



[Course](#) > [Clasific...](#) > [Labora...](#) > Lab-01-2

## Lab-01-2

Abrir notebook

<https://github.com/lab-ml-itba/Laboratorio-01>

### Pregunta 1

0.0/1.0 point (graded)

Utilizar el dataset `generate_gaussians_distributions()` y entrenar la regresión logística.

Cuanto da el accuracy?

Submit

You have used 0 of 5 attempts

### Pregunta 2

0.0/1.0 point (graded)

Utilizar el dataset `generate_ellipse_data()` y entrenar la regresión logística.

Cuanto da el accuracy?



You have used 0 of 5 attempts

---

## Pregunta 3

0.0/1.0 point (graded)

Indicar cuales de las afirmaciones son correctas para el dataset `generate_ellipse_data()`

- ☐ El modelo entrenado detecta todos como rojos
- ☐ El accuracy es igual a la fracción de observaciones rojas
- ☐ El accuracy es igual a la fracción de observaciones azules
- ☐ Si hubiese una recta que pueda separar al menos una observación azul, esa sería mejor que la obtenida

You have used 0 of 2 attempts

---

## Problema 3.1

0.0/1.0 point (graded)

Usar el método `predict_proba` para calcular la probabilidad de una observación para el dataset `generate_ellipse_data()`.

Tener en cuenta que este método espera una array de dimensión  $(N, 2)$ , donde 2 es la dimensión de las observaciones y N es la cantidad de observaciones cuya probabilidad quiero obtener.

Si quiero obtener la probabilidad de una sola observación tendré que pasarle un numpy array de tamaño  $(2, 1)$ .

Para castearlo puede hacer lo siguiente: `np.array([x1, x2]).reshape(1, 2)`

Cual es la probabilidad de que la observación dada por el punto  $[0.5, -0.75]$  sea rojo?

You have used 0 of 5 attempts

---

## Problema 3.2

0.0/1.0 point (graded)

Cual es la probabilidad de que la observación dada por el punto [0.5, -0.75] sea azul?

You have used 0 of 5 attempts

---

## Problema 3.3

0.0/1.0 point (graded)

¿Donde está la recta?

Puede usar el siguiente código para graficar la recta:

```
x1 = np.linspace(-2, 2, 100)
x2 = -(model.coef_[0][0] * x1 + model.intercept_[0]) / model.coef_[0][1]
plt.plot(x1, x2, color='#414e8a', linewidth=2)
```

Analizar y ejecutar el código en la notebook y responder

Cual es la pendiente de la recta?

Cual es la ordenada al origen de la recta?

¿Cuántos parámetros aprendió el modelo?

Submit

You have used 0 of 5 attempts

---

## Problema 3.4

0.0/1.0 point (graded)

Importar la función `log_loss` de `sklearn`: [https://scikit-learn.org/stable/modules/generated/sklearn.metrics.log\\_loss.html](https://scikit-learn.org/stable/modules/generated/sklearn.metrics.log_loss.html)

Leer la documentación y calcularla.

El primer parámetro que recibe son los valores de  $y$ , y el segundo son las predicciones de probabilidad para cada  $X$ . (Usar `model.predict_proba` para calcularlas)

Esta función calcula el promedio de las entropías cruzadas entre los valores verdaderos ( $y$ ) y la estimación de las probabilidades de cada observación ( $X_i$ ).

Cuanto vale la entropía cruzada?

Submit

You have used 0 of 5 attempts

---

## Problema 4

0.0/1.0 point (graded)

Agregue columnas con el cuadrado de cada columna, el producto y una columna de unos. Eso esta indicado en la función `add_columns` de la notebook. Es importante mantener el orden que se pide en la función para que la librería de graficación funcione correctamente

Cuanto da el accuracy?

Submit

You have used 0 of 5 attempts

---

## Pregunta 5

0.0/1.0 point (graded)

Indicar cuales de las siguientes afirmaciones son correctas

- ☐ La opción `fit_intercept` esta en `False` ya que le estoy pasando una columna de todos unos que hace las veces del bias (ordenada al origen)
- ☐ El resultado de accuracy de la regresión logística teoricamente debería haber sido 1 para el dataset `generate_elipse_data` ya que una cuadrática puede fraccionar el plano con una elipse
- ☐ El resultado de accuracy de la regresión logística teoricamente debería haber sido 1 para el dataset `generate_gaussians_distributions` (Sin agregar columnas) ya que es linealmente separable

Submit

You have used 0 of 2 attempts

---

## Pregunta 6

0.0/1.0 point (graded)

La regresión logística en `sklearn` tiene un parámetro `C` que se explicará en teoría mas adelante. Modifique ese parámetro y busque si encuentra algún valor de `C` con el que se obtienen mejores accuracies

Indicar cual de las opciones es correcta

- ☐ Para valores de  $C$  menores a 0.001 se logra un accuracy de 1 para ambos datasets
- ☐ Para un valor de  $C$  igual a 0 se logra un accuracy de 1 para ambos datasets
- ☐ Para valores de  $C$  mayores a 100000 se logra un accuracy de 1 para ambos datasets
- ☐ No hay forma de lograr accuracies de 1 en ambos datasets

Submit

You have used 0 of 2 attempts