**Exercises:**

**2-a: 1-(1/2)^2-(1/2)^2 =0.5**

**2-b: 0**

**2-c:**

**1 - ( 0.6 )2 - ( 0.4 )2 = 0.48**

**1 - ( 0.4 )2 - ( 0.6 )2 = 0.48**

**0.5 × 0.48 + 0.5 × 0.48 = 0.48**

**2-d:**

**1-(1/4)2-(3/4)2 = 0.375**

**1-(0/8)2-(8/8)2 = 0**

**1-(1/8)2-(7/8)2 = 0.218**

**4/20\*0.375+8/20\*0.218 = 0.16252**

**2-e:**

**1-(3/5)2-(2/5)2 = 0.48**

**1-(3/7)2-(4/7)2 = 0.4898**

**1-(2/4)2-(2/4)2 = 0.5**

**1-(2/4)2-(2/4)2 = 0.5**

**5/20\*0.48+7/20\*0.4898+4/20\*0.5+4/20\*0.5 = 0.4914**

**2-f: Car Type**

**2-g: Everyone is different and has no commonality with others**

**3-a：Entropy = - 4/9 × log2( 4/9 ) - 5/9 × log2( 5/9 ) = 0.9911**

**3-b:**

**Entropy = 4/9 × [ - 1/4 × log2( 1/4 ) - 3/4 × log2( 3/4 ) ] + 5/9 × [ - 1/5 × log2( 1/5 ) - 4/5 × log2( 4/5 ) ] = 0.7616  
a1：0.9911 - 0.7616 = 0.2294**

**Entropy = 5/9 × [ - 2/5 × log2( 2/5 ) - 3/5 × log2( 3/5 ) ] + 4/9 × [ - 2/4 × log2( 2/4 ) - 2/4 × log2( 2/4 ) ] = 0.9839  
 a2：0.9911 - 0.9839 = 0.0072**

**3-c:**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **a3** | **Class label** | **Split ponit** | **Entropy** | **Info Gain** |
| **1.0** | **+** | **2.0** | **0.848** | **0.143** |
| **3.0** | **-** | **3.5** | **0.989** | **0.003** |
| **4.0** | **+** | **4.5** | **0.918** | **0.073** |
| **5.0** | **-** | **5.5** | **0.984** | **0.007** |
| **5.0** | **-** |
| **6.0** | **+** | **6.5** | **0.973** | **0.018** |
| **7.0** | **+** | **7.5** | **0.889** | **0.102** |
| **7.0** | **-** |

**3-d: a1**

**3-e:**

**a1's classification error rate = 2/9**

**a2's classification error rate = 4/9**

**So a1 is the best division**

**3-f:**

**a1=0.167+0.178=0.345**

**a2=0.267+0.222=0.489**

**So a1 is the best division**

**5-a:**

**E=−0.4log2​0.4−0.6log2​0.6=0.9710**

**ΔA​=E−107​EA=T​−103​EA=F​ = 0.2813**

**ΔB​=E−104​EB=T​−106​EB=F​ = 0.2565**

**Therefore, the decision tree induction algorithm selects the A attribute**

**5-b:**

**GINI ​：G=1−(104​)2−(106​)2=0.48**

**GINIA=T​:1−(74​)2−(73​)2=0.4898**

**GINIA=F​：1−(30​)2−(33​)2=0**

**EA​=GINI ​−107​GINIA=T​−103​GINIA=F​=0.1371**

**GINIB=T​：1−(43​)2−(41​)2=0.3750**

**GINIB=F​：1−(61​)2−(65​)2=0.2778**

**EB​=GINI ​−104​GINIB=T​−106​GINIB=F​=0.1633**

**Therefore, the decision tree induction algorithm selects the B attribute**

**5-c:**

**Information gain examines the contribution of features to the entire data, not to specific categories, so generally it can only be used for global feature selection**

**The Gini coefficient is a feature selection method similar to the information entropy, which is used for the impurity of the data. When making feature selection, we can choose the one with the largest ΔGini(X).**

**18-a：0.5 ×0 + 0.5 ×1 = 50% (The number of positive examples and negative examples are equal)**

**18-b: 0.5×0.8 + 0.5×0.2 = 50%**

**18-c: 2/3 × 0 + 1/3 ×1 = 1/3 = 33.3%**

**18-d: 2/3 × 1/3 + 1/3 × 2/3 = 4/9 = 44.4%**