



AZƏRBAYCAN ELM FONDU

Azərbaycan Elm Fondunun
Ümummilli Lider Heydər Əliyevin 100-illik
yubileyinə həsr olunmuş
“Əsas qrant müsabiqəsi-2023” ün
(AEF-MCG-2023-1(43)) qalibi olmuş
layihənin yerinə yetirilməsi üzrə aralıq
(rüblük olaraq 1-ci mərhələ)

ELMİ-TEXNİKİ HESABAT

Layihənin adı: **Sosial kiber fiziki sistemlərdə fərdi məlumatların qorunması üçün süni intellekt üsullarının işlənməsi**

Layihə rəhbərinin soyadı, adı və atasının adı: **Abdullayeva Fərqanə Cabbar qızı**

Layihənin nömrəsi: **AEF-MCG-2023-1(43)-13/04/1-M-04**

Müqavilənin imzalanma tarixi: **04 dekabr 2023-cü il**

Qrant layihəsinin yerinə yetirilmə müddəti: **24 ay**

Layihənin icra müddəti (başlama və bitmə tarixi): **01 yanvar 2024-cü il – 01 yanvar 2026-cı il**

Layihənin **I mərhələ** üzrə (rüb) məbləği:

Hesabatda aşağıdakı məsələlər işıqlandırılmalıdır:

1 Layihənin həyata keçirilməsi üzrə cari rübdə yerinə yetirilmiş **elmi işlər**

Industry 4.0 mühitində insanın genderinin avtomatik tanınması biometrik və demoqrafik tədqiqatların aktual məsələlərindən biridir. Demoqrafik statistikanın Industry 4.0 mühitində aparılmasını avtomatlaşdırmaq üçün genderin tanınması texnologiyalarından istifadə edirlər. İnsanlar əvvəlki biliklərinə görə danışanın genderini asanlıqla təyin edə bilər. Lakin şəxsin kişi və ya qadın cinsinə məxsus olduğunu kompüter sistemləri asanlıqla müəyyən edə bilmir. Məqalədə səs siqnallarının gender siniflərinə kateqoriyalaşdırılmasını həyata keçirən alqoritmlər işlənmişdir. Səs yazısından insanın genderinin tanınması çətin məsələ hesab olunur. Bu məqsədlə işdə genderin tanınması üçün çox sayda fazalardan ibarət kompleks yanaşma təklif olunub. Burada səs siqnallarından əlamətlərin çıxarılması üçün 20 ədəd

MFCC parametri hesablanmışdır. MFCC-lərin CNN, GaussianNB və Knn alqoritmlərinin daxilində istifadə edilməsi insanın səs siqnallarında genderin tanınması dəqiqliyini olduqca artırmışdır.

Fişinq kiberhücumlarını aşkarlamaq üçün URL və HTML-dən istifadə etməklə təlim edilmiş dərin neyron şəbəkəsinə əsaslanan model təklif edilmişdir. Bu model mətn məlumatından əlamətləri çıxarmaq üçün təbii dil emalından (Natural Language Processing, NLP) istifadə edir. Onun semantik asılılıqlarını modelləşdirmək üçün konvolyusiya laylarından istifadə olunur. Modelin yoxlanılması üçün müxtəlif performans metrikaları (precision, recall, F-measure) istifadə olunur. Experimentlər fişinq səhifələrinin müəyyən edilməsində yüksək təsnifat dəqiqliyi əldə edən real fişinq verilənləri ilə aparılmışdır.

Tibbi informasiya məkanının elektron sağlamlıq qeydləri, ictimai səhiyyə məlumatları, tibbi ədəbiyyat, səhiyyə standartları və tibbi məlumatlarının mübadiləsi infrastrukturunu kimi müxtəlif aspektləri araşdırılmışdır. Fərdi tibbi məlumatların idarə edilməsi sahəsindəki dəyişikliklərin təhlili aparılmışdır. Fərdi tibbi məlumatların rəqəmsallaşdırılması ilə bağlı kibertəhlükəsizlik riskləri və təhdidləri araşdırılmışdır. Fərdi tibbi məlumatların kibertəhlükəsizliyinin təmin edilməsi üçün bütövlüyünün və məxfiliyinin təmin edilməsi və giriş nəzarət kimi məsələlər araşdırılmışdır.

Süni intellektə əsaslanan RNN, LSTM və GRU kimi dərin təlim modelləri və onların iki istiqamətli (bidirectional) variantları üzrə müqayisəli təhlil aparılmışdır. Bu tədqiqatın əsas məqsədi dərin təlim modellərini müqayisə etmək və hər bir modelin performansını dəqiq olaraq qiymətləndirməkdir. Modellərin performansını təhlil etmək üçün dəqiqlik (accuracy) qiymətləndirmə indeksindən istifadə edilmişdir. Eksperimentlər seçilmiş müxtəlif ölçülü verilənlərdən istifadə etməklə aparılmışdır.

Elektron ticarət platformasında müştərilərin fərdi məlumatlarına müdaxilə edə biləcək fişinq (phishing), kobud güc (brute-force) , “man-in-the-middle” və SQL inyeksiyaları hücumlarına baxılmışdır. E-ticarətdə fərdi məlumatlara kibertəhdidlərin aşkarlanmasında beynəlxalq təcrübələrə əsasən ən çox istifadə olunan maşın təlimi və dərin təlim metodları araşdırılmışdır.

Veb səhifələrin ranqlaşdırma parametrlərinin hesablanması üçün zona bölünməsi strategiyasına əsaslanan yanaşma təklif edilib. Bu yanaşmanın məqsədi qraf strukturlarında dərəcə məlumatlarını klassifikasiya etmək üçün çox sayda neyron şəbəkədən istifadə etməkdir. Yanaşmaya görə veb səhifələr üç fərqli zonaya bölünür. Burada əsas zona GNN-lərin öyrədilməsi üçün istifadə edilir. Bu zonada təlim

	verilənlərinin sinifləri əvvəlcədən məlum olur. Növbəti zona kəşf edilməmiş zonadır. Burada klassifikatorlar qovşaq parametrlərini sinifləndirir. Növbəti zona kəşf etmə zonasıdır, çoxsaylı kəşf edilməmiş zonalara aid qovşaqları və tilləri özündə birləşdirir.
2	Layihənin həyata keçirilməsi üzrə planda nəzərdə tutulmuş işlərin yerinə yetirilmə dərəcəsi (cari rüb üçün, faizlə qiymətləndirməli)
	Planda nəzərdə tutulmuş işlər 100 % yerinə yetirilmişdir.
3	Hesabat dövründə alınmış elmi nəticələr , onların yenilik dərəcəsi
	<p>Yanaşmanın əsas elmi yeniliyi genderin effektiv tanınması üçün CNN-lərin MFCC parametrləri ilə birgə işləməsini təmin edən hibrid modelin işlənməsidir. Bu məqsədlə ilk öncə səs siqnallarından MFCC parametrləri çıxarılır daha sonra alınmış informativ əlamətlər əsasında verilənlərin qadın və kişi siniflərinə klassifikasiyası həyata keçirilir. Təklif edilmiş yanaşmanın elmi tədqiqatlar üçün açıq olan VoxCeleb_gender verilənlər bazası üzərində test nəticələri onun mövcud metodlarla müqayisədə daha üstün olduğunu göstərir. Təklif edilmiş sistemi bir neçə faza təşkil edir. 1) səs verilənlərindən MFCC əlamətlərinin çıxarılması; 2) generasiya edilmiş verilənlərin Qaus paylanması şəklində normallaşdırılması; 3) verilənlərin klassifikasiyası. Səs siqnallarına MFCC üsulunu tətbiq etdikdən sonra alınmış çıxış matrisi bütün kadrlardan çıxarılmış əlamətlər vektorunu əmələ gətirir. Bu matrisdə sətirlər uyğun kadrların nömrəsini, sütunlar isə uyğun əlamət vektorunun əmsallarını göstərir. Bizim sistemdə bu məqsədlə 20 ədəd MFCC-dən ibarət əlamətlər vektoru istifadə olunub. Klassifikasiyanı həyata keçirmək üçün formalaşmış bu matris klassifikatorun girişinə ötürülür. Klassifikasiya mərhələsini təlim və test mərhələləri təşkil edir. MFCC əmsallarını hesablamaq üçün siqnalların analizində tətbiq olunan standart alqoritmədən istifadə edilmişdir.</p> <p>Mətn məlumatlarında fişinq kiberhücumlarının aşkarlanması üçün dərin neyron şəbəkə modelinə əsaslanan metod təqdim olunur və NLP-dən istifadə etməklə onun həllinə yanaşma təklif edilmişdir. Logistik reqressiya, dayaq vektor maşınları, qərar ağacları, Naive Bayes kimi mövcud yanaşmalardan fərqli olaraq, bu yanaşma mətn tipli məlumatlarda fişinqi yüksək dəqiqliklə aşkarlaya bilmişdir.</p> <p>Tibbi informasiya məkanının elektron sağlamlıq qeydləri, ictimai səhiyyə məlumatları, tibbi ədəbiyyat, səhiyyə standartları və tibbi məlumatlarının mübadiləsi infrastrukturunu kimi müxtəlif aspektləri araşdırılmışdır. Fərdi tibbi məlumatların idarə edilməsi sahəsindəki dəyişikliklərin təhlili aparılmışdır. Fərdi tibbi məlumatların rəqəmsallaşdırılması ilə bağlı kibertəhlükəsizlik riskləri və təhdidləri araşdırılmışdır.</p>

Fərdi tibbi məlumatların kibertəhlükəsizliyinin təmin edilməsi üçün bütövlüyünün və məxfiliyinin təmin edilməsi və giriş nəzarət kimi məsələlər araşdırılmışdır.

Süni intellektə əsaslanan RNN, LSTM və GRU kimi dərin təlim modelləri və onların iki istiqamətli (bidirectional) variantları üzrə müqayisəli təhlil aparılmışdır. Bu tədqiqatın əsas məqsədi dərin təlim modellərini müqayisə etmək və hər bir modelin performansını dəqiq olaraq qiymətləndirməkdir. Modellərin performansını təhlil etmək üçün dəqiqlik (accuracy) qiymətləndirmə indeksindən istifadə edilmişdir. Eksperimentlər seçilmiş müxtəlif ölçülü verilənlərdən istifadə etməklə aparılmışdır. Təcrübənin nəticələri göstərir ki, iki istiqamətli LSTM modeli digər dərin təlim modelləri ilə müqayisədə ən yüksək dəqiqlik nəticələri əldə etmişdir. Həmçinin, iki istiqamətli GRU modelidə nəzərəcarpacaq dəqiqlik nəticələri göstərmişdir.

Elektron ticarət platformasında müştərilərin fərdi məlumatlarına müdaxilə edə biləcək fişinq (phishing), kobud güc (brute-force), "man-in-the-middle" və SQL inyeksiyaları hücumlarına baxılmışdır. E-ticarətdə fərdi məlumatlara kibertəhdidlərin aşkarlanmasında beynəlxalq təcrübələrə əsasən ən çox istifadə olunan maşın təlimi və dərin təlim metodları araşdırılmışdır.

İnternet mühitində istifadəçilərin sorğularının rəqləşdirilməsi mühüm məsələ hesab olunur. Buyuk strukturlu verilənlərdə rankların hesablanması mürəkkəb məsələdir və iterativ hesablamalar tələb edir. Bu səbəbdən müvcud rəqləşdirma alqoritmlərini optimallaşdırmaq üçün əlavə parametrlərdən və ya evristik yanaşmalardan istifadə edirlər. Bu yolla işlənmiş üsullarda da böyük iterativ hesablamaların aparılması qaçılmaz olur. Təklif olunan yanaşmada nüvə adlanan bəzi təsadüfi seçilmiş subqraflar (zonalar) üçün iterativ hesablama tələb edən hibrid üsul işlənir. Təklif olunan metodda çoxsaylı klassifikatorlar qrafı üç zona kateqoriyasına bölürlər: Əsas zona; Kəşf edilməmiş zona; Kəsişmə zonası.

Kəşf edilməmiş və kəsişmə zonaları yalnız proqnozlaşdırma üçün istifadə olunur. Əsas zona sinifləri məlum olan və ya ənənəvi deterministik alqoritmlər vasitəsilə hesablanan altqrafadan ibarətdir. Kəşf edilməmiş zona adlı altqrafda hər bir qovşaq üçün ranklar qraf əsaslı neyron şəbəkələri vasitəsilə hesablanır. Təcrübələr çoxlu təsnifatlayıcıların nəticələrinin toplanması yolu ilə kəsişmə zonalarında həqiqi etiketlərin təsnif edilməsi ehtimalının kəşf edilməmiş zonalara nisbətən daha yüksək olduğunu göstərmişdir.

4 Layihənin yerinə yetirilməsi zamanı istifadə olunan üsul və yanaşmalar

Təbii dilin emalı (NLP), dərin təlim, maşın təlimi metodları, səs siqnallarının emalı,

	Data Mining, Text Mining, K-Nearest Neighbors, Gaussian Naive Bayes metodları, Qraf Neyron Şəbəkələri, Convolutional Neural Network, Recurrent Neural Network, Long Short-Term Memory, Gated Recurrent Unit (GRU)
5	Layihə üzrə elmi nəşrlər (məqalələr, monoqrafiyalar, icmaller, konfrans materialları, tezislər) (dərc olunmuş, çapa qəbul olunmuş və çapa göndərilmişləri ayrılıqda qeyd etməklə) (surətlərini əlavə etməli!)
	<p>Çapa göndərilmiş</p> <ol style="list-style-type: none"> 1. F.J. Abdullayeva, "Gender recognition for the collecting valuable demographic information in a social environment in Industry 4.0 transformations", Data and Information Management. (Scopus) (rəydədir) 2. F.J. Abdullayeva, S.N. Suleymanzade, "Estimating page ranks with inductive capability of graph neural networks and zone partitioning", Automatic Control and Computer Sciences. (Web of Science, Scopus, IF: 0.9) (rəydədir).
6	İxtira və patentlər, səmərələşdirici təkliflər
	Yoxdur
7	Layihə üzrə ezamiyyətlər
	Dokuz Eylül Universiteti, İzmir, Türkiyə, 25-29 Mart 2024.
8	Layihə üzrə elmi ekspedisiyalarda iştirak
	Yoxdur
9	Layihə üzrə digər tədbirlərdə iştirak
	Yoxdur
10	Layihə mövzusu üzrə elmi məruzələr (seminarlar, konfranslar, dəyirmi masalar və s. çıxışlar)
	Azərbaycan Respublikası Elm və Təhsil Nazirliyi İnformasiya Texnologiyaları İnstitutunda layihə rəhbərinin iştirakı ilə dəfələrlə seminarlar, dəyirmi masalar keçirilmişdir.
11	Layihə üzrə əldə olunmuş cihaz, avadanlıq və qurğular, mal və materiallar
	Yoxdur
12	Yerli həmkarlarla əlaqələr
	Azərbaycan Respublikası Elm və Təhsil Nazirliyi İnformasiya Texnologiyaları İnstitutunun müvafiq şöbələrinin əməkdaşları (layihə icarçısı olmayan) ilə birgə seminarlar və dəyirmi masalar keçirilmişdir.
13	Xarici həmkarlarla əlaqələr
	Türkiyənin İzmir şəhərində yerləşən Dokuz Eylül Universitetinin Kompüter Elmləri kafedrasının əməkdaşları prof. Dr. Efendi Nasiboğlu, prof. Dr. Çağın Kandemir Çavaş,

	prof. Dr. Emel Kuruoğlu Kandemir, Doç. Dr. Resmiye Nasiboğlu, dr. Erdem Alkım, dr. Övgü Kınay, doktorant Mikayıl Sadıgzade, Araştırma Görevlisi Süheyla Uygur ilə müzakirələr aparılmışdır.
14	Layihə mövzusu üzrə kadr hazırlığı Elmlər doktoru, fəlsəfə doktoru və magistr hazırlığı proqramları üzrə doktorantlar və magistrantlar hazırlanır.
15	Sərgilərdə iştirak Yoxdur
16	Təcrübəartırmada iştirak və təcrübə mübadiləsi Yoxdur
17	Layihə mövzusu ilə bağlı elmi-kütləvi nəşrlər, kütləvi informasiya vasitələrində çıxışlar, yeni yaradılmış internet səhifələri və s. www.ict.az saytında informasiya verilmişdir.

Layihə rəhbərinin imzası _____ Abdullayeva Fərqanə Cabbar qızı

Tarix _____

QEYD: bütün hallarda uyğun olan bəndlər doldurulmalıdır.