

Decision Tree Examples

Q1.

1. What is the entropy of this collection of training examples with respect to the target function classification?
2. What is the information gain of $a1$ and $a2$ relative to these training examples?
3. Draw decision tree for the given dataset.

Instance	Classification	a1	a2
1	+	T	T
2	+	T	T
3	-	T	F
4	+	F	F
5	-	F	T
6	-	F	T

Q2. Draw the Decision tree using ID3 algorithm.

Instance	a1	a2	a3	Classification
1	True	Hot	High	No
2	True	Hot	High	No
3	False	Hot	High	Yes
4	False	Cool	Normal	Yes
5	False	Cool	Normal	Yes
6	True	Cool	High	No
7	True	Hot	High	No
8	True	Hot	Normal	Yes
9	False	Cool	Normal	Yes
10	False	Cool	High	Yes

Q3. NASA wants to be able to discriminate between Martians (M) and Humans (H) based on the following characteristics: Green $\in \{N, Y\}$, Legs $\in \{2,3\}$, Height $\in \{S, T\}$, Smelly $\in \{N, Y\}$ Our available training data is as follows:

	<u>Species</u>	<u>Green</u>	<u>Legs</u>	<u>Height</u>	<u>Smelly</u>
1	M	N	3	S	Y
2	M	Y	2	T	N
3	M	Y	3	T	N
4	M	N	2	S	Y
5	M	Y	3	T	N
6	H	N	2	T	Y
7	H	N	2	S	N
8	H	N	2	T	N
9	H	Y	2	S	N
10	H	N	2	T	Y

- Greedy learn a decision tree using the ID3 algorithm and draw the tree.
- Write the learned concept for Martian as a set of conjunctive rules (e.g., if (green=Y and legs=2 and height=T and smelly=N), then Martian; else if ... then Martian;...; else Human).