

2017(平成 29)年 2 月 25 日

西南学院大学 大学院

経営学研究科長 伊藤 龍峰 様

審 査 委 員

主査 吉武春光



副査 史 一華



副査 藤岡 豊



博士論文審査報告書

氏 名： 経営学研究科経営学専攻 15DB001 野間 利博

論文題目： 「テキストマイニングによるブランド評価の研究」

1. 本論文の目的

本論文は、電子化された新聞記事の新たな可能性を見出したいという野間氏の強い意志の下、日本経済新聞の 2003 年から 2012 年までの 10 年間という多量の記事を計算機処理することにより、新聞記事の新たな活用法を探ったものである。

2. 本論文の概要

文字列になっている記事を計算機処理するアプローチは、一般にテキストマイニングと呼ばれている手法を用いる。マイニング(Mining)とは、本来は、岩石の中から貴金属を発掘するという由来を持つ用語で、データの中から何か新しい特徴を見出す処理のことである。特に、テキストマイニング(Text Mining)は文字化されているデータの中から特徴を見出すことであり、現在のマイニングでは、最も一般的な処理である。テキストマイニングを行うためには、まずは情報を文字化することから始める必要がある。しかし、新聞の場合、編集作業の電子化が 1990 年代に始まっており、2000 年代に入ると電子化された新聞記事そのものが手に入るようになったので、容易にテキストマイニングを行える状況にある。テキストマイニングの手法は、文を語に分割して出現頻度を数えることが前提である。その出現頻度を統計学的に処理して、その文の特徴を表す語を見出すのである。しかしながら、2010 年代に入ると、処理対象とするテキストデータの量が増え、統計処理ソフトで処理できる量を超えてしまったために、処理対象をいかに絞り込んで統計処理ソフトにかけるのかという技に結果が左右されるという状況になっている。

2012 年頃から、別のマイニング手法が現れている。それは、処理対象を絞り込むのではなく、処理対象全体を語のベクトルに変換し、そのベクトル間の関係を見出そうというものである。

この手法は従来からあるニューラルネットワークを発展させたディープラーニングと呼ばれる手法で、その代表ソフトが 2013 年に発表された word2vec である。word2vec は、自然科学分野で注目を浴びている手法であるが、未だ、社会科学の分野での利用は数少ない。

本論文は、日本経済新聞の 2003 年から 2012 年までの 10 年間という多量の記事に対して、まず、処理対象を絞り込みながら統計処理ソフトを用いて分析を行ったが、結果が絞り込みに左右されることと、代表的な特徴しか見い出せないことを明らかにしている。次の段階として word2vec を使って 1 年分の記事を処理単位にして 10 年間の移り変わりに対して、特徴を見い出そうとした。その結果、企業のブランド戦略が見えるという結果を得た。そこで、その結果の確認のために、アンケート調査を行い、その結果と照らしあわせて、word2vec を使った結果が、ほぼアンケート結果と一致していることを見い出している。

3. 本論文の構成と内容

本論文は、2 部構成となっている。

第 1 部は、4 つの章で構成されている。第 1 章は、本論文の目的を述べた上で、処理対象として新聞記事を選んだ理由を述べている。第 2 章では、テキストマイニングと新聞記事分析の先行研究を多数、レビューしている。テキストマイニングの研究では、使用してある手法は様々であるが、基本概念は、語の分解、形態素解析、構文解析という手順をたどり、得られた語の出現頻度や構文情報などを統計学的に処理して特徴を見い出している。新聞記事分析の研究では、新聞が統一された観点の元で編集されているという特徴を活かしており、手軽に入手できるという利便性からも行われている研究が多い。第 3 章では、「新聞記事には新しい製品開発のヒントが存在するのか」という仮説を立てて、10 年分の新聞記事を解析し、語の出現頻度を用いたヒット製品の特徴を発見することを試みている。試みの手順としては、まず、日本語形態素解析プログラム MeCab を用いて、日本経済新聞の 2003 年から 2012 年までの 10 年間の記事の本文（各年約 18 万記事）を形態素解析し、品詞が付与された語を得ている。得られた語数は各年当たり約 5 千万語となった。解析単位を年とした場合、その約 5 千万語の中から特徴を見つける必要があるため、テキストマイニングに使えるような色々なソフトを試行している。それらは TTM というテキストマイニングソフト、日本語ワードネット、統計解析ソフト SPSS を用いた因子分析、統計解析ソフト SPSS を用いたクラスター分析、そして、ディープラーニングを用いるソフト word2vec である。word2vec は前述の約 5 千万語を一度に処理出来るのであるが、それ以外のソフトでは処理できる量を超えているために、ビッグデータ処理用のソフトウェア群 Hadoop 中の高水準言語 Apache Pig を用いて語の整理を行った後で解析処理を行っている。TTM は、語の係り受け関係を考慮した分析を行うことも出来るが、語の統計処理での解釈が難しくなるために、語の出現頻度のみの解析を行っている。この試行の後に、次の第 4 章において、日経 MJ(流通新聞)が毎年発表している「ヒット商品番付」の製品を、日本経済新聞の記事からテキストマイニング出来るかどうか挑戦している。その分析手順は、年ごとに Apache Pig を用いて、記事から製品名を含んだ複数文を抽出し、次に予めリストアップしておいたその製品の特徴を表す語（共起語幹

と称する)も含んだ複数文のみに絞り込んでいる。そうして得られた複数文に対して TTM を用いて分析を行い、記事の特徴を示す語(理由語幹と称する)を見い出そうとした。分析対象年は、「ヒット商品番付」が発表された年を含む前後3年間としている。こうして得られた理由語幹が妥当であるかどうかを検証した。以上の分析を「iPod」について行ったが、製品の特徴は検出できたものの、その製品のヒットした理由は見い出せなかった。ここで問題になったのは、絞り込みにより記事数が激減したことと、記事を絞り込むのに使った語の妥当性であった。

第1部の結果を受けて、次に第2部は、記事の絞り込みを行わずに分析処理を行える word2vec を用いて新聞記事分析の可能性を探っている。従来の統計手法が語の出現頻度を分析しているのに対して、word2vec は語と語の出現の関係を分析しベクトル化している。word2vec は positive word として指定した語のベクトルに近いベクトルを持つ語を、類似の程度を示すコサイン類似度を示しながら出力するようになっている。その際、negative word を指定しておくことで、その negative word のベクトルに近い語は除かれるようになっている。word2vec のベクトル化の仕組みは「同じ文脈で利用される単語は、同じ意味を持つ」という仮説に基づいているために、word2vec の結果として得られた類義語は新聞の記事において同じような文脈で用いられている表現の特徴の類似性を表していると言える。第2部は4つの章で構成されており、まず、第5章では、アップルの iPod とソニーの WALKMAN について word2vec を用いて試行を行い、特徴となる語が出現していることを確認している。それにより、企業のイメージ(コーポレート・ブランド)と製品のイメージ(プロダクト・ブランド)の関係を見い出すことができるのではないのか、という仮説をたてた。その際、word2vec の結果には企業名や製品名などが混在するので、考察の結果、企業名、製品名、製品カテゴリーという3つの観点から検討を加えるのが良いと判断された。次に、第6章では、その仮説を検証するために、2003年から2012年までの各年について、アップルの iPod とソニーの WALKMAN の分析を行っている。その結果、アップルの iPod に関しては、「アップル」という語で得られた word2vec の類似度の上位に「iPod」が出現することから、「アップル」と「iPod」は関連性が強い、ことを見い出している。一方、ソニーの WALKMAN については、「WALKMAN」という語で得られた word2vec の類似度に、「ソニー」よりも上位に「アップル」が得られたことから、「ソニー」と「WALKMAN」の関連性は弱い、ことを見い出している。次に、引き続いて、2003年から2012年までの各年について、国産自動車メーカー4社(トヨタ、日産、ホンダ、マツダ)について、企業名で得られた word2vec の類似度を分析している。トヨタとマツダの製品に関する新聞記事には、それぞれその新聞記事表現にトヨタあるいはマツダという企業に共通の新聞記事表現の特徴が表れていることが判った。日産については日産という企業に共通の新聞記事表現の特徴がはっきりしない。ホンダについては、ホンダという企業に共通の新聞表現の特徴がシビックや FIT というホンダの製品には現れているが、それ以上に Vitz、カローラ、Prius といったトヨタの製品にも現れていることから、ホンダという企業に共通の新聞表現の特徴はトヨタという企業に共通の新聞表現の特徴に似ていると言える。さて、自動車メーカーのコーポレート・ブランドとプロダクト・ブランドの関連性については、マツダが最も関連性が強いことが判った。トヨタもコーポレート・ブランドとプロ

ダクト・ブランドが一体化しているが、ハイブリッド車のプリウスや高級車のレクサスの評価が影響しているようである。更には、国産ビールメーカー4社（アサヒ、キリン、サントリー、サッポロ）について、企業名で得られた word2vec の類似度を分析している。その結果は、サッポロ以外の3社は企業名と製品名の関連性が強いと判った。最後に、日本の大手化粧品メーカー3社（資生堂、花王、コーセー）について、企業名で得られた word2vec の類似度を分析している。その結果は、花王は企業名と製品名の関連性が最も強く、次に、コーセーが関連性が強かった。しかし、最も大手である資生堂は、関連性が弱いと判った。

次に第6章で得られたブランド評価に関する結果が実際の各社のブランドの実態と合致しているかどうかを検証するために、広告会社を対象に「企業のブランドイメージについての調査」を実施している。168カ所を対象にアンケートを実施して49件の回答を得ている。第7章では、このアンケート調査の詳細と、回答内容に対して行った SPSS を用いた因子分析の結果を詳述している。そして第8章では、第6章で得られたブランド評価に関する結果が実際の各社のブランドの実態と合致しているかどうかについて、第7章で述べたアンケート調査結果と、製品に関連する新聞以外の文献、新聞記事、各社の決算資料などを用いて検証している。家電・情報機器の場合、自動車メーカーの場合、ビールメーカーの場合、化粧品メーカーの場合の各々についてコーポレート・ブランドとプロダクト・ブランドの関連性を考察した結果、word2vec の結果とアンケート調査の結果が大筋で一致することを見出した。

最後に、終章では、第1部と第2部を通して行った、新聞記事活用の新たな可能性を見出す研究のまとめを行っている。第1部では、大規模テキストに対する従来のテキストマイニング手法は一般的な製品の特性を見出すことが出来ることを確認したが、新製品開発のヒントになるようなものを見出すことは出来なかった。また、問題点としては統計処理ソフトにおける変数の絞り込みが結果を大きく左右することを確認した。第2部では、新しいテキストマイニングのツール word2vec を用いて大規模テキストを処理することで、企業のコーポレート・ブランドとプロダクト・ブランドの関係性を導き出すことが出来た。

4. 本論文の評価

近年、企業は自社のブランド戦略として、①コーポレート・ブランド戦略、②メガ・ブランド戦略、③プロダクト・ブランド戦略を推進している。コーポレート・ブランド戦略とは、企業が個々の製品に重きを置くのではなく、個々の製品を超えた企業ブランドとしての一貫性を追求する戦略である。メガ・ブランド戦略とは、企業の複数の中核的製品に重きを置いて、当該ブランドを中心にして企業としてのブランドを形成する戦略である。プロダクト・ブランド戦略とは、企業ブランドとしての一貫性を追求するよりもむしろ個々の製品の存在感を最大限に追求するブランド戦略である。三者の関係はコーポレート・ブランド戦略とプロダクト・ブランド戦略を両極にして、メガ・ブランド戦略がその中間に位置するものと考えても差し支えない。従来、企業は自社のブランド戦略がどの程度社会に浸透しているかを確認するために、消費者に対するアンケート調査に依拠しながら、データを収集するために多大な苦労を重ねてきた。しかし、もし新

聞記事を活用してブランド戦略の浸透度を確認する方法が存在するのであれば、企業にとって当該データの収集負担は画期的に低減し、企業のブランド戦略の実行速度も大きく変化する可能性が出てくる。

本論文は新聞記事の新たな活用法を探るという問題意識の下、従来、自然科学分野で活用されてきた word2vec というテキストマイニング・ツールを使って、まずはブランドの両極にあるコーポレート・ブランドとプロダクト・ブランドに絞って両ブランドの関係性を業界や企業ごとに評価し、各企業におけるブランド戦略を把握しようとした。本論文は新聞記事を消費者の声に見立てて、新聞記事のデータから消費者の意向を確認するという手法を採用している。しかし、そのようなことが本当に可能なのか？本論文はこのような疑問に答えるべく広告代理店へのアンケート調査も行って、新聞記事の分析結果とアンケート調査結果の分析結果を照合し、両者に大きな齟齬がないことを確認している。この発見事実は消費者の意向把握を主にアンケート調査に依拠してきた従来のマーケティング研究にとっては驚きであろう。マーケティング研究の分野にとっても今後、本論文を発端として同様の研究が積み重ねられていく可能性があるのではないか？本論文が分析対象とした新聞記事データが単一新聞の 10 年分であったという研究上の限界があるとはいえ、本論文が企業のブランド評価の新たな方法を提示し、社会科学分野に新たな知見をもたらしたことは学術的に高く評価される。新聞記事データの種類や量を拡大しても果たして word2vec というテキストマイニング手法が効果をもつのか研究の興味は尽きることなく、野間氏の今後の一層の研究も期待される。

審査員一同は、以上の厳正かつ慎重な審査及び野間氏の公聴会での質疑応答の結果に基づいて、本論文が博士（経営学）の学位にふさわしい十分な内容と水準を備えていると判断する。

以上