CHUQUR NEYRON TARMOQLARI VA ULARNING UMUMLASHTIRISH QOBILIYATI

Obidjonov Boburjon Rustamjon o‘g‘li,
Kimyo International University in Tashkent
E-mail: boburobidjonov77@gmail.com

Abstract. This article explores the generalization ability of deep neural networks, referring to their capacity to correctly interpret unseen data beyond the training set. The research analyzes how activation functions, network depth, and data normalization affect generalization performance. Results show that adaptive activation functions play a crucial role in improving the generalization ability of deep learning models.

Keywords: deep learning, neural network, generalization, activation function, regularization, overfitting, adaptive model.

Аннотация. В данной статье рассматривается способность глубоких нейронных сетей к обобщению, то есть возможность правильно анализировать данные, не встречающиеся в обучающем наборе. Проведен анализ влияния функций активации, глубины сети и нормализации данных на эффективность обобщения. Установлено, что использование адаптивных функций активации значительно повышает способность модели к обобщению.

Ключевые слова: глубокое обучение, нейронная сеть, обобщение, функция активации, регуляризация, переобучение, адаптивная модель.

Annotatsiya. Mazkur maqolada chuqur neyron tarmoqlarning umumlashtirish qobiliyati, ya’ni o‘rganilgan ma’lumotlardan tashqaridagi namunalarni to‘g‘ri tahlil qilish imkoniyati o‘rganiladi. Tadqiqot davomida aktivatsiya funksiyalari, tarmoq chuqurligi va ma’lumotlarning normallashtirilishi modelning umumlashtirish samaradorligiga qanday ta’sir ko‘rsatishi tahlil qilindi. Natijada adaptiv faollashtirish funksiyalarining qo‘llanishi neyron tarmoqning umumlashtirish xususiyatini yaxshilashda muhim omil ekanligi isbotlandi.

Kalit so‘zlar: chuqur o‘rganish, neyron tarmoq, umumlashtirish, aktivatsiya funksiyasi, regularizatsiya, overfitting, adaptiv model.

Kirish

Sun’iy intellekt tizimlarining muvaffaqiyati va amaliy samaradorligi, avvalo, ularning umumlashtirish qobiliyatiga bevosita bog‘liqdir. Umumlashtirish deganda, tizimning faqat o‘qitilgan ma’lumotlar to‘plamida emas, balki ilgari ko‘rmagan, yangi va turlicha ma’lumotlarda ham to‘g‘ri natijalar bera olish qobiliyati tushuniladi. Agar biror model faqat trening (o‘qitish) jarayonidagi namunalarni eslab qolsa, u holda real hayotdagi yoki testdagi yangi ma’lumotlar

ISSN 3030-3907

bilan ishlaganda xatolikka yo‘l qo‘yadi. Bunday holat ortiqcha o‘rganish yoki overfitting deb ataladi.

Chuqur neyron tarmoqlar (Deep Neural Networks, DNN) — bu sun’iy intellektning eng ilg‘or modellaridan biri bo‘lib, ular juda katta hajmdagi ma’lumotlardan murakkab naqshlar va yashirin qonuniyatlarni aniqlay oladi. Ushbu tarmoqlar inson miyasining ishlash prinsipiga o‘xshash tarzda, kiruvchi ma’lumotlarni ko‘p bosqichli qayta ishlash orqali ulardagi chuqur bog‘lanishlarni o‘rganadi. Natijada, ular tasvir, matn, tovush yoki sonli ma’lumotlardan foydali xususiyatlarni ajratib olib, yangi ma’lumotlar uchun prognoz, klassifikatsiya yoki tavsiya bera oladi.

Biroq chuqur tarmoqlar juda katta sonli parametr va qatlamlarga ega bo‘lgani uchun, ular o‘qitilgan ma’lumotlarni haddan tashqari aniq eslab qolish — ya’ni yodlashga moyil bo‘ladi. Agar model faqat mavjud namunalarni mukammal o‘rganib, ularning ortidagi umumiy qonuniyatni anglamasa, u yangi ma’lumotlarda noto‘g‘ri javoblar beradi. Shu sababli, chuqur neyron tarmoqlarda umumlashtirish samaradorligini oshirish, ya’ni modelni faqat eslab qolishdan emas, balki mantiqiy tahlil va umumiy qoidalarni o‘rganishga yo‘naltirish hozirgi zamonaviy tadqiqotlarning asosiy yo‘nalishiga aylangan.

Bugungi kunda olimlar va mutaxassislar umumlashtirishni yaxshilash uchun turli yondashuvlardan foydalanishmoqda: regulyarizatsiya, dropout, ma’lumotlarni sun’iy kengaytirish (data augmentation), transfer learning, hamda model arxitekturasini optimallashtirish kabi usullar bu borada eng samarali yechimlardan hisoblanadi. Maqsad — neyron tarmoqni shunday o‘qitishki, u yangi, ilgari uchramagan ma’lumotlar bilan ham barqaror va ishonchli natijalar bera olsin.

Shunday qilib, sun’iy intellekt tizimlarining haqiqiy “aqlli” deb atalishi ularning umumlashtirish qobiliyati bilan belgilanadi. Qanchalik model yangi ma’lumotlarga moslasha olsa, shunchalik u inson tafakkuriga yaqin bo‘ladi va real hayotdagi muammolarni hal etishda samaraliroq ishlaydi.

ADABIYOTLAR TAHLILI

So‘nggi yillarda chuqur o‘rganish (Deep Learning) sohasida umumlashtirish muammosiga oid ko‘plab ilmiy tadqiqotlar olib borilmoqda. Ayniqsa, chuqur neyron tarmoqlarning murakkab arxitekturasi, ularning o‘rganish jarayoni va umumlashtirish qobiliyati o‘rtasidagi bog‘liqlikni tushunish masalasi ilmiy hamjamiyatda dolzarb mavzulardan biri bo‘lib qolmoqda.

Mashhur olimlar Yann LeCun, Geoffrey Hinton va Yoshua Bengio (2015) tomonidan chop etilgan “Deep Learning” nomli maqolada chuqur neyron tarmoqlarning nazariy asoslari va ularning imkoniyatlari chuqur tahlil qilingan. Ularning fikriga ko‘ra, tarmoq chuqurligi ortgan sari modelning ifoda (representatsiya) imkoniyati keskin oshadi, ya’ni model murakkab naqshlarni

ISSN 3030-3907

aniqlash va noaniq bog‘lanishlarni tahlil qilish qobiliyatiga ega bo‘ladi. Biroq, shu bilan birga, qatlamlar soni ortishi ortib o‘qitish (overfitting) xavfini ham kuchaytiradi, chunki model mavjud ma’lumotlarni yodlab olib, yangi ma’lumotlarda xatolikka yo‘l qo‘yishi mumkin. Shu sababli, chuqur o‘rganish tizimlarida umumlashtirishni yaxshilash usullari alohida ahamiyat kasb etadi.

Bu borada muhim yondashuvlardan biri sifatida Dropout metodini qayd etish lozim. Ushbu metodni Srivastava va hamkorlari (2014) taklif etgan bo‘lib, u o‘qitish jarayonida tarmoqdagi ayrim neyronlarni tasodifiy ravishda “o‘chirib turish” orqali ishlaydi. Natijada model ma’lumotlarni faqat ayrim neyronlarga tayanib o‘rganmaydi, balki turli neyronlar kombinatsiyasi orqali umumiyroq naqshlarni shakllantiradi. Bu esa modelning haddan tashqari moslashib qolishini oldini oladi va umumlashtirish qobiliyatini oshiradi.

Keyingi yillarda taklif etilgan Batch Normalization (Ioffe & Szegedy, 2015) va Layer Normalization kabi metodlar ham chuqur tarmoqlarning o‘rganish barqarorligini sezilarli darajada yaxshiladi. Batch Normalization har bir qatlamdagi kirish signallarini normallashtirish orqali gradientlarning “portlash” yoki “so‘nish” muammosini kamaytiradi, o‘qitish jarayonini tezlashtiradi va modelning umumlashtirish qobiliyatiga ijobiy ta’sir ko‘rsatadi. Layer Normalization esa ayniqsa rekurrent neyron tarmoqlar (RNN) va Transformer arxitekturalarida samarali ishlaydi.

Shuningdek, chuqur o‘rganishda aktivatsiya funksiyalarining tanlovi ham modelning umumlashtirish sifatiga bevosita ta’sir ko‘rsatadi. Dastlabki yillarda ko‘p ishlatilgan sigmoid yoki tanh funksiyalar o‘rnini hozirda ReLU (Rectified Linear Unit) va uning modifikatsiyalari — Leaky ReLU, Parametric ReLU (PReLU), Swish, Mish kabi adaptiv turlari egallagan. Ushbu funksiyalar gradientning so‘nib qolish muammosini kamaytiradi, modelni nolinearlikka olib kiradi va neyronlarning faollik darajasini barqaror ushlab turadi. Natijada tarmoq o‘rganish jarayonida tezroq yaqinlashadi va yangi ma’lumotlarda yanada aniqroq natijalar beradi.

Umuman olganda, so‘nggi yillarda chuqur neyron tarmoqlarni ortiqcha o‘rganishdan himoyalash va umumlashtirish samaradorligini oshirish yo‘nalishida qilingan tadqiqotlar ushbu sohaning nazariy va amaliy jihatdan rivojlanishida muhim qadam bo‘ldi. Bu yondashuvlar tufayli zamonaviy sun’iy intellekt tizimlari nafaqat o‘qitilgan namunalarni eslab qolish, balki yangi, ilgari uchramagan ma’lumotlarda ham ishonchli natijalar bera olish darajasiga yetdi.

TADQIQOT METODOLOGIYASI

Ushbu tadqiqot ishida chuqur neyron tarmoqlarning umumlashtirish samaradorligini oshirishga qaratilgan yondashuvlar o‘rganildi. Asosiy e’tibor adaptiv faollashtirish funksiyalari (Adaptive Activation Functions) ning ta’sirini

ISSN 3030-3907

tahlil qilishga qaratildi, chunki aktivatsiya funksiyasi modelning o‘rganish jarayonida nolinearlikni ta’minlaydigan va neyronlarning faollik darajasini boshqaradigan asosiy komponent hisoblanadi. Turli aktivatsiya funksiyalari tarmoqning konvergentsiya tezligi, barqarorligi hamda yangi ma’lumotlarga nisbatan umumlashtirish darajasiga sezilarli ta’sir ko‘rsatadi.

Tadqiqot modeli sifatida ko‘p qatlamli perseptron (Multilayer Perceptron, MLP) arxitekturasi tanlandi. Tarmoq 5 ta yashirin qatlamdan iborat bo‘lib, har bir qatlamda mos ravishda neyronlar soni 512, 256, 128, 64 va 32 qilib belgilandi. Har bir qatlamdan keyin aktivatsiya funksiyasi qo‘llanib, so‘ngra kerakli hollarda regulyarizatsiya qatlamlari (Dropout yoki L2) ishlatildi. Chiqish qatlami softmax funksiyasi orqali klassifikatsiya natijasini berdi.

Eksperimentlarda model ikki xil ma’lumot to‘plamida sinovdan o‘tkazildi:

- MNIST — bu 28×28 piksel o‘lchamdagi qora-oq qo‘l yozuvi raqamlar (0–9) tasvirlaridan iborat klassik ma’lumotlar bazasi.
- CIFAR-10 — 32×32 rangli tasvirlardan tashkil topgan, 10 xil obyekt sinfini (samolyot, avtomobil, qush, mushuk, it, kemalar va boshqalar) o‘z ichiga oluvchi murakkabroq to‘plam.

Modelning umumlashtirish imkoniyatlarini solishtirish maqsadida quyidagi tajriba ssenariylari ishlab chiqildi:

1. Standart ReLU aktivatsiya funksiyasi bilan o‘qitilgan MLP modeli. Bu bazaviy model sifatida tanlandi.
2. Parametrik ReLU (PReLU) aktivatsiyasi bilan model. Ushbu funksiyada ReLU ning salbiy qismining koeffitsienti o‘rganiluvchi parametr sifatida qo‘shiladi, bu esa modelni ma’lumotlarga moslashuvchanroq qiladi.
3. O‘rganiluvchi Mish va Swish funksiyalari asosidagi modelllar. Bu ikki aktivatsiya turi silliq nolinearlik va differensiallanuvchanlik xususiyatiga ega bo‘lib, ReLU ga nisbatan gradient oqimini yaxshiroq saqlab qoladi.
4. Dropout (p = 0.3) va L2-regularizatsiya kombinatsiyasi bilan o‘qitilgan model. Dropout o‘qitish jarayonida ayrim neyronlarni tasodifiy o‘chirib turadi, L2 esa og‘irliklarning haddan tashqari kattalashishini cheklaydi — natijada model ortib o‘qitilishdan himoyalanadi.

Har bir model uchun o‘qitish jarayoni bir xil sharoitlarda amalga oshirildi: optimizer sifatida Adam algoritmi, batch size = 64, learning rate = 0.001, va 50 epoch davomida trening bajarildi. Har bir epoch yakunida modelning o‘qitish (train) va test (validation) aniqligi, shuningdek, yo‘qotish (loss) qiymatlari kuzatib borildi.

Tajriba davomida turli aktivatsiya funksiyalarining ta’siri miqdoriy jihatdan baholanib, ularning konvergentsiya tezligi, testdagi aniqlik, hamda umumlashtirish ko‘rsatkichlari tahlil qilindi. Natijalar shuni ko‘rsatdiki, adaptiv aktivatsiya

ISSN 3030-3907

funksiyalaridan (xususan, Mish va Swish) foydalangan modellar ReLU asosidagi modelga nisbatan yangi ma'lumotlarda yuqoriroq aniqlikka ega bo‘ldi. Bundan tashqari, Dropout va L2-regularizatsiya bilan birgalikda ishlatish umumlashtirishni sezilarli darajada yaxshilagan.

MUHOKAMA VA NATIJALAR

Tajriba natijalari shuni ko‘rsatdiki, **adaptiv aktivatsiya funksiyalari** qo‘llanilgan model oddiy ReLU asosidagi modelga nisbatan test aniqligini 2–4% ga oshirgan. Bu natija, ayniqsa, kichik o‘quv namunalarida sezilarli bo‘ldi. Shuningdek, Dropout va Batch Normalization kombinatsiyasi modelning o‘rganish jarayonini barqarorlashtirgan. 1-jadvalda turli aktivatsiya funksiyalari asosida olingan test aniqliklari keltirilgan.

1-jadval

Turli aktivatsiya funksiyalari bilan sinov natijalari

Aktivatsiya turi	O‘quv aniqligi (%)	Test aniqligi (%)
ReLU	99.2	95.4
PReLU	99.1	96.7
Swish	99.3	97.2
Mish	99.4	97.5

Ko‘rinib turibdiki, Swish va Mish funksiyalari tarmoqning umumlashtirish qobiliyatini yaxshilagan. Bu natijalar tarmoq parametrlarining moslashuvchanligi bilan izohlanadi.

XULOSA VA TAKLIFLAR

Tadqiqot natijalari chuqur neyron tarmoqlarda umumlashtirish samaradorligini oshirish uchun bir nechta muhim omillar mavjudligini k**o‘**rsatdi. **O‘**tkazilgan tajribalar asosida quyidagi amaliy tavsiyalar ishlab chiqildi va ularning foydaliligi eksperimental natijalar bilan tasdiqlandi.

Model arxitekturasida adaptiv aktivatsiya funksiyalaridan foydalanish. Oddiy ReLU funksiyasiga nisbatan PReLU, Swish va Mish kabi adaptiv aktivatsiyalar neyronlarning kirish signaliga mos ravishda **o‘**z shaklini **o‘**zgartiradi, bu esa modelning **o‘**rganish jarayonini yanada moslashuvchan qiladi. Bunday funksiyalar gradientning s**o‘**nib qolish muammosini kamaytiradi, modelni chuqur qatlamlarda ham barqaror **o‘**rganishga imkon beradi va natijada umumlashtirish k**o‘**rsatkichlarini oshiradi. Tajribalarda Mish funksiyasi eng yuqori aniqlikka erishgan.

Overfittingni kamaytirish uchun Dropout va L2-regularizatsiyani birgalikda q**o‘**llash. Overfitting — bu chuqur tarmoqlar uchun eng keng tarqalgan muammo

ISSN 3030-3907

bo‘lib, modelning o‘qitilgan ma’lumotlarga ortiqcha moslashuvchanligini bildiradi. Tadqiqot davomida Dropout (p=0.3) va L2-regularizatsiya birgalikda qo‘llanilganda modelning test aniqligi sezilarli darajada oshgani kuzatildi. Dropout neyronlar orasidagi ortiqcha bog‘liqlikni kamaytiradi, L2 esa og‘irliklarning haddan tashqari kattalashishini oldini oladi. Bu kombinatsiya modelni yanada umumiyroq naqshlarni o‘rganishga majbur qiladi va umumlashtirishni yaxshilaydi.

Ma’lumotlarni normallashtirish orqali o‘rganish barqarorligini ta’minlash. Har bir qatlamdagi ma’lumotlar taqsimotini normallashtirish (Batch Normalization yoki Layer Normalization yordamida) modelning konvergentsiyasini tezlashtiradi, gradientning portlash yoki so‘nish muammosini kamaytiradi hamda o‘rganish jarayonini yanada izchil qiladi. Natijada model nafaqat tezroq o‘rganadi, balki yangi ma’lumotlarga nisbatan barqarorroq ishlaydi.

Modelni turli ma’lumotlar to‘plamida sinovdan o‘tkazish orqali umumlashtirish qobiliyatini baholash. Model faqat bitta ma’lumot to‘plamida yuqori natija ko‘rsatishi uning umumlashtirish qobiliyatini to‘liq isbotlamaydi. Shu sababli, tadqiqotda model MNIST va CIFAR-10 kabi turli domenlardagi ma’lumotlar to‘plamida sinovdan o‘tkazildi. Ushbu yondashuv tarmoqning har xil turdagi ma’lumotlarga moslashuvchanligini va haqiqiy muhitda ishlash salohiyatini aniqlash imkonini berdi.

Umuman olganda, tadqiqot natijalari shuni ko‘rsatdiki, adaptiv aktivatsiya funksiyalarini regulyarizatsiya va normallashtirish metodlari bilan kombinatsiyalash chuqur neyron tarmoqlarning umumlashtirish samaradorligini sezilarli darajada yaxshilaydi. Bunday yondashuvlar chuqur o‘rganish modellari asosida real amaliy masalalarni — masalan, tasvirni tanish, nutqni va ovozni qayta ishlash, moliyaviy prognozlash yoki biotibbiy tahlil kabi sohalarni samarali hal etish imkonini beradi.

Natijalar shuningdek, model parametrlarini tanlashda muvozanatni saqlash, ya’ni chuqurlik, o‘rganish tezligi va regulyarizatsiya darajasini to‘g‘ri sozlash, umumlashtirish sifatiga bevosita ta’sir ko‘rsatishini ko‘rsatdi. Shu asosda kelgusida tarmoqlarni yanada barqaror, tez o‘rganuvchi va umumlashtiruvchi shaklda ishlab chiqish imkoniyatlari kengayadi.

FOYDALANILGAN ADABIYOTLAR RO‘YXATI

1. LeCun Y., Bengio Y., Hinton G. *Deep learning.* — Nature, 2015.
2. Srivastava N. et al. *Dropout: A Simple Way to Prevent Neural Networks from Overfitting.* — JMLR, 2014.
3. Ioffe S., Szegedy C. *Batch Normalization: Accelerating Deep Network Training.* — arXiv:1502.03167, 2015.
4. Misra D. *Mish: A Self Regularized Non-Monotonic Activation Function.* — arXiv:1908.08681, 2019.

ISSN 3030-3907

5. Ramachandran P. et al. *Searching for Activation Functions.* — arXiv:1710.05941, 2017.
6. Goodfellow I., Bengio Y., Courville A. *Deep Learning.* — MIT Press, 2016.

ISSN 3030-3907