



Süni intellekt sistemlərində dil texnologiyaları: formalaşma mərhələləri və tətbiqi perspektivləri

Language Technologies in AI Systems: Evolution Stages and Application Prospects

Arzu Novruz Revina / Arzu Novruz Revina 

Azerbaijan University,

Department of Foreign Languages, Senior Lecturer

arzu.revina@au.edu.az

ORCID ID: 0000-0002-6868-5849 <i>Daxil oldu</i> Received: 10.02.2026 <i>Qəbul olundu</i> Accepted: 11.03.2026	Xülasə. Məqalədə süni intellekt (Sİ) və təbii dilin emalı (TDE) texnologiyalarının təkamül mərhələləri ilə fundamental elmi əsasları araşdırılır. Tədqiqatda Sİ-nin ilkin fəlsəfi-mexaniki modellərindən tutmuş, A. Tyurin qin hesablama nəzəriyyəsi, Tyuring testi və XX əsrin fundamental TDE hadisələrinə qədər olan dövr təhlil süzgəcindən keçirilir. Həmçinin N. Chomskynin universal qrammatika konsepsiyasının erkən qayda-əsaslı sistemlərə təsiri, XXI əsrdə isə statistik və neyron şəbəkə modellərinin inkişaf trayektoriyası nəzərdən keçirilir. Məqalədə TDE-nin virtual assistentlər, sentiment analizi, tibbi-hüquqi mətnlərin avtomatlaşdırılmış emalı, maşın tərcüməsi və chatbotlar kimi müasir praktik tətbiqləri qruplaşdırılır. Yekun olaraq, intellektual linqvistika inteqrasiyasının gələcək innovasiyalardakı rolu qiymətləndirilir, eləcə də Azərbaycan dilinin rəqəmsal resurslarının mövcud vəziyyəti, rəqəmsallaşma çətinlikləri və inkişaf imkanları şərh olunur. Açar sözlər. Süni intellekt, təbii dilin emalı, neyron şəbəkə modelləri, NLP tətbiqləri, Azərbaycan dilinin rəqəmsallaşdırılması. Abstract. The article examines the evolutionary stages and fundamental scientific foundations of Artificial Intelligence (AI) and Natural Language Processing (NLP) technologies. The study analyzes the historical trajectory ranging from the early philosophical and mechanical models of AI to A. Turing's computability theory, the Turing test, and the seminal NLP milestones of the 20th century. Furthermore, it reviews the impact of N. Chomsky's universal grammar concept on early rule-based systems, as well as the developmental path of statistical and neural network models during the transition into the 21st century. The paper categorizes contemporary practical applications of NLP, such as virtual assistants, sentiment analysis, automated processing of medical and legal texts, machine translation, and chatbots. Conclusively, the role of intelligent linguistic integration in future innovations is evaluated, while the current state of digital resources for the Azerbaijani language, its digitalization challenges, and development prospects are discussed. Keywords. Artificial intelligence, natural language processing, neural network models, NLP applications, digitalization of the Azerbaijani language. Резюме. В статье исследуются эволюционные этапы и фундаментальные научные основы технологий искусственного интеллекта (ИИ) и обработки естественного языка (NLP). В исследовании анализируется историческая траектория — от ранних философско-механических моделей ИИ до теории вычислений А. Тьюринга, теста Тьюринга и основополагающих событий в области NLP XX века. Также рассматривается влияние концепции универсальной грамматики Н. Хомского на ранние системы, основанные на правилах, и траектория развития статистических и нейросетевых моделей в XXI веке. В работе классифицируются такие современные практические приложения NLP, как виртуальные ассистенты, анализ тональности (сентимент-анализ), автоматизированная обработка медицинских и юридических текстов, машинный перевод и чат-боты. В заключение оценивается роль интеграции интеллектуальной лингвистики в будущие инновации, а также комментируются текущее состояние цифровых ресурсов азербайджанского языка, трудности его цифровизации и перспективы развития. Ключевые слова: Искусственный интеллект, обработка естественного языка, нейросетевые модели, приложения NLP, цифровизация азербайджанского языка..
--	---

Giriş

XXI əsrdə süni intellekt texnologiyalarının sürətli inkişafı təbii dilin emalı (TDE) sahəsində də ciddi irəliləyişlərə səbəb olmuşdur. Hesablama gücünün artması və alqoritmik yanaşmaların təkmilləşməsi nəticəsində kompüterlərin insan dilini anlama və cavablandırma qabiliyyəti nəzəri konsepsiyadan çıxaraq praktiki reallığa çevrilmişdir. TDE informatika, dilçilik, psixologiya və süni intellektin qovşağında yerləşən, mürəkkəb və multidissiplinar bir sahə olaraq, həm elmi tədqiqatlara, həm də müxtəlif tətbiq sahələrinə mühüm təsir göstərməkdədir.

Tarixi inkişaf baxımından TDE əvvəlcə qaydalara əsaslanan yanaşmalarla başlamış, sonralar statistik modellərlə davam etmiş və nəhayət, dərin öyrənmə texnologiyalarının tətbiqi ilə yeni mərhələyə qədəm qoymuşdur. Bu inkişaf prosesi yalnız texniki deyil, həm də nəzəri və fəlsəfi səviyyədə bu sahənin əsaslarını dəyişdirmişdir. Turing testi və süni intellektin mahiyyəti ilə bağlı aparılan müzakirələr də bu sahənin elmi çərçivəsinin formalaşmasında mühüm rol oynamışdır. Məqalədə TDE-nin tarixi mərhələləri, əsas modelləri və çağdaş texnologiyalar əsasında yaranan problemlər və imkanlar araşdırılır. Tədqiqatın məqsədi süni intellekt və təbii dilin emalı sahələrinin paralel inkişafını elmi və tarixi aspektlərdən təhlil etmək, eyni zamanda Azərbaycan dilinin bu sahədəki mövcud vəziyyətinə və inkişaf perspektivlərinə də diqqət yetirməkdir.

Süni intellektin formalaşmasında tarixi və elmi əsaslar

Süni intellekt anlayışı müasir dövrün məhsulu kimi görünə də, bu konsepsiyanın əsası qədim fəlsəfi və mexaniki düşüncələrə dayanır. Hələ XVII əsrdə mexaniki sistemlərin insan və heyvan davranışlarını təqlid edə biləcəyi fikri geniş yayılmışdı. Məşhur riyaziyyatçı və filosof Rene Dekart insanı mürəkkəb bir mexanizmə bənzədərək onun davranışlarını saat mexanizmi ilə izah etməyə çalışmışdır. Bu dövrdə “ağıllı maşın” fikri yalnız texnoloji deyil, həm də fəlsəfi baxışların predmetinə çevrilmişdi. XIX əsrin əvvəllərində ingilis riyaziyyatçısı Çarlz Bebbic tərəfindən hazırlanan “Fərq Maşını” (Difference Engine) isə bu ideyaların praktikaya çevrilməsi yolunda atılmış əhəmiyyətli addımlardan biri oldu. Bu mexanizm yalnız riyazi əməliyyatlar aparmaqla kifayətlənmirdi, həm də müəyyən dərəcədə yaddaş tutumu və oyun strategiyalarını modelləşdirmək qabiliyyətinə malik idi (Schultz və Ellen-Schultz, 2007). Bu, insan aqlının texnoloji modelləşdirilməsi istiqamətində ilkin təcrübələrdən sayılırdı.

XX əsrdə süni intellekt sahəsində əsas dönüş nöqtəsi İkinci Dünya Müharibəsi zamanı müşahidə olundu. Məhz bu dövrdə Alan Tyuring tərəfindən hazırlanan kodqırma cihazı

“Bombe” müharibənin nəticəsinə ciddi təsir göstərdi (Gladwin, 1997). Daha sonra 1950-ci ildə Tyuringin “Hesablama Maşınları və İntellekt” adlı məqaləsində səsləndirdiyi məşhur “Maşınlar düşünə bilərmidi?” sualı süni intellektin nəzəri əsaslarını formalaşdırdı (Turing, 1950).

“Süni intellekt” (Artificial Intelligence) termini isə ilk dəfə 1956-cı ildə Dartmut Kollecinə keçirilən elmi seminar zamanı Con Makkarti tərəfindən irəli sürüldü (Dick, 2019). Bu tədbir, süni intellekt sahəsində akademik tədqiqatların intensivləşməsinə səbəb oldu. Sonrakı onilliklərdə Eliza, Stajyer, Bilginkimi sistemlər yaradıldı (Kutlusoy, 2019) və 1970-ci illərdə Yaponiya tərəfindən WABOT-I adlı ilk insana bənzər robot təqdim olundu (Acar, 2020). Bununla belə, 1974–1980-ci illərdə sahəyə olan marağın azalması və maliyyələşdirmənin dayandırılması nəticəsində “Süni İntellekt Qışı” adlandırılan dövr yaranmışdır (Öztürk və Şahin, 2018). 1980-ci illərin ortalarından etibarən bir sıra dövlətlər, xüsusən də Böyük Britaniya və Yaponiya, süni intellekt sahəsinin inkişafını yenidən dəstəkləməyə başladılar. 1997-ci ildə IBM tərəfindən yaradılan Deep Blue proqramının şahmat üzrə dünya çempionu Qari Kasparovu məğlub etməsi süni intellektin potensialını geniş ictimaiyyətə nümayiş etdirdi (Schultz və Ellen-Schultz, 2007).

XXI əsrin əvvəllərində süni intellekt gündəlik həyatımıza daxil olmağa başladı. “Roomba” kimi ağıllı məişət cihazları, sosial media platformalarının (Facebook, Netflix, Twitter və s.) alqoritmlərində süni intellektin istifadəsi bu texnologiyanın əhatə dairəsini daha da genişləndirdi (Acar, 2020). 2011-ci ildə IBM-in Watson sistemi televiziya intellektual yarışmada insanları məğlub etdi (Sariel, 2017), Google isə 2018-ci ildə səsli danışqla telefon zəngi edən və insan kimi davranan “Duplex” sistemini təqdim etdi (Acar, 2020).

Bu inkişafılar süni intellektin yalnız nəzəri bir tədqiqat sahəsi olmaqdan çıxaraq, tibbdə, nəqliyyatda, təhsildə və bir çox başqa sektorlarda tətbiq olunan praktik bir texnologiyaya çevrildiyini göstərir. Lakin bütün bu nailiyyətlərə baxmayaraq, süni intellektin insan aqlını nə dərəcədə təqlid etdiyi və ya gələcəkdə onu ötüb keçə biləcəyi məsələsi hələ də müzakirə mövzusudur (Kırpınar, 2019).

Göründüyü kimi, süni intellektin inkişafı yalnız texnoloji tərəqqi ilə məhdudlaşmır, o, eyni zamanda insan düşüncəsinin, məntiq və davranış modellərinin də dərin şəkildə öyrənilməsinə tələb edir. Bu prosesdə dilin və ünsiyyətin rolu xüsusilə əhəmiyyətlidir, çünki intellektual sistemlərin insanlar kimi başa düşməsi və cavab verməsi üçün dil əsas vasitədir. Elə bu səbəbdən, süni intellektin ən mühüm istiqamətlərindən biri kimi Təbii dilin emalı (TDE) sahəsi ön plana çıxır.

Təbii dilin emalının tarixi inkişaf mərhələləri və elmi əsasları

Dil yalnız düşüncə və informasiyanın deyil, həm də hisslərin, qeyri-müəyyənliklərin və natamamlıqların ötürülməsində mühüm vasitə kimi çıxış edir. Onu sadəcə riyazi qaydalar toplusu kimi təsəvvür etmək çətindir. Cümlə və ya ifadənin yaranması, dilin daxili məntiqi strukturuna əsaslanarsa da, bu strukturu texnoloji şəkildə modelləşdirərək kompüterlərə başa salmaq və onların da eyni şəkildə cavab verməsini təmin etmək mürəkkəb məsələ olaraq qalır.

Təbii dilin maşınlar tərəfindən işlənməsi ilə bağlı ilk konseptual yanaşmalar XX əsrin 1930-cu illərinə təsadüf edir. Həmin dövrdə “tərcümə edən maşın” anlayışı ilk dəfə patentləşdirilmişdi. Bu istiqamətdə iki fərqli patent təqdim olunmuşdu. Birincisi, fransalı mühəndis Jorj Artsruni tərəfindən irəli sürülmüş və ikidilli lüğətdən istifadə etməklə bir dildəki sözləri digər dilə kağız lent üzərində çevirməyə əsaslanmışdı. Bu yanaşma sadə idi və dilin qrammatik aspektlərini nəzərə almırdı. Digər yanaşmanı isə rusiyalı Piter Troyanski təklif etmişdi. O, ikidilli lüğətə əlavə olaraq, Esperanto dili üzərində qurulmuş dil qrammatikasının emalını təmin edən daha sistemli bir metod irəli sürmüşdü (Hutchins, 1997). Hər iki yanaşma yenilikçi olsa da, praktiki tətbiq imkanları məhdud idi və nəzəri səviyyədən kənara çıxmamışdı.

Təbii dilin emalı ilə bağlı ilk real tətbiqlər isə İkinci Dünya Müharibəsi dövründə müşahidə olunmağa başlandı. Almanların məxfi yazışmalarını şifrələmək üçün istifadə etdikləri “Enigma” adlı cihaz, Alman hərbi hissələrinin Avropa boyunca əlaqələndirilməsi və hərbi əməliyyatların planlaşdırılmasında mühüm rol oynayırdı. Bu texnologiya, o dövrdə gizli rabitənin təmin olunmasında böyük nailiyyət hesab olunurdu. 1946-cı ildə isə Böyük Britaniya tərəfindən “Colossus” adlı cihaz yaradıldı. Bu cihaz, “Tunny” adlandırılan şifrə sistemini uğurla deşifrə edə bilirdi. Bu hadisə, müharibənin gedişatında dönüş nöqtəsi oldu, çünki Britaniya qüvvələri almanların mövqeləri, strategiyaları və qüvvələri barədə öncədən məlumat əldə edərək təcəbirlər görə bildilər. Bütün bu əməliyyatlar Britaniyanın kriptografik mərkəzi olan Bletchley Parkda həyata keçirilirdi. Burada Alan Tyuring və digər mütəxəssislərin əməkdaşlığı nəticəsində Enigma sisteminin açılması mümkün olmuşdu. Bu hadisə yalnız müharibə tarixində deyil, eyni zamanda müasir kompüter elminin və təbii dilin emalının inkişafında da mühüm mərhələ hesab olunur (Gladwin, 1997). 1950-ci ildə Alan Turing məşhur “Turing testi” (yəni təqlid oyunu) konsepsiyasını təqdim etdi. Bu testdə üç iştirakçı olur: kişi (oyunçu A), qadın (oyunçu B) və sorğuçu (oyunçu C). Sorğuçu yalnız yazılı mesajlar vasitəsilə oyunçuların cinsiyyətini müəyyən etməyə çalışır. Burada oyunçu B doğru cavab verməklə sorğuçuya kömək edir, oyunçu A isə onu çaşdırmağa çalışır. Turingin təklifi ondan ibarət idi ki, oyunçu A-nın yerinə bir maşın qoyulsun.

Əgər sorğuçu maşını digər oyunçudan fərqləndirə bilməzsə, maşın testi keçmiş sayılır, əks halda isə uğursuzluq kimi qiymətləndirilir. Test çərçivəsində maşından mürəkkəb məsələləri həll etməsi və ya sənət əsəri yaratması tələb olunmur – əsas meyar gündəlik söhbətləri insan kimi aparmaq və dialoqun kontekstini başa düşmək qabiliyyətidir. Turing üçün belə bir testi keçmək maşının düşünə bilmə potensialına malik olduğuna dəlalət edirdi (Turing, 1950).

Turingin bu fikirləri onun daha əvvəlki işləri – Gödelin natamliq teoremi və hesablama nəzəriyyəsində irəli sürdüyü prinsiplərlə sıx bağlı idi. Turing testi əslində bu sualı ortaya qoyur: görəsən, şüurlu görünmək kimi bir qabiliyyət Turing maşınının hesablaya biləcəyi problemlər sinfinə daxildirmi, yoxsa bu, formal olaraq sübut edilə bilməyən, lakin doğru olan şeylər sinfinə aiddir? Test sadədir, amma ortaya qoyduğu sual son dərəcə fundamental və dərinidir. Bu sual yalnız şüur və hesablama arasındakı əlaqəni deyil, həm də maşınların insan düşüncəsinə ən yaxın fəaliyyət sahələrindən biri olan təbii dili başa düşmək qabiliyyətini ön plana çıxarır.

Süni intellektin əsas problemlərindən biri insan dilini başa düşmək və emal etməkdir. Bu isə təbii dilin strukturlarını və qrammatikasını dərinləndirən öyrənməyi tələb edir. Hər hansı bir dilin tam və düzgün işləməsi üçün qrammatikanın tətbiqi zəruridir. Bir cümlənin mənası yalnız qrammatik qaydaların doğru şəkildə tətbiqi nəticəsində tam şəkildə anlaşılır. 1957-ci ilə qədər maşınlar qrammatik strukturların daxil edilməsi və onların mənanı başa düşməsi istiqamətində ciddi irəliləyiş əldə olunmamışdı. Məhz həmin ildə Noam Chomsky-nin təqdim etdiyi “Sintaktik Strukturlar” əsəri bu sahədə inqilabi dönüş nöqtəsi oldu. Chomsky “dil strukturunun formal nəzəriyyəsi”ni əsas götürərək universal qrammatika əsasında güclü və sistemli dil qaydalarının formalaşdırılmasına çalışdı (Lees, 1957). Lakin daha sonra Charles Hockett Chomsky-nin yanaşmasının bəzi məhdudiyyətlərini vurğuladı. O, Chomsky-nin dili sabit və tam müəyyən edilmiş formal sistem kimi qəbul etməsinin reallıqda ideal şəraitə uyğun olmadığını bildirdi (Hockett, 1966).

Təbii dilin emalı sahəsində erkən istiqamətlərdən biri maşın tərcüməsi olmuşdur. Maşın tərcüməsinin əsas məqsədi, bir təbii dildən digərinə mətn və ya nitqin avtomatik şəkildə tərcüməsini həyata keçirə biləcək proqramlar hazırlamaqdır. Məsələn, 1954-cü ildə Corcaun Universiteti ilə IBM şirkəti arasında həyata keçirilən birgə layihə nəticəsində avtomatik mətn tərcüməsi proqramı hazırlanmışdır. Aparılan təcrübələr göstərmişdir ki, həmin proqram altmış rus cümləsini uğurla ingilis dilinə tərcümə edə bilmişdir. Bu sistem cümlələri birbaşa xəritələndirərək və müəllif tərəfindən xüsusi hazırlanmış lüğətdən istifadə etməklə işləyirdi.

Müəllif iddia edirdi ki, yaxın gələcəkdə maşın ardıcıl və etibarlı tərcümələr edə biləcək. Lakin inkişaf gözlənilmədiyindən xeyli yavaş gedirdi, bu da layihəyə ayrılan maliyyənin kəskin azalmasına səbəb oldu.

1960-cı illərin sonlarında Massaçusets Texnologiya İnstitutunda (MIT) Terri Vinograd tərəfindən SHRDLU adlı maşın hazırlanmışdır. Bu proqram təbii dilin emalı sahəsində ilk kompüter proqramı kimi obyektləri hərəkət etdirmək, cari vəziyyəti müəyyənləşdirmək və adları yadda saxlamaq kimi funksiyaları yerinə yetirə bilirdi. Bütün bu əməliyyatlar “bloklar dünyası” adlı sadələşdirilmiş mühitdə həyata keçirilirdi. SHRDLU süni intellekt sahəsində mühüm irəliləyiş hesab olunaraq tədqiqatçıların diqqətini bu texnologiyanın inkişafına cəlb etdi. Lakin proqramın uğuru daha mürəkkəb və qeyri-müəyyən real dünya situasiyalarını təhlil etməkdə məhdudlaşdı və bu səbəbdən inkişafı müəyyən qədər dayanmışdır (Winograd, 1980). 1969-cu ildə Roger Şank cümlənin mənasının daha yaxşı başa düşülməsi üçün “token” konsepsiyasını təqdim etdi. Bu tokenlərə real dünya obyektləri, hadisələr, zaman və məkan daxildir. Hər bir cümlə üçün bu tokenlər maşına hadisələrin gedişi və orada iştirak edən obyektlər barədə daha dəqiq anlayış qazandırır (Schank, 1969).

İndiyədək maşının cümləni başa düşməsi üçün tətbiq olunan bütün qaydalar əsasən ifadələrin quruluşuna (frazə strukturlarına) əsaslanırdı. Tədqiqatçılar maşının ifadənin mənasını anlamaq üçün izləyə biləcəyi aydın və dəqiq qaydalar toplusu yaratmağa çalışırdılar. 1970-ci ildə Uilyam Vuds təbii dillərin təmsili üçün artırılmış keçid şəbəkələri (Augmented Transition Networks — ATN) konseptini təqdim etdi. Bu yanaşmada sonlu avtomatlar və rekursiya istifadə olunaraq, mövcud məlumat əsasında ifadənin mənası müəyyən edilir. ATN-lər, yalnız cümlənin müəyyən hissəsi məlum olduqda belə mümkün olan nəticəni verə bilər və əlavə məlumat daxil olduqca ifadənin mənasında dəyişikliklər edə bilər. Əgər verilən ifadə qeyri-müəyyənlik yaradırsa, sistem ehtimal əsaslı nəticə çıxarır və lazım gəldikdə, əlavə məlumat əldə olunana qədər qərar qəbul etməyi təxirə salır. Bu yolla ATN-lər, kifayət qədər məlumat olmadıqda problemi rekursiv şəkildə həll edir və daha çox məlumat daxil edilənə qədər qərarı saxlayır.

Hazırda təbii dilin emalı sahəsində ənənəvi qaydalara əsaslanan yanaşmalardan uzaqlaşılıb və dərin öyrənməyə (deep learning) üstünlük verilir. Bu dəyişikliyin səbəbi, dərin öyrənmənin insan tərəfindən əvvəlcədən müəyyən edilmiş qaydalarla və sabit meyarlarla həll olunması çətin olan məsələləri daha uğurla yerinə yetirməsidir. Dilin təbii xüsusiyyəti ondan ibarətdir ki, bir söz müxtəlif mənə daşıya bilər və bu mənalar onun ətrafındakı sözlərdən asılı

olaraq dəyişə bilər. Bu cür qeyri-müəyyənlikləri ənənəvi qaydalarla və ya qərar ağacı kimi üsullarla əhatə etmək mümkün deyil. Dərin öyrənmə isə bu problemi effektiv şəkildə həll edir, çünki o, proqramçının ayrıca qaydalar yazmasına ehtiyac duymur – alqoritm özü verilmiş girişə əsaslanaraq çıxışı öyrənir və qərar vermə prosesini avtomatik qurur.

XXI əsrə keçid ilə təbii dilin emalı sahəsində klassik nəzəri yanaşmalardan uzaqlaşaraq, daha çox statistik və neyron modellərə əsaslanan müasir yanaşmalar ön plana çıxmağa başladı. Bu dəyişikliklər növbəti mərhələdə yeni texnologiyaların formalaşmasına səbəb oldu.

Təbii dilin emalının müasir tətbiq sahələri və yanaşmaları

Təbii Dilin Emalı (NLP) sahəsi süni intellektin ən sürətlə inkişaf edən istiqamətlərindən biridir və son illərdə onun tətbiq imkanları əhəmiyyətli dərəcədə artmışdır. Xüsusilə dərin öyrənmə texnologiyalarının geniş yayılması NLP-nin müxtəlif sahələrdə praktiki istifadəsini mümkün etmişdir. Məsələn, səsli köməkçilər – Siri, Alexa, Google Assistant kimi – artıq gündəlik həyatımızın bir parçasına çevrilib. Onlar evdə, iş yerində, hətta avtomobillərdə insanların suallarını cavablandırır, tapşırıqları yerinə yetirir və beləliklə, gündəlik fəaliyyətləri asanlaşdırırlar.

Digər tərəfdən, sosial media platformalarında hər gün minlərlə fikir və rəy paylaşılır. Bu böyük məlumat kütləsini əl ilə analiz etmək praktiki deyil və vaxt tələb edir. NLP burada bizneslər üçün qiymətli bir vasitə olaraq ortaya çıxır. Sentiment analizinin köməyi ilə şirkətlər müştəri rəylərini müsbət, mənfi və neytral kateqoriyalara ayıraraq, bazar haqqında ümumi mənzərə əldə edir, məhsullarını və xidmətlərini təkmilləşdirirlər. Bu proses yalnız fərdi rəylərin yox, ümumi fikirlərin kəmiyyətə ölçülməsini də təmin edir ki, bu da biznes strategiyasının düzgün qurulmasında əvəzsizdir.

NLP-nin digər mühüm tətbiq sahələrindən biri isə spam filtrləmədir. İnternetin genişlənməsi ilə e-poçt, sosial media və müxtəlif onlayn platformalarda çoxlu saxta məlumat və spam yaranır. Bu məlumatların seçilməsi və düzgün filtrasiya edilməsi üçün maşın öyrənməsi metodları – qərar ağacları, Naiv Bayes, SVM və digər alqoritmlərdən istifadə olunur. Məsələn, YouTube-un şərhlərində TubeSpam kimi vasitələr şərhlərin saxta və ya real olduğunu ayırd edir. Veb səhifələrdə spamın qarşısını almaq üçün daha mürəkkəb metodlar tətbiq edilir ki, bu da internet mühitinin təhlükəsizliyini artırır (Alberto et al., 2015). Tibbi sahədə də NLP-nin rolu böyükdür. Xəstəliklərin diaqnozu və müalicəsi zamanı yaranan tibbi məlumatlar çox zaman struktursuz olur və onların əl ilə təhlili həm çətin, həm də vaxt aparan prosesdir. NLP və maşın

öyrənməsi texnologiyaları tibbi mətnləri avtomatik analiz edərək, diaqnostik qərarların verilməsini sürətləndirir və dəqiqləşdirir. Həmçinin, xəstələrin fərdi davranışlarını və müalicə reaksiyalarını qiymətləndirmək üçün aspekt əsaslı sentiment analizi kimi metodlardan istifadə olunur ki, bu da şəxsiləşdirilmiş terapiyanın inkişafına yardım edir (Chen, 2019). Bundan başqa, NLP insanların yazılı ünsiyyətini də asanlaşdırır. Məsələn, Grammarly kimi proqramlar yazılan mətnin qrammatik səhvlərini aşkar edir, mətnin ümumi mövzusunı anlayır və daha yaxşı ifadə yolları təklif edir. Hüquq sahəsində isə onlayn müqavilələrdə şəxsi məlumatların qorunması məqsədilə Named Entity Recognition (NER) texnologiyası tətbiq edilir. Bu alətlər fərdi identifikasiyaya yönəlik məlumatları avtomatik tanıyır və qoruyur (Hiremath & Patil, 2020).

Axtarış sistemləri, o cümlədən Google və Bing, veb məzmununu müxtəlif dillərə tərcümə etmək üçün maşın tərcümə texnologiyalarından istifadə edir, bununla da istifadəçilərə müxtəlif dillərdə məlumat əldə etmə imkanı yaradırlar. Veb saytlar üçün chatbotlar isə istifadəçilərlə interaktiv əlaqə quraraq sorğulara tez cavab verir, xidmət keyfiyyətini artırır və insan əməyini azaldır.

Beləliklə, NLP sahəsindəki müasir yanaşmalar və tətbiqlər yalnız texnoloji tərəqqini deyil, həm də cəmiyyətin müxtəlif sahələrində fəaliyyətin effektivliyini artıran mühüm vasitə kimi özünü doğruldur. Bu sahənin gələcəyi daha da parlaqdır və yeni texnologiyalarla birlikdə daha geniş tətbiq sahələri meydana çıxacaq.

Azərbaycan dilinin Təbii Dilin Emalı (NLP) sahəsindəki mövcud vəziyyəti və inkişaf perspektivləri

Təbii dilin emalı (TDE / NLP) sahəsi son onilliklərdə süni intellekt texnologiyalarının sürətli inkişafı ilə paralel olaraq geniş vüsət almışdır. İngilis dili başda olmaqla bir çox dünya dilləri üçün yüksək keyfiyyətli dil modelləri, açıq mənbəli resurslar və tətbiq alətləri formalaşmışdır. Azərbaycan dili isə zəngin tarixə və ifadə imkanlarına malik olsa da, bu sahədə yetərinə təmsil olunmamışdır. Nəticədə həm elmi tədqiqat səviyyəsində, həm də praktiki tətbiqlərin hazırlanmasında müəyyən boşluqlar müşahidə olunmaqdadır. Bu vəziyyət Azərbaycan dilinin süni intellekt mühitində daha geniş istifadəsini və rəqəmsal transformasiyasını məhdudlaşdırır.

Azərbaycan dilinin riyazi-statistik metodlarla və texnoloji vasitələrlə öyrənilməsinə dair ilk təşəbbüslər ötən əsrin ortalarına, yəni 1960-70-ci illərə təsadüf edir. Bu dövrdə Sovet elmi mühitində formalaşmış və xaricdə təhsil almış azərbaycanlı mütəxəssislər — o cümlədən V.Y.

Pines, M.Ə. Mahmudov, K. Vəliyeva və C.Ə. Rəhmanov — dilin formal modelləşdirilməsi və avtomatlaşdırılmış təhlili istiqamətində ilkin elmi əsasları yaratmağa başlamışdılar. 1980-ci ildə M.Ə. Mahmudov və V.Y. Pines tərəfindən nəşr olunmuş “Maşın tərcüməsi və avtomatlaşdırılmış məlumat sistemlərində Azərbaycan mətnlərinin morfoloji təhlil alqoritmı” adlı əsər bu sahədə ilk nəzərəcərpacaq elmi-praktik töhfələrdən biri sayılır (Mahmudov & Pines, 1980).

Bu mərhələdə mühüm rol oynamış digər bir tədqiqat isə akademik A.A. Axundovun “Riyazi dilçilik” (1979) kitabıdır. Sözügedən əsər Azərbaycan dilinə riyazi yanaşma prinsiplərini təqdim etməklə yanaşı, saitlərin statistik təhlili, morfoloji modelləşdirmə üsulları və bu sahədə istifadə olunan əsas anlayışları sistemli və anlaşılan formada izah etmişdir (Axundov, 1979). Bununla belə, 1970–1980-ci illərdə bu sahəyə dövlət səviyyəsində yetərli dəstək verilmədiyindən, mövcud potensial tam şəkildə realizə olunmamış, tədqiqatların davamlılığı təmin edilməmişdir.

Müstəqillik dövründə, xüsusilə 1990-cı illərdən sonra, Azərbaycan dilinin informasiyalaşdırılması və rəqəmsallaşdırılması dövlət siyasətinin bir hissəsinə çevrildi. Bu dövrdə “Azərbaycan dilinin milli korpusu”, “Orfoqrafik və izahlı elektron lüğətlər”, “İmla yoxlayıcı”, “Nıtq tanıma sistemi” kimi layihələr reallaşdırıldı. Bu təşəbbüslər mühüm addımlar olsa da, müasir NLP meyarları baxımından hələ də ilkin mərhələ kimi qiymətləndirilə bilər.

Bugünkü vəziyyətə gəldikdə isə, Azərbaycan dilinin TDE sahəsindəki təmsilçiliyi hələ də məhdud səviyyədədir. Məsələn, dilimiz Universal Dependencies (UD) platformasında yer alsada, mövcud resurslar məhduddur və beynəlxalq NLP layihələrində istifadəyə geniş imkan yaratmır. Halbuki bu platforma spaCy, Stanza və digər məşhur NLP kitabxanaları üçün əsas sintaktik “treebank” bazası rolunu oynayır. Hazırda 200-dən çox dil UD-də təmsil olunur və Azərbaycan dili üçün bu sahədə daha geniş resursların hazırlanması zəruri olaraq qalır.

Bundan əlavə, GPT, BERT, T5 kimi qabaqcıl transformator əsaslı modellərin Azərbaycan dili üçün effektiv şəkildə tətbiqi və ya təlimi üçün zəruri olan genişhəcmli və keyfiyyətli korpuslar ya ümumiyyətlə mövcud deyil, ya da ictimaiyyət üçün əlçatan deyil. Mövcud təşəbbüslər — məsələn, AzerBERT və oxşar modellər — əsasən fərdi və qeyri-kommersiya səyləri çərçivəsində inkişaf etdirilib. Bu sahədə resursların artırılması və koordinasiyanın gücləndirilməsi gələcəkdə elmi və texnoloji tərəqqinin daha da sürətlənməsinə imkan yaradacaqdır.

Problemin aradan qaldırılması üçün elmi-tədqiqat institutlarının və dövlət qurumlarının bu sahədə fəal rol alması vacibdir. Xüsusilə, korpusların hazırlanması, açıq mənbəli verilənlərin paylaşılması (məsələn, CC BY-NC-ND 4.0 lisenziyası əsasında), terminoloji verilənlər bazalarının yaradılması və yeni model təlimləri üçün beynəlxalq platformalarla integrasiya əsas hədəf olmalıdır. Hugging Face, GitHub kimi platformalarda Azərbaycan dilinə dair resursların yerləşdirilməsi və beynəlxalq əməkdaşlıqların qurulması bu istiqamətdə mühüm rol oynaya bilər.

Beləliklə, Azərbaycan dilinin NLP sahəsindəki inkişafı yalnız texnoloji deyil, həm də milli və strateji əhəmiyyət daşıyır. Ana dilimizin rəqəmsal transformasiya dövründə layiqli şəkildə təmsil olunması üçün həm elmi, həm texnoloji, həm də institusional səviyyədə ardıcıl və məqsədyönlü səylər göstərməlidir.

Nəticə

Süni intellekt və təbii dilin emalı sahəsi elmi biliklərin, texnoloji tərəqqinin və interdisiplinar yanaşmaların kəsişməsində formalaşaraq dinamik inkişaf yolu keçmişdir. Tarixi təməllərini məntiq, dil fəlsəfəsi və formal sistemlərdən alan bu sahə zamanla statistik metodların, neyron şəbəkələrin və dərin öyrənmə alqoritmlərinin integrasiyası ilə yeni mərhələyə qədəm qoymuşdur. Klassik dillər nəzəriyyəsindən tutmuş, çağdaş süni intellekt sistemlərinə qədər olan inkişaf trayektoriyası, həm nəzəri, həm də praktiki səviyyədə insan dilinin avtomatlaşdırılmış təhlilinə və emalına dair mühüm nəticələrin əldə edilməsinə şərait yaratmışdır.

Hazırda təbii dilin emalı yalnız akademik tədqiqatların predmeti olmaqla qalmır, eyni zamanda sosial, iqtisadi və hüquqi sistemlərin strukturuna integrasiya olunan, real problemlərin həllinə yönələn texnoloji həllər toplusudur. Səsli köməkçilər, avtomatik tərcümə sistemləri, hüquqi sənədlərin analizi və tibbi diaqnostika sistemlərində NLP texnologiyalarının tətbiqi bu sahənin çoxşaxəli potensialını nümayiş etdirir.

Nəticə etibarilə, süni intellekt və təbii dilin emalı sahəsindəki inkişafın linqvistik modellərlə yanaşı, mühəndislik və hesablama elmlərinin qarşılıqlı sintezini zəruri edir. Gələcəkdə bu sahənin inkişaf istiqamətləri etik tənzimləmə, çoxdilli sistemlərin təkmilləşdirilməsi və insan-kompüter qarşılıqlı əlaqəsinin daha təbii səviyyəyə çatdırılması ilə səciyyələnməkdir. Beləliklə, bu sahə yalnız texnologiyanın deyil, həm də insan düşüncəsinin və dilin mahiyyətinin daha dərinə dərk edilməsinə yönələn fundamental bir tədqiqat istiqaməti olaraq önəmini qoruyub saxlayacaqdır. Bu kontekstdə Azərbaycan dilinin təbii dilin emalı sahəsindəki mövqeyi xüsusi diqqət tələb edir. Tarixi baxımdan zəngin elmi irsə malik olsa da,

dilimiz müasir NLP texnologiyalarına integrasiya baxımından hələ də əhəmiyyətli boşluqlarla üz-üzədir. Azərbaycan dilinin beynəlxalq platformalarda (məsələn, Universal Dependencies) təmsil olunmaması, açıq mənbəli resursların və müasir dil modellərinin çatışmazlığı onun rəqəmsal texnologiyalarda tətbiqini əngəlləyir. Bu problemlərin aradan qaldırılması üçün dövlət səviyyəsində təşviqlər, elmi-tədqiqat mərkəzlərinin dəstəklənməsi və dil resurslarının açıq şəkildə paylaşılması zəruridir. Yalnız bu yolla Azərbaycan dili süni intellekt əsrində layiqli şəkildə mövcudluğunu təmin edə bilər.

ƏDƏBİYYAT

1. Acar, E. (2020). Yapay Zekâya Giriş. İstanbul: Akademik Yayınlar.
2. Alberto TC, Lochter JV, Almeida TA (2015) TubeSpam: comment spam filtering on youtube. In: 2015 IEEE 14th International conference on machine learning and applications (ICMLA)
3. Axundov, A. A. (1979). Riyazi dilçilik: Xülasələr və təfriqələr [Dərs vəsaiti]. Bakı: ADU nəşriyyatı.
4. Chen, P.-H. (2019). Essential elements of natural language processing: What radiology should know. The Association of University Radiologists.
5. Dick, S. (2019). Artificial intelligence. Harvard Data Science Review, 1(1), 1-8. doi: 10.1162/99608f92.92fe150c
6. Gladwin, L. A. (1997). Alan Turing, Enigma, and the breaking of German machine ciphers in World War II. Prologue Magazine, Fall 1997, 203–217 <https://www.archives.gov/files/publications/prologue/1997/fall/turing.pdf>
7. Hiremath, B. N., & Patil, M. M. (2020). Enhancing optimized personalized therapy in clinical decision support system using natural language processing. Journal of King Saud University – Computer and Information Sciences, 34(6), 2840–2848. <https://doi.org/10.1016/j.jksuci.2020.03.006>
8. Hutchins, J. (1997a): 'From first conception to first demonstration: the nascent years of machine translation, 1947-1954. A chronology.' Machine Translation 12, 195-252.
9. Kırpınar, İ. (2019). Bilişsel psikoloji kognitif yetiler yapay zeka, epistemoloji, felsefe, dilbilim, matematik, nörobilimler ilişkisi. İstanbul: Psikonet Yayınları Uzman Serisi.
10. Kutlusoy, Z. (2019). Felsefe açısından yapay zeka. G. Telli (Ed.), Yapay zeka ve gelecek içinde (ss. 25-43). İstanbul: Doğu Kitapevi.
11. Mahmudov, M. Ə., & Pines, V. Y. (1980). Maşın tərcüməsi və avtomatlaşdırılmış məlumat sistemlərində Azərbaycan mətnlərinin morfoloji təhlil alqoritmi. İnformasiya vərəqi. Energetika və avtomatika seriyası, (2).
12. McCarthy, J. (2007). From here to human-level AI. Artificial Intelligence, 171(18), 1174-1182.
13. Öztürk, K. ve Şahin, M. E. (2018). Yapay sinir ağları ve yapay zekaya genel bir bakış. Takvim-i Vekayi, 6(2), 25-36.
14. Sarıel, S. (2017). Günümüzde yapay zeka. M. Karaca (Ed.) İnsanlaşan makineler ve yapay zekâ içinde (ss. 21-25). İstanbul: İstanbul Teknik Üniversitesi Vakfı Dergisi.

15. Schank, R. (1969). A conceptual dependency parser for natural language. In Proceedings of the Conference on Computational Linguistics, Sång-Säby, Sweden pp. 1–3.
16. Schultz, D. P. ve Ellen-Schultz, S. (2007). Modern psikoloji tarihi (Y. Aslay, çev.). İstanbul: Kaknüs Yayınları
17. Turing, A. M. (1950). Computing Machinery and Intelligence. *Mind*, 59(236), 433–460.
18. Winograd, T. (1980). What does it mean to understand language? *Cognitive Science*, 4(2), 209–241.
19. Revina A.N. (2025) Historical development and modern approaches in the field of artificial intelligence and natural language processing. *Universidad y Sociedad*, 17(S1). e5505