İnsan Sağlamlığına Yönəlmiş Kibertəhlükələr və Onların Rəqəmsal Ekspertizası

Sənan Vəlizadə
ETN İnformasiya Texnologiyaları İnstitutu, Bakı, Azərbaycan
senanvalizade@gmail.com

***Xülasə*—Rəqəmsal mühitdə insan sağlamlığına yönəlmiş təhlükələrin, xüsusilə intihar həddinə çatdırma, intihara meyillilik, kiberbullinq və emosional təzyiqlərin aşkarlanması və təhlili üçün çoxkomponentli yanaşma nəzərdən keçirilmişdir. Bu yanaşma çərçivəsində təbii dilin emalı əsaslı sentiment analizi, sosial şəbəkələrin qraf modelləşdirilməsi və loq fayl təhlili vasitəsilə istifadəçi davranışlarının analizi məsələlərinə baxılmışdır. Sözü gedən modellər vasitəsilə yüksək dəqiqliklə riskli məzmunların aşkarlanması və Rəqəmsal Risk İndeksinin hesablaması ilə fərdi psixoloji vəziyyətləri proqnozlaşdırmağın mümkün olduğu müəyyənləşdirilmişdir. Sistem həm profilaktik müdaxilə, həm də hüquqi sənədləşdirmə üçün effektiv alət təqdim edir.**

Açar sözlər— rəqəmsal ekspertiza; kibertəhlükələr; sentiment analiz; sosial şəbəkə analizi; onlayn psixoloji zorakılıq

I. GİRİŞ

İnformasiya-kommunikasiya texnologiyalarının (İKT) sürətli inkişafı və sosial media istifadəsinin geniş yayılması rəqəmsal həyatın insan sağlamlığına və psixikasına olan təsirlərini daha da aktuallaşdırmışdır. Xüsusilə yeniyetmələr onlayn mühitdə kiberbullinq, intihara təşviq, psixoloji manipulyasiya və internet asılılığı kimi kibertəhlükələrə məruz qalırlar [1, 2].

BMT və UNICEF-in hesabatlarına əsasən, 13–17 yaş arası yeniyetmələrin 50%-dən çoxu onlayn aqressiya və psixoloji təzyiqlə qarşılaşdıqlarını bildirirlər [3]. Bu kimi hallar bəzi hallarda depressiya, davranış pozuntuları və intihar cəhdləri ilə nəticələnir [4].

Ənənəvi psixoloji qiymətləndirmə metodlarının bu halları vaxtında aşkarlamaq üçün yetərsiz olduğu müasir dövrdə, yeni texnoloji yanaşmaların işlənməsi tələb olunur. Bu baxımdan, rəqəmsal ekspertiza metodları – o cümlədən loq faylların təhlili, təbii dilin email (TDE) əsaslı sentiment analizləri, sosial şəbəkələrin qraf modelləri və s. – psixoloji risk faktorlarını avtomatik müəyyənləşdirmək üçün perspektivli vasitələrdir [5].

II. ƏDƏBİYYATIN İCMALI

Onlayn psixoloji təhlükələr üzrə tədqiqatlar son illərdə nəzərəçarpacaq dərəcədə artmışdır. Hinduja və Patchinin [1] araşdırmalarında kiberbullinqin yeniyetmələrdə depressiya və intihar riski ilə əlaqəsi empirik sübutlarla ortaya qoyulmuşdur. Bu cür hallar, sosial rifaha və psixoloji təhlükəsizliyə ciddi təsirlər göstərir.

Livingstone və Stoilova [6] Avropanın 19 ölkəsində apardıqları araşdırmada yeniyetmələrin təxminən 27%-nin onlayn psixoloji zorakılığa məruz qaldığını bildirmişlər. Bu məlumatlar, rəqəmsal müdaxilə alətlərinə olan ehtiyacın artdığını göstərir.

Sosial qraf modellərinin tətbiqi də bu kontekstdə əhəmiyyətlidir. Barabasi [5] tərəfindən təklif edilən şəbəkə nəzəriyyəsi, onlayn sosial strukturların psixoloji təzyiqlərin yayılmasında necə rol oynaya biləcəyini izah edir.

Digər tərəfdən, təbii dilin emalı və sentiment analiz son dövrlərdə intihara təşviq və depressiv məzmunların avtomatik aşkarlanmasında geniş istifadə olunmağa başlanmışdır. Renjith və digərlərinin təqdim etdiyi dərin təlim yanaşması bu sahədə effektiv nəticələr təqdim edir. Bundan əlavə, süni intellekt sistemlərinin davranış analitikası ilə birləşdirilməsi psixoloji diaqnostika və erkən müdaxilə imkanlarını artırır

III. METODOLOGİYA

Tədqiqat çərçivəsində insan sağlamlığına yönəlmiş kibertəhlükələrin rəqəmsal ekspertiza metodları ilə təhlili məqsədilə çoxkomponentli metodoloji yanaşma təklif olunmuşdur və istifadə olunan texniki əsaslar dörd istiqamətdə sistemləşdirilmişdir: TDE əsaslı sentiment analiz, sosial şəbəkələrin qraf vasitəsilə modelləşdirilməsi, loq fayl davranış təhlili və rəqəmsal ekspertiza prinsipləri.

Təklif olunan model real istifadəçi məlumatları üzərində deyil, psixoloji risklər və onlayn davranışlarla bağlı mövcud elmi ədəbiyyat, nəzəri yanaşmalar və ictimai platforma müşahidələrindən çıxarılan simulyasiya məlumatları əsasında formalaşdırılmışdır. Məqsəd bu metodların real verilənli tədqiqatlara inteqrasiyası üçün konseptual və texniki yanaşma təqdim etməkdir.

A. Təbii dilin emalı və sentiment analizi

Bu tədqiqatın əsas məqsədlərindən biri, sosial media platformalarında yayımlanan və intihara meyllilik, kiberbullinq və emosional yüklü gərginlik ehtiva edən məzmunların avtomatik identifikasiyasıdır. Bu məqsədə çatmaq üçün Təbii Dil Emalı (TDE) texnologiyaları tətbiq olunmuş, xüsusi olaraq sentiment analizi üsullarından istifadə edilmişdir. Bu üsullar, istifadəçi tərəfindən yazılmış mətnlərdə emosional və semantik xüsusiyyətləri analiz edərək psixoloji risk səviyyəsini qiymətləndirməyə imkan verir.

Bunun üçün transformer əsaslı modellərdən, xüsusilə BERT kimi iki istiqamətli kontekstual təlim modellərindən

DOI: 10.25045/SPCDH4.0.2025.28

istifadə olunmuşdur. Bu texnologiyalar mətnin mənasını kontekstdə anlamaqla yüksək dəqiqliklə mənfi, neytral və yüksək riskli məzmunları təsnifləndirməyə imkan yaradır. Modellər həmçinin zamanla dəyişən istifadəçi davranışlarını təhlil edərək, psixoloji vəziyyətdə baş verən dəyişikliklərin proqnozlaşdırılmasını da dəstəkləyir.

Bu yanaşma tək-tək paylaşımlar üzərində deyil, ümumi davranış və emosional meyillər üzərində qurulan davamlı və adaptiv analiz sistemi formalaşdırmağa şərait yaradır. Paylaşımlar “neytral”, “mənfi”, “intihara meylli” və “kiberbullinq” olmaqla dörd sinifə ayrılmış və BERT, RoBERTa modelləri ilə analiz edilmişdir. Bu modellər əvvəlcədən təlim görmüş etiketli məlumatlar əsasında çoxsinifli təsnifat modeli kimi öyrədilmişdir. Hər bir paylaşım, mətnin kontekstual semantikasına əsasən uyğun risk sinifinə aid edilmişdir. Modelin təlimində aşağıdakı parametrlər tətbiq edilmişdir:

- Optimallaşdırma (Optimizer): AdamW
- İtki funksiyası (Loss Function): Categorical Cross-Entropy
- Dropout: 0,3
- Təlim/Test nisbəti: 80%/20%

Həmçinin siniflərə verilmiş etiketlər və onlara uyğun çəki qiymətləri aşağıdakı cədvəldəki kimidir:

CƏDVƏL I. SINIFLƏRƏ VERILMIŞ ETIKETLƏR VƏ ÇƏKI DƏYƏRLƏRI

Sinif	Etiket	Çəki
Neytral	0	0
Mənfi	1	0,4
İntihara meyilli	2	0,8
Kiberbullinq	3	1

Bu çəkilər əsasında hər bir istifadəçinin paylaşımlarının çəkili ortalaması alınaraq psixoloji risk indeksi hesablanmış və aşağıdakı aralığa əsasən kateqoriyalaşdırılmışdır:

- 0,00–0,40: Aşağı risk
- 0,41–0,70: Orta risk
- 0,71–1,00: Yüksək risk

Beləliklə, bu texnologiyalar psixoloji təhlükələrin aşkarlanması və qarşısının alınması üçün real vaxtlı və etibarlı bir vasitə olaraq istifadə oluna bilər [8]. Hər bir istifadəçi üzrə bu paylaşımların emosional yükləri nəzərə alınaraq 0 ilə 1 arasında dəyişən bir risk indeksi hesablanmışdır.

Müəyyən etik dəyərlərə əsasən 0.65 və yuxarı nəticələr “yüksək psixoloji risk” olaraq qəbul edilmişdir və bu istifadəçilər sistem tərəfindən erkən xəbərdarlıqla işarələnmişdir. Modelin effektivliyi Dəqiqlik = 0,87, Həssaslıq = 0,81, F1-ölçüsü = 0,84 metrikaları ilə qiymətləndirilmişdir. Hər istifadəçiyə görə ümumi emosional davranış indeksi çıxarılmış və müəyyən dəyərlərə əsasən “yüksək risk” statusu müəyyən olunmuşdur.

B. Qraf analizi

Sosial şəbəkə strukturlarının təhlili üçün qraf nəzəriyyəsinə əsaslanan modellər geniş istifadə olunur. Bu modellərdə şəbəkə G = (V, E) formasında təqdim edilir, burada V istifadəçiləri, E isə onların qarşılıqlı əlaqələrini (məsələn, şərh, cavab, paylaşım və s.) ifadə edir. Bu cür strukturlaşdırma sosial şəbəkədəki əlaqələrin sıxlığını və informasiya axınının dinamikasını qiymətləndirməyə imkan verir [9].

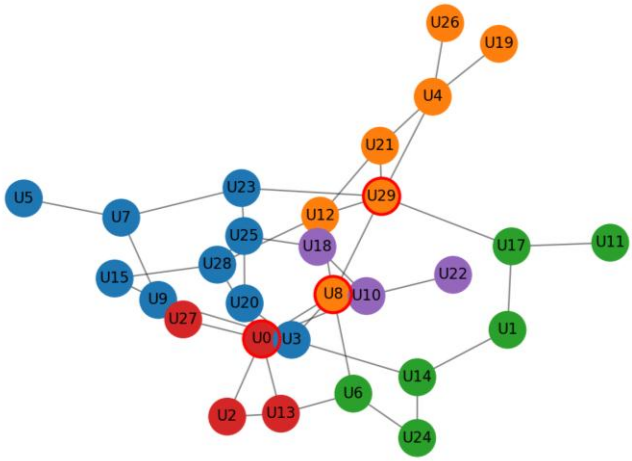
Qraf üzərində aparılan təhlillər aşağıdakı əsas mərkəzilik göstəricilərinə əsaslanırlar:

- Dərəcə mərkəziliyi (Degree Centrality) – istifadəçinin birbaşa əlaqələrinin ümumi sayı, onun şəbəkədəki aktivliyini göstərir.
- Yaxınlıq mərkəziliyi (Closeness Centrality) – bir qovşağın bütün digər qovşaqlara orta məsafəsinin göstəricisidir və məlumat axınına çatma sürətini əks etdirir.
- Vasitəçilik mərkəziliyi (Betweenness Centrality) – bir qovşağın digər istifadəçilər arasındakı ən qısa yol üzərində neçə dəfə yerləşdiyini ölçür və bu onun şəbəkədəki vasitəçi rolunu müəyyən edir.

Sosial şəbəkə daxilində icmaların aşkarlanması üçün müxtəlif metodlardan istifadə olunur. Bu məqsədlə sıx tətbiq olunan Louvain [10] və Girvan–Newman [11] alqoritmləridir. Louvain alqoritmi modulyarlıq göstəricisinə əsasən şəbəkəni yüksək daxili əlaqəyə malik klasterlərə bölür. Girvan–Newman alqoritmi isə “edge betweenness” prinsipinə əsasən əlaqələrin ayrılmasını modelləşdirir.

Mövcud tədqiqatlarda göstərilir ki, bəzi şəbəkə icmaları daxilində daha intensiv şəkildə yayılmış depressiv və aqressiv məzmunlar müşahidə olunmuşdur. Belə klasterlər psixoloji risk ehtiva edən istifadəçilərin müəyyən edilməsi baxımından əhəmiyyətlidir. Bu strukturlar üzərində zaman kəsiyində aparılan müşahidələr istifadəçi davranışında risk faktorlarını əvvəlcədən proqnozlaşdırmağa imkan verir.

Aşağıdakı şəkildə Qraf üzərində Louvain klasterləmə nəticəsində 5 əsas istifadəçi icması göstərilmişdir.



Şəkil 2. Sosial şəbəkə qraf strukturu – istifadəçilərarası əlaqələrin klasterləşməsi.

C. Loq fayl analizi və davranış profilləri

İstifadəçilərin informasiya məkanında fəaliyyətlərinə dair göstəricilər – o cümlədən orta sessiya müddəti, klik sıxlığı, istifadəçi sorğuları və emosional məzmunla qarşılıqlı əlaqə – sistem loqları vasitəsilə qeydə alınan əsas parametrlər kimi

qiymətləndirilir [12, 13]. Bu göstəricilərdən istifadə etməklə davranış dinamikasına dair analiz aparmaq və istifadəçinin psixoloji vəziyyətinə dolayı yanaşma tətbiq etmək mümkündür. Təhlil prosesində həm klassik statistik metodlara, həm də maşın təlimi əsaslı modellərə istinad edilir.

Əsas davranış göstəriciləri aşağıdakı kimi qruplaşdırılır:

- session_time_avg: İstifadəçinin gündəlik sessiyalar üzrə platformada keçirdiyi orta vaxt (dəqiqə ilə ölçülür). Bu göstərici, istifadəçinin ümumi aktivliyinin və bağlılıq səviyyəsinin göstəricisidir.
- emotional_content_rate: İstifadəçinin baxdığı məzmunlar arasında emosional (depressiv, aqressiv və s.) məzmunların nisbəti (% ilə). Bu parametr emosional yönümlü məzmunlara marağı əks etdirir.
- post_frequency: Gündəlik paylaşım sayı. Aktivlik səviyyəsinin ölçülməsi və motivasiya dərəcəsinin qiymətləndirilməsi üçün istifadə olunur.
- quiet_periods: Aktiv olmayan (offline) və ya paylaşım etmədiyi passivlik intervalları. Bu dövrlər psixoloji təcrid və ya dəyişiklik göstəricisi ola bilər.
- interaction_ratio: Digər istifadəçilərlə şərhlər, bəyənmələr və cavablar vasitəsilə qurulan qarşılıqlı əlaqənin nisbəti. Sosial dəstəyin və inteqrasiyanın göstəricisidir.

Bu göstəricilər ilkin mərhələdə "MinMaxScaler" alqoritmi vasitəsilə eyni intervala gətirilir (məsələn, 0 ilə 1 arasında) və bu da fərqli vahidlərdə olan göstəricilərin bir model daxilində müqayisəsinə imkan yaradır. Daha sonrisə logistic regression modeli tətbiq edilərək hər bir istifadəçi üçün Rəqəmsal Risk İndeksi (RRİ) hesablanır.

Modelin təlimi üçün simulyasiya yolu ilə yaradılmış 450 fərqli istifadəçi davranış profili əsasında şərti məlumat bazası formalaşdırılır. Bu model yeni istifadəçi üçün risk indeksini təxmin edərkən öyrəndiyi bu nümunələrdən istifadə edir.

Əldə edilən RRİ dəyərləri aşağıdakı risk kateqoriyalarına uyğunlaşdırılır:

CƏDVƏL II. RƏQƏMSAL RISK INDEKSI DƏYƏRLƏRI

Risk indeksi aralığı	Risk kateqoriyası
0 – 0,40	Aşağı risk
0,41 – 0,7	Orta risk
0,71 - 1	Yüksək risk

Əgər bir istifadəçinin RRİ dəyəri 0,71-dən yuxarı olarsa, bu şəxs sistem tərəfindən "yüksək risk" kateqoriyasına daxil edilir. Bu cür istifadəçilər üçün sistem avtomatik olaraq erkən xəbərdarlıq siqnalı yaradır. Nəticədə bu istifadəçilər real vaxtda izlənə bilər və ehtiyac olduqda psixoloji dəstək və ya hüquqi müdaxilə planlaşdırıla bilər.

Bu yanaşma, psixoloji risklərin aşkarlanmasında davranış təhlilinin necə effektiv istifadə oluna biləcəyini nümayiş etdirir və rəqəmsal mühitdə təhlükə siqnallarını vaxtında müəyyən etməyə imkan verir.

D. Rəqəmsal ekspertiza və kibertəhlükələrin hüquqi sənədləşdirilməsi

İnsan sağlamlığına yönəlmiş kibertəhlükələrin, xüsusilə intihara təşviq, kiberbullinq və emosional təzyiq kimi halların identifikasiyası yalnız texniki aşkarlanma ilə məhdudlaşmır. Bu hallar həm də hüquqi, sosial və psixoloji aspektlərdən sənədləşdirilməli, qiymətləndirilməli və arxivləşdirilməlidir. Rəqəmsal ekspertiza (digital forensics) bu prosesin mühüm komponentidir və istifadəçi davranışlarının, mətnlərdəki emosional risklərin, o cümlədən qarşılıqlı əlaqələrin texniki və hüquqi dəyərləndirmə üçün dəlil kimi çıxarılmasına xidmət edir [14, 15].

Ənənəvi rəqəmsal ekspertiza yanaşmalarında sübutlar əsasən texniki parametrlərə – fayl tarixçəsi, IP ünvanı, giriş vaxtı və s. əsaslanır. Lakin psixoloji təhlükələr kontekstində bu yanaşma bəzən qeyri-kafi qalır. Emosional və sosial davranışların hüquqi dəyərinin qiymətləndirilməsi üçün mətnin məzmunu və istifadəçi profili ilə əlaqəsi də nəzərə alınmalıdır.

Burada konkret kontekstə bağlı məzmun təhlili ön plana çıxır. Məsələn,"Məni rahat buraxın" ifadəsi neytral görünə bilər, lakin əgər bu cümlə ardıcıl təhqir və təhdid mesajlarından sonra paylaşılıbsa, bu, təzyiq nəticəsində yaranmış emosional müdafiə reaksiyası kimi qiymətləndirilə bilər.

Hər bir mətnin öncəki və sonrakı paylaşım konteksti, o cümlədən sosial qrafdakı yerləşməsi, mətnin intensivliyini və məqsədini dəyişə bilər.

Eyni ifadənin fərqli kontekstdə müxtəlif hüquqi dəyərlərə malik ola biləcəyi nəzərə alınaraq, sistem mətnləri yalnız onun çərçivəsində deyil, öncəki və sonrakı məzmunla birlikdə analiz edir. Bu yolla sübutun mənşəyi, təhrikedici elementi və emosional yükü daha obyektiv qiymətləndirilir.

Kontekstual təhlil sübutun niyyət yönümlü (intent-based) qiymətləndirilməsinə şərait yaradır ki, bu da psixoloji zorakılığın hüquqi təsbitində ənənəvi metodlardan daha yüksək dəyərləndirmə dəqiqliyi təqdim edir.

Sistem xüsusilə təkrarlanan və sistemli emosional zorakılıq hallarını identifikasiya etmək üçün kontekstual oxşarlıq analizi tətbiq edir və bu halları hüquqi istintaq üçün ayrıca sübut klasteri şəklində təqdim edir.

Bu proses aşağıdakı bir neçə addımda göstərilə bilər:

A. Rəqəmsal izlərin qorunması və hüquqi uyğunluq

Analiz üçün ilk və əsas prinsip məlumatın tamlığının, yəni onun hər hansı bir dəyişikliyə məruz qalınmasından qorunmasıdır. Sosial şəbəkə platformalarında istifadəçi fəaliyyətləri (paylaşımlar, şərhlər, qarşılıqlı əlaqələr, sessiya məlumatları və s.) avtomatik şəkildə sistem loqları vasitəsilə qeydə alınır. Bu loqlar yalnız oxuna bilən formatda (read-only) saxlanılır və SHA-256 kimi kriptoqrafik heş (hashing) metodları ilə müdaxiləyə qarşı qorunur. Hər bir mənbə elementi üçün yaradılan heş dəyəri sübutun dəyişdirilmədiyini təsdiq edən rəqəmsal barmaq izi rolunu oynayır [16].

B. Zaman oxu və hadisələrin xronologiyası

Rəqəmsal sübutların əsasını zaman etiketləri (timestamp) təşkil edir. Hər paylaşım, baxış və reaksiya əməliyyatının hansı zaman kəsiyində baş verdiyi dəqiq qeyd olunur. Bu məlumat analitikə "kiberhadisə xəritəsi" qurmağa imkan verir. Belə bir xəritə vasitəsilə, məsələn, bir istifadəçiyə qarşı yönəlmiş sistemli təhqir və təhdid halları zamana görə düzülərək, hüquqi baxışda təkrarlanan və məqsədli zorakılıq nümunəsi kimi təqdim edilə bilər [17].

C. Mənbə və kimliyin müəyyən olunması texnologiyaları

Platforma üzərində yayımlanan şübhəli və riskli məzmunların mənbə analizi üçün istifadəçi IP ünvanları, cihaz identifikatorları (MAC address, IMEI), brauzer izləri (user-agent string) və sessiya məlumatları toplanır. Bu texniki atributlar rəqəmsal iz kimi ekspertiza hesabatlarına əlavə edilir və konkret cihazla konkret hərəkət arasında əlaqə qurmağa imkan verir. Bu cür analiz hüquqi dəlillərin mənbəyi və etibarlılığı baxımından kritik əhəmiyyət daşıyır [15].

D. Emosional və davranış risklərinin arxivləşdirilməsi

Tədqiqat çərçivəsində tətbiq olunan sentiment analizi, loq fayl davranış modelləri və sosial qraf təhlilləri vasitəsilə aşkarlanan yüksək riskli məzmunlar ayrıca rəqəmsal ekspertiza üçün təyin olunmuş arxivdə toplanması nəzərdə tutulur. Bu arxivdə hər bir qeyd üçün aşağıdakılar təqdim olunur:

- Mətnin tam məzmunu və onun təhlil nəticələri (risk növü, ehtimal göstəricisi),
- Paylaşımın tarixi və zamana dair göstəricilər,
- İstifadəçinin nəzərdə tutulmuş zaman kəsiyində davranış göstəriciləri (sessiya uzunluğu, aktivlik səviyyəsi),
- Sistem tərəfindən avtomatik verilmiş risk statusu.

Bu məlumatlar hüquqi orqanlara və psixoloji müdaxilə qruplarına təqdim edilə biləcək tam sübut bazası formalaşdırır [14, 17].

E. Beynəlxalq standartlarla uyğunluq

Rəqəmsal sübutların hüquqi dəyərə malik olması üçün NIST SP 800-86, ISO/IEC 27037, Sübutların ötürülmə zənciri (Chain of Custody) kimi beynəlxalq standartlara uyğunluq təmin edilməlidir. Bu standartlar sübutların necə toplanmalı, saxlanmalı və təqdim edilməli olduğunu müəyyən edir. Sistemdə tətbiq edilən sübut toplama metodlarının bu yanaşmaya uyğunlaşdırılması nəzərdə tutulmuşdur [18, 19].

Bu yanaşma təkcə texnoloji əsaslara deyil, eyni zamanda hüquqi və sosial məsuliyyətlərə əsaslanır. Kibertəhlükələrin və onlar nəticəsində yarana biləcək risklərin yalnız aşkarlanması deyil, həm də etibarlı şəkildə sənədləşdirilməsi üçün prosesə rəqəmsal ekspertizanın inteqrasiyası kritik rol oynayır. Sistem, həm profilaktik xəbərdarlıq həlli, həm də hüquqi istintaq məqsədilə istifadə oluna biləcək tam hüquqi-süni intellekt əsaslı əsaslı təqdim edə bilər.

V. NƏTİCƏ

Bu tədqiqatda insan sağlamlığına yönəlmiş kibertəhlükələrin erkən aşkarlanması və proqnozlaşdırılması məqsədilə rəqəmsal ekspertiza yanaşması təklif olunmuşdur. Təklif olunan model Təbii Dil Emalı (TDE), sosial şəbəkə qraf analizi, loq fayl davranış modelləri və rəqəmsal ekspertiza metodlarını birləşdirərək çoxkomponentli və qarşılıqlı əlaqəli analitik yanaşma təqdim edir.

TDE əsaslı sentiment analiz modulu, xüsusilə depressiv və intihara meyilli məzmunun yüksək dəqiqliklə identifikasiyasını təmin edə bilir (F1-ölçüsü > 0,80). Bu nəticələr sistemin emosional risk faktorlarını mətn üzərindən effektiv şəkildə aşkar edə bildiyini göstərir. Paralel olaraq, sosial qraf strukturlarının təhlili vasitəsilə istifadəçilərarası əlaqələrin intensivliyi və təsir radiusu qiymətləndirilmiş, klasterlərdə mənfi emosional məzmunun lokallaşmış şəkildə yayılması müşahidə olunmuşdur. Yüksək mərkəzilik göstəricilərinə malik istifadəçilərin aşkarlanması isə psixoloji təsirin necə və kimlər tərəfindən yayılmasına dair mühüm dəlillər təqdim edə bilir.

Loq fayllar əsasında qurulan davranış modelləri, istifadəçilərin sessiya müddətləri, paylaşım sıxlığı və məzmun istiqamətləri kimi dəyişənlərə əsaslanaraq avtomatlaşdırılmış psixoloji risk qiymətləndirmə sistemi yaratmağa imkan vermişdir. "Logistic regression" tətbiqi ilə hesablanan Rəqəmsal Risk İndeksi (RRİ) fərdi və klaster səviyyəsində risk proqnozlarını hesablamaq üçün səmərəli mexanizm kimi çıxış etmişdir.

Simulyasiya məlumatları üzərində həyata keçirilmiş bu yanaşma, metodoloji baxımdan sistemin texniki effektivliyini və nəzəri əsaslandırmasını təsdiqləyir. Gələcəkdə bu model real zamanlı sosial media məlumatları və inteqrasiya olunmuş loq sistemləri ilə sınaqdan keçirilərsə, həm elmi, həm də praktik baxımdan önəmli tətbiq perspektivləri yarada bilər.

Təklif olunan yanaşmanın tətbiq imkanlarını aşağıdakı ssenari üzərindən modelləşdirmək olar. Məsələn, bir yeniyetmə sosial mediada ardıcıl şəkildə təhqir və hədə-qorxu məzmunlu paylaşımlara məruz qalır. Onun emosional vəziyyəti zamanla pisləşir, paylaşım intensivliyi azalır, sessiya müddəti artır və daha çox depresiv mövzularda məzmunlarla qarşılıqlı əlaqə qurur. Nəhayət, "həyatdan bezdim", "yaşamaq istəmirəm" kimi paylaşımlar edir ki, buna isə ətraf mühit tərəfindən gec reaksiya verilir.

Təklif olunan sistem bu halı daha erkən, aşağıdakı şəkildə aşkarlaya bilərdi:

- TDE modulu – intihara meyilli və emosional yüklü ifadələri tanıyaraq risk siqnalı yaradır;
- Loq fayl təhlili – istifadəçi davranışında anomaliyaları qeydə alır;
- Sosial şəbəkə qraf analizi – zərərli əlaqələri və kiberbullinqə məruz qalma hallarını aşkara çıxarır;
- Rəqəmsal ekspertiza modulu – bu prosesin izlərini sənədləşdirərək hüquqi istintaq üçün istifadə edilə biləcək sübutlar yaradır.

Beləliklə, təqdim olunan çoxkomponentli rəqəmsal ekspertiza modeli həm profilaktik müdaxilə üçün erkən xəbərdarlıq sistemini, həm də hüquqi çərçivədə istifadə oluna biləcək sübut əsaslı yanaşmanı təmin etməklə insan sağlamlığına yönəlmiş onlayn təhdidlərə qarşı effektiv cavab strategiyası təklif edir.

İSTİNADLAR

[1] S. Hinduja and J. W. Patchin, “Cyberbullying: Identification, Prevention, and Response”, Cyberbullying Research Center, 2020.

[2] S. Livingstone and M. Stoilova, “Cyberbullying, online harms and regulation,” London School of Economics, EU Kids Online, 2019.

[3] UNICEF, “Children in a Digital World,” “The State of the World's Children”, United Nations, 2017.

[4] World Health Organization, “Preventing suicide: A global imperative,” WHO Report, 2014.

[5] A.-L. Barabási, “Linked: The New Science of Networks”, Cambridge, MA: Perseus Publishing, 2002.

[6] S. Livingstone and M. Stoilova, “The 4Cs: Classifying Online Risk to Children”, EU Kids Online, London School of Economics and Political Science, London, UK, 2021.

[7] S. Renjith, A. Abraham, S. B. Jyothi, L. Chandran, and J. Thomson, “An ensemble deep learning technique for detecting suicidal ideation from posts in social media platforms,” arXiv:2112.10609, 2021.

[8] J. Devlin, M. Chang, K. Lee, and K. Toutanova, “BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding,” “Proc. of NAACL-HLT”, 2019

[9] L. C. Freeman, “Centrality in social networks conceptual clarification,” “Social Networks”, vol. 1, no. 3, pp. 215–239, 1979.

[10] V. D. Blondel, J. L. Guillaume, R. Lambiotte, and E. Lefebvre, “Fast unfolding of communities in large networks,” “J. Stat. Mech. Theory Exp.”, vol. 2008, no. 10, P10008, 2008.

[11] M. Girvan and M. E. J. Newman, “Community structure in social and biological networks,” “Proc. Natl. Acad. Sci. USA”, vol. 99, no. 12, pp. 7821–7826, 2002.

[12] Y. Ren, Y. Zhang, X. Zhang, J. Li, and H. Lin, “Depression Detection Based on Multi-Source Heterogeneous Data from Social Media,” “IEEE Trans. Knowl. Data Eng.”, vol. 35, no. 1, pp. 631–644, Jan. 2023.

[13] J. Guntuku, D. Preotiuc-Pietro, M. Eichstaedt, D. Ungar, and L. H. Schwartz, “Personality, demographic, and mental health predictors of posting behavior on social media,” “npj Digital Medicine”, vol. 2, no. 1, p. 43, 2019.

[14] B. Carrier, “Digital Forensics,” in “Handbook of Information Security”, vol. 1, H. Bidgoli, Ed. Hoboken, NJ: Wiley, 2006, pp. 1001–1017.

[15] E. Casey, “Digital Evidence and Computer Crime: Forensic Science, Computers and the Internet”, 3rd ed. Burlington, MA: Academic Press, 2011.

[16] M. K. Rogers and K. Seigfried, “The future of computer forensics: A needs analysis survey,” “Computers & Security”, vol. 23, no. 1, pp. 12–16, 2004.

[17] S. Garfinkel, “Digital forensics research: The next 10 years,” “Digital Investigation”, vol. 7, pp. S64–S73, 2010.

[18] ISO/IEC 27037:2012, “Guidelines for identification, collection, acquisition and preservation of digital evidence”, Int. Organization for Standardization, Geneva, Switzerland, 2012.

[19] K. Kent, S. Chevalier, T. Grance, and H. Dang, “Guide to Integrating Forensic Techniques into Incident Response,” “NIST Special Publication 800-86”, Natl. Inst. of Standards and Technology, Gaithersburg, MD, USA, Aug. 2006.