Assignment6

October 2025

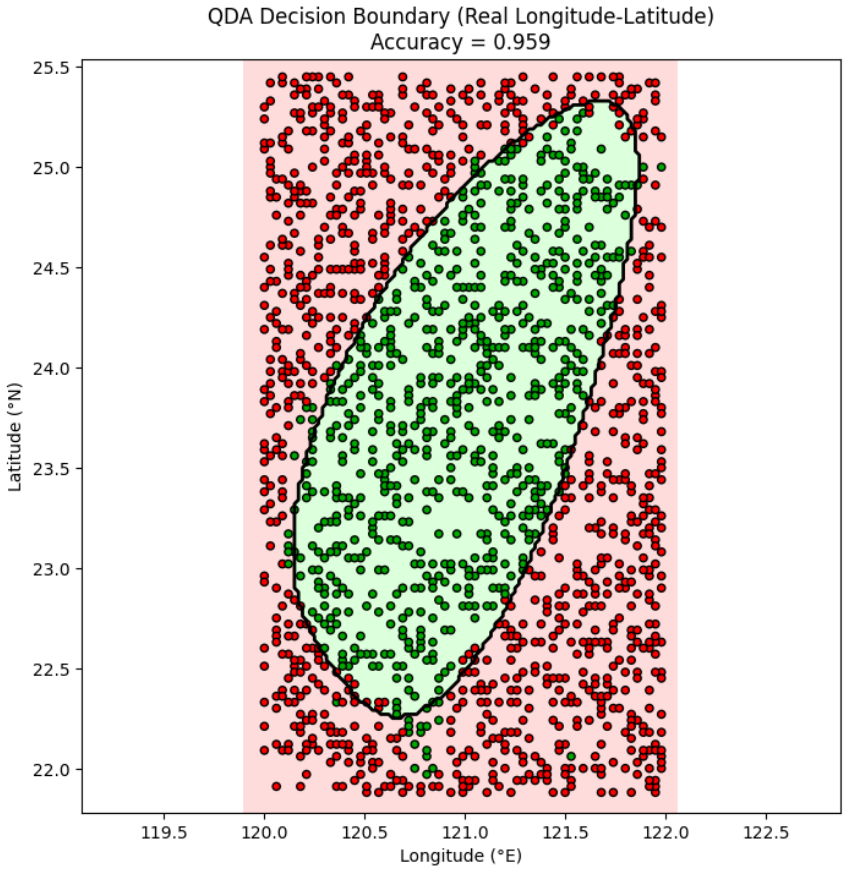
1 Question

I found that the speed of running the regression code in this assignment is way slower than the classification code.
I suspect that it is because of the neural network I use in the regression code since the code which I also use neural network in previous assignments is also slow.
How to make it faster?

2 Progamming report

1.
(b)The GDA model assumes that each class has sampled from Gaussian normal distribution and has its own covariance matrix. Also, it assumes a Bernoulli distribution for each class.
Thus, we have $P(x|y=i) = \frac{1}{(2\pi)^{n/2}|\sum_i|^{1/2}} exp(-\frac{1}{2}(x-\mu_i)^T \sum_i^{-1}(x-\mu_i)$, and $P(y) = \phi^y(1-\phi)^{1-y}$, i=0,1.
Then we learn the parameters $\mu_i, \sum_i, \phi$ by MLE using data in the training set.
Then by Baye's Theorem, we have $P(y=1|x) = \frac{P(x|y=1)P(y=1)}{P(x|y=1)P(y=1)+P(x|y=0)P(y=0)}$ which is the probability for classification.
Now look at the decision boundary, we have $\frac{1}{(2\pi)^{n/2}|\sum_i|^{1/2}} exp(-\frac{1}{2}(x-\mu_0)^T \sum_0^{-1}(x-\mu_0 = \frac{1}{(2\pi)^{n/2}|\sum_1|^{1/2}} exp(-\frac{1}{2}(x-\mu_1)^T \sum_i^{-1}(x-\mu_1))$.
Taking log, then we will get $-\frac{1}{2}x^T(\sum_0^{-1} - \sum_1^{-1})x + (\mu_0^T \sum_0^{-1} - \mu_1^T \sum_1^{-1})x - \frac{1}{2}(\mu_0^T\mu_0 - \mu_1^T\mu_1) - \frac{1}{2}ln\frac{|\sum_0|}{|\sum_1|}$ which is a quadratic form.
So the decision boundary is of a quadratic curve shape.
Since most of the data that are 1 in the data set should form a boundary line like the boundary of Taiwan which is a elliptic shape like, GDA can be a good choice doing this classification problem.
(c)Accuracy is approximately 96 percent, and the accuracy is given by $\frac{\sum 1\{y_{pred}=y_{test}\}}{\sum 1\{y_{test}\}}$
(d)The black line is the decision boundary.

```
Accuracy: 0.9590
```



2.

(b) I choose three points and output their value in previous regression model and classification model and h checking if h outputs the value we want as the definition of h.

(c) I choose 0.5 as the value making any result of the classification model in assignment 4 greater than it outputs 1. Then define h by using if the value of cls model is greater than 0.5, outputs the value of reg model else outputs -999.

(d)Plot that demonstrate the behavior of h

```
Example outputs of h(lon, lat):
R(120.30, 22.50) = 25.78,  C(120.30, 22.50) = 0.00,  h(120.30, 22.50) = -999.00
R(121.20, 23.80) = 13.38,  C(121.20, 23.80) = 1.00,  h(121.20, 23.80) = 13.38
R(121.80, 24.50) = 20.29,  C(121.80, 24.50) = 0.36,  h(121.80, 24.50) = -999.00
```