Practical 2

Write a program for Candidate Elimination algorithm for finding the consistent version space based on a given set of training data samples. The training data is read from a .CSV file.

Program:-

```
import numpy as np
import pandas as pd

df = pd.read_csv('dataset.csv')
data = df.drop('EnjoySport',axis='columns')
target = df.EnjoySport

def candidate_elimination(concepts, target):
    specific_h = concepts[0].copy()
    print("Initialization Of Specific_h And General_h\n")
    print(specific_h)
    general_h = [["?" for i in range(len(specific_h))] for i in
range(len(specific_h))]
    print(general_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] ='?'
        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] = '?'

    indices = [i for i, val in enumerate(general_h) if val == ['?',
'?', '?', '?', '?', '?']]
    for i in indices:
        general_h.remove(['?', '?', '?', '?', '?', '?'])
    return specific_h, general_h
```

```
s_final, g_final = candidate_elimination(data, target)
print("\n\nFinal Specific_h:", s_final, sep="\n")
print("\n\nFinal General_h:", g_final, sep="\n")
```

OUTPUT

```
Initialization Of Specific_h And General_h

['Sunny' 'Warm' 'Normal' 'Strong' 'Warm' 'Same']
[['?', '?', '?', '?', '?', '?'], ['?', '?', '?', '?', '?', '?'], ['?', '?', '?', '?', '?', '?'], ['?', '?', '?',
'?', '?', '?'], ['?', '?', '?', '?', '?', '?'], ['?', '?', '?', '?', '?', '?']]


Final Specific_h:
['Sunny' 'Warm' '?' 'Strong' '?' '?']


Final General_h:
[['Sunny', '?', '?', '?', '?', '?'], ['?', 'Warm', '?', '?', '?', '?']]
```