CS 6375

ASSIGNMENT _INDUCTIVE LEARNING__________

Names of students in your group:

Pallavi Pandey-PXP170009

Number of free late days used: _0___________________

Note: You are allowed a **total** of 4 free late days for the **entire semester**. You can use at most 2 for each assignment. After that, there will be a penalty of 10% for each late day.

Please list clearly all the sources/references that you have used in this assignment.

1. Classroom slides

Q1.

```python
'''

theta0=0.0
theta1=1.0
r=0.0
temp=0.0
num=0.0

x=[3,1,0,4]
y=[2,2,1,3]


for itr in range(5):

    for z,p in zip(x,y):

        r+=(theta0+(theta1*z))-p
        temp+=((theta0+(theta1*z))-p)*z
        num+=((theta0+(theta1*z))-p)**2

    theta0=(theta0-((0.1/4)*(r)))
    theta1=(theta1-((0.1/4)*(temp)))

    error=((1/8)*num)
    print("Theta0 is" ,theta0)
    print("Theta1 is", theta1)
    print("Error is",error)
    r=0.0
    temp=0.0
    num=0.0
```

```
<terminated> test.py [C:\Users\user\AppData\Local\Programs\Python\Python37-32\python.exe]
Theta0 is 0.0
Theta1 is 0.85
Error is 0.5
Theta0 is 0.030000000000000006
Theta1 is 0.7975
Error is 0.3481249999999999
Theta0 is 0.0675
Theta1 is 0.7731250000000001
Error is 0.3178203124999999
Theta0 is 0.106125
Theta1 is 0.75709375
Error is 0.2986223632812499
Theta0 is 0.14409375
Theta1 is 0.7437578125
Error is 0.28147623168945307
```

In the first image, when the value of step size is 0.1, the error decreases for each iteration but the point of local minima is not achieved.

```python
'''
Created on Sep 8, 2018

@author: user
'''

theta0=0.0
theta1=1.0
r=0.0
temp=0.0
num=0.0

x=[3,1,0,4]
y=[2,2,1,3]


for itr in range(5):

    for z,p in zip(x,y):

        r+=(theta0+(theta1*z))-p
        temp+=((theta0+(theta1*z))-p)*z
        num+=((theta0+(theta1*z))-p)**2

    theta0=(theta0-((0.294/4)*(r)))
    theta1=(theta1-((0.294/4)*(temp)))

    error=((1/8)*num)
    print("Theta0 is" ,theta0)
    print("Theta1 is", theta1)
    print("Error is",error)
    r=0.0
    temp=0.0
    num=0.0
```

```
<terminated> test.py [C:\Users\user\AppData\Local\Programs\Python\Python37-32\python.exe]
Theta0 is 0.0
Theta1 is 0.559
Error is 0.5
Theta0 is 0.2593079999999999
Theta1 is 0.9607509999999999
Error is 0.4705632499999999
Theta0 is 0.20614986000000005
Theta1 is 0.44228273500000026
Error is 0.4593982330512497
Theta0 is 0.4734795529799999
Theta1 is 0.9458643107349998
Error is 0.4656340926473543
Theta0 is 0.3661083496917001
Theta1 is 0.3299116357681754
Error is 0.48914831251029606
```

When the step size is 0.294, the error first decreases and then increases. Thus, an increase in step size helps to achieve the point of local minima faster.

Q2.
Let the total no of people be =100
Total no of people who test positive and also have the disease (True Positive)= 80% or 80 people.
Total no of people who test negative and also don't have the disease(True Negative)= 90% or 90 people
Total no of people who test negative and also have the disease(False negative) = 100-90= 10 people
Or (10/100)X100
False Negative =10%

Total no of people who test positive and also don't have the disease(False positive)= 100-80= 20 people
Or (20/100)X100
False Positive =20%

Q3.
Selecting the most specific hypothesis(S) based on the training data:
Pros: All the positive consistent hypothesis are included in S as it includes all positive examples.
Cons: Since the negative examples are ignored in this hypothesis, it can't tell us if the training data is inconsistent. There is no scope for generalization.

Selecting the most general hypothesis (G) based on the training data:

Pros: It helps us fit many hidden instances as it is the most generalized value.

Cons: A very generalized hypothesis can increase chances of false positive values.

Q4.

Consistent Hypothesis: A hypothesis h can be said to be consistent training examples D of the target concept c if it is of the form h(x)=c(x).

So, for each training example (x,c(x)) in D, hypothesis is consistent when h(x)=c(x)

(h, D) = for every(x, c(x)) belongs to D :: h(x) = c(x)

Version Space: The version space w.r.t Hypothesis H and the training data D will be the subset of the hypothesis from H consistent with D.

VS= {hIh belong to H and is consistent {h,D}}

Each member of Version Space lies between the general and specific boundary

Q5.

The most general hypothesis has **don't care (?)** value for each attribute.

Q6.

a) For the first instance x1, 2 values are possible (GPA>3.5 or GPA<3.5).
Similarly,

For the second instance x2, 2 values are possible (student has taken CS 6375 or not).
For the third instance x3, 2 values are possible (student has taken CS 6350 or not).
For the fourth instance x4, 2 values are possible(if work experience>2 or work experience<2)
Thus, the total number of instances can be =2X2X2X2= **2^4 =16**

b) Each instance has 2 possible labeling available, that is, so total labeling possible = **2^16**
c) For each attribute, there can be 3 possible choices, that is, 0 or 1 or ?. Thus, total choices (c) can be 3^4=81. Now, each hypothesis can accept the choices as both positive or negative, that is, c or –c. So, total hypothesis can be **2^(3^4) or 2^81.**
d) We can select or arrange only2 attributes out of 4. So, 4P2= 4!/2!=**12**
e) **2^4=16** ways to label the leaf nodes of each of the decision tree.

Q7.

x1=1

~x1=0

S={null,null,null,null,null}

S1={x1, x1,~ x1,x1,x1 }

The iterations 2nd is ignored as it has the negative values.

S3={x1,x1,~x1,x1,~x1}

The iteration 4th is ignored as it has the negative values.

S5={x1,x1,~x1,x1,~x1}

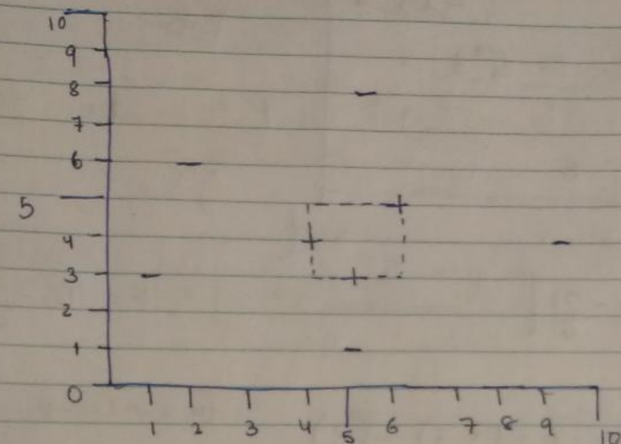
Q8.

Positive: (GPA<3.5 /\Exp>=3)V(GPA>=3.5 /\Exp>=1)

Negative: (GPA<3.5 /\Exp<3)V(GPA>=3.5 /\Exp<1)

Q9.

Q9

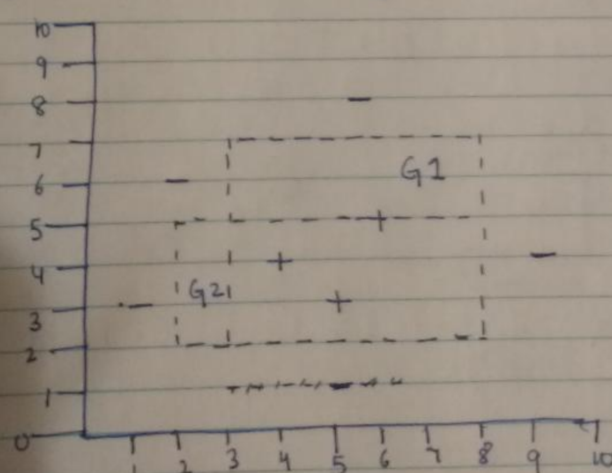


a) $S \equiv (4 \le x \le 6) \wedge (3 \le y \le 5)$

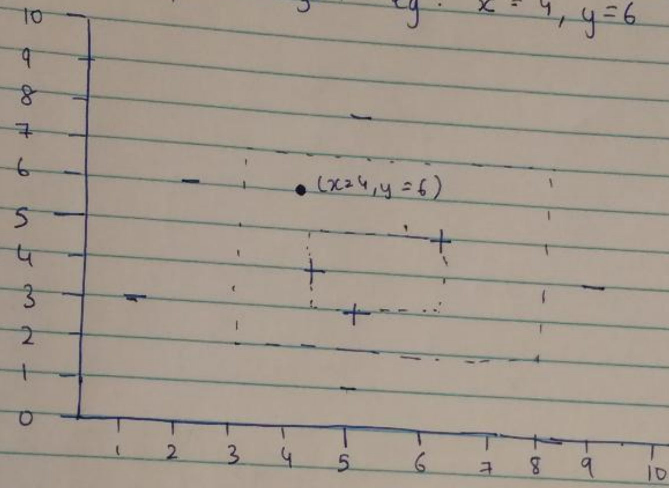
b) ~~Theare~~ There will be 2 general boundaries

$G_1 \equiv (3 \le x \le 8) \wedge (2 \le y \le 7)$

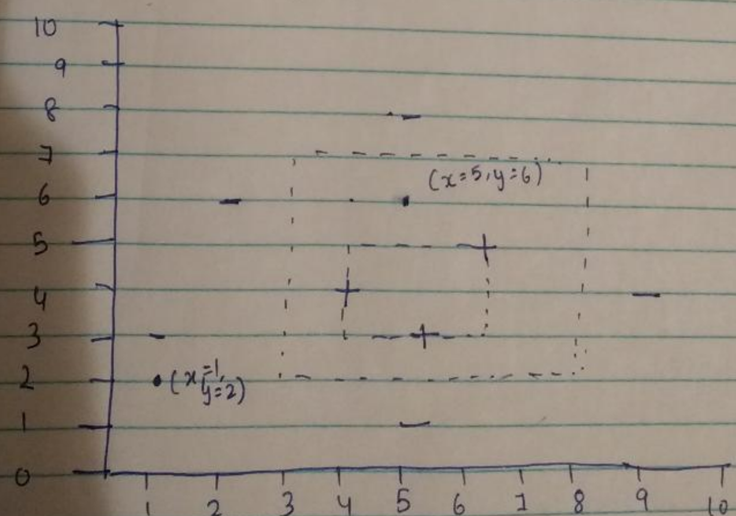
$G2 \equiv (2 \le x \le 8) \wedge (2 \le y \le 5)$



c) A query lying b/w S and G will reduce the version space size. eg: $x = 4, y = 6$



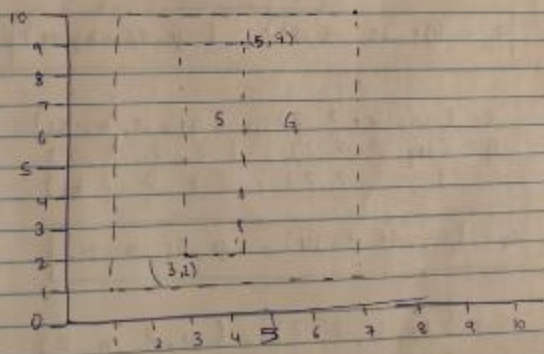
A query outside G or inside S would not reduce the size of the V.S eg: $x = 1, y = 2$ and $x = 5, y = 6$



$y=2$ and $x=5$, $y=4$

d) For a concept like $3 \le x \le 5$, $2 \le y \le 9$ as an example. 4 training examples would be needed and would be sufficient

+ (3, 2)
+ (5, 9)
− (1, 1)
− (7, 10)



Q10.

Q10.

a) $S = \{(\phi, \phi, \phi, \phi), (\phi, \phi, \phi, \phi)\}$
$G = \{(?, ?, ?, ?), (?, ?, ?, ?)\}$

$\boxed{x_1 = (ug, se, l, hs), (gr, cs, h, hs)}$ +

$S = (ug, se, l, hs), (gr, cs, h, hs)$
$G = (?, ?, ?, ?), (?, ?, ?, ?)$

$\boxed{x_2 = (ug, se, h, fr), (gr, cs, h, hs)}$ +

$S = (ug, se, ?, ?), (gr, cs, h, hr)$
$G = (?, ?, ?, ?), (?, ?, ?, ?)$

$\boxed{x_3 = (gr, se, l, so), (gr, cs, h, se)}$ −

$S = (ug, se, ?, ?), (gr, cs, h, hs)$
$G = (ug, ?, ?, ?), (?, ?, ?, ?)$
$(?, ?, ?, ?), (?, ?, ?, hs)$

$\boxed{x_4 = (ug, se, l, ju), (gr, se, h, ju)}$ +

$S = (ug, se, ?, ?), (gr, ?, h, ?)$
$G = (ug, ?, ?, ?), (?, ?, ?, ?)$

b) Total consistent hypothesis on running candidate elimination for S and G are:-

(ug, se, ?, ?) (gr, ?, h, ?)
(ug, ?, ?, ?) (?, ?, h, ?)
(ug, se, ?, ?) (gr, ?, ?, ?)
(ug, ?, ?, ?) (gr, ?, h, ?)
(ug, se, ?, ?) (?, ?, h, ?)
(ug, se, ?, ?) (?, ?, ?, ?)
(ug, ?, ?, ?) (gr, ?, ?, ?)
(ug, ?, ?, ?) (?, ?, ?, ?)

For the following data
(ug, cs, h, do), (gr, ma, l, se) ⊕
only 2 hypothesis are consistent

(ug, ?, ?, ?), (?, ?, ?, ?)
(ug, ?, ?, ?), (gr, ?, ?, ?)