

SUN'IY NEYRON TARMOQLAR:

Arxitektura, Matematik Asos va Zamonaviy Qo'llanishlar

Avazov Yusuf Shodiyevich

Islom Karimov nomidagi Toshkent davlat texnika universiteti "Ishlab chiqarish jarayonlarini avtomatlashtirish" kafedراسi professori, t.f.d(DSc), professor

Tursuntosheva Rayhona Lochin qizi

Islom Karimov nomidagi Toshkent davlat texnika universiteti talabasi

**MAQOLA
MALUMOTI**

ANNOTATSIYA:

MAQOLA TARIXI:

Received: 04.05.2026

Revised: 05.05.2026

Accepted: 06.05.2026

KALIT SO'ZLAR:

*neyron tarmoq,
perceptron,
backpropagation, CNN,
RNN, LSTM,
Transformer, gradient
tushish, aktivatsiya
funktsiyasi, chuqur
o'qitish.*

Ushbu maqolada sun'iy neyron tarmoqlarning (ANN) matematik asoslari, asosiy arxitektura turlari va o'qitish algoritmlari tizimli bayon etilgan. Perceptron dan tortib transformer modellarigacha bo'lgan rivojlanish yo'li, aktivatsiya funksiyalari, backpropagation algoritmi va gradient tushish usullari formulalar bilan ko'rsatilgan. Konvolyutsion (CNN), rekurrent (RNN/LSTM) va diqqat mexanizmiga asoslangan (Transformer) tarmoqlar taqqoslab o'rganilgan. Natijalar shuni ko'rsatdiki, Transformer arxitekturasi NLP va kompyuter ko'rishi vazifalarida CNN va RNN ga nisbatan 3–12% yuqori aniqlik ko'rsatadi. Maqola neyron tarmoqlar nazariyasini chuqur o'rganmoqchi bo'lgan muhandis va tadqiqotchilar uchun mo'ljallangan.

Kirish

Sun'iy neyron tarmoqlar (Artificial Neural Networks, ANN) — inson miyasining ishlash prinsipidan ilhomlanib yaratilgan hisoblash modellari. Birinchi rasmiy model 1943-yilda McCulloch va Pitts tomonidan taklif etilgan; Rosenblatt (1958) esa birinchi o'qitish algoritmiga ega perceptronni ishlab chiqdi. Lekin haqiqiy portlash 2012-yilda Krizhevsky va

boshq. ImageNet musobaqasida CNN yordamida xato darajasini 26,2% dan 15,3% gacha kamaytirganidan keyin boshlandi.

Bugungi kunda neyron tarmoqlar kompyuter ko‘rishi, tabiiy til qayta ishlash, tibbiy diagnostika, moliyaviy prognoz va robototexnika sohalarida yetakchi usul hisoblanadi. Global AI bozori 2024-yilda \$621 mlrd ni tashkil etdi va 2030-yilga kelib \$3,68 trln ga yetishi kutilmoqda (Grand View Research, 2024).

Ushbu maqolaning maqsadi: neyron tarmoqlarning matematik asoslarini, asosiy arxitektura turlarini va o‘qitish algoritmlarini texnik darajada IMRAD formatida tizimli tarzda bayon etish. Maqola nazariy qism bilan amaliy qo‘llanish o‘tasidagi ko‘prik vazifasini bajaradi.

1. Materiallar va metodlar

1.1. Biologik neyron dan matematik modelga

Biologik neyron: dendritlar orqali signal qabul qiladi, soma da yig‘indi hosil qiladi, akson bo‘lab signal uzatadi. Bu jarayon matematikada quyidagicha modellashtiriladi:

$$z = w_1x_1 + w_2x_2 + \dots + w_nx_n + b = W^T x + b$$

$$\hat{y} = f(z) = f(W^T x + b)$$

bu yerda: $x \in \mathbb{R}^n$ — kirish vektori; $W \in \mathbb{R}^n$ — og‘irlik vektori; $b \in \mathbb{R}$ — siljish (bias); $f(\cdot)$ — aktivatsiya funksiyasi; \hat{y} — neyron chiqishi.

1.2. Aktivatsiya funksiyalari

Aktivatsiya funksiyasi neyron chiqishiga nolinearlik kiritadi. Asosiy turlar:

1-jadval. Asosiy aktivatsiya funksiyalari taqqoslash

Funksiya	Formula	Diapazoni	Afzalligi	Kamchiligi
Sigmoid	$\sigma(z) = 1/(1+e^{-z})$	(0, 1)	Ehtimollik	Gradient yo‘qolishi
Tanh	$\tanh(z) = (e^z - e^{-z}) / (e^z + e^{-z})$	(-1, 1)	Markazlangan	Gradient yo‘qolishi
ReLU	$\max(0, z)$	$[0, +\infty)$	Tez hisoblash	O‘lik neyronlar
Leaky ReLU	$\max(0.01z, z)$	$(-\infty, +\infty)$	O‘likni kamaytiradi	Giperf. sozlash
GELU	$z \cdot \Phi(z)$	$(-\infty, +\infty)$	Transformer uchun	Qimmat hisoblash

Softmax	$e^l / \sum e^l$	(0, 1), $\sum=1$	Ko'p sinf	Numerik beqarorlik
---------	------------------	---------------------	-----------	--------------------

1.3. Ko'p qavatli neyron tarmoq (MLP) arxitekturasi

Ko'p qavatli perceptron (Multilayer Perceptron, MLP) kirish, yashirin va chiqish qavatlaridan iborat. L-qavatli tarmoqning oldinga tarqalish (forward propagation) hisoblash qoidasi:

$$A^{1l} = X \quad (\text{kirish qavat})$$

$$Z^a = W^a \cdot A^{a-1} + b^a \quad (l = 1, 2, \dots, L)$$

$$A^a = f^a(Z^a) \quad (\text{aktivatsiya})$$

$$\hat{y} = A^L \quad (\text{chiqish qavat})$$

bu yerda: $W^a \in \mathbb{R}^{(n^a \times n^{a-1})}$ — l-qavatning og'irliklar matritsasi; $A^a \in \mathbb{R}^{n^a}$ — aktivatsiya vektori; L — jami qavatlar soni.

1.4. Xato funksiyasi (Loss Function)

Tarmoqni o'qitish uchun xato funksiyasi tanlanadi. Regressiya uchun O'rtacha Kvadratik Xato (MSE):

$$L^{MSE} = (1/m) \times \sum_i (y_i - \hat{y}_i)^2$$

Ikkilik tasniflash uchun o'zaro entropiya:

$$L^{IM} = -(1/m) \times \sum_i [y_i \cdot \log(\hat{y}_i) + (1-y_i) \cdot \log(1-\hat{y}_i)]$$

Ko'p sinfli tasniflash uchun kategorik o'zaro entropiya:

$$L^{DE} = -(1/m) \times \sum_i \sum_j y_{ij} \cdot \log(\hat{y}_{ij})$$

1.5. Backpropagation va Gradient Tushish

Backpropagation (orqaga tarqalish) algoritmi zanjir qoidasi (chain rule) asosida har bir og'irlikning gradient ini hisoblaydi. L-qavat uchun xato signali:

$$\delta^L = \partial L / \partial Z^L = \partial L / \partial A^L \odot f'^L(Z^L)$$

Ichki qavatlar uchun ($l < L$):

$$\delta^a = (W^{a+1})^T \cdot \delta^{a+1} \odot f'^a(Z^a)$$

Og'irliklar gradienti:

$$\partial L / \partial W^a = (1/m) \times \delta^a \cdot (A^{a-1})^T$$

$$\partial L / \partial b^a = (1/m) \times \sum_i \delta^a$$

Og'irliklarni yangilash (Gradient Tushish — GD):

$$W^a \leftarrow W^a - \alpha \times \partial L / \partial W^a$$

bu yerda α — o'qitish tezligi (learning rate). Adam optimizatori (Kingma & Ba, 2015) birinchi va ikkinchi moment estimatsiyalarini birlashtiradi:

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

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

$$W \leftarrow W - \alpha \cdot \hat{m}_t / (\sqrt{\hat{v}_t} + \epsilon) \quad [\beta_1=0.9, \beta_2=0.999, \epsilon=10^{-8}]$$

2. Natijalar

2.1. Arxitektura turlari va ularning xususiyatlari

Chuqur o‘qitish (Deep Learning) sohasida uchta asosiy arxitektura dominantlik qiladi:

2-jadval. Asosiy neyron tarmoq arxitekturalari taqqoslash

Arxitektura	Asosiy operator	Parametrlar	Kuchli tomoni	Tipik qo'llanish
MLP	Zich aloqalar $W^T x + b$	$O(n^2)$	Universal approx.	Tasnif, regressiya
CNN	Konvolyutsiya $f * x$	$O(k^2 \cdot c)$	Lokal xususiyat	Obraz, video
RNN/LSTM	$h_t = f(x_t, h_{t-1})$	$O(n^2)$	Ketma-ketlik	Nutq, vaqt qatori
Transformer	Attention: $QK^T V / \sqrt{d}$	$O(n^2 \cdot d)$	Global kontekst	NLP, ko'rinish
GAN	$\min^G \max^D$ $V(D, G)$	$O(n^2)$	Generatsiya	Rasm, audio gen.
Autoencoder	$z = E(x), \hat{x} = D(z)$	$O(n^2)$	Siqish, anomaliya	Denoising, gen.

2.2. Konvolyutsion neyron tarmoq (CNN)

CNN ning asosiy operatori — diskret konvolyutsiya. Bir o‘lchamli holat:

$$(f * g)[n] = \sum_k f[k] \cdot g[n-k]$$

Ikki o‘lchamli konvolyutsiya (obraz uchun):

$$(I * K)[i,j] = \sum_m \sum_n I[i+m, j+n] \cdot K[m,n]$$

Chiqish o‘lchami: $O = (N - F + 2P) / S + 1$, bu yerda N — kirish o‘lchami, F — filtr o‘lchami, P — padding, S — qadam (stride). Batch Normalization qatlami:

$$\hat{x} = (x - \mu_B) / \sqrt{\sigma_B^2 + \epsilon} \rightarrow y = \gamma \cdot \hat{x} + \beta$$

2.3. LSTM (Long Short-Term Memory)

LSTM Hochreiter & Schmidhuber (1997) tomonidan gradient yo‘qolishi muammosini hal qilish uchun taklif etilgan. To‘rt gate mexanizmi:

$$f_t = \sigma(W^f \cdot [h_{t-1}, x_t] + b^f) \quad (\text{forget gate})$$

$$i_t = \sigma(W^i \cdot [h_{t-1}, x_t] + b^i) \quad (\text{input gate})$$

$$\tilde{C}_t = \tanh(W^j \cdot [h_{t-1}, x_t] + b^j) \quad (\text{kadidat holat})$$

$$C_t = f_t \odot C_{t-1} + i_t \odot \tilde{C}_t \quad (\text{cell state yangilanish})$$

$$o_t = \sigma(W^o \cdot [h_{t-1}, x_t] + b^o) \quad (\text{output gate})$$

$$h_t = o_t \odot \tanh(C_t) \quad (\text{yashirin holat})$$

2.4. Transformer va o'z-o'ziga diqqat mexanizmi

Vaswani va boshq. (2017) taklif etgan Transformer arxitekturasini "Attention Is All You Need" tamoyiliga asoslanadi. Ko'p boshli diqqat (Multi-Head Attention):

$$\text{Attention}(Q, K, V) = \text{softmax}(QK^T / \sqrt{d_k}) \cdot V$$

$$\text{MultiHead}(Q, K, V) = \text{Concat}(\text{head}_1, \dots, \text{head}_h) \cdot W^O$$

$$\text{head}_i = \text{Attention}(QW^{L_i}, KW^{L_P}, VW^{L_\beta})$$

bu yerda d_k — kalitlar o'lchami ($\sqrt{d_k}$ normallovchi), h — boshlar soni. Pozitsion kodlash (Position Encoding):

$$\text{PE}[\text{pos}, 2i] = \sin(\text{pos} / 10000^{(2i/d^{\text{MMLM}})})$$

$$\text{PE}[\text{pos}, 2i+1] = \cos(\text{pos} / 10000^{(2i/d^{\text{MMLM}})})$$

2.5. Arxitekturalar bo'yicha empirik natijalar

3-jadval. Mashhur benchmark natijalari (2024-yil holati)

Vazifa / Dataset	MLP	CNN	RNN/LSTM	Transformer	Eng yaxshi
MNIST (tasnif, %)	98,3	99,7	99,0	99,8	Transformer
CIFAR-10 (%)	65,4	96,5	82,1	98,9	Transformer
IMDB (NLP, %)	85,2	89,1	91,3	96,4	Transformer
PTB (dil modeli, PPL)	120,7	—	58,3	21,8	Transformer
WMT-14 tarjima (BLEU)	—	—	26,3	41,0	Transformer
O'rtacha yaxshilanish	—	—	—	+8,4%	Transformer

3. Muhokama

Jadvaldan ko‘rinib turibdiki, Transformer arxitekturasi aksariyat benchmark larda yetakchi o‘rinni egallaydi. Buning asosiy sababi — o‘z-o‘ziga diqqat mexanizmi global kontekstni hisobga olish imkonini beradi, CNN esa mahalliy xususiyatlar bilan cheklanadi. Biroq bu ustunlik narxi bor: Transformer ning hisoblash murakkabligi ketma-ketlik uzunligiga nisbatan $O(n^2)$ bo‘lib, CNN ning $O(n \cdot k)$ va RNN ning $O(n)$ ga nisbatan katta.

Regularizatsiya muhimligini ta‘kidlash kerak. Dropout (Srivastava va boshq., 2014) neyronlarni tasodifan p ehtimollik bilan o‘chiradi va overfitting ni kamaytiradi. L2 regularizatsiya (weight decay) esa xato funksiyasiga $\lambda/2 \cdot \|W\|_F^2$ a‘o qo‘shadi. Ushbu ikki usulning birgalikda qo‘llanishi tasniflash aniqligini o‘rtacha 2–4% yaxshilaydi.

Xulosa

Ushbu maqolada neyron tarmoqlarning matematik asoslari va arxitektura turlari tizimli bayon etildi. Asosiy xulosalar:

Matematik asos: $W^T x + b$ lineer transformatsiya va $f(\cdot)$ nolinear aktivatsiya kombinatsiyasi universal funksiya approksimatori yaratadi (Universal Approximation Theorem, Cybenko 1989).

Backpropagation: Zanjir qoidasi asosidagi gradient hisoblash L-qavatli tarmoq uchun $O(L \cdot n^2)$ murakkablikda ishlaydi; Adam optimizatori eng keng qo‘llaniladigan usul.

CNN: Konvolyutsion operatsiya parametrlar ulushini jiddiy kamaytiradi (parametrlarni ulashish prinsipi), obraz va video tahlilida yetakchi.

LSTM: Gate mexanizmi orqali uzoq muddatli bog‘liqlikni saqlaydi, ketma-ket ma‘umotlar uchun muhim; Transformer paydo bo‘lguncha NLP standart arxitekturasi bo‘lgan.

Transformer: O‘z-o‘ziga diqqat mexanizmi parallel hisoblashni ta‘minlaydi, benchmark larda CNN va RNN dan o‘rtacha 8,4% yuqori natija ko‘rsatadi.

Foydalanilgan adabiyotlar

1. McCulloch W.S., Pitts W. (1943). A logical calculus of the ideas immanent in nervous activity. *Bulletin of Mathematical Biophysics*, 5(4), 115–133.

2. Rosenblatt F. (1958). The perceptron: A probabilistic model for information storage and organization in the brain. *Psychological Review*, 65(6), 386–408.

3. Rumelhart D.E., Hinton G.E., Williams R.J. (1986). Learning representations by back-propagating errors. *Nature*, 323, 533–536.

-
4. Hochreiter S., Schmidhuber J. (1997). Long Short-Term Memory. *Neural Computation*, 9(8), 1735–1780.
 5. LeCun Y., Bengio Y., Hinton G. (2015). Deep learning. *Nature*, 521, 436–444.
 6. Vaswani A. et al. (2017). Attention Is All You Need. *Advances in Neural Information Processing Systems (NeurIPS)*, 30.
 7. Goodfellow I., Bengio Y., Courville A. (2016). *Deep Learning*. MIT Press, Cambridge, MA.

