1. Consider the hypothesis class of homogeneous half-spaces in the $\mathbb{R}^d$ (i.e. each instance $x \in \mathbb{R}^d$ and $H$ consists of all $h_w$ having the form $h(x) = +1$ if and only if $w \cdot x > 0$, so $0$ is always mapped $-1$). Determine the VC-dimension of homogeneous half-spaces in $\mathbb{R}^d$. In other words, for each $d \geq 1$ find a set $S \subset \mathbb{R}^d$ is shattered by $H$ and show that no set $S' \subset \mathbb{R}^d$ with $|S'| > |S|$ is shattered by $H$. (Since if $S'$ is shattered, then every subset of $S'$ is also shattered, it suffices to show that no set $S'$ with $|S'| = |S| + 1$ is shattered). You may use the fact that in any set $S$ of $d + 1$ points in $\mathbb{R}^d$, there is at least one point $x \in S$ that can be expressed as a linear combination of the other points in $S$.

Recall that the VC-dimension is the size of the largest shattered set. To show the VC-dimension is some value $k$, you need to show that there is some set of size $k$ that is shattered, and that no set of size $k + 1$ is shattered.

For significant partial credit, find the VC-dimension of homogeneous half-spaces on the real line ($d = 1$) and the real plane ($d = 2$).

2. Consider the following experiment. A magician has three coins in his pocket, a two-headed coin, a two-tailed coin, and a fair coin. The magician picks a coin from his pocket with each coin equally likely. The magician then flips the coin twice, and sees what the comes up (either $hh$, $ht$, $th$, or $tt$). To make this more formal, consider three random variables $coin \in \{0, \frac{1}{2}, 1\}$, $flipA \in \{0, 1\}$, and $flipB \in \{0, 1\}$ where the value of $coin$ gives the probability that a "head" results when the coin is flipped (and $1 - coin$ is the probability that a "tail" results) and $flipA$ and $flipB$ are indicator functions for the events "the first flip is a head" and "the second flip is a head" respectively. Let each triple of values $(coin, flipA, flipB)$ correspond to a point (atomic event) in $\Omega$, and assume that $P(coin = 0) = P(coin = 1/2) = P(coin = 1) = 1/3$.

First, what is $|\Omega|$? Second, how many points in $\Omega$ have zero probability? Third (main part), what is $P(flipB = 1|flipA = 1)$?

3. Use Weka (http://www.cs.waikato.ac.nz/ml/weka/) to run logistic regression (logistic) and naive bayes on the iris2.ARFF data set (see web page). Train on the entire data set, and extract the logistic regression decision boundary and the Naive bayes distributions. Implement LDA (perhaps in MATLAB or something that inverts matrices for you). Compare the accuracies of the three classifiers and the LDA and logistic regression decision boundaries. (For Naive Bayes, the decision boundary is a hyperquadratic, and can be difficult to find analytically.)