

SNS 及び, RSS と株価変動の相関に関する研究

山野剛志^{†1} 横川隼也^{‡2} 山下海^{‡3} 広瀬啓雄^{‡4}[†] 公立諏訪東京理科大学

1. はじめに

2021年1月に世界最大のコンピュータゲーム小売店である米国のゲームストップの株価の乱高下が社会問題となった。これは、ヘッジファンドが空売りしていた銘柄を SNS 等を介して結束した個人投資家が共同で買い上げ、株価が乱高下し、複数のヘッジファンドが多額の損失を計上した。

この事案から近年、株価の変動に SNS の影響が無視できなくなっていると考察される。

Twitter を用いた研究では, "Twitter mood predicts the stock market" という先行研究がある。

これは, "I feel", "I am feeling" などを含む感情が表れているツイートのみを取得し, Opinion Finder (OF) と GPOMS ツールを用いて, 分析を行い数値化する。この数値を用いて, グレンジャー因果分析と自己組織型ファジィニューラルネットワークを使用し, DJIA を予測する。

その結果, 気分状態を活用しない予測に対して, 予測の正確率を 73.3% から 87.6% と約 14% 向上させることができた。

本研究の目的は SNS 及び, RSS と株価変動の関係性について着目し, ネットで流れる情報量及び, 感情が株価変動と相関があるかを明らかにすることである。

2. 実験

本研究の構成としては, 日経平均株価に関連するワードを含むツイート, ツイート数とネット記事のタイトルをそれぞれ日ごとに取得し, 感情分析して数値化した。

そして, 取得したデータと日経平均株価の前日比との関係を調べた。

(1) 関連ワード取得

日経平均株価に関連するワードの取得手法については, 2018 年から 2021 年を対象に行った。

2018 年から 2021 年 6 月 30 日, 2021 年 7 月 14 日から 2021 年 11 月 10 日の間である, 1 年ごとに "日経平均株価" というワードの関連上位ワードを Google 上のサービスである Google Trends を用いて, 取得限界まで取得した。

(2) ツイート取得

上記で取得したワードを含むツイートを 1 日区切りとして, 1 年ごと約 4 年間取得した。この際に, 既存の API による取得方法は実行や取得数に制限が存在したため, "TWINT" ライブラリを用いた。

(3) ネット記事取得

取得方法としては, Web サイトのニュースやブログなどの, 更新情報の日付やタイトル, その内容の要約などを配信するため技術である RSS を用いた。

RSS については産経新聞, 日経新聞等の RSS が集約されている RSS 愛好会というサイトから日本市場に関するカテゴリのみを取得し, 分析対象とした。

分析対象期間は 2021 年 7 月 14 日から 11 月 10 日であり, 12 時間ごとに RSS を入手した。

(4) ポジティブ, ネガティブ分析

janome を用いて形態素解析を行ったツイート, ネット記事のタイトルを金融辞書を用いて, ポジティブ (+), ネガティブ (−) で数値化を行った。

(5) RSS, ツイート, 前日比との関係性

日経平均株価の "前日比とツイートの感情分析結果", "前日比と記事のタイトルの感情分析結果", "ツイートの感情分析結果と記事タイトルの感情分析結果" との相関を計算した後, グレンジャー因果検定により, 因果関係の有無を調べた。

また, 上記に関しては, 2 つの方法で調査した。

(i) 特定のキーワードを含む全て用いて, 感情分析した結果の日ごとの平均値

(ii) 特定のキーワードを含む全て用いて, 感情分析した結果をポジティブ, ネガティブに分類し, それぞれの日ごとの件数

3. 実験結果・考察

3-1. 前日比とツイートの感情分析結果について

表 1 前日比と相関係数が高いワードを含むツイート上位

関連ワード	SCORE
日経平均株価	0.694
株価	0.657

株式	0.617
ドル	0.499
円	0.487
村田製作所	0.240

表 1 より, 日経平均株価は, “日経平均株価” という特定のキーワードに限らず, 経済ワードをツイートする人の感情と連動していることが判明した.

また, 一番強い相関を持っていた関連ワードを含むツイートの感情分析結果と株価の前日比で因果検定を行うと, 年度によって棄却される方向が異なっていた.

つまり, 株価の変動とツイートの感情分析結果因果関係があることが判明したが, 因果を持つ方向が年度によって異なるので, 株価の変動とツイッターは双方向の因果関係は観測されなかったものの交互に影響を与えあっていることが示された.

3-2. 前日比と新聞記事とカテゴリ内の記事のタイトルの感情分析結果について

表 2 前日比と相関係数が高い新聞記事とカテゴリ内の記事のタイトルの感情分析結果上位

新聞記事／カテゴリ	SCORE
ロイター／株式市場	0.420
日経新聞／マーケット	0.345
日経新聞／速報	0.318
日経新聞／国際	0.276

表 2 より, 記事のタイトルの感情分析結果と前日比で特に相関が強く表れていたのは, 記事の中でも市場のカテゴリに分類されているものであるということが判明した.

一番強い相関を持つ新聞記事とカテゴリに含まれる記事のタイトルの感情分析結果と前日比を因果検定したが, 有意ではなかった. つまり, 連動はしているが, 双方に影響を及ぼしているかは確認されなかった.

3-3. ツイートの感情分析結果と記事タイトルの感情分析結果について

表 3 “日経平均株価” を含むツイートと相関係数が高い新聞記事とカテゴリ内の記事のタイトルの感情分析結果上位

新聞記事／カテゴリ	SCORE
日経新聞／速報	0.499
日経新聞／マーケット	0.420
ロイター／マーケット	0.393

表 3 より, 前日比と記事のタイトルよりも強く相関が表れていた.

さらに, 新聞記事とカテゴリ内の記事のタイトルの感情分析結果, ツイートの感情分析結果で因果検定を行ったところ前者から後者へは有意水準 5%で棄却された.

これにより, ニュース記事はツイートの感情分析結果に因果関係をもっており, ニュース記事がツイッターに影響を与えていることが示された.

4. まとめ

本研究では, 日経平均株価に関連するワードに関するツイートとニュース記事のタイトルの両方を感情分析した結果が共に, 株価の前日比と相関があることを示した. 加えて, ツイッターの方がより強く相関を持っていることが表 1, 2 より示された. だが, 今回ツイッター側の方が圧倒的にデータ数が多いのに対し, 記事タイトルの方は取得期間がかなり短かったため, 偏りを持つ可能性がある. 今後さらにデータ数を増やして取り組みたい.

株価の変動とニュース記事の両方からツイッターに因果関係があることが判明したため, ツイッターは様々なチャネルから影響を受けていることが示された. また, ツイッターから株価の変動に関しては, 年度によっては因果関係を持ち, そのツイッターはニュース記事からの影響を受けるため, ツイッター及び RSS の感情分析結果は株価予測の 1 要因であると考えられる.

参考文献

- [1] Ito T., Sakaji H., Tsubouchi K., Izumi K., Yamashita T. (2018)
Text-Visualizing Neural Network Model: Understanding Online Financial Textual Data.
In: Phung D., Tseng V., Webb G., Ho B., Ganji M., Rashidi L. (eds)
Advances in Knowledge Discovery and Data Mining. PAKDD 2018.
Lecture Notes in Computer Science, Springer, vol 10939, pp 247-259.
- [2] Johan Bollen, HuinaMao, Xiaojun Zengb.
“Twitter mood predicts the stock market” .
Journal of Computational Science, Volume 2, Issue 1, March 2011, Pages 18