***UOT-811.11-112***
***Səhifə: 101-108***
*https://doi.org/10.59849/2663-8967.2026.2.101*

***Rəna Mahmudova,***
***Maya Heydərova***
***AMEA Nəsimi adına Dilçilik İnstitutu,***
***filologiya üzrə fəlsəfə doktoru***
*rena.memmedova.1991@inbox.ru*
*mayahaciyeva.mh@gamil.com*

# TƏBİİ DİL EMALINDA SENTİMENT ANALİZ

## XÜLASƏ

Təbii dil emalı kompüterlərin insan dilini anlamaq, şərh etmək və emal etmək qabiliyyətini inkişaf etdirən mühüm bir sahədir. Bu sahənin əsas istiqamətlərindən biri olan sentiment analiz mətnlərdə ifadə olunan emosional tonun müəyyənləşdirilməsinə yönəlmişdir. Sentiment analiz vasitəsilə mətnlər müsbət, mənfi və ya neytral kateqoriyalara bölünür.

Sentiment analizin əsas məqsədi insan fikirlərini avtomatik şəkildə anlamaqdır. Sosial şəbəkələrdə paylaşımlar, rəylər, müzakirələr və xəbərlər kimi böyük həcmli mətn məlumatlarının artması bu sahənin əhəmiyyətini artırdı. Müasir dövrdə şirkətlər müştəri məmnuniyyətini ölçmək, istifadəçi davranışlarını analiz etmək və bazar tendensiyalarını müəyyənləşdirmək üçün sentiment analizindən geniş istifadə edirlər. Bu sahədə istifadə olunan yanaşmalar əsasən iki qrupa bölünür: lüğət əsaslı və maşın öyrənməsi əsaslı metodlar. Lüğət əsaslı yanaşmada əvvəlcədən hazırlanmış emosional söz bazalarından istifadə olunur. Hər bir söz müəyyən emosional dəyərə malik olur və mətnin ümumi tonu bu dəyərlərin cəmi əsasında müəyyən edilir. Digər tərəfdən, maşın öyrənməsi yanaşmaları böyük həcmli etiketlənmiş məlumatlar üzərində öyrədilmiş modellərə əsaslanır. Bu modellər mətnlərdəki gizli nümunələri aşkar edərək daha dəqiq nəticələr əldə etməyə imkan verir.

Son illərdə dərin öyrənmə metodlarının inkişafı sentiment analizin keyfiyyətini əhəmiyyətli dərəcədə artırmışdır. Neyron şəbəkələr, xüsusilə təkrarlanan neyron şəbəkələri, uzun, qısa müddətli yaddaş modelləri və transformerlər mətnin kontekstini daha yaxşı anlamağa kömək edir. Bu texnologiyalar sayəsində yalnız ayrı-ayrı sözlər deyil, bütöv cümlə və mətn səviyyəsində emosional məna daha dəqiq təhlil olunur.

Bununla belə, sentiment analizdə müəyyən çətinliklər də mövcuddur. Dilin mürəkkəbliyi, çoxmənalılıq, sarkazm və ironiyanın mövcudluğu analiz prosesini çətinləşdirir. Məsələn, eyni söz müxtəlif kontekstlərdə fərqli emosional məna ifadə edə bilər. Bundan əlavə, müxtəlif dillər və mədəniyyətlər arasında ifadə fərqləri də modellərin ümumiləşdirmə qabiliyyətinə təsir göstərir.

Nəticə etibarilə, sentiment analiz təbii dil emalının mühüm istiqamətlərindən biri olmaqla, həm nəzəri, həm də praktik sahələrdə geniş tətbiq olunur. Bu sahənin inkişafı ilə daha dəqiq və kontekstual analizlərin aparılması mümkün olacaq və bu da insan-kompüter qarşılıqlı əlaqəsinin daha səmərəli qurulmasına şərait yaradacaqdır.

***Açar sözlər*: *Süni intellekt, sentiment analiz, N-gram, NLP, dilçilik, böyük dil modelləri***

**Giriş:** Süni intellektin insan zəkası ilə yarışdığı bir dövrə qədəm qoyuruq. Çağdaş dövrümüzdə süni intellektlə yaradılan proqramlar artıq insanların bir nömrəli köməkçisinə çevrilib. Müasir dövrümüzdə insanların əksəriyyəti internet və sosial şəbəkələrdən istifadə edərək öz hisslərini, fikirlərini ifadə edir, təcrübələrini bölüşürlər. Bu, adətən internetdən istifadə edərək böyük miqdarda məlumatların ötürülməsinə səbəb olur. Lakin bu məlumatların əksəriyyəti təhlil edildikdə faydalıdır. Məsələn, əksər sənaye şirkətləri və seçki kampaniyaları kommunikasiya saytları vasitəsilə insanların fikirlərini bilməyə əsaslanır və onların müsbət, mənfi və ya neytral olub-olmadığını öyrənir. Sentiment analiz internetdə nəhəng məlumat mübadiləsi nəticəsində yaranmışdır.

"Sentiment analiz böyük verilənlərdə subyektiv məlumatların sistemli təhlilinə imkan verən təbii dilin emalı üsuludur" [1, s.128]. Sentiment analiz bir çox tədqiqatçılar üçün geniş sahədir və onu rəyin öyrənilməsi də adlandırmaq olar, çünki o, ideya və fikirləri müsbət, mənfi və ya neytral kimi təsnif etməyə kömək edir. Sentiment analizi internetdə və sosial mediada rəy və sorğularda geniş istifadə olunan mətn tədqiqatıdır. O, kommersiya saytlarında onların məhsulun qəbul və ya rədd edilməsini bilmək üçün cavabları və müştəri rəylərini idarə edir; bu, müştərinin seçimini izah etdiyi üçün şirkətin satışlarını yaxşılaşdırmağa kömək edir. Sosial şəbəkələr vasitəsilə müxtəlif fikirlərin partlaması ilə sistemlər, siyasətçilər, psixoloqlar, istehsalçılar və tədqiqatçılar tərəfindən ən yaxşı qərarları həyata keçirmək üçün təhlil etmək üçün yeni ideyalar yarandı. Sentiment analizi mətn vahidində hiss məzmununu çıxarmaq və müəyyən etmək üçün NLP, statistika və maşın öyrənmə yanaşmalarından istifadə edərək yüksək effektivliyə malikdir.

"Sentiment analiz və ya rəylərin intellektual analizi məhsullar, xidmətlər, təşkilatlar, fərdlər, problemlər, hadisələr, mövzular və onların xüsusiyyətləri kimi obyektləri araşdıran insanların fikirlərini, hisslərini, duyğularını, qiymətləndirmələrini və münasibətlərini təhlil edən tədqiqat sahəsidir" [2, s.112]. Tərifinə görə sentiment analiz təbii dilin emalı sahəsi olsa da, insanların ruh vəziyyətini bir sıra sahələrə aid qiymətləndirmək, təhlil etmək və aşkarlamaqla mübarizə aparan mətnlərin təhlilinə diqqət yetirir, tədqiqatların əksəriyyəti bunu rəyin öyrənilməsi ilə məhdudlaşdırır. Bununla belə, rəyin öyrənilməsi, fikir analizi, emosiyaların öyrənilməsi və qeyri-müəyyənliyin aşkarlanması olan sentiment analizi çətiri altındakı digər üç sahənin yalnız bir alt sahəsidir. Qeyd edək ki, qeyri-müəyyənliyin aşkarlanması, statistik və yaxud sintaktik-semantik təhlil səviyyələrinin insan dilini aydınlaşdırmaq üçün qənaətbəxş səviyyəyə çatmaq üçün adekvat olmadığını hesab edən linqvistik təbiəti sayəsində digər iki alt sahənin birləşməsi hesab olunur. Bundan sonra, hazırkı məqalə qeyri-müəyyənliyi aradan qaldırmaq və mətnlər və sosial media yazıları üzərində nifrət nitqinin, istehzanın və sarkazmın aşkarlanması və qeyri-müəyyənləşdirilməsi alt vəzifələrində yanlış mühakimələrin qarşısını almaq üçün təhlilin praqmatik və sosial-praqmatik səviyyələrinə çatmaq üçün dərindən araşdırma aparmağı təklif edir.

Sentiment analizdə və ya rəy hasilatında (opinion mining) məqsəd insanların yazılı dildə (mətn) ifadə olunan fikirlərini kəşf etməkdir. Hiss termini "bir şey haqqında hiss etdiyi şey", "şəxsi təcrübə, öz hissi", "nəyəsə münasibət" və ya "fikir" deməkdir. Rəylər demək olar ki, bütün insan fəaliyyətlərinin mərkəzidir və bizim davranışlarımıza əsas təsir edən amillərdir. Bizim inanclarımız və reallıq qavrayışımız və etdiyimiz seçimlər əhəmiyyətli dərəcədə başqalarının dünyanı necə görüb qiymətləndirdiyinə bağlıdır. Bu səbəbdən qərar verməli olduğumuz zaman tez-tez başqalarının fikirlərini öyrənirik. Bu, təkcə fərdlər üçün deyil, həm də təşkilatlar üçün də doğrudur. Ümumi müştəri məmnuniyyətinin əhəmiyyətli komponentlərini və ya aspektlərini müəyyən etmək üçün ənənəvi olaraq, qapalı formada müştəri məmnuniyyəti anketlərindən istifadə ediləcəkdir. Bununla belə, sorğu vərəqələrinin hazırlanması və icrası bahalıdır və ya əlçatan olmaya bilər. Bəzi hallarda dövlət qurumlarına müştərilərdən məmnunluq anketləri toplamaq hətta qanunla qadağan edilir [3, s.2].

Sosial media və digər saytlarda artan fikirlərin daha sürətli öyrənilməsi üçün sentiment analizi çox vacib hala gəlir. Son illərdə rabitə, hava trafiki və alternativ bazarlar saytlarında böyük informasiya partlayışı, bütün bu nəhəng informasiyanın idarə edilməsi və təhlili ənənəvi üsullardan istifadə edilmir, ona görə də elm adamları və tədqiqatçılar bu məlumatların öhdəsindən gəlmək üçün yüksək effektiv üsullar işləyib hazırlamışlar.

Bing Liu sentiment analizinin 3 səviyyəsinin olduğunu bildirir: 1) Sənəd səviyyəsi, 2) cümlə səviyyəsi və 3) obyekt və aspekt səviyyəsi [4]. Sentiment analizinin sənəd səviyyəsində hər bir sənədin yalnız bir obyekt haqqında rəy bildirdiyini fərz edir. Sənəd səviyyəsi birdən çox obyekti qiymətləndirən və ya müqayisə edən sənədlərə tətbiq edilə bilməz. Cümlə səviyyəsi sentiment analizi cümlələrə fokuslanır və hər bir cümlənin müsbət, mənfi və ya neytral rəy ifadə edib-etmədiyini müəyyənləşdirir. Obyekt və aspekt səviyyəsi isə obyektin xüsusiyyətlərinə əsaslanır.

Məlumatların toplanması sentiment analizdə ilk addımdır. Bunlara istifadəçi qrupları, Twitter, Facebook, bloqlar və s. kimi mənbələrdən məlumatların toplanmasını aid etmək olar. Bu

məlumatlar skan, mətn təhlili və təsnifatdan istifadə edilən dil emalı kimi ənənəvi üsullarla təhlil edilə bilməz. Mətnin hazırlanması məlumatları təhlil etməzdən əvvəl araşdırır. Ünsiyyət saytlarındakı bəzi rəy və söhbətlərdə təhqiredici və nalayiq sözlər var, ona görə də onlar yoxlanılır və nəticə daha etibarlı təhlil üçün hazırlanır. Bu proses təhlillə əlaqəli olmayan məzmunu seçir və sonra onu silir. Prosesin məqsədi avtomatik təhlilə göndərilməmişdən əvvəl spam və uyğunsuz rəylərin silinməsidir. Bu halda, təhlildən əvvəl mətn hazırlamaq üçün istifadə olunan nitq hissəsi texnikasından istifadə edə bilərik.

Sentiment analizi yalnız kommersiya məhsullarının nəzərdən keçirilməsinə tətbiq edilmir; o, həmçinin bütün növ sosial kommunikasiya saytlarına və birja bazarlarına tətbiq edilə bilər. Üç mövzu əhval-ruhiyyənin təhlili emosiyaların aşkarlanması, qaynaqların qurulması və öyrənmənin ötürülməsi çətiri altında işləyir. Duyğuların aşkarlanması Sentiment analiz ilə yaxından əlaqəli olan yeni bir araşdırma sahəsidir. Sentiment analizin məqsədi mətndən müsbət, mənfi və ya neytral hissləri aşkar etməkdir, emosiya aşkarlanması isə qəzəb, ikrah, qorxu, xoşbəxtlik, kədər və təəccüb kimi mətnlərin ifadəsi vasitəsilə hiss növlərini aşkar etmək və tanımaq məqsədi daşıyır.

Resursların yaradılması bir leksikondur, müsbət, mənfi və ya neytral qütblərə görə rəy bildirmək üçün istifadə olunan lüğətdir. Öyrənmənin ötürülməsi biliyin öyrənilmiş bir tapşırıqdan maşında öyrənmənin yeni tapşırığına köçürülməsi kimi qəbul edilir. Aşağıdakı meyarlara uyğun olaraq mətn təsnifatı aşağıdakı kimidir: Birinci standart hisslərin (müsbət, mənfi, neytral) polaritesidir. İkinci standart, aşağıdakı kimi xəstəlik məlumatlarını idarə etmək üçün əksər siyasi məqalələrə və tibb müəssisələrinə tətbiq olunan nəticənin polaritesidir.

İnternet düzgün mühitdir və istifadəçilər tərəfindən paylaşılan əksər məlumat və ideyalar üçün əsas mənbədir. Bu, sentimentin formasiya baxımından ehtiyatlı bir yerdir. Çoxlu fikir və məqalələri izləyərək, insanların öz məzmunlarını forumlar, mikrobloqlar və ya onlayn sosial şəbəkə saytları kimi müxtəlif sosial media vasitəsilə dərc etməyə üstünlük verildiyi müəyyənləşir. Məlumat mənbəyinin seçilməsi sentimental analizin ilk addımıdır. Bütün kommunikasiya saytları məlumatların toplanması üçün münbit mühitdir; məlumatların əksəriyyəti faydalıdır və çox vaxt faydalı olmayan sui-istifadə xarakterli məlumatlar var, bu məlumatlar təhlil zamanı avtomatik olaraq xaric edilir.

Təbii dil emalı (Natural Language Processing): Yazılı və ya şifahi olaraq emal etmək və tətbiq etmək üçün bütün insan dilləri ilə məşğul olur. NLP-nin əsas təklifi budur. NLP mətnin bir hissəsinə aiddir; mətnin bir hissəsinə fel, sifət və isim daxildir. Bu halda biz qeyri-mütəşəkkil məlumat əldə edə bilərik və alqoritmlərdən birini (nitq hissəsi və ya N-qram) istifadə edərək əlavə emal tələb edə bilərik. Nitqin bir hissəsi cümlələri kiçik sözlərə bölə bilər və hər sözün bir mənası var; məsələn, ingilis dilində hər bir kiçik söz fərqli mənaya malik ola bilər və bu, istifadə və funksiyalarına görə gəlir və isim, əvəzlik, feil, zərf, sifət, bağlayıcı, ön söz və kəsilmə kimi bir neçə növə və ya kiçik nitq hissələrinə təsnif edilir.

N-qram sadəcə işarələrin ardıcıllığıdır. Hesablama dilçiliyi kontekstində bu əlamətlər adətən sözlər olur, baxmayaraq ki, onlar simvollar və ya simvolların alt dəstləri ola bilərlər. N sadəcə işarələrin sayına aiddir. N qram söz ardıcıllığının özü və ya ona ehtimal təyin edən proqnozlaşdırıcı model üçün istifadə olunur. Qram hərflərin birləşməsidir; n qram cümləni bir neçə hissəyə bölməyə işarə edir (sözü saymaq); məsələn, "Cümə günü bayram olacaq"—bu 4 qramdan ibarətdir; "sabah bayramdır" - bu 2 qramdan ibarətdir.

Kommunikasiya saytlarında böyük həcmdə məlumatlar saxlanılır. Bu məlumatlar hər il sürətlə artır. İnternetdə saxlanılan məlumatların əksəriyyəti son illərdə istehsal edilmişdir. Lakin istehsal edilən məlumatların əksəriyyəti emal edildikdə faydalı ola bilər. Biznes, siyasi və sosial saytlar məlumat artımının əsas mənbəyi və ən böyük datainkubatordur.

Bu sahədə tədqiqatçı müəyyən etmək çətin olan qeyri-müəyyən cümlələr və ya sözlər şəklində gələn bir sıra məhdudiyyətlər və problemlərlə üzləşir. Bu məhdudiyyətlər hədəflənmiş məlumatları təhlil etmək üçün maneə yaradır və etibarsız nəticəyə gətirib çıxara bilər. Məlum anketlərdən, sadə anketlərdən və ya rol əsaslı anketlərdən birini istifadə edərək əhval-ruhiyyənin təhlili üçün çətinlik yaradan bu maneələrin müxtəlif növləri var.

Son zamanlar geniş yayılan informasiya qurğusu ChatGPT proqramında, eləcə də VADER, Transformers, TextBlob və sairdə sentiment analiz aparmaq mümkündür.

Süni intellekt sahəsində böyük dil modellərinin (BDM) inkişafı böyük həcmli mətn məlumatlarından zəngin semantik təsvirlər çıxarmaq üçün dərin öyrənmə üsullarından istifadə etməklə sentiment analiz sahəsində əhəmiyyətli dəyişikliklərə səbəb olmuşdur. BDM-lərin əvvəlcədən öyrədilmiş, məsələn BERT, GPT, XLNet və s. kimi müxtəlif variantları sentiment analiz daxil olmaqla təbii dil emalı sahəsində bir çox tapşırıqları yerinə yetirərkən yüksək performans nümayiş etdirmişdir.

BDM-lər mətnlərlə işləməyi, başa düşməyi, tərcümə etməyi, suallara cavab verməyi və nitq yaratmağı bacaran süni intellek proqramlarıdır. Bu modellər böyük həcmdə etiketlənməmiş mətnlər üzərində nəzarətsiz (unsupervised) ilkin öyrənmə vasitəsilə mürəkkəb dil nümunələrini, kontekstual məlumatları və semantik əlaqələri öyrənərək yeni mətnlər yaradır, suallara cavab verir.

Bu sahədə sentiment analizin ilkin metodları qaydalara əsaslanan sistemlərə və ya daha sadə maşın öyrənməsi alqoritmlərinə əsaslanırdı və bu metodlar çox vaxt insan dilinin mürəkkəbliyini tam əhatə edə bilmirdi. Son illər dərin öyrənmə texnologiyalarının inkişafı bu sahəni əsaslı şəkildə dəyişdirmişdir. Bu modellər əvvəlcədən öyrədilmiş arxitekturalara əsaslanaraq sentiment analizi yüksək dəqiqliklə edir və hətta kinayə, jarqon və kontekstdən asılı incəliklər mövcud olduqda belə müsbət, mənfi və neytral tonları effektiv şəkildə ayırd edə bilir. Eyni zamanda təbii dil emalı çərçivəsində "self-attention" (özünə diqqət) mexanizmləri və transformer modelləri kimi qabaqcıl texnikalardan istifadə edərək sözlər və ifadələr arasındakı mürəkkəb əlaqələri aşkar edir. Bu isə daha dərin sentiment analizə imkan yaradır və böyük həcmli məlumatlardan emosiyaların və fikirlərin daha effektiv çıxarılmasını təmin edir.

Tədqiqatlarda sentiment analiz metodologiyalarını üç qrupa ayırırlar: lüğət (leksikon) əsaslı metodologiyalar, maşın öyrənməsi metodologiyaları və hibrid metodologiyalar [5]. Lüğət əsaslı texnika hiss və emosiyaları ifadə edən sözlərdən ibarət olan "emosiya lüğəti"ndən istifadə edir. Bu lüğət müsbət və ya mənfi emosiyaları təsvir etmək üçün tez-tez istifadə olunan ifadələri əhatə edir [5]. Bu zaman sentiment analizdə mətnin ümumi emosional dəyəri, həmin mətn daxilindəki sözlərin sentiment ballarının toplanması yolu ilə müəyyən edilir. Bu metodun əsas üstünlüyü onun sadə tətbiq olunması və maşın öyrənməsi yanaşmaları ilə müqayisədə daha az təlim məlumatı tələb etməsidir. Lakin bu üsul bəzi hallarda daha az dəqiq olur, çünki sözlərin kontekstini və cümlə daxilindəki qarşılıqlı əlaqələri nəzərə almır. Lüğət əsaslı metodlar əsasən sosial media monitorinqi və siyasi analiz kimi tətbiq sahələrində istifadə olunur.

Hazırda təbii dilin emalı daxilində sentiment analiz sahəsində geniş istifadə olunan maşın öyrənməsi yanaşmasıdır. Bu yanaşma da öz növbəsində nəzarətli, nəzarətsiz və dərin öyrənmə kateqoriyalarına bölünür. Birinci halda model etiketlənmiş məlumatlar üzərində öyrədilir, yəni giriş mətninə uyğun sentiment analiz (müsbət, mənfi və s.) edir. Bu zaman model müxtəlif sentiment kateqoriyalarına uyğun mətn nümunələrini öyrənir və sonradan yeni, etiketlənməmiş mətnləri bu biliklərə əsasən təsnif edə bilir.

Nəzarətsiz öyrənmə zamanı isə modelə etiketlənmiş məlumat təqdim olunmur və o, verilənlərdəki struktur və nümunələri özü aşkar etməlidir. Bu yanaşma oxşar mətnlərin və ya mövzuların qruplaşdırılması kimi tapşırıqlarda istifadə olunur. Nəzarətli və nəzarətsiz öyrənmə üsullarının geniş tətbiqi ilə yanaşı, son illərdə süni intellekt sahəsində ən mühüm inkişaf istiqamətlərindən biri dərin öyrənmə texnologiyalarının formalaşması olmuşdur. Bu yanaşma ənənəvi maşın öyrənməsi metodlarından fərqli olaraq verilənlərdən xüsusiyyətlərin avtomatik çıxarılmasını təmin edir.

Dərin öyrənmə zamanı çoxqatlı neyron şəbəkələrdən istifadə olunur və onlar verilənlərdə mürəkkəb əlaqə və nümunələri öyrənə bilir. Dərin öyrənmə modelləri sentiment analiz, təbii dilin emalı və nitqin tanınması kimi sahələrdə geniş tətbiq edilir [6].

Hibrid texnika isə maşın öyrənməsi və lüğət əsaslı metodları birləşdirərək sentiment analizin performansını artırır.

Bundan əlavə, sentiment analiz sənəd səviyyəli sentiment analizdən tutmuş incə aspekt əsaslı sentiment analizə (sənəd daxilindəki müəyyən aspektlərə və ya varlıqlara münasibətdə hissləri təhlil edən) qədər müxtəlif səviyyələrdə işləyir. Bu müxtəlif səviyyəli təhlillər, müxtəlif tapşırıqların həllində sentiment analizə uyğunlaşa bilməsi və tətbiq olunmasına kömək edir [7].

Əvvəlcədən öyrədilmiş BDM-lərdə sadaladığımız yanaşmaları tamamlayan- fine-tuning (yenidən tənzimləmə), transfer learning (transfer öyrənmə), domain adaptation (sahəyə uyğunlaşma), Zero-shot və Few-shot Learning (sıfır nümunəli və az nümunəli öyrənmə) kimi konkret sahə və dillərə uyğunlaşdırılmış daha çevik müxtəlif öyrənmə strategiyaları mövcuddur.

Fine-tuning – əvvəlcədən öyrədilmiş BDM-lərin parametrlərinin konkret sahəyə aid etiketlənmiş məlumatlar üzərində yenidən öyrədilməsi prosesidir. Bu, modeli konkret sentiment analiz tapşırığına uyğunlaşdırmağa imkan verir. Bunun üçün domenə aid işarələnmiş məlumatlar toplanır (müştəri rəyləri, sosial media postları, tvitlər və s.) və model bu məlumatlar əsasında incə tənzimləmə üsulu ilə öyrədilir. Bu yanaşma vasitəsilə kontekstual anlama və kinayənin aşkarlanması mümkün olur.

Transfer learning – bir tapşırıqdan öyrənilən biliklərin digər əlaqəli tapşırığa tətbiqidir. Böyük dil modelləri ilə sentiment analiz kontekstində transfer öyrənmə əvvəlcədən öyrədilmiş modellərin çoxlu etiketlənmiş məlumatlara malik mənbə tapşırıq üzərində öyrədilməsi edilməsi və əldə olunan dil təsvirlərin məhdud etiketlənmiş məlumata malik hədəf sentiment analiz tapşırığına ötürülməsini nəzərdə tutur.

Domain adaptation – mənbə və hədəf sahələr arasındakı fərqlərin azaldılması və model performansının artırılması təmin edir [8].

Zero-shot və Few-shot Learning – metodları əlavə etiketlənmiş məlumat tələb etmədən, yaxud az sayda nümunə ilə sentiment analiz aparmağa imkan verir. Bu yanaşmalar BDM-lərin öyrənilmiş ümumi biliklərini istifadə edərək, yeni tapşırıqlara tez uyğunlaşmasını təmin edir [9].

Bu yanaşmaların praktik nümunələri BERT, ChatGPT, Gemini, DeepSeek və s. kimi böyük dil modellərində aydın görünür.

ChatGPT OpenAI tərəfindən hazırlanmış ən çox istifadə olunan böyük dil modellərindən biridir. İnsan dilini anlamaqda və təbii dilə yaxın mətnlər yaratmaqda yüksək səviyyədə səlisliyə və dəqiqliyə malikdir. ChatGPT, insan dilinin mürəkkəbliyini və incəliklərini başa düşdüyü üçün sentiment analiz aparmaq üçün ideal vasitədir. İstifadəçi mətn daxil etdikdə, dilin semantik və kontekstual xüsusiyyətlərini mənimsəyir, sorğuya əsasən mətnin emosional tonunu təhlil edir. ChatGPT ilə birlikdə sentiment analiz şirkətlərə və müştərilərə bir çox üstünlük təqdim edir. Bu üstünlüklərdən bəzilərinə real vaxt rejimində rəy, avtomatlaşdırılmış təhlil, müştəriyə yönəlmə, xərc-səmərəlilik və rəqabət üstünlüyü daxildir.

2018-ci ildə istifadəyə verilmiş BERT modeli sentiment analizi sahəsində mətnlərin emosional tonunu müəyyən etmək imkanı olan güclü dil modelidir. O, böyük həcmli mətn korpusları üzərində əvvəlcədən öyrədilir və daha sonra fine-tuning yanaşmasından istifadə edərək sentiment analizi həyata keçirir.

Modelin əsas üstünlüyü konteksti iki istiqamətli şəkildə öyrənməsi və attention mexanizmi vasitəsilə sözlər arasındakı əlaqələri daha dəqiq qavramasıdır. Bu səbəbdən sentiment analizi zamanı cümlənin ümumi mənasını daha düzgün başa düşülür.

Gemini Google şirkətinin DeepMind bölməsi tərəfindən hazırlanmış ən son nəsil multimodal böyük dil modelləri ailəsidir. O, transformer əsaslı arxitekturaya malik olub kontekstual məlumatı yüksək dəqiqliklə analiz edə bilir və həm zero-shot, həm də few-shot yanaşmaları ilə əlavə xüsusi təlim olmadan sentiment təsnifatı aparmağa qadirdir. Geminin multimodal olması, yəni yalnız mətn deyil, eyni zamanda şəkil və digər məlumat növlərini də analiz edə bilməsi sentiment analizini daha geniş kontekst daxilində aparmağa imkan verir.

Bundan əlavə, sosial şəbəkə dilinin qeyri-müəyyənliyini həll etmək və insan ünsiyyətini tam başa düşmək üçün daha dərin linqvistik səviyyədən baxır. Müvafiq olaraq, o, sentiment təhlili sahəsinin tədqiqatlarını təhlilin semantik və sintaktik səviyyələrindən kənara çıxmağa və sentiment təhlilində qeyri-hərfi dilin anlaşılmazlığını həll etmək üçün sosial şəbəkələr dilinin linqvistik təhlilinin praqmatik və sosial-praqmatik səviyyələri olan daha yüksək təhlil

səviyyələrinə daha dərindən nəzər salmağa dəvət edir. Bundan əlavə, o, sosial şəbəkələrdəki yazıların yanlış qiymətləndirilməsinin qarşısını almaq üçün nitq aktları nəzəriyyəsi və nəzakətdən istifadə edərək nifrət nitqinin avtomatik aşkarlanması üçün təhlilin sosial-praqmatik modelini təklif edir [10, s.49].

**Nəticə:** Bununla belə, bu alətlər əsasən müəyyən bir dildə yazılmış mətn məlumat dəstlərinə tətbiq olunur. Bundan əlavə, sentiment analizi modellərinin böyük əksəriyyəti ingilis dilində işləyir. İnternetdə ən çox istifadə olunan dil ingilis dili olsa da, müxtəlif digər dillərdə yazılmış çoxlu mətnlər var. Dildən asılı olmayan çox yönlü əhval-ruhiyyə təhlili aləti müəyyən bir coğrafi bölgə və ya bir qrup insanla məhdudlaşmayan geniş tətbiqlər sahəsi açacaqdır. O, həmçinin, mahiyyətcə bütün dünyaların ən yaxşılarından istifadə edəcək, xüsusən də ingilis dili üçün təbii dil emalı və mətn analitikası icması tərəfindən hazırlanmış alətlərin geniş çeşidi digər dillərə də genişləndirilə bilər.

## Ədəbiyyat siyahısı


1. İskəndərli, G. (2023). *Sentiment analiz texnologiyaları və Azərbaycan dili*. "Azərbaycan dilinin İKT problemləri, İKT-nin Azərbaycan dili problemləri" respublika elmi-praktiki konfransı. s. 128-133.
2. Hacırəhimova, M.Ş., İsmayılova, M.İ. (2020). *Sentiment analiz: problemləri və həlli yolları.* "İnformasiya texnologiyaları problemləri". №2, s. 112–124.
3. Farhadloo, M., Rolland, E. (2016). *Fundamentals of Sentiment Analysis and Its Applications.* Sentiment Analysis and Ontology Engineering. Pp.1-24.
4. Bing Liu. Sentiment Analysis and Opinion Mining, Morgan & Claypool Publishers, May 2012, 168 p.
5. Sharma, H. D., & Goyal, P. (2023). An Analysis of Sentiment: Methods, Applications, and Challenges. Engineering Proceedings, 59(1), 68. https://doi.org/10.3390/ engproc 2023059068
6. Gunasekaran, K.P. (2023). Exploring Sentiment Analysis Techniques in Natural Language Processing: A Comprehensive Review. ArXiv, abs/2305.14842
7. Buscemi, A., & Proverbio, D. (2024). ChatGPT vs Gemini vs LLaMA on Multilingual Sentiment Analysis. ArXiv, abs/2402.01715
8. Upadhye, A. (2024). Sentiment analysis using large language models: Methodologies, applications, and challenges. International Journal of Computer Applications, 186(20), 30–34. https://doi.org/10.5120/ijca2024923625
9. Brown, T., Mann, B., Ryder, N., Subbiah, M., Kaplan, J., Dhariwal, P., ... & Amodei, D. (2020). Language models are few-shot learners. Advances in Neural Information Processing Systems, 33, 1877–1901
10. Assem, S., Alansary, S. (2022). *The Development of Sentiment Analysis from a Linguistic Perspective.* The Egyptian Journal of Language Engineering. 9(2). Pp. 40-52.


***Rana Mahmudova,***
***Maya Heydarova***

***Sentiment analysis in natural language processing***

***Summary***

Natural Language Processing (NLP) is a crucial field that develops the capability of computers to understand, interpret, and process human language. Sentiment analysis, one of the primary branches of this field, focuses on determining the emotional tone expressed in texts. Through sentiment analysis, texts are classified into positive, negative, or neutral categories.

The main goal of sentiment analysis is to automatically understand human opinions. The rise of large-scale text data, such as social media posts, reviews, discussions, and news, has increased the significance of this field. Today, companies widely utilize sentiment analysis to measure customer satisfaction, analyze user behavior, and identify market trends.

Approaches used in this area are mainly divided into two groups: lexicon-based and machine learning-based methods. The lexicon-based approach utilizes pre-defined emotional word databases. Each word has a specific emotional value, and the overall tone of the text is determined based on the sum of these values. On the other hand, machine learning approaches rely on models trained on large volumes of labeled data. These models enable more accurate results by uncovering hidden patterns in texts.

In recent years, the advancement of deep learning methods has significantly improved the quality of sentiment analysis. Neural networks, particularly Recurrent Neural Networks (RNNs), Long Short-Term Memory (LSTM) models, and transformers, help better understand the context of the text. Thanks to these technologies, emotional meaning is analyzed more precisely not just at the level of individual words, but across entire sentences and texts.

However, certain challenges still exist in sentiment analysis. The complexity of language, ambiguity, and the presence of sarcasm and irony complicate the analysis process. For instance, the same word can carry different emotional meanings in different contexts. Additionally, variations in expression across different languages and cultures affect the generalization capability of the models.

In conclusion, as one of the essential directions of natural language processing, sentiment analysis is widely applied in both theoretical and practical fields. With the development of this area, more accurate and contextual analyses will become possible, creating opportunities for more efficient human-computer interaction.

***Key words:*** *Artificial intelligence, sentiment analysis, N-gram, NLP, linguistics, large language models*

***Рана Махмудова,***
***Мая Гейдарова***

***Сентимент-анализ в обработке естественного языка***

***Резюме***

Обработка естественного языка (NLP) — это важнейшая область, развивающая способность компьютеров понимать, интерпретировать и обрабатывать человеческий язык. Сентимент-анализ (анализ тональности текста), являющийся одним из ключевых направлений этой области, направлен на определение эмоционального тона, выраженного в текстах. С помощью сентимент-анализа тексты делятся на положительные, отрицательные или нейтральные категории.

Основная цель сентимент-анализа — автоматическое понимание мнений людей. Рост объемов больших текстовых данных, таких как публикации в социальных сетях, отзывы, обсуждения и новости, повысил значимость этой сферы. В современную эпоху компании широко используют анализ тональности для оценки удовлетворенности клиентов, анализа поведения пользователей и определения рыночных тенденций.

Подходы, используемые в этой области, в основном делятся на две группы: методы на основе словарей и методы на основе машинного обучения. В подходе на основе словарей используются заранее подготовленные базы эмоционально окрашенных слов. Каждое слово имеет определенное эмоциональное значение, и общий тон текста определяется на основе суммы этих значений. С другой стороны, подходы машинного обучения опираются на модели, обученные на больших объемах размеченных данных. Эти модели позволяют получать более точные результаты, выявляя скрытые закономерности в текстах.

В последние годы развитие методов глубокого обучения значительно повысило качество сентимент-анализа. Нейронные сети, в частности рекуррентные нейронные сети (RNN), модели долговременной краткосрочной памяти (LSTM) и трансформеры, помогают лучше понимать контекст текста. Благодаря этим технологиям эмоциональный смысл анализируется более точно не только на уровне отдельных слов, но и на уровне целых предложений и текстов.

Тем не менее, в сентимент-анализе существуют определенные трудности. Сложность языка, многозначность, наличие сарказма и иронии усложняют процесс анализа. Например, одно и то же слово в разных контекстах может выражать разное эмоциональное значение. Кроме того, различия в выражении мыслей между разными языками и культурами также влияют на способность моделей к обобщению.

В заключение, будучи одним из важных направлений обработки естественного языка, сентимент-анализ широко применяется как в теоретических, так и в практических областях. С развитием этой сферы станет возможным проведение более точного и контекстуального анализа, что создаст условия для более эффективного взаимодействия между человеком и компьютером.

***Ключевые слова:*** *Искусственный интеллект, анализ настроений, N-грамма, НЛП, лингвистика, большие языковые модели*

***Çapa tövsiyə edən:*** ***AMEA, Nəsimi adına Dilçilik İnstitutu***

***Rəyçilər:*** ***professor M.Mahmudov***
***dosent K.Həbibiova***