Տվյալների գաղտնիության և էթիկայի խնդիրները big data analytics-ում

**Թամարա Խաչատրյան**

Թվային դարաշրջանում տվյալների քանակը աճում է աննախադեպ տեմպերով՝ ձևավորելով նոր տնտեսական ու սոցիալական իրականություն։ Մեծ տվյալների վերլուծությունը դարձել է առանցքային գործիք տարբեր ոլորտներում՝ սկսած բիզնես գործընթացների օպտիմալացումից ու առողջապահական համակարգերի բարելավումից, մինչև ֆինանսական ոլորտում խարդախության հայտնաբերում ու կանխարգելում։ Մեծ տվյալների 5 հիմնական բնութագրերը՝ ծավալը (Volume), բազմազանությունը (Variety), արագությունը (Velocity), ճշգրտությունը (Veracity) և արժեքը (Value), թույլ են տալիս կազմակերպություններին իրական ժամանակում մշակել հսկայական ու բազմաբնույթ տվյալներ՝ ստանալով արժեքավոր պատկերացումներ ու ընդունելով հիմնավորված որոշումներ։ Այնուամենայնիվ, մեծ տվյալների հնարավորությունների հետ մեկտեղ առաջանում են նաև գաղտնիության ու էթիկայի լուրջ խնդիրներ։

Հոդվածում բացահայտել ենք մեծ տվյալների վերլուծության կիրառմամբ ֆինանսական խարդախության հայտնաբերման ու կանխարգելման արդյունավետ մեթոդները՝ օգտագործելով քանակական ու որակական տվյալներ։ Աշխատանքում ներկայացված են տվյալների գաղտնիության ու էթիկայի խնդիրների հիմնական մարտահրավերները, ինչպես նաև դրանց հաղթահարման տեխնոլոգիական ու իրավական լուծումները։ Հոդվածում մշակել ենք Fraud Detection համակարգ, որն ավտոմատացված կերպով հետևում ու հայտնաբերում է կասկածելի գործարքները: Արդյունքները հնարավորություն կտան բարձրացնել թվային գործարքների անվտանգությունը և վստահելիությունը, ընդլայնել տվյալների պատասխանատու կառավարման պրակտիկան ու խթանել էթիկական մոտեցումները մեծ տվյալների վերլուծության ոլորտում։

**Բանալի բառեր։** տվյալների գաղտնիություն, տվյալների էթիկա, անվտանգություն, Մեծ տվյալ, խարդախություն

**Ներածություն:** Բրիտանացի գիտնական, գյուտարար, համաշխարհային ցանցի ստեղծող Թիմ Բերներս-Լի-ն նշել է․ «Տվյալները թանկարժեք բան են ու կգոյատևեն ավելի երկար, քան համակարգերը»՝ ընդգծելով տվյալների արժեքն ու դրանց հետ պատասխանատվությամբ վարվելու շատ ավելի մեծ անհրաժեշտությունը։ Տվյալների մեծ քանակի հավաքագրումը, պահպանումն ու վերլուծությունը առաջացնում են ռիսկեր, որոնք կապված են անհատական կյանքի անձեռնմխելիության խախտման, տվյալների չարաշահման և ալգորիթմական կողմնակալության հետ։ Ինչպես նշել է Օնտարիոյի տեղեկատվության ու գաղտնիության նախկին հանձնակատար Դոկտոր Անն Քավուկյանը 2010 թվականին․ «Գաղտնիությունը միայն գաղտնիության մասին չէ, այլ մեր տվյալների հետ պատահող իրավիճակի վերահսկողության մասին է»։

Այսօր ավելի հաճախ են արձանագրվում տվյալների արտահոսքի դեպքեր, վերահսկողական տեխնոլոգիաների չարաշահումներ, ինչպես նաև որոշումների ընդունման գործընթացներում թափանցիկության բացակայություն։ Այս երևույթներն առաջացնում են հասարակության վստահության անկման վտանգ ու ընդգծում են տվյալների էթիկական օգտագործման անհրաժեշտությունը։ **Հոդվածի նպատակն է** ուսումնասիրվել են մեծ տվյալների վերլուծության կիրառման հնարավորություններն ու սահմանափակումները՝ տվյալների գաղտնիության ու էթիկայի համատեքստում, հատուկ շեշտադրումով ֆինանսական խարդախության կանխարգելման վրա։Անհրաժեշտ է բացահայտել մեծ տվյալների պաշտպանման տեխնոլոգիական, իրավական ու էթիկական խնդիրները, ներկայացնել դրանց համահունչ գործիքներն ու մեթոդները, ինչպես նաև ուսումնասիրել տվյալների անվտանգ ու պատասխանատու մշակման կարևորությունը ֆինանսական խարդախության հայտնաբերման ու կանխարգելման գործում։ Հոդվածում համադրելով supervised learning-ի տարբեր մեթոդներ՝ կառուցել ենք խարդախությունների հայտնաբերման արդյունավետ մոդել, իսկ կիրառված վիճակագրական ու գրաֆիկական գործիքները ապահովել են արդյունքների ճշգրիտ վերլուծություն և համեմատություն։

Թվային դարաշրջանում տվյալները դարձել են համաշխարհային տնտեսության, նորարարության, որոշումների կայացման գործընթացի բարելավման և ոլորտների վերաձևավորման կենսական նշանակություն ունեցող աղբյուրը:[[1]](#footnote-1) Վերլուծություններում տվյալների օգտագործման էթիկական ենթատեքստը վերաբերում է թափանցիկությանը, արդարությանը և հաշվետվողականությանը՝ ապահովելու համար, որ տվյալները ճիշտ օգտագործվեն և յուրաքանչյուր անձի իրավունքը պաշտպանված լինի։[[2]](#footnote-2) Ժամանակակից իրավաբանները առաջ են քաշել անձնական կյանքի փիլիսոփայական մոտեցումը՝ հիմնվելով այն գաղափարի վրա, որ սոցիալական անհատականության, ինքնության և քաղաքացիության պատշաճ զարգացումը, կարևորագույն առումով, կախված է անձնական կյանքի պաշտպանությունից, ներառյալ թվային և տեղեկատվական ազդեցությունը:[[3]](#footnote-3) Մեծ տվյալները ավելին են, քան պարզապես փոխկապակցված տվյալների բազայի աղյուսակներ, դրանք փող են և իշխանություն:[[4]](#footnote-4) Մեծ Տվյալներն ունեն տեղեկատվական անվտանգությանը ազդող հատուկ բնութագրեր՝ բազմազանություն, ծավալ, արագություն, արժեք, փոփոխականություն ու ճշգրտություն:[[5]](#footnote-5) Մեծ տվյալների վերլուծությունում օգտագործվում են տարբեր գործիքներ՝ ընկերություններին օգնելու խնայել ժամանակ և գումար, որոնք նրանց հնարավորություն են տալիս օպտիմալացնել մեքենայական ուսուցման համակարգերը և լուծել իրենց մեծ տվյալների կարիքները նորարարական եղանակներով, միաժամանակ ստանալով բիզնես որոշումները կողմնորոշող պատկերացումներ:[[6]](#footnote-6) Հայտնաբերելով կազմակերպության մեծ ծավալի տվյալների խորքում թաքնված չարամիտ գործունեության օրինաչափությունները, մեծ տվյալների անվտանգության գործիքները կարող են իսկապես հնարավորություն տալ բիզնեսներին ավելի լավ հասկանալ, թե արդյոք և ինչպես են ենթարկվել հարձակման:[[7]](#footnote-7) Խարդախ գործողություններին արձագանքելով՝ կազմակերպությունները կարող են արգելափակել կասկածելի գործարքները, սառեցնել խարդախ գործունեության հետ կապված հաշիվները և տեղեկացնել շահագրգիռ կողմերին համապատասխան գործողություններ ձեռնարկելու համար:[[8]](#footnote-8)

Մեծ տվյալները այսօր ստացել են **«նոր նավթ»** մականունը՝ բիզնեսի աճի ու նորարարության խթանման գործում իրենց դերի համար:

**Մեծ տվյալների առավելություններն** են․ 1/ **բարելավված որոշումների կայացումը** 2/ **հաճախորդների վարքագծի ուսումնասիրությունը,** 3/ **գործառնական արդյունավետության բարձրացում**ը, 4/ **օպտիմիզացված գնագոյացում**ը, **5/ ռիսկերի կառավարումն ու խարդախության հայտնաբերումը**:[[9]](#footnote-9)

**Քեթի Օ'Նիլն** /ամերիկացի մաթեմատիկոս, տվյալագետ և հեղինակ/ իր ելույթում ընդգծել է մեծ տվյալների վերլուծության էթիկական և գաղտնիության հետ կապված **մտահոգությունները**:

**1. Ալգորիթմները օբյեկտիվ չեն -** ալգորիթմները չեզոք գործիքներ չեն, այլ «կոդի մեջ ներկառուցված կարծիքներ են», որոնք արտացոլում են իրենց ստեղծողների կողմնակալություններն ու ընտրությունները՝ հավերժացնելով առկա անհավասարությունները։

**2. Թափանցիկության և հաշվետվողականության պակաս,**

**3. Ալգորիթմական կողմնակալություն -** խտրականությունը այնպիսի ոլորտներում, ինչպիսիք են վարձումը, վարկային պատմությունը և քրեական արդարադատությունը, կարող են հավերժացնել սոցիալական անհավասարությունները և զգալի բացասական ազդեցություն ունենալ մարդկանց կյանքի վրա,

4․ Նպատակի սահմանափակում - մի նպատակով հավաքված տվյալները չպետք է օգտագործվեն մեկ այլ նպատակով՝ առանց պատշաճ համաձայնության։

**5. Ալգորիթմական աուդիտների անհրաժեշտությունը** - Օ'Նիլը առաջարկում է ալգորիթմական աուդիտները որպես ալգորիթմների արդարացիության ու ճշգրտության գնահատման միջոց: Աուդիտների ընթացքում պետք է ուսումնասիրվեն օգտագործված տվյալները, հաջողության սահմանումը, սխալների հավանականությունն ու հասարակության համար երկարաժամկետ հետևանքները:[[10]](#footnote-10)

**Մեծ տվյալների պաշտպանության** համար անհրաժեշտ է համադրել **տեխնիկական**, **կազմակերպչական** ու **իրավական** մեթոդներ։

**1․** Անվտանգության **տեխնիկական միջոցառումներն են․**

1/ **Encryption /գաղտնագրում/[[11]](#footnote-11)**

2/ **Access Control /մուտքի վերահսկողություն/**[[12]](#footnote-12)

3/ **Anonymization and Pseudonymization /անանունացում և կեղծանունավորում/,**

Տվյալների վերանույնականացումը կանխելու համար ներդրվել են K-anonymity /K-անանունության/, L-diversity /L-բազմազանության/ և T-closeness /T-մոտության/ հասկացությունները:

*Աղյուսակ 1․ Հիվանդների տվյալները պարունակող ոչ անանունացված աղյուսակ*[[13]](#footnote-13)

|  |  |  |  |
| --- | --- | --- | --- |
| **Անուն** | **Տարիք** | **Փոստային ինդեքս** | **Բժշկական վիճակ** |
| Վլադիսլավ | 29 | 47677 | Սրտի հիվանդություն |
| Ալեքսեյ | 24 | 47602 | Սրտի հիվանդություն |
| Զախար | 28 | 47905 | Գրիպ |
| Ռամիլ | 27 | 47909 | Գրիպ |
| Օլգա | 24 | 47607 | Գրիպ |

*К-anonymity*- 1/ ընդհանրացում։ «տարիք» ատրիբուտը կարող է գրվել ավելի ընդհանուր ու լայն իմաստով։ 2/ Ճնշում։ որոշ ատրիբուտային արժեքներ թաքցվում են \* նշանի միջոցով։

*Աղյուսակ 2․ Հիվանդների տվյալները պարունակող К-anonymity աղյուսակ*[[14]](#footnote-14)

|  |  |  |  |
| --- | --- | --- | --- |
| **Անուն** | **Տարիք** | **Փոստային ինդեքս** | **Հիվանդություն** |
| Վլադիսլավ | 20 < տարիք ≤ 30 | 476\*\* | Սրտի հիվանդություն |
| Ալեքսեյ | 20 < տարիք ≤ 30 | 476\*\* | Սրտի հիվանդություն |
| Զախար | 20 < տարիք ≤ 30 | 479\*\* | Գրիպ |
| Ռամիլ | 20 < տարիք ≤ 30 | 479\*\* | Գրիպ |
| Օլգա | 20 < տարիք ≤ 30 | 476\*\* | Գրիպ |

*L-diversity*– եթե խումբը ունի 10 գրառում, և բոլորն ունեն «սրտի հիվանդություն», ապա նույնիսկ եթե k-anonymity ապահովված է, գաղտնիությունը խախտվում է, հետևաբար L-diversity-ն պահանջում է, որ այդ խմբում լինեն գոնե L տարբեր հիվանդություններ։

*Աղյուսակ 3․ Հիվանդների տվյալները պարունակող L-diversity աղյուսակ*[[15]](#footnote-15)

|  |  |  |
| --- | --- | --- |
| **Տարիք** | **Աշխատավարձ** | **Հիվանդություն** |
| 29 | 3k | Գրիպ |
| 22 | 4k | Գաստրիտ |
| 23 | 5k | Թոքաբորբ |
| 52 | 6k | Գրիպ |
| 43 | 11k | Գաստրիտ |
| 36 | 8k | Թոքաբորբ |

*T-closeness* **-** սա L-diversity-ի հետագա կատարելագործում է, որը հիմնված է անանունացման վրա ու օգտագործվում է տվյալների բազմություններում գաղտնիությունը պահպանելու համար՝ նվազեցնելով տվյալների ներկայացման մանրամասնությունը։

*Աղյուսակ 4․ Հիվանդների տվյալները պարունակող T-closeness աղյուսակ*[[16]](#footnote-16)

|  |  |  |
| --- | --- | --- |
| **Տարիք** | **Աշխատավարձ** | **Հիվանդություն** |
| 2\*  2\*  2\* | 3k  4k  5k | Գրիպ  Գաստրիտ  Թոքաբորբ |
| ≥40  ≥40  ≥40 | 6k  11k  8k | Գրիպ  Գաստրիտ  Թոքաբորբ |

Դիֆերենցիալ գաղտնիությունը գաղտնիության սպառնալիքների դեմ պայքարի արդյունավետ մեթոդներից մեկն է, որտեղ վերլուծաբանը չունի տեղեկատվությանը անմիջական մուտք, փոխարենը օգտագործվում է միջնորդ, որը ծառայում է որպես գաղտնիության պահապան, վերջինս էլ ընդունում է վերլուծաբանի հարցումները և վերադարձնում արդյունքը՝ փոքր աղավաղումներով։[[17]](#footnote-17)

4/ **Integrity /տվյալների ամբողջականության վերահսկում/ -** Hash ֆունկցիաներ (օրինակ՝ SHA-256)՝ ստուգելու համար տվյալների չփոփոխված լինելը։

5/ **Anomaly Detection /անոմալիաների հայտնաբերում/,**[[18]](#footnote-18)

6/ Ցանցային ակտիվության մոնիթորինգ[[19]](#footnote-19)

7/ Backup and Recovery /տվյալների պահուստավորում և վերականգնում /[[20]](#footnote-20)

8/ **Հոմոմորֆ կոդավորում** - հնարավորություն է տալիս հաշվարկներ կատարել կոդավորված տվյալների վրա՝ թույլ տալով վերլուծություն կատարել առանց անհատական ​​գրառումները վերծանելու ու բարձրացնելով տվյալների անվտանգությունը։[[21]](#footnote-21)

**2․** Մեծ տվյալների պաշտպանության **կազմակերպչական միջոցառումներն են․**

**1/ Data Governance /տվյալների կառավարում/,**

2/ **Աշխատակիցների ուսուցում,**

3/ **Աուդիտ և մոնիթորինգ,**

4/ Ֆիզիկական անվտանգություն – նպատակն է սահմանափակել սերվերներին ու տվյալների կենտրոններին ֆիզիկական մուտքը և իրականացնել անվտանգության միջոցառումներ՝ սարքավորումները պաշտպանելու համար:[[22]](#footnote-22)

**3․** Անվտանգության **իրավական մեթոդներ են․**

1/ **Տվյալների պաշտպանության ընդհանուր կանոնակարգ (GDPR),**

2/ **Կալիֆոռնիայի սպառողների գաղտնիության մասին օրենք (CCPA),**

3/ **Տվյալների պաշտպանության այլ տարածաշրջանային և ազգային օրենքներ։**[[23]](#footnote-23)

**GDPR (General Data Protection Regulation)՝** Ընդհանուր տվյալների պաշտպանության կանոնակարգը ո**ւժի մեջ մտել է** 2018 թվականի մայիսի 25-ին, ԵՄ տարածքում։[[24]](#footnote-24) **Ըստ GDPR կանոնակարգի Հոդված 5-ի՝** **անձնական տվյալների մշակմանը վերաբերող սկզբունքներն են՝ 1/ օրինականություն, արդարություն և թափանցիկություն**, 2/ ն**պատակի սահմանափակում**, 3/ **data minimization, 4/ ճշտություն**, 5/ պ**ահպանման ժամկետի սահմանափակում**, 6/ ա**նվտանգություն և ամբողջականություն**, 7/ պ**ատասխանատվություն**։[[25]](#footnote-25)

Երբ կազմակերպությունները խախտում են տվյալների պաշտպանության կանոնները, GDPR-ը նախատեսում է տուգանքներ՝ կախված խախտման ծանրությունից։ Ամենածանր խախտումների դեպքում /օրինակ՝ տվյալների իրավունքների անտեսում/՝ կազմակերպությունը կարող է տուգանվել մինչև €20 միլիոնով, կամ մինչև իր տարեկան համաշխարհային շրջանառության 4%-ով՝ կախված նրանից, թե որն է ավելի մեծ։ Փոքր խախտումների համար /օրինակ՝ ընթացակարգային կանոնների խախտումը/ տուգանքը կարող է լինել մինչև €10 միլիոն, կամ շրջանառության 2%-ը։[[26]](#footnote-26)

GDPR-ի թերություններն են՝

1/ **կիրառման բարդությունը - փ**ոքր ու միջին բիզնեսների համար չափազանց բարդ ու թանկ է բոլոր պահանջները բավարարելը,

2/ **Big Data-ի հետ անհամատեղելիություն որոշ դեպքերում -** GDPR-ը նախատեսված է տվյալների հավաքման հստակ նպատակների համար, բայց Big Data-ում տվյալները հաճախ հավաքվում են ապագա վերլուծությունների համար, որոնց նպատակը նախապես հստակ չէ,

3/ տ**եխնոլոգիական առաջընթացից հետ մնալը - օ**րենքը չի հասցնում հարմարվել նոր AI և մեքենայական ուսուցման պրոֆիլավորման մեթոդներին, որոնք կարող են անուղղակիորեն բացահայտել անձնական տվյալներ,

4/ կ**իրառման անհավասարություն ԵՄ երկրներում - չ**նայած GDPR-ը միասնական փաստաթուղթ է, տարբեր երկրներում վերահսկող մարմինների խստությունը և կարողությունները տարբեր են,

5/ հ**սկողության թույլ մեխանիզմներ ԵՄ-ից դուրս գտնվող ընկերությունների համար - ե**թե տվյալները մշակվում են ԵՄ-ից դուրս, վերահսկողությունը գրեթե անհնար է,

6/ հ**ստակ մեթոդաբանության բացակայություն** - չկա կոնկրետ տեխնիկական ցուցակ, թե ինչն է համարվում «բավարար» անվտանգություն (օր՝ կոդավորման ստանդարտներ, անոնիմիզացիայի մեթոդներ)։

Big Data համատեքստում այս թերությունների լուծման համար անհրաժեշտ է՝

1/ ***Big Data-ի համար հստակ ուղղեցված կանոններ* - մ**շակել հավելյալ ուղեցույցներ, որոնք թույլ կտան տվյալների հավաքում ու վերլուծություն ապագա կանխատեսումների համար՝ առանց օրենքի խախտման,

2/ *տ****եխնիկական ստանդարտների ներդրում* - կ**ազմակերպություններին տրամադրել պարտադիր տեխնիկական ստանդարտների ցանկ՝ տվյալների կոդավորման, անանունացման ու դեանոնիմիզացիայի դեմ պաշտպանության համար։

3/ ***ՓՄՁ-ների համար պարզեցված ռեժիմ* - փ**ոքր կազմակերպությունների համար սահմանել պարզեցված ու մատչելի պահանջներ, որպեսզի նրանք կարողանան պահպանել օրենքը առանց չափազանց մեծ ծախսերի,

4/ ***ավտոմատացված վերահսկողության համակարգեր* - մ**շակել AI/ML գործիքներ, որոնք ավտոմատ կստուգեն տվյալների մշակման համապատասխանությունը GDPR-ին,

5/ *տ****եղեկացվածության բարձրացում և ուսուցում* – կազմակերպել պ**ետական և մասնավոր հատվածում աշխատող անձանց համար տվյալների պաշտպանության մասին պարբերական ուսուցումներ։

6/ *տարածաշրջանային ու միջազգային վերահսկողության ամրապնդում* - ԵՄ ներսում ներդնել համաժամեցված վերահսկողության մեխանիզմներ՝ ապահովելով ֆինանսավորման հավասարեցում ու համատեղ վերապատրաստման ծրագրեր։

**CCPA** **(California Consumer Privacy Act)՝** Կալիֆորնիայի Սպառողների Գաղտնիության Օրենքը ԱՄՆ-ի Կալիֆորնիա նահանգի անձնական տվյալների պաշտպանության մասին օրենք է, որը ուժի մեջ է մտել 2020 թվականի հունվարին։ Այն նման է ԵՄ-ի GDPR-ին, բայց մշակվել է Կալիֆորնիայի բնակիչների համար։

Ըստ CCPA օրենքի բաժին 1798.140.-ի՝ օրենքը կիրառելի է բոլոր այն կազմակերպությունների նկատմամբ, որոնք համապատասխանում են հետևյալ շեմերից մառնվազն մեկին՝

1. տարեկան համախառն եկամուտը գերազանցում է **$25 միլիոն**,
2. հավաքել են կամ օգտագործում են **50,000+ Կալիֆորնիայի բնակիչների տվյալներ**,
3. հավաքած տվյալների վաճառքից ստանում են **50%-ից ավել** եկամուտ։[[27]](#footnote-27)

Օրենքը նախատեսում է վարչական տուգանքներ, երբ ընկերությունը խախտում է CCPA-ի դրույթները․ մինչև $2,500 ոչ դիտավորյալ խախտման համար, և $7,500՝ դիտավորյալ խախտումների համար։[[28]](#footnote-28)

CCPA-ի թերություններն են՝

1/ Հստակ սահմանումների բացակայություն,[[29]](#footnote-29)

2/ Կ**իրառման սահմանափակում - փ**ոքր և միջին բիզնեսները կարող են մեծ քանակությամբ տվյալներ հավաքել, բայց օրենքի շեմերը չեն տարածվում նրանց վրա։

3/ **Մեծ տվյալների համար ոչ բավարար հարմարվողականություն - օ**րենքը թույլ է տալիս տվյալների երկրորդային օգտագործում առանց լրացուցիչ հստակ համաձայնության, եթե այն համատեղելի է «սկզբնական նպատակին»։

4/ Հսկողության և կիրառման թույլ մեխանիզմներ - գլխավոր վերահսկող մարմինը Կալիֆոռնիայի գլխավոր դատախազն է, որը չունի բավարար ռեսուրսներ ամբողջությամբ վերահսկելու հազարավոր ընկերությունների համապատասխանությունը։

5/ Թվային հետևման (cookies, tracking pixels) թույլ կարգավորում։

Այս թերությունների լուծման համար անհրաժեշտ է՝

1/ Սահմանումների հստակեցում,

2/ Կիրառման շեմերի նվազեցում,

3/ Երկրորդային օգտագործման խստացում,

4/ Հստակ վերահսկող մարմնի ստեղծում,

5/ Թվային հետևման խիստ կանոններ։

**GDPR** ու **CCPA** կարևոր քայլեր են տվյալների պաշտպանության և էթիկայի ուղղությամբ, սակայն դրանք դեռևս չեն հասցնում լուծել **մեծ տվյալների վերլուծության** ու **արհեստական բանականության** հետ կապված բոլոր մարտահրավերները։ Առաջարկվող բարելավումները կարող են նպաստել ավելի արդար, թափանցիկ ու անվտանգ տվյալային համակարգերի ստեղծմանը։

Շատ երկրներ, ինչպիսիք են Կանադան, Ճապոնիան, Ավստրալիան, Սինգապուրը, Բրազիլիան և այլք, ունեն տվյալների պաշտպանության համապարփակ օրենքներ, որոնք բավականին նման են GDPR-ին:[[30]](#footnote-30)

ՀՀ **«Տեղեկատվության ազատության մասին» օրենքը**՝ ընդունված 2003 թվականի սեպտեմբերի 23-ին, ամրապնդում է ՀՀ քաղաքացիների՝ **հասարակական տեղեկատվություն ստանալու իրավունքը**։[[31]](#footnote-31) Այս օրենքում մեծ տվյալները չեն կարգավորվում որպես առանձին կատեգորիա, այն հիմնականում կենտրոնանում է **ինֆորմացիայի մատչելիության** վրա, բայց ոչ **մշակման, վերլուծման կամ ավտոմատ որոշումների էթիկայի** վրա։

**«Անձնական տվյալների պաշտպանության մասին»** ՀՀ օրենքը՝ ընդունված 2015 թվականի մայիսի 18-ին, կարգավորում է պետական կառավարման կամ տեղական ինքնակառավարման մարմինների, հիմնարկների և կազմակերպությունների, իրավաբանական ու ֆիզիկական անձանց կողմից անձնական տվյալները մշակելու, դրանց նկատմամբ պետական հսկողություն իրականացնելու կարգն ու պայմանները:[[32]](#footnote-32) Այս օրենքը սահմանափակում է այն դեպքերը, երբ անձնական տվյալները կարող են մշակվել, սակայն **չի կարգավորում AI վերլուծությունների, ավտոմատ որոշումների, կամ ապանձնացված տվյալների վերադարձի ռիսկերը**, և չկան տվյալների վերանույնականացման արգելքի վերաբերյալ դրույթներ։

**Ըստ «Բանկային գաղտնիքի մասին» ՀՀ օրենքի՝** ընդունված 7 հոկտեմբերի 1996թ․, բանկային գաղտնիք են համարվում բանկի հաճախորդին սպասարկելու կապակցությամբ տվյալ բանկին հայտնի դարձած հաճախորդի հաշիվների վերաբերյալ տեղեկությունները, հաճախորդի հանձնարարությամբ կամ հօգուտ հաճախորդի կատարված գործառնությունների վերաբերյալ տեղեկությունները, որոնց հրապարակումը արգելվում է:[[33]](#footnote-33)

Այսպիսով՝ ՀՀ իրավական դաշտը անձնական տվյալների պաշտպանության մասով ունի բավականին լավ հիմքեր՝ համաչափության, օրինականության, հավաստիության սկզբունքներով։ Սակայն մեծ տվյալների, արհեստական բանականության ու ապագայում կանխատեսումների բարոյագիտական խնդիրները դեռևս ամբողջությամբ չեն կարգավորվում։ Դրա համար պետք է հստակեցնել սահմանները, որ **ավտոմատացված վերլուծությունները չգերակայեն մարդու իրավունքներին ու ազատությանը**։

Քանի որ տվյալները փոխում են մեր կյանքի ու բիզնեսի վարման ձևը, դիտարկենք տվյալագիտության վիճահարույց փորձերը, որոնք հարցեր են առաջացրել մեծ տվյալների էթիկայի վերաբերյալ։

**1. Target-ի հղիության կանխատեսումը -** Մինեապոլիսի մի տղամարդ իմացավ իր դստեր հղիության մասին, քանի որ Target-ը արդյունավետորեն թիրախավորել էր իր դստեր մանկական ապրանքների գնումները և կանխատեսել հղիությունը։ Կանխատեսողական վերլուծության նախագիծը հաջողված էր, բայց այն ընդգծեց տեղեկատվության հնարավոր չարաշահումը։

**2. Մեծ տվյալների էթիկան և արհեստական ​​բանականության «Գեղեցկության մրցույթը»** - 2016 թվականին արհեստական բանականության կողմից գնահատված առաջին գեղեցկության մրցույթը ինտերնետից ընտրեց 44 հաղթողի: Ավելի քան 100 երկրներից ներկայացված 6000 լուսանկարներից միայն մի քանիսն էին ոչ սպիտակամորթ: Ոչ սպիտակամորթ հաղթողները ասիացիներ էին: Ակնհայտ խնդիրն այն էր, որ ներկայացված լուսանկարների մեծ մասը Աֆրիկայից ու Հնդկաստանից էին,։ Այս ինտերնետային գեղեցկության մրցույթը կազմակերպած էր Beauty․AI ընկերության կողմից, որի ​​գլխավոր տնօրեն Ալեքս Ժավորոնկովը պնդեց, որ օգտագործված ալգորիթմը կողմնակալ է, քանի որ այն տվյալները, որոնց վրա նրանք ուսուցանել են, բավականաչափ բազմազան չեն եղել: Ապագա նախագծերի համար հույս կա շտկել կողմնակալության խնդիրը՝ օգտագործելով տվյալների ավելի շատ հավաքածուներ ու մշակելով ալգորիթմներ, որոնք կարող են բացատրել ցանկացած կողմնակալություն:

**3. Ինքնավար մեքենաներ -** մեծ տվյալների էթիկայի դեպք տեղի ունեցավ 2018 թվականի սկզբին, երբ Uber-ի ինքնագնաց մեքենան վրաերթի ենթարկեց ու սպանեց մի արիզոնացի կնոջ։ Ինքնագնաց մեքենաները նախագծվում են նման վթարներից խուսափելու համար, ուստի սա լուրջ էթիկական դիլեմաներ առաջացրեց այդ մեքենաների համար նախագծվող ալգորիթմների վերաբերյալ։

**4. Նորթփոյնթի ռիսկի գնահատում** - Միացյալ Նահանգներում դատական ​​համակարգերը կախված են ալգորիթմներից՝ հանցագործների շրջանում հանցագործության կրկնության հավանականությունը որոշելու համար: «Նորթփոյնթ» ընկերությունն ունի COMPAS (Correctional Offender Management Profiling for Alternative Sanctions) կոչվող ծրագիր, որը հանցագործներին գնահատում է ապագայում նրանց կրկին հանցագործություն կատարելու հավանականության հիման վրա. որքան բարձր է միավորը, այնքան մեծ է կրկին հանցագործություն կատարելու հավանականությունը: ProPublica-ի կողմից 2016 թվականին կատարված վերլուծության մեջ, որը կատարվել է 10,000 հանցագործների ուսումնասիրությունից հետո, սևամորթ մեղադրյալներին «սխալմամբ» ավելի հաճախ են բարձր միավորներ տրվել, քան սպիտակամորթներին: Ավելին, սպիտակամորթ հանցագործներին հաճախ տրվում էր ավելի ցածր գնահատական, քան պետք է ստանային: Նորթփոյնթը հերքել է իրենց ալգորիթմներում առկա ցանկացած ռասայական կողմնակալություն, սակայն այս թեմայի հետ կապված վիճահարույց հարցերը մտահոգություններ են առաջացնում:[[34]](#footnote-34)

**5․ Facebook-ի և Cambridge Analytica-ի սկանդալը** - Cambridge Analytica անունով բրիտանական քաղաքական տվյալների վերլուծական ընկերությունը օգտագործեց միլիոնավոր Facebook օգտատերերի անձնական տվյալներ՝ առանց նրանց համաձայնության, որպեսզի ստեղծի հոգեբանական պրոֆիլներ ու օգտագործի դրանք քաղաքական քարոզչության համար։ Մոտ 270,000 օգտատեր ներբեռնել էին "This Is Your Digital Life" անունով թեստային հավելված, որը հավաքագրում էր ոչ միայն իրենց տվյալները, այլև իրենց ընկերների տվյալները՝ առանց այդ ընկերների գիտության և թույլտվության։ Այս մեթոդով հավաքվել էր 87 միլիոն օգտատերերի անձնական տվյալ, որոնք փոխանցվել էին Cambridge Analytica-ին։ Տվյալները օգտագործվել էին ԱՄՆ-ի 2016թ. նախագահական ընտրությունների ընթացքում՝ նպաստելու Դոնալդ Թրամփի քարոզարշավին, ինչպես նաև Բրեքզիթի քարոզչությանը։ Այս սկանդալը ընդգծեց տվյալներին չարտոնված մուտք գործելու ու չարաշահելու ռիսկերը, ինչպես նաև ցույց տվեց, թե ինչ է պատահում, երբ տվյալների էթիկան չի պահպանվում: Արդյունքում տեղի ունեցավ **Facebook-ի նկատմամբ վստահության անկում, Cambridge Analytica-ի փակում, GDPR ու CCPA կարգավորումների ակտիվացում և Facebook-ը վճարեց** **5 միլիարդ դոլար տուգանք ԱՄՆ FTC-ին։**

**6. Yahoo-ի տվյալների արտահոսք** - Yahoo-ն 2013-2014թթ․ ժամանակահատվածում բազմաթիվ տվյալների արտահոսքի զոհ է դարձել, որոնք ազդել են նրա բոլոր 3 միլիարդ օգտատիրոջ հաշիվների վրա։ Գողացված տեղեկությունները ներառում էին անուններ, էլեկտրոնային փոստի հասցեներ, ծննդյան ամսաթվեր, ինչպես նաև հեշավորված գաղտնաբառեր։ Դեպքն ավելի վատացավ Yahoo-ի կողմից տեղեկատվության ուշացված բացահայտման պատճառով, քանի որ ընկերությունը հրապարակայնորեն չհաստատեց խախտման մասին մինչև 2016 թվականը։

**7․ TikTok-ի դեմ դատական գործեր՝ կապված անչափահասների տվյալների հետ** - պարզվել է, որ հարթակը հաճախ հավաքել է երեխաների տվյալները՝ առանց ծնողների համաձայնության։ ԱՄՆ-ում 2021 թվականին TikTok-ը վճարեց 92 միլիոն դոլար, որպեսզի կարգավորի տվյալների պաշտպանության խախտումների շուրջ դատական գործը։ Իսկ 2023 թվականին Մեծ Բրիտանիայում երկրի տեղեկատվական հանձնակատարի գրասենյակը (ICO) TikTok-ի նկարմամբ կիրառեց 12.7 միլիոն ֆունտ ստերլինգի տուգանք՝ նույն խախտման պատճառով։

Այս օրինակները ընդգծում են տվյալների էթիկայի անտեսման աղետալի հետևանքները:

Ի հակադրություն վերոնշյալ օրինակների՝ **Apple**-ը վաղուց ի վեր դիրքավորվել է որպես գաղտնիության վրա կենտրոնացած ընկերություն: Ընկերությունը նվազագույնի է հասցնում անձնական տվյալների հավաքագրումը, տվյալների մեծ մասը մշակում է օգտատիրոջ սարքում՝ ամպային պահեստում չօգտագործելու փոխարեն, տրամադրում է թափանցիկության հաշվետվություններ ու օգտատիրոջը տալիս է իր տվյալների նկատմամբ զգալի վերահսկողություն: Այս գործելակերպը հանգեցրել է նրան, որ Apple-ը գովաբանվում է որպես գաղտնիության մոդել, որը տվյալների էթիկայի անկյունաքարն է:[[35]](#footnote-35)

**Carnegie Endowment-ի կողմից մշակված “FinCyber Timeline”-ը** բացառիկ աղբյուր է, որը ներկայացնում է ֆինանսական ոլորտի վրա ուղղված կիբերհարձակումների փաստագրված պատկերը։ Այնտեղ կան մոտ **200 փաստագրված դեպքեր**, որոնք վերաբերում են **բանկերի, ապահովագրական ընկերությունների ու բորսաների վրա հարձակումներին**։

1. 2016-ի փետրվարին հաքերները ներթափանցեցին **Բանգլադեշի Կենտրոնական բանկի** ցանց, գողացան SWIFT մուտքային տվյալներն ու փորձեցին փոխանցել մոտ 1 միլիարդ դոլար։ Չորս փոխանցում հաջողվեց՝ 81 մլն դոլար տեղափոխվեց Ֆիլիպիններ, բայց մեծ մասը կասեցվեց սխալների պատճառով։ Հարձակման հետևում, ըստ ՄԱԿ-ի, կանգնած էին UNSC-ի հետ կապ ունեցող անձինք։ Վնասակար ծրագիրը ոչ միայն գողացավ տվյալներ, այլև ջնջեց հետքերն ու մանիպուլյացրեց SWIFT համակարգը։
2. 2018թ․ օգոստոսին հաքերները հարձակվեցին **Հնդկաստանի Cosmos Bank**-ի վրա՝ գողանալով մոտ 13.5մլն դոլար։ Նրանք գողացան քարտերի տվյալներ, տեղադրեցին պրոքսի սերվեր ու ընդամենը մի քանի ժամում 28 երկրների բանկոմատներից կատարեցին շուրջ 15,000 կանխիկացման գործարք, նաև չարտոնված SWIFT փոխանցումներ։ Հարձակման հետևանքով Cosmos-ի առցանց բանկային ծառայությունը մեկ շաբաթից ավելի անջատված էր, ու գումարները չվերականգնվեցին։
3. 2018թ․ հոկտեմբերի 29-ին **Պակիստանի Bank Islami**-ն հայտնաբերեց կիբերհարձակում իր միջազգային վճարային քարտերի ցանցի վրա։ Բանկը հայտնաբերեց կասկածելի գործարքներ Պակիստանից դուրս վճարային քարտերով ու անմիջապես փակեց իր միջազգային վճարային սխեման։ Բանկը հաստատեց, որ հաճախորդների հաշիվներից հանվել է մոտ 2.6միլիոն պակիստանյան ռուփի (համարժեք 19,500 դոլարի)։ Միջադեպից հետո Պակիստանի պետական բանկը (SBP) հրահանգնել է բոլոր բանկերին, որ ապահովեն բոլոր վճարային քարտերի անվտանգությունն ու վերահսկել քարտերի գործունեությունը իրական ժամանակում։[[36]](#footnote-36)

Վերոնշյալ օրինակները ցույց են տալիս, որ բանկերը ենթարկվում են տարբեր տեսակի հարձակումների՝ սկսած մասշտաբային համակարգային գրոհներից մինչև ֆիզիկական խարդախություն ATM-ների միջոցով։

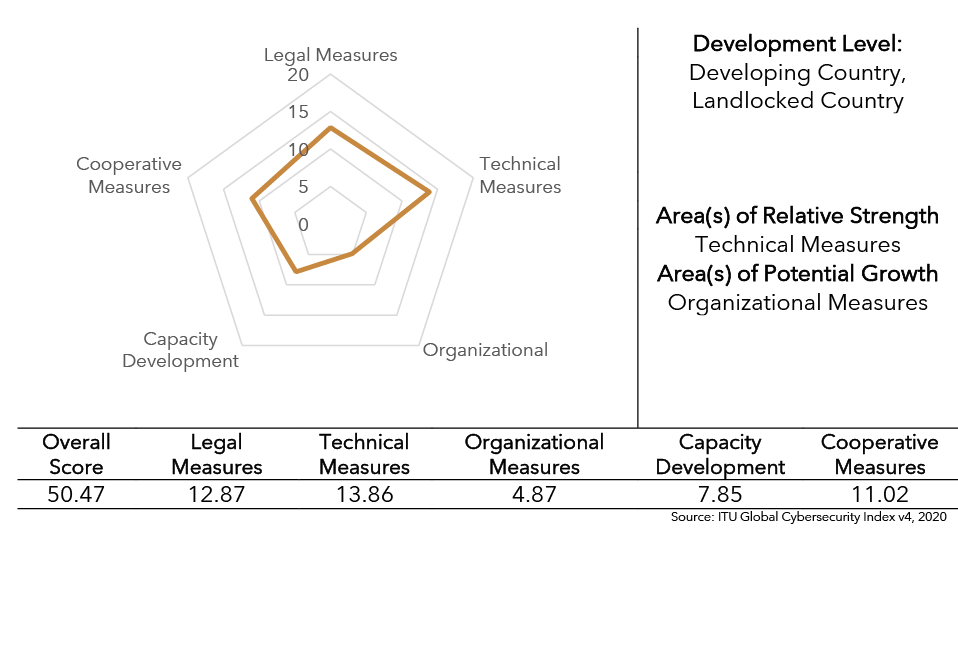
**Global Cybersecurity Index (GCI)**-ը ՄԱԿ-ի (ITU) կողմից մշակված ինդեքս է, որը գնահատում է, թե որքանով են երկրները պատրաստ կիբերանվտանգության սպառնալիքներին։ Ինդեքսը հիմնված է մի քանի հիմնական ուղղությունների վրա՝

1. **Իրավական շրջանակ (Legal Measures) -** օրենքներ, կանոնակարգեր,
2. **Տեխնիկական միջոցառումներ (Technical Measures),**
3. **Կազմակերպչական մոտեցումներ (Organizational Measures) -** պետական ռազմավարություններ, կառույցներ,
4. **Միջազգային համագործակցություն (Cooperation),**
5. **Թիմերի կարողությունների զարգացում (Capacity Development)։**

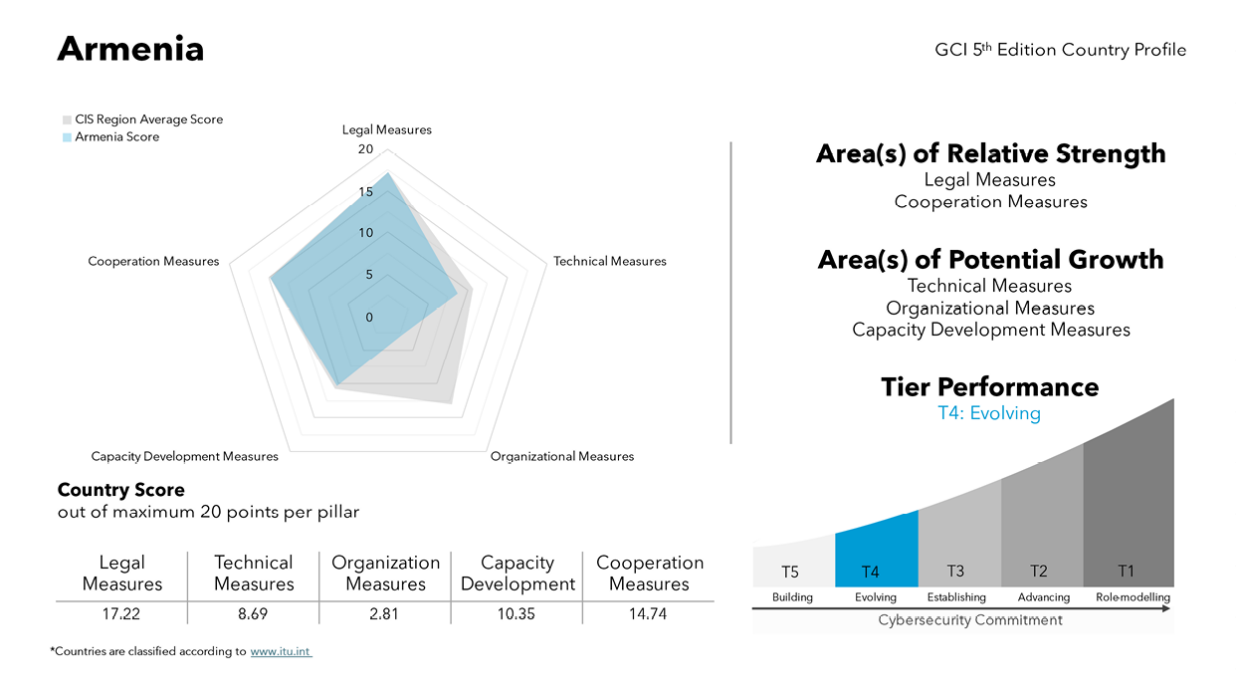
GCI-ն միավորում է հարցումների, վիճակագրության և ազգային քաղաքականությունների տվյալները՝ յուրաքանչյուր երկրի համար ձևավորելով ընդհանուր գնահատական։ 2024 թվականին այն ընդգրկել է Tier մոդել, որը երկրներին դասակարգում է կոնկրետ մակարդակներով։[[37]](#footnote-37)

2015 թվականի գնահատմամբ, Հայաստանի GCI-ը եղել է 0․176, իսկ առաջին հորիզոնականը եղել է ԱՄՆ-ն` 0.824 ինդեքսով։[[38]](#footnote-38) 2017 թվականի գնահատմամբ, Հայաստանի GCI-ը եղել է 0.196՝ զբաղեցնելով 110-րդ հորիզոնականը։[[39]](#footnote-39) 2018 թվականին պատկերը փոխվել է, և Հայաստանի GCI-ը եղել է 0.495՝ աճելով 2015 թվականի նկատմամբ 181․25 տոկոսով և զբաղեցնելով 79-րդ հորիզոնականը։[[40]](#footnote-40)

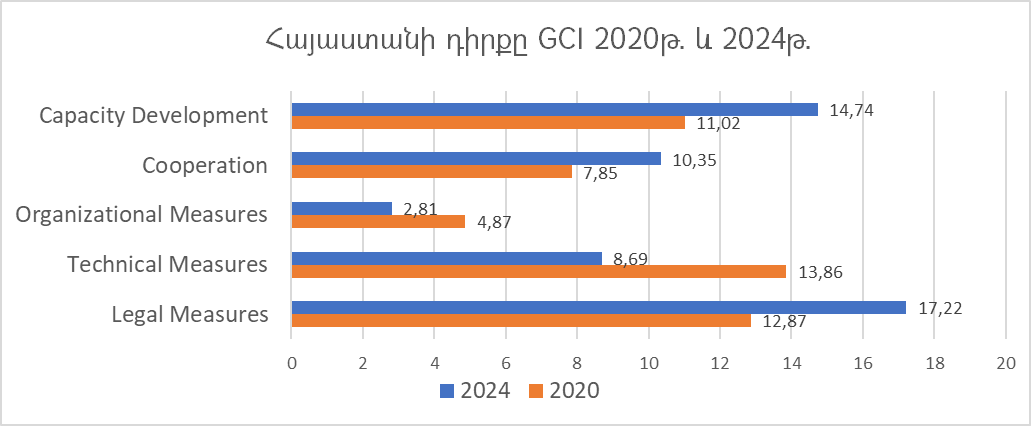
2020 թվականին Հայաստանի GCI-ը ընդհանուր գնահատմամբ՝ 50.47/100 է և ունի հետևյալ տեսքը՝



*Գծապատկեր 1․ Հայաստանի դիրքը GCI 2020-ում[[41]](#footnote-41)*



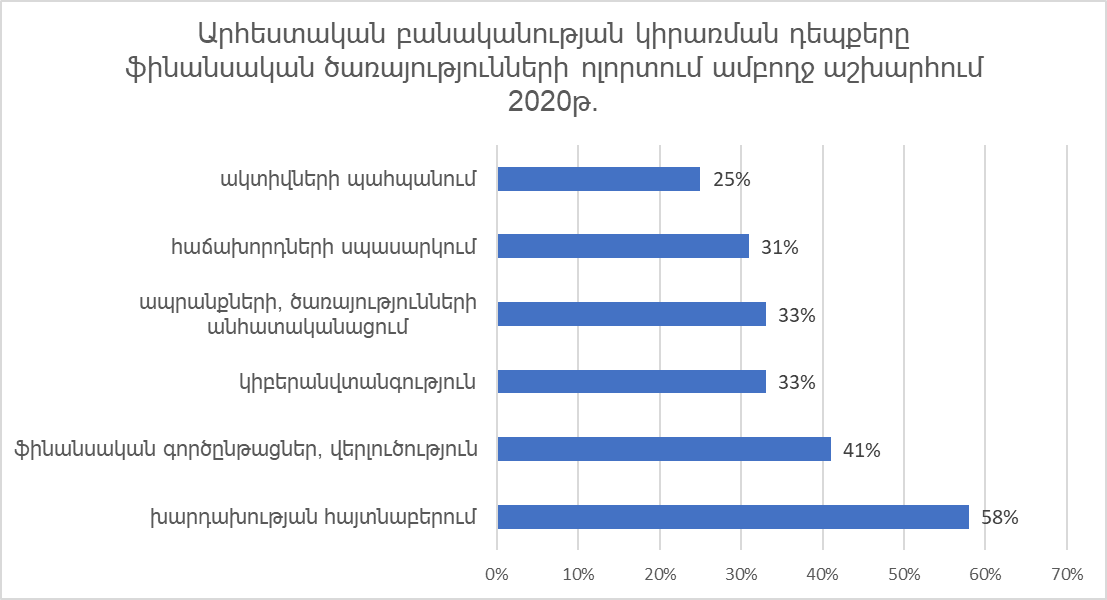
*Գծապատկեր 2․ Հայաստանի դիրքը GCI 2024-ում[[42]](#footnote-42)*



*Գծապատկեր 3․ Հայաստանի դիրքը GCI 2020 և 2024[[43]](#footnote-43)*

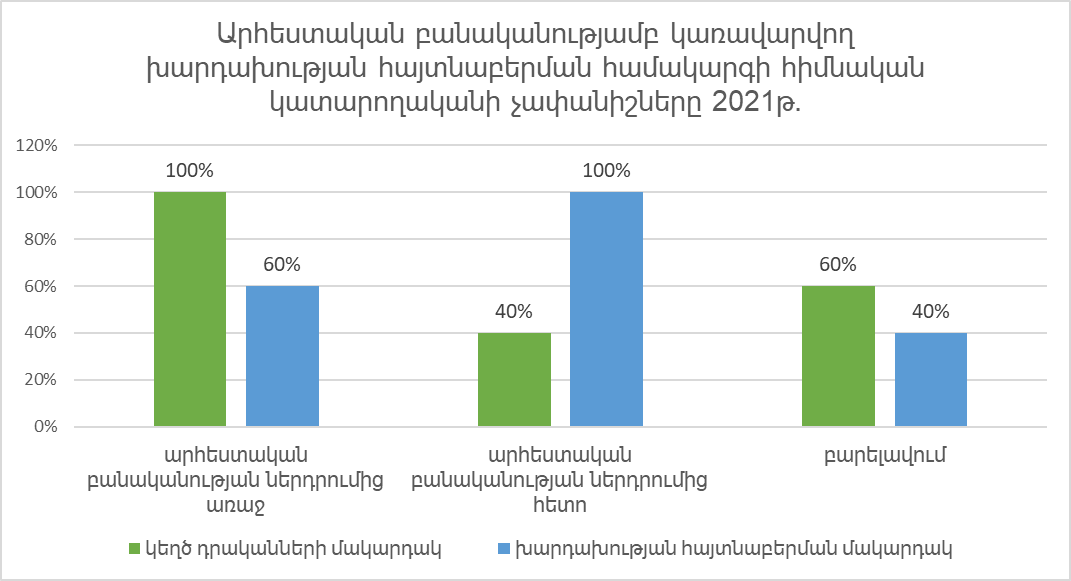
Հայաստանի դիրքը GCI 2024-ում ընդհանուր գնահատմամբ 53.81/100 է՝ աճելով 2015 թվականի նկատմամբ 205․74 տոկոսով, իսկ 2020 թվականի նկատմամբ՝ 6․61 տոկոսով։ 2024 թվականին 2020 թվականի նկատմամբ աճել են Legal Measures /33․79 տոկոս/, Capacity Development /31.84 տոկոս/, Cooperation Measures /33․75 տոկոս/, իսկ Technical Measures և Organizational Measures նվազել են համապատասխանաբար 37․3 և 42․2 տոկոսներով։ Հայաստանի կարգավիճակը՝ **Tier 4 (Evolving)**, ինչը նշանակում է, որ Հայաստանը գտնվում է «ձևավորվող» փուլում ու դեռ չի հասել բարձր մակարդակների։ Այս թիվը ցույց է տալիս, որ Հայաստանը շարունակում է զարգացնել իր կիբերանվտանգության ենթակառուցվածքները, սակայն դեռ զգալի բացեր կան, հատկապես կազմակերպչական և տեխնիկական ոլորտներում, ու խնդիրը հատկապես ծրագրերի իրականացման և կադրերի ժամանակակից ուսուցման պակասի մեջ է:[[44]](#footnote-44)

Այսօր, երբ ֆինանսական գործարքների մեծ մասը կատարվում է թվային եղանակով, բիզնեսները ավելի են ենթարկվում խարդախության: Ֆինանսական ծառայությունների ոլորտի 58%-ը ապավինում է արհեստական ​​բանականության միջոցով խարդախության հայտնաբերմանը, քանի որ այն կարող է գտնել մանրուքներ, որոնք մարդկանց կողմից անտեսանելի է։ Մեքենայական ուսուցման գեղեցկությունն այն է, որ այն թույլ է տալիս բիզնեսներին ավելի արագ ու մասշտաբային կերպով հայտնաբերել խարդախությունը /Գծապատկեր 4․/:**[[45]](#footnote-45)**



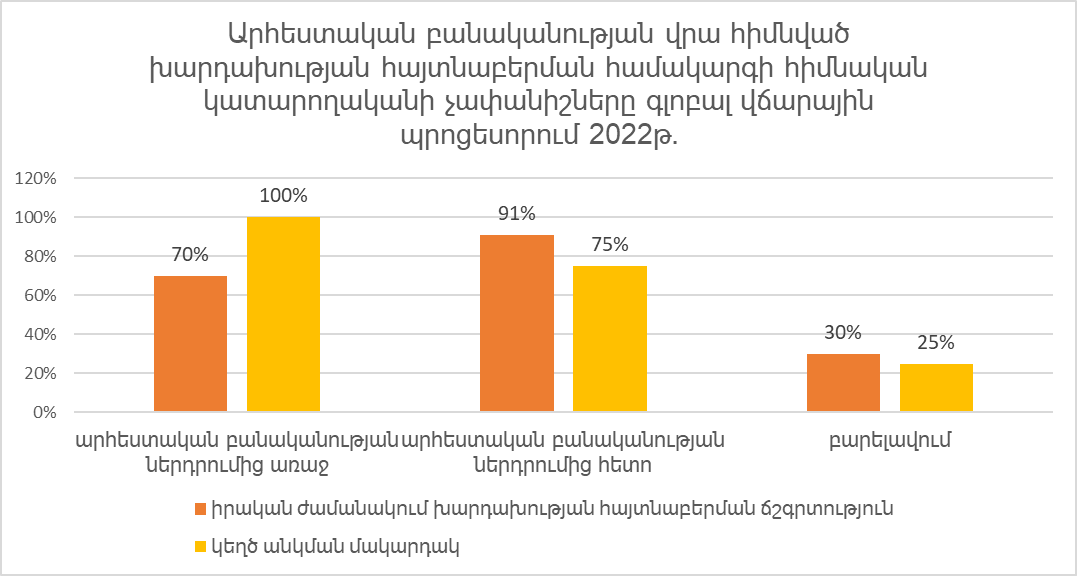
*Գծապատկեր 4․ Արհեստական ​​բանականության կիրառման դեպքերը ֆինանսական ծառայությունների ոլորտում ամբողջ աշխարհում 2020թ.[[46]](#footnote-46)*

Միացյալ Նահանգների 10 լավագույն բանկերից մեկը 2021 թվականին ներդրել է խորը ուսուցման վրա հիմնված խարդախության հայտնաբերման համակարգ, որի արդյունքները զգալի են և ցուցադրում են արհեստական ​​բանականության ներուժը խարդախության հայտնաբերման գործում /Գծապատկեր 5․/։



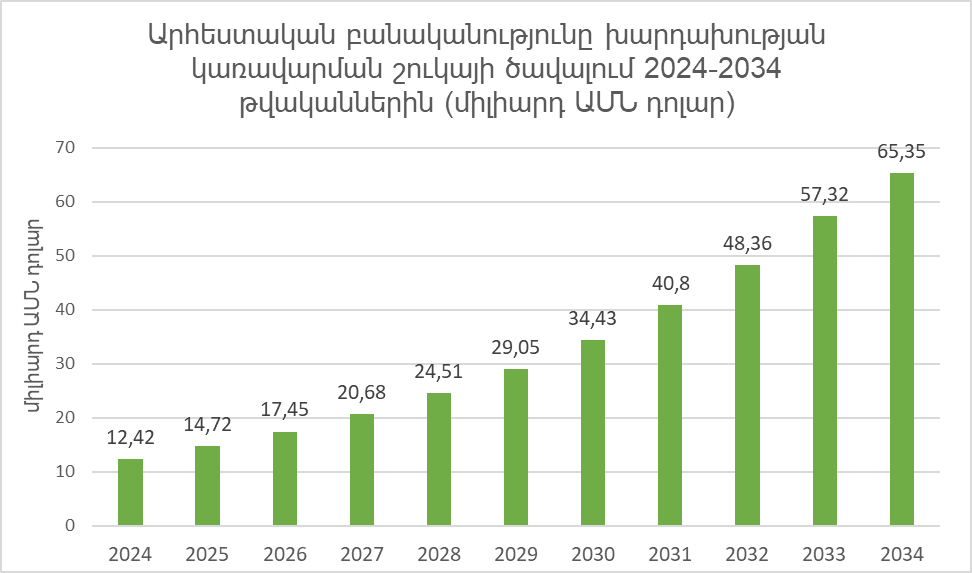
*Գծապատկեր 5. Արհեստական ​​բանականությամբ կառավարվող խարդախության հայտնաբերման համակարգի հիմնական կատարողականի չափանիշները 2021թ․***[[47]](#footnote-47)**

2022 թվականին տարեկան ավելի քան 65 միլիարդ գործարք մշակող համաշխարհային վճարային պրոցեսորը իր խարդախության հայտնաբերման համակարգում ինտեգրել է մեքենայական ուսուցման ալգորիթմներ /Գծապատկեր 6․/:**[[48]](#footnote-48)**



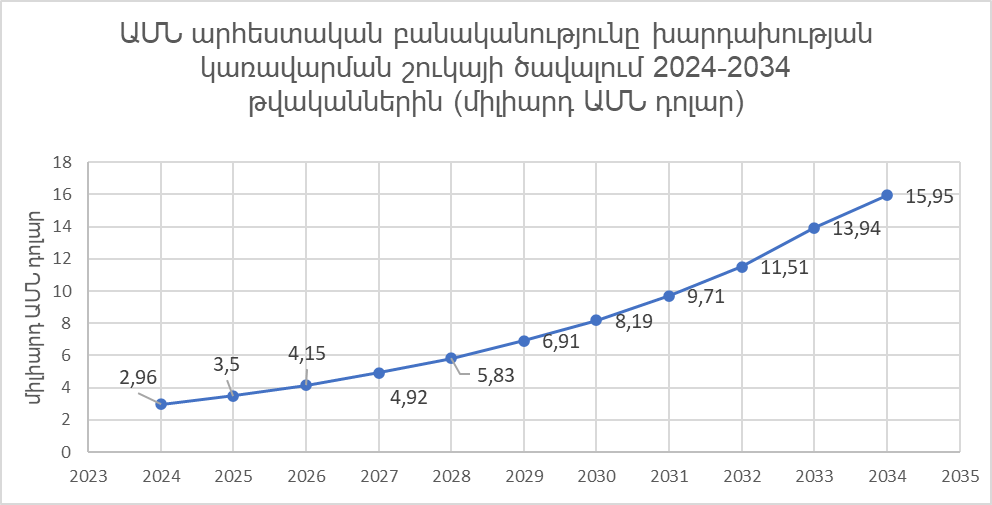
*Գծապատկեր 6. Արհեստական ​​բանականության վրա հիմնված խարդախության հայտնաբերման համակարգի հիմնական կատարողականի չափանիշները գլոբալ վճարային պրոցեսորում 2022թ․***[[49]](#footnote-49)**

Խարդախության կառավարման ոլորտում արհեստական ​​բանականության համաշխարհային շուկայի չափը 2024 թվականին գնահատվել է 12.42 միլիարդ ԱՄՆ դոլար և կանխատեսվում է, որ մինչև 2034 թվականը այն կհասնի մոտ 65.35 միլիարդ ԱՄՆ դոլարի՝ 2025-2034 թվականներին աճելով 18.06% տարեկան աճի տեմպով։ Խարդախության կառավարման արդյունավետ համակարգի նկատմամբ պահանջարկի աճը խթանում է խարդախության կառավարման ոլորտում արհեստական ​​բանականության շուկայի աճը /Գծապատկեր 7./։



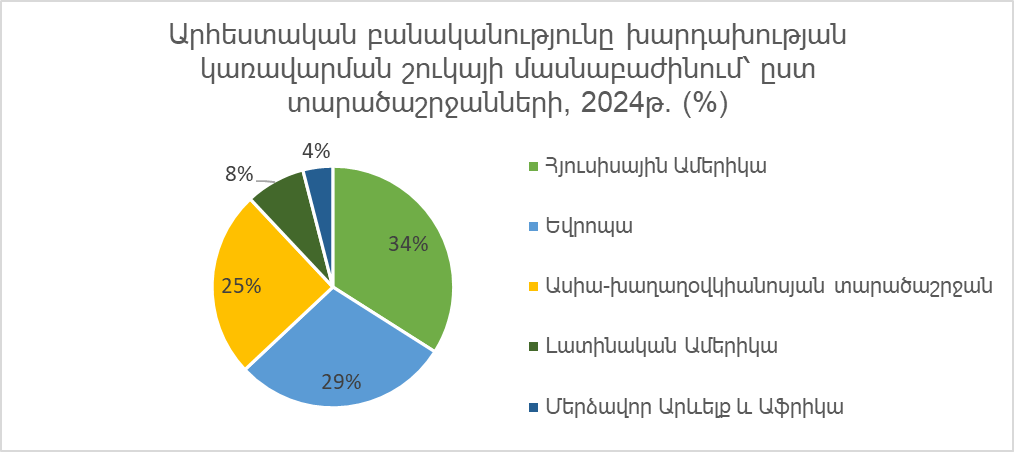
*Գծապատկեր 7․ Արհեստական ​​բանականությունը խարդախության կառավարման շուկայի ծավալում 2024-2034 թվականներին (միլիարդ ԱՄՆ դոլար)* **[[50]](#footnote-50)**

ԱՄՆ-ում խարդախության կառավարման ոլորտում արհեստական ​​բանականության շուկայի ծավալը 2024 թվականին կազմել է 2.96 միլիարդ ԱՄՆ դոլար, իսկ կանխատեսումների համաձայն՝ մինչև 2034 թվականը այն կկազմի մոտ 15.95 միլիարդ ԱՄՆ դոլար, որը 2025-ից 2034 թվականներին կկազմի 18.34% տարեկան աճի տեմպ /Գծապատկեր 8./։ **[[51]](#footnote-51)**



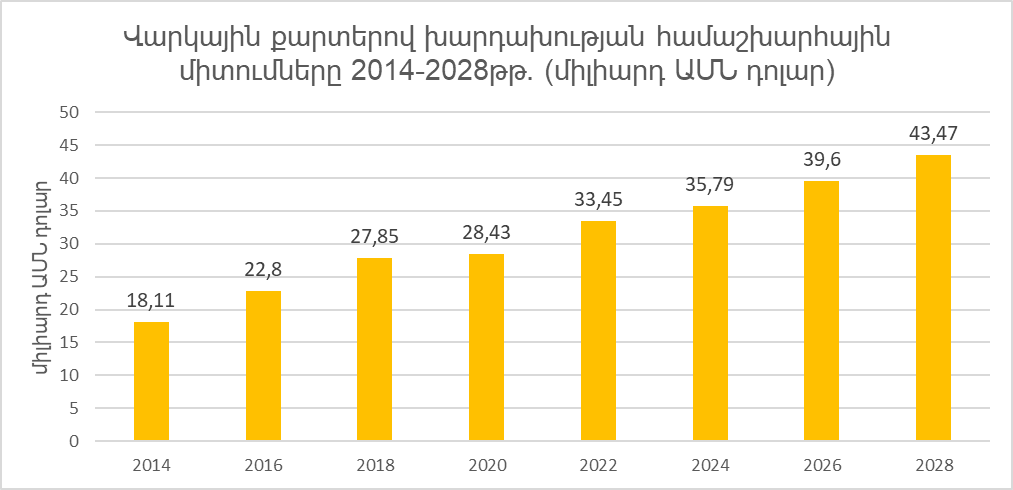
*Գծապատկեր 8․* *ԱՄՆ արհեստական ​​բանականությունը խարդախության կառավարման շուկայի ծավալում 2024-2034 թթ․ (միլիարդ ԱՄՆ դոլար)* **[[52]](#footnote-52)**

Հյուսիսային Ամերիկան աշխարհի ամենաշատ տուժած տարածաշրջաններից մեկն է փողերի լվացման ու ահաբեկչության ֆինանսավորման անօրինական գործողությունների պատճառով:[[53]](#footnote-53) Պայմանավորված ցանցային հանցագործությունների, խարդախությունների ու առաջադեմ կիբեռհարձակումների աճով՝ Եվրոպայում նույնպես արագ աճ է կանխատեսվում / Գծապատկեր 9․/:[[54]](#footnote-54)



*Գծապատկեր 9․ Արհեստական ​​բանականությունը խարդախության կառավարման շուկայի մասնաբաժինում՝ ըստ տարածաշրջանների, 2024թ. (%)* ***[[55]](#footnote-55)***

Ըստ վիճակագրության՝ 2022 թվականին ամբողջ աշխարհում խարդախության պատճառով կորուստը կազմել է 33.45 միլիարդ դոլար, որից գրեթե 41%-ը տեղի է ունեցել միայն ԱՄՆ-ում: Նիլսոնի զեկույցի համաձայն՝ վարկային քարտերի խարդախությունը 2028 թվականին կհասնի մինչև 43.47 միլիարդ դոլարի:*[[56]](#footnote-56)*



*Գծապատկեր 10. Վարկային քարտերով խարդախության համաշխարհային միտումները 2014-2028թթ․ (միլիարդ ԱՄՆ դոլար)[[57]](#footnote-57)*

Ներկայումս ֆինանսական հաստատությունների **ավելի քան 50%-ը** պատրաստվում է ներդնել արհեստական ​​բանականության լուծումներ՝ խարդախության նոր ու անհայտ տեսակները հայտնաբերելու համար:**[[58]](#footnote-58)**

Կառուցենք մոդել, որով կհայտնաբերենք ֆինանսական խարդախությունը։ Մենք կառուցել ենք սինթետիկ dataset գաղտնի տվյալների պաշտպանման նպատակով, ինչպես նաև GDPR և էթիկայի պահանջներով, անձնական տվյալները թաքցնելու, խարդախության կանխարգելման ռազմավարությունները չհրապարակելու համար։ Dataset-ն ունի 150000 տող, որտեղ յուրաքանչյուր տող ներկայացնում է մեկ օգտատիրոջ գործարք։ Սյուները ներառում են օգտատիրոջ տվյալները, քարտի և անվտանգության ստուգումների արդյունքները, հաշվարկված ռիսկի և խարդախության լեյբլները /Հավելված Աղյուսակ 5․/։

Կառուցել ենք Supervised learning-ի **Logistic Regression, Random Forest, XGBoost, CatBoost մոդելները։**

**Logistic Regression-ը** մաթեմատիկական մոդել է, որը հաշվարկում է յուրաքանչյուր գործարքի խարդախ լինելու «հավանականությունը»։

* **Random Forest**-ը Ensemble մոդել է, որը ստեղծում է բազմաթիվ Decision Tree-եր ու վերցնում նրանց միջին կանխատեսումը։
* **XGBoost-ը** Gradient boosting-ի վրա հիմնված հզոր ալգորիթմ է, որը ստեղծում է շատ Decision Tree-եր հերթականությամբ, այնպես, որ ամեն հաջորդ Tree-ն փորձում է ուղղել նախորդների սխալները։ Մոդելն ապահովում է բարձր ճշգրտություն և արագություն։
* CatBoost-ը հատուկ մշակված է **դասակարգման** և **ռեգրեսիայի** խնդիրների համար, որտեղ ունենք ներմուծվող հատկանիշներ և թիրախային փոփոխական։

Մոդելների կանխատեսումների աղյուսակային պատկերը ներկայացնելու համար կիրառվել է **Confusion Matrix**-ը, որը ցույց է տալիս իրական ու կանխատեսված դասերի համադրությունը՝ բաժանված True Positive, False Positive, True Negative ու False Negative արժեքների /Հավելված Գծապատկեր 11/։

**TP (True Positives) -** Փաստացի դրական տվյալներ, որոնք ճիշտ կանխատեսվել են որպես դրական մեր մոդելի կողմից։

**TN (True Negatives) -** Փաստացի բացասական տվյալներ, որոնք ճիշտ կանխատեսվել են որպես բացասական մեր մոդելի կողմից։

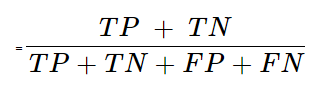
**FP (False Positives) -** Փաստացի բացասական տվյալներ, բայց մեր մոդելը դրանք կանխատեսել է որպես դրական:

**FN (False Negatives): -** Փաստացի դրական տվյալներ, բայց մեր մոդելը դրանք կանխատեսել է որպես բացասական:

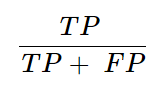
Ցանկալի սցենարները TPուTNէ։ Այստեղ FP-ը կարող է ավելացնել բանկի ծախսերը՝ կեղծ ահազանգի հետաքննության պատճառով, իսկ FN-ը՝ ֆինանսական կորուստներ՝ չբացահայտված խարդախության պատճառով։

Արդյունքները վերլուծելու համար օգտագործում ենք հետևյալ մետրիկները /Հավելված Գծապատկեր 12/՝

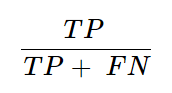
1. **Accuracy** - չափում է մոդելի ընդհանուր ճշտությունը՝



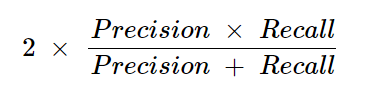
1. **Precision –** ցույց է տալիս, թեքանի կանխատեսված positive դեպքեր իրականում positive են։ Օգտակար է, երբ կարևոր է նվազեցնել կեղծ ահազանգերի քանակը՝



1. **Recal** - ցույց է տալիս, թե բոլոր փաստացի դրականներից քանիսն է ճիշտ կանխատեսել մոդելը։ Օգտակար է, երբ կարևոր է բացահայտել բոլոր խարդախ դեպքերը՝



1. **F1 Score** - հավասարակշռում է Precision ու Recall արժեքները։ Օգտակար է, երբ դասի բաշխումը անհավասարակշռված է, քանի որ այն հավասարակշռում է երկու չափումները:



1. **ROC-AUC** - ցույց է տալիս, թե մոդելը որքան լավ է տարբերում positive ու negative դասերը․ եթե 1.0 է՝ կատարյալ է, եթե 0.5 է՝ պատահական գուշակում։
2. **PR-AUC** – սա Precision-Recall կորի տակ գտնվող տարածքն է։ Ավելի օգտակար է, քան ROC-AUC-ը, երբ տվյալները շատ անհավասարակշռված են։

Հետազոտության ընթացքում կառուցել ենք ROC գծապատկեր /Հավելված Գծապատկեր 13/։ AUC արժեքը ցույց է տալիս, թե որքան լավ է մոդելը տարբերակում ընդհանուր առմամբ positive ու negative օրինակները։ Գծապատկերի դիագոնալ գիծը ցույց է տալիս պատահական գուշակման մակարդակը՝ AUC=0.5, և բոլոր մոդելների ROC կորերը պետք է այս գծից վերև լինեն։ Սա թույլ է տալիս տեսողականորեն համեմատել մոդելների արդյունավետությունը և ցույց տալ խարդախության հայտնաբերման մակարդակը։

Արդյունքներից տեսնում ենք, որ Logistic Regression-ը ունի ավելի շատ False Positives, Random Forest-ը՝ ավելի շատ missed fraud, XGBoost-ի False Positives-ը նորից շատ է, իսկ CatBoost-ն ունի ավելի լավ Recall և համեմատաբար ավելի լավ Precision։ Քանի որ մեր նպատակը այն է, որ ոչ մի fraud չմնա չհայտնաբերված, ուստի հետագայում կայքի պատրաստման համար կկիրառենք CatBoost-ը, քանի որ Recall-ը մոտավորապես 99% է։ Ստեղծելով խարդախությունների հայտնաբերման Fraud Detection համակարգ՝ շատ կարևոր է նշել, որ կայքում խարդախության հայտնաբերման հատվածում մենք վերբեռնված տվյալների ֆայլը պաշտպանել ենք՝ գաղտնագրելով, ապա անմիջապես հեշավորել անձնական ու զգայուն սյուները՝ օգտագործողի տվյալները նույնականացնել չկարողանալու նպատակով։

Այսպիսով՝ համակարգում խարդախության հայտնաբերումը հնարավորություն տվեց հայտնաբերել կասկածելի գործարքները իրական ժամանակում՝ ապահովելով տվյալների անվտանգությունը, նվազեցնելով կորուստները և բարձրացնելով վստահությունը: Ավտոմատացված ռիսկային վերլուծությունը թույլ տվեց արագ ու ճշգրիտ որոշումներ կայացնել՝ առանց մարդկային սխալների։

Հայաստանում դեռ չկա նման Fraud Detection համակարգ, որը ավտոմատացված կերպով հետևում ու հայտնաբերում է կասկածելի գործարքները: Նմանատիպ կայքը հանդիսանալու է լավագույն լուծումը, որն ապահովում է արագ, անվտանգ և ճշգրիտ ֆինանսական գործարքների գնահատում՝ բարձրացնելով հաճախորդների վստահությունն ու նվազեցնելով հնարավոր կորուստները։

Հոդվածում պարզ դարձավ, որ խարդախությունների բացահայտումը մեծապես հիմնված է անձնական ու ֆինանսական զգայուն տվյալների վրա։ Այսպիսի տվյալների խորքային վերլուծությունը, թեև արդյունավետ է կիբերհանցագործությունները կանխելու համար, միաժամանակ ստեղծում է գաղտնիության խախտման, տվյալների չարաշահման ու անհատի իրավունքների սահմանափակման վտանգներ։ Այդ պատճառով big data analytics-ում անհրաժեշտ է հավասարակշռել անվտանգության ապահովումն ու տվյալների էթիկայի պահպանությունը։

Вопросы конфиденциальности данных и этики в аналитике больших данных

**К. Хачатрян, Т. Хачатрян**

В цифровую эпоху объём данных растёт беспрецедентными темпами, формируя новую экономическую и социальную реальность. Анализ больших данных стал ключевым инструментом в различных областях: от оптимизации бизнес-процессов и совершенствования систем здравоохранения до выявления и предотвращения мошенничества в финансовом секторе. 5 основных характеристик больших данных: объём, разнообразие, скорость, достоверность и ценность, позволяют организациям обрабатывать огромные и разнообразные данные в режиме реального времени, получая ценную информацию и принимая обоснованные решения. Однако, наряду с возможностями больших данных, возникают и серьёзные вопросы конфиденциальности и этики.

В статье мы выявили эффективные методы выявления и предотвращения финансового мошенничества с помощью аналитики больших данных, используя количественные и качественные данные. В статье представлены основные проблемы, связанные с конфиденциальностью данных и этикой, а также технологические и правовые решения для их преодоления. В статье мы разработали систему Fraud Detection, которая автоматически отслеживает и обнаруживает подозрительные транзакции. Результаты исследования позволят нам повысить безопасность и надёжность цифровых транзакций, расширить практики ответственного управления данными и продвигать этические подходы к аналитике больших данных.

**Ключевые слова:** конфиденциальность данных, этика данных, безопасность, большие данные, мошенничество

Data Privacy and Ethics Issues in Big Data Analytics

**K. Khachatryan, T. Khachatryan**

In the digital age, the amount of data is growing at an unprecedented pace, shaping a new economic and social reality. Big data analysis has become a key tool in various fields, from optimizing business processes and improving healthcare systems to detecting and preventing fraud in the financial sector. The 5 main characteristics of big data: volume, variety, velocity, veracity, and value, allow organizations to process huge and diverse data in real time, obtaining valuable insights and making informed decisions. However, along with the capabilities of big data, serious privacy and ethics issues also arise.

In the article, we have identified effective methods for detecting and preventing financial fraud using big data analytics, using quantitative and qualitative data. The paper presents the main challenges of data privacy and ethics issues, as well as technological and legal solutions to overcome them. In the article, we have developed a Fraud Detection system that automatically tracks and detects suspicious transactions. The results will enable us to increase the security and trustworthiness of digital transactions, expand responsible data management practices, and promote ethical approaches to big data analytics.

**Keywords:** data privacy, data ethics, security, Big Data, fraud

**Օգտագործված գրականության ցանկ**

Harsha Patil, Vikas Mahandule, Juber Fakir, Omprasad Ajgaonkar «Balancing data privacy and ethics in the age of big data: challenges and solutions», 18.05.2024 <https://jibi.aspur.rs/archive/v3/n1/1.pdf>

Dr. Dinesh Kalla «Data Privacy and Ethics in Analytics», 2024 <https://www.irejournals.com/formatedpaper/1706428.pdf>

Masood Mortazavi, Khaled Salah «Privacy and Big Data», October 2015, էջ 6-19 <https://www.researchgate.net/publication/292984270_Privacy_and_Big_Data>

Маклакова Юлия Александровна «Возможные принципы этики Big Data» <https://interactive-plus.ru/e-articles/225/Action225-13928.pdf>

José Moura, Carlos Serrão «Security and Privacy Issues of Big Data», January 2019 <https://www.researchgate.net/publication/332996823_Security_and_Privacy_Issues_of_Big_Data>

Christopher B. Davison, Edward J. Lazaros, Jensen J. Zhao, Allen D. Truell, Brianna Bowles «Data privacy in the age of big data analytics», 2021 էջ 177-180 <https://iacis.org/iis/2021/2_iis_2021_185-195.pdf>

Hervais Simo «Big Data: Opportunities and Privacy Challenges» էջ 6-7 <https://arxiv.org/pdf/1502.00823>

Ezekiel Onyekachukwu Udeh, Prisca Amajuoyi, Kudirat Bukola Adeusi, Anwulika Ogechukwu Scott «The role of big data in detecting and preventing financial fraud in digital transactions», August 2024 <https://www.researchgate.net/publication/383397952_The_role_of_big_data_in_detecting_and_preventing_financial_fraud_in_digital_transactions>

Annie Badman, Matthew Kosinski «What is big data?», 18 November 2024 <https://www.ibm.com/think/topics/big-data>

Rishika Patel «Balancing Act: Ethics and Privacy in the Age of Big Data Analytics», February 20, 2024 [https://cioinfluence.com/it-and-devops/balancing-act-ethics-and-privacy-in-the-age-of-big-data-analytics/#:~:text=Ethical%20and%20Privacy%20Concerns:%20\*%20Mass%20Surveillance:,processes%20and%20hold%20actors%20accountable%20for%20misuse](https://cioinfluence.com/it-and-devops/balancing-act-ethics-and-privacy-in-the-age-of-big-data-analytics/#:~:text=Ethical%20and%20Privacy%20Concerns:%20*%20Mass%20Surveillance:,processes%20and%20hold%20actors%20accountable%20for%20misuse).

«Big Data Analytics – Ethical And Privacy Issues» Big Data Analytics – Ethical And Privacy Issues <https://ebsedu.org/blog/data-analytics-ethical-privacy-issues#:~:text=Privacy%20Concerns,identifiable%20information%20to%20protect%20privacy>.

«Top 5 Methods of Protecting Data» <https://www.titanfile.com/blog/5-methods-of-protecting-data/>

А. В. Ситников, С. Н. Нестеренков, А. Н. Марков «Методы защиты Больших Данных», 11-12 мая 2022 год <https://libeldoc.bsuir.by/bitstream/123456789/48384/1/Sitnikov_Metody.pdf>

GDPR՝ ընդունված 2018 թվականի մայիսի 25-ին Chapter 2 Principles Article 5 Principles relating to processing of personal data, Chapter 8 Remedies, liability and penalties Article 83 General conditions for imposing administrative fines <https://gdpr-info.eu/>

CCPA՝ ընդունված 2020 թվականի հունվարին, Section 1798.140. Definitions, Section 1798.155. Civil penalties, Section 1798.140. Definitions <https://www.consumerprivacyact.com/>

Հայաստանի իրավական տեղեկատվական կենտրոն՝ «Տեղեկատվության ազատության մասին» ՀՀ օրենք՝ ընդունված 2003 թվականի սեպտեմբերի 23-ին, Հոդված 8․ Տեղեկատվության ազատության սահմանափակումները <https://www.arlis.am/hy/acts/1372>

Հայաստանի իրավական տեղեկատվական կենտրոն՝ «Անձնական տվյալների պաշտպանության մասին» ՀՀ օրենքը՝ ընդունված 2015 թվականի մայիսի 18-ին, Գլուխ 1 Ընդհանուր դրույթներ Հոդված 1 Օրենքի կարգավորման առարկան <https://www.arlis.am/hy/acts/98338>

Հայաստանի իրավական տեղեկատվական կենտրոն՝ Բանկային գաղտնիքի մասին ՀՀ օրենք՝ ընդունված Ազգային ժողովի կողմից 7 հոկտեմբերի 1996 թ., Հոդված 4. Բանկային գաղտնիքը, Հոդված 7. Բանկային գաղտնիքի հրապարակման արգելումը <https://www.arlis.am/hy/acts/51416>

«What is data privacy?» <https://www.cloudflare.com/learning/privacy/what-is-data-privacy/>

«Big Data Ethics and 10 Controversial Experiments» June 9, 2022 <https://datasciencedojo.com/blog/big-data-ethics-10-experiments/>

«7 Essential Data Ethics Examples for Businesses in 2025» December 26th, 2024 <https://atlan.com/data-ethics-examples/>

«Timeline of Cyber Incidents Involving Financial Institutions» <https://carnegieendowment.org/features/fincyber-timeline?lang=en>

Global Cybersecurity Index <https://www.itu.int/en/ITU-D/Cybersecurity/Pages/Global-Cybersecurity-Index.aspx>

Brianna Valleskey «AI Fraud Detection: How It Works», March 26, 2022 <https://www.inscribe.ai/fraud-detection/ai-fraud-detection>

Saloni Thakkar «Enhancing Fraud Detection in Financial Transactions through Advanced AI Algorithms», October 2024 <https://www.researchgate.net/publication/384654151_Enhancing_Fraud_Detection_in_Financial_Transactions_through_Advanced_AI_Algorithms>

Shivani Zoting, Aditi Shivarkar «AI in Fraud Management Market Size and Forecast 2025 to 2034» 15 Jul 2025 <https://www.precedenceresearch.com/ai-in-fraud-management-market>

«Fraud Detection and Prevention Market – Global Industry Analysis and Forecast (2023-2029)» June 2023 <https://www.maximizemarketresearch.com/market-report/global-fraud-detection-and-prevention-market/62819/>

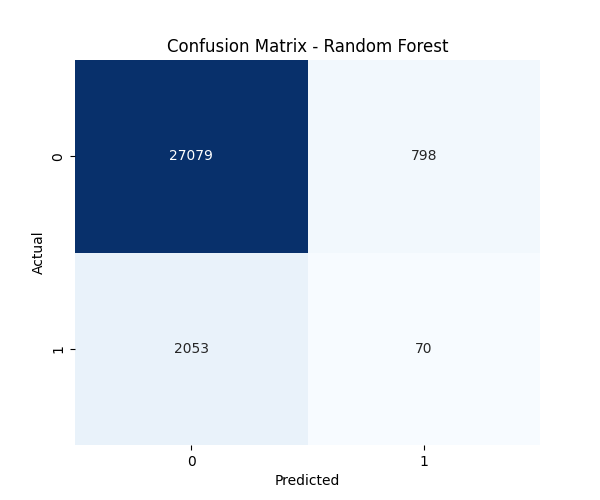
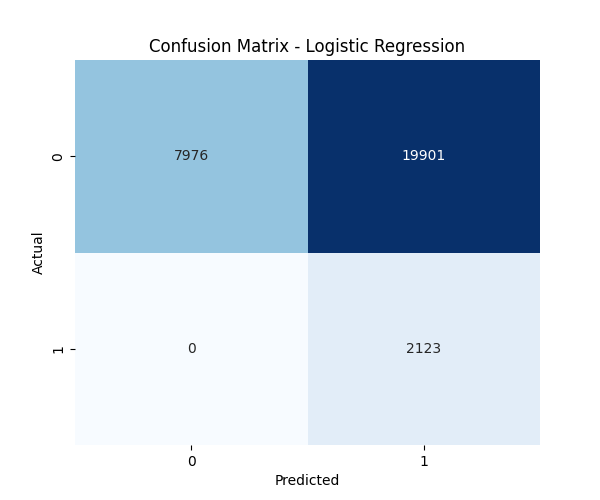
Dr. Arunkumar M. Sampath «AI/ML- and Generative AI-enabled Financial Fraud Detection» Feb 15, 2024 <https://medium.com/@arunsampath_47187/ai-ml-and-generative-ai-enabled-financial-fraud-detection-ba4700a8e2a1>

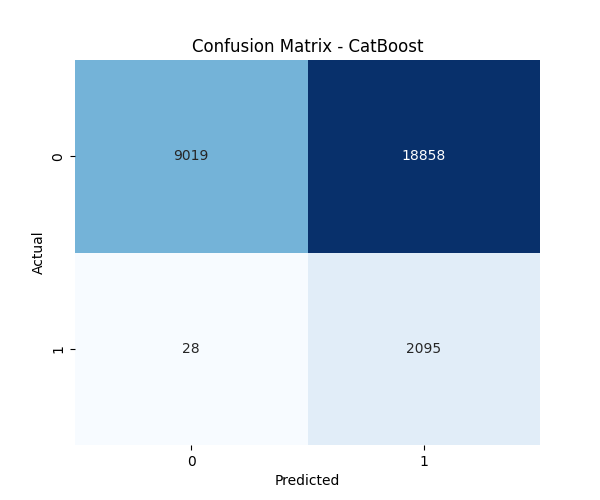
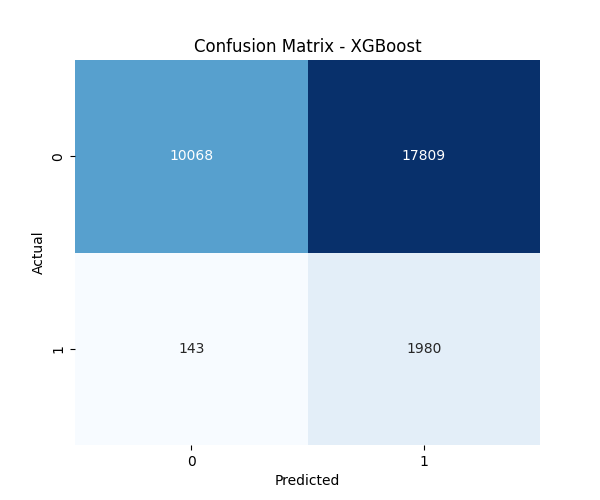
«Global AI In Fraud Detection Market Report» August 2024 <https://market.us/report/ai-in-fraud-detection-market/>

**Հավելված**

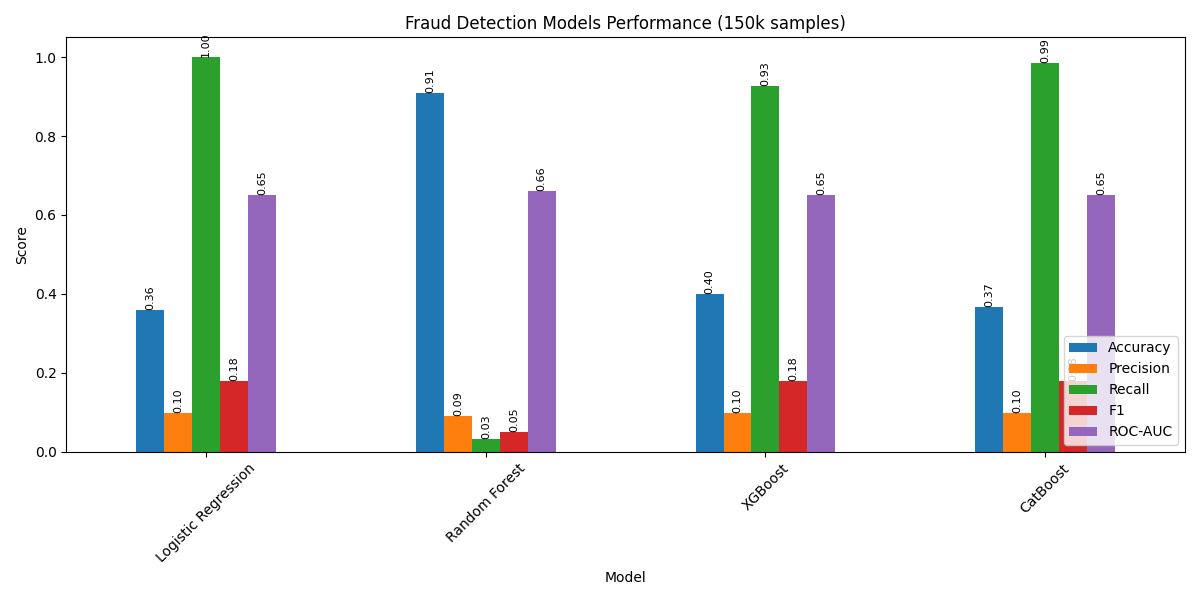
*Աղյուսակ 5․* *150000 տող սինթետիկ տվյալներ՝ օգտատերերի խարդախ և ոչ խարդախ գործարքների վերաբերյալ[[59]](#footnote-59)*

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **user\_id /x1/** | | **user\_country /x2/** | **user\_city /x3/** | | | **user\_account\_status /x4/** | | | | | **user\_balance /x5/** | | | **user\_verification\_level /x6/** | | | **transaction\_id /x7/** | |
| 57a1b2a7-e02e-4bf0-9028-a7dc685dafca | | IN | City\_113 | | | active | | | | | 3942.0 | | | full | | | 48d9546a-5884-49c8-b172-ba042ee9ec6f | |
| 63b9a0e4-902f-4d17-8c4d-c487bef70b24 | | DE | City\_315 | | | suspended | | | | | 1462.03 | | | full | | | b547f22c-956a-404d-bc9d-44609de90293 | |
| af8edc26-01e2-473c-8ae9-4bb8fa72f84c | | FR | City\_51 | | | suspended | | | | | 8216.29 | | | full | | | 44119e1d-ca80-48d5-883d-c18ef2272b84 | |
| 41128a6e-97fd-4eb7-9cea-553344326029 | | FR | City\_146 | | | suspended | | | | | 642.16 | | | full | | | 59bc34a6-4561-43c2-8478-9819a74cfe99 | |
| d1586964-64f2-4509-b21e-1301c167d298 | | AM | City\_127 | | | active | | | | | 2811.63 | | | basic | | | 40570491-9768-4323-848d-645e83d2c0c6 | |
| 56d24d3c-67a6-415b-b60e-a9070e74aea4 | | DE | City\_329 | | | suspended | | | | | 2267.65 | | | full | | | ea6ec740-5e8d-4ee3-8a30-83b2c1065913 | |
| 9425a2c0-b9b9-4909-ac37-86a04dc397a2 | | DE | City\_219 | | | active | | | | | 1777.87 | | | none | | | 8db5b58d-09f0-4f24-b55d-33c1876cf824 | |
| 875b2c9b-f60e-4d20-9c30-58dd7c296499 | | US | City\_453 | | | closed | | | | | 3090.45 | | | basic | | | 986fbebd-de7e-473d-b2f9-96e27bcad55f | |
| 0dff7f68-c329-431f-bc07-70254a406cb2 | | FR | City\_379 | | | active | | | | | 4794.23 | | | none | | | 8879b82c-f705-45c1-a9f6-5b457d846b67 | |
| c5ab9bbc-e81e-4101-83e7-13aa11b879e5 | | IN | City\_72 | | | closed | | | | | 3724.49 | | | none | | | 087bebd9-9f33-4462-9957-420571db513a | |
| 0337590a-8c7b-4a91-90f7-f7be454b3a5a | | IN | City\_271 | | | closed | | | | | 1983.45 | | | full | | | 33c3e19f-0973-4ee1-b431-f4262d561ab5 | |
| 90279313-fcbd-473e-8508-a0cf9e6bdff3 | | CN | City\_174 | | | closed | | | | | 100.09 | | | none | | | 35130d92-85bb-467f-89cf-8a88fe6167b7 | |
| e1212e70-ecec-493e-9a3f-2b3ea1d1f63c | | IN | City\_209 | | | active | | | | | 5097.42 | | | basic | | | e9bde7a6-d14d-4fbf-ae09-e8e5b681fc74 | |
| **transaction\_amount /x8/** | **transaction\_currency /x9/** | | | | **transaction\_type /x10/** | | **card\_bin /x11/** | | **card\_last /x12/** | | | **status\_3d /x13/** | | | **cvv\_result /x14/** | **Ip /x15/** | |
| 1958.04 | EUR | | | | deposit | | 447160 | | 2008 | | | passed | | | False | 212.32.195.224 | |
| 4882.31 | EUR | | | | withdrawal | | 476490 | | 8787 | | | failed | | | False | 65.74.241.111 | |
| 26.18 | EUR | | | | refund | | 470976 | | 5835 | | | not\_enrolled | | | False | 26.23.2.40 | |
| 412.46 | USD | | | | deposit | | 488700 | | 7033 | | | passed | | | False | 225.13.0.243 | |
| 3636.58 | USD | | | | refund | | 422024 | | 3072 | | | failed | | | False | 219.42.62.55 | |
| 3454.21 | AMD | | | | deposit | | 484393 | | 5959 | | | passed | | | False | 41.165.188.226 | |
| 3530.89 | USD | | | | purchase | | 442973 | | 3418 | | | not\_enrolled | | | True | 9.29.42.108 | |
| 1854.7 | EUR | | | | withdrawal | | 447951 | | 6596 | | | not\_enrolled | | | True | 78.66.255.224 | |
| 1981.75 | RUB | | | | purchase | | 470978 | | 7117 | | | failed | | | False | 150.228.140.180 | |
| 1815.12 | RUB | | | | refund | | 424393 | | 6097 | | | not\_enrolled | | | True | 96.9.244.227 | |
| 4138.8 | USD | | | | purchase | | 401359 | | 5972 | | | failed | | | False | 165.94.124.11 | |
| 3522.17 | USD | | | | withdrawal | | 413856 | | 1630 | | | not\_enrolled | | | True | 130.233.235.96 | |
| 280.58 | EUR | | | | withdrawal | | 495021 | | 2672 | | | failed | | | True | 213.225.230.23 | |
| **device\_id /x16/** | | | | **shipping\_country /x17/** | | | | **billing\_country /x18/** | | **is\_fraud /y/** | | |
| 92d5394d-66d3-4250-903b-06498a66921d | | | | FR | | | | RU | | 0 | | |
| 769b9eff-2b1a-4d72-9908-34268f15a0d1 | | | | IN | | | | DE | | 0 | | |
| 3a906524-2a86-4afc-a969-f3efbc1dcfdb | | | | FR | | | | US | | 0 | | |
| 2c6f8a77-91de-4d6f-9445-6711d232c289 | | | | RU | | | | US | | 0 | | |
| 04f4f656-5afc-43d3-906d-a34718b67e91 | | | | AM | | | | FR | | 0 | | |
| a0fc9bc6-a37e-4a73-bfb4-4198ff0526ec | | | | IN | | | | US | | 0 | | |
| d974048f-6d1c-494e-85b4-f18a2ee12784 | | | | CN | | | | RU | | 0 | | |
| 6a01b4db-7ec0-4ebe-9f91-9ec4e078ee00 | | | | FR | | | | US | | 0 | | |
| 0b5cd03c-3a69-4b6f-aad2-16abb825cef4 | | | | FR | | | | US | | 0 | | |
| 7b8b071a-f3f0-4ed8-bcc6-1debab32a5cc | | | | DE | | | | CN | | 0 | | |
| ffa6b5b4-50b8-48d2-a030-4968f4d1611c | | | | AM | | | | IN | | 0 | | |
| 2cda7d7e-611d-4e9a-a67c-8f986979e558 | | | | FR | | | | RU | | 1 | | |
| b3d3879b-cc8b-48cb-992b-62a1f553fb2d | | | | US | | | | US | | 0 | | |

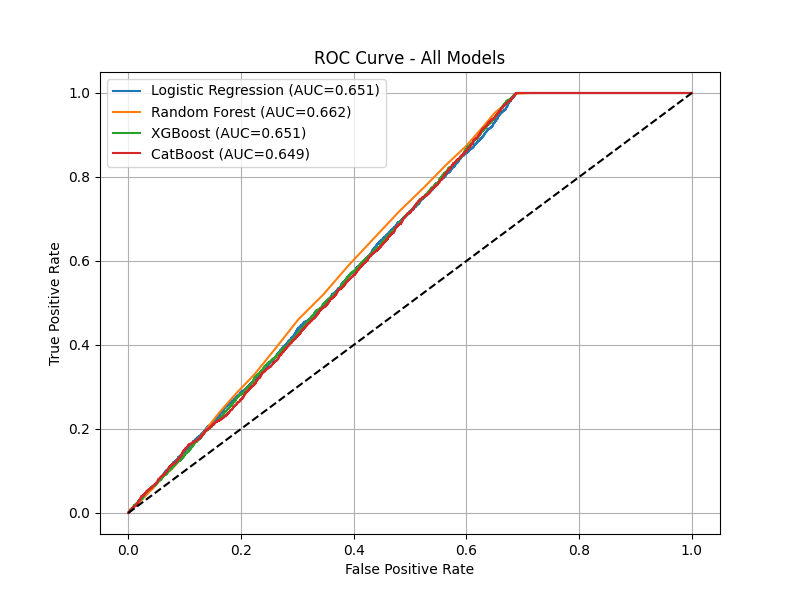
**

**

*Գծապատկեր 11․ Confusion Matrix՝ Supervised learning-ի մոդելների համար[[60]](#footnote-60)*



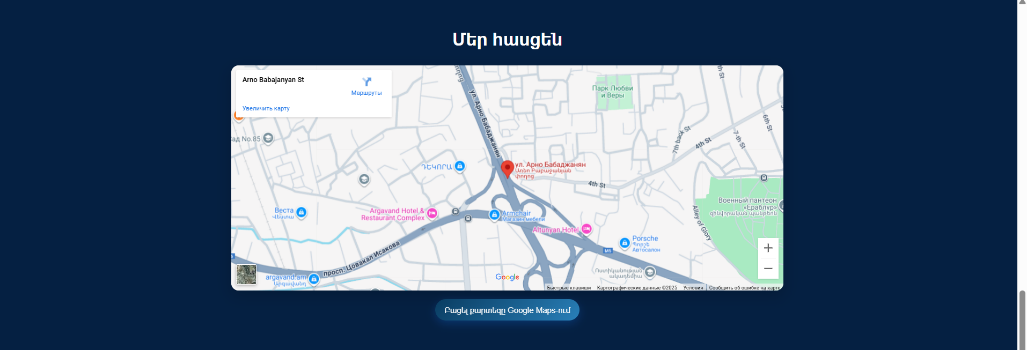
*Գծապատկեր 12․ Supervised learning-ի մոդելների ճշգրտության ցուցանիշներ[[61]](#footnote-61)*

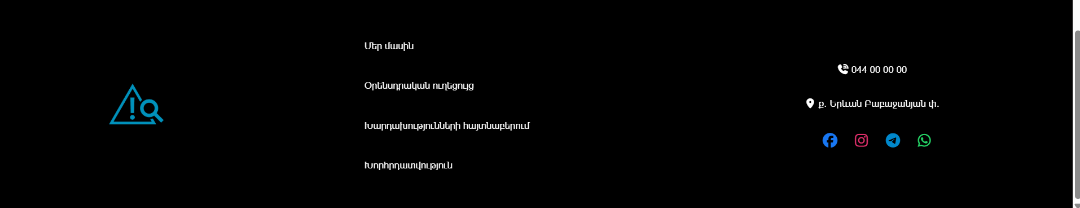


*Գծապատկեր 13․ ROC գծապատկեր՝ Supervised learning-ի մոդելների համար[[62]](#footnote-62)*

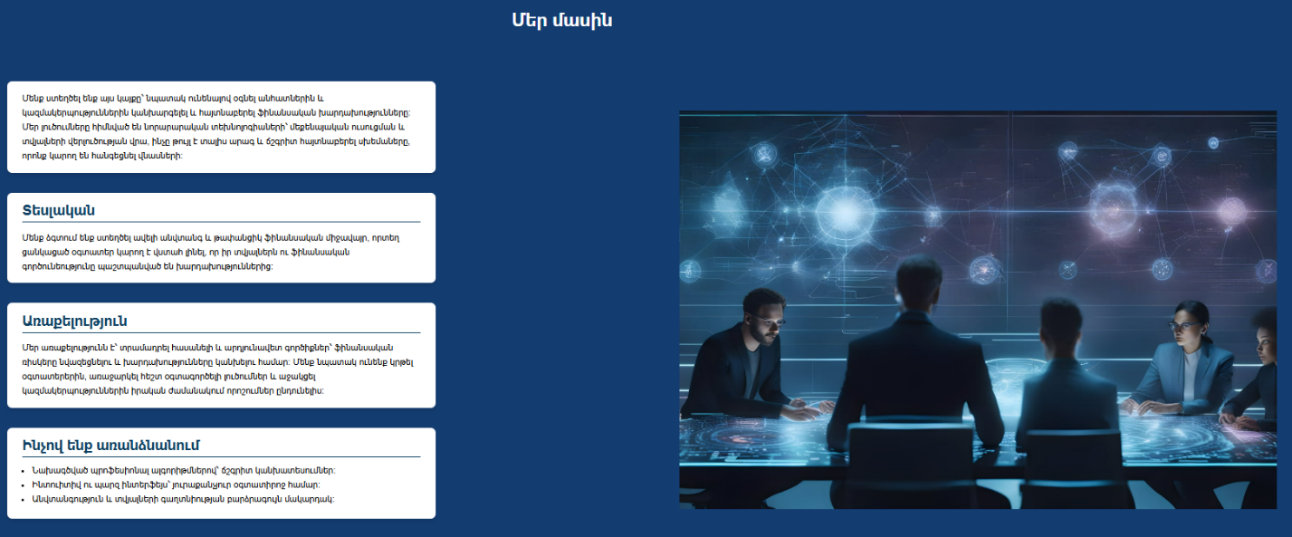








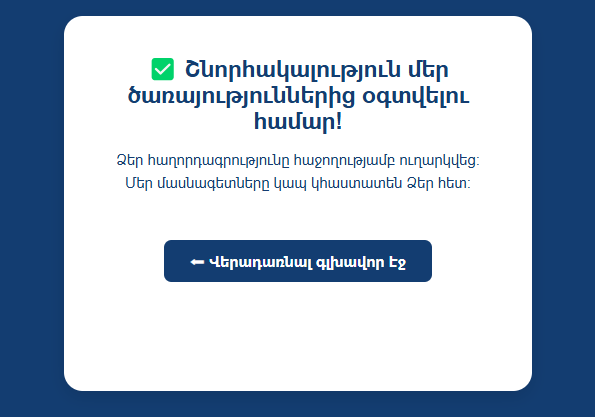
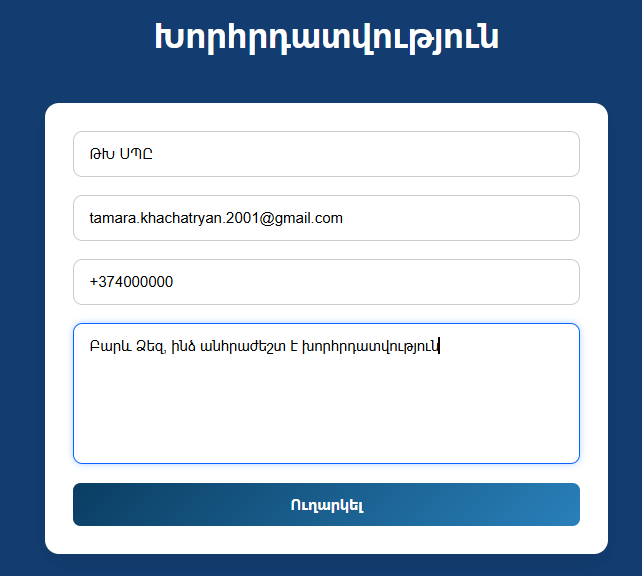
*Նկար 1․ Fraud Detection կայքի գլխավոր էջ*



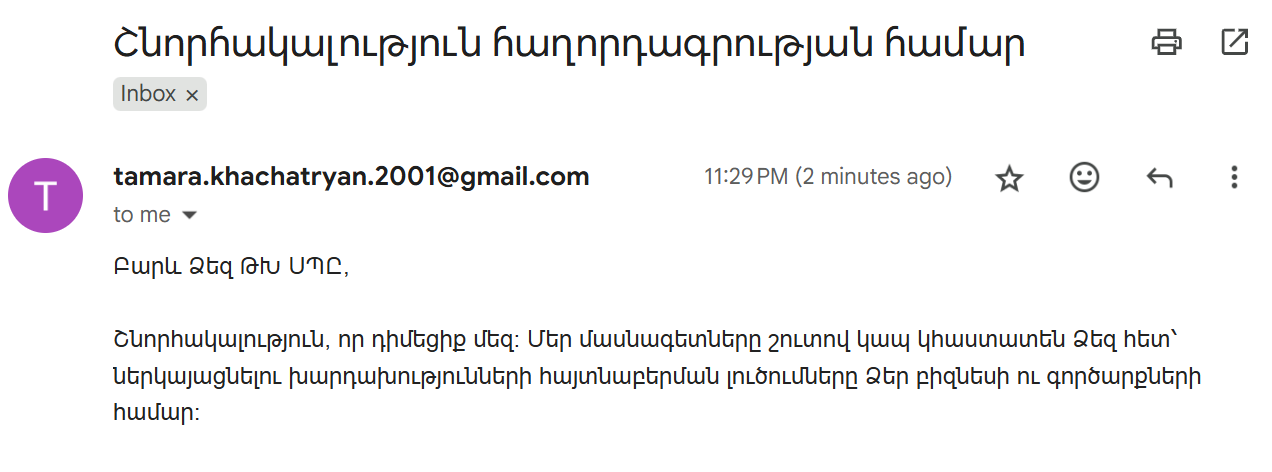
*Նկար 2․ Fraud Detection կայքի մեր մասին էջ*



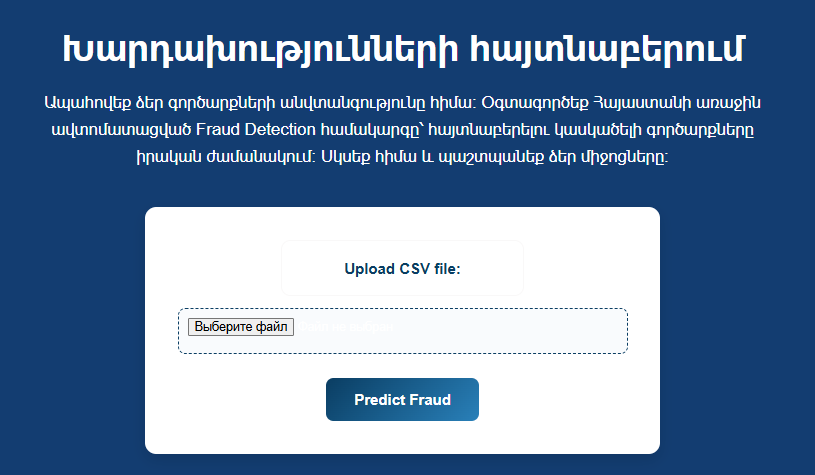
*Նկար 3․ Fraud Detection կայքի Օրենսդրական ուղեցույց էջ*

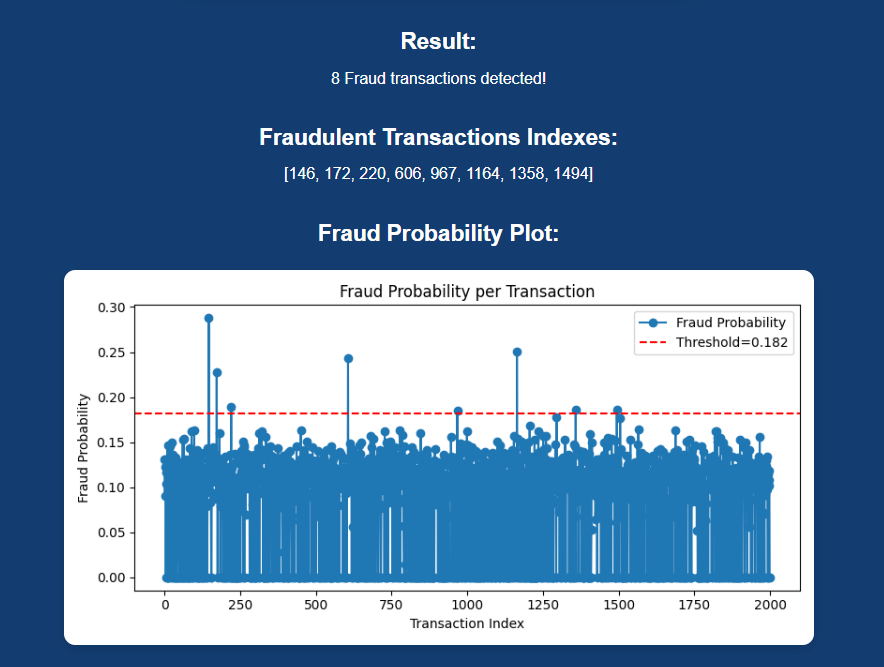


*Նկար 4․ Fraud Detection կայքի խորհրդատվություն էջ*



*Նկար 5․ Fraud Detection կայքի հետադարձ կապ*





*Նկար 6․ Fraud Detection կայքի խարդախության հայտնաբերման էջ*

1. Harsha Patil, Vikas Mahandule, Juber Fakir, Omprasad Ajgaonkar «Balancing data privacy and ethics in the age of big data: challenges and solutions», 18.05.2024 <https://jibi.aspur.rs/archive/v3/n1/1.pdf> /հուլիսի 10 12։00/ [↑](#footnote-ref-1)
2. Dr. Dinesh Kalla «Data Privacy and Ethics in Analytics», 2024 <https://www.irejournals.com/formatedpaper/1706428.pdf> /հուլիսի 10 13։10/ [↑](#footnote-ref-2)
3. Masood Mortazavi, Khaled Salah «Privacy and Big Data», October 2015, էջ 6-19 <https://www.researchgate.net/publication/292984270_Privacy_and_Big_Data> /հուլիսի 10 14։00/ [↑](#footnote-ref-3)
4. Маклакова Юлия Александровна «Возможные принципы этики Big Data» <https://interactive-plus.ru/e-articles/225/Action225-13928.pdf> /հուլիսի 20 12։00/ [↑](#footnote-ref-4)
5. José Moura, Carlos Serrão «Security and Privacy Issues of Big Data», January 2019 <https://www.researchgate.net/publication/332996823_Security_and_Privacy_Issues_of_Big_Data> /հուլիսի 20 13։00/ [↑](#footnote-ref-5)
6. Christopher B. Davison, Edward J. Lazaros, Jensen J. Zhao, Allen D. Truell, Brianna Bowles «Data privacy in the age of big data analytics», 2021 էջ 177-180 <https://iacis.org/iis/2021/2_iis_2021_185-195.pdf> /հուլիսի 17 11։00/ [↑](#footnote-ref-6)
7. Hervais Simo «Big Data: Opportunities and Privacy Challenges» էջ 6-7 <https://arxiv.org/pdf/1502.00823> /հուլիսի 14։00/ [↑](#footnote-ref-7)
8. Ezekiel Onyekachukwu Udeh, Prisca Amajuoyi, Kudirat Bukola Adeusi, Anwulika Ogechukwu Scott «The role of big data in detecting and preventing financial fraud in digital transactions», August 2024 <https://www.researchgate.net/publication/383397952_The_role_of_big_data_in_detecting_and_preventing_financial_fraud_in_digital_transactions> /հուլիսի 10 15։00/ [↑](#footnote-ref-8)
9. Annie Badman, Matthew Kosinski «What is big data?», 18 November 2024 <https://www.ibm.com/think/topics/big-data> /հուլիսի 10 15։20/ [↑](#footnote-ref-9)
10. Rishika Patel «Balancing Act: Ethics and Privacy in the Age of Big Data Analytics», February 20, 2024 [https://cioinfluence.com/it-and-devops/balancing-act-ethics-and-privacy-in-the-age-of-big-data-analytics/#:~:text=Ethical%20and%20Privacy%20Concerns:%20\*%20Mass%20Surveillance:,processes%20and%20hold%20actors%20accountable%20for%20misuse](https://cioinfluence.com/it-and-devops/balancing-act-ethics-and-privacy-in-the-age-of-big-data-analytics/#:~:text=Ethical%20and%20Privacy%20Concerns:%20*%20Mass%20Surveillance:,processes%20and%20hold%20actors%20accountable%20for%20misuse). /հուլիսի 10 17։00/ [↑](#footnote-ref-10)
11. «Big Data Analytics – Ethical And Privacy Issues» Big Data Analytics – Ethical And Privacy Issues <https://ebsedu.org/blog/data-analytics-ethical-privacy-issues#:~:text=Privacy%20Concerns,identifiable%20information%20to%20protect%20privacy>. /հուլիսի 10 19։00/ [↑](#footnote-ref-11)
12. «Top 5 Methods of Protecting Data» <https://www.titanfile.com/blog/5-methods-of-protecting-data/> [↑](#footnote-ref-12)
13. Աղյուսակը կազմվել է հեղինակի կողմից, որի համար հիմք են հանդիսացել А. В. Ситников, С. Н. Нестеренков, А. Н. Марков «Методы защиты Больших Данных», 11-12 мая 2022 год <https://libeldoc.bsuir.by/bitstream/123456789/48384/1/Sitnikov_Metody.pdf> տվյալները /հուլիսի 12 12։00/ [↑](#footnote-ref-13)
14. Աղյուսակը կազմվել է հեղինակի կողմից, որի համար հիմք են հանդիսացել А. В. Ситников, С. Н. Нестеренков, А. Н. Марков «Методы защиты Больших Данных», 11-12 мая 2022 год <https://libeldoc.bsuir.by/bitstream/123456789/48384/1/Sitnikov_Metody.pdf> տվյալները /հուլիսի 12 12։00/ [↑](#footnote-ref-14)
15. Աղյուսակը կազմվել է հեղինակի կողմից, որի համար հիմք են հանդիսացել А. В. Ситников, С. Н. Нестеренков, А. Н. Марков «Методы защиты Больших Данных», 11-12 мая 2022 год <https://libeldoc.bsuir.by/bitstream/123456789/48384/1/Sitnikov_Metody.pdf> տվյալները /հուլիսի 12 12։00/ [↑](#footnote-ref-15)
16. Աղյուսակը կազմվել է հեղինակի կողմից, որի համար հիմք են հանդիսացել А. В. Ситников, С. Н. Нестеренков, А. Н. Марков «Методы защиты Больших Данных», 11-12 мая 2022 год <https://libeldoc.bsuir.by/bitstream/123456789/48384/1/Sitnikov_Metody.pdf> տվյալները /հուլիսի 12 12։00/ [↑](#footnote-ref-16)
17. А. В. Ситников, С. Н. Нестеренков, А. Н. Марков «Методы защиты Больших Данных», 11-12 мая 2022 год <https://libeldoc.bsuir.by/bitstream/123456789/48384/1/Sitnikov_Metody.pdf> /հուլիսի 13 12։00/ [↑](#footnote-ref-17)
18. «Top 5 Methods of Protecting Data» <https://www.titanfile.com/blog/5-methods-of-protecting-data/> /հուլիսի 25 12։00/ [↑](#footnote-ref-18)
19. «Big Data» <https://www.securityvision.ru/info/big_data/#:~:text=%D0%B2%20%D0%B8%D0%BD%D1%84%D0%BE%D1%80%D0%BC%D0%B0%D1%86%D0%B8%D0%BE%D0%BD%D0%BD%D0%BE%D0%B9%20%D0%B1%D0%B5%D0%B7%D0%BE%D0%BF%D0%B0%D1%81%D0%BD%D0%BE%D1%81%D1%82%D0%B8:-,1.,%D0%BF%D0%BE%D0%B2%D0%B5%D0%B4%D0%B5%D0%BD%D0%B8%D0%B5%20%D0%B8%20%D0%BF%D0%BE%D1%82%D0%B5%D0%BD%D1%86%D0%B8%D0%B0%D0%BB%D1%8C%D0%BD%D1%8B%D0%B5%20%D0%B8%D0%BD%D1%86%D0%B8%D0%B4%D0%B5%D0%BD%D1%82%D1%8B%20%D0%B1%D0%B5%D0%B7%D0%BE%D0%BF%D0%B0%D1%81%D0%BD%D0%BE%D1%81%D1%82%D0%B8>. [↑](#footnote-ref-19)
20. «Top 5 Methods of Protecting Data» <https://www.titanfile.com/blog/5-methods-of-protecting-data/> /հուլիսի 25 12։00/ [↑](#footnote-ref-20)
21. «Rishika Patel «Balancing Act: Ethics and Privacy in the Age of Big Data Analytics», February 20, 2024 [https://cioinfluence.com/it-and-devops/balancing-act-ethics-and-privacy-in-the-age-of-big-data-analytics/#:~:text=Ethical%20and%20Privacy%20Concerns:%20\*%20Mass%20Surveillance:,processes%20and%20hold%20actors%20accountable%20for%20misuse](https://cioinfluence.com/it-and-devops/balancing-act-ethics-and-privacy-in-the-age-of-big-data-analytics/#:~:text=Ethical%20and%20Privacy%20Concerns:%20*%20Mass%20Surveillance:,processes%20and%20hold%20actors%20accountable%20for%20misuse). /հուլիսի 18 14։00/ [↑](#footnote-ref-21)
22. «Top 5 Methods of Protecting Data» <https://www.titanfile.com/blog/5-methods-of-protecting-data/> /հուլիսի 25 12։00/ [↑](#footnote-ref-22)
23. Rishika Patel «Balancing Act: Ethics and Privacy in the Age of Big Data Analytics», February 20, 2024 [https://cioinfluence.com/it-and-devops/balancing-act-ethics-and-privacy-in-the-age-of-big-data-analytics/#:~:text=Ethical%20and%20Privacy%20Concerns:%20\*%20Mass%20Surveillance:,processes%20and%20hold%20actors%20accountable%20for%20misuse](https://cioinfluence.com/it-and-devops/balancing-act-ethics-and-privacy-in-the-age-of-big-data-analytics/#:~:text=Ethical%20and%20Privacy%20Concerns:%20*%20Mass%20Surveillance:,processes%20and%20hold%20actors%20accountable%20for%20misuse). /հուլիսի 20 13։30/ [↑](#footnote-ref-23)
24. <https://gdpr-info.eu/> GDPR՝ ընդունված 2018 թվականի մայիսի 25-ին /հուլիսի 19 15։00/ [↑](#footnote-ref-24)
25. <https://gdpr-info.eu/art-5-gdpr/> GDPR՝ ընդունված 2018 թվականի մայիսի 25-ին, Chapter 2 Principles Article 5 Principles relating to processing of personal dat /հուլիսի 19 15։00/ [↑](#footnote-ref-25)
26. <https://gdpr-info.eu/art-83-gdpr/> GDPR՝ ընդունված 2018 թվականի մայիսի 25-ին, Chapter 8 Remedies, liability and penalties Article 83 General conditions for imposing administrative fines /հուլիսի 19 15։00/ [↑](#footnote-ref-26)
27. <https://www.consumerprivacyact.com/section-1798-140-definitions/> CCPA՝ ընդունված 2020 թվականի հունվարին, Section 1798.140. Definitions /հուլիսի 19 15։00/ [↑](#footnote-ref-27)
28. <https://www.consumerprivacyact.com/section-1798-155-civil-penalties/> CCPA՝ ընդունված 2020 թվականի հունվարին, Section 1798.155. Civil penalties /հուլիսի 19 15։00/ [↑](#footnote-ref-28)
29. <https://www.consumerprivacyact.com/section-1798-140-definitions/> CCPA՝ ընդունված 2020 թվականի հունվարին, Section 1798.140. Definitions /հուլիսի 19 15։00/ [↑](#footnote-ref-29)
30. «What is data privacy?» <https://www.cloudflare.com/learning/privacy/what-is-data-privacy/> /հուլիսի 25 17։00/ [↑](#footnote-ref-30)
31. Հայաստանի իրավական տեղեկատվական կենտրոն՝ «Տեղեկատվության ազատության մասին» ՀՀ օրենք՝ ընդունված 2003 թվականի սեպտեմբերի 23-ին, Հոդված 8․ Տեղեկատվության ազատության սահմանափակումները <https://www.arlis.am/hy/acts/1372> /հուլիսի 19 16։00/ [↑](#footnote-ref-31)
32. Հայաստանի իրավական տեղեկատվական կենտրոն՝ «Անձնական տվյալների պաշտպանության մասին» ՀՀ օրենքը՝ ընդունված 2015 թվականի մայիսի 18-ին, Գլուխ 1 Ընդհանուր դրույթներ Հոդված 1 Օրենքի կարգավորման առարկան <https://www.arlis.am/hy/acts/98338> /հուլիսի 19 16։00/ [↑](#footnote-ref-32)
33. Հայաստանի իրավական տեղեկատվական կենտրոն՝ Բանկային գաղտնիքի մասին ՀՀ օրենք՝ ընդունված Ազգային ժողովի կողմից 7 հոկտեմբերի 1996 թ., Հոդված 4. Բանկային գաղտնիքը, Հոդված 7. Բանկային գաղտնիքի հրապարակման արգելումը <https://www.arlis.am/hy/acts/51416> /հուլիսի 19 16։00/ [↑](#footnote-ref-33)
34. «Big Data Ethics and 10 Controversial Experiments» June 9, 2022 <https://datasciencedojo.com/blog/big-data-ethics-10-experiments/> /օգոստոսի 1 10։00/ [↑](#footnote-ref-34)
35. «7 Essential Data Ethics Examples for Businesses in 2025» December 26th, 2024 <https://atlan.com/data-ethics-examples/> /օգոստոսի 1 12։00/ [↑](#footnote-ref-35)
36. «Timeline of Cyber Incidents Involving Financial Institutions» <https://carnegieendowment.org/features/fincyber-timeline?lang=en> /օգոստոսի 1 12։00/ [↑](#footnote-ref-36)
37. Global Cybersecurity Index <https://www.itu.int/en/ITU-D/Cybersecurity/Pages/Global-Cybersecurity-Index.aspx> [↑](#footnote-ref-37)
38. GCI 2015 V1 <https://www.itu.int/dms_pub/itu-d/opb/str/D-STR-SECU-2015-PDF-E.pdf> /օգոստոսի 2 14։00/ [↑](#footnote-ref-38)
39. GCI 2017 V2 <https://www.itu.int/dms_pub/itu-d/opb/str/D-STR-GCI.01-2017-R1-PDF-E.pdf> /օգոստոսի 2 14։00/ [↑](#footnote-ref-39)
40. GCI 2018 V3 <https://www.itu.int/dms_pub/itu-d/opb/str/D-STR-GCI.01-2018-PDF-E.pdf> /օգոստոսի 2 14։00/ [↑](#footnote-ref-40)
41. GCI 2020 V4 <https://www.itu.int/epublications/publication/D-STR-GCI.01-2021-HTM-E> /օգոստոսի 2 14։00/ [↑](#footnote-ref-41)
42. GCI 2024 V5 <https://www.itu.int/epublications/publication/global-cybersecurity-index-2024> /օգոստոսի 2 14։00/ [↑](#footnote-ref-42)
43. Գծապատկերը կազմվել է հեղինակի կողմից, որի համար հիմք են հանդիսացել GCI 2020 V4 [https://www.itu.int/epublications/publication/D-STR-GCI.01-2021-HTM-E GCI 2020 V4](https://www.itu.int/epublications/publication/D-STR-GCI.01-2021-HTM-E%20GCI%202020%20V4) և GCI 2024 V5 [https://www.itu.int/epublications/publication/global-cybersecurity-index-2024 GCI 2024 V5](https://www.itu.int/epublications/publication/global-cybersecurity-index-2024%20GCI%202024%20V5) տվյալները/օգոստոսի 2 14։00/ [↑](#footnote-ref-43)
44. GCI 2024 V5 <https://www.itu.int/epublications/publication/global-cybersecurity-index-2024> /օգոստոսի 2 14։00/ [↑](#footnote-ref-44)
45. Brianna Valleskey «AI Fraud Detection: How It Works», March 26, 2022 <https://www.inscribe.ai/fraud-detection/ai-fraud-detection> /հուլիսի 15 14։00/ [↑](#footnote-ref-45)
46. Գծապատկերը կազմվել է հեղինակի կողմից, որի համար հիմք են հանդիսացել Brianna Valleskey «AI Fraud Detection: How It Works», March 26, 2022 <https://www.inscribe.ai/fraud-detection/ai-fraud-detection> տվյալները /հուլիսի 15 14։00/ [↑](#footnote-ref-46)
47. Գծապատկերը կազմվել է հեղինակի կողմից, որի համար հիմք են հանդիսացել Saloni Thakkar «Enhancing Fraud Detection in Financial Transactions through Advanced AI Algorithms», October 2024 <https://www.researchgate.net/publication/384654151_Enhancing_Fraud_Detection_in_Financial_Transactions_through_Advanced_AI_Algorithms> տվյալները /հուլիսի 15 14։00/ [↑](#footnote-ref-47)
48. Saloni Thakkar «Enhancing Fraud Detection in Financial Transactions through Advanced AI Algorithms», October 2024 <https://www.researchgate.net/publication/384654151_Enhancing_Fraud_Detection_in_Financial_Transactions_through_Advanced_AI_Algorithms> /հուլիսի 15 14։00/ [↑](#footnote-ref-48)
49. Գծապատկերը կազմվել է հեղինակի կողմից, որի համար հիմք են հանդիսացել Saloni Thakkar «Enhancing Fraud Detection in Financial Transactions through Advanced AI Algorithms», October 2024 <https://www.researchgate.net/publication/384654151_Enhancing_Fraud_Detection_in_Financial_Transactions_through_Advanced_AI_Algorithms> տվյալները /հուլիսի 15 18։00/ [↑](#footnote-ref-49)
50. Գծապատկերը կազմվել է հեղինակի կողմից, որի համար հիմք են հանդիսացել Shivani Zoting, Aditi Shivarkar «AI in Fraud Management Market Size and Forecast 2025 to 2034» 15 Jul 2025 <https://www.precedenceresearch.com/ai-in-fraud-management-market> տվյալները /հուլիսի 15 15։00/ [↑](#footnote-ref-50)
51. Shivani Zoting, Aditi Shivarkar «AI in Fraud Management Market Size and Forecast 2025 to 2034» 15 Jul 2025 <https://www.precedenceresearch.com/ai-in-fraud-management-market> /հուլիսի 15 15։00/ [↑](#footnote-ref-51)
52. Գծապատկերը կազմվել է հեղինակի կողմից, որի համար հիմք են հանդիսացել Shivani Zoting, Aditi Shivarkar «AI in Fraud Management Market Size and Forecast 2025 to 2034» 15 Jul 2025 <https://www.precedenceresearch.com/ai-in-fraud-management-market> տվյալները /հուլիսի 15 15։00/ [↑](#footnote-ref-52)
53. «Fraud Detection and Prevention Market – Global Industry Analysis and Forecast (2023-2029)» June 2023 <https://www.maximizemarketresearch.com/market-report/global-fraud-detection-and-prevention-market/62819/> /հուլիսի 15 15։00/ [↑](#footnote-ref-53)
54. «Fraud Detection and Prevention Market – Global Industry Analysis and Forecast (2023-2029)» June 2023 <https://www.maximizemarketresearch.com/market-report/global-fraud-detection-and-prevention-market/62819/> /հուլիսի 15 15։00/ [↑](#footnote-ref-54)
55. Գծապատկերը կազմվել է հեղինակի կողմից, որի համար հիմք են հանդիսացել Shivani Zoting, Aditi Shivarkar «AI in Fraud Management Market Size and Forecast 2025 to 2034» 15 Jul 2025 <https://www.precedenceresearch.com/ai-in-fraud-management-market> տվյալները /հուլիսի 15 15։00/ [↑](#footnote-ref-55)
56. Dr. Arunkumar M. Sampath «AI/ML- and Generative AI-enabled Financial Fraud Detection» Feb 15, 2024 <https://medium.com/@arunsampath_47187/ai-ml-and-generative-ai-enabled-financial-fraud-detection-ba4700a8e2a1> /հուլիսի 15 15։00/ [↑](#footnote-ref-56)
57. Գծապատկերը կազմվել է հեղինակի կողմից, որի համար հիմք են հանդիսացել Dr. Arunkumar M. Sampath «AI/ML- and Generative AI-enabled Financial Fraud Detection» Feb 15, 2024 <https://medium.com/@arunsampath_47187/ai-ml-and-generative-ai-enabled-financial-fraud-detection-ba4700a8e2a1> տվյալները /հուլիսի 15 15։00/ [↑](#footnote-ref-57)
58. «Global AI In Fraud Detection Market Report» August 2024 <https://market.us/report/ai-in-fraud-detection-market/> /հուլիսի 15 15։00/ [↑](#footnote-ref-58)
59. Աղյուսակը կազմվել է հեղինակի կողմից Python-ի միջոցով սինթետիկ տվյալների գեներացմամբ [↑](#footnote-ref-59)
60. Գծապատկերի համար հիմք են հանդիսացել Աղյուսակ 5-ի տվյալները [↑](#footnote-ref-60)
61. Գծապատկերի համար հիմք են հանդիսացել Աղյուսակ 5-ի տվյալները [↑](#footnote-ref-61)
62. Գծապատկերի համար հիմք են հանդիսացել Աղյուսակ 5-ի տվյալները [↑](#footnote-ref-62)