**U19EC046 | ML | LAB 1**

**AIM**

Implement and demonstrate the FIND-S algorithm for finding the most specific hypothesis based on a given set of training data samples. Read the training data from a .CSV file.

**THEORY:**

In Machine Learning, concept learning can be termed as “*a problem of searching through a predefined space of potential hypothesis for the hypothesis that best fits the training examples”* ​

In order to understand Find-S algorithm, you need to have a basic idea of the following concepts as well:​

* Concept Learning​
* General Hypothesis​
* Specific Hypothesis​

**Concept Learning​**

Machines can also learn from concepts to identify whether an object belongs to a specific category or not. Any algorithm that supports concept learning requires the following:​

* Training Data​
* Target Concept​
* Actual Data Objects

**General Hypothesis​**

Hypothesis, in general, is an explanation for something. The general hypothesis basically states the general relationship between the major variables. For example, a general hypothesis for ordering food would be *I want a burger.*​

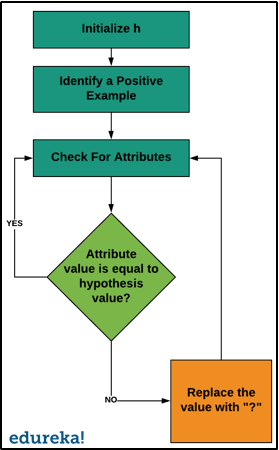
* G = { ‘?’, ‘?’, ‘?’, …..’?’}

**Specific Hypothesis​**

The specific hypothesis fills in all the important details about the variables given in the general hypothesis. The more specific details into the example given above would be *I want a cheeseburger with a lot of lettuce.* ​

* S = {‘Φ’,’Φ’,’Φ’, ……,’Φ’}

**ALGORIHM**



**CODE**

|  |
| --- |
| **import numpy as np**  **import pandas as pd**  **data = pd.read\_csv('./Data.csv')**  **def seriesToList(series):**  **return [value for \_, value in series.items()]**  **def initialHypothesis(df):**  **for i, r in df.iterrows():**  **if r[-1] == 'Yes':**  **return seriesToList(r)**  **def updateHypothisis(h, curr):**  **currList = seriesToList(curr)**  **if currList[-1] == 'Yes':**  **for i in range(len(h)):**  **if h[i] != '?' and h[i] != currList[i]:**  **h[i] = '?'**  **h = initialHypothesis(data)**  **print(f"initial Hypothesis : {h}")**  **for i, r in data.iterrows():**  **updateHypothisis(h, r)**  **print(f"final hypothesis : {h[:-1]}")** |

**OUTPUT**

|  |
| --- |
| initial Hypothesis :  ['Morning', 'Sunny', 'Warm', 'Yes', 'Mild ', 'Strong', 'Yes']  final hypothesis :  ['?', 'Sunny', '?', 'Yes', '?', '?'] |

**CONCLUSION**

In this practical we have implemented Python code for FIND-S algorithm for finding most specific hypothesis from given training samples.