

XATOLIK FUNKSIYALARI MASHINANI O'QITISH TIZIMLARIDA: MSE, CROSS-ENTROPY VA ZAMONAVIY YONDASHUVLAR

Tojimamatov Israil Nurmatovich

*Farg'ona davlat universiteti Amaliy matematika
va informatika kafedrası katta o'qituvchisi*

E-mail: israiltojimatov@gmail.com

Mirzayev Asadbek Marifjon o'g'li

*Farg'ona davlat universiteti Amaliy matematika
yo'nalishi 3-bosqich 23.10-guruh talabasi*

E-mail: mirzayevasadbek7177@gmail.com

ANNOTATSIYA: Ushbu maqola zamonaviy mashinani o'qitish tizimlarida va chuqur neyron tarmoqlarida xatolik funksiyalarining keng qamrovli tadqiqotini taqdim etadi. Xatolik funksiyalari o'qitish jarayonining asosiy komponenti bo'lib, model parametrlarini optimallashtirish yo'nalishini belgilaydi. Ishda klassik xatolik funksiyalari, jumladan, o'rtacha kvadratik xatolik va kross-entropiya, shuningdek, turli mashinani o'qitish vazifalar uchun zamonaviy modifikatsiyalar va ixtisoslashtirilgan funksiyalar batafsil ko'rib chiqiladi. Xatolik funksiyalarining matematik xususiyatlariga, ularning optimallashtirish jarayoniga ta'siriga va muayyan vazifalar uchun mos funksiyani tanlashning amaliy jihatlariga alohida e'tibor qaratiladi. Turli xatolik funksiyalarining raqamli barqarorligi, chetga chiqishlarga sezgirligi va yaqinlashish tezligi masalalari o'rganiladi. Regressiya, tasniflash, segmentatsiya va generativ modellash vazifalarida turli funksiyalarning samaradorligiga qiyosiy tahlil taqdim etiladi. Ish shuningdek, adaptiv xatolik funksiyalari va ko'p vazifali o'qitishni o'z ichiga olgan zamonaviy tendentsiyalarni qamrab oladi.

Kalit so'zlar: xatolik funksiyalari, o'rtacha kvadratik xatolik, kross-entropiya, mashinani o'qitish, chuqur neyron tarmoqlari, optimallashtirish, gradient tushish, regressiya, tasniflash.

АННОТАЦИЯ: Данная статья представляет собой комплексное исследование функций потерь в современных системах машинного обучения и глубоких нейронных сетях. Функции потерь являются фундаментальным компонентом процесса обучения, определяющим направление оптимизации параметров модели. В работе детально рассматриваются классические функции потерь, включая среднеквадратичную ошибку и кросс-энтропию, а также современные модификации и специализированные функции для различных задач машинного обучения. Особое внимание уделяется математическим свойствам функций потерь, их влиянию на процесс оптимизации и практическим аспектам выбора подходящей функции для конкретных задач. Исследуются

вопросы численной стабильности, чувствительности к выбросам и скорости сходимости различных функций потерь. Представлен сравнительный анализ эффективности различных функций в задачах регрессии, классификации, сегментации и генеративного моделирования. Работа также охватывает современные тенденции, включая адаптивные функции потерь и многозадачное обучение.

Ключевые слова: функции потерь, среднеквадратичная ошибка, кросс-энтропия, машинное обучение, глубокие нейронные сети, оптимизация, градиентный спуск, регрессия, классификация

ANNOTATION: This article presents a comprehensive investigation of loss functions in modern machine learning systems and deep neural networks. Loss functions constitute a fundamental component of the training process, determining the direction of model parameter optimization. The work examines in detail classical loss functions, including mean squared error and cross-entropy, as well as modern modifications and specialized functions for various machine learning tasks. Particular attention is devoted to mathematical properties of loss functions, their influence on the optimization process, and practical aspects of selecting appropriate functions for specific tasks. Issues of numerical stability, sensitivity to outliers, and convergence rates of different loss functions are investigated. A comparative analysis of the effectiveness of various functions in regression, classification, segmentation, and generative modeling tasks is presented. The work also covers contemporary trends, including adaptive loss functions and multi-task learning.

Keywords: loss functions, mean squared error, cross-entropy, machine learning, deep neural networks, optimization, gradient descent, regression, classification

KIRISH

Mashinani o'qitish va sun'iy intellekt sohasida xatolik funksiyalari modelning ishlash sifatini baholash va uni yaxshilash yo'nalishini belgilovchi asosiy mexanizm hisoblanadi. Xatolik funksiyasi modelning bashorati va haqiqiy natijalar o'rtasidagi farqni miqdoriy jihatdan ifodalaydi va optimallashtirish jarayonida minimallashtirilishi kerak bo'lgan maqsad funksiyasini aniqlaydi. To'g'ri tanlangan xatolik funksiyasi modelning samarali o'qitishini ta'minlaydi, noto'g'ri tanlov esa yomon natijalarga yoki hatto modelning umuman o'rganmasligiga olib kelishi mumkin.

Xatolik funksiyalarining tarixi mashinani o'qitishning dastlabki kunlariga borib taqaladi. O'rtacha kvadratik xatolik statistika va optimallashtirishning klassik usullaridan kelib chiqqan bo'lib, eng kichik kvadratlar usuli bilan chambarchas bog'liq. Kross-entropiya esa axborot nazariyasidan kelib chiqqan va tasniflash masalalarida keng qo'llaniladi. Zamonaviy chuqur o'qitish davri yanada murakkab va ixtisoslashtirilgan xatolik funksiyalarini yaratishga turtki berdi.

Xatolik funksiyasini tanlashda bir qator muhim omillarni hisobga olish zarur. Birinchidan, muammoning tabiati muhim - regressiya masalalari uchun bir turdagi funksiyalar, tasniflash uchun boshqa turlar mos keladi. Ikkinchidan, ma'lumotlarning xususiyatlari, masalan, shovqin darajasi, chetga chiqishlarning mavjudligi va sinflarning muvozanati ta'sir qiladi. Uchinchidan, optimallashtirish jarayonining xususiyatlari, masalan, gradientlarning xatti-harakati va yaqinlashish tezligi muhimdir.

Zamonaviy chuqur o'qitish tizimlarida xatolik funksiyalari nafaqat modelni o'qitish uchun, balki model sifatini baholash, turli arxitekturalarni solishtirish va giperparametrlarni sozlash uchun ham ishlatiladi. Xatolik funksiyasining qiymati o'qitish jarayonini monitoring qilish va to'xtatish mezonlarini aniqlash uchun muhim ko'rsatkichdir. Shuningdek, u regularizatsiya texnikalari va boshqa yordamchi komponentlar bilan birgalikda ishlatiladi.

Xatolik funksiyalarining nazariy asoslari matematikaning turli sohalarini, jumladan, optimallashtirishni, ehtimollik nazariyasini va axborot nazariyasini o'z ichiga oladi. Konvekslik, differentsiallanuvchanlik va Lipschitz uzluksizligi kabi matematik xususiyatlar xatolik funksiyasining optimallashtirish algoritmlari bilan qanchalik yaxshi ishlashini belgilaydi. Bu xususiyatlarni tushunish to'g'ri funksiyani tanlash va uni muvaffaqiyatli qo'llash uchun zarur.

Hozirgi zamon tadqiqotlari yanada murakkab va moslashuvchan xatolik funksiyalarini yaratishga qaratilgan. Adaptiv xatolik funksiyalari o'qitish jarayonida o'zlarini avtomatik ravishda sozlaydi. Ko'p vazifali o'qitishda bir nechta xatolik funksiyalari birgalikda qo'llaniladi va ularning nisbati muhim ahamatga ega. Meta-learning yondashuvlari xatolik funksiyasining o'zini ham o'rgatishga harakat qiladi.

Ushbu maqolaning maqsadi xatolik funksiyalarining keng qamrovli tahlilini taqdim etish, klassik va zamonaviy yondashuvlarni solishtirish va amaliyotchilar uchun foydali tavsiyalar berishdir. Maqola tuzilishi quyidagicha: ikkinchi bo'limda o'rtacha kvadratik xatolik va uning variantlari ko'rib chiqiladi, uchinchi bo'limda kross-entropiya va tasniflash uchun boshqa funksiyalar tahlil qilinadi, to'rtinchi bo'limda maxsus vazifalar uchun ixtisoslashtirilgan funksiyalar muhokama qilinadi, beshinchi bo'limda zamonaviy tendentsiyalar va kelajak yo'nalishlari ko'rsatiladi.

O'rtacha kvadratik xatolik mashinani o'qitishda eng ko'p qo'llaniladigan xatolik funksiyalaridan biri bo'lib, ayniqsa regressiya masalalarida keng tarqalgan. Ushbu funksiya bashorat qilingan va haqiqiy qiymatlar o'rtasidagi farqni kvadratga ko'taradi va barcha namunalar bo'yicha o'rtachani hisoblaydi. Kvadratga ko'tarish katta xatoliklarga katta jazolar berishni ta'minlaydi, bu esa modelni barcha namunalar uchun yaxshi ishlashga undaydi.

O'rtacha kvadratik xatolikning asosiy afzalligi uning matematik xususiyatlaridir. Funksiya silliq va hamma joyda differentsiallanuvchi, bu gradientni hisoblashni osonlashtiradi. Funksiya konveks bo'lganligi sababli, lokal minimum global minimum

bilan mos tushadi, bu esa optimallashtirishni sodda qiladi. Gradientlar bashoratlardan haqiqiy qiymatlarga chiziqli bog'liq, bu esa optimallashtirish algoritmlarining barqaror ishlashini ta'minlaydi.

Biroq, o'rtacha kvadratik xatolik ba'zi cheklovlariga ega. Eng muhim cheklov - chetga chiqishlarga yuqori sezgirlik. Agar ma'lumotlarda bir nechta juda katta xatoliklar mavjud bo'lsa, ular xatolik funksiyasi qiymatiga katta ta'sir qiladi va model bu chetga chiqishlarga moslanishga harakat qiladi, natijada umumiy ishlash yomonlashadi. Kvadratga ko'tarish bu muammoni yanada kuchaytiradi, chunki katta xatoliklar eksponent ravishda o'sadi.

O'rtacha absolut xatolik o'rtacha kvadratik xatolikka alternativa bo'lib, farqning absolut qiymatini oladi. Bu funksiya chetga chiqishlarga kamroq sezgir, chunki xatoliklar chiziqli ravishda jazalanadi. Biroq, o'rtacha absolut xatolik nolda differentsiallanmaydi, bu esa ba'zi optimallashtirish algoritmlari uchun muammo tug'dirishi mumkin. Amalda subgradient usullari yoki silliq yaqinlashtirishlar qo'llaniladi.

Huber xatoligi o'rtacha kvadratik va o'rtacha absolut xatoliklarning afzalliklarini birlashtiradi. Kichik xatoliklar uchun u kvadratik, katta xatoliklar uchun esa chiziqli xatti-harakatga ega. Bu gibridd yondashuv chetga chiqishlarga kam sezgirlikni ta'minlaydi va bir vaqtning o'zida barcha joylarda differentsiallanadi. Chegara qiymati giperparametr sifatida sozlanadi va muammoning xususiyatlariga bog'liq.

Quantile regression xatoligi bashoratlarning turli kvantillarini modellashtirish imkonini beradi. Oddiy regressiya shartli o'rtachani bashorat qiladi, quantile regression esa shartli medianani yoki boshqa kvantillarni bashorat qilishi mumkin. Bu noaniqlikni baholash va bashoratlarni oralig'ini yaratish uchun foydali. Asymmetric xususiyat turli yo'nalishdagi xatoliklarga turlicha jazolar berishni ta'minlaydi.

Log-cosh xatoligi o'rtacha kvadratik xatolikka o'xshash, lekin chetga chiqishlarga kamroq sezgir. U kichik xatoliklar uchun taxminan kvadratik, katta xatoliklar uchun esa chiziqli. Funksiya hamma joyda ikki marta differentsiallanadi, bu esa ikkinchi tartibli optimallashtirish usullarini qo'llash imkonini beradi. Amalda, log-cosh o'rtacha kvadratik va Huber o'rtasidagi muvozanatni ta'minlaydi.

Regressiya masalalarida xatolik funksiyasini tanlashda ma'lumotlarning xususiyatlarini hisobga olish muhimdir. Agar ma'lumotlarda ko'p chetga chiqishlar bo'lsa, o'rtacha absolut yoki Huber xatoligini tanlash maqsadga muvofiq. Agar ma'lumotlar toza va Gauss taqsimotiga yaqin bo'lsa, o'rtacha kvadratik xatolik yaxshi natija beradi. Noaniqlikni baholash zarur bo'lsa, quantile regression qo'llaniladi.

Xatolik funksiyasining shkala muammosi ham muhim. Agar maqsad o'zgaruvchining qiymatlari juda katta yoki juda kichik bo'lsa, xatolik funksiyasi qiymatlari ham shu shkala bo'yicha o'zgaradi. Bu optimallashtirish uchun muammo tug'dirishi mumkin. Normalizatsiya yoki standartlashtirish, yoki nisbiy xatolik funksiyalaridan foydalanish bu muammoni hal qiladi.

Ko'p o'lchovli regressiyada har bir o'lchov uchun alohida xatolik hisoblash va ularni yig'ish mumkin. Yig'indini olish eng oddiy yondashuv bo'lsa-da, ba'zan og'irlikli yig'indi yoki boshqa agregatsiya usullari qo'llaniladi. Agar turli o'lchovlar turli ahamiyatga ega bo'lsa yoki turli shkalada bo'lsa, og'irlikli yondashuv zarur.

Kross-entropiya xatoligi tasniflash masalalarida eng keng qo'llaniladigan funksiya bo'lib, axborot nazariyasidan kelib chiqqan. Ushbu funksiya modelning bashorat qilgan ehtimollik taqsimoti va haqiqiy taqsimot o'rtasidagi farqni o'lchaydi. Kross-entropiya maksimum ehtimollik printsipligiga asoslangan bo'lib, model haqiqiy sinfga yuqori ehtimollik berganida xatolik kichik bo'ladi.

Binary kross-entropiya ikki sinfli tasniflash uchun ishlatiladi. Har bir namuna uchun model bitta sinfga tegishli bo'lish ehtimolligini bashorat qiladi. Agar haqiqiy sinf birinchi sinf bo'lsa va model bunga yuqori ehtimollik bergan bo'lsa, xatolik kichik bo'ladi. Agar model noto'g'ri bashorat qilsa, xatolik katta bo'ladi. Logarifm ishlatilishi nolga yaqin ehtimolliklar uchun katta jazolarni ta'minlaydi.

Categorical kross-entropiya ko'p sinfli tasniflash uchun binary kross-entropiyaning umumlashtirilishi hisoblanadi. Model har bir sinf uchun ehtimollik beradi va ehtimolliklar yig'indisi birga teng bo'lishi kerak. Softmax aktivatsiya funksiyasi odatda oxirgi qatlamda qo'llaniladi va bu ehtimolliklar taqsimotini ta'minlaydi. Kross-entropiya haqiqiy sinf uchun logarifmik ehtimollikni o'lchaydi.

Sparse categorical kross-entropiya categorical kross-entropiya bilan bir xil, lekin kirish formatida farq qiladi. Oddiy categorical kross-entropiyada haqiqiy yorliqlar one-hot encoded bo'lishi kerak, sparse versiyasida esa oddiy butun sonlar ishlatiladi. Bu xotiradan tejankor foydalanishni ta'minlaydi, ayniqsa ko'p sinfli masalalarda. Matematik jihatdan ikkala yondashuv bir xil natija beradi.

Focal loss imbalanced dataset muammosini hal qilish uchun yaratilgan. Oddiy kross-entropiyada oson namunalar ham xatolikka hissa qo'shadi va ko'p bo'lgan sinflar uchun bu hissa katta bo'ladi. Focal loss oson namunalarning ta'sirini kamaytiradi va qiyin namunalarga ko'proq e'tibor beradi. Modulating factor yordamida to'g'ri tasniflangan namunalarning og'irligi kamaytiriladi.

Label smoothing xatoligi modelning haddan tashqari ishonchini kamaytirish uchun ishlatiladi. Oddiy tasniflashda haqiqiy yorliqlar bir va nol qiymatlarga ega. Label smoothing bu qat'iy qiymatlarni yumshatadi, masalan, bir o'rniga to'qson to'qqiz foiz va nol o'rniga bir foiz ishlatiladi. Bu modelning regularizatsiyasi bo'lib, overfitting ni kamaytiradi va umumlashtirish qobiliyatini oshiradi.

Weighted kross-entropiya sinflarning imbalance muammosini hal qilish uchun qo'llaniladi. Har bir sinf uchun og'irlik belgilanadi va xatolik hisoblashda bu og'irliklar qo'llaniladi. Kam uchraydigan sinflar uchun yuqori og'irliklar beriladi, ko'p uchraydigan sinflar uchun past og'irliklar ishlatiladi. Bu model kam namunali sinflarga ko'proq e'tibor berishini ta'minlaydi.

Hinge loss support vector machines bilan bog'liq bo'lib, binary tasniflash uchun ishlatiladi. Bu funksiya to'g'ri tasniflangan va chegara-dan uzoq namunalar uchun nol xatolik beradi. Chegara yaqinidagi yoki noto'g'ri tasniflangan namunalar uchun xatolik chiziqli ravishda o'sadi. Hinge loss maksimal margin prinsipiga asoslangan va robust classifier larni yaratadi.

Triplet loss metric learning va yuz tanish kabi vazifalar uchun ishlatiladi. Bu funksiya uchta namuna bilan ishlaydi: anchor, positive va negative. Maqsad anchor va positive orasidagi masofani kamaytirish va anchor va negative orasidagi masofani oshirishdir. Margin parametr ikkalasi o'rtasidagi minimal farqni belgilaydi. Bu yondashuv samarali embedding larni o'rgatadi.

Contrastive loss ham metric learning da qo'llaniladi va juft namunalar bilan ishlaydi. Agar juft bir sinfdan bo'lsa, ularning embedding lari yaqin bo'lishi kerak. Agar turli sinflardan bo'lsa, ularning embedding lari uzoq bo'lishi kerak. Margin parametr minimal ajralish masofasini belgilaydi. Bu usul siamese network larda keng qo'llaniladi.

Zamonaviy mashinani o'qitishda ko'plab maxsus vazifalar mavjud bo'lib, ular o'ziga xos xatolik funksiyalarini talab qiladi. Tasvirni segmentatsiya qilish, ob'ektlarni aniqlash, generativ modellash va boshqa murakkab vazifalar uchun standart xatolik funksiyalari yetarli emas. Tadqiqotchilar bu vazifalar uchun ixtisoslashgan funksiyalar yaratdilar va ular ancha yaxshi natijalar beradi.

Dice loss tasvirlarni segmentatsiya qilish uchun keng qo'llaniladigan funksiya hisoblanadi. Dice koeffitsienti bashorat qilingan va haqiqiy segmentatsiya maskalari o'rtasidagi o'xshashlikni o'lchaydi. Funksiya piksellar ustidan kesishma va birlashma nisbatini hisobga oladi. Dice loss imbalanced segmentatsiya masalalarida binary kross-entropiyadan yaxshiroq ishlaydi, chunki u global overlap ga e'tibor beradi.

IoU loss ham segmentatsiya uchun ishlatiladi va Intersection over Union metrikasiga asoslangan. Bu metrika ob'ektlarni aniqlash va segmentatsiya qilishda standart baholash ko'rsatkichidir. IoU loss ni to'g'ridan-to'g'ri optimallashtirish metrikaga moslashtirilgan bashoratlarni yaratadi. Biroq, IoU differentsiallanuvchi emas, shuning uchun smooth yaqinlashtirishlar qo'llaniladi.

Tversky loss Dice loss ning umumlashirilishi bo'lib, false positive va false negative xatolariga turli og'irliklar berish imkonini beradi. Parametrlar orqali bu xatolarning nisbatini sozlash mumkin. Tibbiy tasvirlarda, masalan, false negative xatolar false positive lardan ko'ra xavfli bo'lishi mumkin. Tversky loss bunday vazifalar uchun moslashtirilgan sozlamalarni taqdim etadi.

Perceptual loss generativ modellarda va uslub o'tkazishda ishlatiladi. Oddiy piksel-wise xatoliklar o'rniga, bu funksiya tasvirlarning yuqori darajadagi xususiyatlarini solishtiradi. Oldindan o'rgatilgan neyron tarmoq, masalan VGG, xususiyatlar ekstraktori sifatida ishlatiladi. Bashorat qilingan va haqiqiy tasvirlarning

xususiyatlari orasidagi farq hisoblanadi. Bu perceptual jihatdan sifatli tasvirlarni yaratadi.

Adversarial loss generativ adversarial network larda asosiy komponent hisoblanadi. Generator tasvirlar yaratadi va discriminator ularni haqiqiy yoki soxta deb tasniflashga harakat qiladi. Generator discriminator ni aldashga harakat qiladi. Bu ikki tarafa o'yini min-max optimallashtirish muammosiga aylanadi. Adversarial loss haqiqiy ma'lumotlar taqsimotini o'rganishda juda samarali.

Style loss uslub o'tkazish vazifalarida qo'llaniladi va tasvir uslubini aniqlaydi. Gram matritsa xususiyatlar orasidagi korrelyatsiyani o'lchaydi va uslubni ifodalaydi. Bashorat qilingan tasvir gram matritsasi reference uslub gram matritsasiga yaqinlashtiriladi. Content loss bilan birgalikda, bu kombinatsiya content ni saqlab, uslubni o'tkazadi.

Reconstruction loss autoencoders da ishlatiladi va kirishni qayta tiklash qobiliyatini o'lchaydi. Encoder kirish ma'lumotlarini past o'lchovli latent space ga siqadi va decoder uni qayta tiklaydi. Xatolik funksiyasi asl va qayta tiklangan ma'lumotlar o'rtasidagi farqni o'lchaydi. Odatda MSE yoki binary kross-entropiya ishlatiladi, ma'lumot turiga qarab.

KL divergence variational autoencoder da regularizatsiya komponenti bo'lib, latent space taqsimotini normal taqsimotga yaqinlashtiradi. Bu posterior va prior taqsimotlari o'rtasidagi farqni o'lchaydi. KL divergence bilan reconstruction loss ni muvozanatlash muhim va bu beta-VAE kabi variantlarda sozlanadi. To'g'ri muvozanat disentangled representation larni o'rgatishga yordam beradi.

CTC loss ketma-ketlik labeling vazifalarida, masalan, nutqni tanish va matn tanishda qo'llaniladi. Bu funksiya kirish va chiqish ketma-ketliklari turli uzunlikda bo'lishi mumkin bo'lgan vazifalar uchun mo'ljallangan. CTC barcha mumkin bo'lgan alignment larni hisobga oladi va ularning ehtimolliklarini yig'adi. Bu alignment ni aniq belgilash zaruratini bartaraf etadi.

Attention-based loss transformer arxitekturalarida attention mexanizmini o'rgatish uchun qo'shimcha komponent bo'lishi mumkin. Attention weights ning ma'lum xususiyatlarga ega bo'lishini ta'minlash uchun qo'shimcha jazolar qo'shiladi. Masalan, sparse attention yoki local attention ni rag'batlantirish mumkin. Bu xatolik asosiy tasniflash yoki regressiya xatoligi bilan birgalikda ishlatiladi.

Xatolik funksiyalari sohasidagi tadqiqotlar doimiy ravishda rivojlanmoqda va yangi yondashuvlar paydo bo'lmoqda. Zamonaviy tendentsiyalar adaptiv funksiyalar, meta-learning va ko'p vazifali o'qitishni o'z ichiga oladi. Bu yo'nalishlar yanada samarali va moslashuvchan o'qitish tizimlarini yaratishga qaratilgan.

Adaptiv xatolik funksiyalari o'qitish jarayonida o'zlarini avtomatik ravishda sozlaydi. Odatda xatolik funksiyasi va uning parametrlari oldindan belgilanadi va o'qitish davomida o'zgarmaydi. Adaptiv yondashuvda esa funksiya parametrlari yoki

hatto funksiyaning o'zi o'qitish jarayonida o'rganiladi. Masalan, Huber xatoligidagi chegara qiymati dinamik ravishda sozlanishi mumkin.

Curriculum learning bosqichma-bosqich murakkablashtirish strategiyasini qo'llaydi. Modelni avval oson namunalarda o'rgatish va asta-sekin qiyinroq namunalarga o'tish. Xatolik funksiyasi namunalarning qiyinlik darajasiga qarab og'irlik berishi mumkin. Bu yondashuv ayniqsa murakkab vazifalar va katta ma'lumotlar to'plamlarida samarali. Model tobora murakkab pattern larni o'rganadi.

Self-paced learning curriculum learning ga o'xshash, lekin model o'zi qaysi namunalarni o'rganishni tanlaydi. Xatolik funksiyasi har bir namunaga dinamik og'irlik beradi va model avval oson deb hisoblagan namunalarni o'rganadi. Vaqt o'tishi bilan barcha namunalar o'rganiladi, ammo tartib model tomonidan belgilanadi. Bu avtomatik curriculum yaratadi.

Meta-learning yoki learning to learn xatolik funksiyasining o'zini ham o'rgatishga harakat qiladi. Umumiy task-agnostic xatolik funksiyalari o'rniga, ma'lum bir vazifalar oilasi uchun optimallashtirilgan funksiya o'rganiladi. Bu yuqori darajadagi optimallashtirish muammosi bo'lib, meta-model past darajadagi modelni o'rgatish uchun eng yaxshi xatolik funksiyasini topadi.

Xulosa

Ushbu keng qamrovli tadqiqot mashinani o'qitish tizimlaridagi xatolik funksiyalarining fundamental rolini va ular optimallashtirish jarayonining mohiyatini qanday belgilashini tasdiqladi. O'rtacha kvadratik xatolik (MSE) regressiya vazifalarida konveksligi va differentsiallanuvchanligi tufayli asosiy vosita bo'lib xizmat qilsa-da, chetga chiqishlarga yuqori sezgirlik kabi cheklavlarga ega. Ushbu muammolar Huber va Log-cosh kabi robust funksiyalar orqali hal etiladi.

Tasniflash masalalarida kross-entropiya xatoligi, axborot nazariyasi asosida, bashorat qilingan va haqiqiy ehtimollik taqsimotlari o'rtasidagi farqni samarali o'lchaydi. Uning Focal loss kabi modifikatsiyalari sinflar nomutanosibligi kabi murakkab amaliy muammolarni hal qilish uchun yaratilgan.

Tahlil shuni ko'rsatadiki, zamonaviy chuqur o'rganish doirasida ixtisoslashtirilgan vazifalar, masalan, segmentatsiya uchun Dice loss va generativ modellar uchun Perceptual yoki Adversarial loss kabi noyob funksiyalar standart yondashuvlarga qaraganda yuqori samaradorlikni ta'minlaydi.

Foydalanilgan Adabiyotlar

1. Goodfellow I, Bengio Y, Courville A. Deep Learning. Chuqur O'rganish. MIT Press tomonidan nashr etilgan fundamental darslik.
2. Bishop CM. Pattern Recognition and Machine Learning. Naqshlarni Tanib Olish va Mashinada O'rganish. Springer nashriyoti tomonidan chiqarilgan klassik manba.

3. Ng AY. Feature Selection, L-one regularization, and support vector machines. Xususiyatlarni Tanlash, L-bir regulyarizatsiya va Qo'llab-Quvvatlash Vektor Mashinalari. Mashinada O'rganish Jurnali.
4. Kingma DP, Welling M. Auto-Encoding Variational Bayes. Avto-Kodlash Variatsion Bayes. Xalqaro O'rganish Vakilliklari Konferensiyasi materiallari.
5. Ronneberger O, Fischer P, Brox T. U-Net: Convolutional Networks for Biomedical Image Segmentation. U-Net: Biomedikal Tasvir Segmentatsiyasi Uchun Konvolyutsion Tarmoqlar. Tibbiy Tasvirlar Bo'yicha Xalqaro Konferensiya materiallari.
6. Lin TY, Goyal P, Girshick R, He K, Dollár P. Focal Loss for Dense Object Detection. Zich Ob'ektni Aniqlash Uchun Focal Loss. Xalqaro Kompyuter Ko'rish Konferensiyasi materiallari.
7. Reddi S J, Hande SS, Sreepathi S. A Comprehensive Study of Loss Functions for Deep Learning. Chuqur O'rganish Uchun Xatolik Funktsiyalarining Keng Qamrovli Tadqiqoti. Neyron Hisoblash Jurnali.
8. Hadsell R, Chopra S, LeCun Y. Dimensionality Reduction by Learning an Invariant Mapping. Invariant Xaritalashni O'rganish Orqali O'lchamni Qisqartirish. Kompyuter Ko'rish Konferensiyasi materiallari.
9. Hinton GE, Salakhutdinov R R. Reducing the Dimensionality of Data with Neural Networks. Neyron Tarmoqlar Yordamida Ma'lumotlar O'lchamini Qisqartirish. Science Jurnali.