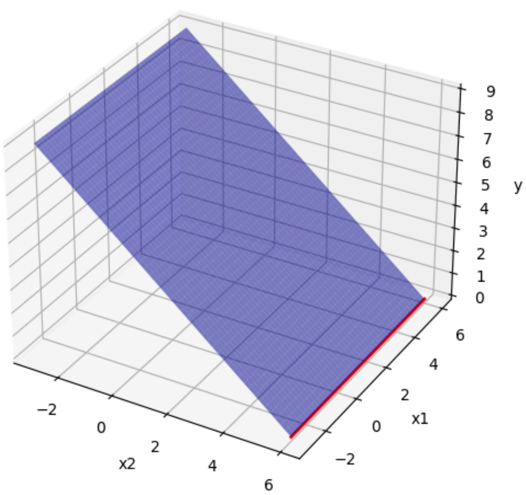
طراح: محمد امانلو، مهدی جمال‌خواه زمان آزمون: ۲۰ دقیقه

پرسش اول

فرض کنید که یک طبقه‌بند logistic regression را آموزش می‌دهید که تابع فرض آموزش دیده در آن به صورت $h_\theta(x) = \sigma(\theta_0 + \theta_1 x_1 + \theta_2 x_2)$ است که داریم $\theta_0 = 6, \theta_1 = 0, \theta_2 = -1$ است. مرز تصمیم برای $h_\theta(x)$ را در فضای دوبعدی x_1, x_2 رسم کنید و توضیح دهید که چگونه آن مرز بدست آمده است.

پاسخ:



خط تقاطع منحنی $y = 0$ و منحنی که به عنوان ورودی به تابع σ داده می‌شود همان مرز جدا کننده دو کلاس است. زیرا تابع σ در نقطه 0 برابر $\frac{1}{2}$ می‌شود یعنی احتمال متعلق بودن به هر دو کلاس برابر است، و این همان مرز جداکننده است. (به نظرتان چگونه می‌توانیم این مرز را غیر خطی کنیم؟)

پرسش دوم

مجموعه داده ساده زیر را در نظر بگیرید، می‌خواهیم از این مجموعه داده برای تشخیص آنکه هر دانشجو با مقدار مطالعه در طول ترم داده شده، آیا درس هوش مصنوعی را پاس خواهد کرد و یا نه، استفاده کنیم. ابتدا مقادیر انتروپی‌های داده شده را محاسبه کرده و سپس درخت تصمیم مربوطه را تشکیل دهید. (راهنمایی: $(\log_2 3 = 1.6$

1. $H(Passed)$
2. $H(Passed|GPA)$
3. $H(Passed|Studied)$

GPA	Studied	Passed
L	F	F
L	T	T
M	F	F
M	T	T
H	F	T
H	T	T

پاسخ:

ابتدا به محاسبه انتروپی‌ها می پردازیم:

.1

$$H(Passed) = -(\frac{2}{6}\log_2\frac{2}{6} + \frac{4}{6}\log_2\frac{4}{6})$$
$$H(Passed) = -(\frac{1}{3}\log_2\frac{1}{3} + \frac{2}{3}\log_2\frac{2}{3})$$
$$H(Passed) = \boxed{\log_2 3 - \frac{2}{3} \approx 0.92}$$

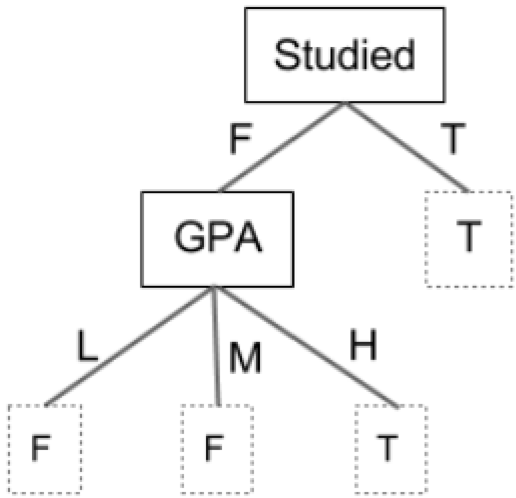
.2

$$H(Passed|GPA) = -\frac{1}{3}(\frac{1}{2}\log_2\frac{1}{2} + \frac{1}{2}\log_2\frac{1}{2}) - \frac{1}{3}(\frac{1}{2}\log_2\frac{1}{2} + \frac{1}{2}\log_2\frac{1}{2}) - \frac{1}{3}(1\log_2 1)$$
$$H(Passed|GPA) = \frac{1}{3}(1) + \frac{1}{3}(1) + \frac{1}{3}(0)$$
$$H(Passed|GPA) = \boxed{\frac{2}{3} \approx 0.66}$$

.3

$$H(Passed|Studied) = -\frac{1}{2}(\frac{1}{3}\log_2\frac{1}{3} + \frac{2}{3}\log_2\frac{2}{3}) - \frac{1}{2}(1\log_2 1)$$
$$H(Passed|Studied) = \frac{1}{2}(\log_2 3 - \frac{2}{3}$$
$$H(Passed|Studied) = \boxed{\frac{1}{2}\log_2 3 - \frac{1}{3} \approx 0.46}$$

ما می‌خواهیم ابتدا بر روی متغیری تقسیم کنیم که بیشترین Information gain را دارد، یعنی باید مقدار $H(Passed) - H(Passed|A)$ را به ازای تمامی A های ممکن محاسبه کنیم. این بدان معنی است که ما کمترین مقدار $H(Passed|A)$ را در نظر بگیریم. بنابراین باید ابتدا بر اساس متغیر "Studied" تقسیم کنیم. در نتیجه درخت نهایی به شکل زیر خواهد شد.



پرسش سوم

مجموعه داده زیر را در نظر بگیرید که در آن متغیر y برچسب است و متغیرهای A, B, C ویژگی‌های باینری هستند. طبقه‌بند بیز ساده (Naive bayes)، نمونه‌های (0,0,1) و (1,1,1) را چگونه دسته‌بندی می‌کند؟ (در حالت تساوی احتمال، برچسب صفر ارجح است.)

A	B	C	y
0	0	1	0
0	1	0	0
1	1	0	0
0	0	1	1
1	1	1	1
1	0	0	1
1	1	0	1

پاسخ:

$$P(y=0)=\frac{3}{7} \qquad P(y=1)=\frac{4}{7}$$

$$P(A=0|y=0)=\frac{2}{3} \qquad P(A=0|y=1)=\frac{1}{4}$$

$$P(B=0|y=0)=\frac{1}{3} \qquad P(B=0|y=1)=\frac{1}{2}$$

$$P(C=0|y=0)=\frac{2}{3} \qquad P(C=0|y=1)=\frac{1}{2}$$

$$P(0,0,1|y=0)=\frac{2}{3}\times\frac{1}{3}\times\frac{1}{3} \qquad P(0,0,1|y=1)=\frac{1}{4}\times\frac{1}{2}\times\frac{1}{2}$$

$$P(1,1,1|y=0)=\frac{1}{3}\times\frac{2}{3}\times\frac{1}{3} \qquad P(1,1,1|y=1)=\frac{3}{4}\times\frac{1}{2}\times\frac{1}{2}$$

$$\frac{P(y=0|0,0,1)}{P(y=1|0,0,1)}=\frac{\frac{2}{3}\times\frac{1}{3}\times\frac{1}{3}\times\frac{3}{7}}{\frac{1}{4}\times\frac{1}{2}\times\frac{1}{2}\times\frac{4}{7}}=\frac{8}{9}<1\Rightarrow y^*=1$$

$$\frac{P(y=0|1,1,1)}{P(y=1|1,1,1)}=\frac{\frac{1}{3}\times\frac{2}{3}\times\frac{1}{3}\times\frac{3}{7}}{\frac{3}{4}\times\frac{1}{2}\times\frac{1}{2}\times\frac{4}{7}}=\frac{8}{27}<1\Rightarrow y^*=1$$