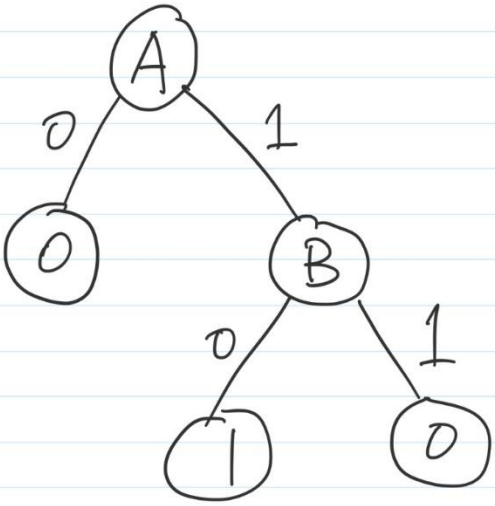
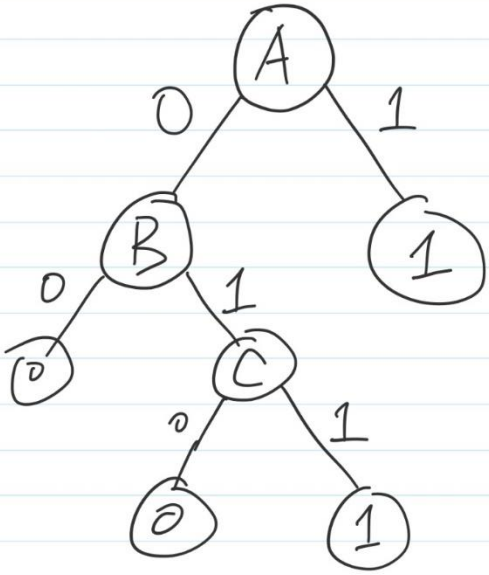
Yiran Liu
CIS 472, Spring 2021
Homework 1

Q1

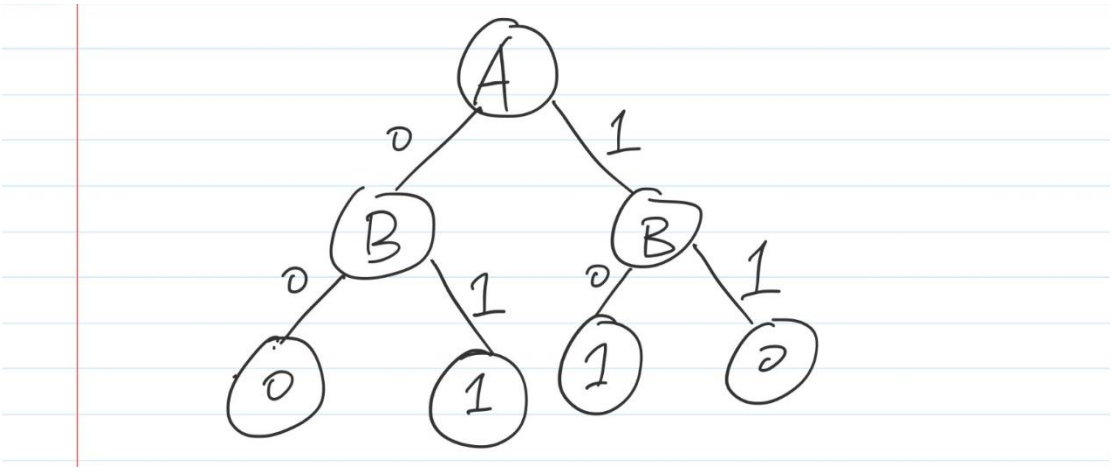
(a)



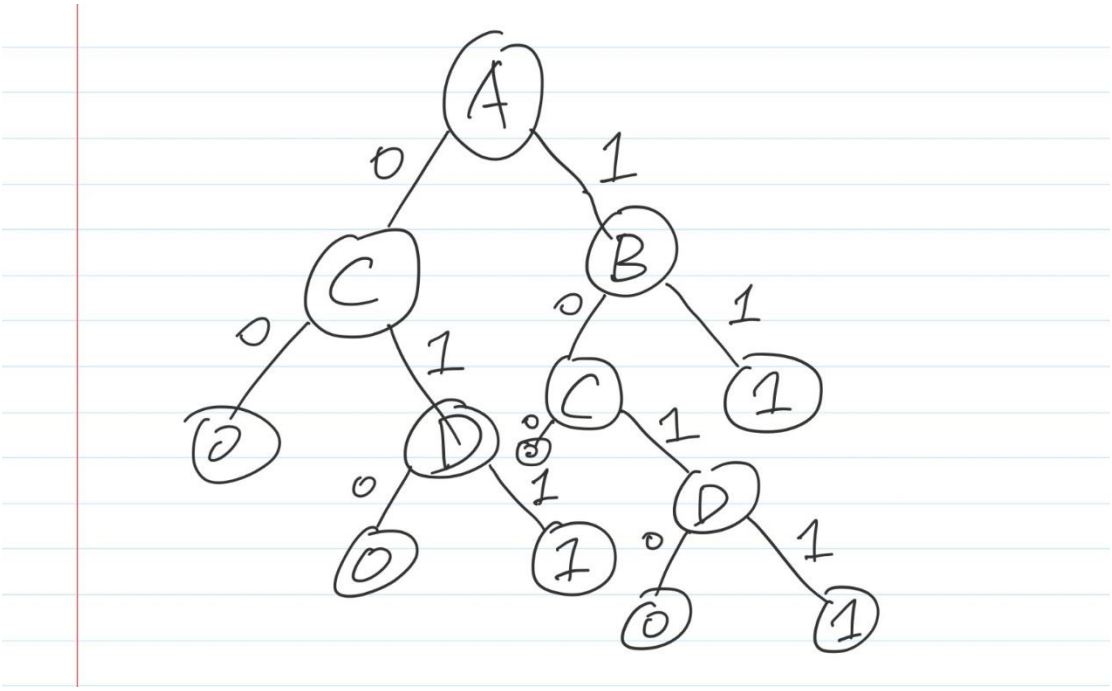
(b)



(c)

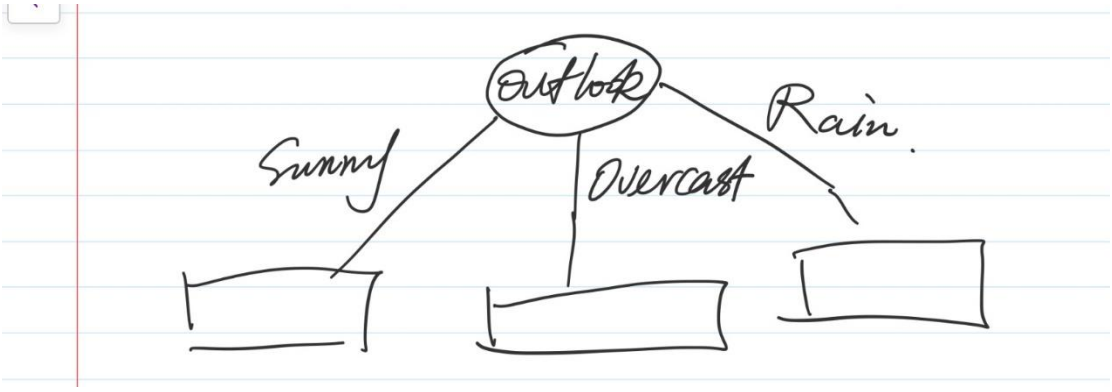


(d)



Q2

(1)



(2) **<u>Sunny: 1, 2, 8, 9, 11</u>**

Overcast: 3, 7, 12, 13

Rain: 4, 5, 6, 10, 14

(3) Sunny: 2 +, 3 -

E = (-2/5) * log2 (2/5) – (3/5) * log2 (3/5) = 0.971

Gain(temperature) = 0.971 – 0 – 1 * (2/5) – 0 = 0.571

Gain(humidity) = 0.971 – 0 – 0 = 0.971

Gain(windy) = 0.971 – 1 * (4/6) – 1 * (2 / 6) = -0.029

Thus, humidity is chosen.

Overcast: all yes, skip

Rainy: 3 +, 2 –

E = (-3/5) * log2(3/5)-(2/5) * log2(2/5) = 0.971

Gain(temperature) = 0.971 – 0 – (3/5) * 0.918 – (2/5) * 1 = 0.0202

Gain(humidity) = 0.971 – (2/5) * 1 – (3/5) * 0.918 = 0.0202

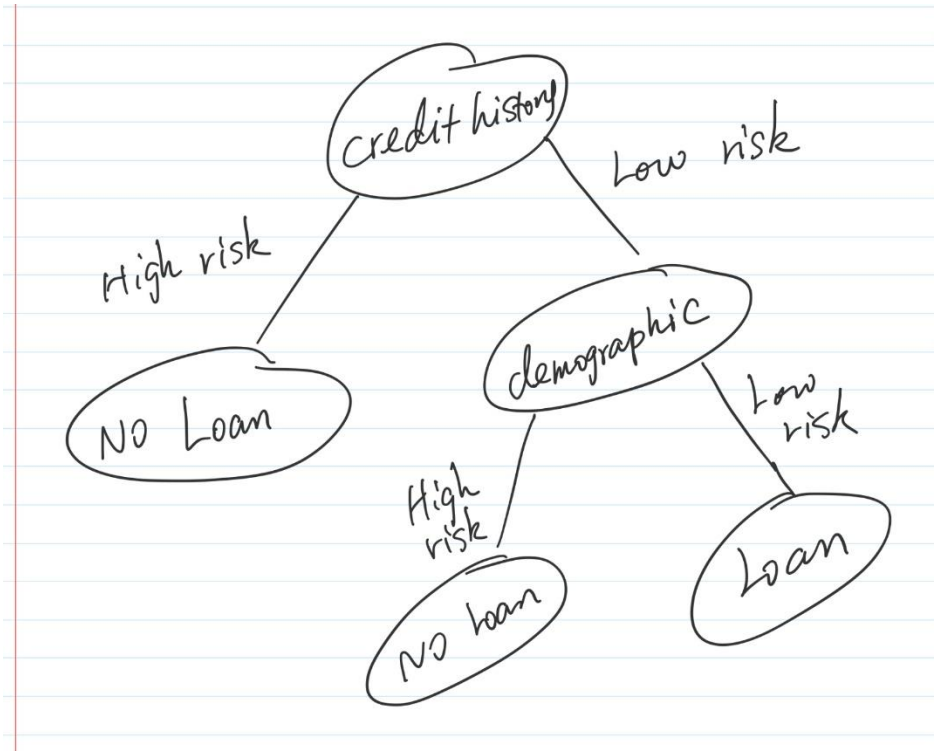
Gain(windy) = 0.971 – 0 – 0 = 0.971

Thus, windy is chosen.

Ans: Under sunny, humidity is chosen, there's no further leaf of overcast, windy under rain is chosen.

Q3

(1) Run the first decision tree first, only run the second decision tree under the "low risk" leaf. Loan under the low risk leaf.



(2) N = (n1) – 1 + n2