Khadija Jumani (SP23-BDS-023)

Assignment #01-ML

Dr. Shaneela

Question # 01

Traning set given :-

	#	Label	Instance (pair)
X_1	1	+	{(male brown tall US) (female black short US)}
X_2	2	+	{(male brown short French) (female black short US)}
X_3	3	-	{(female brown tall German) (female black short Indian)}
X_4	4	+	{(male brown tall Irish) (female brown short Irish)}

@ Candidate - Elimination Algorithm :

$S_0 : \{(\phi, \phi, \phi, \phi), (\phi, \phi, \phi, \phi)\}$; $G_0 : \{(?,?,?,?) (?,?,?,?)\}$

X_1 :

$S_1 : \{(\text{male brown tall US}), (\text{female black short US})\}$

$G_1 : \{(?,?,?,?) (?,?,?,?)\}$

X_2 :

$S_2 : \{(\text{male, brown}, ?, ?) (\text{female black short US})\}$

$G_2 : \{(?,?,?,?), (?,?,?,?)\}$

X_3 : neg $(S_2 = S_3)$ remains same

S_3 : {(male, brown, ?, ?) (female black short US)}.

G_3 : {(male, ?, ?, ?) (?, ?, ?, ?), (?, ?, ?, ?) (?, ?, ?, US)} (pairwise *)

∴ S3 has (male) but neg has female

So, gender = male in g will exclude the neg & still cover S3

∴ S3 has (US) but neg has indian = US

So, nationality in g will exclude neg & still cover S3.

X_4 :

S_4 : {(male, brown, ?, ?) (female, ?, short, ?)}.

G_4 : {(male, ?, ?, ?) (?, ?, ?, ?)}

So, S_4 and G_4 are the final boundaries!

P.T.O

(b) How many distinct hypotheses from the given hypothesis space are consistent with the following single positive training example?

+ ((male black short Portuguese)
(female blonde tall Indian))

As for each person it is either ? or a specific value So;
Combined consistent hypothesis are:

$$2^4 \times 2^4 = 2^8 = 256$$

(c) As no. of queries = 8 ;
and consistent hypothesis = 256 ;
So, minimal # of binary answers required to pick one hypothesis out of 256 possibilities is:

$$\boxed{\log_2(256)}$$

thus:-

Question # 02

Consider the given example :

Example	citations	size	inLib	price	editions	buy
1	some	small	no	afford	many	no
2	many	big	no	expen	one	Yes
3	some	big	always	expen	few	no
4	many	medium	no	expen	many	Yes
5	many	small	no	afford	many	Yes

(a) Concepts are possible for this Ins. Space?

Instance Space = 2 x 3 x 2 x 2 x 3 = 72

(b) Hypothesis expressed?

Syntactically Distinct hyp = 4 x 5 x 4 x 4 x 5

= 1600

(c) Semantically Distinct Hyp?

S.D. hyp = 1 + (3 x 4 x 3 x 3 x 4)

= 433

(d) Find S Algorithm :

$h_0 = \langle \phi, \phi, \phi, \phi, \phi \rangle$

$\therefore h_1$ ignored due to being neg.

$h_2 = \langle$ many, big, no, exp, one $\rangle$

$\therefore h_3$ ignored

$h_4 = \langle$ many, ?, no, exp, ? $\rangle$

$h_5 = \langle$ many, ?, no, ?, ? $\rangle$ → maximally specified hypothesis

Question # 03

Consider the following example

	Size	Color	Shape	Class/Label
X_1	Big	Red	Circle	No
X_2	Small	Red	Triangle	No
X_3	Small	Red	Circle	Yes
X_4	Big	Blue	Circle	No
X_5	Small	Blue	Circle	Yes

(a) Find all possible hypotheses for the given dataset using Candidate Elimination Algo

$S_0 : (\phi, \phi, \phi)$; $G_0 : (?, ?, ?)$

- X_1 :- (Neg : $G \rightarrow S$)

$G_1 : \{ < (small, ?, ?), (?, blue, ?), (?, ?, triangle) > \}$

write the opposite ones

$S_1 : < \phi, \phi, \phi >$ → remains same

It will be consistent as class is No

- X_2 :- (Neg : $G \rightarrow S$)

hai (small, ?, ?) inconsistent

$G_2 : < (small, blue, ?), (small, ?, circle), (?, blue, ?), (big, ?, triangle), (?, blue, triangle) >$

consistent

check each hypo of G_1 if consistent with X_2 write same else

$S_2 : < \phi, \phi, \phi >$

- X_3 :- (Positive → G → S)

S_3 : (small, red, circle)

G_3 : (Small, ?, Circle)

[∴ compare X_3 with G_3 and write the consistent ones]

- X_4 :- (Neg : S → G)

consistent

G_4 : (Small, ?, Circle)

[∴ Compare G_3 with X_4 and write consistent same & non-consistent in sub-groups]

S_4 : (Small, Red, Circle)

↳ consistent (S_3 with X_4)

So written same

- X_5 :- (pos : G → S)

(X_5 not consistent with S_4 so write ?)

S_5 : (Small, ?, Circle)

G_5 : (Small, ?, Circle) ∴ G_3 consistent to X_5

S = G : (small, ?, Circle) Ans

Question # 04

Consider the following :

	X_1	X_2	X_1 OR X_2
1.	0	0	0
2.	0	1	1
3.	1	0	1

(a) Find the version Space for the given example using list-then-eliminate Algo.

$X_1 \rightarrow 0, 0, 1$

$X_2 \rightarrow 0, 1, 0$

⊙ **Instance Space** : → 2 classifications of each features!

$2 \times 2 =$ 4 possibilities → 0, 1 and 0, 1

$X = \{(0,0), (0,1), (1,0), (1,1)\}$

⊙ **Hypothesis Space** : → $4 \times 4 = 16$ possibilities !

$X_1 \rightarrow 0, 1, ?, \phi$

$X_2 \rightarrow 0, 1, ?, \phi$

$(0,0), (0,1), \ldots (0,?), (0,\phi), \ldots$

$\ldots, (1,0), (1,1), \ldots, (1,?), (1,\phi),$

$(?,0), (?,1), \ldots, (?,?), (?,\phi), (\phi,0), (\phi,1)$

$(\phi,?), (\phi,\phi)$

(a) Version Space :

(0,0), (0,1), ..., (0,?), ..., (1,1), (1,0), (1,?), (?,0), (?,1), ..., (?,?), (ϕ,ϕ)

Now, training Set Given :-

X_1	X_2	X_1 OR X_2
0	0	0 (neg)
0	1	1 (pos)
1	0	1 (pos)

So the version Space consisting of all consistent hypotheis will be Version Space

$$\boxed{< (?, ?) >}$$

Ans

(b) Find the outcome of each hypothesis in version space for (1,1) input.

The outcome for the (1,1) can be both 0 or 1 as;

(?,?) can accept any outcome.

$$\boxed{? = 0 \text{ or } 1}$$