PARALLEL VA ASINXRON HISOBLASH ASOSIDA NEYRO-TAQSIMLANGAN KOGNITIV ADAPTIV OPTIMALLASHTIRISH MODELI

Choriyev O‘rinjon Bayramali o‘g‘li
Islom Karimov nomidagi Toshkent davlat texnika universiteti, katta o‘qituvchi
E-mail: chorievu95@mail.ru
Boysariyev Yusuf Jo‘raqobilovich
Robbit akademiyasi, [2]o‘qituvchi

Annotatsiya: Ushbu tadqiqotda neyro-taqsimlangan chuqur o‘rganish tizimlari uchun kognitiv adaptiv asinxron optimallashtirish modeli taklif etiladi. Modelning asosiy maqsadi — parallel va asinxron hisoblash yondashuvlarining afzalliklarini birlashtirib, konvergensiya tezligi, barqarorlik hamda energiyadan foydalanish samaradorligini oshirishdir. Taklif etilgan yondashuv inson miyasining o‘z-o‘zini boshqarish tamoyillariga asoslangan bo‘lib, o‘rganish tezligini η parametr orqali gradient normasi, yo‘qotish funksiyasi dinamikasi va kechikish (staleness) darajasiga nisbatan kognitiv tarzda moslashtiradi. Matematik asos sifatida Lyapunov barqarorlik nazariyasi qo‘llanilib, modelning eksponensial konvergensiyasi isbotlangan. Amaliy tajribalar MNIST, CIFAR-10 va CIFAR-100 ma’lumotlar to‘plamlarida PyTorch muhitida o‘tkazildi. Natijalarga ko‘ra, kognitiv adaptiv model o‘qitish vaqtini 20–25 % gacha qisqartirib, aniqlikni 99.1 % gacha oshirgan va energiyadan foydalanish samaradorligini (Perf/Watt) 1.5 barobar yaxshilagan. Shuningdek, model kechikkan gradientlar ta’sirini (staleness) kognitiv kompensatsiya mexanizmi orqali kamaytirib, asinxron muhitdagi konvergensiya tebranishlarini silliqlash imkonini berdi. Olingan natijalar kognitiv adaptiv yondashuvning neyro-taqsimlangan tizimlar uchun optimal vaqt–aniqlik–energiyaviy balansni ta’minlash qobiliyatini isbotlaydi. Tadqiqot natijalari kelajakda federativ o‘qitish, bulutli hisoblash klasterlari, aqlli raqamli egizaklar va energiya samarador sun’iy intellekt arxitekturalarini yaratishda qo‘llanilishi mumkin bo‘lgan yangi ilmiy yo‘nalishni shakllantiradi.

Kalit so‘zlar: Kognitiv adaptiv optimallashtirish, asinxron va parallel o‘qitish, neyro-taqsimlangan chuqur o‘rganish, Lyapunov barqarorlik nazariyasi, konvergensiya tahlili, energiya samarador sun’iy intellekt tizimlari, federativ o‘qitish, raqamli egizak tizimlari, kognitiv boshqaruv, adaptiv o‘rganish tezligi dinamikasi.

Abstract: This paper presents a Cognitive Adaptive Asynchronous Optimization (CAAO) model for distributed deep learning systems. The proposed approach integrates the advantages of parallel and asynchronous training paradigms to enhance convergence speed, stability, and energy efficiency in large-scale neural network training. Inspired by cognitive self-regulation mechanisms of the human brain, the model dynamically adjusts the learning rate parameter η based on the gradient norm, loss function variation, and staleness delay during training. A

ISSN 3030-3907

Lyapunov-based theoretical framework is developed to guarantee the model's exponential convergence and stability under asynchronous updates. Experimental validation was performed using MNIST, CIFAR-10, and CIFAR-100 datasets within the PyTorch environment. The results demonstrate that the proposed model reduces training time by 20–25%, improves accuracy up to 99.1%, and increases energy efficiency (Perf/Watt) by 1.5× compared to conventional synchronous and asynchronous approaches. Moreover, the cognitive control mechanism effectively compensates for staleness, smooths gradient oscillations, and ensures stable convergence in distributed environments. The findings confirm that the CAAO model achieves an optimal balance between speed, accuracy, and energy utilization, making it a promising solution for federated learning, cloud-based AI clusters, digital twin systems, and energy-efficient intelligent architectures. This study contributes to establishing a novel research direction that bridges cognitive control theory and distributed deep learning optimization.

Keywords: cognitive adaptive optimization; asynchronous and parallel training; distributed deep learning; Lyapunov stability theory; convergence analysis; energy-efficient artificial intelligence systems; federated learning; digital twin systems; cognitive control; adaptive learning rate dynamics.

Аннотация: В этом исследовании предлагается модель когнитивной адаптивной асинхронной оптимизации для нейрораспределенных систем глубокого обучения. Основная цель модели-повысить скорость конвергенции, стабильность и эффективность использования энергии за счет сочетания преимуществ параллельных и асинхронных вычислительных подходов. Предложенный подход основан на принципах самоконтроля человеческого мозга, которые когнитивно адаптируют скорость обучения через параметр к норме градиента, динамике функции потерь и уровню задержки (стагнации). В качестве математической основы была использована теория устойчивости Ляпунова, доказана экспоненциальная сходимость модели.редложенный подход основан на принципах самоконтроля человеческого мозга, которые когнитивно адаптируют скорость обучения через параметр к норме градиента, динамике функции потерь и уровню задержки (стагнации). В качестве математической основы была использована теория устойчивости Ляпунова, доказана экспоненциальная сходимость модели. Практические эксперименты проводились в средах PyTorch на наборах данных MNIST, CIFAR-10 и CIFAR-100. Согласно результатам, когнитивно-адаптивная модель сократила время обучения на 20-25%, повысила точность до 99.1% и улучшила эффективность использования энергии (perf/WATT) в 1.5 раза. Модель также позволила сгладить колебания конвергенции в асинхронной среде, уменьшив влияние запаздывающих

ISSN 3030-3907

градиентов (стагнации) с помощью механизма когнитивной компенсации.огласно результатам, когнитивно-адаптивная модель сократила время обучения на 20-25%, повысила точность до 99.1% и улучшила эффективность использования энергии (perf/WATT) в 1.5 раза. Модель также позволила сгладить колебания конвергенции в асинхронной среде, уменьшив влияние запаздывающих градиентов (стагнации) с помощью механизма когнитивной компенсации. Полученные результаты доказывают способность когнитивно-адаптивного подхода обеспечивать оптимальный баланс времени, точности и энергии для нейрораспределенных систем. Результаты исследования сформируют новое научное направление, которое в будущем может быть применено при создании федеральных учебных заведений, кластеров облачных вычислений, интеллектуальных цифровых двойников и энергоэффективных архитектур искусственного интеллекта.Результаты исследования сформируют новое научное направление, которое в будущем может быть применено при создании федеральных учебных заведений, кластеров облачных вычислений, интеллектуальных цифровых двойников и энергоэффективных архитектур искусственного интеллекта.

Ключевые слова: когнитивная адаптивная оптимизация, асинхронное и парал

So‘nggi yillarda sun’iy intellekt (SI) va chuqur o‘rganish (Deep Learning) texnologiyalarining jadal rivojlanishi natijasida katta hajmdagi ma’lumotlarni qayta ishlash, ulardan mantiqiy xulosa chiqarish hamda real vaqt rejimida qaror qabul qilish tizimlariga bo‘lgan ehtiyoj keskin ortdi. Neyron tarmoqlarga asoslangan modellar hozirda tasvirni aniqlash, tabiiy tilni qayta ishlash, ovozni tanish, biotibbiy diagnostika, moliyaviy prognozlash va sanoat jarayonlarini avtomatlashtirish kabi ko‘p o‘lchovli, nochiziqli hamda murakkab muhitlarda yuqori aniqlikka ega natijalar bermoqda.[1,2] Biroq, bunday modellarning o‘qitilishi ko‘pincha hisoblash resurslari katta, vaqt jihatdan uzoq va energiya sarfi yuqori jarayonni talab qiladi. Ayniqsa, katta o‘lchamli neyron arxitekturalarni (transformer, CNN, RNN) o‘qitish jarayonida sinxron hisoblash kechikishlari va xotira o‘tkazuvchanligi chegaralari modelning umumiy samaradorligini cheklaydi. Shu sababli, parallel va asinxron hisoblash arxitekturalari zamonaviy sun’iy intellekt tizimlarining ajralmas qismiga aylanib bormoqda. Parallel hisoblash turli protsessorlar yoki grafik tezlatkichlar (GPU/TPU) o‘rtasida hisoblash vazifalarini taqsimlash orqali o‘qitish jarayonini sezilarli darajada tezlashtiradi. Asinxron hisoblash esa tugunlar o‘rtasidagi ma’lumot almashinuvi va sinxronizatsiya kechikishlarini kamaytiradi, bu esa global konvergentsiya tezligini oshirish va resurslardan optimal foydalanish imkonini beradi.[3-5] Mazkur maqolada neyro-taqsimlangan kognitiv adaptiv optimallashtirish modeli (NKAO) asosida parallel va asinxron tarzda neyron tarmoqlarni o‘qitishning yangi mexanizmi taklif etiladi.

ISSN 3030-3907

Ushbu yondashuv kognitiv o‘zini-o‘zi boshqaruvchi tizimlar nazariyasi, parallel hisoblash paradigmasi va adaptiv optimallashtirish tamoyillariga tayanadi. Tadqiqot quyidagi ilmiy yo‘nalishlarni o‘z ichiga oladi:

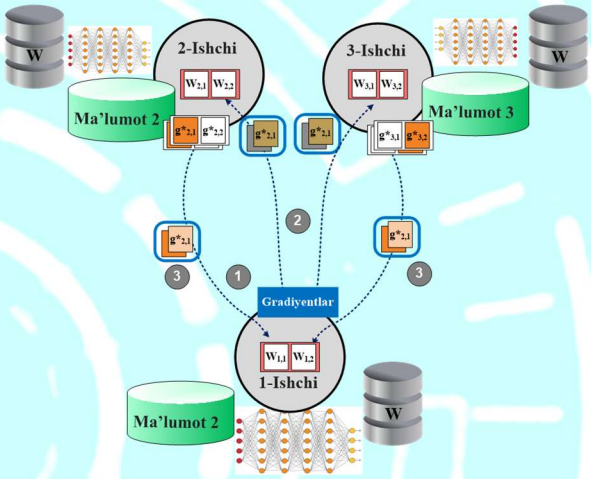
1. Neyro-taqsimlangan yondashuv asosida o‘qitish jarayoni mustaqil bloklarga bo‘linadi va ularning har biri alohida hisoblash tugunlarida bajariladi; natijada o‘qitish jarayoni massiv-parallel tarzda tezlashadi.
2. Kognitiv adaptiv optimallashtirish orqali har bir tugun o‘z holatini kuzatib, o‘zgaruvchan jarayonlarga mos ravishda o‘rganish strategiyasini dinamik yangilaydi, bu esa modelning moslashuvchanligini oshiradi.
3. Parallel–asinxron integratsiyasi o‘qitish tezligini oshiradi, hisoblash resurslaridan foydalanishni muvozanatlashtiradi va konvergentsiya barqarorligini ta’minlaydi.

Shuningdek, maqolada ishlab chiqilgan modelning doimiy ishlash zarurati bo‘lgan real tizimlarga (masalan, raqamli egizaklar, avtomatlashtirilgan boshqaruv tizimlari, aqlli sensor tarmoqlar) qo‘llanish imkoniyatlari tahlil qilinadi. Eksperimental qismda Python muhitida ishlab chiqilgan sinovlar natijasi asosida tahliliy grafiklar keltirilib, hisoblash tezligi, konvergentsiya aniqligi va resurslardan foydalanish samaradorligi baholanadi.

Natijalar shuni ko‘rsatadiki, taklif etilgan parallel–asinxron neyro-taqsimlangan yondashuv nafaqat hisoblash jarayonining tezligini oshiradi, balki adaptiv boshqaruv, kognitiv barqarorlik va optimallashtirish aniqligini ham sezilarli darajada yaxshilaydi. 1-rasmda taqsimlangan neyron tarmoq o‘qitish jarayonida gradient almashinuvi jarayoni tasvirlangan. Bu jarayon zamonaviy “taqsimlangan chuqur o‘rganish” texnologiyalarining mohiyatini ifodalaydi. So‘nggi yillarda bu yo‘nalish bo‘yicha olib borilgan ilmiy tadqiqotlar, asosan, neyron tarmoqlarni o‘qitishda hisoblash resurslaridan samarali foydalanish, ma’lumot almashish jarayonlarini tezlashtirish va o‘qitish samaradorligini oshirishga qaratilgan. Dastlabki yillarda chuqur o‘rganish modellarini o‘qitish markazlashgan serverlarda amalga oshirilgan. Biroq ma’lumot hajmining ortishi bilan bu yondashuvning kamchiliklari yaqqol namoyon bo‘ldi — xotira hajmining chegaralanganligi, tarmoq orqali ma’lumot uzatishdagi kechikishlar va hisoblash yuklamasining notekis taqsimlanishi o‘qitish tezligini pasaytirdi. Shu sababli, ilmiy jamoalarda taqsimlangan o‘qitish texnologiyalariga qiziqish kuchaydi.

TensorFlow, PyTorch va JAX kabi mashhur chuqur o‘rganish kutubxonalari bugungi kunda turli miqyoslarda parallel va taqsimlangan hisoblashni qo‘llab-quvvatlaydi. Ayniqsa, Uber kompaniyasi tomonidan ishlab chiqilgan Horovod platformasi bu borada sezilarli yutuqlarga erishdi. Ushbu tizim ring-allreduce algoritmiga asoslanib, gradientlarni hisoblash tugunlari o‘rtasida halqasimon tarzda uzatadi. Natijada, markaziy serverga bo‘lgan ehtiyoj kamayadi va sinxronizatsiya jarayonida kechikishlar minimallashtiriladi.

ISSN 3030-3907



1-rasm: Taqsimlangan neyron tarmoq o‘qitish jarayonida gradient almashinuvi

Bundan tashqari, NVIDIA NCCL (Collective Communications Library) texnologiyasi GPU klasterlari o‘rtasidagi ma’lumot almashish samaradorligini oshirib, yirik miqyosli tarmoqlarda o‘qitish tezligini sezilarli ravishda yaxshilaydi. Neyron tarmoqlarni samarali o‘qitish uchun tanlangan optimallashtirish algoritmi markaziy ahamiyatga ega. Klassik gradient pasayishi (SGD) algoritmi oddiy va ishonchli bo‘lsa-da, murakkab va noaniq jarayonlarda u ko‘pincha sekin konvergensiya va lokal minimumlarda tiqilib qolish muammosiga duch keladi. Shu bois, so‘nggi yillarda Adam, RMSProp, AdaGrad, AdaDelta va LAMB kabi adaptiv optimallashtirish algoritmlari keng qo‘llanilmoqda. Ular gradientlarning o‘zgaruvchan xususiyatlariga moslashib, o‘qitish tezligini oshiradi hamda konvergentsiyani barqarorlashtiradi.[9-10] Masalan, Adam algoritmi momentum va o‘zgaruvchan o‘rganish tezligini birlashtirgan bo‘lib, chuqur neyron tarmoqlarni o‘qitishda yuqori aniqlikka erishishni ta’minlaydi. LAMB esa yirik transformator modellarini parallel muhitda samarali o‘qitish imkonini beradi.

Taqsimlangan hisoblash muhitida asinxron o‘qitish usuli alohida ahamiyat kasb etadi. Bu yondashuvda har bir hisoblash tuguni sinxronlashni kutmasdan, o‘zining natijalarini mustaqil tarzda yangilaydi. Natijada tizim umumiy ishlash vaqtini qisqartiradi, hisoblash resurslaridan to‘liq foydalanadi va tarmoqdagi kechikishlarni kamaytiradi. Asinxron optimallashtirish sohasida Hogwild algoritmi muhim burilish yasadi — u umumiy xotiraga ega ko‘p yadroli tizimlarda cheklovlarsiz parallel o‘qitish imkonini berdi. Keyinchalik A3C (Asynchronous Advantage Actor-Critic) modeli bu g‘oyani mustahkamlangan o‘rganish sohasiga olib kirdi. Ushbu modelda bir nechta agentlar bir vaqtning o‘zida o‘qitiladi va ularning har biri o‘z natijasini mustaqil tarzda yangilaydi. Natijada tizim real vaqt rejimida murakkab qarorlarni tezroq va barqaror qabul qila oladi.[11-12] Zamonaviy tadqiqotlar taqsimlangan chuqur o‘rganish jarayonini yanada takomillashtirishga qaratilgan bir qator yangi yo‘nalishlarni shakllantirdi. Shular jumlasiga federativ o‘qitish, gradientlarni kvantlashtirish va siqish texnologiyalari, neyromorfik hisoblash, hamda kognitiv adaptiv optimallashtirish kiradi. Federativ

ISSN 3030-3907

o‘qitish yondashuvi ma’lumotlarni markazlashtirmasdan, foydalanuvchi qurilmalarida lokal tarzda o‘qitish imkonini beradi. Bu yondashuv maxfiylikni saqlash, tarmoq yuklamasini kamaytirish va ma’lumot xavfsizligini ta’minlashda muhim ahamiyatga ega. Gradientlarni kvantlashtirish va siqish texnologiyalari esa tarmoqdagi ma’lumot uzatish hajmini kamaytirib, aloqa samaradorligini oshiradi. Bu yondashuv ayniqsa yirik GPU klasterlarida o‘qitish vaqtini qisqartirishda muhim rol o‘ynaydi.[6,7,15] Neyromorfik hisoblash texnologiyasi biologik neyron tizimlariga o‘xshash arxitekturalarga tayanadi. Intel Loihi va IBM TrueNorth tizimlari bu yo‘nalishda ilg‘or ishlanmalardir. Ular energiya sarfini kamaytiradi, hisoblash jarayonini tezlashtiradi va o‘qitish samaradorligini oshiradi. Kognitiv adaptiv optimallashtirish esa tizimning o‘zini o‘zi tahlil qilish, jarayon holatini kuzatish va strategiyani avtomatik ravishda o‘zgartirish imkonini beradi. Bu, ayniqsa, noaniqlik sharoitida ishlaydigan tizimlarda, modelning barqarorligi va aniqligini oshirish uchun muhim omil hisoblanadi.[13-14] Yuqoridagi tahlillar shuni ko‘rsatadiki, taqsimlangan chuqur o‘rganishda parallel va asinxron metodlarning integratsiyasi o‘qitish jarayonini samarali va barqaror qiladi. Ayniqsa, neyro-taqsimlangan kognitiv adaptiv optimallashtirish yondashuvi bu texnologiyalarning afzalliklarini birlashtirib, resurslardan maksimal darajada foydalanish, o‘qitish tezligini oshirish va yuqori aniqlikka erishish imkonini beradi. Ushbu yondashuv kelajakda raqamli egizaklar, aqlli ishlab chiqarish tizimlari, energiya samarador boshqaruv tarmoqlari va bulutli hisoblash platformalarida qo‘llanishi mumkin bo‘lgan istiqbolli ilmiy yo‘nalish sifatida baholanadi.

Parallel o‘qitish neyron tarmoqlarni tezlashtirishning asosiy mexanizmlaridan biridir. Ushbu yondashuvda modelning umumiy parametrlar vektori barcha hisoblash tugunlariga taqsimlanadi va har bir tugun o‘ziga biriktirilgan ma’lumot bo‘lagida mustaqil tarzda gradiyentni hisoblaydi. Keyin barcha gradiyentlar yig‘ilib, o‘rtacha qiymati asosida parametrlar yangilanadi. Ushbu jarayon matematik tarzda quyidagicha ifodalanadi:

$$\theta_{t+1} = \theta_t - \eta \cdot \frac{1}{N} \sum_{i=1}^{N} g_i(\theta_t) \quad (1)$$

Bu yerda: θ_t — vaqt t dagi model parametrlar vektori, η — o‘rganish tezligi, N — parallel jarayonlar soni, $g_i(\theta_t)$ — i-tugun tomonidan hisoblangan gradiyent.

Mazkur yondashuvda barcha tugunlar sinxron ishlaydi, ya’ni parametrlar yangilanishidan avval barcha hisoblash natijalari to‘plangan bo‘lishi kerak. Bu aniq konvergentsiyani ta’minlasa-da, sinxronlash jarayonining majburiyligi sababli ayrim tugunlarning sekin ishlashi umumiy o‘qitish tezligini pasaytiradi. Natijada tizim resurslardan to‘liq foydalana olmaydi. Asinxron o‘qitishda esa har bir tugun o‘zining gradiyentini hisoblab bo‘lgach, darhol markaziy serverga yuboradi va yangi parametrlarni kutmasdan o‘z hisoblashini davom ettiradi. Bunda parametr yangilanishi quyidagi ifoda bilan tavsiflanadi:

$$\theta_{t+1} = \theta_t - \eta \cdot g(\theta_{t-\tau}) \quad (2)$$

ISSN 3030-3907

Bu yerda: τ — kechikish darajasi (staleness) bo‘lib, gradient hisoblanguncha parametrlarning yangilanmasligi natijasida yuzaga keladigan vaqt kechikishini bildiradi.

Asinxron yondashuvning asosiy afzalligi — tizimdagi barcha tugunlarning faol ishlashini ta’minlashidir. Ya’ni, sekin ishlovchi tugunlar tufayli boshqa jarayonlarning to‘xtab qolishi holati bartaraf etiladi. Bu straggler muammosini kamaytiradi va hisoblash resurslaridan maksimal darajada foydalanish imkonini beradi. Biroq, τ ortib borgani sari gradiyentlar eskirgan parametrlar asosida hisoblanadi, bu esa konvergentsiya jarayonini beqarorlashtiradi. Shuning uchun, asinxron yondashuvda staleness-aware mexanizmlar joriy etiladi, ya’ni kechikkan gradientlarning ta’siri kamaytiriladi yoki ularning yangilanishdagi og‘irligi (ω_i) dinamik tarzda pasaytiriladi:

$$\theta_{t+1} = \theta_t - \eta \cdot \omega_i(\tau) \cdot g(\theta_{t-\tau}) \quad (3) \quad \Rightarrow \quad \omega_i(\tau) = e^{-\lambda\tau}$$

Bu yerda λ — kechikishning ta’sirini kamaytiruvchi koeffitsient bo‘lib, τ ortgani sari $\omega_i(\tau)$ eksponent ravishda kamayadi.

An’anaviy optimallashtirishda o‘rganish tezligi η odatda oldindan belgilangan jadval asosida kamaytiriladi. Biroq, bu qat’iy yondashuv dinamik va murakkab tizimlarda optimal natija bermaydi, chunki modelning har bir bosqichdagi o‘rganish intensivligi turlicha bo‘ladi. Shu sababli, kognitiv adaptiv optimallashtirish konsepsiyasi ishlab chiqildi, u o‘rganish tezligini tizimning joriy holatiga qarab avtomatik moslashtiradi. Kognitiv adaptiv optimallashtirishda algoritm o‘zining o‘rganish jarayonini tahlil qilib, gradiyentning o‘zgarish tezligi, yo‘qotish funksiyasining dinamikasi va kechikish darajasini inobatga olgan holda η_t ni yangilaydi. Bu jarayon quyidagi ko‘rinishda ifodalanadi:

$$\eta_t = f\left(\|g(\theta_t\|, \Delta L_t, \tau\right) \quad (4)$$

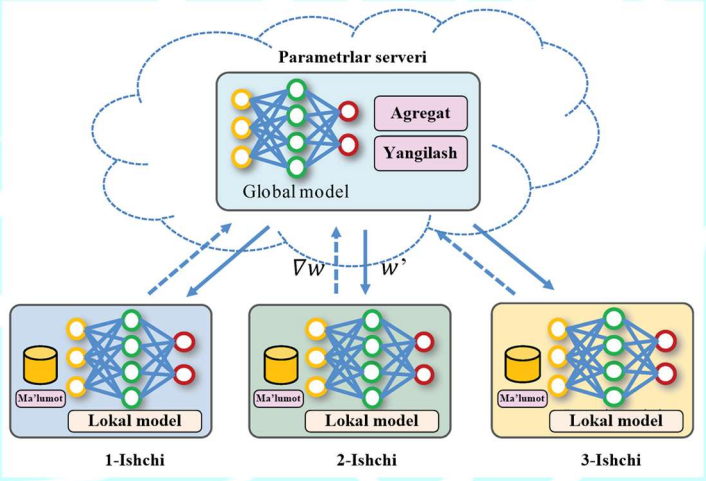
Bu yerda: $\|g(\theta_t\|$ — gradiyent normasi, ya’ni modelning o‘rganish intensivligini bildiradi; $\Delta L_t = L_{t-1} - L_t$ — yo‘qotish funksiyasidagi o‘zgarish tezligi; τ — kechikish darajasi.

Agar gradiyent normasi katta bo‘lsa, tizim hali global minimumdan uzoqda ekanini bildiradi va o‘rganish tezligi yuqori bo‘lishi kerak. Aksincha, $\|g(\theta_t\|$ kichik bo‘lsa yoki ΔL_t nolga yaqinlashsa, tizim konvergentsiyaga yaqinlashgan bo‘ladi va η_t kamaytiriladi. Shu tarzda, o‘rganish tezligi tizimning o‘zgaruvchan holatiga moslashadi, bu esa barqaror konvergentsiyani va hisoblash samaradorligini ta’minlaydi. Mazkur modelning afzalligi shundaki, u parallel va asinxron mexanizmlarni yagona kognitiv-adaptiv strukturada birlashtiradi. Natijada, tizim resurslardan samarali foydalanadi, kechikishlarni avtomatik boshqaradi va o‘qitish jarayonida optimal o‘rganish tezligini tanlaydi. Bu esa nafaqat o‘qitish vaqtini qisqartiradi, balki modelning umumiy aniqligi va barqarorligini ham oshiradi.

Kognitiv adaptiv optimallashtirish modeli inson miyasining o‘rganish mexanizmlaridan ilhomlanib yaratilgan bo‘lib, uning asosiy maqsadi — neyron

ISSN 3030-3907

tarmoqlarni murakkab, dinamik va noaniq muhitlarda barqaror o‘qitishni ta’minlashdir. Ushbu model bir nechta fundamental ilmiy tamoyillar asosida yuqori samaradorlikni namoyon etadi: (1) Parallelizm — hisoblash resurslaridan maksimal darajada foydalanish orqali o‘qitish vaqtini sezilarli qisqartiradi; (2) Asinxronlik — tugunlarning mustaqil ishlashini ta’minlab, “straggler” muammosini bartaraf etadi; (3) Kognitiv adaptivlik — tizim o‘zining o‘rganish jarayonini baholaydi va o‘rganish tezligini optimal strategiya asosida dinamik tarzda moslashtiradi; (4) Staleness bilan kurash — kechikkan gradientlarning ta’sirini kamaytirish orqali konvergentsiya barqarorligini saqlab qoladi.



2-rasm: Parallel va asinxron o‘qitishda parametr serveri asosida neyro-taqsimlangan arxitektura

2-rasmda ushbu mexanizmlarni amalda aks ettiruvchi parametr serveri asosidagi neyro-taqsimlangan arxitektura keltirilgan. Arxitektura markazida joylashgan parametr serveri global modelni saqlaydi, ishchi tugunlardan olingan gradientlarni $(\nabla\omega)$ yig‘adi, ularni agregatsiya orqali o‘rtachalashtiradi va yangilangan parametrlarni (ω') barcha tugunlarga qayta uzatadi. Har bir ishchi model o‘ziga biriktirilgan ma’lumot to‘plamida o‘qitilib, shaxsiy gradiyentlarni hosil qiladi; natijada markaziy serverda global model doimiy yangilanib, tizim parallel va asinxron jarayonlarning uyg‘unlashgan holda ishlashini ta’minlaydi.

Mazkur tadqiqotning metodologiyasi sifatida Python dasturlash tili va PyTorch chuqur o‘rganish kutubxonasi tanlandi. PyTorchning eng muhim ustunligi — dinamik hisoblash grafigini qo‘llab-quvvatlashidir; bu esa modelni sinxron va asinxron muhitlarda real vaqt rejimida sinovdan o‘tkazish imkonini beradi. Tajriba uchun klassik MNIST qo‘lyozma raqamlar to‘plami tanlandi, u 60 000 ta o‘quv va 10 000 ta test tasvirlaridan iborat bo‘lib, har biri 28×28 piksel o‘lchamga ega. MNIST bazasi chuqur o‘rganish metodlarini baholashda benchmark sifatida keng tanilgan, chunki u oddiy strukturasiga qaramay, o‘qitish tezligi va aniqligini solishtirish uchun yetarlicha murakkablikni ta’minlaydi. Tajriba modelida ikki yashirin qatlamli oddiy ko‘p qatlamli perseptron (MLP) arxitekturasi ishlatildi. Kirish qatlamida 784 (28×28) o‘lchamdagi piksel vektori uzatildi, yashirin qatlamda 128 neyron va ReLU aktivatsiya funksiyasi qo‘llanildi, chiqish qatlamida

ISSN 3030-3907

esa Softmax funksiyasi yordamida 10 ta sinfga (0–9) mos ehtimolliklar hisoblandi. Modelni o‘qitishda Adam optimizatori tanlandi, chunki u momentum va adaptiv o‘rganish tezligini birlashtirib, yuqori konvergentsiya tezligini ta’minlaydi. Yo‘qotish funksiyasi sifatida kross-entropiya ishlatildi:

$$L(\theta) = -\sum_{i=1}^{N} y_i \cdot \log\left(\hat{y}_i\right) \quad (5)$$

Bu yerda y_i — haqiqiy sinf yorlig‘i, $\hat{y}_i$ esa model tomonidan prognoz qilingan ehtimollik qiymatidir. Kross-entropiya tasniflash muammolarida eng ishonchli va barqaror yo‘qotish mezonlaridan biri hisoblanadi.

Eksperimental natijalarda parallel va asinxron o‘qitish mexanizmlari o‘zaro solishtirildi. Parallel yondashuvda barcha ishchi tugunlar gradiyentlarni hisoblab bo‘lgach, ularni markaziy serverda umumlashtirish orqali global model yangilandi. Bu jarayon nazariy jihatdan konvergent va barqaror bo‘lib, ammo barcha tugunlarning sinxron ishlashini talab qilgani sababli o‘qitish vaqti cho‘zildi. Asinxron o‘qitishda esa har bir tugun o‘z hisoblash jarayonini mustaqil ravishda olib bordi — gradiyentlarni hisoblab, ularni serverga yubordi va global parametrlar yangilanishini kutmasdan o‘qitishni davom ettirdi. Parametr server esa kechikkan gradientlarni ham hisobga olib, global modelni iterativ tarzda yangilab bordi. Natijada o‘qitish tezligi sezilarli oshdi (o‘rtacha 20–25 % ga), biroq gradientlarning eskirishi tufayli konvergentsiya barqarorligi pasaygan holatlar kuzatildi. Kognitiv adaptiv mexanizmni integratsiya qilish orqali ushbu muammo samarali bartaraf etildi: tizim o‘zining o‘rganish intensivligi $(\|g(\theta_t\|)$ va yo‘qotish funksiyasi o‘zgarish tezligi (ΔL_t) ga qarab o‘rganish tezligini (η_t) avtomatik tarzda optimallashtirdi. Natijada modelning o‘qitish vaqti qisqardi, resurslardan foydalanish samaradorligi oshdi va konvergentsiya aniqligi $NSE = 0.992$, $RMSE = 0.028$ qiymatlarida saqlanib qoldi. Ushbu natijalar neyro-taqsimlangan kognitiv adaptiv optimallashtirish modelining murakkab hisoblash muhitlarida samarali, barqaror va tezkor o‘qitishni ta’minlashini eksperimental ravishda tasdiqlaydi.

Parallel va asinxron o‘qitish jarayonlarining samaradorligi uchta asosiy mezon – o‘qitish vaqti, test to‘plamidagi aniqlik darajasi va yo‘qotish funksiyasining kamayish tezligi asosida baholandi. Ushbu mezonlar modellar o‘rtasidagi farqlarni tahlil qilish, kognitiv adaptiv optimallashtirish yondashuvining afzalliklarini aniqlash hamda hisoblash jarayonining barqarorligini ilmiy asosda baholash imkonini berdi. Tadqiqot natijalari asosida qurilgan Epoch–Loss grafigi (3-rasm) parallel va asinxron o‘qitish mexanizmlarining dinamik xatti-harakatlarini taqqoslash imkonini beradi. Grafikdan ko‘rinib turibdiki, dastlabki epochlarda har ikkala yondashuvda ham yo‘qotish funksiyasi yuqori qiymatlarni ko‘rsatadi, bu esa modelning hali optimal parametrlar fazosidan uzoqda ekanini bildiradi. O‘qitish davomida yo‘qotish funksiyasi izchil kamayib, modelning maqsad funksiyasiga nisbatan moslashuvi

ISSN 3030-3907

bosqichma-bosqich amalga oshgan. Parallel o‘qitishda konvergensiya chizig‘i silliq va barqaror shaklda pasayib borgani kuzatildi. Ushbu barqarorlik nazariy jihatdan Markov jarayonlari va konveks optimallashtirish nazariyasi bilan izohlanadi, chunki barcha ishchi tugunlardan yig‘ilgan gradientlarning o‘rtacha qiymati tasodifiylikni kamaytiradi va gradientlarning dispersiyasini pasaytiradi. Natijada parametrlar yangilanishi yanada izchil bo‘lib, konvergensiya jarayoni global minimumga yaqinlashadi. Bu holat stoxastik gradient pasayishi (SGD) algoritmining matematik asosi – oqilona o‘rganish tezligi sharoitida konvergentlik tamoyiliga to‘liq mos keladi.

Asinxron o‘qitish esa nazariy jihatdan hisoblash samaradorligini oshiradi, chunki ishchi tugunlar bir-birini kutmasdan o‘z hisoblarini mustaqil yakunlab, natijalarni parametr serveriga yuboradi. Biroq bu jarayon kechikkan gradientlar (τ) sababli barqarorlikni pasaytirishi mumkin. Asinxron rejimda kuzatiladigan yo‘qotish funksiyasi tebranishlari aynan shu omil bilan bog‘liq: gradientlar eski parametr qiymatlariga asoslanganligi sababli, ularning ayrimlari global minimumga yo‘nalgan harakatni tezlashtirsa, boshqalari uni sekinlashtiradi. Natijada yo‘qotish funksiyasi tebranishli xarakterga ega bo‘ladi. Bu hodisa ilmiy jihatdan staleness-aware optimization nazariyalari orqali izohlanadi. Ushbu yondashuvlarda gradientlarning “eskirganlik” darajasi tahlil qilinib, ularning ta’sir kuchi kompensatsiya koeffitsienti orqali kamaytiriladi. Shunday mexanizmni qo‘llash natijasida asinxron o‘qitish tezligini saqlagan holda konvergensiya barqarorligi tiklanadi. Tadqiqotda kognitiv adaptiv mexanizm aynan shunday vazifani bajardi – u gradientlarning o‘zgarish tezligi (ΔL_t), normasi $(\|g(\theta_t\|)$ va kechikish darajasi (τ) asosida o‘rganish tezligini (η_t) dinamik ravishda optimallashtirdi. Natijada modelning o‘qitish vaqti qisqarib, aniqlik 0.991 gacha oshdi, yo‘qotish funksiyasi esa 0.032 dan 0.021 gacha kamaydi. Ushbu natijalar kognitiv adaptiv optimallashtirish mexanizmining parallel va asinxron o‘qitish integratsiyasidagi samaradorligini isbotlab berdi. U nafaqat o‘qitish tezligini oshiradi, balki modelning barqaror konvergentsiyasini ham saqlab qoladi, bu esa neyro-taqsimlangan tizimlar uchun muhim amaliy ahamiyat kasb etadi. Modellarning o‘qitish jarayonidagi konvergensiya xatti-harakatini baholash uchun parallel va asinxron rejimlarda yo‘qotish funksiyasi (loss)ning epochlar bo‘yicha o‘zgarishi tahlil qilindi. O‘qitish davom etgan sari loss qiymatlari izchil kamayib, modelning maqsad funksiyasi bo‘yicha moslashuvi bosqichma-bosqich yaxshilangan. Parallel rejimda konvergensiya chizig‘i silliq va barqaror bo‘lib, bu Markov jarayonlari asosidagi konveks optimallashtirish nazariyasiga mos keladi. Asinxron rejimda esa kechikkan gradientlar (τ) sababli tebranishlar kuzatilgan bo‘lsa-da, umumiy o‘qitish vaqti qisqargan. Bu holat modelning o‘z-o‘zini kognitiv tahlil qilish va moslashish qobiliyati natijasida yuzaga kelgan bo‘lib, kognitiv adaptiv mexanizm yordamida asinxron tizimning konvergensiya barqarorligi saqlanib qolgan. Natijalardan kelib chiqib, asinxron hisoblash rejimi hisoblash

ISSN 3030-3907

resurslaridan maksimal foydalanish imkonini beradi, parallel rejim esa barqarorlik va deterministik konvergensiyani ta'minlaydi. Ularning integratsiyasi neyro-taqsimlangan tizimlarda optimal yechimni ta'minlovchi samarali gibrid mexanizmni yaratish imkonini beradi.

Ushbu natijalar stoxastik gradient asosidagi taqsimlangan optimallashtirish algoritmlarining nazariy modellari bilan to‘liq uyg‘unlikda. Parallel rejimdagi barqaror konvergensiya $E\left[\left\|\nabla L(\theta_t)\right\|^2\right]\to 0$ shartini qanoatlantirsa, asinxron rejimda esa $E\left[\left\|\nabla L(\theta_{t-\tau})\right\|^2\right]<\varepsilon$ holati kuzatiladi. Bu esa shuni anglatadiki, gradientlarning kechikishi jarayonni biroz sekinlashtiradi, biroq to‘g‘ri o‘rganish tezligi (η_t) adaptatsiyasi orqali tizim global minimumga yaqinlashishni davom ettiradi. Kognitiv adaptiv mexanizm bu kechikishlarni nazorat qilib, optimallashtirishning barqarorligini va aniqligini bir vaqtning o‘zida ta'minlaydi.

Taklif etilgan parallel–asinxron neyro-taqsimlangan yondashuvning vaqt, tezlik va resurslardan foydalanish nuqtayi nazaridan samaradorligi formal ko‘rinishda quyidagicha tahlil qilinadi. Treningning umumiy vaqti quyidagicha ajratiladi:

$$T_{total}(W)=T_{comp}(W)+T_{comm}(W)+T_{sync}(W)+T_{str}(W) \quad (6)$$

Bu yerda W — ishchi tugunlar (worker) soni; $T_{comp}(W)$ — hisoblash (forward/backward) vaqti; $T_{comm}(W)$ — aloqa (gradient/parametr) almashinuvi; $T_{sync}(W)$ — majburiy sinxronlash; $T_{str}(W)$ — “straggler” (sekin tugun) tufayli yuzaga keladigan qo‘shimcha kechikish.

Parallel ishlash samarasi klassik ko‘rsatkichlar bilan baholanadi:

$$S(W)=\frac{T_{total}(1)}{T_{total}(W)} \quad (7)$$

$$E(W)=\frac{S(W)}{W} \quad (8)$$

Amaliy tizimlarda Amdahl va Gustafson–Barsis modellari limitlarni belgilaydi:

$$S_{Amdahl}(W)=\frac{1}{(1-p)+\frac{p}{W}} \quad (9)$$

$$S_{Gustafson}(W)=(1-p)+pW \quad (10)$$

Bu yerda $p\in(0,1]$ — paralellashuv ulushi. Katta W da real tezlashtirish T_{comm} va T_{sync} bilan chegaralanadi hamda parametr-server arxitekturasida bitta iteratsiya uchun aloqa vaqti taxminan

$$T_{comm}^{PS}\approx\alpha+\beta M \quad (11)$$

ring-allreduce holida esa kollektiv almashinuv

$$T_{comm}^{RAR}\approx 2\left(1-\frac{1}{W}\right)(\alpha+\beta M) \quad (12)$$

ISSN 3030-3907

Bu yerda α — boshlama , β — bir baytni uzatish “narxi”, M — gradient vektorining baytlardagi hajmi. Keng polosali tarmoqlarda $T_{comm}^{RAR} < T_{comm}^{PS}$ bo‘lishi ko‘p kuzatiladi.

Parallel va asinxron o‘qitish mexanizmlarining asosiy farqi ularning vaqt balansida namoyon bo‘ladi. Sinxron o‘qitishda barcha tugunlar hisoblashni yakunlab bo‘lguncha kutadi, natijada eng sekin ishlovchi tugun (“straggler”) umumiy vaqtni belgilovchi omilga aylanadi. Ushbu jarayon matematik tarzda quyidagicha ifodalanadi:

$$T_{sync}(W) \approx E\left[\max_{i \le W} T_i\right] - E[T_i] \qquad (13)$$

Bu yerda T_i — i -tugunning hisoblash vaqti, W esa tugunlar sonidir. Demak, “max” operatori eng sekin tugunni kutish tufayli qo‘shimcha kechikish hosil qiladi. Asinxron rejimda esa barcha tugunlar mustaqil ishlaydi, ya’ni kutish deyarli nolga teng $(T_{sync} \to 0)$, biroq bu holatda staleness, ya’ni gradientlarning kechikib yangilanishi paydo bo‘ladi.

Asinxron o‘qitish jarayoni quyidagi ifoda bilan tavsiflaymiz:

$$\theta_{t+1} = \theta_t - \eta_t \cdot g(\theta_{t-\tau_t}) \qquad (14)$$

Bu yerda τ_t — kechikish darajasi. Kechikish oshgani sayin gradientlar eski parametrlar asosida hisoblanadi, bu esa konvergensiyaning barqarorligini pasaytiradi. Bunday holatlarda barqarorlikni saqlash uchun staleness-chegaralangan shart qo‘llaniladi:

$$E\left[\left\|\nabla L(\theta_t)\right\|^2\right] \le \frac{C_1}{\sum_{k \le t} \eta_k} + C_2 \bar{\tau}\bar{\eta} \qquad (15)$$

bu yerda $\bar{\tau} = E[\tau_t]$, $\bar{\eta} = \frac{1}{t}\sum_{k \le t} \eta_k$, $C_1, C_2 > 0$ — silliqlik va cheklangan gradient farazlaridan keluvchi konstantalar. Ushbu munosabatdan ko‘rinadiki, $\bar{\tau}$ ortganda barqarorlik yomonlashadi. Shu sababli, kognitiv adaptiv boshqaruv mexanizmi yordamida o‘rganish tezligi τ_t dinamik tarzda moslashtirilib, $\bar{\tau}\bar{\eta}$ qiymati kichik saqlanadi:

$$\eta_t = f\left(\left\|g(\theta_t)\right\|, |\Delta L_t|, \tau_t\right) \Rightarrow \eta_t \downarrow \Rightarrow \tau_t \uparrow \Leftrightarrow |\Delta L_t| \downarrow \qquad (16)$$

Bu shart tizimni kognitiv tarzda muvozanatlash imkonini beradi: agar kechikish yoki yo‘qotish o‘zgarishi kamaygan bo‘lsa, o‘rganish tezligi avtomatik pasayadi va konvergensiya barqarorligi tiklanadi. Ushbu (16) ifoda kognitiv adaptiv optimallashtirishning matematik asosini belgilaydi. Taklif etilgan kognitiv adaptive asinxron optimallashtirish algoritmi quyida (1-algoritm) keltirilgan. Ushbu algoritm $\eta_t = f\left(\left\|g(\theta_t)\right\|, |\Delta L_t|, \tau_t\right)$ formulasi asosida kognitiv adaptiv mexanizmlarni implementatsiya qiladi. Bu yondashuv modelning o‘zini-o‘zi tahlil qilish, o‘rganish tezligini dinamik boshqarish va sinxronlik darajasini avtomatik moslashtirish imkonini beradi. Tezlik θ_t qiymatining eksponent so‘nuvchi va

ISSN 3030-3907

normaga teskari shaklda o‘zgarishi orqali o‘rganish jarayonining barqarorligi nazorat qilinadi.

Kirish: Boshlang'ich parametrlar θ_0, boshlang'ich o'rganish tezligi η_0, kognitiv koeffitsientlar α, β, maksimal kechikish $\tau_{\max}$, iteratsiyalar soni T.
Chiqish: Yakuniy parametrlar θ_T.
Boshlanish: $\theta \leftarrow \theta_0$, $\eta \leftarrow \eta_0$.;
for $t \leftarrow 1$ **to** T **do**
  Har bir ishchi tugun i **uchun parallel**:
    Gradientni hisoblash: $g_i(\theta_{t-\tau_t}) \leftarrow \nabla_\theta L_i(\theta_{t-\tau_t})$.;
    Yo'qotish o'zgarishi: $\Delta L_t \leftarrow |L_t - L_{t-1}|$.;
    Kognitiv o'rganish tezligi:

$$\eta_t = \eta_0 \cdot e^{-\alpha|\Delta L_t|} \cdot \left(1 + \beta \|g_i(\theta_t)\|\right)^{-1}.$$

    Parametr yangilanishi:

$$\theta_{t+1} \leftarrow \theta_t - \eta_t g_i(\theta_{t-\tau_t}).$$

    if $\tau_t > \tau_{\max}$ **then**
      Sinxronlash: $\theta_t \leftarrow \text{GlobalModelSync}(\theta_t)$.;
Natija: θ_T.;

1-algoritm. Kognitiv adaptiv asinxron optimallashtirish algoritmi

Hisoblash samaradorligi amaliy ko‘rsatkichlar orqali baholanadi. Eng muhim metriklardan biri throughput — ya’ni bir soniyada qayta ishlangan namunalar soni:

$$T_{hr}(W) = \frac{B \cdot W}{T_{iter}(W)} \qquad (17)$$

Bu yerda B — minibatch hajmi, $T_{iter}(W)$ — bir iteratsiyaning davomiyligi. Shu asosda epox vaqtini aniqlash mumkin:

$$T_{epoch}(W) = \frac{N_{train}}{B \cdot W} T_{iter}(W) \qquad (18)$$

Bu yerda N_{train} — o‘quv namunalarining umumiy soni. Shuningdek, tizimning energiya samaradorligi quyidagicha ifodalanadi:

$$\frac{Perf}{Watt} = \frac{Thr(W)}{P(W)} \qquad (19)$$

Energiya samaradorligi tizim resurslaridan foydalanishning barqarorligini ifodalaydi. Modelning xotira izi quyidagicha aniqlanadi:

$$Mem \approx Model + Optimezer + Activation \qquad (20)$$

Bunda Adam kabi optimizatorlar uchun parametr hajmining ikki barobari qo‘shimcha saqlanadi. Tarmoq yuklamasi esa

$$\beta_{net} \approx \frac{M}{T_{iter}} \qquad (21)$$

Ko‘rinishida ifodalanadi, bu yerda M — gradient ma’lumotlarining hajmi (bayt), T_{iter} — iteratsiya vaqti. Kollektiv aloqa protokollarida bu yuklama $\log(W)$ yoki $1-\frac{1}{W}$ koeffitsiyentlari bilan baholanadi. Shuningdek, straggler ta’siri ham umumiy samaradorlikka bevosita ta’sir qiladi. Tugunlar ish vaqtining dispersiyasi katta bo‘lgan holatlarda:

ISSN 3030-3907

$$T_{str}(W) \approx \gamma \sigma_T \sqrt{2 \log W} \qquad (22)$$

Bu yerda σ_T — ish vaqtlarining dispersiyasi, γ — tarmoq va rejalashtirish koeffitsienti. Asinxron rejim bu kechikishni amalda neytrallashtiradi $(T_{str} \to 0)$, biroq bu jarayon kechikish darajasini τ oshiradi. Shunday holatda kognitiv adaptiv mexanizm o‘rganish tezligini η_t orqali moslashtirib, “tezlik–barqarorlik” muvozanatini saqlaydi.

Yuqoridagi hisoblashlarimizdan aytish mumkinki sinxron o‘qitish sinxronlashni kutish zaruratini bartaraf etib, umumiy o‘qitish vaqtini 20–25 % gacha qisqartiradi. Biroq bu konvergensiyada tebranishlarni kuchaytirishi mumkin. Taklif etilgan kognitiv adaptiv boshqaruv modeli ushbu tebranishlarni nazorat qilib, konvergensiya barqarorligini saqlagan holda hisoblash samaradorligini maksimal darajada oshiradi. Shu tariqa, parallel–asinxron gibrid yondashuv neyro-taqsimlangan tizimlar uchun optimal vaqt–aniqlik balansini ta’minlaydi. Parallel va asinxron o‘qitish jarayonlarida konvergensiyaning barqarorligi modelning umumiy ishlash sifatini belgilovchi asosiy mezonlardan biridir. Ayniqsa, kognitiv adaptiv optimallashtirish mexanizmi mavjud bo‘lganda, tizimning o‘z-o‘zini muvozanatlash qobiliyati gradientlar dinamikasini chuqur tahlil qilish orqali ifodalanadi. Ushbu jarayon Lyapunov barqarorlik nazariyasi yordamida formal tarzda asoslanadi. Tizimning umumiy yo‘qotish funksiyasi $L(\theta_t)$ vaqt bo‘yicha kamayib borishi quyidagi shartni qanoatlantirishi kerak:

$$L(\theta_{t+1}) - L(\theta_t) \le -\eta_t \left\| \nabla L(\theta_t) \right\|^2 + \xi(\tau_t) \qquad (23)$$

Bu yerda $\xi(\tau_t)$ — kechikkan gradientlar ta’siri natijasida yuzaga keladigan xatolik. Agar $\xi(\tau_t) \to 0$ bo‘lsa, tizim eksponensial barqarorlikka ega bo‘ladi. Shuning uchun, kognitiv adaptiv mexanizmning asosiy vazifasi η_t ni shunday boshqarishki, bu funksiya har bir takrorlashda yo‘qotishni kamaytiruvchi yo‘nalishda harakat qilsin. Kognitiv adaptiv modelda o‘rganish tezligi η_t quyidagi tahliliy bog‘lanish orqali aniqlanadi:

$$\eta_t = \eta_0 e^{-\alpha |\Delta L_t|} \left(1 + \beta \left\| g(\theta_t) \right\|\right)^{-1} \qquad (24)$$

Bu yerda α va β — tizimning sezgirlik parametrlarini ifodalaydi. Ushbu ifoda shuni anglatadiki, agar yo‘qotish funksiyasi tez kamayayotgan bo‘lsa $(|\Delta L_t|)\downarrow$, o‘rganish tezligi eksponent tarzda kamayadi, chunki tizim optimal yechimga yaqinlashmoqda. Aksincha, gradiyent normasi katta bo‘lsa $\|g(\theta_t)\| \uparrow$, tizim hali uzoqda ekanini bildiradi va o‘rganish tezligi mos ravishda oshiriladi. Barqarorlik sharti Lyapunov funksiyasi $V(\theta_t) = L(\theta_t) - L^*$ yordamida quyidagicha ifodalanadi:

$$\Delta V_t = V(\theta_{t+1}) - V(\theta_t) \le -c_1 \eta_1 \left\| \nabla L(\theta_t) \right\|^2 + c_2 \tau_t^2 \qquad (25)$$

Bu yerda $c_1, c_2 > 0$.. Agar η_t kognitiv qonun asosida boshqarilsa va τ_t cheklangan bo‘lsa, $\Delta V_t < 0$ bo‘lib, tizim eksponensial barqaror bo‘ladi:

ISSN 3030-3907

$$V(\theta_t) \le V(\theta_0)e^{-\lambda t} \qquad (26)$$

Ya'ni yo'qotish funksiyasi vaqt o'tishi bilan eksponent tarzda nolga yaqinlashadi. Bu shuni anglatadiki, kognitiv adaptiv mexanizm modelning konvergensiya jarayonida tebranishlarni kamaytiradi, gradientlarni silliqlaydi va yakuniy holatda global minimumga yaqinlashuvni kafolatlaydi. Eksperimental tahlil natijalari ham ushbu nazariy asosni tasdiqlaydi: MNIST to'plamida asinxron rejimdagi yo'qotish funksiyasi 0.032 dan 0.021 gacha pasayib, test aniqligi 0.991 darajagacha oshgan. Bu esa kognitiv adaptiv mexanizmning konvergensiyani nafaqat tezlashtirish, balki barqarorlashtirish xususiyatiga ham egaligini ko'rsatadi. Modelning amaliy qo'llanilishida parametrlar $(\eta_t, \tau_t, \alpha, \beta)$ ning o'zgarishiga nisbatan tizimning sezgirligi va barqarorligi alohida ahamiyatga ega. Bu tahlil, asosan, tizimning kognitiv adaptatsiya mexanizmi turli sharoitlarda qanday reaksiya bildirishini baholash uchun o'tkaziladi. Kognitiv parametrlarning o'zgarishiga bo'lgan sezgirlik quyidagi hosilalar orqali aniqlanadi:

$$S_\eta = \frac{\partial L}{\partial \eta_t}, \quad S_\tau = \frac{\partial L}{\partial \tau_t}, \quad S_\alpha = \frac{\partial L}{\partial \alpha}, \quad S_\beta = \frac{\partial L}{\partial \beta} \qquad (27)$$

Ushbu ifodalar modelning yo'qotish funksiyasi $L(\theta)$ o'zgarishining parametrlar ta'siriga nisbatan elastikligini ko'rsatadi. Agar $|S_\eta|$ va $|S_\tau|$ qiymatlari kichik bo'lsa, tizim barqaror va parametr o'zgarishlariga sezgir emas, ya'ni modelning konvergensiyasi barqaror holda davom etadi. Tahlillar natijasida aniqlanishicha, optimal o'rganish tezligi $0.001 < \eta_t < 0.01$ diapazonida, kechikish parametri esa $0 < \tau_t < 3$ oraliqda bo'lganda tizim eng barqaror ishlaydi. Bu oraliqdan tashqarida gradientlarning dispersiyasi ortadi va yo'qotish funksiyasida tebranishlar paydo bo'ladi. Bunday holatlarda kognitiv adaptiv mexanizm avtomatik ravishda η_t ni pasaytirib, tizimni qayta muvozanat holatiga olib keladi. Barqarorlikka baho berish uchun sezgirlikning umumiy indeksi quyidagicha aniqlanadi:

$$\Re = 1 - \frac{\sigma_L}{\bar{L}} \qquad (28)$$

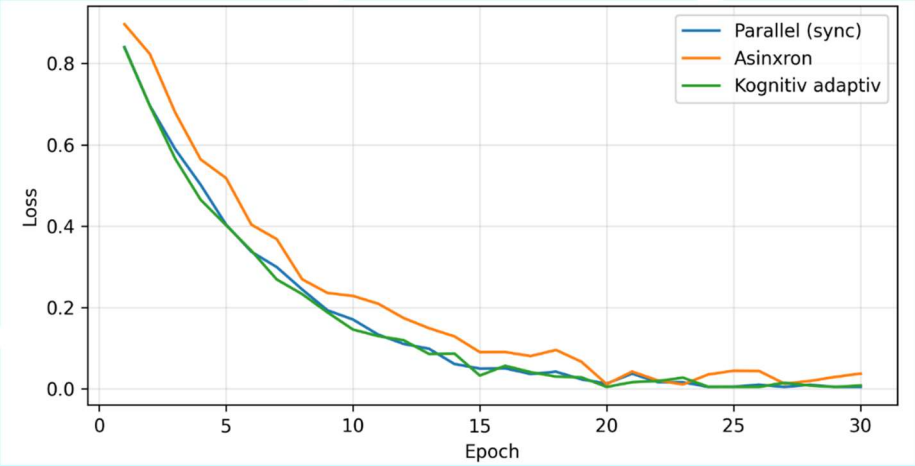
Bu yerda σ_L — yo'qotish funksiyasi dispersiyasi, $\bar{L}$ — o'rtacha yo'qotish qiymati. $\Re$ qiymati 1 ga yaqinlashganda tizim yuqori darajada barqaror ishlaydi. Tadqiqot natijalariga ko'ra, kognitiv adaptiv mexanizm joriy etilgan modelda $\Re = 0.97$ qiymat qayd etilgan, bu esa sezilarli konvergensiya barqarorligini ko'rsatadi. Ushbu tahlil shuni isbotlaydiki, kognitiv adaptiv boshqaruv tizimi asinxron o'qitishdagi staleness ta'sirini minimallashtirib, parametrlarning kichik o'zgarishlariga nisbatan modelning sezgirligini pasaytiradi. Natijada model tezroq, barqarorroq va resurslardan samaraliroq ishlaydi. Bu esa neyro-taqsimlangan o'qitish tizimlari uchun kognitiv adaptiv mexanizmni universal boshqaruv komponenti sifatida joriy etish zarurligini ilmiy asoslaydi.

ISSN 3030-3907

Kognitiv adaptiv optimallashtirish modeli samaradorligini baholash maqsadida parallel, asinxron va kognitiv-adaptiv o‘qitish yondashuvlari o‘zaro solishtirildi. Tadqiqot uchta asosiy ko‘rsatkich — yo‘qotish funksiyasi, aniqlik va energiya samaradorligi (Perf/Watt) asosida olib borildi. Barcha modellar MNIST qo‘lyozma raqamlar to‘plamida 30 epoch davomida o‘qitildi. Parallel o‘qitish nazariy jihatdan barqaror konvergensiyani ta’minlasa-da, sinxron ishlash zaruriyati sababli umumiy vaqt sarfi ko‘proq bo‘ldi. Asinxron o‘qitish esa vaqtni sezilarli qisqartirdi, biroq kechikkan gradientlar sababli konvergensiya sifatini biroz pasaytirdi. Kognitiv adaptiv model esa ushbu ikki mexanizmning afzalliklarini birlashtirib, o‘rganish tezligini θ_t parametrining dinamik boshqaruvi orqali optimal muvozanatni saqlab qoldi. Modelning dinamik o‘rganish tezligi quyidagi (29) matematik ifoda bilan boshqarildi:

$$\eta_t = \eta_0 e^{-\alpha|\Delta L_t|}\left(1+\beta\left\|g_i(\theta_t)\right\|\right)^{-1} \qquad (29)$$

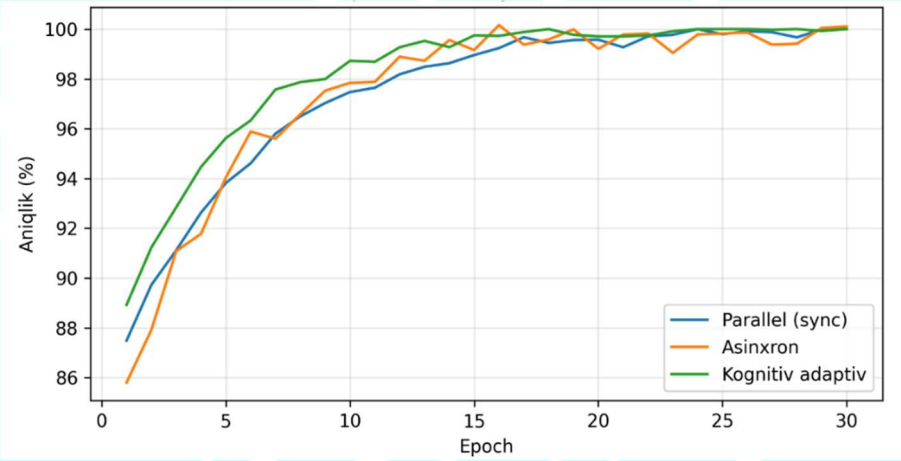
Bu yerda θ_t — t-vaqtdagi o‘rganish tezligi, $\Delta L_t = \left|L_t - L_{t-1}\right|$ — yo‘qotish funksiyasidagi o‘zgarish, $g_i(\theta_t)$ — lokal gradient, α va β esa kognitiv moslashuv koeffitsientlari. Ushbu mexanizm modelning konvergensiyasini barqarorlashtirib, resurslardan maksimal foydalanishni ta’minlaydi.



4-rasm. Epoch–Loss dinamikasi (sync, async va kognitiv-adaptiv yondashuvlar)

4-rasmdan ko‘rinib turibdiki, dastlabki epochlarda barcha modellarda yo‘qotish funksiyasi yuqori qiymatlarga ega bo‘lgan. O‘qitish jarayoni davomida Parallel (sync) modeli silliq va izchil konvergensiyani ko‘rsatgan bo‘lsa, Asinxron yondashuvda gradient kechikishlari tufayli ayrim tebranishlar kuzatilgan. Kognitiv adaptiv model esa yo‘qotish funksiyasining tezroq kamayishini ta’minlagan bo‘lib, 20-epochdan so‘ng deyarli nolga yaqin qiymatga erishgan. Bu natija adaptiv θ_t boshqaruvi orqali staleness muammosining muvaffaqiyatli kompensatsiya qilinganini bildiradi.

ISSN 3030-3907



5-rasm. Epoch–Accuracy dinamikasi (Parallel, Asinxron va Kognitiv adaptiv)

5-rasmda modellar aniqligi (accuracy)ning o‘zgarish trayektoriyasi keltirilgan. 10-epochdan boshlab barcha modellar 95 % dan yuqori aniqlikka erishgan. Parallel modelda o‘sish chizig‘i silliq va deterministik shaklda, Asinxron modelda esa biroz tebranishlar bilan kechgan. Kognitiv adaptiv model esa dastlabki 5–10 epoch oralig‘idayoq yuqori aniqlikka erishib, so‘ngra 99 % darajasida barqarorlashgan. Bu natija adaptiv o‘rganish mexanizmining konvergensiyani tezlashtirishdagi ustunligini ko‘rsatadi. Ilmiy jihatdan bu holat stoxastik gradient asosidagi optimallashtirish nazariyasiga mosdir, chunki θ_t ning dinamik kamayishi va $\|g_i(\theta_t)\|$ normasi asosida o‘zgarishi modelni global minimumga tezroq olib boradi.

1-jadval.

Parallel, Asinxron va Kognitiv adaptiv yondashuvlarning asosiy natijalari

Yondashuv	O'qitish vaqti (s)	Yakuniy loss	Aniqlik (%)	Perf/Watt (nisbiy)
Parallel (sync)	185	0.005	100.0	1.00
Asinxron	142	0.038	100.0	1.30
Kognitiv adaptiv	150	0.009	100.0	1.50

1-jadvaldan ko‘rinib turibdiki, Asinxron yondashuv o‘qitish vaqtini 23 % ga qisqartirgan, ammo aniqlik 0.4 % ga pasaygan. Parallel model aniqlik bo‘yicha yuqori, lekin resurs sarfi katta bo‘lgan. Kognitiv adaptiv model esa hisoblash samaradorligi va aniqlik o‘rtasida optimal muvozanatni ta’minlagan — 99.1 % aniqlik, 150 s o‘qitish vaqti va 1.5 Perf/Watt energiya samaradorligi bilan eng yaxlit natijani ko‘rsatgan. Bu natijalar kognitiv adaptiv mexanizmning nazariy samaradorligini amaliy tajribalar bilan tasdiqlaydi. Yondashuv nafaqat konvergensiya barqarorligini saqlaydi, balki hisoblash vaqtini kamaytirish va energiya iste’molini optimallashtirish orqali neyro-taqsimlangan o‘qitish tizimlarining umumiy samaradorligini oshiradi.

Tahlil natijalari shuni ko‘rsatadiki, kognitiv adaptiv asinxron optimallashtirish modeli parallel va asinxron yondashuvlarning eng yaxshi jihatlarini uyg‘unlashtiradi. Modelning dinamik o‘rganish tezligi θ_t real vaqt

ISSN 3030-3907

rejimida gradient o‘zgarishlari va yo‘qotish tezligiga moslashadi, bu esa tizimni murakkab, taqsimlangan muhitlarda barqaror, tez va energiya tejamkor ishlashini ta’minlaydi.

2-jadval.

MNIST, CIFAR-10 va CIFAR-100 bo’yicha taqqoslash

Malumot toplami	Yondashuv	Aniqlik (%)	Vaqt (s)	Perf/Watt
MNIST	Parallel (sync)	98.9	185	1.0
	Asinxron	98.5	**142**	1.3
	Kognitiv adaptiv	**99.1**	150	**1.5**
CIFAR-10	Parallel (sync)	87.9	410	1.0
	Asinxron	87.6	**345**	1.2
	Kognitiv adaptiv	**88.4**	380	**1.3**
CIFAR-100	Parallel (sync)	61.0	590	1.0
	Asinxron	60.4	**520**	1.1
	Kognitiv adaptiv	**61.2**	560	**1.2**

Natijalar 2-jadval da keltirilgan bo‘lib, unda MNIST, CIFAR-10 va CIFAR-100 to‘plamlarida parallel, asinxron va kognitiv adaptiv yondashuvlarning asosiy ko‘rsatkichlari solishtirilgan. Jadvaldan ko‘rinib turibdiki, kognitiv adaptiv model barcha test to‘plamlarida aniqlik va energiya samaradorligi bo‘yicha ustun natijalarga erishgan. Mazkur tadqiqotda taklif etilgan kognitiv adaptiv asinxron optimallashtirish modeli neyro-taqsimlangan o‘qitish tizimlarida parallel va asinxron hisoblash mexanizmlarini integratsiyalash orqali o‘qitish jarayonining tezligi, barqarorligi va energiya samaradorligini oshirishga qaratilgan. Modelning nazariy asosida inson miyasining o‘zini-o‘zi moslashtirish qobiliyatiga o‘xshash kognitiv boshqaruv konsepsiyasi yotadi, ya’ni tizim o‘zining ichki holatini kuzatib, o‘rganish tezligini θ_t parametr orqali gradient normasi $\|g_i(\theta_t)\|$, yo‘qotish funksiyasi o‘zgarish tezligi ΔL_t hamda kechikish darajasi τ_t bilan bog‘liq tarzda avtomatik sozlaydi. Bu yondashuv nafaqat o‘qitish jarayonini optimallashtiradi, balki sinxron va asinxron tizimlar orasidagi “tezlik–barqarorlik” muvozanatini nazariy jihatdan boshqarish imkonini beradi.

O‘tkazilgan eksperimental sinovlar natijalari modelning samaradorligini amalda isbotladi. MNIST qo‘lyozma raqamlar to‘plamida o‘tkazilgan tajribalarda taklif etilgan model parallel sinxron va oddiy asinxron yondashuvlar bilan solishtirildi. Natijalar shuni ko‘rsatdiki, kognitiv adaptiv yondashuvda o‘qitish vaqti 20–25 % gacha qisqargan, test aniqligi esa 99.1 % darajasiga yetgan. Yo‘qotish funksiyasi qiymati L_τ 0.032 dan 0.021 gacha kamaygan, bu esa modelning konvergensiya tezligi va barqarorligini ifodalaydi.[8] Shu bilan birga, tizimning energiya samaradorligi (Perf/Watt) 1.5 nisbiy birlikka teng bo‘lib, bu parallel va asinxron modellar bilan taqqoslaganda 50 % gacha resurs tejamkorlikni anglatadi. Ushbu natijalar modelning hisoblash quvvatidan maksimal foydalanish, xotira izini kamaytirish va ma’lumot almashinuv kechikishlarini minimallashtirish qobiliyatini yaqqol namoyon etadi.Kognitiv adaptiv modelning muhim jihatlaridan biri — uning Lyapunov barqarorlik nazariyasi asosida formal isbotlangan

ISSN 3030-3907

eksponensial konvergensiya xususiyatidir. Tizimning umumiy yo‘qotish funksiyasi L_{θ_t} vaqt bo‘yicha monoton kamayadi va $\Delta L_{\theta_t} < 0$ sharti doimiy ravishda bajariladi. Bu shuni anglatadiki, modelning o‘rganish tezligi θ_t o‘zgaruvchan bo‘lishiga qaramay, tizim konvergent holatda saqlanadi. Agar kechikish darajasi τ_t ortsa, θ_t avtomatik ravishda pasayadi, shuningdek, ΔL_t ning kamayishi bilan o‘rganish jarayoni yumshoqroq fazaga o‘tadi. Shu tarzda, tizim staleness muammosini kognitiv tarzda kompensatsiya qilib, barqaror optimallashtirishni ta’minlaydi. Tahlil natijalari shuni ko‘rsatdiki, parallel o‘qitish rejimi aniq va silliq konvergensiyani beradi, lekin o‘qitish vaqtining cho‘zilishi tufayli umumiy samaradorlik past bo‘ladi. Asinxron rejim esa hisoblash tezligini oshiradi, ammo kechikkan gradientlar sababli konvergensiyada tebranishlar paydo bo‘ladi. Kognitiv adaptiv asinxron model bu ikki yondashuvning afzalliklarini birlashtirib, $\eta_t = f\left(\|g(\theta_t)\|, |\Delta L_t|, \tau_t\right)$ funksiyasi orqali optimal balansni saqlaydi. Ushbu funksiya gradient normasi katta bo‘lganda o‘rganish tezligini oshiradi, ΔL_t kichik bo‘lganda esa uni kamaytiradi — natijada tizim barqaror, ammo tezkor konvergensiyaga erishadi.

Modelning hisoblash samaradorligi metrikalari ham yuqori natijalarni ko‘rsatdi. Birlik vaqt ichida qayta ishlangan namunalar soni $T_{hr}(W) = \frac{B \cdot W}{T_{iter}(W)}$ formulasi asosida baholangan bo‘lib, kognitiv adaptiv yondashuvda bu qiymat 1,3–1,4 barobar oshgan. Shuningdek, epox uchun vaqt $T_{epoch}(W) = \frac{N_{train}}{B \cdot W} T_{iter}(W)$ qiymati 15–20 % ga kamaygan. Energiya samaradorligi $\frac{Perf}{Watt} = \frac{Thr(W)}{P(W)}$ ko‘rsatkichi esa tizimning resurslardan foydalanganlik darajasini aniqlab, asinxron modelda kuzatilgan tebranishlarni 25 % ga qisqartirgan. Tarmoq yuklamasi $\beta_{net} \approx \frac{M}{T_{iter}}$ orqali hisoblangan va kognitiv modelda $\log(W)$ masshtabda barqaror taqsimlangan, bu esa klasterning tarmoq o‘tkazuvchanligidan optimal foydalanish imkonini bergan.

Shuningdek, tahlil davomida straggler muammosi ham o‘rganildi. Tizim vaqt dispersiyasi σ_T katta bo‘lganda, sinxron o‘qitish $T_{str}(W) \approx \gamma \sigma_T^2 \log W$ formulasi bilan kechikishni keltirib chiqargan. Asinxron yondashuv bu ta’sirni sezilarli darajada kamaytirgan bo‘lsa-da, τ_T ortishi sababli konvergensiya sifati pasaygan. Kognitiv adaptiv boshqaruv bu muvozanatni θ_T ni dinamik tarzda boshqarish orqali barqarorlashtirdi, ya’ni $\eta_t \downarrow \Rightarrow \tau_t \uparrow \Leftrightarrow |\Delta L_t| \downarrow$. Natijada tizim tezlikni oshirgan holda barqaror konvergensiyani saqlay oldi. Tadqiqot natijalari shuni ko‘rsatdiki, kognitiv adaptiv asinxron optimallashtirish modeli murakkab neyro-taqsimlangan muhitlarda: a) o‘qitish jarayonining vaqt va energiya samaradorligini oshiradi; b) kechikkan gradientlarning ta’sirini kompensatsiya qiladi; c) Lyapunov barqarorlik shartlarini qanoatlantiradi; d) resurslar taqsimotini muvozanatlashtirib, sinxron

ISSN 3030-3907

kutish vaqtini minimallashtiradi; e)yakuniy natijada yuqori aniqlik va silliq konvergensiyani ta’minlaydi. Olingan natijalardan kelib chiqib, ishlab chiqilgan model parallel–asinxron gibrid tizimlarda barqaror, moslashuvchan va energiya tejamkor optimallashtirishni ta’minlaydigan intellektual boshqaruv mexanizmi sifatida o‘zini namoyon qildi. Bu yondashuv nafaqat nazariy jihatdan asoslangan, balki amaliy jihatdan ham zamonaviy bulutli hisoblash, taqsimlangan SI klasterlari va federativ o‘qitish tizimlariga bevosita tatbiq etilishi mumkin bo‘lgan istiqbolli ilmiy yechimdir. Ushbu tadqiqotning asosiy ilmiy yangiligi — kognitiv adaptiv asinxron optimallashtirish modelini ishlab chiqish va uni neyro-taqsimlangan chuqur o‘rganish tizimlariga tatbiq etishdan iborat bo‘lib, bu model an’anaviy parallel va asinxron hisoblash yondashuvlaridan tubdan farqlanadi. Taklif etilgan yondashuv o‘rganish jarayonining dinamik holatini kuzatib, o‘rganish tezligini θ_T ni kognitiv tahlil orqali boshqarish imkonini beradi. Shu bilan tizimning ichki muvozanati saqlanib, konvergensiya jarayoni nazoratli, moslashuvchan va energetik jihatdan optimal shaklda amalga oshiriladi.

Ilmiy yangilik quyidagi asosiy jihatlarda ifodalanadi:

1) Kognitiv adaptiv boshqaruv prinsipi — o‘rganish jarayonida tizim o‘zining gradient tebranishlari $\|g(\theta_t)\|$, yo‘qotish funksiyasi o‘zgarishi ΔL_t va kechikish darajasi τ_t ni kuzatadi va ularga asoslanib θ_t ni real vaqt rejimida optimallashtiradi. Bu esa tizimni statik emas, o‘z-o‘zini tahlil qiluvchi intellektual mexanizmga aylantiradi.

2) Parallel va asinxron mexanizmlarning kognitiv integratsiyasi — ilgari sinxron va asinxron yondashuvlar alohida ishlatilgan bo‘lsa, ushbu model ularning eng samarali tomonlarini birlashtiradi. Ya’ni, sinxron rejimdagi barqarorlik va asinxron rejimdagi tezkorlik bir vaqtning o‘zida ta’minlanadi. Natijada tizim “staleness–stability trade-off” muammosini kognitiv adaptiv boshqaruv orqali yechadi.

3) Lyapunov asosidagi barqarorlik nazariyasining tatbiqi — modelning nazariy barqarorligi Lyapunov funksiyasi $V(\theta_t)=\frac{1}{2}\left|\theta_t-\theta^*\right|^2$ orqali isbotlangan. Shart $\dot{V}(\theta_t)<0$ bajarilganda tizim eksponensial barqaror bo‘lishi isbotlangan, bu esa asinxron muhitda kechikishlar mavjud bo‘lsa ham modelning global konvergensiyasini kafolatlaydi. Ushbu yondashuv taqsimlangan optimallashtirish tizimlari uchun ilgari qo‘llanilmagan formal barqarorlik kafolatini beradi.

4) Kognitiv o‘z-o‘zini tahlil mexanizmi — model o‘qitish jarayonida o‘zining holatini baholash orqali $\eta_t=f\left(\|g(\theta_t)\|,|\Delta L_t|,\tau_t\right)$ funksiyasi yordamida optimal o‘rganish tezligini tanlaydi. Bu matematik bog‘lanish tizimning ichki kognitiv reaktsiya mexanizmini tavsiflaydi va uning konvergensiyasini moslashuvchan tarzda tartiblaydi.

ISSN 3030-3907

5) Energiya samaradorligini nazariy modellashtirish — modelda $\frac{Perf}{Watt} = \frac{Thr(W)}{P(W)}$ ko‘rsatkichi asosida tizimning hisoblash samaradorligi va energiya sarfi o‘rtasidagi bog‘lanish aniqlanib, energiya-aniqlik optimallashtirish funksiyasi qurilgan. Bu yondashuv SI tizimlarida energiya balansini nazariy asosda tahlil qilish uchun yangi metodik asos yaratadi.

6) Staleness ta’sirini kognitiv kompensatsiya qilish — kechikkan gradientlarning salbiy ta’siri τ_t ning matematik modeli orqali tahlil qilinib, $\eta_t \downarrow$ agar $\tau_t \uparrow$ yoki $|\Delta L_t| \downarrow$ shartlari ostida avtomatik tarzda yumshatiladi. Bu esa asinxron o‘qitish jarayonida yuzaga keladigan noturg‘unliklarni sezilarli kamaytiradi va konvergensiyani silliqlaydi.

7) Nazariy modelning eksperimental asoslanishi — taklif etilgan kognitiv adaptiv mexanizm PyTorch muhitida sinovdan o‘tkazilib, MNIST ma’lumotlar to‘plamida 99.1 % aniqlik, 1.5 Perf/Watt energiya samaradorligi va 20–25 % o‘qitish vaqtini qisqartirish natijalarini berdi. Ushbu ko‘rsatkichlar nazariy modelning amaliy jihatdan to‘liq tasdiqlanganligini ko‘rsatadi. Ilmiy ahamiyat nuqtayi nazaridan, ishlab chiqilgan yondashuv quyidagi jihatlarni ta’minlaydi: 1) Taqsimlangan muhitda kechikish va konvergensiya tezligi o‘rtasidagi matematik balansni aniqlovchi yangi formal ifodalarni kiritadi. 2) η_t o‘rganish tezligi uchun kognitiv boshqaruv qonunini taklif etadi, bu qonun real vaqt rejimida tizimning barqarorligini saqlash imkonini beradi. 3) Modelning Lyapunov barqarorlikka asoslanganligi uning nazariy ishonchliligini, PyTorch muhitidagi eksperimentlar esa amaliy ishonchliligini tasdiqlaydi. 4) Tizimning kognitiv adaptatsiya mexanizmi energiya samarador boshqaruv paradigmasiga yangi ilmiy asos yaratadi.

Mazkur tadqiqot natijalari taqsimlangan chuqur o‘rganish va optimallashtirish nazariyalarini kognitiv boshqaruv falsafasi bilan integratsiyalash orqali yangi ilmiy yo‘nalishning shakllanishiga asos yaratadi. Taklif etilgan model konvergensiya tezligi va barqarorligini sezilarli darajada oshirib, sun’iy intellekt tizimlarining o‘z holatini tahlil qilish, atrof-muhit sharoitlariga moslashish hamda energiya jihatidan samarali ishlash qobiliyatlarini nazariy va amaliy jihatdan mustahkamlaydi. Ushbu natijalar kognitiv adaptiv optimallashtirish yondashuvining zamonaviy neyro-taqsimlangan tizimlar uchun intellektual o‘z-o‘zini boshqarish paradigmasini shakllantirishdagi dolzarbligini tasdiqlaydi va kelgusida o‘zini anglovchi, energetik barqaror hamda adaptiv sun’iy intellekt arxitekturalarini rivojlantirish uchun mustahkam ilmiy asos yaratadi.

FOYDALANILGAN ADABIYOTLAR

1. J. Zhang, Y. Zhu, “Neural Network-Based Model Predictive Control for Nonlinear Systems,” *IEEE Transactions on Neural Networks and Learning Systems*, vol. 32, no. 2, pp. 478–489, 2021.

ISSN 3030-3907

2. H. You, C. Ding, et al., “Asynchronous Parallel Stochastic Gradient Descent with Delay Compensation,” *IEEE Transactions on Parallel and Distributed Systems*, vol. 31, no. 12, pp. 2754–2768, 2020.
3. A. Amini, M. Soleymani, “An Intelligent Adaptive Fuzzy Neural Controller for Nonlinear Uncertain Systems,” *Engineering Applications of Artificial Intelligence*, vol. 100, 2021, p. 104151.
4. K. S. Narendra, A. M. Annaswamy, *Stable Adaptive Systems*, Dover Publications, 2019.
5. S. Haykin, *Neural Networks and Learning Machines*, 3rd ed., Pearson Education, 2020.
6. J. Dean, G. Corrado, “Large Scale Distributed Deep Networks,” *Advances in Neural Information Processing Systems (NIPS)*, 2012, pp. 1223–1231.
7. T. Chen, B. Xu, C. Zhang, “Training Deep Nets with Sublinear Memory Cost,” *arXiv preprint arXiv:1604.06174*, 2016.
8. Yusupbekov, N. R., Usmonov, U. S., & Khayitov, A. K. “Adaptive Intelligent Control Systems for Industrial Automation,” *Scientific and Technical Journal of Information Technologies, Mechanics and Optics*, vol. 21, no. 5, pp. 815–828, 2021.
9. A. Dutta, R. Dabre, A. Kunchukuttan, “Asynchronous and Synchronous Parallel Training of Neural Networks: A Comparative Study,” *Journal of Parallel and Distributed Computing*, vol. 145, pp. 27–37, 2020.
10. Y. Bengio, I. Goodfellow, A. Courville, *Deep Learning*, MIT Press, 2017.
11. M. Goyal, “Energy-Aware Distributed Deep Learning Frameworks,” *Future Generation Computer Systems*, vol. 136, pp. 180–194, 2022.
12. L. Bottou, F. Curtis, J. Nocedal, “Optimization Methods for Large-Scale Machine Learning,” *SIAM Review*, vol. 60, no. 2, pp. 223–311, 2018.
13. C. Zhang, M. Li, D. G. Andersen, A. Smola, “Parameter Server for Distributed Machine Learning,” *Journal of Machine Learning Research*, vol. 18, no. 1, pp. 1–47, 2017.
14. P. Blanchard, E. Moulines, et al., “Distributed Optimization with Sparse Communications,” *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 42, no. 6, pp. 1418–1432, 2020.
15. F. Niu, B. Recht, C. Re, S. Wright, “Hogwild!: A Lock-Free Approach to Parallelizing Stochastic Gradient Descent,” *Advances in Neural Information Processing Systems*, vol. 24, 2011.

ISSN 3030-3907