

# Time Series Prediction Based on Machine Learning Techniques

柴田 大輝<sup>1</sup> 藤原 真吾<sup>1</sup> 佐々木 洋<sup>1</sup>

Daiki Shibata<sup>1</sup>, Shingo Fujiwara<sup>1</sup>, and Hiroshi Sasaki<sup>1</sup>

<sup>1</sup>みずほ第一フィナンシャルテクノロジー株式会社  
<sup>1</sup>Mizuho-DL Financial Technology Co., Ltd.

**Abstract:** 近年、機械学習の金融分野における活用が注目されており、関連する研究が多くなされている。一方、機械学習の不安定性、とりわけ乱数による影響に関する懸念も、金融以外の分野では指摘されている。上記状況を踏まえ、本研究では、株式リターン予測に機械学習を用いた場合、どの程度乱数の影響を受けるのか調べた。その結果、乱数を変えるだけで大きく運用シミュレーションのパフォーマンス（シャープレシオ）が変わり、全く異なる結果となることが分かった。すなわち、単純に機械学習を用いることは非常に危険であり、予測への機械学習活用には注意が必要であることが示唆される。一方で、方向適合率は比較的乱数の影響が小さく、すべての性質に対して乱数依存性が大きいわけではないことも分かった。このような比較的安定的な性質をうまく活用した、より安定性のある結果を得る方法の模索や、不安定であっても全体的に高いパフォーマンスを常に示すようなより予測力の高いファクターを見つけることが、株式リターン予測に対する機械学習活用への足掛かりとなるのではないかと考えられる。

## 1. 導入

近年、様々な分野で機械学習が活用されており、金融分野においても機械学習の活用が注目され、研究がされてきた[1-5]。

一方、機械学習の不安定性に対する懸念も示されている。例えば、[6]では、強化学習において乱数シードを変えるだけで全く異なる結果を作りうることを指摘している。また、[7]では、結果の優位性を示すために必要な乱数シードの試行回数を評価し、少なくとも 20 パターン以上を試み、安定性を確認しなければ有意な結果とは言い難いことを示している。

しかし、これまで金融分野においてはこのような機械学習の不安定性に関する検証はほとんど行われていない。多くの論文では、機械学習を用いた分析について、単一の結果を示しているのみで、その安定性に関する検証は行われていない。

## 2. 研究結果

本研究では、上記状況を踏まえ、株式リターン予測という観点から、機械学習の乱数シードに対する

安定性に関する実証実験を行った。

### 2.1 シミュレーション条件

本研究では、機械学習を用いて TOPIX, S&P500, DAX のリターン予測を行い、その安定性に関する検証を行った。Python のライブラリ scikit-learn ([8]) を用い、decision tree, bagging, adaboost, random forest による予測を行った。予測ファクターは、表 1 を用いる。ここで、予測ファクターによる影響を調べるため、各資産 2 パターンの予測ファクターで実験を行う。予測ファクターは前々週の金曜日から前週の金曜日のリターン、ないしは変化を用い、月曜日から次週の月曜日にかけてのリターンを予測する。ローリング幅、およびバックテスト期間は、表 2 の通りである。

ここで、ハイパーパラメータは以下のように決定する(図 1)。まず、各手法のハイパーパラメータを表 3 で定義された範囲から、ランダムで一つずつ選択する。次に、選択されたハイパーパラメータを用い、過去 52 週に渡り、予測対象資産のリターンを予測し、その正答率を計算する。上記計算を、乱数を変えて 50 パターン行い、最も正答率の高かった

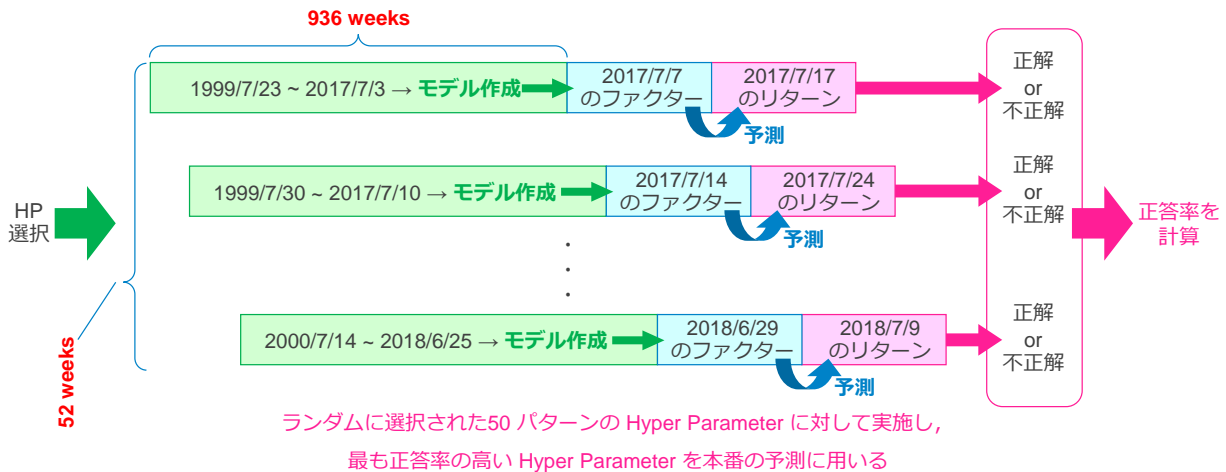


図 1: ハイパーパラメータ (HP) の決定方法 (TOPIX の Pattern 1 の場合. 2018/7/13 基準のファクターから 2018/7/23 のリターンを予測する場合を例として表示). 上記方法を 50 パターンの HP に対して行い、最も正答率の高い HP を予測に用いる.

表 1: 予測に用いるファクター

| TOPIX     |  |
|-----------|--|
| Pattern 1 | <ul style="list-style-type: none"> <li>前週の TOPIX リターン</li> <li>JGB 10 YR - JGB 2 YR 金利差の変化</li> <li>TOPIX 配当利回り変化</li> <li>VIX 変化</li> <li>JGB 3 M 金利変化</li> <li>USDJPY リターン</li> <li>S&amp;P500 リターン</li> </ul> |
| Pattern 2 | Fama & French の 5 ファクター [9]  |

| S&P 500   |   |
|-----------|---|
| Pattern 1 | <ul style="list-style-type: none"> <li>前週の S&amp;P500 リターン</li> <li>UST 10 YR - UST 2 YR 金利差の変化</li> <li>S&amp;P500 配当利回り変化</li> <li>VIX 変化</li> <li>UST 3 M 金利変化</li> <li>EURUSD リターン</li> <li>DAX リターン</li> </ul> |
| Pattern 2 | Fama & French の 5 ファクター [9]   |

| DAX       |  |
|-----------|--|
| Pattern 1 | <ul style="list-style-type: none"> <li>前週の DAX リターン</li> <li>BUNDS 10 YR - BUNDS 2 YR 金利差の変化</li> <li>DAX 配当利回り変化</li> <li>VIX 変化</li> <li>BUNDS 3 M 金利変化</li> <li>EURUSD リターン</li> <li>S&amp;P500 リターン</li> </ul> |
| Pattern 2 | Fama & French の 5 ファクター [9]  |

ハイパーパラメータを本番の予測で使用する. ハイパーパラメータは, 予測のたびに決定しなおす.

このようにして決定されたハイパーパラメータを用い, 予測対象資産のリターンを予測する. この予測が, 株価上昇予測であればロング, 低下予測であればショートとするシミュレーションを行う. シミュレーションは, 乱数シードを変えて 30 パターン

表 2: シミュレーション条件

| 資産     | パターン      | ローリング幅 | バックテスト期間               |
|--------|-----------|--------|------------------------|
| TOPIX  | Pattern 1 | 936 週  | 2014/12/29 ~ 2020/6/29 |
|        | Pattern 2 | 1040 週 | 2014/12/29 ~ 2020/8/31 |
| S&P500 | Pattern 1 | 990 週  | 2014/12/29 ~ 2020/6/29 |
|        | Pattern 2 | 1040 週 | 2014/12/29 ~ 2020/8/31 |
| DAX    | Pattern 1 | 990 週  | 2014/12/29 ~ 2020/6/29 |
|        | Pattern 2 | 1040 週 | 2014/12/29 ~ 2020/8/31 |

表 3: ハイパーパラメータ選択範囲

| 予測手法          | ハイパーパラメータ             | 下限 | 上限 |
|---------------|-----------------------|----|----|
| Decision Tree | 最大深さ                  | 1  | 5  |
|               | Leaf 内の最小数            | 1  | 50 |
| Bagging       | 予測器数                  | 1  | 50 |
| AdaBoost      | 予測器数                  | 1  | 50 |
|               | 各予測器 (決定木) の最大深さ      | 1  | 5  |
| Random Forest | 作成する Decision Tree の数 | 1  | 50 |
|               | 最大深さ                  | 1  | 5  |
|               | Leaf 内の最小数            | 1  | 50 |

行う.

## 2.2 結果

前節の方法で, TOPIX, S&P500, DAX に対してバックテストシミュレーションを行った. ここでは, シミュレーションのパフォーマンスをシャープレシオ (リターン/リスク) で評価する. 30 パターンの乱数シードでシャープレシオを計算し, その違いをまとめた. その結果が, 図 2 である.

この結果を見ると, 資産, 手法に依らず, シャープレシオが大きくばらついていることが分かる. すなわち, シャープレシオでパフォーマンスを評価する場合, そのパフォーマンスは乱数シードによって

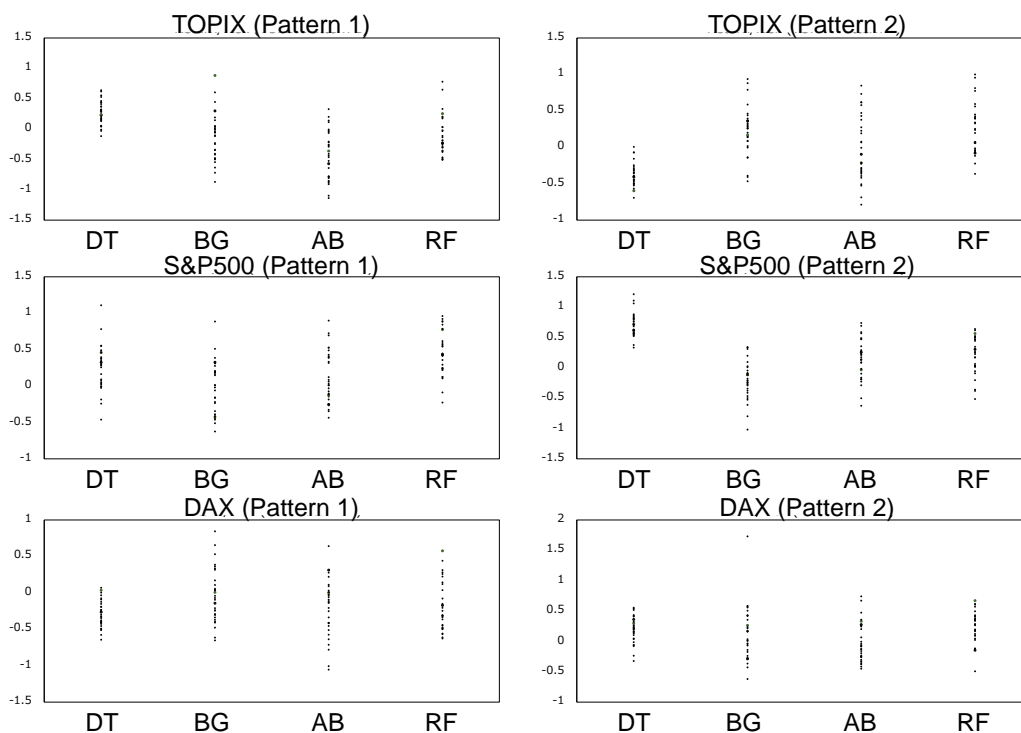


図 2: シャープレシオ. それぞれ, DT は decision tree, BG は bagging, AB は adaboost, RF は random forest の結果を表す.

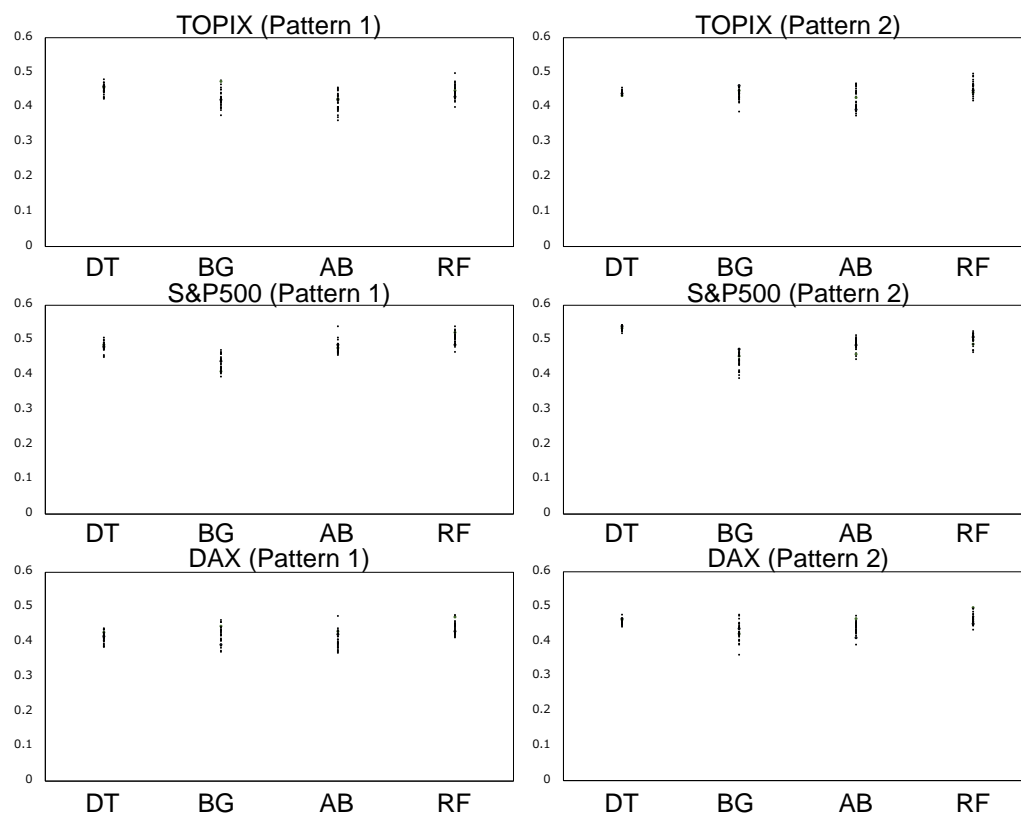


図 3: 方向適合率. それぞれ, DT は decision tree, BG は bagging, AB は adaboost, RF は random forest の結果を表す.

全く異なる結果となることを示している。このことは、乱数シードを変えるだけで異なる結果を生む危険性を示している。

### 2.3 ディスカッション

これまでの結果は、機械学習を用いた運用シミュレーションを行った場合、乱数による影響が大きく、安定した結果が得られにくいということを示している。このことを乗り越え、機械学習を資産運用へと活用するための解決策の一つとして、乱数に依存しない性質を見つけ出し、その性質を運用方法へ応用することが考えられる。

図 3 は、Pattern 1, 2 の TOPIX, S&P500, DAX に対する方向適合率を示したものである。この結果を見ると、方向適合率は機械学習手法に依らず似た値を示すとともに、乱数シードによる影響も小さく、比較的安定していることが分かる。さらに、その方向適合率を上昇方向の方向適合率と、低下方向の方向適合率に分解したものが、図 4, 5 である。乱数シードによるばらつきは見られるものの、機械学習によって方向適合率に違いが見られる。例えば、TOPIX においては、decision tree, random forest が上昇方向の方向適合率が高く、bagging, adaboost が低下方向の方向適合率が高くなっている。実際に、平均値の差

の検定を行うと、その差が有意であることが分かる(表 5, 6)。この結果を、運用へと応用することを考えたい。

以下の、4 パターンのシミュレーションを行う。まず、上昇方向の方向適合率が高い decision tree の予測結果に注目し、decision tree が株価リターン上昇を予測した場合には、ロングポジションをとる。次に、低下方向の方向適合率が高い bagging に注目し、bagging が株価リターン低下を予測した場合には、ショートポジションをとる。それ以外の場合には、ニュートラルポジションをとる (DT + BG)。同様のシミュレーションを、decision tree と adaboost を組み合わせた場合 (DT + AB)、random forest と bagging を組み合わせた場合 (RF + BG)、random forest と adaboost を組み合わせた場合 (RF + AB) に対しても行う。その結果が、図 6 である。

この結果を見ると、TOPIX の Pattern 1 において、DT + BG が decision tree 単品で用いた場合よりも、若干水準が良くなっている。また、RF + BG も、random forest よりは若干水準が高い。また、Pattern 2 においても、decision tree, random forest 単品よりも、組み合わせたほうが若干良くなるように見える。他の資産に対しても若干パフォーマンスが改善するような傾向は見られるものの、いずれも統計的に有意

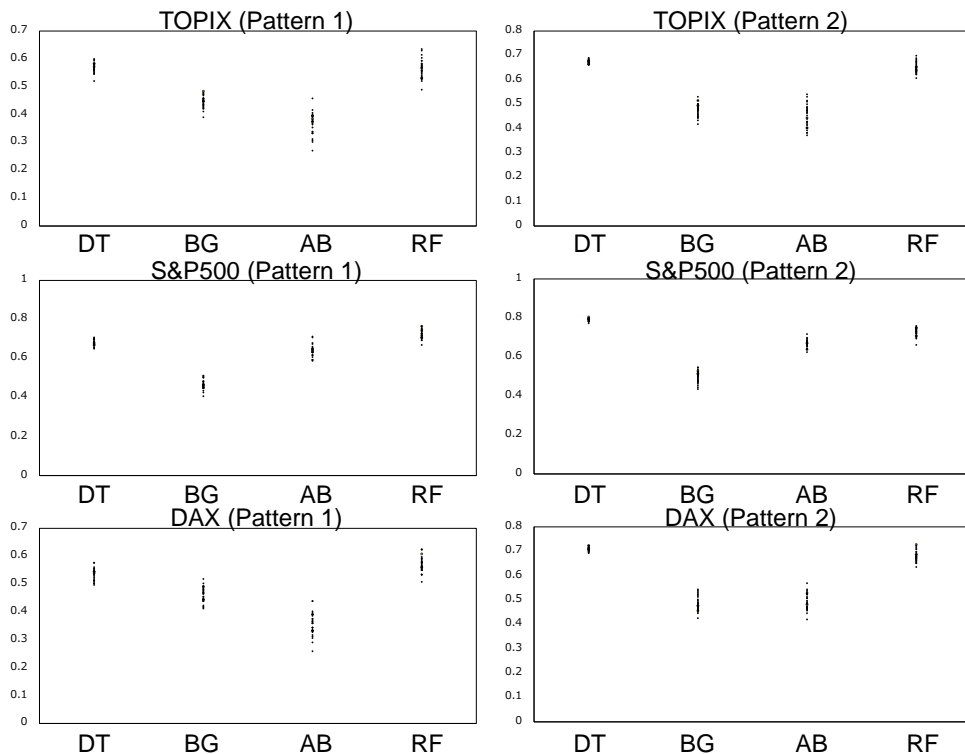


図 4: 上昇方向の方向適合率。それぞれ、DT は decision tree, BG は bagging, AB は adaboost, RF は random forest の結果を表す。

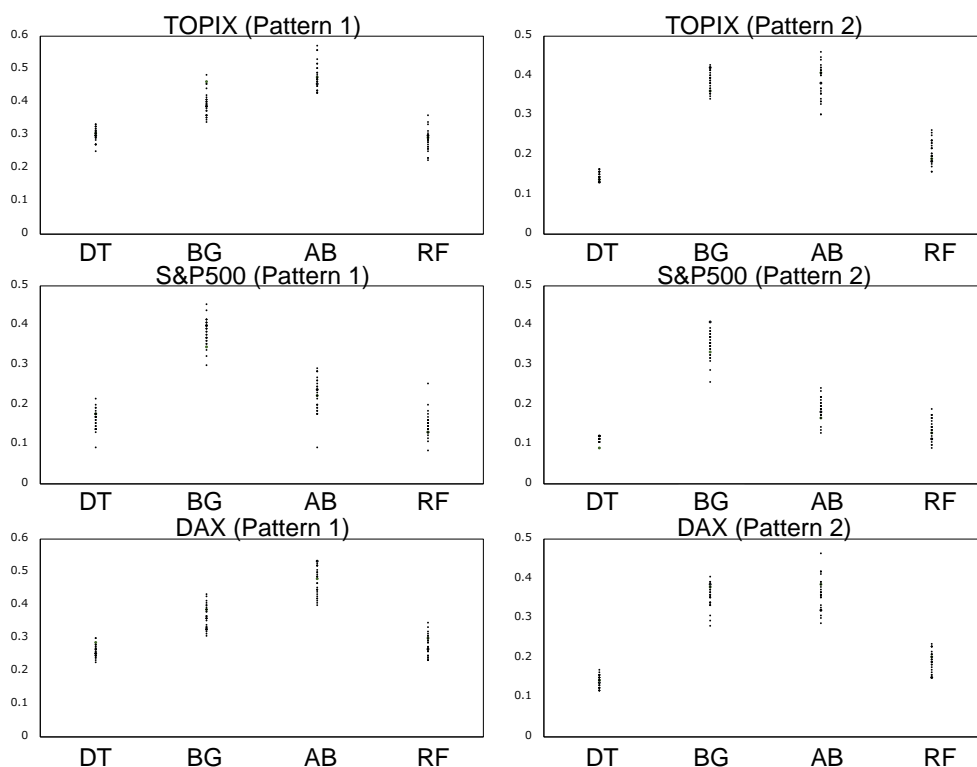


図 5: 低下方向の方向適合率. それぞれ, DT は decision tree, BG は bagging, AB は adaboost, RF は random forest の結果を表す.

表 5: 上昇方向の方向適合率に関する平均値の差の検定. それぞれ, DT は decision tree, BG は bagging, AB は adaboost, RF は random forest の結果を表す.

| TOPIX (Pattern 1) |          |          |          |          | TOPIX (Pattern 2) |          |          |          |          |
|-------------------|----------|----------|----------|----------|-------------------|----------|----------|----------|----------|
|                   | DT vs BG | DT vs AB | RF vs BG | RF vs AB |                   | DT vs BG | DT vs AB | RF vs BG | RF vs AB |
| t                 | 23.6     | 26.3     | 16.6     | 21.8     | t                 | 39.7     | 26.3     | 27.8     | 21.4     |
| 自由度               | 54       | 40       | 52       | 56       | 自由度               | 34       | 31       | 56       | 42       |
| P (両側)            | 3.9E-30  | 7.6E-27  | 1.8E-22  | 4.4E-29  | P (両側)            | 4.5E-30  | 9.2E-23  | 1.9E-34  | 3.2E-24  |

| S&P500 (Pattern 1) |          |          |          |          | S&P500 (Pattern 2) |          |          |          |          |
|--------------------|----------|----------|----------|----------|--------------------|----------|----------|----------|----------|
|                    | DT vs BG | DT vs AB | RF vs BG | RF vs AB |                    | DT vs BG | DT vs AB | RF vs BG | RF vs AB |
| t                  | 41.1     | 7.0      | 42.7     | 12.8     | t                  | 53.1     | 32.1     | 34.3     | 10.6     |
| 自由度                | 47       | 42       | 57       | 54       | 自由度                | 33       | 37       | 53       | 58       |
| P (両側)             | 1.8E-38  | 1.5E-08  | 5.4E-45  | 5.4E-18  | P (両側)             | 1.5E-33  | 1.3E-28  | 7.3E-38  | 3.6E-15  |

| DAX (Pattern 1) |          |          |          |          | DAX (Pattern 2) |          |          |          |          |
|-----------------|----------|----------|----------|----------|-----------------|----------|----------|----------|----------|
|                 | DT vs BG | DT vs AB | RF vs BG | RF vs AB |                 | DT vs BG | DT vs AB | RF vs BG | RF vs AB |
| t               | 11.7     | 21.7     | 15.3     | 24.0     | t               | 40.0     | 32.9     | 28.7     | 24.4     |
| 自由度             | 55       | 44       | 58       | 53       | 自由度             | 32       | 31       | 53       | 49       |
| P (両側)          | 1.6E-16  | 4.2E-25  | 4.4E-22  | 3.7E-30  | P (両側)          | 7.0E-29  | 1.1E-25  | 5.5E-34  | 4.7E-29  |

表 6: 低下方向の方向適合率に関する平均値の差の検定. それぞれ, DT は decision tree, BG は bagging, AB は adaboost, RF は random forest の結果を表す.

| TOPIX (Pattern 1) |          |          |          |          | TOPIX (Pattern 2) |          |          |          |          |
|-------------------|----------|----------|----------|----------|-------------------|----------|----------|----------|----------|
|                   | DT vs BG | DT vs AB | RF vs BG | RF vs AB |                   | DT vs BG | DT vs AB | RF vs BG | RF vs AB |
| t                 | 12.4     | 21.9     | 12.7     | 21.2     | t                 | 49.0     | 30.7     | 26.3     | 19.5     |
| 自由度               | 45       | 43       | 56       | 54       | 自由度               | 39       | 33       | 57       | 51       |
| P (両側)            | 4.6E-16  | 6.5E-25  | 4.1E-18  | 7.8E-28  | P (両側)            | 1.2E-36  | 7.8E-26  | 1.5E-33  | 2.7E-25  |

| S&P500 (Pattern 1) |          |          |          |          | S&P500 (Pattern 2) |          |          |          |          |
|--------------------|----------|----------|----------|----------|--------------------|----------|----------|----------|----------|
|                    | DT vs BG | DT vs AB | RF vs BG | RF vs AB |                    | DT vs BG | DT vs AB | RF vs BG | RF vs AB |
| t                  | 29.3     | 6.8      | 28.5     | 7.6      | t                  | 39.2     | 15.1     | 28.4     | 7.8      |
| 自由度                | 54       | 46       | 58       | 52       | 自由度                | 33       | 35       | 55       | 58       |
| P (両側)             | 8.1E-35  | 1.8E-08  | 8.0E-36  | 4.7E-10  | P (両側)             | 2.9E-29  | 6.6E-17  | 1.4E-34  | 1.2E-10  |

| DAX (Pattern 1) |          |          |          |          | DAX (Pattern 2) |          |          |          |          |
|-----------------|----------|----------|----------|----------|-----------------|----------|----------|----------|----------|
|                 | DT vs BG | DT vs AB | RF vs BG | RF vs AB |                 | DT vs BG | DT vs AB | RF vs BG | RF vs AB |
| t               | 15.1     | 25.8     | 10.8     | 20.9     | t               | 36.4     | 28.7     | 23.5     | 19.9     |
| 自由度             | 44       | 40       | 58       | 54       | 自由度             | 45       | 38       | 57       | 49       |
| P (両側)          | 5.2E-19  | 1.4E-26  | 1.8E-15  | 1.7E-27  | P (両側)          | 4.8E-35  | 2.2E-27  | 5.8E-31  | 3.8E-25  |

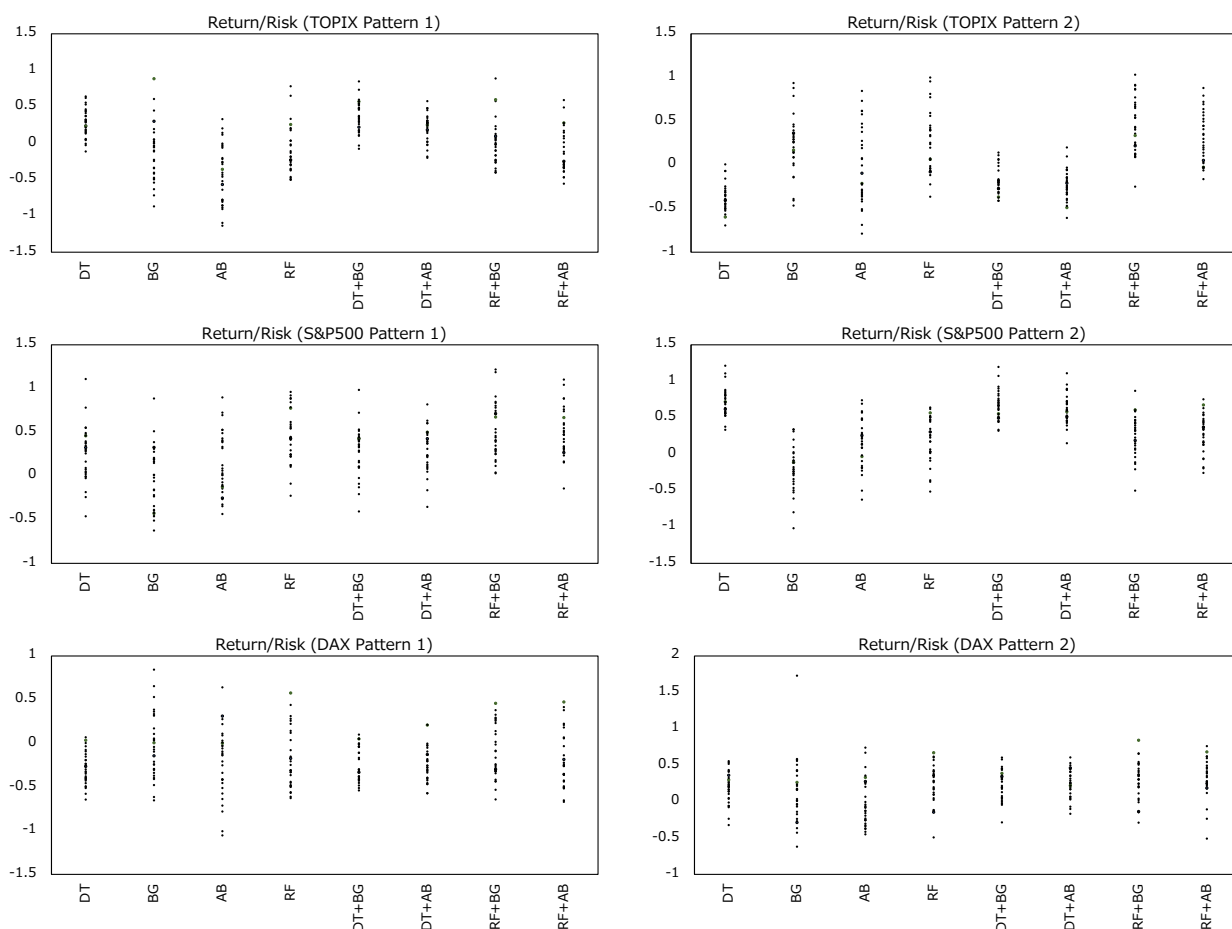


図 6: 複数の手法を組み合わせた場合のパフォーマンス. それぞれ, DT は decision tree, BG は bagging, AB は adaboost, RF は random forest の結果を表す.

なほどの改善は見られない. やはり, 機械学習を用いての運用が困難であることを示しており, 有効な運用方法に関する研究は, 今後の課題となる.

### 3. まとめと今後の展望

これまで, 機械学習を金融へ活用することに関し, 多くの研究がされてきた. しかし, いずれの研究も結果の安定性に関する議論はされていなかった. 今回, 我々は, 運用へと機械学習を活用した場合における安定性に関して, 乱数依存性という観点から研究を行った. その結果, シャープレシオは乱数シードによって大きく変動することが分かった. すなわち, 時系列方向を機械学習によって予測することは容易ではない. 評価, 実利的な有効性という観点から, 単純に機械学習を用いることは問題であることが分かる.

一方で, 方向適合率は乱数に大きくは依存しない

ことも分かった. このような乱数に依存しないような性質の調査や, 安定である理由の解明, 活用方法に関する研究が今後は必要になると考えられる. また, 今回は比較的一般的なファクターを用いて研究を行ったが, 機械学習に対してより有効に効くようなファクターを探すことも, 重要な研究となるであろう. さらに, この不安定性が何に由来するものなのか, 単純にサンプル数を増やすことで解決できることなのか, より詳細な研究が必要となる.

### 謝辞

本研究を進めるにあたり, みずほ第一フィナンシャルテクノロジー株式会社の井口亮氏との議論が参考になりました. 特に, 同氏からはハイパーパラメータの決定方法等や機械学習の不安定性に関する先行研究等, 分析の基礎にあたる部分において貴重な助言を頂きました.

## 参考文献

- [ 1 ] Patel, J., Shah, S., Thakker, P.: Predicting stock and stock price index movement using Trend Deterministic Data Preparation and machine learning techniques, *Expert Systems with Applications*, Vol. 42, Issue, 1, pp. 259-268, (2015)
- [ 2 ] Patel, J., Shah, S., Thakker, P., Kotecha, K.: Predicting stock market index using fusion of machine learning techniques, *Expert Systems with Applications*, Vol. 42, Issue, 4, pp. 2162-2172, (2015)
- [ 3 ] Heaton, J., B., Polson, N., G. and Witte, J., H.: Deep learning for finance: deep portfolios, *Applied Stochastic Models in Business and Industry*, Vol. 33, Issue, 1, pp. 3-12, (2017)
- [ 4 ] Basak, S., Kar, S., Saha, S., and Khaidem, L., Dey, S., R.: Predicting the direction of stock market prices using tree-based classifiers, Vol. 47, pp. 552-567 (2019)
- [ 5 ] Gu, S., Kelly, B., and Xiu, D.: Empirical asset pricing via machine learning, Vol. 33, issue, 5, pp. 2223–2273 (2020)
- [ 6 ] Henderson, P., Islam, R., Bachman, P., Joelle Pineau, J., Precup, D., Mege, D.: Deep Reinforcement Learning That Matters, *Thirty-Second AAAI Conference on Artificial Intelligence*, Vol., 31, No. 1 (2018)
- [ 7 ] Colas, C. , Sigaud, O., Oudeyer, P., Y.: How Many Random Seeds? Statistical Power Analysis in Deep Reinforcement Learning Experiments, *arXiv: 1806.08295v2* (2018)
- [ 8 ] Pedregosa, F., G. Varoquaux, A. Gramfort, V. Michel, B. Thirion, O. Grisel, M. Blondel, P. Prettenhofer, R. Weiss, V. Dubourg, and J. Vanderplas. *Scikit-Learn: Machine Learning in Python*”, *Journal of Machine Learning Research*, Vol. 12, pp. 2825–2830 (2011)
- [ 9 ] [https://mba.tuck.dartmouth.edu/pages/faculty/ken.french/data\\_library.html](https://mba.tuck.dartmouth.edu/pages/faculty/ken.french/data_library.html)