

そして、1日あたりのスコア S_i を以下に示す

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

N :1日あたりのツイート数

また、

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

M 日間ツイートを取得した場合のデータセットは、

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

ここで、 z は為替の変化状態を表す

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。

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

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

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

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

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、 $Q(\mu_{(z)}^{(c)}, \Lambda_{(z)}^{(c)} | \mathbf{X}^{(c)})$ は事後分布である。

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

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

z_1^* を投資判断に活用する

5 為替予測

抽出したツイートから辞書を用いた求めたスコアを入力として為替を予測する。従来研究では1日単位での長期予測を行っていた。

長期予測では多くのツイートを対象として分析できるため、予測を行いやすい。本研究では、Twitter を用いているためリアルタイム性を活用して短期予測を行えばいいと考えている。しかし、トランプ大統領など多数のユーザを対象とするならば1日単位で考慮する必要がある

実際に、Fig. 5 のような流れで為替予測を行う。

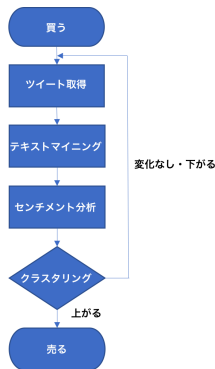


Fig. 5: 自動売買の流れ

参考文献

- [1] 三井健史, 伊藤智也, 中西勇人, 濱川礼, SNS の投稿を用いた感情記録ライフログシステム, 情報処理学会研究報告, Vol. 2014-EC-32 No. 1, pp.1-6, 2014.6.6.