

Title	【課題研究報告書】ソーシャルメディア感情推定と機械学習を用いた仮想通貨の価格変動予測
Author(s)	Zhou, You
Citation	
Issue Date	2023-03
Type	Thesis or Dissertation
Text version	author
URL	<a href="http://hdl.handle.net/10119/18363">http://hdl.handle.net/10119/18363</a>
Rights	
Description	Supervisor : 岡田 将吾, 先端科学技術研究科, 修士(情報科学)

## 概要

注目度が高まっている仮想通貨市場は、株式市場より価格の変化が激しく、短期の価格予測は課題である。一般に、短期的な値動きの予測は、株式や FX と同様に、市場指標のみに基づいて、各指標の組み合わせで判断されることが多い。しかし、市場の外には、常に値動きを左右する様々な要因が存在する。ソーシャルメディアもその要因の一つである。

先行研究により、仮想通貨の代表のビットコイン (BTC) の値動きは、Twitter, Reddit などのソーシャルメディア上に投稿されたテキスト内容と関連度が高いことが分かっている。先行研究では関連 Tweet の感情推定結果や仮想通貨の市場データを特徴量として機械学習を用い価格を予測する手法が提案されている。その実験結果によると、感情推定とマーケット情報を使用する特徴量は、マーケット情報のみを使用する特徴量と比べて、一部のモデルではより良い精度を得られた。

しかし、当日のデータのみから特徴量を抽出しており、市場動向が時系列で変化することを考慮していないため、予測精度はまだ十分ではない。また、全ての Tweet を同じ影響度で扱っているため、真実の感情推定結果は反映されていない。

本研究では、仮想通貨 BTC, ETH, XRP, 及び LTC に対して、2020/1/1~2020/12/31 の Tweet とマーケット情報を取得して研究対象とする。VADER を用いて Tweet ごとの感情推定スコアを算出する。本研究では、感情推定スコアに対して、Follower 数, Reply 数, Retweet 数, Like 数, 及び当日の Tweet 取得総数などを考慮することで各 Tweet の重みを算出する方法や移動平均線及び傾きの導入を提案する。

Follower 数, Reply 数, Retweet 数, Like 数を考慮することで感情推定精度を向上させる。本研究の実験では重みを 7 種を使用した。感情推定結果とマーケット情報の移動平均線及び傾きで、時系列の特性を特徴量に反映させる。本研究の実験では感情推定スコア及びマーケット情報に対して、それぞれの 3 日, 5 日, 7 日の単純移動平均線, 指数平滑移動平均線及び移動平均線の傾きを使用した。この 2 点の組み合わせた特徴量を抽出し、機械学習モデル MLP, SVM, RF で実験し、一番良い組み合わせ (感情推定の重みと移動平均線の種類) を抽出した。そして、この組み合わせを使ったモデルの予測結果を分析し、本研究で提案した方法の有効性が明らかになった。

最後に、本研究での予測結果と先行研究の予測結果を比較した。本研究のデータ数及び期間は、先行研究とは違うため、定量的に比較できず、各感情推定導入前後の予測精度の差異などを比較した。本研究で提案した方法は、先行研究より多くの場合で有効であることが明らかになった。