

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

富山県立大学情報システム工学専攻

麻生 到

1 はじめに

現在、ツイッターなどのマイクロブログには、様々なニュースやそれに対する人々の反応が書かれており、その情報量は膨大かつ、増加し続けている。この膨大な情報を実世界の動きを観測するためのソーシャルセンサーとして利用する研究の数は増加しており、観測する対象を予め設定し、それについて詳細な分析を行ったものが多く見られる。特に、経済動向を分析対象としたものとして、ツイッターからキーワードを用いて株式に関する情報を収集し、株価動向との関連の分析に取り組んだ事例があるなど、ツイッター情報は経済動向の分析に大いに用いられている。

また、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。その分析は金融経済現象へ応用されている。現在では、オルタナティブ・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。

既存研究としては、Bollenらが、ツイートを対象に OpinionFinder(OF)と Google-Profile of Mood States を用いて、「calmness」などの6つの心的状態を表す指数を抽出し、ダウ平均株価の予測を行った。しかし、分析対象となるツイートは「I feel」、「I'm」といった心的状態を明言したものに限られていることに加えて、ツイート情報はダウ平均株価の過去の数値データによる予測を補うものとして用いられている。

本研究では、Twitter の情報を取得し、その情報を活用することで金融マーケットの状況を予測する。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判明している。そこで、Twitter から感情に関わる情報を引き出すことで、マーケットの状況を予測出来るという仮定に基づき為替予測を行うシステムを開発する。

2 分析手法

本研究では、Twitter からの情報を収集し、収集した情報をセンチメント分析・クラスタリングを経て為替取引を行う。

2.1 今回の取り組み

今回は、サーバへのティックデータの蓄積と LSTM を用いて為替予測を行った。

2.1.1 ティックデータの蓄積

サーバへ Oanda API を用いてティックデータの蓄積を行った。図1にデータの一部を示す。

```
2019-08-07 05:10:46,106.130
2019-08-07 05:10:54,106.128
2019-08-07 05:10:54,106.126
2019-08-07 05:10:59,106.123
2019-08-07 05:11:01,106.126
2019-08-07 05:11:01,106.126
2019-08-07 05:11:11,106.123
2019-08-07 05:11:17,106.126
2019-08-07 05:11:18,106.130
2019-08-07 05:11:28,106.128
```

図1: 1 ティックデータの一部

蓄積しているデータは、タイムスタンプとドル円の価格の CSV 形式のデータである。

2.1.2 LSTM を用いた為替予測

とりあえず、ティックのデータを蓄積している間に予測に用いるモデルを構築した。分析手法として、時系列データの予測が得意な LSTM を用いる。

2.1.3 LSTM

LSTM は、従来の RNN において問題視されていた勾配消失や発散を解決するために提案されたセル内には、主に忘却ゲートと入力ゲート、出力ゲートがある。忘却ゲートでは、通過させる情報と通過させない情報を制御する。入力ゲートでは、セルの状態を更新する。出力ゲートでは、隠れユニットの値を更新する。図2に LSTM の概要図を示す。

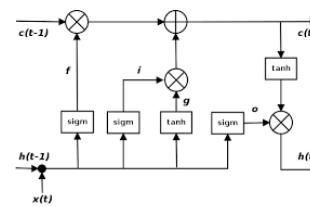


図2: 2 LSTM の図

2.1.4 分析結果

訓練データには、1ヶ月分の為替データ(分足)を用いて、テストデータとして2週間分のデータを用いた。分析結果を図3に示す。

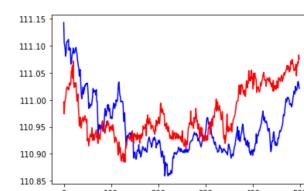


図3: 3 分析結果

赤チャートは予測結果であり、青チャートは正解のデータである。

3 おわりに

今後の展望としては、予測手法の改善や変更であったり、対象がトランプ大統領だけであったので分析対象を増やす。