在し、その中で最も自分の使用目的(価格の安さや旅行向きの軽さ等も含む)に向けた機能を持った商品を選ぼうとすると、こういった推薦システムでは不十分が多い。機能を重視した推薦技術として、レ

ビューから商品の機能を自動的に抽出する手法や、機能に基づいて自動的に商品をランキングする手法はいくつか提案されている[2,3,9,10,]. これらを用いることで、例えばデジタルカメラなら、連写速度・重量などの各機能について、商品をランキングすることができる。

商品の機能と自身の使用目的の関係を熟知している消費者であれば、上記のように機能に基づいたランキングを用いて購買すべき商品を決定することがおそらくは可能である。しかしながら、使用目的に必要な機能に関する知識のない初心者は、そのようなランキングやショッピングサイトで提供される情報を用いても自分の使用目的に合致した商品を選ぶのは容易なことではない。これが顕著なのは、PC やスマホのように本質的に多くの使用目的がある商品であるが、カメラのように目的が一見単一(撮影)に見えるものであっても、実際は必要とされる機能が異なる数多くの目的がある商品もある。例えば Canon のデジタル一眼レフであれば、鳥などの高速動物撮影に向くのは連写速度と AF 性能に優れる EOS 1DX, 1DX2 や EOS 7D2 であり、細部が重要な風景の撮影には高解像度の EOS 5Ds が、幅広い分野の撮影には連写速度・AF・解像度・高感度ノイズのバランスに優れた EOS 5D3 が向いているなどである。他のメーカーのカメラについてもこのような向き不向きがある。こういった知識がない初心者が、

機能に基づいたランキングを与えられても、自身の目的に合致した商品を選択するのは困難である。他に参考になる情報としては、売り手側の情報とオンラインレビューであるが、その不十分な点について以下に詳しく述べる。

まず売り手側の情報は機能(特徴、性能等を含む)の記述が多く、使用目的に関してはユーザーによる違いもあるのでカバーしきれないことや、向いていない目的があえて記されていないことも多い。カメラで言えば風景(および夜景)・スポーツ・ポートレートなどの一般的目的に関してはメーカーの記述があることも多いが、他にも暗い体育館の球技・子供の運動会・カワセミ・昆虫・花・天体・建造物・ドッグラン・カーレース・航空機等々の数多くの目的が存在し、それらに関して必要な機能は微妙に異なってくるのだが、それら全てについて記述しているメーカーはないし、機能からこれらの目的に向いているかどうかを判断することは初心者には至難の業である。

次にオンラインレビューであるが、これは購買意思決定に非常に重要であることが知られている。Dimensional Research の調査[1]によれば、このようなサイトにおける購買意思決定の90%がオンラインレビューに影響されているとのことである。これは、売り手側の提供する情報と比較して、オンラインレビューの方が宣伝目的の記述が少なく、実際の使用感に関する情報が豊富であると消費者たちが考えているということであろう。確かに、上記の例に挙げたカメラやPCの様々な目的に関しても、それに必要な機能について、レビューで言及されていることは珍しくない。一方で消費者がレビューを読んで複数の商品の中から自分の目的に合致したものを選ぼうとすると、読むべきレビューの数は非常に多くなり、時間と労力がかかってしまう。

このような理由から、消費者の目的に適合した商品を自動的に選ぶ、またはその選択に役立つ推薦理由などの情報を発見・整理してくれるような、「目的指向推薦システム」があれば消費者の利便性を大いに向上させられると考えられるが、その実現は自明ではなく、現在でも挑戦的な課題となっている。本研究では、この「目的指向推薦システム」の実現に必要と考えられる技術として、オンラインレビューから「商品の使用目的」と「商品の機能」を関連づける知識をマイニングする手法を提案する。

提案手法は、まずその商品の属するカテゴリに含まれる全ての商品のレビューから、機能に関する記述を抽出する。次に、各機能について、それを含む記述から使用目的と機能を関連づけて説明している部分を抽出する。さらに、この手法を Amazon のレビューデー

タセット[12]のデジタルカメラに関する部分に適用し、その実験結果について考察する。

本論文の構成は次のようになっている。2 節では関連研究について述べる。3 節では目指すべき「目的指向推薦システム」の概要を提案する。「4 節ではオンラインレビューから「商品の使用目的」と「商品の機能」を関連づける知識をマイニングする手法について述べる。5 節では実験とその結果に関する考察を行い、6 節で結論を述べる。

2. 関連研究

消費者の意見、特にレビューから有用な潜在的知識を抽出することは学術的にも産業的にも重要な課題となってきた。レビューの中には、ユーザーの商品使用経験が書かれていることもあるし、商品の特定の特徴(機能)が良かったのか悪かったのかが書かれていることもある。こういった知識を抽出するのが特徴に基づくセンチメント分析である。これは特徴抽出、センチメント予測、センチメント分類、要約作成等を含む研究分野である [3]。特徴抽出は語彙統語(または文法)、あるいは統語と話法に基づいてレビューのテキストを分析する手法が多い[4, 5]。そのうち、教師ありの手法の例としては、Wong と Lam [6,7]による、隠れマルコフモデルと CRF(Conditional Random Fields)を学習法に用いてオークションサイトの文章から特徴抽出を行う手法が挙げられる。教師なしの手法としては、Hu と Liu [8] によるもの等がある。彼らは、商品の特徴は名詞か名詞句であると仮定し、アソシエーションルールマイニングを用いてレビューの集合から頻出するアイテム集合として特徴を抽出する。そして、この特徴を表す名詞または名詞句に隣接する形容詞がユーザーの意見を表すものとしている。さらにこれを用いてさほど頻出しない特徴を探索する。

特徴に基づく要約生成手法としては、Kangale ら [9] の、レビューのテキストから各特徴に関して正または負のレーティングとレビューの要約を生成する手法がある。また、Kamal [10]は OSVS(Opinion Summarization and Visualization System)と名付けられた、棒グラフと円グラフによって商品の特徴に関する多くのユーザーのセンチメントの要約を一目で見取ることができるシステムを提案している。Filho ら[11]は、各特徴に関して異なる意見を容易に比較できるグラフィカルユーザーインタフェースシステムである XOpin を提案している。

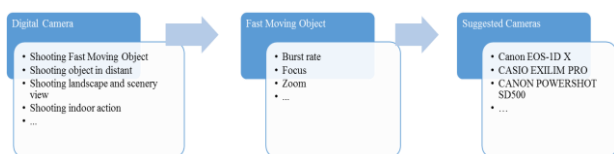
本研究と最も関連があると考えられるのは、Zhang ら[2]による、各特徴ごとに推薦商品のランキングを生成する手法である。この手法は、特徴に関する語を含む各文の中から、あらかじめ与えた構文パターン(as

<word> as, the same as 等)によって、比較される商品を発見する。前節で述べたように、本研究は彼らのものと以下の点で異なっている。(1) 本研究で対象としているのは、商品の特徴(機能)に関してあまり知識のない初心者である。したがって彼らはどの特徴に基づいて商品を選択すれば良いかわからないため、彼らの手法だけでは商品を選択できないと考えられる。(2) 各特徴や商品全体のレビューの要約を生成する代わりに、特徴(機能)がユーザーの使用目的実現にどう役立っているかに着目する。(3) この、特徴(機能)と使用目的の関連づけに基づいて、将来は「目的指向推薦システム」を構築していく予定である。

3. 目的指向推薦システム

前述のように、商品が技術的特徴や機能を数多く備えていると、初心者は自身の目的を達成するため買うべき商品を決定する際に困難に直面する。図1は、我々が将来において目指している「目的指向推薦システム」のフローを図示したものである。

図1 「目的指向推薦システム」のフロー



システムはまず、与えられた商品のカテゴリ(例:デジタルカメラ)に対して、可能な目的の候補のリストを提示する。ユーザーはその中から自身の目的をいくつか選択する(テキストで入力された独自の目的に対応することも別の実現目標として考えられる)。するとシステムがその目的に最も関連した特徴・機能を出力する。ここで、それぞれの特徴や機能が、どうして目的に重要なのか、その理由と程度(最低限必要な値や、これだけあれば十分等)も提示することができればさらに理解に役立つと考えられる。最終的に、システムはこれらの特徴・機能に基づいて、推薦する商品を価格・種別(デジタルカメラであれば一眼レフ、ミラーレス一眼、レンズ一体式)等でグループ分けして表示する。また、ユーザーに自身のカテゴリを入力してもらい、それをグループ分けに役立てるという方法も考えられる。ユーザーのカテゴリとしては、その商品を短期的に使うのか、あるいは長期的に使うのか、等が考えられる。長期的に商品を使用することを考えているユーザーは、現在はその商品カテゴリに関して入門レベルであっても、将来的にはプロレベル等に達するまで経験を積み重ねたいと考えられる。したがって、そのようなユーザー

には、現在の目的を満たすだけでなく、その先の発展まで考慮した商品を勧めた方が良いかもしれない。例えばデジタルカメラであれば、そのようなユーザーには、レンズ一体式のカメラよりも、撮影そのものに詳しくなるために必要な各種のマニュアル撮影モードと、拡張性を確保するレンズ交換可能なマウントを備えた一眼レフを勧める方が良いと考えられる。一方で短期的使用であれば、その目的を達成するのに十分であれば、安価なレンズ一体式カメラを勧めた方が良いでしょう。

このシステムの肝となる部分は、レビューデータに基づいて、商品の特徴(機能)とユーザーの目的を、関連づけるところにある。次節では、その実現のために、我々が現在開発を進めている手法を説明する。

4. 商品の特徴とユーザーの使用目的を関連づけるアルゴリズム

4.1 商品の特徴の識別

ここで対象にする特徴とは、商品に機能を与える属性である。異なる商品特徴は異なる機能を持ち、異なる使用目的に関連づけられる。商品の特徴を抽出すること自体は前節で述べた関連研究等を用いればある程度可能なので、これらの手法ないし手動で商品の特徴を抽出することを前提とする。他にも、kakaku.com、cnet.com、dpreview.com、consumerreports.org 等のようなデジタルカメラの機能を紹介するサイトまたは商品の特徴から商品を検索できるサイトの情報を用いて、自動的に特徴を抽出することも可能と考えられる。本稿の実験では、Zhangら[2]が手動で選んだデジタルカメラの特徴を用いている。表1は、その特徴と同義語である。

表1: デジタルカメラの特徴と類義語

| |
|---|
| Digital Camera |
| resolution/ pixel/ megapixel |
| lens/ wide angle/ normal range |
| optical/ zoom/ optical zoom/ digital zoom |
| memory/ megabytes/ MB |
| burst/ continuous/ shutter/ recovery/ motion/ sport |
| battery/ batteries/ power |
| focus/ exposure/ manual/ iso |
| LCD/ screen |
| compression/ compress/ jpeg |
| flash/ light |

これらの語には、撮影に重要な要素である撮像素子のサイズ、ダイナミックレンジ、ローパスフィルター

の有無、手ぶれ補正、レンズ交換式のカメラを考慮したレンズの各種性能(レンズ自身の AF 速度や解像度に加え、F 値・ぼけ味や歪曲・周辺減光・各種収差等)、カメラが搭載する現像アルゴリズム(超解像技術, HDR, 収差補正技術)などを含んでいないが、それでも今回の手法のテストには十分な種類の特徴語を含んでいると考えられる。

4.2 使用目的に関係する文の抽出

今回使用したレビューのデータは 1,494,938 文を含むが、その約 80% が上記の特徴またはその類似語を 1 個以上含んでいた。これらの文から、特徴と使用目的を両方含んでいるような文を抽出する手法を提案する。手法を上記のレビューデータに適用した結果として、約 1,000 の文が得られた。

- まず、各レビューは複数の文が含まれていることが多いので、Punkt Sentence Tokenizer [15] を用いてレビューを文に分割する。このうち、特徴を含むものだけを残す。
- 特徴を含むレビュー文は、しばしばその特徴が特定の目的にとって良いか悪いかを示す表現を含んでいる。実際にデータに含まれていた例としては、“*The photo burst mode is also nice for catching action shots.*” という文がある。ここで、“*Catching action shots*” の部分が、“burst” (連射)機能を用いる目的となっている。このレビューの断片は、高速で移動する物体を撮影するためのカメラを購入することを考えているなら、連射速度の速いカメラを選ぶべきであるということを示唆している。したがって、特徴を含む文から、その特徴の利用に関する表現を含む文を抽出することが次のステップとなる。そのために、商品の特徴について述べた文章の各単語の part-of-speech (POS)を解析することで、これを実現する。本手法では、あらかじめ手動で構築された構文パターン {Adjective Modifier (JJ, VBD, VBG, VBN) or Adverb Modifier (RB, RBR, RBS) + “for” + Gerund/Noun (VBG, NN)} を含む文を抽出する。POS タグを各単語に割り当てるためには、NLTK3.0 [14] を用いた。

なお、他のパターンも検討したが、現在のところはこれが最良の結果となった。将来的には、手動で構築されたパターンからブートストラップのような方法を用いて自動的にパターンを拡張してゆく方法や、使用目的や特徴の活用法を暗に含んでいるような文を発見する方法も検討する予定である。

4.3 商品特徴と使用目的の関連

上記で得られた文には、様々な使用目的が含まれているが、ユーザーによる表記の違いや、上記の構文パターンに当てはまらない文に含まれる目的等もあるため、かなりの間違いをも含んでしまっていることがわかった。これを整理するために、ある特徴を他の特徴と区別できるような重要語を取得する方法を提案する。例えば、“family” と “group” という語は “wide angle” (広角) という特徴と強い関連がある。これは、広角レンズが家族写真や集合写真のように、撮影対象が数多い場合に適しているからだと解釈できる。そこで、提案手法では、以下のように Labeled LDA (Latent Dirichlet Allocation) および Word2Vec を利用することを考えた。

- Labeled LDA (LLDA) による重要語の取得: 4.2 節の方法で得られた文を LLDA における「文書」、各特徴を「ラベル」とみなして、トピックモデリングを行う。これによって、各特徴について重要な語を得ることができる。LLDA の実装としては、Stanford Topic Modeling Toolbox [16, 17] を用いた。
- Word2Vec によるクリーニング: 上記で得られた重要語の中には雑音が含まれることもある。これらの雑音の多くが、他の重要語とは関連のあまりない語であった。そこで、Tomas Mikolov らによって開発された、単語をベクトル化する手法である Word2Vec [18, 19] を利用して、これらの雑音を除去する手法を考案した。例えば LLDA の結果で得られた単語集合が {breakfast cereal dinner lunch} であったとすると、Word2Vec を使って互いに似ていない “cereal” を除くことができるからである。モデルの訓練データとしては、レビューデータセットのテキストから一般的すぎる単語やストップワードを NLTK3.0 [14] を用いて除いたものを使った。

5. 実験と考察

本節では、Amazon.com の英文レビューのうちデジタルカメラのみからなるデータセットを用いた実験結果とそのケーススタディについて述べる。表 2 は、4.2 節に述べた手法を適用して得られた、“burst(連写)” 機能と使用目的の対を含む文に関するもの 29 個のうち、一部を表している。

表 2: "burst"機能の使用について消費者が何を述べているか

| No. | Functions or Purposes |
|-----|---|
| 1 | also excellent for capturing that first baby smile |
| 2 | also nice for catching action shots |
| 3 | awesome for capturing goofy expressions |
| 4 | awesome for taking pano shots which I do a frequently |
| . | ... |
| 28 | was great for taking pictures of our eight month old daughter |
| 29 | worked great for action |

得られた結果は、連写機能が重要になってくる使用目的を表しているものがほとんどであった。しかしながら、#1と#28のように、同じ目的(この場合は子供の撮影)について異なる表現を用いているものもいくつか存在した。

これらの結果を改良するために、4.3節で述べたように labeled LDA と word2vec を用いた結果を以下で説明する。

- **Labeled LDA**

表 3 は、"burst"トピックを構成する重要な単語とそのスコアである。左が 4.2 節の手法を適用した結果得られた、使用目的を表す部分文の集合に、Labeled LDA を適用した結果であり、右がそれ以外の部分も含めた文全体に対する結果である。この結果を見ると、(1) Labeled LDA は、文の一部のように短いドキュメントにもうまく対応できている (2) 4.2 節の手法を適用した方が、良い結果が得られている、ということがわかった。以下、その詳細を述べる。

表 3: Labeled LDA – 使用目的に関係する文と、それ以外も含めたすべての文に対する結果

| Features | Purpose-related sentences | Whole sentence |
|----------|---|---|
| Burst | capturing 5.0 kids 3.0 catching 2.0 getting 1.0 recording 0.0 | burst 14.0 feature 5.0 love 3.0 second 2.0 photo 1.0 batteries 0.0 |

Zhang ら[2] の研究によれば、"burst"機能は高速で移動する物体を撮影するのに使われるものである。表 3 の左(使用目的に関係する部分文の結果)で得られている"kids"は、実際に高速で移動する物体である(走る動きや、移り変わる表情等)。しかしながら、表 3 の右(使用目的に関係する文全体の結果)では、そのように高速移動物体に関する語が現れていない。したがって、上記(1)(2)の結果が観察さ

れたと言える。一方、表 3 の左に出現する "capturing", "catching", "getting" という単語は、それら自身だけでは、何を意味するのかがわかりにくいという問題がある。そこで、4.3 節で述べた、Word2Vec によるクリーニングが必要になると考えられる。以下では、その結果について述べる。

- **Labeled LDA + Word2Vec**

表 4 は先ほどの Labeled LDA の結果(左列)と、その結果から、Word2Vec を利用して外れ値を除いたもの(右列)である。

Table 4: LLDA と LLDA+Word2Vec の結果

| Features | LLDA | LLDA + W2V |
|----------|---|-------------------------------|
| burst | capturing kids catching getting recording | capturing kids catching |

この手法で得られた(除かれずに残った)単語を評価するために、表 4 の右列の 2 単語以上を含む文の一部を手動抽出して、ケーススタディを行う。表 5 は "burst" 機能について得られた結果である。

表 5: "burst"機能に関して得られた使用目的

| Features | Terms combination | Purposes |
|----------|-------------------|--|
| burst | capturing, kids | -fantastic for capturing my kids' sports activities -nice for capturing kids running wild |
| | kids, catching | -great for catching kids who change their expression quickly |

このように、上記の 2 単語以上を含む文は、使用目的を良く説明していることが多かった。この結果を熟練者やプロに評価してもらうことも検討している。

6. むすび

本論文では、使用目的に必要な機能に関する知識のない初心者を助ける「目的指向推薦システム」の必要性について述べた後、その実現に必要な技術として、商品の特徴と使用目的を関連づける知識を抽出する手法を提案した。この手法では、まず商品の特徴の集合を与え、それを含むレビューから使用目的に関係する文または文の一部を抽出する。提案手法を Amazon.com のデジタルカメラのユーザーレビューに適用し、その結果に関する考察を行った。結果として、初心者の購

入決定に役立つような知識が得られた。

将来的には、商品の特徴を非明示的に記述しているような文章からも特徴と目的を自動的に抽出する手法を提案したい。例えば

This camera is too small to store printable pictures.

という文は、カメラの記憶容量という特徴に関する文であると考えられるが、“storage”や“memory”という明示的な単語は出現していない。これらの特徴も抽出できれば、さらに得られる知識が豊かになると考えられる。さらに究極的な目標としては、今回の結果をさらに発展させて、目的指向推薦システムを実際に構築することが挙げられる。

謝辞

本研究はJSPS 科研費 15K00423 および 栢森情報科学振興財団の助成を受けたものです。

参考文献

- [1] Marketing Land, April, 2014 <http://marketingland.com/survey-customers-more-frustrated-by-how-long-it-takes-to-resolve-a-customer-service-issue-than-the-resolution-38756> (visited on: Jan 6, 2016)
- [2] K. Zhang, R. Narayanan, and A.N. Choudhary, "Voice of the customers: mining online customer reviews for product feature-based ranking." In Proceedings of the 3rd Conference on Online Social Networks, pp. 1-9. 2010.
- [3] M. Hu and B. Liu, "Mining opinion features in customer reviews." In Proceedings of the Nineteenth National Conference on Artificial Intelligence, Sixteenth Conference on Innovative Applications of Artificial Intelligence, pp. 755-760. 2004.
- [4] A. Das and S. Bandyopadhyay, "Topic-based Bengali opinion summarization." In Proceedings of the 23rd International Conference on Computational Linguistics: Posters, pp. 232-240. Association for Computational Linguistics, 2010.
- [5] A. Abbasi, H. Chen and A. Salem, "Sentiment analysis in multiple languages: Feature selection for opinion classification in Web forums." ACM Transactions on Information Systems (TOIS) Vol. 26, No. 3, pp.12:1-12:34, 2008.
- [6] T.L. Wong and W. Lam, "Hot item mining and summarization from multiple auction Web sites." In Proceedings of the 5th IEEE International Conference on Data Mining, pp. 797-800, 2005.
- [7] T.L. Wong and W. Lam, "Learning to extract and summarize hot item features from multiple auction web sites." Knowledge and Information Systems 14, no. 2, pp.143-160, 2008.
- [8] M. Hu and B. Liu, "Mining and summarizing customer reviews." In Proceedings of the tenth ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, pp. 168-177, 2004.
- [9] A. Kangale, S.K. Kumar, M.A. Naeem, M. Williams and M.K. Tiwari, "Mining consumer reviews to

generate ratings of different product attributes while producing feature-based review-summary." International Journal of Systems Science (2015): 1-15.

- [10] A. Kamal, "Review Mining for Feature Based Opinion Summarization and Visualization." arXiv preprint arXiv:1504.03068 (2015).
- [11] C. Brun and G. Rondeau, "A graphical user interface for feature-based opinion mining." In Proceedings of the 2012 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies: Demonstration Session, pp. 5-8. Association for Computational Linguistics, 2012.
- [12] J. McAuley, R. Pandey and J. Leskovec, "Inferring networks of substitutable and complementary products." In Proceedings of the 21th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, pp. 785-794. ACM, 2015.
- [13] C.P. Wei, Y.M. Chen, C.S. Yang and C.C. Yang, "Understanding what concerns consumers: a semantic approach to product feature extraction from consumer reviews." Information Systems and E-Business Management, Vol. 8, No. 2, pp. 149-167, 2010.
- [14] Natural Language Toolkit, <http://www.nltk.org/>
- [15] Punkt Sentence Tokenizer, <http://www.nltk.org/modules/nltk/tokenize/punkt.html>
- [16] Stanford Topic Modeling Toolbox, <http://nlp.stanford.edu/software/tmt/tmt-0.4/>
- [17] D. Ramage, D. Hall, R. Nallapati and C.D. Manning, "Labeled LDA: A supervised topic model for credit attribution in multi-labeled corpora." In Proceedings of the 2009 Conference on Empirical Methods in Natural Language Processing, pp. 248-256. Association for Computational Linguistics, 2009.
- [18] Gensim Word2Vec, <https://radimrehurek.com/gensim/models/word2vec.html>
- [19] T. Mikolov, K. Chen, G. Corrado and J. Dean, "Efficient Estimation of Word Representations in Vector Space". In Proceedings of Workshop at International Conference on Learning Representations (ICLR), 2013.

付録

表 6: Part-Of-Speech タグとその意味

| Tags | Description |
|------|------------------------------------|
| JJ | Adjective |
| NN | Noun, singular or mass |
| RB | Adverb |
| RBR | Adverb, comparative |
| RBS | Adverb, superlative |
| VBD | Verb, past tense |
| VBG | Verb, gerund or present participle |
| VBN | Verb, past participle |