MACHINE LEARNING LAB

July 30, 2020

Manish Singh CS3
Roll No. 17/466

Computer Science and Engineering Department
Rajasthan Technical University, Kota

Contents

1 Problem Statement

For a given set of training data examples stored in a .CSV file, implement and demonstrate the Candidate Elimination algorithm to output a description of the set of all hypotheses consistent with the training examples.(enjoysport.csv)

2 Candidate Elimination algorithm

1. Initialize h to the most specific hypothesis in S
2. Initialize h to the most General hypothesis in G
3. For each example in data set
    If (e is a positive example) then
        Elements of G that classify e as negative are removed from G
        Each element s of S that classifies e as negative is removed and replaced by the minimal generalizations of s that classify e as positive and are less general than some member of G
            Non-maximal hypotheses are removed from S;
    Else Elements of S that classify e as positive are removed from S
        Each element g of G that classifies e as positive is removed and replaced by the minimal specializations of g that classifies e as negative and are more general than some member of S
            Non-minimal hypotheses are removed from G

3 Training Set

sky	airtemp	humidity	wind	water	forcast	enjoysport
sunny	warm	normal	strong	warm	same	yes
sunny	warm	high	strong	warm	same	yes
rainy	cold	high	strong	warm	change	no
sunny	warm	high	strong	cool	change	yes

4 Program

```python
import numpy as np
import pandas as pd

data = pd.DataFrame(pd.read_csv('enjoysport.csv'))
concepts = np.array(data.iloc[:,:-1])
target = np.array(data.iloc[:,-1])

def learn(concepts, target):
specific_h = concepts[0].copy()
general_h = [["?" for i in range(len(specific_h))] for i in range(len(specific_h))]

for i, h in enumerate(concepts):
if target[i] == "yes":
for x in range(len(specific_h)):
if h[x] != specific_h[x]:
specific_h[x] = '?'
general_h[x][x] = '?'
print(" \n\nFor Training instance No:{0} the hypothesis is\n".format(i))
print("Specific Hypothesis: ",specific_h)
print("General Hypothesis: ",general_h,)
if target[i] == "no":
for x in range(len(specific_h)):
if h[x] != specific_h[x]:
general_h[x][x] = specific_h[x]
else:
general_h[x][x] = '?'
print(" \n\nFor Training instance No:{0} the hypothesis is\n".format(i))
print("Specific Hypothesis: ", specific_h)
print("General Hypothesis: ",general_h,)
indices = [i for i,val in enumerate(general_h) if val == ['?', '?', '?', '?', '?', '?']]

for i in indices:
general_h.remove(['?', '?', '?', '?', '?', '?'])

return specific_h, general_h
print("*"*20,"Costumer Elimination Algorithm","*"*20)
s_final, g_final = learn(concepts, target)
print("Final Specific hypothesis:", s_final)
print("Final General hypothesis:", g_final)
```

5 Output


```
******************** Candidate Elimination Algorithm ********************


For Training instance No:0 the hypothesis is

Specific Hypothesis:  ['sunny' 'warm' 'normal' 'strong' 'warm' 'same']
General Hypothesis:  [['?', '?', '?', '?', '?', '?'], ['?', '?', '?', '?', '?', '?'],
'?', '?', '?', '?', '?'], ['?', '?', '?', '?', '?', '?'], ['?', '?', '?', '?', '?', '


For Training instance No:1 the hypothesis is

Specific Hypothesis:  ['sunny' 'warm' '?' 'strong' 'warm' 'same']
General Hypothesis:  [['?', '?', '?', '?', '?', '?'], ['?', '?', '?', '?', '?', '?'],
'?', '?', '?', '?', '?'], ['?', '?', '?', '?', '?', '?'], ['?', '?', '?', '?', '?', '


For Training instance No:2 the hypothesis is

Specific Hypothesis:  ['sunny' 'warm' '?' 'strong' 'warm' 'same']
General Hypothesis:  [['sunny', '?', '?', '?', '?', '?'], ['?', 'warm', '?', '?', '?'
['?', '?', '?', '?', '?', '?'], ['?', '?', '?', '?', '?', '?'], ['?', '?', '?', '?',


For Training instance No:3 the hypothesis is

Specific Hypothesis:  ['sunny' 'warm' '?' 'strong' '?' '?']
General Hypothesis:  [['sunny', '?', '?', '?', '?', '?'], ['?', 'warm', '?', '?', '?'
['?', '?', '?', '?', '?', '?'], ['?', '?', '?', '?', '?', '?'], ['?', '?', '?', '?',
Final Specific hypothesis: ['sunny' 'warm' '?' 'strong' '?' '?']
Final General hypothesis: [['sunny', '?', '?', '?', '?', '?'], ['?', 'warm', '?', '?'
```

6 Result

We have found the most specific hypothesis and most genral hypothesis based on a given set of training data samples using Candidate Elimination algorithm.