labset-4

March 19, 2025

For a given set of training data examples stored in a .CSV file, implement and demonstrate the Find-S

[4]:
```
import pandas as pd

data = pd.read_csv('./training_data.csv')
```

[5]:
```
print(data)
```

```
  Experience Qualification   Skill  Age Hired
0        Yes       Masters  Python   30   Yes
1        Yes     Bachelors  Python   25   Yes
2         No     Bachelors    Java   28    No
3        Yes       Masters    Java   40   Yes
4         No       Masters  Python   35    No
```

[6]:
```
def find_s_algorithm(data):
    """Implements the Find-S algorithm to find the most specific hypothesis."""
    # Extract feature columns and target column

    attributes = data.iloc[:, :-1].values# All columns except last
    #feature  = data[['Experience','Qualification','Skill','Age']].values
    #target = data[['Hired']].values
    target = data.iloc[:, -1].values  # Last column (class labels)

    # Step 1: Initialize hypothesis with first positive example
    for i in range(len(target)):
        if target[i] == "Yes":  # Consider only positive examples
            hypothesis = attributes[i].copy()
            break

    # Step 2: Update hypothesis based on other positive examples
    for i in range(len(target)):
        if target[i] == "Yes":
            for j in range(len(hypothesis)):
                if hypothesis[j] != attributes[i][j]:
                    hypothesis[j] = '?'  # Generalize inconsistent attributes
```

```
    return hypothesis



# Run Find-S Algorithm
final_hypothesis = find_s_algorithm(data)

# Print the learned hypothesis
print("Most Specific Hypothesis:", final_hypothesis)
```

```
Most Specific Hypothesis: ['Yes' '?' '?' '?']
```

[ ]:

[ ]: