

Machine Learning Homework 1

B12508026 戴偉璿

October 29, 2025

Contents

1	Problem 1	3
1.1	Bayesian Decision Model	3
1.1.1	Model Description	3
1.1.2	Parameter Estimation	3
1.1.3	Feature Selection	3
1.1.4	Result	8
1.1.5	Others	10
1.2	Random Forest Classifier	10
1.2.1	Model Description	10
2	Problem 2	10

1 Problem 1

1.1 Bayesian Decision Model

1.1.1 Model Description

In this homework, I implement a Bayesian Decision Model to classify data points based on their features. The model assumes that the data points are generated from a mixture of Gaussian distributions, each corresponding to a different class.

A multivariate Gaussian distribution is defined as follows:

$$p(\mathbf{x}|\mu, \Sigma) = \frac{1}{(2\pi)^{d/2}|\Sigma|^{1/2}} \exp\left(-\frac{1}{2}(\mathbf{x} - \mu)^T \Sigma^{-1}(\mathbf{x} - \mu)\right) \quad (1)$$

where μ is the mean vector, Σ is the covariance matrix, and d is the dimensionality of the data.

1.1.2 Parameter Estimation

To estimate the parameters of the Gaussian distributions for each class, I need the prior probabilities $P(C_i)$, the mean vectors μ_i , and the covariance matrices Σ_i for each class C_i .

The prior probabilities are estimated as follow, N_i is the number of samples in class C_i and N is the total number of samples.

$$P(C_i) = \frac{N_i}{N} \quad (2)$$

The mean vector for each class is estimated as:

$$\mu_i = \frac{1}{N_i} \sum_{x \in C_i} x \quad (3)$$

The covariance matrix for each class is estimated as:

$$\Sigma_i = \frac{1}{N_i} \sum_{x \in C_i} (x - \mu_i)(x - \mu_i)^T \quad (4)$$

1.1.3 Feature Selection

In this homework, I implement a forward feature selection algorithm to select the best subset of features that maximizes the model's performance. The algorithm starts with an empty set of features and iteratively adds the feature that results in the highest increase in performance, measured by the Area Under the Curve (AUC) of the Receiver Operating Characteristic (ROC) curve.

The algorithm for the forward feature selection algorithm is as follows:

Algorithm 1 Forward Feature Selection (FFS)

Require: Dataset $D = \{(X_{\text{train}}, X_{\text{test}}, y_{\text{train}}, y_{\text{test}})\}$, feature dimension d **Ensure:** Best feature mask `best_mask`, best AUC `best_auc`, best posterior `best_prior`

```

1: Initialize best_mask  $\leftarrow [0, 0, \dots, 0]$ 
2: Initialize best_auc  $\leftarrow 0.0$ 
3: repeat
4:   local_best_auc  $\leftarrow -1$ 
5:   local_best_mask  $\leftarrow \text{best\_mask}$ 
6:   for  $i = 1$  to  $d$  do
7:     if best_mask $[i] = 1$  then
8:       continue
9:     end if
10:    local_mask  $\leftarrow \text{best\_mask}$ ; local_mask $[i] \leftarrow 1$ 
11:    Train model with local_mask to get AUC  $\text{auc}_i$  and posterior  $\text{prior}_i$ 
12:    if  $\text{auc}_i > \text{local\_best\_auc}$  then
13:      local_best_auc  $\leftarrow \text{auc}_i$ 
14:      local_best_mask  $\leftarrow \text{local\_mask}$ 
15:      local_best_prior  $\leftarrow \text{prior}_i$ 
16:    end if
17:  end for
18:  if local_best_auc  $\leq \text{best\_auc}$  then
19:    break
20:  else
21:    best_auc  $\leftarrow \text{local\_best\_auc}$ 
22:    best_mask  $\leftarrow \text{local\_best\_mask}$ 
23:    best_prior  $\leftarrow \text{local\_best\_prior}$ 
24:  end if
25: until no improvement in AUC
26: return best_auc, best_mask, best_prior

```

Following is the feature selected in each fold:

```

1 Fold 1: ['FaceWM', 'FaceWmax', 'Na-Chin', 'LFaceH', 'MouthW', 'LVermlionH', 'LVermlionC', 'NoseVol',
  ↪ 'NoseW', 'NasoFacialA']
2 Fold 2: ['FaceWM', 'FaceWL', 'FaceWmax', 'Na-Chin', 'LFaceH', 'MouthW', 'LVermlionH', 'LVermlionC',
  ↪ 'NoseSurfA', 'NoseW', 'NasoFacialA']
3 Fold 3: ['FaceWM', 'FaceWmax', 'Na-Chin', 'LFaceH', 'MouthW', 'LVermlionH', 'LVermlionC', 'NoseVol',
  ↪ 'NoseW', 'NasoFacialA']
4 Fold 4: ['FaceWM', 'FaceWL', 'FaceWmax', 'Na-Chin', 'Subn-Chin', 'LFaceH', 'LipH', 'UVermlionH',
  ↪ 'LVermlionH', 'UVermlionC', 'LVermlionC', 'NoseVol', 'NoseW', 'NasoFacialA']
5 Fold 5: ['FaceWM', 'FaceWmax', 'Na-Chin', 'Subn-Chin', 'LFaceH', 'LipH', 'UVermlionH', 'LVermlionH',
  ↪ 'UVermlionC', 'LVermlionC', 'NoseVol', 'NoseW', 'NasoFacialA']
6 Fold 6: ['FaceWM', 'FaceWmax', 'Subn-Chin', 'MouthW', 'LVermlionH', 'NoseVol', 'NoseW', 'NasoFacialA']

```

```

7 Fold 7: ['FaceWM', 'FaceWL', 'FaceWmax', 'Na-Chin', 'Subn-Chin', 'LFaceH', 'LipH', 'UVermlionH',
  ↳ 'LVermlionH', 'UVermlionC', 'LVermlionC', 'NoseVol', 'NoseW', 'NasoFacialA']
8 Fold 8: ['FaceWM', 'FaceWmax', 'MFaceH', 'LFaceH', 'MouthW', 'LVermlionH', 'LVermlionC', 'NoseVol',
  ↳ 'NoseW', 'NasoFacialA']
9 Fold 9: ['FaceWM', 'FaceWmax', 'Na-Chin', 'LFaceH', 'MouthW', 'LVermlionH', 'LVermlionC', 'NoseVol',
  ↳ 'NoseW', 'NasoFacialA']
10 Fold 10: ['FaceWM', 'FaceWmax', 'Na-Chin', 'LFaceH', 'MouthW', 'LVermlionH', 'LVermlionC', 'NoseVol',
  ↳ 'NoseW', 'NasoFacialA']
11 Fold 11: ['FaceWM', 'FaceWmax', 'Subn-Chin', 'LipD', 'MFaceH', 'LFaceH', 'MouthW', 'LipH',
  ↳ 'UVermlionH', 'LVermlionH', 'UVermlionC', 'LVermlionC', 'NoseVol', 'NasoFacialA']
12 Fold 12: ['FaceWM', 'FaceWmax', 'Na-Chin', 'LFaceH', 'MouthW', 'LVermlionH', 'LVermlionC', 'NoseVol',
  ↳ 'NoseW', 'NasoFacialA']
13 Fold 13: ['FaceWM', 'FaceWmax', 'Na-Chin', 'LFaceH', 'MouthW', 'LVermlionH', 'LVermlionC', 'NoseVol',
  ↳ 'NoseW', 'NasoFacialA']
14 Fold 14: ['FaceWM', 'FaceWmax', 'Na-Chin', 'LFaceH', 'MouthW', 'LVermlionH', 'LVermlionC', 'NoseVol',
  ↳ 'NoseW', 'NasoFacialA']
15 Fold 15: ['FaceWM', 'FaceWmax', 'Na-Chin', 'LFaceH', 'MouthW', 'LVermlionH', 'LVermlionC', 'NoseVol',
  ↳ 'NoseW', 'NasoFacialA']
16 Fold 16: ['FaceWM', 'FaceWL', 'FaceWmax', 'MFaceH', 'LFaceH', 'MouthW', 'LVermlionH', 'LVermlionC',
  ↳ 'NoseVol', 'NoseW', 'NasoFacialA']
17 Fold 17: ['FaceWM', 'FaceWmax', 'Na-Chin', 'LFaceH', 'MouthW', 'LVermlionH', 'LVermlionC', 'NoseVol',
  ↳ 'NoseW', 'NasoFacialA']
18 Fold 18: ['FaceWM', 'FaceWmax', 'Na-Chin', 'LFaceH', 'MouthW', 'LVermlionH', 'LVermlionC', 'NoseVol',
  ↳ 'NoseW', 'NasoFacialA']
19 Fold 19: ['FaceWM', 'FaceWmax', 'Na-Chin', 'LFaceH', 'MouthW', 'LVermlionH', 'LVermlionC', 'NoseVol',
  ↳ 'NoseW', 'NasoFacialA']
20 Fold 20: ['FaceWM', 'FaceWmax', 'Na-Chin', 'LFaceH', 'MouthW', 'LVermlionH', 'UVermlionC', 'LVermlionC',
  ↳ 'NoseVol', 'NoseW', 'NasoFacialA']
21 Fold 21: ['FaceWM', 'FaceWmax', 'Na-Chin', 'LFaceH', 'MouthW', 'LVermlionH', 'LVermlionC', 'NoseVol',
  ↳ 'NoseW', 'NasoFacialA']
22 Fold 22: ['FaceWL', 'FaceWmax', 'Na-Chin', 'LFaceH', 'LVermlionH', 'LVermlionC', 'NoseVol', 'NoseW',
  ↳ 'NasoFacialA']
23 Fold 23: ['FaceWM', 'FaceWmax', 'Na-Chin', 'LFaceH', 'MouthW', 'LVermlionH', 'LVermlionC', 'NoseVol',
  ↳ 'NoseW', 'NasoFacialA']
24 Fold 24: ['FaceWL', 'FaceWmax', 'Na-Chin', 'LFaceH', 'LVermlionH', 'NoseVol', 'NoseW', 'NasoFacialA']
25 Fold 25: ['FaceWM', 'FaceWmax', 'Na-Chin', 'LFaceH', 'MouthW', 'LVermlionH', 'LVermlionC', 'NoseVol',
  ↳ 'NoseW', 'NasoFacialA']
26 Fold 26: ['FaceWM', 'FaceWL', 'FaceWmax', 'Na-Chin', 'Subn-Chin', 'LFaceH', 'LipH', 'UVermlionH',
  ↳ 'LVermlionH', 'UVermlionC', 'LVermlionC', 'NoseVol', 'NoseW', 'NasoFacialA']
27 Fold 27: ['FaceWM', 'FaceWmax', 'Na-Chin', 'LFaceH', 'MouthW', 'LVermlionH', 'LVermlionC', 'NoseVol',
  ↳ 'NoseW', 'NasoFacialA']
28 Fold 28: ['FaceWM', 'FaceWL', 'FaceWmax', 'MFaceH', 'LFaceH', 'MouthW', 'LVermlionH', 'LVermlionC',
  ↳ 'NoseVol', 'NoseW', 'NasoFacialA']
29 Fold 29: ['FaceWM', 'FaceWL', 'FaceWmax', 'Na-Chin', 'Subn-Chin', 'LFaceH', 'LipH', 'UVermlionH',
  ↳ 'LVermlionH', 'UVermlionC', 'LVermlionC', 'NoseVol', 'NoseW', 'NasoFacialA']
30 Fold 30: ['FaceWM', 'FaceWL', 'FaceWmax', 'Na-Chin', 'Subn-Chin', 'LFaceH', 'LipH', 'UVermlionH',
  ↳ 'LVermlionH', 'UVermlionC', 'LVermlionC', 'NoseVol', 'NoseW', 'NasoFacialA']
31 Fold 31: ['FaceWM', 'FaceWL', 'FaceWmax', 'Na-Chin', 'Subn-Chin', 'MFaceH', 'LVermlionH', 'LVermlionC',
  ↳ 'NoseVol', 'NoseW', 'NasoFacialA']
32 Fold 32: ['FaceWM', 'FaceWL', 'FaceWmax', 'Na-Chin', 'Subn-Chin', 'LFaceH', 'LipH', 'UVermlionH',
  ↳ 'LVermlionH', 'UVermlionC', 'LVermlionC', 'NoseVol', 'NoseW', 'NasoFacialA']
33 Fold 33: ['FaceWM', 'FaceWL', 'FaceWmax', 'Na-Chin', 'Subn-Chin', 'LFaceH', 'LipH', 'UVermlionH',
  ↳ 'LVermlionH', 'UVermlionC', 'LVermlionC', 'NoseVol', 'NoseSurfA', 'NoseW', 'NasoFacialA']
34 Fold 34: ['FaceWM', 'FaceWL', 'FaceWmax', 'LipD', 'MFaceH', 'LFaceH', 'MouthW', 'LipH', 'UVermlionH',
  ↳ 'LVermlionH', 'UVermlionC', 'LVermlionC', 'NoseVol', 'NoseSurfA', 'NoseW', 'NasoFacialA']
35 Fold 35: ['FaceWM', 'FaceWmax', 'MFaceH', 'LFaceH', 'MouthW', 'LVermlionH', 'LVermlionC', 'NoseSurfA',
  ↳ 'NoseW', 'NasoFacialA']
36 Fold 36: ['FaceWL', 'FaceWmax', 'Na-Chin', 'Subn-Chin', 'MFaceH', 'LVermlionH', 'NoseVol', 'NoseW',
  ↳ 'NasoFacialA']

```

```

37 Fold 37: ['FaceWM', 'FaceWL', 'FaceWmax', 'Subn-Chin', 'MFaceH', 'LFaceH', 'LipH', 'UVermillionH',
  ↳ 'LVermillionH', 'UVermillionC', 'LVermillionC', 'NoseVol', 'NoseW', 'NasoFacialA']
38 Fold 38: ['FaceWM', 'FaceWL', 'FaceWmax', 'Na-Chin', 'LFaceH', 'LVermillionH', 'LVermillionC', 'NoseVol',
  ↳ 'NoseW', 'NasoFacialA']
39 Fold 39: ['FaceWM', 'FaceWL', 'FaceWmax', 'Na-Chin', 'Subn-Chin', 'LFaceH', 'MouthW', 'LVermillionH',
  ↳ 'LVermillionC', 'NoseVol', 'NoseSurfA', 'NoseW', 'NasoFacialA']
40 Fold 40: ['FaceWM', 'FaceWL', 'FaceWmax', 'Subn-Chin', 'MFaceH', 'LFaceH', 'LipH', 'UVermillionH',
  ↳ 'LVermillionH', 'UVermillionC', 'LVermillionC', 'NoseVol', 'NoseW', 'NasoFacialA']
41 Fold 41: ['FaceWM', 'FaceWL', 'FaceWmax', 'Na-Chin', 'Subn-Chin', 'MFaceH', 'LipH', 'UVermillionH',
  ↳ 'UVermillionC', 'LVermillionC', 'NoseVol', 'NoseW', 'NasoFacialA']
42 Fold 42: ['FaceWM', 'FaceWL', 'FaceWmax', 'Na-Chin', 'Subn-Chin', 'LFaceH', 'LipH', 'UVermillionH',
  ↳ 'LVermillionH', 'UVermillionC', 'LVermillionC', 'NoseVol', 'NoseW', 'NasoFacialA']
43 Fold 43: ['FaceWM', 'FaceWL', 'FaceWmax', 'Na-Chin', 'Subn-Chin', 'LFaceH', 'LipH', 'UVermillionH',
  ↳ 'LVermillionH', 'UVermillionC', 'LVermillionC', 'NoseVol', 'NoseW', 'NasoFacialA']
44 Fold 44: ['FaceWM', 'FaceWmax', 'LFaceH', 'LipH', 'UVermillionH', 'LVermillionH', 'LVermillionC', 'NoseVol',
  ↳ 'NoseW', 'NasoFacialA']
45 Fold 45: ['FaceWL', 'FaceWmax', 'Na-Chin', 'LFaceH', 'MouthW', 'LVermillionH', 'LVermillionC', 'NoseSurfA',
  ↳ 'NoseW', 'NasoFacialA']
46 Fold 46: ['FaceWL', 'FaceWmax', 'MFaceH', 'LFaceH', 'MouthW', 'LVermillionH', 'LVermillionC', 'NoseSurfA',
  ↳ 'NoseW', 'NasoFacialA']
47 Fold 47: ['FaceWM', 'FaceWmax', 'MFaceH', 'LFaceH', 'MouthW', 'LipH', 'LVermillionH', 'LVermillionC',
  ↳ 'NoseVol', 'NoseW', 'NasoFacialA']
48 Fold 48: ['FaceWU', 'FaceWM', 'FaceWL', 'LFaceH', 'LVermillionH', 'LVermillionC', 'NoseW', 'NasoFacialA']
49 Fold 49: ['FaceWM', 'FaceWmax', 'Na-Chin', 'LFaceH', 'MouthW', 'LVermillionH', 'LVermillionC', 'NoseVol',
  ↳ 'NoseW', 'NasoFacialA']
50 Fold 50: ['FaceWL', 'FaceWmax', 'MFaceH', 'LFaceH', 'MouthW', 'LVermillionH', 'LVermillionC', 'NoseVol',
  ↳ 'NoseW', 'NasoFacialA']
51 Fold 51: ['FaceWM', 'FaceWmax', 'Na-Chin', 'LFaceH', 'MouthW', 'LVermillionH', 'LVermillionC', 'NoseSurfA',
  ↳ 'NoseW', 'NasoFacialA']
52 Fold 52: ['FaceWM', 'FaceWL', 'FaceWmax', 'Na-Chin', 'Subn-Chin', 'LFaceH', 'LipH', 'UVermillionH',
  ↳ 'LVermillionH', 'UVermillionC', 'LVermillionC', 'NoseVol', 'NoseW', 'NasoFacialA']
53 Fold 53: ['FaceWM', 'FaceWmax', 'Na-Chin', 'LFaceH', 'MouthW', 'LVermillionH', 'LVermillionC', 'NoseVol',
  ↳ 'NoseW', 'NasoFacialA']
54 Fold 54: ['FaceWL', 'FaceWmax', 'MFaceH', 'LFaceH', 'MouthW', 'LVermillionH', 'LVermillionC', 'NoseVol',
  ↳ 'NoseW', 'NasoFacialA']
55 Fold 55: ['FaceWM', 'FaceWmax', 'Na-Chin', 'LFaceH', 'MouthW', 'LVermillionH', 'LVermillionC', 'NoseVol',
  ↳ 'NoseW', 'NasoFacialA']
56 Fold 56: ['FaceWM', 'FaceWmax', 'Na-Chin', 'LipD', 'MFaceH', 'LFaceH', 'MouthW', 'LipH', 'UVermillionH',
  ↳ 'LVermillionH', 'UVermillionC', 'LVermillionC', 'NoseVol', 'NasoFacialA']
57 Fold 57: ['FaceWM', 'FaceWL', 'FaceWmax', 'Na-Chin', 'LFaceH', 'LVermillionH', 'LVermillionC', 'NoseVol',
  ↳ 'NoseW', 'NasoFacialA']
58 Fold 58: ['FaceWM', 'FaceWL', 'FaceWmax', 'Na-Chin', 'MFaceH', 'LFaceH', 'LipH', 'UVermillionH',
  ↳ 'LVermillionH', 'UVermillionC', 'LVermillionC', 'NoseVol', 'NoseW', 'NasoFacialA']
59 Fold 59: ['FaceWM', 'FaceWL', 'FaceWmax', 'Subn-Chin', 'MFaceH', 'LFaceH', 'LipH', 'UVermillionH',
  ↳ 'LVermillionH', 'UVermillionC', 'LVermillionC', 'NoseVol', 'NoseSurfA', 'NoseW', 'NasoFacialA']
60 Fold 60: ['FaceWM', 'FaceWmax', 'MFaceH', 'LFaceH', 'MouthW', 'LVermillionH', 'LVermillionC', 'NoseVol',
  ↳ 'NoseW', 'NasoFacialA']
61 Fold 61: ['FaceWM', 'FaceWmax', 'Na-Chin', 'LFaceH', 'MouthW', 'LVermillionH', 'LVermillionC', 'NoseVol',
  ↳ 'NoseW', 'NasoFacialA']
62 Fold 62: ['FaceWM', 'FaceWL', 'FaceWmax', 'Na-Chin', 'Subn-Chin', 'LFaceH', 'LipH', 'UVermillionH',
  ↳ 'LVermillionH', 'UVermillionC', 'LVermillionC', 'NoseVol', 'NoseW', 'NasoFacialA']
63 Fold 63: ['FaceWM', 'FaceWmax', 'Subn-Chin', 'MouthW', 'LVermillionH', 'NoseVol', 'NoseW', 'NasoFacialA']
64 Fold 64: ['FaceWM', 'FaceWL', 'FaceWmax', 'Na-Chin', 'Subn-Chin', 'LFaceH', 'LipH', 'UVermillionH',
  ↳ 'LVermillionH', 'UVermillionC', 'LVermillionC', 'NoseVol', 'NoseW', 'NasoFacialA']
65 Fold 65: ['FaceWM', 'FaceWL', 'FaceWmax', 'Na-Chin', 'Subn-Chin', 'LFaceH', 'LipH', 'UVermillionH',
  ↳ 'LVermillionH', 'UVermillionC', 'LVermillionC', 'NoseVol', 'NoseW', 'NasoFacialA']
66 Fold 66: ['FaceWM', 'FaceWL', 'FaceWmax', 'Na-Chin', 'Subn-Chin', 'LFaceH', 'LipH', 'UVermillionH',
  ↳ 'LVermillionH', 'UVermillionC', 'LVermillionC', 'NoseVol', 'NoseW', 'NasoFacialA']

```

```

67 Fold 67: ['FaceWU', 'FaceWM', 'FaceWL', 'FaceWmax', 'MFaceH', 'LFaceH', 'MouthW', 'LipH', 'UVermilionH',
  ↳ 'UVermilionC', 'LVermlionC', 'NoseVol', 'NasoFacialA']
68 Fold 68: ['FaceWM', 'FaceWmax', 'Na-Chin', 'LFaceH', 'MouthW', 'LVermlionH', 'LVermlionC', 'NoseVol',
  ↳ 'NoseW', 'NasoFacialA']
69 Fold 69: ['FaceWM', 'FaceWL', 'FaceWmax', 'Na-Chin', 'Subn-Chin', 'LFaceH', 'LipH', 'UVermilionH',
  ↳ 'LVermlionH', 'UVermilionC', 'LVermlionC', 'NoseVol', 'NoseW', 'NasoFacialA']
70 Fold 70: ['FaceWM', 'FaceWL', 'FaceWmax', 'Na-Chin', 'Subn-Chin', 'LFaceH', 'LipH', 'UVermilionH',
  ↳ 'LVermlionH', 'UVermilionC', 'LVermlionC', 'NoseVol', 'NoseW', 'NasoFacialA']
71 Fold 71: ['FaceWM', 'FaceWmax', 'Na-Chin', 'LFaceH', 'MouthW', 'LVermlionH', 'LVermlionC', 'NoseVol',
  ↳ 'NoseW', 'NasoFacialA']
72 Fold 72: ['FaceWM', 'FaceWL', 'FaceWmax', 'Na-Chin', 'MFaceH', 'LFaceH', 'LipH', 'UVermilionH',
  ↳ 'LVermlionH', 'UVermilionC', 'LVermlionC', 'NoseVol', 'NoseW', 'NasoFacialA']
73 Fold 73: ['FaceWM', 'FaceWmax', 'Na-Chin', 'LFaceH', 'MouthW', 'LVermlionH', 'LVermlionC', 'NoseVol',
  ↳ 'NoseW', 'NasoFacialA']
74 Fold 74: ['FaceWL', 'FaceWmax', 'Na-Chin', 'LFaceH', 'MouthW', 'LVermlionH', 'LVermlionC', 'NoseSurfA',
  ↳ 'NoseW', 'NasoFacialA']
75 Fold 75: ['FaceWM', 'FaceWmax', 'LFaceH', 'MouthW', 'LVermlionH', 'NoseVol', 'NoseW', 'NasoFacialA']
76 Fold 76: ['FaceWM', 'FaceWmax', 'Na-Chin', 'LFaceH', 'MouthW', 'LVermlionH', 'LVermlionC', 'NoseVol',
  ↳ 'NoseW', 'NasoFacialA']
77 Fold 77: ['FaceWL', 'FaceWmax', 'MFaceH', 'LFaceH', 'LVermlionH', 'NoseVol', 'NoseW', 'NasoFacialA']
78 Fold 78: ['FaceWM', 'FaceWL', 'FaceWmax', 'Na-Chin', 'Subn-Chin', 'LFaceH', 'LipH', 'UVermilionH',
  ↳ 'LVermlionH', 'UVermilionC', 'LVermlionC', 'NoseVol', 'NoseW', 'NasoFacialA']
79 Fold 79: ['FaceWM', 'FaceWL', 'FaceWmax', 'Na-Chin', 'Subn-Chin', 'LFaceH', 'LipH', 'UVermilionH',
  ↳ 'LVermlionH', 'UVermilionC', 'LVermlionC', 'NoseVol', 'NoseW', 'NasoFacialA']
80 Fold 80: ['FaceWM', 'FaceWL', 'FaceWmax', 'Subn-Chin', 'MFaceH', 'LFaceH', 'MouthW', 'LipH', 'UVermilionH',
  ↳ 'LVermlionH', 'UVermilionC', 'LVermlionC', 'NoseVol', 'NoseW', 'NasoFacialA']
81 Fold 81: ['FaceWM', 'FaceWL', 'FaceWmax', 'Na-Chin', 'LFaceH', 'LVermlionH', 'LVermlionC', 'NoseVol',
  ↳ 'NoseW', 'NasoFacialA']
82 Fold 82: ['FaceWM', 'FaceWmax', 'Na-Chin', 'LFaceH', 'MouthW', 'LVermlionH', 'LVermlionC', 'NoseVol',
  ↳ 'NoseW', 'NasoFacialA']
83 Fold 83: ['FaceWM', 'FaceWmax', 'Na-Chin', 'LFaceH', 'MouthW', 'LVermlionH', 'LVermlionC', 'NoseVol',
  ↳ 'NoseW', 'NasoFacialA']
84 Fold 84: ['FaceWU', 'FaceWM', 'Na-Chin', 'LFaceH', 'MouthW', 'LipH', 'LVermlionH', 'LVermlionC', 'NoseW',
  ↳ 'NasoFacialA']
85 Fold 85: ['FaceWM', 'FaceWmax', 'Na-Chin', 'LFaceH', 'MouthW', 'LVermlionH', 'LVermlionC', 'NoseVol',
  ↳ 'NoseW', 'NasoFacialA']
86 Fold 86: ['FaceWL', 'FaceWmax', 'MFaceH', 'LFaceH', 'LVermlionH', 'NoseVol', 'NoseW', 'NasoFacialA']
87 Fold 87: ['FaceWM', 'FaceWmax', 'Na-Chin', 'LFaceH', 'MouthW', 'LVermlionH', 'LVermlionC', 'NoseVol',
  ↳ 'NoseW', 'NasoFacialA']
88 Fold 88: ['FaceWM', 'FaceWL', 'FaceWmax', 'Na-Chin', 'MFaceH', 'LFaceH', 'LVermlionH', 'LVermlionC',
  ↳ 'NoseVol', 'NoseW', 'NasoFacialA']
89 Fold 89: ['FaceWM', 'FaceWmax', 'MFaceH', 'LFaceH', 'MouthW', 'LVermlionH', 'LVermlionC', 'NoseVol',
  ↳ 'NoseW', 'NasoFacialA']
90 Fold 90: ['FaceWM', 'FaceWL', 'FaceWmax', 'Na-Chin', 'Subn-Chin', 'LFaceH', 'LipH', 'UVermilionH',
  ↳ 'LVermlionH', 'UVermilionC', 'LVermlionC', 'NoseVol', 'NoseW', 'NasoFacialA']
91 Fold 91: ['FaceWM', 'FaceWmax', 'Na-Chin', 'LFaceH', 'MouthW', 'LVermlionH', 'LVermlionC', 'NoseVol',
  ↳ 'NoseW', 'NasoFacialA']
92 Fold 92: ['FaceWM', 'FaceWL', 'FaceWmax', 'MFaceH', 'LFaceH', 'MouthW', 'LVermlionH', 'LVermlionC',
  ↳ 'NoseVol', 'NoseW', 'NasoFacialA']
93 Fold 93: ['FaceWM', 'FaceWmax', 'Na-Chin', 'LFaceH', 'MouthW', 'LVermlionH', 'LVermlionC', 'NoseVol',
  ↳ 'NoseW', 'NasoFacialA']
94 Fold 94: ['FaceWM', 'FaceWmax', 'Na-Chin', 'Subn-Chin', 'LFaceH', 'LipH', 'UVermilionH', 'LVermlionH',
  ↳ 'UVermilionC', 'LVermlionC', 'NoseVol', 'NoseW', 'NasoFacialA']
95 Fold 95: ['FaceWU', 'FaceWM', 'MFaceH', 'LFaceH', 'MouthW', 'LVermlionH', 'LVermlionC', 'NoseVol', 'NoseW',
  ↳ 'NasoFacialA']
96 Fold 96: ['FaceWM', 'FaceWmax', 'MFaceH', 'LFaceH', 'MouthW', 'LVermlionH', 'LVermlionC', 'NoseVol',
  ↳ 'NoseW', 'NasoFacialA']
97 Fold 97: ['FaceWM', 'FaceWL', 'FaceWmax', 'Na-Chin', 'LFaceH', 'MouthW', 'LVermlionH', 'LVermlionC',
  ↳ 'NoseSurfA', 'NoseW', 'NasoFacialA']

```

```

98 Fold 98: ['FaceWM', 'FaceWL', 'FaceWmax', 'MFaceH', 'LFaceH', 'MouthW', 'LVermlionH', 'LVermlionC',
  ↪ 'NoseSurfA', 'NoseW', 'NasoFacialA']
99 Fold 99: ['FaceWM', 'FaceWL', 'FaceWmax', 'Na-Chin', 'Subn-Chin', 'LFaceH', 'LipH', 'UVermlionH',
  ↪ 'LVermlionH', 'UVermlionC', 'LVermlionC', 'NoseVol', 'NoseW', 'NasoFacialA']
100 Fold 100: ['FaceWU', 'FaceWM', 'MFaceH', 'LFaceH', 'MouthW', 'LipH', 'UVermlionH', 'UVermlionC',
  ↪ 'LVermlionC', 'NoseVol', 'NoseSurfA', 'NoseW', 'NasoFacialA']
101 Fold 101: ['FaceWL', 'FaceWmax', 'MFaceH', 'LFaceH', 'MouthW', 'LVermlionH', 'LVermlionC', 'NoseSurfA',
  ↪ 'NoseW', 'NasoFacialA']
102 Fold 102: ['FaceWM', 'FaceWL', 'FaceWmax', 'MFaceH', 'LFaceH', 'MouthW', 'LVermlionH', 'LVermlionC',
  ↪ 'NoseVol', 'NoseSurfA', 'NoseW', 'NasoFacialA']
103 Fold 103: ['FaceWL', 'FaceWmax', 'MFaceH', 'LFaceH', 'LVermlionH', 'NoseVol', 'NoseW', 'NasoFacialA']

```

I'd tried AIC and BIC for feature selection, but they did not perform as well as AUC-based selection in this study (with $AUC < 0.7$ and selecting the same features repeatedly).

To discuss this phenomenon, take AIC as an example, AIC is defined as

$$AIC = 2k - 2 \ln(L), \quad (5)$$

where k is the number of parameters and L is the maximized likelihood of the model. Because the number of features (and thus parameters) in the dataset is relatively small, the penalty term $2k$ exerts only a weak influence on the criterion.

As a result, AIC becomes dominated by the likelihood term and fails to effectively penalize model complexity, leading to the selection of suboptimal feature subsets.

In contrast, using AUC directly as the objective function better reflects the model's discriminative ability and yields feature combinations that improve classification performance.

1.1.4 Result

Metrics with Feature Selection: The AUC of loocv with feature selection is 0.9233, Fig. 1 shows the ROC curve.

With threshold = 0.5, the accuracy, sensitivity, and specificity are listed below:

- Accuracy: 0.8350 (86/103)
- Sensitivity: 0.8537 (35/41)
- Specificity: 0.8226 (51/62)

Features Selected Frequency The frequency of selected features across all folds is summarized in Table 1.

The top 2 features selected are NasoFacialA, LVermlionH, and NoseW with LVermlionH and NoseW tied at 100 times, so I choose NasoFacialA and LVermlionH(alphabetical) to plot the decision boundary in Fig. 2.

Table 1: Frequency of Selected Features Across All Folds

Feature	Frequency	Feature	Frequency
NasoFacialA	103	FaceWL	55
LVermlionH	100	LipH	37
NoseW	100	UVermlionC	35
FaceWmax	99	UVermlionH	35
LFaceH	98	MFaceH	34
LVermlionC	95	Subn-Chin	33
NoseVol	92	NoseSurfA	15
FaceWM	91	FaceWU	5
Na-Chin	71	LipD	3
MouthW	62		

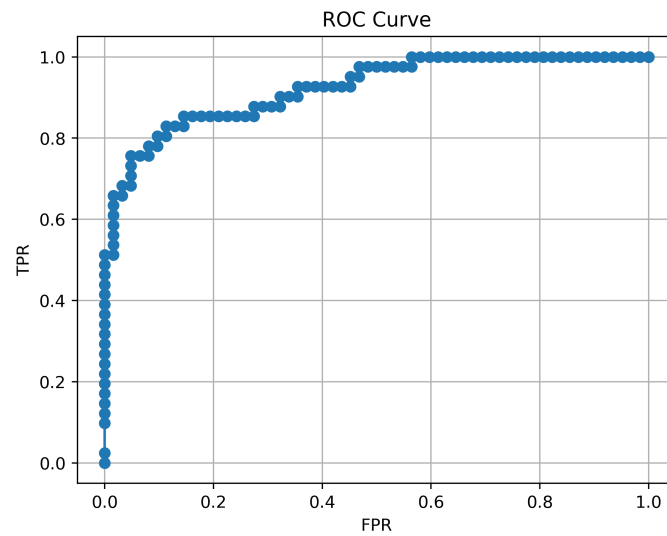


Figure 1: ROC curve with feature selection

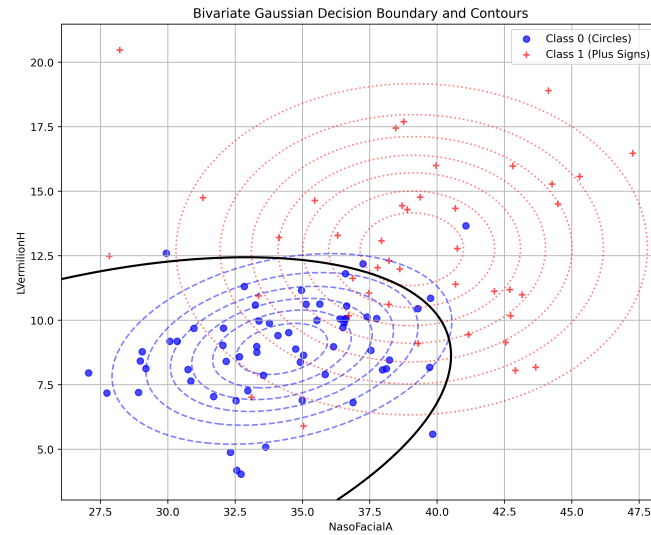


Figure 2: Decision boundary with feature selection

1.1.5 Others

1.2 Random Forest Classifier

1.2.1 Model Description

In this homework, I implement a Random Forest Classifier to classify data points based on their features. The model constructs a multitude of decision trees during training and outputs the mode of the classes (classification) of the individual trees.

2 Problem 2