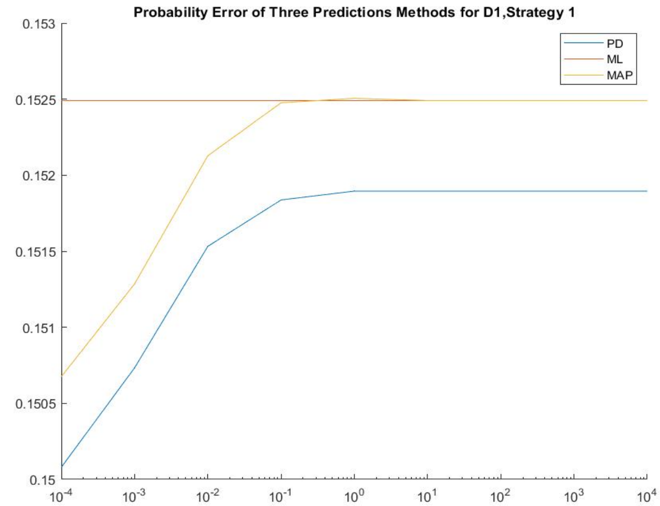
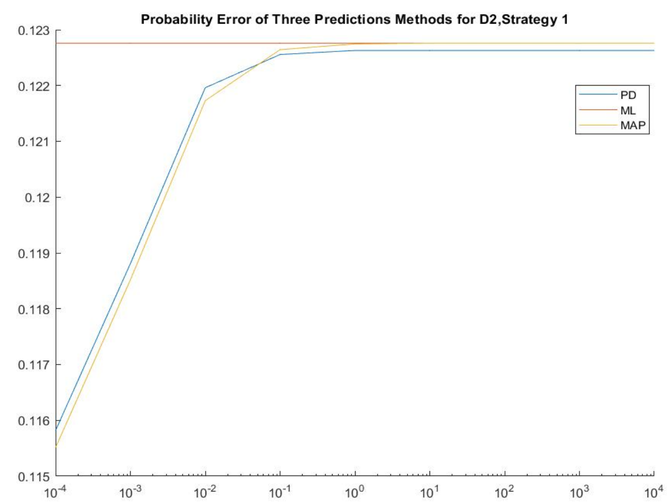
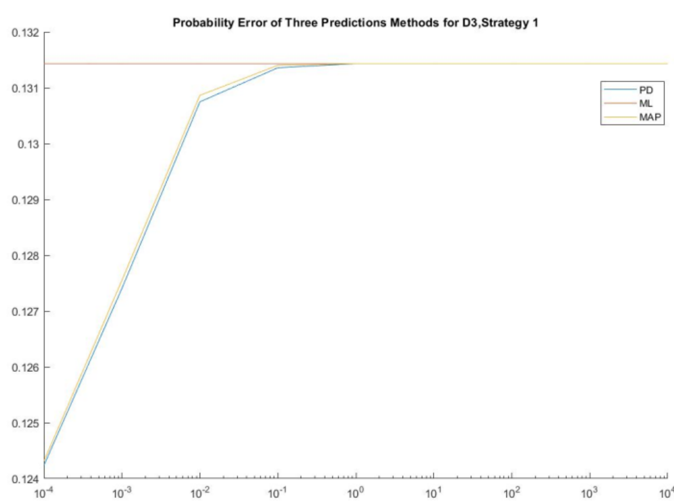
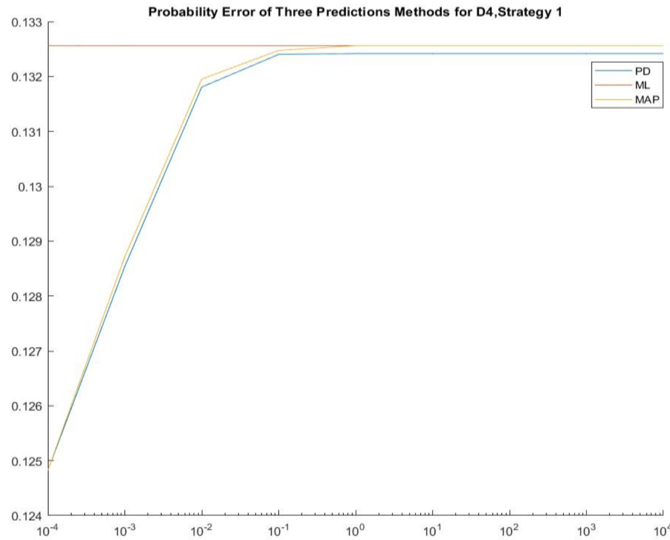
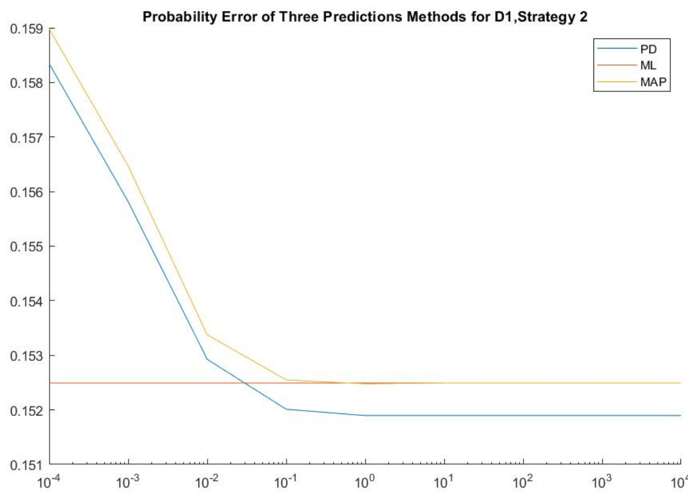
Homework 3 and 4 Report

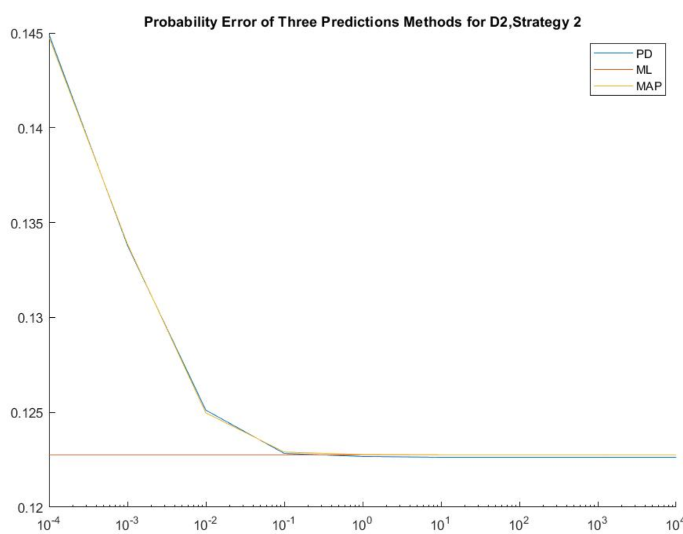


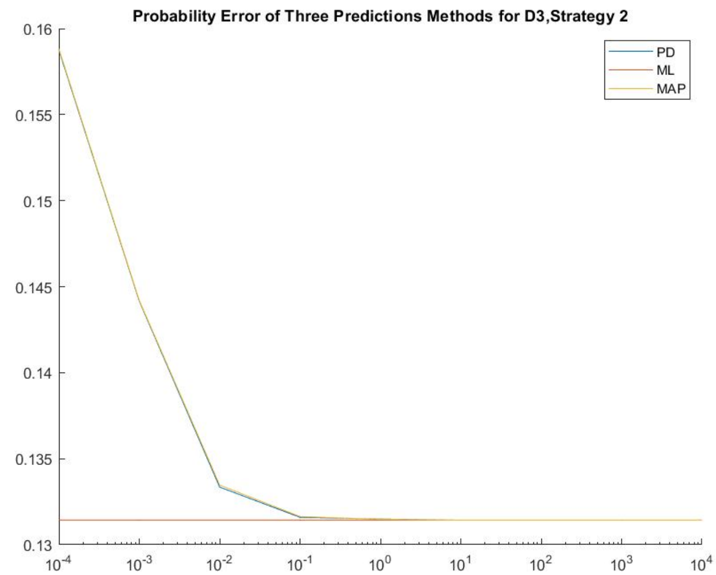


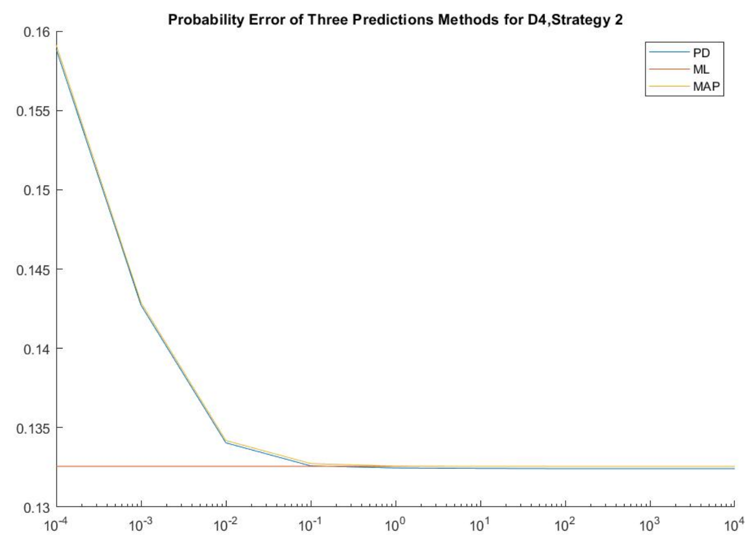



Probability Error of Three Predictions Methods for D4,Strategy 1
PD
ML
MAP
0.133
0.132
0.131
0.13
0.129
0.128
0.127
0.126
0.125
0.124
10^{-4} 10^{-3} 10^{-2} 10^{-1} 10^{0} 10^{1} 10^{2} 10^{3} 10^{4}


Probability Error of Three Predictions Methods for D1,Strategy 2
PD
ML
MAP
0.159
0.158
0.157
0.156
0.155
0.154
0.153
0.152
0.151
10^{-4} 10^{-3} 10^{-2} 10^{-1} 10^{0} 10^{1} 10^{2} 10^{3} 10^{4}


Probability Error of Three Predictions Methods for D2,Strategy 2
PD
ML
MAP
0.145
0.14
0.135
0.13
0.125
0.12
10^{-4} 10^{-3} 10^{-2} 10^{-1} 10^{0} 10^{1} 10^{2} 10^{3} 10^{4}





a) In this part, the goal is to compute the predictive distribution $P_{X|T}(X|D_1)$. Before that, we first need to solve $P_{\mu|T}(\mu|D_1) = G(\mu, \mu_1, \Sigma_1)$, where $\mu_1 = \frac{n\sigma_0^2}{\sigma^2+n\sigma_0^2}\mu_{ML} + \frac{\sigma^2}{\sigma^2+n\sigma_0^2}\mu_0, and\ \Sigma_1 = \sigma_n^2 = \frac{\sigma^2\sigma_0^2}{\sigma^2+n\sigma_0^2}$.
Therefore, $P_{X|T}(X|D_1) = \int P_{X|\mu}(X|\mu)P_{\mu|T}(\mu|T)d\mu = \int f(x-\mu)h(\mu)d\mu = G(x,0,\sigma^2) * G(x,\mu_n,\sigma_n^2)$, with $f(x) = G(x,0,\sigma^2), and\ h(x) = G(x,\mu_n,\sigma_n^2)$. $P_{X|T}(X|D_1) = G(x,\mu_n,\sigma^2+\sigma_n^2)$.

b) For ML estimation, we just do what we have done in homework 2.

c) To find the solution of MAP estimation, we follow the formula $P_{X|T}(\mathrm{X}|D_1) = P_{X|\mu}(\mathrm{X}|\mu_{MAP})$, where $\mu_{MAP} = \arg max_\mu P_{\mu|T}(\mu|D_1)$.

By only comparing the α and the error rate of the estimation of the predictive distribution with strategy 1 ($\mu_0 = 1\ for\ the\ cheetah\ class, and\ \mu_0 = 3\ for\ the\ grass\ class$), it can be seen that the classification has the lowest error rates when α is as lowest as 10e-4, since the weight of the prior mean μ_0 is largest. As the α increases, the error rate gets larger as well. After α = 1, the prior mean becomes ignorable which results in a flat error rate as the α become larger. This relationship between α and the error rate of the estimation also exists for the MAP solution, since we also using the diagonal matrix $(\Sigma_0)_{ii} = \alpha w_i$ to compute μ_{MAP}, which bringing the mean μ_0. However, since the ML solution does not use the prior mean, it does not have such relationship with α. Because of that, the error rate of the ML solution is always flat.

By comparing the plot of different sets of train data, we can see only a large drop of error rate between data set 1 and data set 2. The number of foreground samples is 125 and the number of background samples is 500 in data set 2 while the number of foreground samples is 75 and the number of background samples is 300 in data set 1. It means that the observations are insufficient in data set 1while being sufficient in data set 2. Although in data set 3 and 4, the number of samples becomes larger than data set 2, these samples only provide different observations under the sufficient condition which only giving us a very small change of error rate.

When we implement the strategy 2, the change of error rates of the predictive distribution and MAP with respect to α gets reversed that the error rate goes down as α get larger. Under this strategy, the prior means for both classes are the same which perform poorly in prediction. Therefore, as the α gets larger, the weight of the prior means becomes lower which bringing a lower error rate even it is still not good. Compared to both two strategies, the performance of strategy 1 is better than strategy 2, even the difference of error rate is only about 0.002.

```
load('TrainingSamplesDCT_subsets_8.mat');
load("Alpha.mat");
%load("Prior_1.mat");
load("Prior_2.mat");
img = im2double(imread('cheetah.bmp'));
mask = im2double(imread('cheetah_mask.bmp'));

%read Zig-Zag Pattern.txt file
zz = fopen('Zig-Zag Pattern.txt','r');
zzPat = fscanf(zz,'%d',[8,8])+1;
fclose(zz);

% obtain the DCT of the image
[row,colm] = size(img);
img_zzs = zeros(row-8,colm-8,64);
for i = 1:row-8
    for j = 1:colm-8
        dctImg = dct2(img(i:i+7,j:j+7));
        for x = 1:8
            for y = 1:8
                img_zzs(i,j,zzPat(x,y)) = dctImg(x,y);
            end
        end
    end
end
[r,m] = size(img_zzs,1,2);
%%
% mean_FG = mean(D1_FG);
% mean_BG = mean(D1_BG);
% cov_FG = cov(D1_FG);
% cov_BG = cov(D1_BG);
% len_FG = length(D1_FG);
% len_BG = length(D1_BG);
% PY_FG = length(D1_FG)/(length(D1_FG)+length(D1_BG));
% PY_BG = length(D1_BG)/(length(D1_FG)+length(D1_BG));


% mean_FG = mean(D2_FG);
% mean_BG = mean(D2_BG);
% cov_FG = cov(D2_FG);
% cov_BG = cov(D2_BG);
% len_FG = length(D2_FG);
% len_BG = length(D2_BG);
% PY_FG = length(D2_FG)/(length(D2_FG)+length(D2_BG));
% PY_BG = length(D2_BG)/(length(D2_FG)+length(D2_BG));

% mean_FG = mean(D3_FG);
% mean_BG = mean(D3_BG);
% cov_FG = cov(D3_FG);
% cov_BG = cov(D3_BG);
% len_FG = length(D3_FG);
% len_BG = length(D3_BG);
% PY_FG = length(D3_FG)/(length(D3_FG)+length(D3_BG));
% PY_BG = length(D3_BG)/(length(D3_FG)+length(D3_BG));
```

```
mean_FG = mean(D4_FG);
mean_BG = mean(D4_BG);
cov_FG = cov(D4_FG);
cov_BG = cov(D4_BG);
len_FG = length(D4_FG);
len_BG = length(D4_BG);
PY_FG = length(D4_FG)/(length(D4_FG)+length(D4_BG));
PY_BG = length(D4_BG)/(length(D4_FG)+length(D4_BG));

errorPD = zeros(1,9); % Error of Predictive Distribution
errorML = zeros(1,9); % Error of Maximum Likehood
errorMAP = zeros(1,9); % Error of Maximun Per
%%
% Predictive Distribution
for a = 1:length(alpha)
    sigma_0 = diag(alpha(a)*W0);

    part1_FG = (len_FG*sigma_0/(cov_FG+len_FG*sigma_0))*mean_FG';
    part2_FG = (cov_FG/(cov_FG+len_FG*sigma_0))*mu0_FG';
    mu_n_FG = part1_FG+part2_FG;
    sigma_n_FG = (cov_FG*sigma_0)/(cov_FG+len_FG*sigma_0);
    sigma_n_FG_Comb = sigma_n_FG+cov_FG;

    part1_BG = (len_BG*sigma_0/(cov_BG+len_BG*sigma_0))*mean_BG';
    part2_BG = (cov_BG/(cov_BG+len_BG*sigma_0))*mu0_BG';
    mu_n_BG = part1_BG+part2_BG;
    sigma_n_BG = (cov_BG*sigma_0)/(cov_BG+len_BG*sigma_0);
    sigma_n_BG_Comb = sigma_n_BG+cov_BG;

    % BDR
    img_BDR = zeros([r,m]);
    X = zeros([1,64]);
    count = 0;
    for i = 1:row-8
        for j = 1:colm-8
            X(1,:) = img_zzs(i,j,:);
            PX_T_FG = log(sqrt((2*pi)^64*det(sigma_n_FG_Comb))^(-1)*exp(-(X-
mu_n_FG')/sigma_n_FG_Comb*(X-mu_n_FG')'/2)*PY_FG);
            PX_T_BG = log(sqrt((2*pi)^64*det(sigma_n_BG_Comb))^(-1)*exp(-(X-
mu_n_BG')/sigma_n_BG_Comb*(X-mu_n_BG')'/2)*PY_BG);

            if PX_T_FG > PX_T_BG
               img_BDR(i,j) = 1;
            end

            if mask(i,j) ~= img_BDR(i,j)
                count = count+1;
            end
        end
    end
%     figure(1)
%     subplot(3,3,a)
%     imagesc(img_BDR);
%     colormap(gray(255));
```

```
    errorPD(a) = count/(row*colm);
end
% Mximum Likehood
for a = 1:length(alpha)
    img_ML = zeros([r,m]);
    X = zeros([1,64]);
    count = 0;
    for i = 1:row-8
        for j = 1:colm-8
            X(1,:) = img_zzs(i,j,:);
            PX_FG = log(sqrt((2*pi)^64*det(cov_FG))^(-1)*exp(-(X-mean_FG)/cov_FG*(X-
mean_FG)'/2)*PY_FG);
            PX_BG = log(sqrt((2*pi)^64*det(cov_BG))^(-1)*exp(-(X-mean_BG)/cov_BG*(X-
mean_BG)'/2)*PY_BG);

            if PX_FG > PX_BG
               img_ML(i,j) = 1;
            end

            if mask(i,j) ~= img_ML(i,j)
                count = count+1;
            end
        end
    end
%     figure(2)
%     subplot(3,3,a)
%     imagesc(img_ML);
%     colormap(gray(255));

    errorML(a) = count/(row*colm);
end
% Maximum a posteriori
for a = 1:length(alpha)
    sigma_0 = diag(alpha(a)*W0);

    part1_FG = (len_FG*sigma_0/(cov_FG+len_FG*sigma_0))*mean_FG';
    part2_FG = (cov_FG/(cov_FG+len_FG*sigma_0))*mu0_FG';
    mu_n_FG = part1_FG+part2_FG;

    part1_BG = (len_BG*sigma_0/(cov_BG+len_BG*sigma_0))*mean_BG';
    part2_BG = (cov_BG/(cov_BG+len_BG*sigma_0))*mu0_BG';
    mu_n_BG = part1_BG+part2_BG;

    % BDR
    img_MAP = zeros([r,m]);
    X = zeros([1,64]);
    count = 0;
    for i = 1:row-8
        for j = 1:colm-8
            X(1,:) = img_zzs(i,j,:);
            PX_FG_MAP = log(sqrt((2*pi)^64*det(cov_FG))^(-1)*exp(-(X-
mu_n_FG')/cov_FG*(X-mu_n_FG')'/2)*PY_FG);
            PX_BG_MAP = log(sqrt((2*pi)^64*det(cov_BG))^(-1)*exp(-(X-
mu_n_BG')/cov_BG*(X-mu_n_BG')'/2)*PY_BG);
```

```
            if PX_FG_MAP > PX_BG_MAP
               img_MAP(i,j) = 1;
            end

            if mask(i,j) ~= img_MAP(i,j)
                count = count+1;
            end
        end
    end
%     figure(3)
%     subplot(3,3,a)
%     imagesc(img_MAP);
%     colormap(gray(255));

    errorMAP(a) = count/(row*colm);
end

%
figure(1)
hold on;
plot(alpha,errorPD);
plot(alpha,errorML);
plot(alpha,errorMAP);
hold off;
set(gca,'XScale','log');
legend('PD','ML','MAP')
title('Probability Error of Three Predictions Methods for D4,Strategy 2');
```