Experiment-2

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

THEORY:

**Candidate Elimination**

* You can consider this as an extended form of Find-S algorithm.
* Consider both positive and negative examples.
* Actually, positive examples are used here as Find-S algorithm (Basically they are generalize  from the specification).
* While negative example is specified from generalize form.

**Terms Used**

* **Concept learning:**Concept learning is basically learning task of machine (Learn by Train data)
* **General Hypothesis:**Not Specifying features to learn machine.
* G= {'?', '?','?','?'...} -Number of attributes
* **Specific Hypothesis:**Specifying features to learn machine (Specific feature)
* S= {'pi','pi','pi'...} -Number of pi depends on number of attributes.
* **Version Space:** It is intermediate of general hypothesis and Specific hypothesis.It not only just written one hypothesis but a set of all possible hypothesis based on training data-set.

**Algorithm Concept:**

**Step1:**Load Data set

**Step2:** Initialize General Hypothesis  and Specific  Hypothesis.

**Step3:** For each training example

**Step4:**If example is positive example

if attribute\_value == hypothesis\_value:

Do nothing

**else:**

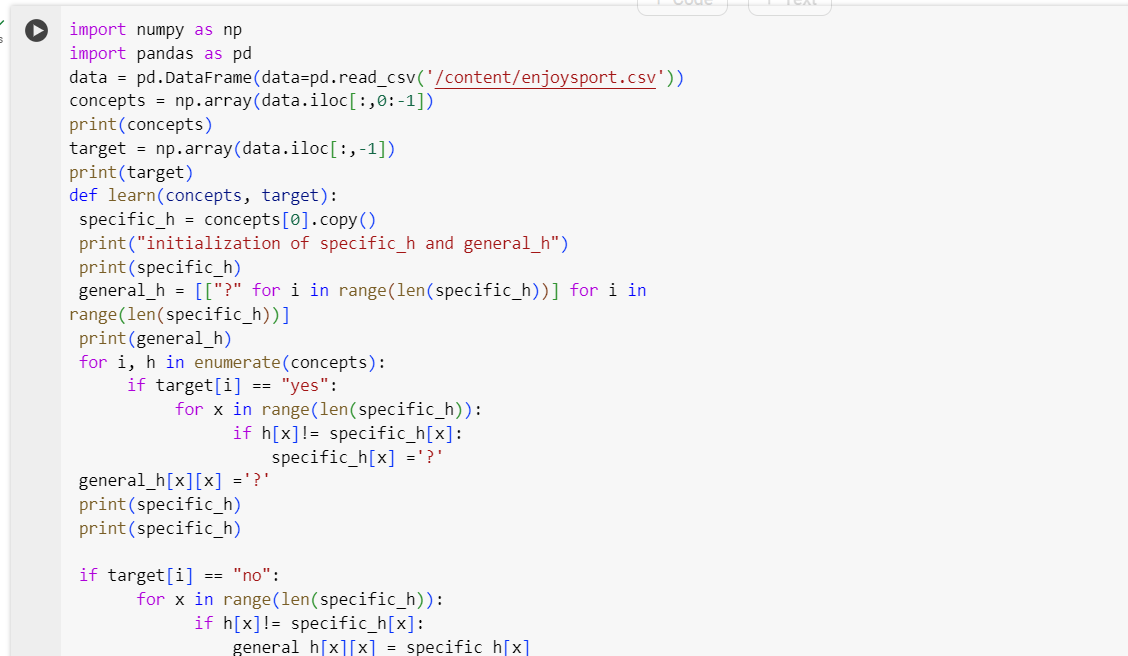
replace attribute value with '?' (Basically generalizing it)

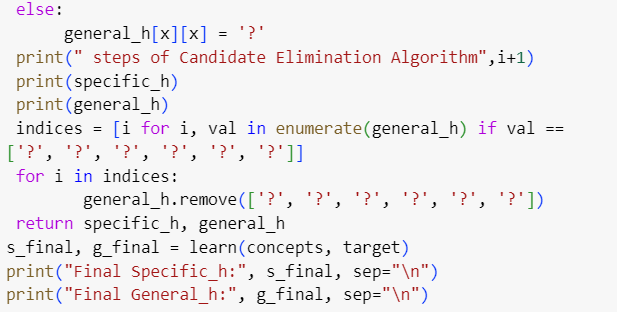
**Step5:**If example is Negative example

Make generalize hypothesis more specific.

**Dataset:**

Enjoy Sport Dataset is saved as. CSV (comma separated values) file in the current working directory otherwise use the complete path of the dataset set in the program:





Output:

