

SÜNİ INTELEKTİN ƏSASLARI: İDEYALARIN TƏKAMÜLÜNDƏN MÜƏYYƏNLİK AXTARIŞINA QƏDƏR

AKİF ƏLİZADƏ, geologiya-mineralogiya elmləri doktoru, AMEA-nın həqiqi üzvü, professor, AR Elm və Təhsil Nazirliyinin Geologiya və Geofizika İnstitutu. E-mail: a.alizade@gia.science.az
<https://orcid.org/0009-0003-5328-2320>

MİSİR MƏRDANOV, fizika-riyaziyyat elmləri doktoru, AMEA-nın müxbir üzvü, professor, Bakı Dövlət Universiteti.
E-mail: misirmardanov@yahoo.com
<https://orcid.org/0000-0003-3901-0719>

RAMİN RZAYEV, texnika elmləri doktoru, professor, AR Elm və Təhsil Nazirliyinin Riyaziyyat İnstitutu. E-mail: raminrza@yahoo.com
<https://orcid.org/0000-0001-7658-2850>

Məqaləyə istinad:

Əlizadə A., Mərdanov M., Rzayev R. (2026). Süni intellektin əsasları: ideyaların təkamülündən müəyyənlik axtarışına qədər. *Azərbaycan məktəbi*. № 2 (715), səh. 9–23

DOI:

10.30546/32898065.2026.2.0175.10

Məqalə tarixçəsi

Göndərilib: 15.12.2025

Qəbul edilib: 13.04.2026

ANNOTASIYA

Süni intellektin inkişafının əsasını təşkil edən ideyaların təkamülünün təhlilinə həsr olunan bu məqalədə müəyyənliyin təmin olunması problemi həm metodoloji, həm də fəlsəfi aspektdən tədqiq edilir. Məqalədə süni intellektin tarixi mərhələləri erkən fəlsəfi və riyazi konsepsiyalardan, o cümlədən David Hilbert və Alan Turingin işlərindən başlayaraq təhlil edilir. Bu təhlil XX və XXI əsrlərin praktiki mühəndis yanaşmalarına qədər olan dövrü sistemli şəkildə əhatə edir. Xüsusi diqqət abstrakt ideyaların formal modellərə və alqoritmik strukturlara çevrilməsi prosesinə yetirilir. Göstərilir ki, bu modellər öyrənmə, mühakimə, qərar qəbuletmə və qeyri-müəyyən mühitdə adaptasiya kimi intellektual funksiyaların modelləşdirilməsinə imkan verir. Məqalədə həmçinin biliklərin formallaşdırılması üsulları, eləcə də intellektual sistemlərin etibarlılığını və proqnozlaşdırılma imkanlarını artırmağa yönəlmiş yanaşmalar təhlil olunur. Nəticə olaraq göstərilir ki, süni intellektə müəyyənliyin təmin olunması yalnız texniki məsələ deyil. Bu problem həm də idrak proseslərinin formal modelləşdirilməsi ilə bağlı elmi problemlər kompleksinə daxildir.

Açar sözlər: Süni intellekt, ideyaların təkamülü, Hilbertin 13-cü problemi, Kolmoqorov-Arnold teoremi, KAN (Kolmoqorov-Arnold Şəbəkəsi), neyron şəbəkəsi, Veyerştras yaxınlaşma teoremi.

FOUNDATIONS OF ARTIFICIAL INTELLIGENCE: FROM THE EVOLUTION OF IDEAS TO THE PURSUIT OF CERTAINTY

AKIF ALIZADEH, Doctor of Geological and Mineralogy Sciences, Full Member of ANAS, Professor, Institute of Geology and Geophysics of the Ministry of Science and Education of AR. E-mail: a.alizade@gia.science.az
<https://orcid.org/0009-0003-5328-2320>

MISIR MARDANOV, Doctor of Physical and Mathematical Sciences, Corresponding Member of ANAS, Professor, Baku State University. E-mail: misirmardanov@yahoo.com
<https://orcid.org/0000-0003-3901-0719>

RAMIN RZAYEV, Doctor of Technical Sciences, Professor, Institute of Control Systems of the Ministry of Science and Education of AR. E-mail: raminrza@yahoo.com
<https://orcid.org/0000-0001-7658-2850>

To cite this article:

Alizadeh A., Mardanov M., Rzayev R. (2026). Foundations of Artificial Intelligence: From the Evolution of Ideas to the Pursuit of Certainty. *Azerbaijan Journal of Educational Studies*. Vol. 715, Issue II, pp. 9–23

DOI:

10.30546/32898065.2026.2.0175.10

Article history

Received: 15.12.2025

Accepted: 13.04.2026

ABSTRACT

This article examines the evolution of ideas underlying the development of artificial intelligence and addresses the problem of achieving certainty as a key research direction in both methodological and philosophical contexts. The historical stages of artificial intelligence are systematically examined, ranging from early philosophical and mathematical concepts, including the works of David Hilbert and Alan Turing, to the practical engineering approaches of the twentieth and twenty-first centuries. Particular attention is paid to the transformation of abstract ideas into formal models and algorithmic structures. It is demonstrated that these models enable the representation of intellectual functions such as learning, reasoning, decision-making, and adaptation to uncertain environments. Additionally, the study analyzes methods of knowledge formalization and approaches aimed at improving the reliability and predictability of intelligent systems. The article concludes that achieving certainty in artificial intelligence is not only a technical issue but also part of a broader set of scientific problems related to the formal modeling of cognitive processes.

Keywords: Artificial intelligence, evolution of ideas, Hilbert's 13th problem, Kolmogorov–Arnold theorem, KAN (Kolmogorov–Arnold Network), neural network, Weierstrass approximation theorem.

GİRİŞ

İnsan düşüncəsinin tarixi ideyaların inkişaf tarixi kimi xarakterizə oluna bilər. Bəzi ideyalar qısa müddət mövcud olur və tez unudulur, digərləri isə uzun müddət ərzində elmi müzakirələrin və yeni tədqiqat istiqamətlərinin formalaşmasına təsir göstərir. Lakin yalnız məhdud sayda ideya elmi inkişafın ümumi istiqamətinə əhəmiyyətli təsir edir. Bu ideyalardan biri insan intellektinə bənzər sistemlərin yaradılması ilə bağlıdır. Tarix boyu əfsanələr, fəlsəfi yanaşmalar və elmi tədqiqatlar insan şüurunun öyrənilməsi və onun süni şəkildə modelləşdirilməsi ideyası ətrafında formalaşmışdır. Piqmaliyon və Qolem haqqında miflərdən (Naso, 2009; Scholem, 1990) başlayaraq, İntibah dövrünün avtomatlarına, Leybnitsin mexaniki qurğularına və XX əsrin kibernetik modellərinə qədər bu istiqamətdə müxtəlif yanaşmalar irəli sürülüb (Kang, 2011).

Müasir mərhələdə bu ideyalar elmi və texnoloji inkişaf nəticəsində praktik tətbiq sahəsinə keçib. Süni intellekt (Sİ) bu kontekstdə insan idrak proseslərinin modelləşdirilməsi və analizi üçün istifadə olunan alətlər və metodlar toplusu kimi nəzərdən keçirilir. Bu meyil yalnız texnoloji maraqlardan deyil, həm də insan idrakının dərinədən öyrənilməsi ehtiyacından irəli gəlir. Sİ ideyası müxtəlif dövrlərdə, mədəniyyətlərdə və elmi məktəblərdə qorunub saxlanıb və inkişaf edib. Bu maraqlar təkcə praktik problemlərin həlli ilə bağlı deyil, həm də insanın öz idrak mexanizmlərini anlamasına yönəlib. Bu kontekstdə Sİ-nin inkişafı yalnız texnoloji məqsədlərlə məhdudlaşmır, həm də müəyyənliyin əldə olunması və idrak proseslərinin formal təsviri ilə bağlıdır.

Hesablama texnologiyalarının inkişafı və böyük həcmdə verilənlərin toplanması nəticəsində Sİ artıq nəzəri ideya olmaqdan çıxaraq geniş tətbiq sahəsinə malik texnologiyaya çevrilib. Hazırda o, bir çox sahələrdə strateji əhəmiyyət kəsb edir. Lakin bu inkişaf fonunda intellektin təbiəti, quruluşu və modelləşdirilməsi ilə bağlı fundamental suallar aktuallığını saxlayır. Texnoloji inkişafın nəzəri izahdan daha sürətli olması yeni problem yaradır: modellərin yüksək effektivliyi ilə onların interpretasiya olunmasının çətinliyi arasında ziddiyyət meydana çıxır. Eyni zamanda maşın öyrənməsinin praktiki uğurları ilə izah oluna bilən və şəffaf modellərə olan ehtiyac arasında fərq müşahidə olunur.

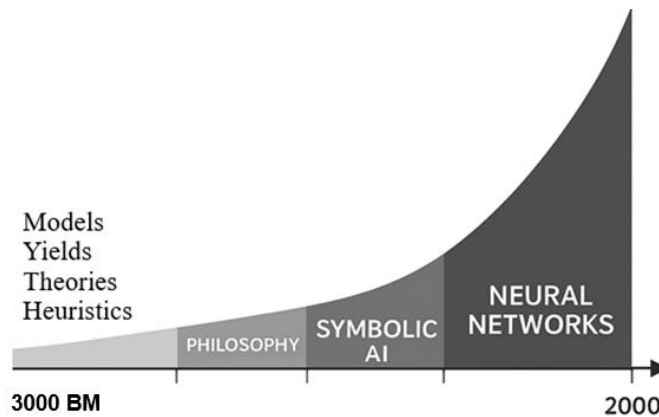
Müasir Sİ sistemləri mürəkkəb problemlərin həllində yüksək nəticələr göstərsə də, onların işləmə mexanizmləri çox vaxt kifayət qədər şəffaf deyil. Bu sistemlər tez-tez “qara qutu” kimi xarakterizə olunur, çünki qəbul etdikləri qərarların daxili məntiqini izah etmək, hətta onları yaradanlar üçün belə çətin olur. Bu kontekstdə əsas sual yaranır: Sİ sistemləri yalnız hesablama baxımından effektiv olmaqla kifayətlənməlidir? Yoxsa onların qərarları həm də interpretasiya oluna bilən və izah edilə bilən olmalıdır? Əgər modellərin daxili strukturu və işləmə prinsipləri aydın deyilsə, bu sistemlər tərəfindən əldə olunan nəticələrin etibarlılığı və mənalandırılması məsələsi açıq qalır. Bu səbəbdən müəyyənliyin təmin olunması problemi müasir dövrdə həm elmi, həm də metodoloji baxımdan aktualdır. Bu problem məntiq, koqnitiv elmlər, semantika və etik yanaşmaların inteqrasiyasını tələb edir. Beləliklə, Sİ-nin tədqiqi yalnız mühəndislik məsələsi deyil, həm də fundamental elmi araşdırma sahəsinə çevrilir.

Bu işin məqsədi Sİ-nin inkişaf tarixini, metodoloji əsaslarını və əsas elmi-fəlsəfi problemlərini sistemli şəkildə təhlil etməkdir. Həmçinin onun nəzəri və praktik inkişafı üçün əsas prinsiplərin müəyyən edilməsi nəzərdə tutulur.

SÜNİ İNTELLEKTİN YÜKSƏLİŞİ: ƏFSANƏDƏN ELMİ REALLIĞA

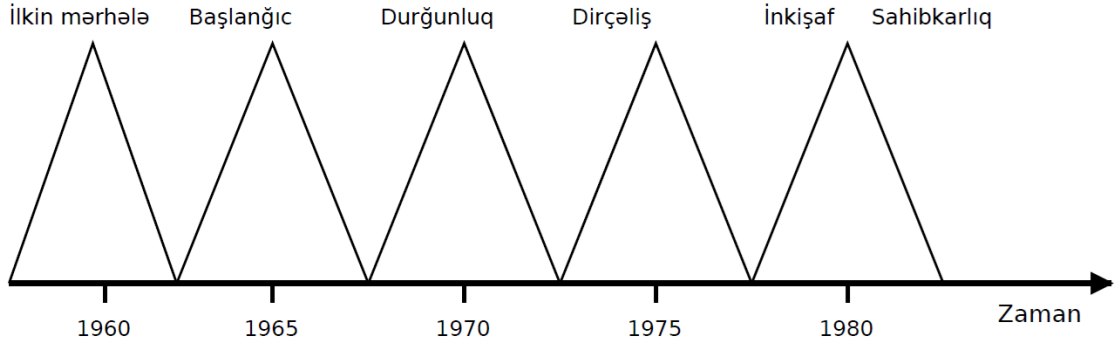
İnsan fəaliyyətinin mühüm xüsusiyyətlərindən biri onun yaradıcılığa yönəlməsidir. Tarix boyu insan yalnız alətlər yaratmaqla kifayətlənməyib, həm də daha mürəkkəb funksiyaları yerinə yetirə bilən sistemlərin qurulmasına çalışıb. Bu cəhdlər insan idrak proseslərinin anlaşılması və onların süni şəkildə modelləşdirilməsi ilə bağlı olub.

Bu istiqamətdə formalaşan ideyalar uzun müddət, əsasən, fəlsəfi və konseptual xarakter daşıyıb. Lakin onların elmi əsasda öyrənilməsi yalnız XX əsrdə mümkün olub. Məhz bu dövrdə “düşünən maşınlar” ideyası fəlsəfi mülahizələr səviyyəsindən çıxaraq riyazi modellər və mühəndislik yanaşmaları çərçivəsində tədqiq olunmağa başlanıb (McCulloch and Pitts, 1943; Rosenblatt, 1958) (Şəkil-1).



Şəkil 1. Süni intellektin inkişaf mərhələləri

Sİ sahəsinin tanınan tədqiqatçısı Patrik Uinston bu dövrü “inkişafın yeni mərhələsi” kimi xarakterizə edib. O, bu yanaşma ilə Sİ sahəsində abstrakt ideyalardan real modellərin və sistemlərin yaradılmasına keçidi vurğulayıb. Uinston 1987-ci ildə nəşr olunan əsərində xüsusi zaman miqyası təqdim edib (Şəkil-2). O göstərib ki, uzun müddət, əsasən, nəzəri xarakter daşıyan tədqiqatlar bu mərhələdə intellektual sistemlərin məqsədyönlü layihələndirilməsi və tətbiqi ilə əvəz olunub (Dormehl, 2017).



Şəkil 2. Süni intellektin inkişafına yanaşmaların təkamülü

Uinstonun təqdim etdiyi sxemdə idrakın mifoloji və fəlsəfi izah mərhələlərindən elmi yanaşmaya keçidi vizual şəkildə göstərilir. Bu mərhələdə Sİ yalnız texnoloji istiqamət kimi deyil, həm də insan idrak proseslərinin modelləşdirilməsi ilə bağlı elmi sahə kimi qiymətləndirilir (Dormehl, 2017). Bu keçid Sİ sahəsində nəzəri yanaşmalardan praktik modellərin və sistemlərin yaradılmasına keçidi ifadə edir. Bu kontekstdə insan ilk dəfə idrak proseslərini məqsədyönlü şəkildə modelləşdirməyə və texniki sistemlərdə reallaşdırmağa cəhd edib. Beləliklə, Sİ ideyası tarixi inkişafın nəticəsi kimi formalaşmış və insan düşüncəsinin müşahidədən analitik yanaşmaya və sistemli modelləşdirməyə doğru inkişafını əks etdirib. Bu baxımdan, müəyyənliyin təmin olunması problemi Sİ-nin inkişafında əsas elmi məsələlərdən biri kimi çıxış edir və idrak proseslərinin formal təsviri ilə bağlıdır (Hilbert, 1900; Kolmogorov, 1957; Arnold, 1957; Shannon, 1948).

Zehnin təbiəti haqqında fəlsəfi mülahizələrdən mühəndislik yanaşmalarına keçid insan idrakı tarixində mühüm mərhələ kimi qiymətləndirilir. XIX–XX əsrlər ərzində əvvəllər, əsasən, metafizik və məntiqi kontekstdə müzakirə olunan məsələlər eksperimental və riyazi formalarda ifadə olunmağa başlayıb. Fəlsəfi müzakirələrin predmeti olan bir çox məsələlər tədricən idrak proseslərinin formallaşdırılması və alqoritmləşdirilməsi üçün əsas təşkil edib. Bu prosesdə riyazi məntiq, informasiya nəzəriyyəsi və kibernetika elmlərinin formalaşması həlledici rol oynayıb. Corc Bul, Qotlob Frege və Kurt Gödelin tədqiqatları mühakimələrin formal işarələr sistemi vasitəsilə ifadə olunmasını təmin edib (Hilbert, 1900; Kolmogorov, 1957; Arnold, 1957). Bu yanaşma məntiqi yalnız təsviri vasitə kimi deyil, həm də hesablana bilən formal sistem kimi nəzərdən keçirməyə imkan verib.

Beləliklə, idrakın məntiqi proses kimi fəlsəfi təsviri riyazi formal modelə çevrilib.

Bu ideyanın mühəndislik konsepsiyasına çevrilməsində həlledici rolu Alan Turing oynaymışdır. Onun məşhur (Turing, 1950) əsəri yalnız universal hesablama maşını modelini təqdim etməyib, həm də fundamental bir sual qoyub: “Maşın idrak funksiyalarını yerinə yetirə bilirmi?” (McCulloch and Pitts, 1943). Turing göstərmişdir ki, mühakimə və qərar qəbul etmə kimi idrak prosesləri alqoritmik şəkildə modelləşdirilə bilər və buna görə də texniki sistemlərdə reallaşdırılması mümkündür. Bu yanaşma idrak probleminin fəlsəfi müstəvidən hesablama modelləri çərçivəsinə keçirilməsinə imkan verib, idrak proseslərini informasiya emalı kimi şərh etməyə əsas yaradıb. Turingin ideyaları sonrakı tədqiqatçılar tərəfindən inkişaf etdirilərək “süni intellekt” istiqamətinin formalaşmasına gətirib

çıxarıb. 1956-cı ildə Dartmut konfransında C.Mak-Karti, M.Minski, K.Şennon və N.Roçester “Artificial Intelligence” terminini rəsmi şəkildə təqdim ediblər (Brown, et al., 2020). Onlar bu sahəni idrak funksiyalarını yerinə yetirə bilən sistemlərin yaradılması ilə bağlı elmi istiqamət kimi müəyyən ediblər. Bu hadisə Sİ-nin sistemli elmi tədqiqat sahəsi kimi formalaşmasının başlanğıcı hesab olunur.

Sonrakı onilliklər Sİ-nin sürətli inkişafı və yanaşmaların dəyişməsi ilə xarakterizə olunur. Məntiq və çıxarış qaydalarına əsaslanan simvolik yanaşma idrak proseslərini işarələr və formal strukturlar sistemi kimi təsvir edirdi. Lakin zaman keçdikcə aydın oldu ki, insan idrakını yalnız formal konstruksiyalar vasitəsilə tam izah etmək mümkün deyil. Bu səbəbdən neyron şəbəkələrə, evristik və koqnitiv modellərə əsaslanan yeni istiqamətlər formalaşdı (McCulloch and Pitts, 1943; Rosenblatt, 1958; Rumelhart et al., 1986). Bu yanaşmalar idrak və öyrənmə proseslərinin bioloji mexanizmlərinə daha yaxın modellərin qurulmasına yönəlib.

Beləliklə, fəlsəfədən mühəndisliyə keçid kəskin qırılma deyil, yanaşmaların tədrici transformasiyası kimi baş verib. İdrakın mahiyyəti ilə bağlı fəlsəfi ideyalar Sİ sistemlərinin konseptual əsasını təşkil edib, mühəndislik yanaşmaları isə bu ideyaların formal modellər və texniki sistemlər şəklində reallaşdırılmasına imkan verib. Xüsusilə Norbert Viner və Con fon Neymanın işlərindən sonra bu sahəyə maraq əhəmiyyətli dərəcədə artıb (Shannon, 1948). Nəticədə, tədqiqatçılar iki əsas yanaşma üzrə fərqlənib (Dormehl, 2017; Goodfellow et al., 2016; LeCun et al. 2015).

Birinci yanaşmaya görə, Sİ mövcud proqram təminatı və aparat vasitələrinin təkmilləşdirilməsi ilə əldə oluna bilər və bunun üçün hesablama resurslarının artırılması kifayətdir. İkinci yanaşmaya görə isə Sİ sistemlərinin yaradılması üçün yeni metodlar və texnologiyalar tələb olunur. Bu yanaşma idrak proseslərinin daha dərinə öyrənilməsinə və yeni alqoritmik prinsiplərin işlənilib hazırlanmasını zəruri hesab edir.

Elmi nailiyyətlər mühüm rol oynasa da, Sİ sahəsində müşahidə olunan inkişaf göstərir ki, son illərdə sistemlərin imkanlarının artması, əsasən, miqyaslandırma ilə bağlıdır. Başqa sözlə, Sİ modellərinin effektivliyinin artırılması onların ölçüsünün və istifadə olunan resursların genişləndirilməsi hesabına əldə olunur.

Sİ-də miqyaslandırma təlim prosesində üç əsas komponentin eyni vaxtda artırılmasını nəzərdə tutur:

- təlim üçün istifadə olunan verilənlərin həcmi;
- modelin ölçüsü (parametrlərin sayı ilə müəyyən olunur);
- hesablama resursları.

Böyük həcmli verilənlər əsasında öyrədilən Sİ modelləri, bir qayda olaraq, daha yüksək effektivlik göstərir. Bundan əlavə, müəyyən miqyaslandırma səviyyəsindən sonra modellər əvvəlki ölçülərdə müşahidə olunmayan yeni funksional imkanlar göstərə bilər. Bu cür keçidlər çox vaxt qeyri-xətti xarakter daşıyır və modelin imkanlarında kəskin dəyişikliklərlə müşayiət olunur. Bu isə sistemlərin davranışının əvvəlcədən proqnozlaşdırılmasını çətinləşdirir və onların etibarlılığı ilə bağlı əlavə tədqiqatların aparılmasını zəruri edir.

SÜNI İNTELLEKTİN RİYAZİ ƏSASLARI

1900-cü ildə David Hilbert Parisdə keçirilən II Beynəlxalq Riyaziyyatçılar Konqresində təqdim etdiyi məşhur məruzəsində 13-cü problemi irəli sürüb (Hilbert, 1900). Bu problem cəbr tənliklərinin həllinin hansı formalarda ifadə oluna bilməsi ilə bağlıdır. Hilbert 7-ci dərəcəli tənliyin həllinin mümkünlüyü barədə sual irəli sürüb. O araşdırırdı ki, tənliyin həllini sadə əməliyyatların məhdud sayı və yalnız iki dəyişəndən asılı funksiyalar vasitəsilə ifadə etmək mümkündürmü?

Bu problem cəbr funksiyalarının strukturunun və onların analitik təqdimat imkanlarının sərhədlərini araşdırır və funksional kompozisiya ilə çoxölçülü analiz sahələrində mühüm istiqamətlərin formalaşmasına təsir göstərmiş (Kolmogorov, 1957; Arnold, 1957). Beləliklə, Hilbertin qoyduğu məsələ mürəkkəb funksiyaların daha sadə funksional asılılıqlar vasitəsilə təsvir olunmasının mümkünlüyünü araşdırır və XX əsr riyaziyyatının inkişafına əhəmiyyətli təsir göstərmiş.

4-cü dərəcəyə qədər olan cəbr tənlikləri radikallar vasitəsilə həll oluna bilər. Bu o deməkdir ki, onların kökləri toplama, çıxma, vurma, bölmə və kök çıxarma əməliyyatlarının məhdud sayı ilə ifadə edilə bilər. Bu tənliklər üçün açıq analitik həll formulları mövcuddur: məsələn, 3-cü dərəcəli tənliklər üçün Kardanı formulu, 4-cü dərəcəli tənliklər üçün isə Lodoviko Ferrarinin metodu tətbiq olunur. Bu yanaşmalar bütün köklərin dəqiq analitik ifadəsini əldə etməyə imkan verir. Lakin 5-ci dərəcədən etibarən vəziyyət dəyişir. XIX əsrdə Nils Abel göstərmiş ki, ümumi halda 5-ci dərəcəli tənliyin köklərini radikallar vasitəsilə ifadə edən formul mövcud deyil. Daha sonra Evarist Galois bu nəticəni izah edib və qrup nəzəriyyəsi çərçivəsində tənliyin strukturunu onun radikallarla həll olunma imkanları ilə əlaqələndirib.

Beləliklə, müəyyən olunub ki, 5-ci dərəcədən yuxarı tənliklərin həlli daha mürəkkəb analitik və ya ədədi metodlar tələb edir və sadə cəbr əməliyyatlarının çərçivəsini aşır. Bu nəticə Abel-Ruffini teoremi ilə ifadə olunur. Bununla yanaşı, Çirnhauz çevrilmələri ümumi dərəcəli tənliyi daha sadə formaya gətirməyə imkan verir; belə çevrilmələr nəticəsində bəzi yüksək dərəcəli hədlər (məsələn, x^{n-1} , x^{n-2} və x^{n-3}) aradan qaldırıla bilər (Hilbert, 1900; Kolmogorov, 1957; Arnold, 1957). $n=5$ halı üçün bu nəticə ilk dəfə 1786-cı ildə Bring tərəfindən əldə edilib, ümumi hal isə 1834-cü ildə Jerrard tərəfindən göstərilmiş.

Əlavə normallaşdırma çevrilmələrindən sonra 5-ci, 6-cı və 7-ci dərəcəli cəbr tənliklərinin həlli müvafiq olaraq 1, 2 və 3 parametrdən asılı olan formal tənliklərin həllinə gətirilə bilər.

$$\begin{aligned}x^5+ax+1=0, x=f(a); \\x^6+ax^2+bx+1=0, x=f(a, b); \\x^7+ax^3+bx^2+cx+1=0, x=f(a, b, c).\end{aligned}$$

Bu yanaşma yüksək dərəcəli tənliklərin köklərinin tapılması məsələsini sadələşdirərək onu parametrik formalarda verilmiş funksiyaların öyrənilməsi probleminə gətirir. 7-ci dərəcəli tənlik kontekstində Hilbertin 13-cü problemi çərçivəsində xüsusi sual formalaşdırılıb (Hilbert, 1900). Burada $x=f(a, b, c)$ funksiyasının yalnız iki dəyişəndən asılı olan funksiyaların məhdud sayda superpozisiyası şəklində ifadə olunma imkanı araşdırılır. Bu məsələ çoxölçülü cəbr funksiyalarının

strukturunun öyrənilməsi ilə bağlıdır və mürəkkəb funksional asılılıqların daha sadə komponentlər vasitəsilə təsvir olunmasının mümkünlüyünü araşdırır.

Bu məsələ uzun müddət açıq qalıb və həm cəbr, həm də çoxdəyişənli funksiyalar sahəsində çalışan tədqiqatçıların diqqətini cəlb edib. Onun həlli Kolmoqorov-Arnold superpozisiya teoremi ilə verilib (Kolmogorov, 1957; Arnold, 1957). Teoremə görə, bir neçə dəyişəndən asılı olan istənilən kəsilməz funksiya sonlu sayda tək dəyişənli kəsilməz funksiyaların kompozisiyası kimi təqdim oluna bilər. Ümumi halda, Kolmoqorov-Arnold teoremi aşağıdakı şəkildə ifadə olunur:

$$f(\mathbf{x}) = f(x_1, x_2, \dots, x_n) = \sum_{q=1}^{2n} \Phi_q \left(\sum_{p=1}^n \phi_{q,p}(x_p) \right). \quad (1)$$

Aşağıda kəsilməz 2 dəyişənli $f(x, y) = x^2 + y^2 + xy$ funksiya üçün bu teoremin tətbiqinə dair nümunə göstərilib:

$$f(x, y) = \sum_{i=1}^m \sum_{j=1}^n a_{ij} x^i y^j.$$

Beləliklə, Kolmoqorov-Arnold teoremi Hilbertin qoyduğu suala müsbət cavab verib və göstərib ki, çoxdəyişənli funksiyalar müəyyən şərtlər daxilində daha sadə funksiyaların kompozisiyası vasitəsilə təqdim oluna bilər. Bu nəticə çoxölçülü funksiyaların strukturunun və onların funksional parçalanma imkanlarının öyrənilməsində mühüm rol oynayıb. Kolmoqorov-Arnold teoremi neyron şəbəkələrinin approksimasiya qabiliyyətinin nəzəri əsaslarından birini təşkil edir (Kolmogorov, 1957; Arnold, 1957). Bu yanaşmaya görə, çoxlaylı neyron şəbəkələri kəsilməz funksiyaları istənilən dəqiqliklə yaxınlaşdırma bilər (Funahashi, 1989; Cybenko, 1989; Hornik et al., 1989). Bu xüsusiyyət proqnozlaşdırma, idarəetmə, klasterləşdirmə və təsnifat kimi məsələlərdə neyron şəbəkələrinin effektiv tətbiqini təmin edir. Bu kontekstdə Kolmoqorov-Arnold yanaşması neyron şəbəkələrinin strukturunun formalaşdırılmasında mühüm rol oynayır və şəbəkə elementlərinin qarşılıqlı əlaqələrinin təşkilinə nəzəri əsas verir. Nəticə etibarilə, neyron şəbəkələri giriş və çıxış dəyişənləri arasında mürəkkəb asılılıqları onların analitik formasını əvvəlcədən müəyyən etmədən modelləşdirməyə imkan verir (Funahashi, 1989; Cybenko, 1989; Hornik et al., 1989).

Sonrakı mərhələdə Kolmoqorov-Arnold yanaşmasına əsaslanan ilk neyron şəbəkə arxitekturaları işlənilib hazırlanıb (Kolmogorov, 1957; Arnold, 1957; Funahashi, 1989; Cybenko, 1989; Hornik et al., 1989). Bu arxitekturalarda gizli laylar Kolmoqorovun daxili funksiyalarına analogi olaraq giriş dəyişənlərini birölçülü komponentlərə çevirmək funksiyasını yerinə yetirir. Tədqiqatçılar (xüsusilə Hecht-Nielsen, 1987) göstəriblər ki, qeyri-xətti aktivasiya funksiyalarına malik üçlaylı feedforward neyron şəbəkəsi Kolmoqorov teoreminin praktik interpretasiyası kimi nəzərdən keçirilə bilər (Funahashi, 1989). Bu yanaşma aşağıdakı nəticələrə gətirib çıxarıb:

- neyron şəbəkələrinin universal approksimasiya qabiliyyətinin riyazi əsaslarının formalaşması;
- çoxlaylı perseptronların, eləcə də sigmoid, tanh və ReLU tipli aktivasiya funksiyalarından istifadə edən modellərin inkişafı.

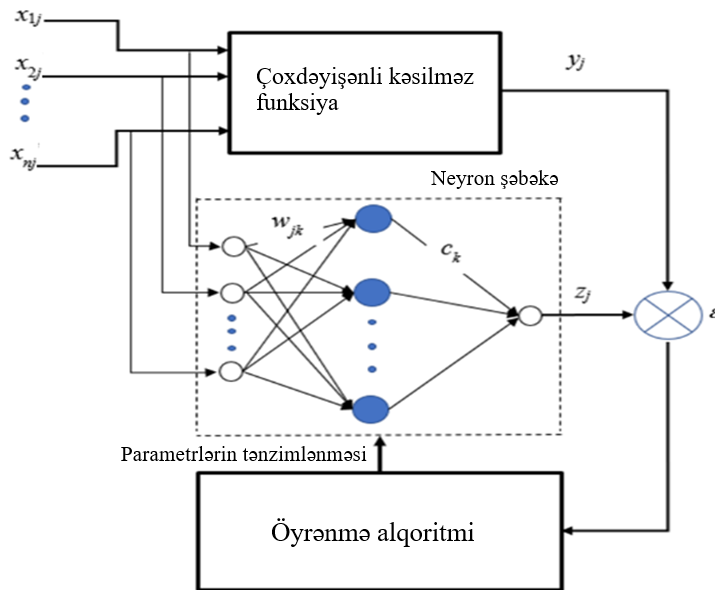
Sonrakı inkişaf mərhələsində Kolmoqorov-Arnold yanaşmasının təsiri bir sıra istiqamətlərdə müşahidə olunur:

- dərin öyrənmə arxitekturalarında (Deep Learning), burada funksiyaların superpozisiyası çoxsəviyyəli təqdimatlar vasitəsilə həyata keçirilir (Goodfellow et al., 2016; LeCun et al., 2015);
- radial bazis funksiyalı şəbəkələrdə (RBFN) və wavelet çevrilmələrində, hansı ki funksional parçalanmanın xüsusi halları kimi şərh olunur (Funahashi, 1989);
- simvolik və hibrid şəbəkələrdə, burada funksional komponentlər mürəkkəb asılılıqların modelləşdirilməsi və interpretasiya olunan təqdimatların qurulması üçün istifadə olunur (Funahashi, 1989; Cybenko, 1989; Hornik et al., 1989).

Kolmoqorov-Arnold teoreminin sxematik təsviri Şəkil 3-də təqdim olunub. Feedforward neyron şəbəkəsinin çıxış siqnalı aşağıdakı formada ifadə olunur:

$$z_j = \sum_{k=1}^m c_k \varphi \left(\sum_{i=1}^n w_{ki} x_{ij} - \theta_k \right), \quad (2)$$

burada n – giriş vektorunun komponentlərinin sayı; m – “gizli” laydakı qeyri-xətti neyronların sayı (modelin qurulması zamanı təyin olunur); w_{ki} və c_k – müvafiq olaraq giriş və çıxış sinaptik əlaqələrinin çəkilişi; θ_k – “gizli” laydakı k -cı neyronun hədd qiyməti (threshold); $\varphi(\cdot)$ – “gizli” laydakı neyronun aktivasiya funksiyasıdır (məsələn, sigmoid funksiya $\varphi(t)=1/(1+e^{-t})$). (2) ifadəsi (1) düsturunun alternativ simvolik təqdimatıdır.



Şəkil 3. Kolmoqorov-Arnold teoreminin neyron şəbəkəsi vasitəsilə realizasiyası

Corc Cybenko tərəfindən verilmiş approksimasiya teoremi (Cybenko, 1989) neyron şəbəkələri nəzəriyyəsinin əsas nəticələrindən biridir. Teoremə görə, Evklid fəzasının kompakt alt çoxluğunda müəyyən edilmiş istənilən kəsilməz funksiya yaxınlaşdırıla bilər. Bu proses qeyri-xətti və kəsilməz aktivasiya funksiyasına malik gizli laylı üçlaylı feedforward neyron şəbəkəsi vasitəsilə həyata keçirilir. Bu nəticə klassik Veyştras teoreminin (1885) müəyyən mənada ümumiləşdirilmiş forması kimi qəbul olunur. Veyştras teoreminə görə, qapalı intervalda verilmiş istənilən kəsilməz funksiya həqiqi əmsallara malik çoxhədli funksiyalar ardıcılığı ilə müntəzəm şəkildə yaxınlaşdırıla bilər (Cybenko, 1989):

$$\forall f(x) \in C[a, b] \text{ və } \varepsilon > 0 \exists P_n(x), |f(x) - P_n(x)| < \varepsilon.$$

Başqa sözlə, Cybenko teoremi Veyştrasın ideyasını analitik funksiyalar kontekstindən hesablama modelləri kontekstinə keçirir (Cybenko, 1989). Veyştras teoremində yaxınlaşdırıcılar çoxhədli funksiyalar olduğu halda, Cybenko teoremində bu rolu neyronların qeyri-xətti aktivasiya funksiyalarının çəkili kombinasiyaları yerinə yetirir.

Beləliklə, neyron şəbəkələri universal approksimator kimi nəzərdən keçirilir: kifayət qədər neyron və uyğun seçilmiş çəkilər olduqda, onlar kəsilməz funksiyaları istənilən dəqiqliklə yaxınlaşdırıla bilər. Bu nəticə Sİ və öyrənmə nəzəriyyəsi üçün fundamental əhəmiyyətə malikdir. O göstərir ki, hətta sadə arxitektura malik neyron şəbəkələri giriş və çıxış dəyişənləri arasındakı mürəkkəb asılılıqları modelləşdirmək qabiliyyətinə malikdir (Hornik et al., 1989). Sonrakı mərhələdə bu yanaşma Hornik, Cybenko və digər tədqiqatçıların işlərində inkişaf etdirilib, universal approksimasiya teoremi daha ümumi formada ifadə olunub. Bu nəticələr neyron şəbəkələrinin adaptiv idarəetmə, reqressiya, təsnifat, klasterləşdirmə və proqnozlaşdırma kimi məsələlərdə tətbiqi üçün riyazi əsas yaradıb (Cybenko, 1989; Hornik et al., 1989). Bu baxımdan, Cybenko teoremi Veyştras teoreminin funksional-konstruktiv interpretasiyası kimi qiymətləndirilə bilər: əgər Veyştras teoremi yaxınlaşdırıcı funksiyanın mövcudluğunu göstərsə, Cybenko teoremi belə bir yaxınlaşdırmanın neyron şəbəkəsi vasitəsilə necə reallaşdırıla biləcəyini izah edir (Cybenko, 1989; Hornik et al., 1989).

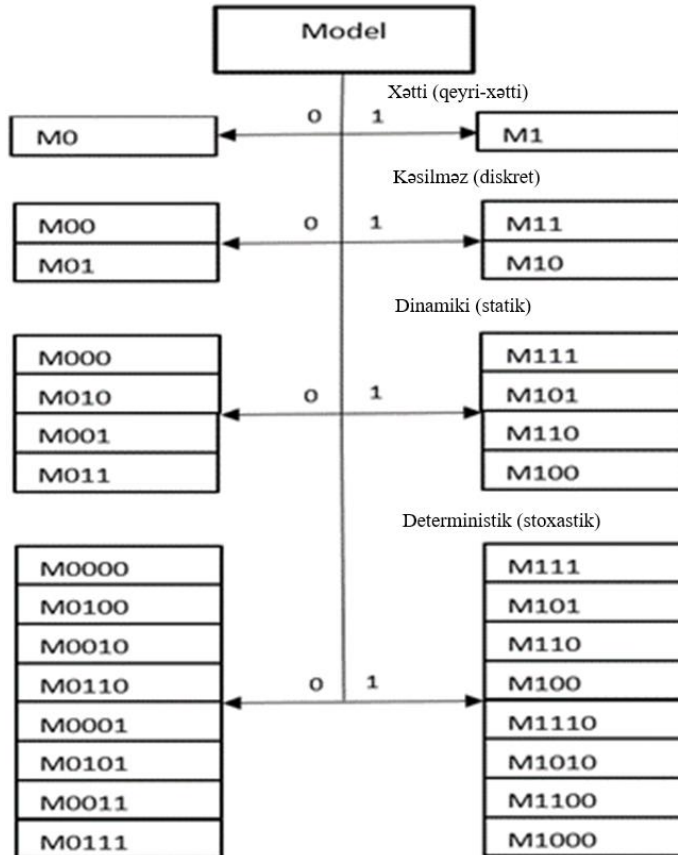
KLASSİK MODELƏRİN MÜASİR ROLU

Müasir mərhələdə klassik yanaşmalar hibrid və neyron şəbəkə strukturlarına inteqrasiya olunaraq aşağıdakı imkanları təmin edir (Funahashi, 1989; Cybenko, 1989; Hornik et al. 1989; Goodfellow et al. 2016; LeCun et al., 2015):

- qərarların interpretasiya oluna bilməsi (məsələn, qaydalara əsaslanan neyron şəbəkələri);
- modellərin dayanıqlığının və fiziki uyğunluğunun təmin edilməsi (physics-informed yanaşmalar);
- məhdud verilənlər şəraitində dəqiqliyin və interpretasiya imkanlarının artırılması.

Bu səbəbdən klassik modellər Sİ ekosisteminin mühüm tərkib hissəsi olaraq qalır və formal məntiq, ehtimal yanaşmaları və adaptiv öyrənmə arasında balans yaradır (Funahashi, 1989; Cybenko, 1989; Hornik et al., 1989). Təlim verilənlərinin həcmnin artırılması Sİ modellərinə daha geniş məlumat əsasında öyrənmək imkanı verir. Bu isə daha böyük verilənlər toplusunun emalı üçün modellərin daha çox parametərə malik olmasını tələb edir və onların effektiv optimallaşdırılmasını şərtləndirir (Goodfellow et al., 2016; LeCun et al., 2015).

Modelləşdirilən obyekt və ya prosesin xüsusiyyətlərindən (məsələn, xətti/qeyri-xətti, kəsilməz/diskret, dinamik/statik və deterministik/stoxastik) asılı olaraq mövcud modellər 16 sinfə bölünür. Konseptual modellərin təsnifatı Şəkil 4-də təqdim olunub. Məsələn, M1010 indeksi xətti, diskret, dinamik və stoxastik modellər sinfini ifadə edir (Funahashi, 1989; Cybenko, 1989; Hornik et al., 1989).



Şəkil 4. Konseptual modellərin təsnifatı

Xətti modellər giriş və çıxış dəyişənləri arasında əlaqəni xətti kombinasiyalar vasitəsilə ifadə edir. Bu modellərə xətti reqressiya, loqistik reqressiya və xətti nüvəyə malik dəstək vektor maşınları (SVM) kimi klassifikatorlar daxildir (Goodfellow et al., 2016; LeCun et al., 2015). Onlar proqnozlaşdırma, ilkin təsnifat və verilənlərin operativ təhlili kimi məsələlərdə geniş istifadə olunur. *Qeyri-xətti modellərdə* isə giriş və çıxış dəyişənləri arasında əlaqə mürəkkəb qeyri-xətti funksiyalar vasitəsilə qurulur. Bu kateqoriyaya neyron şəbəkələri (Funahashi, 1989; Cybenko, 1989; Hornik et al., 1989), qərar ağacları, polinomial və radial nüvəli SVM modelləri, eləcə də Random Forest və Gradient Boosting kimi yanaşmalar daxildir. Bu modellər mürəkkəb asılılıqların modelləşdirilməsi üçün tətbiq olunur və təsvirlərin, nitqin və mətnin emalı kimi sahələrdə effektiv nəticələr verir (Goodfellow et al., 2016; LeCun et al., 2015).

Kəsilməz modellərdə dəyişənlər və funksiyalar müəyyən interval daxilində istənilən qiymət qəbul edə bilər. Bu kateqoriyaya davamlı aktivasiya funksiyalarına malik neyron şəbəkələri (Funahashi, 1989; Cybenko, 1989; Hornik et al., 1989), reqressiya modelləri və diferensiallaşdırıla bilən optimizasiya üsulları daxildir. *Diskret modellərdə* isə dəyişənlər sonlu və ya sayılabilən qiymətlər çoxluğunu qəbul edir. Bu modellərə qərar ağacları, qraf modelləri (məsələn, Bayesian şəbəkələr), Markov zəncirləri və gücləndirici öyrənmədə (Reinforcement Learning) diskret vəziyyət və hərəkətlər üçün tətbiq olunan Q-learning kimi yanaşmalar daxildir (Hochreiter and Schmidhuber, 1997; Sutton and Barto, 2018).

Statik modellər zaman faktorunu nəzərə almır və çıxış yalnız cari girişdən asılı olur. Bu tip modellərə xətti reqressiya, müxtəlif klassifikatorlar və feedforward neyron şəbəkələri daxildir (Funahashi, 1989; Cybenko, 1989; Hornik et al., 1989). *Dinamik modellər* isə zaman asılılığını nəzərə alır və əvvəlki vəziyyətlərin cari nəticəyə təsirini modelləşdirir. Bu kateqoriyaya rekurrent neyron şəbəkələri (RNN), LSTM və GRU arxitekturaları, gizli Markov modelləri (HMM), Kalman filtri və vəziyyətlərə əsaslanan gücləndirici öyrənmə modelləri daxildir (Hochreiter and Schmidhuber, 1997; Sutton and Barto, 2018).

Deterministik modellər eyni giriş məlumatları üçün həmişə eyni nəticəni verir. Bu kateqoriyaya xətti reqressiya, SVM modelləri, stoxastik komponentlərsiz neyron şəbəkələri və təsadüfi elementlərdən istifadə etməyən qərar ağacları daxildir. *Stoxastik modellərdə* isə ehtimal mexanizmlərinin mövcudluğu səbəbindən eyni girişlər üçün fərqli nəticələr əldə oluna bilər. Bu modellərə Bayesian şəbəkələr, Gaussian proseslər, neyron şəbəkələrdə Dropout texnikası və ehtimalı siyasətlərə əsaslanan gücləndirici öyrənmə modelləri daxildir (Hochreiter and Schmidhuber, 1997; Sutton and Barto, 2018).

Cədvəl 1-də Sİ tətbiqlərində istifadə olunan birləşdirilən model təsnifatının bəzi geniş yayılan nümunələri təqdim olunub. Model seçimi müstəqil və “şüurlu” qərar deyil, əvvəlcədən müəyyən edilmiş kriteriyalar və alqoritmlər əsasında həyata keçirilən sistemli prosesdir. Bu proses vəzifənin məqsədi, verilənlərin xüsusiyyətləri, performans göstəriciləri və modellərin müqayisəsi kimi amillərə əsaslanır (Goodfellow et al., 2016; LeCun et al., 2015; Hochreiter and Schmidhuber, 1997).

Beləliklə, Sİ-də model seçimi məsələ tipindən, verilənlərin xüsusiyyətlərindən, keyfiyyət kriteriyalarından və hesablama məhdudiyyətlərindən asılı olaraq həyata keçirilir. Təcrübədə bu proses çox vaxt avtomatlaşdırılmış müqayisə və qiymətləndirmə üsulları (AutoML) vasitəsilə reallaşdırılır (Goodfellow et al., 2016; Hochreiter and Schmidhuber, 1997).

Cədvəl 1. Birləşdirilmiş model təsnifatından bəzi nümunələr

Konseptual model	Xətti	Kəsilməz	Dinamik	Deterministik	Nümunələr
Xətti reqressiya	Xətti	Kəsilməz	Statik	Deterministik	LASSO və Ridge reqressiya
Feedforward neyron şəbəkə	Qeyri-xətti	Kəsilməz	Statik	Deterministik / Stoxastik	Çoxlaylı perceptron
RNN / LSTM	Qeyri-xətti	Kəsilməz	Dinamiki	Deterministik / Stoxastik	Zaman sıraları
Həllər ağacı	Qeyri-xətti	Diskret	Statik	Deterministik	Təsnifat, reqressiya
HMM	Qeyri-xətti	Diskret	Dinamiki	Stoxastik	Sıra seqmentasiyası, nitqin tanınması
Gaussian Process	Qeyri-xətti	Kəsilməz	Statik	Stoxastik	Ehtimallı qiymətləndirmə ilə reqressiya
Ehtimallı siyasətə malik RL	Qeyri-xətti	Diskret / Kəsilməz	Dinamiki	Stoxastik	Oyunlar, robotlarla idarəetmə

BELLMANA GÖRƏ “ƏDƏDLƏRİN LƏNƏTİ”

“Ölçülərin lənəti” termini Richard Bellman tərəfindən təklif edilib və yüksək ölçülü fəzalarda məlumatların emalı zamanı yaranan çətinlikləri xarakterizə edir. Ölçülərin sayı artdıqca fəzanın həcmi eksponentsial şəkildə böyüyür, bu isə məlumatların seyrəkliyinə, approksimasiya keyfiyyətinin azalmasına, hesablama xərclərinin artmasına və analiz metodlarının effektivliyinin aşağı düşməsinə səbəb olur. Bu hallar təsnifat, klasterləşdirmə və reqressiya məsələlərini mürəkkəbləşdirir və çoxölçülü məlumatların emalı üçün ölçülərin azaldılması və regularizasiya kimi metodların tətbiqini zəruri edir.

Neyron şəbəkələrinin ilkin inkişaf mərhələsində bu fenomen mühüm məhdudiyyətlərdən biri olub. Giriş dəyişənlərinin sayı artdıqca təlim fəzasının ölçüsü eksponentsial şəkildə böyüyür, nəticədə, məlumatların seyrəkliyi artır, nöqtələr arasındakı məsafələr dəyişir və etibarlı ümumiləşdirmə üçün daha böyük həcmdə təlim verilənləri tələb olunur. Bu amillər ilkin neyron şəbəkə modellərinin mürəkkəb çoxölçülü asılılıqları modelləşdirmə qabiliyyətini məhdudlaşdırırdı. Bu problem xüsusilə perseptron modellərində aydın şəkildə müşahidə olunurdu. 1950–1960-cı illərdə təklif olunmuş ilk neyron şəbəkə modelləri sadə, birlaylı strukturlardan ibarət idi və yalnız xətti ayırıcı funksiyaları reallaşdırırdı. Giriş xüsusiyyətlərinin sayının artması ilə modelin effektiv öyrənməsi üçün tələb olunan təlim nümunələrinin sayı sürətlə artırdı ki, bu da onların praktik tətbiq imkanlarını məhdudlaşdırırdı.

Bu fenomeni üçlaylı perseptron nümunəsində nəzərdən keçirək: gizli layda 50, giriş layda isə 100 neyron olsun. Tutaq ki, modelin öyrədilməsi üçün 1000 təlim nümunəsi mövcuddur. Bu halda

müşahidələrin model parametrlərinə nisbəti $1000/5000=0.2$ təşkil edir. Belə aşağı nisbət modelin ümumiləşdirmə qabiliyyətini məhdudlaşdırır, yenidən öyrənmə (overfitting) riskini artırır və öyrənmə prosesinin effektivliyini azaldır, çünki mövcud məlumat həcmi bütün parametrlərin etibarlı qiymətləndirilməsi üçün kifayət etmir. Əgər giriş dəyişənlərinin sayını 1000-ə qədər artırıbsaq, çəkirlərin sayı $1000 \times 50 = 50\,000$ olacaq. Bu halda məlumat-parametr nisbətini, məsələn, 10:1 səviyyəsində saxlamaq üçün təxminən 500000 təlim nümunəsi tələb olunur. Bu həcm həmin dövrdə həm texniki, həm də praktik baxımdan çətin əldə edilən resurs idi.

Bundan əlavə, ilkin arxitekturaların məhdudluğu mürəkkəb çoxölçülü funksiyaların səmərəli approksimasiyasına imkan vermirdi. Bu isə yüksək təsnifat səhvlərinə və zəif ümumiləşdirmə qabiliyyətinə səbəb olurdu. Beləliklə, ölçülərin lənəti və birlaylı perseptronların struktur məhdudyyətləri erkən neyron modellərinin əsas çətinliklərini üzə çıxarıb və çoxlaylı şəbəkələrin, eləcə də yüksək ölçülü verilənlərlə işləyə bilən öyrənmə metodlarının inkişafını stimullaşdırıb. Müasir modellərdə isə parametr sayı artıq çox böyük miqyaslara çatır (Goodfellow et al., 2016; LeCun et al., 2015).

Müasir dərin öyrənmə (Deep Learning) üsulları ölçülərin lənətinin təsirini azaltmaq üçün müxtəlif yanaşmalardan istifadə edir. Bu yanaşmalara ölçülərin azaldılması, regularizasiya, normallaşdırma, eləcə də konvolyusion və rekurrent arxitekturalar daxildir (Goodfellow et al., 2016; LeCun et al., 2015; Shannon, 1948). Bu metodlar yüksək ölçülü fəzalarda mürəkkəb asılılıqların effektiv şəkildə modelləşdirilməsinə imkan verir. Böyük həcmli verilənlər üzərində öyrənmə zamanı modellərin əlavə parametrlərə ehtiyacı yaranır ki, bu da məlumatdakı qanunauyğunluqların daha dəqiq aşkar edilməsini və öyrənmə prosesinin effektivliyini təmin edir.

NƏTİCƏ

David Hilbertin 13-cü məsələsindən müasir böyük dil modellərinə qədər olan inkişaf xətti çoxdəyişənli funksiyaların nəzəri tədqiqindən onların praktik modelləşdirilməsinə keçidi göstərir. Əgər Hilbertin qoyduğu məsələ mürəkkəb funksiyaların strukturunun öyrənilməsi ilə bağlı idisə, müasir neyron şəbəkə modelləri bu tip asılılıqların hesablamada reallaşdırılmasına imkan verib. Kolmoqorov, Arnold, Funahaşi və Rumelhartın ideyalarının inkişafı universal approksimasiya konsepsiyasının formalaşmasına gətirib çıxarıb. Bu nəticələr göstərir ki, uyğun arxitekturalara və kifayət qədər parametrlə malik neyron şəbəkələri kəsilməz funksiyaları verilmiş dəqiqliklə yaxınlaşdırıla bilər və bu, dərin öyrənmə modellərinin nəzəri əsasını təşkil edir. Müasir böyük dil modelləri bu yanaşmanın tətbiq sahələrindən biridir. Bu modellər mətn məlumatları əsasında öyrənərək dil strukturlarını modelləşdirir, kontekstual asılılıqları nəzərə alır və müxtəlif tipli idrak funksiyalarını (məsələn, mətnin analizi, generasiyası və interpretasiyası) reallaşdırır.

Beləliklə, Sİ inkişafı nəzəri riyaziyyat və mühəndislik yanaşmalarının qarşılıqlı təsiri nəticəsində formalaşmışdır. Bu proses idrakın modelləşdirilməsi istiqamətində ardıcıl inkişaf mərhələlərini əks etdirir və müasir Sİ sistemlərinin yaranmasına gətirib çıxarıb.

İstifadə edilmiş ədəbiyyat

1. Arnold, V.I. (1957). On functions of three variables. *Doklady Akademii Nauk SSSR*, vol. 114, no 4, pp. 679–681,
2. Brown, T. et al. (2020). “Language models are few-shot learners” in *Adv. Neural Inf. Process. Syst.*
3. Cybenko, G. (1989). Approximation by superpositions of a sigmoidal function. *Math. Control Signal Systems 2*, pp. 303–314. <https://doi.org/10.1007/BF02551274>
4. Devlin, J., Chang, M.W., Lee, K., and Toutanova, K. (2019). “BERT: Pre-training of deep bidirectional transformers for language understanding” in *Proc. NAACL-HLT*.
5. Dormehl, L. (2017). *Thinking Machines: The Quest for Artificial Intelligence and Where It's Taking Us Next*. Tarcher.
6. Funahashi, K. (1989). “On the approximate realization of continuous mappings by neural networks”. *Neural Networks*, vol. 2, no. 3, pp. 183–192.
7. Goodfellow, I., Bengio, Y., and Courville, A. (2016). *Deep Learning*. MIT Press.
8. Graves, A. (2012). “Supervised sequence labelling”. in *Studies in Computational Intelligence*. Springer.
9. Hilbert, D. (1900). Mathematical Problems. *Bulletin of the American Mathematical Society*, 8 (10): 437–479. doi:10.1090/S0002-9904-1902-00923-3.
10. Hinton, G. (1989). “Connectionist learning procedures”. *Artif. Intell.*, vol. 40, no. 1–3, pp. 185–234.
11. Hochreiter, S., and Schmidhuber, J. (1997). “Long short-term memory”. *Neural Computing*, vol. 9, no. 8, pp. 1735–1780.
12. Hornik, K., Stinchcombe, M., and White, H. (1989). “Multilayer feedforward networks are universal approximators”. *Neural Networks*, vol. 2, no. 5, pp. 359–366.
13. Kang, M. (2011). *Sublime Dreams of Living Machines: The Automaton in the European Imagination*. Harvard University Press.
14. Kolmogorov, A.N. (1957). On the representation of continuous functions of many variables by superposition of continuous functions of one variable and addition. *Doklady Akademii Nauk SSSR*, vol. 114, no. 5, pp. 953–956.
15. LeCun, Y., Bengio, Y., and Hinton, G. (2015). “Deep learning”. *Nature*, vol. 521, pp. 436–444.
16. McCulloch, W., and Pitts, W. (1943). “A logical calculus of the ideas immanent in nervous activity”. *Bull. Math. Biophys.*, vol. 5, no. 4, pp. 115–133.
17. Publius Ovidius Naso. (2009). *Metamorphoses*. Translated by A.D. Melville; introduction and notes by E. J. Kenney. Oxford: Oxford University Press.
18. Rosenblatt, F. (1958). “The perceptron: A probabilistic model for information storage and organization in the brain”. *Psychol. Rev.*, vol. 65, no. 6, pp. 386–408.
19. Rumelhart, D.E., Hinton, G.E., and Williams, R.J. (1986). “Learning representations by back-propagating errors”. *Nature*, vol. 323, no. 6088, pp. 533–536.
20. Scholem, G. (1990). *The Golem: Jewish Magical and Mystical Traditions*. State University of New York Press, Alban.
21. Shannon, C.E. (1948). “A mathematical theory of communication”. *Bell Syst. Tech. J.*, vol. 27, pp. 379–423, 623–656.
22. Sutton, R., and Barto, A. (2018). *Reinforcement Learning: An Introduction*, 2nd ed. MIT Press, 2018.
23. Turing, A. (1950). Computing Machinery and Intelligence. *Mind*, 59(236), 433–460.
24. Vaswani, A. et al. (2017). “Attention is all you need” in *Proc. Adv. Neural Inf. Process. Syst. (NeurIPS)*, pp. 5998–6008.