NYCU Pattern Recognition, Homework 2

312553024, 江尚軒

Part. 1, Coding (60%):

(25%) Logistic Regression w/ Gradient Descent Method (slide ref)

1. (0%) Show the hyperparameters (learning rate and iteration, etc) that you used

```
LR = LogisticRegression(
    learning_rate=1e-2,
    num_iterations=10000,
)
```

2. (5%) Show your weights and the intercept of your model.

```
LR: Weights: [-0.50076954  0.10884352  0.60821965  0.06066269  0.13523584], Intercep: -1.8148004330971121
```

3. (5%) Show the AUC of the classification results on the testing set.

```
AUC=0.8500
```

4. (15%) Show the accuracy score of your model on the testing set

```
LR: Accuracy=0.8095
```

(25%) Fisher Linear Discriminant, FLD (slide_ref)

5. (0%) Show the mean vectors mi (i=0, 1) of each class of the training set.

```
FLD: m0=[ 0.35994138 -0.04560139], m1=[0.32519126 0.04435118]
```

6. (5%) Show the within-class scatter matrix *Sw* and between-class scatter matrix *Sb* of the training set.

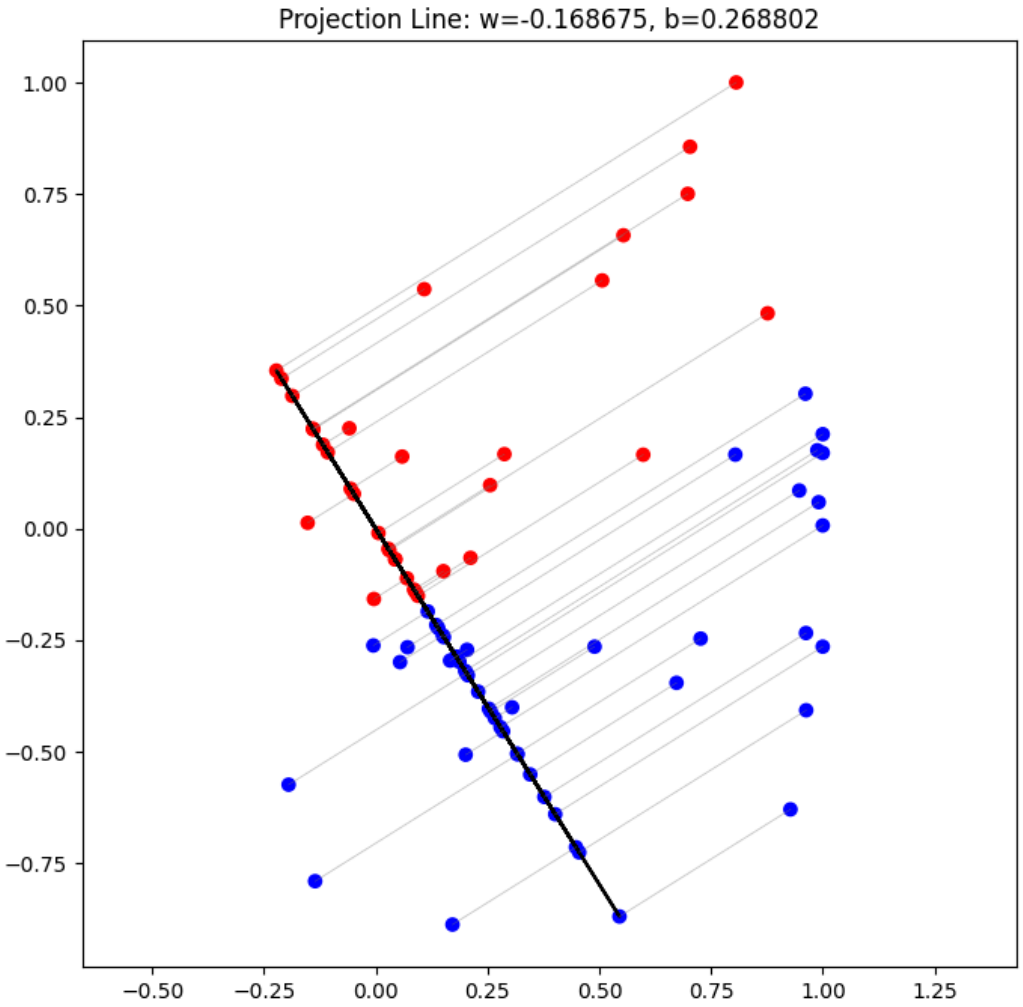
```
Sw=
[[0.50298617 0.18634845]
 [0.18634845 0.45157657]]
```

```
Sb=
[[ 0.00120757 -0.00312586]
 [-0.00312586  0.00809147]]
```

7. (5%) Show the Fisher's linear discriminant *w* of the training set.

```
w=
[[-0.16867464]
 [ 0.26880233]]
```

8. (15%) Obtain predictions for the testing set by measuring the distance between the projected value of the testing data and the projected means of the training data for the two classes. Show the accuracy score on the testing set.




```
FLD: Accuracy=0.7381
```

(10%) Code Check and Verification

9. (10%) Lint the code and show the PyTest results.

   Lint:

```
adsl-1-2@adsl-1-2:~/Andy/NYCU/碩一/碩一下/圖形識別/1122-pattern-recognition/Homework/HW2/code$ flake8 main.py
adsl-1-2@adsl-1-2:~/Andy/NYCU/碩一/碩一下/圖形識別/1122-pattern-recognition/Homework/HW2/code$
```

   PyTest:

```
(1122-pattern-recognition) (base) adsl-1-2@adsl-1-2:~/Andy/NYCU/碩一/碩一下/圖形識別/1122-pattern-recognition/Homework/HW2/code$ pytest ./test_main.py -s
============================= test session starts ==============================
platform linux -- Python 3.12.2, pytest-8.1.1, pluggy-1.4.0
rootdir: /home/adsl-1-2/Andy/NYCU/碩一/碩一下/圖形識別/1122-pattern-recognition/Homework/HW2/code
collected 2 items

test_main.py (395, 2) (395,)
2024-04-19 16:27:00.489 | INFO     | test_main:test_logistic_regression:35 - accuracy=0.9517
.(395, 2) (395,)
2024-04-19 16:27:00.490 | INFO     | test_main:test_fld:45 - accuracy=0.8759
.

============================== 2 passed in 6.87s ===============================
```

Part. 2, Questions (40%):

1. (10%) Is it suitable to use Mean Square Error (MSE) as the loss function for Logistic Regression? Please explain in detail.

Mean Square Error (MSE) is not suitable for Logistic Regression because:

1. MSE is designed for regression tasks, not classification.
2. It assumes continuous output, while logistic regression predicts probabilities.
3. MSE leads to non-convex optimization, making training difficult.
4. It's sensitive to outliers, which can skew the model's performance.

2. (15%) In page 31 of the lecture material (linear_classification.pdf), we introduce two methods for performing classification tasks using Fisher's linear discriminator: 1) Determining a threshold, 2) Using the k-NN (k-nearest neighbors) rule. Please discuss at least three aspects, either advantages or disadvantages, of using the k-NN method compared to determining a threshold (resources, performance, etc.).

    1. Adaptability: k-NN is more adaptable to complex data distributions, making it suitable for non-linear decision boundaries or multi-class classification.
    2. Interpretability: Determining a threshold provides a clear decision rule, while k-NN lacks interpretability due to reliance on neighboring data points.
    3. Computational Complexity: k-NN's computational complexity increases with dataset size, potentially leading to slower predictions, whereas determining a threshold is computationally efficient.

3. (15%) In logistic regression, what is the relationship between the sigmoid function and the softmax function? In what scenarios will the two functions be used respectively?

    In logistic regression, the sigmoid function and the softmax function are both activation functions for classification. However, the sigmoid function is used for binary classification, while the softmax function is used for multi-class classification.