COMP4431 Artificial Intelligence
Quiz 04
Suggested Solution

1.

(a) [2.5 marks]

Step 1: Calculate the Entropy of the Target Attribute (Go Hiking)

The target attribute "Go Hiking" has the following distribution:

- Yes: 4
- No: 4

The entropy $H(whole\_table)$ is calculated as:

$H(whole\_table) = -((4/8)\log(4/8)+(4/8)\log(4/8))=1$

Step 2: Calculate the Entropy for Attribute Temperature

- **C_Hot**: ID {1,2,3} -> Go_Hiking { No, No, Yes}
  - Entropy = $-((2/3)\log(2/3)+(1/3)\log(1/3))\approx 0.918$
- **C_Mild**: ID {4,8} -> Go_Hiking { No, Yes}
  - Entropy = $-((1/2)\log(1/2)+(1/2)\log(1/2))= 1$
- **C_Cool**: ID {5,6,7} -> Go_Hiking { Yes, No, Yes}
  - Entropy = $-((2/3)\log(2/3)+(1/3)\log(1/3))\approx 0.918$

Weighted Entropy for Temperature:

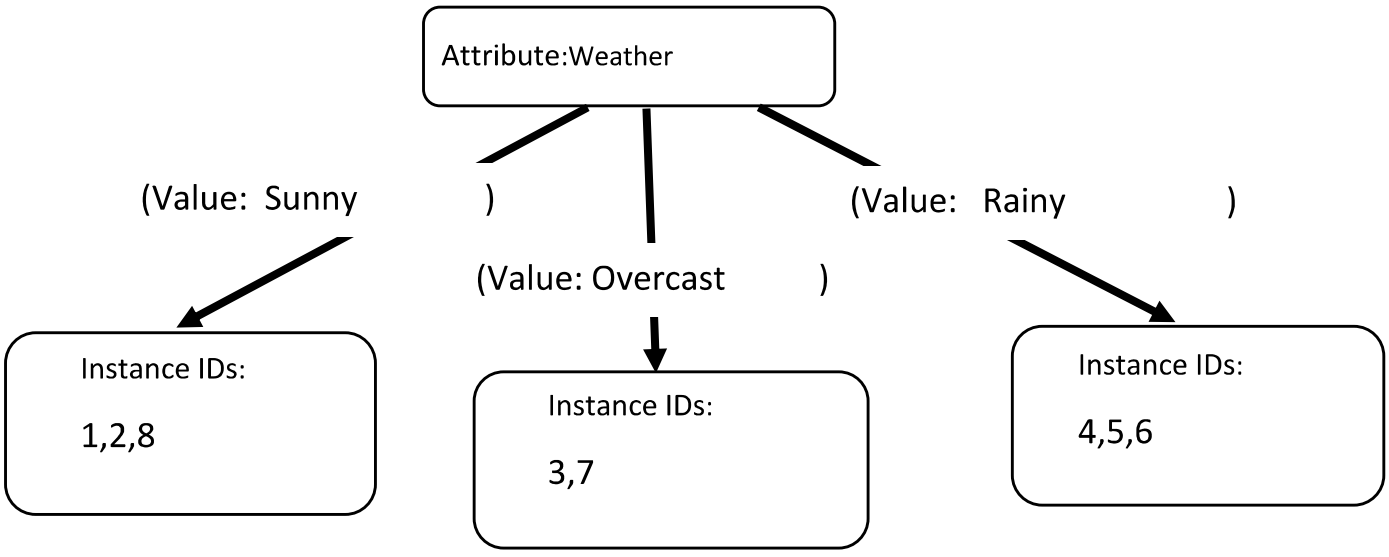
$H(\text{Temperature})=3/8\times 0.918+2/8\times 1+3/8\times 0.918\approx 0.9385$

Information Gain for Temperature:

$IG(\text{Temperature})= 1-0.9385 = 0.0615$

(b) [1 mark]

The attribute with the highest information gain is "Weather" with an information gain of 0.656. Therefore, "Weather" should be chosen as the root node.



2.

a) [2.5 marks]

Step1. Calculate the weighted sum (z):

The weighted sum z is calculated using the formula:

$$s=w_1 \cdot x_1+w_2 \cdot x_2+b$$

Given:

- $w_1=0.5$
- $w_2=-0.3$
- $b=0.1$
- $x_1=0.8$
- $x_2=0.4$

Substitute these values into the formula:

$$s=(0.5\times0.8)+(-0.3\times0.4)+0.1$$

$$=0.4-0.12+0.1=\ 0.38$$

Step2. Apply the sigmoid activation function:

The sigmoid function $\sigma(x)$ is defined as:

$$\sigma(x)=\frac{1}{1+e^{-x}}$$

Substitute $s=0.38$ into the sigmoid function:

$$\sigma(0.38)=\frac{1}{1+e^{-0.38}}$$

$$\sigma(0.38)=11+0.684\approx11.684\approx0.593\sigma(0.38)=1+0.6841\approx1.6841\approx0.593$$

Calculate $e^{-0.38}$

$e^{-0.38}\approx0.684$

Therefore:

$\sigma(0.38)=1/(1+0.684)=1/1.684\approx0.594$

Thus, the output of the neuron after applying the sigmoid activation function is approximately **0.594**.

b) **[1 mark]**
The sign function is not suitable for use in neural networks during backpropagation because:
(either one** of the following, or other acceptable answers)
Non-Differentiability: The sign function is not differentiable at zero and has a derivative of zero elsewhere, which prevents the calculation of gradients needed for weight updates.

Lack of Gradient Information: With a zero gradient, the network cannot learn effectively, as there is no information on how to adjust the weights to minimize the error.