

NEYRON TARMOQLARDA CHIZIQLI BO'LMAGAN FAOLLASHTIRISH FUNKSIYALARI VA ULARDAN FOYDALANISH

<https://doi.org/10.5281/zenodo.8331867>

S.S.Radjabov

S.U.Jo'rayev

Texnika fanlari doktori (DSc), katta ilmiy xodim, "Toshkent irrigatsiya va qishloq xo'jaligini mexanizatsiyalash muhandislari instituti" milliy tadqiqot universiteti, huzuridagi Fundamental va amaliy tadqiqotlar instituti, Toshkent, O'zbekiston
Katta o'qituvchi, Namangan davlat universiteti. Namangan, O'zbekiston,
jurayevs261@gmail.com

Annotatsiya

Chiziqli bo'lmagan faollashtirish funksiyalari chiziqli faollashtirish funksiyalarining quyidagi cheklovlarini hal qiladi: Ular orqaga tarqalishga ruxsat beradi, chunki endi hosila funksiyasi kirish bilan bog'liq bo'ladi va orqaga qaytib, kirish neyronlaridagi qaysi og'irliklar yaxshiroq bashorat qilishini tushunish mumkin; Ular neyronlarning bir nechta qatlamlarini joylashtirish imkonini beradi, chunki chiqish endi bir nechta qatlamlardan o'tgan kirishning chiziqli bo'lmagan kombinatsiyasi bo'ladi. Har qanday chiqish neyron tarmog'ida funksional hisoblash sifatida taqdim etilishi mumkin.

Kalit So'zlar

Sigmoid, Tanh funksiyasi, ReLU, Leaky ReLU, Parametrik ReLU, ELU, Softmax

I.Kirish

Faollashtirish funksiyasi neyronning eng asosiy darajada faollashishi kerakmi yoki yo'qligini aniqlaydi. Har qanday faollashtirish funksiyasi kirishlarning og'irlashtirilgan yig'indisini va egilishni kiritish sifatida qabul qilishi mumkin. Faollashtirish funksiyalariga qadam funksiyasi, Sigmoid, ReLU, Tanh va Softmax kabilar kiradi.

Faollashtirish funksiyasi neyronni faollashtirishni hal qiladi. Bu oddiy matematik operatsiyalar yordamida bashorat qilish jarayonida neyronning tarmoqqa kirishi muhim yoki muhim emasligini hal qiladi

Bir necha xil chiziqli bo'lmagan neyron tarmoqlarni faollashtirish funksiyalarini va ularning xususiyatlarini ko'rib chiqishimiz mumkin.

Sigmoid / Logistik faollashtirish funksiyasi

Bu funksiya har qanday haqiqiy qiymatni kirish sifatida qabul qiladi va 0 dan 1 gacha bo'lgan qiymatlarni chiqaradi.

Kirish qanchalik katta bo'lsa (musbatroq), chiqish qiymati 1 ga yaqinroq bo'ladi, kirish qanchalik kichik bo'lsa (manfiyroq), quyida ko'rsatilganidek, chiqish 0 ga yaqinroq bo'ladi.



1-rasm. Sigmoid/Logistik faollashtirish funksiyasi

Matematik jihatdan uni quyidagicha ifodalash mumkin:

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

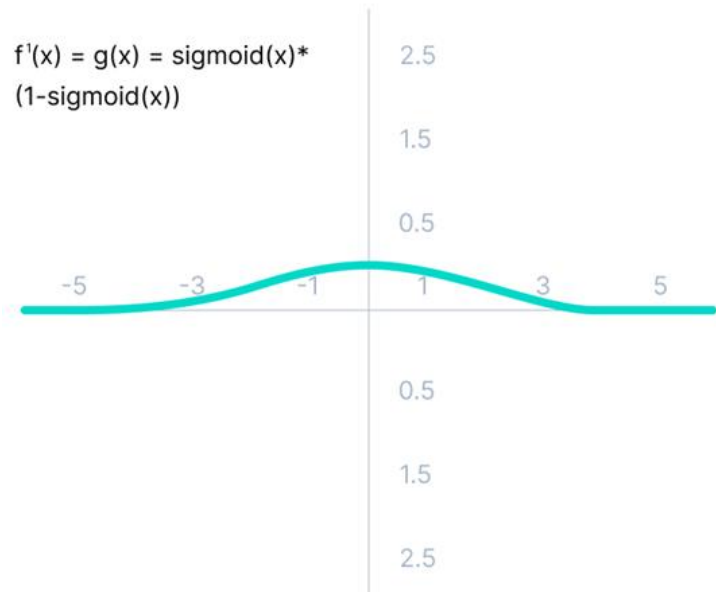
Sigmasimon/logistik faollashtirish funksiyasi nima uchun eng keng tarqalgan funksiyalardan biri hisoblanadi:

- U odatda chiqish sifatida ehtimollikni bashorat qilishimiz kerak bo'lgan modellar uchun ishlatiladi. Har qanday narsaning ehtimoli faqat 0 va 1 oralig'ida mavjud bo'lganligi sababli, sigmasimon egrilik o'z diapazoni tufayli to'g'ri tanlovdir.

- Funksiya farqlanadi va silliq gradientni ta'minlaydi, ya'ni chiqish qiymatlarida sakrashni oldini oladi. Bu sigmasimon shaklni faollashtirish funksiyasining S-shakli bilan ifodalaydi.

Sigmasimon/logistik funksiyasining cheklovlari quyidagicha muhokama qilinadi:

- Funksiyaning hosilasi $f'(x) = \text{sigmoid}(x) * (1 - \text{sigmoid}(x))$ dir.

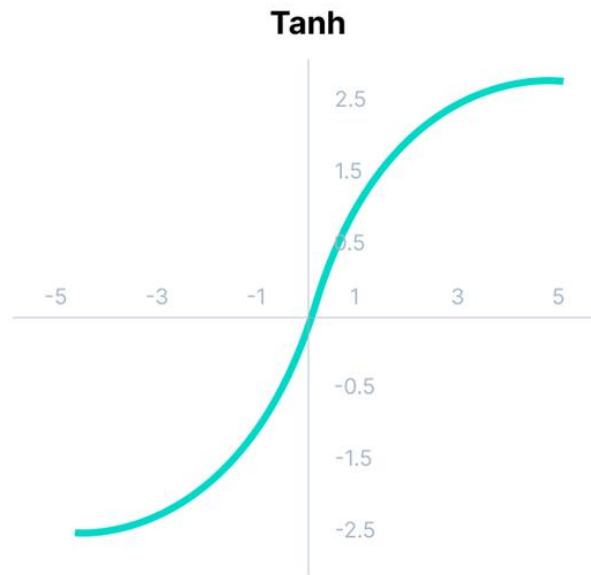


2-rasm. Sigmasimon faollashtirish funksiyasining hosilasi

Yuqoridagi rasmdan ko'rib turganimizdek 3 dan katta yoki -3 dan kichik qiymatlar uchun funksiya juda kichik gradientlarga ega bo'ladi. Gradient qiymati nolga yaqinlashganda, tarmoq o'rganishni to'xtatadi va *Vanishing gradient* muammosidan yo'qotishga uchraydi. Logistik funksiyaning chiqishi nolga yaqin simmetrik emas. Shunday qilib, barcha neyronlarning chiqishi bir xil belgida bo'ladi. Bu neyron tarmoqni o'qitishni qiyinlashtiradi va beqaror qiladi.

Tanh funksiyasi (giperbolik tangens).

Tanh funksiyasi sigmasimon/logistik faollashtirish funksiyasiga 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 1 ga yaqinroq bo'ladi, holbuki kirish qanchalik kichik bo'lsa (manfiyroq), chiqish -1 ga yaqinroq bo'ladi.



3-rasm. Tanh funksiyasi (giperbolik tangens)

Matematik jihatdan uni quyidagicha ifodalash mumkin:

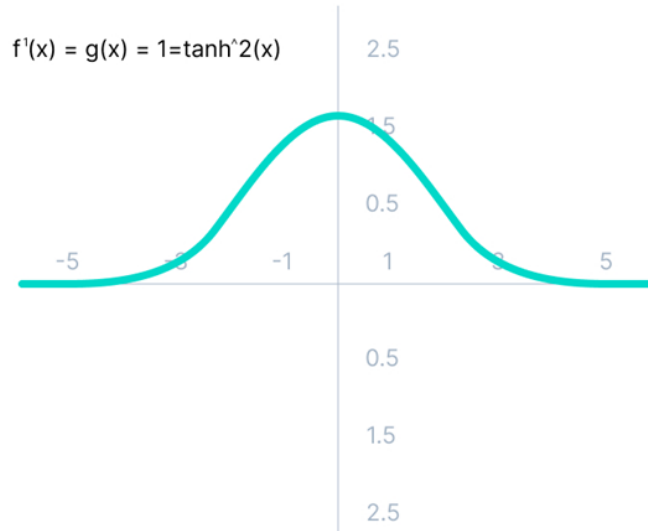
$$f(x) = \frac{(e^x - e^{-x})}{(e^x + e^{-x})}$$

Ushbu faollashtirish funksiyasidan foydalanishning afzalliklari quyidagilardan iborat:

- Tanh faollashtirish funksiyasining chiqishi Nolga markazlashtirilgan, shuning uchun biz chiqish qiymatlarini kuchli salbiy, neytral yoki kuchli ijobiy deb osongina xaritalashimiz mumkin.

- Odatda neyron tarmoqning yashirin qatlamlarida ishlatiladi, uning qiymatlari yashirin qatlam uchun o'rtacha 0 yoki unga juda yaqin bo'ladi. Bu ma'lumotlarni markazlashtirishga yordam beradi va keyingi qatlam uchun o'rganishni ancha osonlashtiradi.

Tanh (derivative)



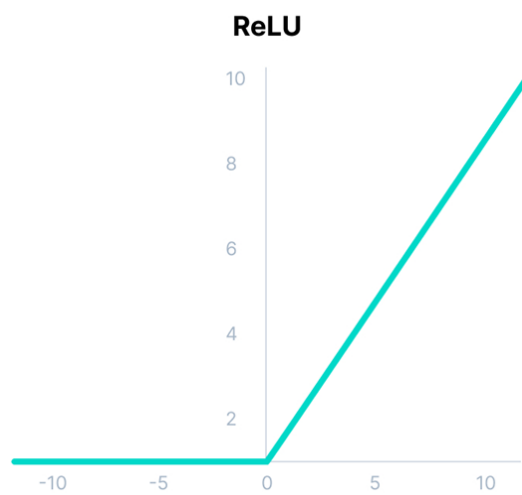
4-rasm. Tanh faollashtirish funksiyasining gradienti (hosilasi).

Ko'rib turganingizdek, u sigmasimon faollashtirish funksiyasiga o'xshash, gradientlarning yo'qolishi muammosiga ham duch keladi. Bundan tashqari, tanh funksiyasining gradienti sigmasimon funksiyaga nisbatan ancha tik.

Sigmasimon va tanhning ikkalasi ham gradientning yo'qolishi muammosiga duch kelgan bo'lsada, tanh nolga markazlashgan va gradientlar ma'lum bir yo'nalishda harakatlanishi bilan cheklanmagan. Shuning uchun amalda tanh har doim sigmasimon faollashtirish funksiyasidan afzaldir.

ReLU funksiyasi

ReLU to'g'rilangan chiziqli birlikni anglatadi. Garchi u chiziqli funksiya haqida taassurot qoldirsada, ReLU hosilaviy funksiyaga ega va bir vaqtning o'zida uni hisoblash samaradorligini oshirishga imkon beradi. Bu yerda asosiy narsa ReLU funksiyasi barcha neyronlarni bir vaqtning o'zida faollashtirmaydi. Neyronlar faqat chiziqli transformatsiyaning chiqishi 0 dan kam bo'lsa, o'chiriladi.



5-rasm. ReLU faollashtirish funksiyasi

Matematik jihatdan uni quyidagicha ifodalash mumkin:

$$f(x) = \max(0, x)$$

ReLU-dan faollashtirish funksiyasi sifatida foydalanishning afzalliklari quyidagilardan iborat:

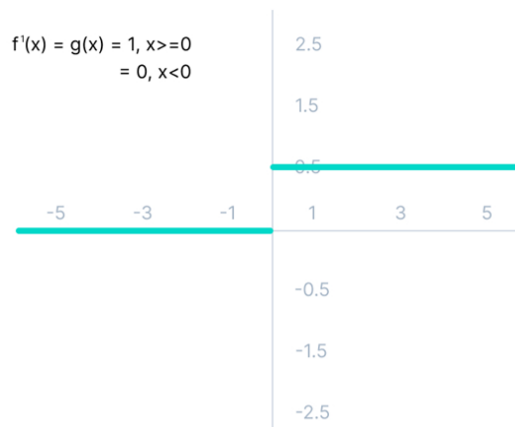
- Faqatgina ma'lum miqdordagi neyronlar faollashtirilganligi sababli, ReLU funksiyasi sigmasimon va tanh funksiyalariga qaraganda ancha samarali hisoblash imkonini beradi.

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

Ushbu funksiya duch keladigan cheklovlar Dying ReLU muammosi sifatida quyida tushuntiriladi.

Dying ReLU muammosi

Dying ReLU muammosi



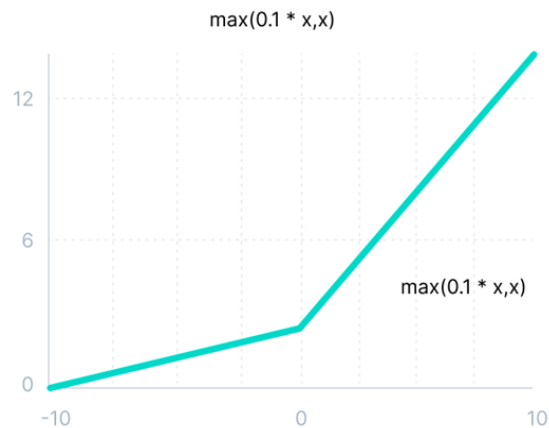
6-rasm. Dying ReLU muammosi

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 funksiyasi

Leaky ReLU - bu Dying ReLU muammosini hal qilish uchun ReLU funksiyasining takomillashtirilgan versiyasi, chunki u salbiy sohada kichik ijobiy nishabga ega.

Leaky ReLU



7-rasm. Leaky ReLU grafigi

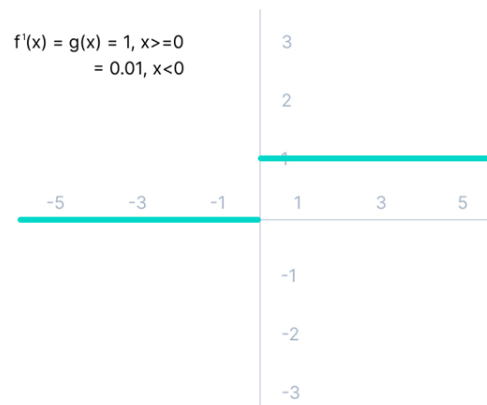
Matematik jihatdan uni quyidagicha ifodalash mumkin:

$$f(x) = \max(0.1x, x)$$

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 (derivative)



8-rasm. Leaky ReLU funksiyasining hosilasi

Ushbu funksiya duch keladigan cheklovlarga quyidagilar kiradi:

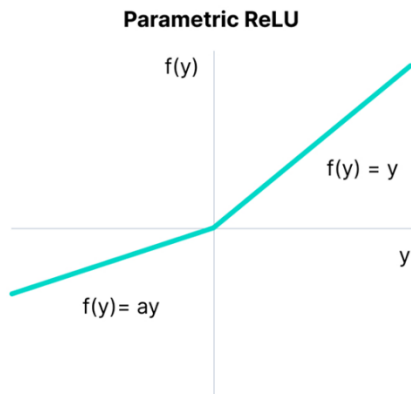
Prognozlar manfiy kirish qiymatlari uchun mos kelmasligi mumkin.

Manfiy qiymatlar uchun gradient model parametrlarini o'rganishni ko'p vaqt talab qiladigan kichik qiymatdir.

Parametrik ReLU funksiyasi

Parametrik ReLU ning yana bir variantidir, u o'qning chap yarmida gradientning nolga aylanishi muammosini hal qilishga qaratilgan.

Bu funksiya a argumenti sifatida funksiyaning manfiy qismining qiyaligini ta'minlaydi. Orqaga tarqalish orqali a ning eng mos qiymati o'rganiladi.



9-rasm. Parametrik ReLU grafigi

Matematik jihatdan uni quyidagicha ifodalash mumkin:

$$f(x) = \max(ax, x)$$

Bu yerda " a " - salbiy qiymatlar uchun nishab parametri.

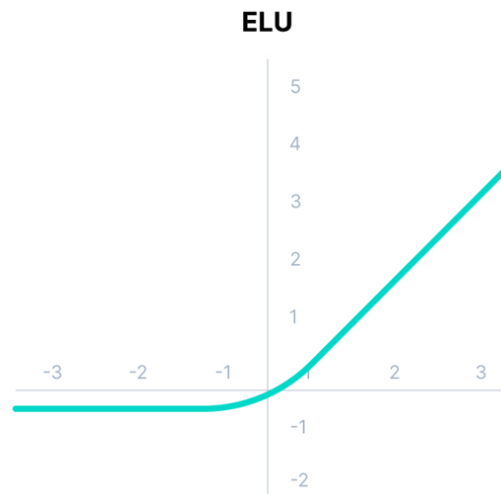
Parametrlangan ReLU funksiyasi o'lik neyronlar muammosini hal qilishda hali ham samarali. ReLU funksiyasi muvaffaqiyatsizlikka uchraganda va tegishli ma'lumotlar keyingi qatlamga muvaffaqiyatli o'tmaganda ishlatiladi.

Bu funksiyaning cheklanishi shundaki, u qiyalik parametri a qiymatiga qarab turli muammolar uchun turlicha ishlashi mumkin.

Ekspontensial chiziqli birliklar (ELUs) funksiyasi

Ekspontensial chiziqli birlik yoki qisqacha ELU ham funksiyaning salbiy qismining qiyaligini o'zgartiruvchi ReLU varianti hisoblanadi.

ELU to'g'ri chiziq bilan o'tmagan ReLU va Parametrik ReLU funksiyalaridan farqli o'laroq, salbiy qiymatlarni aniqlash uchun logorifmik egri chizig'idan foydalanadi.



10-rasm. ELU faollashtirish funksiyasi

Matematik jihatdan uni quyidagicha ifodalash mumkin:

$$\begin{cases} x & \text{for } x \geq 0 \\ \alpha(e^x - 1) & \text{for } x < 0 \end{cases}$$

ELU quyidagi afzalliklari tufayli ReLU uchun kuchli alternativ hisoblanadi:

ELU chiqishi -a ga teng bo'lguncha asta-sekin silliqdashib boradi, RELU esa keskin silliqdashadi.

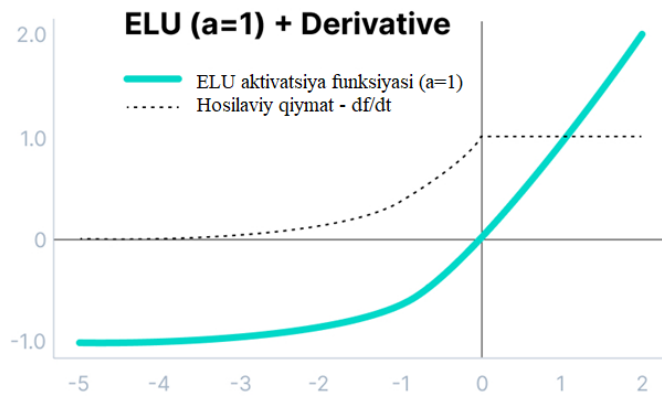
Kirishning salbiy qiymatlari uchun logarifmik egri chizig'ini kiritish orqali o'lik ReLU muammosidan qochadi. Bu tarmoqqa og'irliklar va noto'g'ri yo'nalishlarni to'g'ri yo'naltirishga yordam beradi.

ELU funksiyasining cheklovlari quyidagilardan iborat:

Ekspensial operatsiya kiritilganligi sababli hisoblash vaqtini oshiradi

"a" qiymatini o'rganish amalga oshirilmaydi

Portlash gradient muammosi



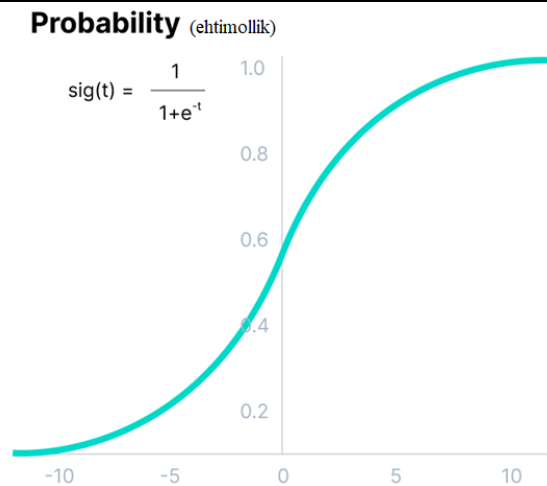
11-rasm. ELU faollashtirish funksiyasi va uning hosilasi

Matematik jihatdan uni quyidagicha ifodalash mumkin:

$$f'(x) = \begin{cases} 1 & \text{for } x \geq 0 \\ f(x) + \alpha & \text{for } x < 0 \end{cases}$$

Softmax funksiyasi

Softmax faollashtirish funksiyasining tafsilotlarini o'rganishdan oldin, biz uning qurilish blokiga - ehtimollik qiymatlarini hisoblashda ishlaydigan sigmasimon / logistik faollashtirish funksiyasiga e'tibor qaratishimiz kerak.



12-rasm. Ehtimollik aktivatsiya funksiyasi grafigi

II.Xulosa

Ushbu maqolada biz ortga tarqalish algoritmidagi ishlatiladigan chiziqli bo'lmagan faollashtirish funksiyalari haqida umumiy ma'lumot berdik.

Matematik funksiyalar va ularning xarakteristikalarining oddiy ro'yxatini tuzish o'rniga, biz masalaga izchil yondashishga harakat qildik, mantiqiy nuqtai nazardan ma'lumotni qayta ishlashga olib keladigan muammolar va ehtimolliklarni ta'kidlab o'tdik.

Nazariy funksiyalar kabi ba'zi kam ma'lum bo'lgan ifodalar neyron tarmoqlarida oldinga va orqaga tarqalish algoritmini yangi muammolarga qo'llashni kengaytirish imkonini berdi.

Tarmoqlarni orqaga tarqalishdan tashqari algoritmlar bilan o'rgatish uchun boshqa ko'plab nochiziqli faollashtirish funksiyalarini ham ko'rishimiz mumkin. Masalan, Radial Basis Functions (RBF) Gauss funksiyalaridan foydalanadi; ammo, bu maqolada, biz orqaga tarqalish mexanizmiga e'tibor qaratdik.

FOYDALANILGAN ADABIYOTLAR:

1. Jo'rayev S.U. Reografik tahlilning rivojlanish istiqbollari haqida //Educational Research in Universal Sciences ISSN: 2181-3515 VOLUME 2 | SPECIAL ISSUE 5 | 2023
2. Niyozmatova N.A. Mamatov N.S. Samijonov A.N. Raximov E. Method for selecting informative and non-informative features // IOP Conf. Series: Materials Science and Engineering 919 (2020) 042013

3. Mamatov N.S. Niyozmatova N.A. Samijonov A.N. Abdullayeva B. Jurayev S.U. The choice of informative features based on heterogeneous functionals// IOP Conf. Series: Materials Science and Engineering 919 (2020) 042009
4. Jo'rayev S. U. PYTHON KUTUBXONALARI //GOLDEN BRAIN. - 2023. - T. 1. - №. 16. - C. 249-260.
5. Jo'rayev S. U. Python programming language and its features // American Journal of Pedagogical and Educational Research Volume 14, | July, 2023
6. Jo'rayev S. U. Performing an arithmetic operation in the python programming language // American Journal of Pedagogical and Educational Research Volume 14, | July, 2023