



経済に関する オルタナティブ・データを考慮した 金融マーケット予測手法の開発

麻生 到

富山県立大学大学院情報システム工学専攻 情報基盤工学講座

平成 30 年 11 月 30 日

- 1.1 はじめに
- 1.2 本研究の概要
- 2.1 Web 情報からのテキスト・マイニング
- 2.2 経済用語とセンチメント分析
- 2.3 センチメント分析の感情スコア
- 3.1 為替のファンダメンタル分析
- 3.2 為替のテクニカル分析
- 3.3 本研究の位置付け
- 4.1 ルールベース構築のためのクラスタリング (1/2)
- 4.2 ルールベース構築のためのクラスタリング (2/2)
- 4.3 洗練化のための自動売買システムの構築



1.1 はじめに

本研究の背景

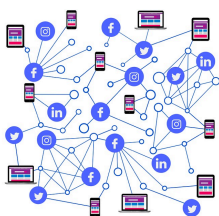
計算機科学の発展により、ビッグデータの蓄積や蓄積したデータを機械学習を用いて分析することが可能となっている。

ビッグデータを用いたデータ分析は、金融経済現象にも応用されている。

近年では、オルタナティブ・データと呼ばれるデータを活用することで新たな金融工学の地平が切り開かれている。

本研究の目的

Web 上の景気に関する情報や SNS からの情報のテキストからセンチメント分析を行うことによって、金融マーケットの状況を予測する。





1.2 本研究の概要

1.1 はじめに

1.2 本研究の概要

2.1 Web 情報からのテキスト・マイニング

2.2 経済用語とセンチメント分析

2.3 センチメント分析の感情スコア

3.1 為替のファンダメンタル分析

3.2 為替のテクニカル分析

3.3 本研究の位置付け

4.1 ルールベース構築のためのクラスタリング (1/2)

4.2 ルールベース構築のためのクラスタリング (2/2)

4.3 洗練化のための自動売買システムの構築

収集する Web 情報

本研究では、経済に関する Twitter からの情報を取得し活用する

金融市場における知識発見

行動経済学では、感情が個人の判断に大きく影響することが分かっている。

そこで、Twitter から感情に関わる情報を引き出す事で、マーケットの状況を予測出来るという仮定する。

特定のアカウントからツイートを取得して、**センチメント分析**の分析結果を予測に役立てる。

マーケット状況の予測

センチメント分析によって得られた結果を入力として、**為替の予測**を行うことでマーケット状況を予測できているかの検証を行う。



2.1 Web 情報からのテキスト・マイニング

Twitter 上の特定のアカウントからツイートを取得する

特定のアカウント

政治家やニュース・メディアなど各ジャンルにおいて、フォロワーが多いアカウント (8320 アカウント) のツイートを取得した

```
Sat Jun 30 09:45:11 +0000 2018,sacai 2019年リゾートコレクション、ペンドルトンとナイキのコラボでより自由なハイブリッドヘ  
Sat Jun 30 09:30:07 +0000 2018,ドルチェ&ガッバーナ"クマ"のメンズバッグ&サンダル、"ゼブラ"モチーフのキッズバッグ - https:  
Sat Jun 30 09:15:09 +0000 2018,オニツカタイガーの新スニーカー「TSUNAHIKI」ユニークな靴底の綱引き専用シューズがモデル - htt  
Sat Jun 30 09:00:08 +0000 2018,ロットホロン 2018年秋冬コレクション公開 - https://t.co/ntbFNW9SzL https://t.co/aBSWj  
Sat Jun 30 08:45:08 +0000 2018,ジェラート ピケのカフェが京都にオープン - オーガニック素材にこだわったクレープやハンバーガー  
Sat Jun 30 08:30:07 +0000 2018,メゾン マルジェラ、メンズ初の「アーティザナル」コレクションを19春夏シーズンで発表 - https:/  
Sat Jun 30 08:15:08 +0000 2018,「におい展」静岡・札幌・福岡で - 嗅覚で楽しむ展覧会、世界一臭い食べ物から最高級の香料まで -
```

Fig. 1: 取得したツイートの例

取得したテキストに対して、形態素解析を行い品詞ごとに分割する

使用する品詞

- 1 名詞
- 2 副詞
- 3 動詞
- 4 形容詞



2.2 経済用語とセンチメント分析

従来のセンチメント分析

極性辞書 (日本語評価極性辞書など) を用いて、それらの単語の頻度を数えてセンチメント分析を行う

従来のセンチメント分析では、単語の頻度しか考慮していないため文脈を考慮することができないといった問題点がある。

そこで、本研究では前後の文脈を考慮することができる Word2vec を用いて辞書を作成する。また、辞書を作成するには経済に関する文章への考慮が必要なため内閣府が公開している「景気ウォッチャー調査」を用いる。

景気ウォッチャー調査

タクシー運転手や小売店の店主など、景気に敏感な人々に景気に対する判断とその理由を回答してもらう調査

▲	食料品製造業	・来月も得意先がなくなる。景況感が悪く、物の動きも悪くなる。	
▲	通信業 (経営)	・地域の景況感からも、悪くなることはあっても良くなることはなさそうである。	
×	出版・印刷・	・利益が出ないため、資金繰りが大変である。	
◎	職業安定所 (・北朝鮮問題などの外的要因がなければ、当面、回復傾向が続くとみている。	

Fig. 4: 景気ウォッチャー調査の例



2.3 センチメント分析の感情スコア

従来のセンチメント分析の感情スコア

従来の極性辞書を用いた分析や機械学習を用いた分析では、感情軸がポジティブ・ネガティブの1つの指標のみである

従来の感情軸が1つのセンチメント分析では、Twitterなどの人が発信しているテキストを分析する場合には感情の表現力が乏しいと考えられる。

本研究の感情スコア

心理学者 Plutick が考案した Plutick の感情の環に基づいて8つの感情軸から感情スコアを算出する





3.1 為替ファンダメンタル分析

ファンダメンタル分析

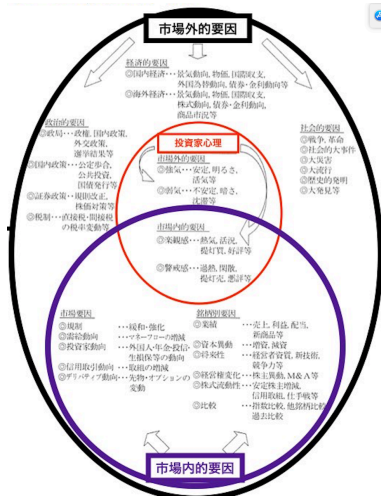
国際的な経済の動きや個別の企業の財務情報を基に分析する方法

ファンダメンタル分析の例

- 1 金利が上がれば通貨も上がる
→ 高金利の通貨に両替して、銀行預金したほうが得なため
- 2 貿易黒字になると通貨が上がる
→ 国内の通貨の需要が増えるので価格が上がる
- 3 景気が良くなると通貨が上がる

ファンダメンタル分析は市場内の要因を考慮して分析を行っている

しかし、要因が多すぎてどの要因を考慮するかが困難なため予測が難しい





3.2 為替のテクニカル分析

テクニカル分析

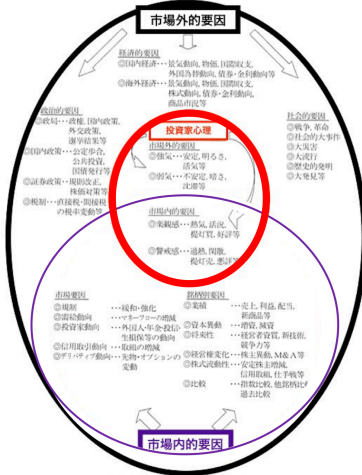
現在の相場の流れは下降トレンドなのか上昇トレンドなのかといったことを把握する方法

テクニカル分析の例

- 1 **トレンド系指標**
→ 「トレンドの勢い」を考慮する
- 2 **オシレータ系指標**
→ 買われすぎや売られすぎであるかを確認する
- 3 **出来高指標**
→ 買う勢いが強いのか売る勢いが強いのかを比較する

テクニカル分析では過去の時系列パターンなどを分析している

しかし、テクニカル分析では論理性に欠けているため根拠が弱い





3.3 本研究の位置付け

1.1 はじめに

1.2 本研究の概要

2.1 Web 情報からのテキスト・マイニング

2.2 経済用語とセンチメント分析

2.3 センチメント分析の感情スコア

3.1 為替のファンダメンタル分析

3.2 為替のテクニカル分析

3.3 本研究の位置付け

4.1 ルールベース構築のためのクラスタリング (1/2)

4.2 ルールベース構築のためのクラスタリング (2/2)

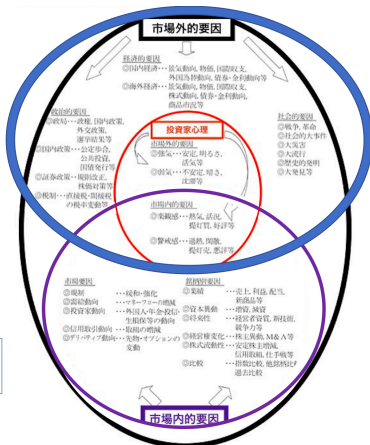
4.3 洗練化のための自動売買システムの構築

提案手法

オルタナティブ・データから景況感や社会情勢の情報を活用することで経済的要因や社会的要因、政治的要因を分析対象に入れることが可能となる

位置付け

本研究では、従来のファンダメンタル分析やテクニカル分析ではカバーしきれなかった要因を考慮することでマーケットの変化にも対応できる分析手法を提案する





4.1 ルールベース構築のためのクラスタリング (1/2)

為替 (ドル円) が上がるか下がるか変化なしといったラベル付けし, 3 クラスのクラスタリングを行う
クラスタリングを行う際の入力としては, Pultick の感情の環に基づいた 8 つの感情スコアを使用する

感情スコアの算出

ある感情 c の 1 ツイートあたりの感情スコア $s_i^{(c)}$ を以下に示す

$$s_i^{(c)} = \frac{\sum_{l=0}^L x_l^{(c)}}{L} \quad (1)$$

$x_l^{(c)}$: 感情 i のスコア, L : 分析対象となる 1 ツイートあたりの単語数
そして, 1 日あたりのスコア S_i を以下に示す

$$S_i^{(c)} = \frac{\sum_{n=0}^N s_n^{(c)}}{N} \quad (2)$$

N : 1 日あたりのツイート数
また,

$$S = \{S_i^{(c)} | c = 0, 1, \dots, 8\} \quad (3)$$

以上のような感情スコアを入力と定義する. そして, その入力に従う確率分布を多変量ガウス分布と仮定し, そのパラメータである平均ベクトルと分散共分散行列を推定することによってクラスタリングを行なっていく



4.2 ルールベース構築のためのクラスタリング (2/2)

問題設定

M 日間ツイートを取得した場合のデータセットは,

$$S_{(z)}^{(c)} \equiv \{S_m^{(c)} \in \mathbb{R}^D | (z_m) = (z), m = 0, 1, \dots, M-1\} \quad (4)$$

ここで, z は為替の変化状態を表す

ベイズ推論による学習則

為替の変化状態 z が与えられたとき, 感情スコア $\{s^{(c)}\}$ が算出されたことを支持する度合いは以下のよう
に示すことができる.

$$\begin{aligned} P(z | S, \{s^{(c)}\}) &\propto \prod_{c=0}^{C-1} \int \int \mathcal{N}_D(s^{(c)} | \mu_{(z)}^{(c)}, (\Lambda_{(z)}^{(c)})^{-1}) \times Q(\mu_{(z)}^{(c)}, \Lambda_{(z)}^{(c)} | X^{(c)}) d\mu_{(z)}^{(c)} d\Lambda_{(z)}^{(c)} \\ &= S_D(s^{(c)} | m^*, Q^*, f^*) \end{aligned} \quad (5)$$

$$m^* = m_{(e,z)}^{(c)} \quad (6)$$

$$Q^* = (\nu_{(e,z)}^{(c)} - D + 1) \frac{\beta_{(e,z)}^{(c)}}{\beta_{(e,z)}^{(c)} + 1} W_{(e,z)}^{(c)} \quad (7)$$

$$f^* = \nu_{(e,z)}^{(c)} - D + 1 \quad (8)$$

$$(9)$$

ここで, $\mathcal{N}_D(s^{(c)} | \mu_{(z)}^{(c)}, (\Lambda_{(z)}^{(c)})^{-1})$ は $s^{(c)}$ の平均ベクトル μ と分散共分散行列 Λ^{-1} の D 変量ガウス分布であり, $Q(\mu_{(z)}^{(c)}, \Lambda_{(z)}^{(c)} | X^{(c)})$ は事後分布である.

1.1 はじめに

1.2 本研究の概要

2.1 Web 情報からのテキスト・マ
イニング

2.2 経済用語とセ
ンチメント分析

2.3 センチメント
分析の感情スコア

3.1 為替のファン
ダメンタル分析

3.2 為替のテクニ
カル分析

3.3 本研究の位置
付け

4.1 ルールベース
構築のためのクラ
スタリング (1/2)

4.2 ルールベース
構築のためのクラ
スタリング (2/2)

4.3 洗練化のため
の自動売買システ
ムの構築



4.3 洗練化のための自動売買システムの構築

自動売買システムに活用する分析結果の算出

自動売買システムにおいて買う・売る、動かないの判断は式 (5) を用いて以下のように求める

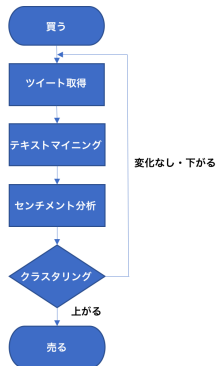
$$z_1^* = \arg \max_{z \in \mathbb{T}_M} P(z | S, \{s^{(c)}\}) \quad (10)$$

z_1^* を投資判断に活用する

z_1^* の分析結果を元に自動売買を行なっていく

右の図のように、ドル円を買った状態からスタートする。

買った日のツイートを取得しテキストマイニングを行い、その日の景況感や社会情勢などの情報からセンチメント分析とクラスタリングを経て投資判断して売るタイミングを決める



1.1 はじめに

1.2 本研究の概要

2.1 Web 情報からのテキスト・マイニング

2.2 経済用語とセンチメント分析

2.3 センチメント分析の感情スコア

3.1 為替のファンダメンタル分析

3.2 為替のテクニカル分析

3.3 本研究の位置付け

4.1 ルールベース構築のためのクラスタリング (1/2)

4.2 ルールベース構築のためのクラスタリング (2/2)

4.3 洗練化のための自動売買システムの構築



4.1 数値実験ならびに考察 (経済用語と為替の関連検証)

1.1 はじめに

1.2 本研究の概要

2.1 Web 情報からのテキスト・マイニング

2.2 経済用語とセンチメント分析

2.3 センチメント分析の感情スコア

3.1 為替のファンダメンタル分析

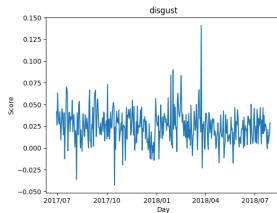
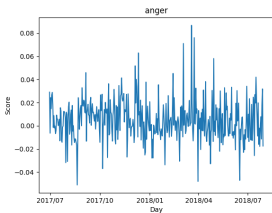
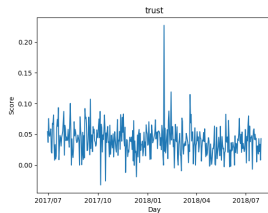
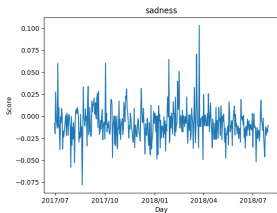
3.2 為替のテクニカル分析

3.3 本研究の位置付け

4.1 ルールベース構築のためのクラスタリング (1/2)

4.2 ルールベース構築のためのクラスタリング (2/2)

4.3 洗練化のための自動売買システムの構築





4.2 数値実験ならびに考察 (獲得されたルールベース)

1.1 はじめに

1.2 本研究の概要

2.1 Web 情報からのテキスト・マイニング

2.2 経済用語とセンチメント分析

2.3 センチメント分析の感情スコア

3.1 為替のファンダメンタル分析

3.2 為替のテクニカル分析

3.3 本研究の位置付け

4.1 ルールベース構築のためのクラスタリング (1/2)

4.2 ルールベース構築のためのクラスタリング (2/2)

4.3 洗練化のための自動売買システムの構築

