For Questions 1- 3, please submit a word file or a PDF file;
For Question 4 (programming question), please submit an .ipynb file.

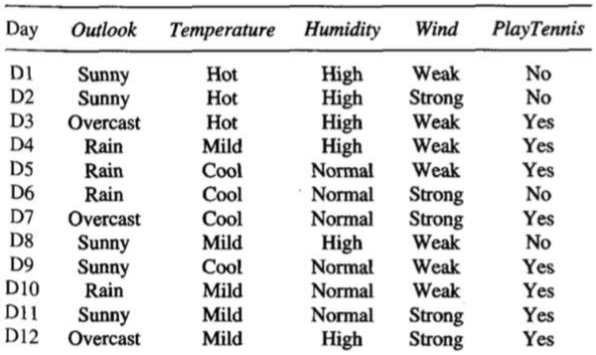
Question 1 (8 points):

Consider again the example application of Bayes rule in Section 6.2.1 of Tom Mitchell's textbook. Suppose the doctor decides to order a second laboratory test for the same patient and suppose the second test returns a positive result as well. What are the posterior probabilities of *cancer* and ¬*cancer* respectively following these two tests? Assume that the two tests are independent.

Question 2 (8 points):

Section 6.9.1 of Tom Mitchell's textbook demonstrates an example using the Naïve Bayes Algorithm to predict a new instance based on a dataset with 14 examples from Table 3.2 of Chapter 3 of the book. If we only have 12 examples as shown below, what is the prediction results for the same new instance? Show your calculation.

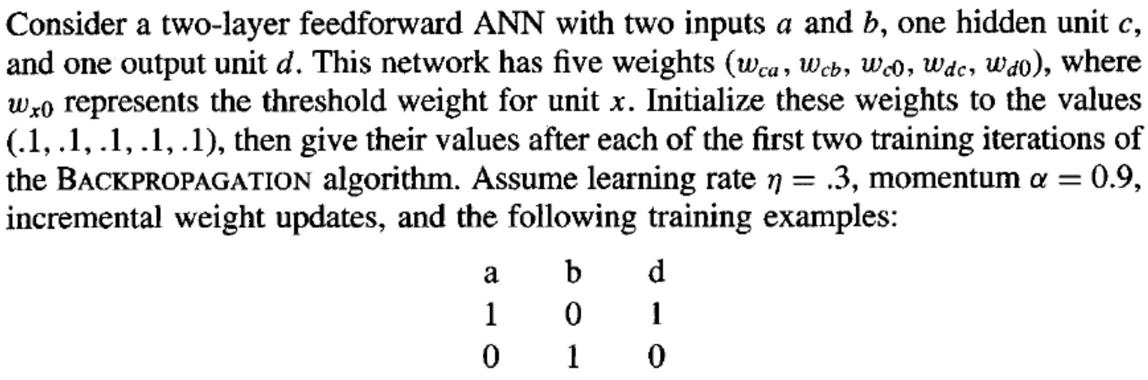
New instance: <Outlook=sun, Temperature=cool, Humidity=high, Wind=strong>



Day	*Outlook*	*Temperature*	*Humidity*	*Wind*	*PlayTennis*
D1	Sunny	Hot	High	Weak	No
D2	Sunny	Hot	High	Strong	No
D3	Overcast	Hot	High	Weak	Yes
D4	Rain	Mild	High	Weak	Yes
D5	Rain	Cool	Normal	Weak	Yes
D6	Rain	Cool	Normal	Strong	No
D7	Overcast	Cool	Normal	Strong	Yes
D8	Sunny	Mild	High	Weak	No
D9	Sunny	Cool	Normal	Weak	Yes
D10	Rain	Mild	Normal	Weak	Yes
D11	Sunny	Mild	Normal	Strong	Yes
D12	Overcast	Mild	High	Strong	Yes

Question 3 (14 points):

Answer question 4.7 (page 125) of Tom Mitchell's textbook as quoted below:



Consider a two-layer feedforward ANN with two inputs a and b, one hidden unit c, and one output unit d. This network has five weights (w_{ca}, w_{cb}, w_{c0}, w_{dc}, w_{d0}), where w_{x0} represents the threshold weight for unit x. Initialize these weights to the values (.1, .1, .1, .1, .1), then give their values after each of the first two training iterations of the BACKPROPAGATION algorithm. Assume learning rate $\eta = .3$, momentum $\alpha = 0.9$, incremental weight updates, and the following training examples:

a	b	d
1	0	1
0	1	0

Question 4 – Programming (40 points):

In this programming problem, you will get familiar with building a neural network using backpropagation. You are supposed to implement the following steps:

Step 1: use our "titanic" dataset in homework #3, and split data in the same way you did in homework #3 – 80% as training and 20% test sets;

Step 2: Fit a neural network using independent variables 'pclass + sex + age + sibsp' and dependent variable 'survived'. Fill in n/a attributes with the average of the same attributes from other training examples. Use 2 hidden layers and set the activation functions for both the hidden and output layer to be the **sigmoid** function. Set "solver" parameter as either **SGD** (stochastic gradient descend) or **Adam** (similar to SGD but optimized performance with mini batches). You can adjust parameter "alpha" for regularization (to control overfitting) and other parameters such as "learning rate" and "momentum" as needed.

Step 3: Check the performance of the model with out-of- sample accuracy, defined as

out-of-sample percent survivors correctly predicted (on test set)
out-of-sample percent fatalities correctly predicted (on test set)

Please try two different network structures (i.e., number of neurons at each hidden layer) and show their respective accuracy.

Step 4: Compare the out-of-sample accuracy (as defined in step 3) with the random forest obtained in homework #3. (You can either use a table or plot the results of the two algorithms in one figure). Explain any difference in accuracy.

Note: There are two options to implement the neural network:

Option 1: use scikit-learn library;

Here is the tutorial: http://scikit-learn.org/stable/modules/neural_networks_supervised.html

Option 2 (bonus: 5 points): implement backpropagation yourself; in your implementation, you better set the following:

(1) the initial weights to be uniformly between [-0.1, +0.1]

(2) the number of iterations to be around 5000 or more (but not tens of thousand)

You can choose either option for this homework. You will get 5 bonus points if you choose option 2. No matter what you choose, make sure you know how to update the weights.