

テクニカル指標と株価変動の関連性に関する実証分析

久保村 光汰 (20X4026) 指導教員 劉 慶豊

1. 序論

近年、日本政府は「貯蓄から投資へ」を推進し、新NISAやiDeCoなどの税制優遇措置を整備している。金融庁[1]による「資産所得倍増プラン」では、現預金を投資に回し、企業価値向上と資産所得拡大を通じた「成長と分配の好循環」を目指す方針が示された。

また、「老後2000万円問題」により、年金に頼らず資産形成を進める必要性が広く認識されるようになった[2]。そのような状況の中で、株へ投資して、効率的に資産形成することが重要な課題となっている。投資者やトレーダーが効率的な資産形成を行うためには、適切な投資判断が不可欠である。そこで本研究は、テクニカル指標を活用した株価予測モデルを構築し、業種ごとのテクニカル指標の有効性を検証し、各指標の重要度の違いを明らかにすることで、投資者やトレーダーが業種特性に応じた戦略を立てるために情報を提供し、効率的な資産形成の実現に貢献することを目指す。

2. データについて

2.1 データの収集方法

本研究では、株価データの収集にyfinanceライブラリを使用する。yfinanceは、無料で広範な金融データを取得可能なPythonライブラリである。本研究では企業ごとに過去1か月分の1時間足データを取得し、株価予測に活用する。

また、収集したデータから特徴量を作成するためにTechnical Analysis Library (ta) を用いて、テクニカル指標を計算し、株価予測モデルの入力データとして利用する。

2.2 データの詳細と分析対象について

本研究では、テクニカル指標として移動平均線 (SMA10, SMA20), ボリンジャーバンド (BB_Upper, BB_Lower), MACD, RSI, MACD Signalを使用する。

本研究では短期間ではノイズ、長期間ではファンダメンタル要素の影響を考慮する。サンプル期間1ヶ月となっている。

そして上記のデータと機械学習の手法を用いて各モデルの株価予測を行った。

分析対象は、業種ごとに選定した流動性の高い日本企業で、直近で不祥事やデータ不足がない企業を選んだ。業種は消費財、金融、テクノロジーの3分類となっている。4時間後と24時間後の株価変動を予測する。

3. 予測に用いた手法

本研究では以下の3つの手法を利用する。

3.1 アンサンブル学習

アンサンブル学習は、複数のモデルを組み合わせることで予測精度を向上させる手法である。主な方法としてバギング、ブースティング、スタッキング、ブレンディング、投票法がある。バギングは過学習を防ぎ、ブースティングは精度を段階的に高め、スタッキングとブレンディングは異なるモデルを組み合わせることで精度を向上させる。投票法は複数モデルの予測結果を集め、最も多く投票された予測を最終結果とする。今回は投票法のハード投票を使用する。

3.2 XGBoost

XGBoostは、決定木を用いた教師あり学習アルゴリズムで、大規模データに対応し、高精度な予測が可能である。先の決定木の予測残差を基に次の決定木を構築し、予測精度を高める。正則化や並列学習を活用し、過学習を抑えつつ効率的に処理を行う。

3.3 ランダムフォレスト

ランダムフォレストは、複数の決定木を組み合わせることで予測精度を向上させるアンサンブル学習の手法で、過学習に強く、高い汎化能力を持つ。非線形な関係や複雑なデータにも対応でき、欠損値があっても安定した予測が可能。計算資源を多く消費し、解釈性が低い点がデメリットであるが、医療や金融などで広く利用されている。

4. 実装

4.1 実装の設定

ターゲット変数は、4時間後または24時間後の株価の終値が現在の終値より高いかどうかで1または0に設定する。前処理として、テクニカル指標の計算、ターゲット変数の生成、

データの分割を行う。欠損値がある行を削除する。最後に、訓練データとテストデータを分割し、時系列の順序を保ちながら80%を訓練データ、20%をテストデータとして使用する。また、モデル内のパラメータと評価基準は表1で確認できる。

表1：各モデルのパラメーターと評価基準

手法	木の数	最大の深さ	評価基準
アンサンブル学習	ランダムフォレスト: 100 XGBoost: 100	6 (XGBoost) 10 (ランダムフォレスト)	正解率 (Accuracy)
XGBoost		100	6 ログ損失 (Logloss)
ランダムフォレスト		100	10 正解率 (Accuracy)

4.2 分析結果と考察

各モデルは訓練データで学習後、テストデータの予測精度で訓練済みモデルを評価する。予測精度の結果は表2のようになる。

表2：各モデルの株価予測精度

業種	モデル	24時間後	4時間後
Technology	ランダムフォレスト	0.9697	0.6456
	XGBoost	0.8561	0.7182
	アンサンブル学習	0.9848	0.6909
Consumer Goods	ランダムフォレスト	0.7636	0.5379
	XGBoost	0.7909	0.5833
	アンサンブル学習	0.8727	0.5833
Finance	ランダムフォレスト	0.5379	0.447
	XGBoost	0.4167	0.5985
	アンサンブル学習	0.6061	0.4242

24時間後の予測精度が4時間後より高いのは、[3]の効率的市場仮説 (EMH) によれば、短期的な価格変動はノイズが多く、予測が難しい。一方では長期的な市場がノイズや突発的な変動の影響を受けにくく、長期的なトレンドや市場反応を捉えやすいためである。

今回は予測精度の高い方をテクニカル指標と株価の関連度が高く、テクニカル指標の重要度比較の意義があると考え、24時間後の株価予測での特徴量の比較を図1に示した。

本研究の結果として、業界ごとに株価予測における重要なテクニカル指標が異なることを確認できた。

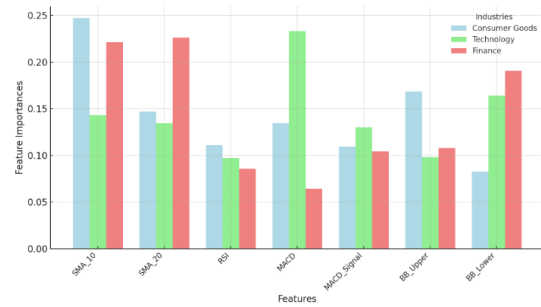


図1：業種毎での特徴量重要度の比較図

テクノロジー業界ではMACDが最も重要であり、急激な価格変動を捉えるために有効であることが示された。またBB_Lowerも反転ポイントの予測に重要な役割を果たす。

金融業界では、安定した市場環境において、SMA_10およびSMA_20が最も高い重要度を示し、BB_Lowerが反転のサインとして機能することが分かった。

消費財業界ではSMA_10が短期的な価格トレンドに有効であり、BB_UpperやMACDも価格の急激な上昇や反転を捉えるために重要な指標であることが分かった。

5. 結論

本研究は機械学習の手法を利用して、テクニカル指標を用いた短期的な株価予測の有効性の確認と業種別における、テクニカル指標の重要度の特徴を捉えることができた。

これにより業種ごとに重要なテクニカル指標が明らかになったため、それぞれの業種に最適化された投資戦略の構築に貢献できると考えられる。

今後の研究では、業種ごとのパラメーターチューニングを行い、特定業種に特化したアプローチで予測精度を向上させることを目指す。また、テクニカル指標の役割を活用し、実際の取引戦略の構築やモデルの解釈性向上を進めることで、実務への応用可能性が高まる。

参考文献

- [1] 藤丸敏(2023), 資産所得倍増プランについて、試算所得倍増プラン始動〜金融ビジネスの今後〜,QUICH 資産運用討論会.
- [2] 金融庁(2019), 高齢社会における資産形成・管理, 金融審議会市場ワーキング・グループ報告書.
- [3] Fama, E. F. (1970). Efficient capital markets. *Journal of finance*, 25(2), 383-417.