2017年MFTA® 合格論文

“ジョン・ブルックス賞”受賞
人工知能の集合知によるアルゴリズム運用

茨城大学大学院理工学研究科 機械システム工学専攻
鈴木 智也

要　旨

テクニカル分析の予測可能性および収益性の向上を目指して、3つのアイディアを導入した。第一に、複雑な株価変動内に隠れた法則性を自動検出すべく、伝統的な機械学習モデルの1種であるニューラルネットワークを導入した。第二に、その予測力を高める工夫として集団学習法を適用し、予測結果の自信度を評価する新しいテクニカル指標（コンセンサスレシオ）を提案した。第三に、その自信度が高い銘柄を選択する枠組みとして2段階選抜法を提案した。これらの妥当性を評価すべく、日本および米国の約1000銘柄に対して4期間の投資シミュレーションを行ったところ、良好なパフォーマンスを得たので報告する。

1. はじめに

人工知能を用いるメリットとして、大量の情報を高速に自動処理し、客観的に物事を判断できる。そのため、複雑な株価変動に隠れた規則的なパターンを検出する用途や、主観的な感情に流されないロジカルな運用を試みる場合、人工知能は有効に機能すると考えられる。人工知能に関する学術分野は多岐に渡るが、本稿では人工知能の頭脳を構成する機械学習（教師あり学習）[1]に着眼する。Alpha Go[2]の台頭によりディープラーニング（深層学習）が一躍有名になったが、その基礎となるニューラルネットワークは以前より最も代表的な機械学習モデルであるため、本稿においてもニューラルネットワークに基づいて運用モデルを構築する。ただし過去の値動きのみを用いて投資判断を行うため、テクニカル分析の範疇を超えない。このようにコンピュータ技術を駆使するテクニカル分析は「モダンテクニカル分析」として認知されつつある。

本稿では、さらに高度化したモダンテクニカル分析を提案する。そのコア概念として「集合知」を導入し、たくさんの人工知能（機械学習）による予測会議を実施する。「3人寄れば文殊の知恵」と言われるように、他人の知恵を集合知として統合することで、物事の判断能力を向上できることが知られている[3, 4]。そこで本研究では、株価予測において集合知の効果を発揮する。さらに株価変動の法則性が高まる瞬間において、それぞれの人工知能は同じような予測結果を示すと考えられる。つまり予測結果の合意度が法則性の強さに関するテクニカル指標として機能する可能性があるため、この指標に基づいた投資銘柄の選択方法についても検討する。以上の妥当性を評価すべく、日本および米国で観測された現実の株価データを用いて投資シミュレーションを実施し、株価予測力や収益性について分析する。

2. 機械学習による非線形予測モデル

伝統的なファイナンスの分野においては変数間の説明性が求められるため、重回帰分析が広く用いられる。しかし機械学習の手法として最も基礎的であるため、線形な関係性しか表現できない。一方、説明性よりも運用成績を重要視するならば、現実の金融市場は線形モデルで表現できるほど単純ではないため、非線形な関係性も表現できる機械学習モデルを積極的に導入したい。そこで本稿では、最も代表的な非線形モデルとしてニューラルネットワークを導入する。

2-1. ニューラルネットワーク

まず導入として、いかなる予測モデルも下記のように表現できる。

$$\text{未来} = F(\text{過去})$$

ここで F は過去と未来の関係を表し、なんらかの関数で記述できる。よって関数 F に「過去」の情報を入力すれば、その出力として「未来」の予測値を得ることができる。たとえば線形回帰式の場合、関数 F は下記のようになる。

$$\begin{aligned}\hat{x}(t+1) &= F(x(t), x(t-1), \cdots, x(t-d)) \\ &= w_0 x(t) + w_1 x(t-1) + \cdots + w_d x(t-d)\end{aligned}$$

ここで $w_0, w_1, \cdots, w_d$ は回帰係数、$\hat{x}$ は予測値である。なお本稿では x を収益率とするため、

$$x(t) = \frac{price(t) - price(t-1)}{price(t-1)}$$

とする。

次に、上記の線形回帰式を非線形モデルに拡張するため、非線形フィルターとしてシグモイド関数 $f(x) = \frac{1}{1+exp(-x)}$ を挿入する。

$$\begin{aligned}\hat{x}(t+1) &= F(x(t), x(t-1), \cdots, x(t-d)) \\ &= f(w_0 x(t) + w_1 x(t-1) + \cdots + w_d x(t-d)) \\ &= \frac{1}{1 + exp[-(w_0 x(t) + w_1 x(t-1) + \cdots + w_d x(t-d))]}\end{aligned}$$

これは「単一ニューロンモデル」に相当し、その概念図を図 1 に示す。なおシグモイド関数の出力値は $\hat{x}(t+1) \in [0,1]$ のように正の値のみであるため、本稿では $2 \cdot \hat{x}(t+1) - 1 \in [-1,1]$ のようにスケール変換を施す。

単一ニューロンモデルは非線形予測モデルに相当するが、単調増加または単調減少といったシンプルな非線形性しか表現できない。そこで図 2 のように、複数のニューロンモデルを結合することで、より複雑で非線形な関数 F を表現できる。これを「ニューラルネットワーク」と呼び、下記のように記述できる。

$$\begin{aligned}\hat{x}(t+1) &= F(x(t), x(t-1), \cdots, x(t-d)) \\ &= w_1 o_1 + w_2 o_2 + \cdots + w_j o_j + \cdots + w_N o_N \\ o_j &= \frac{1}{1 + exp[-(w_{0,j} x(t) + w_{1,j} x(t-1) + \cdots + w_{d,j} x(t-d))]}\end{aligned}$$

ここで $\{w_1, w_2, \cdots, w_N\}$ および $\{w_{0,j}, w_{1,j}, \cdots, w_{d,j} \mid j = 1 \sim N\}$ はモデルパラメーターであり、d は入力層のニューロン数、N は中間層のニューロン数、o_i は中間層における第 j 番目のニューロンからの出力値である。ここで $w \in [-\infty, \infty]$ なので、$\hat{x}(t+1) \in [-\infty, \infty]$ となる。

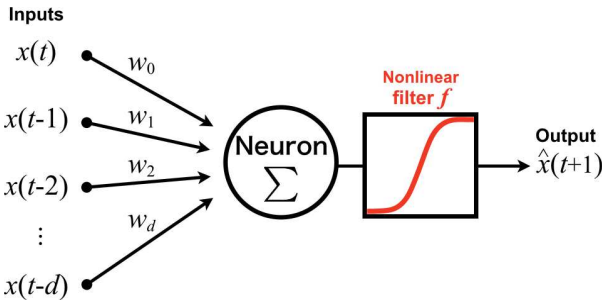

図 1. 単一ニューロンモデルの概念図

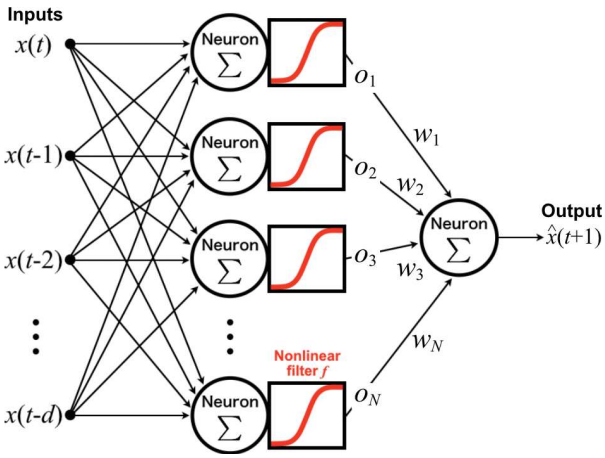

図 2. ニューラルネットワークモデルの概念図

2-2. ハイパーパラメーターの最適化による過学習の防止

予測を開始する前に、まずニューラルネットワークを学習する必要がある。そこで過去の実現データをニューラルネットワークに入力した場合、適切な出力を返せるようにモデルパラメーター w を学習する。重回帰式のように線形モデルであれば最小２乗法で解けるが、非線形性を含む機械学習モデルでは下記のような最急勾配法[1)]を用いるのが一般的である。

$$w_j \leftarrow w_j - \eta \frac{\partial E}{\partial w_j}$$

ここで $j \in \{1, \cdots, N\}$ であり、$\{w_1, w_2, \cdots, w_N\}$ のモデルパラメーターを学習する。これを解くと、

$$w_j \leftarrow w_j + \eta(x^*(t+1) - \hat{x}(t+1))o_j(t)$$

となる。ここで $\hat{x}(t+1)$ は予測値（出力値）、$x^*(t+1)$ は真値（教師信号）である。なお、η は学習係数であり０～１の値をユーザが設定する。

次に、入力層と中間層を繋ぐモデルパラメーター $\{w_{0,j}, w_{1,j}, \cdots, w_{d,j} \mid j = 1 \sim N\}$ は逆誤差伝播法[1)]によって学習できる。図２のように出力層のニューロンが１個である場合は、以下の更新式によって学習できる。

$$w_{i,j} \leftarrow w_{i,j} + \eta w_j(x^*(t+1) - \hat{x}(t+1))x(t-i+1)$$

ここで $i \in \{1, \cdots, d\}$ である。真値 $x^*(t+1)$ と予測値 $\hat{x}(t+1)$ の間の学習誤差が充分に小さくなるまで、上式の更新を繰り返す。図３に示すように、この学習過程を「バックテスト」と呼び、学習誤差 E は以下の平均２乗誤差によって算出される。

$$E = \frac{1}{\beta} \sum_{t=\alpha}^{\alpha+\beta-1} (x^*(t+1) - \hat{x}(t+1))^2$$

ここで、α はバックテストの開始時刻、β はバックテストに用いた学習データ数である。

全てのモデルパラメーターを学習したのち、新規データを予測するために学習済みニューラルネットワークを用いる。学習した入出力データの関係性を $\hat{F}$ と書くと、出力値である $\hat{x}(t+1)$ は予測値となる。なお本稿において、推定値には ^ 印を付ける。

$$\hat{x}(t+1) = \hat{F}(x(t)), x(t-1), \cdots, x(t-d)$$

ここで $\{x(t), x(t-1), \cdots, x(t-d)\}$ は学習に用いていない新規の収益率データである。このように新規データに対する予測のことをフォワードテスト（図３）と呼ぶ。

バックテストにおいてモデルパラメーターを学

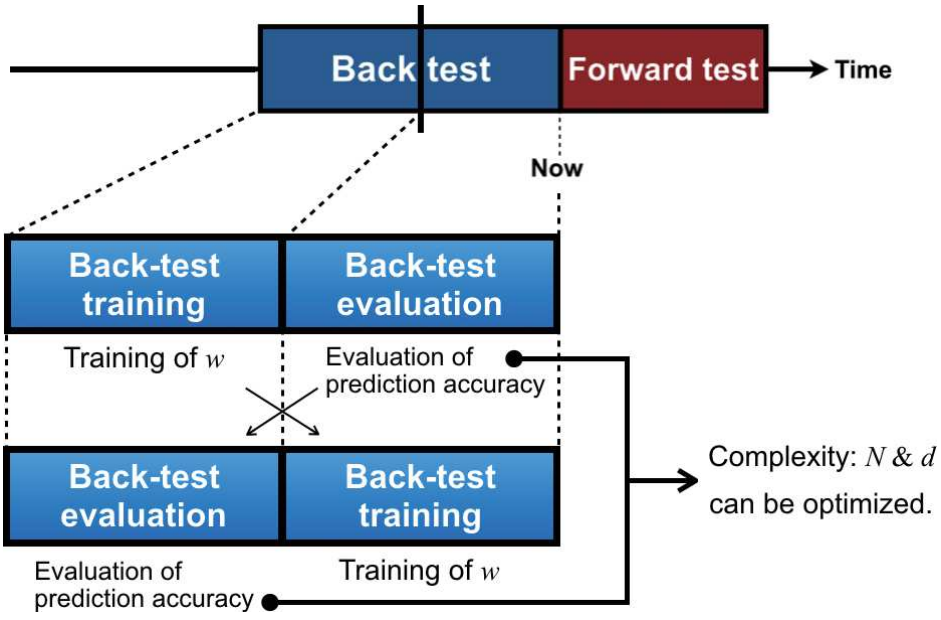

図 3. 交差検証法の概念図

習する際、過学習（カーブフィッティングやオーバーフィッティングとも言う）に注意する必要がある。もし学習済みニューラルネットワークがバックテストのみにおいて機能しても全く無意味である。当然ながら、新規データのフォワードテストにおいて機能する必要がある。図 4 に示すように、予測モデルが複雑になるほど過去の学習データに対する当てはめ力が高まる一方、新規データに対する汎用性を失い新規のテストデータに対する当てはめ力が低下する。このような過学習を防止するために、我々は予測モデルの最適な複雑性を選ぶ必要がある。ニューラルネットワークにおいては、中間層のニューロン数 N や入力層に与えるデータ数 d が複雑性に起因し、これらをハイパーパラメーターと呼ぶ。ハイパーパラメーターもモデルパラメーターと同様に、バックテストで最適化する必要がある。

しかし新規データを適用しない限り過学習の程度が分からないため、図 3 に示すようにバックテストに用いる学習データを 2 分割して訓練部（Training）と評価部（Evaluation）に分ける。まずハイパーパラメーターを適当に決め、訓練部の誤差が最小になるようにモデルパラメーターを最適化し、評価部に対して汎化性能を調べる。次に訓練部と評価部を入れかえて同様にモデルパラメーターの最適化および汎化性能の評価を行う。最後に入れ替え前後の汎化性能を平均化することで、最初に定めたモデルパラメーターの妥当性を評価する。これは「交差検証法」と呼ばれ、機械学習の汎化性能を調べる方法として一般的に用いられる [1)]。この汎化性能が最良となるハイパーパラメーターを特定した後に、バックテスト全体を学習データとして用いてモデルパラメーターを最適化し直すことで、過学習を抑えた学習を行うことができる。なお上述のニューラルネットワークの場合、モデルパラメーターは $\{w_1, w_2, \cdots, w_N\}$ や $\{w_{0,j}, w_{1,j}, \cdots, w_{d,j} \mid j = 1 \sim N\}$、ハイパーパラメーターは N や d に相当する。

2-3. フォワードテストによる予測精度

交差検証法によってニューラルネットワークを学習した後、交差検証法に用いたデータ以降の新規データに対してフォワードテストを実施する。本稿では 1991 年から 2013 年に観測された実際の株価をシミュレーションに用いるが、市場構造

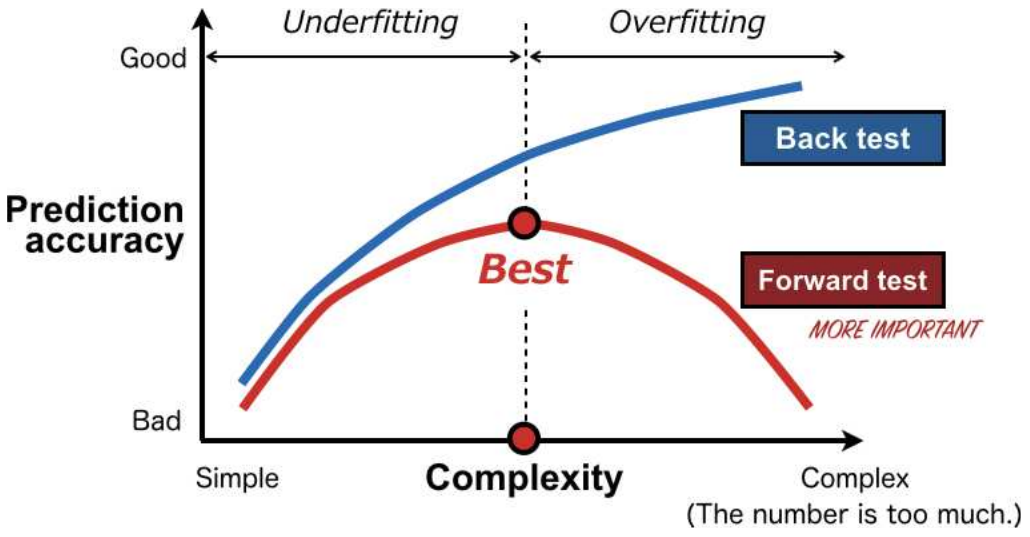

図 4. 未学習（Underfitting）と過学習（Overfitting）の違い
バックテストにおいてニューロン間の結合強度 w（モデルパラメーター）を最適化し、フォワードテストにおいて最適化したニューラルネットワークの汎化性能を未学習と過学習の観点から評価する。この汎化性能が最大になるように、ニューラルネットワークの複雑さ（ハイパーパラメーター）を最適化する。ただし図 3 の交差検証法では、Back-test evaluation がここでのフォワードテストに相当する。

表 1. シミュレーションに用いた株価データの詳細
全期間でデータに欠損が無い銘柄のみを使用したため、途中で上場廃止になった銘柄には投資しない仕組みになる。これにより生存バイアスの影響を受ける可能性がある点に留意する。なお年間の営業日は約 250 日であるため、2.5 年の学習データのサイズは約 1250 点である。

市場		東京証券取引所	ニューヨーク証券取引所
データ入手先		Yahoo! finance Japan [5]	Yahoo! Finance [6]
対象銘柄数		590	500
第 1 期	バックテスト（パラメーター学習）	1991/1 ～ 1995/12（5 年） （Training：2.5 年、Evaluation：2.5 年）	
	フォワードテスト（運用）	1996/1 ～ 1998/6（2.5 年）	
	上昇銘柄の割合	14.9%（強いベア相場）	69.3%（ブル相場）
第 2 期	バックテスト（パラメーター学習）	1996/1 ～ 2000/12（5 年） （Training：2.5 年、Evaluation：2.5 年）	
	フォワードテスト（運用）	2001/1 ～ 2003/6（2.5 years）	
	上昇銘柄の割合	55.9%（Bullish）	60.5%（Bullish）
第 3 期	バックテスト（パラメーター学習）	2001/1 ～ 2005/12（5 年） （Training：2.5 年、Evaluation：2.5 年）	
	フォワードテスト（運用）	2006/1 ～ 2008/6（2.5 年）	
	上昇銘柄の割合	17.1%（ベア相場）	59.1%（ブル相場）
第 4 期	バックテスト（パラメーター学習）	2006/1 ～ 2010/12（5 年） （Training：2.5 年、Evaluation：2.5 年）	
	フォワードテスト（運用）	2011/1 ～ 2013/6（2.5 年）	
	上昇銘柄の割合	69.2%（ベア相場）	82.2%（強いブル相場）

の経年変化を考慮して表 1 に示すように 4 期に分ける。対象とする銘柄として、東京証券取引所第一部より 590 銘柄、ニューヨーク証券取引所より 500 銘柄を選出した。銘柄数については単に全期間で欠損が無い銘柄に厳選したためであり、特に選出において恣意性はない。しかし期間の途中で上場廃止になった銘柄は必然的に投資対象外となるため、生存バイアスの影響を受ける可能性は否めない。

　各 4 期間におけるフォワードテストの予測精度を表 2 に示す。この予測シミュレーションでは、翌日の株価が上昇 $(x(t+1) \geq 0)$ または下降 $(x(t+1) < 0)$ するかの 2 択を予測対象としたため、予測精度が 50% を超過するほど優れている。加えて、2-2 章の交差検証法にてハイパーパラメーターを最適化する際も、この予測精度を用いて汎化性能を評価した。つまりニューラルネットワークの複雑さは、この予測精度を最大化するように決定した。その後、最適化されたニューラルネットワークを新規データに適用（フォワードテストを実施）し、その予測精度を表 2 に示す。結果として、いずれのケースも 50% 以上の予測精度を示しており、実際の株価変動でもわずかに予測の余地があることを示唆している。しかしどの予測精度も充分に高いとは言えない。そこで次章以降、この予測精度を向上させるために様々な工夫を施す。

表 2. フォワードテストにおける予測精度
（全銘柄の平均値）

	東京証券取引所	ニューヨーク証券取引所
第 1 期	54.8%	59.1%
第 2 期	55.8%	52.7%
第 3 期	51.6%	52.0%
第 4 期	54.8%	53.2%

3. 人工知能の集合知

前章では、複雑な株価変動に隠れた法則性を自動検出すべく、伝統的な機械学習モデルの 1 種であるニューラルネットワークを導入したが、その予測精度は充分に高いとは言えない。そこで第二のアイディアとして集団学習法を導入し、集合知の効果により機械学習の予測力を高める工夫を施す。

3-1. 集合知の例

前章の予測シミュレーションと同様に、翌日の株価が上昇または下降するかを 2 択で予測する場合を例題として考えてみたい。例えば、19 人の経済評論家に相場予想を伺うことができるとし、彼らの予測精度は 53% 程度（前章のニューラルネットワークと同程度）とする。もし 19 人の多数決で上昇または下降を予測する場合、その予測精度はどの程度になるであろうか？

これは二項分布を積分することで解くことができる。それぞれの経済評論家の予想が当たる確率を p とすると、当たるか外れるかが独立に決まるならば、19 人中 x 人当たる確率 $P(x)$ は、下記の二項分布になる。

$$P(x) = {}_{19}C_x p^x (1-p)^{19-x}$$

ここで p が当たり確率、$1-p$ が外れ確率である。なお例題では $p = 0.53$ に相当する。多数決が正解するには過半数の 10 人以上 $(x \geq 10)$ が正解すれば良いので、多数決の予測精度（当たり確率）は

$$\sum_{x=10}^{19} P(x) = 0.60$$

となる。このように多数決を取るだけで、予測精度は 53% から 60% へ向上する。これが集合知の効果である。

初期の予測精度 p を様々変えて同様の多数決を取った場合の結果を表 3 に示す。全く予測力がない場合（つまり $p = 0.50\%$）は、たとえ多数決を取ったとしても全く効果はない。しかし $p > 50\%$ であれば、50% からの乖離度（予測力）が大きいほど多数決によって益々予測力を拡大できる。この観点より、機械学習にもこの集合知を適用しようという発想は極めて自然であり、集団学習として様々な方法論が提案されている[7]。

表 3. 19 人の多数決（集合知）によって予測精度が向上する様子

初期の予測精度 p	50%	55%	60%	65%	70%
多数決による予測精度	50%	**67%**	**81%**	**91%**	**97%**

3-2. ニューラルネットワークによる集団学習

先述の例題での計算過程で述べたように、集合知および集団学習の効果を発揮するには、それぞれの予測が独立になされる必要がある。しかしこの条件を完全に満たすのは困難であるが、以下に示すバギング[8]という学習アルゴリズムによってある程度の独立性を担保することができる。

まず始めに、図 5 に示すように、学習データの各行をランダムに復元抽出することで、異なる学習データを複製する。その際、入力データと教師データ（Ideal output x^*）の組合せは破壊せず、元の学習データと同じサイズになるまで復元抽出を繰り返す。複製された学習データは元の学習データと異なるため、それぞれで訓練されたニューラルネットワークもある程度独立な予測器になることが期待される。よって同種の復元抽出を繰り返し、多数の独立なニューラルネットワークを構築することで、集合知（集団学習）の効果を発揮する。

バックテストにおいて全ニューラルネットワークのモデルパラメーターおよびハイパーパラメーターを学習した後、図 6 に示すように実際の運用（フォワードテスト）に活用する。新規の観測データを全ニューラルネットワークに入力し、未来の株価変動に関する予測値を得る。これらの多数決を取ることで集団としての最終予測値を得る。

表 4 にニューラルネットワークの集団学習によって得られた予測精度を示す。なお表 2 と同じデータを用い、1000 体のニューラルネットワークを独立に学習した。結果として、全てのケースにおいて予測精度は約 55 ～ 60% 程度に上昇したが、まだ充分に高いとは言えない。しかし集団学習によって悪化するケースは皆無であるため、計算コストは増加するものの予測精度の向上にとって有効だと言える。

Original training data

No.	Inputs to a neural network	Ideal output (Teacher signal)
#1	$x(t\text{-}1), x(t\text{-}2), \ldots, x(t\text{-}d\text{-}1)$	$x(t)$
#2	$x(t\text{-}2), x(t\text{-}3), \ldots, x(t\text{-}d\text{-}2)$	$x(t\text{-}1)$
#3	$x(t\text{-}3), x(t\text{-}4), \ldots, x(t\text{-}d\text{-}3)$	$x(t\text{-}2)$
#4	$x(t\text{-}4), x(t\text{-}5), \ldots, x(t\text{-}d\text{-}4)$	$x(t\text{-}3)$
#5	$x(t\text{-}5), x(t\text{-}6), \ldots, x(t\text{-}d\text{-}5)$	$x(t\text{-}4)$
#6	$x(t\text{-}6), x(t\text{-}7), \ldots, x(t\text{-}d\text{-}6)$	$x(t\text{-}5)$
#7	$x(t\text{-}7), x(t\text{-}8), \ldots, x(t\text{-}d\text{-}7)$	$x(t\text{-}6)$
#8	$x(t\text{-}8), x(t\text{-}9), \ldots, x(t\text{-}d\text{-}8)$	$x(t\text{-}7)$
#9	$x(t\text{-}9), x(t\text{-}10), \ldots, x(t\text{-}d\text{-}9)$	$x(t\text{-}8)$
#10	$x(t\text{-}10), x(t\text{-}11), \ldots, x(t\text{-}d\text{-}10)$	$x(t\text{-}9)$

Randomized training data

No.	Inputs to a neural network	Ideal output (Teacher signal)
#1	$x(t\text{-}2), x(t\text{-}3), \ldots, x(t\text{-}d\text{-}2)$	$x(t\text{-}1)$
#2	$x(t\text{-}8), x(t\text{-}9), \ldots, x(t\text{-}d\text{-}8)$	$x(t\text{-}7)$
#3	$x(t\text{-}6), x(t\text{-}7), \ldots, x(t\text{-}d\text{-}6)$	$x(t\text{-}5)$
#4	$x(t\text{-}2), x(t\text{-}3), \ldots, x(t\text{-}d\text{-}2)$	$x(t\text{-}1)$
#5	$x(t\text{-}10), x(t\text{-}11), \ldots, x(t\text{-}d\text{-}10)$	$x(t\text{-}9)$
#6	$x(t\text{-}4), x(t\text{-}5), \ldots, x(t\text{-}d\text{-}4)$	$x(t\text{-}3)$
#7	$x(t\text{-}6), x(t\text{-}7), \ldots, x(t\text{-}d\text{-}6)$	$x(t\text{-}5)$
#8	$x(t\text{-}4), x(t\text{-}5), \ldots, x(t\text{-}d\text{-}4)$	$x(t\text{-}3)$
#9	$x(t\text{-}10), x(t\text{-}11), \ldots, x(t\text{-}d\text{-}10)$	$x(t\text{-}9)$
#10	$x(t\text{-}4), x(t\text{-}5), \ldots, x(t\text{-}d\text{-}4)$	$x(t\text{-}3)$

図 5. 集団学習において学習データをランダム復元抽出する様子
実際のサンプル数は 10 以上であり、各サンプルにおける時間順序（行情報）を破壊しないように列情報をランダム化する。

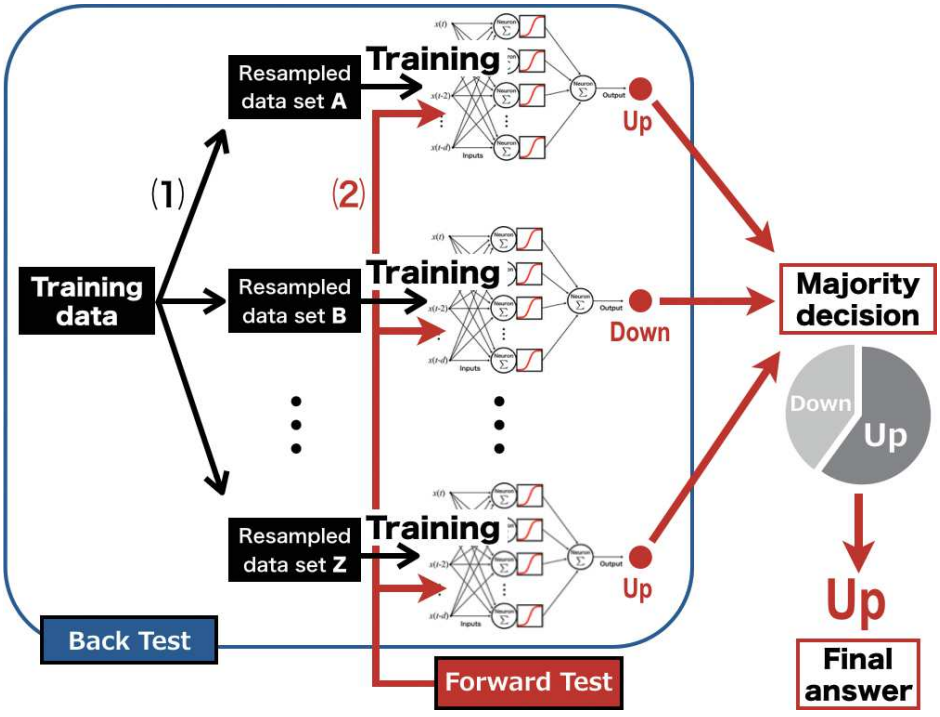

図 6. ニューラルネットワークによる集団学習
まずバックテストにおいて、再サンプルした学習データを用いて各ニューラルネットワークを構築する。その後のフォワードテストにおいて、各ニューラルネットワークが出力した予測値を統合し、その多数決によって最終的な予測値を算出する。

表 4. 表 2 と同様。ただし集団学習を適用した場合。全てのケースにおいて予測精度は向上している。

	東京証券取引所	ニューヨーク証券取引所
第 1 期	59.9%	61.7%
第 2 期	57.9%	55.6%
第 3 期	55.5%	55.3%
第 4 期	57.0%	54.6%

3-3. コンセンサスレシオ

予測精度をさらに向上させる工夫として、新しいテクニカル指標を提案する。集合知の効果を利用する集団学習では主に多数決や平均値に着眼するが[7]、我々の先行研究[9, 10]において、集団のばらつき（標準偏差）に着目することで予測に伴うリスクを評価できることが分かってきた。たとえば標準偏差が大きい場合は、みんなの意見が一致しないためコンセンサスが低い。したがってその場合の集合知は信頼性に乏しく、予測に伴うリスクが高い。この観点より次式のような「コンセンサスレシオ」C を提案する。

$$C(\hat{x}(t+1) \geq 0) = \frac{\hat{x}(t+1) \geq 0 \text{ を示す予測器の個数}}{\text{集団を構成する予測器の個数}}$$

つまり予測器の集団において、何割の予測器が $\hat{x}(t+1) \geq 0$（未来の株価は上昇する）と判断したかを示す。この C が大きいほどコンセンサスが高いため、自信のある予測結果と言える。そこで毎時刻 t において、各投資対象銘柄についてコンセンサスレシオ C を算出し、ある閾値 θ を超過した株式銘柄のみを投資対象とする。一方、$C(\hat{x}(t+1) \geq 0) < \theta$ の場合は充分なコンセンサスが得られていないとみなし、時刻 t においては投資対象外とする。本研究では最も単純な設定として、$\theta = 50\%$ とした。

3-4. 2 段階選択

コンセンサスレシオの活用において注意すべき点がある。全く予測ができない銘柄（ランダムウォークな株価変動）に対してコンセンサスレシオを適用した場合、コンセンサスレシオ自体も確率的にランダムに変動するため、偶発的に高いコンセンサスを示す可能性がある。この場合の高いコンセンサスは単に確率現象なので、予測能力とは無関係である。

このようなノイズ（テクニカル分析流に言えば「だまし」）を避けるためには、予測が困難な銘柄をバックテストの時点で除去しておく必要がある。本稿ではこれを「第 1 次選抜」と呼び、バックテストの予測精度が乏しい（予測可能性が低い）銘柄の下位 75% を除去し、上位 25% のみをフォワードテストに用いる。この様子を図 7 に示す。

次に、フォワードテストにおいてコンセンサスレシオ C を活用する。第 1 次選抜を通過した予測可能性が高い銘柄においても、毎時刻においてコンセンサスが高いとは限らない。図 7 右図のよう

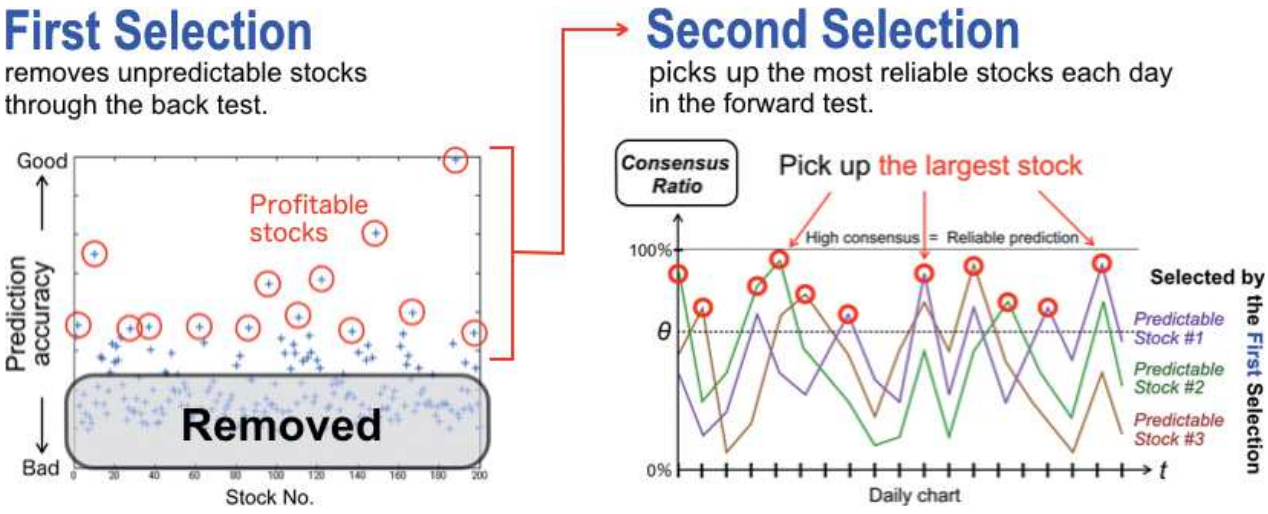

図 7. 2 段階選択の概念図

に激しく変動するため、常に特定の銘柄を投資対象に固定するのは本稿において得策ではない。そこで各時刻 t において、最も高いコンセンサスレシオ C を示す銘柄のみを投資対象とする[11]。これを本稿では「第2次選抜」と呼び、2段階のルールに基づいて最終的な投資銘柄を選出する。直感的に言えば、第 1 次選抜はテクニカル分析[12]が通用しそうな「銘柄」の選択であり、第 2 次選抜は選ばれた銘柄の中で AI らが強く法則性を感じる旬な「タイミング」の選択といえる。

表 5 に 2 段階選択によるフォワードテストの予測精度を示す。なお表 2 や表 4 と同じ株価データに対して予測実験を行った。しかし第 2 次選抜において投資対象銘柄が動的に変化するため、予測対象銘柄も各時刻 t につき 1 銘柄である。結果として、平均的に 70% 程度まで予測精度は上昇している。また表 2 や表 4 と比べて、全てのケースにおいて予測精度は向上しているため、コンセンサスレシオに基づく 2 段階選択は有用だと考えられる。

表 5. 表2と同様。ただし2段階選択を適用した場合。全てのケースにおいて、表 2 および表 4 よりも予測精度は向上している。

	東京証券取引所	ニューヨーク証券取引所
第 1 期	65.9%	84.0%
第 2 期	71.6%	69.0%
第 3 期	60.4%	60.5%
第 4 期	62.1%	57.0%

4. 投資シミュレーション

最後に、実際に運用した場合にどのような収益が得られるのかシミュレーションする。これまで紹介したように、バックテストを通じて 1000 体のニューラルネットワークのモデルパラメーターおよびハイパーパラメーターを集団学習し、多数決による予測精度が高い上位 25% を投資対象とする（第 1 次選抜）。その後のフォワードテストにおいて、各時刻 t でコンセンサスレシオ C が最大の銘柄を投資対象とする。

4-1. 運用パフォーマンス

予測精度を収益に直結させるため、投資方法は最もシンプルにする。本稿では時刻 t を日足で考えるため、毎日に第 2 次選抜された銘柄を購入し、翌日に手仕舞いする。これをフォワードテストにおいて繰り返す。なお本稿では、株価データとして始値を用いるため、毎日の市場開場時に当日分の購入と先日分の手仕舞いを同時に行う。補足として、全投資対象銘柄のコンセンサスレシオが閾値 θ に満たない場合は、その日の購入は行わない。さらに、もしショートポジションも持ちたい場合は、以下のようにコンセンサスレシオを変更すれば良い。

$$C(\hat{x}(t+1)<0)=\frac{\hat{x}(t+1)<0\text{ を示す予測器の個数}}{\text{集団を構成する予測器の個数}}$$

しかし本稿では議論を簡潔にしたいため、ロングオンリーとする。

次に、資産の時間変化を以下の式で評価する。

$$\begin{aligned}M_1(t)&=M_1(1)+M_1\cdot x(1)+\cdots+M_1\cdot x(t)\\&=M_1(t-1)+M_1(1)\cdot x(t)\\M_2(t)&=M_1(1)+M_1\cdot(1+x(1))\cdot(1+x(2))\cdot\cdots\\&\quad\cdot(1+x(t))=M_2(t-1)+M_2(1)\cdot x(t)\end{aligned}$$

ここで $M_1(1)$ および $M_2(1)$ は初期資産額であり、それぞれ 1 とする。さらに $M_1(t)$ は単利運用の場合、$M_2(t)$ は複利運用の場合に相当する。なお $x(t)$ は、時刻 $t-1$ に購入して時刻 t で手仕舞いした時の収益率であり、投資を行わない場合は $x(t)=0$ となる。さらに売買に伴う取引手数料 $cost(t)$ を考慮すると、収益率 $x(t)$ は次式となる。

$$\begin{aligned}x(t)&=\frac{price(t)-price(t-1)-cost(t)}{price(t-1)}\\cost(t)&=\frac{c}{100}\cdot(price(t)+price(t-1))\end{aligned}$$

ここで c は手数料率[%]であり、近年のネット証券を鑑みると 0.1[%]程度である。特に SMBC 日興証券の信用取引（ダイレクトコース）のように $c=0$[%]とみなせる場合もある。

本稿の提案手法に対する比較基準として、買い持ち戦略によるパフォーマンス $M_3(t)$ も評価する。

$$M_3(t)=M_3(t-1)+M_3(t-1)\cdot\bar{x}(t)$$

ここで初期資産額を $M_3(1)=1$ とし、$\bar{x}(t)$ は時刻 t における全銘柄の平均収益率である。つまり等配分ポートフォリオに相当し、買い持ちであるため複利運用となる。なお厳密には等配分を保つためにリバランスに伴う手数料が発生するが、計算が複雑化するため無視する（つまり買い持ち戦略にとって有利な状況である）。

図 8 と図 9 において、単利運用の資産変化 $M_1(t)$ を示す。なおシミュレーションに用いた株価銘柄はこれまでと同様に表 1 に示すユニバースを対象にするが、毎日の第 2 次選抜によって投資対象は動的に変化する。結果として、全ての期間および日米の両市場において、本稿の 2 段階選択の方が買い持ち戦略（Market average）よりも優れた運用パフォーマンスを示している。さらに図 12 と図 13 において、複利運用の資産変化 $M_2(t)$ を示す。本稿はデイリーの運用であるため投資機会が多く、複利効果の恩恵を受けやすい。したがって複利運用も本提案手法のポイントに含まれる。

次により現実的な状況として、$c=0.1$[%] の手数料率を考慮する。単利運用の結果 $M_1(t)$ を図 10 と図 11 に、複利運用の結果 $M_2(t)$ を図 14 と図 15 に示す。日本市場においては、全 4 期において買い持ち戦略（Market average）よりもやや優れた状況を保つが、米国市場においては、第 3 期および第 4 期おいて買い持ち戦略（Market average）とほぼ同等になる。さらに手数料率を $c=0.2$[%] にすると、日本市場においても第 3 期および第 4 期おいて買い持ち戦略（Market average）より劣る。この原因の 1 つとして、市場の構造変化が考えられる。本稿ではバックテストで学習したパラメーターを固定したままフォワードテストの運用に適用したが、新しい株価を観測する毎にパラメーターを再学習（オンライン学習）することも可能である。これにより構造変化の影響を吸収できるが、多大な計算コストを要するため今後の課題とする。その代わり、得られた運用成績の収益性について慎重に検証すべく、統計的仮説検定を実施する。

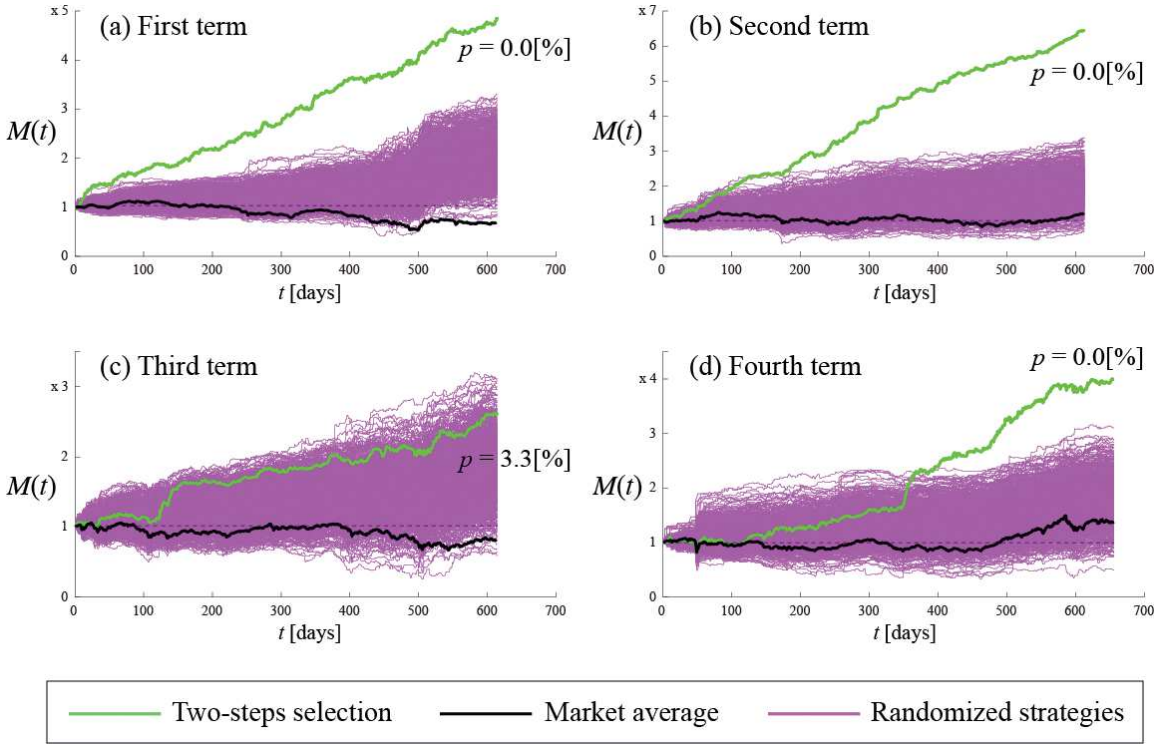

図 8. 単利運用における資産増幅率 $M_1(t)$ の時間推移。ただし日本市場かつ取引手数料率 $c=0$[%] の場合。
図中の Randomized strategies や p 値については 4-2 章の本文を参照されたし。Market average については 4-1 章の $M_3(t)$ によって算出した。

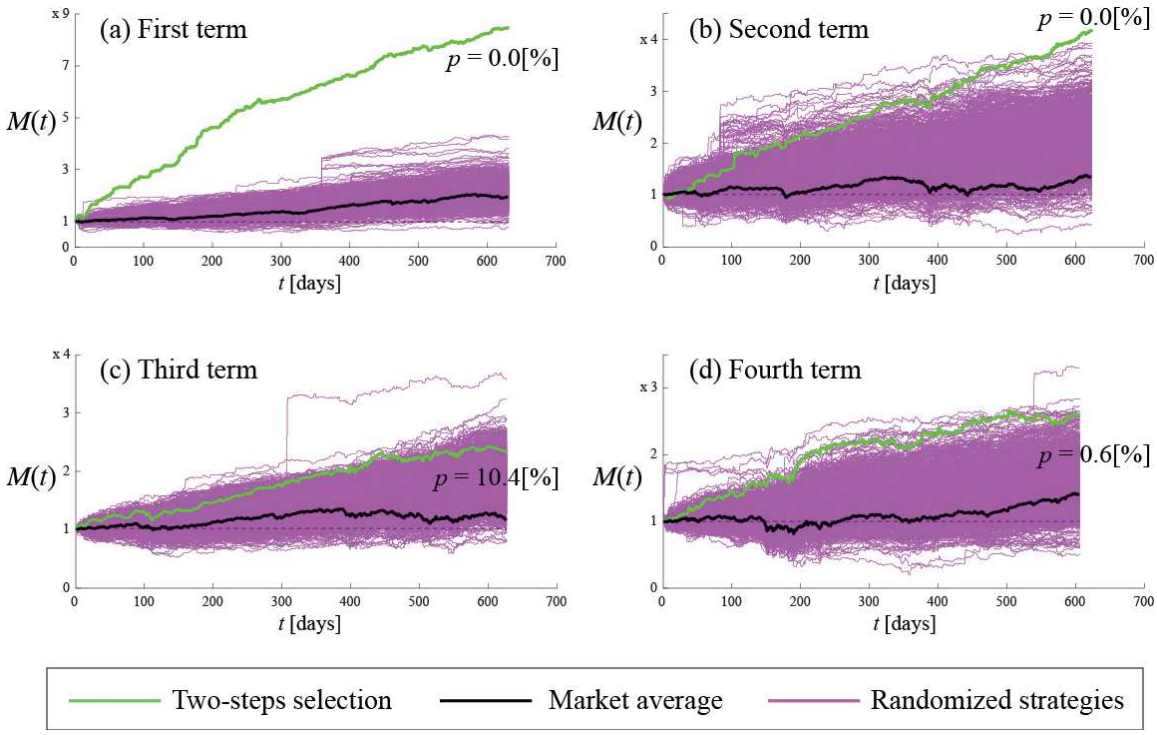

図 9. 図 8 と同様。ただし米国市場かつ取引手数料率 $c = 0$[%]の場合。

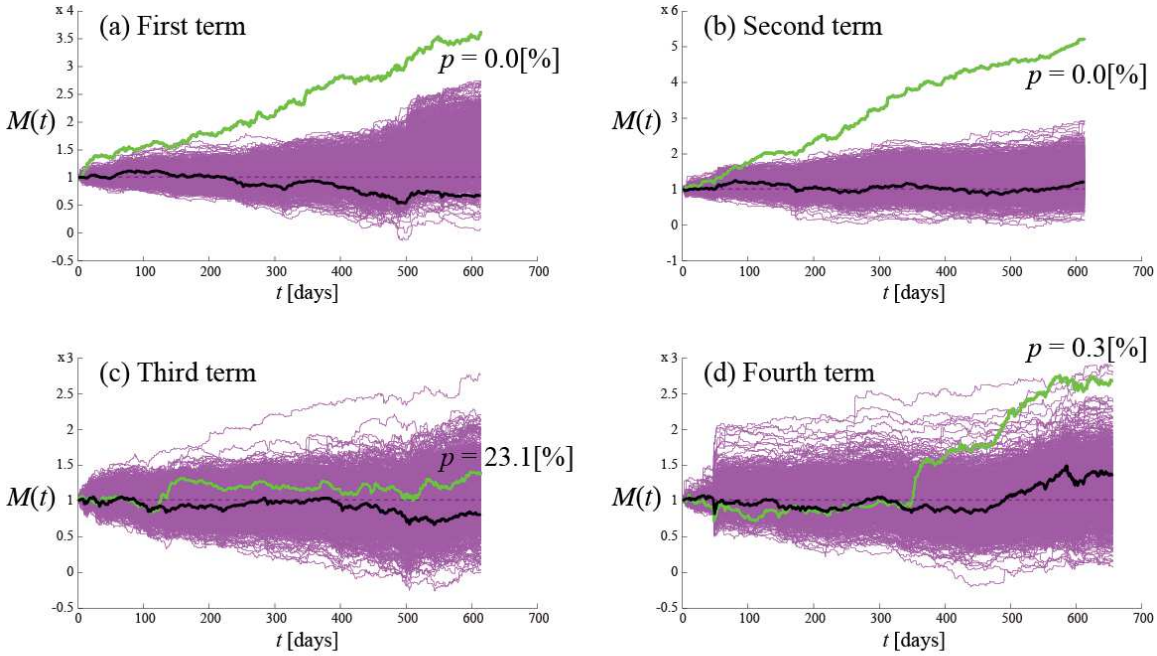

図 10. 図 8 と同様。ただし日本市場かつ取引手数料率 $c = 0.1$[%]の場合。

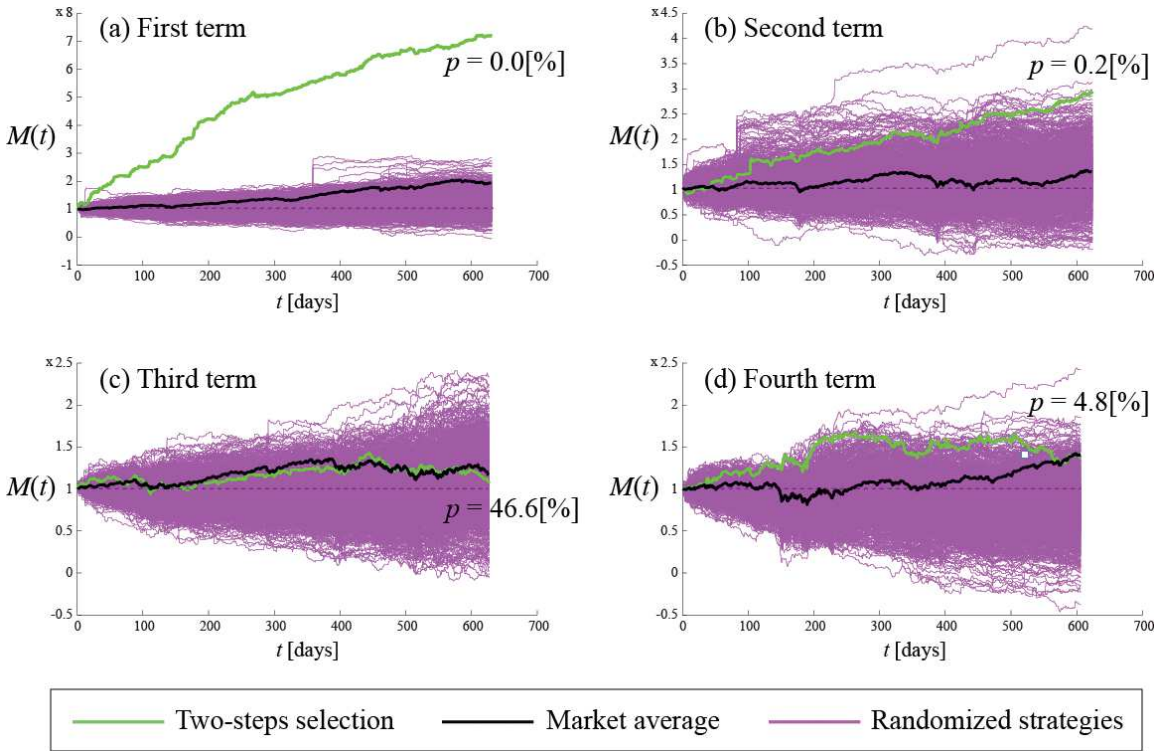

図 11. 図 8 と同様。ただし米国市場かつ取引手数料率 $c = 0.1$ [%] の場合。

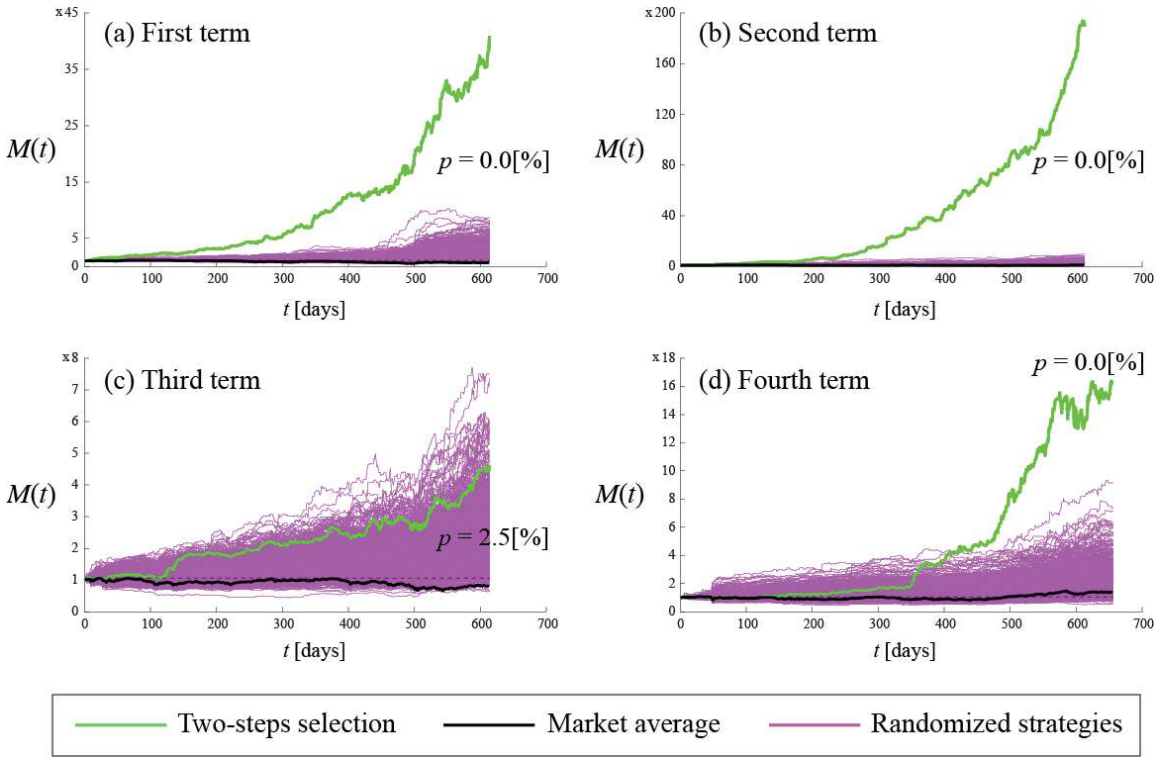

図 12. 複利運用における資産増幅率 $M_2(t)$ の時間推移。ただし日本市場かつ取引手数料率 $c = 0$ [%] の場合。
図中の Randomized strategies や p 値については 4-2 章の本文を参照されたし。Market average については 4.1 章の $M_3(t)$ によって算出した。

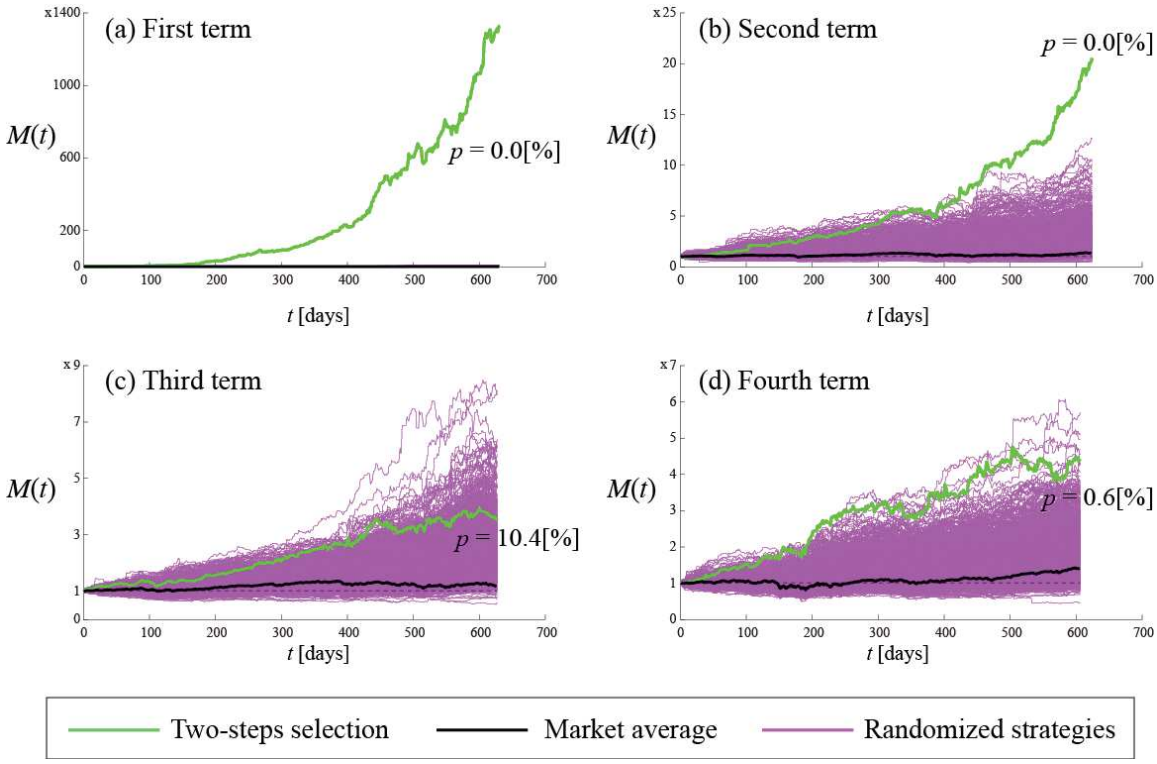

図 13. 図 12 と同様。ただし米国市場かつ取引手数料率 $c = 0$[%]の場合。

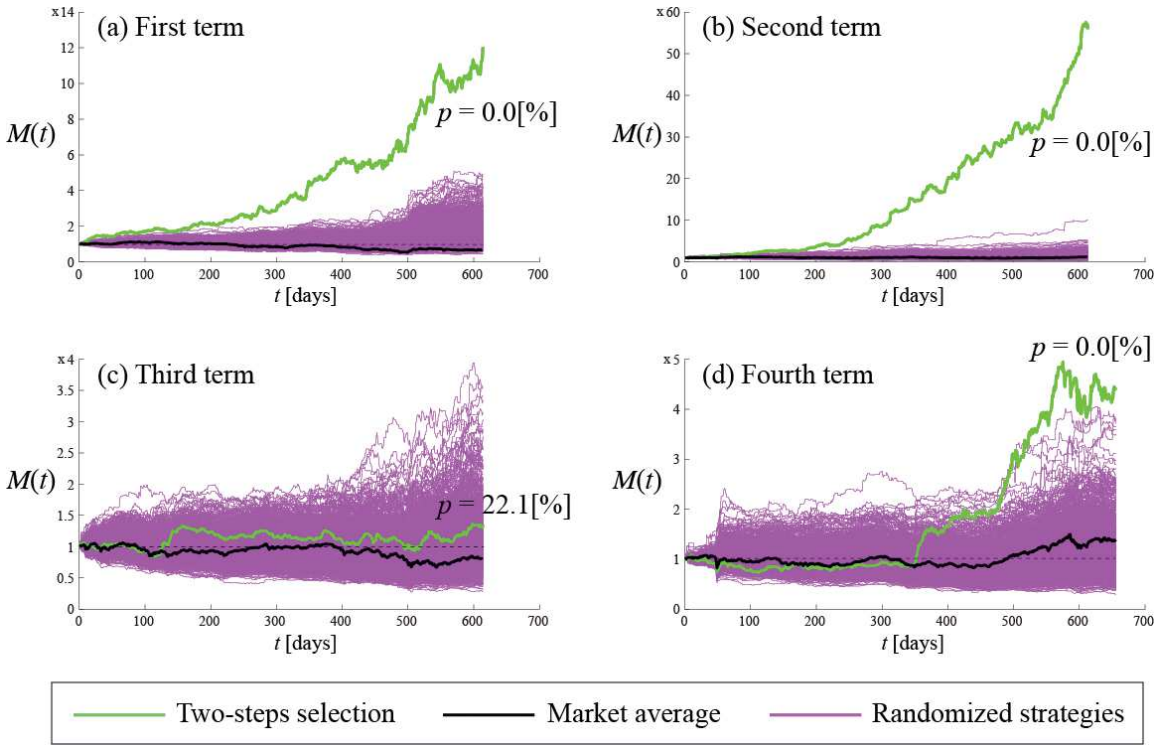

図 14. 図 12 と同様。ただし日本市場かつ取引手数料率 $c = 0.1$[%]の場合。

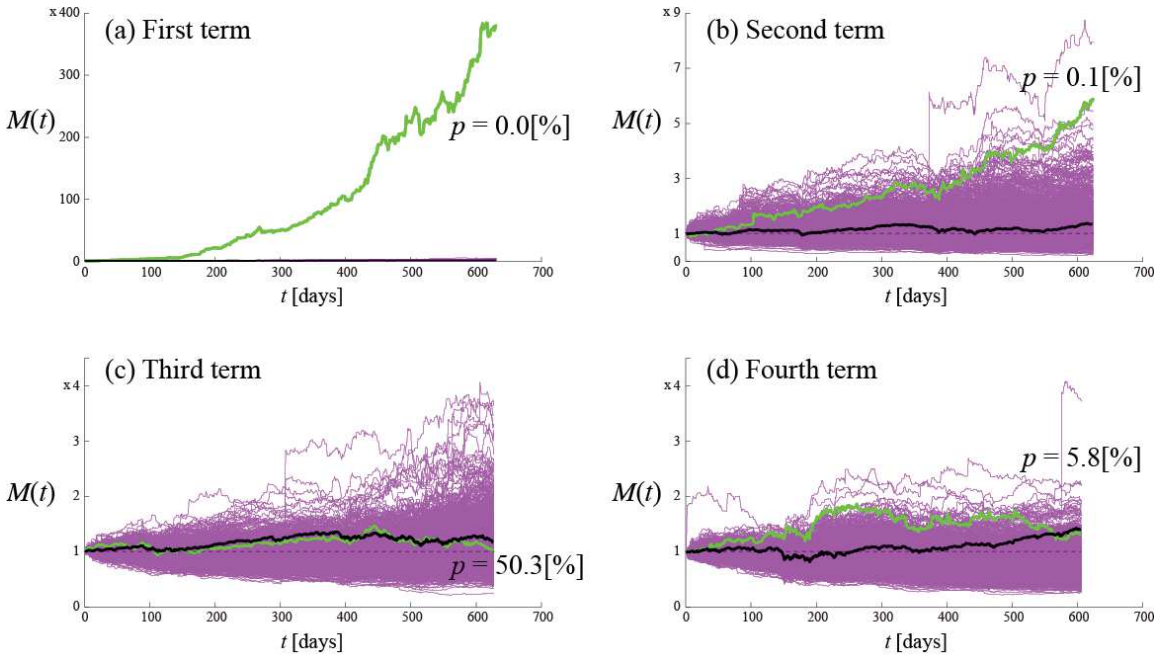

図 15. 図 12 と同様。ただし米国市場かつ取引手数料率 $c = 0.1$[%]の場合。

4–2. 統計的仮説検定による運用成績の妥当性

Evidence-based technical analysis [13]（根拠のあるテクニカル分析）の観点から、図 8 から図 15 において得られた運用成績が単にまぐれによるものではないことを統計的仮説検定によって検証する。

もし本提案手法が真に収益力を持つならば、その収益力の所以は 2 段階選択によって選ばれた投資銘柄にある。一方、得られた運用成績が単にまぐれであるならば、ランダムに投資銘柄を選択した場合においても同等の運用成績が高確率で得られるはずである。この観点より、2 段階選択とランダム選択による運用成績の乖離を検証する。

各期間においてランダム選択による運用を 1000 回実施し、運用成績の確率分布を推定した。その結果を図 8 から図 15 の「Randomized strategies」として示す。統計的仮説検定の観点より、2 段階選択を上回るランダム選択が精々 5% 未満（$p < 5$[%]）しか出現しないならば、有意水準 5% の片側検定において、2 段階選択の収益力は妥当である（まぐれではない）と主張できる。そこで各図において、ランダム選択の優越確率を p 値として示した。なお p 値は、フォワードテスト完了後の最終的な資産に基づいて算出した。結果として、第 3 期以外であれば手数料率 $c = 0.1$[%]であっても概ね $p < 5$[%]を実現している。つまり本稿が提案する 2 段階選択は統計学的に妥当であり、得られた収益は単にまぐれによるものではないことを支持する。しかし第 3 期においては、市場の構造変化を踏まえたオンライン学習が必要であると考えられる。

5. まとめ

本稿では、アルゴリズム運用の収益性を向上すべく、3 つのアイディアを融合した。第一に、複雑な株価変動内に隠れた法則性を自動検出すべく、伝統的な機械学習モデルの 1 種であるニューラルネットワークを導入した。第二に、その予測力を高める工夫として集団学習法を適用し、予測結果の自信度を評価する新しいテクニカル指標（コンセンサスレシオ）を提案した。第三に、その自信度が高い銘柄を選択する枠組みとして 2 段階選抜法を提案した。特にバックテストにおける

第 1 次選抜は、予測可能性の乏しい（ランダム性が高い）銘柄を除外するためのフィルターとして機能する。次にフォワードテストでは予測に自信が持てる銘柄を投資対象にすべく、第 2 次選抜としてコンセンサスレシオが最大となる銘柄を動的に選出した。

これらの妥当性を評価すべく、日本および米国の約 1000 銘柄に対して 4 期間の投資シミュレーションを実施した。その結果、ネット証券のような 0.1% 程度の取引手数料であれば市場平均を上回る収益を得られる可能性があり、効率的市場仮説の反証となりえる。この結果は日本および米国の 2 大市場において観測されたため、株式市場の特性として汎用的であると考えられる。

●プロフィール
鈴木 智也 MFTA®

新潟県新潟市生まれ。IFTA 国際検定テクニカルアナリスト（MFTA）。平成 17 年東京理科大学大学院理学研究科物理学専攻博士課程修了。理学博士。同年東京電機大学工学部電子工学科助手、平成 18 年より同志社大学工学部情報システムデザイン学科専任講師、平成 21 年より茨城大学工学部知能システム工学科准教授を経て、平成 28 年より同大学教授。平成 30 年より茨城大学大学院理工学研究科機械システム工学専攻長および領域長、現在に至る。さらに平成 29 年より大和証券投資信託委託（株）クウォンツ運用部特任主席研究員を兼務。研究分野は、時系列解析、テキスト解析、機械学習、人工知能、金融工学など実践的なデータサイエンスに従事。電子情報通信学会、情報処理学会、人工知能学会、日本テクニカルアナリスト協会、日本証券アナリスト協会 各会員。

＜参考文献＞

1） T. Hastie, R. Tibshirani, J. Friedman (2009): The Elements of Statistical Learning: Data Mining, Inference, and Prediction, (Springer)
2） D. Silver, et. al. (2016): Mastering the Game of Go with Deep Neural Networks and Tree Search, Nature, vol.529, pp.484–489.
3） 西垣 通（2013）：集合知とは何か、（中央公論新社）
4） ジェームズ・スロウィッキー（2009）：「みんなの意見」は案外正しい、（角川書店）
5） Yahoo! Finance Japan: http://finance.yahoo.co.jp/
6） Yahoo! Finance: http://finance.yahoo.com/
7） Z. H. Zhou（著）、宮岡 悦良（翻訳）、下川 朝有（翻訳）（2017）：アンサンブル法による機械学習、（近代科学社）
8） L. Breiman (1996): Bagging Predictors, Mach. Learn, vol.24, pp.123–140.
9） T. Suzuki and K. Nakata (2014): Risk Reduction for Nonlinear Prediction and its Application to the Surrogate Data Test, Physica D, vol.266, no.1, pp.1–12.
10） T. Suzuki, Y. Ohkura (2016): Financial Technical Indicator Based on Chaotic Bagging Predictors for Adaptive Stock Selection in Japanese and American Markets, Physica A, vol.442, pp.50–66.
11） コンセンサスレシオに応じてポートフォリオを組む方法も考えられる。
12） 本稿は AI 技術として機械学習を用いるものの、過去の株価変動のみに基づいて投資判断するため、テクニカル分析の範疇と考える。
13） D. Aronson (2006): Evidence-Based Technical Analysis: Applying the Scientific Method and Statistical Inference to Trading Signals, (John Wiley & Sons)