

企業のTwitter投稿に 対する株式市場の 反応分析

3班

202220533 高橋和暉 202220526 加藤優友

202220529 後藤大河 202220523 衛藤愛羅

アドバイザー教員 三崎広海

目次

1. 研究背景
2. 既往研究
3. 研究目的
4. 手法
5. 結果と考察
6. 今後の予定
7. まとめ

1. 研究背景

投資家は様々な情報を元に投資活動を行っている

- 指標 (PER, PBR)
- ニュース (日銀, 日経新聞, Kabutan)
- SNS (Twitter, 掲示板)



1. 研究背景

投資や資産運用では、"**リスク管理**"が重要な問題

- **価格変動リスク**: 換金する際の受取金額が当初支払った金額より下回る
- **信用リスク**: 融資した会社が破たんする可能性(連鎖倒産など)

金融研究→リスク管理の研究が多い(ボラティリティ予測など)

ニュース内容が株式リターンのボラティリティ予測に有効的(五島ら^[1])

1. 研究背景

板情報とは

- 気配や気配値(けはいね)とも呼ばれ、証券取引所に出された売買注文の価格と数量を表すもの。左側には売りの数量、右側には買いの数量、中央には気配値(注文価格)が表示され、その時点でいくらなら買えるのか、売れるのかがわかる。

ボラティリティとは

- 価格変動の度合いを示す言葉で、大きいとその商品の価格変動が大きいこと、小さいとその商品の価格変動が小さいことを意味するボラティリティが大きいと取引リスクが高く、小さいとリスクが低い

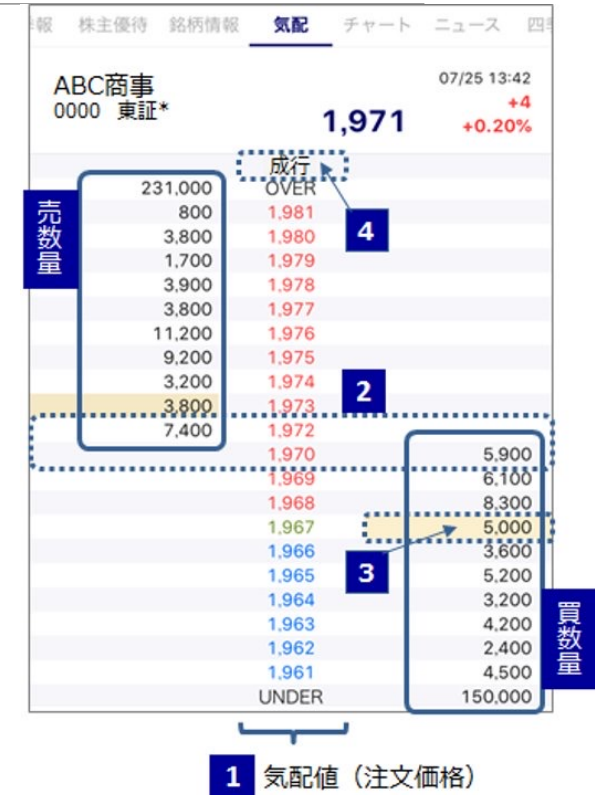


図1: 板情報の例

2. 既往研究(柳瀬ら)

・海外においてはSNS等の情報を用いて株価を予測している研究はあるが、日本国内でインターネット上のテキストデータを用いた株価予測で高い精度での予測が可能といった研究成果は少ない

➡ インターネット上のデータと株価の変動の間に相関関係があるかどうかの調査

- ・インターネット上のテキストデータ・株価データの取得
- ・テキストデータの加工
- ・定量化データの抽出および数値化
→テキストデータに対し感情分析をして感情値を算出
- ・データの分析・株価予測
→感情値と日経平均株価の相関係数を算出し、全ての感情値または相関係数の良い感情値のみを用いて機械学習

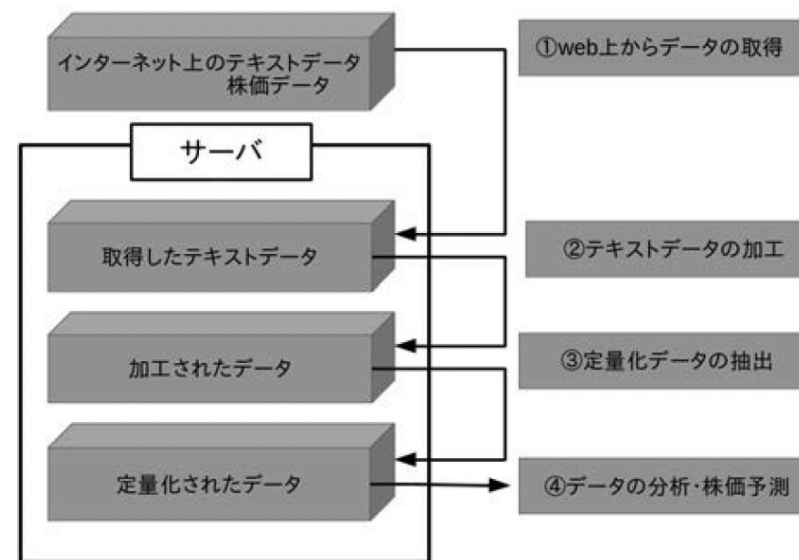


図2: システムの流れ

2. 既往研究(柳瀬ら)

研究結果

- 最も高い相関性を示したものでも0.28
- 相関係数の評価が高かったパラメータと株価上下の相関性は他のパラメータをもちいた場合に比べて高かった



**ニュース記事と日経平均株価の
推移に明確な関係性があるとは言いがたいものの、
相関係数で高い評価を得たパラメータを用いることで
(同じ時点の) 株価の変動を説明できる可能性**

2. 既往研究 (Johanら)

研究概要

既往研究よりオンラインソーシャルメディア(ブログ、ツイッターフィードなど)の情報から指標を抽出し、様々な経済・商業指標の変化を予測できることを示唆



大衆のアメリカ大統領選挙期間中の**ツイート**が、時間の経過とともにダウ平均株価の動向との関連性に関して分析

- ポジティブとネガティブを測定するOpinionFinder
 - 6つの次元で気分を測定するGoogle-Profile of Mood States (GPOMS)
- 2つの気分追跡ツールによってTwitterテキストコンテンツを分析

2. 既往研究 (Johanら)

研究結果

- ツイッターの内容から、簡単なテキスト処理によって人々の気分の変化を追跡できる
- 気分の変化は様々な社会文化的要因に高度に分化して反応する
- 「穏やか」「嬉しい」などの**一部の感情はDJIA株価変動に影響**
- 過去のDJIAの値と気分時系列に基づいて学習した自己組織化ファジーニューラルネットワークは、DJIAの終値を予測する最も基本的なモデルの精度を大幅に向上



**大衆ではなく企業のツイートは
株価変動にどのように影響するかという研究は行われていない**

2. 既往研究

柳瀬ら^[2]

ネットニュースを**7つの感情因子**に分けて感情解析を行い、**株価との相関を分析した**

- 明確な相関関係はなかったものの、相関係数の良いパラメータを用いることで**精度良く同じ時点の株価を説明できる可能性**がある

Johanら^[3]

大衆のアメリカ大統領選挙期間中の**ツイートを6感情に分類し、ダウ平均株価の動向との関連性**に関して分析した

- 「穏やか」「嬉しい」などの**一部の感情は株価変動に影響**する



SNS等から入手した情報と株価との相関を分析した研究は**数多く存在**する

3. 研究目的

鳥海ら^[4]の研究において「価格データよりも板情報に、板情報よりもトレーダによる注文に、より多くの情報が含まれていると考えられる。」と述べられている

ボラティリティを用いた株価変動についての論文はあるが、板情報の反応の研究は少ない

- 高頻度データを用いて、板情報の反応分析をリスクマネジメントに含めることが可能か評価

反応の対象として、Twitterに着目

Twitterは日々膨大な情報が配信されているという理由から

- ツイート情報が板情報に対して本源的な情報を持つか調査

3. 研究目的

板情報の反応から、企業のツイートが投資家の行動に与える影響を検証する

企業のツイートと板情報の反応の間に
相関関係の有無を調査することを目的としている

➤ 企業のTwitter投稿と板情報(10本気配)の相関関係を明らかにする

4. 研究手法

- Twitter を利用して企業のツイート情報を取得
- 日経NEEDSデータから板情報データを取得する
- 上記2つのデータを用いてツイート時刻前後の売買の注文変動率を計算する

Twitterデータの取得

対象：任天堂株式会社

→ Twitter投稿数とフォロワーの量、上場企業という観点から選択

期間：2020年3月～2021年2月の1年間

データ：西暦、日付、時刻、ツイート内容

計887個のツイートを取得

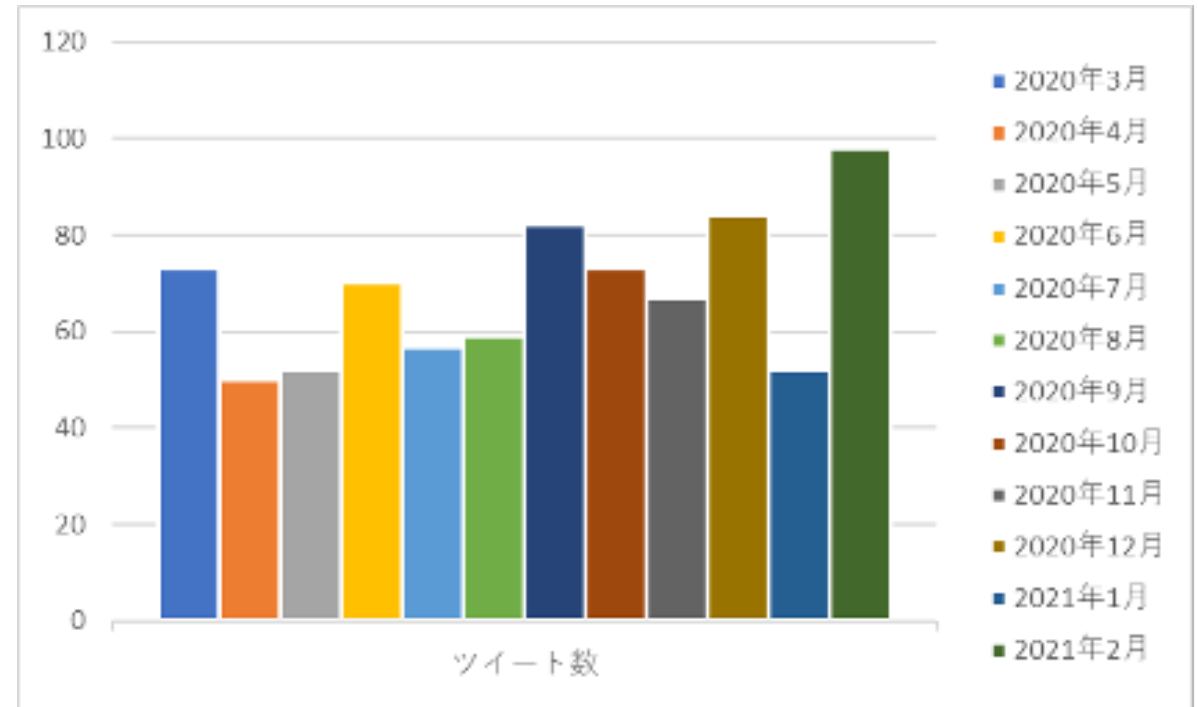


図1 任天堂株式会社のツイート数

ツイート情報の取得方法

Twitter APIや手作業で企業のツイート情報を取得

@nintendoのアカウントのツイート内容、投稿日時を取得出来た。

yyyy	mm	dd	h	m	contents
2020	3	1	18	5	『Dead Cells』の「#
2020	3	2	17	23	[任天堂HP]「キャン
2020	3	4	10	8	[トピックス]「ニャニ
2020	3	4	10	40	[トピックス]『ルイー
2020	3	4	17	19	[任天堂HP]「キャン
2020	3	5	10	2	[トピックス]キミをみ
2020	3	5	17	13	[任天堂HP]「キャン

yyyy/mm/dd h:m = yyyy年mm月dd日 h時m分

Contents: ツイート内容

板情報データ

日経NEEDSデータから、任天堂株式会社の銘柄の情報を抽出

データ: 売買注文の価格、数量、注文時刻など計95項目

1か月で約400万件のデータが存在→Pythonを用いて、データの加工、分析

	0	1	2	3	4	5	6	7	8	9	...	85	86	87	88	89	90	91	92	93	94
0	1200	20200302	11	1	1	7974	NaN	80000.0	NaN	80000050524	...	0	NaN	NaN	0	NaN	NaN	0	NaN	NaN	0
1	1200	20200302	11	1	1	7974	NaN	80000.0	NaN	80000052430	...	0	NaN	NaN	0	NaN	NaN	0	NaN	NaN	0
2	1200	20200302	11	1	1	7974	NaN	80000.0	NaN	80000053219	...	0	NaN	NaN	0	NaN	NaN	0	NaN	NaN	0
3	1200	20200302	11	1	1	7974	NaN	80000.0	NaN	80000100946	...	0	NaN	NaN	0	NaN	NaN	0	NaN	NaN	0
4	1200	20200302	11	1	1	7974	NaN	80000.0	80000.0	80000124476	...	0	NaN	NaN	0	NaN	NaN	0	NaN	NaN	0

5 rows × 95 columns

板情報データの中身の図

出典：株式会社日本経済新聞社 NEEDSデータ

考案した指標

N_buy, N_sell : ツイート投稿前5分間の板情報の買注文数, 売注文数

N'_buy, N'_sell : ツイート投稿後5分間の板情報の買注文数, 売注文数

$Score_buy = (N'_buy - N_buy) / N_buy$ (板情報の買注文数の増減率)

$Score_sell = (N'_sell - N_sell) / N_sell$ (板情報の売注文数の増減率)

5. 結果と考察

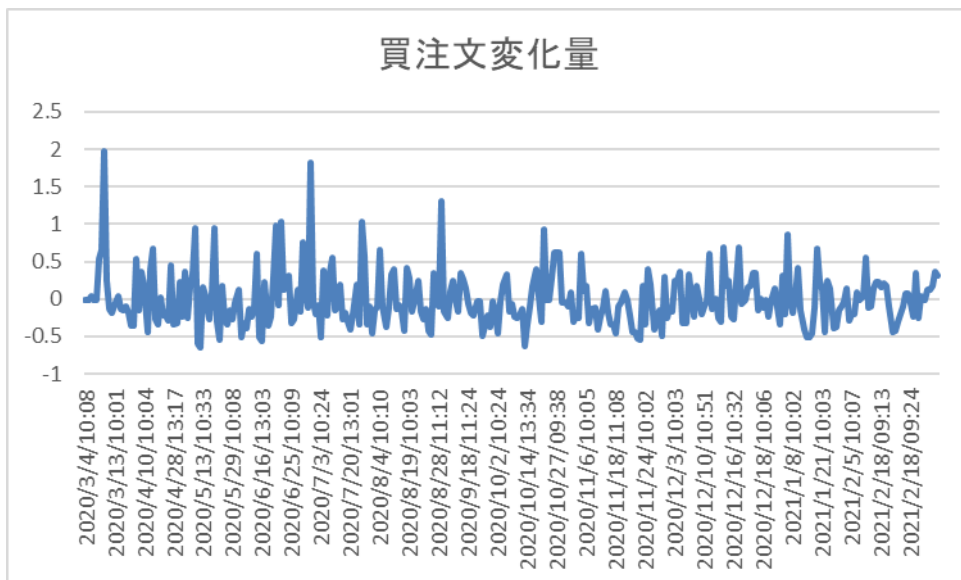
企業のツイート投稿日時のみで板情報の反応を分析

- ・ Twitterアカウントを持つ上場企業の選定
→ **Twitter投稿**及び**フォロワーの量が多く**、
上場企業であるという観点から**任天堂株式会社**
社を選定
- ・ Twitterデータと、板情報データを基に、企業の
Twitter投稿の時刻に関連する2020年3月～2021年2
月の売買注文数の変動率と平時の変動率を求めた。

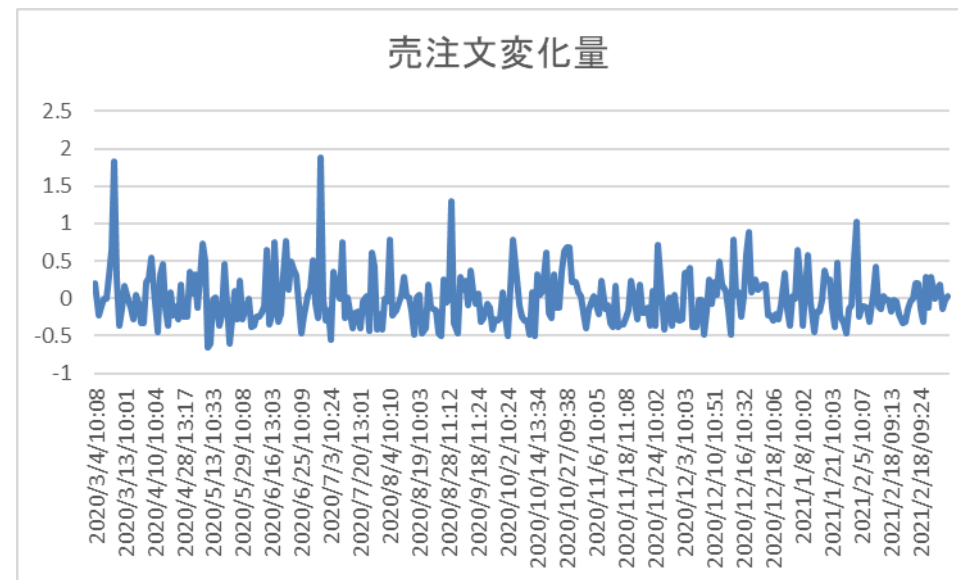


任天堂株式会社 Twitterより

1年間の売買注文の変化量(Twitter)

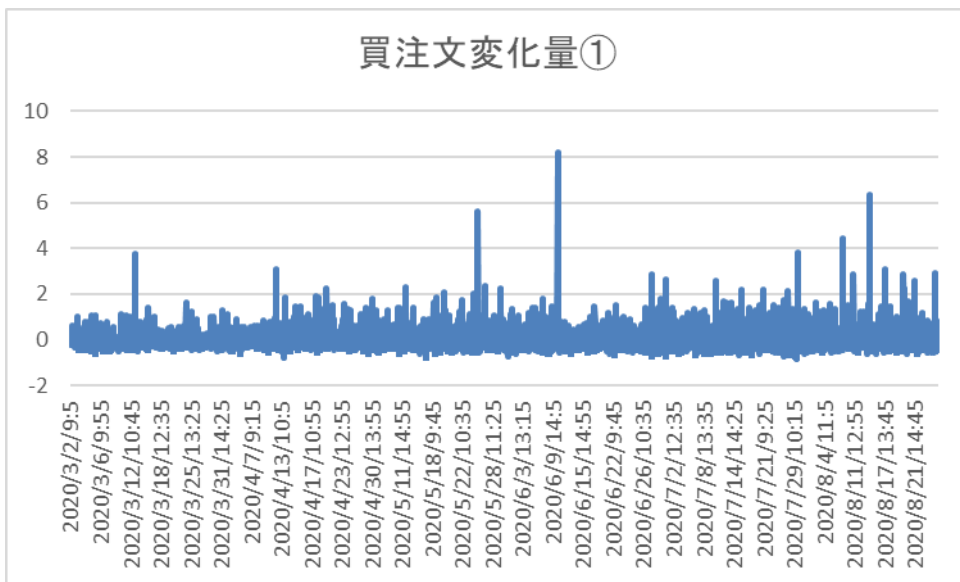


Twitter前後の買い注文変化量

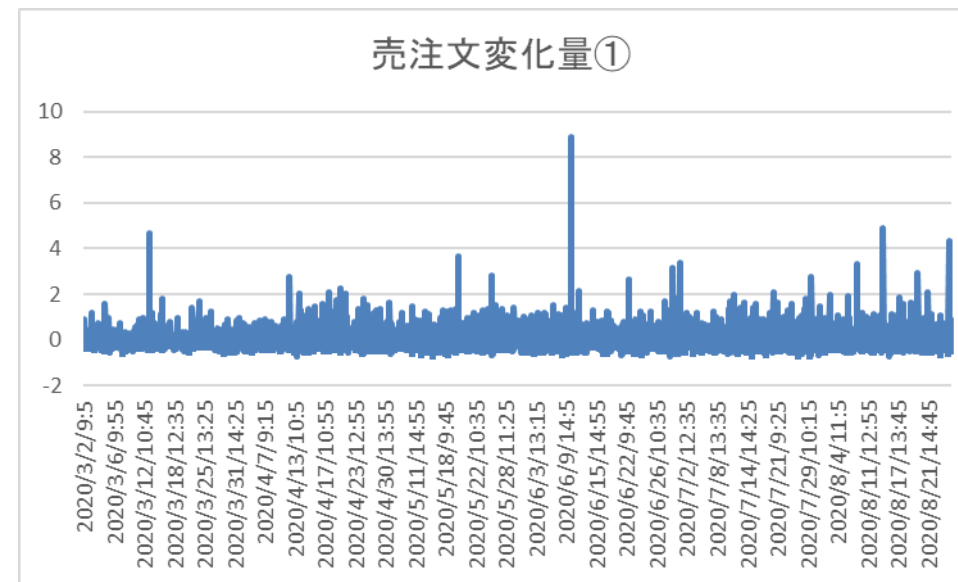


Twitter前後の売り注文変化量

1年間の売買注文の変化量(平時)前半部

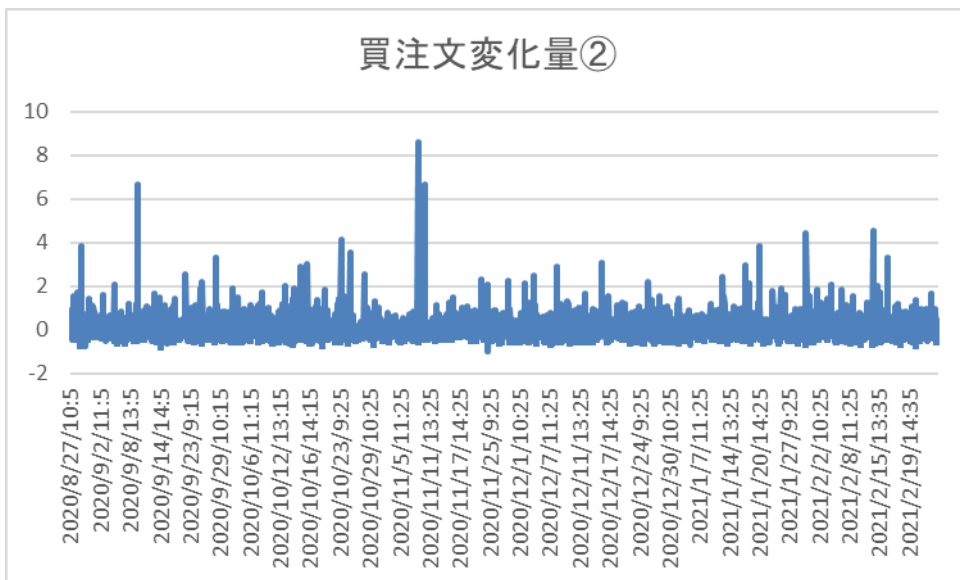


平時の買い注文変化量①

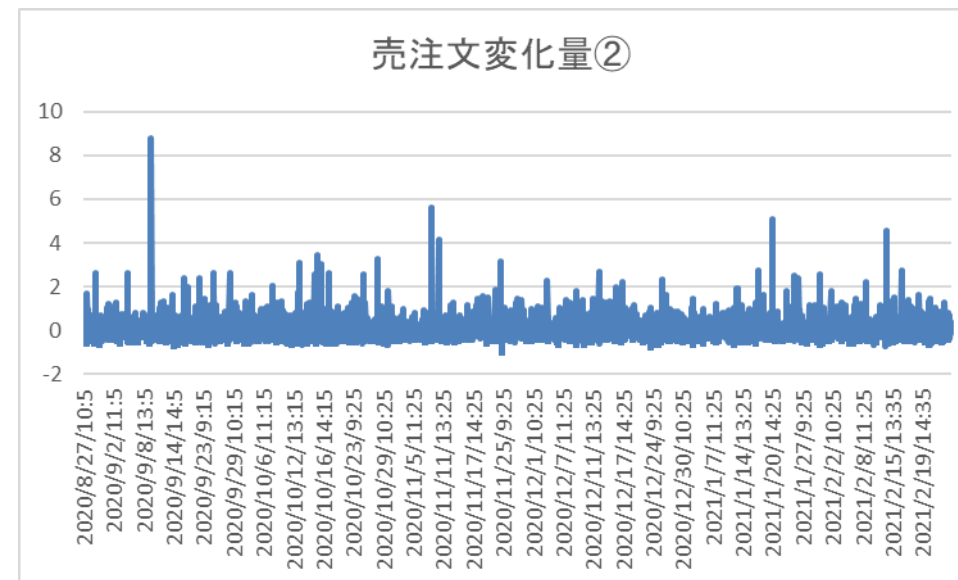


平時の売り注文変化量①

1年間の売買注文の変化量(平時)後半部



平時の買い注文変化量②



平時の売り注文変化量②

基本統計量

基本統計量 Twitter

	買注文變化量	売注文變化量
count	321	321
mean	-0.005083291	-0.003936042
std	0.363418335	0.352367768

基本統計量 平時

	買注文變化量	売注文變化量
count	7260	7260
mean	0.06848971	0.063102486
std	0.509056597	0.493612594

5. 結果と考察

- 企業がツイートを行うと、平時と比べて注文数は減少傾向になり、標準偏差から、株価の反応は小さくなることが分かった。
- しかし、企業のツイートが板情報の変動に影響を及ぼしているという結論には不十分である。
- 公式ツイートだけの研究ではなく、ニュースやアナリストレポートなどのデータを使用した分析が必要だと考える。

6. 今後の予定

- 他の情報サイトなども考慮にいて、板情報の変動を考察し、リスクマネジメントの評価を行う
- 他の時間帯の板情報との増減率の違いを分析
- 他の企業の板情報の増減率も調査し、比較分析する
- 今回の指標では、負の値の最大が-1のため、公平に見るためにも新たな指標を導入する必要がある
- 約定価格に関して分析を行い、本研究の板情報分析との違いを考察する

7. まとめ

- 独自の手法を用いて、企業のTwitter投稿に対する株式市場の反応分析を行った。
- 結果として、売り注文、買い注文ともに、平時と比べて注文数は減少傾向になり、標準偏差から株価の反応は小さくなることが分かったが、リスクマネジメントの評価を行うには不十分であった。
- 今後も調査の対象を増やすこと、また調査期間を延長することで、さらなる考察を行いたい。

8. 参考文献

1. 五島圭一, 高橋大志, 山田哲也: 自然言語処理による景況感ニュース指数の構築とボラティリティ予測への応用
2. 柳瀬 仁洋, 小高 知宏, 黒岩 丈介, 白井 治彦: 株価予測におけるビックデータの利用, 福井大学大学院工学研究科研究報告, vol.68, pp.59-66, 2018
3. Johan Bollena, Huina Maa, Xiaojun Zeng, "Twitter mood predicts the stock market", Journal of Computational Science 2 (2011) 1–8
4. 鳥海 不二夫, 西岡 寛兼, 梅岡 利光, 石井 健一郎: 板情報による市場相違性の検出, 人工知能学会論文誌, 2012, 27 巻, 3 号, pp. 143–150, 2012