



Fundamentos de IA y Machine Learning

Lección 3: Modelos de Regresión y Clasificación I

Modelos de Regresión y Clasificación I

CASO PRÁCTICO

Vamos a dividir el caso práctico en tres problemas similares a los vistos en clase, y en un ejercicio de búsqueda de información.

Problema I

Se desea estimar el valor Y de una casa dada su superficie X_1 y el número de habitaciones X_2 . Hasta ahora, se conocen los siguientes datos de entrenamiento:

Y	120	125	130	112	122	135	128	140	150	160
X_1	60	55	70	60	62	70	60	80	80	85
X_2	1	2	1	2	1	2	3	1	2	3

Se pide hallar:

1. El modelo de regresión lineal simple de Y en función de X_1 .
2. El modelo de regresión lineal simple de Y en función de X_2 .
3. El modelo de regresión lineal múltiple de Y en función de X_1 y X_2 .
4. Validar cada modelo con el siguiente conjunto de *test*.

Y	120	128	140	115	120	145	125	170	135	155
X_1	58	50	80	65	55	75	55	80	70	75
X_2	1	3	1	2	1	2	3	3	1	3

Problema II

Haciendo uso de la base de datos *Iris* vista anteriormente, se ha ejecutado un algoritmo para estimar los parámetros de una regresión logística múltiple. Como el problema tiene tres clases: setosa (1), versicolor (2) y virgínica (3), se han generado los dos modelos siguientes:

$$f_1(\mathbf{X}, \beta_1) = 541.0741 - 33.121X_1 - 10.2824X_2 - 38.5734X_3 - 90.4374X_4$$

$$f_2(\mathbf{X}, \beta_2) = 501.1562 - 26.3497X_1 - 24.2689X_2 - 10.7613X_3 - 130.0915X_4$$

Se pide:

1. Hallar las predicciones del modelo para el siguiente conjunto de *test*.
2. Evaluar el rendimiento del clasificador en dicho conjunto.
3. Comparar el modelo obtenido con el expuesto en la lección 3.

Patrón	X_1	X_2	X_3	X_4	Clase
1	4.6	3.2	1.4	0.2	1
2	5.3	3.7	1.5	0.2	1
3	5.7	4.4	1.5	0.4	1
4	5.0	3.5	1.6	0.6	2
5	5.5	2.5	4.0	1.3	2
6	5.7	3.0	4.2	1.2	2
7	5.7	2.8	4.1	1.3	3
8	5.8	2.7	5.1	1.9	3
9	6.3	2.5	5.0	1.9	3
10	5.9	3.0	5.1	1.8	3

Problema III

Para un problema de clasificación binaria y con tres características se han recogido los siguientes patrones de entrenamiento.

Patrón	X_1	X_2	X_3	Clase
1	4.6	3.2	1.4	1
2	5.3	3.7	1.5	3
3	5.7	4.4	1.5	1
4	5.0	3.5	1.6	2
5	5.5	2.5	4.0	1
6	5.7	3.0	4.2	2
7	5.7	2.8	4.1	2
8	5.8	2.7	5.1	1
9	6.3	2.5	5.0	2
10	5.9	3.0	5.1	3

Se pide:

1. Hallar las predicciones del modelo para el siguiente conjunto de *test*, con el valor de $k = 1$ y $k = 3$ vecinos, y la distancia euclídea.
2. Evaluar el rendimiento del clasificador en dicho conjunto para ambos valores de k .

Patrón	X_1	X_2	X_3	Clase
1	5	3.5	1.7	1
2	4.3	2.8	1.5	1
3	2.7	4.5	1.2	3
4	5.0	4.2	1.3	3
5	6.3	2.5	4.1	1
6	5.2	3.0	4.5	2
7	4.5	3	4.2	2
8	5.9	2.9	5.2	2
9	5	2.4	5.1	1
10	4.5	3.2	5.0	2

Problema IV

Explicar en qué consiste la metodología de *RandomForest*.