

Inductive learning

CS 6375 ASSIGNMENT-1



September 9, 2018

fall -2018

THE UNIVERSITY OF TEXAS AT DALLAS

SAIPRAVEEN VABBILISETTY – SXV165130

LIKITHA KOMMINENI - LXK162430

1. Gradient Descent is decreasing the Error function and finding the local optimum by updating the theta variables.

Equations for updating Theta 0 and Theta 1 are:

Given,

To find Error Function,

To find updated ϴ,

The α value considered is **0.3**

The updated values are obtained from the following program

import math

def gradientDescent1(x, y, t1,t2, alpha, m):

add = 0

for i in range(0, m):

add = add + ((t1+t2\*x[i])- y[i])

mul = float((1/m)\*add)

thetaNew = t1 - alpha \* mul

return thetaNew

def gradientDescent2(x, y, t1,t2, alpha, m):

add = 0

for i in range(0, m):

add = add + (((t1+t2\*x[i])- y[i]) \* x[i])

mul = float((1/m)\*add)

thetaNew = t2 - alpha \* mul

return thetaNew

def find\_error(t1,t2,m,x,y):

add = 0.0

for i in range(0,m):

a = (t1+t2\*x[i])- y[i]

p = math.pow(((t1+t2\*x[i])- y[i]),2)

add = add + p

err = (1/(2\*m))\*float(add)

return err

x = [3,1,0,4]

y = [2,2,1,3]

m = 4

numIterations= 5

alpha = 0.3

theta1 = 0

theta2 = 1

print("alpha value is", alpha)

for i in range(0,numIterations):

error = find\_error(theta1,theta2,m,x,y)

print("------------------ITERATION " + str(i+1) + "------------------")

print ("Error is ", error)

theta1 = gradientDescent1(x, y, theta1,theta2, alpha, m)

theta2 = gradientDescent2(x, y, theta1,theta2, alpha, m)

print ("theta0 is : ", theta1)

print ("theta1 is : ", theta2)

Output for the above program is as follows:

alpha value is 0.3

------------------ITERATION 1------------------

Error is 0.5

theta0 is : 0.0

theta1 is : 0.55

------------------ITERATION 2------------------

Error is 0.48312499999999997

theta0 is : 0.26999999999999996

theta1 is : 0.8154999999999999

------------------ITERATION 3------------------

Error is 0.27070081249999983

theta0 is : 0.2997

theta1 is : 0.545455

------------------ITERATION 4------------------

Error is 0.2621245323312499

theta0 is : 0.482517

theta1 is : 0.6923075499999999

------------------ITERATION 5------------------

Error is 0.16563156915895813

theta0 is : 0.5223773700000001

theta1 is : 0.5288814055000001

We could see that the error function has decreased its value from 0.5 to 0.16 for the given conditions.

-----------------------------------------------------------------------------------------------------------------------------------------

1. False Positive is 10 % and False Negative is 20 %

-----------------------------------------------------------------------------------------------------------------------------------

1. a. Most specific hypothesis (S) based on a training data:

Pros:

* This type of hypothesis covers the observed positive training examples.
* We get more defined hypothetical space

Cons:

* if reduced any further, there are chances to miss out positive training examples, and hence leading to inconsistency
* if novel data that is never seen before is observed by a learner, it assumes it to be negative

b. Most general hypothesis (G) based on a training data are:

Pros:

* covers the observed positive training examples and covers as much of the remaining feature space without including any negative training examples
* Includes all the positive examples

Cons:

* if enlarged any further, there are chances to include negative training examples, and hence leading to inconsistency
* If a negative data already observed, learner assumes it to be positive

-------------------------------------------------------------------------------------------------------------------------------------------------

1. **Consistent Hypothesis:** A hypothesis h is consistent with a set of training examples D if and only if h(x) = c(x) for each example (x, c(x)) in D.

**Version Space:** The version space, denoted , with respect to hypothesis space H and training examples D, is the subset of hypotheses from H consistent with the training examples in D.

------------------------------------------------------------------------------------------------------------------------------------------------------

1. The most generic hypothesis has “**?**” value for each attribute.

-----------------------------------------------------------------------------------------------------------------------------------

1. A. The total number of instances possible are 16. Since 4 attributes, the total number of Boolean combinations possible are 2^4.

B. For conjunctive hypothesis, there are 4 possible choices for each attribute ?, T, F,ᴓ. So possible hypothesis is 2^16.

C. There can be 81 such combinational hypothesis. Four attributes and 3 choices for each attribute.

D. The number of ways for selecting two attributes is 4C2 which is 6. Selecting the root and leaf node is 2 ways. So total combinations are 12 decision trees.

E. The total number of ways for labelling is 2^4 that is 16.

------------------------------------------------------------------------------------------------------------------------------------------

1. s0 = (ɸ, ɸ, ɸ, ɸ, ɸ)

Since first training data is positive, S is

s1 = (1, 1, 0, 1, 1)

Since second training data is negative we ignore that and make no changes in S boundary.

s2 = (1, 1, 0, 1, 1)

Third training data is positive, and we update the S boundary.

s3 = (1, 1, ?, 1, ?)

Since fourth training data is negative we ignore that and make no changes in S boundary.

s4 = (1, 1, ?, 1, ?)

Fifth training data is positive, and we update the S boundary.

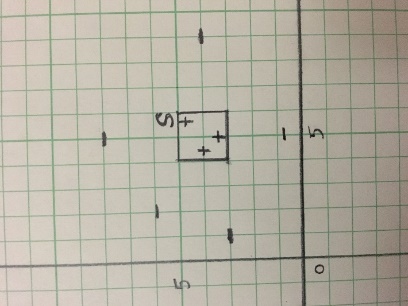
S5 = (1, 1, ?, 1, ?)

------------------------------------------------------------------------------------------------------------------------------------------

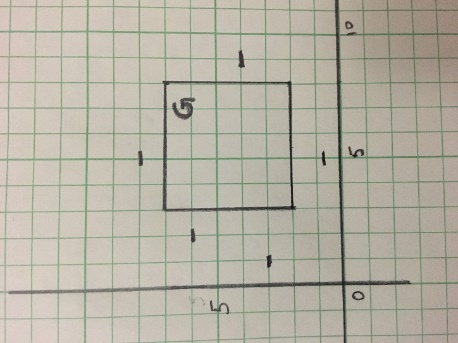
1. DNF f = ( All Data ʌ D2) ∨ (~All Data ʌ D1)

------------------------------------------------------------------------------------------------------------------------------------------

1. a. S-boundary = (4<=x<=6) ʌ (3<=y<=5)



b. G-boundary = (3<=y<=8) ʌ (2<=x<=7)



c. The query that is not is 5≤x≤6, 4≤y≤5 or any point covered in S boundary.

d. 4 is the smallest number of training examples I can provide to learn target concept perfectly

------------------------------------------------------------------------------------------------------------------------------------------

s0 = ((ɸ, ɸ, ɸ, ɸ), ( ɸ, ɸ, ɸ, ɸ))

G0 = ((?,?,?,?),(?,?,?,?))

S1 = ((ug, se, l, hs); (gr, cs,h,hs))

G1 = ((?,?,?,?),(?,?,?,?))

S2 = ((ug, se, ?, ?), (gr, cs,h,hs))

G2 = ((?,?,?,?),(?,?,?,?))

S3 = ((ug, se, ?, ?), (gr, cs,h,hs))

G3 = (((ug,?,?,?),(?,?,?,?)), ((?,?,?,?),(?,?,?,hs)))

S4 = ((ug, se, ?, ?), (gr, ?, h, ?))

G4 = ((ug,?,?,?),(?,?,?,?))

b. Six hypothesis total consistent hypotheses are returned after running the Candidate Elimination algorithm.

Two hypotheses are consistent with the given data point