

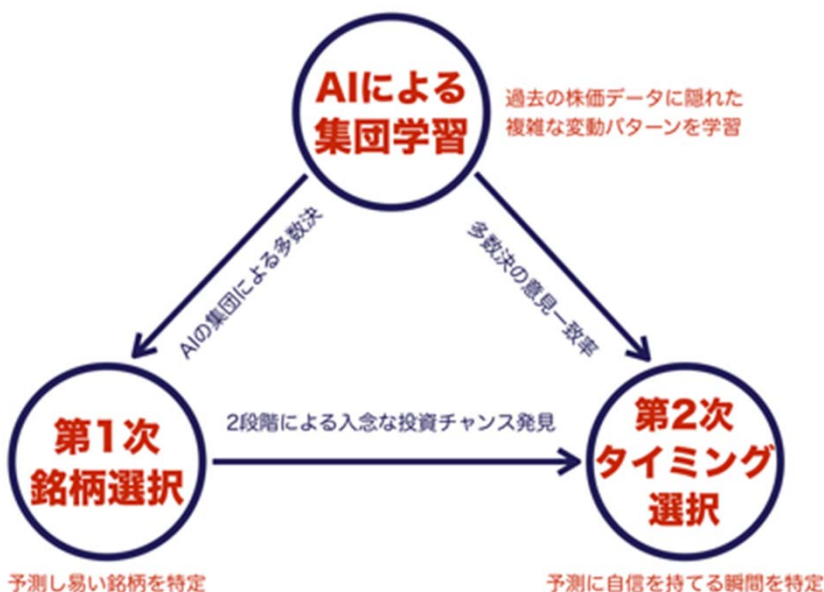
## 【鈴木教授による解説シリーズ13】 集団化する人工知能 集合知AIモデルのシミュレーション(後編)

2019年 9月 6日

### お伝えしたいポイント

- ・ 前回から引き続き、集合知AIモデルが機能するメカニズムを解説します。
- ・ 「第1次銘柄選択」 および「第2次タイミング選択」の効果を検証します。
- ・ シミュレーションによって原著論文の正答率(約60%~80%)を再現します。
- ・ 集合知AIモデルの妥当性を総合的に検証します。

集合知AIモデルは「集団学習・第1次銘柄選択・第2次タイミング選択」の工程に分かれており、前回において集団学習と第1次銘柄選択のメカニズムについて解説しました。今回は、本モデルの核心である第1次銘柄選択および第2次タイミング選択で得られる効果についてシミュレーション実験を行います。なお、集合知AIモデルの原著論文「Consensus Ratio and Two-steps Selection to Detect Profitable Stocks」は、国際テクニカルアナリスト連盟 (IFTA) ホームページ ([https://ifta.org/public/files/journal/d\\_ifta\\_journal\\_18.pdf](https://ifta.org/public/files/journal/d_ifta_journal_18.pdf)) より無料で入手できます。



注：

左図は、集合知AIモデルのコンセプト：  
解説シリーズ第1回 (AI運用に挑む)  
[https://www.daiwa-am.co.jp/specialreport/quants/2017/quantst\\_20171207.html](https://www.daiwa-am.co.jp/specialreport/quants/2017/quantst_20171207.html)

を一部変更して再掲しています。

(出所) 大和証券投資信託委託

## シミュレーション実験の枠組み

原著論文では、実際の株価データに対して集合知AIモデルを適用し、株価変動方向の予測において約60%~80%の正答率を得ました。モデル構築においては集合知の統計的性質を考慮しましたが、モデルの妥当性については運用成績の結果から帰納的に確認したに過ぎません。より深い理解として、株価変動のどのような性質を捉えて高い正答率を実現できたのか説明するには、予測可能性の源泉となる株価の法則性についてシンプルな仮説を用いてモデル化し、そのモデルを通じて演繹的に集合知AIモデルの動作原理を確認する作業が必要です。

### 株価変動の予測しやすさとAIエージェントのモデル化

株価変動の予測しやすさについて、以下のような仮説を立てます。

- 投資できる株式銘柄は充分に多く、それぞれを  $s = 1, 2, \dots$  のように表す。
- 株価変動には弱い法則性  $R_s(t)$  が存在し、 $R_s(t)$  が大きいほど株価変動を予測しやすい。  
(つまり  $R_s(t)$  は株価変動の予測しやすさを意味し、株価変動自体ではありません。)
- 法則性  $R_s(t)$  は、銘柄  $s$  固有の法則性  $E[R_s]$  とタイミング効果  $T_s(t)$  によって構成される。

$$R_s(t) = E[R_s] + T_s(t)$$

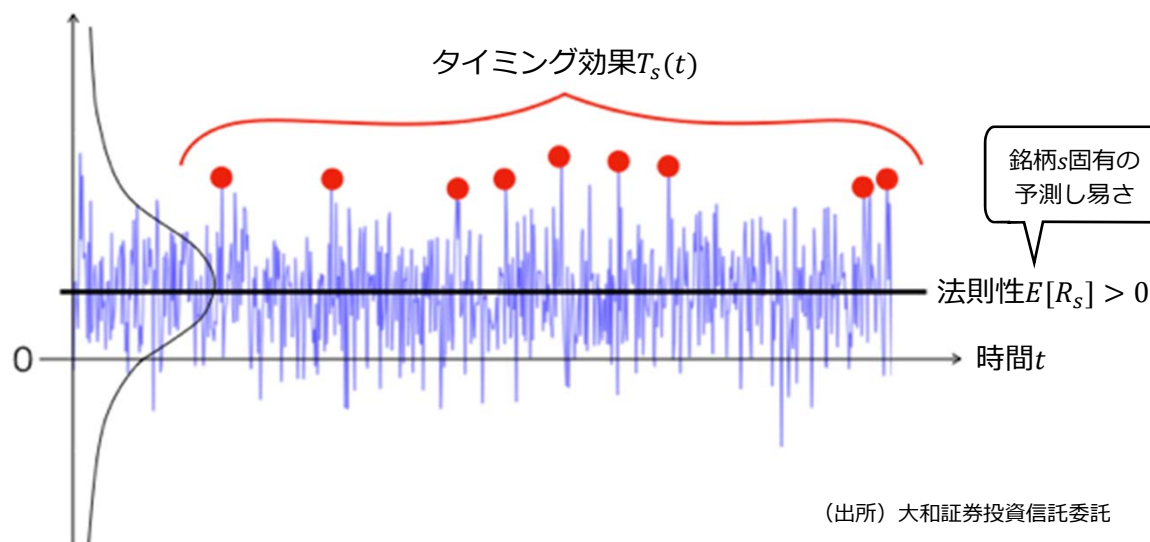
- $T_s(t) > 0$  なら法則性は強まり、 $T_s(t) < 0$  なら法則性は弱まる。
- パラメータ設定(※)として、 $E[R_s]$  は0%~10%の一様乱数、 $T_s(t)$  は平均0%および標準偏差  $\sigma_T$  % の正規乱数とする。なお、 $\sigma_T = 5$  とする。

$$E[R_s] \sim U(0, 10)$$

$$T_s(t) \sim N(0, \sigma_T)$$

銘柄  $s$  の予測しやすさ (法則性の度合い)

$$R_s(t) = E[R_s] + T_s(t)$$



AIエージェントについて、以下のような仮説を立てます。

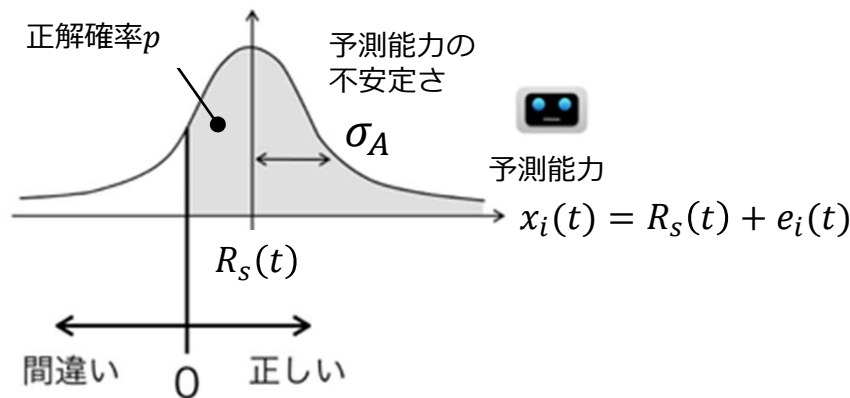
- 投資判断するAIエージェントは充分に多く、それぞれを  $i = 1, 2, \dots$  のように表す。
- 各AIエージェントの予測能力を  $x_i(t)$  とし、 $x_i(t) > 0$  ならば株価変動方向を正しく予測できる。
- AIエージェントは法則性  $R_s(t)$  を頼りに株価変動を予測するが、非法則的な変動要因によって予測能力  $x_i(t)$  は運にも左右される。この不安定さを  $e_i(t)$  とする。

$$x_i(t) = R_s(t) + e_i(t)$$

- $e_i(t) < 0$  ならば不運による予測能力の低下、 $e_i(t) > 0$  ならば幸運 (まぐれ) による予測能力の増加とする。
- $e_i(t)$  は、AIエージェントの個性 (予測方法や参照データの違い) に起因する。
- パラメータ設定※として、 $e_i(t)$  は平均0%および標準偏差  $\sigma_A$  %の正規乱数とする。なお、 $\sigma_A = 50$  とする。

$$e_i(t) \sim N(0, \sigma_A)$$

- $x_i(t) > 0$  となる確率を計算することで、予測が正しい確率 (正解確率)  $p$  を得る。
- 前回紹介した集団学習 (バギング) によって多数のAIエージェントを構築するが、復元されたデータセット間の重複に伴い、実質的に独立な個体数  $n$  は構築されたエージェント数に比べて極めて少ない。



(出所) 大和証券投資信託委託

※本セクションで注意したこと

各パラメータは、集合知AIモデルの正答率が現実の値に近くなる (高くなりすぎない) ように設定しました。法則性  $R_s(t)$  のうち、基本的に内在する法則性  $E[R_s]$  は0%~10%に留め、その他5%のタイミング効果  $T_s(t)$  によってランダムに変動するようにしました。一方、AIエージェントは50%もの不安定さを持つので、株価予測は「運の要素が非常に大きい」ことを前提としています。

第1次銘柄選択のシミュレーション

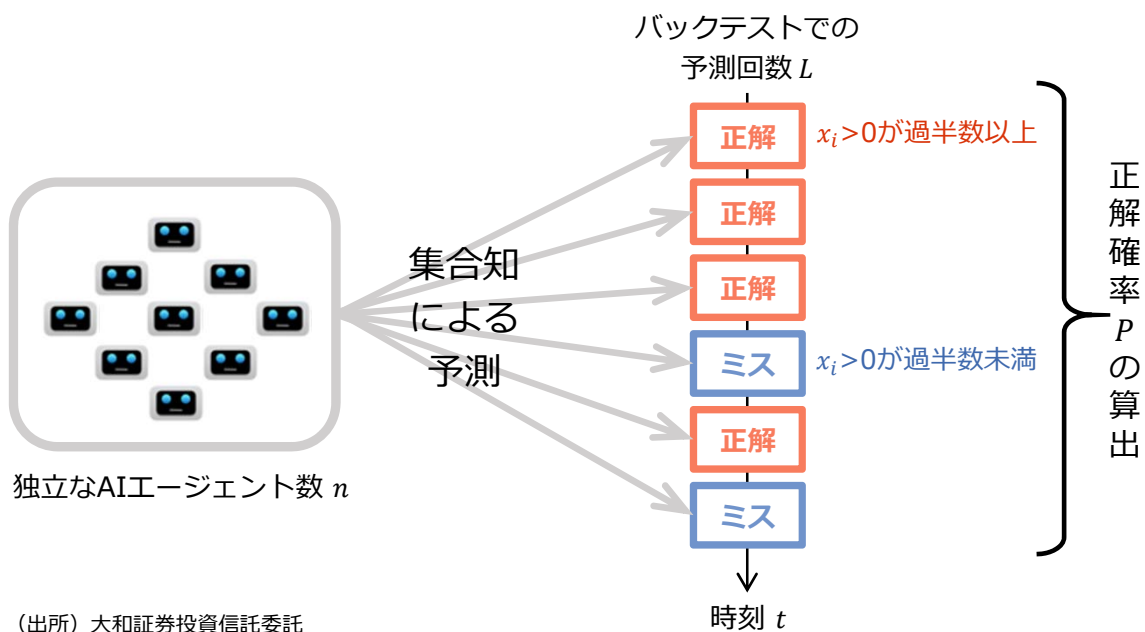
株価銘柄 $s$ 毎に多数のAI (原著論文では500体) によって変動方向を予測し、 $x_i > 0$ を満たすAIエージェントが過半数を越えれば、集合知 (多数決) による予測は正しいとします。バックテストにおいて異なる時刻 $t$ にて上記の予測を $L$ 回を行い、集合知による正解確率 $P$ を求めます。

投資対象の銘柄数を1000とし、第1次銘柄選択によって正解確率 $P$ が高い上位30%の銘柄 $s$ を選択します。前回の考察が正しいならば、集合知によって予測しづらい銘柄 ( $E[R_s] \approx 0$ ) と予測しやすい銘柄 ( $E[R_s] \gg 0$ ) を分離し易くなるため、独立なAIエージェント数 $n$ が多いほど集合知を発揮し、第1次銘柄選択によって $E[R_s]$ が高い銘柄が選ばれるはずで、この是非をシミュレーションによって検証します。

さらに、バックテストにおける予測回数 $L$ が多いほど、正解確率 $P$ の信頼性が高まります。しかし株式市場はイベント等の外乱によって内部構造が変化する可能性があるため、大きな $L$ を用いるのは得策ではありません。また小さな $L$ も統計的な信頼性が低下します。そこで1週間の営業日をおおよそ5日間とみなし、 $L = 40$  (2ヶ月)・ $L = 120$  (6ヶ月)・ $L = 480$  (2年) の3通りでシミュレーションしました。なお集団学習における訓練データの重複に伴い、実質的に独立なAIエージェント数は $n = 19$ 程度とします。

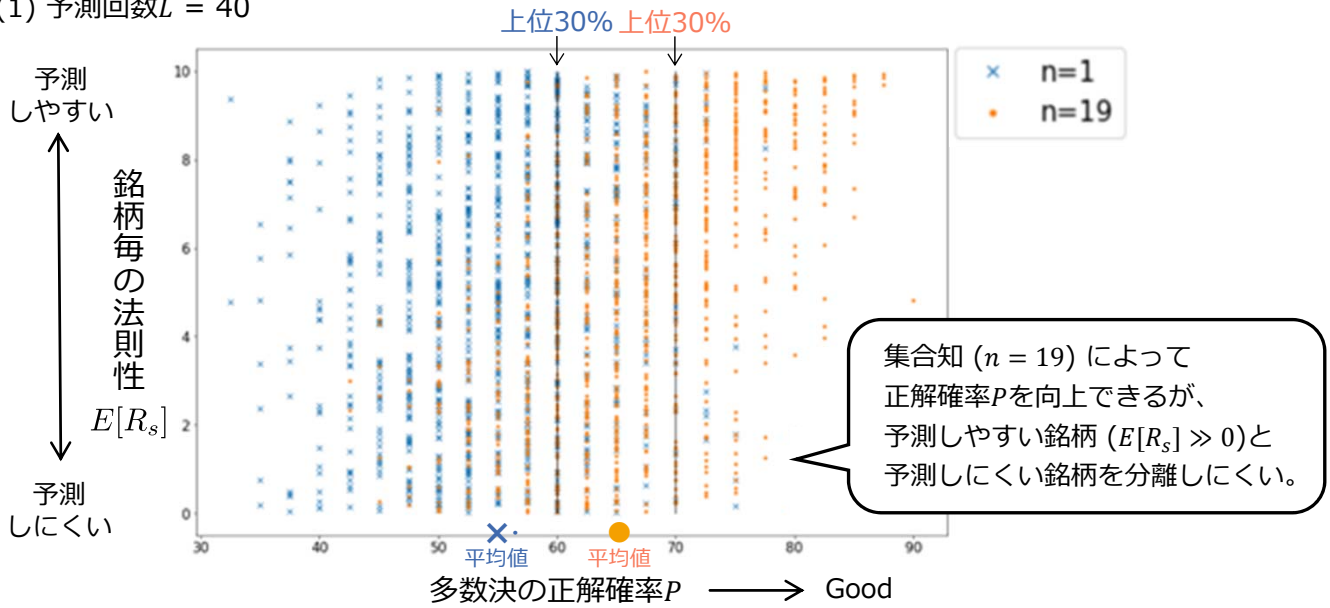
シミュレーション結果を次ページに示します。結果として、次のことが言えます。

- 集合知によって多数決の正解確率 $P$ は向上する。
- 予測回数 $L$ が多いほど、集合知による第1次銘柄選択によって、 $E[R_s]$ が高い銘柄 $s$ を特定できる。
- 予測回数 $L$ が少ないほど正解確率 $P$ が安定しないため、第1次銘柄選択の効果は低下する。
- 予測回数 $L = 120$ かつAIエージェント数 $n = 19$ でも、第1次銘柄選択による効果を確認できる。

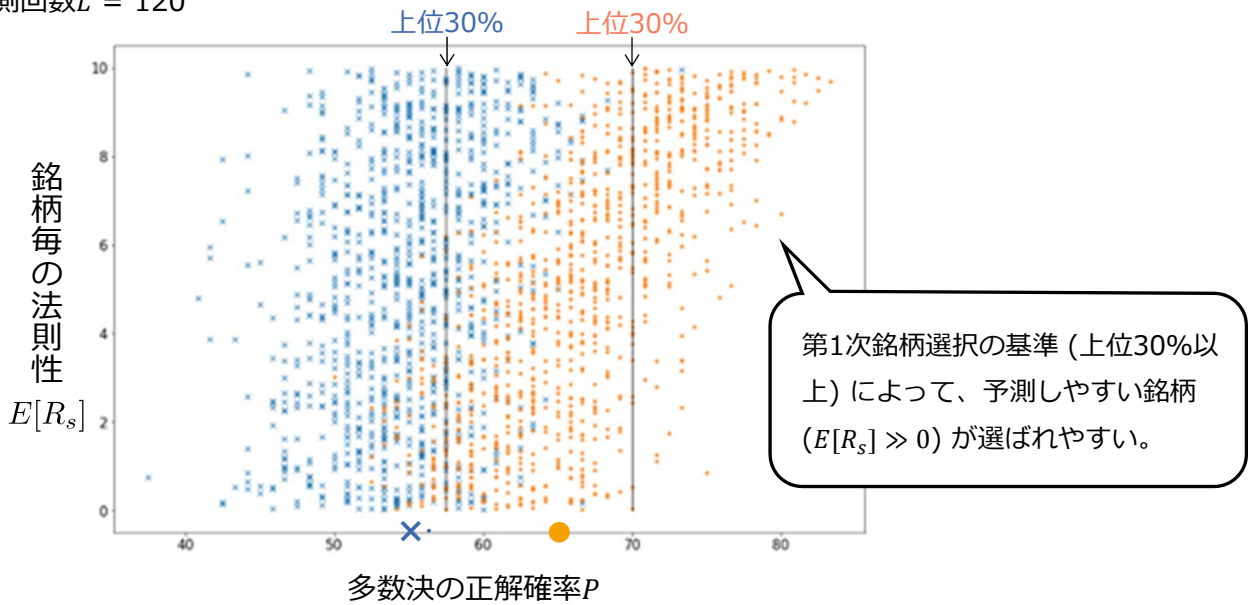




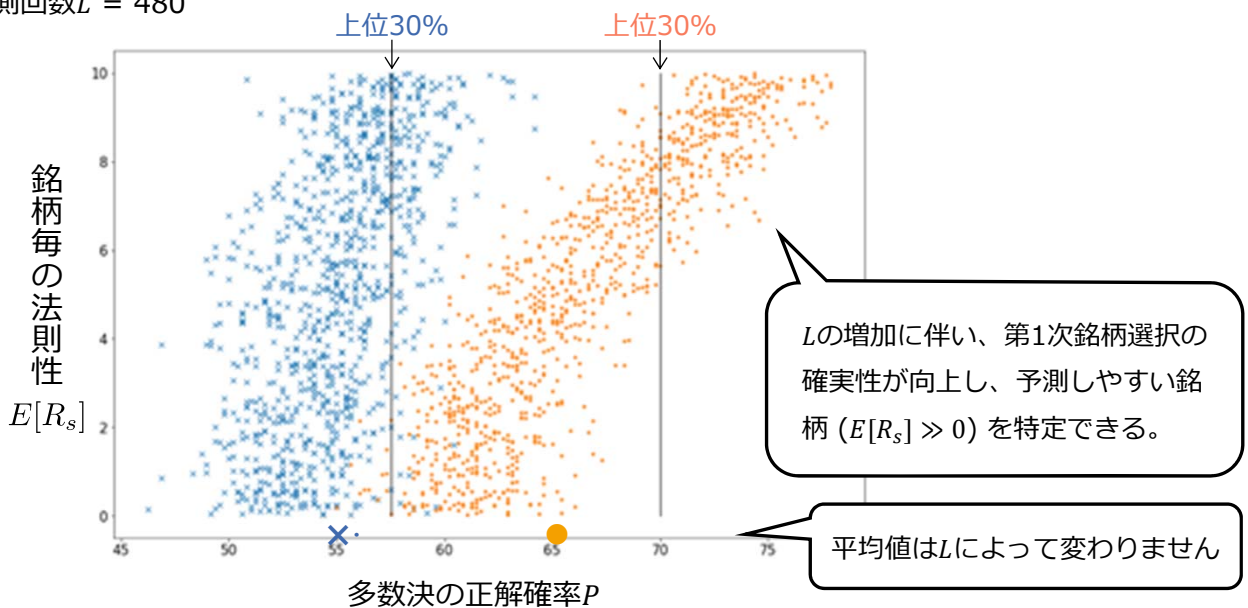
(1) 予測回数  $L = 40$



(2) 予測回数  $L = 120$



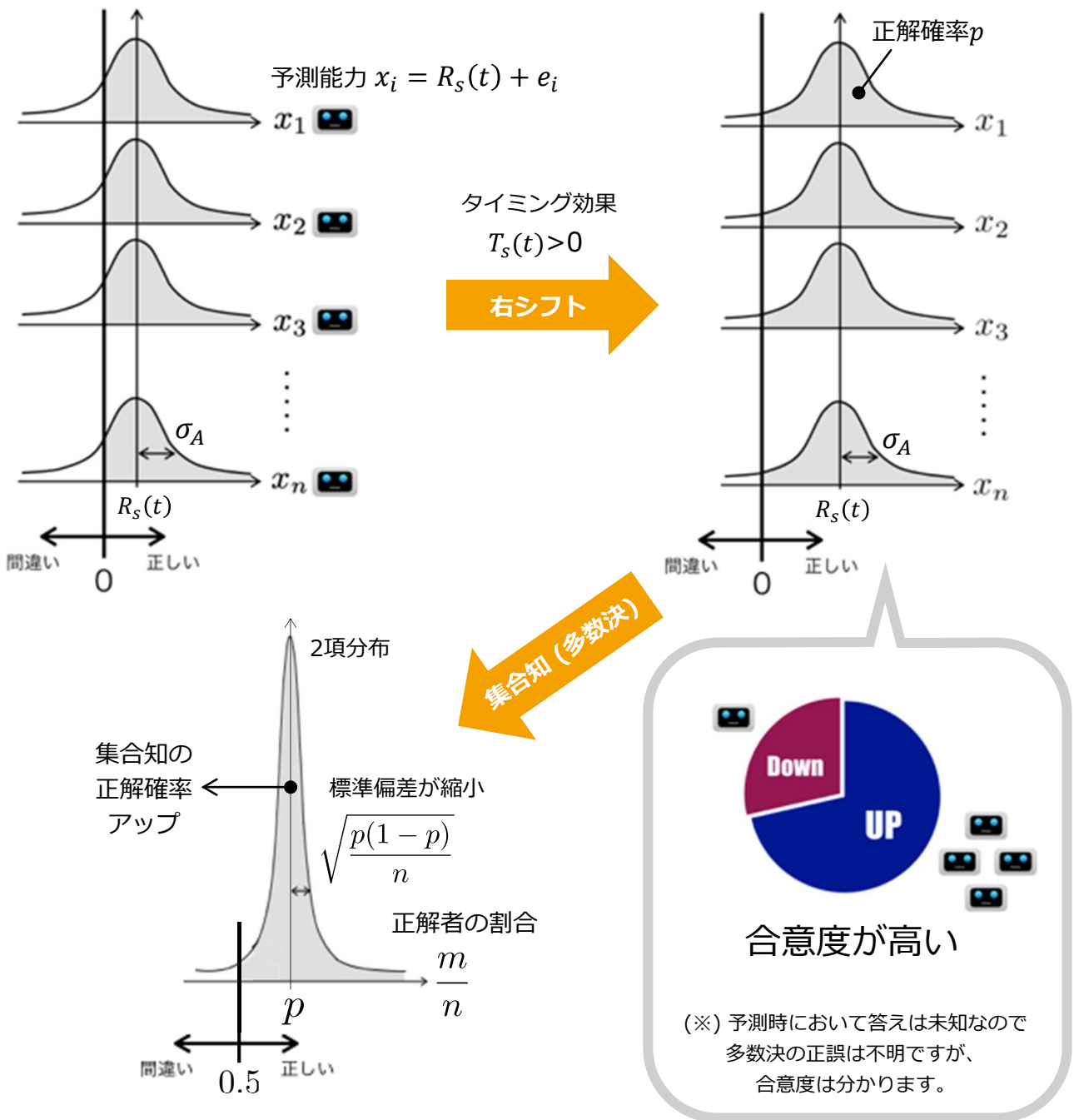
(3) 予測回数  $L = 480$



(出所) 大和証券投資信託委託

第2次タイミング選択のシミュレーション

第1次銘柄選択によって選ばれた上位30%の有望銘柄に対して運用を始めるとします。その際には先に仮定したタイミング効果 $T_s(t)$ によって、各銘柄の法則性 $R_s(t)$  (予測しやすさ) は揺らぎます。もし $T_s(t) > 0$ ならば $R_s(t)$ は大きくなるため、各AIエージェントの予測能力 $x_i$ の分布は右方向にシフトします。すると正解確率 $p$ は大きくなるため、正解するAIエージェントが増えると共に、それぞれの予測結果の一致率 (集合知の合意度) が高くなります。正解確率 $p$ が大きいほど集合知 (多数決) によって更に正解確率を高められるため、このようなGoodタイミングを特定することが第2次タイミング選択の狙いです。

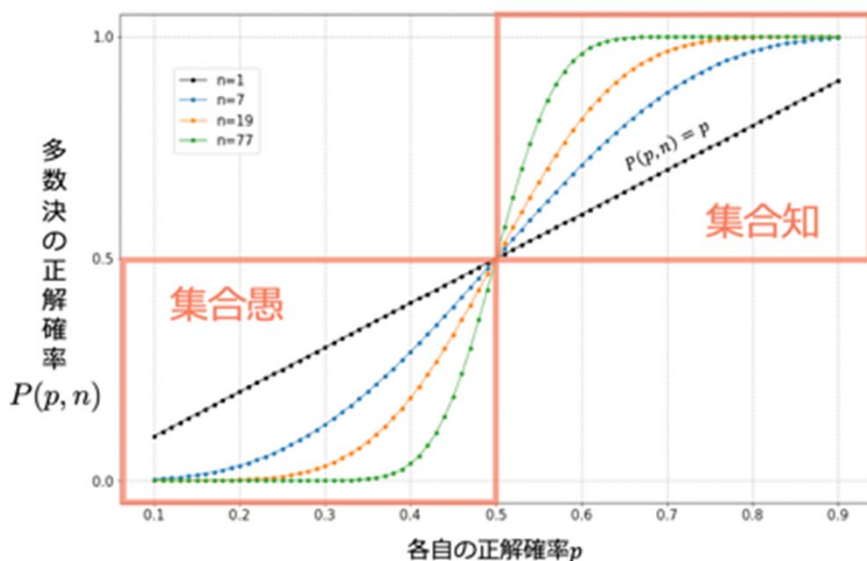


(出所) 大和証券投資信託委託

予測時において多数決の正誤は不明ですが、合意度の計算は可能です。そこで、以下の計算式により多数決の合意度を求め、タイミング効果 $T_s(t)$ が大きな銘柄 $s$ およびタイミング $t$ を特定します。これを「第2次タイミング選択」と呼ぶことにします。

$$\text{多数決の合意度} = \frac{\text{多数派に属するAIエージェント数}}{\text{AIエージェントの総数}}$$

ただし先述の株価変動モデルにおいて、悪いタイミング効果 ( $T_s(t) < 0$ ) の存在も仮定しています。 $T_s(t) < 0$ ならば、各AIエージェントの予測能力 $x_i$ の分布は左方向にシフトし、正解確率 $p$ は低下します。さらに法則性 $R_s(t) < 0$ になれば、正解確率 $p < 0.5$ になるため、多数決による判断は集合知ではなく「集合愚」となります。(集合愚とは、多数決によって判断能力が低下する現象。)

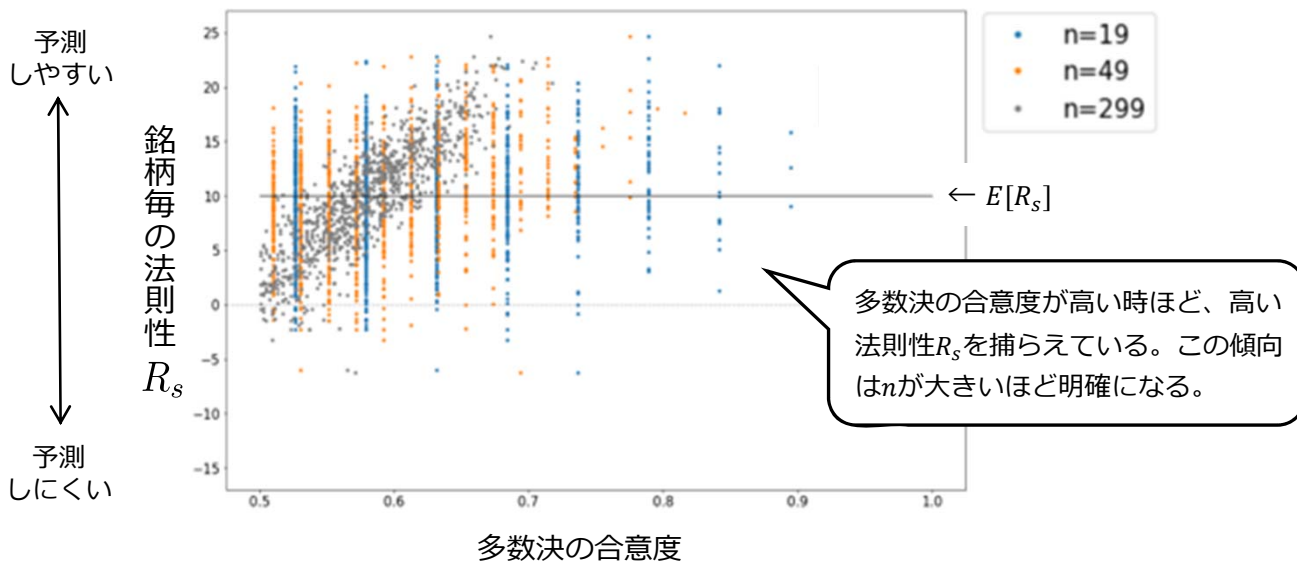


前回の図を再掲。  $p < 0.5$ では集合愚となるため、多数決によって正解確率が低下します。

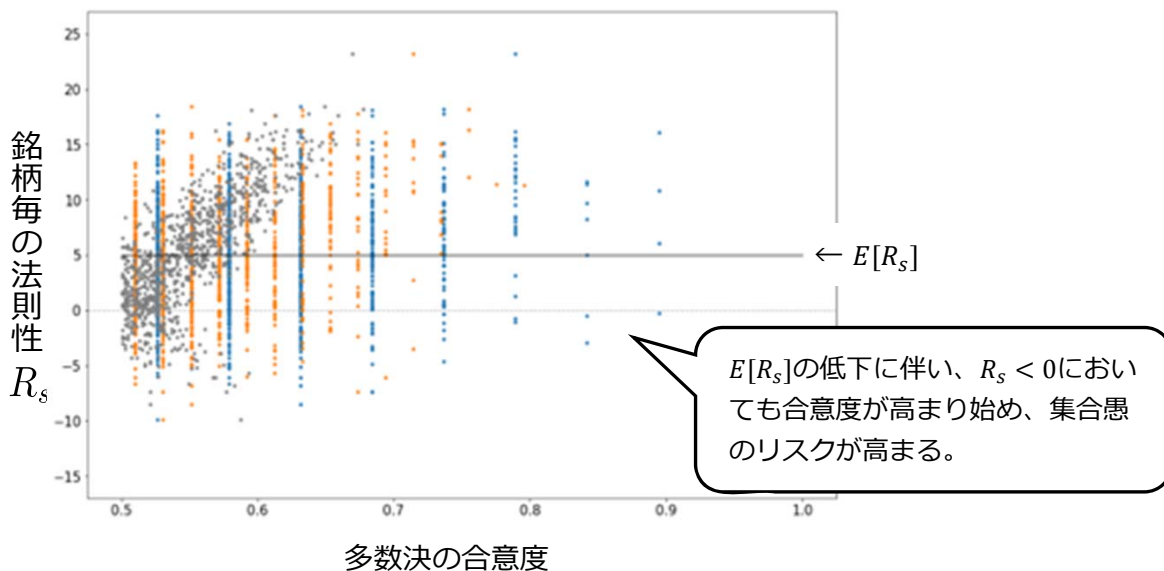
(出所) 大和証券投資信託委託

AIは過去の経験に基づいて意思決定するため、過去の経験に逆らう現象 ( $R_s(t) < 0$ ) については反対に予測します。この場合、多数決の合意度も反対の予測に集中するため、第2次タイミング選択はむしろ悪影響を及ぼします。しかし前処理として第1次銘柄選択によって $E[R_s]$ が高い銘柄に限定すれば $R_s(t) < 0$ になりにくいため、集合愚を選ぶリスクを低減できます。そこで $E[R_s]$ の値を変えながら第2次タイミング選択の効果を検証しました。その結果を次ページに示します。確かに $E[R_s]$ が大きい銘柄ほど、多数決の合意度によって高い法則性 $R_s$ および集合知を抽出できます。

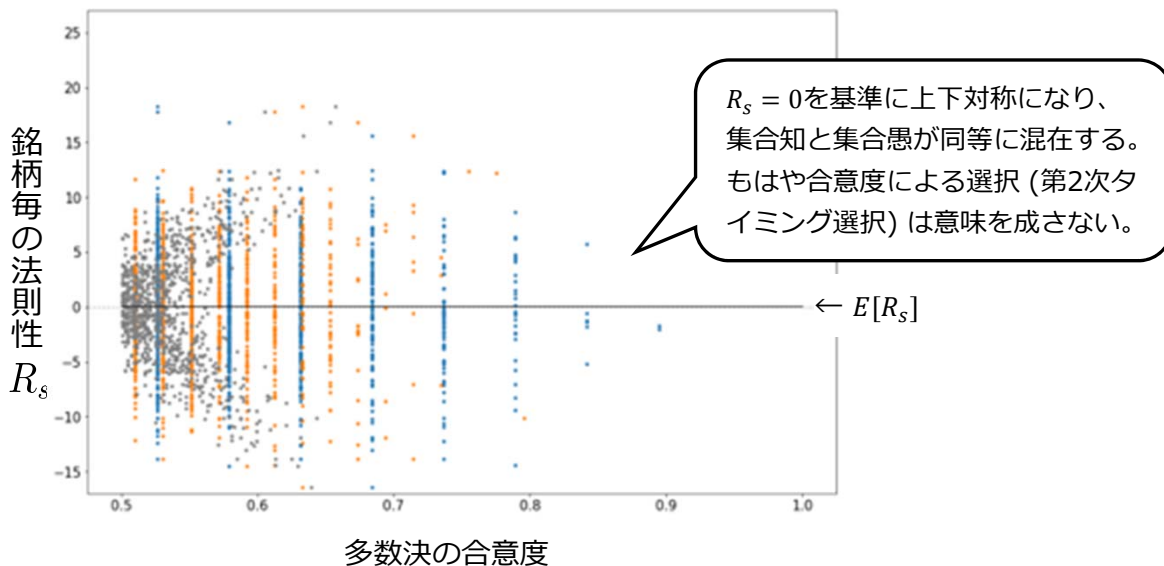
(1)  $E[R_s] = 10\%$  の場合



(2)  $E[R_s] = 5\%$  の場合



(3)  $E[R_s] = 0\%$  の場合





## 集合知AIモデル (第1次選択+第2次選択+集合知) のシミュレーション

### 各時刻で選ばれる投資対象銘柄の法則性

$L$ 回のテスト予測を通じて第1次銘柄選択を行った後、運用を実施する時刻を $t = 1 \sim 1000$ とします。この運用期において、どのような法則性 $R_s(t)$ が選択されるかシミュレーションします。その際に、それぞれの選択法の効果を検証するため、下記4通りの方法を比較します。

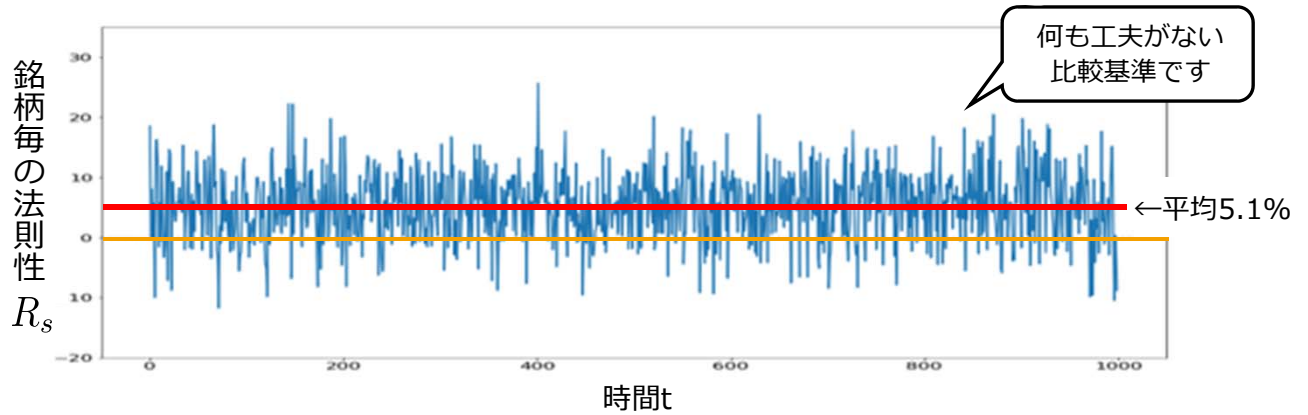
1. **ランダム選択**：全銘柄からランダムに $R_s(t)$ を選択 (何も工夫が無い比較基準)。
2. **第1次銘柄選択のみ**：第1次選択 (上位30%) を通過した有望銘柄からランダムに $R_s(t)$ を選択。
3. **第2次タイミング選択のみ**：全銘柄から集合知の合意度が最も高い $R_s(t)$ を選択。
4. **第1次選択+第2次選択**：第1次選択 (上位30%) を通過した有望銘柄から、集合知の合意度が最も高い $R_s(t)$ を選択。

なお、現実的な状況を見据えて「 $L = 120$  かつ  $n = 19$ 」に設定してシミュレーションしました。結果を次ページに示します。要点は以下のとおりです。

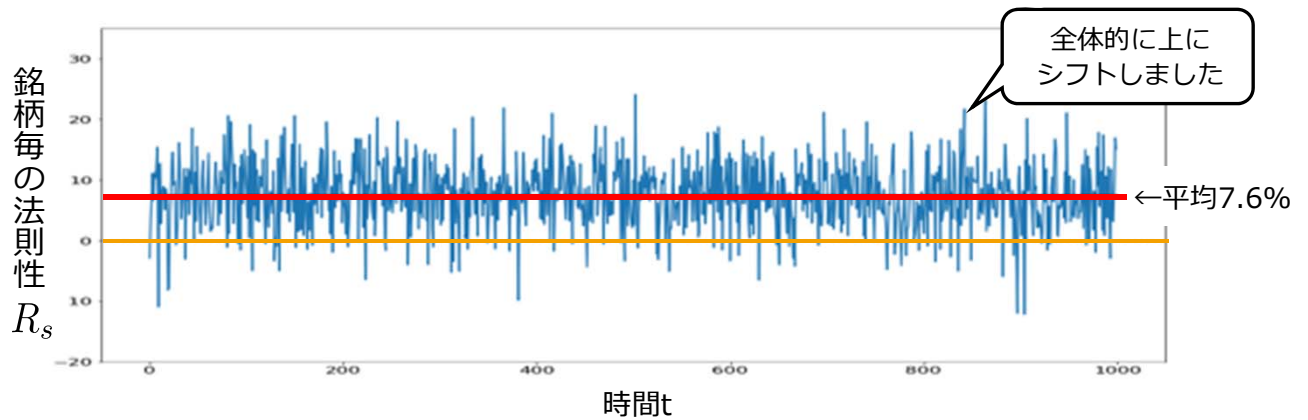
- 第1次銘柄選択および第2次タイミング選択は、単独でも法則性 $R_s(t)$ が高い銘柄 $s$ やタイミング $t$ を選択できる。
- ただし第2次タイミング選択のみでは、集合愚も多く発生する。
- 第1次銘柄選択も併用することで、第2次タイミング選択による集合愚を減らすことができる。

したがって、第1次銘柄選択および第2次タイミング選択を併用することで、集合知AIモデルとしての機能を発揮できると言えます。

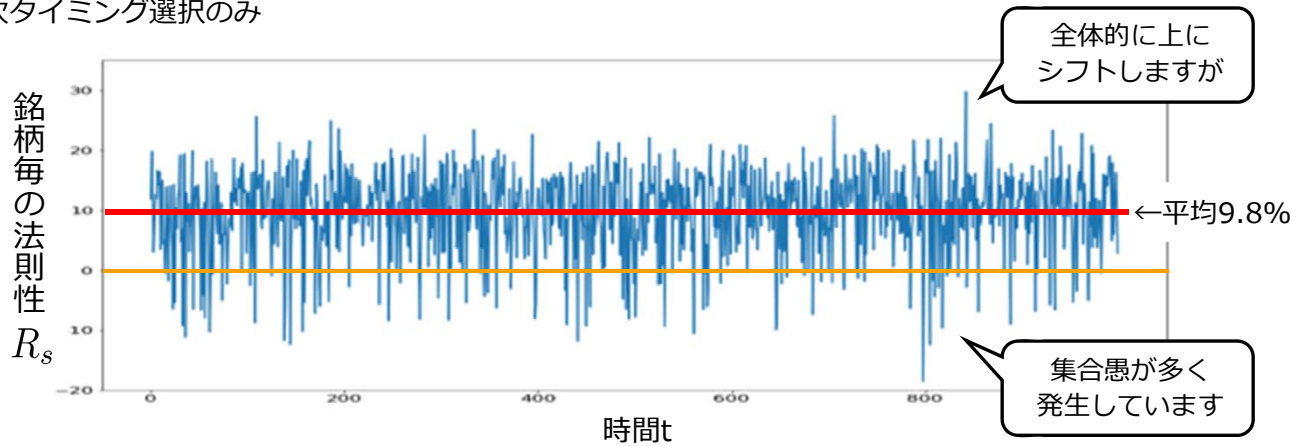
(1) ランダム選択



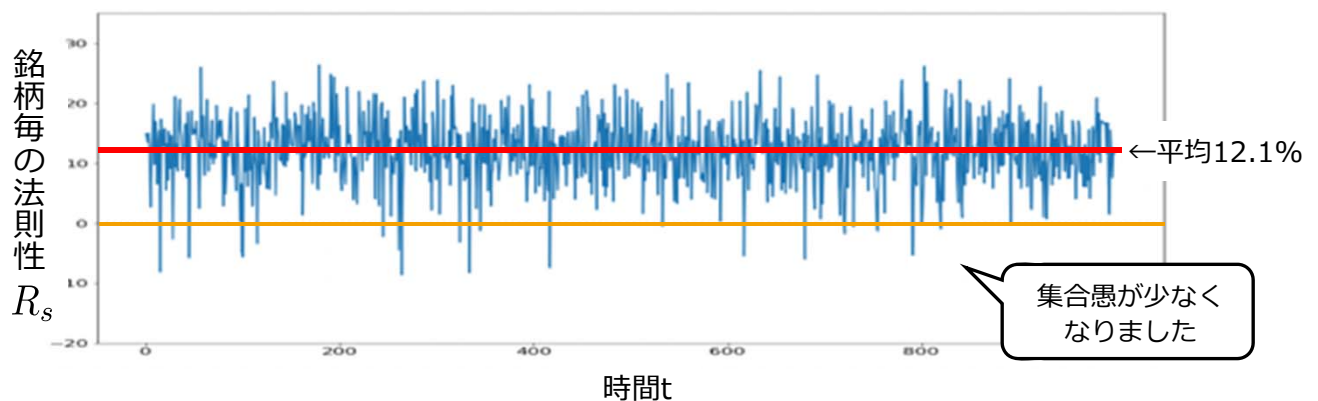
(2) 第1次銘柄選択のみ



(3) 第2次タイミング選択のみ



(4) 第1次選択+第2次選択



## 集合知による正答率の向上

次に、前章で得た $R_s(t)$ を通じてAIエージェントの正答率を求めます。さらに集合知によってどの程度まで正答率を向上できるか検証します。

まず、前章のシミュレーション ( $L = 120$  かつ  $n = 19$ ) によって $R_s(t)$ を決定し、AIエージェントの予測能力 $x_i = R_s(t) + e_i$ を乱数 $e_i$ によって定めます ( $e_i \sim N(0,50)$ )。ここで $x_i > 0$ ならば予測に正解できるとします。集合知を用いない場合は単体 ( $i = 1$ ) による予測ですが、集合知を用いる場合は $i = 1 \sim n$ とします。そして $x_i > 0$ を満たすAIエージェント数 $m$ を求め、 $m/n > 0.5$ ならば予測に正解できるとします。

以上を時刻 $t = 1 \sim 1000$ において実施し、正答率を算出しました。結果を下表に示します。要点として、次のことが言えます。

- 第1次銘柄選択および第2次タイミング選択は、単独でも正答率を向上できる。
- 両方の選択法を併用すると、正答率をより向上できる。
- いずれの場合も、集合知 (多数決) によって、さらに正答率を向上できる。

表中の「第1次選択+第2次選択+集合知あり」は、集合知AIモデルの最終形です。この正答率は80%に近づき、原著論文の結果を概ね再現できました。そこで、株価変動モデルの観点から集合知AIモデルが機能するメカニズムを考察すると、

- 第1次銘柄選択によって、法則性 $E[R_s]$ が高い銘柄を選べる。
- 第2次タイミング選択によって、タイミング効果 $T_s(t)$ が高い銘柄を選べる。
- ただし第2次タイミング選択のみでは集合愚を選ぶ可能性があるため、事前のバックテストを通じて第1次銘柄選択を適用し、 $R_s$ が下がりにくい ( $E[R_s]$ が高い) 銘柄を選ぶ必要がある。
- $R_s > 0$ であれば集合知を発揮できるため、多数決によって更に正答率を向上できる。

と言えます。

	ランダム選択	第1次選択のみ	第2次選択のみ	第1次選択+第2次選択
集合知 なし	51.7%※1	54.2%	57.7%※2	60.8%
集合知 あり	62.5%	71.7%	75.1%	80.7%

(出所) 大和証券投資信託委託

※1 0%~10%の法則性 $E[R_s]$ を仮定しているため、ランダム選択でも50%以上の正答率になります。なお投資対象とする銘柄の選択がランダムであり、売買の選択がランダムという意味ではありません。

※2 第2次選択においては集団学習を用いますが、最終的な投資判断は集合知 (多数決) ではなく、単独のAIエージェントに従った場合です (6ページの図を参照)。

### その他の考察

本稿では原著論文の結果を再現するように、現実を見据えたパラメータを設定しましたが、シミュレーションの利点として自由にパラメータを変えることができます。そこでパラメータを変えながら、様々なシミュレーションを実施した結果を簡単に述べます。もし予測対象が株価と似た特徴を持つならば、その際の指針になると思います。

- バックテストにおいて予測回数 $L$ を増やすほど、第1次銘柄選択において高い $E[R_s]$ を持つ銘柄 ( $R_s$ が下がりにくい銘柄) を選択できる。その結果、第2次タイミング選択において集合愚を選ぶリスクを低減できる。⇒ 次ページの図(1)\*

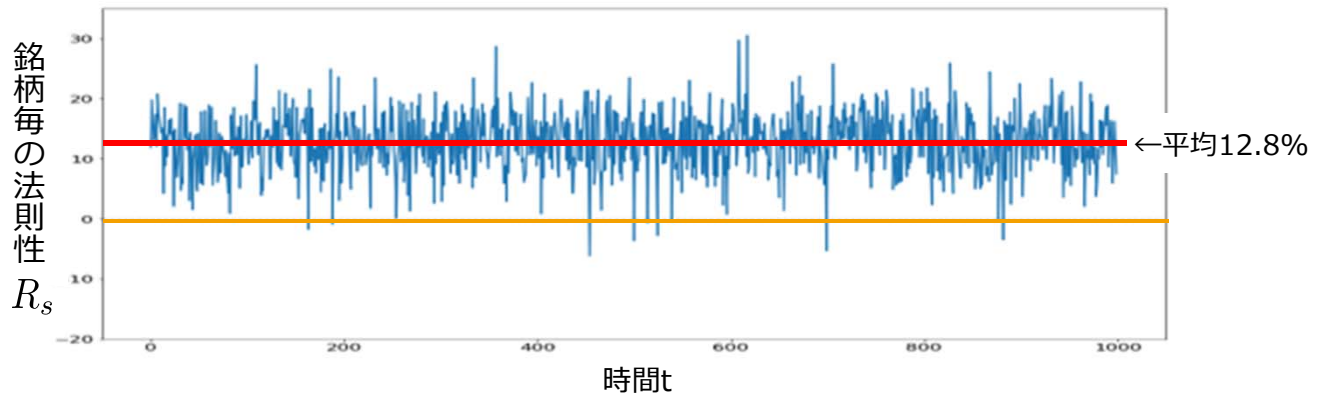
(※10ページの図(4)と比較してください。次ページの図(2)も同様です。)

- AIエージェント同士の独立性が高まるほど、エージェント数 $n$ を増やす効果がある。この $n$ が大きいほど第1次銘柄選択が向上するが、大きな集合愚を拾ってしまう場合もある。⇒ 次ページの図(2)\*
- タイミング効果の標準偏差 $\sigma_T$ が小さいほど、集合知AIモデルは第1次銘柄選択のみの結果に近づく。
- タイミング効果の標準偏差 $\sigma_T$ が大きいほど、集合知AIモデルは第2次タイミング選択のみの結果に近づく。

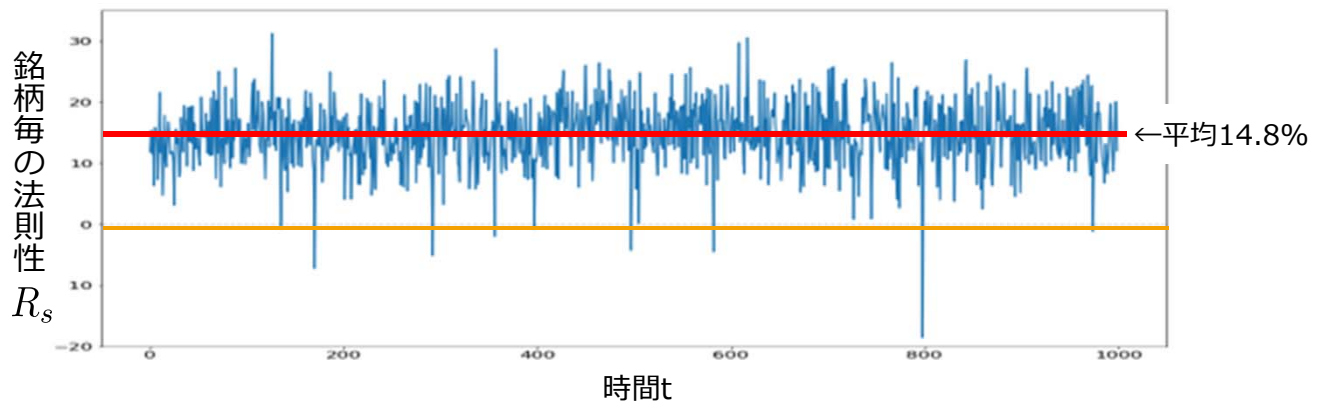
次ページに「第1次選択+第2次選択」によって選ばれた法則性 $R_s(t)$ について、パラメータ $L$ と $n$ を変えた場合を示します。先述の「平均12.1%」より徐々に向上していることを確認できます。



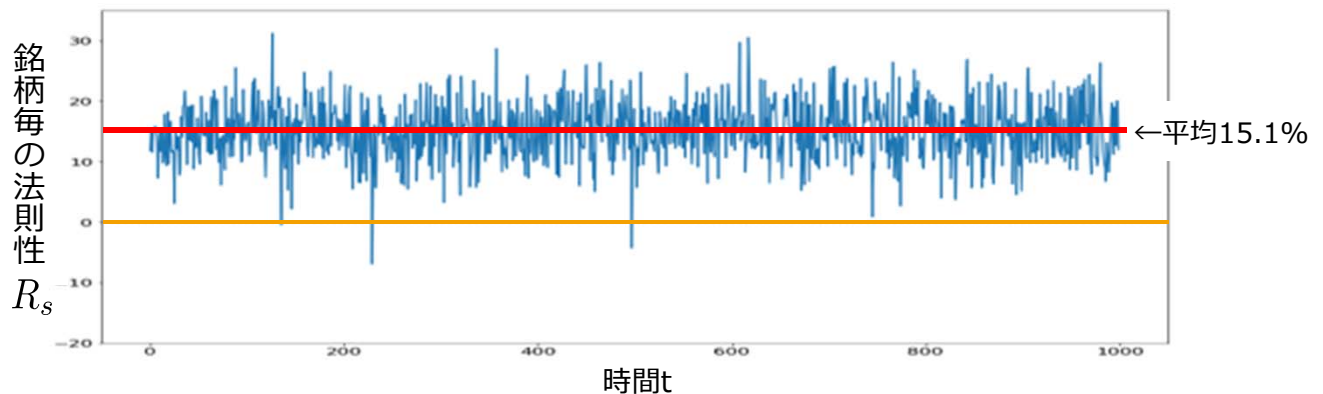
(1) 予測回数 $L$ を増やした場合 ( $L = 480, n = 19$ )



(2) 独立なエージェント数 $n$ を増やした場合 ( $L = 120, n = 49$ )



(3) 両方を増やした場合 ( $L = 480, n = 49$ )



(出所) 大和証券投資信託委託

## 鈴木教授 プロフィール

### 鈴木智也 (すずきともや)

茨城大学大学院理工学研究科機械システム工学領域教授。理学博士。IFTA国際検定テクニカルアナリスト。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，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。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，平成30年よりCollabWiz(株)代表取締役を兼務。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，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。

## バックナンバー

### 鈴木教授による解説シリーズ

- **001 「AI運用に挑む」**  
[https://www.daiwa-am.co.jp/specialreport/quants/2017/quantst\\_20171207.html](https://www.daiwa-am.co.jp/specialreport/quants/2017/quantst_20171207.html)
- **002 「集団化する人工知能」**  
[https://www.daiwa-am.co.jp/specialreport/quants/2018/quantst\\_20180125.html](https://www.daiwa-am.co.jp/specialreport/quants/2018/quantst_20180125.html)
- **003 「2年目のジンクスを集合知AIで緩和」**  
[https://www.daiwa-am.co.jp/specialreport/quants/2018/quantst\\_20180301.html](https://www.daiwa-am.co.jp/specialreport/quants/2018/quantst_20180301.html)
- **004 「時系列データの見えない法則をつかむ」**  
[https://www.daiwa-am.co.jp/specialreport/quants/2018/quantst\\_20180409.html](https://www.daiwa-am.co.jp/specialreport/quants/2018/quantst_20180409.html)
- **005 「愚かな人間心理・カモにするAI」**  
[https://www.daiwa-am.co.jp/specialreport/quants/2018/quantst\\_20180501.html](https://www.daiwa-am.co.jp/specialreport/quants/2018/quantst_20180501.html)
- **006 「ナイトメア★アノマリーを狙え」**  
[https://www.daiwa-am.co.jp/specialreport/quants/2018/quantst\\_20180601.html](https://www.daiwa-am.co.jp/specialreport/quants/2018/quantst_20180601.html)
- **007 「ブルーオーシャンAI戦略」**  
[https://www.daiwa-am.co.jp/specialreport/quants/2018/quantst\\_20180703.html](https://www.daiwa-am.co.jp/specialreport/quants/2018/quantst_20180703.html)
- **008 「深層学習による株価予測 (前編)」**  
[https://www.daiwa-am.co.jp/specialreport/quants/2018/quantst\\_20180802.html](https://www.daiwa-am.co.jp/specialreport/quants/2018/quantst_20180802.html)
- **009 「深層学習による株価予測 (後編)」**  
[https://www.daiwa-am.co.jp/specialreport/quants/2018/quantst\\_20181001.html](https://www.daiwa-am.co.jp/specialreport/quants/2018/quantst_20181001.html)
- **010 「ニュースを読んで投資判断する集合知AI」**  
<https://www.daiwa-am.co.jp/specialreport/quants/2018/1204.html>
- **011 「投資理論とコンピュータの歴史」**  
[https://www.daiwa-am.co.jp/specialreport/quants/20190701\\_01.html](https://www.daiwa-am.co.jp/specialreport/quants/20190701_01.html)
- **012 「集合知AIモデルのシミュレーション(前編)」**  
[https://www.daiwa-am.co.jp/specialreport/quants/20190805\\_01.html](https://www.daiwa-am.co.jp/specialreport/quants/20190805_01.html)

### 佐々木先生による解説シリーズ

- **001 「テキストデータからの特徴抽出 ニュースからの単語による特徴表現」**  
<https://www.daiwa-am.co.jp/specialreport/quants/20190204.html>
- **002 「テキストデータからの特徴抽出 単語の頻度から分かるニュース記事の特徴①」**  
<https://www.daiwa-am.co.jp/specialreport/quants/20190426.html>
- **003 「テキストデータからの特徴抽出 単語の頻度から分かるニュース記事の特徴②」**  
[https://www.daiwa-am.co.jp/specialreport/quants/20190531\\_01.html](https://www.daiwa-am.co.jp/specialreport/quants/20190531_01.html)

#### 当資料のお取扱いにおけるご注意

- 当資料は投資判断の参考となる情報提供を目的として大和投資信託が作成したものであり、勧誘を目的としたものではありません。投資信託のお申込みにあたっては、販売会社よりお渡しする「投資信託説明書(交付目論見書)」の内容を必ずご確認のうえ、ご自身でご判断ください。
- 当資料は信頼できると考えられる情報源から作成しておりますが、その正確性・完全性を保証するものではありません。運用実績などの記載内容は過去の実績であり、将来の成果を示唆・保証するものではありません。記載内容は資料作成時点のものであり、予告なく変更されることがあります。また、記載する指数・統計資料等の知的所有権、その他一切の権利はその発行者および許諾者に帰属します。