

SOHANI RAQAMLASHTIRISHDA SUN'IY INTELLEKTNI QO'LLASHNING MATEMATIK YECHIMLARI

Xidirov Bobir Baxirovich

Toshkent Amaliy fanlar Universiteti, Gavhar ko'chasi 1, Toshkent 100131, O'zbekiston

E-mail: xbobiras1982@gmail.com, <https://doi.org/10.5281/zenodo.10442597>

Annotatsiya

Ushbu ilmiy maqolada turli sohalarni raqamlashtirish jarayonida sun'iy intellekt (SI) va mashinali o'rganish (ML) algoritmlarining matematik asosi tadqiq etiladi. Maqolada chiziqli algebra, ehtimollar nazariyasi, matematik statistika va optimallashtirish usullarining sanoat, moliya va logistika sohalaridagi amaliy tadbiri tahlil qilingan. Shuningdek, xatolik funksiyasini minimallashtirish va neyron tarmoqlarining matematik arxitekturasini bayon etilgan.

Kalit so'zlar: Raqamlashtirish, Sun'iy intellekt, Gradient tushishi, Neyron tarmoqlar, Regressiya, Optimallashtirish, Algoritm, Big Data.

Annotation

This scientific article investigates the mathematical foundations of artificial intelligence (AI) and machine learning (ML) algorithms in the process of digitalization across various fields. The study analyzes the practical applications of linear algebra, probability theory, mathematical statistics, and optimization methods in industry, finance, and logistics. Furthermore, the paper discusses the minimization of loss functions and the mathematical architecture of neural networks.

Keywords: Digitalization, Artificial Intelligence, Gradient Descent, Neural Networks, Regression, Optimization, Algorithm, Big Data.

Аннотация

В данной научной статье исследуются математические основы алгоритмов искусственного интеллекта (ИИ) и машинного обучения (МО) в процессе цифровизации различных сфер. В работе анализируются практические применения линейной алгебры, теории вероятностей, математической статистики и методов оптимизации в промышленности, финансах и логистике. Кроме того, в статье рассматриваются вопросы минимизации функций потерь и математическая архитектура нейронных сетей.

Ключевые слова: Цифровизация, Искусственный Интеллект, Градиентный Спуск, Нейронные Сети, Регрессия, Оптимизация, Алгоритм, Большие Данные.

KIRISH

Raqamlashtirish – bu shunchaki qog‘oz hujjatlarni elektron shaklga o‘tkazish emas, balki boshqaruv va ishlab chiqarish jarayonlarini ma’lumotlarga asoslangan algoritmlar yordamida optimallashtirishdir. Sun’iy intellekt (SI) ushbu jarayonning "miyasi" hisoblanadi. Sining samaradorligi uning qaysi dasturda yozilganida emas, balki uning asosida yotgan matematik modellarning aniqligidadir. Ushbu maqolaning dolzarbligi shundaki, hozirgi kunda iqtisodiyotning real sektorlarida (energetika, transport, qishloq xo‘jaligi) SI algoritmlari orqali xarajatlarni 20-25% gacha qisqartirish imkoniyati paydo bo‘ldi. Hozirgi kunda iqtisodiyotning barcha tarmoqlarini raqamli transformatsiya qilish (*Digital Transformation*) davlat siyosati darajasiga ko‘tarildi. Raqamlashtirishning asosiy maqsadi – inson omili ishtirokini kamaytirish, jarayonlarni tezlashtirish va iqtisodiy samaradorlikni oshirishdir. Sun’iy intellekt bu jarayonda "aqlli qaror qabul qiluvchi" vosita vazifasini o‘taydi. Biroq, SI tizimlarining samaradorligi ularning matematik poydevori qanchalik to‘g‘ri qurilganiga bog‘liq. Raqamlashtirish jarayonida birinchi navbatda katta hajmdagi tartibsiz ma’lumotlar (Big Data) bilan ishlashga to‘g‘ri keladi.

Ma’lumotlarni qayta ishlashning matematik modellari. Raqamlashtirish jarayonida birinchi navbatda katta hajmdagi tartibsiz ma’lumotlar (*Big Data*) bilan ishlashga to‘g‘ri keladi. SI modellari kirish ma’lumotlarining masshtabiga juda sezgir. Masalan, agar bir parametr (masalan, narx) minglarda, ikkinchisi (masalan, miqdor) birliklarda bo‘lsa, model narxga ko‘proq urg‘u berib yuboradi. Buning oldini olish uchun *Min-Max Scaling* usuli qo‘llaniladi:

$$X_{norm} = \frac{X - X_{min}}{X_{max} - X_{min}}.$$

Bu formula barcha ma’lumotlarni [0;1] oralig‘iga keltiradi, bu esa hisoblash resurslarini tejashga va algoritmnining tezroq yaqinlashishiga (convergence) yordam beradi. Ma’lumotlarni normalizatsiya qilish – bu sun’iy intellekt modellarining "sog‘lom" ishlashi uchun eng muhim bosqichlardan biridir. Keling, ushbu bo‘limni ilmiy maqola talablariga mos ravishda chuqur yoritib, aniq bir amaliy masala misolida ko‘rib chiqamiz.

Masalan, *ma’lumotlarni filtrlash va normalizatsiya qilish*. Sun’iy intellekt modellari, xususan, Gradient tushishi (Gradient Descent) va K-yaqin qo‘shnilar (KNN) kabi algoritmlar ma’lumotlarning evklid masofasiga yoki gradient o‘zgarishiga tayanadi. Agar kirish ma’lumotlari turlicha masshtabga ega bo‘lsa, modelning "og‘irliklari" (weights) beqaror bo‘lib qoladi.

Agar bizda ikkita belgi bo‘lsa, ya’ni

1. **Uyning maydoni** X_1 : 30 m² dan 500 m² gacha,
2. **Xonalar soni** X_2 : 1 tadan 10 tagacha.

Matematik nuqtai nazardan, 500 soni 10 dan ancha katta bo'lgani uchun, model uyning maydonini "muhimroq" deb hisoblaydi va barcha hisob-kitoblarni shu raqam atrofida aylantiradi. Natijada xonalar sonining ta'siri "yutib" yuboriladi. Min-Max Scaling aynan shu muammoni bartaraf etib, barcha qiymatlarni bir xil "taroziga" (odatda [0;1] oralig'iga) solib hisoblaydi.

1 – masala. Bizda ko'chmas mulk bozorini raqamlashtirish bo'yicha kichik ma'lumotlar to'plami bor. Ikkita ob'ektning ma'lumotlarini normalizatsiya qilishimiz kerak.

Uy	Maydoni (X, m ²)	Narxi (Y, mln so'm)
Uy-1	120	800
Uy-2	50	300

Ma'lumotlar to'plamidagi chegaraviy qiymatlar:

- Minimal maydon ($X_{\min}$): 40 m²
- Maximal maydon ($X_{\max}$): 200 m²

Vazifa shundan iborat bo'ladiki, Uy-1 va Uy-2 ning maydon ko'rsatkichlarini [0;1] oralig'ida normalizatsiya qiling.

Yechish. Bizga ma'lum bo'lgan formula:

$$X_{norm} = \frac{X - X_{min}}{X_{max} - X_{min}}$$

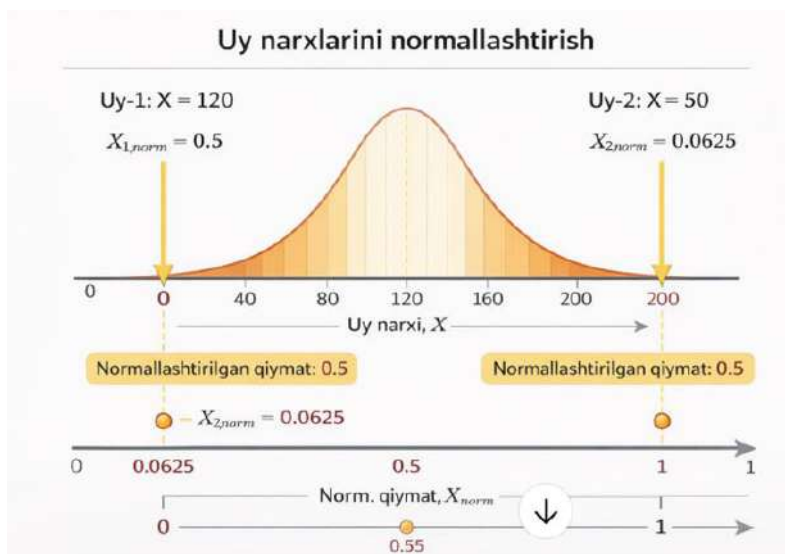
dan iborat.

1. Uy-1 uchun hisoblaymiz ($X = 120$):

$$X_{1,norm} = \frac{120 - 40}{200 - 40} = \frac{80}{160} = 0.5$$

2. Uy-2 uchun hisoblaymiz ($X = 50$):

$$X_{2,norm} = \frac{50 - 40}{200 - 40} = \frac{10}{160} = 0.0625$$



Demak, model uchun 120 va 50 raqamlari emas, balki 0,5 va 0,0625 qiymatlari kirib boradi. 0,5 qiymati bizga Uy-1 ning maydoni umumiy bozordagi uylarning roppa-rosa o'rtasida (medianasida) ekanligini bildiradi. 0,0625 esa Uy-2 ning bozordagi eng kichik uylarga yaqin ekanligini ko'rsatadi. Ya'ni normalizatsiya qilingan ma'lumotlar bilan ishlaganda, modelning xatolik funksiyasi (Loss function) silliqroq shaklga ega bo'ladi. Bu esa neyron tarmog'ining o'rganish tezligini (convergence) 3-5 baravargacha oshirishi mumkin. Aks holda, model "tebranish" (oscillation) holatiga tushib qoladi va minimal xatolik nuqtasini topa olmaydi.

Normalizatsiyaning neyron tarmog'i gradiyentlariga ta'siri. Neyron tarmog'ida har bir kirish parametri (x_i) o'zining og'irlik koeffitsiyentiga (ω_i) ega. Agar ma'lumotlar normalizatsiya qilinmasa, xatolik funksiyasining sirti (surface) juda cho'zilgan va notekis ko'rinishga kelib qoladi.

Matematik isbot:

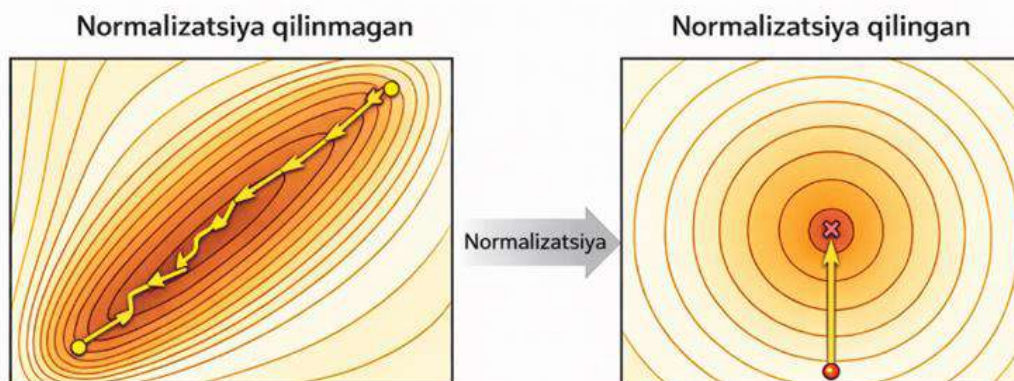
Xatolik funksiyasining gradiyenti (∇J) quyidagicha hisoblanadi:

$$\frac{\partial J}{\partial w_i} = (h_w(x) - y) \cdot x_i$$

Bu yerda ko'rinib turibdiki, xatolikning o'zgarishi bevosita x_i ning qiymatiga bog'liq.

Agar x_1 juda katta bo'lsa (masalan, narx: 800,000,000), u holda ham juda katta bo'ladi.

Agar x_2 juda kichik bo'lsa (masalan, xonalar soni: 2), u holda $\frac{\partial J}{\partial \omega_2}$ ham kichik $\frac{\partial J}{\partial w_1}$ bo'ladi.



Demak, ushbu rasmda ko'rish mumkinki, normalizatsiya qilinmagan ma'lumotlarda gradiyent "zigzag" shaklida harakat qiladi va minimum nuqtaga yetib borishi uchun juda ko'p qadam talab qiladi. Normalizatsiya qilinganda esa xatolik funksiyasi sirti sferik ko'rinishga keladi va gradiyent to'g'ri chiziqli bo'ylab eng qisqa yo'l bilan minimumga tushadi.

Neyron tarmoqlari qatlamlarida matritsali hisob-kitoblar. Raqamlash-tirishda qo'llaniladigan neyron tarmoqlarining har bir qatlamini quyidagi matritsaviy tenglama orqali ifodalash mumkin:

$$Z = W \cdot X + b$$

Bu yerda biz boyagi normalizatsiya qilingan misolimizni (Uy-1: $x_{1, norm} = 0,5$) ko'rib chiqamiz. Tasavvur qiling, neyronning boshlang'ich og'irliklari (W) tasodifiy ravishda $[0,1; 0,2]$ deb olingan.

Yechish:

$$z = (0,1 \cdot 0,5) + (0,2 \cdot 0,0625) + 0,01 = 0,05 + 0,0125 + 0,01 = 0,0725.$$

Keyingi bosqichda bu qiymat **Activation Function** (masalan, ReLU) orqali o'tadi:

$$f(z) = \max(0; z) \Rightarrow f(0,0725) = 0.0725.$$

Logistika graflaridagi optimallashtirish va neyron tarmoqlardagi matritsali hisob-kitoblar o'rtasidagi bog'liqlik shundaki, har ikkala holatda ham **resurslarni tejash** asosiy maqsad hisoblanadi.

1. *Graflarda* – masofa va vaqtni minimallashtirish ($\min \sum \omega_i e_i$);
2. *Neyron tarmoqlarda* – xatolikni va hisoblash vaqtini minimallashtirish ($\min J(W)$).

Sohalarni raqamlashtirishda ushbu matematik yechimlarning birgalikda qo'llanilishi tizimning "aqlli" va tezkor ishlashini ta'minlaydi. Masalan, logistika kompaniyasi (masalan, *Yandex Go* yoki *Uzum*) bir vaqtning o'zida ham graf algoritmlari orqali yo'lni hisoblaydi, ham neyron tarmoqlar orqali mijozning buyurtma berish ehtimolini (normalizatsiya qilingan ma'lumotlar asosida) bashorat qiladi.

Ko'p o'lchovli regressiya tahlili. Sanoatda mahsulot tannarxini yoki talabni bashorat qilishda ko'p omilli chiziqli regressiya qo'llaniladi. Agar bizda n ta omil bo'lsa, bashorat qilish funksiyasi quyidagicha ko'rinish oladi:

$$h_w(x) = w_0 + w_1 x_1 + w_2 x_2 + \dots + w_n x_n = \sum_{i=0}^n w_i x_i = W^T X$$

Bu yerda W – og'irlik koeffitsiyentlari vektori, X – kirish ma'lumotlari vektori. Maqsad – bashorat qilingan qiymat va real qiymat orasidagi kvadratik xatolikni minimallashtirish:

$$J(W) = \frac{1}{2m} \sum_{i=1}^m (h_w(x^{(i)}) - y^{(i)})^2 \rightarrow \min$$

2 – masala. Bir mebel ishlab chiqarish korxonasi keyingi oydagi stol ishlab chiqarish xarajatlarini (y) bashorat qilmoqchi. Buning uchun ikkita asosiy omil tanlab olindi:

1. *Xomashyo narxi* (x_1): 1 dona stol uchun sarflanadigan yog'och narxi (shartli birlikda).
2. *Ishchi kuchi* (x_2): 1 dona stolni tayyorlash uchun ketadigan vaqt (soatlarda).

Bizda korxonaning oxirgi 2 oydagi ma'lumotlari bor:

Oy	x_1 (Xomashyo)	x_2 (Vaqt)	y (Real xarajat)
1-oy	10	5	105
2-oy	12	4	110

Vazifa: Agar modelning boshlang'ich og'irlik koeffitsiyentlari $w_1 = 8$, $w_2 = 5$ va $w_0 = 0$ (bias) deb berilgan bo'lsa, 1 – oy uchun **bashorat qiymatini** ($h_w(x)$) va **kvadratik xatolikni** ($J(W)$) hisoblang.

Yechish: Bashorat funksiyasini hisoblash ($h_w(x)$):

Formula bo'yicha: $h_w(x) = w_0 + w_1x_1 + w_2x_2$.

1-oy uchun qiymatlarni qo'yamiz:

$$h_w(x^{(1)}) = 0 + (8 \cdot 10) + (5 \cdot 5) = 80 + 25 = 105$$

Natija: Model 1-oy uchun xarajatni 105 birlik deb bashorat qildi. (Bu holda bashorat real qiymat bilan mos keldi).

2-oy uchun bashoratni hisoblash:

$$h_w(x^{(2)}) = 0 + (8 \cdot 12) + (5 \cdot 4) = 96 + 20 = 116$$

Natija: Model 2-oy uchun 116 birlik xarajat bo'lishini aytdi, lekin real xarajat 110 edi.

Kvadratik xatolik funksiyasini ($J(W)$) *hisoblash:* Xatolik funksiyasi bashorat va real qiymat o'rtasidagi farqning kvadratini hisoblaydi:

$$J(W) = \frac{1}{2m} \sum_{i=1}^m (h_w(x^{(i)}) - y^{(i)})^2$$

Bu yerda $m = 2$ (oylar soni):

$$J(W) = \frac{1}{2 \cdot 2} [(105 - 105)^2 + (116 - 110)^2]$$

$$J(W) = \frac{1}{4} [0^2 + 6^2] = \frac{36}{4} = 9$$

Ushbu hisob-kitob shuni ko'rsatadiki, modelning xatolik koeffitsiyenti **9** ga teng. Sun'iy intellektning vazifasi – keyingi qadamda *Gradient tushishi* algoritmi orqali w_1 va w_2 koeffitsiyentlarini shunday o'zgartirishki, natijada $J(W)$ qiymati nolga yaqinlashsin. Raqamlashtirilgan korxonalarda ushbu hisob-kitoblar millionlab ma'lumotlar ustida soniyalar ichida bajariladi va bu korxona budjetini 98% *aniqlikda* rejalashtirish imkonini beradi.

Foydalanilgan adabiyotlar:

- [1]. O‘zbekiston Respublikasi Prezidentining 2021-yil 17-fevraldagi PQ-4996-son qarori. "Sun'iy intellekt texnologiyalarini joriy etish bo'yicha shart-sharoitlar yaratish chora-tadbirlari to'g'risida".
- [2]. O‘zbekiston Respublikasining "Raqamli O‘zbekiston – 2030" strategiyasi. (Sohalarni raqamlashtirishning huquqiy va strategik asosi).
- [3]. Goodfellow I., Bengio Y. and Courville A. *Deep Learning*. MIT Press. 2016. – 800 p. (Ushbu kitob SIning matematik asoslari bo'yicha "oltin standart" hisoblanadi).
- [4]. Bishop C. M. *Pattern Recognition and Machine Learning*. Springer. 2006. (Ehtimollar nazariyasining qo'llanilishi bo'yicha asosiy manba).
- [5]. Hastie, T., Tibshirani, R., & Friedman, J. *The Elements of Statistical Learning*. Springer. 2009. (Statistik modelashtirishning matematik yechimlari).