Logistische Regression G 2 D (7) 2 1/1+e-2 Z z wTx+b Log Coss J(w) = 1/2 ((gin log gin) + (1-gin) log 1-gin)) Eine Beubachtung Schaffen, ob es zu einer bestimmten klanse gehört. L Vorhersage für ein binares Éreignis

Lineare Regression $y/\hat{y}^2 = \omega_j \times + b$ $MSE = \frac{1}{n} \cdot RSS$ $J(y,\hat{y}) = \frac{1}{n} \sum_{i=1}^{n} (y_{ii} - \hat{y}_{ii})^2$ $= \frac{1}{n} \sum_{i=1}^{n} (y_{ii} - (\omega_{x_{ii}} + b))^2$

Zusammenhang zwischen unabhängigen Merlimalen Zu einer abhängigen Voriable.

Verlaufspreis anhand von Anz. Fimmen. Alter Gebande oder Wohnlage bestimmen.

| Gradienten ab 8 | rtieg | |
|-----------------|---|---------|
| wiederhole w:z | w-~ (dg(w)) | |
| - zu g | roß zzz instabiles lernen. überspringen des globalen | Uniones |
| - 7 u h | lein 222 brancht zu lange | |
| Momentum | | |
| Annahme: Die | 2 Richtung des Gradienten ungenheit ist eine gete Vooh | |
| in der Verge | ungenhert ist eine gete Vorh | ersage |
| fir die Zuli | inflige Richtung. | |
| Sdrttn+B. | Sehritt n. 1 + R? Schrift n. 2 | e + 1 . |
| Matrix Com | anations | |
| a. shape | 4.3 3.4 4.3 | |
| b. Shape | 1,3 3,1 1.3 3,1 | |
| a+b | 4,3 3,4 X X | |
| a b | 4.3 3.4 4 | |

4,3

X

a+b a·b np. det (a,b)

 $\frac{1}{\alpha}$

3,4 X

Aletivierungsfuntion On einem Neuron werden die gewichteten wete + Rias Zu einem Ausgebewert transforniert, welcher dam an die nächetgelegende Schicht westertranspotiet wird. gut fir Wahrscheinlichtigten als letzte Abhiviengbru. Ausgabefunkan gubb wet Obis 1 -1: Sigmoid, logistische S(2) 2 /1+e-2 - varishing gradient bithdolen L. figmoid 1 - Die weite von den Melestungen bunen bei Signord sehr blein O ber fast Osein, und eine 0,5 wederholde hulhiphilation dieser bleim wete filmt dazu dassdie Gradiets or worderen Schilt leum altudisiet werden. 2: tent, selse vie Signer d'autre Shifted version, d'entre of signer d'entre veil 0 al Ferne Werke -1 bis 1, bisst besser viel 0 al Ferne 3: ReLU(x) z max (0, mxx), eifant + effiziente truttion -4: Lealing RRUL fix) = {x fix x>0 Typischevise 0,01 - Eealy Rely =5: GELLY Gaussian Error line Unit gut in NLP, aber hoher kechenaufvard

Altivienz flubtion Requiremb - Nicht liker = West ansamm eine Lineure rausbunnt, wint man keine kompleken Daten maddiem - Sterig differsiable = be Echen und tearthe Spitzen und frendstrieg millt - monotone Stegny = vehicht Mehrdertighet, bein vernisrends lernen, die Lennithy wind deuthin.

An Urspung fullin & Espisistr, blone throng 3w. + and - wete, A US Gellefution (Alutin mit der Angabeschild,

1. Sigmaid: binane Klues flytion

2. Softmax? melvhlungen Klups flytion

3. Rell = the nicht so get, weil er hin negamin weter

3. Dell = the nicht so get, weil er hin negamin weter

aussist. Bei hegren an vermodet. idephtato Jultion

Neural Wetwork

[5] le could be activation vector a,

k l z row vector w

j is the number pos of hidden larger

l is the neuron

Z= Bevor Minemys Julian

2: Bevor Minemys Julian

3: 2: W (1) A (1-1) + b (1)

4: 2) z g(e) (Z(1)) himst van der lette Schillet

A (2) z g(e) (Z(1))

inems!

nemers!

NEM 2 (k neurons of Athidden layse, features of n-1th layse)

.



