

人口知能
中間レポート

035743A：比嘉雅樹

提出日： 2005/01/11

1 C4.5 プログラムにより、例題を実行し、結果を得よ。

1.1 プログラムの実行結果

C4.5 [release 8] decision tree generator Tue Jan 11 14:09:57 2005

Options:

File stem <../Data/mini>

Read 9 cases (3 attributes) from ../Data/mini.data

Decision Tree:

目の色 = 茶: - (4.0)

目の色 = 青:

| 髪の色 = ブロンド: + (2.0)

| 髪の色 = 黒: - (2.0)

| 髪の色 = 赤: + (1.0)

Tree saved

Evaluation on training data (9 items):

Before Pruning		After Pruning			
Size	Errors	Size	Errors	Estimate	
6	0(0.0%)	6	0(0.0%)	(43.6%)	<<

2 自身で分類問題を設計し、決定木を生成せよ。

2.1 問題意図

現在、工 1-401 教室での娯楽はテレビゲーム、麻雀、漫画等ある。その内最も興味のある麻雀の、どういう状況の時になら和了りやすいかを決定木で生成してみた。

2.2 属性及び属性値の設定

データはテンパイした後のみに焦点を絞り、属性は

- テンパイまでに鳴いた数 「naki」
- 待ちの種類「matishurui」(両面「r」、単騎「t」、シャボ「s」、辺張「pk」、嵌張「pk」、その他の多面張「tmen」)
- 和了牌の筋を自分がいくつ捨てたか「matihai」s0,s1,s2
- 場に見えていない残りの自分の和了牌の数(割合)「nokori」
- 親か子か「o,k」

に設定した。しかし、「待ちの種類」の辺張と嵌張は待ちの性質が似ているため一緒にまとめた。クラスは和了が「+」、和了れない時が「-」とした。

2.3 トレーニングデータの作成

トレーニングデータの sample.data,sample.names,sample.test は

```
/net/home/y03/j03043/public_html/
```

参照

2.4 C4.5 プログラムによる実行結果

```
[nw0343:~/Desktop/R8/Src] higamasa% c4.5 -f ../Data/sample
```

```
C4.5 [release 8] decision tree generator Tue Jan 11 15:01:43 2005
```

```
Options:
```

```
File stem <../Data/sample>
```

Read 83 cases (5 attributes) from ../Data/sample.data

Decision Tree:

```
matihai = s0:
| matishurui = pk: - (6.0/1.0)
| matishurui = t: - (4.0/1.0)
| matishurui = tmen: + (3.0/1.0)
| matishurui = r:
| | nokori <= 0.666 : - (12.0/4.0)
| | nokori > 0.666 : + (5.0)
| matishurui = s:
| | naki = 0: + (3.0/1.0)
| | naki = 1: + (1.0)
| | naki = 2: - (3.0)
| | naki = 3: - (0.0)
| | naki = 4: - (0.0)
matihai = s1:
| naki = 2: - (4.0/1.0)
| naki = 3: - (2.0)
| naki = 4: - (1.0)
| naki = 0:
| | oya or ko = o:
| | | nokori <= 0.4 : - (2.0)
| | | nokori > 0.4 : + (3.0/1.0)
| | oya or ko = k:
| | | matishurui = r: + (4.0)
| | | matishurui = pk: - (1.0)
| | | matishurui = t: + (1.0)
| | | matishurui = s: + (2.0/1.0)
| | | matishurui = tmen: + (1.0)
| naki = 1:
| | matishurui = r: + (4.0/1.0)
| | matishurui = pk: + (2.0/1.0)
| | matishurui = t: - (1.0)
| | matishurui = s: + (0.0)
| | matishurui = tmen: + (0.0)
matihai = s2:
| naki = 0: + (10.0)
```

```

| naki = 1: + (4.0)
| naki = 2: + (2.0)
| naki = 3: - (1.0)
| naki = 4: - (1.0)

```

Simplified Decision Tree:

```

matihai = s2: + (18.0/3.7)
matihai = s0:
| nokori <= 0.666 : - (20.0/8.0)
| nokori > 0.666 : + (17.0/8.9)
matihai = s1:
| naki = 1: + (7.0/4.4)
| naki = 2: - (4.0/2.2)
| naki = 3: - (2.0/1.0)
| naki = 4: - (1.0/0.8)
| naki = 0:
| | oya or ko = k: + (9.0/3.5)
| | oya or ko = o:
| | | nokori <= 0.4 : - (2.0/1.0)
| | | nokori > 0.4 : + (3.0/2.1)

```

Tree saved

Evaluation on training data (83 items):

Before Pruning		After Pruning			
Size	Errors	Size	Errors	Estimate	
40	13(15.7%)	15	22(26.5%)	(42.8%)	<<

2.5 得られた決定木の分析

得られた決定木を見てみると、筋を2種類捨てていれば和了りやすく、筋を捨てていなくても残りの和了牌の割合が0.666以上であれば和了りやすい

と分かる。鳴きに関しては、鳴くのはテンパイには近づく代わりに警戒される為、1回が適当である事が読み取れる。また、子の時に対して親の時は多少和了りにくい事がわかる。これは、親は点数が1.5倍になる為警戒されているからだと推測できる。

2.6 テストデータによる評価

テストデータは実際に東風戦を一回行いそのデータを取った。

Evaluation on test data (11 items):

Before Pruning		After Pruning		
Size	Errors	Size	Errors	Estimate
40	1 (9.1%)	15	2 (18.2%)	(42.8%) <<
(a)	(b)	<-classified as		
6	1	(a): class +		
1	3	(b): class -		

テストデータにより評価してみたところ、高い精度で分類できている事が分かる。

2.7 考察

今回はデータを集める機会が少なくデータの量も少ないため分類の正確性も定かではないが、ある程度予想通りの結果になった。麻雀では運による要素が大きいため、実際勝負の場ではそれなりの誤差もあると考えられるが、これからの麻雀でもデータを集め、より正確な分類を行っていきたいと思う。