

96412248

ياسين عيسى

Hw 2

①

$$w_1 = 1, w_2 = 1, b = 1, \text{Sigmoid}(2) = \frac{1}{1 + e^{-2}}$$

$$\text{log loss} = -\frac{1}{N} \sum_{i=1}^N (y \log(\hat{y}) + (1-y) \log(1-\hat{y}))$$

$$\alpha = 0.05, \text{batch size} = 2$$

$$Z = w_1 x_1 + w_2 x_2 + b$$

$$\hat{y} = \text{Sigmoid}(Z) = \frac{1}{1 + e^{-(w_1 x_1 + w_2 x_2 + b)}}$$

$$\frac{\partial L}{\partial w_1} = \frac{\partial L}{\partial \hat{y}} \frac{\partial \hat{y}}{\partial Z} \frac{\partial Z}{\partial w_1}$$

$$\frac{\partial L}{\partial w_2} = \frac{\partial L}{\partial \hat{y}} \frac{\partial \hat{y}}{\partial Z} \frac{\partial Z}{\partial w_2}$$

$$\frac{\partial z}{\partial w_1} = u_1, \quad \frac{\partial z}{\partial w_2} = u_2$$

$$\frac{\partial \hat{y}}{\partial z} = \frac{d}{dz} \left( \frac{1}{1 + e^{-z}} \right) = s(z)(1 - s(z))$$

$$\frac{\partial L}{\partial \hat{y}} = \frac{d}{d\hat{y}} \left( -\frac{1}{N} \left( \sum y \log \hat{y} + \sum (1-y) \log (1-\hat{y}) \right) \right)$$

$$= \frac{1}{d\hat{y}} \cdot \frac{1}{N} \left( \sum \frac{y}{\hat{y}} + \sum \frac{(1-y)}{(1-\hat{y})} \right)$$

$$= -\frac{1}{N} \sum \left( \frac{y}{\hat{y}} + \frac{(1-y)}{(1-\hat{y})} \right)$$

$$\frac{\partial L}{\partial w_1} = \left( -\frac{y}{\hat{y}} + \frac{1-y}{1-\hat{y}} \right) (s(z)(1-s(z)) u_1$$

$$\frac{\partial L}{\partial w_2} = \left( -\frac{y}{\hat{y}} + \frac{1-y}{1-\hat{y}} \right) (s(z)(1-s(z)) u_2$$

$$\Rightarrow s(z) = \hat{y}$$

$$= \left( \frac{y}{\hat{y}} + \frac{1-y}{1-\hat{y}} \right) (\hat{y})(1-\hat{y})$$

$$= -y(1-\hat{y}) + (1-y)\hat{y}$$

$$= -y + y\hat{y} + \hat{y} - y\hat{y}$$

$$= \hat{y} - y$$

$$\Rightarrow \left\{ \begin{array}{l} \frac{\partial L}{\partial w_1} = \kappa_1 (\hat{y} - y) \\ \frac{\partial L}{\partial w_2} = \kappa_2 (\hat{y} - y) \end{array} \right.$$

First batch

$$\left\{ \begin{array}{ll} \kappa_1 = 22, \kappa_2 = 1 & y_1 = 0 \\ \kappa_1 = 25, \kappa_2 = 0 & y_2 = 0 \end{array} \right.$$

$$\Rightarrow z = w_1 \kappa_1 + w_2 \kappa_2 + b$$

$$\Rightarrow \left\{ \begin{array}{l} z_1 = 22 + 1 + 1 = 24 \\ z_2 = 25 + 1 = 26 \end{array} \right.$$

$$\hat{y} = \sigma(24) \approx 0.99$$

$$\sigma(26) \approx 0.99$$

$$\text{loss} \rightarrow \frac{\partial L}{\partial w_1} \left\{ \begin{array}{l} 22(0.99) = 21.78 \\ 25(0.99) = 24.75 \end{array} \right. \frac{\partial L}{\partial w_2} \left\{ \begin{array}{l} 1(0.99) = 0.99 \\ 0(0.99) = 0 \end{array} \right.$$



$$w_1 = w_1 - \alpha \frac{\partial L}{\partial w_1} = w_1 - 0.05 \left( \frac{21.78 + 24.75}{2} \right) \\ 1.16325$$

$$w_1 = -0.16325$$

$$w_2 = w_2 - \alpha \frac{\partial L}{\partial w_2} = 1 - 0.05 \left( \frac{0.99}{2} \right) = 0.975 \\ 0.975$$

batch 2, itr 1

$$\begin{cases} x_1 = 47, x_2 = 1, y_1 = 1 \\ x_1 = 52, x_2 = 0, y_2 = 0 \end{cases}$$

$$z_{ij} = \begin{cases} 47(-0.16325) + 1(0.975) + 1 = -5.6978 \\ 52(-0.16325) + 0(0.975) + 1 = -7.487 \end{cases}$$

$$\text{Sigmoid}(z) = \begin{cases} \textcircled{1} \approx 0.003 \\ \textcircled{2} \approx 0.00055 \end{cases}$$

$$\frac{\partial L}{\partial w_1} = \begin{cases} (0.003 - 1)47 \\ (0.00055 - 0)52 \end{cases} \quad \frac{\partial L}{\partial w_2} = \begin{cases} 1(0.003 - 1) \\ 0(0.00055 - 0) \end{cases}$$

$$w_1 = -0.16325 - 0.005 \left( \frac{-46.917 + 0.0282}{2} \right)$$

$$w_2 = -0.0459$$

$$w_2 = 0.975 - 0.005 \star 0.977$$

CHITRA

$$\text{batch 3, itr 1} \quad \left\{ \begin{array}{l} x_1 = 46 \quad x_2 = 1 \quad y_1 = 1 \\ x_1 = 56 \quad x_2 = 1 \quad y_2 = 1 \end{array} \right.$$

$$w_1 = 0.1056, w_2 = 0.979$$

$$\text{batch 4, itr 1} \quad \left\{ \begin{array}{l} x_1 = 55 \quad x_2 = 0 \quad y_1 = 0 \\ x_1 = 60 \quad x_2 = 0 \quad y_2 = 1 \end{array} \right.$$

$$w_1 = -0.0316, w_2 = 0.979$$

$$\text{batch 1, itr 2} \quad \left\{ \begin{array}{l} w_1 = -0.109 \\ w_2 = 0.978 \end{array} \right.$$

$$\text{batch 2, itr 2} \quad \left\{ \begin{array}{l} w_1 \approx 0.002 \\ w_2 \approx 0.98 \end{array} \right.$$

$$\text{batch 3, itr 2} \quad \left\{ \begin{array}{l} w_1 \approx 0.030 \\ w_2 \approx 0.98 \end{array} \right.$$

$$\text{batch 4, itr 2} \quad \left\{ \begin{array}{l} w_1 \approx -0.089 \\ w_2 \approx 0.98 \end{array} \right.$$

$$\text{test } x_1 = 46, x_2 = 1 \Rightarrow y_{\text{pred}} = 0.10$$

این را با حاصل امتحان خطای بایستی مقایسه کرد

باقی batch با دستورات عینش را به دستورات  $Hw2.ipynb$  پیوسته می باشد

قسمت ب)

تلاش به Stochastic به GP بنظر قرارداد

Step-Size و Learning Rate در یک ساختار آزمون و خطا

در یک ساختار دیگر نویسنده (اسلاید 25 از Lesson 4)

(ب) این مسئله داریم با Hyperparameter ها تنظیم کردن

نقشه (1)

راه حل: اگر در یک ساختار یاد بگیرد

Step-size کاهش باید و اگر کم بود Step-size افزایش باید

⑤ برای GD به جای ~~چون~~ متوجه شدیم و متوجه می‌شویم

ضرورت می‌کند در این نقاط گیریم و دیگر وزن ها به دست نرود

از آنجا که SGD به جای batch ما داریم عمل می‌کند است برای

batch انتخاب به روش دیگری batch دیگر



سوال ۳ :

در Logistic Regression توانستیم از تقسیم لایین در کلاس واقعی

نتیجه‌ها در Linear Regression که داده‌ها

نزدیک هم بود در یک کلاس قرار داد.

در اینجا ما به نیاز Activation Function و درک می‌کنیم Activation Function

نشان می‌دهد که ما می‌توانیم نتایج خروجی داده‌ها

بسیار ساده هستند و می‌توانیم به سادگی نتایج

تجمع داده‌ها خروجی تقسیم کنیم

به کدهای مربوطه به از #92 قرار دارند

مجموعه داده های Iris ۳۰۰۰  
۴ ویژگی

طول و عرض کاسبرگ و لب برگ ها و طول زنبق با گل زعفرانی

است به ۳ نوع از این گل وجود دارد و به خاطر مصارف تجاری

چون تفاوت قیمت هر نوع از این گل ها خیلی زیاد است از این

داده ها برای جداسازی این دسته ها (فصل بندی) استفاده می شود

توضیح  
بعد از جداسازی داده ها به داده های تست و آموزش

داده های ورودی را به وسیله Standard Scaler نرمال سازی می کنیم

تفاوتی عددی در میان Logistic Regression رخ ندهد

در نهایت حل به روش آموزش داریم و به وقت

82 درصد در آزمون به کار حل با به ویژگی ها

آموزش ما داریم به وقت 100 درصد می بینیم



وجود ویژگی های کافی در نگاره برای یک دل جنگ است

بیاد تأثیر دانه است