

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、「calmess」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事で、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。

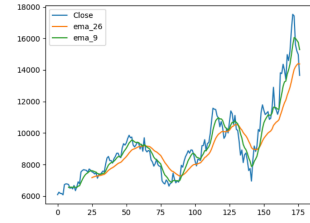


図 1: 1 EMA と為替のグラフ

2.1.2 ボリンジャーバンド

ボリンジャーバンドは、「標準偏差」と「正規分布」に従った考え方である

$$\sigma = \sqrt{\frac{\sum_{i=1}^m (a_i - \bar{a})^2}{m-1}} \quad (4)$$

m :期間, a :為替の価格, \bar{a} :期間中の為替価格の平均
ボリンジャーバンドと為替のグラフを図 2 に示す。

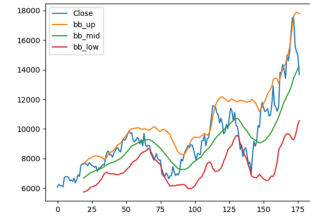


図 2: 2 ボリジャーバンドと為替のグラフ

為替の取引は、正規分布の「68-95-99 ルール」に基づいて行う。

2 今回の取り組み

テクニカル指標の実装とその指標を用いた為替予測

2.1 テクニカル指標

本研究では、EMA(指数移動平均線)、ボリジャーバンド、ストキャスティクス、MACDを用いる

2.1.1 EMA

EMA は、SAM に対して直前の為替を重視するような指標である。SAM は以下のように示される。

$$S = \frac{a_{m-1} + a_{m-2} + \dots + a_{m-m}}{m} \quad (1)$$

S :SMA, m :期間, a :為替の価格

そして、EMA は以下の様になる。

$$E_n = E_{n-1} + \alpha(a_n - E_{n-1}) \quad (2)$$

$$\alpha = \frac{2}{m+1} \quad (3)$$

E :EMA, m :期間, a :為替の価格

EMA と為替のグラフを図 1 に示す。

為替の取引のルールは、短期と長期が交わる瞬間をタイミングとしている。

2.1.3 ストキャスティクス

ストキャスティクスは、現在の相場に対して一定期間の変動幅に基づいて、売られすぎているか・買われすぎているかを判断するための指標である。% K, % D, % SD と呼ばれる指標を用いる。

$$K = \frac{C - L_m}{H_m - L_m} \times 100 \quad (5)$$

K :% K, C :現在の終値, L_m :過去 m 日間の最安値, H_m :過去 m 日間の最高値

% D は、% K の単純移動平均、% SD は、% D の単純移動平均でそれぞれ表される。ルールとしては、% K と % D が 20 % もしくは 30 % 以下になる (買いサイン)、% K が % D を下から上へクロス (買いサイン)、% K と % D が 70 % もしくは 80 % 以下になる (売りサイン)、% K が % D を上から下へクロス (買いサイン) などが考えられる。

2.1.4 MACD

MACD では「MACD」、「短期 EMA」、「長期 EMA」、「シグナル」の主に 4 つの指標を計算する必要がある。

$$MACD = EMA(12) - EMA(26)$$

シグナルは、MACD の指数平滑移動平均で表される。MACD とシグナルのグラフを図 3 に示す。

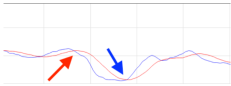


図 3: 3 MACD とシグナルのグラフ

取引のタイミングは、MACD とシグナルがクロスする瞬間である。

2.2 予測手法

前回までは、国際学会で提案したベイズ理論を用いる予定であったが、テクニカル指標を用いるため時系列を考慮しなければならない。そのため、時系列データに効果的である LSTM を用いて予測を行う。入力テーブルを図 4 と図 5 に示す。

Date	Open	High	Low	Close	Volume	ema_26	ema_9	hh_up	hh_mid	hh_low	per_s	per_d	macd	signal
2018-10-31 08:55:00	113.069	113.067	113.069	113.067	2	113.050908	113.053252	113.066771	113.04875	113.030729	1.000000	1.000000	0.000153	-0.004871
2018-10-31 08:56:00	113.067	113.067	113.069	113.067	5	113.052031	113.056715	113.067866	113.04900	113.030132	1.000000	1.000000	0.002058	-0.003485
2018-10-31 08:57:00	113.067	113.069	113.069	113.066	4	113.053007	113.059002	113.069324	113.04940	113.029476	0.970588	0.990186	0.003704	-0.002548
2018-10-31 08:58:00	113.069	113.066	113.069	113.063	9	113.053585	113.060977	113.069674	113.04950	113.029326	0.882353	0.950980	0.004405	-0.000757
2018-10-31 08:59:00	113.078	113.069	113.078	113.069	13	113.054739	113.063295	113.072085	113.05025	113.028415	1.000000	0.950980	0.005544	0.000503

図 4: 4 入力テーブル (テクニカル指標)

rule	rule	rule	rule	rule
0	0	0	0	0
0	0	0	0	0
0	0	0	0	0
0	0	0	0	0
0	0	0	0	0

図 5: 5 入力テーブル (ルール)

LSTM の実装は、まだ途中であるため予測はできていない。

3 おわりに

今後の展望としては、LSTM の実装やリアルタイムへの適応、ほかのテクニカル指標への考慮などが考えられる。