

はじめに  
提案手法  
実験  
実験結果  
分析  
おわりに

# ソーシャルメディアにおける影響関係から 金融市場に対する作用のモデル化と分析

木下大輔

富山県立大学 電子情報工学科

December 4, 2020

## 背景

ソーシャルメディア上の情報は社会を反映しているといえ、そこから投資家が何を認知し、どのような感情を抱いているかを反映するセンサとし金融市場の観測や予測をより精密に行うことが可能になっている。また近年ソーシャルメディア内での相互作用が投資家の心理状態の変化を引き起こし、金融市場の価格変動の決定要因の一つになっている。

## 目的

- ソーシャルメディアにおける影響関係から金融市場に対する作用をモデル化し、暗号資産をはじめとする金融資産の価格変動をより精緻にかつ安定した予測をする
- ソーシャルメディアにおける影響関係から金融市場に対する作用を定式化し、end-to-end のニューラルネットワークである Return Neural Network(RNN) を用いてモデル化する手法を提案する

モデル化するソーシャルメディアの影響関係から金融市場への作用を以下のステップに定式化する。

## ステップ

### ① 観測と投稿

ユーザが金融市場に関する事象を観測し、事象に関する投稿をソーシャルメディアに行う。

### ② 影響

ユーザが他ユーザの事象に関する投稿を閲覧することにより、事象に関する認知、心理状態に影響を受け、場合によっては事象に関する発信を行い情報が拡散する。

### ③ 認知と心理変化

情報が拡散されることにより様々なユーザの事象に関する認知高めたり、心理状態が変化する。

### ④ 行動

ソーシャルメディアから認知や心理状態に対して影響を受けたユーザが実際に事象に関する投資行動を行い、金融市場が変動する。

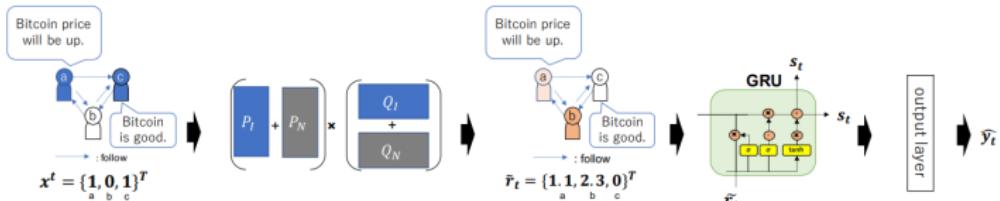
# end-to-end のネットワークを用いたモデル化

4/15

定式化したソーシャルメディアの影響関係から金融市場に対する作用を、end-to-end のニューラルネットワークである RNN を用いてモデル化する手法を説明する。

## 要素

- ユーザ:  $u_n$
- ユーザごとの特定金融商品に関する投稿の有無を表す one-hot ベクトル:  $x^t$
- ユーザ  $u_j$  からユーザ  $u_i$  に対する影響度:  $e_j, i$
- 影響関係行列:  $E$
- 時刻:  $t$
- ユーザが受けける影響ベクトル:  $r_t$
- ユーザの状態:  $s_t$
- 金融市場の状態:  $h_t$



a) ユーザごとの特定金融商品に関する投稿の有無を表す one-hot ベクトル  $x^t$  を得る。

b)  $x^t$  と影響関係行列  $E$  との行列積を計算。影響関係行列  $E$  は、学習パラメタ行列  $I$  と隣接行列  $N$  の和として定義し、計算においてそれぞれの行列にPMFを適用した行列  $(P_I, Q_I, P_N, Q_N)$  を用いて近似計算する。

c) 時刻  $t$  にユーザが受けける影響ベクトル  $\tilde{r}_t$  を獲得。

d) GRUを用いて、ソーシャルメディアにおけるユーザの状態  $s_t$  を獲得。

e) 金融市場の状態を表す  $h_t$  やソーシャルメディアにおけるユーザの状態を表す  $s_t$  を全結合層の入力として、予測を出力する。

図 1: end-to-end のネットワークを用いたモデルの概要

既存手法では直接観測されないユーザの影響関係や影響度は直接観測できないので、フォロワー数やネットワーク指標で代用するが、今回のように時間を通したユーザの影響関係を取り扱う際に、ネットワーク指標のみを影響関係指標とする際は以下の課題がある。

## 課題

- (1) ユーザ同士の信用度、または金融市場に影響を与え合うほどの相互作用をもたらすほどの関係を持たないときの相互作用の過大評価
- (2) ユーザ同士がフォローしていないく直接閲覧をしない場合でも、投稿する内容が興味関心に類似する場合、他のユーザを介して時間を通して影響関係にある可能性がある

対象とするユーザ  $M$  人のうち、ユーザ  $u_i$  がユーザ  $u_j$  の投稿を閲覧することをソーシャルメディアのフォロー関係を表す隣接行列  $\mathbf{N}$  によって表現し、ユーザ  $u_i$  がユーザ  $u_j$  の投稿を閲覧したときに受ける影響度を、学習パラメータ行列  $\mathbf{I}$  によって表す。そして、隣接行列  $\mathbf{N}$  と学習パラメータ行列  $\mathbf{I}$  の和によって、影響関係行列  $\mathbf{E} \in \mathbf{R}^{M \times M}$  を定義する。

## 影響関係行列の定義

$$\mathbf{E} = \mathbf{N} + \mathbf{I} \quad (1)$$

ある時刻  $t$ において、 $K$  人のユーザが対象としている金融資産に関連する投稿をしたかどうかを表した、one-hot ベクトル  $\mathbf{x}_t \in \mathbf{R}^M$  と影響関係行列  $\mathbf{E}$  の影響行列  $\mathbf{E}$  の行列積によって、ある時刻  $t$  にユーザが受ける影響ベクトル  $\mathbf{r}_t \in \mathbf{R}^M$  を獲得する。

## 影響ベクトル

$$\mathbf{r}_t = \mathbf{E} \mathbf{x}_t \quad (2)$$

隣接行列  $\mathbf{N}$ , 行列  $\mathbf{P}_N \in \mathbf{R}^{M \times K}$  と行列  $\mathbf{Q}_N \in \mathbf{R}^{K \times M}$  に分解し、 $\mathbf{I}$  に基づく学習パラメタ行列  $\mathbf{P}_I \in \mathbf{R}^{M \times K}$  と学習パラメタ行列  $\mathbf{Q}_I \in \mathbf{R}^{K \times M}$  を定義することで式 (1) の影響関係行列  $\mathbf{E}$  を以下のように近似する。

## 影響関係行列の近似

$$\mathbf{E} \sim (\mathbf{P}_I + \mathbf{P}_N)(\mathbf{Q}_I + \mathbf{Q}_N)$$

よって、式 (2) は以下のように近似され表される。

## 影響ベクトルの近似

$$\mathbf{r}_t \simeq (\mathbf{P}_I + \mathbf{P}_N)(\mathbf{Q}_I + \mathbf{Q}_N)\mathbf{x}_t \quad (3)$$

今回の実験では、end-to-end のアルゴリズムの時系列モデルを取り扱う RNN の一種である GRU を用いる。最初に式 (3) で獲得された影響ベクトル  $r_t$  に対して次の処理を行い埋め込みベクトル  $\tilde{r}_t \in \mathbf{R}^{hidden}$  を得る。

### 埋め込みベクトル

- 学習時に用いる Dropout 層 :  $dropout$
- 対象とするソーシャルメディアのユーザ数 :  $M$
- 学習パラメタ:  $W_{in}, b_{in}$

$$\tilde{r}_t = dropout(\tanh(W_{in} \frac{r_t}{\sqrt{M}} + b_{in}))$$

GRU を用いてユーザの認知・心理状態を獲得する。

### ユーザの認知・心理状態

- ユーザの認知・心理状態 :  $s_t$

$$s_t = GRU(\tilde{r}_i) \tag{4}$$

はじめに

提案手法

実験

実験結果

分析

おわりに

獲得した  $s_t$  を用いて、金融市場の状態を表す特徴量  $h_t$  を獲得する。

## 金融市場の状態を表す特徴量

- 金融市場の価格変動を表す特徴量 :  $m_t$
- 金融市場の状態を表す特徴量 :  $h_t$

$$h_t = GRU(m_t) \quad (5)$$

ソーシャルメディアが金融市場の予測にどれだけ寄与しているのかの可能性を高めるため、 $s_t$  および  $h_t$  を一次元の値に変換し、予測を行う。

## 予測向上

- $s_t$  と  $h_t$  の連結ベクトル:  $concat(s_t, h_t)$
- 学習パラメタ:  $\mathbf{W}_{s-out}, b_{s-out}, \mathbf{W}_{h-out}, b_{h-out}, \mathbf{W}_{out}, b_{out}$
- 時刻  $t$  に出力される予測値:  $\hat{y}_t$

$$s_t = \tanh(\mathbf{W}_{s-out}s_t + b_{s-out})$$

$$h_t = \tanh(\mathbf{W}_{h-out}h_t + b_{h-out})$$

$$\hat{y}_t = \tanh(\mathbf{W}_{out}concat(s_t, h_t) + b_{out}) \quad (6)$$

## 実験対象

今回の研究ではツイッターのデータを用いて暗号資産の将来の価格変動のボラティリティを予測する実験を行い、提案手法の学習モデルの精度を評価する。対象とする暗号資産はビットコイン、イーサリアム、リップルの暗号資産の市場で時価総額が大きい3つの暗号資産について、将来のボラティリティの予測を行う。

## - タセット

暗号資産のデータは、2017年1月1日から2017年12月31日までの1年間の価格の日足データを用いる。1タイムステップの $\Delta t = 1$ は1日とする。また以下の要素を計算し利用する。

- (1)  $ChangePrice_{\Delta t}$ : 時刻  $t$  から  $\Delta t = 1, 3, 5, 10, 25$  日前からの終値変化率
- (2)  $DeviationRate_{\Delta t}$ : 時刻  $t$  における  $\Delta t = 3, 5, 10, 25$  日前からの終値移動平均からの乖離率
- (3)  $Volatility_{\Delta t}$ : 時刻  $t$  における  $\Delta t = 3, 5, 10, 25$  日前からの終値移動平均からの乖ボラティリティ(過去  $\Delta t$  日間の、1日前からの終値変化率  $ChangePrice_1$  の標準偏差)

# 実験設定 2

11/15

## 評価方法

評価方法として暗号資産市場の将来のボラティリティの予測においては、RMSE を用いる。 $\hat{y}$  を予測値  $y$  を実際の値とする  $RMSE$  は以下の式で与えられる。

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (\hat{y}_i - y_i)^2}$$

## 比較モデル

評価の正当性を保つために、比較手法は提案手法のモデルをベースに既存研究との差分の構造を変えた以下の 3 つのモデルとする。

- (1) *Market*: 提案手法のモデルにおいてソーシャルメディアから得られる情報を用いた金融市場のデータのみを扱う時系列モデル
- (2) *SocialSensor*: 提案手法のモデルにおいて式 (4) にソーシャルメディアにおける投稿に関する統計量を用いた、ソーシャルセンサの時系列モデル
- (3) *LTM*: 提案手法のモデルにおいて式 (4) にソーシャルメディアの影響関係として LTM を利用した時系列モデル

	Market	Social Sensor	LTM	Proposed Model
金融市場データの利用	○	○	○	○
ソーシャルメディアの投稿に関する統計量の利用	×	○	○	○
ネットワークにおける閲覧を伴う影響関係の利用	×	×	○	○
金融資産に対する時間を通じた影響関係と心理状態の利用	×	×	×	○

図 2: 提案手法と比較手法のデータ利用・構造の違い

## 結果

- ① ほとんどの比較検証実験の場合において、提案手法が他の手法よりも精緻に将来のボラティリティの予測を行えている
- ② イーサリアム、リップルの場合において、予測時点から 1 日後のボラティリティの予測では提案手法は *Market, SocialSensor* の手法を上回ることが出来なかった
- ③ ソーシャルメディアの情報を用いない手法は、用いるモデルと比べて安定的に優れた予測を出すことが出来ていない

	1日後	3日後	5日後	10日後
Market	0.0230	0.0309	0.0323	0.0420
Social Sensor	0.0394	0.0416	0.0307	0.0459
LTM	0.0317	0.0323	0.0392	0.0431
Proposed model	<b>0.0215</b>	<b>0.0220</b>	<b>0.0267</b>	<b>0.0414</b>

	1日後	3日後	5日後	10日後
Market	<b>0.0146</b>	0.0200	0.0245	0.0269
Social Sensor	0.0176	0.0218	0.0244	0.0239
LTM	0.0160	0.0196	0.0253	0.0243
Proposed model	0.0184	<b>0.0184</b>	<b>0.0226</b>	<b>0.0237</b>

	1日後	3日後	5日後	10日後
Market	0.0654	0.0813	0.0931	0.1042
Social Sensor	<b>0.0572</b>	0.0730	0.0957	0.0993
LTM	0.0617	0.0728	0.0900	0.0994
Proposed model	0.0627	<b>0.0655</b>	<b>0.0790</b>	<b>0.0938</b>

図 3: ビットコインの比較  
結果

図 4: イーサリアムの比  
較結果

図 5: リップルの比較結  
果

## 考察

- ソーシャルメディアのネットワークにおける直接的な関係がなくとも、お互いの興味関心と情報発信の内容が類似しているユーザは、時間を通して影響関係になる可能性を示唆している。

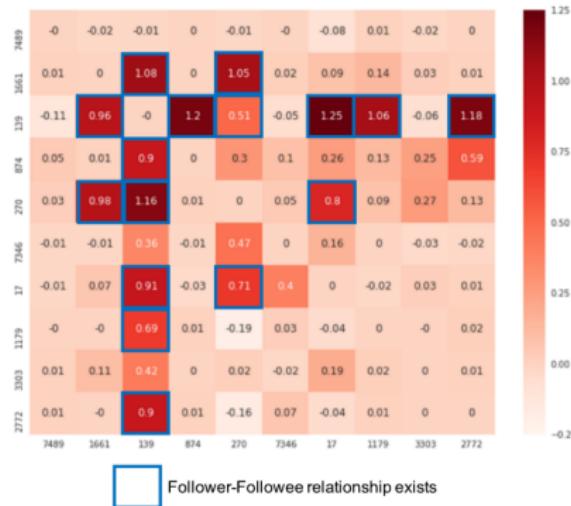


図 6: ビットコインの影響関係を表す行列  $E$  のおいて、サンプリングされた 10 アカウントの影響関係のヒートマップ

## 考察

- プレスリリースなどの1つのニュースでも価格変動が比較的簡易に起こる金融資産は、ソーシャルメディアのユーザで個人投資家が好む金融資産であるので、それぞれの性質によってソーシャルメディアから金融市場への影響の強さが異なることが示唆される。

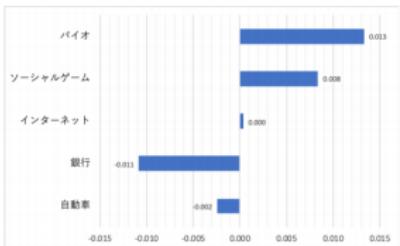


図 7: 金融市場における価格変動を表す変数への重みに対して、ソーシャルメディアにおける事象を表す変数への重みの差分

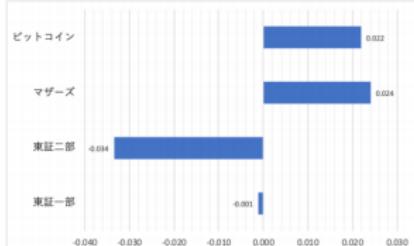


図 8: 金融市場における価格変動を表す変数への重みに対して、ソーシャルメディアにおける事象を表す変数への重みの差分

## 結果

- ① ソーシャルメディアにおける影響関係から金融市場への作用をモデル化する手法が、既存手法に比べて暗号資産の価格変動をより精緻にかつ安定して予測できることを示した。
- ② ソーシャルメディアのネットワークが直接接続されていなくとも、時間を通して影響関係になる可能性がある。
- ③ 金融資産・市場の性質によってソーシャルメディアから金融市場への影響度が異なることを示した。