

Rezolvarea problemelor cu ajutorul metodelor de învățare



Obiective

Dezvoltarea sistemelor care învață singure. Algoritmi de învățare. Specificarea, proiectarea și implementarea sistemelor care învață singure cum să rezolve probleme de regresie.



Aspecte teoretice

Proiectarea și dezvoltarea sistemelor care învață singure.

Algoritmi de învățare de tipul:

- metoda celor mai mici pătrate
- *stochastic gradient descent*
- algoritmi evolutivi



Probleme abordate

1. Scurta prezentare a problemei
 - a. ce se da (input X , output Y , un input x_{nou}), ce se cere (funcția care transformă X în Y : $f(X) = Y$, astfel încât să poată fi calculat $y_{nou} = f(x_{nou})$)
 - b. ce poate fi X ? -->
 - i. o listă de valori numerice (regresie simplă) $X = (x_1)$, $x_1 = (x_{11}, x_{21}, \dots, x_{n1})$, unde n e nr de exemple de antrenare),
 - ii. vector cu mai multe dimensiuni de valori numerice (regresie multiplă): dacă avem 2 dimensiuni: $X = (x_1, x_2)$, $x_1 = (x_{11}, x_{21}, \dots, x_{n1})$, $x_2 = (x_{12}, x_{22}, x_{32}, \dots, x_{n2})$, unde n e nr de exemple de antrenare
 - c. ce poate fi Y ? -->
 - i. o listă de valori numerice (pt un exemplu, trebuie prezis un singur output), $Y = (y_1)$, $y_1 = (y_{11}, y_{21}, \dots, y_{n1})$, unde n e nr de exemple de antrenare),
 - ii. vector cu mai multe dimensiuni de valori numerice: dacă avem 3 dimensiuni: $Y = (y_1, y_2, y_3)$, $y_1 = (y_{11}, y_{21}, \dots, y_{n1})$, $y_2 = (y_{12}, y_{22}, y_{32}, \dots, y_{n2})$, $y_3 = (y_{13}, y_{23}, \dots, y_{n3})$, unde n e nr de exemple de antrenare (pt un exemplu, trebuie prezise mai multe (3) output-uri)
2. Metode de identificare a funcției f pt cazul în care f este funcție liniară, $X = (x_i)_{i=1,n}$, $x_i = (x_{i,1})$ - un exemplu are un singur atribut, $Y = (y_i)_{i=1,n}$, $y_i = (y_{i,1})$ - un exemplu are un singur output numeric
 - a. metoda celor mai mici pătrate (least square root)
 - b. gradient descent
 - c. algoritmi evolutivi
3. Exemplu de problema

Enunț

Se cunosc următoarele n ($n = 5$) informații aferente unei anumite perioade de timp: numărul de ore înșorite dintr-o zi și numărul de beri consumate pe o terasă.

Nr exemplu	Nr ore înșorite (X)	Nr beri (Y)
$i = 1$	2	4
$i = 2$	3	5
$i = 3$	5	7

i = 4	7	10
i = 5	9	15

Să se aproximeze (folosind un model liniar) câte beri se vor consuma într-o zi cu 8 ore însorite.

Rezolvare:

Se identifică dreapta $Y = aX + b$ (trebuie calculați coeficienții a și b)

a. Metoda least square root - identificare (exacta) a coeficienților a și b

1. Pentru fiecare cuplu (x, y) se calculează x^2 și xy .

Nr exemplu	Nr ore însorite (X)	Nr beri (Y)	x^2	xy
i = 1	2	4	4	8
i = 2	3	5	9	15
i = 3	5	7	25	35
i = 4	7	10	49	70
i = 5	9	15	81	135

2. Se însumează, pe rând, valorile fiecărei coloane

Nr exemplu	Nr ore însorite (X)	Nr beri (Y)	x^2	xy
i = 1	2	4	4	8
i = 2	3	5	9	15
i = 3	5	7	25	35
i = 4	7	10	49	70
i = 5	9	15	81	135
	$\Sigma=26$	$\Sigma=41$	$\Sigma=168$	$\Sigma=263$

3. Se calculează a și b

$$a = (n\sum xy - \sum x \sum y) / (n(\sum x^2) - (\sum x)^2)$$

$$a = (5 * 263 - 26 * 41) / (5 * 168 - 26 * 26)$$

$$a = 1.51$$

$$b = (\sum y - a(\sum x)) / n$$

$$b = (41 - 1.51 * 26) / 5$$

$$b = 0.3$$

deci funcția de aproximare va fi $y = 1.51 * x + 0.3$

4. predicția consumului de bere pentru ziua cu 8 ore însorite va fi:

$$1.51 * 8 + 0.3 = 12.38$$

b. Metoda SGD

Modelarea coeficienților (a, b) ai drepte de regresie:

- la iterația 0: valori random (sau 0) pt a și b

- la iterația $t + 1$ ($t = 0, 1, 2, \dots$)

$$a(t+1) = a(t) - \text{learning_rate} * \text{error}(t) * x(t)$$

$$b(t+1) = b(t) - \text{learning_rate} * \text{error}(t)$$

unde $\text{error}(t) = \text{computed} - \text{realOutput} = a(t) * x + b(t) - \text{realOutput}$, iar learning_rate e parametrul al SGD

Discutie: asemanari si diferente intre Stochastic GD (erorarea se calculeaza pentru un singur exemplu din setul de antrenare) si Batch GD (erorarea se calculeaza pentru mai multe exemple din setul de antrenare)

c. Metoda bazata pe Algoritmi Evolutivi

Cromozom

- reprezentare: cromozom = (a,b), a,b sunt nr reale

- fitness: patratul erorii calculată pe tot setul de antrenare