

~~どきゅん~~

DQNTレーディング

ヤフー株式会社
金政 実

自己紹介

趣味

- 機械学習×何か

仕事

- 機械学習×広告

年齢

- 18歳と143.9ヵ月



殴りたい, この笑顔

ギリギリU30

AIブーム

人工知能ブーム

- AIで**週休3日**※
- AIで**仕事**
なくなる

(´ω`)
ハノ|
くく

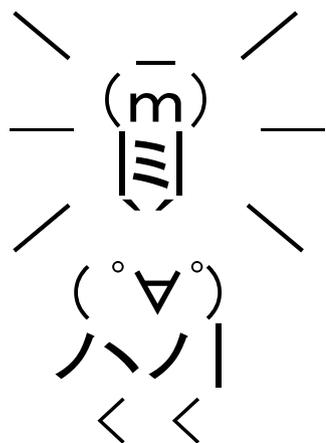
※ヤフーでは介護, 育児を理由とした制度として週4日勤務というものがある
※将来的に導入を目指す
※それ以外にどこでもオフィスという**神制度**がある

機械学習 x 株

予測で儲ける

- AIで**週休7日**！
- AIで**仕事(する必要が)なくなる**
- AI**最高ツ**！

一刻も早く実現するべき



ピコーン
ひらめいた

機械学習 x 株



データを取ってくる



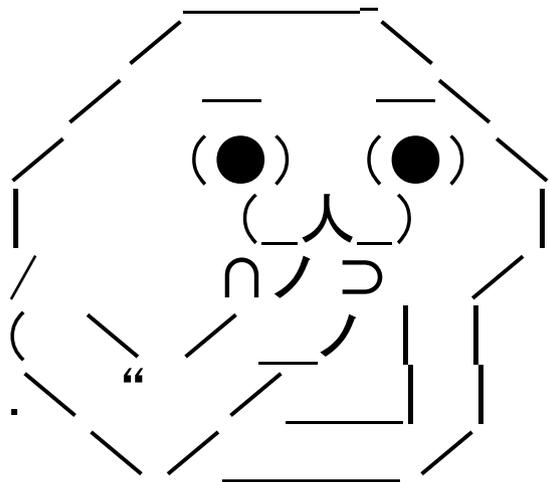
アルゴリズムにぶち込む



予測する

ねっ, 簡単でしょう?

ところが...



以外と
面倒だな

そう単純でもない

1. **手数料が発生**
2. **スリッページもある**
3. **ポジション建てるのはOK**
解消タイミングが不明
4. **AUC0.5...(ランダム)**

単純にぶち込むのは**アウト**

DQN(強化学習+N.N.)

機械学習

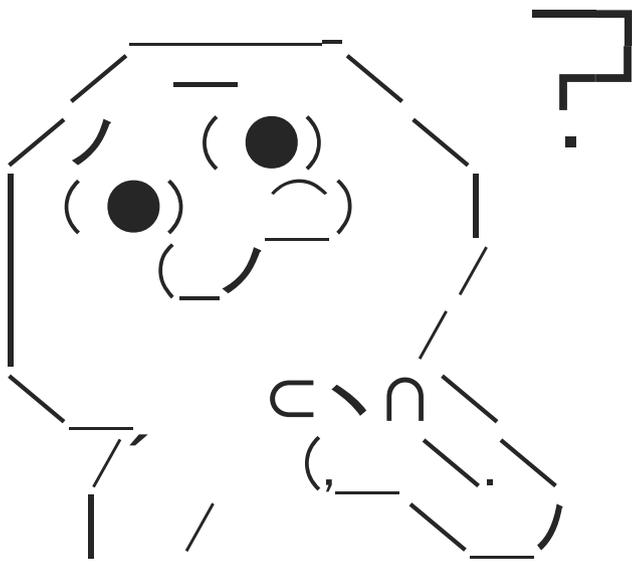
- 次の瞬間に結果が決まる

例) 競馬

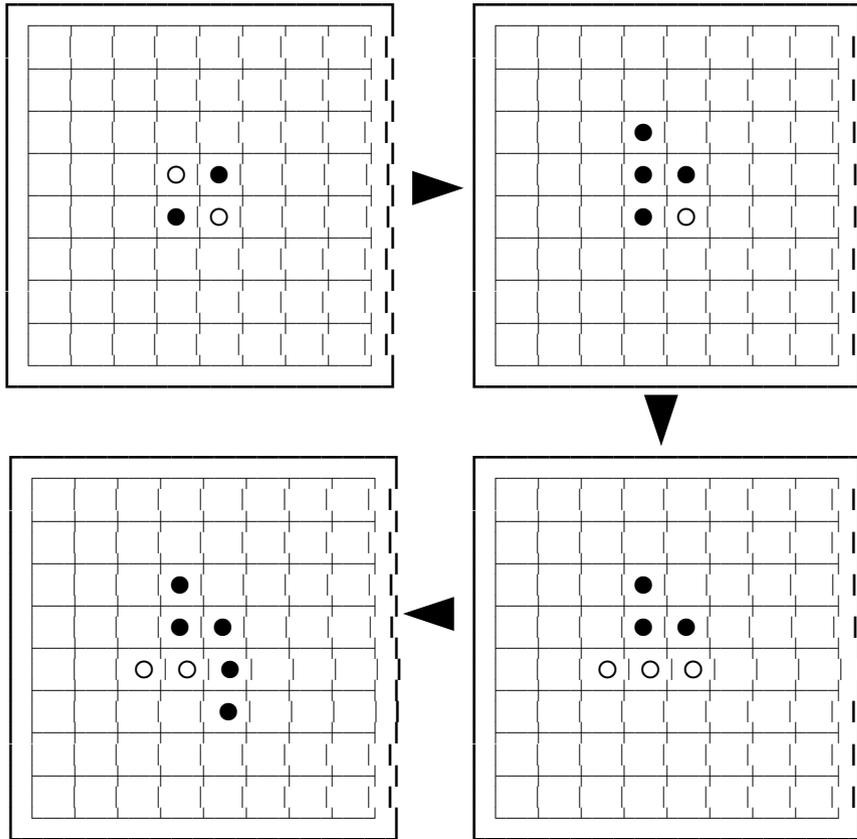
強化学習

- 結果がわかるのはまだ先
- 操作, 運用が必要な問題

例) 囲碁, マリオ, 運転



強化学習



状態を評価するQ関数

最善の1手を指し続ける

盤の状態

$$s = (1, 0, -1, -1, \dots, 1)^T$$

次の1手

行動aの良さ $Q(s, a)$

最良の行動 $a^* = \arg \max_a Q(s, a)$

Q関数の更新

コイントス

当: 100円

外: 0円

Q関数の更新

コイントス

当: 100円

外: 0円

$$Q(s, a) \leftarrow (1 - \alpha)Q(s, a) + \alpha r$$

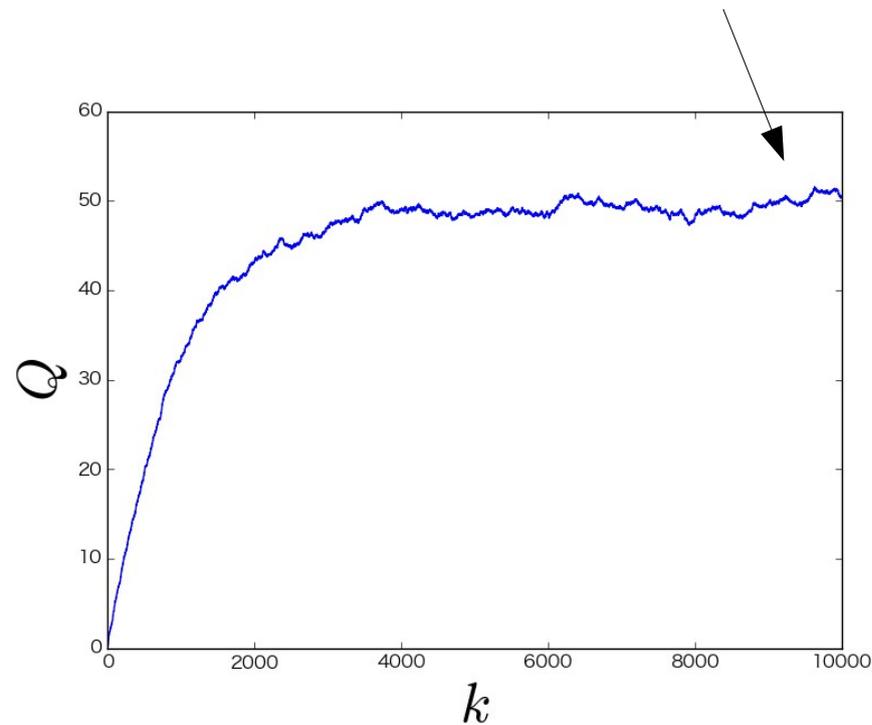
s: 状態(ゲームの状況)

r: 報酬

Qが徐々にE(r)に近づく!

Q関数の更新

50円に収束



$$\alpha = 0.001, Q(0) = 0.0$$

コイントス

当: 100円

外: 0円

$$Q(s, a) \leftarrow (1 - \alpha)Q(s, a) + \alpha r$$

s: 状態(ゲームの状況)

r: 報酬

s, aは**実質的にない**

Qが徐々に**E(r)**に近づく!

$$Q(k+1) = (1 - \alpha)Q(k) + \alpha r(k)$$

実際のQ関数

王銀王



王王

+10点



王

-1000点

× $Q(s, a) \leftarrow (1 - \alpha)Q(s, a) + \alpha r$

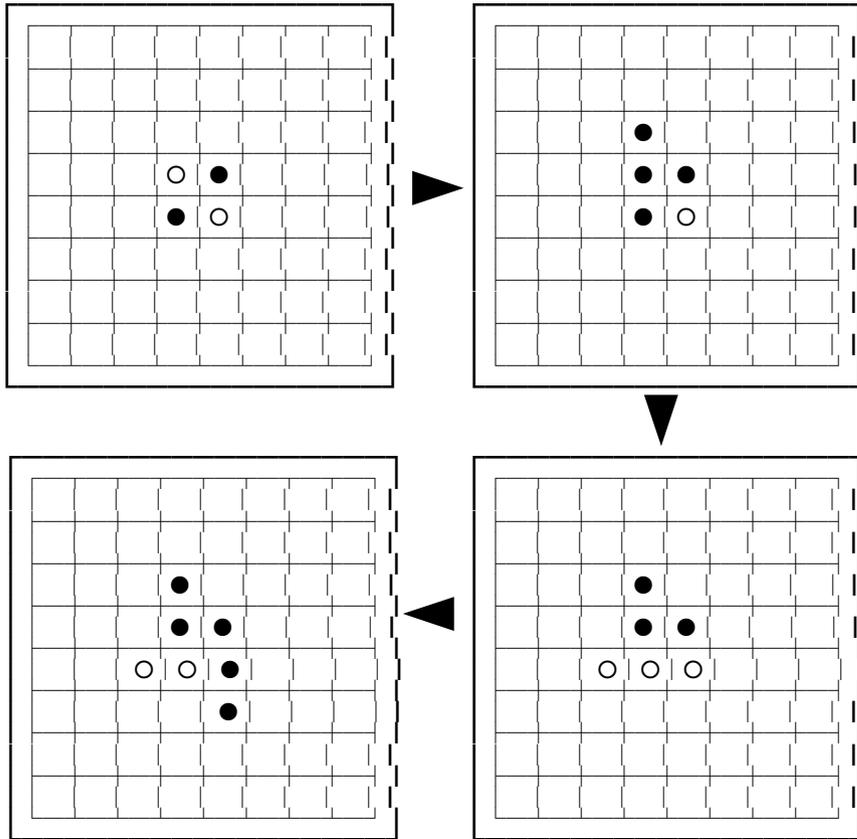
○ $Q(s, a) \leftarrow (1 - \alpha)Q(s, a) + \alpha \left(\underbrace{r}_{\text{今回貰える報酬}} + \gamma \underbrace{\max_{a'} Q(s', a')}_{\text{将来的に貰える報酬}} \right)$

今回貰える報酬

将来的に貰える報酬

目先の報酬に囚われない！

強化学習



Q: 行動を起こした時の良さ
最善の1手を指し続ける

盤の状態

$$s = (1, 0, -1, -1, \dots, 1)^T$$

次の1手

行動aの良さ $Q(s, a)$

最良の行動 $a^* = \arg \max_a Q(s, a)$

問題設定

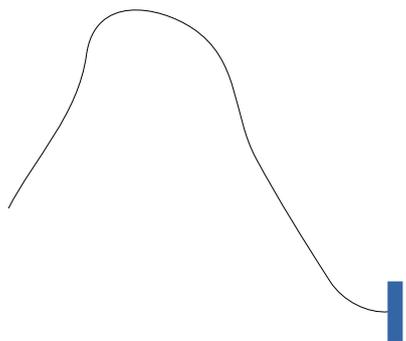
トヨタ株だけを考える

- 流動性がある
(売買したい時に可)
- 安定
(急激な変動がない)



コイツを売り買いするッ！

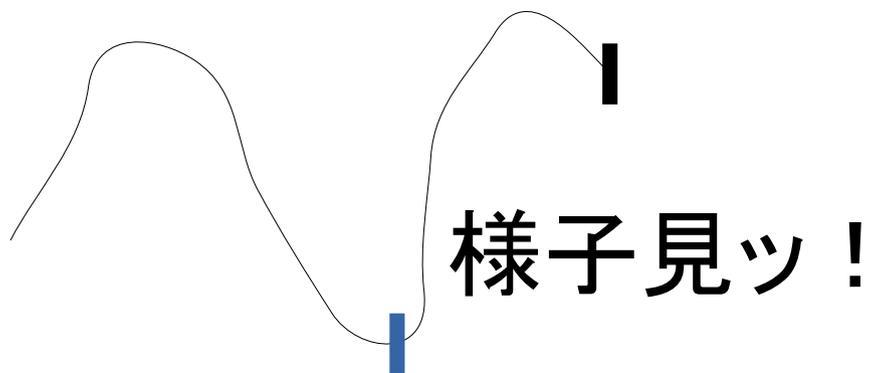
問題設定



買いッ！



問題設定

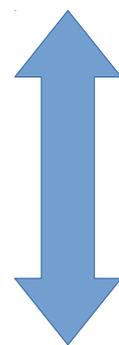
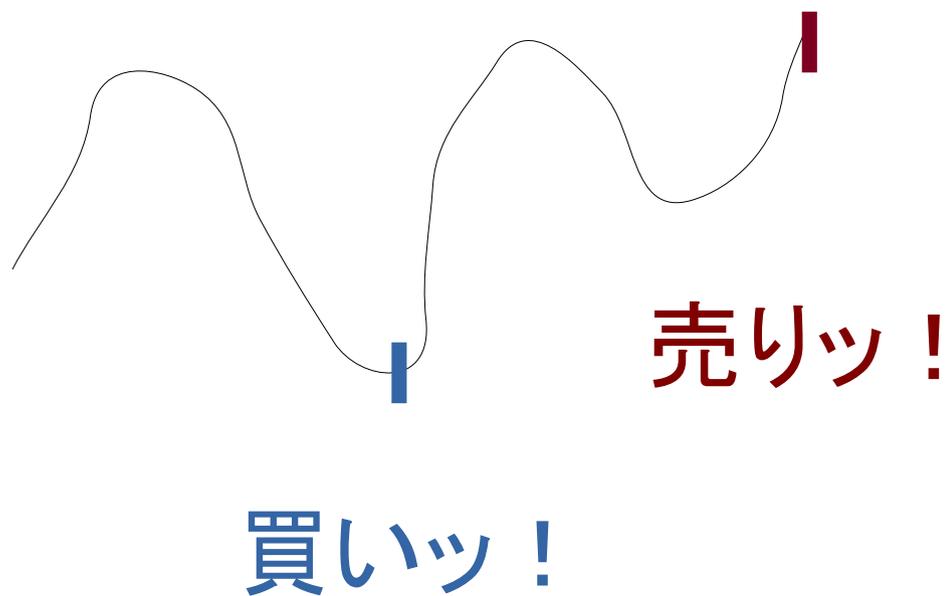


様子見ッ！

買いッ！

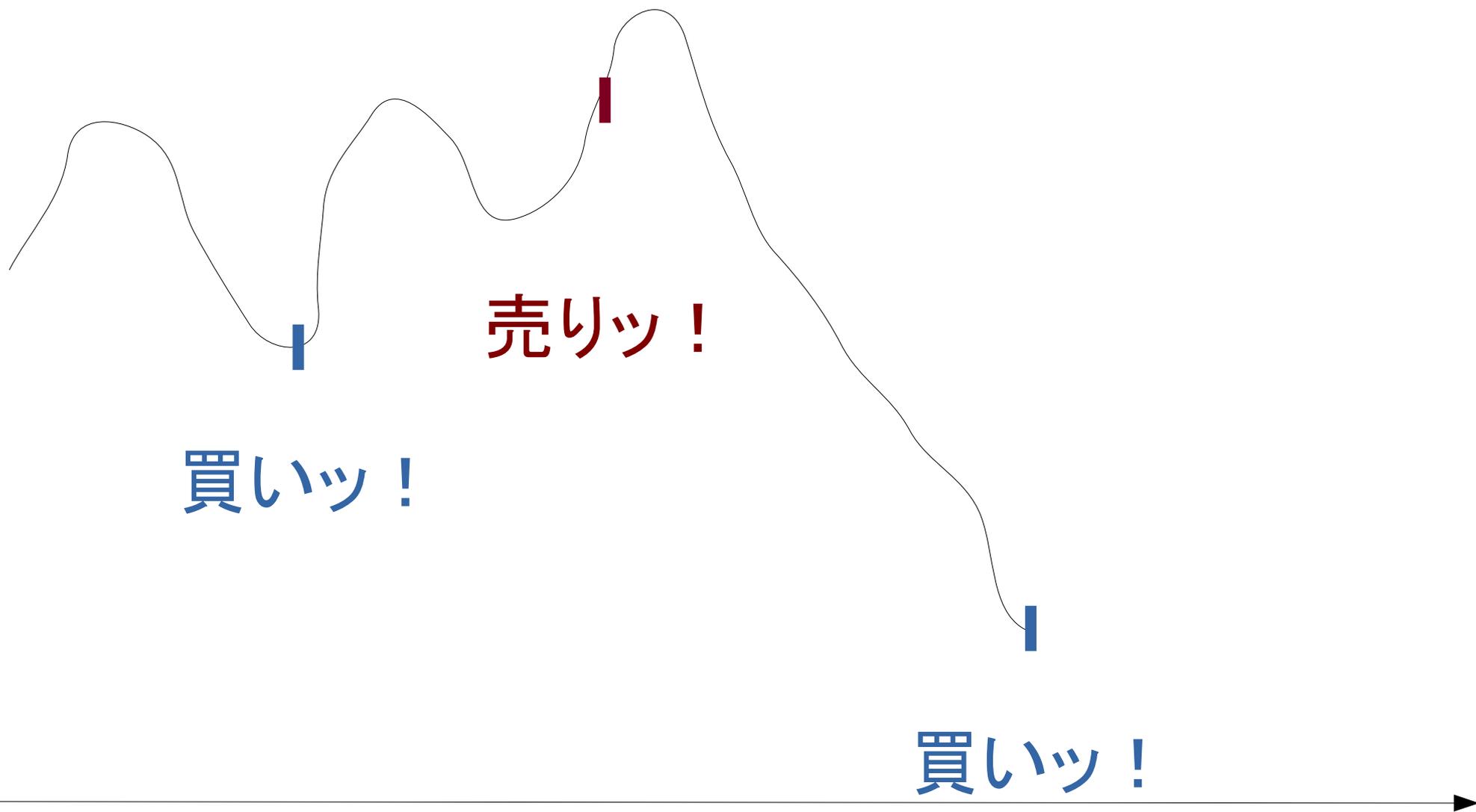


問題設定

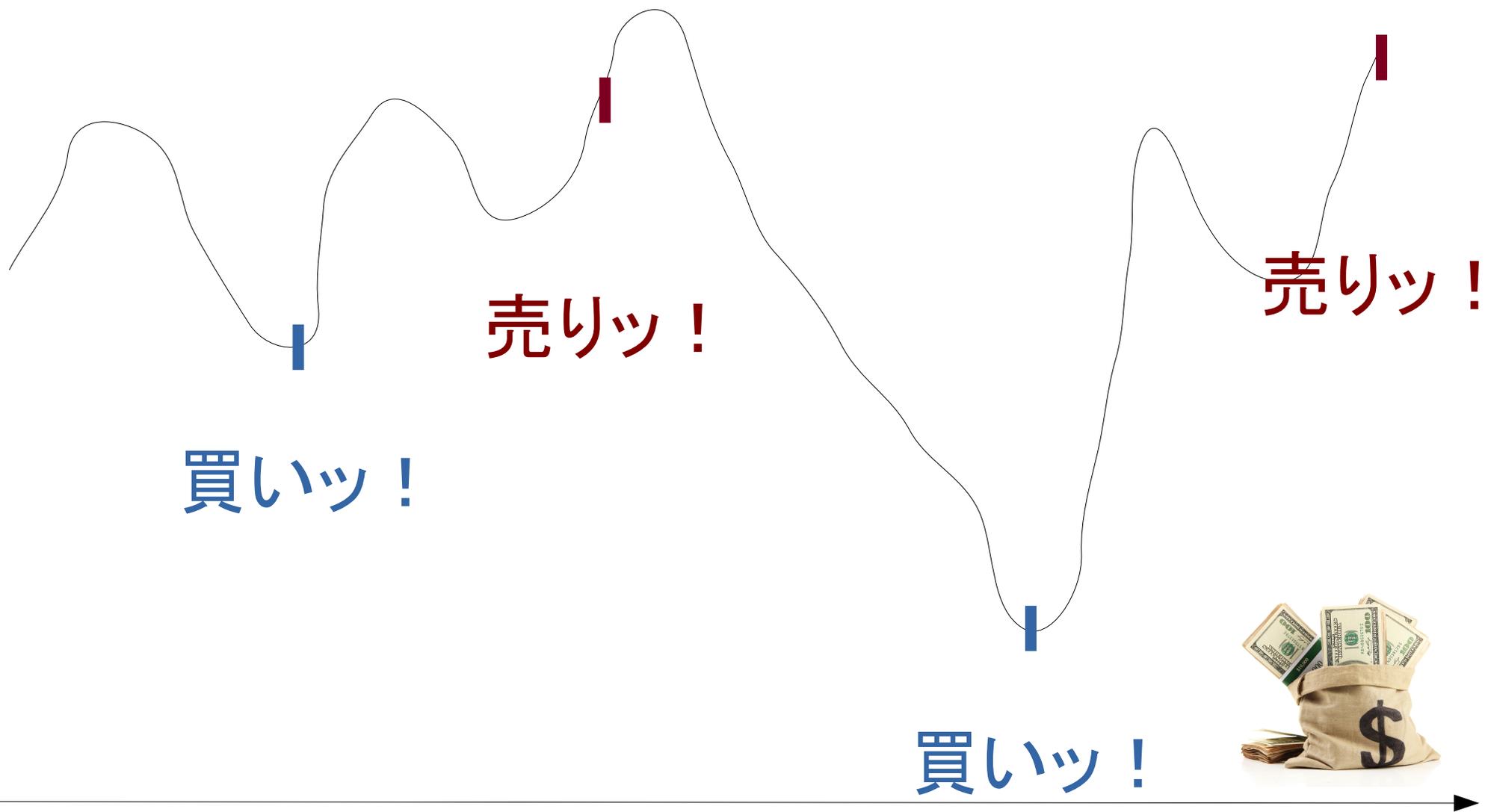


利益確定の際に報酬発生

問題設定



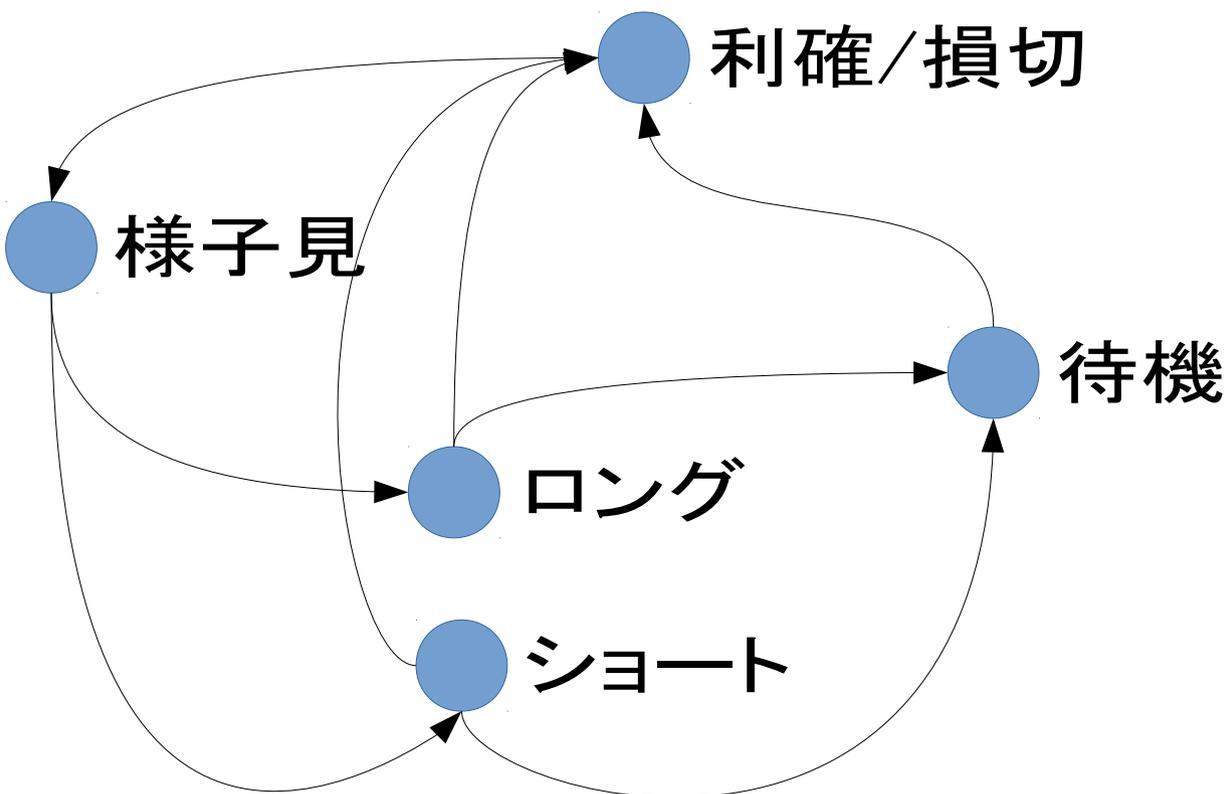
問題設定



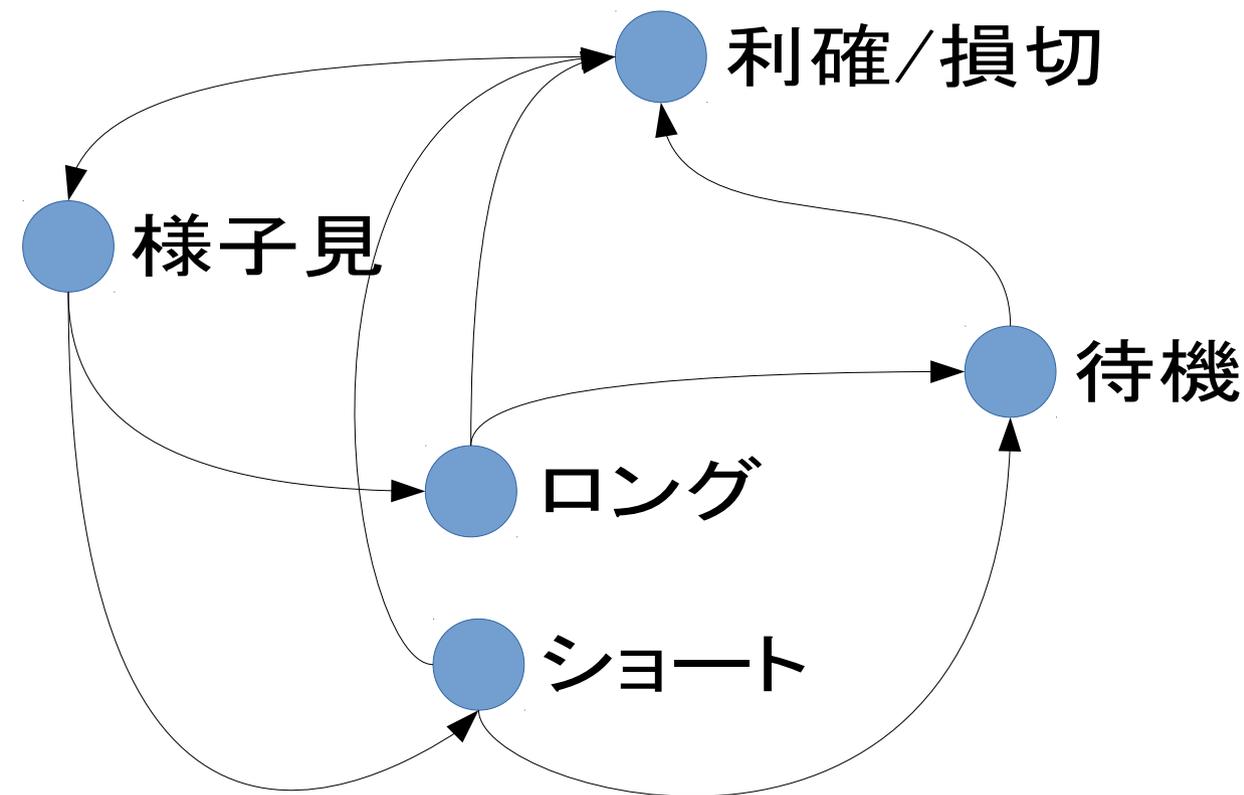
状態/行動遷移

報酬

- 無効な行動は罰則
- 利確/損切において
対数変化率が報酬



状態/行動遷移



報酬

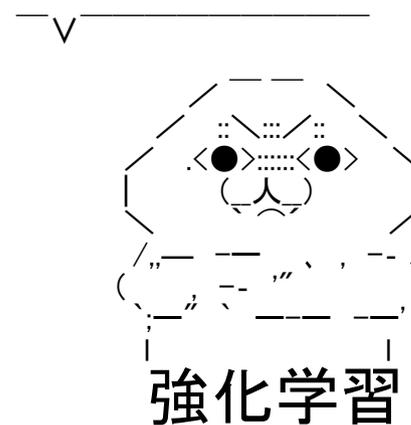
- 無効な行動は罰則
- 利確/損切において
対数変化率が報酬

糞ポジを保持し続ける
糞AIに

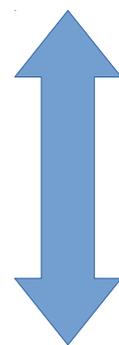
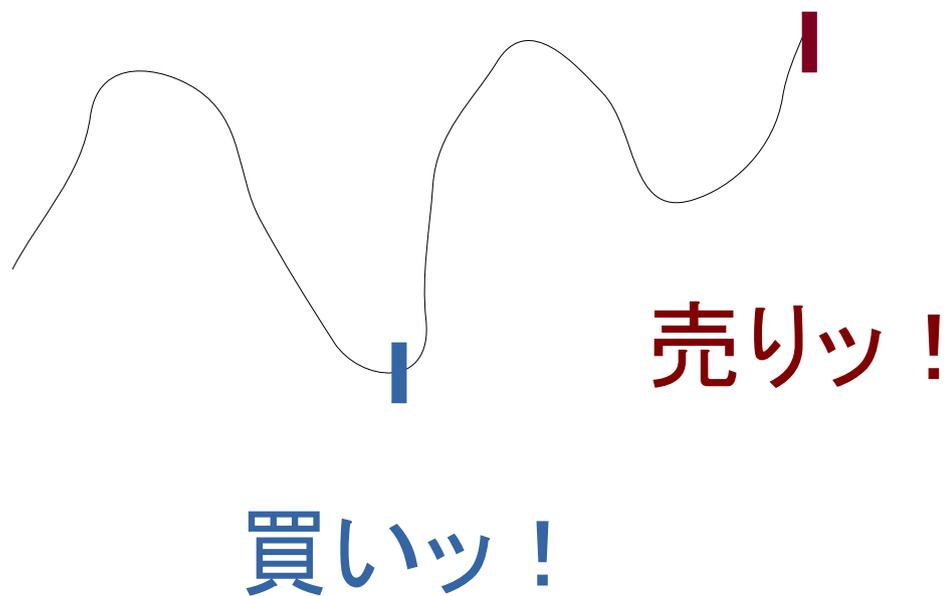
問題設定



糞ポジでも損切しなければ
損ではないッ

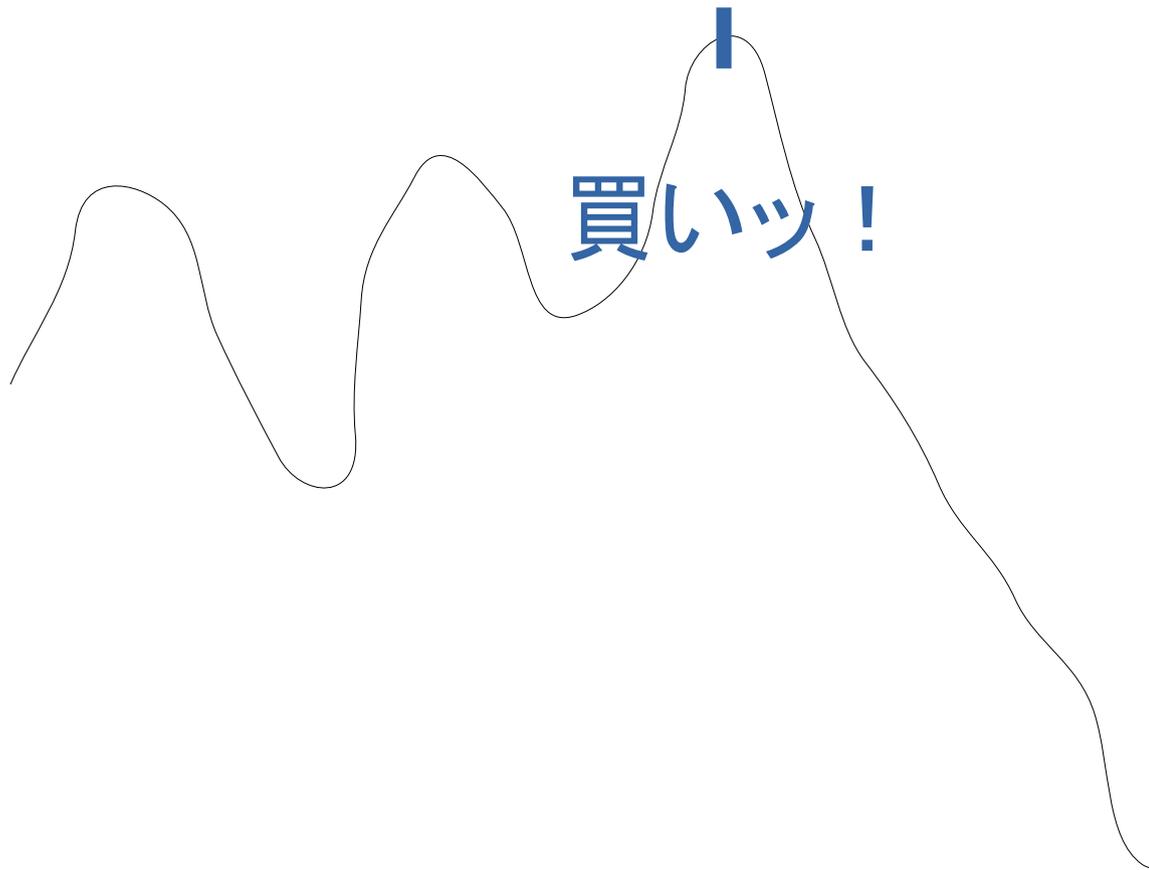


問題設定



利益確定の際に報酬発生

問題設定



糞ポジでも損切しなければ
損ではないッ



おまえはそれでも
現実の俺は死ぬ

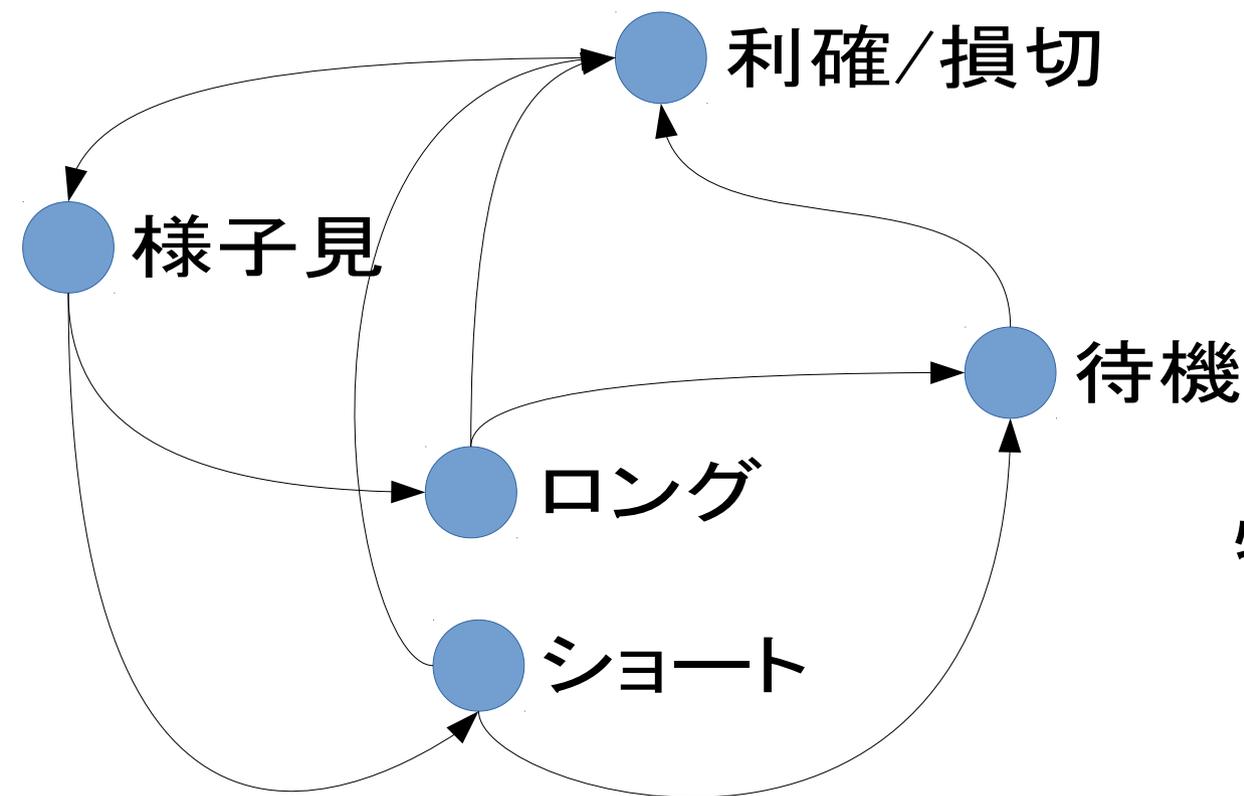
状態/行動遷移

報酬

- 無効な行動は罰則
- 利確/損切において対数変化率が報酬

特別ルール

- 10%で損切り



シミュレーション設定

細かいパラメーター/Feature

- 企業秘密

リプレイ回数

- 10000回

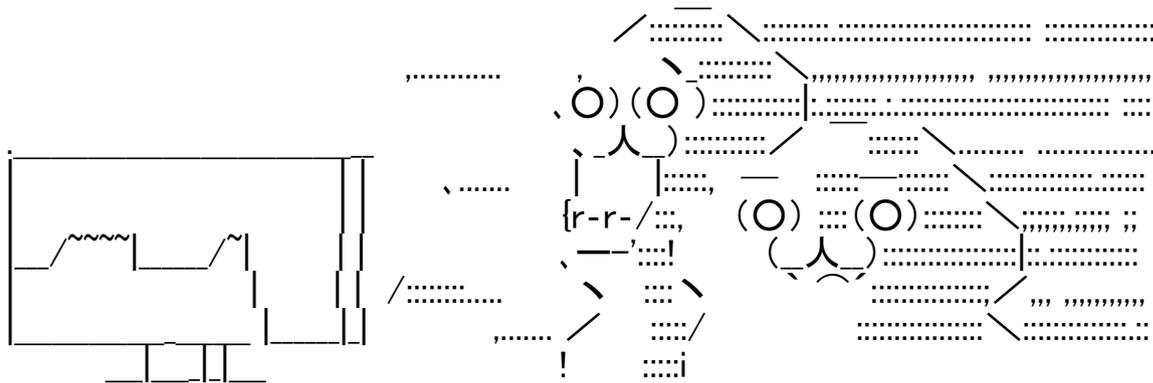
訓練データ

- 2000営業日(約8年)

テストデータ

- 371営業日(約1.5年)

50試行



結果

```
0.31978881  
0.31978881  
-0.46217226  
0.44811595  
0.42227679  
0.52553291  
0.47076625  
0.3314136  
0.33007195
```

...

平均対数変化率: 0.1114
t値: 2.146
P値: 3.670%

1年あたりの利益: $\exp(0.1114/1.5)=1.077$
税金考慮: $(1.077-1)*0.8+1=1.061 = +6.1\%$

年+6%の複利

反省

```
0.31978881  
0.31978881  
0.46217226  
0.44811595  
0.42227679  
0.52553291  
0.47076625  
0.3314136  
0.33007195
```

...

テストデータが1つしかない
→他の銘柄/期間でもやるべき

期間が同じ
→期間を変えるべき

銘柄が同じ
→銘柄を変えるべき

検定だけでは不十分
俺たちの戦いはこれからだッ！

質問コーナー