

## ニュースが株式市場に与える影響の測定方法 A Method of Measurement of The Impact of Japanese News on Stock Market

片山 大輔<sup>†‡</sup> 津田 和彦<sup>‡</sup>  
Daisuke Katayama Kazuhiko Tsuda

### 1. はじめに

近年、ビックデータという言葉が注目を浴びている。ビックデータへの関心の高まりは、資産運用業界においても例外ではなく、今日では様々な取り組みが行われている(諏訪部 [2015])。実際に、ファンドマネジャーの元には、投資の意思を決定するための大量の情報が届けられる。財務情報やマーケットの動き、マクロ指標などを示す数値情報と共に、ニュースやレポートといったテキスト情報も届けられる。これら全ての情報に目を通し、投資行動に移すことは、人間の処理能力では不可能に近い。その中でも大量のテキスト情報は数値情報のように定量的な判断が難しい。特にニュースなどは、株式市場に影響を与える情報に曖昧性や多義性があり判断に時間を要する。そのため、これらのニュースを漏れなく効率的に処理することが求められている。

そこで本研究では、日本の上場企業に関するニュースが株式市場にどのように影響を与えるのかを検証する。具体的には大量のニュース情報を利用して、どのような属性をもつニュースがどのようにその企業の出来高や株価に影響を与えるのかを明らかにする。

### 2. 資産運用分野における自然言語処理

資産運用分野(ファイナンス分野)におけるテキスト情報分析は内外で活発である。例えば Tetlock [2007] では、Wall street Journal column から、記事中に含まれている単語を用いて記事の悲観度を抽出し、その増減が将来の株価の予測可能性を持つと指摘している。続く Tetlock [2008] においては、個別企業に着目し、ニュース記事の情報が将来の企業の利益およびリターンを予測できるのかを検証し、ニュースにネガティブな単語が多いほど、将来の当該企業の利益およびリターンが低くなるとしている。Tetlock のこれらの分析は、ネガティブなニュースが株価に負のインパクトを与えるというものであるが、ポジティブなニュースについての検証は行われていない

日本においても、既に 2003 年に高橋・津田 [2003] はアナリストレポートに着目し、企業の株価を予測するモデルを提案している。近年テキスト情報を利用した分析がさらに増えてきており、沖本・平澤 [2014] は Tetlock [2007] を参考に、QUICK 端末で配信される日経ニュースを定量化したニュース指標を用いて、株式市場の予測可能性について検証している。この分析では、個別銘柄ベースでどのような影響があるかは検証されていない。五島・高橋 [2016] は、ディープラーニング(深層学習)の手法を用いて、英文

の金融ニュースのセンチメントを測定し、ニュースが日本株式市場にどのような影響を与えるかを調査している。

このように、ニュースを利用した分析は比較的多く存在するものの、短期的な市場全体への影響を調査したものが多く。しかしながら、資産運用においては、市場全体の動きの予測も重要であるものの、基本的にはポートフォリオを構築する、個別銘柄毎の予測が最も重要である。けれども、そのような研究は稀であるのが現状である。そこで本研究では、個別銘柄に対応したニュースが個別銘柄の株価の動きにどのように影響を与えるのかというところに焦点を当てる。

### 3. 株価に対するニュースの影響

ここでは、個別銘柄において、当該銘柄のニュースが株価に与える影響を調査するため、以下に示す仮説を構築し、検証する。

- i. ポジティブ(ネガティブ)なニュースが出たら、その後の当該企業の株価は増加する。多くの投資家は、ニュースを見てその内容を判断し投資行動を決定する
- ii. i の影響は 1 面の記事ほどその影響が大きくなると考えられる。なぜなら、1 面の記事ほど、多くの投資家の目に止まりやすいからである
- iii. i は時価総額が小さい企業ほどその影響が大きくなる。時価総額の大きな企業は、多くの投資家やアナリストが注目しているため、時価総額の小さな企業と比較して、そのニュースを事前に察知している可能性があるため、インパクトが落ちる可能性がある

### 4. ニュースの極性

#### 4.1 分析対象データ

分析対象は、2017 年 1 月末時点の TOPIX500 の構成銘柄とする。理由としては、時価総額が小さい企業は、ニュースがそもそも存在しない可能性があることに加えて、実際の投資において流動性の問題などが生じるためである。

分析期間は 1983 年から 2016 年末までとする。

ニュース情報は、日経新聞朝刊の記事情報を日経テレコンから取得し利用した。各記事には、メタ情報としてどの企業に関するニュースかという情報が付与されている。これを用いて、ニュースと企業とのマッチングを行う。

株価を始めとする個別企業に関するデータについては SPEEDA から取得した。

<sup>†</sup>野村アセットマネジメント株式会社

<sup>‡</sup>筑波大学大学院

記事の極性を評価するための極性辞書は、東北大学で公開されている日本語評価極性辞書の用言編 (小林 [2005]), 名詞編 (東山 [2008]) を利用した。

## 4.2 ニュースの極性評価

ニュースの極性は、先述の極性辞書を用いて評価する。用言編, 名詞編ともに, ポジティブもしくはネガティブと極性が付与されている単語のみを抽出する。ポジティブと評価される単語は約 4000 単語, ネガティブと評価される単語は約 6000 単語となった。さらに, ポジティブな単語には+1 のスコアを, ネガティブな単語には-1 のスコアを付与した。その後, 個々のニュース記事についてポジティブな単語の数, ネガティブな単語の数を数えあげ, 下記の式によってスコアを付与する。

$$x = \frac{(+1) \times w_p + (-1) \times w_n}{w_a}$$

ここで,  $w_p$  はニュース記事に含まれるポジティブな単語の数,  $w_n$  はネガティブな単語の数,  $w_a$  は記事に含まれる単語の総数である。このスコアをニュースの極性と定義する。

## 4.3 仮説の検証方法

仮説の検証については, 基本的にはパネルデータを用いた回帰分析を行う。被説明変数を個別企業のニュース発表後のリターン等とし, 説明変数は個別企業に関するニュースの極性を用いる。また仮説毎に必要な変数を適宜説明変数として加える。

## 5. 結果

仮説 i を以下の回帰分析で検証した。

$$r_{i,t} = \beta_1 x_{i,t} + \beta_0$$

ここで,  $i$  は個別企業を,  $t$  は時点を表す。  $r$  はニュース発表日の企業への TOPIX 超過プライスリターンである。  $x$  はニュースの極性スコアである。同日に同一企業に対する記事が, 新聞に複数掲載される場合がある。その場合, 極性スコアは全ての記事の平均値を利用した。

この回帰分析の結果を表(1)に示す。

表(1) 回帰分析の結果

説明変数	係数	t値
ニュース極性	0.0187	8.103 ***
切片	0.0004	13.074 ***
自由度修正済み決定係数		0.001
データ数		310232

\*は5%, \*\*は1%, \*\*\*は0.1%有意であることを示す

結果より, ニュースの極性の係数はプラスに統計的に有意である。仮説の通り, 企業に関するニュースがポジティ

ブ(ネガティブ)な場合は記事発表後の株価リターンにプラス(マイナス)の影響を与えることがわかる。

仮説 ii は 1 面であるかのダミー変数を用いた回帰分析を行った。結果は, 交差項の係数は有意に負となり 1 面の記事ほどリターンが下がるという結果となった。そこで, 1 面記事の発表前後の株価リターンの推移のイベントスタディを行った。ポジティブなニュースは, 1 面に発表以前には株価は大きく変動せず, 1 面に発表されると, 株価は上昇することが分かった。一方, ネガティブなニュースは, 1 面に発表される以前から株価が低下してきており, 1 面に発表されると, 株価が逆に上昇するということがわかった。1 面に発表されるネガティブなニュースは, スクープを除き, 1 面に発表される以前から 1 面以外の部分で発表されていることが多い。株価はこれに反応して低下していると考えられる, そして, 最後に 1 面に大きく出ると, 悪材料出尽くしとなり, 逆に株価が上昇すると考えられる。このためネガティブニュースであっても 1 面記事は株価が上昇し, 仮説と異なる結果となったと思われる。

仮説 iii は, 仮説 i で用いた回帰式に時価総額の変数を加えて, 極性スコアとの交差項を加えたところ, 係数は有意に負となり, 時価総額が小さい企業ほど, 極性の影響が強くなるという結果をえた。したがって仮説通り, 時価総額が小さい企業ほど極性スコアの影響が強いことを確認できた。

## 6. おわりに

本研究により, ニュースを極性辞書で判定したセンチメントが株式市場に一定の影響を与えることを確認した。投資家は, 大量のデータを同じ方法で処理することにより, マーケットに打ち勝つことができる可能性がある。

今後は, 今回は一般的な極性辞書を用いたが, 株式市場に関するニュースには特融の単語や言い回しなどがあるため, 金融用の極性辞書の作成する。また, 実際にセンチメント指数でポートフォリオを構築した場合の収益率の計測なども行い, 実務に利用可能かについても検証する予定である。

### 参考文献

- [1] 高橋悟, 津田和彦 “アナリストレポートからの資産運用知識の学習システム, FIT2003 (2003)
- [2] 沖本竜義, 平澤英司, “ニュース指標による株式市場の予測可能性”, 証券アナリストジャーナル, 第 52 巻, 第 4 号 (2014).
- [3] 五島圭一, 高橋大志, “ニュースと株価に関する実証分析 - デイブローニングによるニュース記事の評判分析”, 証券アナリストジャーナル, 第 54 巻, 第 3 号 (2016).
- [4] 諏訪部貴嗣, “データ革命と株式運用”, 証券アナリストジャーナル, 第 53 巻, 第 4 号 (2015)
- [5] 小林のぞみ, 乾健太郎, 松本裕治, 立石健二, 福島俊一. “意見抽出のための評価表現の収集”. 自然言語処理, Vol.12, No.3, (2005).
- [6] 東山昌彦, 乾健太郎, 松本裕治, “述語の選択選好性に着目した名詞評価極性の獲得”, 言語処理学会第 14 回年次大会論文集, (2008).
- [7] Tetlock, Paul C, “Giving content to investor sentiment: The role of media in the stock market”, Journal of Finance, 62 (2007).
- [8] Tetlock, Paul C., Maya L. Saar-Tsechansky, Sofus Macskassy, “More than words: Quantifying language to measure firms’ fundamentals”, Journal of Finance, 63 (2008)