

۱) توابع فنی‌السازی نظریه عی پندلیه (a)

ایجاد مرده این تابع غریبی بودن را به سبب اضافه مرده و

استار (یارگی) آنکه بسیار را به سببی (هدف آن

تابع فنی‌السازی) نداشته و خوبی دارد این متناسب با وروی لایه بعید است

می‌شد \rightarrow در مارکت محل خالی بود و توانایی شستی اردی بیکده را

- زانو -

۶) خیر برای انتداب دیگر تابع فنی‌السازی با مواردی

شامل شسته نمایی \rightarrow سوپرایز جلوگیری

از اسباب (جلوگیری از توقف محل با وزن عی نانلب) رادرنگر چون

هر تابع فنی‌السازی غریبی است و برای رانن اشتباه باشدی تواند
خشن

باعث شکل ای که مکاری شود که مکاری را خراب

شد. صحیح این سرمه بلفع مسلسل مایز است دارد.

۷) مرادا:

۱- مردی لایه پیر باشد و روابط پیشیده و دفعی تری راسته ی ترازن

پادری مرد.

۲- عی بیشتری تراست و بیشتری عی مختلف را به صورت سلسه رایج

نایس (عنز کاری) - $D_N N$ - اجتماعی (عنز ب ای نونهارت).

میابی:

۱- عی توانی بای - Vanishing gradient $\nabla_{\theta} L$ و وزن θ در

۲- عی ابتدایی نشوند یا جنی جنی θ update شوند

دستگذ

۱- تواندیاری را نزولی موقت نمایم

۲- بینی که اینجا شرکت و اخراجات را کنمایم

از اصولی ترین و سادگی قدری backpropagation feed forward

train

MLP میتواند طولانی ترین شرکت و دلیل دیرگیری شود

۳- داده‌ی بسیار بسیار بسیار باشد و حواسی تواند

باشد و بجلی تواند اتفاقی نماید Overfitting

جزئیات - و نویزی داده آموزش را برآورده و علاوه بر احتفاظ

رسانی به داده‌ی زیاد هم جالش دارد

تابع d : Sigmoid

$$P(x) = \frac{1}{1 + e^{-x}}$$

دستگاه

هزایا: خروجی را بیازهی (۱ و ۰) فنرمه داریم

باعث شده بسته دید است (استالان) به سهل داشت باشد و برای سادگی

از جن خروجی از جن خروجی binary classification

$f'(x) = f(x)(1-f(x))$ مسئله آن از روی خوش یعنی

$y = \alpha$

میابیم: صفر نباید وای باید همراهی نداشته در

$f'(x)$ شرط آن به حریم training rule

ذیه یعنی بین صفر و ۱ باید شده باشند

ورودی بسی بزرگ باشند و این حفظ شده و مامون update

وزیری و در نزدیکی باید شده (مانندیاری)

gradient

درست شده

$$P(x) = \frac{e^x - e^{-x}}{e^x + e^{-x}}$$

: tanh پاک

مزایا: صفر کر تر بوده و خوبی بی دارد \rightarrow بجزء آغاز

ترنر (Binary classification) میگیرد. برای مسائلی که دو دسته دارند خوبی داشته باشند.

برای خوبی داشت این معنی است آنکه ایجاد شوندگان

$$f'(x) = (1-f(x))(1+f(x))$$

~ در امداد نکات صفر \rightarrow تراویل سیری داشته وی تواند

باشد \rightarrow sigmoid \rightarrow منحنی converges شود.

$$\lim_{x \rightarrow -\infty} f'(x) = 0 \quad \text{و} \quad \lim_{x \rightarrow +\infty} f'(x) = 0$$

مواردی هستند که دارند

و اگر مقادیر ورودی کی بزرگ باشد تراویل آن صفر شود.

نشان

حصَّ درایی سوئیپ بدرآز (Sigmoid) بعد از بازگشته از به

وزرسانی و وزن جلوگیری شدنه خود بین:

$$P(x) = \max(0, x) = \begin{cases} x & x \geq 0 \\ 0 & x < 0 \end{cases} : \text{ReLU تابع}$$

هزایا: مسئله اشباع در پیشترانه از مرا که $P(x)$

برای $x > 0$ درست است. برخلاف آنها بقیه درون

مقادیر کمی بزرگ ورودی باشند جلوگیری از به وزن و

بنی شوند. ساره بعدن بسیار بسیار بقیه و حاسوس نیزه نیزه

نیزه.

معایب: به طور نهی مشخص نیزه هستند که در نقطه صفر

مشخص وجود ندارند (نمایمایه) $\lim_{x \rightarrow 0^-} f'(x) \neq \lim_{x \rightarrow 0^+} f'(x)$

مشخص

بازایی صیغه وردی یعنی متفق با برصری یاد

بروزساختن وزیر شرط

$$P(z_1, z_2, \dots, z_k) = \frac{e^{z_i}}{\sum_{j=1}^k e^{z_j}} \quad ; \text{Soft max تابع}$$

$$= P(y=i | X)$$

مزایا برای مسائل چند کلasse بیمارستان است

زیرا دارای یکدیگر احتال ارزش داره مرتبط به کلاس اول است

حقدار جوابت max احتال است

محابی بمحاسبات بیشتر نیاز داشت به بقیه و صیغه

نیاز داشت استاد عذر از آن دلایلی خصی داشت

مشترک

اپٹر ا مکروہ وی خروجی شبکہ MLP را کیسے کوں کوں

(۲)

و داریم:

$$f_1(z) = \text{ReLU}(z) = \max(0, z) \rightarrow f_1(z) \geq 0$$

$$f_x(z) = \text{Sigmoid}(f_1(z)) \rightarrow f_x(z) \geq 0$$

$$\text{Sigmoid}(f_1(z)) \geq 0,1 \rightarrow f_x(z) = y \geq 0,1$$

حال آنکہ آستانہ ۰،۱ را برای طبقہ بندر نزدیکی میں بھون

$y \geq 0,1$ - در نزدیکی خروجی طبقہ بنڈی برای هر داد و ورودی

ایم خوب سنت جیا - اصلانی توان مالی

(۱۰۴) سربو طب بے کالاس صفر را درس - پس بنی نہ

loss برای SE از درایو سوال از دقت نیز در مورد سئون ۳

$$L = \frac{1}{\tau} [(g_1^1 + g_1)^2 + (g_r^1 + g_r)^2] \quad \text{استاده کرد سئون}$$

ابیرا برای داده ورودی (۰ و ۱) \rightarrow (۱ و ۰) کاسیات را انجام

داده ورداری :

: Feedforward مدخل

ایسا ورودی وزن دار نورون هایی را حساب کرد و داریم

$$\begin{bmatrix} x_1^{(1)} & x_r^{(1)} \end{bmatrix} \begin{bmatrix} w_{11}^{(1)} & w_{r1}^{(1)} \\ w_{12}^{(1)} & w_{r2}^{(1)} \end{bmatrix} = \begin{bmatrix} x^{(1)} \\ w^{(1)} \end{bmatrix}$$

$$\begin{bmatrix} w_{11} & 1 \end{bmatrix} \begin{bmatrix} 9 & -1 \\ -1 & \omega \end{bmatrix} = \begin{bmatrix} 14 & -\varepsilon \end{bmatrix}$$

$$Z^{(1)} = \begin{bmatrix} z_1^{(1)} & z_r^{(1)} \end{bmatrix} = \begin{bmatrix} 14 & -\varepsilon \end{bmatrix}$$

جستجو

برابر درودی یعنی وزن دار نورول یعنی لایه اتم است.

برابر درودی یعنی وزن دار نورول (روت نسیم) $\approx h_2, h_1$

حال خروجی نورول یعنی لایه پخته را که بسته به آن $\sim 60\text{ dy}$ لایه بعدی

$$Q^{(1)} = \sigma([z_1^{(1)}, z_2^{(1)}]) = \sigma(z^{(1)}) =$$

$$\sigma([14 - \epsilon]) = [1 \quad 0,02] \xrightarrow{\text{یا ارقام اکثراً تخمین زده شده}} \text{یا ارقام اکثراً تخمین زده شده}$$

و درودی وزن دار نورول یعنی خروجی را حساب می‌کنیم

$$Z^{(1)} = X^{(1)} W^{(1)} = [1 \quad 0,02] \begin{bmatrix} 1 & -2 \\ \frac{1}{\epsilon} & 2 \end{bmatrix} =$$

$$[1,000 \quad -1,99]$$

حال خروجی نزدیکی اور را بود - \hat{y}_i آور - y_i همان

مقادیر برابر \hat{y}_i و y_i باشند $\hat{y}_i = y_i$ مقدار

$$\sigma(z^{(l)}) = \hat{Y} = [g_1^1, g_r^1] = [e^{-\lambda} / (1 + e^{-\lambda})] \quad l = 1, 2, \dots, L$$

درایی مثال همانطور که بینشید احتمال آن دستا متعلق به کلاس 1،
 t_1

باشد λ_1 و λ_2 متعلق به t_1 . در جزو \hat{Y} احتمال بالاتر باشد

و باشد λ_1 او در غیر ای λ_2 صورت - صفر. بنی فنا در داریں

بنی

را اس بود، $\frac{\partial L}{\partial \hat{y}^1}$ ابتدا ما ترسیم $\frac{\partial L}{\partial \hat{y}^1}$ backpropagation را در

$$g = \frac{\partial L}{\partial \hat{y}^1} = \left[\frac{\partial L}{\partial \hat{y}_1^1}, \frac{\partial L}{\partial \hat{y}_r^1} \right] \quad \text{و داریم}$$

$$g = \left[-(y_1 - \hat{y}_1^1) \quad -(y_r - \hat{y}_r^1) \right]$$

جذب

حال ماتریس و نتایج را محاسب کنید

$$g = \begin{bmatrix} g_1^1 - y_1 & g_1^1 - y_2 \\ g_2^1 - y_1 & g_2^1 - y_2 \end{bmatrix} = \begin{bmatrix} -0.2 & 0.12 \\ 0.12 & -0.2 \end{bmatrix}$$

حال برای این بتوانیم را محاسب کنیم

$\frac{\partial L}{\partial w^{(1)}}$ را محاسب کنیم:

$$g \leftarrow g \odot \frac{\partial y^1}{\partial z^{(1)}} = \left[\frac{\partial L}{\partial g_1^1} \times \frac{\partial g_1^1}{\partial z_1^{(1)}} \quad \frac{\partial L}{\partial g_2^1} \times \frac{\partial g_2^1}{\partial z_2^{(1)}} \right]$$

$$\rightarrow g = [-0.2 \times \sigma(z_1^{(1)}) \times (1 - \sigma(z_1^{(1)})) \quad 0.12 \times \sigma(z_2^{(1)}) \times (1 - \sigma(z_2^{(1)}))]$$

$$g = \begin{bmatrix} -0.1 & 0 \\ 0 & 0.1 \end{bmatrix}$$

Feed Forward - زیرا در عکس $z_1^{(1)}$ و $z_2^{(1)}$ قبل از شرک

آنچه دادیم. حال برای حسابی $\frac{\partial L}{\partial w^{(1)}}$ باید

$$\frac{\partial L}{\partial w^{(1)}} = O^{(1)T} \times g$$

جستجو

خروجی نورون (عصب) ای و درود (۱)

$$\frac{\partial L}{\partial w^{(r)}} = \begin{bmatrix} O_1^{(r)} \\ O_r^{(r)} \end{bmatrix} \times \left[\frac{\partial L}{\partial z_1^{(r)}} \quad \frac{\partial L}{\partial z_r^{(r)}} \right] =$$

$$\begin{bmatrix} 1 \\ 0/0 \end{bmatrix} \times \begin{bmatrix} -0/0 & 0/0 \\ 0/0 & 0/0 \end{bmatrix} = \begin{bmatrix} -0/0 & 0/0 \\ -0/0 & 0/0 \end{bmatrix}$$

حال تاریخ سے باہر بہلیں چلیں یا ہمارے لایہ دکت برهیں

رادر وی ریزیم و داریم : $f \in \mathcal{G} \times W^{(r), T}$

$$g = \left[\frac{\partial L}{\partial z_1^{(r)}} \times w_{11}^{(r)} + \frac{\partial L}{\partial z_p^{(r)}} \times w_{p1}^{(r)} \right] \frac{\partial L}{\partial z_1^{(r)}} \times w_{1r}^{(r)} + \frac{\partial L}{\partial z_p^{(r)}} \times w_{pr}^{(r)}$$

$$\rightarrow g = \begin{bmatrix} 0, \omega \\ 0, 1 \end{bmatrix} \times \begin{bmatrix} 1 & \frac{1}{\epsilon} \\ -\nu & \nu \end{bmatrix} =$$

$$[-0.1^{\circ}V \quad 0.1^{\circ}V(\omega)]$$

دشنه

حال برای یافتن $\frac{\partial L}{\partial w^{(1)}}$ باید کو را طبق رابطه زیر حساب

$$g = g \circ \left(\frac{\partial O^{(1)}}{\partial z^{(1)}} \right)$$

$$[g_1, g_2] \circ \left[\frac{\partial O_i^{(1)}}{\partial z_i^{(1)}} \quad \frac{\partial O_r^{(1)}}{\partial z_r^{(1)}} \right] =$$

$$\left[-0.18 \quad 0.001 \right] \circ [O'(14) \quad O'(-1)]$$

$$= \begin{bmatrix} 0 & 0.001 \end{bmatrix} \approx \begin{bmatrix} 0 & 0 \end{bmatrix}$$

حال برای یافتن $\frac{\partial L}{\partial w^{(1)}}$ داریم

$$\frac{\partial L}{\partial w^{(1)}} = X^{(1)\top} \times g \quad g = \begin{bmatrix} 0 & 0 \end{bmatrix} \quad \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}$$

ماتریس $X^{(1)}$ سینه وزیری (4×2) می‌باشد

$$W^{(2)} \leftarrow W^{(2)} - \alpha \frac{\partial L}{\partial w^{(2)}}$$

برابر است با

شمش

حُرْضِي سِيَرْ بِراَبِرَ اَرْهَمْ وَدَارِي :

$$W^{(2)} = \begin{bmatrix} 1 & -1 \\ \frac{1}{\epsilon} & 2 \end{bmatrix} + (-0.1) \begin{bmatrix} -0.05 & 0.1 \\ -0.100 & 0.1000 \end{bmatrix}$$

$$= \begin{bmatrix} 1.000 & -0.100 \\ 0.1200 & 1.9999 \end{bmatrix}$$

صَارِخُورَهُ بِينِدِ تَرَادِ بِسَارِجْيَهُ اَرْهَمْ

حَالِ بِراَبِي دَادِي بِدرِي صَرْفَاً كَاسِيَادِ رَالِبَاهِ دَادِه

$$\text{ودَارِي : } (-1, 0) \rightarrow (0, 0)$$

اَبِرِي بِراَبِي دَادِه وَنَدِي بِراَبِرَ اَرْهَمْ دَادِه feed forward

$$Y^1 = [y_1^1 \quad y_2^1] = [0.05 \quad 0.11]$$

$$\underline{Z}^1 = [-1.5, 1.5] \quad Z^1 = [0.15 \quad 1.99]$$

$$O^1 = [0.17]$$

$\frac{\partial L}{\partial y^1} = 1, \frac{\partial L}{\partial y^1} |_{\mu=1}$ backprop جا بک پرپ

$$\frac{\partial L}{\partial y^1} = [0, 0.4 \quad -0.12]$$

$\frac{\partial L}{\partial w^{(1)}} =$ حال تعداد وزن عیّن برابر است با

$$\frac{\partial L}{\partial w^{(1)}} = [\begin{matrix} 0 & 0 \\ 0 & 0 \end{matrix}]$$

$$\frac{\partial L}{\partial w^{(1)}} = [\begin{matrix} 0 & 0 \\ 0.13V & -0.1012 \end{matrix}]$$

$\frac{\partial L}{\partial w^{(1)}} =$ نتیجه نتایج وزن برابر است با

$$w^{(1)} = [\begin{matrix} 9 & -1 \\ -2 & \omega \end{matrix}]$$

$$w^{(1)} = [\begin{matrix} 1,00\omega & -2,001 \\ 0,13VGF & V,001P \end{matrix}]$$

دستگیر

اگر همه وزر (4) را ابتدأ حفظ باشند، چنین ورن.

$\text{درست} \rightarrow \text{لای خروجی} \rightarrow \text{سازن} \rightarrow \text{را تو می‌سین} \rightarrow \text{بن}$

حاصل بین وزر دار برای چنین ورن نیز خروجی

در سار برای چنین ورن (طبیعتاً تابع مقام اعزامی) در هم لای برای چنین ورن

آن لای در سار است. -) وانی باعث شدن - حلوی بن

از یادگیری Feature می‌توان درسته وانی عمل

یک شر نورون کی ریلای صفتی هم محلی شد و حلقی ندارد.

6) خضر. سیبی بہ تابع ضرور نزخ یادگیری کی تو انداز این اتفاق

لیغند. در - ام - ام - صوف کی این الگوریتم مدلینگ و دن

تابع خضراء - آتائی تو انداز دارای نوساناتی شود - صفت ناهن

دستگذ

نیابد. به طورشال فرض کنیم تابع هزینه برابر $J(w) = (w - 1)^2$

است. طبق الگوریتم گرادیان داریم:

$$w^{i+1} \leftarrow w^i - \alpha \times \nabla J(w^i)$$

هزینه کنترل مقدار $\alpha = 1$ و $w_0 = -2$ باز از

گرادیان برابر $-1 = 1 - (-2)$ است. حال داریم:

$$w' \leftarrow w_0 - \nabla J(w_0) = -2 + 4 = 2$$

از این نتیجه نتیجه شده است که در این نقطه هزینه برابر $16 = 2^2$ است.

چنان‌که اگر α را بزرگ تر کنیم بزرگ‌ترین مقدار گرادیان نیزی شود.

و زمانی که مدل نتواند Underfitting (C)

آرگوی وجود در داده‌ای آموزش را می‌داند و در نتیجه

دستگاه

دعاً (و) (ادع) آنرا زدنی و نه باندی باش.

راهنمایی: افراد اسکن باید در مدل آسنا (آزار) به عنوان پرداز و نهاد را داشته باشند.

بِسْمِ اللَّهِ الرَّحْمَنِ الرَّحِيمِ

ویرجیئی بئر : افزوون ویرجیئی تولۇز باىن شۇدىل بىرماڭىر . ھىنى

• دیاں بیٹھ رہے تواندھ سو (دل سو) سو

Oret Pitting

زنان رخی رهه سرل بخوبی روی داده ای آموزش حمل و بار

روی دار خود نیز باشد بتوانی

توانم نت دا(۴۰) آمرزش را خطا ده و به جزوئیاند - و فریزدانه

۴۵ آموزشی حساسیت

دشنه

اھٰر، ۴

سالوسازی مدل ماسّ حزف لایه یا نورون‌های اضافی.

کلما \rightarrow regularization \rightarrow باعث شدن تبدیل کوی سازی

برای راشه باشد در درگیر Dropout

در طی خرآند آموزش به طور تصادفی بخشی از نورون‌های این مدل دور می‌شوند.

نتیجه \rightarrow شده نسبت به وابستگی زیاد به نورون‌های خاص جدا شده و

کوی سازی برای راشه باشد. صنیعی توان ارزیابی cross validation نیز استفاده شده است. در مرحله داده را به کمپیوچر لغیم و دو هزار داده محاسبه عذران داده شده است. اگر به فایل زیر برای داده دسترسی داشته باشد روی داده کلیک کنید.

در آخرین Epoch برا بر ۹۸٪، ۹۷٪ برا برای داده آموزشی.

صنیع نورولار loss برا برای هم ۲ تعریف نزولی است و نزد درس بهم.

(رسانه) وقت روی داده آموزشی بالا است -
درخت

آنالیز روی دارهای ترتیبی خود را - در نظر داشت

دلیل Fit

Q5

کرد این سوال در فایل Q5.ipynb موجود است (1)

لرچیسی مینم این سوال را با استفاده از google colab

اجرا نمایم / سئوچن در بارگذاری دستگاهی

نیاز است باشد. همین نتایج حامل در عصبی وجود در پوسته ۰.۵

قابل دیر است - همانطور ملاحظه شد وقتی دل روی

داره آوزشی برابر ۸۷.۴% است - در این مسئلہ از

۱۲۰ تغییر با توابع فعالسازی ReLu و تعداد نورون ۳۰

مشترک

بر تردید استاده در دارد. همین لایه خروجی ها را

بر اساس نتایج آن دارند (اندروید بدلیل وجود آن مالس است) (دارد)

تابع فعال سازی Softmax (فوت نیز)

برای آن داده های صفر تا ۱ باشد هر پرسش آن را

categorical-crossentropy نام نهاده. تابع خروجی هم نزدیک استاده

از جمله سعی های سعی هایی که نیز هستند

بر تردید استاده در دارد این تابع فعال سازی با

تابع categorical-crossentropy باعث شود

برای نتایج حاصل از Softmax بروزگیری شود و برابر

برای نتایج خروجی از مقدار y_k^* بعنوان y_k و مقدار y_k بعنوان y_k' داشته باشیم

سازمانی را دارا نشاند) یعنی تابع هزینه این را برآورده

کرده. این تعداد لایه ای است - آورده بنت به این سیر یا

کسر از این باشد بتر کل کوچک \downarrow مرا که اگر لایه را بتر کنیم وقتی

اماً بتر خامی زیاد نمی کرد هنوز هم بود وقت داده ام اعزامی بتر

کسر اما وقت روی داده شد - تغییر زیاد نمی کند و همچنان بود که

از این هم بسته همین این تعداد نورونی - هم انتقام بده ایم

با خط و آریاشت ساخت انتقام - تعداد لایه ۴ بود آمد. همین

update کوچک \downarrow بین درجه ۱۲۸ ۱، batch size

داده باید تعداد شر و تعداد ۱، Epoch ۱، آورده.

دلیل این که این اگر تعداد ۱، Epoch ۱، بتر نمی شود

"آیا 6 نه شبایه است بسیار داده است شده است"

باعث شدن Validation loss برای نمودار برای تأیید شدن بدراز

با هدف وظیفه loss کاهش داده آورده ای

برای مسئله از بین سازنده از overfitting

در هر رحله صرفاً از گاریان فعل استفاده نماید و به نوعی از

سیاریکی هست گاریان ع استفاده خوده شبایه است در عین

از کلی گردش و سر و میری تر هست راسود در این سوال توابع

متوازنی ReLu در لایه عقی خفی استفاده شده در مقایسه

با سینکوید و tanh از لحاظ وقت بسیار بتر

کل عدد آن خانوادے زیادی ندارد جس سا سے با آن قرائیں ۷

(فَتَحْلُولَ الْمُكَبَّرَةِ وَسَعْيَهُ لِلرَّحْمَةِ) ۖ ۗ ۗ ۗ ۗ ۗ ۗ ۗ ۗ ۗ ۗ ۗ ۗ ۗ

فرموده ای متابعد از اینجا

سیدار خروشیل زیارد باشد مسق آنها صفری شود و کمالاً رجارتی بخشد

نیز دیرگ وزن عسیق تیرین شت از لهذا لغودار خطا تابع سیسوید

لیکراز آن حمل کرد و میرزا = loss دراده اعزمن و سان به

هم نزدیک بوده بسیار و جنی خود کا حصہ ہے یا بنے۔ ٹھانٹلور کے لفڑاری

را بینی های سینه لغزدار دارد و آن را آنچه ای دارد برای داده آموزشی Accuracy و gloss

کمپرسن - آن برای Validation تعریف شده است

نَاتِلِي وَجْه دارنَدَه درَانِ نَزُوكِه بِنَاشِنَه اَحَدَه فَ اَبِي سَدِّوْمَاه دَرِعْنَه

یا مارپیچ برای Validation داده ها را در یک Epoch تقریباً ۶۰٪ و تقریباً ۴۰٪ درست را با آن خواهد داشت.

و تقریباً نزدیک این میزان دقت در داده های (ولیدیشن) است.

بنت ب داده آموزشی (لرنر) در داده آزمونی (تست) دقت در داده آزمونی ۹۰٪ است.

و برای داده آزمایشی (تست) مدل fit کرده باشیم.

سرمه