

Experiment 1 - Find S Algorithm

April 24, 2023

1 Problem Statement

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.

2 Import Libraries

```
[ ]: import pandas as pd
import numpy as np
```

3 Implement Algorithm

Declare a function to calculate the final specific hypothesis given a vector of concepts (tuples) and a vector of targets.

```
[ ]: def train(concepts, targets, specificHypothesis):
    for i, val in enumerate(targets):
        if val == 'Yes':
            specificHypothesis = concepts[i]
            break

    for i, val in enumerate(concepts):
        if targets[i] == 'Yes':
            for i in range(len(specificHypothesis)):
                if val[i] != specificHypothesis[i]:
                    specificHypothesis[i] = '?'

    return specificHypothesis
```

4 Import Dataset

```
[ ]: data = pd.read_csv('../data/find-s.csv',
    ↪names=['Sky', 'Temperature', 'Humidity', 'Wind', 'Water', 'Forecast', 'Enjoy Sport?']
    ↪))
```

Here's what the data looks like in a data frame.

```
[ ]: data
```

```
[ ]:      Sky Temperature Humidity   Wind Water Forecast Enjoy Sport?
0  Sunny          Warm   Normal  Strong  Warm     Same       Yes
1  Sunny          Warm   High    Strong  Warm     Same       Yes
2  Rainy          Cold   High    Strong  Warm     Change      No
3  Sunny          Warm   High    Strong  Cool     Change      Yes
```

Retrieving the data points as the vector of concepts (tuples) of the data set.

```
[ ]: dataPoints = np.array(data)[: , :-1]
phiLength = dataPoints.shape[0] + 1
```

Retrieving the target vector.

```
[ ]: dataTarget = np.array(data)[: , -1]
dataTarget
```

```
[ ]: array(['Yes', 'Yes', 'No', 'Yes'], dtype=object)
```

5 Calculate Hypotheses

Defining the specific hypothesis to be all zeros initially.

```
[ ]: specificHypothesis = np.zeros(phiLength)
specificHypothesis
```

```
[ ]: array([0., 0., 0., 0., 0.])
```

Calling the function defined above and obtaining our final hypothesis.

```
[ ]: print(f'The final hypothesis is: {train(dataPoints, dataTarget,
↪specificHypothesis)}')
```

The final hypothesis is: ['Sunny' 'Warm' '?' 'Strong' '?' '?']
