

DOI: <https://doi.org/10.36719/2789-6919/56/116-120>

Nuranə İsmaylova
Azərbaycan Dövlət Pedaqoji Universiteti
<https://orcid.org/0009-0007-5795-4605>
nurana.smaylova@gmail.com

Süni intellekt və maşın öyrənməsində funksiyaların rolu: nəzəri əsaslar və tətbiqlər

Xülasə

Müasir süni intellekt (Sİ) sistemlərinin əksər əsas komponentlərinin təməlinə riyazi funksiya anlayışı dayanır. Neyron şəbəkələrinin qeyri-xətti qabiliyyətini təmin edən aktivləşdirmə funksiyalarından tutmuş, optimallaşdırmanın hədəfini müəyyən edən itki funksiyalarına qədər hər bir arxitektura qərarı funksional seçimdən ibarətdir.

Bu məqalədə funksiya anlayışının tarixi inkişafı, riyazi təməli, eləcə də müasir dərin öyrənmə, transformerlər, diffuziya modelləri və kvant maşın öyrənməsindəki konkret tətbiqləri sistemli şəkildə araşdırılır. Araşdırma göstərir ki, funksiyanın xassələri — xüsusilə diferensiallıq, qeyri-xəttilik və hesabi effektivlik — birbaşa modelin öyrənmə keyfiyyətini müəyyən edir. Aktivləşdirmə funksiyasının seçimi neyron şəbəkəsinin öyrənmə sürətinə təsir edir, itki funksiyası öyrənmənin riyazi məqsədini göstərir və qradiyent hesabı isə törəmə anlayışına əsaslanır.

Açar sözlər: *funksiya, aktivləşdirmə funksiyası, maşın öyrənməsi, süni intellekt, neyron şəbəkəsi, itki funksiyası, optimallaşdırma, transformer, diffuziya modeli*

Nurana Ismaylova
Azerbaijan State Pedagogical University
<https://orcid.org/0009-0007-5795-4605>
nurana.smaylova@gmail.com

The Role of Functions in Artificial Intelligence and Machine Learning: Theoretical Foundations and Applications

Abstract

The concept of mathematical function underlies every core component of modern artificial intelligence (AI) systems. From activation functions that enable nonlinear capability in neural networks to loss functions that define the objective of optimization, every architectural decision amounts to a functional choice. This paper systematically examines the historical development of the function concept, its mathematical foundations, and its concrete applications in deep learning, transformer architectures, diffusion models, and quantum machine learning.

The research demonstrates that the properties of functions – particularly differentiability, nonlinearity, and computational efficiency – directly determine the quality of model learning. The choice of activation function affects the learning rate of a neural network, the loss function defines the mathematical objective of learning, and the gradient computation is based on the concept of derivatives.

Keywords: *function, activation function, machine learning, artificial intelligence, neural network, loss function, optimization, transformer, diffusion model*

Giriş

Süni intellekt (Sİ) kompüterlərə və maşınlar insan kimi düşünmə, öyrənmə və qərar vermə qabiliyyətini qazandırmağa yönəlmiş bir elmi sahədir. Onun müxtəlif alt sahələri var: maşın öyrənməsi, dərin öyrənmə, təbii dilin işlənməsi və görüntü tanıma.

1943-cü ildə U.Mak-Kallok və U.Pitts öz məqalələrində süni neyron sistemi anlayışını təklif etdilər. D.Xebb 1949-cu ildə “Davranış Orqanizasiyası” əsərində neyronların öyrənilməsinin əsaslarını göstərdi. Bu fikirləri amerikan neyrofizioloq Frenk Rozenblatt inkişaf etdirərək insan davranışlarını modelləşdirən mexanizm yaratdı və onu “Perseptron” adlandırdı.

Süni intellekt insan beyninin koqnitiv funksiyalarının öyrənildiyi, maşın və sistemdə təkrarlandığı kompüter elmi sahəsidir. O, rəqəmsal texnologiyaları idarə etmək, böyük məlumatlar əsasında şəffaf və dəqiq qərar qəbul etmək üçün qurulur.

Maşın öyrənməsi modelinin özəyi isə funksiyalardır: modelin daxilindən keçən hər bir məlumat parçası müxtəlif funksiyaların ardıcıl tətbiqi ilə çevrilir, xətalınır, düzəldilir və nəhayət, proqnoza çevrilir. Süni intellekt və maşın öyrənməsində funksiyalar, həm nəzəri təməl, həm də praktik tətbiqlər üçün mərkəzi anlayışdır; əslində bütün modellər “funksiya seçimi və qurulması” ətrafında formalaşır.

Tədqiqat

İstənilən neyron şəbəkəsinin əsas məqsədi qeyri-xətti ayrılmayan giriş məlumatlarını xətti ayrılabilir bilən mücərrəd xüsusiyyətlərə çevirməkdir. Qatlar xətti və qeyri-xətti funksiyaların kombinasiyasından ibarətdir (Dubey və b., 2022, s. 95). Ən populyar qeyri-xətti qatlar aktivasiya funksiyalarıdır: Logistik Sigmoid, Tanh, ReLU, ELU, Swish və Mish. Neyron şəbəkələrinin uğurunun əsas şərtlərindən biri modelə qeyri-xəttiliyi gətirən uyğun aktivasiya funksiyasının seçilməsidir. Son otuz ildə ədəbiyyatda 400-dən çox aktivasiya funksiyası təklif edilmişdir (Kunc və Kléma, 2024).

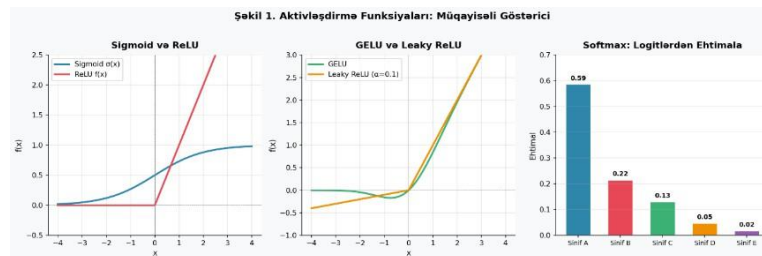
ReLU (Rectified Linear Unit) bu günün aparıcı aktivasiya funksiyasıdır. ReLU son illərdə böyük populyarlıq qazanmış, onun effektivliyi və sadəliyi onu dərin öyrənmə ictimaiyyətində üstün standart aktivasiya funksiyasına çevirmişdir (Zhang və b., 2024, s. 2). Neyron şəbəkələri mürəkkəb nümunələri öyrənmək və modelləşdirmək üçün qeyri-xəttiliyi modelə daxil edən aktivasiya funksiyalarına böyük etibar göstərir. Tədqiqatlar ReLU, Sigmoid, Tanh, Leaky ReLU, ELU və Swish funksiyalarını dəqiqlik, öyrənmə sürəti kimi əsas göstəricilər üzrə müqayisə edir (Agarwal, 2024, s. 11).

Zərər funksiyası modelin nə qədər yanlış proqnoz verdiyini riyazi olaraq ölçür. Öyrənmə prosesinin bütün məntiqini bu funksiya müəyyən edir:

- Orta Kvadratik Səhv (MSE): $L = (1/n) \cdot \sum (y_i - \hat{y}_i)^2$ — reqressiya tapşırıqlarında
- Çarpaz-Entropiya: $L = -\sum y_i \cdot \log(\hat{y}_i)$ — klassifikasiya tapşırıqlarında

Sİ sistemlərinin işləmə prinsipi müxtəlif alqoritmlər və statistik modellər üzərində qurulmuşdur. Ən çox istifadə olunan metodlardan biri maşın öyrənməsidir ki, bu da verilən məlumatlar əsasında sistemlərin özlərini təkbaşına inkişaf etdirməsinə imkan verir.

Optimallaşdırma funksiyaları zərəri azaltmaq üçün modelin ağırlıqlarını dəfələrlə yeniləyir. Əsas prinsip qradiyent enişidir:



$$w_{t+1} = w_t - \eta \cdot \nabla L(w_t)$$

Burada η öyrənmə sürətidir (learning rate). Bu prosesin müxtəlif variantları mövcuddur: SGD (Stokastik Qradyent Enişi), Adam, RMSprop. Dərin neyron şəbəkə metodları böyük həcmdə

verilənlər tələb edir; bəzən daha sadə klassik metodları tətbiq etmək lazım gəlir, çünki dərin neyron şəbəkələri həmişə kömək edə bilmirlər (AZƏRTAC, 2021).

Reqlarlaşma, modelin yalnız öyrənmə məlumatına deyil, yeni məlumatlara da uyğunlaşmasını təmin edir. L1, L2 normları və Dropout mexanizmi bu məqsədlə istifadə olunur.

Maşın öyrənməsi müxtəlif sahələrdə proqnozlaşdırıcı modellər üçün əsas rol oynayır. Dərin öyrənmə isə daha mürəkkəb verilənlər üzərində nümunələri aşkar etmək üçün istifadə olunur — xüsusən səs tanıma və görüntü analizləri sahəsində geniş tətbiq tapır. NLP kompüterlərə insan dilini başa düşmək və cavablandırmaq imkanı verir. CNN arxitekturasında aktivasiya funksiyasının seçimi şəbəkənin effektivliyinə birbaşa təsir edir. Tədqiqatlar göstərir ki, müxtəlif aktivasiya funksiyaları dəqiqlik, konvergensiya sürəti və təsnifat etimadı göstəricilərini əhəmiyyətli dərəcədə fərqləndirə bilər (Emanuel və b., 2024, s. 293).

Süni intellekt maşın öyrənməni özündə bir alt sahə kimi birləşdirir və bu çərçivədə funksiyalar təkcə statistik modellərdə deyil, həm də məntiqi qayda sistemlərində, ekspert sistemlərində və planlaşdırma alqoritmlərində istifadə olunur. Klassik Sİ-də məntiqi çıxarış qaydaları da “funksiya” kimi işləyir: verilmiş bilik bazasından yeni faktlar çıxaran xəritələndirmədir. Müasir Sİ sistemləri çox vaxt hibriddir: maşın öyrənmə funksiyaları (neyron şəbəkələr) üstündə əlavə qərar və idarəetmə funksiyaları (məsələn, qayda əsaslı modul) işləyir (Unite.AI, n.d.).

Məsələn, spam filtr sistemi e-poçtun mətnini giriş vektoru kimi alır və bir funksiya vasitəsilə “spam” və ya “normal” etiketi proqnozlaşdırır; təlim zamanı bu funksiyanın parametrləri keçmiş etiketli e-poçtlar əsasında optimallaşdırılır. Bu sadə misalda belə, həm model funksiyası, həm itki funksiyası, həm də qərar funksiyası (məsələn, ehtimal > 0.5 isə “spam”) eyni vaxtda mühüm rol oynayır (Metsis, 2006, s. 7).

Spam filtr sistemləri maşın öyrənməsinin klassik tətbiq nümunəsidir. E-poçt mətni giriş vektoru kimi qəbul edilir və model funksiyası vasitəsilə “spam” və ya “normal” etiketi proqnozlaşdırılır; təlim zamanı parametrlər etiketli məlumatlarla optimallaşdırılır (Graham, 2002).

Model funksiyası. E-poçt mətni söz-tezlikləri və ya TF-IDF vektoru kimi x şəklində təsvir olunur. Model funksiyası $f(x) = P(\text{spam} | x)$ ehtimalını hesablayır – məsələn, logistik reqressiya ilə $f(x) = \sigma(w \cdot x + b)$, burada $\sigma(z) = \frac{1}{1+e^{-z}}$ siqmoid funksiyasıdır. Təlimdə bu w və b dəyərləri keçmiş spam/normal e-poçtlara uyğunlaşdırılır (Bishop, 2006, s. 212).

İtki (zərər) funksiyası. Modelin düzgünlüyünü ölçmək üçün cross-entropy itki funksiyası istifadə olunur:

$$L = -\frac{1}{N} \sum [y \log(\hat{y}) + (1 - y) \log(1 - \hat{y})]$$

Burada y real etiket (1=spam, 0=normal), $\hat{y} = f(x)$ proqnoz ehtimalıdır (Géron, 2019, s. 112).

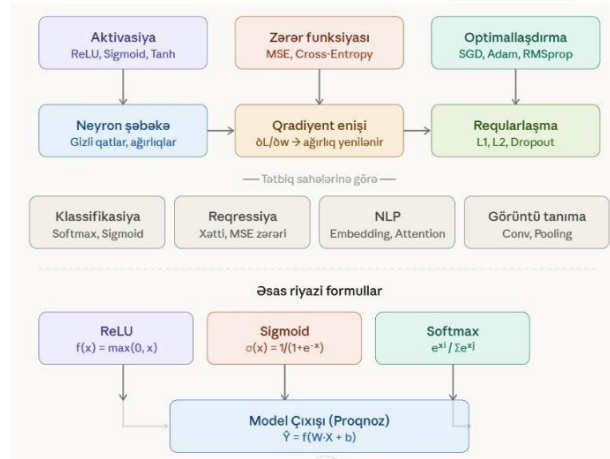
Qərar funksiyası. Proqnoz ehtimalından qərar çıxmaq üçün threshold funksiyası tətbiq edilir:

$$g(\hat{y}) = \begin{cases} \text{spam,} & \text{əgər } \hat{y} > 0.5 \\ \text{normal,} & \text{digər hallarda} \end{cases}$$

Tam proses sxemi

Komponent	Funksiya	Rol
Giriş	x (mətn vektoru)	E-poçt təsviri
Model	$f(x) = \sigma(wx + b)$	Spam ehtimalı
İtki	Cross-entropy L	Optimallaşdırma
Qərar	$g(\hat{y}) > 0.5$	Etiket proqnozu

Bu nümunədə üç funksiya sinxron işləyir: model proqnoz verir, itki keyfiyyəti ölçür, qərar hərəkətə çevirir. Praktikada Bayesian filtr və ya dərin neyron şəbəkələr (LSTM) daha mürəkkəb funksiyalar tətbiq edir, lakin prinsip eynidir.



Şəkil 2. Süni intellekt və maşın öyrənməsində funksiyaların qarşılıqlı əlaqəsi

Nəticə

Süni intellektin bugünkü tətbiqlərində aparıcı komponent kimi maşın öyrənməsi metodları çıxış edir. Hazırda süni intellekt mürəkkəb texnoloji və proqram həlləri kompleksinə çevrilib və bir çox sahədə insanın intellektual fəaliyyətini üstələyən nəticələr əldə etməyə imkan verir. Funksiyalar – aktivasiya, zərər, optimallaşdırma və reqlarlaşma – maşın öyrənməsi sistemlərinin dörd əsas sütununu təşkil edir.

Bu funksiyaların düzgün seçilməsi modelin həm öyrənmə sürətini, həm dəqiqliyini, həm də ümumiləşmə qabiliyyətini birbaşa müəyyən edir. Tədqiqat və inkişaf işlərində “Süni İntellekt” iki qola bölünür: “tətbiq olunan Sİ” xüsusi tapşırığı insan təfəkkürünün təqlidi hesabına yerinə yetirən sistemlər; “ümumiləşdirilmiş SI” isə maşının zəkasını istənilən işdə heç bir əvvəlcədən proqramlaşdırma olmadan bir insan kimi inkişaf etdirməkdir. Gələcək perspektivlər baxımından aktivasiya funksiyaları sahəsindəki tədqiqatların miqyası artmaqdadır – son onillikdə 400-dən çox yeni funksiya təklif edilmişdir, bu da sahənin nə qədər dinamik inkişaf etdiyini göstərir. Funksiyaların seçimi və optimallaşdırılması bundan sonra da Sİ tədqiqatlarının mərkəzində qalacaqdır.

Sahənin açıq problemlərindən biri daha interpretasiya olunan aktivləşdirmə funksiyalarının işlənilib hazırlanmasıdır: böyük dil modelləri hazırda “qara qutu” kimi fəaliyyət göstərir. Digər aktual istiqamət enerji effektivliyidir: milyardlarla parametri hesablayan neyron şəbəkələrinin enerji istehlakı ekoloji baxımdan ciddi narahatlıq yaradır; hesabi cəhətdən daha yüngül, lakin eyni dərəcədə güclü funksional alternativlər axtarışı davam edir.

Ədəbiyyat

1. Agarwal, M. (2024). Comparison of activation functions in neural networks. *International Journal of Advanced Education and Research*, 9(3), 10–16. <https://www.alleducationjournal.com/assets/archives/2024/vol9issue3/9025.pdf>
2. AZƏRTAC. (2021, 9 dekabr). *Süni intellekt texnologiyalarının inkişafında əsas trendlər*. https://azertag.az/xeber/suni_intellekt_texnologiyalarinin_inkisafinda_esas_trendler-1946778/ <https://alas.edu.az/suni-intellekt-muhendisliyi/>
3. Bishop, C. M. (2006). *Pattern recognition and machine learning*. Springer.
4. Dubey, S. R., Singh, S. K. və Chaudhuri, B. B. (2022). Activation functions in deep learning: A comprehensive survey and benchmark. *Neurocomputing*, 503, 92–108. <https://doi.org/10.1016/j.neucom.2022.06.111>

5. Emanuel, R.H.K., Docherty, P.D., Lunt, H. və b. (2024). The effect of activation functions on accuracy, convergence speed, and misclassification confidence in CNN text classification: a comprehensive exploration. *J Supercomput* 80, 292–312.
<https://doi.org/10.1007/s11227-023-05441-7>
6. Géron, A. (2019). *Hands-on machine learning with Scikit-Learn, Keras, and TensorFlow*.
7. Graham, P. (2002). *A plan for spam*. <http://www.paulgraham.com/spam.html>
8. Kunc, V. və Kléma, J. (2024). *Three Decades of Activations: A Comprehensive Survey of 400 Activation Functions for Neural Networks* (arXiv:2402.09092). arXiv.
<https://doi.org/10.48550/arXiv.2402.09092>
9. Metsis, V., Androutsopoulos, I. və Paliouras, G. (2006). Spam filtering with naive Bayes, which naive Bayes? *Proceedings of the Third Conference on Email and Anti-Spam (CEAS)*.
10. Süni İntellekt Mühəndisliyi: Gələcəyin Texnologiyası və Tətbiq Sahələri. (2025). ALAS Academy. <https://alas.edu.az/suni-intellekt-muhendisliyi/>
11. Unite.AI. (n.d.). *Artificial intelligence news and analysis*. <https://www.unite.ai/>
12. Zhang, S., Lu, J. və Zhao, H. (2024). Deep Network Approximation: Beyond ReLU to Diverse Activation Functions. *Journal of Machine Learning Research*, 25, 1–39.
<https://jmlr.org/papers/volume25/23-0912/23-0912.pdf>

Daxil oldu: 23.11.2025

Qəbul edildi: 04.03.2026