Homework Assignment N°4

BML36
Thibault Douzon
Rajavarman Mathivanan

September 26th, 2018

Contents

1 Exercise 1: Decision Trees

1.1 Part a

0.51996

1.2 Part b

entropy of the dataset: 0.991

Feature a_1	+	-	p_+	p_-	entropy
T	3	1	$\frac{3}{4}$	$\frac{1}{4}$	0.811
F	1	4	$\frac{1}{5}$	$\frac{4}{5}$	0.722

new entropy for a_1: $0.811 \times \frac{4}{9} + 0.722 \times \frac{5}{9} = 0.762$
information gain of a1 = 0.229

Feature a_2	+	-	p_+	p_-	entropy
T	2	3	$\frac{2}{5}$	$\frac{3}{5}$	0.971
F	2	2	$\frac{2}{4}$	$\frac{2}{4}$	1

new entropy for a_2: $0.971 \times \frac{5}{9} + 1 \times \frac{4}{9} = 0.762$
information gain of a1 = 0.007