

機械学習を用いた為替取引における適した時間粒度の分析

加藤 大樹

李 嘉誠

能登 正人

神奈川県大学大学院工学研究科工学専攻電気電子情報工学領域

1 はじめに

近年、市場価格予測に関する研究は日々進歩を遂げている。よりよい成果を得るために機械学習を用いた研究も行われており、学習モデルに入力するための様々な指標や特定の問題に特化した機械学習モデルが開発されている [1]。予測に利用される情報は基本となる 4 本値や出来高の他にも、投資家が利用するテクニカル指標やニュース記事などを用いたものまで多数存在する。機械学習モデルを予測に用いる際、学習やテストに用いるデータの質や状態により予測結果は変化する。多くの研究では 1 日ごとのデータを使用しているが、データの時間粒度（どれほどの期間の値動きをまとめるか）と予測結果との関係性に着目した研究は非常に少ない。データの粒度が異なる場合、データ量や価格変動のトレンドに明確な違いが出現する。

本研究では、時系列データの異なる時間粒度による予測結果の違いについて考察する。まず機械学習ベースの予測モデルを用いて価格を予測し、その後簡易的な取引のシミュレーションを行うことで異なる時間粒度間の結果に差があるかを検証し報告する。

2 関連研究

時系列予測を行う際に時間粒度に注目した研究も少ないが存在する。Shin ら [2] は、分、時間、日のそれぞれの尺度を用いたアンサンブルモデルを提案している。ここでは、作成したアンサンブルモデルがよい結果を示したことで、そして利用する粒度によって精度が異なることが示されている。Atkins ら [3] は異なる粒度を利用し、価格の上昇下降という二値分類に対する予測精度について議論し、予測モデルの有用性を示している。しかしながら、これらの研究でも検証に使われている時間粒度はわずかであり、多くのサンプルを対象とし深く掘り下げた研究は行われていない。

Analysis of Suitable Time Interval for Exchange Trading Using Machine Learning

Daiki Kato, Jiacheng Li and Masato Noto

Field of Electrical, Electronics and Information Engineering,
Course of Engineering, Graduate School of Engineering, Kanagawa University

3 実験

本研究では、機械学習モデルを用いて価格予測を行う。予測した結果を元にした取引シミュレーションより損益を計算し、得られた収益に対して評価を行う。

3.1 データ

実験に使用するデータは、為替通貨ペアよりドル/円を選択する。2015 年 1 月から 2020 年 12 月までの期間を使用し、このデータを時間粒度ごとに計 12 種類に分割する。本研究では、1, 2, 5, 10, 15, 30 分, 1, 2, 4, 6, 12, 24 時間を対象とする。全期間の傾向を掴むため、ランダムな時点の連続する 1000 ステップを取得し、うち 9 割を学習データ、1 割をテストデータとして使用する。機械学習モデルに入力する際に行う前処理としては、原時系列に対して対数差分変換を用いる。

3.2 価格予測

機械学習モデルとして広く使われている LSTM を用いて価格の予測を行う。入力は終値のみの場合と、4 本値（始値、終値、高値、安値という価格の変動を代表する 4 つの値）の場合の 2 通りとし、次ステップの終値の予測を繰り返す。同一のモデルを用いて各粒度に対して 200 回の予測を行い平均を求め、その粒度の結果とする。予測結果の評価は、真値と予測値との RMSE を用いて行う。

3.3 取引シミュレーション

ここでは、簡易的な取引モデルを作成し、得られた損益に対し評価を行う。各地点での実際の価格と 1 ステップ先の予測値を基にし、次に価格が上がると予測すれば買い、価格が下がると予測すれば売りを行う。この処理を前の機械学習モデルで予測した全ての予測値に適用する。開始時点での所有資産を 160,000 円、投資資産をそのうちの 150,000 円に設定する。資産が減少した際も投資金額を抑えず運用するために、初期の所有資産を取引資産より 10,000 円多く設定している。取引は投資資産を n 回に分けた金額で行い、投資回数を分けることによる効果を確認する。今回は $n = 1, 3, 5, 10$ を使用する。取引終了時点でドルを保有していた場合、

終了時点でのドル/円価格を基準に全て円に換算し評価を行う。それぞれの n ごとに、投資資産額に対する最終資産を計算し売買益の利益を求める。

4 結果と考察

まず、各粒度の RMSE を比較する。図 1 に、真値と予測値より算出された RMSE を示す。

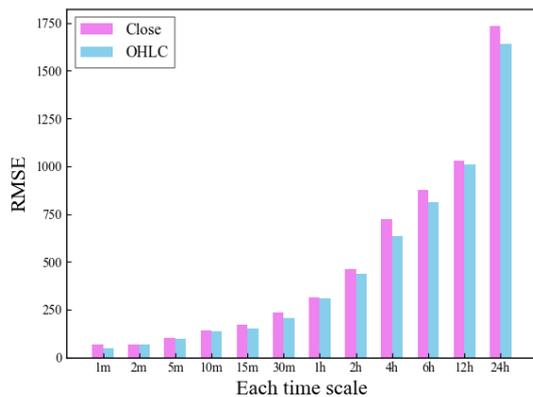


図 1: 各時間粒度ごとの RMSE の比較

ここで、Close は終値、OHLC は 4 本値を表している。図 1 より、時間粒度が大きくなるにつれ予測値と真値との RMSE も大きくなる事が確認できた。これは粒度が大きいほど 1 ステップ間の価格変動が大きくなるためであると考えられる。本研究では、各粒度ごとの予測のしやすさを評価する。RMSE から分かるように、単純な利益の比較では大小からなる価格変動の影響を受けるため、厳密な比較ができない。そこで、価格変動の影響を抑えるために式 (1) を用いてスケールの調整を行い、改めて比較する。

$$\text{Profitability}_i = \text{Returns}_i \times \frac{\text{RMSE}_{12}}{\text{RMSE}_i} \quad (1)$$

ここで $i (1 \leq i \leq 12)$ は、 i 番目の時間粒度であることを表す。投資量に注目すると、 n が大きいほどリスクが分散され損益は小さく収まっていたため、各粒度の $n = 1$ 同士を比較する。取引シミュレーションより得られた利益に対して補正を行ったものを図 2 に示す。

図 2 より、今回用意した粒度では、小さいものと大きいものでの資産変動が大きく、真ん中付近では小さくなった。これより、二値分類の精度が高いほど損益の変動が大きくなる可能性を示している。図 2 から、特に 4 時間、12 時間では利益が多く、6 時間、24 時間では損失が大きくなっていることが読み取れる。比較的大きい粒度のほうが、損益も安定しており利益が出や

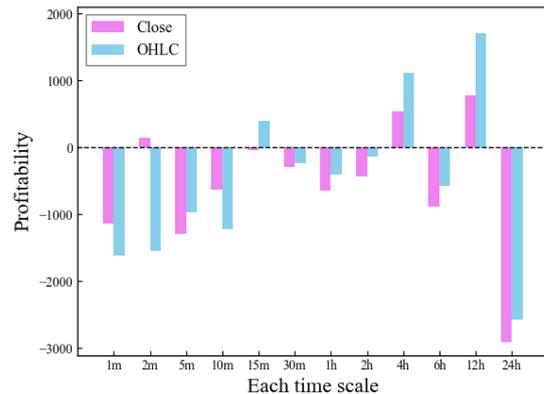


図 2: RMSE により補正した取引による利益

すい傾向がある一方、多くの人が 1 日尺を用いて取引を行っているために利益が伸びなかった可能性がある。これは 1 日ごとの取引を行いたい場合であっても 4 時間ごと、12 時間ごとのデータを元に予測、取引を行った方が期待収益が高いことを表しているといえる。

5 おわりに

本研究では、異なる時間粒度のデータを予測に使用することにより取引の収益に有意差が現れるかを確認した。実験の結果、利益が出やすい時間粒度を確認できた。1 日ごとや 1 時間ごとの予測を行いたい場合にも、適切な粒度を活用することでより高い収益を得られることを示している。一方で、本研究で用いた通貨ペアや学習モデルと異なるものを使用することで異なる結果が得られる可能性も存在する。より複雑なルールに基づいた取引を行うことでも改善の余地があるため、より多方面から検証する必要がある。

参考文献

- [1] Sako, K., Mpinda, B. N. and Rodrigues, P. C.: Neural Networks for Financial Time Series Forecasting, *Entropy*, Vol. 24 (2022).
- [2] Shin, M., Mohaisen, D. and Kim, J.: Bitcoin Price Forecasting via Ensemble-based LSTM Deep Learning Networks, *Proceedings of The 35th International Conference on Information Networking*, pp. 603–608 (2021).
- [3] Atkins, A., Niranjana, M. and Gerding, E.: Financial News Predicts Stock Market Volatility Better Than Close Price, *The Journal of Finance and Data Science*, Vol. 4, No. 2, pp. 120–137 (2018).