NEYRON TARMOQLARIDA FAOLLASHTIRISH FUNKSIYALARINING TAQQOSIY TAHLILI: SIGMOID, TANH VA RELU FUNKSIYALARINING NAZARIY VA AMALIY JIHATLARI

Tojimamatov Israil Nurmamatovich
Farg'ona davlat universiteti Amaliy matematika va informatika kafedrasi katta o'qituvchisi
E-mail: israiltojimamatov@gmail.com
Ibrohimova Zulhumor Shavkatjon qizi
Farg'ona davlat universiteti Amaliy matematika yo'nalishi 3-bosqich 23.09-guruh talabasi
E-mail: zulhumoribrohimova89@gmail.com

Annotatsiya: *Ushbu maqola neyron tarmoqlardagi faollashtirish funksiyalarining keng qamrovli tadqiqotiga bag'ishlangan bo'lib, sigmoid funksiya, giperbolik tangens va to'g'rilangan chiziqli birlikka alohida e'tibor qaratilgan. Ish har bir faollashtirish funksiyasining matematik asoslarini ochib beradi, ularning hisoblash xususiyatlarini tahlil qiladi va zamonaviy chuqur o'rganish arxitekturalarida amaliy qo'llanilishini o'rganadi. Yo'qolib borayotgan gradient muammosi, hisoblash samaradorligi va turli faollashtirish funksiyalarining ma'lumotlardagi nochiziqli bog'liqliklarni modellashtirish qobiliyatiga alohida e'tibor beriladi. Tadqiqot faollashtirish funksiyasini tanlash yaqinlashish tezligi, model aniqligi va neyron tarmoqning umumiy ishlashiga sezilarli ta'sir ko'rsatishini ko'rsatadi. Tahlil natijalari har bir funksiyaning turli mashinali o'rganish vazifalarida, jumladan, tasvirlarni tasniflash, tabiiy tilni qayta ishlash va naqshlarni tanishda afzalliklari va cheklovlarini ko'rsatadi. Ish sun'iy intellekt sohasidagi tadqiqotchilar va amaliyotchilar uchun hal qilinayotgan muammoning o'ziga xosligi va neyron tarmoq arxitekturasiga qarab faollashtirish funksiyalarini optimal tanlash bo'yicha amaliy tavsiyalar beradi.*

Kalit so'zlar: *faollashtirish funksiyalari, neyron tarmoqlar, sigmoid funksiya, giperbolik tangens, to'g'rilangan chiziqli birlik, chuqur o'rganish, gradient tushish, yo'qolib borayotgan gradient muammosi*

Аннотация: *Данная статья посвящена комплексному исследованию функций активации в нейронных сетях с особым акцентом на сигмоидальную функцию, гиперболический тангенс и выпрямленную линейную единицу. Работа раскрывает математические основы каждой функции активации, анализирует их вычислительные характеристики и исследует практическое применение в современных архитектурах глубокого обучения. Особое внимание уделяется проблемам затухающего градиента, вычислительной эффективности и способности различных функций активации моделировать нелинейные зависимости в данных. Исследование демонстрирует, что выбор функции активации существенно влияет на скорость сходимости, точность модели и* *общую производительность нейронной сети. Результаты анализа показывают преимущества и ограничения каждой функции в контексте различных задач машинного обучения, включая классификацию изображений, обработку естественного языка и распознавание образов. Работа предоставляет практические рекомендации для исследователей и практиков в области искусственного интеллекта относительно оптимального выбора функций активации в зависимости от специфики решаемой задачи и архитектуры нейронной сети.*

Ключевые слова: *функции активации, нейронные сети, сигмоидальная функция, гиперболический тангенс, выпрямленная линейная единица, глубокое обучение, градиентный спуск, проблема затухающего градиента*

Annotation: *This article presents a comprehensive investigation of activation functions in neural networks with particular emphasis on the sigmoid function, hyperbolic tangent, and rectified linear unit. The work reveals the mathematical foundations of each activation function, analyzes their computational characteristics, and explores practical applications in modern deep learning architectures. Special attention is devoted to the problems of vanishing gradient, computational efficiency, and the ability of various activation functions to model nonlinear dependencies in data. The study demonstrates that the choice of activation function significantly affects convergence speed, model accuracy, and overall neural network performance. The analysis results show the advantages and limitations of each function in the context of different machine learning tasks, including image classification, natural language processing, and pattern recognition. The work provides practical recommendations for researchers and practitioners in the field of artificial intelligence regarding the optimal selection of activation functions depending on the specifics of the problem being solved and the neural network architecture.*

Keywords: *activation functions, neural networks, sigmoid function, hyperbolic tangent, rectified linear unit, deep learning, gradient descent, vanishing gradient problem*

KIRISH

Sun'iy intellekt va mashinali o'rganish sohasining jadal rivojlanishi zamonaviy texnologiyalar landshaftida inqilobiy o'zgarishlarni keltirib chiqarmoqda. Neyron tarmoqlar ushbu rivojlanishning markaziy elementi sifatida ko'plab ilmiy va amaliy muammolarni hal qilishda muhim rol o'ynamoqda. Zamonaviy chuqur o'rganish modellari tasvirlarni tanish, tabiiy tilni qayta ishlash, avtomatlashtirilgan haydash va tibbiy diagnostika kabi turli sohalarda insonning kognitiv qobiliyatlariga yaqinlashgan natijalarni namoyish etmoqda.

Neyron tarmoqlarning muvaffaqiyati ko'p jihatdan ularning murakkab nochiziqli bog'liqliklarni modellashtirish qobiliyatiga bog'liq. Bu qobiliyat asosan faollashtirish funksiyalarining to'g'ri tanlanishi va qo'llanilishi orqali amalga oshiriladi. Faollashtirish funksiyalari biologik neyronlarning impulsli xatti-harakatini taqlid qilish uchun yaratilgan bo'lib, sun'iy neyron tarmoqlarga nochiziqli xususiyatlarni kiritadi. Ushbu nochiziqlilik neyron tarmoqlarga murakkab funksiyalarni yaqinlashtirish va real dunyo ma'lumotlarining murakkab tuzilmalarini o'rganish imkonini beradi.

Faollashtirish funksiyalari neyron tarmoqlarning har bir qatlamida neyronlarning chiqish qiymatlarini hisoblash uchun qo'llaniladi. Ular kirish signallarining og'irlikli yig'indisini ma'lum bir oraliqqa transformatsiya qiladi va shu orqali tarmoqning ifodaviy kuchini oshiradi. Faollashtirish funksiyasisiz neyron tarmoq qanchalik chuqur bo'lmasin, u faqat chiziqli transformatsiyalarni amalga oshirishi mumkin bo'ladi. Bu esa tarmoqning foydaliligi va amaliy qo'llanilishi jihatidan jiddiy cheklovlarni keltirib chiqaradi.

Tarixiy jihatdan, sigmoid va giperbolik tangens kabi klassik faollashtirish funksiyalari neyron tarmoqlar nazariyasining dastlabki bosqichlarida keng qo'llanilgan. Ushbu funksiyalar biologik asoslangan va matematik jihatdan yaxshi o'rganilgan xususiyatlarga ega. Biroq, chuqur neyron tarmoqlarning rivojlanishi bilan ushbu funksiyalarning ba'zi kamchiliklari, xususan, yo'qolib borayotgan gradient muammosi, aniq namoyon bo'la boshladi. Bu muammo chuqur tarmoqlarni samarali o'qitishni qiyinlashtirdi va yangi yechimlarni izlashga olib keldi.

Yigirma birinchi asrning boshlarida to'g'rilangan chiziqli birlik funksiyasi kiritilishi neyron tarmoqlar sohasida muhim qadam bo'ldi. Ushbu funksiyaning soddaligi va hisoblash samaradorligi uni zamonaviy chuqur o'rganish arxitekturalarining asosiy komponenti sifatida tez sur'atda mashhurlashtirdi. To'g'rilangan chiziqli birlik ko'plab amaliy vazifalarida klassik funksiyalarga nisbatan ustunlikni namoyish etdi va chuqur tarmoqlarni o'qitishni sezilarli darajada tezlashtirdi.

Hozirgi vaqtda faollashtirish funksiyalari bo'yicha tadqiqotlar davom etmoqda va yangi variantlar taklif etilmoqda. Har bir funksiya o'zining afzalliklari va kamchiliklariga ega bo'lib, ma'lum bir vazifa va arxitektura uchun eng mos keladigan variantni tanlash muhim ahamiyatga ega. Ushbu maqolada sigmoid, giperbolik tangens va to'g'rilangan chiziqli birlik funksiyalarining chuqur tahlili amalga oshiriladi va ularning nazariy asoslari, matematik xususiyatlari, amaliy qo'llanilishi va cheklovlari batafsil o'rganiladi.

Tadqiqotning asosiy maqsadi bu uchta asosiy faollashtirish funksiyasining to'liq taqqosiy tahlilini taqdim etish va turli xil mashinali o'rganish vazifalarida ulardan samarali foydalanish bo'yicha amaliy tavsiyalar berishdir. Ushbu bilimlar sun'iy intellekt sohasidagi tadqiqotchilar, muhandislar va talabalar uchun neyron tarmoqlarni loyihalash va optimallashtirish jarayonida asosiy qo'llanma bo'lib xizmat qilishi mo'ljallangan.

FAOLLASHTIRISH FUNKSIYALARINING NAZARIY ASOSLARI

Neyron tarmoqlarda faollashtirish funksiyalarining zaruriyati biologik neyron tarmoqlarning ishlash printsiplaridan kelib chiqadi. Biologik neyronlar elektrokimyoviy signallarni qabul qiladi va ma'lum bir chegaraviy qiymatdan oshganda impuls hosil qiladi. Ushbu nochiziqli xatti-harakatni sun'iy neyron tarmoqlarda takrorlash uchun matematik transformatsiyalar ishlab chiqilgan bo'lib, ular faollashtirish funksiyalari deb ataladi.

Matematik nuqtai nazardan, faollashtirish funksiyasi har bir neyronning chiqish qiymatini uning kirish signallarining og'irlikli yig'indisidan hisoblash uchun qo'llaniladigan transformatsiyadir. Berilgan neyron uchun kirish vektori va og'irliklar vektori bo'lsin. Neyronning chiqishi faollashtirish funksiyasi orqali hisoblanadi va bu transformatsiya kirish bo'shlig'ini chiqish bo'shlig'iga xaritalashtiradi. Ushbu xaritalashtirish tarmoqning ifodaviy qobiliyatini belgilaydi va murakkab funksiyalarni yaqinlashtirish imkonini beradi.

Faollashtirish funksiyalari bir qator muhim matematik xususiyatlarga ega bo'lishi kerak. Birinchidan, ular nochiziqli bo'lishi zarur, chunki faqat nochiziqli funksiyalar neyron tarmoqlarga murakkab nochiziqli bog'liqliklarni o'rganish imkonini beradi. Agar faollashtirish funksiyasi chiziqli bo'lsa, bir nechta qatlamlardan iborat chuqur tarmoq ham faqat chiziqli transformatsiyani amalga oshiradi. Ikkinchidan, ular uzluksiz va farqlanuvchi bo'lishi kerak, chunki gradientga asoslangan optimallashtirish algoritmlari gradientlarni hisoblashni talab qiladi.

Uchinchidan, faollashtirish funksiyalarining hosilasi hisoblash jihatidan samarali bo'lishi muhim ahamiyatga ega. Neyron tarmoqlarni o'qitish jarayonida gradientlar orqaga tarqalish algoritmi yordamida hisoblanadi va bu jarayon barcha qatlamlarda faollashtirish funksiyalarining hosilalarini hisoblashni o'z ichiga oladi. Shuning uchun, hosilani tez va samarali hisoblash mumkin bo'lgan funksiyalar afzalroqdir.

To'rtinchidan, faollashtirish funksiyalari gradient oqimini saqlab qolish qobiliyatiga ega bo'lishi kerak. Yo'qolib borayotgan gradient muammosi chuqur neyron tarmoqlarning o'qitilishida jiddiy to'siq hisoblanadi. Ushbu muammo gradientlarning orqaga tarqalish jarayonida qiymatlarining keskin kamayishi natijasida yuzaga keladi va tarmoqning dastlabki qatlamlarini o'qitishni deyarli imkonsiz qiladi. Shuning uchun, zamonaviy faollashtirish funksiyalari gradientlarni saqlash yoki hech bo'lmaganda yo'qolib borishini sekinlashtirish uchun maxsus ishlab chiqilgan.

Beshinchidan, faollashtirish funksiyalarining qiymatlari oralig'i muhim ahamiyatga ega. Ba'zi funksiyalar cheklangan oraliqda qiymatlar qabul qiladi, masalan, noldan birga yoki minus birdan plyus birgacha. Boshqa funksiyalar esa chegaralanmagan bo'lishi mumkin. Cheklangan faollashtirish funksiyalari chiqish qiymatlarini normallashtirishda foydali bo'lishi mumkin, lekin ba'zi holatlarda gradientlarning yo'qolishiga olib kelishi mumkin.

Oltinchidan, faollashtirish funksiyalarining simmetriyasi ham muhim xususiyatdir. Ba'zi funksiyalar koordinatalar boshiga nisbatan simmetrik bo'lib, ularning o'rtacha qiymati nolga teng. Bu xususiyat o'qitish jarayonini tezlashtirishi va tarmoqning yaqinlashishini yaxshilashi mumkin. Boshqa funksiyalar esa nosimmetrik bo'lib, ularning o'rtacha qiymati noldan farq qiladi.

Faollashtirish funksiyalarining yana bir muhim jihati ularning salbiy kirishlarga nisbatan xatti-harakatidir. Ba'zi funksiyalar salbiy kirishlarni butunlay yo'q qiladi, ya'ni nolga aylantiradi. Boshqalari esa salbiy kirishlarni ma'lum bir transformatsiya orqali o'tkazadi. Bu xususiyat tarmoqning o'rganish qobiliyatiga va xususiyatlarni saqlash imkoniyatiga ta'sir qiladi.

Faollashtirish funksiyalarining hisoblash murakkabligi ham amaliy jihatdan muhim hisoblanadi. Oddiy arifmetik amallardan iborat funksiyalar eksponentsial yoki trigonometrik funksiyalarga nisbatan tezroq hisoblanadilar. Zamonaviy chuqur o'rganish modellari millionlab parametrlarni o'z ichiga olishi mumkin va hisoblash samaradorligi muhim ahamiyat kasb etadi.

Oxir-oqibat, faollashtirish funksiyasining tanlovi tarmoqning arxitekturasi, hal qilinadigan vazifa turi va mavjud hisoblash resurslari kabi ko'plab omillarga bog'liq. Universal eng yaxshi faollashtirish funksiyasi mavjud emas va har bir holat uchun tajriba asosida optimal variant tanlanishi kerak. Keyingi bo'limlarda sigmoid, giperbolik tangens va to'g'rilangan chiziqli birlik funksiyalarining har biri batafsil tahlil qilinadi va ularning nazariy xususiyatlari, amaliy qo'llanilishi va cheklovlari muhokama qilinadi.

SIGMOID FUNKSIYASI: KLASSIK YONDASHUV VA UNING XUSUSIYATLARI

Sigmoid funksiyasi neyron tarmoqlar tarixida eng qadimgi va eng ko'p o'rganiladigan faollashtirish funksiyalaridan biridir. Uning nomi lotincha sigma harfiga o'xshash S shaklidan kelib chiqqan. Sigmoid funksiyasi logistik funksiya deb ham ataladi va biologik neyronlarning faoliyatini modellashtirish uchun tabiiy tanlov hisoblanadi.

Sigmoid funksiyasining matematik ta'rifi sodda lekin samarali. Bu funksiya har qanday haqiqiy sonni nol va bir orasidagi qiymatga transformatsiya qiladi. Funksiyaning grafigi S shaklidagi egri chiziqni hosil qiladi va u kirish qiymati nolga teng bo'lganda yarim qiymatni qabul qiladi. Kirish qiymati oshgan sari funksiya birga intiladi va kamaygan sari nolga intiladi.

Sigmoid funksiyasining muhim matematik xususiyatlari uning uzluksizligi va cheksiz marta farqlanuvchiligidir. Funksiyaning hosilasi ham matematik jihatdan oddiy ko'rinishga ega va u o'zi orqali ifodalanishi mumkin. Bu xususiyat orqaga tarqalish algoritmini amalga oshirishda hisoblashlarni sezilarli darajada soddalashtiradi.

Sigmoid funksiyasining hosilasi maksimal qiymatiga kirish nolga teng bo'lganda erishadi va bu qiymat chorakka teng. Kirish qiymati noldan uzoqlashgan sari hosila tez sur'atda kamayadi va cheksizlikka intilganda nolga yaqinlashadi. Bu xususiyat yo'qolib borayotgan gradient muammosining asosiy sabablaridan biridir. Chuqur tarmoqlarda sigmoid funksiyalarining hosilalari ko'paytirilganda natija tez sur'atda nolga intiladi va tarmoqning dastlabki qatlamlarini o'qitish deyarli to'xtaydi.

Sigmoid funksiyasining yana bir muhim xususiyati uning ehtimollik talqiniga ega ekanligidir. Noldan birga qadar bo'lgan oraliqda qiymatlar qabul qilishi tufayli sigmoid funksiyasi chiqishini ehtimollik sifatida talqin qilish mumkin. Bu xususiyat ikkilik tasniflash vazifalarida juda foydalidir. Masalan, spam filtrlarida, tibbiy diagnostikada yoki tasvirlarni tasniflashda sigmoid funksiya neyron tarmoqning oxirgi qatlamida qo'llanilib, ma'lum bir sinfga tegishlilik ehtimolligini beradi.

Sigmoid funksiyasining afzalliklaridan biri uning silliqligidir. Funksiya barcha nuqtalarda uzluksiz va farqlanuvchi bo'lib, gradient asosidagi optimallashtirish algoritmlari uchun qulay. Bundan tashqari, sigmoid funksiyaning chiqish qiymatlari normallashtirilgan oraliqda bo'lganligi sababli tarmoqning chiqish qatlamida qo'llanilganda natijalarni talqin qilishni osonlashtiradi.

Biroq, sigmoid funksiyasining bir qator jiddiy kamchiliklari mavjud. Birinchi va eng muhimi, yuqorida aytib o'tilgan yo'qolib borayotgan gradient muammosi. Chuqur neyron tarmoqlarda sigmoid funksiyalari yashirin qatlamlarda qo'llanilganda gradientlar orqaga tarqalish jarayonida tez sur'atda kichrayadi. Bu dastlabki qatlamlarning og'irliklarini yangilashni sekinlashtiradi yoki butunlay to'xtatib qo'yadi. Natijada tarmoq samarali o'rganolmaydi va yaxshi natijalar bermaydi.

Ikkinchi kamchilik sigmoid funksiyasining nosimmetrikligi va noldan farqli o'rtacha qiymatga ega ekanligi. Sigmoid funksiyaning chiqish qiymatlari doimo musbat bo'lib, yarim atrofida markazlashgan. Bu esa keyingi qatlamlarga o'tadigan signallarning doimo bir xil ishorali bo'lishiga olib keladi. Natijada og'irliklarning yangilanishi bir tomonlama bo'ladi va o'qitish jarayoni sekinlashadi. Bu hodisa zigzag harakati deb ataladi va optimallashtirish jarayonini samarasiz qiladi.

Uchinchi muammo sigmoid funksiyasining hisoblash murakkabligi. Eksponentsial funksiyani hisoblash sodda arifmetik amallardan ko'ra ko'proq vaqt va resurs talab qiladi. Zamonaviy chuqur tarmoqlarda millionlab neyronlar bo'lishi mumkin va har bir neyron uchun sigmoid funksiyani hisoblash umumiy hisoblash vaqtining sezilarli qismini tashkil qiladi. Bu ayniqsa resurs cheklangan qurilmalarda, masalan, mobil telefonlar yoki kiritilgan tizimlarda muhim ahamiyatga ega.

To'rtinchi kamchilik to'yinganlik muammosi deb ataladi. Sigmoid funksiyning kirish qiymati juda katta yoki juda kichik bo'lganda funksiya to'yinish zonasiga kiradi. Bu zonalarda funksiyaning hosilasi nolga yaqin bo'lib, gradientlar juda kichik bo'ladi. Natijada og'irliklar deyarli yangilanmaydi va o'qitish samaradorligi keskin pasayadi. Bu muammo ayniqsa noto'g'ri initsializatsiya qilingan og'irliklar bilan yuzaga keladi.

Ushbu kamchiliklarga qaramay, sigmoid funksiya hali ham ba'zi vazifalar uchun foydali bo'lib qolmoqda. Ayniqsa, ikkilik tasniflash muammolarida va neyron tarmoqning chiqish qatlamida sigmoid funksiya keng qo'llaniladi. Bundan tashqari, sayoz neyron tarmoqlarda yoki qatlamlar soni kam bo'lgan arxitekturalarda sigmoid funksiyaning kamchiliklari unchalik sezilarli emas.

Zamonaviy tadqiqotlarda sigmoid funksiyaning kamchiliklarini bartaraf etish uchun turli modifikatsiyalar taklif etilgan. Masalan, sof sigmoid funksiya o'rniga uni boshqa funksiyalar bilan kombinatsiyalash yoki adaptiv versiyalarini qo'llash. Biroq, chuqur o'rganish sohasida sigmoid funksiya asosan boshqa zamonaviy faollashtirish funksiyalari, xususan to'g'rilangan chiziqli birlik bilan almashtirildi.

Xulosa qilib aytganda, sigmoid funksiya neyron tarmoqlar tarixida muhim rol o'ynagan va hali ham ma'lum vazifalar uchun qo'llaniladi. Uning ehtimollik talqini va silliq xususiyatlari afzallik hisoblanadi. Biroq, yo'qolib borayotgan gradient muammosi, hisoblash murakkabligi va to'yinganlik muammolari tufayli zamonaviy chuqur tarmoqlarda sigmoid funksiya yashirin qatlamlarda kamdan kam qo'llaniladi va asosan chiqish qatlamida ikkilik tasniflash uchun ishlatiladi.

GIPERBOLIK TANGENS: SIMMETRIK ALTERNATIVA

Giperbolik tangens funksiyasi sigmoid funksiyaning takomillashtirilgan versiyasi sifatida qaralishi mumkin. Bu funksiya sigmoid bilan ko'plab umumiy xususiyatlarga ega bo'lsa ham, ba'zi muhim farqlar tufayli amaliyotda ko'pincha afzalroq hisoblanadi. Giperbolik tangens ham S shaklidagi egri chiziqni hosil qiladi, lekin uning qiymatlari oralig'i minus birdan plyus birgacha o'zgaradi.

Giperbolik tangens funksiyasining matematik ta'rifi giperbolik sinus va giperbolik kosinusning nisbati sifatida berilishi mumkin. Shuningdek, u eksponentsial funksiyalar orqali ham ifodalanishi mumkin. Funksiyaning grafigi koordinatalar boshiga nisbatan simmetrik bo'lib, kirish qiymati nolga teng bo'lganda funksiya ham nolga teng bo'ladi. Bu muhim farq sigmoid funksiyaga nisbatan sezilarli afzallik yaratadi.

Giperbolik tangensning hosilasi ham o'zi orqali sodda ifoda bilan ifodalanadi. Bu xususiyat orqaga tarqalish algoritmini amalga oshirishda hisoblashlarni osonlashtiradi. Hosilaning maksimal qiymati birga teng va kirish nolga teng bo'lganda erishiladi. Bu sigmoid funksiyaga nisbatan to'rt marta katta qiymatdir va gradientlarning tezroq tarqalishiga yordam beradi.

Giperbolik tangensning koordinatalar boshiga nisbatan simmetrik ekanligi uning asosiy afzalliklaridan biridir. Funksiyaning o'rtacha qiymati nolga teng bo'lib, bu sigmoid funksiyada mavjud bo'lgan nol bo'lmagan o'rtacha qiymat muammosini hal qiladi. Natijada, giperbolik tangens ishlatilgan tarmoqlarda signallar muvozanatlashgan tarzda tarqaladi va og'irliklar yangilanishi samarali bo'ladi.

Giperbolik tangensning minus birdan plyus birgacha qiymatlar oralig'i ham foydalidir. Bu oraliq sigmoid funksiyaning noldan birgacha oralig'idan kengroq bo'lib, tarmoqqa kattaroq dinamik diapazon beradi. Salbiy va musbat qiymatlarning mavjudligi tarmoqqa xususiyatlarni yaxshiroq ifodalash imkoniyatini beradi va nochiziqli munosabatlarni modellashtirish qobiliyatini oshiradi.

Giperbolik tangensning hosilasi sigmoid funksiyaga nisbatan kattaroq maksimal qiymatga ega ekanligi gradient oqimini yaxshilaydi. Biroq, bu funksiya ham to'yinganlik muammosidan aziyat chekadi. Kirish qiymatlari katta yoki kichik bo'lganda hosila nolga yaqinlashadi va yo'qolib borayotgan gradient muammosi yuzaga keladi. Lekin, simmetriyasi va kengroq dinamik diapazoni tufayli giperbolik tangens bu muammolarni sigmoid funksiyaga nisbatan yaxshiroq boshqaradi.

Amaliy jihatdan, giperbolik tangens ko'plab vazifalar uchun sigmoid funksiyaga nisbatan afzalroq tanlov hisoblanadi. Tabiiy tilni qayta ishlashda, vaqt qatorlari tahlilida va takrorlanuvchi neyron tarmoqlarda giperbolik tangens keng qo'llaniladi. Masalan, uzoq qisqa muddatli xotira tarmoqlarida giperbolik tangens asosiy faollashtirish funksiyasi sifatida ishlatiladi va bu arxitekturalarning muvaffaqiyatida muhim rol o'ynaydi.

Giperbolik tangensning hisoblash murakkabligi sigmoid funksiya bilan deyarli bir xildir. Ikkala funksiya ham eksponentsial hisoblarni talab qiladi va zamonaviy apparatda ularning bajarilish tezligi taxminan bir xil. Biroq, simmetriyasi va gradient oqimini yaxshilash xususiyatlari tufayli giperbolik tangens ko'pincha o'qitish vaqtini qisqartiradi va yaxshiroq yaqinlashishga erishadi.

Giperbolik tangensning kamchiliklari sigmoid funksiya kamchiliklariga o'xshaydi. To'yinganlik muammosi, yo'qolib borayotgan gradient va hisoblash murakkabligi hali ham mavjud. Chuqur tarmoqlarda bu muammolar ayniqsa sezilarli

XULOSA

Faollashtirish funksiyalari neyron tarmoqlarning muhim tarkibiy qismi bo'lib, ularning ishlashi va samaradorligiga bevosita ta'sir ko'rsatadi. Ushbu tadqiqotda sigmoid, giperbolik tangens va to'g'rilangan chiziqli birlik funksiyalarining keng qamrovli tahlili amalga oshirildi. Har bir funksiyaning o'ziga xos afzalliklari, kamchiliklari va qo'llanilish sohalari aniqlandi.

Sigmoid funksiyasi neyron tarmoqlar tarixida muhim o'rin tutadi va ikkilik tasniflash vazifalarida hali ham qo'llaniladi. Uning ehtimollik talqini va silliq xususiyatlari qimmatlidir. Biroq, yo'qolib borayotgan gradient muammosi, to'yinganlik va hisoblash murakkabligi tufayli zamonavir chuqur tarmoqlarda uning qo'llanilishi cheklangan. Sigmoid asosan neyron tarmoqning chiqish qatlamida ikkilik tasniflash uchun ishlatiladi.

Giperbolik tangens sigmoid funksiyaga nisbatan yaxshilangan alternativa hisoblanadi. Uning simmetriyasi, kengroq dinamik diapazoni va yaxshiroq gradient oqimi ko'plab vazifalar uchun afzallik yaratadi. Takrorlanuvchi neyron tarmoqlarda, ayniqsa uzoq qisqa muddatli xotira arxitekturalarida giperbolik tangens keng qo'llaniladi. Biroq, u ham to'yinganlik va yo'qolib borayotgan gradient muammolaridan aziyat chekadi.

To'g'rilangan chiziqli birlik zamonaviy chuqur o'rganishning asosiy faollashtirish funksiyasiga aylandi. Uning soddaligi, hisoblash samaradorligi va yo'qolib borayotgan gradient muammosini hal qilish qobiliyati uni eng mashhur tanlovga aylantirdi. ReLU konvolyutsion neyron tarmoqlarda, tasvirlarni qayta ishlashda va ko'plab boshqa vazifalarida ustun natijalar ko'rsatadi. O'lik neyronlar muammosiga qaramay, ReLU zamonavir arxitekturalarda asosiy standart hisoblanadi.

FOYDALANILGAN ADABIYOTLAR:

1. Goodfellow I, Bengio Y, Courville A. Deep Learning. MIT Press, Cambridge, Massachusetts, 2016. 775 p.

2. LeCun Y, Bengio Y, Hinton G. Deep learning. Nature. 2015; 521(7553): 436-444.

3. Krizhevsky A, Sutskever I, Hinton GE. ImageNet classification with deep convolutional neural networks. Advances in Neural Information Processing Systems. 2012; 25: 1097-1105.

4. Glorot X, Bordes A, Bengio Y. Deep sparse rectifier neural networks. Proceedings of the Fourteenth International Conference on Artificial Intelligence and Statistics. 2011; 15: 315-323.

5. He K, Zhang X, Ren S, Sun J. Delving deep into rectifiers: Surpassing human-level performance on ImageNet classification. Proceedings of the IEEE International Conference on Computer Vision. 2015; 1026-1034.

6. Nair V, Hinton GE. Rectified linear units improve restricted Boltzmann machines. Proceedings of the Twenty-seventh International Conference on Machine Learning. 2010; 807-814.

7. Maas AL, Hannun AY, Ng AY. Rectifier nonlinearities improve neural network acoustic models. Proceedings of the International Conference on Machine Learning. 2013; 30(1): 3.

8. He K, Zhang X, Ren S, Sun J. Deep residual learning for image recognition. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. 2016; 770-778.

9. Clevert DA, Unterthiner T, Hochreiter S. Fast and accurate deep network learning by exponential linear units. International Conference on Learning Representations. 2016.

10. Klambauer G, Unterthiner T, Mayr A, Hochreiter S. Self-normalizing neural networks. Advances in Neural Information Processing Systems. 2017; 30: 971-980.

11. Ramachandran P, Zoph B, Le QV. Searching for activation functions. arXiv preprint arXiv:1710.05941. 2017.