**What is Find-S Algorithm in Machine Learning?**

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

1. Concept Learning
2. General Hypothesis
3. Specific Hypothesis

**1. Concept Learning**

Let’s try to understand concept learning with a real-life example. Most of human learning is based on past instances or experiences. For example, we are able to identify any type of vehicle based on a certain set of features like make, model, etc., that are defined over a large set of features.

These special features differentiate the set of cars, trucks, etc from the larger set of vehicles. These features that define the set of cars, trucks, etc are known as concepts.

Similar to this, machines can also learn from concepts to identify whether an object belongs to a specific category or not. Any [algorithm](https://www.edureka.co/blog/machine-learning-algorithms/) that supports concept learning requires the following:

* Training Data
* Target Concept
* Actual Data Objects

*If you want to learn AI-ML in-depth, come to us and sign up for this*[*Post Graduate Diploma Artificial Intelligence Online Course*](https://www.edureka.co/executive-programs/machine-learning-and-ai)*at Edureka.*

**2. 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 = { ‘?’, ‘?’, ‘?’, …..’?’}

**3. 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 chicken pepperoni filling with a lot of lettuce.*

S = {‘Φ’,’Φ’,’Φ’, ……,’Φ’}

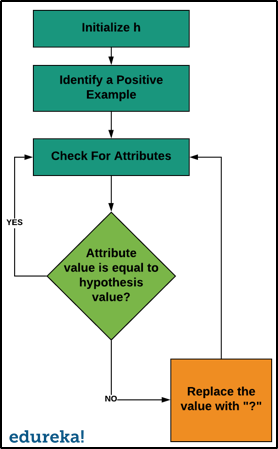
Now ,let’s talk about the Find-S Algorithm in Machine Learning.

The Find-S algorithm follows the steps written below:

1. Initialize ‘h’ to the most specific hypothesis.
2. The Find-S algorithm only considers the positive examples and eliminates negative examples. For each positive example, the algorithm checks for each attribute in the example. If the attribute value is the same as the hypothesis value, the algorithm moves on without any changes. But if the attribute value is different than the hypothesis value, the algorithm changes it to ‘?’.

Now that we are done with the basic explanation of the Find-S algorithm, let us take a look at how it works.

**How Does It Work?**



1. The process starts with initializing ‘h’ with the most specific hypothesis, generally, it is the first positive example in the data set.
2. We check for each positive example. If the example is negative, we will move on to the next example but if it is a positive example we will consider it for the next step.
3. We will check if each attribute in the example is equal to the hypothesis value.
4. If the value matches, then no changes are made.
5. If the value does not match, the value is changed to ‘?’.
6. We do this until we reach the last positive example in the data set.

**Limitations of Find-S Algorithm**

There are a few limitations of the Find-S algorithm listed down below:

1. There is no way to determine if the hypothesis is consistent throughout the data.
2. Inconsistent training sets can actually mislead the Find-S algorithm, since it ignores the negative examples.
3. Find-S algorithm does not provide a backtracking technique to determine the best possible changes that could be done to improve the resulting hypothesis.

Now that we are aware of the limitations of the Find-S algorithm, let us take a look at a practical implementation of the Find-S Algorithm.

**Implementation of Find-S Algorithm**

To understand the implementation, let us try to implement it to a smaller data set with a bunch of examples to decide if a person wants to go for a walk.

The concept of this particular problem will be on what days does a person likes to go on walk.

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **Time** | **Weather** | **Temperature** | **Company** | **Humidity** | **Wind** | **Goes** |
| Morning | Sunny | Warm | Yes | Mild | Strong | Yes |
| Evening | Rainy | Cold | No | Mild | Normal | No |
| Morning | Sunny | Moderate | Yes | Normal | Normal | Yes |
| Evening | Sunny | Cold | Yes | High | Strong | Yes |

Looking at the data set, we have six attributes and a final attribute that defines the positive or negative example. In this case, yes is a positive example, which means the person will go for a walk.

So now, the general hypothesis is:

h0 = {‘Morning’, ‘Sunny’, ‘Warm’, ‘Yes’, ‘Mild’, ‘Strong’}

This is our general hypothesis, and now we will consider each example one by one, but only the positive examples.

h1= {‘Morning’, ‘Sunny’, ‘?’, ‘Yes’, ‘?’, ‘?’}

h2 = {‘?’, ‘Sunny’, ‘?’, ‘Yes’, ‘?’, ‘?’}

We replaced all the different values in the general hypothesis to get a resultant hypothesis.