Week-5, Practice, Solutions

Question-1

The loss function used for binary classification using logistic regression is the binary cross-entropy loss.

Question-2

We are talking about logistic regression without feature transform. To understand the decision boundary, we see the following:

$$P(y = 1 \mid x) = \begin{cases} 1, & \sigma(w^T x) > 0.5 \\ 0, & \text{otherwise} \end{cases}$$

When is $\sigma(w^T x) > 0.5$? From the graph of the sigmoid function, we can see that this happens when $w^T x > 0$. From this, we see that the decision boundary in the case of a logistic regression model is linear. The other part is that the activation function is the sigmoid function, which is non-linear.

Question-3

This is something you must be aware of by now.

Question-4

While using gradient descent in logistic regression, we need to compute the derivatives of the cross-entropy loss with respect to the weights. But this process involves taking the derivative of the sigmoid function by chain rule. Therefore, we need to compute both derivatives.

Question-5

The formula for precision and recall are:

$$P = \frac{TP}{TP + FP}$$

$$R = \frac{TP}{TP + FN}$$

So, the values become:

$$P = \frac{35}{35 + 15} = 0.7$$

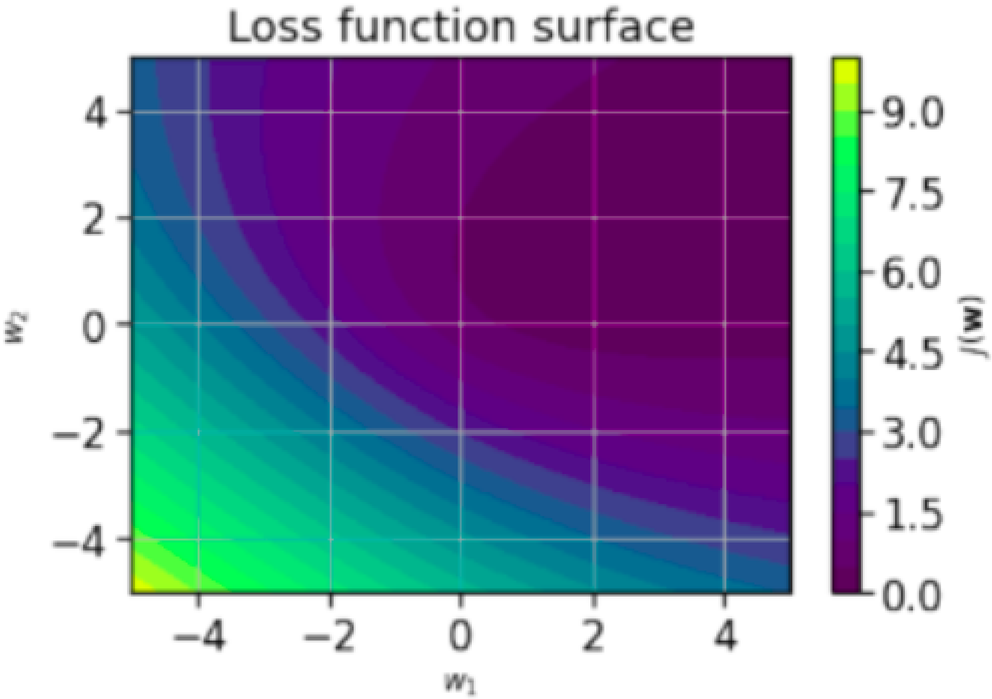
$$R = \frac{35}{35 + 15} = 0.7$$

Question-6

The correct Python expression for the gradient of the loss function with respect to w is:

```
np.transpose(X) @ (y_prob - y) + reg_rate * w
```

Question-7



From the contours, we see that the deep-blue regions have the lowest loss. So, after performing gradient descent for a sufficient number of epochs, the coordinates are likely to be in the deep-blue region. So, $(4, 4)$ is an ideal choice.

Question-8

The derivative of the sigmoid function is given by:

$$\sigma(z) \cdot (1 - \sigma(z))$$

Question-9

The logistic regression model outputs the following probability:

$$P(y = 1 \mid x) = \sigma(w^T x)$$