Use MATLAB’s built in cancer dataset and linear regression to create a simple discriminant function, similar to the following snippet:

[X,d] = cancer\_dataset; %Type help cancer\_dataset for more info

w=X'\d(2,:)'; %Training/MSE linear model creation

y=X'\*w; %Activation/testing

[X,Y,T,AUC] = perfcurve(d(2,:),y',1);

figure,plot(X,Y) %Visualize

xlabel('False positive rate')

ylabel('True positive rate')

title(['2D ROC, AUC=' num2str(AUC)])

A- Find a subset of input variables for the linear regressor to see if a reduced input space performs better. Test at least 5 subsets (including the full 9-dimensional input) and use ROC AUC as your measure of success.

B- Keep the first half of the data for creating the linear regressor (training) and the second half for testing. Repeat the above for the best subset found in A and report the AUC for train/test. Summarize your observations.