Bayesian Classification Worksheet

Table 7.1 Training data tuples from the *AllElectronics* customer database.

RID	age	income	student	credit_rating	Class: buys_computer
1	<=30	high	no	fair	no
2	<=30	high	no	excellent	no
3	31 . . . 40	high	no	fair	yes
4	>40	medium	no	fair	yes
5	>40	low	yes	fair	yes
6	>40	low	yes	excellent	no
7	31 . . . 40	low	yes	excellent	yes
8	<=30	medium	no	fair	no
9	<=30	low	yes	fair	yes
10	>40	medium	yes	fair	yes
11	<=30	medium	yes	excellent	yes
12	31 . . . 40	medium	no	excellent	yes
13	31 . . . 40	high	yes	fair	yes
14	>40	medium	no	excellent	no

1. (By hand) Leave the first four samples out, and then use the remaining 10 samples to calculate the probability that each of the first four samples buys a computer. Use the naïve Bayes classifier that we have been discussing in class.

2. Create a notebook called bayes.ipynb. Using what you learned in 3.1 and 3.2, replicate your work using pandas. Add comments and write-up as necessary. If we can't follow your calculations, we can't grade your work.