WHAT THE PAPER SAYS · WHAT IT DOESN'T · WHY IT MATTERS

Medical AI-ն կարող է average-ում private լինել և միաժամանակ expose անել կոնկրետ patients-ին — հատկապես underrepresented groups-ին

Medical-AI model-ի «privacy-preserving» claim-ը սովորաբար հենվում է average risk-ի վրա: Այս study-ն membership inference risk-ը չափել է per patient՝ յոթ medical datasets և բազմաթիվ models-ի վրա: Pattern-ը անհարմար է. aggregate-ում safe models-ը որոշ individuals-ի membership-ը կարող են գրեթե perfect leak անել (attack AUC ≥0.95), exposed-ը disproportionally underrepresented groups-ից են, և risk-ը worsens է model size-ի հետ: Authors-ը առաջարկում են per-patient assessment, access control և differential privacy:

Written by Lucio Vaglio · Cross-checked by Michele Renda · Edited by Michele Renda · Figures by Laura Nesso and Nova Vela · AI-assisted, human-reviewed

Paper: *Disparate privacy risks from medical AI*, Moritz Knolle, Georg Kaissis, Daniel Rückert and colleagues (Technical University of Munich; Imperial College London; Hasso Plattner Institute), Nature (2026) · DOI: 10.1038/s41586-026-10688-0

Գաղտնիության երաշխիքը խոստում է յուրաքանչ մարդու, ոչ միջին ցուցանիշին

Երբ բժշկական ԱՔ մոդելը ներկայացվում է որպես «գաղտնիությունը պահպանող», այդ պնդումը սովորաբար հիմնվում է մեկ թվի վրա. բոլոր այն հիվանդների միջին հաշվով, որոնց տվյալներով մոդելը ուսուցվել է, հավանականությունը փոքր է, որ հնարավոր կլինի պարզել որևէ կոնկրետ մարդու տվյալների մասնակցությունը ուսուցման հավաքածուին: Սա հանգստացնող է հնչում: Այս հոդվածի անհարմար, բայց կարևոր միտքն այն է, որ սխալ ցուցանիշ ենք դիտում:

Գաղտնիությունը միջին արժեք չէ: Այն խոստում է յուրաքանչյուր անհատի, որ իր տվյալների ընդգրկումը ուսուցման հավաքածուում հետագայում չի բացահայտի իրեն: Իսկ միջինը կարող է այդ խոստումը կատարել գրեթե բոլորի համար և ամբողջովին խախտել այն մի քանի մարդու դեպքում: Հետազոտողները գաղտնիության ռիսկը չափել են հիվանդ առ հիվանդ՝ նախկինից շատ ավելի մանրամասն մակարդակով, և հենց դա էլ գտել են. ընդհանուր ցուցանիշներով անվտանգ թվացող մոդելները որոշ մարդկանց մասնակցությունը կարող են գրեթե անսխալ բացահայտել, և այդ մարդիկ անհամաչափ հաճախ հենց այն խմբերից են, որոնք արդեն ամենաքիչն են պաշտպանված:

Ի՞նչ արեցին հեղինակները

Թիմը՝ Moritz Knolle-ի ղեկավարությամբ, Daniel Rückert-ի, Georg Kaissis-ի և գործընկերների մասնակցությամբ, վերցրել է գաղտնիության հայտնի սպառնալիք՝ **ուսուցման հավաքածուին անդամակցության որոշման հարձակումը** (*membership inference attack*) և փոխել հարցը: «Միջին հաշվով այս հարձակումը որքան հաճախ է հաջողվում ամբողջ տվյալների հավաքածուի վրա» հարցի փոխարեն նրանք հարցրել են. «**Այս կոնկրետ հիվանդը** որքանո՞վ է բացահայտված»:

Նրանք այդ անհատական վերլուծությունն անցկացրել են **յոթ հաստատված բժշկական տվյալների հավաքածուների** վրա, որոնք ներառում էին տարբեր տեսակի տվյալներ՝ բժշկական պատկերներ, էլեկտրասրտագրություններ և էլեկտրոնային առողջապահական գրառումներ: Յուրաքանչյուր հավաքածուի համար ուսումնասիրել են 200 մոդել: Ամեն հիվանդի համար գնահատվել է, թե հարձակվողը որքան վստահ կարող է որոշել՝ տվյալ մարդու գրառումն ընդգրկվա՞ծ է եղել ուսուցման տվյալներում, ապա արդյունքները բաժանվել են ըստ հիվանդության կարգավիճակի, ինքնահայտարարված ռասայի, ապահովագրության, սեռի և պատկերավորման արձանագրության:

> **Ի՞նչ է իրականում membership inference attack-ը**
>
> Այստեղ արտահոսքը ավելի նուրբ է, քան «մոդելը դուրս է տալիս ձեր բժշկական

գրառումը»: Խոսքը **անդամակցության** մասին է՝ հենց այն փաստի, որ ձեր տվյալները եղել են ուսուցման հավաքածուում:

Սկսենք սովորական կիրառությունից: Մոդելին տալիս եք, օրինակ, կրծքավանդակի ռենտգեն և ստանում հավանականություններ՝ թոքաբորբ՝ 78%, ինչպես նաև գնահատումներ սրտի մեծացման, այտուցի, կոնսոլիդացիայի և այլ ախտանիշների համար: Հարձակումը օգտագործում է միայն այս սովորական ախտորոշիչ պատասխանը: Հատուկ գաղտնիության միջերես պետք չէ:

Թուլությունը հետևյալն է. մոդելը հաճախ փոքրինչ **ավելի վստահ** է հենց այն օրինակների նկատմամբ, որոնց վրա ուսուցվել է, քան բոլորովին նոր օրինակների նկատմամբ: Պատկերացրեք ուսանողի, որը նախապես տեսել է քննության հարցերը. ծանոթ հարցերի դեպքում պատասխանը մի փոքր չափազանց արագ ու վստահ է գալիս: Այդ տարբերությունից պարզել, թե կոնկրետ գրառումն ուսուցման օրինակներից մեկն էր, հենց membership inference-ի էությունն է:

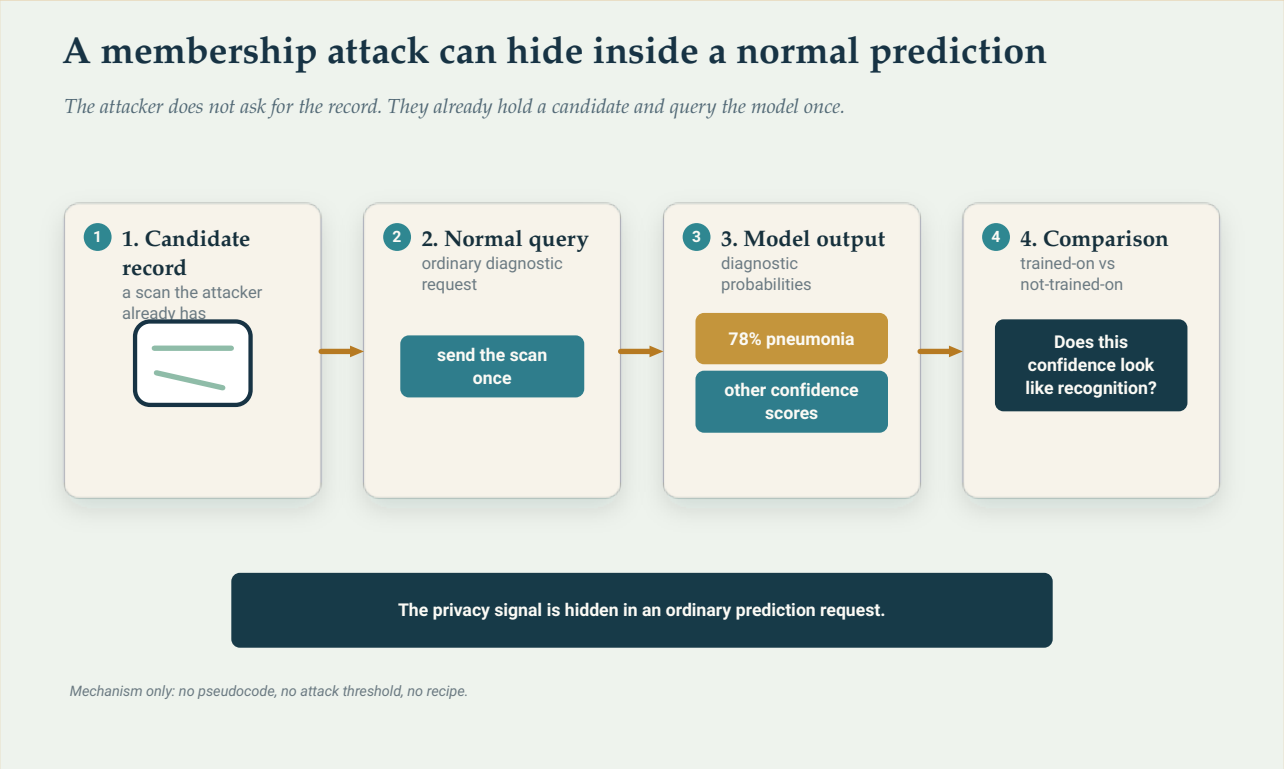
Հարձակվողն ուզում է իմանալ՝ որևէ մարդու պատկերն օգտագործվե՞լ է մոդելի ուսուցման համար: Նա մոդելից չի խնդրում բժշկական գրառումը. թեկնածու պատկերը արդեն ունի՝ տվյալ մարդու իրական պատկերը կամ դրա մոտ տարբերակը: Այն ուղարկում է մոդելին որպես սովորական ախտորոշիչ հարցում և գրանցում պատասխանի վստահությունը: Հետո համեմատում է՝ այդ վստահությունն ավելի նման է մոդելի, որը **տեսել է** տվյալ պատկերը ուսուցման ժամանակ, թե մոդելի, որը **չի տեսել**: Այդ տարբերությունը կարելի է սովորել սեփական «հղումային մոդելի» միջոցով՝ առանց հատուկ սարքավորման: Եթե իրական մոդելը տվյալ պատկերի նկատմամբ անսովոր վստահ է, դա կարող է ուժեղ հուշում լինել, որ պատկերը եղել է ուսուցման տվյալներում:

Անհանգստացնողն այն է, որ հարցումն արտաքինից հարձակման նման չէ. նույնպիսի սովորական ախտորոշիչ հարցում է, ինչ կարող էր անել բժիշկը: Գաղտնիության ազդանշանը թաքնված է սովորական կանխատեսման ներսում: Այդ պատճառով ռիսկը գործնական նշանակություն ունի, թեև այս աշխատանքի շրջանակում այն ցուցադրվել է վերահսկվող լաբորատոր պայմաններում, ոչ իրական գրանցված հարձակումներով:

Ինչո՞ւ է անդամակցությունն ինքնին կարևոր, եթե բժշկական գրառման բովանդակությունը չի արտահոսում: Որովհետև **անդամակցությունն էլ փաստ է**: Եթե մոդելն ուսուցվել է, օրինակ, որոշակի քաղցկեղային բուժում ստացած հիվանդների տվյալներով, ապա հաստատելը, որ ձեր գրառումն այնտեղ է, կարող է բացահայտել, որ դուք հավանաբար ունեցել եք այդ հիվանդությունը: Դա զգայուն տեղեկություն է, նույնիսկ եթե իրական գրառումը երբեք չի ցուցադրվում:

Հեղինակները բաց ձևով թվարկում են հարձակման ենթադրությունները՝ մոդելի սովորական կանխատեսումներին հասանելիություն, թեկնածու գրառում և հարձակվողի սեփական հղումային մոդել: Նպատակը գործողությունների հրահանգ տալը չէ, այլ սպառնալիքը չափելի և քննարկելի

> դարձնելը:



Հարձակման համար հատուկ գաղտնիության API պետք չէ: Սովորական ախտորոշիչ հարցումը կարող է անդամակցության ազդանշան արտահոսեցնել, եթե մոդելը նախկինում տեսած գրառման նկատմամբ անսովոր վստահ է:

Ի՞նչ գտան

Երեք հիմնական արդյունք կա, և յուրաքանչյուր հաջորդը ավելի սուր է:

Միջինները թաքցնում են ամենաբացահայտված մարդկանց: Սովորական՝ ընդհանուր չափումով, շատ մոդելներ բավական անվտանգ են թվում. հիվանդների մեծ մասի դեպքում հարձակումը պատահական գուշակությունից լավ չի աշխատում: Բայց հիվանդ առ հիվանդ պատկերը լրիվ այլ է: Knolle-ն ուսումնասիրության հայտարարության մեջ նշել է, որ նախկին գնահատումները չափել են միայն բոլոր հիվանդների միջին ռիսկը, մինչ այս աշխատանքն առաջին անգամ նայում է անհատական մակարդակին: Որոշ մարդկանց դեպքում հարձակումը հաջողվում է **գրեթե կատարյալ**՝ AUC **0,95 կամ ավելի** (0,5-ը պատահական գուշակություն է, 1,0-ը՝ կատարյալ տարբերակում), նույնիսկ երբ տվյալների հավաքածուի միջինը մոտ է պատահականությանը:

Միջին ցուցանիշը կարող է մնալ պատահականության մակարդակին մոտ, մինչ գրառումների փոքր «պոչը» շատ ավելի բացահայտված է: Հոդվածի հիմնական միտքն այն է, որ գաղտնիությունը պետք է ստուգել հիվանդի մակարդակով, ոչ միայն ընդհանուր միջինով:

Original Aurora diagram — The Clean Paper CC BY 4.0

Բացահայտվածությունը հավասար չի բաշխվում, և ամենաշատը տուժում են թերներկայացված խմբերը: Գրեթե կատարյալ հարձակման նկատմամբ ամենախոցելի հիվանդները համակարգված կերպով պատկանում էին տվյալներում **թերներկայացված խմբերի**՝ էթնիկ փոքրամասնություններ, հազվադեպ հիվանդության ֆենոտիպեր կամ անսովոր պատկերային հատկանիշներ: Էլեկտրոնային գրառումների հավաքածուում սևամորթ հիվանդները ամենախոցելի գրառումների մեջ հանդիպել են սպասվածից **31% ավելի հաճախ**: Մամոգրաֆիայի հավաքածուում չարորակության կասկածով նշված պատկերները ծայրահեղ ռիսկի խմբում գերներկայացված էին **+1 179%**: Այս թվերը օրինակներ են, ոչ ամբողջ պատկերը, և դրանք նկարագրում են հարաբերական գերներկայացվածությունը փոքր՝ ամենաբարձր ռիսկի պոչում, ոչ այն, թե տվյալ խմբի քանի տոկոսն է բացահայտված: Եթե մոդելը նման մարդկանց համար քիչ համադրելի օրինակներ ունի, նրանց գրառումները ավելի հեշտ են առանձնանում, իսկ հենց «առանձնանալը» membership inference-ի օգտագործած ազդանշանն է: Այսպիսով գաղտնիության ձախողումը պատահական չէ. այն կենտրոնանում է արդեն սահմանային դիրքում գտնվող մարդկանց վրա:

Ավելի մեծ մոդելներն այս խնդիրը վատացնում են: Գրեթե կատարյալ հարձակման ենթակա հիվանդների թիվը աճում է մոդելի կարողության հետ: Մաշկաբանական տվյալների մեկ հավաքածուում բաժինը փոքրագույն մոդելի դեպքում գրեթե զրոյից հասել է մոտ **տասից մեկին** ամենամեծ մոդելի դեպքում: Քանի որ բժշկական ԱԲ

համակարգերը ընդհանուր առմամբ դառնում են ավելի մեծ և կարող, հեղինակները զգուշացնում են, որ այս կոնկրետ ռիսկը կարող է աճել, ոչ նվազել։

Rückert-ի ամփոփումը կտրուկ է. «Սա ընդունելի ռիսկ չէ։ Առողջապահական տվյալները չափազանց զգայուն են»։

Ինչու «միջինում անվտանգ»-ը սխալ թեստ է

Ցածր միջին ռիսկը հեշտ է ընկալել որպես ընդհանուր անվտանգության վկայագիր։ Հոդվածի հիմնական դասը այն է, որ այստեղ միջինացումը պարզապես ոչ ճշգրիտ չէ. այն չափում է սխալ բանը։

Գաղտնիության երաշխիքը իմաստ ունի միայն այն դեպքում, երբ գործում է նաև ամենաբարձր ռիսկ ունեցող մարդու համար, ոչ միայն միջին հիվանդի։ Եթե 1000 հիվանդից 999-ը գործնականում անճանաչելի են, բայց մեկի մասնակցությունը կարելի է գրեթե անսխալ որոշել, այդ մոդելը տվյալ մեկ մարդու համար «99,9% գաղտնի» չէ։ Իսկ եթե այդ մարդը կանխատեսելիորեն հազվադեպ հիվանդություն ունեցող կամ էթնիկ փոքրամասնության անդամ հիվանդն է, չափումը ոչ միայն թերի է, այլ նաև անհավասար։ Այն մեծամասնության անվտանգությունը ներկայացնում է որպես բոլորի անվտանգություն։

Հոդվածի առաջարկած հայացքի փոփոխությունը սա է՝ «միջինում որքանո՞վ է գաղտնի այս մոդելը» հարցից անցնել «ո՞վ է ամենաբացահայտված հիվանդը, և ո՞ր խմբին է նա պատկանում» հարցին։ Սրանք տարբեր հարցեր են, և գաղտնիության համար երկրորդն է որոշիչ։

Ի՞նչ սա *չի* ապացուցում կամ պնդում

- Սա **չի ասում, որ բժշկական ԱԲից պետք է հրաժարվել**։ Հեղինակների նպատակը ռիսկը չափել և նվազեցնելն է, ոչ տեխնոլոգիան դադարեցնելը։
- Սա **չի նշանակում, որ յուրաքանչյուր բժշկական մոդել արտահոսք ունի կամ որ որևէ կոնկրետ կիրառվող մոդել արդեն հարձակման է ենթարկվել**։ Աշխատանքը ցույց է տալիս, որ ռիսկը գոյություն ունի, անհավասար է բաշխված և ընդհանուր չափումները կարող են այն չտեսնել։
- Սա **չի նշանակում, որ ձեր բժշկական տվյալներն արդեն բացահայտված են**։ Հարձակումը պահանջում է հատուկ պայմաններ՝ մոդելի հասանելիություն, թեկնածու գրառում և հարձակվողի սեփական ենթակառուցվածք։
- Սա **չի ցույց տալիս հիվանդի գրառման բովանդակության արտահոսք**։ Արտահոսող տեղեկությունը անդամակցությունն է՝ տվյալ հավաքածուում ընդգրկված լինելու փաստը, որը կարող է վտանգավոր լինել այլ պատճառով։
- Սա **անհավասարությունները չի ամփոփում մեկ գլխավոր թվով**։ Կրկնվող ուժեղ արդյունքը օրինաչափությունն է՝ ընդհանուր չափումները թերագնահատում

են անհատական ռիսկը, և թերներկայացված հիվանդները ամենաշատն են տուժում:

Որքա՞ն ուժեղ է ապացույցը

Սա էմպիրիկ և մեթոդաբանական արդյունք է, և բավական ամուր: Այն օրինաչափությունը, որ ընդհանուր գաղտնիության չափումները համակարգված կերպով թերագնահատում են առանձին հիվանդների ռիսկը, մնացորդային ռիսկը կենտրոնանում է թերներկայացված խմբերում և մեծանում է մոդելի կարողության հետ, կրկնվել է **յոթ տարբեր բժշկական տվյալների հավաքածուներում** և յուրաքանչյուրում մեծ թվով մոդելների վրա: Հենց այս լայնությունն է արդյունքը դարձնում ավելի վստահելի, քան մեկ տվյալների հավաքածուի պատահականություն:

Երկու կարևոր վերապահում կա: Նախ՝ սա **ռիսկի ցուցադրում** է՝ հեղինակների իրականացրած հարձակումներով: Այն գնահատում է, թե իրենց ենթադրությունների պայմաններում ինչ կարող է անել կարող հարձակվողը, ոչ թե իրական աշխարհում նման հարձակումների հաճախականությունը: Երկրորդ՝ գրեթե կատարյալ հաջողության տպավորիչ թվերը դիտավորյալ վերաբերում են ամենախոցելի անհատներին: Դա հենց աշխատանքի նպատակը է. դրանք պետք է կարդալ որպես «ռիսկի պոչը շատ ավելի ծանր է, քան միջինը հուշում է», ոչ «հիվանդների մեծ մասը բացահայտված է»:

Ճիշտ վերաբերմունքը ոչ խուճապն է, ոչ անտեսումը: Սա բազմահավաքածու ուսումնասիրություն է, որը ցույց է տալիս, որ բժշկական ԱԲի գաղտնիության ստանդարտ ստուգումը կարող է կույր լինել իր վատագույն դեպքերի նկատմամբ, և այդ վատագույն դեպքերը հաճախ ընկնում են արդեն առավել խոցելի հիվանդների վրա:

Ինչո՞ւ է սա կարևոր

Երկու պատճառ այս աշխատանքը դարձնում է ավելին, քան տեխնիկական մանրուք:

Առաջինը՝ այն փոխում է, թե ինչ պետք է թույլ տանք կոչել «գաղտնիությունը պահպանող»: Եթե մոդելը թողարկվում է միայն միջին գաղտնիության ցուցանիշով, այդ ցուցանիշը կարող է իրապես ցածր լինել, բայց միաժամանակ թաքցնել հիվանդների փոքր ենթախումբ, որոնց անդամակցությունը գրեթե անսխալ որոշելի է: Հեղինակների գործնական պահանջը կոնկրետ է՝ մինչև թողարկումը գաղտնիության ռիսկը գնահատել **յուրաքանչյուր հիվանդի մակարդակով**, վերահսկել կիրառվող մոդելներին հասանելիությունը և օգտագործել մեթոդներ, օրինակ՝ **դիֆերենցիալ գաղտնիություն**: Վերջինը ուսուցման ժամանակ փոքր, ճշգրիտ կարգաբերված աղմուկ է ավելացնում՝ թուլացնելով membership inference հարձակումները, բայց որոշ գին վճարելով մոդելի օգտակարության մեջ: Գաղտնիության և օգտակարության

փոխզիջումը իրական է, ոչ անվճար: «Մենք ստուգել ենք միջինը» այլևս չպետք է համարվի լիարժեք ստուգում:

Երկրորդը՝ գաղտնիությունն այստեղ կապվում է արդարության հետ: Նույն խմբերը, որոնք բժշկական տվյալներում թերներկայացված են և հետևաբար արդեն կարող են ԱԲից ավելի վատ սպասարկում ստանալ, նաև այն խմբերն են, որոնց գաղտնիությունը համակարգը ամենաքիչն է պաշտպանում: Այս երկու անհավասարությունները տարբեր բաժինների խնդիրներ չեն. խոսքը նույն մարդկանց մասին է:

Բժշկական ԱԲ գաղտնիության հանգստացնող պատմությունը՝ «չափեցինք, միջինը ցածր է, ուրեմն ամեն ինչ լավ է», հենց այն պատմությունն է, որը հոդվածը քանդում է: Ոչ թե տեխնոլոգիայից վախեցնելու համար, այլ չափանիշը ճիշտ տեղ տեղափոխելու՝ գաղտնիությունը խոստում է, որը կարող ես հայտարարել միայն այն դեպքում, երբ ստուգել ես նաև ամենախոցելի մարդու համար:

Կարճ ամփոփում

Membership inference հարձակումները փորձում են որոշել՝ կոնկրետ մարդու գրառումը եղե՞լ է ԱԲ մոդելի ուսուցման տվյալներում: Քանի որ անդամակցության փաստն ինքնին կարող է զգայուն տեղեկություն բացահայտել, սա իրական գաղտնիության արտահոսք է նույնիսկ այն դեպքում, երբ բժշկական գրառման բովանդակությունը երբեք չի ցուցադրվում: Նախկին աշխատանքները նման հարձակումների հաջողությունը հիմնականում չափել են տվյալների հավաքածուի **միջինով**: Այս ուսումնասիրությունը չափել է այն **հիվանդ առ հիվանդ**՝ յոթ բժշկական տվյալների հավաքածուներում՝ պատկերներ, ԷՍԳ և էլեկտրոնային գրառումներ, բազմաթիվ մոդելների վրա: Ընդհանուր չափումները զգալիորեն թերագնահատում են ռիսկը. որոշ անհատներ կարող են գրեթե կատարյալ ճանաչվել (հարձակման AUC $\geq$ 0,95), նույնիսկ երբ ընդհանուր միջինը անվտանգ է թվում: Ամենախոցելի հիվանդները համակարգված կերպով թերներկայացված խմբերից են՝ էթնիկ փոքրամասնություններ, հազվադեպ հիվանդություններ, անսովոր պատկերային հատկանիշներ, և խնդիրը մեծանում է մոդելի չափի հետ: Հեղինակները չեն առաջարկում հրաժարվել բժշկական ԱԲից. առաջարկում են գաղտնիությունը չափել անհատական մակարդակով, վերահսկել մոդելների հասանելիությունը և կիրառել դիֆերենցիալ գաղտնիություն: Հիմնական միտքը՝ «միջինում գաղտնի» լինելը գաղտնիության երաշխիք չէ, և այդ միջինի թաքցրած ձախողումները հաճախ ընկնում են ամենաքիչ պաշտպանված մարդկանց վրա:

Առանց չափազանցության

Ի՞նչ է ցույց տալիս աշխատանքը: Յոթ բժշկական տվյալների հավաքածուներում և բազմաթիվ մոդելների վրա անհատական membership inference ռիսկը որոշ

հիվանդների համար շատ ավելի բարձր է, քան ընդհանուր չափումները ցույց են տալիս՝ մինչև գրեթե կատարյալ հարձակման հաջողություն (AUC ≥ 0,95)։ Մնացորդային ռիսկը անհամաչափ ընկնում է թերներկայացված խմբերի վրա և աճում է մոդելի կարողության հետ։

Ի՞նչն է հավանական, բայց ապացուցված չէ։ Որ իրական հարձակվողները արդեն այս եղանակով թիրախավորում են կիրառվող կլինիկական մոդելները։ Ուսումնասիրությունը ցուցադրում է **հնարավորությունն ու դրա բաշխումը** իր նշված ենթադրությունների ներքո, ոչ իրական հարձակումների հաճախականությունը։

Ի՞նչ չի ցույց տալիս։ Որ բժշկական ԱԲից պետք է հրաժարվել, որ յուրաքանչյուր մոդել տվյալ է արտահոսեցնում կամ որևէ կոնկրետ կիրառվող համակարգ արդեն կոտրվել է, որ բացահայտվում է հիվանդի գրառման բովանդակությունը, կամ որ անհավասարությունը կարելի է ամփոփել մեկ թվով։ Ուժեղ արդյունքը կրկնվող օրինաչափությունն է։

Հիմնական սահմանափակումները։ Ռիսկը չափվում է հեղինակների կառուցած հարձակումներով, ոչ իրական դեպքերի դիտարկմամբ։ Ամենատպավորիչ թվերը դիտավորյալ նկարագրում են ամենաբացահայտված անհատներին, իսկ խմբային անհավասարությունները ներկայացվում են տարբեր տվյալների հավաքածուներում կրկնվող օրինաչափությամբ, ոչ մեկ միասնական ցուցանիշով։

Որքա՞ն վստահություն պետք է ունենա ընդհանուր ընթերցողը։ Բարձր՝ որ ընդհանուր գաղտնիության չափումները կարող են թերագնահատել անհատական ռիսկը, որ մնացորդային ռիսկը կենտրոնանում է թերներկայացված հիվանդների վրա և որ այն մեծանում է մոդելի կարողության հետ։ Միջին՝ իրական աշխարհում նման հարձակումների հաճախականության նկատմամբ, որը այս ուսումնասիրությունը չի չափում։ Ճիշտ ընթերցումը ոչ «բժշկական ԱԲն արտահոսեցնում է ձեր տվյալները» է, ոչ «գաղտնիությունը լուծված է», այլ՝ «ստանդարտ ստուգումը կարող է չտեսնել իր վատագույն դեպքերը, և այդ դեպքերը հաճախ վերաբերում են ամենախոցելի հիվանդներին, ուստի ստուգումը պետք է փոխվի»։

Աղբյուրներ

Knolle et al., Disparate privacy risks from medical AI, Nature (2026)
https://doi.org/10.1038/s41586-026-10688-0

Technical University of Munich — press release, ’Study reveals privacy risks in medical AI’ (2026)
https://www.tum.de/en/news-and-events/all-news/press-releases/details/study-reveals-privacy-risks-in

Medical Xpress (2026) — coverage of the study
https://medicalxpress.com/news/2026-06-patient-groups-vulnerable-privacy-medical.html