1. Determine the class (Yes/No) of stolen for a red domestic SUV.
   a.

	Yes	No	P(yes)	P(no)
red	3	2	3/5	2/5
yellow	2	3	2/5	3/5
Total	5	5	5/5	5/5

	Yes	No	P(yes)	P(no)
Sports	4	2	4/5	2/5
SUV	1	3	1/5	3/5
Total	5	5	5/5	5/5

	Yes	No	P(yes)	P(no)
Domestic	2	4	2/5	4/5
Imported	3	1	3/5	1/5
Total	5	5	5/5	5/5

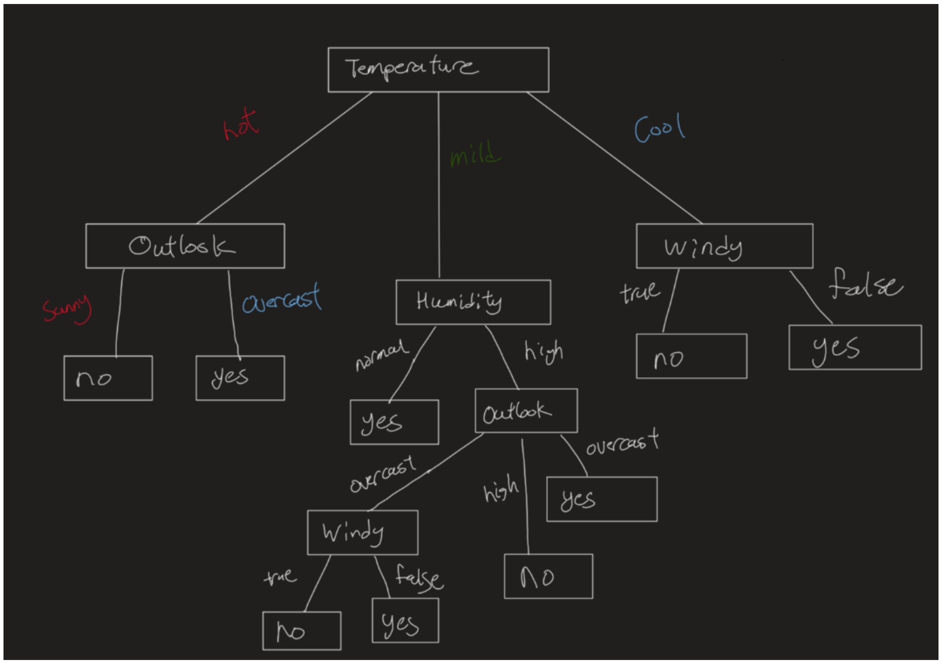
	Stolen	P(yes)/P(no)
Yes	5	5/10
No	5	5/10
Total	10	10/10

   b. $condition = (Red, SUV, Domestic)$,
      i. $P(Yes|condition) = \frac{P(Red|Yes)P(SUV|Yes)P(Domestic|Yes)P(Yes)}{P(condition)} \propto \frac{3}{5} * \frac{1}{5} * \frac{3}{5} * \frac{5}{10} \approx 0.036$
      ii. $P(No|condition) = \frac{P(Red|No)P(SUV|No)P(Domestic|No)P(No)}{P(condition)} \propto \frac{2}{5} * \frac{3}{5} * \frac{4}{5} * \frac{5}{10} \approx 0.096$
      iii. $Since\ P(Yes|condition) + P(No|condition) = 1$
         1. $P(Yes|condition) = \frac{0.036}{0.036+0.096} = 0.27$
         2. $P(No|condition) = \frac{0.096}{0.036+0.096} = 0.72$
      iv. $Since\ P(No|condition)\ is\ higher, this\ will\ get\ classified\ as\ No$
2. This question is about the Naïve Bayes Classifier.
   a. State the simplifying assumption made by the Naïve Bayes classifier
      i. Each feature is independent and have same weight (how important that feature is.)
         1. Also called as conditionally independent.
   b. Given a binary-class classification problem in which the class labels are binary, the dimension of features is d, and each attribute can take k different values.
      i. Provide the number of parameters to be estimated with AND without the simplifying assumption.
         1. With simplification
            a. $2d$
         2. Without simplification
            a. $2(2^d - 1)$
      ii. Briefly justify why the simplifying assumption is necessary.

1. It gets too complicated and complex when there are so many features and each of them are affecting each other.

3. Compute the probability for each of the possible categories
   a. A: The carbon atom is the foundation of life on earth
      i. Words: carbon, atom, life, earth
      ii. $P(Physics|Words) = \frac{P(carbon|Physics)P(atom|Physics)P(life|Physics)P(earth|Physics)P(Physics)}{P(Words)} \propto$
      $0.005 * 0.1 * 0.001 * 0.005 * 0.35 = 8.75 * 10^{-10}$
         1. $P(Physics|Words) = \frac{8.75*10^{-10}}{1.32875*10^{-7}} = 0.0065851$
      iii. $P(Biology|Words) = \frac{P(carbon|Biology)P(atom|Biology)P(life|BioBiology)P(earth|Biology)P(Biology)}{P(Words)} \propto$
      $0.03 * 0.01 * 0.1 * 0.006 * 0.4 = 7.2 * 10^{-8}$
         1. $P(Biology|Words) = \frac{7.2*10^{-8}}{1.32875*10^{-7}} = 0.54186$
      iv. $P(Chem|Words) = \frac{P(carbon|Chem)P(atom|Chem)P(life|Chem)P(earth|Chem)P(Chem)}{P(Words)} \propto$
      $0.05 * 0.2 * 0.008 * 0.003 * 0.25 = 6 * 10^{-8}$
         1. $P(Chem|Words) = \frac{6*10^{-8}}{1.32875*10^{-7}} = 0.45155$
      v. Since P(Bio|Words) has the highest possibility, this sentence will be classified as Biology text.
   b. B: The carbon atom contains 12 protons
      i. Words: carbon, atom, proton
      ii. $P(Physics|Words) = \frac{P(carbon|Physics)P(atom|Physics)P(proton|Physics)P(Physics)}{P(Words)} \propto$
      $0.005 * 0.1 * 0.05 * 0.35 = 8.75 * 10^{-6}$
         1. $P(Physics|Words) = \frac{8.75*10^{-10}}{1.3387*10^{-4}} = 0.06536$
      iii. $P(Biology|Words) = \frac{P(carbon|Biology)P(atom|Biology)P(proton|Biology)P(Biology)}{P(Words)} \propto$
      $0.03 * 0.01 * 0.001 * 0.4 = 1.2 * 10^{-7}$
         1. $P(Biology|Words) = \frac{1.2*10^{-7}}{1.3387*10^{-4}} = 0.00089$
      iv. $P(Chem|Words) = \frac{P(carbon|Chem)P(atom|Chem)P(proton|Chem)P(Chem)}{P(Words)} \propto 0.05 * 0.2 *$
      $0.05 * 0.25 = 1.25 * 10^{-4}$
         1. $P(Chem|Words) = \frac{1.25*10^{-4}}{1.3387*10^{-4}} = 0.93374168$
      v. Since P(Chem|Words) has the highest possibility, this sentence will be classified as Chemistry text.

4. From the classified examples in the above table, construct two decision trees by hand for the classification *Play Gold*.
   a.



   b. $E = -\frac{p}{p+n}\log\left(\frac{p}{p+n}\right) - \frac{n}{p+n}\log\left(\frac{n}{p+n}\right)$
      i. When outlook as the root
         1. When sunny
            a. P = 2, N = 3
            b. $E = -\frac{2}{5}\log 2\left(\frac{2}{5}\right) - \frac{3}{5}\log 2\left(\frac{3}{5}\right) = 0.9709$
         2. When overcast
            a. P = 4, N = 0
            b. $E = -\frac{4}{4}\log 2\left(\frac{4}{4}\right) - \frac{0}{4}\log 2\left(\frac{0}{4}\right) = 0$
         3. When rain
            a. P = 3, N = 2
            b. $E = -\frac{3}{5}\log 2\left(\frac{3}{5}\right) - \frac{2}{5}\log 2\left(\frac{2}{5}\right) = 0.9710$
         4. $IG(O) = 1 - \left[\frac{5}{14} * 0.9709 + \frac{4}{14} * 0 + \frac{5}{14} * 0.9710\right] = 0.3064$
      ii. When Temperature as root
         1. When Hot
            a. P = 2, N = 2
            b. $E = -\frac{2}{4}\log 2\left(\frac{2}{4}\right) - \frac{2}{4}\log 2\left(\frac{2}{4}\right) = 1$
         2. When mild
            a. P = 4, N = 2

b. $E = -\frac{4}{6}\log2\left(\frac{4}{6}\right) - \frac{2}{6}\log2\left(\frac{2}{6}\right) = 0.9183$

3. When cool
   a. P = 3, N = 1
   b. $E = -\frac{3}{4}\log2\left(\frac{3}{4}\right) - \frac{1}{4}\log2\left(\frac{1}{4}\right) = 0.8113$
4. $IG(T) = 1 - \left[\frac{4}{14} * 1 + \frac{6}{14} * 0.9183 + \frac{4}{14} * 0.8113\right] = 0.0889$

iii. When Humidity as root
1. When high
   a. P = 3, N = 4
   b. $E = -\frac{3}{7}\log2\left(\frac{3}{7}\right) - \frac{4}{7}\log2\left(\frac{4}{7}\right) = 0.9852$
2. When normal
   a. P = 6, N = 1
   b. $E = -\frac{6}{7}\log2\left(\frac{6}{7}\right) - \frac{1}{7}\log2\left(\frac{1}{7}\right) = 0.5917$
3. $IG(H) = 1 - \left[\frac{7}{14} * 0.9852 + \frac{7}{14} * 0.5917\right] = 0.21155$

iv. When Windy as root
1. When true
   a. P = 3, N = 3
   b. $E = -\frac{3}{6}\log2\left(\frac{3}{6}\right) - \frac{3}{6}\log2\left(\frac{3}{6}\right) = 1$
2. When false
   a. P = 6, N = 2
   b. $E = -\frac{6}{8}\log2\left(\frac{6}{8}\right) - \frac{2}{8}\log2\left(\frac{2}{8}\right) = 0.8113$
3. $IG(W) = 1 - \left[\frac{6}{14} * 1 + \frac{8}{14} * 0.8113\right] = 0.1078$

v. Since IG(O) has the highest gain, outlook is best for the root.
1. Since E(overcast) = 0, calculating for overcast is unnecessary.
2. When sunny
   a. Temp as the node
      i. When Hot
         1. P = 0, N = 2
         2. $E = -0\log2(0) - 1\log2(1) = 0$
      ii. When mild
         1. P = 1, N = 1
         2. $E = -\frac{1}{2}\log2\left(\frac{1}{2}\right) - \frac{1}{2}\log2\left(\frac{1}{2}\right) = 1$
      iii. When cool
         1. P = 1, N = 0
         2. $E = -1\log2(1) - 0\log2(0) = 0$
      iv. $IG(T) = 1 - \left[\frac{2}{5} * 0 + \frac{2}{5} * 1 + \frac{1}{5} * 0\right] = 0.6$
   b. Humidity as the node
      i. When high

1. P = 0, N = 3
2. $E = -0\log2(0) - 1\log2(1) = 0$

ii. When normal
1. P = 2, N = 0
2. $E = -1\log2(1) - 0\log2(0) = 0$

iii. $IG(H) = 1 - \left[\frac{3}{5} * 0 + \frac{2}{5} * 0\right] = 1$

c. Windy as the node

i. When true
1. P = 1, N = 1
2. $E = -\frac{1}{2}\log2\left(\frac{1}{2}\right) - \frac{1}{2}\log2\left(\frac{1}{2}\right) = 1$

ii. When false
1. P = 1, N = 2
2. $E = -\frac{1}{3}\log2\left(\frac{1}{3}\right) - \frac{2}{3}\log2\left(\frac{2}{3}\right) = 0.9183$

iii. $IG(W) = 1 - \left[\frac{2}{5} * 1 + \frac{3}{5} * 0.9183\right] = 0.04902$

d. Since IG(H) has the highest, humidity will be the node

3. When rain

a. Temp as the node

i. When Hot
1. P = 0, N = 0
2. $E = 0$

ii. When mild
1. P = 2, N = 1
2. $E = -\frac{2}{3}\log2\left(\frac{2}{3}\right) - \frac{1}{3}\log2\left(\frac{1}{3}\right) = 0.9183$

iii. When cool
1. P = 1, N = 1
2. $E = -\frac{1}{2}\log2\left(\frac{1}{2}\right) - \frac{1}{2}\log2\left(\frac{1}{2}\right) = 1$

iv. $IG(T) = 1 - \left[\frac{0}{5} * 0 + \frac{3}{5} * 0.9183 + \frac{2}{5} * 1\right] = 0.04902$

b. Humidity as the node

i. When high
1. P = 1, N = 1
2. $E = -\frac{1}{2}\log2\left(\frac{1}{2}\right) - \frac{1}{2}\log2\left(\frac{1}{2}\right) = 1$

ii. When normal
1. P = 2, N = 1
2. $E = -\frac{2}{3}\log2\left(\frac{2}{3}\right) - \frac{1}{3}\log2\left(\frac{1}{3}\right) = 0.9183$

iii. $IG(H) = 1 - \left[\frac{2}{5} * 1 + \frac{3}{5} * 0.9183\right] = 0.049$

c. Windy as the node

i. When true
   1. P = 0, N = 2
   2. $E = -0\log2(0) - 1\log2(1) = 0$
ii. When false
   1. P = 3, N = 0
   2. $E = -1\log2(1) - 0\log2(0) = 0$
iii. $IG(W) = 1 - \left[\frac{2}{5} * 0 + \frac{3}{5} * 0\right] = 1$

d. Since IG(W) has the highest, Windy will be the node

vi.