

Examen parcial 1

Puntos totales 26/27 ?

Introducción al aprendizaje de máquina

0 de 0 puntos

Nombre(s) *

Mariana

Apellidos *

Ávalos Arce

ID *

0197495

Lee con atención y contesta. 26 de 27 puntos

✓

1. Supón que has entrenado un algoritmo de clasificación binaria utilizando regresión logística. Para un nuevo dato de entrada x , el algoritmo genera como salida una predicción $h(x)=0.7$. Esto significa que (elige todas las opciones que sean correctas):

2/2

☒

La probabilidad de que el dato x pertenezca a la clase 0 es 0.3

✓

☐

La probabilidad de que el dato x pertenezca a la clase 0 es 0.7

☐

La probabilidad de que el dato x pertenezca a la clase 1 es 0.3

☒

La probabilidad de que el dato x pertenezca a la clase 1 es 0.7.

✓

✓

2. Considera el dato 1 del set de datos de la figura. En regresión logística, cuál de las siguientes expresiones es verdad? Selecciona todas las que sean verdaderas.

2/2

x_1	x_2	y
1	0.5	0
1	1.5	0
2	1	1
3	1	0

☐

Durante el entrenamiento, si la hipótesis tiende a 0, el error se incrementará

☒

Durante el entrenamiento, si la hipótesis tiende a 0, el error disminuirá

✓

☐

Durante el entrenamiento, si la hipótesis tiende a 0, es porque la probabilidad de que el dato pertenezca a la clase 0 es menor

☒

Durante el entrenamiento, si la hipótesis tiende a 0, entonces el dato de entrenamiento está siendo bien clasificado

✓



✓

3. Observa el set de datos de la figura. ¿Qué valor tiene el error para el dato 3 si la hipótesis es 0.3. Considera la definición de J(a):

1/1

x_1	x_2	y
1	0.5	0
1	1.5	0
2	1	1
3	1	0

$$J(a) = \frac{1}{m} \sum_{i=1}^m \left[-y^{(i)} \ln(h_a(x^{(i)})) - (1 - y^{(i)}) \ln(1 - h_a(x^{(i)})) \right]$$

1.20

✓

✓

4. Imagina que trabajas en la predicción del clima y quieres predecir si el clima es soleado, nublado o lluvioso. ¿Qué tipo de algoritmo de aprendizaje de máquina utilizarías para predecir el clima de mañana?

1/1

☐ Regresión

☒ Clasificación

✓

✓

5. Imagina que trabajas en la predicción del mercado de valores y quieres predecir si cierta compañía se declarará en quiebra dentro de los próximos 7 días (mediante datos de empresas similares que anteriormente habían estado en riesgo de quiebra). ¿Con qué tipo de algoritmo de aprendizaje de máquina abordarías este problema?

1/1

☒ Clasificación

☐ Regresión

✓

✓

6. ¿Cuál o cuáles de los siguientes problemas utilizan algoritmos de aprendizaje supervisado? Selecciona todos los que apliquen.

2/2

☒ Dado el ADN de un conjunto de personas, predecir la probabilidad de que desarrollen diabetes en los próximos 10 años

✓

☐ Dado un set de datos del estado de salud de pacientes que sufren enfermedades del corazón, identificar si tal vez estos pacientes se puedan agrupar en distintos grupos para darles un tratamiento en particular dependiendo del grupo en el que se encuentren.

☒ Examinar (mediante un programa de computadora) un conjunto de imágenes de perros y gatos para clasificarlas.

✓

☐ Dado un set de 1000 datos de pacientes con la información de cómo han respondido a un tratamiento, identificar si hay diferentes "tipos" de pacientes en cuanto a cómo han respondido al tratamiento y qué categorías serían.

✓

7. Observa el set de datos de la figura. En regresión logística, ¿Qué valor tiene el error para el dato 2 si $a=[0, 1, 1]$?. Considera la definición de J(a):

1/1

x_1	x_2	y
1	0.5	0
1	1.5	0
2	1	1
3	1	0

$$J(a) = \frac{1}{m} \sum_{i=1}^m \left[-y^{(i)} \ln(h_a(x^{(i)})) - (1 - y^{(i)}) \ln(1 - h_a(x^{(i)})) \right]$$

2.57

✓

https://docs.google.com/forms/d/e/1FAIpQLSfmohG_lew-p-gQNDayt6RxbWzxSqoGDNFougUbAvkOhVl-g/viewscore?vc=0&c=0&w=1&flr=0&viewscore=AE0zAgC1SZPKKhBiqZ3rU2ly_hl12O_ay6nALtx6EZZ6zJLv... 2/6

✓ 8. En regresión logística, ¿Cuál de las opciones representa la forma correcta para actualizar los valores de "a" con el método del gradiente descendiente? Selecciona todas las opciones que sean correctas. Además, considera que en regresión logística, la derivada parcial de la función de costo se define como:

$$\frac{\partial J(a)}{\partial a_j} = \frac{1}{m} \sum_{i=1}^m (h_a(x^{(i)}) - y^{(i)}) x_j^{(i)}$$

$$a_j = a_j - \beta \frac{1}{m} \sum_{i=1}^m \left(\frac{1}{1 + e^{-a^T x^{(i)}}} - y^{(i)} \right) x_j^{(i)}$$

✓ ☒ Opción 1 ✓

$$a_j = a_j - \beta \frac{1}{m} \sum_{i=1}^m (h_a(x^{(i)}) - y^{(i)}) x_j^{(i)}$$

✓ ☒ Opción 2 ✓

$$a_j = a_j - \beta \frac{1}{m} \sum_{i=1}^m (a^T x^{(i)} - y^{(i)}) x_j^{(i)}$$

☐ Opción 3

$$a_j = a_j - \beta \frac{1}{m} \sum_{i=1}^m (h_a(x^{(i)}) - y^{(i)}) x^{(i)}$$

☐ Opción 4

✗ 9. Selecciona todas las sentencias que sean verdaderas. 1/2

- ☒ La función sigmoide nunca genera valores más grandes que 1 ✓
- ☐ En regresión logística, la función de costo J siempre será mayor que cero para un conjunto de datos de entrenamiento igual o mayor que uno
- ☒ En regresión logística, si el valor de a' * X es menor que cero, entonces el valor de la hipótesis será menor que 0.5 ✓
- ☐ La función de costo del error cuadrático medio es utilizada en regresión logística

Respuesta correcta

- ☒ La función sigmoide nunca genera valores más grandes que 1
- ☒ En regresión logística, la función de costo J siempre será mayor que cero para un conjunto de datos de entrenamiento igual o mayor que uno
- ☒ En regresión logística, si el valor de a' * X es menor que cero, entonces el valor de la hipótesis será menor que 0.5



✓ 10. Considera el set de datos de entrenamiento de la figura. Si "n" es el número de características, ¿Qué valor tiene "n" en este set de datos? 1/1

x	y
3	2
1	2
0	1
4	3

- ☒ 1 ✓
- ☐ 2
- ☐ 3
- ☐ 4
- ☐ Otro:

✓ 11. Considera el set de datos de la figura. ¿Qué valor tiene el error $J(a_0,a_1)$ si $a_0=1$ y $a_1=0$? Recuerda que en regresión lineal univariable, la hipótesis y la función de costo se definen como: 1/1

$$J(a_0,a_1) = \frac{1}{2m} \sum_{i=1}^m (h_a(x^{(i)}) - y^{(i)})^2$$
$$h_a(x^{(i)}) = a_0 + a_1x^{(i)}$$

x	y
3	2
1	2
0	1
4	3

0.75 ✓

✓ 12. Considera el set de datos de entrenamiento de la figura. En regresión lineal, si $a_0=1$ y $a_1=2$, ¿Qué valor tiene la hipótesis $h_a(x_i)$ para el dato $i=3$? 1/1

$$h_a(x^{(i)}) = a_0 + a_1x^{(i)}$$

x	y
3	2
1	2
0	1
4	3

1 ✓

✓ 13. En regresión logística, ¿Qué valor tiene la hipótesis para el dato $x=[1, 1.5]$ si $a=[0, -1, -1]$? Considera que la hipótesis está definida como: 1/1

$$h_a(x^{(i)}) = g(a^T x^{(i)})$$
$$z = a^T x^{(i)}$$
$$g(z) = \frac{1}{1 + e^{-z}}$$

- ☐ 0.09
- ☐ 0.05
- ☐ 0.07



- ☒

0.07

✓
- ☐

0.03
- ☐

Otro:

- ✓

14. En regresión lineal univariable, el método del gradiente descendiente es utilizado para minimizar la función de costo $J(a_0,a_1)$ que representa el error. Considerando esto, selecciona todas las opciones que sean falsas.

2/2
- ☐

No importa qué valores iniciales tengan a_0 y a_1 . Mientras que el valor del factor de aprendizaje beta sea adecuado, el algoritmo convergerá a la misma solución.
- ☒

Entre más grande sea el valor del factor de aprendizaje beta, más lento converge al error mínimo el método del gradiente descendiente.

✓
- ☐

Si a_0 y a_1 son inicializados con los valores que generan el error mínimo, entonces los valores de a_0 y a_1 nunca cambiarán durante el proceso iterativo del algoritmo.
- ☒

Si las primeras iteraciones del método del gradiente descendiente causan que el error J se incremente, lo más seguro es que el factor de aprendizaje beta sea muy pequeño.

✓

- ✓

15. Supón que después de aplicar regresión logística a un set de datos de entrenamiento, obtienes ciertos valores de "a" tal que $J(a)>0$. Considerando esto, selecciona todas las opciones que sean verdaderas.

2/2
- ☒

Es probable que los datos de entrenamiento no sean linealmente separables

✓
- ☐

No es posible obtener valores de "a" tal que $J(a)>0$
- ☒

Es probable que se requieran más iteraciones para que $J(a)$ disminuya

✓
- ☐

Con los valores obtenidos de "a" se ha logrado clasificar correctamente todos los datos de entrenamiento

- ✓

16. Imagina que estás implementando el método del gradiente descendiente con $\beta=0.3$. Al ejecutar el algoritmo durante 5 iteraciones, te das cuenta que el error J disminuye muy poco. ¿Qué valor de beta elegirías para tratar de disminuir más rápido el error?

1/1
- ☒

0.9

✓
- ☐

0.1
- ☐

No cambiarías el valor de beta ya que sólo es cuestión de tiempo para que el error disminuya.
- ☐

10

- ✓

17. Considera el set de datos de la figura. ¿Cuál es el valor normalizado de la característica 1 del dato 3? Utiliza normalización estándar.

1/1
- | x1 | x2 | y |
|----|------|----|
| 89 | 7921 | 96 |
| 72 | 5184 | 74 |
| 94 | 8836 | 87 |
| 69 | 4761 | 78 |
- ☐

0.74
- ☐

-0.84
- ☐

-0.12
- ☐

1.21
- ☒

1.05

✓
- ☐

Otro:



✓

18. Se tiene un set con 14 datos de entrenamiento, cada dato con 3 características. ¿De qué tamaño son "X", y "y"?

1/1

☒

"X" es de 14x3 y "y" es de 14x1

✓

☐

"X" es de 3x14 y "y" es de 1x14

☐

"X" es de 14x3 y "y" es de 1x14

☐

"X" es de 3x14 y "y" es de 14x1

✓

19. Observa el set de datos de entrenamiento de la figura. Si se quiere aplicar regresión logística, ¿consideras que los datos deben ser normalizados antes?

2/2

x1	x2	y
89	7921	0
72	5184	1
94	8836	1
69	4761	0

☐ Si, pero sólo la columna de x2

☐ No es necesario normalizar

☐ Si, todas las columnas deben ser normalizadas, incluyendo la de "y"

☐ Si, pero sólo la columna de x1

☒ Si, pero sólo las columnas de x1 y x2

☐ Si, pero sólo la columna de "y"

✓

✓

20. Considera el set de datos de la figura. En regresión logística, ¿Qué valor tiene la hipótesis para el dato 1 si "a" es un vector de unos?

1/1

x1	x2	y
89	7921	0
72	5184	1
94	8836	1
69	4761	0

0.90

✓

Opcional: puedes subir una imagen o un archivo PDF con los procedimientos y cálculos de los ejercicios.

 cap - Mariana Áv...

