

# LSTM による時系列予測と株式投資戦略への応用

## Time Series Prediction with LSTM Networks and its Application for Investment Strategy

松本 健<sup>1</sup> 牧本 直樹<sup>1</sup>

Ken Matsumoto<sup>1</sup> and Naoki Makimoto<sup>1</sup>

<sup>1</sup>筑波大学大学院ビジネス科学研究科

<sup>1</sup>Graduate School of Business Sciences, University of Tsukuba

**Abstract:** Researches for financial time series in stock or foreign exchange markets, have been one of traditional themes of financial market analysis. Statistical model approaches such as ARMA and GARCH were mainstream of conventional analysis. However, it is difficult to understand and predict financial time series structures, which are generally characterized by high noise level and low autocorrelation. Meanwhile, researches to capture the structure by artificial intelligence has been increasing in recent years. In particular, Long Short-Term Memory (LSTM), which can capture time series structure, is already widely used in the field of natural language processing and speech recognition. Therefore, in this study, we investigated the model performance in each TOPIX core30 constituent stock by using logistics regression (LOG), random forest (RAF), gradient boosting (GBT), support vector machine (SVM), and LSTM. The performance was evaluated by metrics such as prediction accuracy, F1 measure, AUC, and return. As a result, LSTM showed the best performance in the models. Moreover, we discussed the effectiveness of the stock market neutral strategy by applying the above prediction models. 10-quantile portfolios using the predicted probability outputted by the model, remarks higher accuracy and returns than individual stock trading in all models. Furthermore, LSTM outperformed the others and it is consistent with the result of S&P500 constituent stocks analysis.

## 1. 序論

株式市場や外国為替市場等における金融時系列データを分析対象とする研究は、金融市場分析の伝統的なテーマの1つである。しかしノイズが支配的である金融資産のリターン系列は、一般に自己相関が低いという特徴があり、時系列構造の理解とその予測は難しい。時系列解析は統計モデルによるアプローチが主流であり、金融時系列の実証分析は既に膨大な研究例がある。線形モデルである ARMA やボラティリティ変動モデルである GARCH や確率的ボラティリティモデル [1]、さらにレジームスイッチングモデル [2]等が提案されている。ただし、何れにおいても時系列の構造理解が目的であるため、株式リターンの予測精度は決して高いとは言えない。

一方、人工知能のアプローチで時系列構造を捉える研究は近年増加傾向にある。ロジスティック回帰やランダムフォレスト、勾配ブースティング、サポートベクターマシンといった機械学習モデルやディ

ープニューラルネットワークモデルが時系列解析に応用されつつある。特に時系列データを取り扱うことができるリカレントニューラルネットワーク型の LSTM (Long Short-Term Memory) は、既に自然言語処理や音声認識の分野において、単語列や音波が持つ特徴量を捉えるために有効活用されており、その認識精度の向上に大きく貢献している。

LSTM と他モデルによる金融時系列予測の研究例として、ここでは Fischer and Krauss [3]を取り上げる。S&P500を構成する500銘柄の日次リターンを LSTM、ディープニューラルネットワーク、ランダムフォレスト、ロジスティック回帰のモデルで予測する実験を行っている。構成銘柄毎に、過去の日次リターン240日分から1日先の超過リターンの正負を予測する2値分類モデルを構築している。さらにその分類モデルから算出される予測確率から、分位ポートフォリオを構築して、パフォーマンスを比較している。結果として、正解率や実現リターン、シャープレシオといった指標において、LSTMモデルが他のモデ

ルをアウトパフォームすることを報告している。

このように LSTM の有効性が示されつつある一方で、未だ金融時系列データの実証研究例が十分にあるとは言えず、どの予測モデルが一番有効であるかというコンセンサスは形成されていない状況にある。そこで本研究では、TOPIX core30 構成銘柄の日次時系列データを対象とした LSTM を含む複数のリターン予測モデルを構築し、そのパフォーマンスを評価する。また、予測モデルを活用した株式投資戦略を立案し、その有効性を議論する。

## 2. 分析データ

本研究では 2017 年 10 月時点で編成されていた TOPIX Core30 の構成銘柄を対象とした分析を行う。TOPIX Core30 は、TOPIX ニューインデックスシリーズの 1 つで、東京証券取引所の市場第一部全銘柄の内、時価総額と流動性の特に高い 30 銘柄で構成された株価指数を指す。30 銘柄の日次リターン(2009 年から 2017 年までの約 2000 日分)を分析対象とした。

予測モデルの正解ラベルは各個別銘柄の対ベンチマーク超過リターンの正負  $B_{t+1}^s$  と設定する。個別銘柄  $s$  の株価から算出される対数差リターン系列を

$$R_t^s = \log\left(\frac{P_t^s}{P_{t-1}^s}\right) \quad (1)$$

と定義すると、個別銘柄の対ベンチマーク超過リターンは、個別銘柄とインデックス指標のリターン系列の差で計算できる。インデックス指標を TOPIX 株価指数とすると、1 期先の超過リターンとその正負は、

$$y_{t+1}^s = R_{t+1}^s - R_{t+1}^{TOPIX} \quad (2)$$

$$B_{t+1}^s = \begin{cases} 0 & \text{if } y_{t+1}^s < 0 \\ 1 & \text{if } y_{t+1}^s \geq 0 \end{cases} \quad (3)$$

となる。次に個別銘柄の超過リターンを予測するための特徴ベクトルとしてヒストリカルデータを用意する。過去の時点  $t - m$  との累積日次リターンを

$$CR_{t,m}^s = \log\left(\frac{P_t^s}{P_{t-m}^s}\right) \quad (4)$$

と定義する。このとき個別銘柄と TOPIX 株価指数それぞれの過去 1 年間 (240 日) 以内の累積日次リターン  $CR_{t,m}^s, CR_{t,m}^{TOPIX}$  を特徴ベクトルとした。なお各特徴ベクトルは最大 240 時点分を設定可能であるが、Krauss et. al, [4] の変数重要度を参考に 15 時点として、時点  $m$  を以下のように設定した。

$$m \in \{1, 2, \dots, 9, 10, 20, 40, 60, 120, 240\} \quad (5)$$

## 3. 予測モデル

LSTM モデルと比較検証するために、他の機械学習モデルも併せて分析する。表 1 に利用するモデルとその略称名を記す。今後のモデル名は略称名で統一する。

表 1 予測モデル

略称名	モデル
LOG	正則化ロジスティック回帰
RAF	ランダムフォレスト分類器
GBT	アダブースティング決定木
SVM	サポートベクターマシン分類器
LSTM	LSTM 分類器

過学習の防止と汎化性能向上のために各予測モデルのハイパーパラメータチューニングを行った。データセットをトレーニングデータ、バリデーションデータ、テストデータの 3 種類に分別した。テストデータについては、モデルの最終的な精度評価用として直近の時系列データ 20% を利用する。残りの 80% のデータはランダムシャッフルと層化抽出法によりトレーニングデータ 60% とバリデーションデータ 20% に分割した。ある 1 組のハイパーパラメータ群を所与とした場合にトレーニングデータでモデルのパラメータを学習し、その評価をバリデーションデータで行うことで学習済みモデルを選定した。

なおハイパーパラメータの探索は、グリッドサーチやランダムサーチといった方法論が確立されている。しかしグリッドサーチは、ハイパーパラメータ空間を全探索できるが、ハイパーパラメータの次元数に学習時間が指数関数的に増加するため、ディープラーニングでは現実的な方法論とは言えない。またランダムサーチは現実的な試行回数で探索を打ち切れるが、ハイパーパラメータの大域での最適値が探索できるとは限らない。そこで本研究ではベイズ最適化手法の 1 つである SMBO (Sequential Model-based Global Optimization) のアルゴリズムを採用することとした [5]。

## 4. 個別銘柄のリターン予測

各モデルのパフォーマンスは 30 銘柄毎の予測精度と実現リターンの観点から評価した。予測精度については、予測ラベルをラベル 1 に属する確率  $\hat{P}(B_{t+1}^s = 1)$  と境界値 0.5 より

$$\hat{B}_{t+1}^s = \begin{cases} 0 & \text{if } \hat{P}(B_{t+1}^s = 1) < 0.5 \\ 1 & \text{if } \hat{P}(B_{t+1}^s = 1) \geq 0.5 \end{cases} \quad (6)$$

として正解率、F1 値、AUC を評価尺度とする。また

実現リターンは予測ラベルが 1 のときロングポジション、予測値が 0 のときショートポジションを取った場合のリターン  $\hat{R}_{t+1}^s$  を

$$\hat{R}_{t+1}^s = \hat{B}_{t+1}^s R_{t+1}^s - (1 - \hat{B}_{t+1}^s) R_{t+1}^s \quad (7)$$

として算出する。ここで第 1 項はロングポジションのリターン、第 2 項はショートポジションのリターンを表す。

各評価指標の内、30 銘柄毎の正解率と実現リターンの結果をそれぞれ図 1 と図 2 に示す。30 銘柄の平均正解率は LSTM が最も良い水準にあり、次いで GBT、SVM、RAF が並び、LOG が最も水準が低い結果となった。正解率の標準偏差については GBT と LSTM が低いため、個別銘柄に因らず、安定した予測結果が得られていると言える。また実現リターンにおいても正解率の順位と一致する形で LSTM が最も良いパフォーマンスが得られた。また各銘柄の実現リターンの分散についても低水準となっている。

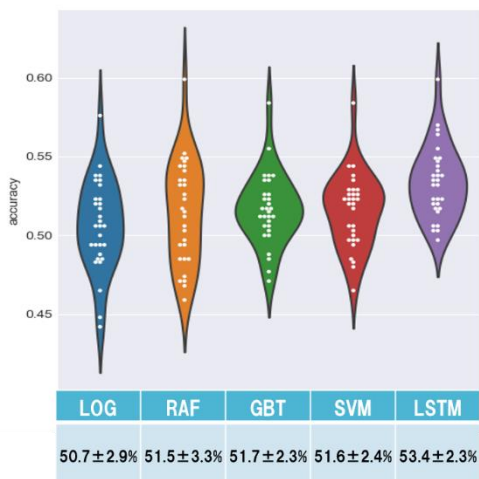


図 1 30 銘柄分の正解率の比較

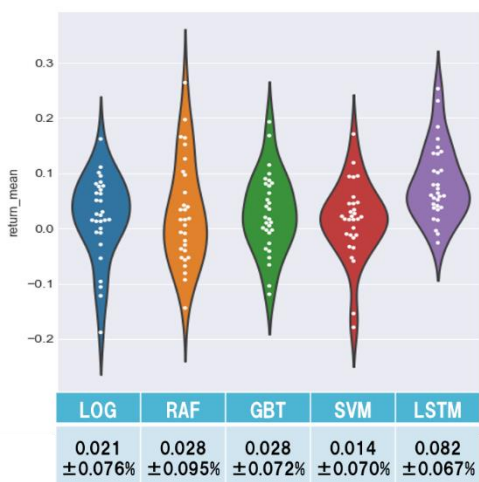


図 2 30 銘柄分のリターンの比較

これらの正解率と実現リターンは、式(6)のように予想確率  $\hat{P}(B_{t+1}^s = 1)$  が、固定された境界値 0.5 を超えるか否かでポジションを取った結果となっている。そのため、例えば予測確率が 51% である場合と 100% である場合を等価に取り扱っており、予測モデルの出力結果を最大限活用できているとは言い難い。そこで次章では、予測確率の大きさを踏まえた投資戦略を検討する。

## 5. 分位ポートフォリオへの応用

本研究で利用する予測モデルは、日次リターン予測かつ超過リターンの正負予測という特性を踏まえると、長期資産運用を目的とした一般的な資産運用より、短期の絶対収益型運用スタイルに親和性が高い。そこで各モデルの予測確率を利用した分位ポートフォリオを考える。各 30 銘柄を予測確率  $\hat{P}(B_{t+1}^s = 1)$  の昇順に並べ直し、この銘柄群から十分位ポートフォリオを構築する。任意の分位数を  $k$  とすると、予測確率が低い第  $k$  十分位までをショートポジションを、予測確率が高い第  $k$  十分位までをロングポジションを取った場合のパフォーマンスを調べる。

例えば  $k = 1$  のとき、第 1 十分位(3 銘柄)をショートポジション、第 10 十分位(3 銘柄)をロングポジションとするため、銘柄群の内 20% が取引対象銘柄となる。  $k = 5$  の場合では、第 1~第 5 十分位までをショートポジション、第 6~第 10 十分位までをロングポジションとなるため、30 銘柄すべてが投資対象となる。個々のロング・ショートポジションはすべてダラーニュートラルとなるように 1 単位ずつ保有するが、個別銘柄の分析結果と比較可能にするため  $1/2k$  倍して単位を揃えた。また本分析では各売買に関わる手数料は考慮しない。

なおこの分位ポートフォリオは、投資戦略の中でも株式ダブルアルファ・ダラーニュートラル戦略に該当する。相対的な価格上昇が見込まれる銘柄のロングポジションと上昇が見込まれない銘柄のショートポジションを保有することで、出来る限り市場の影響を排除して、両ポジションのアルファの獲得を目指す。

各予測モデルの分位ポートフォリオにおいて、 $k$  を変化させた場合の正解率を図 3 に示す。個別銘柄の結果と比較すると、 $k$  に因らずほぼすべての場合において正解率の向上が見られる。分位を取ることでパフォーマンスが向上したという事実は、予想確率を予測の確度とする妥当性を示唆している。

一方で、 $k = 1$  のときでは個別銘柄の結果とは異なり SVM の正解率が一番低い結果が得られた。本研究では SVM の予測確率は Platt's Scaling [6] より求めているが、正解率を向上させるために予測確率への

キャリブレーション方法を再検討する価値があると言える。

次に、 $k$ を変化させた場合の実現リターンの平均値と標準偏差をそれぞれ図4と図5に示す。個別銘柄のリターンと比較すると、実現リターンの水準については特に大きな差異は見られない。しかしどの予測モデルにおいても、投資対象の分散効果が顕著で  $k = 5$  のとき標準偏差が一番小さくなっている。この投資対象の分散効果はリスク調節後リターンに大きく寄与する。このように予測確率を利用した分位ポートフォリオの構築は個別銘柄投資よりもリターンとリスクの意味で効率的であることがわかる。

以上のように、正解率及び実現リターンにおいて、LSTM が他の機械学習モデルをアウトパフォームする結果が得られた。この結果は S&P500 の実証分析 [3] と整合している。

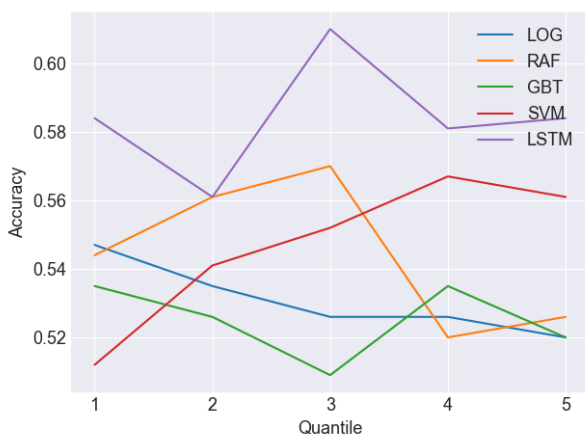


図3 各分位ポートフォリオの正解率の比較

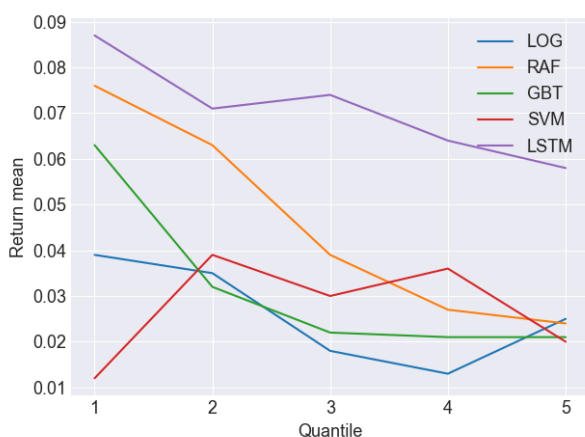


図4 各分位ポートフォリオのリターンの平均

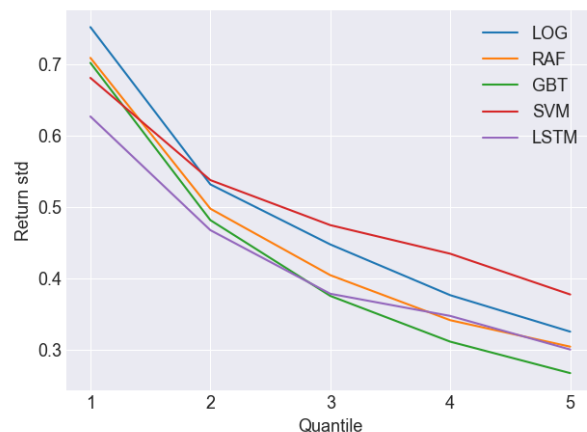


図5 各分位ポートフォリオのリターンの標準偏差

## 6. 結論

TOPIX core30 構成銘柄を対象とした株式リターン予測と投資戦略への応用可能性を議論した。個別銘柄の予測結果では、正解率や実現リターンといった観点で LSTM が最も良いパフォーマンスを示した。またモデルの予測確率を活用した分位ポートフォリオを構築することで、すべてのモデルで正確率向上が見られた。その中でも LSTM は他モデルをアウトパフォームしており、S&P500 構成銘柄を対象とした実証分析結果と整合する結果が得られた。

## 参考文献

- [1] 渡部敏明, ボラティリティ変動モデル, 朝倉書店, 2000.
- [2] A. Ang and G. Bekaert, "International Asset Allocation with Regime Shifts," *The Review of Financial Studies*, vol. 15, no. 4, p. 1137, 2002.
- [3] T. Fischer and C. Krauss, "Deep learning with long short-term memory networks for financial market predictions," *FAU Discussion Papers in Economics*, vol. 11, 2017.
- [4] C. Krauss, X. A. Do and N. Huck, "Deep neural networks, gradient-boosted trees, random forests: Statistical arbitrage on the S&P 500," *European Journal of Operational Research*, vol. 259, no. 2, p. 689, 2017.
- [5] J. Bergstra and B. Kégl, "Algorithms for Hyper-Parameter Optimization," in *NIPS'11 Proceedings of the 24th International Conference on Neural Information Processing Systems*, 2011.
- [6] J. Platt, "Probabilistic Outputs for Support Vector Machines and Comparisons to Regularized Likelihood Methods," MIT press, 1999.