CS4044D- Machine Learning

Submitted By- Vishnu Sajith, B180474CS

1. Write a function (in Python or any language of your choice) to calculate the discriminant function for the given normal density equation (as given below) and prior probabilities.

$$g_i(\mathbf{x}) = -\frac{1}{2}(\mathbf{x}-\boldsymbol{\mu}_i)^t\boldsymbol{\Sigma}_i^{-1}(\mathbf{x}-\boldsymbol{\mu}_i) - \frac{d}{2}\ln 2\pi - \frac{1}{2}\ln|\boldsymbol{\Sigma}_i| + \ln P(\omega_i).$$

Design

- A discriminant function g() is defined which takes the following parameters:
  - x - Sample Points
  - mean - Mean of the Class
  - cov - Covariance Matrix of the Class
  - d - Dimension of the features
  - p - Prior Probability
- The data is imputed as an array along each feature and is then stacked to form a matrix using the numpy stack function, np.stack(). Mean, Transpose, Covariance, Inverse of Covariance Matrix and Determinant of Covariance Matrix is calculated using inbuilt functions from numpy viz; np.mean(), np.transpose() np.cov(), np.linalg.inv(), np.linalg.det(), respectively. Matrix multiplication is done using the np.dot() function and natural logarithm is taken using np.log() function
- The function accounts for all feature dimensions and all Prior Probabilities as they are passed as parameters.
- The above formula is generated and the calculated value of the discriminant function is returned.

```
def g(x,mean,cov,d,p):
    a = np.transpose(x-mean)
    if d == 1:
        cov_inverse = 1/cov
        det_cov = cov
    else:
        cov_inverse = np.linalg.inv(cov)
        det_cov = np.linalg.det(cov)
    a = np.dot(cov_inverse,a)
    a = np.dot((x-mean),a)
    a *= -0.5
    b = -0.5*d*np.log(2*math.pi)
    c = -0.5*np.log(det_cov)
    d = np.log(p)
    return a + b + c + d
```

2. Using the above discriminant function,
    a. Assume the prior probabilities of the first two categories are equal and are equal to 1/2 and that of the third category is zero. Design a dichotomizer for those two categories using the feature x1 alone.

Design

- A dichotomizer is is designed which takes following as parameters:
    - w - The sample data point to be classified.
    - x - The features used to classify the sample data point
    - m1- Mean of class W1
    - m2 - Mean of Class W2
    - m3 - Mean of Class W3
    - c1 - Covariance Matrix of Class W1
    - c2 - Covariance Matrix of Class W2
    - c3 - Covariance Matrix of Class W3
    - d - Dimension of features
- The function compares the discriminant value of the sample data point with the corresponding features and classifies the point to the class with maximum discriminant value.The discriminant value is calculated using the function created in the previous question.
- In this question the sample data points are classified with feature x1 alone and prior probabilities 0.5 for W1 and W2 and 0 for W3.

```
--------------------------------To be classified as W1-----------------------

[-5.01 -8.12 -3.68] is classified as W1 considering 1 features -5.01
[-5.43 -3.48 -3.54] is classified as W2 considering 1 features -5.43
[ 1.08 -5.52  1.66] is classified as W1 considering 1 features 1.08
[ 0.86 -3.78 -4.11] is classified as W1 considering 1 features 0.86
[-2.67  0.63  7.39] is classified as W1 considering 1 features -2.67
[4.94 3.29 2.08] is classified as W2 considering 1 features 4.94
[-2.51  2.09 -2.59] is classified as W1 considering 1 features -2.51
[-2.25 -2.13 -6.94] is classified as W1 considering 1 features -2.25
[ 5.56  2.86 -2.26] is classified as W2 considering 1 features 5.56
[ 1.03 -3.33  4.33] is classified as W1 considering 1 features 1.03

--------------------------------To be classified as W2-----------------------

[-0.91 -0.18 -0.05] is classified as W1 considering 1 features -0.91
[ 1.3  -2.06 -3.53] is classified as W1 considering 1 features 1.3
[-7.75 -4.54 -0.95] is classified as W2 considering 1 features -7.75
[-5.47  0.5   3.92] is classified as W2 considering 1 features -5.47
[ 6.14  5.72 -4.85] is classified as W2 considering 1 features 6.14
[3.6  1.26 4.36] is classified as W1 considering 1 features 3.6
[ 5.37 -4.63 -3.65] is classified as W2 considering 1 features 5.37
[ 7.18  1.46 -6.66] is classified as W2 considering 1 features 7.18
[-7.39  1.17  6.3 ] is classified as W2 considering 1 features -7.39
[-7.5  -6.32 -0.31] is classified as W2 considering 1 features -7.5

--------------------------------To be classified as W3-----------------------

[5.35 2.26 8.13] is classified as W2 considering 1 features 5.35
[ 5.12  3.22 -2.66] is classified as W2 considering 1 features 5.12
[-1.34 -5.31 -9.87] is classified as W1 considering 1 features -1.34
[4.48 3.42 5.19] is classified as W2 considering 1 features 4.48
[7.11 2.39 9.21] is classified as W2 considering 1 features 7.11
[ 7.17  4.33 -0.98] is classified as W2 considering 1 features 7.17
[5.75 3.97 6.65] is classified as W2 considering 1 features 5.75
[0.77 0.27 2.41] is classified as W1 considering 1 features 0.77
[ 0.9  -0.43 -8.71] is classified as W1 considering 1 features 0.9
[ 3.52 -0.36  6.43] is classified as W1 considering 1 features 3.52
```

b. Determine the percentage of points misclassified.

Design

- A function missclassification is created which accepts the following parameters
  - w1 - Class W1
  - w2 - Class W2
  - w3 - Class W3
  - m1 - Mean of Class W1
  - m2 - Mean of Class W2
  - m3 - Mean of Class W3
  - c1 - Covariance Matrix of Class W1

- ➢ c2 - Covariance Matrix of Class W2
- ➢ c3 -Covariance Matrix of Class W3
- ➢ d - Dimension of the features

❖ The function checks if the discriminant value with respect to all classes of a sample point from a particular class and see if the point is misclassified. It'll increment a counter upon encountering a misclassification and then find the percentage of points misclassified in each class and in the whole dataset.

❖ In this question only x1 feature is used and the prior probabilities for class W1 and W2 is 0.5 and for class W3 is 0.

```
----------Percentage of Missclassification----------

% of Missclassified Points in W1 is: 30.0%
% of Missclassified Points in W2 is: 30.0%
% of Missclassified Points in W3 is: 100.0%
% of Missclassified Points in the dataset: 53.333333333333336%
```

c. Repeat the above two steps, but now use the two features x1 and x2

Design

❖ The above procedure is repeated by calling the dichotomizer function and the missclassification function and passing the appropriate parameter for 2 features.

```
--------------------------------To be classified as W1--------------------------------

[-5.01 -8.12 -3.68] is classified as W1 considering 2 features [-5.01 -8.12]
[-5.43 -3.48 -3.54] is classified as W2 considering 2 features [-5.43 -3.48]
[ 1.08 -5.52  1.66] is classified as W1 considering 2 features [ 1.08 -5.52]
[ 0.86 -3.78 -4.11] is classified as W1 considering 2 features [ 0.86 -3.78]
[-2.67  0.63  7.39] is classified as W2 considering 2 features [-2.67  0.63]
[4.94 3.29 2.08] is classified as W2 considering 2 features [4.94 3.29]
[-2.51  2.09 -2.59] is classified as W2 considering 2 features [-2.51  2.09]
[-2.25 -2.13 -6.94] is classified as W1 considering 2 features [-2.25 -2.13]
[ 5.56  2.86 -2.26] is classified as W2 considering 2 features [5.56 2.86]
[ 1.03 -3.33  4.33] is classified as W1 considering 2 features [ 1.03 -3.33]

--------------------------------To be classified as W2--------------------------------

[-0.91 -0.18 -0.05] is classified as W1 considering 2 features [-0.91 -0.18]
[ 1.3  -2.06 -3.53] is classified as W1 considering 2 features [ 1.3  -2.06]
[-7.75 -4.54 -0.95] is classified as W2 considering 2 features [-7.75 -4.54]
[-5.47  0.5   3.92] is classified as W2 considering 2 features [-5.47  0.5 ]
[ 6.14  5.72 -4.85] is classified as W2 considering 2 features [6.14 5.72]
[3.6  1.26 4.36] is classified as W1 considering 2 features [3.6  1.26]
[ 5.37 -4.63 -3.65] is classified as W2 considering 2 features [ 5.37 -4.63]
[ 7.18  1.46 -6.66] is classified as W2 considering 2 features [7.18 1.46]
[-7.39  1.17  6.3 ] is classified as W2 considering 2 features [-7.39  1.17]
[-7.5  -6.32 -0.31] is classified as W1 considering 2 features [-7.5  -6.32]

--------------------------------To be classified as W3--------------------------------

[5.35 2.26] is classified as W2 considering 2 features [5.35 2.26]
[5.12 3.22] is classified as W2 considering 2 features [5.12 3.22]
[-1.34 -5.31] is classified as W1 considering 2 features [-1.34 -5.31]
[4.48 3.42] is classified as W1 considering 2 features [4.48 3.42]
[7.11 2.39] is classified as W2 considering 2 features [7.11 2.39]
[7.17 4.33] is classified as W2 considering 2 features [7.17 4.33]
[5.75 3.97] is classified as W2 considering 2 features [5.75 3.97]
[0.77 0.27] is classified as W1 considering 2 features [0.77 0.27]
[ 0.9  -0.43] is classified as W1 considering 2 features [ 0.9  -0.43]
[ 3.52 -0.36] is classified as W1 considering 2 features [ 3.52 -0.36]

-------------------------Percentage of Missclassification-------------------------

% of Missclassified Points in W1 is: 50.0%
% of Missclassified Points in W2 is: 40.0%
% of Missclassified Points in W3 is: 100.0%
% of Missclassified Points in the dataset: 63.333333333333336%
```

d. Repeat again, with all the three features taken.

Design

- The above procedure is repeated by calling the dichotomizer function and the missclassification function and passing the appropriate parameter for 3 features.

```
--------------------------------To be classified as W1--------------------------------

[-5.01 -8.12 -3.68] is classified as W1 considering 3 features [-5.01 -8.12 -3.68]
[-5.43 -3.48 -3.54] is classified as W1 considering 3 features [-5.43 -3.48 -3.54]
[ 1.08 -5.52  1.66] is classified as W1 considering 3 features [ 1.08 -5.52  1.66]
[ 0.86 -3.78 -4.11] is classified as W1 considering 3 features [ 0.86 -3.78 -4.11]
[-2.67  0.63  7.39] is classified as W2 considering 3 features [-2.67  0.63  7.39]
[4.94 3.29 2.08] is classified as W1 considering 3 features [4.94 3.29 2.08]
[-2.51  2.09 -2.59] is classified as W1 considering 3 features [-2.51  2.09 -2.59]
[-2.25 -2.13 -6.94] is classified as W1 considering 3 features [-2.25 -2.13 -6.94]
[ 5.56  2.86 -2.26] is classified as W2 considering 3 features [ 5.56  2.86 -2.26]
[ 1.03 -3.33  4.33] is classified as W1 considering 3 features [ 1.03 -3.33  4.33]

--------------------------------To be classified as W2--------------------------------

[-0.91 -0.18 -0.05] is classified as W2 considering 3 features [-0.91 -0.18 -0.05]
[ 1.3  -2.06 -3.53] is classified as W2 considering 3 features [ 1.3  -2.06 -3.53]
[-7.75 -4.54 -0.95] is classified as W2 considering 3 features [-7.75 -4.54 -0.95]
[-5.47  0.5   3.92] is classified as W2 considering 3 features [-5.47  0.5   3.92]
[ 6.14  5.72 -4.85] is classified as W2 considering 3 features [ 6.14  5.72 -4.85]
[3.6  1.26 4.36] is classified as W1 considering 3 features [3.6  1.26 4.36]
[ 5.37 -4.63 -3.65] is classified as W2 considering 3 features [ 5.37 -4.63 -3.65]
[ 7.18  1.46 -6.66] is classified as W2 considering 3 features [ 7.18  1.46 -6.66]
[-7.39  1.17  6.3 ] is classified as W2 considering 3 features [-7.39  1.17  6.3 ]
[-7.5  -6.32 -0.31] is classified as W2 considering 3 features [-7.5  -6.32 -0.31]

--------------------------------To be classified as W3--------------------------------

[5.35 2.26 8.13] is classified as W1 considering 3 features [5.35 2.26 8.13]
[ 5.12  3.22 -2.66] is classified as W2 considering 3 features [ 5.12  3.22 -2.66]
[-1.34 -5.31 -9.87] is classified as W1 considering 3 features [-1.34 -5.31 -9.87]
[4.48 3.42 5.19] is classified as W1 considering 3 features [4.48 3.42 5.19]
[7.11 2.39 9.21] is classified as W1 considering 3 features [7.11 2.39 9.21]
[ 7.17  4.33 -0.98] is classified as W2 considering 3 features [ 7.17  4.33 -0.98]
[5.75 3.97 6.65] is classified as W1 considering 3 features [5.75 3.97 6.65]
[0.77 0.27 2.41] is classified as W1 considering 3 features [0.77 0.27 2.41]
[ 0.9  -0.43 -8.71] is classified as W1 considering 3 features [ 0.9  -0.43 -8.71]
[ 3.52 -0.36  6.43] is classified as W1 considering 3 features [ 3.52 -0.36  6.43]

-------------------------Percentage of Missclassification-------------------------

% of Missclassified Points in W1 is: 20.0%
% of Missclassified Points in W2 is: 10.0%
% of Missclassified Points in W3 is: 100.0%
% of Missclassified Points in the dataset: 43.333333333333336%
```

e. Compare your results and conclude.

- ❖ The misclassification is highest when it comes to 2 features - 63.33%
- ❖ The lowest misclassification occurs in the case of 3 features - 43.33%
- ❖ The misclassification in the case of 1 feature is 53.33%, which is the intermediate of the above two cases.

f. Classify the points (1,2,1), (5,3,2), (0,0,0), (1,0,0) using each feature vector mentioned above and compare the results.

- ❖ A function classify is created with takes the following parameters:
  - ➢ x - The sample point to be classified.
  - ➢ d - The dimension of the features
- ❖ The input given in the question is added into a list.
- ❖ Using a loop I have iterated through the list and passed the points to the classify function.
- ❖ A nested loop which takes values in the range 1-3, is also used to appropriately pass the dimension of the features to the classify function.
- ❖ The classify function calls the dichotomizer function with the appropriate parameter to classify the points.

```
[[1, 2, 1]] is classified as W1 considering 1 features 1
[[1, 2, 1]] is classified as W1 considering 2 features [1, 2]
[[1, 2, 1]] is classified as W2 considering 3 features [[1, 2, 1]]
------------------------------------------------------------------
[[5, 3, 2]] is classified as W2 considering 1 features 5
[[5, 3, 2]] is classified as W2 considering 2 features [5, 3]
[[5, 3, 2]] is classified as W1 considering 3 features [[5, 3, 2]]
------------------------------------------------------------------
[[0, 0, 0]] is classified as W1 considering 1 features 0
[[0, 0, 0]] is classified as W1 considering 2 features [0, 0]
[[0, 0, 0]] is classified as W1 considering 3 features [[0, 0, 0]]
------------------------------------------------------------------
[[1, 0, 0]] is classified as W1 considering 1 features 1
[[1, 0, 0]] is classified as W1 considering 2 features [1, 0]
[[1, 0, 0]] is classified as W1 considering 3 features [[1, 0, 0]]
------------------------------------------------------------------
```

Inference and Conclusion

- From our observation above, we know that **percentage of misclassification is least when we consider 3-features**
- Hence we can conclude that the sample [1,2,1] and [5,3,2] are misclassified as W1 and W2 respectively when we consider 1 feature and 2 features.
- Hence, the sample **[1,2,1] most likely belongs to class W2 and sample [5,3,2] most likely belongs to W1.**
- The sample **[0,0,0] and [1,0,0] is classified as W1 considering 1,2, and 3 features.**