

Assignment 2 - Classification

Chris Uzokwe

CS 383

5/5/2021

Part 1:

CS 383 Assignment 2
Chris Uzokwe
5/5/2021

1) a. Sample entropy with (log base 2)

Y	X ₁	X ₂	Count	Prob
+	T	T	3	3/21
+	T	F	4	4/21
+	F	T	4	4/21
+	F	F	1	1/21
-	T	T	0	0/21
-	T	F	1	1/21
-	F	T	3	3/21
-	F	F	5	5/21

$H(Y) = -P(+|Y) \log_2 P(+|Y) - P(-|Y) \log_2 P(-|Y)$
 $= -(4/21) \log_2 (4/21) - (7/21) \log_2 (7/21)$
 $= 0.46 + 0.52 = 0.98 = H(Y)$

b. Information gain on X₁
 $IG(A) = H(\frac{P}{P+0}) - H(\frac{P}{P+0})$
 $E(H(A)) = 0.169 + 0.375 = 0.544$
 $H(A) = -\frac{7}{21} \log_2 \frac{7}{21} - \frac{1}{21} \log_2 \frac{1}{21}$
 $= 0.91$
 $IG(X_1) = 0.98 - 0.544 = 0.436$

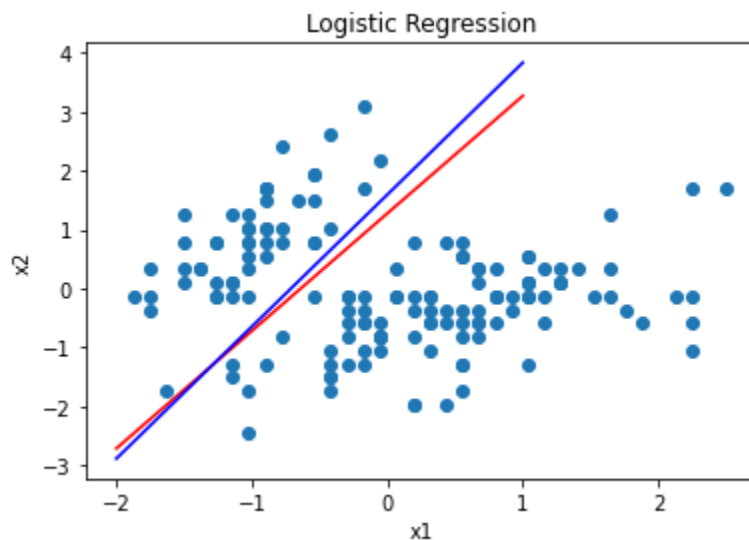
c. $H(X_1) = -\frac{12}{21} \log_2 \frac{12}{21} - \frac{9}{21} \log_2 \frac{9}{21}$
 $= 0.91$

2. a) $P(A=1|X_1) = 0.6$
 $P(A=0) = 0.4$
b) $Chrs = N(21, 1)$
 $\log_2 21 = 4.392$

$H(X_1) = -\frac{12}{21} \log_2 \frac{12}{21} - \frac{9}{21} \log_2 \frac{9}{21}$
 $= 0.91$

$IG(X_1) = 0.98 - 0.544 = 0.436$

Part 2:



thetas = array([8.34583606, 13.06817275, -6.55300008])

Part 3:

```
acc = 0.9272905119008803  
precision = 0.9267676767676768  
recall = 0.8900565885206144  
f1 = 0.9080412371134021
```

Part 4:

```
acc = 0.6733333333333333  
precision = 0.6711409395973155  
recall = 1.0  
f1 = 0.8032128514056226
```

Part 5:

Stats not recorded in time but partial code available