

Laborator 6

Drd. Limboi Sergiu

Agenda

- Probleme de regresie- Metoda gradientului descrescator
- Probleme de clasificare-Regresie logistica

Metoda gradientului descrescator (gradient descent) pentru probleme de regresie

- Modelarea coeficientilor a si b ai dreptei de regresie astfel incat sa minimizam suma erorilor patratice
- Iteratia 1: valori random sau 0 pt a si b
- Iteratia $t+1$ ($t=0,1,\dots$):
$$a(t+1)=a(t)-\text{learningRate}*\text{error}(t)*x(t)$$
$$b(t+1)=b(t)-\text{learningRate}*\text{error}(t)$$

Unde $\text{error}(t)=\text{computed-realOutput}$

Exemplu-regresie

x	y
1	1
2	3
4	3
3	2
5	5

$$y=ax+b$$

Iteratia 1: $a=0$ $b=0 \rightarrow y=0*x+0=0$

Error = predicted (i)-real(i)= $0-1=-1$, pentru ca $real(1)=f(1)=1$

Iteratia 2: $learningRate=0.01$

$$a(t+1)=0-0.01*(-1)*1=0.01$$

$$b(t+1)=0-0.01*(-1)=0.01$$

$$y=0.01x+0.01$$

Tipuri de gradient descent

- Stochastic Gradient descent = eroarea se calculeaza pe un singur exemplu
- Batch Gradient descent = eroarea se calculeaza pe mai multe exemple din setul de training

Clasificarea

- Rezultatul este categorial sau discret
- Input X , output Y , input X_{nou} si se cere functia care transforma pe X in Y : $f(X)=Y$
a.i. sa poata fi calculate $Y_{\text{nou}}=f(X_{\text{nou}})$
- Clasificare binara \rightarrow prezicerea unui output binar (1/0, yes/no, true/false)
- Algoritmul stochastic gradient descent (regresie logistica) \rightarrow tip special de regresie liniara unde rezultatul este categorial
- $f(x)=a_0+a_1x_1+a_2x_2+\dots+a_nx_n$
- Scop: gasirea coeficientilor (a_0, a_1, \dots) care maximizeaza probabilitatea clasificarii corecte

Gradient descent pt clasificare

- Functie logistica $f(x) = \frac{1}{1+e^{-x}}$ (functie sigmoid) -> scop: inputurile sunt transformate in $[0,1]$

X	Xtransformat
-5	0,00669
-4	0,0179
0	0,5
3	0,9525

Modelul de regresie logistica prezice probabilitatea ca un anumit input sa apartina clasei pozitive.

Daca Predictie > 0.5 -> outputul este predictie a clasei 0, unde threshold = 0.5
Predictie < 0.5 -> outputul este predictie a clasei 1

Pt clasificarea cu mai multe clase se pot fixa mai multe praguri-ex.
Threshold = nrClase - 1.

Stochastic gradient descent

- Estimarea coeficientilor
- Pt fiecare instanta din setul de training:
 - Calcularea predictiei folosind valorile curente ale coeficientilor
 - Calcularea noilor coeficienti pe baza erorii de predictie

Se repeta pana cand modelul este suficient de precis.

Exemplu

x1	x2	y
2.781083	2.5505	0
6		
1.456	2.363	0
3.396	4.4002	0
1.388	1.85	0

Calcul predictie

$$y = a_0 + a_1x_1 + a_2x_2$$

$$\text{Pas1: } a_0 = 0, a_1 = 0, a_2 = 0$$

$$X_1 = 2.781$$

$$X_2 = 2.5505$$

$$Y = 0$$

$$\text{Predictia} = \frac{1}{1 + e^{-(a_0 + a_1x_1 + a_2x_2)}} = \frac{1}{1 + e^{-(0.0 + 0.0 \cdot 2.781 + 0.0 \cdot 2.5505)}} = 0.5$$

Calcul noi coeficienti

$$a(t) = a(t-1) - \text{learningRate} \cdot \text{error}(t) \cdot x(t)$$

$$a_0 = 0.0 - 0.3(0.5 - 0.0) \cdot 1$$

$$a_1 = 0.0 - 0.3(0.5 - 0.0) \cdot 2.781$$

-
- O iteratie prin setul de date = o epoca
 - Dupa 10 epoci -> de exemplu $a_0 = -0.406605$ $a_1 = 0.85257$ $a_2 = -1.1047$
 - Realizarea predictiei
 - If(output < 0.5) => clasa 0
 - Else -> clasa 1
 - Folosind coeficientii calculam ecuatia: $-0.406 + 0.85 * 2.78 + 2.55 * (-1.104) = -0.853$
 - Folosim functia sigmoid si avem outputul: $predictia = \frac{1}{1 + e^{0.853}}$