

ブログの言及記事数を用いた売上予測手法の提案

稲本 伸二[†] 手塚 太郎[†] 木村 文則[†] 前田 亮[†]

[†]立命館大学 情報理工学部 〒525-8577 滋賀県草津市野路東 1-1-1

E-mail: [†] cm001062@ed.ritsumei.ac.jp, {tezuka, amaeda}@media.ritsumei.ac.jp, fkimura@is.ritsumei.ac.jp

あらまし ブログはリアルタイムに消費者の意向を把握することや社会における関心の推移の詳細な分析をするのに適した媒体として注目されている．そこで本研究ではブログと実世界の動向との相関分析法の一つとして，ある商品に対するブロガーの注目情報の移り変わりと，商品の売上の相関の傾向を分析し，その結果から商品のおおよその売上高の推移を予測する手法を提案する．

キーワード ブログマイニング，ウェブログ，マーケティング，相関分析

Sales Prediction Method Using the Quantity of Reference Articles on Blogs

Shinji INAMOTO[†] Taro TEZUKA[†] Fuminori KIMURA[†] Akira MAEDA[†]

[†] College of Information Science and Engineering, Ritsumeikan University

1-1-1 Nojihigashi, Kusatsu-shi, Shiga, 525-8577 Japan

E-mail: [†] cm001062@ed.ritsumei.ac.jp, {tezuka, amaeda}@media.ritsumei.ac.jp, fkimura@is.ritsumei.ac.jp

Abstract Blogs are focused as media that is suitable for figuring out real-time interests of customers and to analyze social trends. We therefore observe trends of blogger's focus on items and analyze mutual relation with sales. We also propose a prediction method of rough sales trends using the analyzed results.

Keyword Blog mining, Weblog, Marketing, Correlated analysis

1. はじめに

インターネットの普及に伴い，一般の多くの人々からの情報発信が盛んになり，その発信されている大量の情報を有効に活用したいという要求も高まっている．近年注目されている情報源としてブログがある．ブログは個人の日記としての情報が記されており，商品やサービスに関してマーケティングを行う上で有益なメディアとして注目を集めている[1]．ブログは，リアルタイムに消費者の意向を把握することや社会における関心の推移の詳細な分析をできる可能性があり，ブログマイニングの研究は，Web研究の代表的な研究領域のひとつとして確立しつつある．

ブログマイニングとは，巨大なブログデータの集合やデータベースから潜在する構造やパターンの規則性を抽出する技術である．それにより得た情報を販売戦略や商品企画など，実世界の身近なところに生かすことが期待されている．実際に，水口らはeHyobanにより，ブログ上の評判情報を検索する評判情報を分析する手法を提案した[2]．その手法では，自然文テキストから「良い」「悪い」などの評価表現だけでなく，「液晶」が良いなど，注目情報が抽出できる．

ブログが研究領域として発展してきた背景として，

リアルタイム性という特徴が注目されてきた点が挙げられる．リアルタイム性とは，ブラウザ上で記述した内容が即時に記事として反映される特徴である．奥村は，blogWatcherにより，ブログ空間上にある話題となるキーワードが急激に出現する様子を視覚化する手法を提案した[3]．それにより，社会で話題になっているキーワードの推移をリアルタイムに追跡することができるようになった．それに加え近年ブログが普及したことも考慮すると，ブログ集合は豊富に記事が揃っているため，ある時点で話題になった語を抽出し傾向を計測するのに適している媒体であるといえる．

一方，今日の社会において社会の関心動向を知るとは，多くの店舗の販売活動などにおいて重要である．川添らは，消費者が書くブログは，匿名性が高いという意識があるために本音を書き易く，消費者ニーズを発見するための低コストかつ効率的な情報入手の手段として適していると述べている[4]．

商品の売上には，商品の質の高さはもちろんのこと，著者・宣伝効果などのいくつかの要因が関係している．しかしながらどんな商品であっても，一定の期間がたつと，やがて売上は頭打ちとなり，市場から消えていく．その売上の推移と，ブログ上の商品に対して言及

した記事数(以下、言及記事数とする)との対応付けを調査してみると、一定のパターン化が出来る。それぞれに対し相関があることが分かる。そこで、もし商品の売上の推移を事前に知ることができれば、店舗のマーケティング担当者は適切なタイミングで適切な数の商品を仕入れることが可能となる。本手法では、言及記事数を用いて、商品の売上動向を予測する手法を提案する。

2. 関連研究

Gruhl らの論文“The predictive power of online chatter”[5]では、Amazon における書籍の売上ランキングとブログ上での言及記事数の推移の相関を分析している。実験結果から、書籍の売上の盛り上がりと言及記事数の盛り上がりには相関がみられることを示しており、特に商品の発売直後は強い相関が見られることを発見した。

しかし、その手法が適用できる商品は、言及記事数の推移が突発的に盛り上がったものに限定されているという課題があった。商品の売上の推移パターンは、突発的に盛り上がる以外にもさまざまなパターンがある。例えば商品の発売後何ヶ月もわたりロングヒットするパターン等がある。そのため、売上の推移パターン別に分析する必要があると考えられる。そこで、本手法ではさまざまな売上推移パターンにも柔軟に対応できる予測手法の確立を目指す。

次に福原らの論文“Weblog から社会の関心を探る”[6]ではブログ記事に含まれる単語の出現頻度の推移がパターン分類できることを示した。その分類結果を図 1 に掲載した。

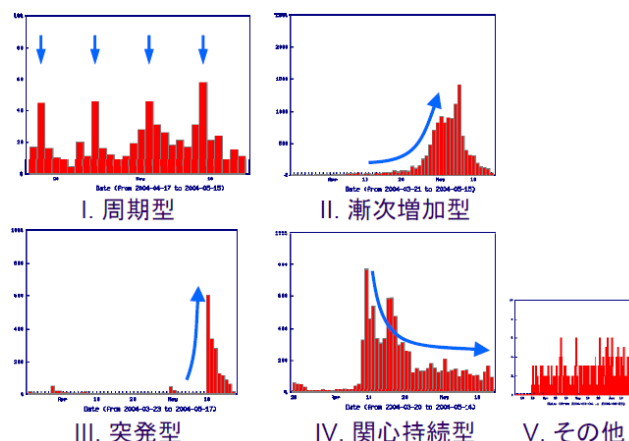


図 1 ブログ記事に含まれる単語の出現頻度の推移の分類結果[6]

単語の出現頻度の傾向によって分類した 5 つのパターンが挙げられている。それぞれのキーワード例を紹介する。I.では「給料日」、「休日」、「BBQ」など周期的な出来事、II.では「クリスマス」、「選挙」、「GW」など予期された出来事、III.では「地震」、「津波」、「脱線」など不測の出来事、IV.では「年金問題」、「イラク」、「ニッポン放送」など関心が続く出来事、V.では「酸性雨」、「大気汚染」、「水不足」など関心の薄い問題や「事故」、「セキュリティ」など一般的な話といったものがある。このパターン分類できる性質を利用し、商品の売上の推移パターンごとに言及記事数の推移パターンが異なる傾向がみられると仮定し分類することで、本提案手法では一定の精度での売上予測を可能にしている。

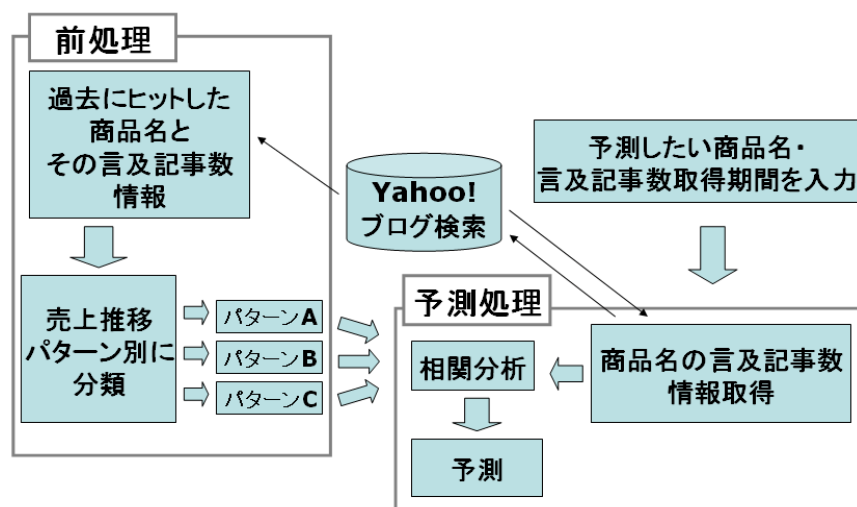


図 2 提案手法の枠組み

3. 売上推移の予測手法の流れ

本章では売上推移の予測手法の流れを簡単に述べた後、それぞれ個別の処理の詳細を述べる。まず、提案手法の枠組みを図2に示す。

3.1 前処理

前処理では、予測したい商品ジャンルで過去にヒットした商品名の一覧を用意し、それぞれの発売日前後の言及記事数の推移を Yahoo!ブログ検索[7]から取得する。そして、その言及記事数の推移をいくつかの売上推移パターン別に分類する。それが終わると、前処理が終了する。

3.1.1 売上推移パターン分類の定義

本手法の予測結果として出力する売上推移パターンを定義する。定義方法は売上ランキングに何週間ランクインしたかで判別する。これは、商品の人気がどの程度まで持続するかを予測するためである。売上の推移パターン分類の定義は以下のとおりとする。

(1)特大ヒット型

発売後 9～12 週間ランクイン

(2)大ヒット型

発売後 5～8 週間ランクイン

(3)中ヒット型

発売後 2～4 週間ランクイン

(4)小ヒット型

発売後 1 週間ランクイン

なお、ここで用いる売上ランキングのデータ取得元や(1)～(4)の型のランクイン継続時間の定義は予測したい商品ジャンルによって適宜変更することが望ましい。

3.1.2 言及記事数の推移パターンの分類方法

福原らはブログ記事に含まれる単語の出現頻度の推移がパターン分類できることを示した[6]。分類は、単語の出現頻度が周期性を持つもの、序々に増加し続けるもの、急激に増加するものなどが挙げられている。このパターン分類できる性質を利用し、商品の売上の推移パターンごとに言及記事数の推移パターンが異なる傾向がみられると仮定し分類する。

まず過去にヒットした商品群について、先に 3.1.1 で分類した 4 つの売上推移パターンごとに商品を分類しておく。その後、それぞれのパターン別に言及記事数の推移に平均をとった値の推移を、言及記事数の推移パターンとする。

3.2 予測処理

予測処理では、ある商品に対して予測する際には、言及記事数を基に該当する定義した売上パターンの特定を行う。特定された売上パターンから、その商品の売上予測を行う。

3.2.1 相関分析

二つの標本列 $\{x_i \mid i = 1, \dots, N\}$ および

$\{y_i \mid i = 1, \dots, N\}$ が与えられた場合、これらの標本列の類似度を測る統計量として相関係数 ρ がある。それは、

$$\rho = \frac{\frac{1}{N} \sum_{i=1}^N (x_i - \bar{x})(y_i - \bar{y})}{\sigma_x \sigma_y}$$

で定義される。ただし、 σ_x および σ_y は、それぞれ、

$\{x_i \mid i = 1, \dots, N\}$ および $\{y_i \mid i = 1, \dots, N\}$ の標準偏差である。

相関係数は、推移データの類似度合いを分析するのに適しており、さまざまな言及記事数の推移パターンの類似度合いを分析する本手法に適しているといえる。

そこで本手法では、相関係数を用いて売上推移のパターン分類を行う。

3.3 Yahoo! ブログ検索

Yahoo!ブログ検索とは、ヤフー株式会社が運営しているブログ検索エンジンである。Yahoo!ブログ検索では、クローラと呼ばれる Web ページを自動取得していくエージェントを利用して大量のブログ記事を逐次取得している。ユーザは、キーワードを入力して検索ボタンを押すと、キーワードが含まれる記事群の情報がブラウザに返される。このキーワード検索にはオプション機能があり、検索対象期間が指定できる。

本手法では、ブログから商品の言及記事情報を取得するために Yahoo!ブログ検索を利用し、記事内に商品名を含む記事数を、必要となる投稿時間ごとに取得する。

4. 評価実験

本手法を用いて、過去の商品の売上と言及記事数をもとに、予測したい商品の売上の推移パターンがどれだけ予測できるかを検証するため、また、予測できないケースはどのような場合かを明確化するため、実験を行った。

4.1 実験 1(発売 3 ヶ月前～発売日/発売後 1 ヶ月)

今回の実験では、2008 年 4 月～2009 年 2 月までに週間売上 TOP30 に入った商品をもとに、2009 年 3 月～12 月までに発売された商品の発売 3 ヶ月前～発売 1 ヶ月後までのデータと発売 3 ヶ月前～発売日までのデータの 2 パターンで売上推移パターンを予測する。

4.1.1 実験に用いたデータ

実験で扱う商品ジャンルとして、文庫本を用いた。文庫本は、そのタイトルがブログ上で語られやすいという性質を持っており、ブロガー同士で商品を紹介しあうことや、商品レビューすることが頻繁に行われている。また取得できる商品名の数が多かったため、発売日前後の言及記事数を重視する本手法に適していると考えられたので実験データとして用いた。以下に、実験に用いたデータの詳細を述べる。

(1)過去にヒットした商品とその売上

売上データは、ORICON STYLE[8]で提供されているものから取得して用いた。

ORICON STYLE とは、エンターテインメントに関する商品の売上ランキング情報を掲載しているサイトである。ORICON STYLE には大きくわけて楽曲・映像・本・ゲームソフトの 4 つのジャンルのランキングデータがある。それぞれ一週間ごとの売上の上位 30 位までの商品名や売上などの情報が掲載されている。2010 年 1 月現在では、2008 年 4 月までの売上データが参照できる。

そこに掲載されている 2008 年 4 月から 2009 年 2 月までに週間売上 TOP30 に入った商品の約 600 本の名前と発売後一週間ごとの売上を用いた。なお、TOP30 のランク外の商品の売上本数は取得することができなかったため 0 とした。

(2)言及記事数

(1)の売上データで取得した 2008 年 4 月から 2009 年 2 月までに発売され週間売上 TOP30 に入った商品それぞれに対し、発売前後 12 週間の言及記事数を Yahoo!ブログ検索から週単位で取得した。取得した結果を売上推移パターンごとに分類したあと、それぞれに対し言及数の推移をグラフ化した結果を図 3 に示す。

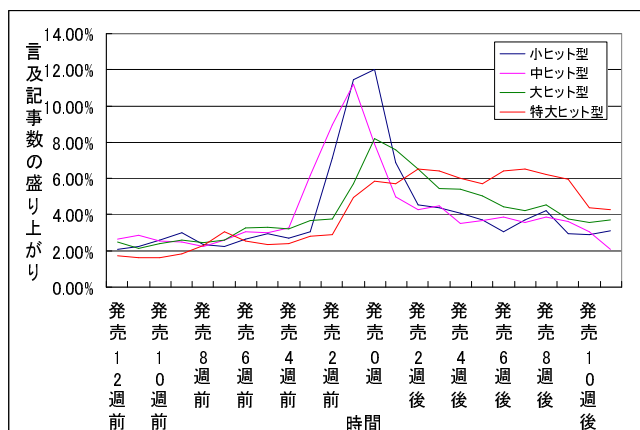


図 3 売上推移パターンごとの言及記事数の推移

ここで、Y 軸として用いている言及数の盛り上がりとは、各計測期間における全商品の言及記事数の総数に対する割合を示している。

(3)売上予測する商品

2009 年 3 月から 12 月までに発売され週間売上 TOP30 に入った商品約 450 本に対し、それぞれの売上の推移パターンを予測した。

(4)検索語の決定方法

ORICON STYLE の文庫本のランキングページに表示されている商品名の多くは、「メインタイトル(読み仮名)」「副タイトル(読み仮名)」「ナンバリング」で構成されており、それらが順に記述されている。

ブロガーは商品に対して言及するとき、商品名をフルネームで語ることは少なく、ある程度省略された形で語る傾向がある。そのため、この手法において商品ごとに言及記事数の取得する際には「メインタイトル」のみを抽出し用いた。その抽出作業で除去した文字列の一覧を以下に示す。

(a) 商品名の末尾に書かれたナンバリング文字

- ・巻数を示すローマ数字、アラビア数字、漢数字
- ・本の位置付けを示す「上」、「中」、「下」の文字

(b) 一度目に出現するスペース以降の文字列

- ・副タイトル
- ・バージョン情報

(c) 括弧とそれに囲まれた文字

- ・読み仮名
- ・副タイトル

4.1.2 実験結果

それぞれの商品に対し売上推移パターンの予測を行い、その正解率を表 1 に示す。

表 1 予測した結果の正解率

	正解率	
	発売 3 ヶ月前から 発売後 1 ヶ月まで	発売 3 ヶ月前 から発売日まで
小ヒット型	54%	53%
中ヒット型	65%	61%
大ヒット型	53%	27%
特大 ヒット型	50%	13%

4.2 実験 1 に対する考察

4.2.1 正解率に対する考察

表 1 の結果により、あまり高い確率で予測できるとはいえないものの、言及記事数の推移データから売上の推移パターンがある程度予測できることが判明した。また、発売後 1 ヶ月までのデータと発売日までのデータそれぞれにおける正解率を比較すると、データ予測期間が長くなるほど精度が大きく上がることがわかった。これは、ブログ上での言及記事数の推移の傾向をとる上で、商品の発売後も重要な要素であるといえる。しかし表 1 の結果では、ビジネスシーンで活用するに足る確率で予測できているとはいえない。そこで、より精度を向上させるために、過去商品の言及記事数の推移を分析し考察を行った。

4.2.2 過去商品の売上推移パターンごとの言及記事数の推移についての考察

図 3 の過去商品の売上推移パターンごとの言及記事数の推移について考察する。その結果、売上推移パターンそれぞれのもっとも言及記事数が盛り上がる時期が異なることが読み取れる。また、小ヒット型・中ヒット型は盛り上がり型が急激であり、それに対し大ヒット型・特大ヒット型の盛り上がり方は緩やかだった。これらのことから、売上推移パターンそれぞれの言及記事数の推移には一定のパターン化が出来ることが分かった。

しかし、発売日の 12 週間前から数週間は、どの売上推移パターンも同じような言及記事数の推移を示す結果になった。その推移の様子を具体化するため、予測に用いた 2008 年 4 月から 2009 年 2 月までに週間売上 TOP30 に入った商品群それぞれの発売日の 12 週間前から発売日の 3 週間後までの言及記事数の推移を平均化しグラフ化した。それを図 4 に示す。

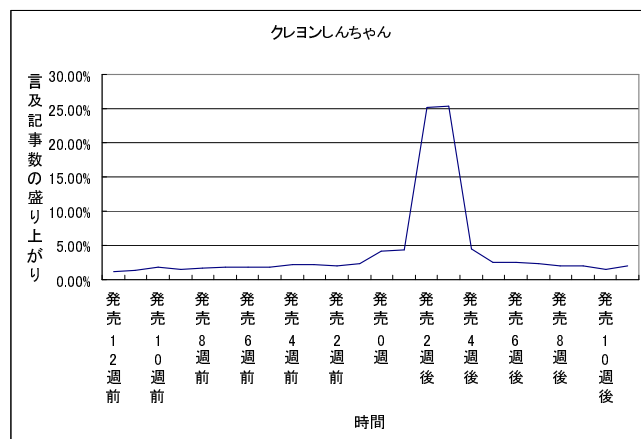


図 4 過去商品の言及記事数の盛り上がりの推移

多くの商品の言及記事数は、発売日の 12 週間前から発売日の 4 週間前まで序々に上昇していき、その後、発売の 4 週間前から発売の週にかけて突発的に盛り上がり、そして発売 1 週間後から急激に下降していく傾向があることが図 4 から読み取れる。このことから、売上予測において、言及記事数の推移の変化があまりみられない期間のデータを予測に用いていたことが正解率を下げている要因だと考えられる。そこで、言及記事数の推移の変化があまりみられない期間の言及記事数の推移を予測データから取り除き、言及記事数の推移の変化の激しい期間のみに着目し、実験を行うことにした。

4.3 実験 2(発売 4 週間前～発売後 3 週間)

実験 1 で得た結果と考察から、さらに精度を上げるため実験方法を変更し検証を行った。今回の実験では、2008 年 4 月から 2009 年 2 月までに週間売上 TOP30 に入った商品をもとに、2009 年 3 月から 12 月までに週間売上 TOP30 に入った商品の発売日の 4 週間前から発売日の 3 週間後までのデータで売上推移パターンを予測する。

4.3.1 実験に用いたデータ

商品ジャンル・過去にヒットした商品とその売上・売上予測する商品は実験 1 と同じものを用いた。

4.3.2 実験結果

それぞれの商品に対し売上推移パターンの予測を行い、その正解率とその比較のグラフを図 5 に示す。

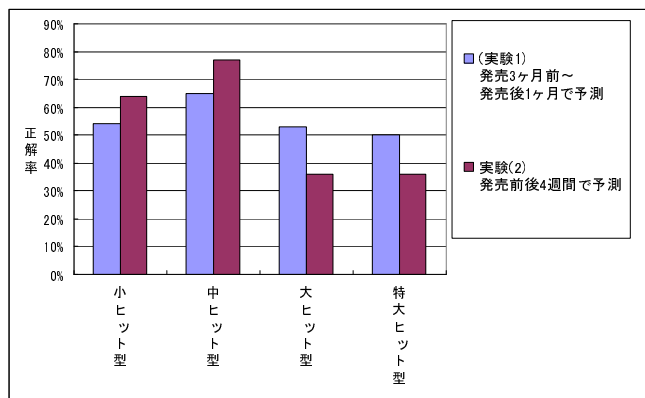


図5 実験1,2の予測結果の正解率とその比較

4.4 実験2に対する考察

4.4.1 正解率に対する考察

実験1と実験2の結果を比較すると、実験3の小ヒット型の正解率は+10%、中ヒット型の正解率は+12%となった。これは、4.2.2で挙げた考察のとおり、言及記事数の推移の変化があまりみられない期間のデータを予測から除外したことが正解率を上げる原因になったのだと考えられる。

しかし、大ヒット型の正解率は-17%、特大ヒット型の正解率は-14%であり、大きく減少した（正解率減少に対する考察は5章で行う）。

4.4.2 言及記事数が参考にならない商品に対する考察

本手法の課題として、予測において発売日前後の言及記事数の推移のみの判断では、言及記事数が参考にならないと思われる商品が多く見られたことが挙げられる。その例として「漢検」を挙げる。「漢検」を検索語として商品の発売前後3ヶ月の言及記事数の推移を取得しグラフ化したものを図6に示す。

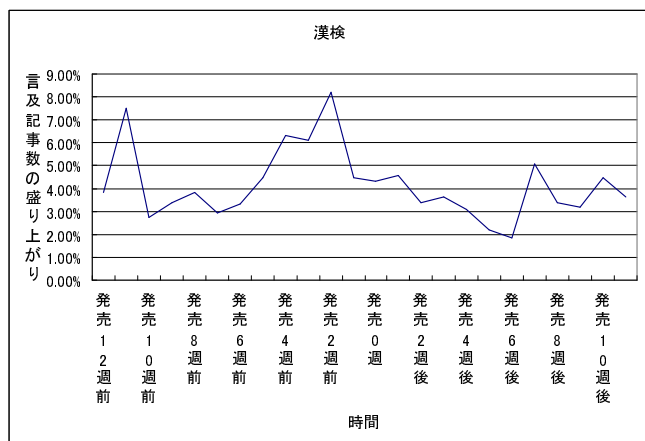


図6 「漢検」の言及記事数の推移

「漢検」という単語は一般的に広く使われるため、常に話題として盛り上がっている。商品の発売前後の盛り上がりと通常時の盛り上がりで明確な違いがでないため、こういった単語が商品名として用いられると、予測がしづらくなることがわかる。このような商品は予測対象から除外できれば、正解率を向上させることができると考えられる。

そこで、言及記事数の推移に対する標本分散を用い、常に話題として盛り上がっている商品を除外する処理を提案する。

4.4.3 常に話題として盛り上がっている商品名を除外する処理

一般的に広く使われる単語は、図6のように不規則かつ頻繁に言及記事数が盛り上がる。このような言及記事数の推移がみられた場合、その標本分散は大きい値をとる。この性質を利用し、常に話題として盛り上がっている商品名を除外する。

まず全商品に対し、商品の発売日の4週間前から発売日の3週間後までの言及記事数の推移の標本分散と、発売日の12週間前から発売日の5週間前までの言及記事数の推移の標本分散を算出する。そして、前者と後者の標本分散の値を比較し、後者のほうが大きい商品を除外する。

実験2にこの手法を適用し、実験3を行う。

4.5 実験3(常に盛り上がっている商品の除外)

実験2で得た結果と考察から、さらに精度を上げるため実験方法を変更し検証を行った。今回の実験では、常に話題として盛り上がっている商品をあらかじめ除外しておき、2008年4月から2009年3月までに週間売上TOP30に入った商品をもとに、2009年3月から12月までに週間売上TOP30に入った商品の発売日の4週間前から発売日の3週間後までのデータで売上推移パターンを予測する。

4.5.1 実験に用いたデータ

商品ジャンル・過去にヒットした商品とその売上・売上予測する商品は実験2と同じものを用いた。そしてそれらのすべての商品に対して4.4.3で提案した手法を適用し、常に話題として盛り上がっている商品152件を除外した。

4.5.2 実験結果

それぞれの商品に対し売上推移パターンの予測を行い、その正解率とその比較のグラフを図7に示す。

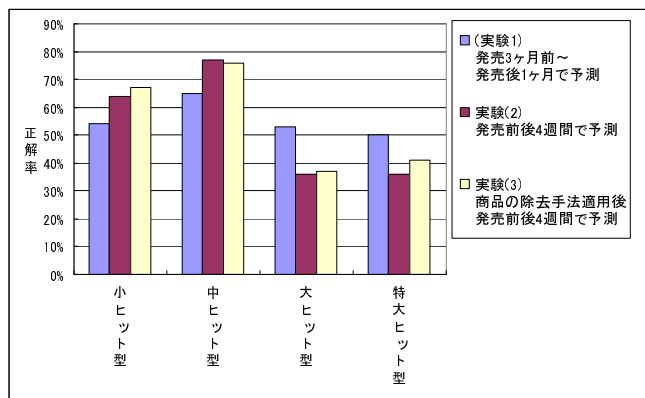


図 7 実験 1,2,3 で予測した結果の正解率とその比較

4.6 実験 3 に対する考察

実験 2 と実験 3 の結果を比較すると、実験 4 の小ヒット型の正解率は+3%、中ヒット型の正解率は-1%、大ヒット型の正解率は+1%、特大ヒット型の正解率は+5%となり、全体の正解率は微量ながら向上する結果となった。このことから、常に話題として盛り上がっている商品名を除外する手法はある程度効果があったといえる。

5. 考察

5.1 考察

本研究により、商品の発売日の 4 週間後の時点において、その後の売上の推移の伸びがある程度予測できることが判明した。

しかし、図 7 からわかるように、大ヒット型と特大ヒット型の正解率は低い結果となった。ここで、正解率が大きく減少した原因を調べるため、誤予測した結果の内訳を表 2 に示した。

表 2 発売前後 4 週間の言及記事数の推移で予測した結果の内訳

	予測結果			
	小ヒット型	中ヒット型	大ヒット型	特大ヒット型
大ヒット型の商品群	11%	3%	36%	50%
特大ヒット型の商品群	18%	14%	32%	36%

表 2 より、大ヒット型は特大ヒットと誤予測し、逆に特大ヒット型は大ヒット型と誤予測しているものが多いことがわかった。このことから、大ヒット型以上を予測するには高い正解率が示せるが、大ヒット型と特大ヒット型それぞれの的確に予測することは難しいといえる。そのため、本手法をビジネスシーンで活

用するためには、本手法が得意とする予測結果を活用することが望ましい。

そこで、得意とする予測結果を示すため、実験 3 の方法を用いて、特大ヒット型と分類されていた商品を大ヒット型の商品としてみなし、実験を行った。具体的には、商品の発売日の 4 週間後の時点で、発売日から 4 週以上の間ランクインし続けるかどうかを予測した。その結果を表 3 に示す。

表 3 言及記事数が参考にならない商品名を除去後、予測した結果

	正解率	
	ランクインする	ランクインしない
小ヒット型	8%	92%
中ヒット型	10%	90%
大ヒット型 または 特大ヒット型	87%	13%

表 3 の結果から、商品の発売日の 4 週間後の時点で、商品の売れ行きが失速するかどうかを判断する場合において高い確率で予測を的中させることができるといえる。

5.2 課題

本手法の課題として、予測において発売日前後の言及記事数の推移のみの判断では、言及記事数が参考にならないと思われる商品が多く見られたことが挙げられる。その例として、「クレヨンしんちゃん」を挙げる。「クレヨンしんちゃん」を検索語として商品の発売前後 3 ヶ月の言及記事数の推移を取得しグラフ化したものを図 8 に示す。

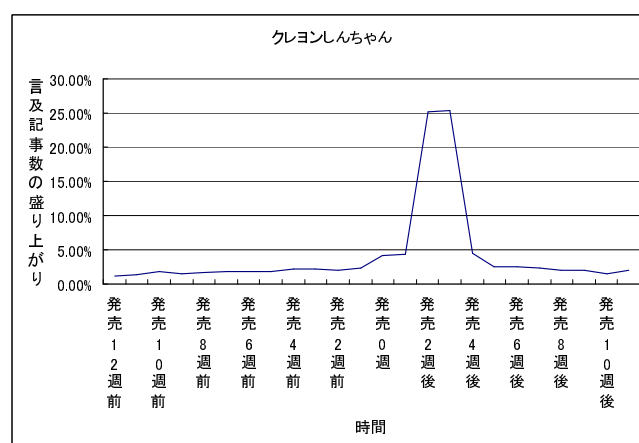


図 8 「クレヨンしんちゃん」の言及記事数の推移

「クレヨンしんちゃん」は、ゲーム・文庫本だけに限らず、アニメ・映画などさまざまなメディアで商品が販売されている。図 8 では、商品の発売週に少し盛

り上がったが、その二週間後にはその数倍の盛り上がりがあった。ここで商品の発売から二週間後に何があったのか調査をしてみると、過去に放映されていた映画がテレビ放送されたことが原因だとわかった。そのため、言及記事数が盛り上がった原因が何であったのか判断できず、こういった単語が商品名として用いられると、予測がしづらくなることがわかる。

このことから、予測したい商品と関係のない言及記事数の除去が課題として挙げられる。

6. まとめ

本論文では、ブログ上の商品に対する言及記事数を用いて売上の推移パターンを予測する手法について述べ、本手法を用いて行った実験の結果について述べた。

その手法とは、まず予測したい商品ジャンルで過去にヒットした商品名の一覧を用意し、それぞれの発売日前後の言及記事数の推移データを検索エンジンから取得する。そして、その言及記事数の推移をいくつかの売上推移パターン別に分類する。そして、その分類されたパターンそれぞれに対し、予測したい商品の言及記事数の推移データと相関分析を行う。最後に、その相関分析の結果から、もっとも相関が強かったものの売上推移パターンを予測結果として出力する。

実験の結果により、特定の条件に当てはまる商品に対し、発売日の4週間後の時点で、商品の売れ行きが続くかどうかを判断する場合において高い確率で予測を的中させることができた。予測対象となる商品の条件は、以下の3つである。

1. ブログ上で頻繁に話題として取り上げられる商品分野であること。
2. 売上の高い商品であること。
3. 商品名が別の意味でも使われるなど、商品への関心の度合いと言及記事数が直接連動しないもの。

しかし、まだ予測の正解率を向上させるために工夫すべき点がいくつかあった。今後、正解率の向上のため課題として挙げた、商品と関係のない言及記事数をどう排除していくかについて解決策を見つけ出し、実験により正解率の上昇が見込めれば理想的である。

謝 辞

本研究の一部は文部科学省私立大学戦略的研究基盤形成支援事業「芸術・文化分野の資料デジタル化と活用を軸とした研究資源共有化研究」の支援を受けています。

文 献

- [1] 奥村 学. “blog マイニング——インターネット上のトレンド, 意見分析を目指して——”, 人工知能誌, vol.21, no.4, pp.424-429, 2006.
- [2] 水口 弘紀, 土田 正明, 久寿居 大. “Weblogを対象にしたリアルタイム評判情報分析システム eHyouban”, 第19回データ工学ワークショップ (DEWS2008) 論文集, 12-27, 2008.
- [3] 奥村学. “ブログマイニング技術の最新動向”. 電子情報通信学会誌, Vol. 91, No. 12, pp. 1054-1059, 2008.
- [4] 川添 恭平, 木村 義紀. “消費者ニーズを発見・獲得するブログ解析技術の研究開発”, INTEC TECHNICAL JOURNAL, 第8号, pp. 36-41, 2008.
- [5] D. Gruhl, R. Guha, R. Kumar, J. Novak. And A. Tomkins, “The predictive power of online chatter” Proc. Of the 11th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, pp.78-87, 2005.
- [6] 福原知宏, 村山敏泰, 中川裕志, 西田豊明. “Weblog から社会の関心を探る” 人工知能学会全国大会論文集 (CD-ROM), 20 (3D2-1), 2006.
- [7] Yahoo! ブログ検索, <http://blog-search.yahoo.co.jp/>
- [8] ORICON STYLE, <http://www.oricon.co.jp/>