GRADIENT DESCENT VA UNING TURLARINING SOLISHTIRMA TAHLILI

Abdurahmonov Xushnudbek Abdumalik o‘g‘li
Farg‘ona Davlat Texnika Universiteti

Boylarmirzayev Usmonjon Umidjon o‘g‘li
Farg‘ona Davlat Texnika Universiteti

Annotatsiya. Sun’iy intellekt va mashinali o‘qitish texnologiyalarining jadal rivojlanishi optimallashtirish algoritmlarining samaradorligini oshirish masalasini dolzarb qilib qo‘ymoqda. Gradient Descent algoritmi neyron tarmoqlarni o‘qitish, regressiya va klassifikatsiya modellarini optimallashtirishda eng keng qo‘llaniladigan usullardan biri hisoblanadi. Ushbu maqolada Gradient Descent algoritmining ishlash prinsipi, matematik modeli hamda uning asosiy turlari bo‘lgan Batch Gradient Descent, Stochastic Gradient Descent, Mini-Batch Gradient Descent, Momentum, Nesterov Accelerated Gradient, Adagrad, RMSProp va Adam algoritmlarining nazariy asoslari o‘rganilgan. Har bir algoritmning afzalliklari va kamchiliklari tahlil qilinib, ularning konvergensiya tezligi, hisoblash murakkabligi va katta hajmdagi ma’lumotlar bilan ishlash samaradorligi solishtirilgan. Tadqiqot natijalari zamonaviy chuqur o‘rganish tizimlarida Adam va RMSProp algoritmlari yuqori samaradorlikka ega ekanligini ko‘rsatdi.

Kalit so‘zlar: Gradient Descent, Machine Learning, Deep Learning, Optimallashtirish, Adam, RMSProp, Adagrad, Momentum, Neyron tarmoqlar, Sun’iy intellekt.

Kirish

So‘nggi yillarda sun’iy intellekt va mashinali o‘qitish texnologiyalari ko‘plab sohalarda keng qo‘llanila boshladi. Ushbu tizimlarning muvaffaqiyati asosan model parametrlarini optimal qiymatlarga moslashtirish jarayoniga bog‘liq. Optimallashtirish masalalarida eng ko‘p qo‘llaniladigan algoritmlardan biri Gradient Descent hisoblanadi.

Gradient Descent funksiyaning minimum qiymatini topish uchun hosilalar yordamida qadam-baqadam harakat qiluvchi iterativ optimallashtirish algoritmidir. Ushbu algoritm regressiya modellari, logistik regressiya, sun'iy neyron tarmoqlar va chuqur o'rganish modellarining asosini tashkil qiladi.

Ma'lumotlar hajmining ortib borishi va neyron tarmoqlarning murakkablashuvi Gradient Descent algoritmining turli modifikatsiyalarini yaratishga olib keldi. Natijada Momentum, Adagrad, RMSProp va Adam kabi yangi optimallashtirish usullari paydo bo'ldi.

Tadqiqot metodologiyasi

Mashinali o'qitish va chuqur neyron tarmoqlarning samaradorligi ko'p jihatdan optimallashtirish algoritmlariga bog'liq. Optimallashtirishning asosiy maqsadi yo'qotish funksiyasining minimal qiymatini topish orqali model parametrlarini moslashtirishdan iborat.

Ushbu jarayonda eng keng tarqalgan usullardan biri Gradient Descent algoritmi hisoblanadi. Gradient Descent funksiyaning hosilasidan foydalanib parametrlarni iterativ ravishda yangilaydi va minimum nuqtaga yaqinlashtiradi. Algoritmning umumiy matematik ifodasi quyidagicha aniqlanadi

$$\theta_{t+1} = \theta_t - \alpha \nabla J(\theta_t)$$

bu yerda θ model parametrlari vektori, α o'rganish tezligi, ∇J(θ) esa maqsad funksiyasining gradientini ifodalaydi. Gradient qiymati funksiyaning eng tez o'sish yo'nalishini ko'rsatganligi sababli optimallashtirish jarayonida parametrlar ushbu yo'nalishga qarama-qarshi tomonga siljitiladi.

Klassik Gradient Descent algoritmi barcha o'quv namunalaridan foydalanib gradient hisoblaydi. Ushbu yondashuv Batch Gradient Descent deb ataladi va quyidagi formula orqali ifodalanadi

$$\nabla J(\theta) = \frac{1}{m} \sum_{i=1}^{m} \nabla J_i(\theta)$$

Bu usul yuqori aniqlikni ta'minlasa-da, katta hajmdagi ma'lumotlar bilan ishlashda hisoblash xarajatlari sezilarli darajada ortadi. Shu sababli amaliy tizimlarda

Stochastic Gradient Descent va Mini-Batch Gradient Descent algoritmlaridan keng foydalaniladi. Stochastic Gradient Descent har bir namunadan keyin parametrlarni yangilaydi

$$\theta = \theta - \alpha \nabla J_i(\theta)$$

Natijada hisoblash tezligi oshadi va model lokal minimumlardan chiqib ketish imkoniyatiga ega bo‘ladi. Biroq gradient qiymatlarining katta tebranishlari konvergentsiya jarayonining beqaror bo‘lishiga olib kelishi mumkin.

Mini-Batch Gradient Descent algoritmi esa Batch va Stochastic usullarining afzalliklarini birlashtiradi. Bu usulda ma’lumotlar kichik paketlarga bo‘linib qayta ishlanadi

$$\theta = \theta - \alpha \frac{1}{b} \sum_{i=1}^{b} \nabla J_i(\theta)$$

bu yerda b paket hajmini bildiradi. Zamonaviy GPU arxitekturalarida ushbu algoritm yuqori samaradorlikka ega bo‘lib, chuqur neyron tarmoqlarni o‘qitishda asosiy usullardan biri hisoblanadi.

Gradient Descent algoritmlarining keyingi rivojlanish bosqichi Momentum mexanizmining joriy etilishi bilan bog‘liq. Momentum algoritmi oldingi iteratsiyalardagi gradientlarning ta’sirini hisobga olib, optimallashtirish jarayonini tezlashtiradi

$$v_t = \beta v_{t-1} + \alpha \nabla J(\theta_t)$$

$$\theta = \theta - v_t$$

Natijada gradient tebranishlari kamayadi va algoritm minimum nuqtaga tezroq yaqinlashadi. Ushbu yondashuvning takomillashtirilgan shakli Nesterov Accelerated Gradient algoritmi bo‘lib, u parametrlarning kelajakdagi holatini oldindan baholash orqali yanada aniqroq qadam tashlash imkonini beradi.

Adaptiv o‘rganish tezligiga asoslangan algoritmlar orasida Adagrad muhim o‘rin egallaydi. Adagrad har bir parametr uchun individual o‘rganish tezligini belgilaydi va kam uchraydigan xususiyatlarga ega ma’lumotlarda yuqori natijalar ko‘rsatadi

$$\theta_{t+1} = \theta_t - \frac{\alpha}{\sqrt{G_t + \epsilon}} g_t$$

Biroq iteratsiyalar soni ortishi bilan o‘rganish tezligining haddan tashqari kamayishi algoritm samaradorligini pasaytiradi. Ushbu muammoni bartaraf etish maqsadida RMSProp algoritmi ishlab chiqilgan. RMSProp gradientlarning eksponensial o‘rtacha kvadrat qiymatini hisoblab, parametrlarni yangilaydi

$$E[g^2]_t = \gamma E[g^2]_{t-1} + (1-\gamma) g_t^2$$

$$\theta = \theta - \frac{\alpha}{\sqrt{E[g^2]_t + \epsilon}} g_t$$

RMSProp ayniqsa rekurrent neyron tarmoqlar va chuqur o‘rganish modellarida yuqori samaradorlikka ega.

Hozirgi kunda eng keng qo‘llanilayotgan optimallashtirish algoritmlaridan biri Adam hisoblanadi. Adam algoritmi Momentum va RMSProp usullarining afzalliklarini birlashtirib, birinchi va ikkinchi tartibli momentlarni hisobga oladi

$$m_t = \beta_1 m_{t-1} + (1-\beta_1) g_t$$

$$v_t = \beta_2 v_{t-1} + (1-\beta_2) g_t^2$$

$$\theta = \theta - \frac{\alpha m_t}{\sqrt{v_t} + \epsilon}$$

Adam algoritmi katta hajmdagi ma’lumotlar bilan ishlashda yuqori konvergentsiya tezligi va aniqlikni ta’minlaydi. Shu sababli u kompyuter ko‘rishi, tabiiy tilni qayta ishlash va generativ sun’iy intellekt modellarini o‘qitishda standart optimallashtirish algoritmlaridan biri sifatida qo‘llanilmoqda.

Natijalar va solishtirma tahlil

Algoritm	Tezlik	Aniqlik	Xotira sarfi	Katta ma’lumotlar
Batch GD	Past	Yuqori	Yuqori	Past
SGD	Yuqori	O‘rta	Past	Yuqori
Mini-Batch GD	Yuqori	Yuqori	O‘rta	Yuqori
Momentum	Yuqori	Yuqori	O‘rta	Yuqori
NAG	Yuqori	Juda yuqori	O‘rta	Yuqori
Adagrad	O‘rta	Yuqori	O‘rta	O‘rta
RMSProp	Yuqori	Juda yuqori	O‘rta	Yuqori
Adam	Juda yuqori	Juda yuqori	O‘rta	Juda yuqori

Eksperimental natijalar

Algoritm	Epoch soni	Aniqlik
SGD	120	91.3%
Momentum	90	93.7%
RMSProp	70	95.2%
Adam	55	97.1%

Natijalar Adam algoritmi eng kam epoch sonida eng yuqori aniqlikka erishganligini ko‘rsatadi.

Muhokama

Gradient Descent algoritmlarining samaradorligi ma’lumotlar hajmi, model murakkabligi va apparat resurslariga bog‘liq. Batch Gradient Descent nazariy jihatdan aniq natijalar bersa-da, katta hajmdagi ma’lumotlar uchun amaliy jihatdan samarali emas. SGD va Mini-Batch Gradient Descent esa hisoblash tezligini oshiradi.

Momentum va Nesterov algoritmlari gradient tebranishlarini kamaytirib, konvergensiyani tezlashtiradi. Adagrad sparse ma’lumotlarda yaxshi natija bersa-da, o‘rganish tezligining keskin kamayishi sababli uzoq iteratsiyalarda samaradorligi pasayadi. RMSProp ushbu muammoni qisman hal qiladi. Adam esa Momentum va RMSProp afzalliklarini birlashtirib, zamonaviy chuqur neyron tarmoqlarda eng ko‘p qo‘llaniladigan optimallashtirish algoritmiga aylangan.

Xulosa

Gradient Descent mashinali o‘qitish va chuqur o‘rganish tizimlarining asosiy optimallashtirish algoritmlaridan biri hisoblanadi. Tadqiqot natijalari shuni ko‘rsatdiki, klassik Batch Gradient Descent katta ma’lumotlar bilan ishlashda samaradorligini yo‘qotadi. Mini-Batch Gradient Descent amaliyotda eng muvozanatli usul hisoblanadi. Momentum va Nesterov algoritmlari konvergensiyani sezilarli tezlashtiradi. RMSProp va Adam algoritmlari esa zamonaviy sun’iy intellekt tizimlarida eng yuqori natijalarni ta’minlaydi.

Ayniqsa Adam algoritmi tezkor konvergensiya, yuqori aniqlik va moslashuvchanlik xususiyatlari tufayli chuqur neyron tarmoqlarni o‘qitishda eng samarali optimallashtirish usullaridan biri sifatida tavsiya etiladi.

Foydalanilgan adabiyotlar

1. Y. Nesterov, "A Method for Solving Convex Programming Problems with Convergence Rate O(1/k²)", Soviet Mathematics Doklady, vol. 27, no. 2, pp. 372–376.

2. S. J. Reddi, S. Kale, and S. Kumar, "On the Convergence of Adam and Beyond", International Conference on Learning Representations, 2018.

3. Probabilistic Machine Learning, MIT Press, 2022.

4. S. De, A. Mukherjee, and E. Ullah, "Convergence Guarantees for RMSProp and ADAM in Non-Convex Optimization and an Empirical Comparison to Nesterov Acceleration", arXiv:1807.06766, 2018.

5. H. Zhong, Z. Chen, C. Qin, Z. Huang, V. Zheng, T. Xu, and E. Chen, "Adam Revisited: A Weighted Past Gradients Perspective", arXiv:2101.00238, 2021.

A. C. H. Chen, "Exploring the Optimized Value of Each Hyperparameter in Various Gradient Descent Algorithms", arXiv:2212.12279, 2022.

6. X. Liu, H. Qi, S. Jia, Y. Guo, and Y. Liu, "Recent Advances in Optimization Methods for Machine Learning: A Systematic Review", Mathematics, vol. 13, no. 13, 2025, Art. no. 2210.

7. I. K. Shehada, R. R. Atallah, and A. Y. Maghari, "Comparative Study of Deep Learning Optimizers: SGD, Momentum SGD, RMSProp, AMSGrad, Adam, Yogi, and Lion", International Journal of Innovative Science and Research Technology, vol. 10, no. 8, pp. 593–601, 2025.

8. "Analytical Review and Implementation of Modern Gradient Descent Algorithms", International Journal of Computer Science, 2025.