

SUN'IY NEYRON TO'RLARI VA ULARNING BIOLOGIK NEYRONLARDAN FARQI

Tojimamatov Isroil Nurmatovich

Farg'ona davlat universiteti Amaliy matematika va

informatika kafedrası katta o'qituvchisi

E-mail: israiltrojimamatov@gmail.com

Erkinova Niginabonu Hamidulla qizi

Farg'ona davlat universiteti talabasi

E-mail: enhahaneem2006@gmail.com

Annotatsiya: Ushbu maqolada sun'iy neyron tarmoqlarning rivojlanishi, ularning energiya sarfi va hisoblash resurslariga bo'lgan ehtiyoji muhokama qilinadi. Katta hajmdagi modellar (masalan, GPT, LLaMA) yuqori hisoblash quvvatini talab qilishi va ekologik muammolarni yuzaga keltirishi ko'rsatildi. Maqolada energiya samaradorligini oshirish yo'llari sifatida modelni siqish va kvantizatsiya texnikalari tahlil qilinadi.

Kvantizatsiya yordamida model parametrlarini kichikroq raqamli formatlarga o'tkazish orqali xotira talabini va energiya iste'molini sezilarli darajada kamaytirish mumkinligi isbotlandi. Ushbu yondashuvlar nafaqat texnik, balki iqtisodiy va ekologik jihatdan ham foydali bo'lib, sun'iy intellektning barqaror rivojlanishiga xizmat qiladi.

Аннотация: В данной статье рассматривается развитие искусственных нейронных сетей, их потребность в вычислительных ресурсах и энергопотреблении. Показано, что крупные модели (например, GPT, LLaMA) требуют высокой вычислительной мощности и вызывают экологические проблемы. В статье анализируются методы повышения энергоэффективности, такие как сжатие моделей и квантование. Было доказано, что с помощью квантования параметров модели в более низкие числовые форматы можно значительно уменьшить требования к памяти и энергопотреблению. Эти подходы полезны не только с технической точки

зрения, но и с экономической и экологической, способствуя устойчивому развитию искусственного интеллекта.

Abstract: This article discusses the development of artificial neural networks, their computational resource requirements, and energy consumption. It highlights that largescale models (such as GPT and LLaMA) demand high computational power and pose environmental challenges. The article analyzes energy efficiency improvement methods, specifically model compression and quantization techniques. It is demonstrated that quantizing model parameters into lower-bit numerical formats significantly reduces memory requirements and energy consumption. These approaches are beneficial not only technically but also economically and environmentally, contributing to the sustainable development of artificial intelligence.

Kalit so'zlar: Sun'iy neyron tarmoqlar, energiya samaradorligi, modelni siqish, kvantizatsiya texnikasi, hisoblash resurslari, energiya tejash, ekologik iz, barqaror rivojlanish, yuqori parametrli modellar.

Ключевые слова: Искусственные нейронные сети, энергоэффективность, сжатие модели, техника квантования, вычислительные ресурсы, экономия энергии, экологический след, устойчивое развитие, модели с большим количеством параметров.

Keywords: Artificial Neural Networks, energy efficiency, model compression, quantization technique, computational resources, energy saving, environmental impact, sustainable development, high-parameter models.

KIRISH

Bugungi kunda sun'iy intellekt (AI) texnologiyalari hayotimizning deyarli barcha sohalariga chuqur kirib bormoqda. Sun'iy neyron tarmoqlari esa aynan shu texnologiyalarning eng muhim va faol rivojlanayotgan qismidir. Ular biologik neyronlarning asosiy printsiplariga asoslangan bo'lib, murakkab ma'lumotlarni tahlil qilish, avtomatlashtirilgan qaror qabul qilish va ko'plab boshqa sohalarda keng qo'llanilmoqda. Biroq, sun'iy neyron tarmoqlarining o'ziga xosligi va ularning biologik

neyronlardan farqlari texnologiyaning samaradorligi va energiya sarfi masalalarida muhim rol o'ynaydi.

Sun'iy neyron tarmoqlarining murakkabligi va ulardan talab qilinayotgan hisoblash resurslari har yili ortib bormoqda. Masalan, hozirgi zamon yirik modellar, masalan GPT yoki LLaMA kabi trillionlab parametrga ega bo'lib, ularga ishlov berish uchun katta miqdorda hisoblash kuchi va energiya talab etiladi. Ushbu yuqori energiya sarfi nafaqat iqtisodiy xarajatlarni oshiradi, balki ekologik muammolarni ham keltirib chiqaradi. Shu sababli, sun'iy neyron tarmoqlarning samaradorligini oshirish va ularning energiya iste'molini kamaytirish bugungi kunda juda dolzarb muammo hisoblanadi.

Biologik neyronlar esa asosan tabiiy evolyutsiya jarayonida shakllangan va energiya sarfi jihatidan juda tejamkor tizimdir. Ularning asosiy vazifasi signallarni tez va samarali uzatishdir, bu esa sun'iy neyron tarmoqlarining arxitekturasini va ishlash printsipi bilan bir qator farqlarga ega. Shu bois, biologik neyronlarning xususiyatlarini o'rganish va ularni sun'iy tizimlarga samarali tatbiq etish orqali yanada qulay, energiya tejamkor va yuqori samarali neyron tarmoqlarni yaratish imkoniyati mavjud.

Ushbu maqolada sun'iy neyron tarmoqlarining asosiy tushunchalari, ularning biologik neyronlardan farqlari, energiya samaradorligi masalalari va zamonaviy texnologiyalardagi energiya tejash usullari, xususan, modelni siqish va kvantizatsiya texnikasi tahlil qilinadi. Shu orqali kelajakda yanada rivojlangan va ekologik jihatdan toza sun'iy intellekt tizimlarini yaratishga qaratilgan muhim yondashuvlar ko'rsatib o'tiladi.

METODOLOGIYA

Ushbu maqolada sun'iy neyron tarmoqlari samaradorligini oshirish va ularning energiya sarfini kamaytirish bo'yicha zamonaviy yondashuvlar tahlil qilinadi. Asosiy e'tibor katta hajmdagi neyron tarmoqlarda yuzaga keladigan energiya va hisoblash resurslari talabining oshishi muammosiga qaratilgan. Metodologiya doirasida quyidagi yondashuvlar ko'rib chiqildi:

Tahliliy tadqiqot va adabiyotlarni o'rganish bo'lib, u sun'iy neyron tarmoqlarining ishlash printsipi, ularning hisoblash murakkabligi va energiya talabini

o'rganish uchun so'nggi ilmiy maqolalar, texnik hisobotlar va amaliy ishlanmalar tahlil qilindi. Xususan, yirik modellar (GPT, LLaMA kabi) va ularning hisoblash resurslariga bo'lgan talabi, shuningdek, bu talabning atrof-muhitga ta'siri o'rganildi.

Model siqish texnologiyalarini o'rganish bo'lib, u katta hajmdagi neyron tarmoqlarni energiya tejovchi qilib optimallashtirishda samarali usullardan biri sifatida **modelni siqish** va ayniqsa **kvantizatsiya** texnikasi tanlandi. Bu texnologiyalar yordamida model parametrlarini 32-bit suzuvchi nuqtali formatdan kichikroq (masalan, 8-bit yoki 4bit) raqamli ifodalarga o'tkazish imkoniyati tahlil qilindi.

Energiya samaradorligini baholash bo'lib, u kvantizatsiya jarayonining modelning arxitekturasi va aniqligiga ta'siri, shuningdek, hisoblash samaradorligi va energiya iste'moli kamayishining amaliy ta'siri matematik va eksperimental asosda o'rganildi. Natijalar GPU/TPU qurilmalari ishining soddalashuvi, serverlarning issiqlik chiqishi va energiya sarfini kamaytirishga qanday ta'sir qilishi ko'rsatildi.

Amaliy misollar va tahlil bo'lib, u 1 milliard parametrga ega neyron tarmoq misolida kvantizatsiya texnikasining xotira talabini 4 barobar kamaytirishi va shu bilan birga energiya iste'molini sezilarli darajada pasaytirishi amaliy ko'rsatkichlari asosida tahlil qilindi. Bu misol texnik va iqtisodiy samaradorlikni baholashda asos bo'ldi.

Iqtisodiy va ekologik ta'sirni o'rganish bo'lib, u katta hajmdagi neyron tarmoqlarning energiya iste'moli tufayli yuzaga keladigan iqtisodiy xarajatlar va ekologik iz (karbon izlari) tahlil qilindi. Shuningdek, energiya samaradorligini oshirishga qaratilgan texnologiyalarni joriy etishning jamiyat va sanoatga foydasi ko'rib chiqildi.

Shu tariqa, metodologiya sun'iy neyron tarmoqlarning energiya samaradorligini oshirishda kvantizatsiya texnikasi kabi ilg'or optimallashtirish usullarini amaliy jihatdan o'rganishga qaratilgan. Bu yondashuvlar hozirgi kun talablariga mos keluvchi samarali va iqtisodiy jihatdan maqbul yechimlarni ishlab chiqishga xizmat qiladi.

NATIJA VA MUHOKAMA

Zamonaviy sun'iy neyron tarmoqlarning o'sishi ularning imkoniyatlarini sezilarli darajada kengaytirgan bo'lsa-da, hisoblash resurslari va energiya sarfi jihatidan jiddiy

muammolarni yuzaga keltirmoqda. Tadqiqotlar shuni ko'rsatadiki, yirik modellar, masalan, trillionlab parametrga ega bo'lgan GPT va LLaMA, ko'p energiya talab qiladigan hisoblash qurilmalariga — GPU va TPU klasterlariga bog'liq. Bu nafaqat energiya xarajatlarini oshiradi, balki katta issiqlik chiqarishi tufayli serverlarni sovutish tizimlariga qo'shimcha yuk tushadi, bu esa ekologik izni yanada kuchaytiradi.

Tahlil natijalari shuni ko'rsatdiki, modelni siqish va ayniqsa, kvantizatsiya texnikasi yirik neyron tarmoqlarning energiya samaradorligini sezilarli darajada yaxshilashga yordam beradi. Kvantizatsiya orqali 32-bitlik suzuvchi nuqtali formatdagi model parametrlarini 8-bit yoki hatto 4-bit formatga o'tkazish mumkin bo'lib, bu nafaqat xotira talabini 4-8 barobar kamaytiradi, balki hisoblash jarayonlarini sezilarli darajada yengillashtiradi. Natijada, model ishlash tezligi oshadi va energiya iste'moli kamayadi — ba'zi tadqiqotlarda bu kamayish 6-7 barobar darajasida kuzatilgan.

Bunday optimallashtirishlar serverlarning issiqlik chiqarishini kamaytiradi, bu esa sovutish tizimlari uchun sarflanadigan energiyani ham qisqartiradi. Shu bilan birga, yengil model serverlarga kamroq yuk tushishini ta'minlaydi, bu esa infratuzilma xarajatlarini kamaytirishda muhim rol o'ynaydi.

Tadqiqotda berilgan 1 milliard parametrli neyron tarmoq misoli bu jarayonni aniq ko'rsatadi: 32-bit formatdagi model taxminan 4 GB xotira talab etsa, kvantizatsiya natijasida bu talab 1 GB ga tushadi. Bu nafaqat xotira tejash, balki energiya sarfini sezilarli kamaytirishga olib keladi, chunki GPU kamroq transistorlarni ishga soladi va ma'lumotlarni tezroq qayta ishlaydi.

Shuningdek, ushbu metodologiya iqtisodiy jihatdan ham samarali ekanligi ko'rsatildi. Yuqori energiya iste'moli talab qiladigan yirik neyron tarmoqlar uchun zarur bo'lgan GPU klasterlarining ishlashi ancha qimmatga tushadi, shu bois kvantizatsiya kabi texnologiyalarni qo'llash xarajatlarni kamaytirishga yordam beradi.

Ammo kvantizatsiya texnikasining ba'zi cheklovlari ham mavjud. Model aniqligiga ta'sir qilmasligi uchun parametrlarning aniqlik darajasini me'yorda kamaytirish zarur. Shu bilan birga, ba'zi murakkab tarmoqlarda to'liq kvantizatsiya qo'llanilishi qiyinchilik tug'dirishi mumkin, bu esa keyingi tadqiqotlar va texnologik yangilanishlarni talab qiladi.

Umuman olganda, sun'iy neyron tarmoqlarining energiya samaradorligini oshirish yo'lida kvantizatsiya texnikasi samarali va amaliy yechim sifatida keng qo'llanilmoqda. Bu nafaqat texnik, balki iqtisodiy va ekologik muammolarni kamaytirishda muhim rol o'ynaydi va sun'iy intellekt texnologiyalarining barqaror rivojlanishini ta'minlashga yordam beradi.

XULOSA

Bugungi kunda sun'iy neyron tarmoqlarning tez sur'atlar bilan rivojlanishi va keng qo'llanilishi, ularning hisoblash resurslari va energiya sarfi jihatidan jiddiy muammolarni yuzaga keltirmoqda. Katta hajmdagi modellar, xususan trillionlab parametrga ega bo'lganlar, yuqori darajada hisoblash quvvatini talab qiladi, bu esa iqtisodiy xarajatlarni oshiribgina qolmay, ekologik muammolarni ham kuchaytiradi. Shu bois, sun'iy neyron tarmoqlarning energiya samaradorligini oshirish, ularning ekologik izini kamaytirish va infratuzilma xarajatlarini optimallashtirish dolzarb vazifa hisoblanadi.

Ushbu maqolada ko'rib chiqilgan kvantizatsiya va modelni siqish texnikalari — ayniqsa, energiya tejovchi yondashuv sifatida — bu muammoni hal qilishda samarali vosita ekanligi isbotlandi. Parametrlarni 32-bitdan kichikroq raqamli formatlarga o'tkazish nafaqat xotira talabini kamaytiradi, balki GPU va TPU qurilmalarining ish yukini yengillashtirib, energiya iste'molini sezilarli darajada pasaytiradi. Bu esa, o'z navbatida, serverlarning issiqlik chiqarishini kamaytirib, sovutish tizimlariga ketadigan energiyani ham qisqartiradi.

Bundan tashqari, kvantizatsiya texnikasi yordamida yaratilgan yengil modellarning ishlash samaradorligi oshib, javob berish tezligi yaxshilanadi, shu bilan birga, hisoblash infratuzilmasi uchun zarur bo'lgan xarajatlar sezilarli darajada kamayadi. Shu bilan birga, ushbu texnologiya ekologik barqarorlikni ta'minlashda ham muhim rol o'ynaydi.

Albatta, kvantizatsiyaning ba'zi cheklovlari mavjud bo'lib, ularni yengish uchun yanada chuqurroq tadqiqotlar va texnologik rivojlanish talab etiladi. Biroq, hozirgi kunda energiya samaradorligini oshirish borasidagi yechimlar sun'iy intellektning yanada barqaror va mas'uliyatli rivojlanishiga xizmat qilmoqda.

Kelajakda yanada samarali, energiya tejovchi va ekologik jihatdan xavfsiz sun'iy neyron tarmoqlarini yaratishda ushbu optimallashtirish usullarining roli yanada oshadi. Shu tariqa, ular ilmiy tadqiqotlar, sanoat va kundalik hayotda keng ko'lamda qo'llanilib, texnologik taraqqiyotga katta hissa qo'shadi.

FOYDALANILGAN ADABIYOTLAR:

1. **Goodfellow, I., Bengio, Y., Courville, A.** *Deep Learning*. MIT Press, 2016. — Sun'iy neyron tarmoqlarning nazariy asoslari, struktura va trening jarayonlari bo'yicha asosiy manba.
2. **Jacob, B., Kligys, S., Chen, B., Zhu, M., Tang, M., Howard, A., Adam, H., Kalenichenko, D.** *Quantization and Training of Neural Networks for Efficient IntegerArithmetic-Only Inference*. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2018, pp. 2704-2713. — Kvantizatsiya texnikalari va ularning samaradorligini oshirish usullari haqida.
3. **Han, S., Mao, H., Dally, W.J.** *Deep Compression: Compressing Deep Neural Networks with Pruning, Trained Quantization and Huffman Coding*. International Conference on Learning Representations (ICLR), 2016. — Modelni siqish va energiya tejash bo'yicha amaliy yondashuvlar.
4. **Strubell, E., Ganesh, A., McCallum, A.** *Energy and Policy Considerations for Deep Learning in NLP*. Proceedings of the 57th Annual Meeting of the Association for Computational Linguistics, 2019. — NLP sohasidagi yirik neyron tarmoqlarning energiya sarfi va ekologik ta'siri tahlili.
5. **Sze, V., Chen, Y.H., Yang, T.J., Emer, J.S.** *Efficient Processing of Deep Neural Networks: A Tutorial and Survey*. Proceedings of the IEEE, 2017, Vol. 105, No. 12, pp. 2295-2329. — Neyron tarmoqlarni samarali hisoblash va energiya tejash bo'yicha umumiy ko'rsatmalar.
6. **Gholami, A., Kim, S., Yao, Z., et al.** *A Survey of Quantization Methods for Efficient Neural Network Inference*. arXiv preprint arXiv:2103.13630, 2021. — Kvantizatsiya usullarining keng qamrovli tahlili.

7. **ChatGPT (OpenAI).** Sun'iy neyron tarmoqlar va kvantizatsiya texnikasi bo'yicha zamonaviy yondashuvlar va energiya samaradorligi haqida maslahatlar va tushuntirishlar. 2025. — Sun'iy intellektning amaliy va nazariy jihatlari bo'yicha keng qamrovli manba.
8. Tursunaliyeva M.Z. Optimization of network traffic with the help of artificial intelligence based on edge computing: a new methodological approach // Science and Innovation. International Scientific Journal. - 2021. - Vol. 4, No. 9. - P. 76-79. - doi:10.5281/zenodo.17258224.
9. Tursunaliyeva M. Energy efficiency in neural networks: problems of optimizing large models // Science and Innovation. International Scientific Journal. - 2021. - Vol. 4, No. 11.
- P. 59-61. - DOI: 10.5281/zenodo.17674536.