Demo: Decision trees

Fraida Fund

This is a simple demo notebook that demonstrates a decision tree classifier.

Attribution: Parts of this notebook are slightly modified from this tutorial from "Intro to Data Mining".

```
import pandas as pd
import numpy as np
import matplotlib.pyplot as plt
import seaborn as sns

import sklearn
from sklearn.tree import DecisionTreeClassifier
from sklearn.ensemble import BaggingClassifier, RandomForestClassifier, AdaBoostClassifier
```

```
df = pd.read_csv('http://www.cse.msu.edu/~ptan/dmbook/tutorials/tutorial6/vertebrate.csv')
df
```

					~	
	Name	Warm-blooded		-		
0	human	1		1	0	
1	python	0		0	0	
2	salmon	0		0	1	
3	whale	1		1	1	
4	frog	0		0	1	
5	komodo	0		0	0	
6	bat	1		1	0	
7	pigeon	1		0	0	
8	cat	1		1	0	
9	leopard shark	0		1	1	
10	turtle	0		0	1	
11	penguin	1		0	1	
12	porcupine	1		1	0	
13	eel	0		0	1	
14	salamander	0		0	1	
	Aerial Creature	e Has Legs H	libernates	Class		
0	() 1	0	mammals		
1	(0	1	reptiles		
2	(0	0	fishes		
3	(0	0	mammals		
4	() 1	1	amphibians		
5	() 1	0	reptiles		
6		1 1	1	mammals		
7		1 1	0	birds		
8) 1	0	mammals		
9		0	0	fishes		
10) 1	0	reptiles		
11) 1	0	birds		
12) 1	1	mammals		
13		0 0	0	fishes		
14) 1		amphibians		
1-1	,	, 1	1	ampiriorans		

We'l make it a binary classification problem:

```
df['Class'] = df['Class'].replace(['fishes','birds','amphibians','reptiles'],'non-mammals')
df
```

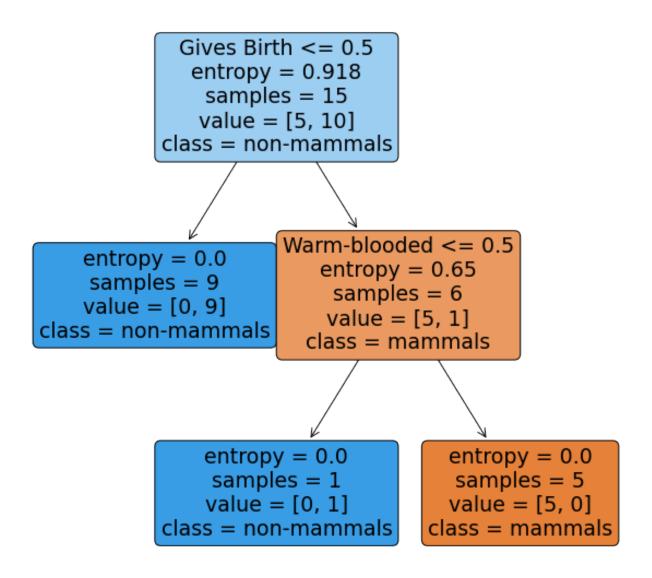
Name Warm-blooded Gives Birth Aquatic Creature 0 human 1 1 1 python 0 0 2 salmon 0 0 3 whale 1 1 4 frog 0 0 5 komodo 0 0 6 bat 1 1 7 pigeon 1 0 8 cat 1 1 9 leopard shark 0 1 1 10 turtle 0 0 0 11 penguin 1 1 1 12 porcupine 1 1 1 13 eel 0 0 0 14 salamander 0 0 mammals 2 0 0 0 mammals 3 0 0 0 non-mammals 4 <td< th=""><th></th><th></th><th></th><th></th><th></th><th></th><th></th></td<>							
1 python 0 0 0 2 salmon 0 0 1 3 whale 1 1 1 4 frog 0 0 0 5 komodo 0 0 0 6 bat 1 1 1 7 pigeon 1 0 0 8 cat 1 1 0 9 leopard shark 0 1 1 10 turtle 0 0 1 11 penguin 1 0 1 12 porcupine 1 1 1 13 eel 0 0 0 14 salamander 0 0 1 0 0 1 non-mammals 1 0 0 non-mammals 2 0 0 0 non-mammals 3		Name	Warm-blooded	Gives Bir	th Aquatic	Creature	!
2 salmon 0 0 1 3 whale 1 1 1 1 1 4 frog 0 0 0 0 1 5 komodo 0 0 0 0 0 6 bat 1 1 1 0 0 7 pigeon 1 0 0 0 1 8 cat 1 1 1 0 0 9 leopard shark 0 1 1 1 0 0 11 penguin 1 0 1 1 0 1 12 porcupine 1 1 0 0 1 13 eel 0 0 0 1 1 14 salamander 0 0 0 1 14 salamander 0 0 0 1 1	0		1		1	0	
3	1	python	0	1	0	0	
4 frog 0 0 1 5 komodo 0 0 0 6 bat 1 1 0 7 pigeon 1 0 0 8 cat 1 1 0 9 leopard shark 0 1 1 1 10 turtle 0 0 1 1 11 penguin 1 0 1 1 12 porcupine 1 1 1 0 1 13 eel 0 0 0 1 1 14 salamander 0 0 1 1 0 1 14 salamander 0 0 0 1 1 0 1 0 1 1 0 1 0 1 0 1 0 1 0 1 0 1 0 1 0 0 0 0 0 0 0 0 0 0 0 <t< td=""><td>2</td><td>salmon</td><td>0</td><td>1</td><td>0</td><td>1</td><td></td></t<>	2	salmon	0	1	0	1	
5 komodo 0 0 0 6 bat 1 1 0 7 pigeon 1 0 0 8 cat 1 1 0 9 leopard shark 0 1 1 10 turtle 0 0 1 11 penguin 1 0 1 12 porcupine 1 1 0 13 eel 0 0 1 14 salamander 0 0 1 0 0 1 0 1 14 salamander 0 0 1 0 0 1 0 mammals 1 0 0 0 0 1 1 0 0 0 0 0 0 1 0 0 0 0 0 0 0 0 0	3	whale	1		1	1	
6 bat 1 1 0 0 7 pigeon 1 0 0 0 8 cat 1 1 0 0 1 9 leopard shark 0 1 <td< td=""><td>4</td><td></td><td>0</td><td>1</td><td>0</td><td>1</td><td></td></td<>	4		0	1	0	1	
7 pigeon 1 0 0 8 cat 1 1 0 9 leopard shark 0 1 1 10 turtle 0 0 1 11 penguin 1 0 1 11 porcupine 1 1 0 13 eel 0 0 1 14 salamander 0 0 1 1 salamander 0 0 1 1 0 0 1 0 1 1 0 0 1 0 1 0 1 0	5	komodo	0	1	0	0	
8 cat 1 1 0 9 leopard shark 0 1 1 10 turtle 0 0 1 11 penguin 1 0 1 11 porcupine 1 1 0 13 eel 0 0 1 14 salamander 0 0 1 0 0 1 0 mammals 1 0 0 1 non-mammals 2 0 0 0 non-mammals 3 0 0 0 non-mammals 4 0 1 1 non-mammals 5 0 1 0 non-mammals 6 1 1 1 mammals 7 1 1 0 non-mammals 8 0 1 0 non-mammals 10 0 1 0	6	bat	1		1	0	
9 leopard shark 0 1 1 1 10 turtle 0 0 0 1 11 penguin 1 0 1 12 porcupine 1 1 1 0 13 eel 0 0 0 1 14 salamander 0 0 0 1 Aerial Creature Has Legs Hibernates Class 0 0 1 non-mammals 1 0 non-mammals 2 0 0 0 mammals 3 0 0 0 mammals 4 0 1 non-mammals 5 0 1 non-mammals 6 1 non-mammals 6 1 non-mammals 7 1 non-mammals 6 1 non-mammals 7 1 non-mammals 7 1 non-mammals 7 1 non-mammals 7 1 non-mammals 1 1 non-mammal		pigeon	1		0	0	
10 turtle 0 0 1 11 penguin 1 0 1 12 porcupine 1 1 0 13 eel 0 0 1 14 salamander 0 0 1 Aerial Creature Has Legs Hibernates Class 0 0 1 non-mammals 1 0 0 1 non-mammals 2 0 0 0 non-mammals 3 0 0 0 mammals 4 0 1 1 non-mammals 5 0 1 0 non-mammals 6 1 1 1 mammals 7 1 1 0 non-mammals 8 0 1 0 non-mammals 10 0 1 0 non-mammals 11 0 1 0 non-mammals 11 0 1 0 non-mammals	8		1		1	0	
11 penguin 1 0 1 12 porcupine 1 1 0 13 eel 0 0 1 14 salamander 0 0 1 Aerial Creature Has Legs Hibernates Class 0 0 1 non-mammals 1 0 0 1 non-mammals 2 0 0 0 non-mammals 3 0 0 0 non-mammals 4 0 1 1 non-mammals 5 0 1 0 non-mammals 6 1 1 1 mammals 7 1 1 0 non-mammals 8 0 1 0 non-mammals 9 0 0 0 non-mammals 10 0 1 0 non-mammals 11 0 1 0 non-mammals 11 0 1 0 non-mammals	9		0	1	1	1	
12 porcupine 1 1 0 13 eel 0 0 1 14 salamander 0 0 1 Aerial Creature Has Legs Hibernates Class 0 0 1 non-mammals 1 0 0 1 non-mammals 2 0 0 0 non-mammals 3 0 0 0 mammals 4 0 1 1 non-mammals 5 0 1 0 non-mammals 6 1 1 1 mammals 7 1 1 0 non-mammals 8 0 1 0 non-mammals 9 0 0 0 non-mammals 10 0 1 0 non-mammals 11 0 1 0 non-mammals 12 0 1 1 non-mammals 13 0 0 0 non-mammals </td <td></td> <td>turtle</td> <td>0</td> <td>1</td> <td>0</td> <td>1</td> <td></td>		turtle	0	1	0	1	
13 eel 0 0 1 14 salamander 0 0 1 Aerial Creature Has Legs Hibernates Class 0 0 1 0 mammals 1 0 0 1 non-mammals 2 0 0 0 non-mammals 3 0 0 0 mammals 4 0 1 1 non-mammals 5 0 1 0 non-mammals 6 1 1 1 mammals 7 1 1 0 non-mammals 8 0 1 0 non-mammals 10 0 1 0 non-mammals 11 0 1 0 non-mammals 12 0 1 1 mammals 13 0 0 0 non-mammals	11	penguin	1		0	1	
14 salamander 0 0 1 Aerial Creature Has Legs Hibernates Class 0 0 1 non-mammals 1 0 0 1 non-mammals 2 0 0 0 non-mammals 3 0 0 0 mammals 4 0 1 1 non-mammals 5 0 1 0 non-mammals 6 1 1 1 mammals 7 1 1 0 non-mammals 8 0 1 0 non-mammals 9 0 0 0 non-mammals 10 0 1 0 non-mammals 11 0 1 0 non-mammals 12 0 1 1 mammals 13 0 0 0 non-mammals	12	porcupine	1		1	0	
Aerial Creature Has Legs Hibernates Class 0 0 1 0 mammals 1 0 0 1 non-mammals 2 0 0 0 non-mammals 3 0 0 0 mammals 4 0 1 1 non-mammals 5 0 1 0 non-mammals 6 1 1 1 mammals 7 1 1 0 non-mammals 8 0 1 0 non-mammals 9 0 0 0 non-mammals 10 0 1 0 non-mammals 11 0 1 0 non-mammals 12 0 1 1 mammals 13 0 0 non-mammals	13	eel	0	1	0	1	
0 0 1 0 mammals 1 0 0 1 non-mammals 2 0 0 0 non-mammals 3 0 0 0 mammals 4 0 1 1 non-mammals 5 0 1 0 non-mammals 6 1 1 1 mammals 7 1 1 0 non-mammals 8 0 1 0 mammals 9 0 0 0 non-mammals 10 0 1 0 non-mammals 11 0 1 0 non-mammals 12 0 1 1 mammals 13 0 0 0 non-mammals	14	salamander	0	1	0	1	
0 0 1 0 mammals 1 0 0 1 non-mammals 2 0 0 0 non-mammals 3 0 0 0 mammals 4 0 1 1 non-mammals 5 0 1 0 non-mammals 6 1 1 1 mammals 7 1 1 0 non-mammals 8 0 1 0 mammals 9 0 0 0 non-mammals 10 0 1 0 non-mammals 11 0 1 0 non-mammals 12 0 1 1 mammals 13 0 0 0 non-mammals							
1 0 0 1 non-mammals 2 0 0 0 mammals 3 0 0 0 mammals 4 0 1 non-mammals 5 0 1 non-mammals 6 1 1 1 mammals 7 1 1 non-mammals 8 0 1 mammals 9 0 0 mammals 9 0 0 non-mammals 10 0 1 0 non-mammals 11 0 non-mammals 12 0 1 mammals 13 0 0 non-mammals			_				
2 0 0 0 non-mammals 3 0 0 0 mammals 4 0 1 1 non-mammals 5 0 1 0 non-mammals 6 1 1 1 mammals 7 1 1 0 non-mammals 8 0 1 0 mammals 9 0 0 0 non-mammals 10 0 1 0 non-mammals 11 0 1 0 non-mammals 12 0 1 1 mammals 13 0 0 0 non-mammals							
3 0 0 0 mammals 4 0 1 1 non-mammals 5 0 1 0 non-mammals 6 1 1 1 mammals 7 1 1 0 non-mammals 8 0 1 0 mammals 9 0 0 0 non-mammals 10 0 1 0 non-mammals 11 0 1 0 non-mammals 12 0 1 1 mammals 13 0 0 0 non-mammals							
4 0 1 1 non-mammals 5 0 1 0 non-mammals 6 1 1 1 mammals 7 1 1 0 non-mammals 8 0 1 0 mammals 9 0 0 0 non-mammals 10 0 1 0 non-mammals 11 0 1 0 non-mammals 12 0 1 1 mammals 13 0 0 0 non-mammals							
5 0 1 0 non-mammals 6 1 1 1 mammals 7 1 1 0 non-mammals 8 0 1 0 mammals 9 0 0 0 non-mammals 10 0 1 0 non-mammals 11 0 1 0 non-mammals 12 0 1 1 mammals 13 0 0 0 non-mammals							
6 1 1 1 1 mammals 7 1 1 1 0 non-mammals 8 0 1 0 mammals 9 0 0 0 non-mammals 10 0 1 0 non-mammals 11 0 1 0 non-mammals 12 0 1 1 mammals 13 0 0 0 non-mammals							
7 1 1 0 non-mammals 8 0 1 0 mammals 9 0 0 0 non-mammals 10 0 1 0 non-mammals 11 0 1 0 non-mammals 12 0 1 1 mammals 13 0 0 0 non-mammals							
8 0 1 0 mammals 9 0 0 0 non-mammals 10 0 1 0 non-mammals 11 0 1 0 non-mammals 12 0 1 1 mammals 13 0 0 0 non-mammals							
9 0 0 0 non-mammals 10 0 1 0 non-mammals 11 0 1 0 non-mammals 12 0 1 1 mammals 13 0 0 0 non-mammals			1 1				
10 0 1 0 non-mammals 11 0 1 0 non-mammals 12 0 1 1 mammals 13 0 0 0 non-mammals				0			
11 0 1 0 non-mammals 12 0 1 1 mammals 13 0 0 0 non-mammals			0 0	0			
12 0 1 1 mammals 13 0 0 0 non-mammals			0 1	0	non-mammals		
0 0 0 non-mammals			0 1	0	non-mammals		
			0 1	1	mammals		
14 0 1 1 non-mammals			0 0	0	non-mammals		
	14		0 1	1	non-mammals		

Decision tree

```
y = df['Class']
X = df.drop(['Name','Class'],axis=1)

clf_dt = DecisionTreeClassifier(criterion='entropy')
clf_dt = clf_dt.fit(X, y)

plt.figure(figsize=(10,10))
```



Feature importance

```
feature importance
0
       Warm-blooded
                       0.283143
        Gives Birth
                       0.716857
2 Aquatic Creature
                       0.000000
                       0.000000
3
   Aerial Creature
4
          Has Legs
                       0.000000
5
                       0.000000
         Hibernates
```