

SUN'IY NEYRON TARMOQLARIDA FAOLLASHTIRUVCHI FUNKSIYALARINING SAMARADORLIGI

Abdullayeva Barnoxon Nuraliyevna

Namangan viloyati norin tumani 39-sonli maktab o'qituvchisi

Madaminjonov Muhammadyusuf Dilshod o'g'li

Namangan shahar Prezident maktabi o'qituvchisi

Sattarov Nosirbek Abdulhodi o'g'li

"University of Business and Science" Nodavlat Universitet o'qituvchisi

Jo'rayev To'xtamurod Ixvoljon o'g'li

"University of Business and Science" Nodavlat Universiteti o'qituvchisi

"Sun'iy neyron tarmoqlarida har bir neyron o'z kirishlarining vaznli yig'indisini hosil qiladi va natijada olingan skalyar qiymatni faollashtirish funktsiyasi deb ataladigan funktsiya orqali o'tkazadi."

Faollashtirish funktsiyasi neyronning faollashtirilishi yoki faollashtirilmasligini aniqlaydi. Bu shuni anglatadiki, u neyronning tarmoqqa kiritishi bashorat qilish jarayonida tegishli yoki ahamiyatsizligini aniqlash uchun ba'zi oddiy matematik operatsiyalardan foydalanadi. [1] ishda asosan bizga faollashtirish funktsiyalari bizga nima uchun kerak degan savolga batafsil javob berilgan.

Faollashtirish funktsiyalarini uch turga bo'lish mumkin:

- Chiziqli faollashtirish funktsiyasi
- Chiziqli bo'lmagan faollashtirish funktsiyalari

Ko'pincha identifikatsiyani faollashtirish funktsiyasi deb ataladigan chiziqli faollashtirish funktsiyasi kirishga proporsionaldir. Chiziqli faollashtirish funktsiyasining diapazoni ($-\infty$ dan ∞ gacha) bo'ladi. Chiziqli faollashtirish funktsiyasi kirishlarning umumiy og'irligini qo'shadi va natijani qaytaradi.[1]

Ijobiy va salbiy tomonlari

Bu ikkilik faollashtirish emas, chunki chiziqli faollashtirish funktsiyasi faqat bir qator faollashtirishni ta'minlaydi. Biz, albatta, bir nechta neyronlarni bir-biriga ulashimiz mumkin va agar bir nechta faollashuv mavjud bo'lsa, biz maksimal (yoki yumshoq maksimal) ni hisoblashimiz mumkin.

Ushbu faollashtirish funktsiyasining hosilasi doimiydir. Ya'ni, gradient x (kirish) bilan bog'liq emas.

Chiziqli bo'lmagan faollashtirish funktsiyalari eng ko'p ishlatiladigan faollashtirish funktsiyalaridir. Nonlinearlik grafikni shunday ko'rinishga olib keladi



Bu modelga turli xil ma'lumotlarni umumlashtirish yoki moslashtirish va chiqishni farqlashni osonlashtiradi.[2]

Hosila yoki differentsial: y o'qining o'zgarishi va x o'qining o'zgarishi. U qiyalik deb ham ataladi.

Monotonik funktsiya: butunlay o'smaydigan yoki kamaymaydigan funktsiya.

Chiziqli bo'lmagan faollashtirish funktsiyalari asosan ularning diapazoni yoki egri chizig'i asosida bo'linadi.



Linear function

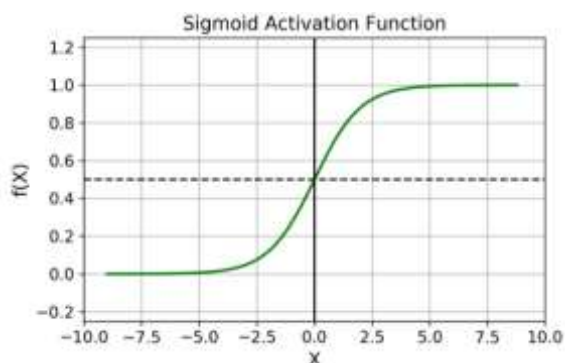


Non-linear function

Best fit linear and non-linear models

Bu funktsiya har qanday haqiqiy qiymatni kirish sifatida qabul qiladi va 0 dan 1 gacha bo'lgan qiymatlarni chiqaradi. Odatda ikkilik tasnifning chiqish qatlamida qo'llaniladi, bunda natija 0 yoki 1 bo'ladi, chunki sigmasimon funktsiya qiymati faqat 0 va 1 oralig'ida bo'ladi, shuning uchun qiymat 0,5 dan katta bo'lsa, natija 1 ga teng bo'lishini oson bashorat qilish mumkin, aks holda 0.[4]

Sigmoid funktsiyasi quyidagi formula asosida ishlaydi: $f(x) = \frac{1}{1 + e^{-x}}$

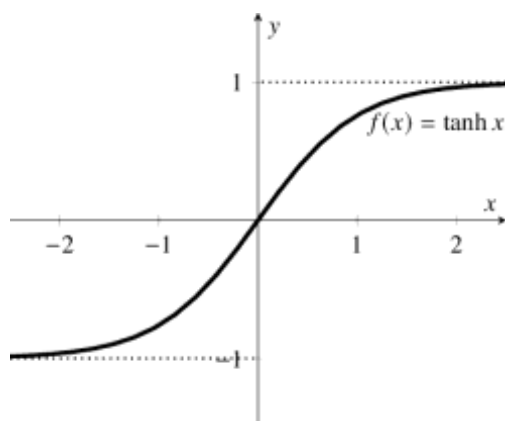


Tanh funktsiyasi sigmasimon/logistik faollashtirish funktsiyasiga juda o'xshaydi va hatto chiqish diapazonidagi farq bilan bir xil S-shakliga ega -1 dan 1 gacha. Tanhda kirish qanchalik katta bo'lsa (musbatroq), chiqish qiymati shunchalik yaqinroq bo'ladi. 1,0 ga teng



bo'ladi, holbuki kirish qanchalik kichik bo'lsa (manfiyroq), chiqish $-1,0$ ga yaqinroq bo'ladi.[4]

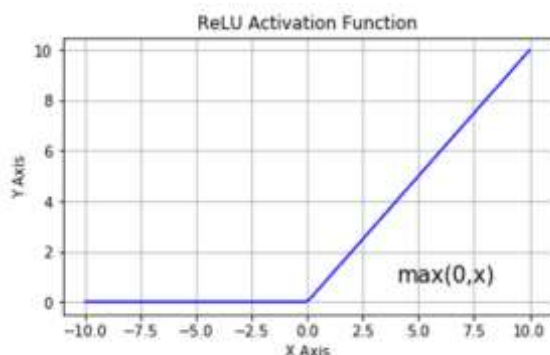
Tanh funksiyasi quyidagi formula asosida ishlaydi: $\tanh(x) = \frac{e^x - e^{-x}}{e^x + e^{-x}}$



ReLU to'g'rilangan chiziqli birlikni anglatadi.

Garchi u chiziqli funktsiya haqida taassurot qoldirsa-da, ReLU hosilaviy funktsiyaga ega va bir vaqtning o'zida uni hisoblash samaradorligini oshirishga imkon beradi.

ReLU funktsiyasi quyidagi formula asosida ishlaydi: $f(x) = \max(0, x)$



Bu yerda asosiy narsa ReLU funktsiyasi barcha neyronlarni bir vaqtning o'zida faollashtirmaydi. Neyronlar faqat chiziqli transformatsiyaning chiqishi 0 dan kam bo'lsa, o'chiriladi.[4]

ReLU-dan faollashtirish funktsiyasi sifatida foydalanishning afzalliklari quyidagilardan iborat:

Faqatgina ma'lum miqdordagi neyronlar faollashtirilganligi sababli, ReLU funktsiyasi sigmasimon va tanh funktsiyalariga qaraganda ancha samarali hisoblash imkonini beradi.

ReLU chiziqli, to'yinmagan xususiyati tufayli gradient tushishining global minimal yo'qotish funktsiyasiga yaqinlashishini tezlashtiradi.

Ushbu funktsiya duch keladigan cheklovlar:

Dying ReLU muammosi.

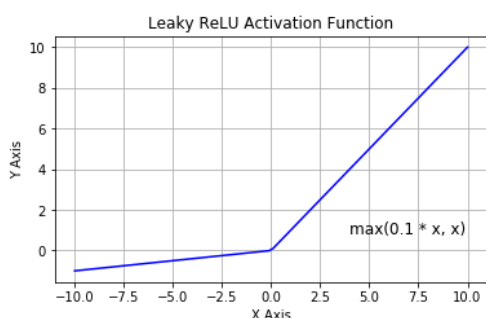


Grafik salbiy bo'lganda, orqaga tarqalish jarayonida ba'zi neyronlar uchun og'irliklar va moyilliklar yangilanmaydi. Bu hech qachon faollashtirilmaydigan o'lik neyronlarni yaratishi mumkin.

Barcha manfiy kirish qiymatlari darhol nolga aylanadi, bu esa modelning ma'lumotlardan to'g'ri moslashish yoki o'qitish qobiliyatini pasaytiradi.

Leaky ReLU - bu Dying ReLU muammosini hal qilish uchun ReLU funktsiyasining takomillashtirilgan versiyasi, chunki u salbiy sohada kichik ijobiy qiymatga ega. Leaky ReLU-ning afzalliklari ReLU-niki bilan bir xil, bundan tashqari, u salbiy kirish qiymatlari uchun ham orqaga tarqalishni faollashtiradi. Salbiy kirish qiymatlari uchun ushbu kichik modifikatsiyani amalga oshirish orqali grafikning chap tomonidagi gradient nolga teng bo'lmagan qiymatga ega bo'ladi. Shuning uchun biz o'sha mintaqada o'lik neyronlarga duch kelmaymiz.

Leaky ReLU funktsiyasi quyidagi formula asosida ishlaydi: $f(x) = \max(0.1x, x)$

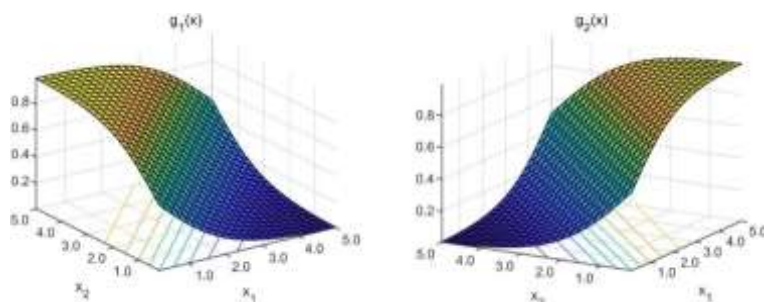


Softmax funktsiyasi kirish sifatida K haqiqiy sondan z vektorini oladi va uni kirish raqamlari ko'rsatkichlariga proporsional K ehtimolliklardan tashkil topgan ehtimollik taqsimotiga normallashtiradi. Ya'ni, softmaxni qo'llashdan oldin, ba'zi vektor komponentlari manfiy yoki bittadan katta bo'lishi mumkin; va 1 ga teng bo'lmasligi mumkin; lekin softmaxni qo'llaganingizdan so'ng, har bir komponent intervalda bo'ladi (0,1) va komponentlar 1 ga qo'shiladi, shuning uchun ular ehtimollik sifatida talqin qilinishi mumkin. Bundan tashqari, kattaroq kirish komponentlari katta ehtimolliklarga mos keladi.

Standart(birlik) Softmax funktsiyasi $\sigma: R^K \mapsto (0,1)^K$, $K \geq 1$ formula bilan aniqlanadi

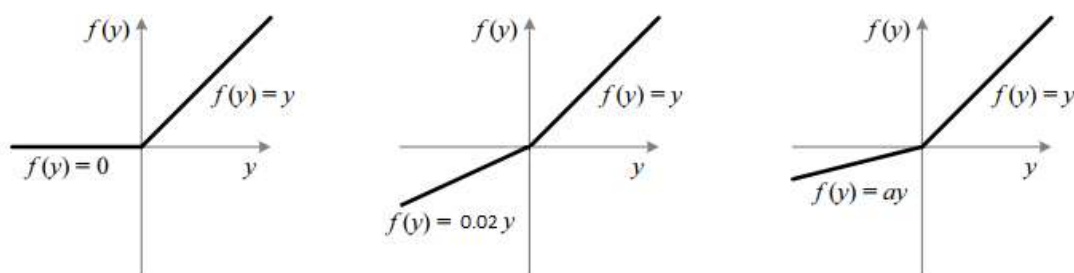
$$\sigma(z)_i = \frac{e^{z_i}}{\sum_{j=1}^K e^{z_j}} \text{ for } i = 1, \dots, K \text{ and } z = (z_1, \dots, z_K) \in R^K.$$





Parametrik ReLU - ReLU ning yana bir variantidir, u o'qning chap yarmida gradientning nolga aylanishi muammosini hal qilishga qaratilgan. Bu funktsiya a argumenti sifatida funktsiyaning manfiy qismining qiyaligini ta'minlaydi. Orqaga tarqalish orqali a ning eng mos qiymati o'rganiladi.[4]

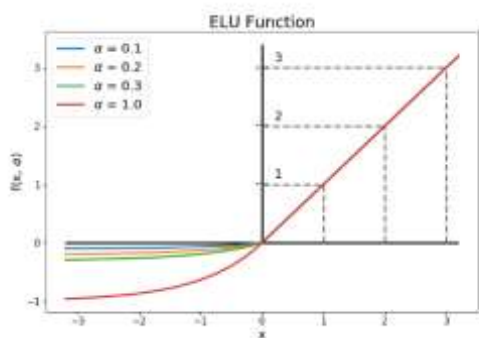
Parametrik funktsiyasi quyidagi formula asosida ishlaydi : $f(y_i) = \begin{cases} y_i, & \text{agar } y_i > 0 \\ a_i y_i, & \text{agar } y_i \leq 0 \end{cases}$



Ekspontensial chiziqli birlik yoki qisqacha ELU ham funktsiyaning salbiy qismining qiyaligini o'zgartiruvchi ReLU variantidir. ELU to'g'ri chiziq bilan o'tmagan ReLU va Parametrik ReLU funktsiyalaridan farqli o'laroq, salbiy qiymatlarni aniqlash uchun log egri chizig'idan foydalanadi.

Ekspontensial chiziqli birliklar (ELU) funktsiyasi quyidagi formula asosida ishlaydi :

$$\begin{cases} f(x) = x & \text{if } x > 0 \\ \alpha(\exp(x) - 1) & \text{if } x \leq 0 \end{cases}$$



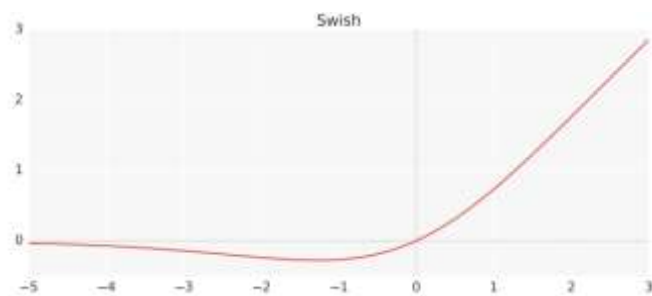
Bu Google tadqiqotchilari tomonidan ishlab chiqilgan o'z-o'zidan himoyalangan faollashtirish funktsiyasi. Swish tasvir tasnifi , mashina tarjimasi va boshqalar kabi turli



qiyin domenlarga qo'llaniladigan chuqur tarmoqlarda ReLU faollashtirish funksiyasiga doimiy ravishda mos keladi yoki undan oshib ketadi.

Swish funksiyasi quyidagi formula asosida ishlaydi : $Swish(x) = x \cdot \sigma(x)$ bu yerda

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

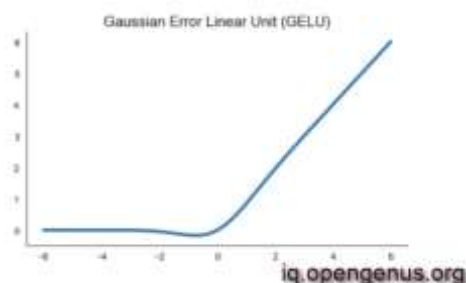


Gauss xatosi chiziqli birligi (GELU) faollashtirish funksiyasi BERT, ROBERT, ALBERT va boshqa eng yaxshi NLP modellari bilan mos keladi. Ushbu faollashtirish funksiyasi tark etish, zonadan chiqish va ReLU xususiyatlarini birlashtirish orqali amalga oshiriladi.

ReLU va tushish birgalikda neyronning chiqishini beradi. ReLU buni aniqlik bilan kiritishni nolga yoki bittaga (kirish qiymatining ijobiy yoki manfiy bo'lishiga qarab) ko'paytirish va o'chirishni stokastik tarzda nolga ko'paytirish orqali amalga oshiradi. Zoneout deb nomlangan RNN muntazamlashtiruvchisi kirish ma'lumotlarini stokastik ravishda bittaga ko'paytiradi. Biz ushbu funktsiyani kirishni nolga yoki stokastik tarzda aniqlangan va kirishga bog'liq bo'lgan birga ko'paytirish orqali birlashtiramiz. Neyron kirishini x ga ko'paytiramiz. Ushbu taqsimot tanlangan, chunki neyron kirishlari odatdagi taqsimotga amal qiladi, ayniqsa ommaviy normalizatsiya bilan.

Gauss xatosi chiziqli birligi (GELU) funksiyasi quyidagi formula asosida ishlaydi :

$$GELU(x) = xP(X \leq x) = x\Phi(x) = x \cdot \frac{1}{2} [1 + \operatorname{erf}(x/\sqrt{2})] \quad \text{agar } X \sim N(0,1)$$



SELU o'z-o'zini normalizatsiya qiluvchi tarmoqlarda aniqlangan va ichki normalizatsiya yo'qolib ketishini oldini oladi, ya'ni har bir qatlam oldingi qatlamlardan



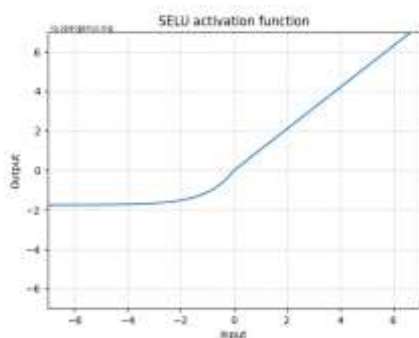
o'rtacha farqni saqlaydi. SELU o'rtacha dispersiyani sozlash orqali ushbu normalizatsiyani ta'minlaydi.

SELU o'rtachani o'zgartirish uchun ijobiy va salbiy qiymatlarga ega, bu ReLU faollashtirish funktsiyasi uchun imkonsiz edi, chunki u salbiy qiymatlarni chiqara olmaydi. Gradientlar dispersiyani sozlash uchun ishlatilishi mumkin. Faollashtirish funktsiyasi uni oshirish uchun birdan kattaroq gradientga ega bo'lgan hududga muhtoj.

Biz quyida ushbu funktsiyalardan foydalanish samarali holatlar jadvalini taqdim etamiz.

Masshtabli eksponensial chiziqli birlik (SELU) funktsiyasi quyidagi formula asosida

$$\text{ishlaydi : } \text{selu}(x) = \lambda \begin{cases} x & \text{if } x > 0 \\ \alpha e^x - \alpha & \text{if } x \leq 0 \end{cases}$$



Dr. Shaik Asif Hussain va Ahlam Salim Abdallah Al Balushilar 2020-yilda ishlab chiqqan modeli inson yuzini aniqlash, tanib olish va tasniflashda 88% aniqlikga erishgan. Ular anaconda va python dasturlarida yuzni aniqlash uchun Viola Jones va Haar Kaskad algoritmidan foydalanishgan. KDEF va VGG 16 ma'lumotlar bazasi hamda Konvolyutsion neyron tarmoqlaridan foydalanishgan. Ishlab chiqilgan algoritmlar mavjud algoritmlardan yaxshiroq yutuqqa erishganini ta'kidlash lozim[5].

Pranav, KB va Manikandanlarning 2020-yilda chop etilgan "Konvolyutsion neyron tarmoqlardan foydalangan holda real vaqt rejimida yuzni aniqlash tizimini loyihalash va baholash" maqolalarida esa real vaqt rejimida yuzni aniqlash loyihasi taklif etilgan hamda ushbu loyihada tanib olish aniqligi 98% ni tashkil etgan ha,da CNN dan foydalanilgan. [6].

Biz ham modelimiz orqali berilgan rasmni tasniflashda faollashtirish funktsiyalarining aniqlik darajalarini ko'rib chiqib holatga qarab faollashtirish funktsiyalarini qanday foydalanish mumkin degan savolga javob berishga harakat qildik. Quyida biz faollashtirish funktsiyalari orqali olgan natijalarimizni keltirib o'tamiz.

FF nomi	Aniqlik ko'rsatkichi (foizlarda)							
Itaratsiyalar	1	1000	2000	3000	5000	7000	9000	10000
ReLU	1%	11%	20%	31%	50%	71%	90%	97%



ELU	1%	10%	20%	30%	50%	70%	91%	95%
GELU	1%	10%	20%	30%	50%	70%	90%	93%
SELU	1%	10%	20%	30%	50%	70%	90%	96%
Sigmoid	1%	10%	18%	29%	50%	69%	87%	91%
Tanh	1%	10%	20%	31%	50%	72%	90%	94%
Softmax	1%	11%	20%	30%	50%	72%	89%	92%
Swish	1%	10%	19%	31%	50%	70%	90%	93%

Biz olingan ushbu natijalar orqali faollashtirish funksiyalari haqida umumiy xulosalarni keltirib o'tmoqchimiz. Demak faollashtirish funksiyalarini turli holatlarda modelni aniqligini o'lchash ham bashorat sifatini oshirishda bizga qo'l keladi. Faollashtirish funksiyalarini turli xil modellarda qo'llab ko'rgan holatda biz quyidagi umumiy jadvalni yaratdik. Ushbu jadvalda biz faollashtirish funksiyalari qay biri qay holatda ishlatishda yuqori samamaraa berishini jadvalda batafsil keltirib o'tdik.

Faollashtirish funksiyasi	Foydalanish holatlari
ReLU (Rectified Linear Unit)	Chuqur neyron tarmoqlarda yashirin qatlamlar, kompyuterni ko'rish vazifalari, mashinani o'rganish, gradient yo'qolish muammolarida
ELU (Exponential Linear Unit)	"O'layotgan ReLU" muammosini hal qilish, yashirin qatlamlardagi murakkablik, matn va nutqni qayta ishlash, Maxsus arxitekturalarda moslashuvchanlik, umumlashtirilgan xususiyatlarni o'rganishlarda
GELU (Gaussian Error Linear Unit)	Tabiiy tilni qayta ishlash (NLP), murakkabroq matematik operatsiyalar, murakkab lingvistik naqshlar va semantik munosabatlar, takomillashtirilgan vakillik o'rganish, ba'zi domenlarda ishlash
SELU (Scaled Exponential Linear Unit)	O'z-o'zini normallashtiradigan neyron tarmoqlari, Chuqur tarmoq barqarorligi, Chuqur tarmoq barqarorligi, Qo'lda ishga tushirishni qisqartirish, Murakkab ma'lumotlar taqdimoti, Chuqur tarmoqlarni o'qitish samaradorligi
Sigmoid	Binar tasnif, logistik regressiya, RNN-larda kirish mexanizmlari, ehtimollarni baholash uchun chiqish qatlami, qatlamni tartibga solishlarda
Tanh	Takroriy neyron tarmoqlari (RNN), chiziqli bo'lmagan yashirin qatlamlar, xususiyatlarni masshtablash, tasvirni qayta ishlash, qoldiq tarmoqlar (ResNets) larda
Softmax	neyron tarmoqlarning chiqish qatlamida ko'p sinfli tasniflash vazifalarida, tasvirlarni tasniflash uchun neyron tarmoqlari, tabiiy tilni qayta ishlash, o'rganishni kuchaytirish, ehtimollarni baholashlarda
Swish	Takomillashtirilgan vakillikni o'rganish, muayyan arxitekturalarda ishlash, tasvirni aniqlash va kompyuterda ko'rish, umumiy maqsadli tarmoqlar, mashg'ulot tezligidagi potentsial yaxshilanishlarda



FOYDALANILGAN ADABIYOTLAR:

1. Marimuthu, P. (n.d.). How activation functions work in deep learning - KDnuggets. KDnuggets. <https://www.kdnuggets.com/2022/06/activation-functions-work-deep-learning.html>.
2. Sharma, S. (2022b, November 20). Activation functions in neural networks - towards data science. Medium. <https://towardsdatascience.com/activation-functions-neural-networks-1cbd9f8d91d6>.
3. GeeksforGeeks. (2023). Activation functions in Neural Networks. GeeksforGeeks. <https://www.geeksforgeeks.org/activation-functions-neural-networks/>.
4. Baheti, P. (2023, April 24). Activation Functions in Neural Networks [12 Types & Use Cases]. V7. <https://www.v7labs.com/blog/neural-networks-activation-functions>.
5. Dr. Shaik Asif Hussain and Ahlam Salim Abdallah Al Balushi 2020 J. Phys.: Conf. Ser. 1432 012087
6. Pranav, KB va Manikandan, J. (2020c). Konvolyutsion neyron tarmoqlardan foydalangan holda real vaqt rejimida yuzni aniqlash tizimini loyihalash va baholash. Procedia Kompyuter fanlari , 171 , 1651–1659. <https://doi.org/10.1016/j.procs.2020.04.177>

