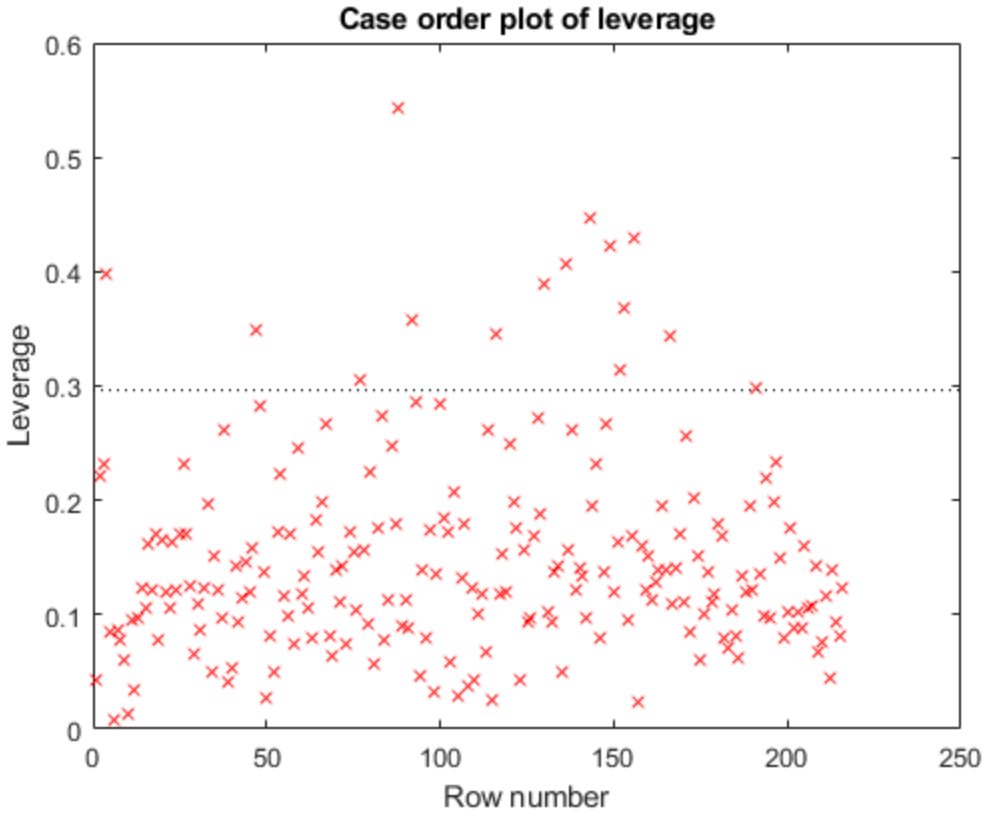
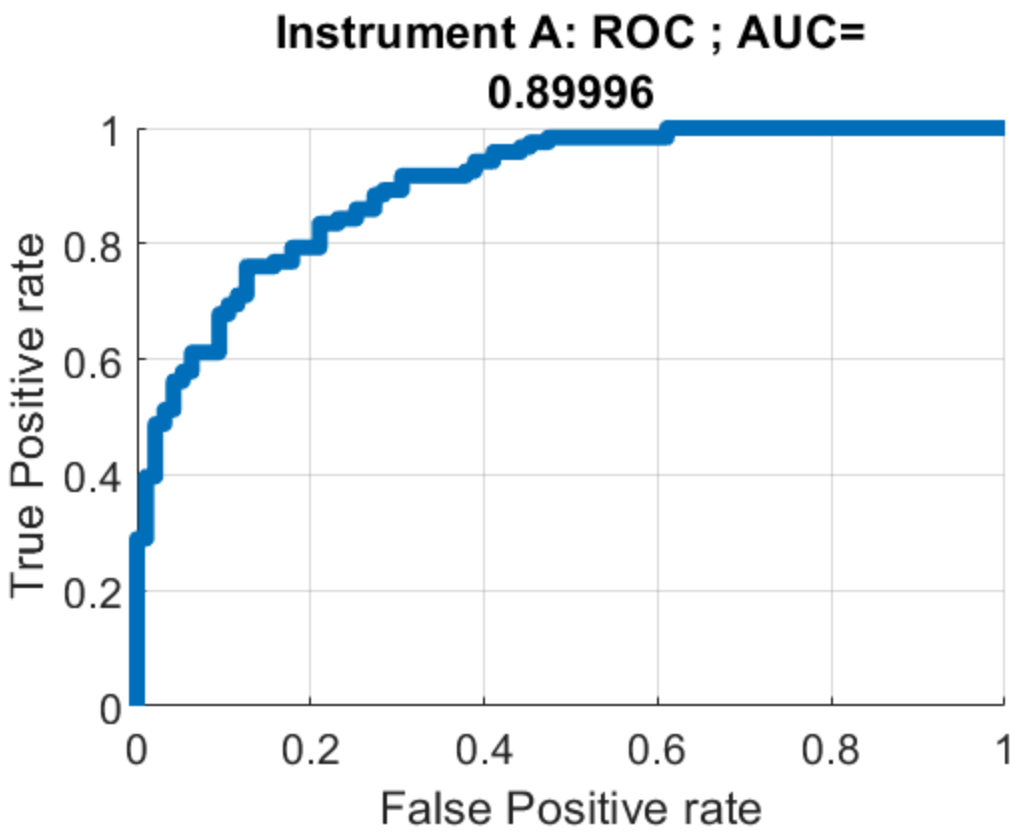
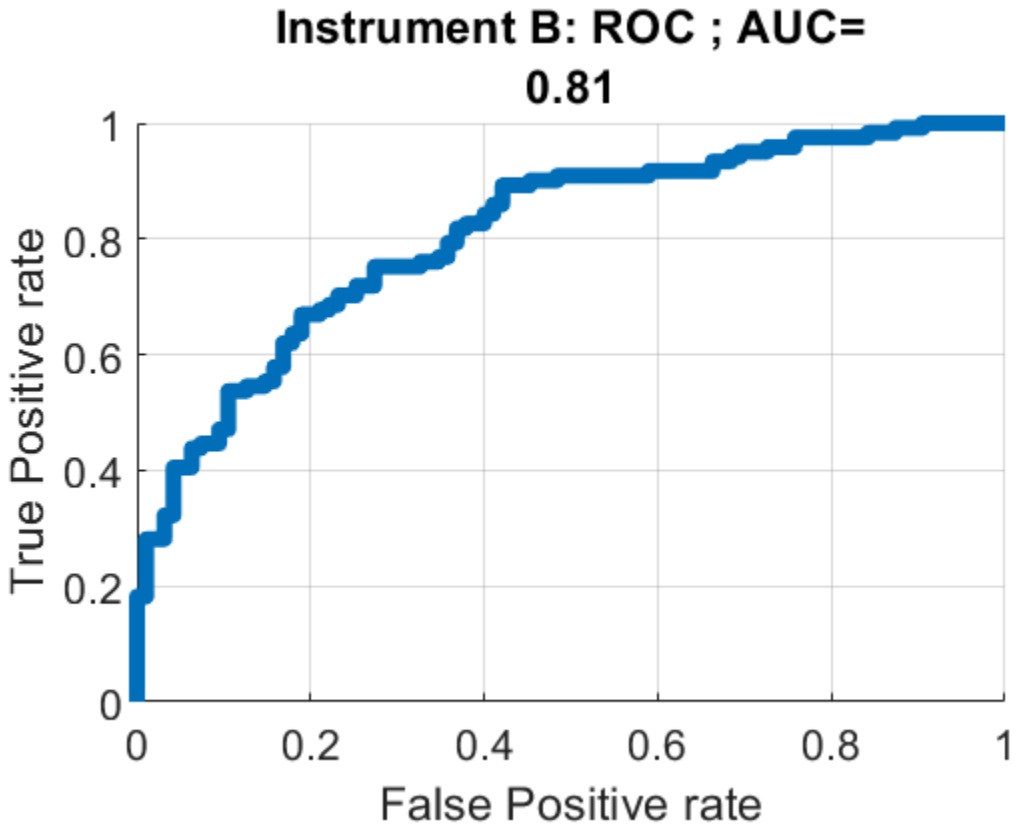
```
%Student Number: 251004930
%Name: David George

  %A)

    %Creating the appropriate arrays for each machine
    T  = readtable("ovarian_diagnostic.csv", 'ReadVariableNames',
 false);
    A_T = readtable("ovarian_A.csv");
    B_T = readtable("ovarian_B.csv");
    A_T= table2array(A_T(:, 1:65));
    B_T= table2array(B_T(:, 1:31));

    %Iterate throught T, converting it to 1 for cancer 0 for no cancer
    for idx = 1:length(T.Var1)

        if T.Var1{idx} == "Cancer"
            T.Var1{idx} = '1';
        end

        if T.Var1{idx} == "Normal"
            T.Var1{idx} = '0';
        end
    end
      T.Var1 = str2double(T.Var1);
      T = table2array(T(:, :));


    %Perfromming multi-var logistic regression
    mod = fitglm(A_T,T,'Distribution', 'binomial');
    modB = fitglm(B_T,T,'Distribution', 'binomial');
    p = mod.Fitted.Probability;
    coef = mod.plotDiagnostics;
    pB = modB.Fitted.Probability;
    coefB = modB.plotDiagnostics;

    %Instrument A

        %ROC analysis for instr A
        [fpr,tpr,thresholds,AUC, opt] = perfcurve(T, p, 1);
        figure
        hold on
        plot(fpr,tpr, 'LineWidth',6);
        grid()
        xlabel("False Positive rate")
        ylabel("True Positive rate")
        title(["Instrument A: ROC ; AUC="; AUC])
        set(gca, 'FontSize', 16)


      % %ROC analysis for instra B
        [fprB,tprB,thresholdsB,AUCB, optB] = perfcurve(T, pB, 1);
```

```
figure
hold on
plot(fprB,tprB, 'LineWidth',6);
grid()
xlabel("False Positive rate")
ylabel("True Positive rate")
title(["Instrument B: ROC ; AUC="; AUCB])
set(gca, 'FontSize', 16)

%The best instrument to detect ovarian cancer would be
instrument A.
%This is due to the fact that the area under the curce (AUC)
for
%intrument A's ROC is GREATER than the AUC for instrument B's
ROC,
%which therefore means A's ROC is the better classifer, and
this
%instrument A is the logical choice.
```




Instrument A: ROC ; AUC=
0.89996
True Positive rate
False Positive rate
0
0.2
0.4
0.6
0.8
1


Instrument B: ROC ; AUC=
0.81
True Positive rate
False Positive rate
0
0.2
0.4
0.6
0.8
1

```
   %B)
        %We want a true postive rate of 90%, therefore 90% is the cut
off
        NinetyTPR = tpr(tpr < 0.90 +.01 & tpr >= 0.90);
        indexPos = find(tpr < 0.90 +.01 & tpr >= 0.90);
        NinetyFPR = fpr(indexPos);
        SmallestFpr= min(NinetyFPR);
        CorrespondingTPR = tpr(find(fpr==SmallestFpr));
        % The closest TRUE psotive rate is 0.09008 and the smallest,
correspondig FPR is
        % 0.3053.

        %For the worst method, the FPR with a TPR of %90 is
        BNinetyTPR = tprB(tprB < 0.90 +.01 & tprB >= 0.90);
        BindexPos = find(tprB < 0.90 +.01 & tprB >= 0.90);
        BNinetyFPR = fprB(BindexPos);
        BSmallestFpr= min(BNinetyFPR);
        BCorrespondingTPR = tprB(find(fprB==BSmallestFpr));
        % The closest TRUE psotive rate is 0.9008 and the smallest,
correspondig FPR is
        % 0.4526.
```

Published with MATLAB® R2019b