

氏名	大森 寛文		
学位の種類	博士（経済学）		
学位記番号	博経済甲第94号		
学位授与年月日	平成26年3月24日		
学位授与の要件	学位規則第3条第3項該当		
学位論文題目	企業・組織における知識発見の実践手法に関する研究 ーテキストマイニングと知識の構造化論の融合による知識の発見ー		
論文審査委員	委員長	教授	菰田 文男
	委員	教授	薄井 和夫
	委員	准教授	丸茂 幸平
	委員	教授	李 潔

論文の内容の要旨

戦後の日本企業は乏しい資金を効率的に利用し、変動相場制度移行以後の円高や石油価格の高騰も産業ロボットや省エネ技術開発により乗り切り、貿易摩擦問題も産業構造の高度化によって克服するなど、直面する困難をその都度企業努力を積み重ねることにより、国際競争力を高めてきた。しかし1991年にバブルが弾けて以後、アメリカ企業やアジア新興国企業が新しいビジネスモデルを構築することによって新たな経済パラダイムに適応努力を続けているのに対し、日本企業はその努力を怠り、次第に国際競争力を失ってきた。したがって、バブル経済の崩壊から今日までを「失われた20年」といわれるのであるが、このままでは「失われた30年」を迎えることになりかねないという状況にある。

「失われた30年」としないためには、マクロ経済政策レベルでもミクロの企業努力のレベルでもさまざまな努力が必要であるが、その一つが企業戦略の立案にとって真に必要な知識を獲得することである。そして、この知識の発見に貢献できると期待されているのがテキストマイニングである。各種のデータの中でも、深く豊かな意味が切り捨てられることなく残されているテキストデータは、企業が事業企画や戦略立案に必要な知識を得る上で最も重要な源泉である。しかし、数値データのように厳密なデータとは異なり、自然言語によって書かれ構造化されていないテキストデータは、曖昧さや多義性を多く含んでいるので、その中に含まれている知識や意味は人間が読むことによって、それを発見する以外にないのが現状であった。

このような現状は過去10年間のテキストマイニングツールの開発と進化により少しずつ克服されつつあり、人間の代わりにコンピュータが読み、知識を抽出することを可能にしつつある。にもかかわらず、コンピュータの情報処理能力の現状は、曖昧さ、冗長さ、多義性を多く含んでいるテキストデータの正確な解析と知識抽出を可能にするにはほど遠い。

日々厳しい市場競争にさらされている企業にとって、ライバル他社に先んじて戦略的な意義のある知識を獲得することが不可欠であるが、テキストマイニングにこのような知識の獲得を期待することは難しいというのが一般的な認識である。したがって、テキストマイニングを有効にするための手法を工夫することが必要である。その一つが「知識の構造化」である。企業の企画部門、研究部門などの各部門には、テキストデータの解析とは無関係に貴重な知識が膨大にストックされている。このような知識を分析の目的に応じて体系的に整理し、テキストマイニングに生かし、さらにテキストマイニングの成果をこの知識ストックに追加してより精緻に構造化するという作業の繰り返し、テキストマイニングを有効にすると考えられる。本稿の目的は、知識の構造化のための作業とテキストマイニングとを組み合わせることによって、テキストデータからの知識抽出するための手法を提案することにある。

この目的を達成するために以下の構成で論じる

第1章 序論

第2章 先行研究レビュー

第3章 研究の方法

第4章 ネットバンキングの利用促進に向けた推進要因と阻害要因に関する知識の発見

第5章 電機業界における経営課題の認識構造と実行動に関する知識の発見

第6章 国会の展開過程にみる議論構造とその変化動向に関する知識の探索

第7章 結論

まず序章で、企業や組織の知識の発見においてテキストデータの分析が有効であるが、しかし失敗が許されない企業の実践の場において真に役立つ知識を得るためにはテキストデータの解析だけでなく、その以外の場面で得られた知識も含めて整理し体系化しておくことが必要であり、知識の構造化とテキストデータの解析とは車の両輪であることを述べる。

第2章では、まず多くのテキストマイニング研究の先行研究のうち、経済学・経営学分野の研究を中心として取り上げ、整理し、その意義と限界を明確にする。すなわち、電子メールログを用いた企業内コミュニケーション分析、有価証券報告書による企業評価分析、会議録の分析などであるが、この限界はマイニングするために単語に切り分けることによって、単語が文全体の中に位置づけられていた際に有していた意味を失ってしまうという限界があると述べる。また膨大な単語群の中から、解析に必要な重要単語を抽出することは容易ではないという限界もある。これらの限界を克服するためには、単語あるいは概念を予め抽出し、その関係を階層化したりネットワーク化しておくことによって関連づけておくことが必要であると述べる。そしてこの構造化のための手法として、小宮山宏、斉藤雄志、畑中洋太郎などの研究成果を摂取する。すなわち、これらの研究が階層関係やネットワーク関係を重視し、ノードとリンクによって知識を体系化して表現する手法の意義を認めつつ、その限界についても論じる。

第3章では筆者の構造化の手法を提起する。それはノードとリンクにより知識を構造化する手法に依拠したうえで、知識発見を静的ではなくダイナミックで継続的な知識創出の仕組みの中に位置づけるという手法である。すなわち、知識の創出とその精緻化は(1)解析フレームの設定、(2)要素の抽出、(3)要素の構造化、(4)意味解釈という4つのステージの円環的で継続的なプロセスによって可能となるという考えにたち、このプロセスの中にノードとリンクで表現される知識体系を作り上げるとともに、それに基づいてテキストデータを解析し、それをさらに知識の構造化の精緻化につなげるということである。このようにすることによって、単に無目的に構造化するのではなく、テキストマイニングに適した構造化を目指すことができると論じる。

以上の分析に基いて、次の3つの章では3つの事例をあげてテキストデータから知識を発見する手法を提起する。

まず第4章では筆者が実施したオンラインバンキングに対するユーザーの意識を知るためのアンケート調査を対象とする。解析のために、オンラインバンクの促進要因と阻害要因を解析のフレームとして設定し、「手数料」という経済的要因が促進要因であり、「通帳」が無いことが阻害要因である、等の意味を発見している。

第5章では、金融庁が提供している有価証券報告書データベースであるEDINETから日本の電機メーカー270社の2007～2011年の有価証券報告書をダウンロードし、企業の業績と事業戦略との間にどのような関係があるかを発見する。そのために、270社を「成長軌道型」「黒字回復型」「赤字突入型」「赤字継続型」の4類型に分け、それぞれがどのような事業戦略を展開しているかを解析する。その結果、「成長軌道型」「赤字突入型」「赤字継続型」がコスト構造の改革を重視していること、「黒字回復型」が収益改善を重視していることなどの知見を得ている。とりわけ興味深い発見は「赤字突入型」「赤字継続型」という業績が低迷ないし悪化している企業が技術力強化を重視している点であった。これは一見すると矛盾しているかに見えるが、有価証券報告書では技術力重視を唄っているものの、実際には実行が伴っていないという貴重な知見を得た。

第6章では、国会会議録検索システムから「中小企業技術革新支援制度(通称、日本版SDIR)」の議事録を1998年～2012年の15年間についてダウンロードし、解析した。解析の目的は日本版SDIRが米国版SDIRを真似て導入されながら、十分な実効が得られていない理由を発見することにある。その結果、同制度が導入され運用されるにつれて、国会での議論も、当初の「基本コンセプト領域」「コアコンセプト領域」から、次第に「運用状況と工夫事項領域」へと次第に具体化されるべきであるにもかかわらず、実際には具体化されず15年間の半ばから後半にかけて「コアコンセプト領域」から「基本コンセプト領域」へと逆行している、等の貴重な知見を得ている。

最後に第7章では、本論文全体のまとめと今後の課題について述べる。

論文審査の結果の要旨

コンピュータの演算処理能力や記憶容量の飛躍的向上は、従来は不可能であった膨大なテキストデータの解析を可能にしつつある。その有効性は次第に知られるようになっており、コールセンターのログから顧客の不満の発見など、少しずつ企業の知識創出において利用されるようになってきている。しかし、未だその利用用途も効果も限定的なものにとどまっているというのが現状である。テキストマイニングは未だ研究段階にとどまっているといわざるをえない。したがって、この限界を克服し、効果が広く認められ、多くの企業によって日常的に利用されるようになるために、実践性あるいは実用性を重視した研究は焦点の課題である。その意味で本研究の意義は大きい。

本論文の独創性あるいは意義は以下にある。

第一に、テキストデータの解析という大きな困難に直面している研究課題に対して、一つの前進の方向性を提示できた。周知のように、自然言語で書かれたテキストは解析に大きな困難がある。最大の困難は、解析ツールの進歩にもかかわらず、抽出された膨大な単語群の中から分析目的に応じて適切な単語を抽出することは困難であるという点にある。筆者はこの困難を克服するために、分析者が過去の経験から有している知識のストックを体系的に整理し構造化しておき、膨大な単語群の中から分析目的にとって重要性の大きい単語を発見しやすくするという手法を提案する。この構造化された知識ストックは、マイニングの結果から学習して不断に豊かになり精緻化される。知識の構造化のための作業とテキストマイニングとは知識創出の両輪として有機的に関連づけられ、不断に進化してゆくものとされる。

第二に、テキストデータの解析を、単に学術研究目的のためにおこなうのではなく、企業の日々の戦略立案や意思決定に実際に利用することを前提にした手法を提起している点にある。例えば上述の学習とはコンピュータアルゴリズムによる自動学習という意味ではなく、分析者自身の判断・思考に媒介されるものである。なぜなら、筆者の想定しているテキストマイニングの利用用途である企業の意思決定の場に利用できるだけの信頼性の高い知識を、アルゴリズムだけに依存して得ることはとうてい不可能だからである。この限界を克服するためには、分析者自身が頭の中で知識を組み替え、進化させ、体系化することが不可欠なのである。本稿がそのための手法を提起し、テキストマイニングと結びつけることに成功している点は高く評価できる。

第三に、テキストデータの解析において、その目的に応じて的確にクラスター分析やネットワーク分析を適用するなど、統計解析の手法にも習熟している。しかも、テキストマイニングから知識を獲得するためには、解析の結果をそのまま信用するのではなく、解析結果の裏に潜んでいる真実を読み解くだけの洞察力を有することで必要であるが、本稿ではこの洞察力が発揮されていることは、有価証券報告書に叙述されている言葉とそれとは異なる企業の実態との間の齟齬の発見の事例からも理解される。

もちろん、不十分な点が無いわけではない。

とくに、知識の構造化についての筆者のアイデアは価値があることは間違いないし、そ

の手法も提示されてはいるのであるが、しかし本稿ではその叙述が抽象的なレベルにとどまっているのではないか、もう少し具体的な事例をあげて論じることによって説得力を高めることが必要ではなかったかとも思われる。本稿のコアになる部分であるだけに惜しまれる。

また、本稿が学術的な研究というよりも、企業の実践の場で利用されることを目指した研究を指向するのであれば、解析対象となったテキストデータ量を大幅に増やし、大量データの中から知識を発見するという試みをおこなうべきであったであろう。

以上のような問題があり改善の余地があることは間違いないが、本稿は企業の事業立案と意志決定という実践の場での利用を指向して、テキストマイニングの新たな手法の確立に向けて大きな一歩を踏み出すことに成功しており、今後のいっそうの研究の進歩を期待することができる。したがって、本論文は、博士論文としての基準である「当該テーマについての研究業績を踏まえ、学術的な独創性」を有するという水準に達していると評価できる。以上により、審査委員会は、本論文が、博士（経済学）の学位を授与するにふさわしいと判定した。