

Parte I

Conceptos

Parte II

Modelos

1. Modelos Lineales

Si disponemos de atributos numéricos y buscamos predecir un atributo también numérico, entonces es natural explorar los modelos lineales. Estos modelos asumen que el atributo a predecir $y^{(k)}$ responde a la forma $w_0 + \sum_i w_i x_i^{(k)}$, es decir a un hiperplano. Como se espera la presencia de errores dentro de las mediciones o simplemente se busca aproximar la forma “real” de y se minimiza alguna métrica de distancia entre $y^{(k)}$ y $y^{(k)'} = w_0 + \sum_i w_i x_i$ como pueden ser según [1]:

- Mínima Suma de Cuadrados

$$\min \sum_i (y^{(i)} - y^{(i)'})^2$$

- Mínimos Valores Absolutos

$$\min \sum_i |y^{(i)} - y^{(i)'}|$$

- M, L y S Estimadores

- Mínima Suma Podada de Cuadrados (Least Trimmed Squares)

$$\min \sum_j (y^{(j)} - y^{(j)'})^2, \quad \text{donde} \quad \{y^{(j)} - y^{(j)'}\} \subset \{y^{(i)} - y^{(i)'}\}$$

- Mínima Media de Cuadrados

$$\min \text{med}_i (y^{(i)} - y^{(i)'})^2$$

Segun [2], tenemos que

2. Ejemplos

```
// Cargamos los datos de alguna manera
Instances dataset = ...
// Preparamos una instancia de prueba
Instance testInstance = ...

// Creamos una instancia del modelo
Classifier model = new SimpleLinearRegression();
```

```
// Entrenamos con los datos de 'dataset'
model.buildClassifier(dataset);

// Podemos imprimir los detalles del modelo
System.out.println(model);
// Tambien podemos predecir una instancia
double prediction = model.classifyInstance(testInstance);
```

También podemos utilizar `LinearRegression()`¹, `MultilayerPerceptron()`² y con algunas modificaciones `Logistic()`.

Parte III

Evaluacion

Referencias

- [1] David Ruppert y Raymond J Carroll. «Trimmed least squares estimation in the linear model». En: *Journal of the American Statistical Association* 75.372 (1980), págs. 828-838.
- [2] I.H. Witten, E. Frank y M.A. Hall. *Data Mining: Practical Machine Learning Tools and Techniques*. The Morgan Kaufmann Series in Data Management Systems. Elsevier Science, 2011. ISBN: 9780080890364. URL: <https://books.google.com.py/books?id=bDtLM8CODsQC>.

¹El criterio para la selección de los atributos a utilizar se decide con `-S` (`attributeSelectionMethod`)

²Si activamos la opción `-G` (GUI) podremos modificar la topología de la red.