

(۱) مراحل ایجاد یک سیستم یادگیر با نظارت را با مثال توضیح دهید؟

(۲) با یک مثال شرح دهید در چه هنگام ممکن است یک طبقه بندی کننده با خطای طبقه بندی (missclassification error) بیشتر را برای حل یک مساله انتخاب کنیم

(۳) فرض کنید یک مجموعه از داده ها بصورت i.i.d با مقادیر حقیقی $x_1, x_2, x_3, \dots, x_n$ از یک pdf به شکل زیر استخراج شده اند

$$f(x) = \begin{cases} \alpha e^{-\alpha x} & x \geq 0 \\ 0 & x < 0 \end{cases}$$

به طوریکه α پارامتر ناشناخته این توزیع می باشد. توسط تخمین متد Maximum Likelihood پارامتر α را تعیین کنید. (فرض کنید که $x \geq 0$ است)

(۴) برای داده های آموزشی زیر یک درخت تصمیم بسازید. جدول entropy بصورت تقریبی بوده و مقادیر احتمالات بدست آمده را با round کردن به اولین کوچکترین عددی که در جدول می باشد، در نظر بگیرید.

x_1	x_2	x_3	class
0	0	1	+
0	0	2	+
0	0	3	+
0	0	4	+
0	1	1	-
0	1	2	-
0	1	3	-
1	0	4	-
1	1	1	+
1	1	2	+

P	Entropy(p)
0.0	0.0
0.1	0.5
0.2	0.7
0.3	0.9
0.4	1.0
0.5	1.0
0.6	1.0
0.7	0.9
0.8	0.7
0.9	0.5

(۵) Version Space چیست؟ اگر در یادگیری یک طبقه بندی کننده False Negative خسارت بیشتری نسبت به False Positive داشته باشد کدام فرضیه از Version Space را انتخاب می کنیم.

(۶) اگر برای داده های آموزشی زیر از دسته بندی کننده ساده استفاده شود کلاس هر یک از داده های زیر را مشخص کنید.

Customer ID	Gender	Car Type	Shirt Size	Class
1	M	Family	Small	C0
2	M	Sports	Medium	C0
3	M	Sports	Medium	C0
4	M	Sports	Large	C0
5	M	Sports	Extra Large	C0
6	M	Sports	Extra Large	C0
7	F	Sports	Small	C0
8	F	Sports	Small	C0
9	F	Sports	Medium	C0
10	F	Luxury	Large	C0
11	M	Family	Large	C1
12	M	Family	Extra Large	C1
13	M	Family	Medium	C1
14	M	Luxury	Extra Large	C1
15	F	Luxury	Small	C1
16	F	Luxury	Small	C1
17	F	Luxury	Medium	C1
18	F	Luxury	Medium	C1
19	F	Luxury	Medium	C1
20	F	Luxury	Large	C1

a) Gender = 'M', car type = 'Luxury', Shirt size = 'medium'

b) Gender = 'F', car type = 'Family', Shirt Size = 'Large'