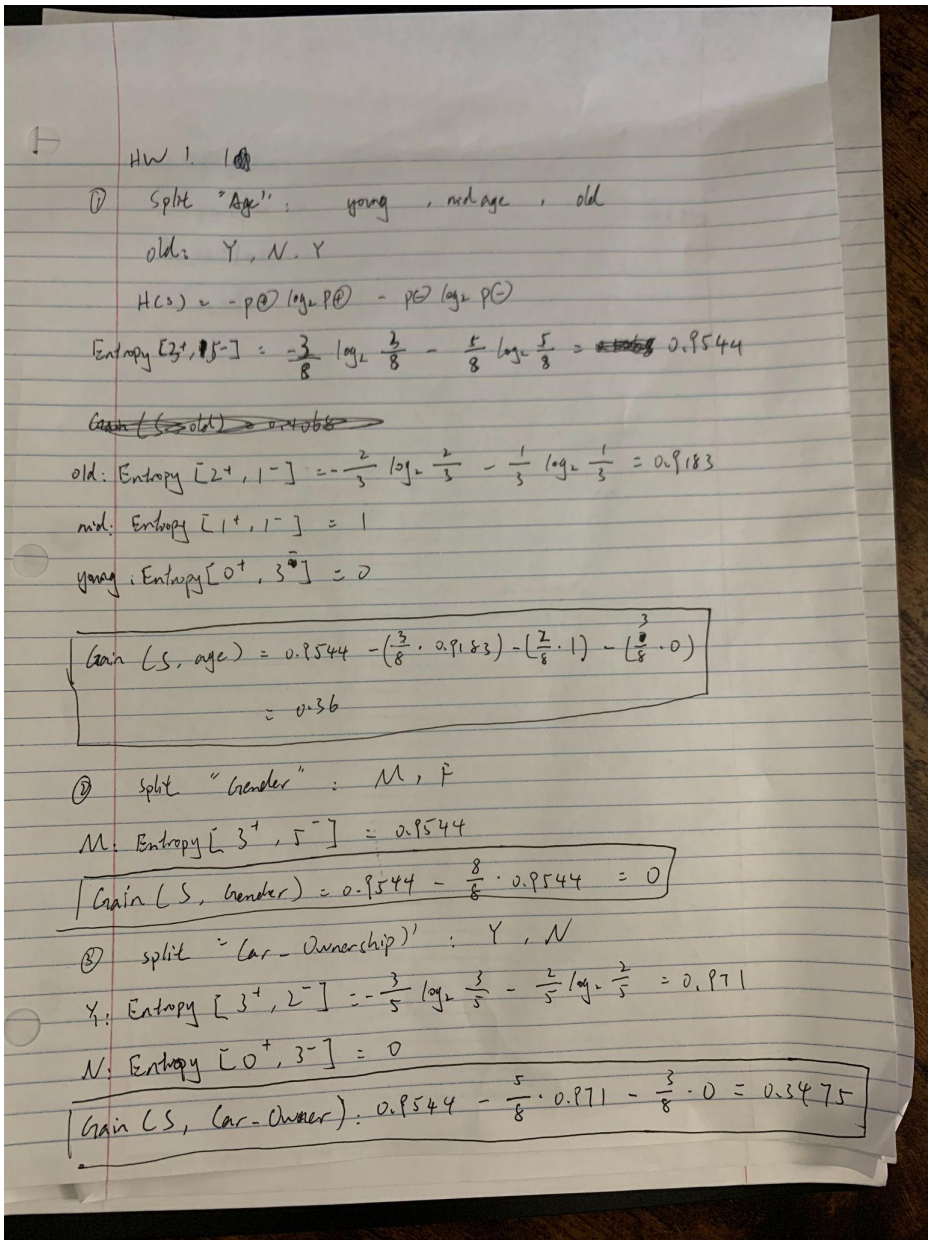
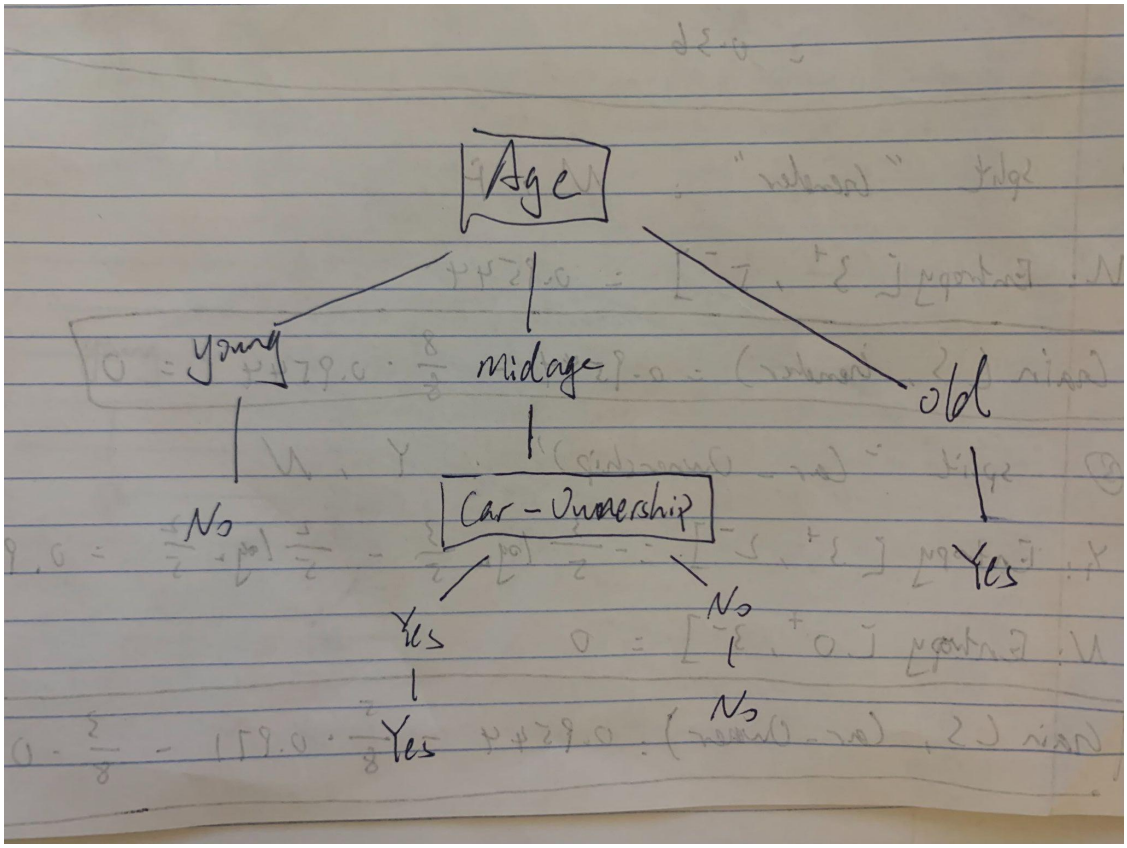
1. At root node, what is the Information Gain if we split by ‘Age‘, ‘Gender‘, and ‘Car Ownership‘ respectively? (10 points)



2. Which column is used to make the first split? (10 points)

Age will be the column that is used to make the first split, because we need to choose the highest information gain to be our first split

3. Once trained, What will be the output on a new sample with Age = 50, Car Ownership = No, Gender = M ? Can you train the same tree without the Gender column, why or why not? (10 points)



The output should be Yes, because when we put [Age = 50, Car Ownership = No, Gender = M], we will first go to the old branch, since we removed all the other attributes in the old branch. Then we can directly go to output Yes.
I think we can train the same tree without a Gender column, because all of the genders are Male, which makes no impact and no influence to the tree. Also, based on the information gain, we know that the information gain for Gender is 0. Since it is too low, we can just drop that.

The calculation of the decision tree: (next page)

Decision tree calculation

Car owner, old

~~Entropy [and old]~~

Gain (old, Car_owner) $= 0.81 - \frac{3}{3} \cdot 0.8181 = 0$

Gain (old, Gender) $= 0.81 - \frac{3}{3} 0.8183 = 0$

Entropy (old) $= -\frac{2}{3} \log_2 \frac{2}{3} - \frac{1}{3} \log_2 \frac{1}{3} = 0.8183$

Entropy (old, car owner) $= 0.8183$

Gain (mid, Car owner) $= 1 - \frac{1}{2} 0 - \frac{1}{2} 0 = 1$

Yes: $E[1^+, 0^-] = 0$

No: $E[0, 1^-] = 0$

Gain (mid, Gender) $= 1 - 0 \frac{2}{2} \cdot 1 = 0$

Male: $E[1^+, 1^-] = 1$

Gain (young, Car) $= 0 - \frac{1}{3} \cdot 0 - \frac{2}{3} \cdot 0 = 0$

Yes: $[0^+, 1^-] = 0$

No: $[0^+, 2^-] = 0$

Gain (young, Gender) $= 0 - \frac{3}{3} \cdot 0 = 0$

M: $[0^+, 3^-] = 0$