



# AZƏRBAYCAN RESPUBLİKASININ PREZİDENTİ YANINDA ELMIN İNKİŞAFI FONDU

**Azərbaycan Respublikasının Prezidenti yanında Elmin İnkişafı Fondu  
və Rusiya Fundamental Tədqiqatlar Fondunun  
1-ci Azərbaycan-Rusiya birgə beynəlxalq qrant  
müsabiqəsinin (EIF-BGM-4-RFTF-1/2017) qalibi olmuş  
layihənin yerinə yetirilməsi üzrə**

## YEKUN ELMİ-TEXNİKİ HESABAT

Layihənin adı: **“Big Data” analitikası texnologiyalarının tətbiqi ilə elektron idarəetmənin səmərəliliyinin artırılması metodları və alqoritmlərinin işlənməsi**

Layihə rəhbərinin adı, atasının adı və soyadı: **Rasim Məhəmməd oğlu Əliquliyev**

Qrantın məbləği: **108 000 manat**

Layihənin nömrəsi: **EIF-BGM-4-RFTF-1/2017-21/08/1-M-11**

Müqavilənin imzalanma tarixi: **17 avqust 2020-ci il**

Qrant layihəsinin yerinə yetirilmə müddəti: **18 ay**

Layihənin icra müddəti (başlama və bitmə tarixi): **01 sentyabr 2020-ci il – 01 mart 2022-ci il**

Diqqət! Bütün məlumatlar 12 ölçülü Arial şrifti ilə, 1 intervalla doldurulmalıdır

Diqqət! Uyğun məlumat olmadığı təqdirdə müvafiq bölmə boş buraxılır

Hesabatda aşağıdakı məsələlər işıqlandırılmalıdır:

- |          |  |
|----------|--|
| <b>1</b> | Layihənin həyata keçirilməsi üzrə yerinə yetirilmiş işlər, istifadə olunmuş üsul və yanaşmalar <ul style="list-style-type: none"><li>E-idarəetmə sistemində toplanmış böyük həcmli verilənlərin emalı üçün metodun işlənməsi; “Big Data” analitikası texnologiyalarından istifadə etməklə e-idarəetmənin səmərəliliyinin artırılması üçün metod və alqoritmlərin işlənməsi;</li><li>“Big Data” analitikası texnologiyalarından istifadə etməklə e-idarəetmənin səmərəliliyinin artırılması üçün metod və alqoritmlərin işlənməsi; G2C (hakimiyyət-vətəndaş) modulunda vətəndaşlara göstərilən e-xidmətlərin analizi;</li><li>E-idarəetmə mühitində hakimiyyət və vətəndaşlar (G2C) arasında münasibətlərin intellektual analizi metodlarının işlənməsi; Sosial medianın intellektual analizi üçün metodların işlənməsi;</li><li>Sosial medianın intellektual analizi üçün metodların işlənməsi; Sosial medianın intellektual analizi əsasında e-idarəetmə sisteminin səmərəliliyinin artırılması üçün əks əlaqə mexanizmlərinin işlənməsi;</li></ul> |
|----------|--|

- Sosial medianın intellektual analizi əsasında e-idarəetmə sisteminin səmərəliliyinin artırılması üçün əks əlaqə mexanizmlərinin işlənməsi; Bulud əsaslı e-idarəetmə sisteminin arxitektur-texnoloji əsaslarının araşdırılması;
- Bulud əsaslı e-idarəetmə sisteminin arxitektur-texnoloji əsaslarının araşdırılması; Big data analitikası texnologiyasının tətbiqi ilə bulud əsaslı e-idarəetmə sisteminin işlənməsi.

*E-idarəetmə sisteminin G2C modulunda toplanmış böyük həcmli məlumatların emalı üçün metodun işlənməsi.*

Məlum olduğu kimi, e-dövlət hazırda istifadə olunan mühüm sosial platformalardan biridir. Bu gün hər bir hökumətin məqsədi vətəndaşları qərarların qəbul edilməsi prosesinə cəlb etməklə, onlarla güclü ittifaq qurmaqdan ibarətdir. Dövlət xidmətlərinə əlçatanlığın təmin olunması və vətəndaşlara məlumatların çatdırılmasını asanlaşdırmaq məqsədi ilə e-dövlət platforması yaradılmışdır.

E-dövlət xidmətləri ictimaiyyətin maraqlarına əsaslanaraq onların ehtiyaclarını maksimum dərəcədə nəzərə almaqla, xidmətlərin səmərəliliyini artırmalı, xidmət xərclərini azaltmalı və keyfiyyətini artırmalıdır. Təklif olunan e-xidmətlər tələbyönlü prinsipə əsaslanmalıdır. Bu məqsədə nail olmaq üçün dövlət qurumları vətəndaşların informasiya ehtiyacları barədə məlumatlarla bağlı interaktiv olmalı və ictimaiyyətin real ehtiyacını öyrənməlidir. Bunun üçün vətəndaş (istifadəçi) şərhlərindən istifadə oluna bilər. Lakin bu cür istifadəçi şərhləri geniş mətn məlumatlarının yaranmasına gətirib çıxara bilər ki, bu da onların mürəkkəb, struktursuz olması baxımından analizində müəyyən çətinliklər yaradır. Mətnlərin analizi, korrelyasiyaların aşkarlanması üçün text mining kimi qabaqcıl hesablama və analitik vasitələrdən istifadə etmək nəzərdə tutulur. Vətəndaşların hansı mövzular ətrafında fikir mübadiləsi apardığını aşkar etmək üçün son zamanlar mövzu modelləşdirmə metodları populyarlıq qazanmışdır. Mövzu modeləşdirmə alqoritmləri mətnlərdə sözləri analiz edərək mövzuları aşkarlayan statistik metodlardır. Maşın təlimi və text mining-də istifadə olunan bu metodlar sənədlər çoxluğunda gizli mövzuların müəyyən olunmasında uğurla tətbiq olunur. Bu alqoritmlər xüsusilə klassifikasiya (təsnifat) probleminə sənədlərin siniflərə “yumşaq” (qeyri-səlis) təsnifatı kimi şərh edilir, yəni sənəd tamamilə bir sinfə yox, bir neçə sinfə müxtəlif mənsubiyyət dərəcəsi ilə daxildir. Bundan başqa bu modellərin nəticələri sənədi yalnız bir qrupa daxil etmək üçün də istifadə oluna bilər. Bunları nəzərə alaraq məqalədə əvvəlcə mövzu modelləşdirmə metodları haqqında məlumat verilmiş və daha sonra e-dövlət mühitində yazılan vətəndaş şərhlərinin hansı mövzulara həsr olunduğunu avtomatik olaraq müəyyən etmək və e-xidmətlərin keyfiyyətinin yüksəldilməsi məqsədilə metod təklif olunmuşdur.

Mövzu modelləşdirmə alqoritmləri sənədlər üçün generativ bir modeldir. Belə ki, burada hər bir sənədə mövzular qarışığı kimi baxılır və hər bir mövzu sözlərin onlar üzərində ehtimal paylanması vasitəsilə təyin olunur. Bir sıra modellər, məsələn, Gizli Semantik Analiz, Ehtimallı Gizli Semantik Analiz, Gizli Dirixle Paylanması (GDP) və s. kimi mövzu modelləri sənədlərdən mövzuların aşkar olunması sahəsində (təsnifatlandırmanın dəqiqliyinin yaxşılaşdırılmasında) uğurla tətbiq olunmuşdur. Təklif olunan metodda Gizli Dirixle Paylanması metodundan istifadə edilmişdir. Metodda əsas məqsəd vektorların ölçüsünü azaltmaqla, klasterləşmənin və mövzu modelləşdirmə metodlarının effektivliyinin və dəqiqliyinin artırılmasından ibarətdir. Təklif etdiyimiz metodda əvvəlcə e-dövlət mühitində istifadəçi şərhləri toplanılır. Toplanan şərhlər

ilkin emal olunur. İlkin emal zamanı sənədlərdən ümumişlək sözlər, rəqəmlər və durğu işarələri təmizlənir. Hər bir söz müxtəlif formalarda şəkilçilər qəbul etdiyi üçün bütün sözlər ilkin variantına (kökünə) qaytarılır. Şərhlərdən terminlər çıxarılır. Daha sonra sənədlər çoxluğu TF-IDF sxeminin köməyiylə təsvir olunur. Məlumdur ki, sənədlər çoxluğunda rast gəlinən terminlərin sayı həddən artıq çox olur və bu say bir sənəddə rast gəlinən terminlərin sayından çox-çox böyükdür. Onda sənədlərin TF-IDF sxemi ilə təsvir olunan vektorlarının elementlərinin çox hissəsi "0"-lar olacaq. Başqa sözlə, vektorlar seyrək olacaq. Bu isə sənədlərin klasterləşməsində iki mühüm problem yaradır:

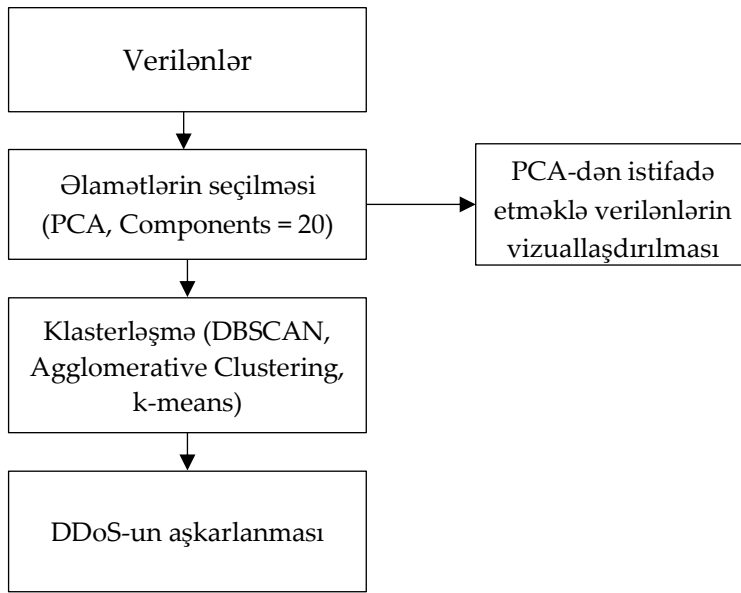
- "Lənətə gəlmiş" ölçü problemi;
- Klasterləşmənin keyfiyyəti.

Bu problemləri aradan qaldırmaq üçün vektordan seyrək terminlər əvvəlcədən təmizlənir. Seyrək terminlər təmizləndikdən sonra yuxarıda göstərilən problemlərə təsir edən digər bir amil ortaya çıxır. Bu da sənədlər çoxluğunda sinonim sözlərin olmasıdır. Belə ki, sənədlər çoxluğunda sinonim sözlər olduqda klasterləşmə zamanı oxşar məzmunlu sənədlər müxtəlif klasterlərə düşə bilər. Bu isə klasterləşmənin keyfiyyətinin aşağı düşməsinə gətirib çıxarır. Bu kimi halları aradan qaldırmaq üçün sənədlər çoxluğunda semantik yaxın sözlərin tapılması və onların kənarlaşdırılması təklif olunur. Sözlərin bir-birinə semantik yaxınlığını tapmaq üçün hər bir terminin genişlənmiş sinonimləri çoxluğundan istifadə olunması təklif olunur. Bunun üçün WordNet şəbəkəsindən istifadə etməklə hər bir terminin sinonimlər çoxluğu tapılır. Qeyd edək ki, WordNet sözlər arasındakı semantik münasibətləri müəyyən etməyə imkan verən şəbəkədir. Məsələn, bu şəbəkənin köməyiylə sinonimləri, hipernimləri, hiponimləri və s. asanlıqla tapmaq mümkündür.

Hər bir terminin genişlənmiş sinonimləri çoxluğu tapıldıqdan sonra sözlər arasında semantik yaxınlıq hesablanır. Əgər sözlər arasında yaxınlıq əvvəlcədən müəyyən olunmuş həddən böyük olarsa, bu sözlər bir termin kimi qəbul olunur. Belə ki, bu sözlərdən yalnız biri saxlanılır, digərləri kənarlaşdırılır. Sənədlər vektor şəklində təsvir olunduqdan sonra klasterləşdirilir. Sənədləri klasterləşdirmək üçün K-means metodundan istifadə olunması təklif olunur. Sənədlər klasterlərə ayrıldıqdan sonra hər bir klaster üzrə mövzuları tapmaq mümkündür. Bunun üçün GDP-dən istifadə olunması təklif olunur. GDP vasitəsilə hər bir klasterdən mövzular çıxarılır və klaster mövzuları müəyyən olunmuş olur.

Beləliklə, təklif olunan metod vasitəsilə vektorun ölçüsünü azaltmaqla, klasterləşmə və mövzu modelləşdirmə metodlarının dəqiqliyini artırmış və sərf olunan zamana qənaət etmiş oluruq. Bu metod vasitəsilə vətəndaşların e-xidmətlərdən narahat olduğu əsas məqamlar tez bir zamanda müəyyən olunmaqla onların məmnuniyyəti təmin oluna və e-xidmətlərin keyfiyyəti artırıla bilər.

PCA və klasterləşmə alqoritmlərinin sintezi əsasında verilənlərin analizi üçün təklif edilmiş DDoS aşkarlama sisteminin arxitekturu şəkil 1-də təsvir edilmişdir.



Şəkil 1. DDoS aşkarlama sisteminin arxitekturu.

Burada ilk əvvəl verilənlərə PCA alqoritmi tətbiq olunaraq ən vacib əlamətlər çıxarılır. Daha sonra seçilmiş əlamətlər əsasında formalaşdırılmış verilənlər bazasına klasterləşdirmə alqoritmləri tətbiq olunaraq DDoS hücumların aşkarlanması həyata keçirilir.

DDoS hücumlarının aşkarlanması yanaşmasının iş axını aşağıdakı addımlardan ibarətdir:

*Addım 1.* X verilənlər matrisinin daxil edilməsi;

*Addım 2.*  $YY^T$  kovariasiya matrisinin daxil edilməsi;

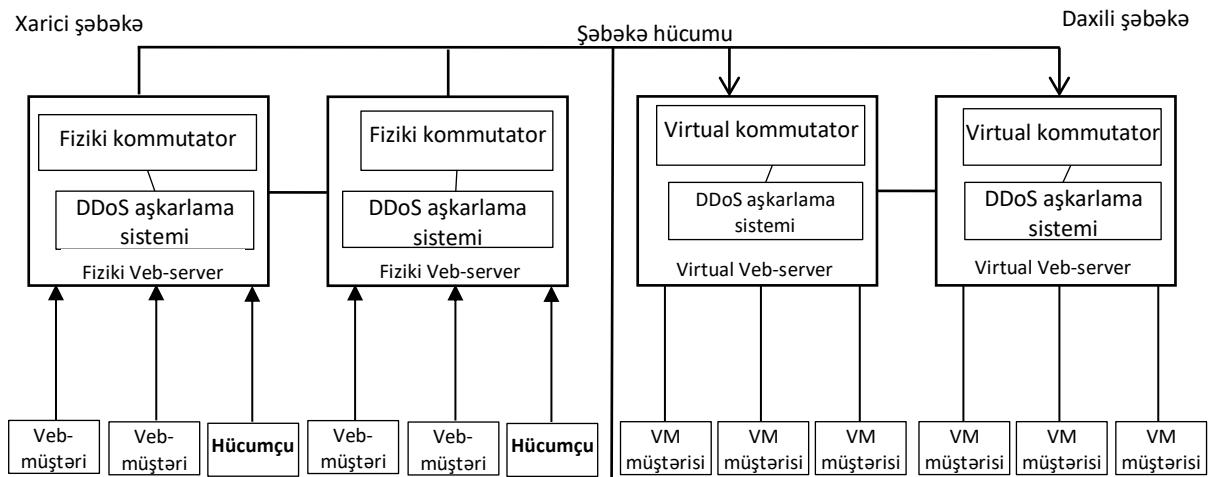
*Addım 3.* Variasiyalari ən böyük olan p sayda əsas komponentin seçilməsi;

*Addım 4.* p sayda əsas komponentdən ibarət W çevirmə matrisinin yaradılması;

*Addım 5.* Ölçüsü azaldılmış D verilənlər çoxluğunun tapılması;

*Addım 6.* Klasterləşmə alqoritmlərindən istifadə etməklə D verilənlər bazasının k sayda klasterə bölünməsi.

Qurulmuş DDoS aşkarlama sistemi buludun daxili və xarici şəbəkəsində yerləşdirilir (şəkil 2).



Şəkil 2. Buludda yerləşdirilmiş DDoS aşkarlama sistemi.

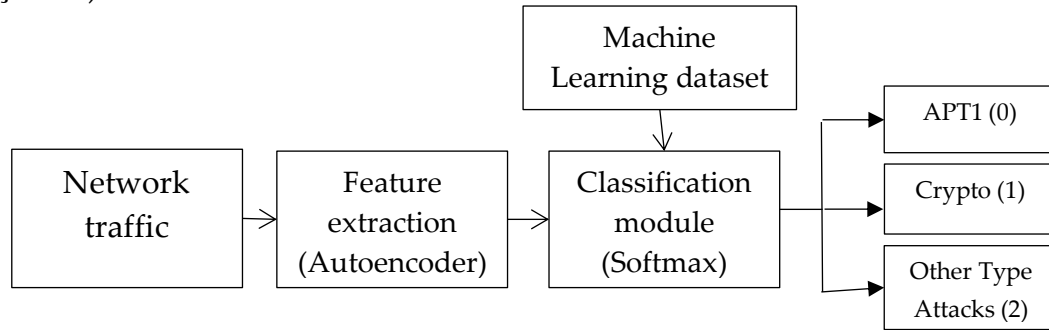
İşlənmiş sistem bulud texnologiyalarının hər bir emal serverində tətbiq edilmişdir. Hər bir serverdə təhlükəsizlik sistemi DDoS hücumunu aşkarlamaq üçün uyğun virtual maşınla əlaqəli

virtual şəbəkədən, daxili şəbəkədən və xarici şəbəkədən gələn və ya çıxan şəbəkə trafikini monitoring edir.

*Bulud infrastrukturuna olan hədəfyönlü APT hücumunun aşkarlanması üsulunun işlənməsi.* Fərz edək ki,  $D = \{(x_i, y_i)\}_{i=1}^n$   $n$  sayda APT nümunələrindən ibarət təlim verilənləri verilmişdir. Burada  $x_i \in R^m$   $m$  ölçülü APT hücum vektorudur,  $y_i \in [1, K]$  hücum siniflərinin uyğun növləridir. Verilmiş bu verilənlər çoxluğunda təlim mərhələsində iştirak etməyən yeni naməlum hücumların aşkarlanması tələb olunur.

$\{x_i\}_{i=1}^n$  təlim verilənlərinin strukturunu öyrənmək üçün dərin avtoenkoder neyron şəbəkəsi istifadə edilmişdir. Avtoenkoder simmetrik neyron şəbəkədir, adətən verilənlər bazasının əlamətlərini supervizorsuz tərzdə öyrənir. Avtoenkoder əlamətlər təsvirini  $x_i$  giriş verilənlərini rekonstruksiya edərək qurur. Avtoenkoder bəzən PCA kimi ölçünün azaldılması üçün də istifadə edilir. PCA verilənlərin transformasiyasını həyata keçirmək üçün xətti funksiya istifadə edir, Avtoenkoder isə qeyri-xətti. Ən sadə halda avtoenkoder enkoder, gizli lay və dekoder hissələrdən təşkil olunur. Modeldə çoxsınıflı klassifikasiyanı həyata keçirmək üçün softmax reqressiya layı istifadə edilmişdir (binar klassifikasiya üçün logistik reqressiya istifadə edilə bilər).

Təklif edilmiş aşkarlama sistemi iki blokdən ibarətdir: əlamətlərin çıxarılması və hücumların təsnifatı (şəkil 3).



**Şəkil 3.** APT hücumunu aşkarlama sistemi.

Öncə APT həyat tsiklində istifadə olunan texnikaları aşkarlamaq üçün şəbəkə trafikinə Avtoenkoder neyron şəbəkəsi tətbiq olunaraq əlamətlər çıxarılır. Bu modulun çıxışında ayrı-ayrı texnikaların generasiya etdiyi hadisələrin əlamətlər vektoru hasil olunur. Generasiya olunmuş əlamətlər vektoru daha sonra klassifikasiya moduluna ötürülür. Bu modulda eyni bir APT hücum ssenarisinə daha çox yaxın olan əlamətlər vektorları kateqoriyalaşdırılır.

*Vətəndaşın davranışı və fərdi məlumatlarına əsasən sosial kreditin qiymətləndirilməsi üzrə dünya təcrübəsinin araşdırılması.* Bu gün müasir intellektual sistemlər, şəbəkə texnologiyalarının imkanlarından istifadə etməklə vətəndaşın qiymətləndirilməsi ənənəvi olaraq bank sistemlərində, şəxsi uçot vərəqələrində və müxtəlif arxivlərdə toplanan statistik göstəricilər əsasında aparılır. Bu tip informasiya əməliyyatları vətəndaşların cəmiyyətdə çəkisi, qanun qarşısında vəzifəsi, sosial normalara necə əməl etməsi və etimadı ilə əlaqədar vəziyyətin öyrənilməsi üçün çox əhəmiyyətlidir. İnformasiya və biliyin cəmiyyətdə rolu artdıqca, qeyri-maddi sistemin formalaşması, işçilərin yaradıcı potensialından maksimum istifadə, sosial tərəfdaşlığın yeni formalarının yaranması və s. kimi sosial hadisələrə daha çox əhəmiyyət verilir.

Bununla əlaqədar vətəndaşların sosial-iqtisadi vəziyyəti, mövcud problemləri və davranışlarının analizi çox aktualdır. Vətəndaşların cəmiyyətdə davranışlarına, təhsilinə, bacarığına, alıcılıq qabiliyyətinə, maraqlarına, etimadına və bir sıra digər xüsusiyyətlərinə görə qiymətləndirilməsi və təsnifatlandırılması sosial kredit adlanır. Sosial kredit vətəndaşın cəmiyyətdəki nüfuzuna təsir edir və cəmiyyətdəki vəziyyəti qiymətləndirmək üçün əhəmiyyətlidir. Sosial kreditin qiymətləndirilməsində əsas məqsəd cəmiyyətdə yüksək intellektə və biliyə malik təbəqənin çəkisini artırmaqdır.

Məlum olmuşdur ki, SKS ilk dəfə 14 iyul 2014-cü ildə Çin Xalq Respublikasında "Sosial kredit sisteminin qurulması planı haqqında (2014-2020)" sənəddə öz əksini tapmışdır. Sənəd "dövlətdə düzgün cəmiyyətin qurulması" məqsədi ilə təklif olunmuşdur və sosial krediti qiymətləndirmək üçün vətəndaşların həyat fəaliyyətlərinin bütün sahələrinin əhatə olunması nəzərdə tutulur: sosial, siyasi və iqtisadi. Layihənin əsas mahiyyətini müəyyən edən sənəddə göstərilir ki, SKS sistemi ətraf mühiti çirkləndirən fabrik və zavod sahiblərindən başlayaraq, korrupsiyaya meyilli dövlət qulluqçularına kimi bütün vətəndaşların fəaliyyətini izləmək imkanına malikdir.

Çin təcrübəsindən məlum olur ki, sosial kreditin qiymətləndirilməsində istifadə olunan məlumatlar əsasən mobil telefonlar, smart-saatlar, planşetlər, smartfonlar və aktivlik izləyicilərindən (wearable activity trackers) əldə edilir. Vətəndaşların gündəlik fəaliyyətini izləyən, yönəldən və müəyyən tapşırıqları xatırladan və yerinə yetirən müasir informasiya texnologiyalarına aid bu cihazlar insanların işini yüngülləşdirmək və dəstəkləmək üçün nəzərdə tutulmuşdur. "Ağıllı əşyalar" adlanan bu avadanlıqlar insanları izləməklə əldə olunan bütün məlumatları müxtəlif serverlərin verilənlər bazasına toplamaq və saxlamaq qabiliyyətinə malikdirlər.

Böyük verilənlərin toplanması və analizi SKS-də "big data" və bulud texnologiyalarından istifadəni labüd edir. SKS-də emal edilən verilənlər vətəndaşa aid ailə, din, sosial durum, maraq və s. haqqında bütün məlumatların cəmlənməsindən, yəni fərdi məlumatlardan əldə edilir. Sosial kreditin qiymətləndirilməsi üçün vətəndaşın həyat fəaliyyətinə aid olan kredit ödəmələrindən virtual məkanda keçirdiyi vaxta kimi bütün məlumatlardan istifadə olunur. Lakin, bu məlumatların necə emal olunması, hansı mənbələrdən toplanması və qorunması haqqında vətəndaşa heç bir məlumat verilmir. Bu baxımdan SKS-nin işində effektivliyi artırmaq üçün və vətəndaşın fərdi məlumatları ilə əlaqədar hüquqlarını nəzərə alaraq fərdi məlumatların e-idarəetmə sistemindən (**FMEİS**) istifadə məqsədəuyğundur.

*FMEİS-nin cəmiyyətdə rolunun müəyyənləşdirilməsi.* Sosial medianın əsas layihələrindən olan Twitter, Facebook, LinkedIn kimi populyar sosial şəbəkələr, Vikipediya virtual ensiklopediyası, Commons və Youtube böyük verilənlərə malik multimedia anbarları, Google və Yahoo kimi nəhəng axtarış serverləri vətəndaşların fərdi məlumatlarını toplayan və emal edən ən nəhəng serverlərdir. Bu serverlərə həm də, FMEİS-in yaranma səbəbləri, inkişaf mərhələsi kimi baxmaq olar. Nəzərə almaq lazımdır ki, bu virtual layihələrdə ayrı-ayrı istifadəçilər və şirkətlər haqqında daxil edilən məlumatlar hər an dəyişir, yenilənir, zənginləşdirilir. Bu serverlərdə toplanan fərdi məlumatların analitik emal hesabına ayrı-ayrı ölkələrdə baş verən demoqrafik prosesləri nəzarətdə saxlamağa, cəmiyyətin nəbzini tutmağa, ölkədəki sosial-siyasi vəziyyəti idarə etməyə çalışırlar.



Dövlət təşkilatları, ayrı-ayrı şəxslər və ya şirkətlər tərəfindən xüsusi texniki vasitələr və proqramlardan istifadə etməklə fərdi məlumatların toplanması gizli və aşkar şəkildə baş verməkdədir. Əsasən də bank sistemləri və sığorta şirkətləri fərdi məlumatları idarə etməklə riskləri azaltmağa və öz iqtisadi göstəricilərini yüksəltməyə çalışırlar.

Çox zaman insanlar toplanan bu fərdi məlumatlardan harada və nə üçün istifadə olunmasından xəbərsiz olur. Bu halda sual yaranır – vətəndaş özü haqqında fərdi məlumatlara necə nəzarət edə bilər? Vətəndaşın fərdi məlumatlarının necə və kim tərəfindən istifadə olunduğuna nəzarət etmək üçün dövlət tərəfindən Milli FMEİS-nin yaradılması məqsəduyğundur. Bu sistemlər fərdi məlumatları toplamaq, təhlükəsizliyini təmin etmək, idarə etmək və nəzarətdə saxlamaq üçün nəzərdə tutulub.

*FMEİS-də big data problemlərinin müəyyənləşdirilməsi.*

FMEİS-də informasiyanın həcmnin böyük olması ilə yanaşı, bu informasiyanın dinamik olaraq hər an dəyişməsi, yenilənməsi, həm də təhlükəsizliyi problemləri məsələlərin həllində “big data” probleminin aktuallığını önə çəkmişdir. Məlum olmuşdur ki, burada problemlər, əsasən verilənlərin müxtəlif mənbələrdən toplanması, dinamikliyi və strukturlaşmaması ilə bağlıdır.

Deyilənləri nəzərə alaraq, əsas problemləri aşağıdakı kimi qruplaşdırılmışdır:

- *Verilənlərin həcmi.* FMEİS ilə əlaqəli işləyən hər bir informasiya sistemində və serverdə, toplanan verilənlərin həcmi onların “big data” kimi analiz olunmasını müəyyən edir. Belə olan halda böyük verilənlərin saxlanması üçün xüsusi şərtlər tələb olunur.
- *Verilənlərin müxtəlif tipdə olması.* FMEİS-də toplanan verilənlər tipinə görə müxtəlifdir müxtəlifdir: mətn, şəkil, audio və video fayllar, cədvəl və s. Bu verilənləri aqreqatlaşdırmadan, lazımı formata gətirmədən bir yerə toplayıb emal etmək mümkün deyil.
- *Verilənlərin emal sürəti.* Fərdi məlumatlar FMEİS-in serverlərinə fərdi şirkətlərdən və dövlət təşkilatlarının informasiya sistemlərindən gətirilir. Serverdə hər bir simvol generasiya olunur. Server üçün konkret zaman çərçivəsində daxil olan böyük verilənlərin sürətli analizini yerinə yetirmək sürətli hesablama yerinə yetirən alqoritmlərlə işləməyi tələb edir. Eyni zamanda nəzərə almaq lazımdır ki, serverdəki bütün verilənlər bazada saxlanıldığına görə zaman keçdikcə verilənlərin həddən artıq çoxalması baş verir. Verilənlərin emalında sürət nə qədər çox olarsa, interaktivlik şərtləri daha yaxşı ödənilir.
- *Verilənlərin dəyişkənliyi.* FMEİS-də hər an dəyişən istənilən informasiyanın bütün köhnə versiyaları bazada saxlanılmalıdır. Bu üsul emal zamanı daha dəqiq nəticə əldə etməyə kömək etsə də, verilənlərin dəyişkənliyi analizi mürəkkəbləşdirir, analiz zamanı verilənlərin idarə olunmasında müəyyən problemlər yaradır.
- *Mürəkkəblilik.* Verilənlər müxtəlif mənbələrdən, verilənlər bazasından daxil olurlar. Bir çox hallarda verilənlər arasında mürəkkəb semantika mövcuddur. Məsələn, vətəndaş haqqında məlumatlar tibb mərkəzindən, təhsil müəssisəsindən, vətəndaşın işlədiyi şirkətin informasiya sistemindən, nəqliyyat agentliyindən, ayrı-ayrı dövlət idarələrindən və s. aid informasiya sistemlərindən toplanır. Belə bir qarışıqlıq verilənlərin emalını çətinləşdirir.

*FMEİS-in vətəndaşa verdiyi hüquqların müəyyənləşdirilməsi:*

- vətəndaşın fərdi məlumatları haqqında informasiya və bu informasiyanın vətəndaş üçün əlçatan olması;
- fərdi məlumatlar dəqiq deyilsə və ya köhnəlmişdirsə vətəndaşın öz fərdi məlumatlarına düzəlişlər etməsinin mümkünlüyü;
- vətəndaş öz fərdi məlumatlarını silmək hüququna malik olmalıdır;
- vətəndaşın məqsəddən asılı olaraq fərdi məlumatlarının emalına qadağa qoyması;
- vətəndaşın fərdi məlumatlarının digər təşkilatlara ötürülməsi ilə bağlı xəbərdar edilməsi;
- vətəndaşın fərdi məlumatlarının digər təşkilatlara ötürülməsinə icazə verməsi və ya qadağa qoyması.

*FMEİS-in müasir e-dövlət mühitində əhəmiyyəti və gələcək perspektivlərinin müəyyənləşdirilməsi:*

- FMEİS bu gün bank sahəsi, ticarət, nəqliyyat, tibb və digər sahələrin radikal dəyişməsinə səbəb olmaqdadır. Sistem informasiya təhlükəsizliyi problemlərinin həlli üçün də əhəmiyyətlidir. SKS-nin səmərəli işinin təmin olunması isə FMEİS-in tətbiqini qaçılmaz edir. Başqa sözlə, SKS FMEİS-nin nüvəsini təşkil edir. Məsələn, FMEİS-in SKS-ə tətbiqindən tibb sahəsi çox əhəmiyyətli dərəcədə faydalana bilər. Sistemdə vətəndaşın səhhəti və həyat şəraiti, iş şəraiti haqqında olan fərdi məlumatlar onun sağlamlığı haqqında proqnozlar verməyə, riskləri aşkarlamağa imkan yaradır.
- Digər tərəfdən, fərdi analitik funksiyanın sistemə daxil edilməsi istifadəçinin şəxsi intellektual köməkçiyə malik olması deməkdir. Bu köməkçi istifadəçinin fərdi məlumatlarının toplanması və idarə edilməsini, informasiya göndəriləcək tərəfin etimadının müəyyən edilməsini və bir çox faydalı işləri həyata keçirə bilər. FMEİS-in funksiyası vətəndaşın bütün fəaliyyəti boyunca (ailədə, iş yerində və hətta sosial mediada) haqqında müxtəlif mənbələrdən toplanan məlumatların aqreqatlaşdırılması və idarə olunmasıdır. Bu məlumatlara qeydiyyat verilənləri, ailə vəziyyəti, maliyyə verilənləri, kredit tarixçəsi və sosial şəbəkələrdəki fəaliyyəti də daxildir.
- FMEİS-ni fərdi məlumatların verilənlər anbarı kimi də təsəvvür etmək olar. Dövlət tərəfindən Milli FMEİS-in yaradılması həm informasiya təhlükəsizliyi, həm də ölkənin iqtisadi inkişafı baxımından çox əhəmiyyətlidir. Milli FMEİS bank, sığorta, istehsalat prosesləri, səhiyyə və s. sahələrin inkişafında mühüm rol oynaya bilər. Milli FMEİS vasitəsilə fərdi məlumatların nəzarətdə saxlanması və idarə olunması vətəndaşlar, vətəndaşla dövlət təşkilatı və şirkətlər arasında münasibətləri idarə etmək üçün bir alətdir. Bu baxımdan hal hazırda dövlət təşkilatları, müəyyən şəxslər və şirkətlər üçün Milli FMEİS-in yaradılması çox böyük əhəmiyyət kəsb edir.
- Onlayn xidmətlərdə daha yüksək qarşılıqlı əlaqələrin əldə edilməsi və gündəlik həyatda fərqli sahələrdə elektron əməliyyatların sayının artması ilə FMEİS-nin əhəmiyyəti özünü daha aydın göstərir. Belə ki, vətəndaşın etimadı, cəmiyyətdə rolu onun fərdi məlumatları ilə xarakterizə edilir.

FMEİS insanlar və təşkilatlar arasındakı münasibətləri dəyişdirə bilər. Məsələn, FMEİS-də məlumatların paylanmış, genişlənmiş şəkildə idarə edilməsi istənilən sayda fənlərarası prosesin optimallaşdırılmasını təmin etməklə müəssisələrin işində məhsuldarlığı yüksəldə və rəqabətədavamlılığını artırmağa imkan verir. Aparıcı FMEİS və SKS istehsalçıları ortaya çıxdıqca, dövlətlər və şirkətlərin bu sistemlərə ehtiyacı çoxaldıqca ehtimal ki, informasiya mübadiləsi üçün



formatların standartlaşdırılmasına, sistemlərin informasiya təhlükəsizliyi şərtlərinə uyğun, etibarlı və çoxfunksiyalı olmasını təmin etmək qarşıya qoyulan ilk vəzifələrdən olacaq.

*Big Data iqtisadiyyatının xüsusiyyətləri və üstünlükləri analiz olunmuşdur.* Bu sahənin e-dövlətin formalaşdırılması və effektiv idarə edilməsi, vətəndaşlarla çevik qarşılıqlı əlaqələrin qurulması, vətəndaş mümnuniyyətinin təmin edilməsi üçün yaratdığı unikal imkanlar şərh olunmuşdur. Big Data analitikası hesabına dövlət idarəetməsinin bütün sahələrində – təhlükəsizlik, müdafiə, hüquq-mühafizə, iqtisadiyyat, kənd təsərrüfatı, təhsil, səhiyyə, sosial təminat, ekologiya, nəqliyyat, rabitə sferalarına uğurlu siyasətin həyata keçirilməsi, optimal qərarların qəbul edilməsi üçün yeni imkanlar açdığı göstərilmişdir.

Həmçinin Big Data iqtisadiyyatının tənzimlənməsini zəruri edən səbəblər araşdırılmışdır. Big Data-nın üstünlükləri ilə yanaşı, vətəndaşlar və cəmiyyət üçün yaratdığı bir sıra risklər analiz edilmişdir. BMT-nin İnsan Hüquqları Bəyannaməsində, Mülki və siyasi hüquqlar haqqında Beynəlxalq Paktda və digər beynəlxalq və regional müqavilələrdə təsbit olunan əsas insan hüquqlarının qorunması ilə bağlı Big Data fəaliyyətinin yarada biləcəyi problemlər göstərilmişdir. Bəzi ölkələrin bu problemlərin qarşısını atmaq üçün həyata keçirdiyi hüquqi tədbirlər araşdırılmışdır.

Eyni zamanda, Big Data-nın hüquqi statusunun müəyyən edilməsi sahəsində mövcud yanaşmalar, baxışlar, problemlər tədqiq edilmişdir. Bəzi beynəlxalq normativ-hüquqi sənədlərdə, bir sıra qabaqcıl dövlətlərin müvafiq qanunvericiliyində bu məsələ ilə bağlı müdələlər şərh olunmuşdur.

Big Data iqtisadiyyatının tənzimlənməsi ilə əlaqədar qanunvericilik aktlarındakı ayrı-ayrı yanaşmalar detallı şəkildə təhlil edilmişdir. O cümlədən, Avropa, Amerika və Çin yanaşmalarının fərqli xüsusiyyətləri göstərilmişdir.

Bunlardan başqa, verilənlərin iqtisadi aktiv kimi tanınması problemləri Big Data iqtisadiyyatının tənzimlənməsi sahəsində ən mühüm problemlərdən biri kimi göstərilmişdir. Müəyyən edilmişdir ki, mövcud qanunvericilik verilənləri əmlak, əmtəə, hüquq obyektinə, yaxud xidmət kimi klassifikasiya etmir. Ənənəvi qanunvericilik verilənlərə, sadəcə, istənilən formada, istənilən daşıyıcıya yazılmış informasiya kimi baxır. Bu da, müvafiq sahənin iqtisadi fəaliyyət istiqaməti kimi tənzimlənməsini çətinləşdirir.

Eləcə də, global Big Data bazarında verilənlərlə bağlı münasibətlərin tənzimlənməsi sahəsində ortaya çıxan problemlərə toxunulmuşdur.

Daha sonra Big Data iqtisadiyyatı sahəsində inhisarçı fəaliyyətin qarşısının alınmasını çətinləşdirən hüquqi boşluqlar araşdırılmışdır. Belə ki, Big Data iqtisadiyyatının tənzimlənməsi sahəsində əsas problemlərdən biri verilənlərdən qeyri-bərabər istifadə, müvafiq bazarın inhisarlaşmasıdır. Bu sahədə tənzimləməni çətinləşdirən əsas problemlərdən biri odur ki, mövcud qanunvericilik aktları verilənləri qiymətli aktiv kimi tanımır.

**Rəqəmsal** əsrdə İnternetin və sosial medianın insan həyatına daxil olması və təsirləri demoqrafik davranışların araşdırılması üçün yeni informasiya mənbələri, verilənlər təqdim edir. Demorqafik tədqiqatlar üçün əhali reyetrlərinin verilənlərindən istifadə olunsada, müəyyən zaman intervalında reyestr insanın həyat dövrü, baş verən hadisələr haqqında bioqrafik verilənlər təqdim etsə də bu məlumatlar olduqca məhdudur. Bu baxımdan effektiv demoqrafik tədqiqatların aparılması üçün yalnız əhali reyestri deyil digər dövlət reyestrlərinin analiz olunması

üçün vahid dövlət reyestri əsasında e-demoqrafiya sisteminin yaradılması olduqca aktual məsələdir. Əhali reyestrlərdə toplanan verilənlərin demoqrafik xarakteristikalarının tədqiq üçün istifadə olunması sahəsində beynəlxalq təcrübə araşdırılmışdır. Təcrübə göstərir ki, hazırda əhali reyestrlərinin verilənlərindən demoqrafik tədqiqatlarda istifadə olunsada digər dövlət reyestrlərin verilənlərindən istifadəyə və e-demoqrafiya sisteminin yaradılmasına dair yanaşmalar mövcud deyil. Qeyd edək ki, ölkələrin prioritetlərindən, e-dövlət quruculuğu modellərindən və strategiyalarından asılı olaraq fərqli e-demoqrafiya modelləri təklif oluna bilər. Konseptual olaraq e-demoqrafiya sisteminin vahid dövlət reyestri əsasında qurulması təklif olunur. Hesab edirik ki, söhbət e-dövlət platformasından gedirsə bütün dövlət reyestrləri inteqrasiya olunmalıdır və reyestrədə toplanan bütün verilənlər demoqrafik tədqiqatlar üçün əhəmiyyətli mənbə rolunu oynayır. Baxılan yanaşmada vahid dövlət reyestri əsasında e-demoqrafiya sistemi yaradılır və e-dövlətin funksiyalarından asılı olmayaraq müstəqil fəaliyyət göstərir. Demoqrafik verilənlərin analizi üçün Big data, OLAP, ETL texnologiyalarından istifadənin üstünlükləri göstərilmiş və təkliflər verilmişdir. Sistemin yaradılması demoqrafik prosesləri daha yaxşı başa düşməyə, daha dərin tədqiqatların aparılmasının sosio-demoqrafiyadan əhali coğrafiyasına, miqrasiya proseslərindən tibbi demoqrafiyaya qədər müxtəlif səviyyələrdə, kəsimlərdə analizlərin aparılması potensialına malikdir.

**Hazırda** istifadəçilərin e-dövlət xidmətlərindən məmnunluq səviyyəsini ölçmək üçün vətəndaşyönümlü yanaşmadan istifadə olunması aktual tədqiqat istiqamətlərindən biri hesab olunur. Vətəndaş məmnuniyyəti e-dövlət layihələrinin uğurlu tətbiqi və genişləndirilməsi üçün əhəmiyyətli göstərici kimi qəbul edilir. E-dövlətin effektivliyinin vətəndaşlar tərəfindən qəbul edilməsi onların xidmətlərdən istifadə niyyətindən, faktiki istifadəsindən asılıdır və eləcə də, davamlı istifadə məmnuniyyət və etibar ilə birbaşa əlaqəlidir. Aparılan tədqiqatlarda dövlət portalından istifadə təcrübəsi olan vətəndaşların e-dövlətdən daha çox məmnun olduqları müəyyən edilmişdir. Bununla yanaşı, nəzəri cəhətdən əlaqəli, lakin empirik olaraq yoxlanılmamış e-dövlət xidmətlərinin reytingləri ilə vətəndaş münasibətləri, qavrayışları arasındakı əlaqəni araşdırmaq olduqca əhəmiyyətlidir. E-xidmətlərin keyfiyyət və əlyetərlik göstəriciləri eləcə də, vətəndaş məmnuniyyətinin ölçülməsinə dair beynəlxalq təcrübə araşdırılmış və ədəbiyyat analizi aparılmışdır. Vətəndaş məmnuniyyəti əsasında e-xidmətlərin qiymətləndirilməsi üçün meyarlar müəyyənləşdirilmişdir. Seçilmiş meyarlara görə vətəndaşların məmnuniyyəti əsasında göstərilən xidmətlərin çoxmeyarlı qiymətləndirilməsi məsələsinə baxılmışdır. Vətəndaş məmnuniyyəti əsasında göstərilən xidmətlərin qiymətləndirilməsi üçün çoxmeyarlı qiymətləndirmə modelindən istifadə olunmuşdur. Ədədi eksperimentdə seçilmiş e-xidmətlər vətəndaşların rəyi əsasında qiymətləndirilmiş və “ən pis hal” metodundan istifadə edilərək rəy qruplaşdırılmışdır. Təklif olunan yanaşma keyfiyyət və əlyetərlik meyarlarını nəzərə almaqla e-dövlət xidmətlərini çoxmeyarlı qiymətləndirməyə imkan verir. Tədqiqatın nəticələri göstərir ki, vətəndaş-hökumət münasibətlərinin səmərəli inkişafı göstərilən e-xidmətlərin keyfiyyətindən və əlyetərliyindən asılı olaraq vətəndaş məmnuniyyətinə əhəmiyyətli dərəcədə təsir edir.

**E-idarəetmədə** fərdi məlumatların qorunması sahəsində beynəlxalq təcrübənin araşdırılmasına həsr olunmuşdur. Fərdi məlumatlar həssas informasiya kateqoriyasına aid edilir və şəxsi məlumatların əldə edilməsi, emalı və istifadəsi qanunla nəzərdə tutulmuş formada,

icazə əsasında olmalı və digər hallarda istifadəsi məhdudlaşdırılmalıdır. E-idarəetmədə məlumatların təhlükəsizliyinin təmin olunması sahəsində bir sır ölkələrin, o cümlədən Estoniya hökumətinin qəbul etdiyi Data səfirliyi konsepsiyası tətbiq olunmuşdur. Data səfirliyi konsepsiyası hökumətin yerli məlumat mərkəzlərinin fəaliyyətinin təbii fəlakət, geniş miqyaslı kibercümm və ya digər böhran vəziyyətində dayandırıldığı bir şəraitdə fəaliyyətini davam etdirməsinə imkan verir. Bununla yanaşı, Data səfirlikləri daha yüksək əlyətərlilik zəmanəti verirsə də, bəzi verilənlərin və ya xidmətlərin qorunması, məlumatın məxfiliyi və bütövlüyü ilə bağlı problemlər səbəbilə hazırda fərdi (xüsusi) buludda yerləşdirilməsi mümkün deyil. Ölkəmizdə fərdi məlumatların qorunması sahəsində mövcud vəziyyət və hökumət buluduna keçid prosesi analiz olunmuşdur. Hökumət buludu konsepsiyası dövlət orqanlarının malik olduğu informasiya sistemlərinin beynəlxalq standartlara cavab verən vahid data mərkəz infrastrukturunda dayanıqlı, çevik, təhlükəsiz və səmərli fəaliyyətinə zəmin yaradır. Hökumət buluduna keçidin mühüm komponentlərindən biri informasiya təhlükəsizliyinin təmin olunmasıdır. Beynəlxalq təcrübədə məlumatların buluda transferi və təhlükəsizliyinin təmin olunmasına dair müxtəlif yanaşmalar olmasına baxmayaraq, rəqəmsal fasiləsizlik təşəbbüsü dövlət xidmətlərin əlyətərlil olmasını, real vaxtda və istənilən vəziyyətdə istifadə oluna bilməsini və yenilənməsini tələb edir.

**E-xidmətlər** haqqında yazılan şərhlərin analizi əsasında və sosial mediada nüfuzlu şəxsləri müəyyən etməklə xidmətlərdən vətəndaşların məmnunluq səviyyəsini qiymətləndirmək üçün yanaşma təklif edilmişdir. Təklif olunmuş metodda əsas məqsəd e-dövlət xidmətləri ilə bağlı mənfi təsirləri minimuma endirmək və müsbət təsiri maksimum dərəcədə artırmaqdır. Bunun üçün şərh yazan vətəndaşlar sosial mediada izlənilə və onların nə dərəcədə nüfuzlu olduğu müəyyən edilə bilər. Belə ki, istifadəçi ilə əlaqəli sosial şəbəkələr, bloqlar, ixtisaslaşmış saytlar, xəbərlər və s.-də bütün məlumatlar izlənə, onunla bağlı ən vacib məsələləri ümumiləşdirən hesabatlar hazırlana bilər. Vətəndaşın e-xidmətdən istifadəsinə təsir edən problemlərin tez bir zamanda müəyyən edilməsi vətəndaş məmnuniyyətinin və dövlət orqanının nüfuzunun artmasına kömək ola bilər.

Təklif olunmuş metodda isə dövlət orqanlarının təklif etdiyi xidmətlər vasitəsilə onların nüfuzunu müəyyənləşdirmək və vətəndaş məmnuniyyətinin qiymətləndirilməsi üçün bu xidmətlərə yazılan şərhərdən istifadə olunur. Belə ki, təklif olunan metod e-dövlət portalında təqdim olunan xidmətlərə yazılan vətəndaş şərhələrini analiz etməklə, hər bir xidmətin orta nüfuz balının hesablanmasına əsaslanır. Bunun vasitəsilə dövlət orqanlarının nüfuzu və onların təqdim etdiyi xidmətlərdən vətəndaş məmnuniyyətinin təyin olunması nəzərdə tutulur.

Təklif olunan metodun əsas mərhələləri bunlardır:

- Şərhələrin toplanması;
- Şərhələrin ilkin emalı;
- Mətnlərin sentiment analizi;
- Vətəndaşların e-xidmətdən orta məmnunluq balı;
- Vətəndaşların dövlət qurumlarından orta məmnunluq balının təyini.

İlkin mərhələdə e-dövlət mühitində hər bir xidmətə yazılan şərhələr vektor şəklində təsvir olunur. Şərhələr toplandıqdan sonra onlar üzərində ilkin emal həyata keçirilir. Mətn ümumişlək sözlərdən, rəqəmlərdən, durğu işarələri və simvollardan təmizlənir. Hər bir söz müxtəlif

formalarda şəkilçilər qəbul etdiyi üçün bütün sözlər ilkin variantına (kökünə) qaytarılır. Daha sonra hər bir sözün sinonimləri tapılır və o, sinonimlər çoxluğu şəklində təsvir olunur. Növbəti mərhələdə hər bir məlumata yazılan şərhlər çoxluğu 3 sinifdə qruplaşdırılır: pozitiv, neqativ və neytral. Yazılan şərhləri bu üç sinifdə qruplaşdırmaq üçün mətnlərin analizində ən qabaqcıl texnologiyalardan biri olan sentiment analizdən istifadə olunması təklif olunur. Hər bir xidmətə yazılan şərhlərin sentiment balı hesablandıqdan sonra hər xidmətin orta məmnunluq balı müəyyən olunur. Xidmətdən orta məmnunluq balı bu xidmətin istifadəçilərinin cəmiyyətdəki nüfuzundan istifadə etməklə təyin olunur.

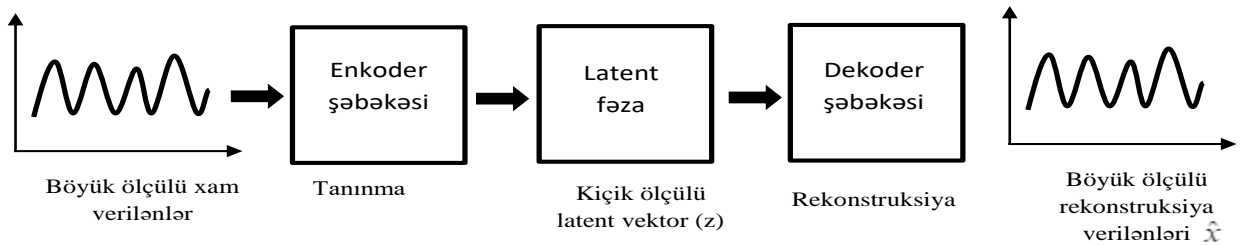
Qeyd edək ki, nüfuzlu istifadəçilər tərəfindən e-xidmət haqqında yazılan hər hansı mənfi və ya müsbət rəy dərhal öz izləyicilərinə ötürülərək yüzlərlə və ya minlərlə istifadəçi arasında dərhal yayıla bilər, bu isə dövlət qurumunun reytinginə birbaşa təsir göstərə bilər. Məhz buna görə də xidmətdən istifadə edən vətəndaşların nüfuzunun və onların xidmətdən məmnunluğunun təyin olunması mühüm əhəmiyyət kəsb edir. Burada iki növ təsir edən şəxs ola bilər: 1) müəyyən sahə (xidmət) haqqında xüsusi biliklərə sahib olan, səlahiyyətləri müəyyən bir sahə ilə məhdudlaşan insanlar, məsələn: maklerlər, avtomobil ustaları və s.; 2) nüfuzları digər sahələri aşan insanlar, məsələn: məşhurlar, idmançılar və s.

İstifadəçilərin nüfuzunu müəyyən etməkdən ötrü onların sosial media hesabından istifadə olunması təklif olunur. Belə ki, istifadəçinin müxtəlif sosial şəbəkələrdə profilləri ola bilər. Bu profillərdəki izləyicilərin sayı, izlənələrin sayı, paylaşdığı postlara olan bəyənmələrin sayı və s.-dən istifadə etməklə hər bir istifadəçinin nüfuzu müəyyən oluna bilər. İstifadəçilərin nüfuzu müəyyən olunduqdan sonra hər bir xidmətin orta nüfuz balı hesablanır. Növbəti mərhələdə e-xidmətlərdən vətəndaşların orta məmnunluq balı əsasında dövlət qurumlarının cəmiyyətdə nüfuzu təyin olunur. Hər bir dövlət qurumunun orta məmnunluq balı müəyyən olunduqdan sonra bu ballara nəzərən dövlət qurumları rəqləşdirilir. Rəqləşdirmə nəticəsində hər bir dövlət qurumunun cəmiyyətdə reytingi təyin olunur.

**E-dövlətin** xidmətlərinin göstərilməsini yaxşılaşdırmağın ən mühüm vasitələrindən biri bulud texnologiyalarıdır. Bulud texnologiyalarının tətbiqi e-dövlətin müxtəlif aspektlərinin yaxşılaşdırılmasına ciddi təsir edə bilər. Bulud əsaslı e-dövlətin effektivliyi elektron dövlət istifadəçilərinin dəyişkən tələblərinə cavab verən virtual resursların effektiv idarə edilməsindən asılıdır. Virtual maşınların resurs utilizasiyası səviyyəsinin qiymətləndirilməsi və müəyyən edilməsi daxil olan tələbat üçün resursların (məsələn, CPU, yaddaş) ayrılmasında bazis hesab olunur.

Bulud infrastrukturunda virtual maşınların iş yükünün Variasiya Avtoenkoderinə əsaslanan proqnozlaşdırılması üsulu təklif edilmişdir.

Fərz edək ki, virtual maşınların yüklənmə dərəcəsini göstərən verilənlər  $x$ , kodlaşdırılmış verilənlər isə  $z$  olsun.  $\varphi$  parametrlərinə malik enkoder şəbəkəsi verilmiş  $x$  giriş verilənlərini  $q_\varphi(z|x)$  şəklində verilmiş paylanmanın vasitəsi ilə  $z$  şəklində kodlaşdırır,  $\theta$  parametrlərinə malik dekoder şəbəkəsi isə  $z$  verilənlərini  $p_\theta(x|z)$  şəklində verilmiş paylanmanın vasitəsi ilə  $x$  verilənləri şəklində dekodlaşdırır. Bu proses aşağıdakı şəkildə təsvir edilmişdir (şəkil 4).



**Şəkil 4.** Virtual maşınların iş yükünün proqnozlaşdırılması.

Fərz edək ki, verilənlərin loqarifmik ehtimalı aşağıdakı kimidir:

$$\log p(x) = \log p(x^{(1)}, \dots, x^{(n)}) = \sum_{i=1}^n \log p(x^{(i)})$$

Onda

$$\log p_{\theta}(x) = D_{KL}(q_{\varphi}(z|x^{(i)})||p_{\theta}(z|x^{(i)})) + E_{q_{\varphi}(z|x^{(i)})}[-\log q_{\varphi}(z|x^{(i)}) + \log p_{\theta}(x, z)]$$

burada  $D_{KL}$  Kulbak-Leibler divergensiya funksiyasıdır (Kullback–Leibler divergence, KL Divergence). Yuxarıdakı ifadənin qiyməti hər zaman müsbət olduğundan aşağıdakı şərt ödənilir:

$$\log p_{\theta}(x) \geq E_{q_{\varphi}(z|x^{(i)})}[-\log q_{\varphi}(z|x^{(i)}) + \log p_{\theta}(x, z)] = L(\theta, \varphi; x^{(i)})$$

burada  $L(\theta, \varphi; x^{(i)})$  variasiya aşağı sərhəddir və aşağıdakı kimi yazıla bilər:

$$L(\theta, \varphi; x^{(i)}) = -D_{KL}(q_{\varphi}(z|x^{(i)})||p_{\theta}(z)) + E_{q_{\varphi}(z|x^{(i)})}[-\log p_{\theta}(x^{(i)}|z)]$$

Bulud texnologiyaları infrastrukturunun mürəkkəbliyini nəzərə alaraq  $q_{\varphi}(z|x)$  və  $p_{\theta}(x|z)$  parametrlərini modelləşdirmək üçün dərin neyron şəbəkəsi istifadə edilmişdir. Enkoder və dekoderin öyrədilməsində Adam optimizasiya funksiyası istifadə edilmişdir. Optimallaşdırmanın məqsədi variasiya aşağı sərhəddi maksimallaşdırmaqdan ibarətdir. Adam – staxostik optimallaşdırma məsələləri üçün yaradılmış staxostik qradiyent üsuludur.

**Rəqəmsal** texnologiyaların, İnternet və sosial medianın insan həyatına daxil olması demoqrafik davranışların araşdırılması üçün yeni informasiya mənbələri, məlumatlar təqdim etmiş olur. Demoqrafik proseslərin idarə edilməsi, demoqrafik davranışların araşdırılması, demoqrafik tədqiqatların aparılması üçün e-demoqrafiya sisteminin yaradılması aktual məsələlərdən biridir. E-demoqrafiya quruculuğu sahəsində mövcud beynəlxalq təcrübə araşdırılmış, vahid əhali reyestrinin yaradılması sahəsində aparılan tədqiqatların mövcud vəziyyəti analiz olunmuşdur. E-demoqrafiya sisteminin qurulması üçün müxtəlif sahələr üzrə elektron reyestrlərin fərdi identifikasiya nömrəsi vasitəsi ilə vahid platformaya inteqrasiya olunması təklif olunur. E-demoqrafiya verilənlər əsasında demoqrafik xarakteristikalar analiz olunmuş və xaricdə təhsil almış məzunların demoqrafik xarakteristikalarının analizi məsələsinə baxılmışdır. Demoqrafik verilənlərin analizi üçün k-means alqoritmindən istifadə olunmuşdur. Məzunların yaşı, cinsi, ailə vəziyyəti, təhsili, ixtisası, təhsil aldığı ölkə və s. göstəricilərinə görə demoqrafik analizlər aparılmışdır. E-demoqrafiya sosial tədqiqatların və əhaliyə dair məlumatların monitorinqinin aparılmasının üçün yeni imkanlar yaradır. E-demoqrafiya sisteminin yaradılması əhali statistikasını, onlayn siyahıyaalma üzrə monitorinqlərin aparılması,



demoqrafik proseslərin dərin təhlil olunması və demoqrafik davranışların öyrənilməsinə imkan verəcəkdir.

Sosial münasibətlərin formalaşmasında və dəyişməsinə virtual mühit çox mühüm rol oynayır. Sosial şəbəkələr sosial proseslərin analizi üçün məlumat mənbəyinə çevrilib. Sosial proseslərin analizi və proqnozlaşdırılması əsasən fərdi məlumatlar əsasında həyata keçirilir. Tədqiqat sosial şəbəkə istifadəçilərinin profillərindəki şərhərdən və şəxsi məlumatlardan istifadə etməklə sosial şəbəkələrdə sosial münasibətlərin analizinə həsr olunub. Tədqiqat hər bir istifadəçinin davranışını aşkarlamaq və onun başqaları ilə münasibətlərinin xarakterini anlamaq məqsədi daşıyır.

Tədqiqatda sosial münasibətlər analiz olunaraq formasına, tipinə, fəsir sahəsinə və xüsusiyyətlərinə görə təsnifatlandırılmışdır. Sosial münasibətlərin analizində istifadə olunan metod və alqoritmlər müəyyənləşdirilmiş, onların qısa icmalı verilmişdir.

Münasibətləri analiz edərkən, ilk növbədə, hər bir insanın virtual məkanda davranışını əks etdirən məlumatlardan və sosial şəbəkələrdə saxlanılan şəxsi məlumatlardan istifadə olunur. Sosial-demoqrafik kontekstin mürəkkəbliyini və çoxşaxəliliyini, onun sosial münasibətlərə təsirini nəzərə alaraq, biz paylaşmış verilənlərin analizi sistemini təklif edirik. Mining alqoritmlərinə əsaslanan metodlar böyük verilənlərin emalının tərkib hissəsidir, yeni biliklər almağa və tədqiqatın səmərəliliyini artırmağa imkan verir. Maşın təlimi metodlarından istifadə təklif olunmuşdur və bu metodlar sosial şəbəkələrdə sosial münasibətlərin müəyyən edilməsi və proqnozlaşdırılması problemlərini daha effektiv həll etməyə imkan verəcəkdir.

Sosial şəbəkələrdəki aktyorlar arasında münasibətləri ölçmək üçün sosial şəbəkə nəzəriyyəsi araşdırılmış, böyük sosial verilənlər (big social data) və böyük sosial media verilənləri (big social media data) ilə əlaqədar tədqiqatlar araşdırılmışdır. Tədqiqatlar əsasında sosial şəbəkələrdə böyük verilənlərlə bağlı problemlər araşdırılaraq müəyyənləşdirilmişdir.

Böyük sosial verilənlərdən istifadə etməklə sosial münasibətlərin analizi üçün təklif olunan metod və yanaşmalar araşdırılmış, böyük sosial verilənlər konsepsiyası əsasında tədqiqat istiqamətləri verilmişdir. Böyük sosial verilənlərdən səmərəli istifadə və sosial şəbəkə mühitində sosial münasibətlərin intellektual analizi üçün metod verilmişdir.

Böyük sosial verilənlər əsasında sosial münasibətlərin intellektual analizi məsələnin səmərəli həlli üçün 3 mərhələli həll üsulu təklif olunmuşdur:

*Sosial münasibətlərin təsnifatı.* Burada əsas göstəricilər aktorlar arasında əlaqələrin sayı, əlaqələrin istiqaməti və çəkisidir. Tədqiqatda məsafə matrisi, dayaq vektor maşını (SVM), KNN və s. istifadə olunur.

*Aktorların klasterləşdirilməsi.* Məsələnin həllində Neyron şəbəkə, genetik alqoritm və s. istifadə olunur.

*Aktyorların rəqləşdirilməsi.* Rəqləşdirərkən aktyorlar arasında liderləri müəyyən etməyə və sosial münasibətlərə təsir edən amilləri aşkarlamağa kömək edir. Tədqiqatda PageRank, SimRank, PeopleRank, HITS alqoritmlərindən istifadə təklif olunmuşdur.

**E-dövlətin** səmərəli idarə olunması, ölkədə sosial-iqtisadi inkişafın və sabitliyin təmini üçün cəmiyyətdə sosial münasibətlərin aşkarlanması və analizi vacib məsələlərdəndir. Sosial münasibətlərin analizi cəmiyyətdə baş verən prosesləri və mövcud sosial problemləri daha aydın şəkildə görməyə imkan yaradır. Bu səbəbdən videotəsvirlərin intellektual analizi

nəticəsində insanları və obyektləri tanımaq, hadisələri və sosial münasibətləri müəyyən etmək üçün bəzi mövcud metod və yanaşmalar araşdırılmışdır. İctimai yerlərdə vətəndaşların davranışlarını müşahidə etməklə sosial münasibətləri aşkarlamaq və baş verə biləcək anomal hadisələri proqnozlaşdırmaq çox mürəkkəb proses olduğu üçün, videomüşahidə sistemləri vasitəsilə əldə olunan təsvirlərdən istifadə etməklə obyektlərin və hadisələrin analizində bəzi mövcud yanaşmalar araşdırılmış, videotəsvirlərin intellektual analizi əsasında sosial münasibətlərin aşkarlanması üçün yeni yanaşma təklif olunmuş, intellektual videomüşahidə sistemlərinin ümumi arxitektura sxemi işlənmişdir. Videomüşahidə kameraları ilə əldə olunan videotəsvirlərin analizi üçün mərhələli həll yolu təklif olunmuşdur.

**Big Data**-nın mühüm mənbəyi olan açıq dövlət məlumatlarının dəyərə çevrilməsi məsələləri ilə bağlı tədqiqat aparılmışdır. Mövzunun elmi aktuallığını şərtləndirən əsas məsələ e-dövlətin fəaliyyəti nəticəsində hər gün böyük həcmdə, müxtəlif xarakterli məlumatlar toplanır. İnformasiya sistemlərinə daxil edilən həmin məlumatlar siyasi, hərbi, təhlükəsizlik, iqtisadiyyat, maliyyə, istehsal, xidmət, energetika, daşınmaz əmlak, təhsil, səhiyyə, nəqliyyat, rabitə, sosial təminat, mədəni sfera, media kimi geniş fəaliyyət sahələrini əhatə edir. Bu məlumatlar istifadə və mühafizə məsələlərinə görə müxtəlif hüquqi kateqoriyaya aid edilir: tam sərbəst (istənilən istifadəçi heç bir autentifikasiya olmadan məlumatları sərbəst şəkildə əldə edə bilər), qismən sərbəst (məlumatları autentifikasiyadan sonra əldə etmək olar), tam qapalı (fərdi məlumatlar, dövlət sirri daşıyan məlumatlar və s.).

2000-ci illərin əvvəllərindən başlayaraq dünyada açıq dövlət məlumatları təşəbbüsləri irəlilə sürülməyə başladı. Bu təşəbbüslərin əsasında açıq dövlət məlumatlarının müxtəlif maraqlı tərəflər üçün əlverişli olması, dövlətin fəaliyyətində şəffaflığın və hesabatlılığın təmin edilməsi, eləcə də, həmin məlumatlardan Big Data analitikası üçün istifadə olunmasına nail olmaq məqsədləri dayanır. Yəni müvafiq təşəbbüslərin mahiyyəti ondan ibarətdir ki, açıq dövlət məlumatları öz-özlüyündə dəyərli deyil, onlar yalnız maraqlı tərəflər üçün əlverişli olduqda və effektiv şəkildə istifadə edildikdə idarəetmə, iqtisadiyyat və sosial sahələrdə dəyər yarada bilər.

Açıq dövlət məlumatlarından səmərəli istifadə əhəmiyyətli dərəcədə iqtisadi faydalar qazandıra bilər. Məsələn, müəyyən edilmişdir ki, ABŞ-da açıq tibbi məlumatlardan səmərəli istifadə etməklə bir ildə 300 milyard dollardan çox dəyər yaratmaq mümkündür.

Açıq dövlət məlumatlarının geniş ictimaiyyət üçün açıq olması, onlardan faydalı məqsədlər üçün istifadə edilməsi ilə bağlı təşəbbüslər nə qədər dəyərli olsa da, bu məlumatların dəyərə çevrilməsi ilə bağlı bir sıra obyektiv və subyektiv problemlər, maneələr mövcuddur. Müvafiq təşəbbüslərin təxminən 20 il öncə meydana gəlməsinə baxmayaraq, bu istiqamətdə bir sıra real addımlar atılsa da, bu sahədə problemlərin, maneələrin aradan qaldırılması üçün hələ çox işlər görülməlidir. Layihə çərçivəsində bu problemlər müəyyən edilmiş və onların aradan qaldırılması yolları göstərilmişdir.

**Xidmətdən imtina hücumları** (Denial of Service, DoS) informasiya təhlükəsizliyi üçün əsas təhdidlərdən biridir, onlar CIA (Confidentiality, Integrity, Availability) triadasında informasiyanın əlverişliliyinin pozulmasına yönəlib. DoS-hücumlar qanuni istifadəçilərin sistemə, şəbəkəyə, tətbiqi proqrama və ya informasiyaya girişini əngəlləmək üçün yerinə yetirilən bədniyyətli hərəkətlərdir. DDoS-hücumlar İnternet infrastrukturunu obyektlərini (router

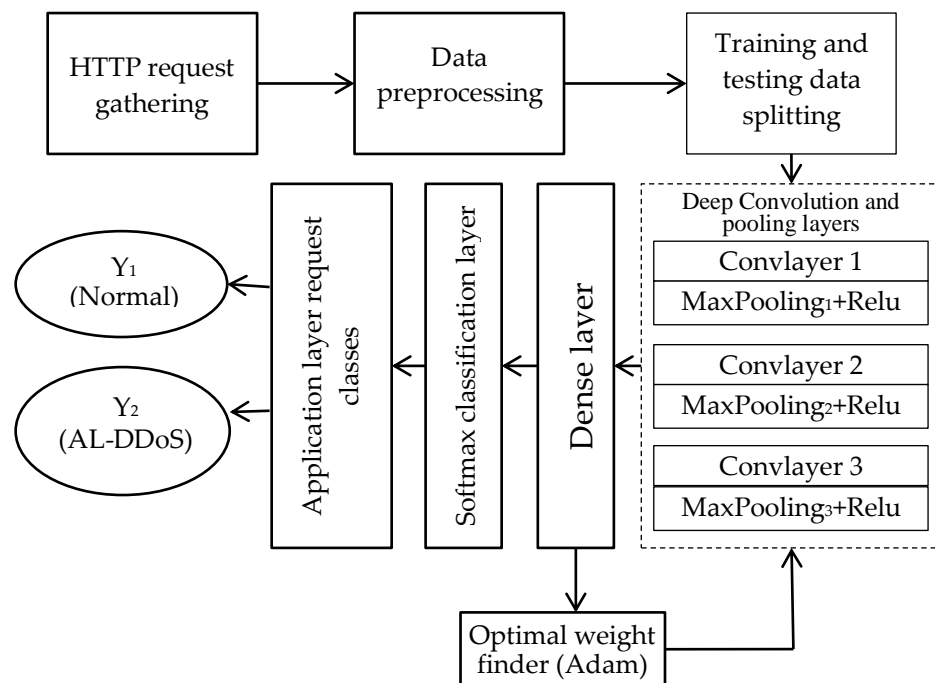
və DNS serverlər daxil olmaqla), buraxma zolağını, serverləri (HTTP, DNS və digər servislər daxil olmaqla) hədəfə alır.

DDoS hücumların aşağıdakı növləri vardır:

- *Həcmə əsaslanan hücumlar* – hədəf serverin buraxma zolağını tutur, ölçü vahidi – saniyədə bitlərdir (bps). Misal – ICMP-, UDP-, TCP-flood hücumlarıdır.
- *Protokola əsaslanan hücumlar* – hədəf serverin resurslarını zəbt edir, ölçü vahidi – saniyədə paketlərdir. Misal – Ping flood, SYN-flood, Smurf hücumlarıdır.
- *Tətbiqi səviyyə hücumları* – tətbiqi səviyyə boşluqlarından istifadə edilməklə hədəf server çökdürülür, ölçü vahidi – saniyədə sorğulardır. Misal – Hash DoS hücumu, Teardrop hücumudur.

Şəbəkə layına olan DDoS hücumları TCP, UDP, ICMP və SYN Floods hücumlarıdır. Tətbiqi səviyyə DDoS hücumları HTTP, HTTPS, və FTP floods hücumlarıdır. Tətbiqi səviyyədə DDoS hücumları hədəf sistemin resurslarını çökdürmək üçün yüksək tezliklə edilən çoxsaylı HTTP-sorğulardan istifadə edirlər. İşdə AL-DDoS hücumların aşkarlanması üçün konvolyusiya neyron şəbəkəsindən istifadə edən avtomatik diaqnostika sistemi təklif edilmişdir.

Tətbiqi səviyyə DDoS hücumların aşkarlanması üçün avtomatik diaqnostika sisteminin arxitekturu şəkil 5-də təsvir edilmişdir.



**Şəkil 5.** DDoS hücumların aşkarlanması üçün avtomatik diaqnostika sisteminin arxitekturu.

Təklif edilmiş yanaşmada CNN modeli konvolyusiya layından, Max-pooling layından, flatten layından və sonuncu tam əlaqəli laydan ibarətdir. Ümumilikdə CNN-nin girişini  $I \in R^k$  şəklində emal olunmamış verilənlər, çıxışını isə klassifikasiyanın  $\hat{y}$  nəticələri təşkil edir.

$$\hat{y} = act(FCN(Flatt(pool(ReLU(conv(I))))))$$

CNN modelinin konvolyusiya layı giriş verilənləri ilə əlaqələnmək üçün çoxsaylı filtrlərdən istifadə edir. Fərz edək ki,  $H \in R^m$  çəki vektoru,  $I \in R^k$  emal olunmamış verilənlərdən əldə

olunmuş verilənlər vektoru və  $b$  sabiti verilmişdir. İki altkonvolyusiya pəncərələri arasında məsafə isə  $d$ -dir.  $I$  vektorunun  $i$ -ci altvektoru  $I_i = [I_{1+(i-1)d}, I_{2+(i-1)d}, \dots, I_{m+(i-1)d}]^T$ ,  $i = 1, 2, \dots, \frac{k-m}{d} + 1$  kimi hesablanır.

1D konvolyusiyanın məqsədi  $H$  vektoru ilə emal olunmamış verilənlərdən əldə olunmuş  $I_i$  verilənlər vektoru arasında hasil hesablamadır. Bu hasil aşağıdakı kimi yazılır:

$$S_i = I_i \times H + b = \sum_{j=1}^m I_{j+(i-1)d} H_j + b,$$

burada,  $H_j$   $H$  vektorunun  $j$ -cu elementidir,  $j = 1, 2, \dots, m$ .

Modelin hər bir konvolyusiya layında aktivləşmə funksiyası kimi Relu (Rectified Linear Unit) istifadə edilmişdir:

$$U_i = \text{Relu}(S_i) \triangleq \max(0, S_i)$$

Optimallaşma funksiyası qradient enişinin hesablanması zamanı digər aktivləşmə funksiyalarından fərqli olaraq Relu qradientin yox olması problemi ilə qarşılaşmır. Bundan başqa Relu funksiyası konvolyusiya şəbəkələrində seyrəkliyi aradan qaldırır, bu isə digər aktivləşmə funksiyaları ilə müqayisədə təlim vaxtını əhəmiyyətli dərəcədə azaldır.

Modeldə birləşmə əməliyyatı kimi Maxpooling istifadə edilmişdir:

$$\text{pool}(U_i) := \max_{l=1, \dots, p} U_{l+(i-1)e}, \forall i = 1, 2, \dots, \frac{k-m}{d} + 1$$

burada  $p$  pulinq ölçüsü,  $e$  addımın ölçüsüdür (stride size).

Konvolyusiya və pulinq layından sonra verilənlər flatten layına ötürülür. Verilənlər Flatten layında 1D struktura çevilir və  $F = [F_1, F_2, \dots, F_q]$  kimi ifadə edilir. Burada  $q$  Flatten funksiyasının tətbiqindən sonra formalaşmış verilənlərin ölçüsüdür. Flatten layından sonra verilənlər tam əlaqəli laya (fully connected layers, FCNs) ötürülür. Daha sonra tam əlaqəli laylarla Relu aktivləşmə funksiyası birləşdirilərək ölçünün azaldılması həyata keçirilir:

$$O = \text{Relu}(W \times F),$$

burada  $W$  tam əlaqəli layların (FCN) çəkilib,  $O = [O_1, O_2, \dots, O_N]$  tam əlaqəli layların çıxış nəticələridir.  $N$  - klassifikasiya məsələsində hücum siniflərinin sayıdır.

Məqələdə klassifikasiya məsələsi həll olunduğu üçün çıxış aktivləşdirmə funksiyası kimi softmax istifadə edilmişdir (reqressiya məsələsində sigmoid istifadə oluna bilər). Belə ki, klassifikasiya məsələsi üçün softmax aşağıdakı kimi ifadə olunur:

$$\hat{y} = \frac{e^{O_n}}{\sum_{j=1}^N e^{O_j}}, \quad n = 1, 2, \dots, N.$$

Proqnozlaşdırılmış qiymətlərlə təlim verilənlərindəki real qiymətlər arasında fərqi minimallaşdırmaq üçün crossentropy funksiyası istifadə edilmişdir və aşağıdakı kimi təyin edilir:

$$L_{\text{crossentropy}} = -\frac{1}{q} \sum_{i=1}^q \sum_{n=1}^N 1\{y_i = n\} \log \hat{y}_i + (1 - 1\{y_i = n\}) \log(1 - \hat{y}_i)$$

burada  $y_i$   $i$ -ci təlim nümunəsinin çıxış qiymətidir və  $q$  təlim nümunələrinin ümumi sayıdır. Yuxarıdakı düsturda  $1\{y_i = n\}$  məntiqi ifadədir, 0 və ya 1 qiymətlərini alır.

İtki funksiyası seçildikdən sonra çəkileri təzələmək məqsədi ilə parametrlərin öyrədilməsi üçün Adam optimallaşdırma funksiyası istifadə edilmişdir. Burada CNN çəkilərini o vaxta qədər təzələyir ki, əvvəlcədən təyin edilmiş maksimum iterasiyada model ən az itki əldə etsin.

**Bulud xidmətlərinə kiberhücumların aşkarlanması üçün üsulun işlənməsi.** Bulud xidmətləri müştərilərə İnfrastruktur Xidmət Kimi (Infrastructure-as-a-service, IaaS), Platforma Xidmət Kimi (Platform-as-a-service, PaaS), və Proqram Təminatı Xidmət Kimi (Software-as-a-service, SaaS) şəklində təqdim edilir. SaaS proqram təminatı tətbiq modelidir, servis provayderi tərəfindən hosting olunur və müştərilərə məsafədən təqdim edilir. SaaS bulud texnologiyalarının başlıca modeli hesab edilir. Verilənlərin itməsi, proqram təminatı boşluqları, əlçatanlıq bu modelin təhdid yaradan ən mühüm faktorlarıdır. Verilənlərin təhlükəsizliyi, şəbəkə təhlükəsizliyi, verilənlərin seqreqasiyası, veb-proqramların təhlükəsizliyi, verilənlərin oğurlanması, virtuallaşmanın boşluqları, əlçatanlıq, identifikatorun idarə edilməsi SaaS modelinin təhlükəsizlik məsələləridir. Bu servis modelində Cross-site scripting, MITM (Man-In-The-Middle) attacks, IP spoofing, port scanning, packet sniffing, DDoS, HTTP attacks, SQL injection flaws kimi kiberhücumlar reallaşa bilər.

Bulud xidmətlərinin kompleks təhlükəsizliyini təmin etmək üçün bu kiberhücumlar vaxtında aşkarlanmalıdır. Bulud texnologiyalarının servis modellərinə olan kiberhücumların aşkarlanması üçün əlamətlərin seçilməsinə əsaslanan yanaşma təklif edilmişdir. Əlamətlərin uyğunluğunu qiymətləndirmək üçün Pirson üsulundan istifadə edilib.

Fərz edək ki,  $[x_{ij}, y_k]$  – giriş verilənləri,  $x_i$  –  $i$ -ci nümunə,  $y_k$  – sinif nişanıdır. Onda Pirson üsulu aşağıdakı kimi təyin edilir:

$$R(i) = \frac{\text{cov}(x_i, Y)}{\sqrt{\text{var}(x_i) * \text{var}(Y)}}$$

burada  $x_i$  –  $i$ -ci dəyişən,  $Y$  çıxış qiymətləri (sınıf nişanları),  $\text{cov}()$  kovariasiya,  $\text{var}()$  variasiyadır. Korrelyasiya üsulu dəyişənlə sinif nişanı (hədəf) arasındakı xətti asılılığı tapır.

Ən yaxşı əlamət altçoxluqları elə əlamətlərdən ibarət olur ki, qrupun əlamətləri ilə hədəf əlamət arasında korrelyasiya yüksək olur, qrupun əlamətlərinin özləri arasında isə heç bir korrelyasiya olmur. Bu əlamətlərin korrelyasiya əsasında seçilməsinin əsas ideyasıdır.

Əlamətlərin korrelyasiya əsasında seçilməsinin iki üsulu vardır: 1) brute force funksiyasının əsasında əlamətlər arasında korrelyasiyanın tapılması; 2) korrelyasiyalı əlamətlər qrupunun tapılması. Adətən verilənlər bazasında ikidən çox əlamət bir-biri ilə korrelyasiyada olur. 3, 4 və daha çox əlamətdən ibarət qruplar tapmaq olur. Bu qruplar tapıldıqdan sonra hər qrupdan hansı əlamətin silinməsi, hansının isə saxlanması həyata keçirilə bilər. Bu prosesi apardıqdan sonra daha dəqiq qərarlar qəbul etmək olur.

**Bulud texnologiyaları mühitində kiberhücumların aşkarlanmasında** maşın təlimi yanaşması çox geniş istifadə olunur. Bu yanaşma hücumları aşkarlama prosesini mərkəzi serverdə həyata keçirir. Bu zaman şəbəkə trafikinin demək olar ki, hamısı analiz edilir. Nəticədə məlumatların konfidensiallığı ilə bağlı ciddi problem ortaya çıxır. Bu problemin aradan qaldırılmasında ən yaxşı həll sensitiv verilənlərin analizinin müvafiq qovşaq kompüterlərdə həyata keçirilməsinə və onların əldə etdiyi nəticələrin koordinator kompüterə ötürülərək mübadiləsinin təmin edilməsinə nail olmaqdır.



Bu məqsədlə bulud texnologiyalarını hədəf alan kiberhücumların aşkarlanması üçün ortalama modelinə (Federated Averaging, FedAvg) əsaslanan CNN (Convolutional Neural Network) alqoritmi üzərində qurulmuş federativ təlim yanaşması işlənmişdir.

Rəqəmsal texnologiyaların, İnternet və sosial medianın insan həyatına daxil olması demoqrafik davranışların araşdırılması üçün yeni informasiya mənbələri, verilənlər təqdim etmiş olur. Tədqiqat işində e-demoqrafiya verilənləri əsasında demoqrafik xarakteristikaların analizi məsələsi araşdırılır. Demoqrafik proseslərin idarə edilməsi, demoqrafik davranışların araşdırılması, demoqrafik tədqiqatların aparılması üçün e-demoqrafiya sisteminin yaradılması aktual məsələlərdən biridir. E-demoqrafiya quruculuğu sahəsində mövcud beynəlxalq təcrübə araşdırılmış, vahid əhali reyestrinin yaradılması sahəsində aparılan tədqiqatların mövcud vəziyyəti analiz olunmuşdur. E-demoqrafiya sisteminin qurulması üçün müxtəlif sahələr üzrə elektron reyestrlərin fərdi identifikasiya nömrəsi vasitəsi ilə vahid platformaya inteqrasiya olunması təklif olunur. E-demoqrafiya verilənləri əsasında demoqrafik xarakteristikalar analiz olunmuşdur. Aparılan eksperimentdə xaricdə təhsil almış məzunların demoqrafik xarakteristikaların analizi məsələsinə baxılmışdır. Tədqiqat işində demoqrafik verilənlərin analizi üçün k-means alqoritmindən istifadə olunmuşdur. Məzunların yaşı cinsi, ailə vəziyyəti, təhsil səviyyəsi, ixtisası, təhsil aldığı ölkə və s. göstəricilərinə görə demoqrafik analizlər aparılmışdır. E-demoqrafiya sosial tədqiqatların və əhaliyə dair məlumatların monitorinqinin aparılmasının üçün yeni imkanlar yaradır. E-demoqrafiya sisteminin yaradılması əhali statistikasını, onlayn siyahıyalma üzrə monitorinqlərin aparılması, demoqrafik proseslərin dərin təhlil olunması və demoqrafik davranışların öyrənilməsinə imkan verəcəkdir.

Beynəlxalq təcrübə göstərir ki, əhali reyestri verilənlərinin davamlı yenilənməsi sosial-iqtisadi göstəricilərin yaxşılaşdırılmasına imkan verən demoqrafik verilənlərin standartının yaradılmasına şərait yaradacaq. Açıq, şəffaf və senzurasız demoqrafik verilənlər demokratik və hesabatlı sistemin yaradılması üçün vacibdir. Şəffalıq və hesabatlılıq hökumətlər tərəfindən qəbul edilən siyasi strategiyaların və proqramların uğurunu qiymətləndirməyə kömək edir. Dünya miqyasında insan inkişafı indeksinə görə öndə olan ölkələrdə demoqrafik verilənlər sistemi daha yaxşı təşkil olunmuşdur. Bu verilənlər bir çox sosial-iqtisadi problemləri demərkəzləşmiş şəkildə həll etməyə imkan verir. Tədqiqat işində inteqrasiya olunmuş əhali reyestri verilənləri əsasında demoqrafik xarakteristikaların analizi məsələsinə baxılır. Əhali məlumatların əldə edilməsi üçün kritik mənbə hesab edilən əhali reyestrlərinin rolu tədqiq olunur. Vahid əhali reyestri verilənlərindən sosio-demoqrafik tədqiqatlarda daha geniş istifadə olunması imkanları araşdırılır. Eksperimental qiymətləndirmə üçün vahid reyestrə inteqrasiya olunmuş fərdlərin məlumatları əsasında COVID-19 pandemiyası şəraitində demoqrafik xarakteristikaların analizi aparılmışdır. Alınmış nəticələr göstərir ki, əhali reyestri əsasında vahid interqasiya olunmuş reyestr – e-demoqrafiya sisteminin yaradılması əldə edilə bilən statistik göstəricilərinin sayının artırılması, daha təfərrüatlı analizlərin aparılması üçün müxtəlif dövlət reyestrlərinin inteqrasiyasını tələb edir. Bu işə daha böyük zaman kəsiyində və daha da detallaşdırılmış səviyyələrdə verilənlərin emalı və statistik təhlilə aparmağa imkan verəcək.

İnternetdəki sosial veb-saytların müxtəlifliyi və sosial medianın insanların gündəlik həyatına təsiri səbəbindən sosial media araşdırmaları əhəmiyyətli dərəcədə aktual olmuşdur. Bu gün virtual mühit sosial münasibətlərin formalaşmasında və dəyişdirilməsində çox mühüm rol

oynayır. Hal-hazırda sosial şəbəkələr sosial proseslərin təhlili üçün məlumat mənbəyinə çevrilib. Sosial proseslərin təhlili və proqnozlaşdırılması elektron dövlətin səmərəli idarə olunması prosesində mühüm vəzifələrdən biridir və fərdi məlumatlar əsasında həyata keçirilir. Layihədə sosial şəbəkə istifadəçilərinin profillərində yerləşən şərhərdən və şəxsi məlumatlardan istifadə edərək sosial münasibətlər araşdırılır və bunun üçün arxitektura sxem təklif edilir. Sosial münasibətlərin intellektual təhlili sisteminin təklif olunan arxitektura sxemi sosial şəbəkələrin istifadəçiləri arasında sosial münasibətləri daha yaxşı başa düşməyə və müəyyən etməyə imkan verəcək.

*İstifadə olunmuş üsul və yanaşmalar:* E-dövlət texnologiyaları; bulud texnologiyaları; e-demoqrafiya; big data analitika; data mining; text mining; opinion mining; sentiment analysis, PCA (Principal Component Analysis – əsas komponentlərin analizi), klasterləşdirmə, təsnifatlandırma, çoxmeyarlı qərar qəbulətmə; qeyri-səlis məntiq; Bayes metodu; sosial şəbəkə texnologiyaları; maşın təlimi; dərin təlim; neyron şəbəkələr; optimallaşdırma üsulları, mövzu modelləşdirmə (topic modelling), qraflar nəzəriyyəsi.

2 Layihənin həyata keçirilməsi üzrə planda nəzərdə tutulmuş işlərin yerinə yetirilmə dərəcəsi (faizlə qiymətləndirməli)

100%

3 Hesabat dövründə alınmış **elmi nəticələr** (onların yenilik dərəcəsi, elmi və təcrübi əhəmiyyəti, nəticələrin istifadəsi və tətbiqi mümkün olan sahələr aydın şəkildə göstərilməlidir)

Eksperiment üçün *BBC NEWS* verilənlər çoxluğundan istifadə olunmuşdur. Bu verilənlər çoxluğu 2004-2005-ci illəri əhatə edən *Biznes*, *Əyləncə*, *Siyasət*, *İdman* və *Texnologiya* adlanan beş aktual sahəyə uyğun *BBC* xəbər veb saytından toplanmış 2225 sənəddən ibarətdir. Eksperimentdə *Biznes*, *Əyləncə* və *İdman* sahələrindən müxtəlif sayda sənədlər toplanmış və analiz olunmuşdur. Klasterləşmə nəticələrini qiymətləndirmək üçün "Təmizlik" (Purity) ölçüsündən istifadə olunmuşdur.

Eksperiment zamanı toplanan sənədlər əvvəlcə ilkin emal olunmuşdur. İlkin emal zamanı sənədlər çoxluğundan durğu işarələri, rəqəmlər, simvollar, ümumişlək sözlər təmizlənmiş və sənədlər çoxluğu TF-IDF sxeminin köməyiylə vektor şəkildə təsvir olunmuşdur. Daha sonra sənədlərdən seyrək terminlər təmizlənmişdir. İlkin emaldan əvvəl və sonra sənədlər çoxluğunda qalan sözlərin sayı cədvəl 1-də təsvir olunmuşdur.

**Cədvəl 1.** Sənədlər və sözlərin sayı

Sənədlərin sayı	Sözlərin sayı: ilkin emaldan	
	əvvəl	sonra
100	8040	4421
300	18356	8851
500	25490	11766
800	33410	14750
1000	36346	15860

İlkin emaldan sonra sənədlər çoxluğuna təklif etdiyimiz metod tətbiq olunmuşdur. Sözlər arasında semantik yaxınlıq  $\alpha$ -nın müxtəlif qiymətlərində (0.1, 0.2, 0.3, 0.4, 0.5) hesablanmışdır. Qeyd edək ki, burada  $\alpha$  -sərhəd qiymətidir. Metoddan sonra sənədlər çoxluğunda qalan terminlərin sayı cədvəl 2-də təsvir olunmuşdur. Cədvəldə sənədlərin sayı 100 olan varianta

baxdıqda, buradan aydın görünür ki,  $\alpha = 0.1$  qiymətində daha çox sayda semantik yaxın sözlər aşkarlanmış və vektorun ölçüsü seyrək terminlər təmizləndikdən sonra qalan sözlərə nisbətən xeyli dərəcədə (26.42%) azalmışdır.  $\alpha$ -nın qiyməti artdıqca, daha az sayda sözlər atılmışdır. Belə ki,  $\alpha = 0.5$  qiymətinə baxdıqda, vektorun ölçüsünün daha az (1.29%) azaldığını görürük. Sənədlərin sayı artdıqca, semantik yaxın olan sözlərin sayı da uyğun olaraq artmış və vektorun ölçüsü xeyli dərəcədə azalmışdır. Məsələn, sənədlərin sayı 800 olan variantda baxsaq,  $\alpha = 0.1$  qiymətində vektorun ölçüsünün 31.67% azaldığını görürük.

**Cədvəl 2.** Emaldan sonra sənədlərdəki sözlərin sayı

Sənədlərin sayı	Sözlərin sayı					
	Azışlək terminlər təmizləndikdən sonra	Semantik sözlər atıldıqdan sonra ( $\alpha =$ )				
		0.1	0.2	0.3	0.4	0.5
100	772	568 (26.42%)	691 (10.49%)	733 (5.05%)	756 (2.07%)	762 (1.29%)
300	878	633 (27.90%)	777 (11.50%)	838 (4.56%)	856 (2.51%)	865 (1.48%)
500	827	589 (28.77%)	725 (12.33%)	789 (4.59%)	809 (2.17%)	816 (1.33%)
800	821	561 (31.67%)	716 (12.79%)	778 (5.23%)	802 (2.31%)	810 (1.34%)
1000	801	561 (29.96%)	702 (12.36%)	765 (4.49%)	784 (2.12%)	791 (1.25%)

Daha sonra sənədlər çoxluğuna  $k$ -means klasterləşmə metodu tətbiq olunmuş və  $\alpha$  -nın müxtəlif qiymətlərində purity hesablanmışdır (cədvəl 3). Qeyd edək ki, burada  $\alpha = 0$  qiyməti seyrək terminlər təmizləndikdən sonra qalan sənədlər çoxluğunu ifadə edir. Qeyd edək ki, eksperiment zamanı 3 sahədən (biznes, əyləncə, idman) sənədlərə baxıldığından, klasterlərin sayı 3 götürülmüşdür. Cədvəldən görüldüyü kimi sənədlər çoxluğundan semantik yaxın sözlərin atılması klasterləşmənin keyfiyyətinə mənfi təsir göstərməmiş, əksinə təmizlik əmsalı kifayət qədər yüksək qiymət almışdır. Bu da klasterləşmənin keyfiyyətinin yüksək olduğunu göstərir. Cədvəldən də görüldüyü kimi sənədlərin sayı artdıqca, təmizlik əmsalının qiyməti də kifayət qədər yüksək qiymət almışdır. Belə ki,  $\alpha = 0.1$  vəziyyətinə baxsaq, sənədlərin sayı 100 olan halda təmizlik əmsalı 0.88 olduğu halda, sənədlərin sayı 1000 olduqda bu əmsal 0.97 olmuşdur.

Sənədlər klasterləşdirildikdən sonra hər bir klasterdən mövzuları çıxarmaq üçün mövzu modelləşdirmə metodu tətbiq olunmuşdur. Bunun üçün hər bir klasterə Gizli Dirixle Paylanması tətbiq olunmuş və klaster mövzuları çıxarılmışdır. Cədvəl 4 və 5-də klasterlər üzrə hər bir mövzudan çıxarılan top 10 sözlər təsvir olunmuşdur. Top 10 söz dedikdə, yəni hər bir mövzu üzrə sənədlər çoxluğunda ən çox işlənən sözlər nəzərdə tutulur.

**Cədvəl 3.**  $\alpha$  -ın müxtəlif qiymətlərində Purity əmsalı

Sənədlərin sayı	Purity $\alpha =$				
	0.0	0.1	0.2	0.3	0.4
100	0.950	0.880	0.810	0.820	0.820
300	0.996	0.980	0.980	0.980	0.980
500	0.996	0.930	0.940	0.990	0.990
800	0.980	0.810	0.890	0.990	0.990
1000	0.982	0.970	0.980	0.970	0.980

**Cədvəl 4.** Klasterlər üzrə mövzular və top 10 sözlər ( $\alpha = 0$ )

Cluster 1			Cluster 2			Cluster 3		
Topic 1	Topic 2	Topic 3	Topic 1	Topic 2	Topic 3	Topic 1	Topic 2	Topic 3
film	Music	music	Rugbi	İreland	open	cluster	Cluster	Cluster
star	award	band	england	England	play	govern	Growth	Year
year	people	Show	Player	Cluster	cluster	compani	Year	Sale
a.m	Show	Year	Play	Win	win	Firm	Rate	Profit
Role	Year	Album	Cluster	Wale	year	deal	economi	Compani
Actor	Won	number	Year	Side	seed	Tax	Bank	Firm
Cluster	Radio	cluster	Cup	Game	match	plan	Econom	Market
Director	Veto	Chart	Zealand	Tri	Set	india	Oil	Share
İnclud	Years	Top	World	Nation	beat	countri	Price	Car
Festiv	Song	singer	Tour	Scotland	world	foreign	Rise	Euro

**Cədvəl 5.** Klasterlər üzrə mövzular və top 10 sözlər ( $\alpha = 0.3$ )

Cluster 1			Cluster 2			Cluster 3		
Topic 1	Topic 2	Topic 3	Topic 1	Topic 2	Topic 3	Topic 1	Topic 2	Topic 3
music	Film	music	england	England	open	cluster	Cluster	Cluster
Band	Star	Show	Ireland	Play	play	firm	Team	Govern
Album	Year	Song	Cluster	Cluster	cluster	year	Year	Compani
Number	Award	Peopl	Wale	Year	year	sale	Rate	Tax
Top	Role	Year	Zealand	Rugbi	seed	compani	economi	State
Year	Cluster	award	Side	Player	match	profit	Rise	Foreign
Chart	İnclud	Radio	Rugbi	Game	set	share	Price	Countri
Singer	director	Make	Player	Season	beat	market	Bank	Oil
Rock	Play	Years	Tri	Cup	world	a.m	econom	İndia
Cluster	Bbc	british	Game	Week	australian	Car	month	China

Cədvəl 4-də  $\alpha = 0$  halında, yəni semantik yaxın sözlər kənarlaşdırılmamışdan əvvəl sənədlər çoxluğundan çıxarılan top 10 söz təsvir olunmuşdur. Müqayisə üçün cədvəl 5-də  $\alpha = 0.3$  halında, semantik yaxın sözlər kənarlaşdırıldıqdan sonra sənədlərdən çıxarılan top 10 söz təsvir olunmuşdur. Bu iki cədvələ diqqət yetirsək, semantik yaxın sözlərin sənədlər çoxluğundan kənarlaşdırılmasının nəticəyə təsir göstərmədiyini görürük. Yəni, GDP-in keyfiyyəti aşağı düşməmişdir. Belə ki, hər iki cədvəldə klaster 1-ə baxdıqda buradakı terminlərin əyləncə, klaster 2-də idman, klaster 3-də isə biznes ilə əlaqəli olduğunu görürük. Bu isə klaster 1-in Əyləncə, klaster 2-in İdman və klaster 3-ün Biznes sənədlərindən ibarət olduğunu göstərir. Göründüyü kimi hər iki cədvəldə klasterlərdən çıxarılan mövzular üst-üstə düşür. Bu isə təklif olunan metodun yaxşı işlədiyini göstərir. Qeyd edək ki,  $\alpha$ -ın digər qiymətlərində (0.1, 0.2, 0.4, 0.5) də bu nəticə özünü doğruldur. Yəni,  $\alpha$ -nın digər qiymətlərində də klasterlərdən çıxarılan mövzular cədvəl 5 ilə üst-üstə düşür, buna görə də müqayisə üçün yalnız  $\alpha = 0.3$  qiyməti verilmişdir.

Göründüyü kimi mövzular dəqiqliklə çıxarılmış, zamanda isə uduş əldə olunmuşdur. Belə ki, aşağıdakı cədvəldə klasterləşməyə və hər bir klasterdən mövzuların çıxarılmasına sərf olunan zaman və onların müqayisəli analizi təsvir olunmuşdur (cədvəl 6).

Cədvəl 6-dan göründüyü kimi təklif olunmuş metod vasitəsilə zamanda xeyli uduş əldə olunmuşdur. Belə ki, sənədlərin sayı 100 olan variantda zamanda 19.82 % uduş əldə olunursa, sənədlərin sayı artdıqca (sənədlərin sayı 1000)  $\alpha = 0.1$  halında bu faizin 39.31 %-ə qədər yüksəldiyini görə bilərik. Bu isə səmərəliliyin xeyli dərəcədə yaxşılaşması deməkdir.

**Cədvəl 6.  $\alpha$  -in müxtəlif qiymətlərində sərf olunan zaman**

Sənədlərin sayı	Sərf olunan zaman					
	$\alpha =$					
	0	0.1	0.2	0.3	0.4	0.5
100	11.6	9.3 (19.82%)	9.86 (15%)	10.32 (11.04%)	10.65 (8.19%)	11.06 (4.65%)
300	12.39	9.08 (26.71%)	9.74 (21.39%)	10.68 (13.80%)	11.26 (9.12%)	11.51 (7.10%)
500	18.7	11.35 (39.30%)	12.87 (31.18%)	13.28 (28.98%)	14.21 (24.01%)	14.98 (19.89%)
800	22.28	13.49 (39.45%)	14.59 (34.52%)	15.76 (29.26%)	16.57 (25.63%)	17.85 (20.33%)
1000	21.06	12.78 (39.31%)	14.09 (33.09%)	15.01 (28.72%)	16.01 (23.97%)	17.25 (18.09%)

Beləliklə, aparılan eksperiment və nəticələri onu göstərir ki, təklif olunmuş metod vasitəsilə böyük sənədlər çoxluğunun ölçüsünü xeyli dərəcədə azaldaraq, bu verilənlərin analizinə sərf oluna vaxta qənaət etmək və klasterləşmə, GDP alqoritminin keyfiyyətini yaxşılaşdırmaq olar.

**1) Bulud infrastrukturuna olan DDoS hücumların aşkarlanması** üçün əsas komponentlər metodu (Principal Component Analysis – PCA) və klasterləşmə alqoritmlərini birləşdirən hibrid model təklif edilmişdir. İşin əsas elmi yeniliyi lazımsız əlamətləri silməklə əsas əlamətləri saxlamaq üçün əlamətlərin seçilməsi üsulundan istifadə edilməsidir. Bundan əlavə əlamətlərin seçilməsi yanaşmasına əsaslanan təklif edilmiş metodun klasterləşmə indeksləri əsasında iyerarxik klasterləşmə, k-means, DBSCAN kimi digər etalon metodlarla müqayisəli analizi aparılmışdır. Verilənlər bazasının həcmindən bu üsulla azaldılması DDoS hücumlarının aşkarlanması dəqiqliyini kifayət qədər artırmışdır.

Bulud mühitində DDoS hücumların effektiv aşkarlanmasını təmin etmək üçün klassifikasiya modeli böyük həcmdə verilənlərlə öyrədilməlidir. İşdə eksperimentlərin aparılması üçün Kanada Kibertəhlükəsizlik İnstitutunun (Canadian Institute of Cybersecurity) CSE-CIC-IDS2018 verilənlər bazası istifadə edilmişdir. DoS, DDoS, Brute Force, XSS, SQL Injection, Infiltration, Port scan və Botnet CSE-CIC-IDS2018 verilənlər bazasının hücum sinifləridir. Təqdim olunan işdə verilənlər bazasının 0 - Benign, 1 - DDoS attack, 2 - DoS attack siniflərdən ibarət csv faylı götürülmüşdür və bu faylda sətirlərin sayı 1048574-dür.

Təklif edilmiş yanaşmanın CSE-CIC-IDS2018 verilənlər bazasına tətbiqi zamanı mövcud metodlarla müqayisədə üsul müxtəlif klasterləşmə metrikaları üzrə yüksək qiymətlər almışdır.

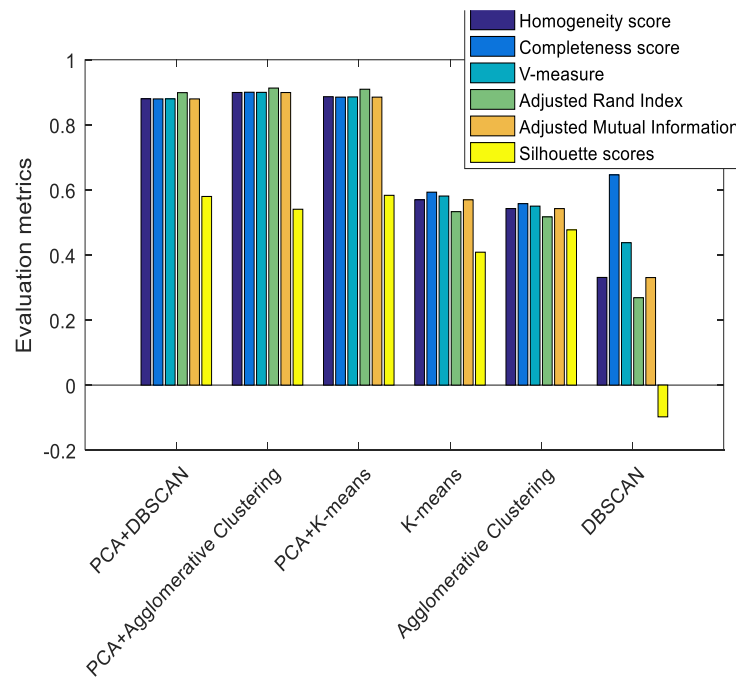
**Cədvəl 7. Müxtəlif qiymətləndirmə indekslərinə görə metodların müayisəsi**

	Homogeneity score	Completeness score	V-measure	Adjusted Rand Index	Adjusted Mutual Information	Silhouette scores	Estimated number of clusters
PCA+DBSCAN	0.8801	0.8797	0.8799	<b>0.8989</b>	0.8797	<b>0.5798</b>	3
PCA+Agglomerative	0.8994	0.9002	0.8998	<b>0.9130</b>	0.8993	<b>0.5405</b>	3
PCA+k-means	0.8864	0.8851	0.8858	<b>0.9094</b>	0.8851	<b>0.5833</b>	3
k-means	0.5696	0.5930	0.5811	0.5331	0.5696	0.4085	3
Agglomerative	0.5426	0.5577	0.5501	0.5171	0.5425	0.4772	3
DBSCAN	0.3309	0.6462	0.4377	0.2686	0.3305	-0.0976	7



Burada ARİ (Adjusted Rand Index) klasterə daxil olan nümunələrin nə dərəcədə oxşar olduğunu müəyyən edir. Bu səbəbdən ARİ indeksi nə qədər yüksək olsa, klasterləşmə modeli bir o qədər yaxşı hesab edilir. Aparılan eksperimentlərdə PCA+DBSCAN, PCA+Agglomerative, PCA+k-means alqoritmləri ARİ indeksi üzrə yüksək qiymətlər almışdır və uyğun olaraq 0.8989, 0.9130, 0.9094 təşkil etmişdir. Lakin ənənəvi k-means, Agglomerative və DBSCAN alqoritmləri ARİ indeksi üzrə aşağı qiymətlər almışdırlar və uyğun olaraq 0.5331, 0.5171, 0.2686 təşkil etmişdir. Qeyd edək ki, ARİ 1 və -1 aralığında qiymətlər alır. İndeksin 0-ra yaxınlaşması bütün nöqtələri təsadüfi seqmentlərdə yerləşdirmək deməkdir. Nöqtələrin təsadüfi seqmentlərdə yerləşdirilməsi isə yaxşı hal hesab edilmir. Yüksək Silhouette qiymətinə malik model yaxşı klasterləşmə hesab edilir. Təklif edilən yanaşma bu metrika üzrə də yüksək qiymətlər almışdır.

Yuxarıdakı cədvələ daxil edilmiş nəticələrin effektivliyini əyani nümayiş etdirmək üçün onun qrafik təsviri şəkil 6-da verilmişdir.



Şəkil 6. Alqoritmlərin müqayisəsi.

Buradan göründüyü kimi təklif edilmiş yanaşma mövcud yanaşmalarla müqayisədə daha üstün nəticələr göstərmişdir.

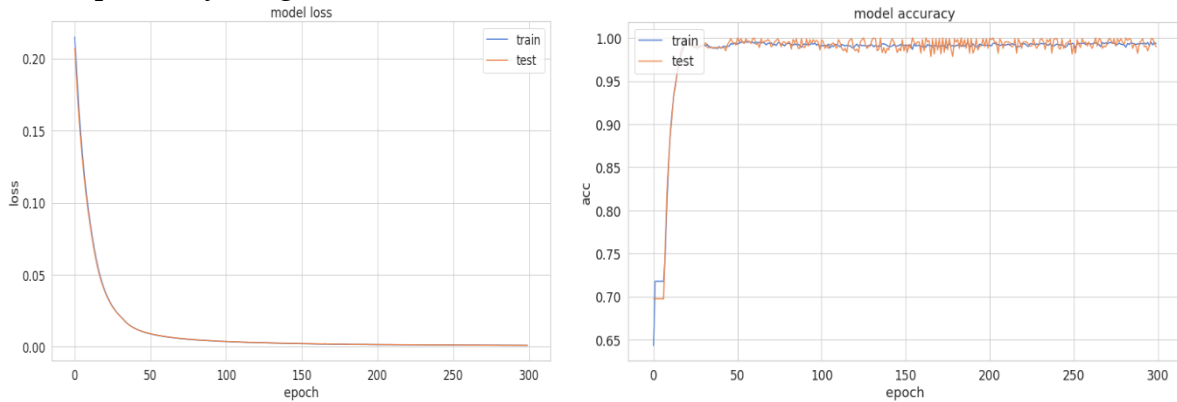
İşin elmi yeniliyi ondan ibarətdir ki, burada ilk əvvəl avtoenkoder neyron şəbəkəsi tətbiq olunaraq şəbəkə trafik verilmələrindən supervizorsuz tərzdə informativ əlamətlər öyrənilmişdir. Əlamətlərin öyrənilməsi mərhələsindən sonra qurulmuş avtoenkoder şəbəkəsinin üst qatına softmax reqressiya layı əlavə edilərək

APT hücumlarının klassifikasiyası həyata keçirilmişdir. İşdə müxtəlif laylar əlavə etməklə qurulmuş dərin neyron şəbəkə modeli elmi tədqiqatlar üçün açıq olan verilənlər bazasının üzərində müxtəlif klassifikasiya indeksləri əsasında test edilmişdir və təklif edilmiş üsul APT hücumlarının aşkarlanmasında mövcud üsullarla müqayisədə daha üstün nəticələr vermişdir. Təklif edilmiş yanaşmanın bulud mühitinə tətbiq modeli işlənmişdir, bulud informasiya

sisteminin APT hücumlarına qarşı təhlükəsizliyini gücləndirmək üçün (One Time Password, OTP) mexanizminə əsaslanan iki faktorlu autentifikasiya sistemi təklif edilmişdir.

Eksperimentlərin aparılması üçün işdə MalwareTrainingSets verilənlər bazası istifadə edilmişdir. Toplanmış verilənlər bazasına APT1, Crypto, Locker, Zeus zərərli proqramları üzrə uyğun olaraq 292, 2024, 434, 2014 nümunə daxil edilmişdir. Eksperimentin aparılması prosesində ilk öncə Avtoenkoder neyron şəbəkəsi MalwareTrainingSets verilənlər bazası üzərində öyrədilmişdir, daha sonra avtoenkoderin çıxışında alınmış nəticələr softmax reqressiya alqoritminin girişinə ötürülərək hücumların klassifikasiyası aparılmışdır.

Qurulmuş avtoenkoder modelinin öyrədilməsi zamanı iterasiyaların sayını artırıdığca model bir o qədər az itkiyə yol verməklə daha dəqiq nəticələr verməyə çalışır. Bu mənzərə şəkil 7-də vizual olaraq daha aydın görünür.



Şəkil 7. Modelin iterasiyaların sayından asılılığı.

Avtoenkoderin nəticələrinin klassik KNN klassifikasiya alqoritmi ilə accuracy metrikası üzrə müqayisəli analizi aşağıdakı cədvəldə göstərilmişdir.

Cədvəl 8. Autoencoder və KNN metodlarının Accuracy metrikasına görə müqayisəsi

Attack types	Autoencoder	KNN	Number of samples
APT1 (0)	<b>0.9832</b>	0.3321	119
Crypto (1)	<b>0.6088</b>	0.5834	795
OtherType Attacks (2)	0.1268	<b>0.7342</b>	1041

Cədvəldən görüldüyü kimi Avtoenkoder KNN alqoritmi ilə müqayisədə daha üstün nəticələr vermişdir.

FMEİS-nin sosial kredit sistemi ilə birgə işinin ümumi arxitekturu işlənmişdir. FMEİS-i dəstəkləyən verilənlər mətn, audio, video-fayl, şəkil, veb-səhifə, elektron cədvəl və sxem ola bilər. Verilənlər xüsusi verilənlər anbarında və ya fayl sistemlərində iyerarxik şəkildə saxlanılır. Bulud hesablamada modelinə əsasən, proqram istifadəçiyə məlumatlarla işləmək üçün bütün lazımi funksiyaları təqdim edir: siyahıların qurulması, məlumatların filtrasiyası, çeşidlənməsi, redaktə edilməsi, dəyişikliklərin tarixçəsinə baxılması, redaktələrin forma və məzmunu.

FMEİS özündə bir sıra alt e-idarəetmə sistemlərinin birləşməsindən təşkil oluna bilər. Bu e-idarəetmə sisteminə sənədlərin və digər məlumatların idarə olunmasını həyata keçirən

sistemlərdən başlayaraq, sosial şəbəkələrdən əldə olunan məlumatlara kimi vətəndaşın davranışını əhatə edən bütün sahələr aid ola bilər.

Təklif olunan arxitekturda prosesin ümumi ardıcılığı aşağıdakı kimidir:

1) FMEİS vətəndaşlar haqqında fərdi məlumatları toplayır (müxtəlif informasiya sistemlərindən və virtual mühitdən).

2) Böyük verilənlər strukturlaşdırılır və emal edilir. Fərdi məlumatların emalı sistemlərini aşağıdakı kateqoriyalara ayırmaq olar:

- Biometrik informasiya emalı sistemi;
- Demografik informasiya emalı sistemi;
- Açıq informasiya emalı sistemi;
- Gizli informasiya emalı sistemi;
- Maliyyə informasiya emalı sistemi;
- Təşkilat daxili informasiya emalı sistemi;
- Virtual davranış informasiya emalı sistemi;
- Xüsusi kateqoriyalı fərdi məlumatların emal sistemi. Xüsusi kateqoriyalı fərdi məlumatlar dedikdə fiziki şəxsin irqi və ya milli mənsubiyyətinə, ailə həyatına, dini etiqadına və əqidəsinə, məhkumluğuna aid olan məlumatlar nəzərdə tutulur.

3) Emal nəticəsi müraciət olunan tərəfə ötürülür. Müraciətə uyğun olaraq fərdi məlumatların və emal nəticələrinin ötürülməsi aşağıda göstərilən hallarda mümkündür:

- müqavilə imzalamaq və icra etmək üçün;
- qanunla tələb olunan məlumatları əldə etmək üçün;
- qanuni (hüquqi) maraqların həyata keçirilməsi üçün;
- müştəridən rəsmi şəkildə razılıq alındığı halda.

*FMEİS-ə gətirilən fərdi məlumatların etibarlılığını və düzgünlüyünü təmin etmək üçün prinsiplər işlənmişdir. Bu prinsiplər aşağıdakılardır:*

- fərdi məlumatların mənbəyi dəqiqləşdirilməli və düzgünlüyünə əminlik olmalıdır;
- fərdi məlumatların istifadəsində şəffaflıq olmalıdır. Məlumatı istifadə etməyə hazırlaşan təşkilat (operator) məqsədi haqqında subyektə xəbərdar etməlidir;
- təşkilat fərdi məlumatların təhlükəsizliyi haqqında zəmanət verməlidir;
- fərdi məlumatların verilənlər bazasının ehtiyat sürətləri yaradılmalı və etibarlı yerdə saxlanmalıdır;
- təşkilatla subyekt arasında qarşılıqlı etimad və inam nəzərə alınmalıdır (subyekt fərdi məlumatların onun sosial-iqtisadi rifahına və ailəsinə qarşı, qərəzli şəkildə istifadə olunmayacağına əmin olmalıdır);
- fərdi məlumatın təhlükəsizliyi və istifadəsi üçün razılaşma mexanizmi qanun çərçivəsində, həmçinin, faktiki davranışlar nəzərə alınmaqla işlənməlidir.

Burada hüquq prinsipləri də gözlənilməlidir. Fərdi məlumatların e-idarə olunması qanuni olmalıdır və verilənlərin emalını qanuni edən hüquqi bir əsas olmalıdır. Başqa sözlə desək, fərdi məlumatların idarə olunmasına icazə verən hüquqi sənəd olmalıdır. Təcrübə göstərir ki, fərdi məlumatların idarə olunmasında bəzən mövcud hüquqi əsaslar yetərinə olmur. Yəni, bəzən

fərdi məlumatların emalı üçün tək-cə vətəndaşın razılığı kifayət etmir, onun ailə üzvləri və ya işlədiyi şirkətin razılığı da tələb olunur.

**Big Data** iqtisadiyyatı sahəsində mövcud olan hüquqi problemlər müəyyən edilmişdir. Bu sahənin tənzimlənməsini zəruri edən səbəblər klassifikasiya olunmuşdur. Big Data-nın hüquqi statusu, bu sahənin tənzimlənməsi sahəsində beynəlxalq təcrübə, mövcud yanaşmalar müqayisəli şəkildə analiz edilmişdir. Verilənlərin iqtisadi aktiv kimi tanınması, qlobal mübadiləsi, Big Data fəaliyyəti sahəsində inhisarçı fəaliyyətin qarşısının alınması problemləri şərh olunmuşdur. Göstərilən problemlərin aradan qaldırılması, müvafiq sahənin normativ-hüquqi bazasının möhkəmləndirilməsi, effektiv tənzimlənməsi üçün təkliflər irəli sürülmüşdür.

Tədqiqat nəticəsində müəyyən olunmuşdur ki, beynəlxalq səviyyədə Big Data analitikası, iqtisadiyyatı üçün istifadə edilən verilənlərin, o cümlədən fərdi məlumatların mühafizəsi haqqında universal, çoxtərəfli müqavilələr mövcud deyil. Həmçinin bu sahədə beynəlxalq səviyyəli nəzarət qurumu yoxdur. Sadəcə, bu sahədə bir neçə beynəlxalq müqavilələr vardır ki, onlar da insan hüquqlarının, o cümlədən şəxsi həyatın toxunulmazlığının təmin edilməsini, eləcə də, şəxsi keyfiyyətlərə görə insanların diskriminasiyasının qadağan olunmasını nəzərə tutur.

Əldə olunan nəticələrə görə, Big Data iqtisadiyyatı ilə bağlı əsas problemlər sırasında fərdi məlumatları istifadə edilən vətəndaşların ictimai etimadı, verilənlərin hamı üçün əlyətərliliyi, onlardan birgə istifadə, verilənlərin etibarlı mühafizəsi məsələləridir. Eyni zamanda, dünya hüquq praktikasında hələlik verilənlərin iqtisadi aktiv kimi tanınmaması Big Data sahəsində inhisarçı fəaliyyətin qarşısının alınması üçün qanuni mexanizmlərin tətbiqinə imkan vermir. Big Data iqtisadiyyatının inkişafını ləngidən səbəblərdən biri də bəzi milli qanunvericilikdə verilənlərin qlobal məkanda mübadiləsini, alqı-satqısını məhdudlaşdıran normaların müəyyən edilməsidir.

Müəyyən olunmuşdur ki, Big Data iqtisadiyyatının tənzimlənməsi sahəsindəki qeyd edilən boşluqları mövcud qanunvericilik vasitəsi ilə aradan qaldırmaq mümkün deyil. Həmin boşluqların aradan qaldırılması, insanların öz fərdi məlumatlarından səmərəli istifadə etmələri üçün Big Data ilə bağlı proseslərin əhatə etdiyi bütün sahələr üzrə ictimai münasibətlərin tənzimlənməsinə yeni fundamental yanaşma tələb edilir.

**İKT-nin** sürətli inkişaf etdiyi, ölkələrin sosial-siyasi vəziyyətinin sürətli dəyişdiyi bir şəraitdə demoqrafik prosesləri ənənəvi statistik üsullarla təhlil etmək mümkün deyil. Ölkədəki effektiv demoqrafik siyasətin həyata keçirilməsi elektron dövlət sisteminin ayrılmaz bir hissəsi kimi qəbul edilir və mövcud demoqrafik vəziyyətin qiymətləndirilməsi, təhlili və effektiv qərar qəbul edilməsi vasitəsilə demoqrafik xüsusiyyətləri əhatə edən reyestr və sistemlərin effektiv istifadəsini tələb edir. Demoqrafik prosesləri idarə etmək, məqsədyönlü demoqrafik siyasətin həyata keçirilməsi üçün elektron demoqrafiyanın formalaşdırılması aktual məsələlərdən biridir.

Demografik proseslərə miqrasiya, doğum və ya ölüm faktlarından əlavə müxtəlif xəstəliklərin, hərbi əməliyyatların, təbii fəlakətlər və s., həmçinin sosial-iqtisadi vəziyyətin də təsir göstərməsi qeyd edilir və bu göstəricilərin demoqrafiyanın analizində nəzərə alınmasının vacibliyi göstərilir. Rəqəmsal demoqrafiya tədqiqatlarının əsas mənbələri kimi isə veb bruzerlərdə axtarış sorğuları, sosial media verilənlər, dövlət reyestrlərində toplanan verilənlər, e-xidmət verilənləri, vətəndaşların məmnunluq göstəriciləri, hökumət-vətəndaş

münasibətlərində əks əlaqə verilənləri və s. aid edilir. Göstərilən mənbələrdə toplanan böyük verilənlərin analizi mövcud tədqiqatları tamamlamağa və demoqrafik davranışlarla əlaqəli yeni təsəvvürlərin yaranmasına, biliklərin əldə olunmasına imkan verəcəkdir. Demoqrafiyanın perspektiv tədqiqat istiqamətləri müəyyənləşdirilmiş, demoqrafik prosesləri idarə etmək üçün e-demoqrafiya sisteminin konseptual model təklif edilmişdir.

Təklif olunan e-demoqrafiya sisteminin konseptual modeli demoqrafik analiz üçün vacib olan informasiya mənbələri araşdırılmış, elektron reyestrlərin yaradılması və inteqrasiyası məsələlərinə baxılmışdır. Aparılan tədqiqatlar göstərir ki, e-demoqrafiya sistemi əhali reyestrinin monitorinqinə və e-dövlət portalına inteqrasiya olunmuş elektron reyestrlərdə müxtəlif xarakterli demoqrafik göstəriciləri təhlil etməyə imkan verəcəkdir.

Vahid əhali reyestrinin yaradılması ənənəvi siyahıyalmanın ləğv edilməsinə və dövlət reyestrində toplanan məlumatların analizi əsasında faktiki olaraq əhalinin fasiləsiz onlayn siyahıyalınmasına və monitorinqinin aparılmasına imkan verəcəkdir. Bu əhali reyestri ölkənin sosial-iqtisadi inkişafını təmin etmək üçün müxtəlif qərar qəbuletmə sistemlərində, ən əsası isə e- demoqrafiya sistemində effektiv şəkildə istifadə edilə bilər ki, bu da sosial tədqiqatların və əhali məlumatlarının monitorinqi üçün səmərəli həllərin işlənməsinə imkan yaradacaqdır. Mövzunun aktuallığı nəzərə alınaraq, gələcək tədqiqatlarda müxtəlif dərəcəli demoqrafik analizlərin aparılması məqsədilə sosial media vasitələrində və dövlət reyestrlərində toplanan böyük verilənlərin intellektual analizi məsələlərinə baxılacaqdır.

**Beynəlxalq** təcrübəyə əsasən vətəndaşların rifah və məmnunluq səviyyəsinin yüksəldilməsi üçün dövlət xidmətlərinin təkmilləşdirilməsinin idarəetmənin səmərəliliyinin artırılması prosesi ilə əlaqəli şəkildə aparılması mühüm əhəmiyyət kəsb edən məsələdir. Dövlət orqanlarının fəaliyyətində şəffaflığın artırılması, vətəndaşlara göstərilən xidmətlərin daha keyfiyyətli, rahat və qabaqcıl texnologiyalar tətbiq etməklə həyata keçirilməsi, vətəndaşlara münasibətdə etik qaydalara, nəzakətli davranışa əməl edilməsi son nəticədə vətəndaş məmnunluğunun yüksəldilməsinə eləcə də, idarəetmənin səmərəliliyinin təmin olunmasına xidmət edir.

Vətəndaş məmnunluğu əsasında e-xidmətlərin çoxmeyarlı qiymətləndirilməsi məsələsi araşdırılmışdır. Beynəlxalq təcrübənin analizi göstərir ki, e-dövlət xidmətlərindən istifadə də vətəndaş məmnuniyyətinin araşdırılması ilə bağlı mövcud tədqiqatlar daha çox müxtəlif xarakterli hesabatların individual səviyyədə analizinə və ümumi mənbələrdən götürülmüş məlumatların subyektiv qiymətləndirilməsinə əsaslanır. İstifadəçilərin dövlət xidmətlərindən məmnunluq səviyyəsini ölçmək üçün xidmətlərin göstərilməsində vətəndaşyönümlü yanaşmadan istifadə olunması effektiv həll hesab olunur. E-dövlət xidmətlərindən istifadə də vətəndaş məmnunluğunun qiymətləndirilməsinə dair tədqiqatların ədəbiyyat analizi aparılmışdır. Vətəndaş məmnunluğu əsasında e-xidmətlərin qiymətləndirilməsinə dair mövcud yanaşmalar analiz olunmuşdur. E-xidmətlərin qiymətləndirilməsi meyarları tədqiq olunmuş və dövlət xidmətlərinin keyfiyyətini xarakterizə edən meyarlar müəyyənləşdirilmişdir. Çoxmeyarlı qiymətləndirmə modeli əsasında e-xidmətlərin rəqləşdirilməsi üçün ən pis hal metodundan istifadə edilmişdir. Ədədi eksperimentdə vətəndaşların rəyi əsasında xidmətlər qiymətləndirilmiş və ən pis hal metodu tətbiq olunaraq rəqləşdirilmişdir. Təklif olunan yanaşma keyfiyyət və əlyetərlilik meyarları nəzərə alınmaqla e-xidmətləri çoxmeyarlı qiymətləndirməyə imkan verir.



Bu baxımdan belə nəticəyə gəlmək olar ki, vətəndaş-hökumət münasibətlərinin səmərəli inkişafı göstərilən e-xidmətlərin keyfiyyətindən və əlyetərliliyindən asılı olaraq vətəndaş məmnuniyyətinə birbaşa təsir edir. İdarəetmədə effektivlik dövlət orqanları arasında fəaliyyətin səmərəsinin, əmək məhsuldarlığının, şəffaflığın və çevikliyin artırılması vasitəsilə vətəndaş məmnunluğunu ən yüksək səviyyədə təmin etməyə imkan verir.

Hazırda ölkələr vətəndaşlarına daha yaxşı e-xidmətlər təklif edə biləcək səmərəli e-idarəetmənin formalaşdırılmasına səy göstərirlər. Bu baxımdan bulud texnologiyaları hökumətlər və vətəndaşlar üçün geniş imkanlar yaradır. E-idarəetmənin inkişaf etdirilməsi fərdi məlumatların qorunması və informasiya təhlükəsizliyinin təmin olunmasını zəruri edir. Fərdi məlumatların təhlükəsizliyi, insan hüquqları və əsas azadlıqlar tək-cə ali ümumbəşəri dəyərlər deyil, həm də iqtisadi və siyasi sabitliyin və inkişafın ən etibarlı təməli hesab olunur.

Məlumatların təhlükəsizliyinin təmin edilməsi sahəsində beynəlxalq təcrübə araşdırılır. Ölkələrin təcrübəsində fərdi məlumatların saxlanılmasına, xaricə ötürülməsinə və bu proseslərin yerinə yetirilməsinə dair yanaşmalara baxılır və qəbul edilmiş qanunlar qeyd olunur. Fərdi məlumatların qorunmasına dair bir sıra ölkələrin, o cümlədən Avropa Birliyi, Rusiya, Estoniya, ABŞ kimi ölkələrin təcrübəsi araşdırılmışdır. Estoniya hökumətinin irəli sürdüyü Data səfirliyi təşəbbüsü dövlətin buludda genişlənməsidir və bu dövlətin ərazi hüdudlarından kənar server resurslarına sahib olması mənasını verir. Rəqəmsal transformasiyalar və e-idarəetmə strategiyasının qəbul edilməsi Estoniyaya gətirdiyi inanılmaz üstünlüklər dövlət administrasiyasının daha çevik, effektiv və əlyətərli olmasına şərait yaratdı. Data səfirliyi - hökumətin mühüm serverlərini diplomatik cəhətdən etibarlı bir yerə transfer edən ilk ölkə nümunəsini təqdim edir. Qeyd edək ki, data səfirlikləri daha yüksək əlyətərlilik zəmanəti verirsə də, bəzi məlumatların (məsələn, dövlət sirri hesab olunan) qorunması, məxfiliyi və bütövlüyü ilə bağlı problemlər səbəbilə hazırda xüsusi bulud xidmətində yerləşdirilməsi mümkün deyil. Ölkəmizdə e-dövlətin inkişafı, rəqəmsal platformalara transformasiya, e-xidmətlərin səmərəliyinin artırılması, informasiya sistemlərinin fəaliyyətinin keyfiyyətli, dayanıqlı və təhlükəsiz infrastrukturda təşkilinin təmin edilməsi, vətəndaşların resurslara fasiləsiz çıxış imkanının yaradılması məqsədilə "bulud" texnologiyasının tətbiqi son dərəcə əhəmiyyətli məsələdir. Hökumət buludunun yaradılması, ölkədə dayanıqlı, təhlükəsiz və səmərəli İKT infrastrukturunun yaradılmasına və onun vasitəsilə dövlət qurumlarının informasiya sistemlərinin istismarına imkan verir. Hökumət buludunun məqsədləri eləcə də, Milli Data mərkəzin arxitekturasının həm fiziki həm də kibertəhlükələrə qarşı təhlükəsizlik tədbirləri qeyd olunmuşdur.

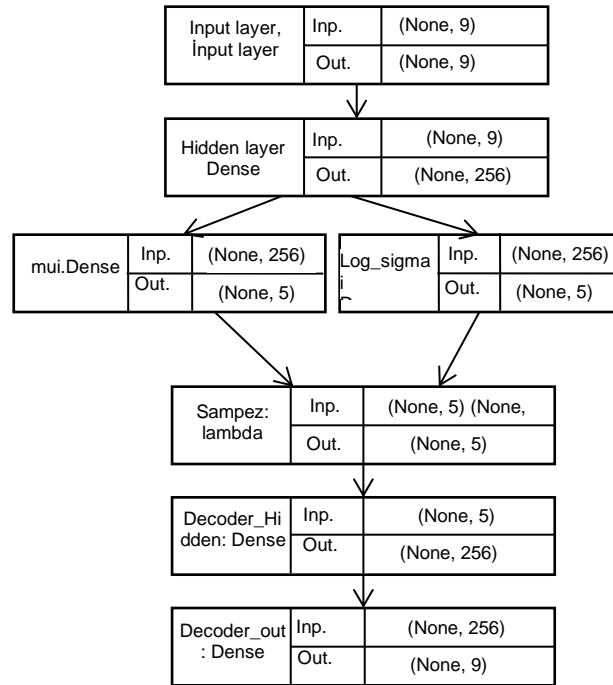
Konsepsiyada kiberhücumlardan qorunma, məlumat itkisinin və sızmasının qarşısının alınması, bulud infrastrukturunun monitorinqi və müntəzəm hesabatlılıq, trafik monitorinqi, müdaxilələrin aşkar edilməsi və qarşısının alınması kimi vacib komponentlərin tətbiq edilməsi göstərilmişdir. Hökumət buludunun yaradılması istifadəçilərin məlumatlarının qorunması, autentifikasiya və çoxfaktorlu identifikasiyanı təmin edən həllərin tətbiqini, virtual serverlərin təhlükəsizliyinin təmin olunması üçün mexanizmlərin işlənməsini zəruri edir. Hökumət buluduna keçidlə yanaşı gələcəkdə kibertəhlükəsizliyinin təmin edilməsi məqsədilə beynəlxalq standartlar əsasında milli standartların hazırlanması və tətbiqi olduqca vacibdir.

**E-dövlət** xidmətlərinin keyfiyyəti vətəndaş və dövlət arasındakı münasibətlərə təsir göstərdiyinə görə, bu istiqamət dövlət orqanları və tədqiqatçılar üçün həmişə maraqlı olmuşdur. Dövlət ilə vətəndaşlar arasında qarşılıqlı əlaqə e-dövlət xidmətlərinin təkmilləşdirilməsi üçün vacibdir. Vətəndaşlardan gələn rəylər dövlət orqanlarının e-dövlət xidmətlərini təkmilləşdirməsinə imkan yaradır. Bu işə vətəndaşların e-dövlət xidmətlərindən məmnunluq səviyyəsinin artmasına köməklik göstərir. Belə ki, e-dövlət platformasında müxtəlif xidmətlər haqqında vətəndaşlar öz fikirlərini şərh bölməsində ifadə edə bilirlər. Bu şərhləri sentiment analiz texnologiyalarının köməyi ilə analiz etməklə, vətəndaşların e-xidmətlərdən nə dərəcədə razı qaldığını müəyyən etmək olar. Bu zaman digər bir faktor da ortaya çıxa bilər. Belə ki, nüfuzlu istifadəçilərin söylədiyi fikir digər istifadəçilərdə xidmət haqqında müəyyən müsbət və ya mənfi fikir formalaşdırıa bilər. Burada nüfuzlu istifadəçi dedikdə, ideyalarına cəmiyyətdən tərəfindən hörmət bəslənən nüfuzlu şəxslər (məşhurlar, siyasi liderlər, elm adamları və s.) nəzərdə tutulur. Nüfuzlu şəxslərdən gələn mesajlar sosial media vasitəsilə minlərlə insana çatıa bilər. Adi istifadəçilərlə müqayisədə nüfuzlu istifadəçilərin mesajlarına daha çox diqqət yetirilir və əhəmiyyət verilir. Qeyd edək ki, sosial medianın meydana gəlməsindən əvvəl müəyyən bir sahədə ictimai fikrə təsir gücünə malik insanlar azlıq təşkil etdiyindən, onları asanlıqla tanımaq mümkün idi. Bu cür insanlara kütləvi informasiya vasitələrindən olan jurnalistlər, elmi dərəcələrə sahib olan şəxslər, siyasətçilər, media sahibləri, məşhurlar və s. daxil idi. Ancaq hazırda sosial mediada nüfuzlu şəxslərin sayı günü-gündən artmaqdadır və fərdlər arasında nüfuzlu olanların müəyyən edilməsi üçün bir sıra tədqiqatlar aparılmışdır.

E-dövlət portalında müxtəlif dövlət qurumları tərəfindən vətəndaşlara müxtəlif e-xidmətlər göstərilir. Göstərilən e-xidmətlərin keyfiyyətli olması və vətəndaşların tələbatını ödəməsi e-dövlətin əsas funksiyalarından biridir. E-xidmətlərin keyfiyyətini ölçməklə dövlət orqanlarının vətəndaşlar arasında nüfuzunu müəyyənləşdirmək olar. O cümlədən dövlət orqanları tərəfindən təqdim edilən çoxsaylı və zəngin informasiya resurslarından istifadə etməklə, insanların ən çox hansı problemlərdən əziyyət çəkdiyini, hansı xidmətlərə daha çox ehtiyacı olduğunu və s. müəyyənləşdirmək olar. Mövcud problemləri müəyyən edib, onları aradan qaldırmaqla vətəndaş məmnuniyyəti təmin oluna bilər. Təklif olunmuş metod vasitəsilə dövlət orqanlarının təqdim etdiyi e-xidmətlərdən vətəndaş məmnuniyyətini qiymətləndirmək və bu orqanlar arasında nüfuzlu olanları, yəni dövlət qurumlarının cəmiyyətdə reytingini təyin etmək mümkündür.

**Virtual** maşınların iş yükünün proqnozlaşdırılması sahəsində olduqca çox sayda yanaşmalar təklif olunmuşdur. Bu metodlar virtual maşınların gələcək iş yükünü əvvəlki t zaman intervalındakı keçmiş iş yükü verilənlərinə əsasən proqnozlaşdırırlar. Bu işlərdə resursların proqnozlaşdırılması SaaS və PaaS səviyyələrində aparılır. Bu səviyyələrdə bulud provayderləri infrastrukturlarında bu tip proqramların olmasından xəbərdar olurlar və onların davranışlarını asanlıqla izləyə bilirlər. Lakin IaaS səviyyəsində virtual maşınlarda hansı proqramın icra olunacağı haqqında, ayrı-ayrı virtual maşınlarda yerinə yetiriləcək proqramlar (tətbiqlər) haqqında heç bir məlumat olmur. Belə olduqda işə virtual maşınların iş yükünün və CPU sərfiyyatının qiymətləndirilməsində proqramların davranışından istifadə etmək mümkün olmur. Bu mühitdə virtual maşınların iş yükünə istifadəçilərin davranışları və istifadəçilərin ani qəbul etdiyi qərarları təsir edir. Bu səbəbdən iş yükündə fluktuasiya əvvəlki iş yükündən və

əvvəlki CPU sərfiyyatından asılı olmur və istifadəçilərin resursa olan dinamik, gözlənilməz, fluktuativ tələbatlarından asılı olaraq dramatik şəkildə dəyişilə bilər. Bu faktorları nəzərə alaraq təqdim olunan məqalədə Variasiya Avtoenkoderinə əsaslanan iş yükünün proqnozlaşdırılması üsulu təklif edilmişdir. Variasiya Avtoenkoderi ehtimal proqnozlaşdırma üsuludur, variasional çıxarışla dərin təlimin birləşməsindən yaranmışdır. Variasiya avtoenkoderinin üstünlüyü ondan ibarətdir ki, model ehtimal qiymətləndirmələr aparır. Bu ehtimal qiymətləri rekonstruksiya ehtimalları adlandırılır. Ehtimalların istifadəsi modeldə hər hansı bir spesifik sərhəd qiymətindən istifadə olunması problemini aradan qaldırır. Variasiya Avtoenkoderi bulud mühitində virtual resursların həddindən artıq yüklənmə ehtimalını qiymətləndirir. Təklif olunan iş yükünün proqnozlaşdırılması üsulunun əsasında həddindən artıq yüklənmiş (zəif işləyən) virtual maşınlar müəyyən olunur, bu maşınların ekstra iş yükünü yerinə yetirə bilən uyğun virtual maşınlar çoxluğu proqnozlaşdırılır və tapşırıqları yerinə yetirmək üçün proqnozlaşdırılan bu virtual maşınlar çoxluğu içindən tapşırıqları yerinə yetirəcək yeni virtual maşınlar seçilir. Variasiya Avtoenkoder modeli Theano kitabxanasından istifadə etməklə qurulub və aşağıdakı şəkildə təsvir edilmişdir (şəkil 8).



**Şəkil 8.** Variasiya Avtoenkoder modeli.

Variasiya Avtoenkoder modelinin arxitekturasının bir hissəsini enkoder, digər hissəsini isə dekoder təşkil edir. Şəbəkənin enkoder hissəsini 6 lay təşkil edir. Dekoder enkoderin əməliyyatına əks əməliyyatı yerinə yetirir və 6 laydan ibarətdir.

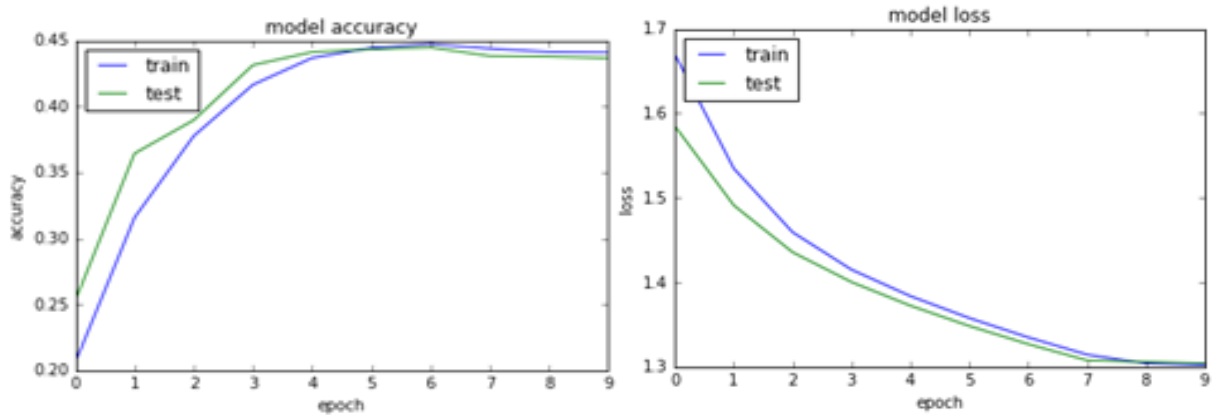
13 bulud qovşağından toplanmış 28147 nümunədən və 9 əlamətdən ibarət iş yükü verilənlər bazası istifadə edilmişdir. 1 dəqiqədə tapşırıqların sayı, 5 dəqiqədə tapşırıqların sayı, 15 dəqiqədə tapşırıqların sayı, yaddaşın tutumu, diskin tutumu, CPU nüvələrinin sayı, hər bir nüvədə CPU-nun sürəti, şəbəkənin buraxma qabiliyyətinə görə ortalama qəbul edilən trafik kilobit həcmi (Kbps), şəbəkənin buraxma qabiliyyətinə görə ortalama ötürülən trafik kilobit həcmi (Kbps) istifadə olunan verilənlər bazasının beş kateqoriyadan ibarət sinif nişanları təşkil

edir. Bu kateqoriyalar çox az (very low, VL), az (low, L), orta (medium, M), yüksək (high, H), çox yüksək (very high, VH) kimi təsnif olunur.

İş yükünün proqnozlaşdırılması variasiya avtoenkoderi və dərin neyron şəbəkəsi əsasında aparılmışdır. Burada ilk öncə neyron şəbəkənin nişanlanmış verilənlər (labeled training data) əsasında 9 iş yükü atributları üzrə öyrədilməsi aparılır, daha sonra nişanlanmamış verilənlər (unlabeled testing data) əsasında iş yükünün proqnozlaşdırılması aparılır.

Variasiya avtoenkoderində verilənlər bazasının siniflərinə uyğun 5 gizli lay istifadə edilmişdir. Modelin effektivliyinin qiymətləndirilməsi üçün precision, recall, F1-score, accuracy metrikaları istifadə edilmişdir.

Şəkil 9-dan görüldüyü kimi variasiya avtoenkoderinin iş yükünün proqnozlaşdırılması üzrə verilənlər bazasına tətbiqi zamanı itki funksiyasının dinamikasında azalma, dəqiqlik funksiyasının dinamikasında isə artım vardır (şəkil 9).



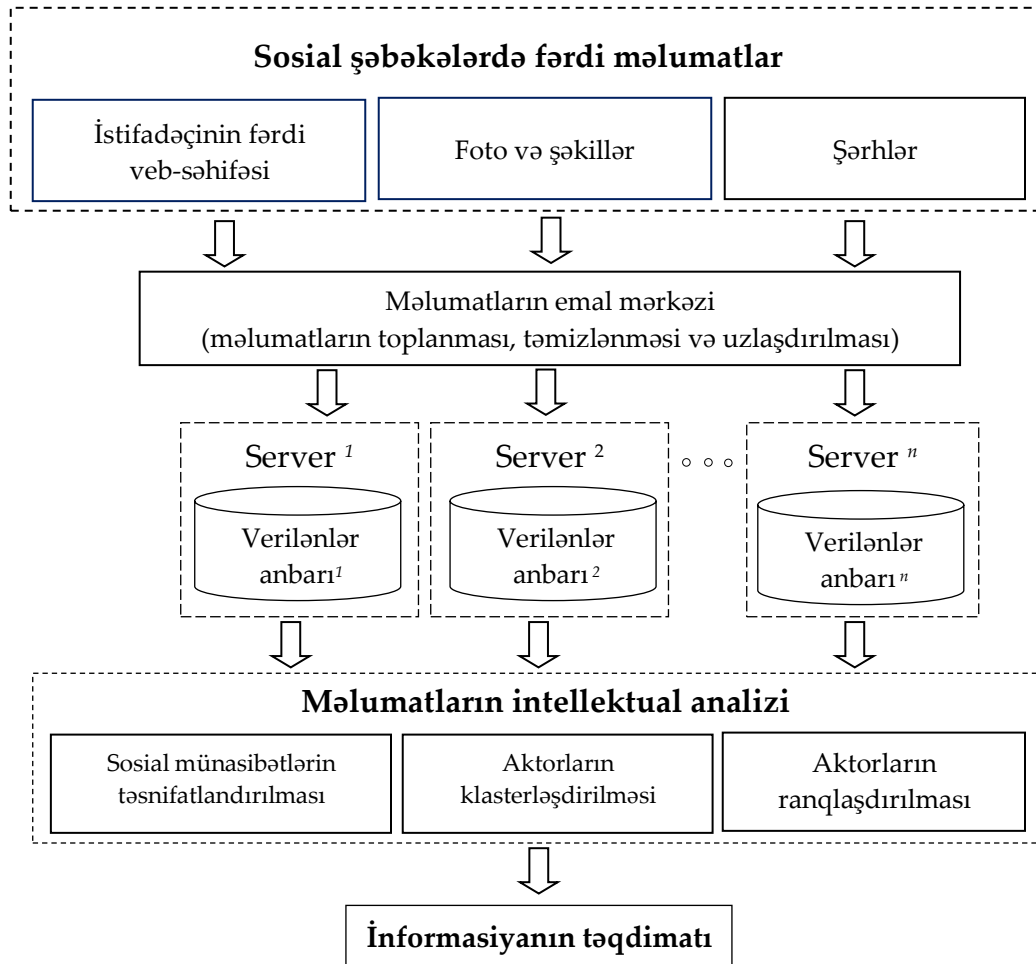
Şəkil 9. İtki və dəqiqlik funksiyalarının iterasiyaların sayından asılılığı.

Bulud resurslarında yükün proqnozlaşdırılması üsulunun test prosesi AMEA İnformasiya Texnologiyaları İnstitutunun Data mərkəzində, Ubuntu 16.04.3 LTS amd64 system, 3312 GB memory, 2933.437 CPU MHz mühitində yerinə yetirilmişdir. Eksperimentlərin aparılması üçün Python proqram paketi və Tensorflow kitabxanası istifadə edilmişdir.

**Sosial-iqtisadi** vəziyyət, doğum, ölüm, miqrasiya, terror hadisələri, müharibə kimi demoqrafiyaya təsir edən göstəricilər daimi olaraq dəyişir. Demoqrafik prosesləri onlayn araşdırmaq və izləmək üçün e-demoqrafiya sistemi qurulmalıdır. E-demoqrafiya sisteminin formalaşması üçün müxtəlif elektron dövlət reyestrləri vahid platformaya inteqrasiya olunmalıdır. Bu reyestrlər inteqrasiya olunduqdan sonra demoqrafik tədqiqatları daha dərin, effektiv tədqiq etmək mümkün olacaqdır. Demoqrafik tədqiqatlar yalnız dövlət reyestrlərindəki məlumatlarla deyil, sosial şəbəkələrdə, mobil telefonlarda, bank sistemləri və sığorta şirkətlərindəki məlumatlarla, müxtəlif axtarış brauzerlərindəki izlər vasitəsi ilə araşdırmalar aparmaq mümkündür. E-demqorafiya verilənləri əsasında demoqrafik göstəricilərin analizi məsələsinə baxılmışdır. Beynəlxalq təcrübədə e-demoqrafiya sahəsində aparılan tədqiqatlar analiz olunmuş, ədəbiyyat icmal aparılmışdır. Jupyter Notebook 6.1.4. platformasında xaricdə təhsil almış məzunların demoqrafik xarakteristikaları analiz edilmişdir. FİN vasitəsilə hər bir fərd haqqındakı reyestrlərdəki məlumatlar inteqrasiya olunaraq götürülmüş məlumatlar çoxluğu (dataset) əsasında ekperiment aparılmış və göstəricilərə görə alınmış nəticələr qrafik

şəkildə vizuallaşdırılmışdır. Bu analizlərin köməyi ilə əhalinin sosial vəziyyəti, iş şəraiti, məşğulluq səviyyəsi, gəliri və s. haqqında daha dərin tədqiqatlar aparmaq mümkündür.

Sosial şəbəkələrdə toplanan fərdi məlumatlardan və istifadəçilərin şərhindən istifadə edərək sosial münasibətlərin analizi üçün arxitektura sxem işlənmişdir. Sosial şəbəkələrdə münasibətlərin analizinin kateqoriyalar üzrə həlli sxemi işlənmişdir. Böyük verilənlərlə əlaqədar yaranan problemin səmərəli həlli üçün sosial münasibətlərin intellektual analizinin 3 mərhələ üzrə həlli təklif olunmuşdur: sosial münasibətlərin təsnifatlandırılması; aktorların klasterləşdirilməsi və aktorların rəqlaşdırılması. Sosial münasibətlərin analizi sisteminin arxitektura sxemi şəkil 10-da verilmişdir:



Şəkil 10. Sosial münasibətlərin analizi sisteminin arxitekturu.

**Videotəsvirlərin** analizi nəticəsində insanların və hadisələrin aşkarlanması çox mürəkkəb məsələdir və real zamanda geniş tətbiq sahəsinə malikdir. Müxtəlif təyinatlı videomüşahidə kameralarından əldə edilən və durmadan artan videotəsvirlərin əlçatan olması obyektlərin və hadisələrin aşkarlanmasında intellektual alqoritmlərdən istifadəni aktuallaşdırır. Videoanalitika sahəsində bəzi mövcud yanaşmalar araşdırılmış, problemlər və həlli vacib məsələlər müəyyənləşdirilmişdir. Müəyyən olunmuşdur ki, obyektin, fərdin və davranışın aşkarlanması, hadisənin qiymətləndirilməsi, anormal və ya normal kimi kateqoriyalara təsnifatlandırılması, sosial münasibətlərin müəyyən olunmasında videoanalitikada həlli vacib məsələlərdəndir.



Videotəsvirin tanınması üçün təklif olunan yanaşmaların əksəriyyətində təsnifatlandırma metodu və dərin təlim arxitekturlarına əsaslanan yanaşmalar istifadə olunur. Dərin təlimin müxtəlif arxitekturları arasında müşahidə analizi üçün ən çox istifadə olunan modellər CNN, avtoenkoderlər və onların kombinasiyasıdır.

Sosial münasibətlərin aşkarlanması məqsədi ilə videotəsvirlərin intellektual analizi sisteminin ümumi arxitektur sxemi işlənmiş və məsələnin həllində mərhələli həll yolu təklif olunmuşdur. Bu yanaşma sosial münasibətlərin analizində istifadə olunan böyük verilənləri səmərəli emal etməyə imkan yaradır. Müəyyənləşdirilmişdir ki, analiz üçün istifadə olunan verilənlər nə qədər çox olarsa nəticə də daha dəqiq alınar. Lakin "big data" dəqiq nəticə əldə etmək üçün əhəmiyyətli olsa da, böyük həcmli verilənlərin saxlanması və emalında çətinliklər yaradır. Problemin həlli üçün məqalədə verilənlər saxlanıcından istifadə olunmaqla verilənlərin paylanmış və paralel emalı təklif olunmuşdur.

Sosial münasibətlər çox mürəkkəb sosial şəbəkədir və sosial münasibətlər şəbəkəsinin təsnifatlandırılaraq iş münasibətləri şəbəkəsi, qohumluq münasibətləri şəbəkəsi və dostluq münasibətləri şəbəkəsi kimi hər birinin ayrı-ayrılıqda analizi e-xidmətlərin daha səmərəli həyata keçirilməsi üçün yeni imkanlar yaradacaqdır. Sosial münasibətlərin intellektual VMS vasitəsilə aşkarlanması sosial media vasitələrindən fərqli olaraq yalnız virtual dostları deyil, fiziki dünyada insanları və qrupları aşkarlamağa, obyektləri tanımağa, baş verəcək hadisələri əvvəlcədən proqnozlaşdırmağa imkan yaradır. Ayrı-ayrı fərdlər arasında sosial bağları müəyyən etmək və proqnozlaşdırmaq cəmiyyəti müəyyən etmək deməkdir. Burada əsas hədəf cəmiyyətin deşifrə olunmasıdır. Cəmiyyətdə gizli münasibətləri aşkarlamaq üçün sensorlarla təchiz olunmuş qurğulardan əldə olunan məlumatların emalı tələb olunur. Bu məlumatların analizində ehtimal qrafından istifadə daha məqsədəuyğundur. Ehtimal qrafına əsasən fərdlər arasında münasibətlərin tipini və xüsusiyyətini müəyyən etmək mümkündür ki, bu məsələyə gələcək tədqiqatlarda baxılması nəzərdə tutulur.

Tədqiqatda əldə edilən nəticələr e-dövlətin daha səmərəli idarə olunması, sosial-iqtisadi proseslərin proqnozlaşdırılması, vətəndaşların təhlükəsizliyinin təmini və bir çox sahələrdə istifadə oluna bilər.

**Big Data**-nın mühüm mənbəyi olan açıq dövlət məlumatlarının mahiyyəti, xüsusiyyətləri, dəyərsənətmə potensialı araşdırılmışdır. Müəyyən edilmişdir ki, açıq dövlət məlumatlarının həcmi (dövlət strukturlarında böyük həcmdə məlumat toplanır), sürətini (dövlət strukturlarında hər gün məlumat toplanır, bu prosesdə sensorlardan da geniş istifadə olunur), müxtəlifliyini (dövlət məlumatları bütün fəaliyyət sahələrini əhatə edir) və mötəbirliyini (rəsmi dövlət məlumatları etibarlı mənbə kimi qəbul edilir) nəzərə alaraq, onları Big Data-ya aid etmək olar. Açıq dövlət məlumatları sahəsində beynəlxalq təşəbbüslər, qabaqcıl ölkələrin, eyni zamanda, Azərbaycanın atdığı addımlar, həyata keçirdiyi tədbirlər analiz edilmişdir. Açıq dövlət məlumatlarının maraqlı olan tərəflərin – dövlətin, biznes və vətəndaş cəmiyyəti qurumlarının, medianın, tədqiqatçıların həmin aktivlərin dəyərə çevrilməsində rolu müəyyən edilmişdir. Həmçinin müvafiq məlumatların dəyərə çevrilməsinə təsir göstərən texniki, iqtisadi, sosial-mədəni və siyasi amillər şərh edilmişdir. Eləcə də, müvafiq metodların analizi əsasında açıq dövlət məlumatlarının dəyərə çevrilməsinin qiymətləndirilməsi meyarları müəyyən olunmuşdur. Eyni zamanda, açıq dövlət məlumatlarının dəyərə çevrilməsinə əngəl olan

problemlər müəyyən edilmiş və onların həllinə dair təkliflər işlənmişdir. Məqalədə əldə edilən elmi nəticələr Azərbaycanda açıq dövlət təşəbbüslərinin effektiv şəkildə həyata keçirilməsi, eləcə də, maraqlı tərəflərin bu məlumatları dəyərə çevirmələri üçün müvafiq strategiyaların hazırlanması və reallaşdırılmasında istifadə edilə bilər.

Açıq dövlət məlumatlarının dəyərə çevrilməsi ilə bağlı problemlərin multidisiplinar yanaşma əsasında kompleks və əlaqəli şəkildə öyrənilməsi və onların aradan qaldırılması üçün təkliflərin işlənməsi tədqiqatın elmi yeniliyini təşkil edir.

**Bulud** mühitində tətbiqi səviyyə DDOS hücumlarının effektiv aşkarlanmasını həyata keçirmək üçün klassifikasiya modeli çoxsaylı verilənlər üzərində test edilməlidir. Eksperimentlərin aparılması üçün CSIC2010 və CSE-CIC-IDS2018 verilənlər bazaları istifadə edilmişdir. CSIC2010 verilənlər bazası 19 əlamətdən, 71484 sətirdən və 2 sinifdən ibarətdir. Modelin aktivləşmə funksiyası softmax, itki funksiyası sparse\_categorical\_crossentropy, optimallaşdırma funksiyası kimi Adam istifadə edilmişdir. CSE-CIC-IDS2018 verilənlər bazası Kanada Kibertəhlükəsizlik İnstitutuna (Canadian Institute of Cybersecurity) məxsusdur. Bu verilənlər bazası istifadəçilərin HTTPS, HTTP, SMTP, POP3, IMAP, SSH və FTP protokolları ilə etdikləri sorğuların formalaşdırdığı davranışlar əsasında yaradılmışdır.

CSE-CIC-IDS2018 verilənlər bazasını 0- Benign, 1- DDoS attack, 2- DoS attack adlı üç sinif təşkil edir və sətirlərin sayı 1048574-dir.

CSIC 2010 və CSE-CIC-IDS2018 verilənlər bazaları üzərində qurulmuş CNN modellərinin strukturu aşağıdakı şəkil 11-də təsvir edilmişdir.

Layer (type)	Output Shape	Param #	Layer (type)	Output Shape	Param #
conv1d_3 (Conv1D)	(None, 18, 64)	192	conv1d_2 (Conv1D)	(None, 75, 64)	192
dense_5 (Dense)	(None, 18, 16)	1040	dense_3 (Dense)	(None, 75, 16)	1040
max_pooling1d_3 (MaxPooling1D)	(None, 9, 16)	0	max_pooling1d_2 (MaxPooling1D)	(None, 37, 16)	0
flatten_3 (Flatten)	(None, 144)	0	flatten_2 (Flatten)	(None, 592)	0
dense_6 (Dense)	(None, 2)	290	dense_4 (Dense)	(None, 3)	1779
Total params: 1,522			Total params: 3,011		
Trainable params: 1,522			Trainable params: 3,011		
Non-trainable params: 0			Non-trainable params: 0		
CSIC 2010 verilənləri üzərində qurulmuş CNN modelinin strukturu			CSE-CIC-IDS2018 verilənləri üzərində qurulmuş CNN modelinin strukturu		

Şəkil 11. CNN modellərinin strukturu.

Burada metodun effektivliyi Accuracy, Precision, Recall və F1-score metrikaları əsasında qiymətləndirilmiş və nəticələr cədvəl 9-da verilmişdir.

**Cədvəl 9.** İki verilənlər çoxluğu üzərində metodun nəticələri.

Dataset	Class	Accuracy	Precision	Recall	F1-score
CSE-CIC-IDS2018 DDoS	Benign (0)	0.9853	0.9923	0.9921	0.9921
	DDoS (1)	0.9999	0.9999	0.9999	0.9999
	DoS (2)	0.9931	0.9903	0.9901	0.9944
CSIC 2010	Anomalous Traffic (0)	0.6331	0.9201	0.6311	0.7531
	Normal Traffic (1)	0.9911	0.9121	0.9924	0.9521

Cədvəldən görüldüyü kimi model CSIC 2010 dataset verilənləri üzərində test edilən zaman Anomalous Traffic (0) sinfindən olan verilənləri aşağı effektivliklə aşkarlaya bilmişdir. Belə ki, accuracy, precision, recall və F1-score metrikaları üzrə model uyğun olaraq 0.63, 0.92, 0.63, 0.75 qiymətlərini almışdır. Normal Traffic (1) verilənlərinin tanınmasında isə model yaxşı nəticələr göstərmişdir və həmin metrikalar üzrə uyğun olaraq 0.99, 0.91, 0.99, 0.95 qiymətlərini almışdır. Modelin CSE-CIC-IDS2018 DDoS dataset verilənləri üzərində test edilməsi zamanı bütün metrikalar üzrə demək olar ki yüksək nəticələr alınmışdır. CNN modelinin ayrı-ayrı verilənlər üzərində fərqli nəticələr göstərməsi verilənlərin həcmindən birbaşa asılıdır. Nəticələrdən görüldüyü kimi CNN modeli böyük həcmli verilənlərdə daha yaxşı işləyir. Bu dərin təlim modelindən gözlənilən nəticədir. Dərin təlim modellərinin xüsusiyyəti odur ki, onlar böyük həcmli verilənlərlə daha yaxşı işləyirlər, kiçik həcmli verilənlərlə işlədikdə isə yalnız nəticələrə daha çox yol verirlər.

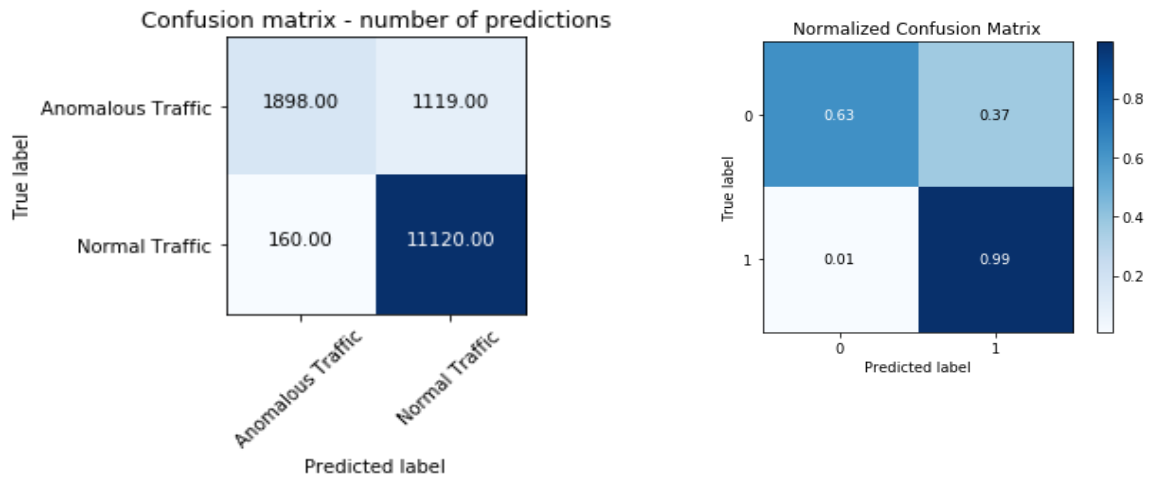
CNN modelinin arxitekturasında itki funksiyası olaraq “sparse\_categorical\_crossentropy” istifadə edilmişdir. Aparılan eksperimentlərdə CNN modeli həm CSE-CIC-IDS2018 DDoS dataset verilənlər bazası həm də CSIC 2010 dataset verilənlər bazası üzərində az itkiyə yol vermişdir və uyğun olaraq 0.0143 və 0.2141 qiymətlərini almışdır və nəticələr cədvəl 10-da verilmişdir.

**Cədvəl 10.** İki verilənlər çoxluğu üzərində metodun dəqiqliyi və itkisi.

Dataset	Accuracy	Loss
CSE-CIC-IDS2018 DDoS	0.9974	0.0143
CSIC 2010	0.9059	0.2141

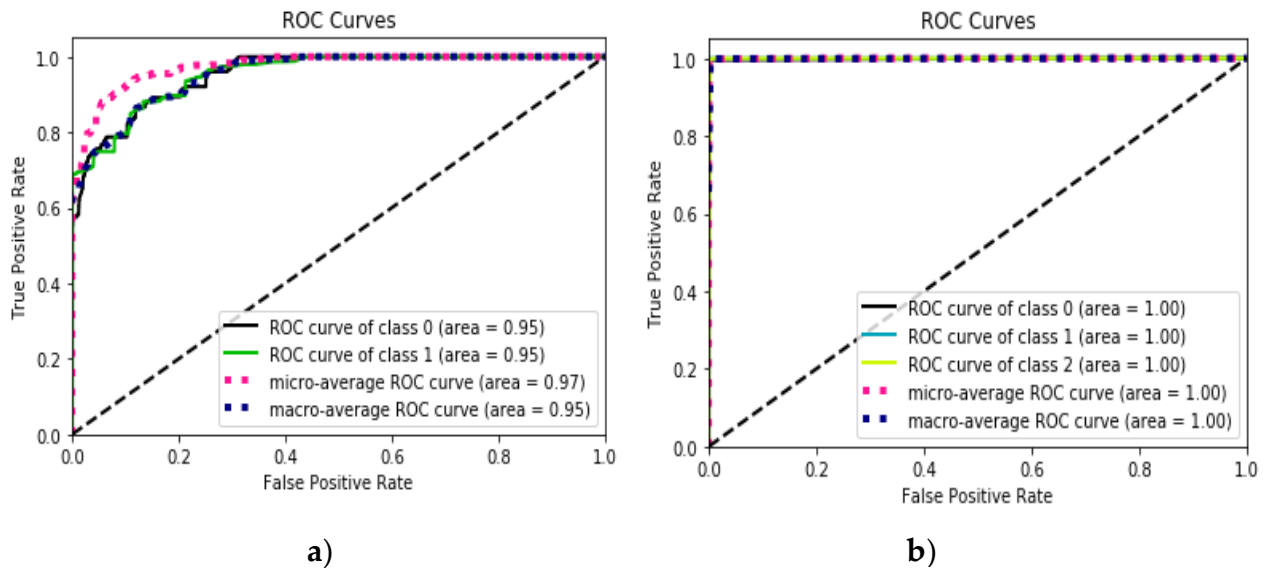
Burada model CSE-CIC-IDS2018 DDoS dataset verilənlər bazasındakı verilənləri 0.9974 dəqiqliklə, CSIC 2010 dataset verilənlər bazasındakı verilənləri isə 0.9059 dəqiqliklə tanıya bilmişdir.

Qurulmuş CNN modeli Anomalous Traffic (0) sinfindən olan 3017 nöqtədən 1898 sayda nöqtəni doğru tanımışdır, 1119 sayda nöqtəni yalnız tanıyaraq Normal Traffic (1) sinfinə aid etmişdir. Alqoritm eyni zamanda Normal Traffic (1) sinfindən olan 11280 sayda nöqtədən 11120 nöqtəni düzgün olaraq Normal Traffic (1) sinfinə aid etmişdir, 160 nöqtəni isə səhv tanıyaraq Anomalous Traffic (0) sinfinə daxil etmişdir. Bu proqnozlaşdırmanın xətalər matrisi şəkil 12-də təsvir edilmişdir.



Şəkil 12. Proqnozlaşdırmanın xətalər matrisi.

CNN modelinin aldığı nəticələri əyani təsvir etmək üçün onun ROC əyrisi çıxarılmışdır və şəkil 13-də təsvir edilmişdir.



Şəkil 13. CNN modelinin CSIC 2010 dataset (a) və CSE-CIC-IDS2018 DDoS dataset (b) verilənlər bazaları üzərində ROC əyrisi

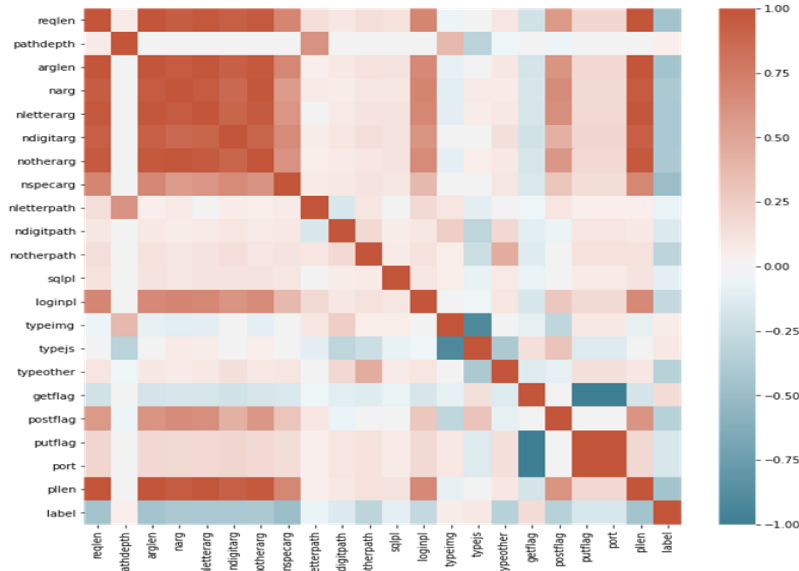
Şəkil 13-dən görüldüyü kimi modelin hər bir sinif üzrə aldığı nəticələr 1-ə çox yaxınlaşmışdır.

Bulud xidmətlərinə olan kiberhüumların aşkarlanmasına həsr olunmuş çox sayda yanaşmalar təklif edilmişdir. Mövcud yanaşmalarda kiberhüumların aşkarlanması zamanı toplanmış verilənlərdəki əlamətlərin vaciblik dərəcəsi nəzərə alınmır. Metodlar hücumların aşkarlanmasının test prosesinə verilənlərin bütün əlamətlərini cəlb edirlər. Bu alqoritmin effektivliyinin azalmasına və alınmış nəticələrin pisləşməsinə səbəb olur. Bu problemi aradan qaldırmaq üçün verilənlər üzərində əlamətlərin seçilməsi əməliyyatı aparılmalıdır.

Təklif edilmiş yanaşma "CSE-CIC-IDS2018" və "HTTP dataset CSIC 2010" verilənləri üzərində test edilmişdir. "HTTP dataset CSIC 2010" verilənlər bazası 19 əlamətdən, 71484 sətirdən və 2 sinifdən ibarətdir. CSE-CIC-IDS2018 verilənlər bazası Kanada Kibertəhlükəsizlik

İnstitutuna (Canadian Institute of Cybersecurity) məxsusdur. Bu verilənlər bazası istifadəçilərin HTTPS, HTTP, SMTP, POP3, IMAP, SSH və FTP protokolları ilə etdikləri sorğuların formalaşdırdığı davranışlar əsasında yaradılmışdır. CSE-CIC-IDS2018 verilənlər bazasını 0-Benign, 1- DDoS attack, 2- DoS attack adlı üç sinif təşkil edir və sətirlərin sayı 1048574-dir.

“HTTP CSIC 2010” verilənlər bazasının əlamətləri arasında qurulmuş korrelyasiya matrisi şəkil 14-də verilmişdir.



Şəkil 14. Korrelyasiya matrisi.

Matrisin sətir və sütunlarını verilənlər bazasının əlamətləri təşkil edir. Burada hər kəsişmədə konkret əlamətlər cütlüyünün korrelyasiyası əks olunmuşdur. Matrisin hər bir xanasındakı rəng bir əlamətin digər əlamətlə korrelyasiya dərəcəsini göstərir. Matrisin diaqonal elementləri hər bir əlamətin özü ilə korrelyasiyasını göstərir və onların qiymətləri hər zaman vahidə bərabərdir. Matrisin açıq rəngli xanaları həmin əlamətlər arasında korrelyasiyanın az olduğunu göstərir. Yuxarıdakı şəkildə 0.75-dən yuxarı qiymət əlamətlər arasında yüksək korrelyasiyanın olduğunu göstərir. Korrelyasiya olunmuş əlamətlər çoxluğunu tamaq üçün 0.75 sərhəd qiyməti təyin edilmişdir bu qiymətdən yuxarı qiymətlərlə korrelyasiya olunan əlamətlər tapılmışdır. Əlamətlər arasında korrelyasiyanı tapmaq üçün Pearson korrelyasiya üsulu tətbiq edilmişdir. Korrelyasiyanın tapılması üçün Kendall və Spearman üsulları da istifadə oluna bilər. Yuxarıdakı şəkildən görüldüyü kimi *reglen* adlı əlamət verilənlər bazasının digər əlamətləri ilə daha çox korrelyasiyaya malikdir. *reglen* əlamətinin ən sıx korrelyasiyada olduğu əlamətlər aşağıdakılardır:

**reglen 1.0**

arglen 0.9940151121065613

narg 0.9511267953115554

nletterarg 0.9817931328235973

ndigitarg 0.931807728802511

notherarg 0.9669999754721563

pllen 0.9937697104142604



Verilənlər bazasına Pirson korrelyasiya üsulunu tətbiq etdikdə 21 əlamətdən 10 əlamət arasında korrelyasiya tapılmışdır və onlar silinmişdir. Beləliklə verilənlər bazasında 11 əlamət saxlanılmışdır.

Yanaşmanın “HTTP CSIC 2010” verilənlər bazasına tətbiqinin təsnifat nəticələri cədvəl 11-də verilmişdir.

**Cədvəl 11.** Təsnifat alqoritmlərinin HTTP CSIC 2010 verilənlər çoxluğu üzərində nəticələri.

Üsullar	Verilənlər	Orijinal verilənlər ROC-AUC	Korrelyasiya verilənləri ROC-AUC
Random Forest	Train set	0.9157	0.9301
	Test set	0.9130	0.9301
Logistic regression	Train set	0.9261	0.9929
	Test set	0.9294	0.9933

Cədvəldən göründüyü kimi Random Forest alqoritm korrelyasiya üsulunun tətbiqindən sonra formalaşmış verilənlər üzərində test etmə ilə müqayisədə orijinal verilənlər üzərində test edildiyi zaman nisbətən zəif nəticələr göstərmişdir. Beləliklə alqoritm orijinal verilənlər üzərində test etdikdə ROC metrikası 0.9157 qiymətini almışdır, yeni formalaşmış verilənlər üzərində bu qiymət 0.9301 olmuşdur. Logistik reqressiya alqoritm də yeni formalaşmış verilənlər üzərində daha yaxşı nəticələr göstərmişdir. Belə ki, alqoritm orijinal verilənlər üzərində test edilən zaman ROC qiyməti 0.9261 olduğu halda, yeni verilənlər üzərində 0.9929 qiymətinə yüksəlmişdir. Bu gözlənilən nəticədir.

Yanaşmanın effektivliyi “CSE-CIC-IDS2018” verilənlər bazasının üzərində də test edilmişdir.

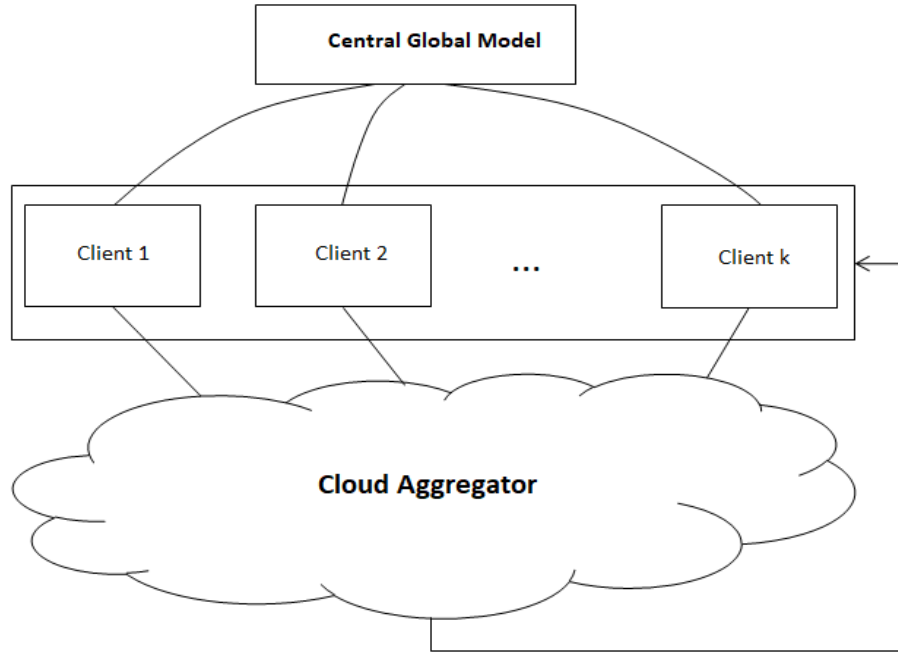
Random Forest alqoritm korrelyasiya üsulunun tətbiqi ilə yeni formalaşmış verilənlər üzərində klassifikasiya nəticələri cədvəl 12-də verilmişdir.

**Cədvəl 12.** Random Forest alqoritm korrelyasiyanın tətbiqindən sonrakı verilənlər çoxluğu üzərindən nəticələri.

Üsul	Verilənlər bazasının sinifləri	Precision	Recall	F1-score
Random Forest	Class 1	0.9999	0.9999	0.9999
	Class 2	1.0000	1.0000	1.0000
	Class 3	0.9999	0.9999	0.9999

Alqoritm effektivliyi Precision, Recall, F1-score metrikaları əsasında qiymətləndirilmişdir. Cədvəldən göründüyü kimi alqoritm hər üç sinifdən olan nümunələri dəqiqliklə tanıya bilmiş və bütün metrikalar üzrə yüksək nəticələr göstərmişdir.

Bulud texnologiyaları mühitində federativ təlim modeli şəkil 15-dəki kimi qurulmuşdur:



Şəkil 15. Federativ təlim modeli

Federativ təlim – neyron şəbəkənin çox sayda qurğularda öyrədilməsi metodudur. Bu hesablama modelində vahid qlobal neyron şəbəkə mərkəzi serverdə saxlanır. Neyron şəbəkəni öyrətmək üçün istifadə olunan verilənlər çoxsaylı heterogen qovşaqlarda lokal saxlanır.

Federativ təlimin iş prinsipi aşağıdakı kimidir:

Hər dövrün başlanğıcında kliyentlərin təsadüfi  $C$  hissəsi seçilir, server cari modelin parametrlərini seçilmiş kliyentlərə göndərir.

Ənənəvi dərin təlim modelinin öyrədilməsində  $n$  sayda nümunədən ibarət  $(x_i, y_i)$ ,  $1 \leq i \leq n$  təlim verilənləri üçün öyrənmə məqsəd funksiyası aşağıdakı kimi təyin edilir:

$$\min_{w \in R^d} f(w)$$

$$f(w) \stackrel{\text{def}}{=} \frac{1}{n} \sum_{i=1}^n f_i(w),$$

burada  $f_i(w) = l(x_i, y_i, w) - (x_i, y_i)$  nümunəsi üçün itki funksiyasıdır.

Dərin təlimdə optimallaşma SGD (stochastic gradient descent – stoxastik qradiyent enmə) üsuluna və onun variantlarına əsaslanır:

$$w_{t+1} \leftarrow w_t - \eta \nabla f(w_t, x_k, y_k)$$

Federativ təlimdə fərz edilir ki,  $n$  sayda təlim nümunələri  $k$  sayda kliyentə paylanmışdır və  $n_k = |P_k|$ , burada  $P_k$   $k$ -cı kliyentdə olan nümunələr çoxluğudur.

Bu modeldə təlim məqsəd funksiyası belə təyin edilir:

$$f(w) = \sum_{k=1}^K \frac{n_k}{n} F_k(w),$$

burada  $F_k(w) \stackrel{\text{def}}{=} \frac{1}{n_k} \sum_{i \in P_k} f_i(w)$ .

Çəkileri aqreqasiya etmək və qlobal modeli yeniləmək üçün ortalama funksiyasından istifadə edilmişdir.

Fərz edək ki,  $\eta$  – modelin öyrənmə sürəti,  $n$  – nümunələrin sayıdır,  $K$  – kliyentlər çoxluğu,  $n_k$  –  $k$ -cı kliyentdə nümunələrin sayıdır,  $C = 1$  kliyentlərin fraksiyasıdır.

$t$ -ci dövrdə mərkəzi server cari modelin  $w_t$  çəkisini hər bir kliyentə paylayır. Hər bir  $k$ -cı kliyent öz lokal verilənləri üzərində qradiyenti aşağıdakı kimi hesablayır:

$$g_k = \nabla F_k(w_k)$$

Hər bir  $k$ -cı kliyent  $k$ -cı qradiyenti mərkəzi serverə göndərir. Mərkəzi server yeni model yaratmaq üçün qradiyentləri aqreqasiya edir:

$$w_{t+1} \leftarrow w_t - \eta \nabla f(w_t) = w_t - \eta \sum_{k=1}^K \frac{n_k}{n} g_k$$

FedAvg halında hər bir  $k$ -cı kliyent  $w_{t+1}^k \leftarrow w_t - \eta \cdot g_k$  çəkisini hesablayır, mərkəzi server aqreqasiyanı aşağıdakı kimi həyata keçirir:

$$w_{t+1} \leftarrow \sum_{k=1}^K \frac{n_k}{n} w_{t+1}^k$$

Təklif edilmiş yanaşma “CSE-CIC-IDS2018” verilənləri üzərində test edilmişdir. CSE-CIC-IDS2018 verilənlər bazası Kanada Kibertəhlükəsizlik İnstitutuna (Canadian Institute of Cybersecurity) məxsusdur. Bu verilənlər bazası istifadəçilərin HTTPS, HTTP, SMTP, POP3, IMAP, SSH və FTP protokolları ilə etdikləri sorğuların formalaşdırdığı davranışlar əsasında yaradılmışdır. CSE-CIC-IDS2018 verilənlər bazasını 0 – Benign, 1 – DDoS attack, 2 – DoS attack adlı üç sinif təşkil edir və sətirlərin sayı 1048574-dir.

Bulud tenologiyalarını hədəf alan kiberhücumların aşkarlanması üçün təklif edilmiş federativ təlim yanaşmasının “CSE-CIC-IDS2018” verilənlər bazasının üzərində test nəticələri cədvəl 13-də verilmişdir.

**Cədvəl 13.** Federativ təlim modelinin nəticələri

Local Machines	Loss	Accuracy (%)
Client 4	0.1908	96
Client 5	0.0938	96
Client 9	0.2044	96
Client 2	0.3006	96
Client 1	0.0791	96
Client 8	0.2780	96
Client 10	0.0935	96
Client 3	0.1124	96
Client 7	0.1834	96
Client 6	0.0452	96
<b>FedAvg</b>	<b>0.3613</b>	<b>97</b>

Cədvəldən göründüyü kimi FedAvg modeli aşkarlama dəqiqliyində ayrı-ayrı kliyentlərin aldığı nəticələrlə müqayisədə uduş əldə edərək Loss və Accuracy metrikaları üzrə uyğun olaraq

0.3613 və 97% qiymətlərini almışdır. Burada verilənlərin analizi ayrı-ayrı kliyentlərdə lokal emal olunduğu üçün onların konfidensiallığının təmin olunmasına nail olunmuşdur.

E-demoqrafiya sisteminin vahid dövlət reyestri əsasında qurulması təklif olunur. Bu halda bütün demoqrafik tədqiqatlar üçün əsas mənbələr olan dövlət reyestrləri, verilənlər bazaları və veb-portallar e-demoqrafiya sisteminə transfer olunmalıdır. Bu tədqiqat işində eksperimental qiymətləndirmə üçün vahid reyestrə inteqrasiya olunmuş xaricdə təhsil almış məzunların məlumatları əsasında demoqrafik təhlilər aparılmışdır.

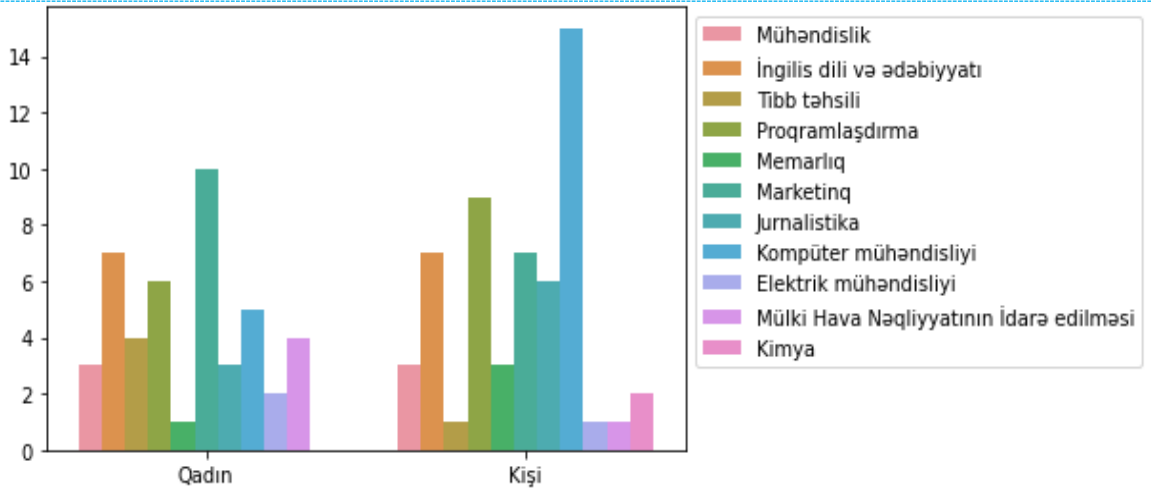
Xaricdə təhsil almış məzunların demoqrafik göstəricilərinə – yaş, cins, ailə vəziyyət və digər xarakteristikalar aid edilir. Bu göstəricilər hipotetik verilənlərdir və eksperiment üçün istifadə olunmuşdur. Eksperimentin aparılması üçün məzunlar haqqındakı məlumatlar Kaggle bazasından götürülmüşdür. Məzunların məlumatları Python proqramlaşdırma dili vasitəsi ilə analiz olunmuşdur. Analizlər Jupyter Notebook 6.1.4. mühitində realizasiya olunmuşdur.

Eksperimental olaraq məzunlar haqqında məlumatlar iki ayrı-ayrı reyestrədə (əhali və təhsil reyestrləri) saxlandığı nəzərdə tutulur. Vahid reyestrə inteqrasiya olunan 100 məzunla bağlı dataset götürülmüşdür. Bu dataset-lərdən məzunların fərdi məlumatları və təhsil haqqında olan məlumatları toplanmışdır. Bu dataset-də hər bir şəxsin fərdi identifikasiya nömrəsi, cins, yaş, vəzifəsi, iş stajı, ailə vəziyyəti, təhsil səviyyəsi, təhsil aldığı ixtisas, təhsil aldığı ölkə və təhsil müddəti mövcuddur. Şəkil 16-da fərdi identifikasiya kodu (PIN) vasitəsi ilə hər iki dataset inteqrasiya olunmuşdur. Burada hər bir məzunun PIN vasitəsi ilə eyni zamanda həm təhsil, həm də fərdi məlumatları əldə etmək mümkündür.

	FIN	CİNS	YAŞ	VƏZİFƏ	STAJ	AİLƏ_VƏZİYYƏTİ	TƏHSİL_SƏVİYYƏSİ	İXTİSAS	TƏHSİL_ALDIĞI_ÖLKƏ	TƏHSİL_MÜDDƏTİ
1	458994	Kişi	23	Tibb işçisi	2	Evli	Bakalavr	İngilis dili və ədəbiyyatı	Türkiyə	4
2	458996	Qadın	24	Aktyor	2	Evli	Magistr	Tibb təhsili	Rusiya	2
3	459000	Kişi	25	İdarəedici	2	Evli	Magistr	Proqramlaşdırma	İtaliya	2
4	458989	Qadın	22	Mühəndis	1	Evli	Bakalavr	Mühəndislik	Çin	4
5	459001	Qadın	24	Hüquqşünas	3	Subay	Bakalavr	Tibb təhsili	Türkiyə	4

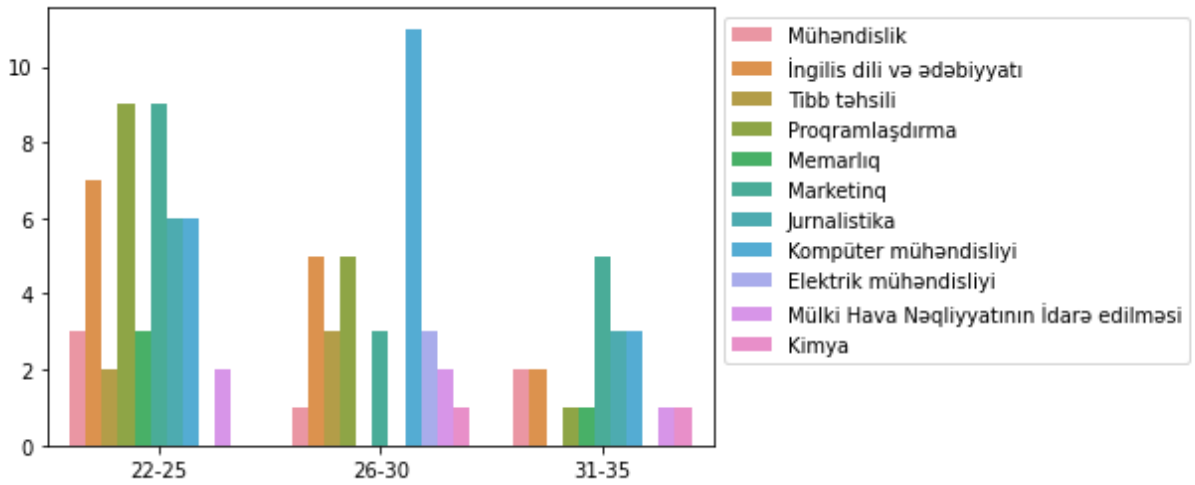
**Şəkil 16.** Məzunlar haqqında məlumatlar

Analiz nəticələri göstərir ki, məzunların təhsil aldığı ixtisaslara və cinsə görə bölgüsü Şəkil 17-də göstərilmişdir. Bu diaqramda qadınların ən çox “Marketing” ixtisası üzrə, kişilərin isə ən çox “Kompüter mühəndisliyi” ixtisası üzrə təhsil aldığı görünür.



Şəkil 17. Məzunların ixtisas seçiminə görə bölgüsü

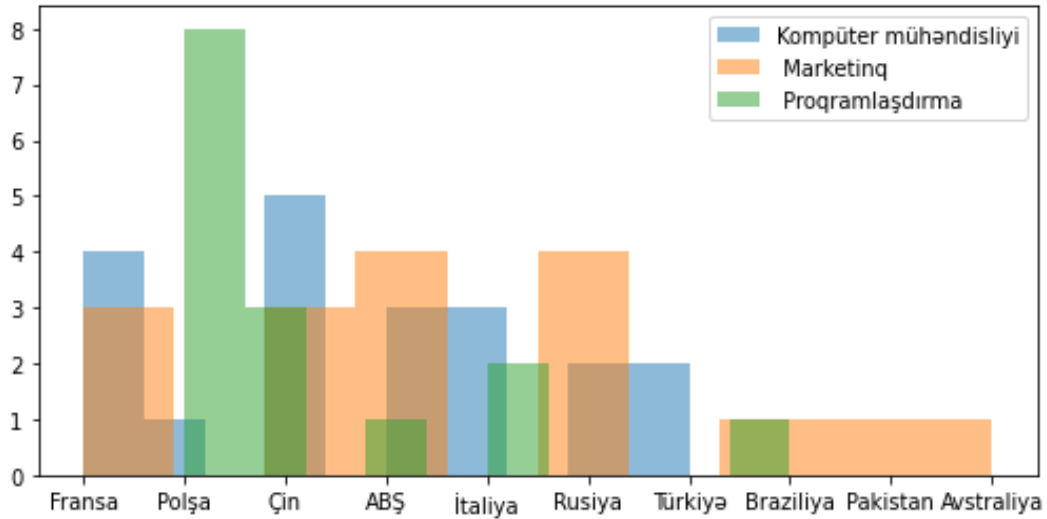
Şəkil 18-də təhsil alan məzunların müxtəlif yaş qruplarına görə bölgüsü verilmişdir. Diaqramdan da görüldüyü kimi 22-25 yaş aralığında olan məzunlar daha çox Proqramlaşdırma və Marketing, 26-30 yaş aralığında Kompüter mühəndisliyi, 31-35 yaş aralığında isə Marketing ixtisası üzrə təhsilini tamamlamışlar. Bu göstəricilərə görə demək olar ki, məzunlar ən çox Kompüter mühəndisliyi, Proqramlaşdırma və Marketing ixtisasları üzrə təhsil almışlar.



Şəkil 18. Məzunların ixtisas seçiminə görə yaş aralığı

Şəkil 8-dəki analiz nəticələrindən istifadə etməklə Kompüter mühəndisliyi, Proqramlaşdırma və Marketing ixtisaslarında təhsil almış məzunların hansı ölkələrdə təhsil aldıklarını analiz edilmişdir. Proqramlaşdırma ixtisası üzrə məzunların ən çoxu Polşada, Kompüter mühəndisliyi ixtisası üzrə Çində, Marketing ixtisası üzrə isə ABŞ və Rusiyada təhsil almışlar. Məzunların ən çox təhsil aldıkları ixtisaslara görə ölkələr üzrə paylanması şəkil 19-da göstərilmişdir.





Şəkil 19. Məzunların ixtisas seçiminə görə ölkələr üzrə paylanması

Aparılan tədqiqatlar göstərir ki, demoqrafik tədqiqatlar üçün əhali reyetrlərinin verilənlərindən istifadə olunsa da, müəyyən zaman intervalında reyestr insanın həyat dövrü, baş verən hadisələr haqqında bioqrafik verilənlər təqdim etsə də bu məlumatlar olduqca məhduddur. Bu baxımdan effektiv demoqrafik tədqiqatların aparılması üçün yalnız əhali reyestri deyil digər dövlət reyestrlərin analiz olunması üçün vahid dövlət reyestri əsasında e-demoqrafiya sisteminin yaradılması olduqca aktual məsələdir. Təcrübə göstərir ki, hazırda əhali reyetrlərinin verilənlərindən demoqrafik tədqiqatlarda istifadə olunsa da digər dövlət reyestrlərin verilənlərindən istifadəyə və e-demoqrafiya sisteminin yaradılmasına dair yanaşmalar mövcud deyil. Tədqiqat işində e-demoqrafiya sisteminin vahid dövlət reyestri əsasında qurulması təklif olunur. Hesab edirik ki, söhbət e-dövlət platformasından gedirsə bütün dövlət reyestrləri inteqrasiya olunmalıdır və reyestrədə toplanan bütün verilənlər demoqrafik tədqiqatlar üçün əhəmiyyətli mənbə rolunu oynayır. Baxılan yanaşmada vahid dövlət reyestri əsasında e-demoqrafiya sistemi yaradılır və e-dövlətin funksiyalarından asılı olmayaraq müstəqil fəaliyyət göstərir. E-demoqrafiya sisteminin yaradılması demoqrafik prosesləri daha yaxşı başa düşməyə, daha dərin tədqiqatların aparılmasını sosio-demoqrafiyadan əhali coğrafiyasına, miqrasiya proseslərindən tibbi demoqrafiyaya qədər müxtəlif səviyyələrdə, kəsimlərdə analizlərin aparılması potensialına malikdir.

Tədqiqat işində eksperimental qiymətləndirmə üçün vahid reyestrə inteqrasiya olunmuş fərdlərin məlumatları əsasında COVID-19 pandemiyası şəraitində demoqrafik xarakteristikaların analizi aparılmışdır.

Fərdlərin demoqrafik göstəricilərinə - yaş, cins, vəzifəsi, ailə üzvlərinin sayı, milliyəti və digər xarakteristikalar aid edilir. Tədqiqat işində bu göstəricilər hipotetik verilənlərdir və eksperiment üçün istifadə olunmuşdur. Şəxslərin məlumatları Python proqramlaşdırma dili vasitəsi ilə analiz olunmuşdur. Analizlər Jupyter Notebook 6.1.4. mühitində realizasiya olunmuşdur.

Tutaq ki, fərdlər haqqında məlumatlar iki ayrı-ayrı reyestrədə (əhali və səhiyyə reyestrləri) saxlanılır. Vahid reyestrə inteqrasiya olunan 1000 fərddə bağlı dataset götürülmüşdür. Bu dataset-lərdən əhalinin fərdi məlumatları və sağlamlıq haqqında olan məlumatları toplanmışdır.

Bu dataset-də hər bir şəxsin fərdi identifikasiya nömrəsi, cins, yaş, vəzifəsi, ailə üzvlərinin sayı, milliyyəti, xronik xəstəliyinin, o cümlədən, şəkərli diabetinin, bronxial astmanın olması, COVID-19-a yoluxma, vəfat edən, sağalan fərdlər haqqında məlumatlar mövcuddur. Şəkil 20-də əhalinin fərdi məlumatları göstərilmişdir. Burada şəxs haqqında 6 kateqoriyada fərdi məlumatlar verilmişdir.

PIN	Gender	Age	Profession	Family_size	Nationality
100213	Man	59	Engineer	1	Azerbaijanis
100214	Woman	56	Healthcare	4	Talyshs
100215	Woman	41	Marketing	1	Kurds
100216	Woman	55	Doctor	2	Russians
100217	Man	54	Marketing	4	Lezgins

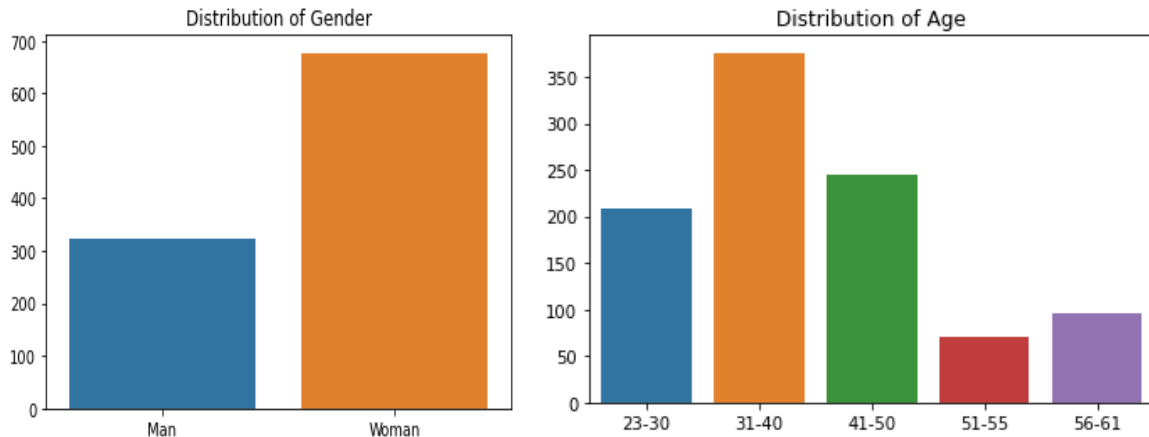
Şəkil 20. Əhalinin fərdi məlumatları

Şəkil 21-də fərdi identifikasiya kodu (PIN) vasitəsi ilə hər iki dataset inteqrasiya olunmuşdur. Burada hər bir fərdin PIN vasitəsi ilə eyni zamanda həm fərdi, həm də sağlamlıq məlumatlarını əldə etmək mümkündür.

PIN	Gender	Age	Profession	Family_size	Nationality	Chronic_disease	Diabet	Bronchial_asthma	Covid_confirmed	Covid_death	Covid_recovered
100213	Man	59	Engineer	1	Azerbaijani	yes	no	no	yes	no	yes
100214	Woman	56	Healthcare	4	Talysh	no	no	no	yes	yes	no
100215	Woman	41	Marketing	1	Kurd	yes	no	no	no	no	no
100216	Woman	55	Executive	2	Russian	yes	no	no	yes	yes	no
100217	Man	54	Marketing	4	Lezgin	no	no	no	yes	yes	no

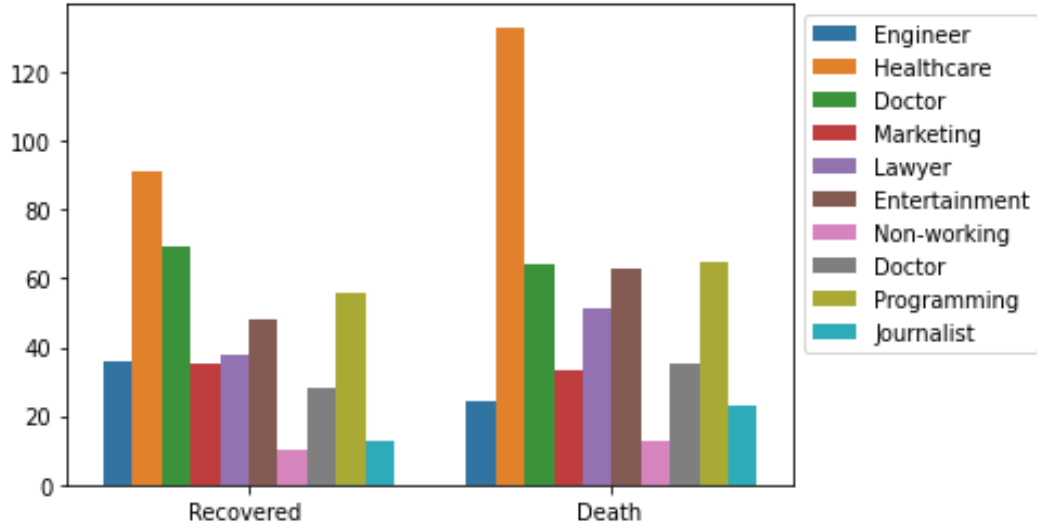
Şəkil 21. Əhali haqqında məlumatlar

Şəkil 22-də fərdlərin yaş aralıqları və cinsə görə bölünmə diaqramı göstərilmişdir. Diaqramlara görə qadın fərdlərin, 31-40 yaş aralığında olan fərdlərin sayının daha çox olduğunu görünür.



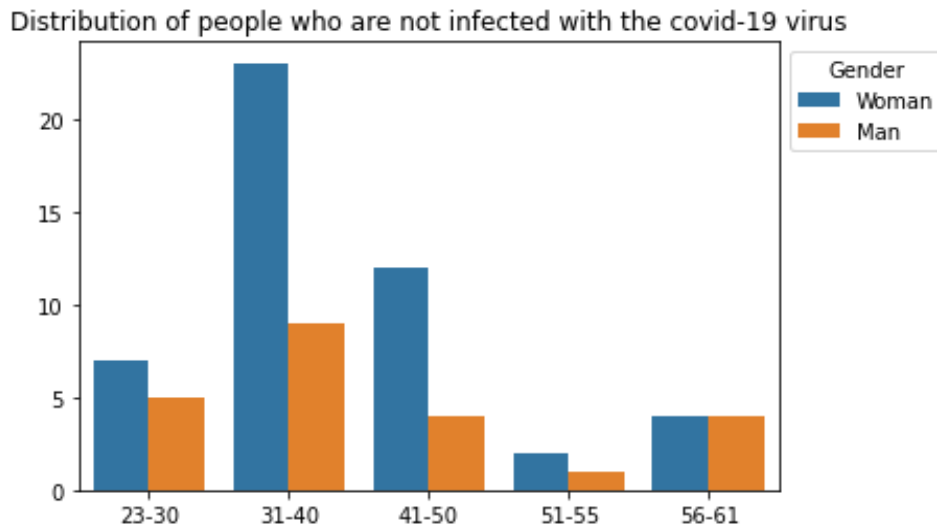
Şəkil 22. Əhalinin cinsə və yaş aralıqlarına görə bölgüsü

Şəkil 23-də Covid-19-dan sağalan və həyatını itirən fərdlərin işlədiyi sahələrə görə bölgüsü verilmişdir. Covid 19-dan sağalan və həyatını itirən fərdlər içərisində tibb işçilərinin sayı daha çoxdur.



Şəkil 23. Covid-19-dan sağalan və həyatını itirən fərdlərin işlədiyi sahələr üzrə bölgüsü

Şəkil 24-də Covid-19-a yoluxmayan fərdlər haqqında məlumatlar qeyd olunub. Şəkil 24-də Covid-19-a yoluxmayanlar 31-40 yaş aralığında qadın cinsində olan fərdlər çoxluq təşkil edir.



Şəkil 24. Covid-19-a yoluxmayan fərdlərin cins və yaş aralığı üzrə bölgüsü

Nəticə olaraq qeyd etmək olar ki, əhali reyestrinin qurulması və yenilənməsi, əldə edilə bilən statistik göstəricilərinin sayının artırılması müxtəlif məlumat mənbələrinin (reyestrlərin) inteqrasiyasını tələb edir. Bu işə əhali reyestri əsasında vahid interqasiya olunmuş reyestr – e-demografiya sisteminin yaradılması deməkdir. Verilənlərin inteqrasiyası (müxtəlif əlyətərli verilən mənbələrindən məlumatların vahid bazaya toplanması) daha çox məlumat və statistik tələbləri ödəmək, daha böyük zaman kəsiyində və daha da detallaşdırılmış səviyyələrdə verilənlərin emalı və statistik analiz üçün geniş imkanlar yaradır. Bununla belə, əsas diqqət

qanunvericilik və siyasi baxışlar, lisenziya, məxfilik və konfidensiallıq, məlumat və metaməlumatlara çıxış və qarşılıqlı informasiya mübadiləsi kimi məsələlərə ön plana çəkilir.

4 Layihə üzrə **elmi nəşrlər** (elmi jurnallarda məqalələr, monoqrafiyalar, icmallar, konfrans materiallarında məqalələr, tezislər) (dərc olunmuş, çapa qəbul olunmuş və çapa göndərilmişləri ayrılıqda qeyd etməklə, uyğun məlumat - jurnalın adı, nömrəsi, cildi, səhifələri, nəşriyyat, indeksi, Impact Factor, həmmüəlliflər və s. bunun kimi məlumatlar - ciddi şəkildə dəqiq olaraq göstərməlidir) (*surətlərini kağız üzərində və CD şəklinə əlavə etməli!*)

#### **Dərc olunmuş:**

1. R. Alguliyev, R. Aliguliyev, G. Niftaliyeva, "Filtration of terrorism-related texts in the e-government environment" // **International Journal of Cyber Warfare and Terrorism**, vol.8, no.4, pp.35-48, 2018. (WoS; Scopus)  
<https://doi.org/10.4018/IJCWT.2018100103>  
ISSN 1947-3435 (print); 1947-3443 (online)
2. R. Alguliyev, F. Yusifov, A. Gurbanli, "Methodology and criteria for evaluating e-services: the case of Azerbaijan" // **eJournal of eDemocracy and Open Government**, vol.10, no.1, pp.106-115, 2018. (EBSCO, DOAJ)  
<https://doi.org/10.29379/jedem.v10i1.493>  
ISSN 2075-9517 (print)
3. F. Abdullayeva, "Bulud əsaslı elektron dövlət xidmətlərinin səmərəliliyinin artırılması metodu" // "**İnformasiya təhlükəsizliyinin aktual multidissiplinar elmi-praktiki problemləri**" IV respublika konfransı, 14 dekabr, Bakı, 2018, s. 13-17.  
<https://doi.org/10.25045/NCInfoSec.2018>
4. İ. Ələkbərova, "Elektron dövlət mühitində fərdi məlumatların mühafizəsi problemləri" // "**İnformasiya təhlükəsizliyinin aktual multidissiplinar elmi-praktiki problemləri**" IV respublika konfransı, s.161-164, Bakı, 14 dekabr 2018.  
<https://doi.org/10.25045/NCInfoSec.2018.38>
5. R. Əliquliyev, R. Alıquliyev, İ. Ələkbərova, "Elektron dövlət mühitində sosial münasibətlərin təhlükəsizliyi" // "**İnformasiya təhlükəsizliyinin aktual multidissiplinar elmi-praktiki problemləri**" IV respublika konfransı, s.64-67, Bakı, 14 dekabr 2018.  
<https://doi.org/10.25045/NCInfoSec.2018.12>
6. Y. Sadiyev, "Evaluating the indicators of web resources operating in the ccTLD domain zone as a key component of national security" // "**İnformasiya təhlükəsizliyinin aktual multidissiplinar elmi-praktiki problemləri**" IV respublika konfransı, s.125-226, Bakı, 14 dekabr 2018.  
<https://doi.org/10.25045/NCInfoSec.2018.27>
7. X. Nurəliyev, "Veb spamlar və onlarla mübarizə metodları" // "**İnformasiya təhlükəsizliyinin aktual multidissiplinar elmi-praktiki problemləri**" IV respublika konfransı, s.233-235, Bakı, 14 dekabr 2018.

<https://doi.org/10.25045/NCInfoSec.2018.56>

8. P.M. Alguлиev, P.M. Alyguлиev, Ф.Ф. Юсифов, И.Я. Алекперова, “Формирование электронной демографии как эффективного инструмента социальных исследований и мониторинга данных о населении” // **Вопросы Государственного и Муниципального Управления (Public Administration Issues)**, №4, с.61-86. 2019. (WoS; Scopus)

<https://vgmu.hse.ru/2019--4/326123454.html>  
ISSN 1999-5431 (print); 2409-5095 (online)

9. R. Əliquliyev, İ. Ələkbərova, “Elektron dövlətin formalaşmasında sosial kapitalın rolunun analizi” // **İnformasiya Texnologiyaları Problemləri**, №1, s.35-46, 2019. (INSPEC, Google Scholar)

<https://doi.org/10.25045/jpit.v10.i1.04>  
ISSN 2077-4001 (print); 2304-0157 (online)

10. R. Əliquliyev, İ. Ələkbərova, “Elektron dövlət mühitində vətəndaşların sosial kreditinin fərdi məlumatlar əsasında qiymətləndirilməsi: problemlər və perspektivlər” // **İnformasiya Cəmiyyəti Problemləri**, №1, s.3-13, 2019. (INSPEC, Google Scholar)

<https://doi.org/10.25045/jpis.v10.i1.01>  
ISSN 2077-964X (print); 2309-7566 (online)

11. R.M. Əliquliyev, R.M. Alıquliyev, R.Ş. Mahmudov, “Azərbaycanda plagiatlıqla mübarizə problemləri və həll yolları” // **İnformasiya Cəmiyyəti Problemləri**, №1, s.34-43, 2019. (CrossRef; Google Scholar)

<https://doi.org/10.25045/jpis.v10.i1.04>  
ISSN 2077-964X (print); 2309-7566 (online)

12. G.Y. Iskandarli, “Applying clustering and topic modeling to automatic analysis of citizens’ comments in e-government” // **International Journal of Information Technology and Computer Science**, vol.12, no.6, pp.1-10, 2020. (INSPEC)

<https://doi.org/10.5815/ijitcs.2020.06.01>  
ISSN 2074-9007 (print); 2074-9015 (online)

13. F.F. Yusifov, A.C. Fərəcova, “Vətəndaş məmnuniyyəti əsasında e-xidmətlərin çoxmeyarlı qiymətləndirilməsi modeli” // **İnformasiya Cəmiyyəti Problemləri**, №2, s.39-50, 2020. (CrossRef; Google Scholar)

<https://doi.org/10.25045/jpis.v11.i2.04>  
ISSN 2077-964X (print); 2309-7566 (online)

14. R.M. Əliquliyev, Y.N. İmamverdiyev, R.Ş. Mahmudov, “İnformasiya təhlükəsizliyi milli təhlükəsizliyin mühüm komponenti kimi” // **İnformasiya Cəmiyyəti Problemləri**, №1, s.3-25, 2020. (CrossRef; Google Scholar)

<https://doi.org/10.25045/jpis.v11.i1.01>  
ISSN 2077-964X (print); 2309-7566 (online)



15. R.Ş. Mahmudov, "Big data iqtisadiyyatının tənzimlənməsi problemləri" // Azərbaycan Texniki Universitetinin 70 illik yubileyinə həsr olunmuş "Dördüncü sənaye inqilabının texnoloji perspektivləri: sənaye interneti, kibernetika sistemləri və intellektual texnologiyalar" respublika elmi-texniki konfransı, s.88-94, Bakı, 26-27 noyabr 2020.  
<http://aztu.edu.az/azp/yubiley/az/main/main.jsp>
16. R.M. Əliquliyev, İ.Y. Ələkbərova, "Fərdi məlumatların e-idarəetmə sisteminin sosial kredit sistemində tətbiqinin perspektivləri haqqında" // **İnformasiya Cəmiyyəti Problemləri**, №1, s.67-76-50, 2021. (CrossRef; Google Scholar)  
<https://doi.org/10.25045/jpis.v12.i1.06>  
ISSN 2077-964X (print); 2309-7566 (online)
17. F.J. Abdullayeva, "Advanced persistent threat attack detection method in cloud computing based on autoencoder and softmax regression algorithm" // **Array**, vol.10, pp.1-11, 2021. (WoS; Scopus)  
<https://doi.org/10.1016/j.array.2021.100067>  
ISSN 2590-0056 (print)
18. F.J. Abdullayeva, "Cloud computing virtual machine workload prediction method based on variational autoencoder" // **International Journal of Systems and Software Security and Protection**, vol.12, no.2, pp. 33-45. 2021. (WoS; Scopus)  
<https://doi.org/10.4018/IJSSSP.2021070103>  
ISSN 2640-4265 (print); 2640-4273 (online)
19. R.M. Əliquliyev, R.M. Alıquliyev, İ.Y. Ələkbərova, "Videotəsvirlərin analizi əsasında sosial münasibətlərin aşkarlanması üçün yanaşma" // **İnformasiya Texnologiyaları Problemləri**, №2, s.3-15, 2021. (CrossRef; Google Scholar)  
<https://doi.org/10.25045/jpit.v12.i2.01>  
ISSN 2077-4001 (print), 2304-0157 (online)
20. R. Mahmudov, "Açıq dövlət məlumatlarının dəyərə çevrilməsi məsələləri", **İnformasiya Cəmiyyəti Problemləri**, №2, s.131-144, 2021.  
<https://doi.org/10.25045/jpis.v12.i2.10>  
ISSN 2077-964X (print), 2309-7566 (online)
21. F.F. Yusifov, A.C. Fərəcova, "E-idarəetmədə fərdi məlumatların qorunması: "Data səfirliyi" konsepsiyası" // **İnformasiya Cəmiyyəti Problemləri**, №2, s.119-130, 2021.  
<https://doi.org/10.25045/jpis.v12.i2.09>  
ISSN 2077-964X (print); 2309-7566 (online)
22. G. Iskandarli, "Detecting the main topics of citizens' comments in e-government" // Proceedings of the **International Symposium on Applied Sciences and Engineering**, pp.539-542, Erzurum, Turkey, 07-09 April 2021.  
<http://kongresempozyum.org/kongre-detay/93412f66-3fcf-4ef7-9817-2777e0d1bcab>

23. Y.N. Imamverdiyev, F.J. Abdullayeva, "Convolutional neural network for detecting application layer distributed denial of service attacks" // Сборник материалов XVI Международной научно-технической конференции "Распознавание-2021", с.22-25, Курск, Россия, 14-17 сентября 2021.

<https://swsu.ru/structura/up/fivt/kvt/recogn21.php>

24. И.Я. Алакбарова, "Разработка общей структуры интеллектуальной системы видеонаблюдения с использованием персональных данных" // Сборник материалов XVI Международной научно-технической конференции "Распознавание-2021", с.37-39, Курск, Россия, 14-17 сентября 2021.

<https://swsu.ru/structura/up/fivt/kvt/recogn21.php>

25. F.J. Abdullayeva, "Detection of cyberattacks in cloud computing service delivery models using correlation-based feature selection" // Proceedings of the 15th IEEE International Conference on Application of Information and Communication Technologies, pp.1-4, Baku, 13-15 October 2021. (WoS)

<https://doi.org/10.1109/AICT52784.2021.9620347>

#### Çара qəbul olunmuş

26. R.M. Aliguliyev, G.Y. Iskandarli, "Measuring citizen satisfaction with e-government services by using sentiment analysis technology" // **International Journal of Electronic Governance**, 2022. (Scopus)

<https://www.inderscience.com/info/ingeneral/forthcoming.php?jcode=ijeg>

ISSN 1740-7494 (print); 1740-7508 (online)

27. F. Yusifov, N. Akhundova, "Analysis of demographic characteristics based on e-demography data" // **Demography and Social Economy**, vol.47, no.1, 2022. (Scopus)

<https://doi.org/10.15407/dse>

ISSN 2072-9480 (print); 2309-2351 (online)

28. Ф.Ф. Юсифов, Н.Е. Ахундова, "Единый реестр населения как источник проведения социально-демографического анализа", **Народонаселение**, т.25, №1, 2022. (WoS; Scopus)

<https://www.jour.fnisc.ru/index.php/index>

ISSN 1562-2495 (print); 1684-1581 (online)

#### Çара göndərilmiş

29. F.J. Abdullayeva, "Distributed Denial of Service Attack Detection in E-Government Cloud via Data Clustering" // **Array**, 2022 (WoS; Scopus).

<https://www.journals.elsevier.com/array/>

ISSN 2590-0056 (print)

30. F.J. Abdullayeva, "Convolutional neural network based automatic diagnostic system for AL-DDoS attacks detection" // **International Journal of Cyber Warfare and Terrorism**, 2021 (WoS; Scopus)

31. И.Я. Алекперова, “Об одном подходе для выявления отношений в социальных сетях”, **Информационное Общество, 2022. (WoS; Scopus)**

<http://infosoc.iis.ru/>

ISSN 1606-1330 (print); 1605-9921 (online)

5	İxtira və patentlər, səmərələşdirici təkliflər
	<b>Yoxdur</b>
6	Layihə üzrə ezamiyyətlər (ezamiyyə baş tutmuş təşkilatın adı, şəhər və ölkə, ezamiyyə tarixləri, həmçinin ezamiyyə vaxtı baş tutmuş müzakirələr, görüşlər, seminarlarda çıxışlar və s. dəqiq göstərməlidir)
	<b>Yoxdur.</b>
7	Layihə üzrə elmi ekspedisiyalarda iştirak (əgər varsa)
	<b>Yoxdur.</b>
8	Layihə üzrə digər tədbirlərdə iştirak
	<b>Yoxdur.</b>
9	Layihə mövzusu üzrə elmi məruzələr (seminar, dəyirmi masa, konfrans, qurultay, simpozium və s. çıxışlar) (məlumat tam şəkildə göstərməlidir: a) məruzənin növü: plenar, dəvətli, şifahi və ya divar məruzəsi; b) tədbirin kateqoriyası: ölkədaxili, regional, beynəlxalq) <ul style="list-style-type: none"><li>▪ “İnformasiya təhlükəsizliyinin aktual multidissiplinar elmi-praktiki problemləri” IV respublika konfransı, Bakı, 14 dekabr 2018.</li><li>▪ Azərbaycan Texniki Universitetinin 70 illik yubileyinə həsr olunmuş “Dördüncü sənaye inqilabının texnoloji perspektivləri: sənaye interneti, kiberfiziki sistemlər və intellektual texnologiyalar” respublika elmi-texniki konfransı, Bakı, 26-27 noyabr 2020.</li><li>▪ International Symposium on Applied Sciences and Engineering, pp.539-542, Erzurum, Turkey, 07-09 April 2021.</li><li>▪ XVI Международная научно-техническая конференция “Распознавание-2021”, Курск, Россия, 14-17 сентября 2021.</li><li>▪ 15th IEEE International Conference on Application of Information and Communication Technologies (Online), Bakı, 13-15 October 2021.</li><li>▪ “Beynəlxalq İnformasiya Təhlükəsizliyi Günü”nə həsr olunmuş onlayn elmi seminar, Bakı, 29 noyabr 2021-ci il.</li><li>▪ Национальный Суперкомпьютерный Форум (НСКФ-2021), Конференция-спутник «Облачные и распределенные вычислительные системы в электронном управлении», Россия, Переславль-Залесский, 30 ноября – 03 декабря 2021.</li></ul>
10	Layihə üzrə əldə olunmuş cihaz, avadanlıq və qurğular, mal və materiallar, komplektləşdirmə məlumatları
	<b>Yoxdur.</b>
11	Yerli həmkarlarla əlaqələr
	<b>Yoxdur.</b>
12	Xarici həmkarlarla əlaqələr

	Rusiya Elmlər Akademiyasının A.A. Xarkeviç adına İnformasiyanın Ötürülməsi Problemləri İnstitutunun "Paylanmış Hesablamalar Mərkəzi"nin və "Paylanmış Hesablama Sistemləri" laboratoriyasının müdiri, f.-r.e.d., prof. Aleksandr Petroviç Afanasyev ilə müzakirələr aparılmışdır.
1 3	Layihə mövzusu üzrə kadr hazırlığı (əgər varsa)
	Elmlər doktoru, fəlsəfə doktoru və magistr hazırlığı proqramları üzrə doktorant və magistrantlar hazırlanır.
1 4	Sərgilərdə iştirak (əgər baş tutubsa)
	<b>Yoxdur.</b>
1 5	Təcrübəartırmada iştirak və təcrübə mübadiləsi (əgər baş tutubsa)
	<b>Yoxdur.</b>
1 6	Layihə mövzusu ilə bağlı elmi-kütləvi nəşrlər, kütləvi informasiya vasitələrində çıxışlar, yeni yaradılmış internet səhifələri və s. (məlumatı tam şəkildə göstərməlidir)
	www.ikt.az, www.science.gov.az və www.azertag.az saytlarında layihə ilə bağlı keçirilən seminarlar və dərc olunan məqalələr haqqında xəbərlər verilmişdir.

**SİFARIŞÇI:**

**Elmin İnkişafı Fondu**

**Baş məsləhətçi**

**Xanım Faiq qızı Daşdəmirova**

\_\_\_\_\_  
(imza)

"09" mart 2022-ci il

**İCRAÇI:**

**Layihə rəhbəri**

**Rasim Məhəmməd oğlu Əliquliyev**

\_\_\_\_\_  
(imza)

"09" mart 2022-ci il



**AZƏRBAYCAN RESPUBLİKASININ PREZİDENTİ YANINDA**  
**ELMİN İNKİŞAFI FONDU**

MÜQAVİLƏYƏ ƏLAVƏ

**Azərbaycan Respublikasının Prezidenti yanında Elmin İnkişafı Fondu**  
**və Rusiya Fundamental Tədqiqatlar Fondunun**  
**1-ci Azərbaycan-Rusiya birgə beynəlxalq grant**  
**müsabiqəsinin (EIF-BGM-4-RFTF-1/2017) qalibi olmuş**  
**layihənin yerinə yetirilməsi üzrə**

**ALINMIŞ NƏTİCƏLƏRİN ƏMƏLİ (TƏCRÜBİ) HƏYATA KEÇİRİLMƏSİ**  
**VƏ LAYİHƏNİN NƏTİCƏLƏRİNDƏN GƏLƏCƏK TƏDQIQATLARDƏ**  
**İSTİFADƏ PERSPEKTİVLƏRİ HAQQINDA**  
**MƏLUMAT VƏRƏQİ**

(Qaydalar üzrə Əlavə 16)

Layihənin adı: **“Big Data” analitikası texnologiyalarının tətbiqi ilə elektron idarəetmənin səmərəliliyinin artırılması metodları və alqoritmlərinin işlənilməsi**

Layihə rəhbərinin adı, atasının adı və soyadı: **Rasim Məhəmməd oğlu Əliquliyev**

Qrantın məbləği: **108 000 manat**

Layihənin nömrəsi: **EIF-BGM-4-RFTF-1/2017-21/08/1-M-11**

Müqavilənin imzalanma tarixi: **17 avqust 2020-ci il**

Qrant layihəsinin yerinə yetirilmə müddəti: **18 ay**

Layihənin icra müddəti (başlama və bitmə tarixi): **01 sentyabr 2020-ci il – 01 mart 2022-ci il**

Diqqət! Bütün məlumatlar 12 ölçülü Arial şrifti ilə, 1 intervalla doldurulmalıdır.

**1. Layihənin nəticələrinin əməli (təcrübi) həyata keçirilməsi**

<b>1</b>	Layihənin əsas əməli (təcrübi) nəticələri, bu nəticələrin məlum analoqlar ilə müqayisəli xarakteristikası
	Təklif edilmiş metodlar və alqoritmlər real məlumatlar üzərində sınaqdan keçirilmiş, onların digər metod və alqoritmlərlə müqayisəli analizi aparılmışdır. Eksperimentlərin nəticələri təsdiq etmişdir ki, təklif olunan metodlar və alqoritmlər həm keyfiyyət, həm də işləmə sürətinə görə müqayisə olunan mövcud metod və alqoritmlərdən daha yaxşı nəticələr nümayiş etdirir.
<b>2</b>	Layihənin nəticələrinin əməli (təcrübi) həyata keçirilməsi haqqında məlumat (istehsalatda tətbiq (tətbiqin aktını əlavə etməli); tədris və təhsildə (nəşr olunmuş elmi əsərlər və s. – təhsil sistemində tətbiqin aktını əlavə etməli); bağlanmış xarici müqavilələr və ya beynəlxalq layihələr (kimlə bağlanıb, müqavilənin və ya layihənin nömrəsi, adı, tarixi və dəyəri); dövlət

proqramlarında (dövlət orqanının adı, qərarın nömrəsi və tarixi); ixtira üçün alınmış patentlərdə (patentin nömrəsi, verilmə tarixi, ixtiranın adı); və digərlərində)

**Yoxdur.**

## 2. Layihənin nəticələrindən gələcək tədqiqatlarda istifadə perspektivləri

1

Nəticələrin istifadəsi perspektivləri (fundamental, tətbiqi və axtarış-innovasiya yönü elmi-tədqiqat layihə və proqramlarında; dövlət proqramlarında; dövlət qurumlarının sahə tədqiqat proqramlarında; ixtira və patent üçün verilmiş ərizələrdə; beynəlxalq layihələrdə; və digərlərində)

Təklif edilmiş metodlar və alqoritmlər səmərəli qərarların qəbulu məqsədilə e-hökumət (e-gov) portalında toplanmış müxtəlif növ məlumatların analizində istifadə oluna bilər. Bu metod və alqoritmlər ümumiyyətlə, böyük həcmli məlumatların toplandığı və analizi böyük əhəmiyyət kəsb edən digər sahələrdə də istifadə oluna bilər. Alınmış nəticələr həm də ali təhsil müəssisələrinin müvafiq ixtisaslarının tədris prosesində, o cümlədən fundamental və innovasiya yönümlü elmi-tədqiqat layihələrinin yerinə yetirilməsində istifadə oluna bilər. Bu metod və alqoritmlər xüsusilə, ASAN sistemində e-xidmətlərin keyfiyyətinin yaxşılaşdırılması məqsədilə elektron dövlət-vətəndaş platformasında toplanmış məlumatların intellektual analizində istifadə oluna bilər.

**SİFARIŞÇI:**

**Elmin İnkişafı Fondu**

**Baş məsləhətçi**

**Xanım Faiq qızı Daşdəmirova**

(imza)

“09” mart 2022-ci il

**İCRAÇI:**

**Layihə rəhbəri**

**Rasim Məhəmməd oğlu Əliquliyev**

(imza)

“09” mart 2022-ci il





**AZƏRBAYCAN RESPUBLİKASININ PREZİDENTİ YANINDA  
ELMİN İNKİŞAFI FONDU**

MÜQAVİLƏYƏ ƏLAVƏ

**Azərbaycan Respublikasının Prezidenti yanında Elmin İnkişafı Fondu  
və Rusiya Fundamental Tədqiqatlar Fondunun  
1-ci Azərbaycan-Rusiya birgə beynəlxalq qrant  
müsabiqəsinin (EIF-BGM-4-RFTF-1/2017) qalibi olmuş  
layihənin yerinə yetirilməsi üzrə**

**ALINMIŞ ELMİ MƏHSUL HAQQINDA MƏLUMAT  
(Qaydalar üzrə Əlavə 17)**

Layihənin adı: **“Big Data” analitikası texnologiyalarının tətbiqi ilə elektron idarəetmənin səmərəliliyinin artırılması metodları və alqoritmlərinin işlənilməsi**

Layihə rəhbərinin adı, atasının adı və soyadı: **Rasim Məhəmməd oğlu Əliquliyev**

Qrantın məbləği: **108 000 manat**

Layihənin nömrəsi: **EIF-BGM-4-RFTF-1/2017-21/08/1-M-11**

Müqavilənin imzalanma tarixi: **17 avqust 2020-ci il**

Qrant layihəsinin yerinə yetirilmə müddəti: **18 ay**

Layihənin icra müddəti (başlama və bitmə tarixi): **01 sentyabr 2020-ci il – 01 mart 2022-ci il**

**Diqqət! Bütün məlumatlar 12 ölçülü Arial şrifti ilə, 1 intervalla doldurulmalıdır**

**1. Elmi əsərlər (sayı)**

№	Tamlıq dərəcəsi	Dərc olunmuş	Çapa qəbul olunmuş və ya çapda olan	Çapa göndərilmiş
1.	Monoqrafiyalar	–	–	–
	həmçinin, xaricdə çap olunmuş	–	–	–
2.	Məqalələr	21	18	3
	həmçinin xarici nəşrlərdə	12	9	3
3.	Konfrans materiallarında məqalələr	6	6	0
	o cümlədən, beynəlxalq konfrans materiallarında	2	2	0
4.	Məruzələrin tezisləri	4	4	0
	həmçinin, beynəlxalq tədbirlərin toplusunda	2	2	0
5.	Digər (icmal, atlas, kataloq və s.)	–	–	–

## 2. İxtira və patentlər (sayı)

No	Elmi məhsulun növü	Alınmış	Verilmiş	Ərizəsi verilmiş
1.	Patent, patent almaq üçün ərizə	–	–	–
2.	İxtira	–	–	–
3.	Səmərələşdirici təklif	–	–	–

## 3. Elmi tədbirlərdə məruzələr (sayı)

No	Tədbirin adı (seminar, dəyirmi masa, konfrans, qurultay, simpozium və s.)	Tədbirin kateqoriyası (ölkədaxili, regional, beynəlxalq)	Məruzənin növü (plənar, dəvətli, şifahi, divar)	Sayı
1.	“İnformasiya təhlükəsizliyinin aktual multidissiplinar elmi-praktiki problemləri” IV respublika konfransı, Bakı, 14 dekabr 2018.	ölkədaxili	şifahi	5
2.	Azərbaycan Texniki Universitetinin 70 illik yubileyinə həsr olunmuş “Dördüncü sənaye inqilabının texoloji perspektivləri: sənaye interneti, kiberfiziki sistemlər və intellektual texnologiyalar” respublika elmi-texniki konfransı, s.88-94, Bakı, 26-27 noyabr 2020.	ölkədaxili	şifahi	1
3.	International Symposium on Applied Sciences and Engineering, pp.539-542, Erzurum, Turkey, 07-09 April 2021	beynəlxalq	divar	1
4.	XVI Международная научно-техническая конференция “Распознавание-2021”, Курск, Россия, 14-17 сентября 2021.	beynəlxalq	şifahi	2
5.	15th IEEE International Conference on Application of Information and Communication Technologies, Baku, 13-15 October 2021	beynəlxalq	şifahi	1

### SİFARIŞÇI:

**Elmin İnkişafı Fondu**

### Baş məsləhətçi

**Xanım Faiq qızı Daşdəmirova**

(imza)

“09” mart 2022-ci il

### İCRAÇI:

### Layihə rəhbəri

**Rasim Məhəmməd oğlu Əliquliyev**

(imza)

“09” mart 2022-ci il