

SUN'IY NEYRONNING MATEMATIK MODEL VA UNING ELEMENTLARI

Tojimamatov Israil Nurmatovich

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

E-mail: israiltojimatov@gmail.com

Iqboljonova Madinabonu Xayrullo qizi

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

E-mail: zayrullayevamadina@gmail.com

Annotatsiya: Ushbu maqola zamonaviy neyron tarmoqlari va sun'iy intellekt tizimlarining asosiy qurilish bloklari bo'lgan sun'iy neyronlarning matematik modellarini keng qamrovli tadqiq etadi. Ishda sun'iy neyronlarni yaratishga ilhom bergan biologik asoslar, shuningdek ularning ishlashining matematik tamoyillari batafsil ko'rib chiqiladi. Neyronning turli komponentlarini, jumladan kirish signallari, sinaptik og'irliklar, faollashtirish funksiyalari, chegara qiymatlari va chiqish signallarini tahlil qilish alohida e'tiborga olingan. Tadqiqot sun'iy neyron modellarining eng oddiy pertseptrondan tortib zamonaviy murakkab arxitekturalargacha bo'lgan evolyutsiyasini qamrab oladi. Turli xil faollashtirish funksiyalari, ularning matematik xususiyatlari va qo'llanish sohalari batafsil tahlil qilinadi. Neyronlarni o'rgatish usullari, jumladan xatolarni orqaga tarqatish va gradient tushish algoritmlari ko'rib chiqiladi. Maqolada shuningdek sun'iy neyronlar ishlashining hisoblash jihatlari, optimallashtirish va miqyoslash muammolari tadqiq etiladi.

Kalit so'zlar: sun'iy neyron, matematik model, faollashtirish funksiyasi, sinaptik og'irliklar, pertseptron, neyron tarmoqlar, mashina o'rganish, orqaga tarqatish, gradient tushish, biologik neyron

Аннотация: Данная статья представляет комплексное исследование математических моделей искусственных нейронов, которые являются фундаментальными строительными блоками современных нейронных сетей и систем искусственного интеллекта. В работе детально рассматриваются биологические основы, послужившие вдохновением для создания искусственных нейронов, а также математические принципы их функционирования. Особое внимание уделяется анализу различных компонентов нейрона, включая входные сигналы, синаптические веса, функции активации, пороговые значения и выходные сигналы. Исследование охватывает эволюцию моделей искусственных нейронов от простейшего перцептрона до

современных сложных архитектур. Подробно анализируются различные типы функций активации, их математические свойства и области применения. Рассматриваются методы обучения нейронов, включая алгоритмы обратного распространения ошибки и градиентного спуска. Статья также исследует вычислительные аспекты работы искусственных нейронов, проблемы оптимизации и масштабирования.

Ключевые слова: искусственный нейрон, математическая модель, функция активации, синаптические веса, перцептрон, нейронные сети, машинное обучение, обратное распространение, градиентный спуск, биологический нейрон

Annotation: This article presents a comprehensive study of mathematical models of artificial neurons, which are fundamental building blocks of modern neural networks and artificial intelligence systems. The work thoroughly examines the biological foundations that inspired the creation of artificial neurons, as well as the mathematical principles of their functioning. Special attention is given to analyzing various neuron components, including input signals, synaptic weights, activation functions, threshold values, and output signals. The research covers the evolution of artificial neuron models from the simplest perceptron to modern complex architectures. Various types of activation functions, their mathematical properties, and application areas are analyzed in detail. Methods of neuron training, including backpropagation algorithms and gradient descent, are examined. The article also explores computational aspects of artificial neuron operation, optimization problems, and scaling issues.

Keywords: artificial neuron, mathematical model, activation function, synaptic weights, perceptron, neural networks, machine learning, backpropagation, gradient descent, biological neuron

KIRISH

Sun'iy neyron zamonaviy sun'iy intellekt va mashina o'rganish tizimlarining asosiy qurilish bloki hisoblanadi. Biologik neyronlardan ilhomlangan bu matematik model insonning miyasidagi neyronlarning ishlash tamoyillarini soddalashtirilgan shaklda aks ettiradi. Sun'iy neyronlar turli xil vazifalarni hal qilishda, jumladan tasvirlarni aniqlash, nutqni tanish, matnlarni qayta ishlash, prognozlash va qarorlar qabul qilishda muhim rol o'ynaydi.

Sun'iy neyron g'oyasi yetti ming o'ttiz yillarning oxirida paydo bo'lgan bo'lsa-da, uning matematik asoslari o'tgan asrning o'rtalarida ishlab chiqilgan. Warren McCulloch va Walter Pitts birinchi marta sun'iy neyronning matematik modelini taqdim etgan va uning mantiqiy operatsiyalarni amalga oshirish qobiliyatini ko'rsatgan. Keyinchalik Frank Rosenblatt pertseptronni yaratdi va uni o'rgatish algoritmini ishlab chiqdi. Bu kashfiyotlar zamonaviy neyron tarmoqlar va chuqur o'rganishning rivojlanishiga asos bo'ldi.

Sun'iy neyron biologik neyronning soddalashtirilgan matematik modelidir. Biologik neyron dendritlar orqali kirish signallarini qabul qiladi, bu signallar neyron tanasida yig'iladi va agar umumiy signal ma'lum bir chegaradan oshsa, neyron akson orqali chiqish signalini yuboradi. Sun'iy neyron ham xuddi shunday tamoyilda ishlaydi, lekin matematik formulalar yordamida ifodalanadi.

Zamonaviy sun'iy intellekt tizimlarining muvaffaqiyati ko'p jihatdan sun'iy neyronning to'g'ri matematik modelini tanlash va samarali o'rgatish algoritmlarini qo'llashga bog'liq. Turli xil vazifalar uchun turli xil neyron modellari va faollashtirish funksiyalari ishlatiladi. Shuning uchun, sun'iy neyronning matematik modelini chuqur tushunish zamonaviy mashina o'rganish mutaxassislari uchun zarur.

Sun'iy neyronning eng oddiy va klassik matematik modeli chiziqli kombinatsiya va faollashtirish funksiyasidan iborat. Har bir sun'iy neyron bir nechta kirish signallarini qabul qiladi, ularni mos og'irliklar bilan ko'paytiradi, natijalarni qo'shadi va olingan yig'indini faollashtirish funksiyasidan o'tkazadi.

Matematik nuqtai nazardan, agar neyron n ta kirish qiymatiga ega bo'lsa, har bir kirish qiymati x indeksi i bo'lgan x bilan belgilanadi. Har bir kirish qiymati mos sinaptik og'irlik w indeksi i bo'lgan w bilan ko'paytiriladi. Bu ko'paytmalarning yig'indisi z harfi bilan belgilangan oraliq qiymatni hosil qiladi.

Oraliq qiymat z barcha kirish qiymatlari va ularning og'irliklari ko'paytmalarining yig'indisiga teng. Bundan tashqari, ko'pincha bias yoki siljish qiymati b qo'shiladi. Bias qiymati neyronning moslashuvchanligini oshiradi va faollashtirish funksiyasini gorizontaal yo'nalishda siljitish imkonini beradi.

Oraliq qiymat z faollashtirish funksiyasidan o'tkaziladi va neyronning chiqish qiymati y hosil bo'ladi. Faollashtirish funksiyasi f harfi bilan belgilanadi va u chiziqli bo'lmagan transformatsiyani amalga oshiradi. Bu chiziqli bo'lmaganlik neyron tarmoqlarning murakkab funksiyalarni o'rganish qobiliyatini ta'minlaydi.

Sun'iy neyronning bu matematik modeli juda sodda bo'lsa-da, ko'plab marta kombinatsiya qilinganda kuchli hisoblash tizimlarini yaratish mumkin. Bir nechta neyronlar qatlamlar shaklida tashkil etilganda, ular murakkab chiziqli bo'lmagan bog'lanishlarni o'rganishi mumkin.

Vektorli yozuvda sun'iy neyronning matematik modeli yanada ixcham ko'rinishga ega. Barcha kirish qiymatlari x vektori, barcha og'irliklar w vektori sifatida yoziladi. Oraliq qiymat z vektorlarning skalyar ko'paytmasi va bias qo'shilishi orqali hisoblanadi. Chiqish qiymati y faollashtirish funksiyasining oraliq qiymatdan olingan natijasidir.

Bu asosiy model turli yo'nalishlarda kengaytirilishi mumkin. Masalan, ko'p chiqishli neyronlar uchun chiqish vektori ishlatiladi. Konvolyutsion neyronlar uchun og'irliklar tensorlar sifatida ifodalanadi. Rekurrent neyronlar uchun vaqt bo'yicha bog'lanishlar qo'shiladi.

Sinaptik og'irliklar sun'iy neyronning eng muhim parametrlari hisoblanadi. Ular har bir kirish signalining neyron chiqishiga qanchalik ta'sir ko'rsatishini belgilaydi. Og'irliklar ijobiy, salbiy yoki nolga teng bo'lishi mumkin. Ijobiy og'irlik kirish signalini kuchaytiradi, salbiy og'irlik esa susaytiradi.

Og'irliklarning qiymatlari o'rgatish jarayonida o'zgartiriladi. Boshlang'ich holatda og'irliklar odatda tasodifiy qiymatlar bilan initsializatsiya qilinadi. Keyin o'rgatish algoritmi og'irliklarni asta-sekin sozlaydi, shunda neyron yoki neyron tarmoq kerakli funksiyani amalga oshiradi.

Og'irliklarning to'g'ri initsializatsiyasi juda muhim. Agar og'irliklar juda katta bo'lsa, neyron darhol to'yinish holatiga tushishi mumkin. Agar og'irliklar juda kichik bo'lsa, o'rgatish jarayoni juda sekin kechadi. Shuning uchun, turli xil og'irlik initsializatsiya usullari ishlab chiqilgan.

Xavier initsializatsiyasi og'irliklarni kirish va chiqish o'lchamlariga qarab sozlaydi. He initsializatsiyasi esa ReLU faollashtirish funksiyasi uchun maxsus optimallashtirilgan. Uniform initsializatsiya og'irliklarni ma'lum bir oraliqda teng ehtimollik bilan tanlaydi. Normal initsializatsiya esa normal taqsimot asosida og'irliklarni tanlaydi.

Og'irliklarning kattaligi regularizatsiya orqali nazorat qilinadi. L1 regularizatsiya og'irliklarning absolyut qiymatlarining yig'indisini cheklaydi va siyrak modellar hosil qiladi. L2 regularizatsiya og'irliklarning kvadratlari yig'indisini cheklaydi va umumiy og'irliklar kichikroq bo'lishini ta'minlaydi.

Og'irliklar matritsasi shaklida ifodalanishi mumkin. Agar neyron qatlami n ta kirishga va m ta chiqishga ega bo'lsa, og'irliklar matritsasi n qator va m ustundan iborat bo'ladi. Bu vektorli hisoblashlarni qo'llash va hisoblashlarni tezlashtirish imkonini beradi.

Og'irliklarning talqini vazifaga bog'liq. Tasvirlarni qayta ishlashda og'irliklar filtrlar yoki xususiyat detektorlari sifatida ishlaydi. Tabiiy tilni qayta ishlashda og'irliklar so'zlar yoki tushunchalar o'rtasidagi semantik bog'lanishlarni ifodalaydi. Vaqt qatorlarini tahlil qilishda og'irliklar temporal bog'lanishlarni yozib oladi.

Faollashtirish funksiyasi sun'iy neyronning muhim komponenti bo'lib, u chiziqli bo'lmagan transformatsiyani amalga oshiradi. Faollashtirish funksiyasiz neyron tarmoqlar faqat chiziqli transformatsiyalarni amalga oshirishi mumkin va bu ularning qobiliyatini jiddiy cheklaydi. Faollashtirish funksiyasi oraliq qiymatni oladi va uni chiqish qiymatiga aylantiradi.

Eng oddiy faollashtirish funksiyasi chegara funksiyasi yoki step function hisoblanadi. Bu funksiya kirish qiymati noldan kichik bo'lsa nol qaytaradi, aks holda bir qaytaradi. Chegara funksiyasi oddiy mantiqiy operatsiyalarni amalga oshirish uchun mos keladi, lekin u farqlanmaydigan nuqtalarga ega va o'rgatishda muammolar tug'diradi.

Sigmoid funksiyasi yoki logistik funksiya keng qo'llaniladigan silliq faollashtirish funksiyasidir. U har qanday haqiqiy sonni nol va bir oralig'idagi qiymatga aylantiradi. Sigmoid funksiya S shaklidagi egri chiziq hosil qiladi va u ehtimolliklarni modellashtirish uchun mos keladi. Biroq, sigmoid funksiya gradientlarning yo'qolishi muammosiga duch keladi.

Giperbolik tangens yoki tanh funksiyasi sigmoid funksiyasiga o'xshash, lekin u kirish qiymatini minus bir va bir oralig'idagi qiymatga aylantiradi. Tanh funksiya nol atrofida simmetrik va gradientlari sigmoid funksiyasiga qaraganda kattaroq. Bu o'rgatish jarayonini tezlashtirishi mumkin.

Rectified Linear Unit yoki ReLU funksiyasi zamonaviy chuqur o'rganishda eng keng tarqalgan faollashtirish funksiyasidir. ReLU juda oddiy: agar kirish qiymati ijobiy bo'lsa, uni o'zgartirmay qaytaradi, aks holda nol qaytaradi. ReLU hisoblash samaradorligi yuqori va gradientlarning yo'qolishi muammosiga duch kelmaydi.

Leaky ReLU funksiyasi ReLU ning modifikatsiyasi bo'lib, u salbiy qiymatlar uchun ham kichik gradient beradi. Bu yotgan neyronlar muammosini hal qilishga yordam beradi. Parametric ReLU yoki PReLU da salbiy qiymatlar uchun koeffitsient o'rganiladigan parametrlar hisoblanadi.

Exponential Linear Unit yoki ELU funksiyasi salbiy qiymatlar uchun eksponensial funksiyadan foydalanadi. U ReLU ning afzalliklarini saqlab qoladi va salbiy qiymatlar uchun silliq gradient beradi. Scaled Exponential Linear Unit yoki SELU o'z-o'zini normallashtirishni ta'minlaydigan maxsus xususiyatlarga ega.

Swish funksiyasi Google tomonidan taklif qilingan va u kirish qiymatini sigmoid funksiya bilan ko'paytiradi. Swish ba'zi vazifalarda ReLU dan yaxshiroq natijalar ko'rsatadi. GELU yoki Gaussian Error Linear Unit funksiyasi transformer arxitekturalarida keng qo'llaniladi.

Softmax funksiyasi ko'p sinfli klassifikatsiya vazifalarida chiqish qatlamida ishlatiladi. U vektorning barcha elementlarini ehtimollik taqsimotiga aylantiradi, ya'ni barcha qiymatlar ijobiy va ularning yig'indisi birga teng. Bu klassifikatsiya natijalarini talqin qilishni osonlashtiradi.

Bias yoki siljish parametri sun'iy neyronning yana bir muhim komponenti hisoblanadi. U faollashtirish funksiyasini gorizontal yo'nalishda siljitish imkonini beradi. Bias qiymati b harfi bilan belgilanadi va u vazn tortilgan yig'indiga qo'shiladi.

Bias parametrining asosiy vazifasi neyronning moslashuvchanligini oshirishdir. Biassiz neyron faqat boshlanish nuqtasi orqali o'tadigan chiziqli funksiyalarni o'rganishi mumkin. Bias qo'shish orqali neyron istalgan nuqtadan o'tadigan chiziqli funksiyalarni o'rganishi mumkin.

Oddiy chiziqli regressiya misolida bias parametrining ahamiyatini ko'rish mumkin. Agar biz og'irlik va bo'yni bog'lashni istasak, biassiz model faqat boshlanish

nuqtasidan o'tadigan chiziqlarni yaratishi mumkin. Bias bilan esa model istalgan vertikal pozitsiyaga ega chiziqni yaratishi mumkin.

Bias parametri ham o'rgatish jarayonida o'zgartiriladi. Uning boshlang'ich qiymati odatda nolga yoki kichik tasodifiy qiymatga o'rnatiladi. Gradient tushish algoritmi bias ni ham og'irliklar bilan birga sozlaydi.

Ko'p neyronli qatlamlarda har bir neyronning o'z bias parametri bo'ladi. Agar qatlamda m ta neyron bo'lsa, m ta bias parametri bo'ladi. Bias parametrlari vektor sifatida ifodalanishi mumkin va bu vektorli hisoblashlarni qo'llash imkonini beradi.

Ba'zi hollarda bias parametrlarini bir xil qiymatga o'rnatish yoki ularni nolga teng deb hisoblash mumkin. Biroq, ko'pchilik hollarda har bir neyron uchun alohida bias parametrini o'rgatish yaxshiroq natijalar beradi.

Batch normalization kabi normalizatsiya usullari qo'llanilganda, bias parametrining roli o'zgarishi mumkin. Ba'zi arxitekturalarda normalizatsiya qatlamidan keyingi bias keraksiz bo'lishi mumkin, chunki normalizatsiya o'z siljish parametriga ega.

Sun'iy neyronlarni o'rgatish neyron tarmoqlarning asosiy vazifalaridan biridir. O'rgatish jarayonida og'irliklar va bias parametrlari shunday sozlanadiki, neyron yoki neyron tarmoq berilgan vazifani optimal bajaradi. O'rgatish uchun o'quv ma'lumotlar to'plami kerak bo'ladi.

Nazorat ostida o'rgatish eng keng tarqalgan o'rgatish usuli hisoblanadi. Bu usulda har bir kirish namunasi uchun to'g'ri chiqish qiymati yoki yorliq beriladi. Neyron tarmoq kirishni qayta ishlaydi va chiqish hosil qiladi. Keyin haqiqiy chiqish va kutilgan chiqish o'rtasidagi xato hisoblanadi.

Xato funksiyasi yoki yo'qotish funksiyasi modelning qanchalik yomon ishlayotganini o'lchaydi. Regressiya vazifalari uchun o'rtacha kvadratik xato keng qo'llaniladi. Klassifikatsiya vazifalari uchun cross-entropy xato funksiyasi ishlatiladi. Xato funksiyasini minimallashtirish o'rgatishning asosiy maqsadidir.

Gradient tushish algoritmi xato funksiyasini minimallashtirish uchun ishlatiladi. Bu algoritim xato funksiyasining har bir parametr bo'yicha gradientini hisoblaydi va parametrlarni gradientga qarama-qarshi yo'nalishda o'zgartiradi. O'rganish tezligi yoki learning rate parametri qadamlar kattaligini belgilaydi.

Stochastic gradient descent yoki SGD algoritmidan har bir iteratsiyada faqat bitta namuna yoki kichik batch ishlatiladi. Bu yondashuv tezroq va xotirani kam talab qiladi. Mini-batch gradient descent bir nechta namunalardan iborat kichik guruhlarida ishlaydi va amaliyotda eng ko'p qo'llaniladi.

Orqaga tarqatish yoki backpropagation algoritmi neyron tarmoqlarda gradientlarni hisoblash uchun ishlatiladi. Bu algoritim zanjir qoidasidan foydalanib, chiqish qatlamidan boshlab orqaga qarab har bir qatlam uchun gradientlarni hisoblaydi.

Backpropagation neyron tarmoqlarning muvaffaqiyatli o'rgatilishini ta'minlaydigan asosiy algoritmdir.

Momentum usuli gradient tushishni tezlashtiradi. U oldingi gradientlarning harakatlanuvchi o'rtachasini hisoblaydi va joriy gradientga qo'shadi. Bu lokal minimumlarga tushib qolishning oldini oladi va tezroq yaqinlashuvni ta'minlaydi.

Adam optimizatori zamonaviy chuqur o'rganishda eng mashhur algoritmardan biridir. U momentum va adaptiv o'rganish tezligini birlashtiradi. Adam har bir parametr uchun alohida o'rganish tezligini saqlaydi va uni gradientlarning statistikasiga qarab sozlaydi.

RMSprop, AdaGrad, AdaDelta kabi boshqa optimizatorlar ham mavjud. Har birining o'z afzalliklari va kamchiliklari bor. Vazifaga qarab eng mos optimizatorni tanlash muhimdir.

Sun'iy neyronlar qatlamlar shaklida tashkil etilganda turli xil arxitekturalar hosil bo'ladi. Eng oddiy arxitektura ilgari ulanuvchi neyron tarmoq yoki feedforward neural network hisoblanadi. Bu arxitekturada neyronlar qatlamlar shaklida joylashgan va har bir qatlam faqat oldingi qatlamdan kirish oladi.

Ilgari ulanuvchi tarmoq kirish qatlami, bir yoki bir nechta yashirin qatlamlar va chiqish qatlamidan iborat. Kirish qatlami ma'lumotlarni qabul qiladi, yashirin qatlamlar xususiyatlarni o'rganadi va chiqish qatlami yakuniy natijani beradi. Har bir qatlamdagi neyronlar oldingi qatlamdagi barcha neyronlar bilan to'liq bog'langan.

Konvolyutsion neyron tarmoqlar yoki CNN tasvirlarni qayta ishlash uchun maxsus arxitektura hisoblanadi. Bu tarmoqlarda konvolyutsion qatlamlar mavjud bo'lib, ular lokal xususiyatlarni aniqlaydi. Konvolyutsion qatlam filtrlar yoki yadrolar yordamida tasvirning turli qismlarini skanerlaydi va xususiyat xaritalarini hosil qiladi.

Pooling qatlamlari konvolyutsion tarmoqlarda o'lchamni kamaytirish va muhim xususiyatlarni saqlash uchun ishlatiladi. Max pooling eng keng tarqalgan pooling usuli bo'lib, u har bir hududdan maksimal qiymatni tanlaydi. Average pooling esa o'rtacha qiymatni hisoblaydi.

Rekurrent neyron tarmoqlar yoki RNN ketma-ketliklarni qayta ishlash uchun mo'ljallangan. Bu tarmoqlarda neyronlar o'z chiqishlarini o'zlariga qaytaradi va bu temporal bog'lanishlarni o'rganish imkonini beradi. RNN matnlarni qayta ishlash, vaqt qatorlarini tahlil qilish va nutqni tanish vazifalarida qo'llaniladi.

Long Short-Term Memory yoki LSTM rekurrent tarmoqlarning kengaytirilgan versiyasidir. LSTM uzun muddatli bog'lanishlarni yaxshiroq o'rganadi va gradient yo'qolishi muammosini hal qiladi. LSTM xotirani boshqarish uchun maxsus gate mexanizmlaridan foydalanadi.

Gated Recurrent Unit yoki GRU LSTM ning soddalashtirilgan versiyasidir. U kamroq parametrlarga ega va tezroq o'rgatiladi, lekin ko'pgina vazifalarda LSTM bilan bir xil samaradorlikni ko'rsatadi.

Transformer arxitekturasida attention mexanizmi asosida ishlaydi va ketma-ketlik vazifalarida eng ilg'or natijalarni ko'rsatadi. Transformer rekurrent bog'lanishlarsiz parallel ravishda ishlaydi va bu uni juda tez qiladi. BERT, GPT va boshqa zamonaviy til modellari transformer arxitekturasiga asoslangan.

Sun'iy neyronlarni amalda amalga oshirish turli xil hisoblash masalalarini hal qilishni talab qiladi. Birinchi muhim jihat matritsali operatsiyalarni samarali bajarishdir. Zamonaviy neyron tarmoqlar millionlab parametrlarga ega bo'lishi mumkin va ularni samarali hisoblash muhim.

Vektorlashtirish yoki vectorization hisoblashlarni tezlashtirish uchun asosiy usul hisoblanadi. Barcha neyronlar uchun hisoblashlarni parallel ravishda bajarish uchun matritsali ko'paytirish va qo'shish operatsiyalari qo'llaniladi. NumPy, TensorFlow, PyTorch kabi kutubxonalar vektorlashtirilgan operatsiyalarni qo'llab-quvvatlaydi.

GPU yoki Graphics Processing Unit neyron tarmoqlarni o'rgatish uchun juda muhim apparatdir. GPU minglab yadroga ega va parallel hisoblashlarda juda samarali. Zamonaviy GPU lar maxsus tensor yadrolariga ega bo'lib, ular matritsali operatsiyalarni yanada tezlashtiradi.

TPU yoki Tensor Processing Unit Google tomonidan ishlab chiqilgan maxsus apparat bo'lib, u neyron tarmoqlarni o'rgatish uchun optimallashtirilgan. TPU GPU dan ham tezroq va energiya samaradorligi yuqori. Bulut platformalarida TPU larga kirish mavjud.

Xotira boshqaruvi katta neyron tarmoqlarda muhim masala hisoblanadi. Har bir neyronning aktivatsiyalari, gradientlari va parametrlari xotirada saqlanishi kerak. Gradient checkpointing kabi usullar xotira iste'molini kamaytiradi va kattaroq modellarni o'rgatish imkonini beradi.

Precision yoki aniqlik darajasi ham muhim jihatdir. Odatda parametrlar float otuz ikki yoki o'ttiz ikki bitli suzuvchi nuqta sonlari sifatida saqlanadi. Biroq, mixed precision training yoki aralash aniqlik o'rgatish o'n olti bitli va o'ttiz ikki bitli sonlarni kombinatsiya qiladi va hisoblashlarni tezlashtiradi.

Distributed training yoki taqsimlangan o'rgatish juda katta modellar va ma'lumotlar to'plamlari uchun zarur. Data parallelism usulida ma'lumotlar bir nechta qurilmalarga bo'linadi. Model parallelism usulida modelning o'zi bir nechta qurilmalarga bo'linadi.

Inference yoki xulosani chiqarish ham optimallashtirishni talab qiladi. O'rgatilgan modeldan foydalanish tezligi muhim, ayniqsa real vaqt ilovalarida. Model quantization, pruning, knowledge distillation kabi usullar modelni kichiklashtiradi va tezlashtiradi.

Sun'iy neyronlar kuchli vosita bo'lsalar-da, bir qator cheklovlar va muammolarga ega. Birinchi muhim muammo interpretatsiya qilish qiyinligidir. Katta neyron

tarmoqlar qora quti sifatida ishlaydi va ularning qarorlarini tushuntirish qiyin. Bu ayniqsa tibbiyot, moliya va huquqiy sohalarida muammo tug'diradi.

Overfitting yoki haddan tashqari moslashtirish neyron tarmoqlarning umumiy muammosi hisoblanadi. Model o'quv ma'lumotlarida juda yaxshi ishlaydi, lekin yangi ma'lumotlarda yomon natijalar ko'rsatadi. Regularizatsiya, dropout, data augmentation kabi usullar overfitting ni kamaytiradi.

Gradient yo'qolishi yoki vanishing gradient chuqur tarmoqlarda kuzatiladi. Orqaga tarqatish jarayonida gradientlar har bir qatlamda kichiklanadi va oxirgi qatlamlarga etib borguncha nolga yaqinlashadi. ReLU faollashtirish funksiyasi va residual connections bu muammoni hal qilishga yordam beradi.

Gradient portlashi yoki exploding gradient aksincha muammo hisoblanadi. Gradientlar juda katta bo'lib qoladi va parametrlar beqaror tarzda o'zgaradi. Gradient clipping usuli bu muammoni hal qilish uchun qo'llaniladi.

Ma'lumotlar etishmovchiligi ko'plab vazifalar uchun muammo hisoblanadi. Neyron tarmoqlar ko'p miqdordagi yorliqli ma'lumotlarga muhtoj. Transfer learning, data augmentation, synthetic data generation kabi usullar bu muammoni engishga yordam beradi.

Hisoblash resurslarining yuqoriligi katta modellarni o'rgatish uchun to'siq bo'lishi mumkin. Zamonaviy til modellari yuzlab GPU da haftalar davomida o'rgatiladi. Bu akademik tadqiqotchilar va kichik kompaniyalar uchun qiyin.

Adversarial attacks yoki dushman hujumlari neyron tarmoqlarning zaifliklarini ko'rsatadi. Kirishga kichik o'zgarishlar kiritish orqali modelni noto'g'ri bashorat qilishga majbur qilish mumkin. Adversarial training modellarni mustahkamroq qiladi.

Bias va adolatlilik masalalari ham muhim. Agar o'quv ma'lumotlari noto'g'ri taqsimlangan bo'lsa, model ham noto'g'ri natijalarga olib keladi. Fairness-aware machine learning bu muammolarni hal qilishga harakat qiladi.

XULOSA

Sun'iy neyronning matematik modeli zamonaviy sun'iy intellekt va mashina o'rganishning asosini tashkil etadi. Biologik neyronlardan ilhomlangan bu sodda matematik model ko'plab marta kombinatsiya qilinganda murakkab vazifalarni hal qilishga qodir kuchli tizimlarni yaratish imkonini beradi.

Ushbu tadqiqot ishida sun'iy neyronning barcha asosiy komponentlari batafsil ko'rib chiqildi. Kirish signallari, sinaptik og'irliklar, bias parametri, vazn tortilgan yig'indi, faollashtirish funksiyalari va chiqish signallari neyronning ishlashini ta'minlovchi asosiy elementlardir. Har bir komponent matematik jihatdan aniq belgilangan va o'z vazifasini bajaradi.

ADABIYOTLAR

1. McCulloch WS, Pitts W. A logical calculus of the ideas immanent in nervous activity. The Bulletin of Mathematical Biophysics. 1943. Vol. 5. No. 4. P. 115-133.

2. Rosenblatt F. The perceptron: A probabilistic model for information storage and organization in the brain. *Psychological Review*. 1958. Vol. 65. No. 6. P. 386-408.
3. Rumelhart DE, Hinton GE, Williams RJ. Learning representations by back-propagating errors. *Nature*. 1986. Vol. 323. P. 533-536.
4. LeCun Y, Bengio Y, Hinton G. Deep learning. *Nature*. 2015. Vol. 521. No. 7553. P. 436-444.
5. Goodfellow I, Bengio Y, Courville A. *Deep Learning*. MIT Press, 2016. 800 p.
6. Haykin S. *Neural Networks and Learning Machines*. Third Edition. Pearson Education, 2009. 906 p.
7. Bishop CM. *Pattern Recognition and Machine Learning*. Springer, 2006. 738 p.