

Ակտիվացիոն ֆունկցիաներ մաս 2

Հայկ Կարապետյան

1 Ածանցյալ

Մոդելը ուսուցանելու համար արդեն գիտենք, որ օգտագործում ենք գրադիենտային վայրեջք: Գրադիենտային վայրեջքի ընթացքում անհրաժեշտ է լինում հաշվել ֆունկցիայի մասնակի ածանցյալները այդ իսկ պատճառով ներքևում կդիտարկենք տարբեր ակտիվացիոն ֆունկցիաների ածանցյալներ:

$$1. \sigma(x) = \frac{1}{1 + e^{-x}}, \quad \sigma'(x) = \frac{e^{-x}}{(1 + e^{-x})^2} = \frac{1}{1 + e^{-x}} \times \frac{e^{-x}}{1 + e^{-x}} = \sigma(x) \times (1 - \sigma(x))$$

Sigmoid-ի լավ կողմերից է այն, որ իր ածանցյալը հենց արտահայտվում է իրենով: Այսինքն հավելյալ հաշվարկների կարիք չկա, մեզ միայն անհրաժեշտ է հիշողության մեջ պահել $\sigma(x)$ արժեքը և օգտագործել այն անհրաժեշտության դեպքում:

$$2. \tanh(x) = 2\text{sig}(2x) - 1$$

Տանգենս հիպերբոլականը կարող ենք արտահայտել sigmoid-ով, այսինքն իր ածանցյալը նույնպես հնարավոր է արագ հաշվել:

$$3. \text{ReLU}(x) = \max(0, x)$$

ReLU-ն ածանցելի չէ 0 կետում, քանի որ 0 կետում նրա ձախակողմյա ածանցյալը հավասար է 0, իսկ աջակողմյանը հավասար է 1: Քանի որ 0 կետում հայտնվելու հավանականությունը 0 է, մենք կարող ենք անտեսել այդ կետը, իսկ եթե նույնիսկ հայտնվենք այդ կետում վերցնենք ածանցյալի արժեքը 0 կամ 1 (tensorflow գրադարանում արժեքը վերցվում է 0):

$$4. \text{Parametric ReLU: } LR(x) = \begin{cases} ax, & \text{երբ } x < 0 \\ x, & \text{երբ } x \geq 0 \end{cases}$$

Parametric ReLU-ն նույնպես զրո կետում ածանցելի չէ և կարող ենք վերցնել այդ կետում արժեքը a կամ 1:

$$5. \text{Exponential linear unit: } ELU(x) = \begin{cases} \alpha(e^x - 1), & \text{երբ } x < 0 \\ x, & \text{երբ } x \geq 0 \end{cases}$$

Եկեք նայենք ELU-ի ձախակողմյա ածանցյալը 0 կետում:

$$(a(e^x - 1))'(x=0) = a \times e^0 = a$$

Աջակողմյա և ձախակողմյա ածանցյալները 0 կետում հավասար կլինեն իրար, եթե $a=0$, հակառակ դեպքում այս ֆունկցիան ածանցելի չէ 0 կետում:

$$6. \text{SoftPlus: } SP(x) = \ln(1 + e^x) \text{ SoftPlus ակտիվացիոն ֆունկցիան նման է ReLU-ին և ածանցելի է բոլոր կետերում:}$$

2 Sigmoid ֆունկցիայի առաջացում

Հավանականությունների տեսությունում կա odds գաղափար: Վերցնենք P հավանականությունը, որ պատահարը տեղի կունենա: Այդ դեպքում $\text{odds} = \frac{P}{1-P}$: P -ն փոխվում է $(0; 1)$ միջակայքում, նշանակում է odds-ը փոխվում է $(0; +\infty)$ միջակայքում: $w x + b$ փոխվում է $(-\infty; +\infty)$ միջակայքում և եթե նրան ուզում ենք բերել $(0; +\infty)$ միջակայք, կարող ենք օգտագործել էքսպոնենտ:

$$\frac{P}{1-P} = e^{wx+b}$$
$$P = \frac{e^{wx+b}}{1 + e^{wx+b}} = \frac{1}{1 + e^{-wx-b}} = \sigma(wx + b)$$