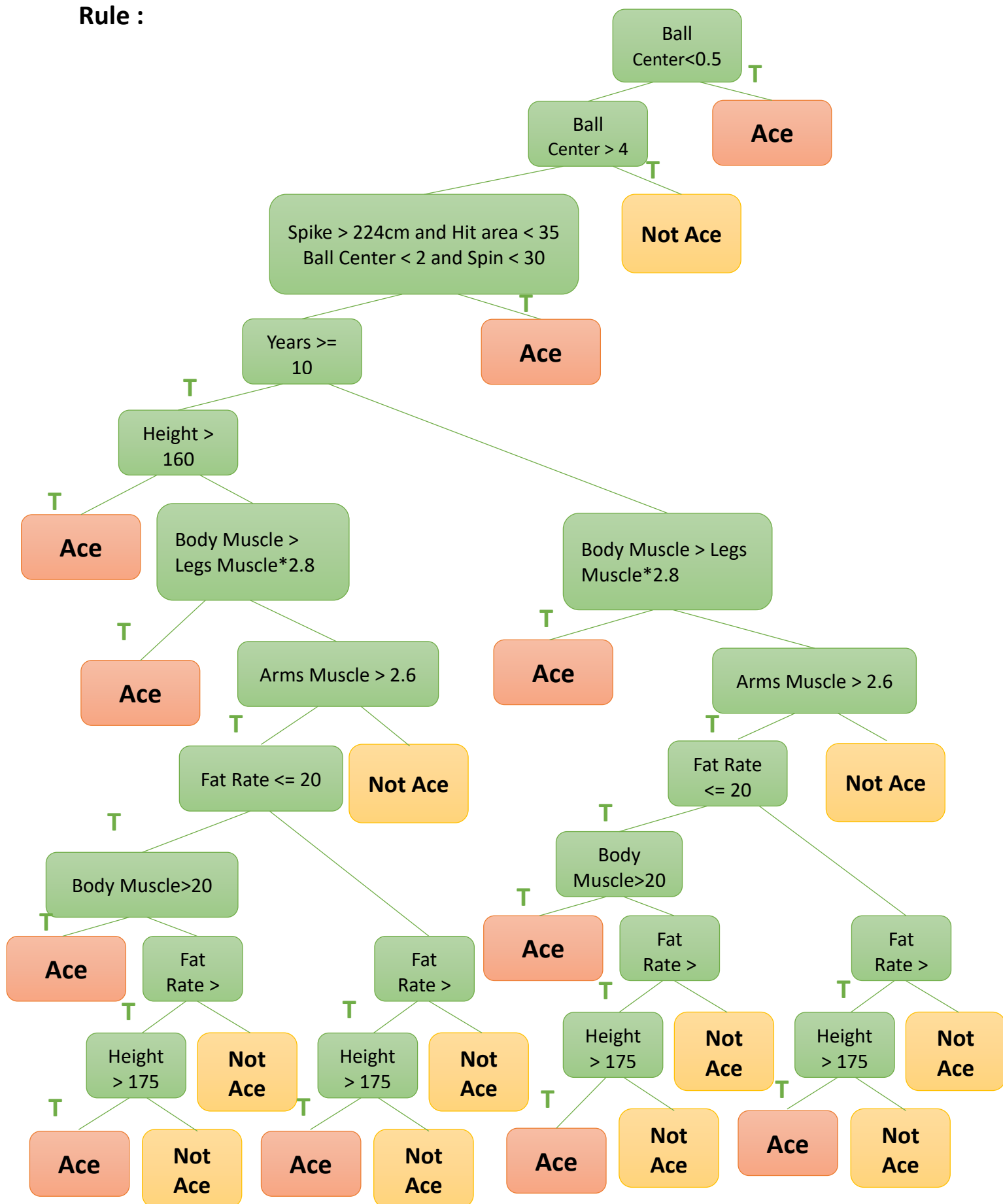
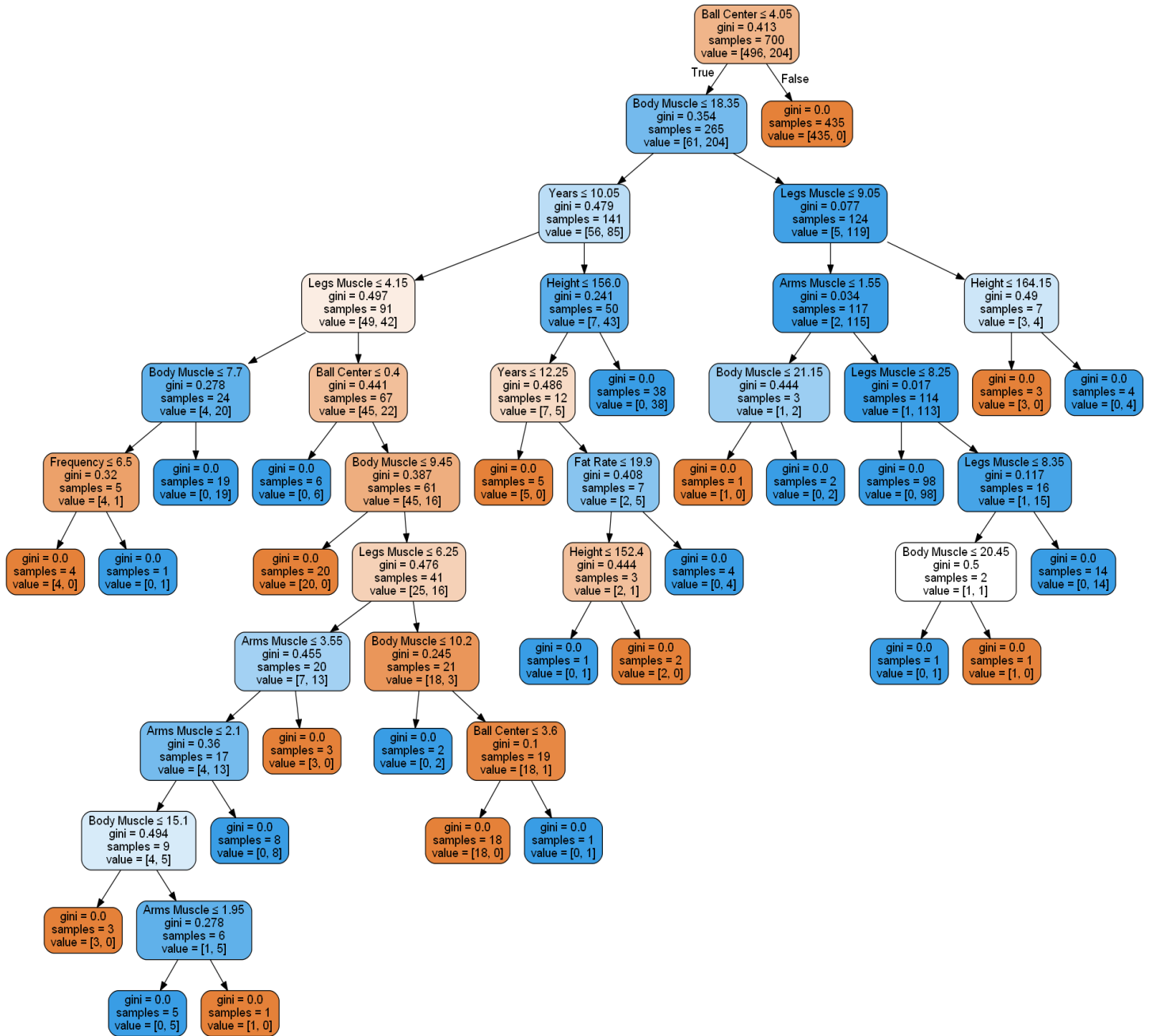


Rule :





資料介紹：這個資料集是根據排球員的身體素質與能力，來推測他能不能在發球時就發出直接得分的 ACE 球。資料集包含打點高度、集球面積、打點與球心距離、球的旋轉速度、身體肌肉量、體脂肪率、球齡、身高與打球頻率。

分類結果：除了第二個規則 “Ball Center > 4” 之外，其他規則並沒有明顯的被決策樹找出來，而除了簡單一點的規則之外，取出來的 value 值幾乎都與設計的規則有些差距，如：feature[i] < value，此類不涉及加減乘除法，也不涉及兩個以上的 feature 比較，value 才會比較接近原始的規則。

Model Comparison

我用了四種分類模型，分別是 Decision Tree、Random Forest、SVM 與 KNN 分類模型。

當規則。分別用(1)1000 筆的資料，前 70%資料預測後 70%資料，(2)500 筆資料，前 30%資料預測後 30%資料，(3)1000 筆資料、前 30%資料預測後 70%資料。

當資料量較多時，Decision Tree 與 Random Forest 做出來的準確度差不多，其次是 SVM，最差的是 KNN 分類器。資料量在 500 筆時，SVM 的表現最好，其次是 KNN，且 KNN 的 K 值設在 1 的時候會得到最佳的準確率。

Training_rate = 0.7

Data size = 1000

Decision Tree

Match = 681 / 700

Confusion_matrix =

```
[[496   9]
```

```
[ 10 185]]
```

	precision	recall	f1-score	support
0	0.98	0.98	0.98	505
1	0.95	0.95	0.95	195
avg / total	0.97	0.97	0.97	700

Random Forest

Match = 680 / 700

Confusion_matrix =

```
[[500   5]
```

```
[ 15 180]]
```

	precision	recall	f1-score	support
0	0.97	0.99	0.98	505
1	0.97	0.92	0.95	195
avg / total	0.97	0.97	0.97	700

SVM

Match = 625 / 700

Confusion_matrix =

```
[[505  0]
```

```
[ 75 120]]
```

	precision	recall	f1-score	support
0	0.87	1.00	0.93	505
1	1.00	0.62	0.76	195
avg / total	0.91	0.89	0.88	700

KNN

Match = 583 / 700

Confusion_matrix =

```
[[441  64]
```

```
[ 53 142]]
```

	precision	recall	f1-score	support
0	0.89	0.87	0.88	505
1	0.69	0.73	0.71	195
avg / total	0.84	0.83	0.83	700

Training_rate = 0.7

Data size = 500

Decision Tree

Match = 289 / 700

Confusion_matrix =

```
[[220  29]
```

```
[ 32  69]]
```

	precision	recall	f1-score	support
0	0.87	0.88	0.88	249
1	0.70	0.68	0.69	101
avg / total	0.82	0.83	0.82	350

Random Forest

Match = 290 / 700

Confusion_matrix =

[[235 14]

[46 55]]

	precision	recall	f1-score	support
0	0.84	0.94	0.89	249
1	0.80	0.54	0.65	101
avg / total	0.82	0.83	0.82	350

SVM

Match = 308 / 700

Confusion_matrix =

[[249 0]

[42 59]]

	precision	recall	f1-score	support
0	0.86	1.00	0.92	249
1	1.00	0.58	0.74	101
avg / total	0.90	0.88	0.87	350

KNN

Match = 299 / 700

Confusion_matrix =

[[225 24]

[27 74]]

	precision	recall	f1-score	support
0	0.89	0.90	0.90	249
1	0.76	0.73	0.74	101
avg / total	0.85	0.85	0.85	350

Training_rate = 0.3

Data size = 1000

Decision Tree

Match = 182 / 700

Confusion_matrix =

[[155 70]

[48 27]]

	precision	recall	f1-score	support
0	0.76	0.69	0.72	225
1	0.28	0.36	0.31	75
avg / total	0.64	0.61	0.62	300

Random Forest

Match = 208 / 700

Confusion_matrix =

[[202 23]

[69 6]]

	precision	recall	f1-score	support
0	0.75	0.90	0.81	225
1	0.21	0.08	0.12	75
avg / total	0.61	0.69	0.64	300

SVM

Match = 225 / 700

Confusion_matrix =

[[225 0]

[75 0]]

	precision	recall	f1-score	support
0	0.75	1.00	0.86	225
1	0.00	0.00	0.00	75
avg / total	0.56	0.75	0.64	300

KNN

Match = 196 / 700

Confusion_matrix =

[[173 52]

[52 23]]

	precision	recall	f1-score	support
0	0.77	0.77	0.77	225
1	0.31	0.31	0.31	75
avg / total	0.65	0.65	0.65	300