

Quiz 2 - Results



Attempt 1 of 1

Written Feb 25, 2024 4:12 PM - Feb 25, 2024 4:57 PM

Released Feb 28, 2024 1:00 PM

Attempt Score **23.5 / 25 - A-**

Overall Grade (Highest Attempt) **23.5 / 25 - A-**

Question 1

1.5 / 1.5 points

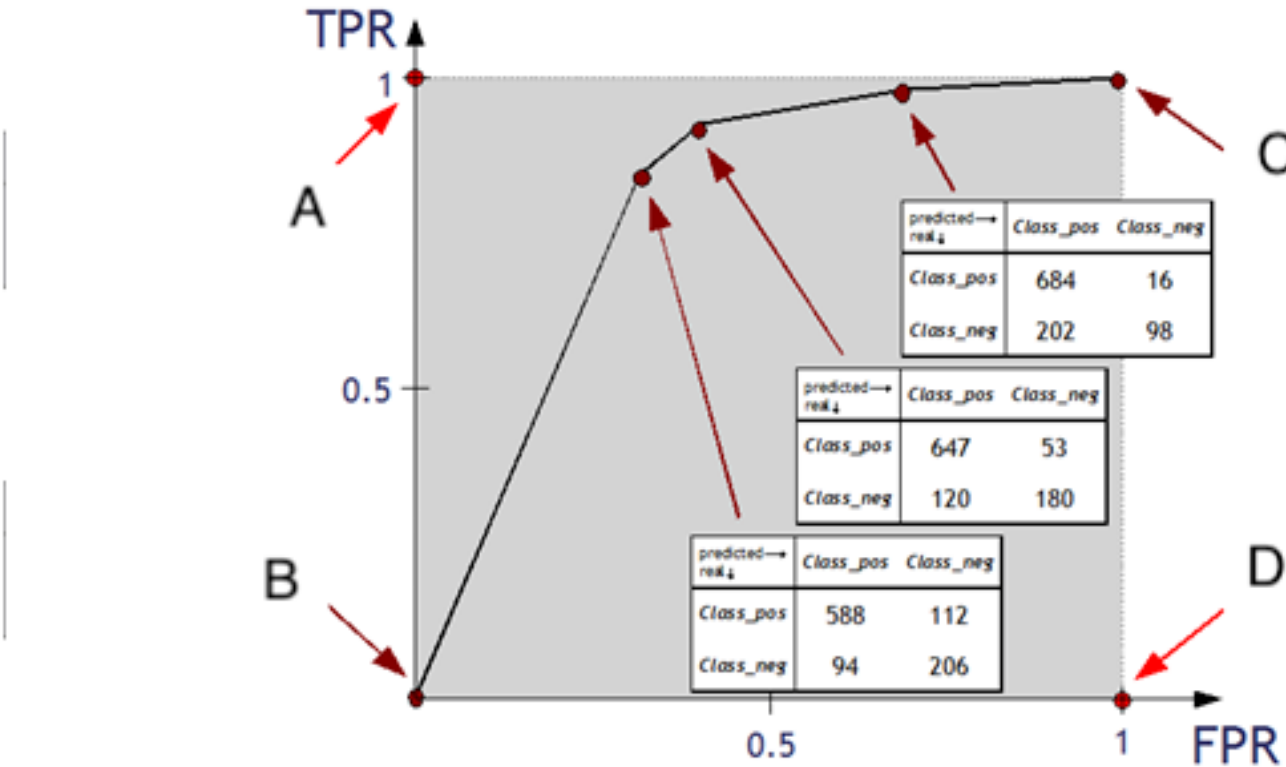
Select all that is true with respect to the F_β score

- ✓ ☐ $\beta < 1$ lends more weight to precision
- ✓ ☐ $\beta > 1$ lends more weight to recall
- ✓ ☐ $\beta < 1$ lends more weight to recall
- ✓ ☐ $\beta > 1$ lends more weight to precision

Question 2

2 / 2 points

Match the labels A, B, C and D with the corresponding confusion matrix



✓ 1

		Predicted	
		Class_Pos	Class_Neg
Real	Class_Pos	700	0
	Class_Neg	0	300

✓ 2

		Predicted	
		Class_Pos	Class_Neg
Real	Class_Pos	0	700
	Class_Neg	0	300

✓ 3

		Predicted	
		Class_Pos	Class_Neg
Real	Class_Pos	700	0
	Class_Neg	300	0

✓ 4

		Predicted	
		Class_Pos	Class_Neg
Real	Class_Pos	0	700
	Class_Neg	300	0

- 1. A
- 2. B
- 3. C
- 4. D

Question 3

1.5 / 3 points

Consider the outcome of a test to detect the presence of a radioactive element in drinking water. Given 700 samples of drinking water from different geographies, the test results are summarized as follows (with positive being the sample has radioactive element present or is contaminated):

		True Value	
		Radioactive Present	Radioactive Absent
Test Outcome	Radioactive Present	300	100
	Radioactive Absent	250	50

Based on the test outcome, answer the following questions.

You just need to fill the right numeric value for each question's answer. **For answers with decimal parts round up to 2 places.** For example, 4.667 should be written as 4.67 and not 4.66 when rounded to 2 decimal places.

How many samples did the test correctly classify?

___350___ ✓ (25 %)

What is the chance (in fraction [0.xx]) that a sample identified as not contaminated is in fact contaminated ?

(Note: if you calculation result is 0.1234, then please enter 0.12)

___0.17___ ✗ (0.83)

What is the chance (in fraction [0.xx]) that a sample identified as contaminated is in fact contaminated ?

(Note: if you calculation result is 0.1234, then please enter 0.12)

___0.25___ ✗ (0.75)

What is the F_1 score (in fraction [0.xx]) for this test ?

(Note: if you calculation result is 0.1234, then please enter 0.12)

___0.63___ ✓ (25 %)

Question 4

1 / 1 point

Select the correct expression for regularization loss in ridge regression where:

$$\hat{y} = a_0 + \sum_{j=1}^p a_j x_j$$



$$\sum_i (y_i - \sum_{j=1}^p x_j a_j)^2 + \lambda \sum_{j=0}^p a_j^2$$



$$\sum_i (y_i - a_0 - \sum_{j=1}^p x_j a_j)^2 + \lambda \sum_{j=1}^p a_j^2$$



$$\sum_i (y_i - a_0 - \sum_{j=1}^p x_j a_j) + \lambda \sum_{j=1}^p a_j^2$$



$$\sum_i (y_i - a_0 - \sum_{j=1}^p x_j a_j)^2 + \lambda \sum_{j=0}^p a_j^2$$

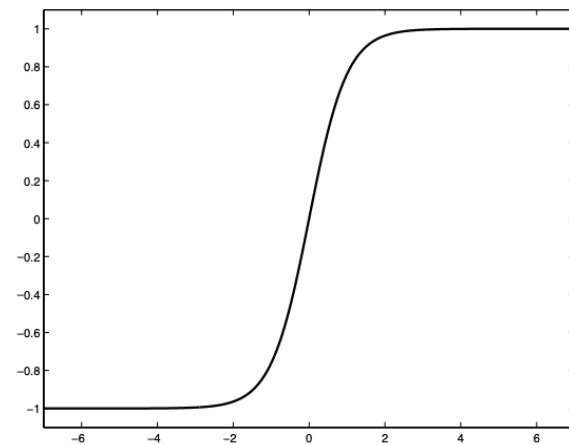
Question 5

2 / 2 points

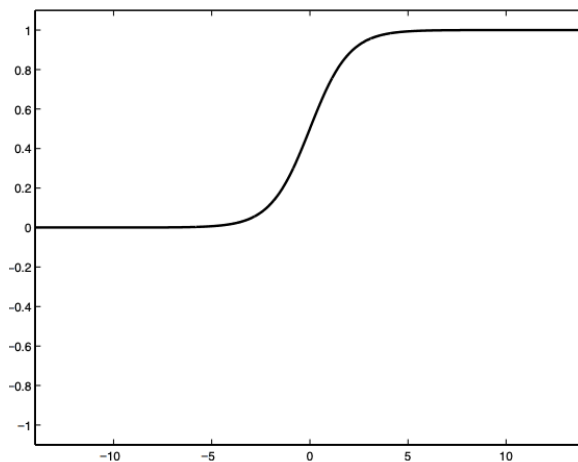
Match the graphs with their corresponding activation functions.

- ✓ 3 Hard TanH
- ✓ 1 TanH
- ✓ 4 Linear
- ✓ 2 Sigmoid

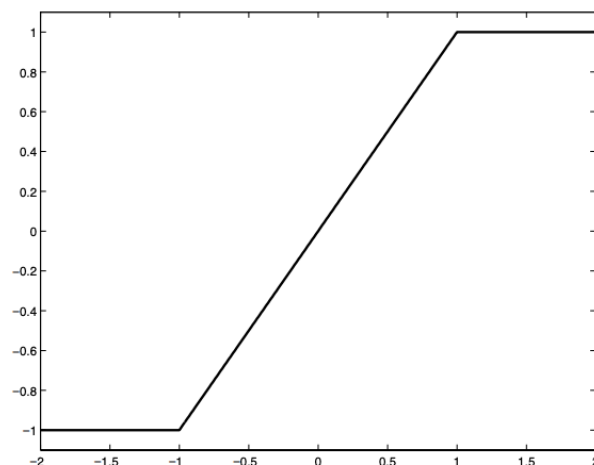
1.



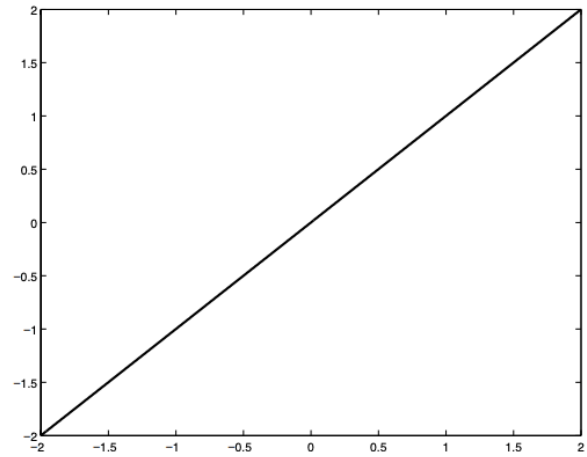
2.



3.



4.



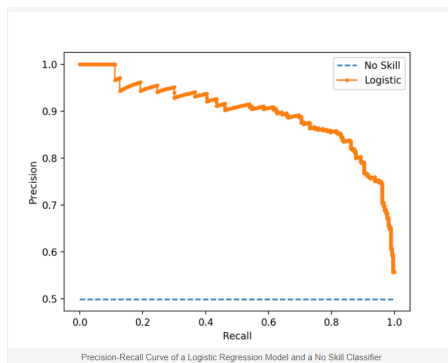
Question 6

2 / 2 points

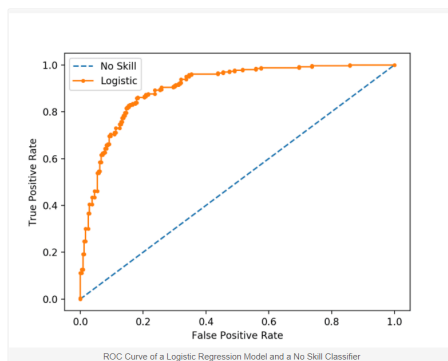
Classify the following graphs as belonging to either a balanced dataset or an imbalanced dataset.

Precision-Recall Curve of a Logistic Regression Model and a No Skill Classifier

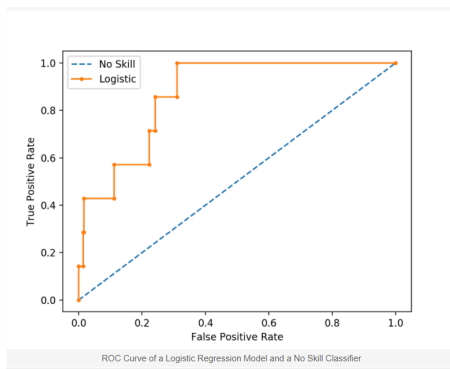
1. Balanced Dataset
2. Imbalanced Dataset

✓ 1

ROC Curve of a Logistic Regression Model and a No Skill Classifier

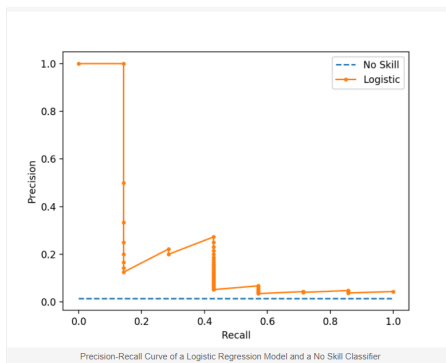
✓ 1✓ 2

ROC Curve of a Logistic Regression Model and a No Skill Classifier



Precision-Recall Curve of a Logistic Regression Model and a No Skill Classifier

✓ 2



Question 7

1.5 / 1.5 points

Select the true statement(s) for a linearly separable binary dataset.

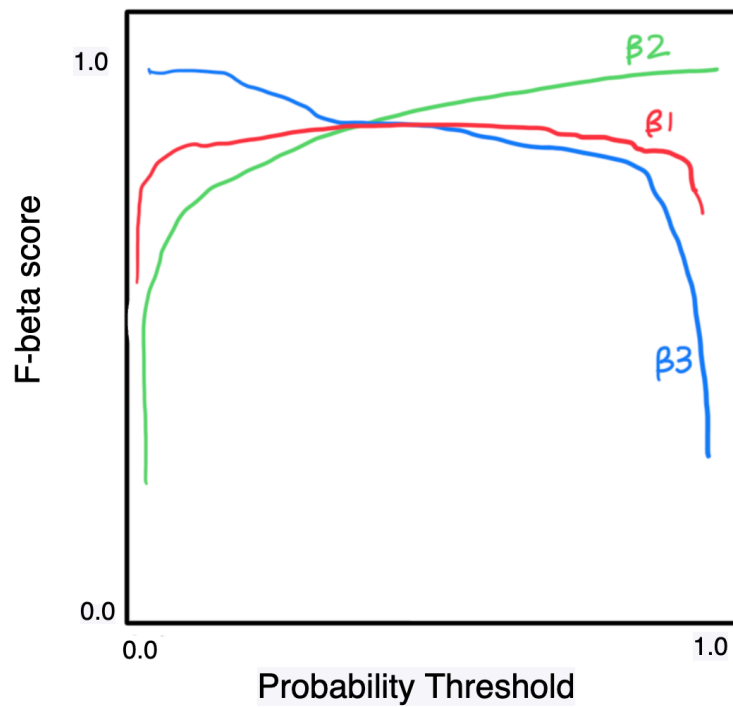
- ✓ ☐ The intersection of the convex hulls of the two classes is empty.
- ✓ ☐ Non-linear transformations can further improve the training accuracy of perceptrons for this dataset.
- ✓ ☐ When a single layer perceptron is trained with this dataset it may or may not converge depending on the separation margin between the two classes.
- ✓ ☐ Only a support vector machine (SVM) with non-linear kernels can guarantee 100% accuracy with this dataset.

Question 8

1.5 / 1.5 points

The following graph was obtained when F_β scores were plotted against probability thresholds for a classifier, where β values = (0.01, 1, 100)

Match the correct β value (**{0.01, 1, 100}**) with their curve labels.



- ✓ 3 β_3 1. 0.01
- ✓ 1 β_2 2. 1
- ✓ 2 β_1 3. 100

Question 9

1 / 1 point

Given $\text{sigmoid}(z) = 0.5$, what is the value of $\tanh(z)$?

- ☐ 0.5
- ☐ -1.0
- ✓ ☒ 0
- ☐ 1.0

Question 10

1 / 1 point

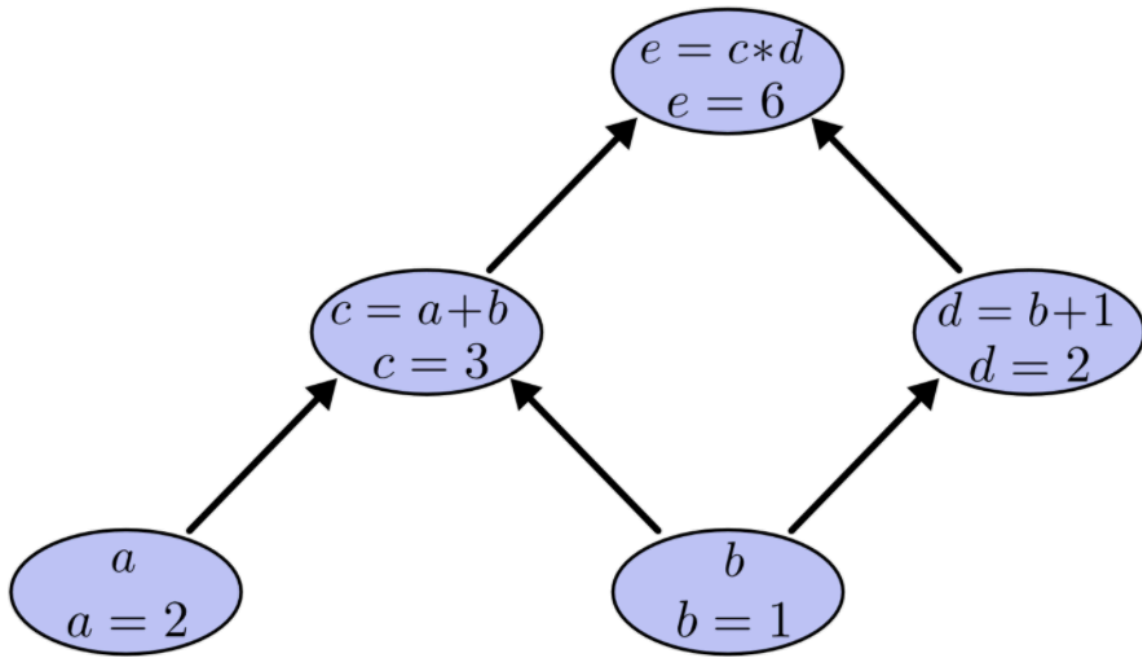
For the purpose of a classifier that can recognize cats ($y=1$) vs. dogs ($y=0$). Which one of these activation functions would you recommend using for the output layer?

- ✓ ☒ Sigmoid
- ☐ TanH
- ☐ Leaky ReLU
- ☐ ReLU

Question 11

1.5 / 1.5 points

Calculate the values for $\partial e / \partial a$ and $\partial e / \partial b$ given the below computational graph.



$\partial e / \partial a$

___2___ ✓(50 %)

$\partial e / \partial b$

___5___ ✓(50 %)

Question 12

1 / 1 point

Which of the following activation functions is most common for hidden layers?

- ☐ TanH
- ☐ Sigmoid
- ✓ ☒ ReLU
- ☐ Linear

Question 13

1 / 1 point

The numerical output of the ReLU function has range



$(-\infty, \infty)$

☐

$[0, 1]$

☐

$(-\infty, 0]$

☒

$[0, \infty)$

Question 14**1 / 1 point**

The F_β score depends on the probability threshold

☒ True

☐ False

Question 15**1 / 1 point**

Select the correct equation for gradient descent, where $\frac{\partial f(x_n)}{\partial x}$ is the gradient of f with respect to x evaluated at x_n and α is the learning rate.

☐ $x_{n+1} \leftarrow \alpha x_n + \alpha \left(\frac{\partial f(x_n)}{\partial x} \right)$

☒ $x_{n+1} \leftarrow x_n - \alpha \left(\frac{\partial f(x_n)}{\partial x} \right)$

☐ $x_{n+1} \leftarrow \alpha x_n - \alpha \left(\frac{\partial f(x_n)}{\partial x} \right)$

☐ $x_{n+1} \leftarrow x_n + \alpha \left(\frac{\partial f(x_n)}{\partial x} \right)$

Question 16**1.5 / 1.5 points**

Which of the following is/are true for Regularization parameter $C (=1/\lambda)$?

- ✓ ☐ Low C leads to complex models
- ✓ ☐ High C leads to complex models
- ✓ ☐ Low C leads to simple models
- ✓ ☐ High C leads to simple models

Question 17**1.5 / 1.5 points**

As the probability threshold increases from 0 to 1, which of the following is always observed?

- ✓ ☐ Precision decreases
- ✓ ☐ Recall increases
- ✓ ☐ Recall decreases
- ✓ ☐ Precision increases

Done