

1. これまでの研究

オセロ対戦A Iの設計を通して、機械学習の研究を積み重ねてきた。

【オセロA I ①】
アルゴリズムA I
手動でデータ分析

【オセロA I ②】
パーセプトロンによる機械学習
データ分析を自動化

【オセロA I ③】
ニューラルネットワーク
による機械学習
+独自の
アルゴリズム

成果

機械学習ができる独自のアルゴリズムを設計できたデータ分析の精度を高めることができた。

課題

教師データなしで学習する強化学習A Iを設計する。

2. 研究の概要

ステージの端からランダムに出てくる隕石を、宇宙船を動かして避ける、弾幕ゲームを作る。A Iは、宇宙船を8方向に動かすことができる。

ポイント① データの蓄積



ステージの状態を毎フレーム8項目で評価し、データを蓄積する。

17フレーム → 0.1秒

ポイント② 決定木の作成



ランダムフォレストを用いた強化学習を作る。回帰木は難易度が高いので、分類木を使った決定木を作る。

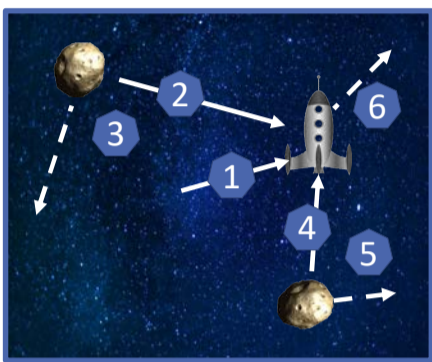
データをとる → 決定木にそって動く → 増えたデータを使って決定木を更に強化する。以上を繰り返し、学習後の宇宙船の動きを評価する。

3. 研究のポイントについて

(1) データをためる

初めからA Iが宇宙船を動かすことも可能だが、実験の効率を上げるため、最初に人が操作したデータをためる。

8項目の要素を設定する



説明変数

- 1 自機のステージ中心との距離
- 2 隕石Aの自機との距離
- 3 隕石Aの進行方向と隕石から見た自機の角度の差
- 4 隕石Bの自機との距離
- 5 隕石Bの進行方向と隕石から見た自機の角度の差

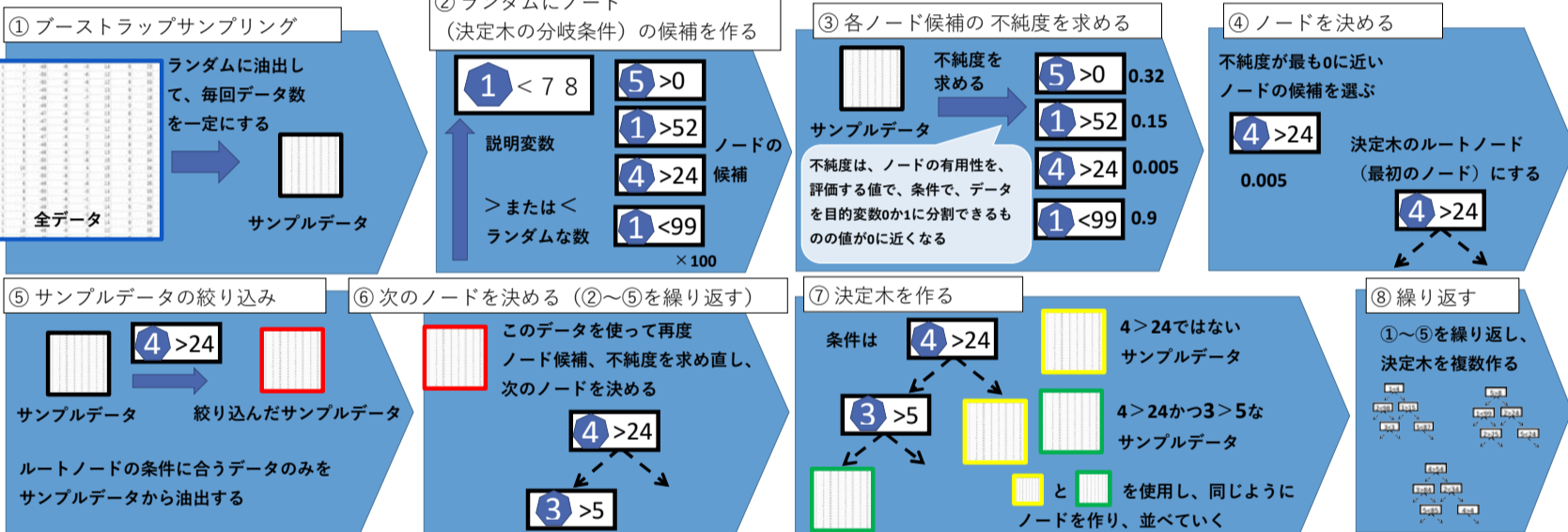
- 6 自機の動いた方向
- 7 隕石にぶつかったかどうか
- 8 ぶつかったときの評価値

目的変数

ぶつかる直前の2秒前までの7を1とするため、75フレームで、ぶつかり、54フレームまでを1にする。さらにその2秒間でも、1~20の評価をつける。(8)76からは2ゲーム目となる。

	1	2	3	4	5	6	7	8
0	54	15	31	14	54	1	0	0
1	64	11	64	58	28	5	0	0
2	31	6	7	14	65	8	0	0
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮
52	55	91	41	15	25	3	0	0
53	48	34	82	25	48	1	0	1
54	73	87	37	26	17	4	1	2
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮
73	54	72	73	58	21	7	1	18
74	46	24	47	73	97	3	1	19
75	38	98	54	95	84	1	1	20
76	5	71	26	18	51	3	0	0
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮

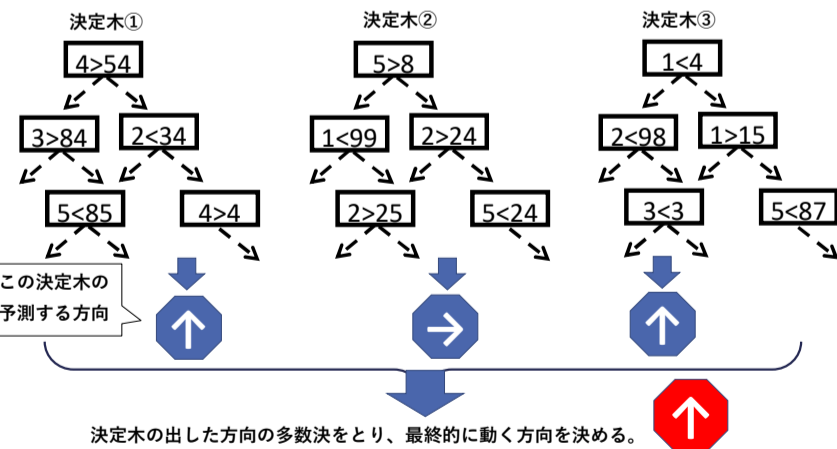
(2) 決定木を作る



※ 不純度が0で、使用したサンプルデータをその条件で、目的変数0か1で完全に分けられる場合、ノードの数が規程の20個に達した場合、その先にノードは作らない。
※ 先にノードを作らない場合、使ったサンプルデータの中で、最も多い方向を選別する。

(3) 決定木に沿って動く

作った決定木をフローチャートの要領で進んでいき、先にノードの続かないところに到達したら、そこで選別された方向がその決定木の予想する方向となる。



始めに人がデータを取った時と同じように8種類のデータを取り、隕石にぶつかるまでこれを繰り返す。

4. 研究結果

- ・データをためてもステージの端によってしまい、出てくる隕石を避けられず当たってしまうことが多く、思うような結果が出せなかった。
- ・データ数が600の時と、6000の時の、避け続けたフレーム数を比較すると、それぞれ100、103となり、データが増えてもほとんど変化はなかった。
- ・実験中、始めは隕石が迫っても、隕石の方に動いたり、動かずにとどまっていたりして、すぐにぶつかっていたが、学習が少し進むと、自機に近づいてくる隕石を、方向転換して避ける動作が、ほぼ毎回みられた。

5. 研究考察

<過学習について>

ブーストランピングにより、各決定木がそれぞれ違う方向を示していたため、過学習は防げていると思われる。

研究結果から、データの数は学習に大きく影響はせず、データのとり方などが大切だと分かった。

ランダムフォレストなどの、基本的な仕組みはできたが、データの中で方向が偏ってしまっていたため、取る説明変数の種類に改善の余地があると思う。

<説明変数の取り方>

今回は、決定木で、最終的に8方向を出すA Iだったが、隕石に当たらない空間を探し出し、そこに向かって動くようにするなどの方法が考えられる。

説明変数は、今回距離と角度だったが、それを隕石同士の位置関係や、自機の座標にするなど、改善することができると思う。

<これまでの研究を通して>

これまでの研究で、パーセプトロンやニューラルネットワーク、ランダムフォレストについて学んで、A Iを使う用途に合わせて、これらの手法を適切に選択することが大切だと分かった。また、今回の研究をとおして、そういった手法だけでなく、A Iに取り込む説明変数などのデータの種類なども同じように大切だと分かった。これからA Iを開発することがあれば、それらも大切にしたい。