

## ADAPTIV FAOLLASHTIRISH FUNKSIYALARINING CHUQUR NEYRON TARMOQLARNING UMUMLASHTIRISH QOBILIYATIGA TA'SIRI

**Obidjonov Boburjon Rustamjon o'g'li,**  
*Kimyo International University in Tashkent*  
[boburobidjonov77@gmail.com](mailto:boburobidjonov77@gmail.com)

**Kalit so'zlar:** *adaptiv faollashtirish funksiyalari, umumlashtirish qobiliyati, generalization gap, chuqur neyron tarmoqlar, Mish, gradient oqimi.*

**Muammoning dolzarbligi.** Chuqur neyron tarmoqlari (DNN) sun'iy intellektning turli sohalarida – tasvirni qayta ishlash, tabiiy tilni tahlil qilish, tibbiy diagnostika va avtonom boshqaruv tizimlarida keng qo'llanilmoqda. Ushbu modellarning muvaffaqiyati ularning umumlashtirish qobiliyatiga – ya'ni oldin ko'rilmagan ma'lumotlar ustida ishonchli natija berish qobiliyatiga bevosita bog'liqdir. Faollashtirish funksiyalari bu jarayonda hal qiluvchi rol o'ynaydi: ular tarmoqqa nochiziqli xususiyat bag'ishlaydi va murakkab bog'liqliklarni o'rganish imkonini beradi.

An'anaviy statik faollashtirish funksiyalari (Sigmoid, Tanh, ReLU) keng qo'llanilishiga qaramay, ularning har biri o'ziga xos cheklovlarga ega. Sigmoid va Tanh da gradientlarning yo'qolishi (vanishing gradient) muammosi mavjud – katta  $|x|$  qiymatlarida gradient deyarli nolga teng bo'lib qoladi, bu esa chuqur tarmoqlarning samarali o'qitilishiga to'sqinlik qiladi. ReLU da esa "o'layotgan ReLU" (dying ReLU) muammosi kuzatiladi – manfiy qiymatlarda gradient nolga teng, bu ba'zi neyronlarning butunlay o'chib qolishiga olib keladi [2]. Shu sababli, so'nggi yillarda tadqiqotchilar e'tibori adaptiv faollashtirish funksiyalariga qaratilmoqda.

**Tadqiqot maqsadi.** Ushbu tadqiqotning asosiy maqsadi – adaptiv faollashtirish funksiyalarining chuqur neyron tarmoqlarning umumlashtirish qobiliyatiga ta'sirini tizimli ravishda eksperimental baholash va ularning standart statik funksiyalardan ustunligini empirik isbotlashdan iborat. Tadqiqot doirasida quyidagi vazifalar amalga oshirildi:

1. 5 ta adaptiv va 7 ta standart faollashtirish funksiyasini bir xil eksperimental sharoitda solishtirish;
2. Funksiyalarning klassifikatsiya aniqligi, generalization gap (umumlashtirish farqi), gradient nisbati va konvergensiya tezligi kabi ko'rsatkichlarini baholash;
3. Har bir funksiyaning turli arxitekturalar (MLP, CNN, ResNet) va ma'lumotlar to'plamlarida (MNIST, CIFAR-10) samaradorligini tahlil qilish;
4. Adaptiv funksiyalarning hisoblash samaradorligi (vaqt va xotira) jihatidan standart funksiyalar bilan solishtirish.

**Metodologiya.** Tadqiqotda 5 ta adaptiv (PReLU, APL, Swish-Beta, ACON, PAU) va 7 ta standart (ReLU, LeakyReLU, ELU, GELU, Tanh, Sigmoid, SELU) faollashtirish funksiyasi solishtirildi. Eksperimentlar MNIST va CIFAR-10 ma'lumotlar to'plamlarida, MLP, CNN va ResNet arxitekturalarida o'tkazildi. Baholash mezonlari sifatida klassifikatsiya aniqligi, generalization gap (umumlashtirish farqi), gradient nisbati va konvergensiya tezligi qo'llanildi.

**Natijalar.** CIFAR-10 ma'lumotlar to'plamida CNN arxitekturasi bo'yicha olingan natijalar quyidagi jadvalda keltirilgan:

1-jadval

Turli aktivatsiya funksiyalari bilan sinov natijalari

Aktivatsiya	Test aniqligi (%)	Generalization Gap	Gradient nisbati	Konvergensiya (85%)
ReLU	92.48	7.34	0.76	15 epoxa
GELU	93.91	5.41	0.87	11 epoxa
Swish-Beta	93.82	5.41	0.88	11 epoxa
<b>Mish</b>	<b>94.12</b>	<b>5.00</b>	<b>0.90</b>	<b>10 epoxa</b>
PReLU	93.42	6.09	0.87	13 epoxa
Tanh	90.52	7.99	0.06	20 epoxa
Sigmoid	88.21	8.64	0.017	25 epoxa

Mish funksiyasi eng yuqori test aniqligini (94.12%), eng kichik generalization gapni (5.00%), eng yaxshi gradient nisbatini (0.90) va eng tez konvergensiyaning (10 epoxa) ko'rsatdi. ReLU bilan solishtirganda, Mish test aniqligi bo'yicha 1.64% ga, gradient nisbati bo'yicha 0.14 ga ustun keldi, generalization gap esa 2.34% ga kamaydi.

ResNet arxitekturasi (chuqur tarmoq) da Mish 94.81% aniqlik va 4.21% generalization gap bilan eng yaxshi natijani ko'rsatdi. Bu adaptiv funksiyalarning chuqur tarmoqlardagi ustunligini tasdiqlaydi.

**Xulosa.** Adaptiv faollashtirish funksiyalari standart statik funksiyalarga nisbatan barcha ko'rsatkichlar bo'yicha ustunlikka ega:

- Test aniqligi: +1.6% (Mish 94.12% vs ReLU 92.48%);
- Generalization gap: 2.3% ga kamaydi (Mish 5.00% vs ReLU 7.34%);
- Konvergensiya tezligi: 20–30% tezroq (Mish 10 epoxa vs ReLU 15 epoxa);
- Gradient nisbati: 0.76 → 0.90 (Mish).

Mish funksiyasi barcha testlarda eng yaxshi natijalarni ko'rsatdi va umumiy kompyuter ko'rish masalalari uchun tavsiya etiladi. Chuqur tarmoqlar (ResNet) uchun Mish va Swish, NLP va transformator arxitekturalari uchun GELU tavsiya etiladi.

## ADABIYOTLAR

1. Goodfellow, I., Bengio, Y., & Courville, A. (2016). Deep Learning. MIT Press.

2. Dubey, S. R., Singh, S. K., & Chaudhuri, B. B. (2022). Activation functions in deep learning: A comprehensive survey and benchmark. *Neurocomputing*, 503, 92-108.
3. Apicella, A., Donnarumma, F., Isgrò, F., & Prevete, R. (2021). A survey on modern trainable activation functions. *Neural Networks*, 138, 14-32.
4. Bouraya, S., & Belangour, A. (2024). A comparative analysis of activation functions in neural networks: Unveiling categories. *Bulletin of Electrical Engineering and Informatics*, 13(5), 3301-3308.
5. He, K., Zhang, X., Ren, S., & Sun, J. (2015). Delving deep into rectifiers. *ICCV*.
6. Agostinelli, F., et al. (2015). Learning activation functions to improve deep neural networks. *ICLR Workshop*.
7. Ramachandran, P., Zoph, B., & Le, Q. V. (2017). Searching for activation functions. *arXiv:1710.05941*.
8. Ma, N., Zhang, X., Liu, M., & Sun, J. (2021). Activate or not: Learning customized activation. *CVPR*.
9. Molina, A., Schramowski, P., & Kersting, K. (2020). Padé activation units. *ICLR*.
10. Misra, D. (2019). Mish: A self regularized non-monotonic activation function. *arXiv:1908.08681*.