УДК 004

Д.Тарковський, магістр гр. КН-21М-1,4,
Центральноукраїнський національний технічний університет

ДОСЛІДЖЕННЯ ТА ПРОГРАМНА РЕАЛІЗАЦІЯ СИСТЕМИ ВИЗНАЧЕННЯ РІВНЯ СТІЙКОСТІ СЕРВІСІВ ЗАБЕЗПЕЧЕННЯ КОНФІДЕНЦІЙНОСТІ НА ОСНОВІ МЕТОДІВ AI

У статті розроблено програмне забезпечення, яке призначено для системи визначення рівня стійкості сервісів забезпечення конфіденційності на основі методів AI. Метою розробки є дослідження та програмна реалізація системи визначення рівня стійкості сервісів забезпечення конфіденційності на основі методів AI. Об'єктом дослідження є процес визначення рівня стійкості сервісів забезпечення конфіденційності на основі методів AI. Предметом дослідження є методи визначення рівня стійкості сервісів забезпечення конфіденційності на основі методів AI. Методи дослідження базуються на методах захисту інформації та штучного інтелекту (AI), методах математичної статистики, методах розробки програмного забезпечення. Результат роботи – програмна реалізація системи визначення рівня стійкості сервісів забезпечення конфіденційності на основі методів AI. В процесі роботи над програмною моделлю виконано аналіз існуючих апаратних та програмних засобів. В повній мірі описані всі компоненти розробленого програмного забезпечення.

комп'ютерні науки, стійкість, конфіденційність, штучний інтеллект

Постановка проблеми. Питання конфіденційності займають перше місце в онлайн-діяльності, бізнес-діях і державних рішеннях. Це здебільшого у відповідь на зломи, скандали та витік особистих даних, які підірвали довіру до технологій та інформаційних систем.

У звіті Консультативного комітету з питань телекомунікацій національної безпеки про кібербезпеку Moonshot йдеться, що конфіденційність є ключовим компонентом кібербезпеки, і що ми повинні змінити наратив, щоб відновити довіру американців до інформаційних систем. Щоб досягти цього, до 2028 року потрібно «гарантувати», що технологічний прогрес більше не загрожуватиме конфіденційності, а натомість підвищить гарантії конфіденційності завдяки безпеці та безпеці їхніх особистих даних.

Одним з найважливіших елементів у майбутніх технологічних досягненнях і онлайн-безпеці є посилений розвиток штучного інтелекту (AI). Проте принципи конфіденційності необхідно враховувати на ранніх етапах процесу розробки штучного інтелекту, щоб збалансувати технологічні переваги та зберегти конфіденційність.

Більшість відомих експертів-криптоаналітиків вважають що, штучний інтелект, тим більше нейрокомп'ютерні мережі, не кращий інструмент криптоаналізу. Підстав у цієї широко поширеної думки дуже багато, наприклад, низька можливість для навчання.

Недоліками названого підходу є, насамперед те, що для кожної нової криптографічної системи необхідно розробляти нову методику навчання елементів нейромережі.

Аналіз останніх досліджень і публікацій. При аналізі останніх досліджень і публікацій [1-10] було виявлено певні прогалини у забезпеченні системи визначення рівня стійкості сервісів забезпечення конфіденційності на основі методів AI.

Мета й завдання дослідження. Метою роботи є дослідження та програмна реалізація системи визначення рівня стійкості сервісів забезпечення конфіденційності на основі методів AI.

Для досягнення поставленої мети визначена програма дослідження, що складається з наступних завдань:

– Огляд існуючих систем визначення рівня стійкості сервісів забезпечення конфіденційності на основі методів AI.

– Дослідження системи визначення рівня стійкості сервісів забезпечення конфіденційності на основі методів AI.

– Програмна реалізація системи визначення рівня стійкості сервісів забезпечення конфіденційності на основі методів AI.

Об'єктом дослідження є процес визначення рівня стійкості сервісів забезпечення конфіденційності на основі методів AI.

Предметом дослідження є методи визначення рівня стійкості сервісів забезпечення конфіденційності на основі методів AI.

Методи дослідження базуються на методах захисту інформації та штучного інтелекту (AI), методах математичної статистики, методах розробки програмного забезпечення.

Виклад основного матеріалу.

Давайте приділимо хвилинку, щоб дослідити наслідки та потенційні наслідки збільшення впровадження штучного інтелекту онлайн. Хоча це здається футуристичним, коли штучний інтелект починає «думати» як люди або навіть замість людей, це може загрожувати трьом основним принципам конфіденційності – точності даних, захисту та контролю:

– Точність даних: щоб штучний інтелект давав точні результати, алгоритми повинні містити великі та репрезентативні набори даних. Недостатня представленість певних груп у наборах даних може призвести до неточних результатів і навіть шкідливих рішень. Це алгоритмічне зміщення часто створюється ненавмисно. Наприклад, дослідники виявили, що розумні мовці не можуть зрозуміти жіночі голоси або голоси меншин, тому що алгоритми побудовані з баз даних, які містять переважно голоси білих чоловіків. З огляду на це, що станеться, якщо ми довіримо AI приймати наші дзвінки в екстрену допомогу?

– Захист даних. Незважаючи на те, що великі набори даних дають більш точні та репрезентативні результати, вони створюють більший ризик конфіденційності, якщо їх порушують. Штучний інтелект може легко деанонімізувати навіть особисті дані, здавалося б, анонімні. Зокрема, дослідники виявили мінімальну анонімність навіть у грубих наборах даних, що призводить до повторної ідентифікації до 95 відсотків. Разом це означає, що ви ризикуєте бути легко ідентифікованим і отримати витік ваших даних, якщо не врахувати міркувань конфіденційності. Використання штучного інтелекту також може призвести до попередження, коли використовується для обробки податків і аналізу права на отримання федеральних пільг.

– Контроль даних: коли штучний інтелект починає бачити та визначати закономірності, він робить висновки та може приймати рішення щодо вас, щоб зробити вашу роботу в Інтернеті легшою або надійнішою. Однак, коли штучний інтелект дає хибні або несприятливі результати, це викликає питання, чи були прийняті рішення справедливими. Наприклад, AI, який використовується для оцінки кредитних ризиків, може ненавмисно скоротити кредитні лінії осіб, які відповідають певним профілям. Ці рішення можуть бути прийняті без вашого відома, згоди чи вибору, особливо якщо дані, які керують цими рішеннями, збираються без вашого відома. Більше того, штучний інтелект може отримати додаткові відомості про вас, наприклад ваші політичні уподобання, расу та релігію, навіть якщо ви ніколи не публікували ці подробиці в Інтернеті.

Суть полягає в тому, що особисті дані можуть використовуватися, а іноді і проти вас, без жодного контролю. Хороша новина полягає в тому, що розробники можуть мінімізувати проблеми конфіденційності на етапі розробки, задовго до виробництва. Таким чином ми все ще можемо реалізувати технологічні переваги AI, не порушуючи конфіденційність людей. Щоб підвищити конфіденційність, ми пропонуємо додати штучний інтелект до стратегії керування даними вашої організації та виділити ресурси не лише на розробку продукту штучного інтелекту, але й на конфіденційність, безпеку та моніторинг.

Інші способи захисту конфіденційності в AI включають:

1. Використовуйте належну гігієну даних. Слід збирати лише ті типи даних, які необхідні для створення AI, а дані мають зберігатися в безпеці та підтримуватися лише стільки часу, скільки необхідно для досягнення мети.

2. Використовуйте хороші набори даних. Розробники повинні створювати AI, використовуючи точні, справедливі та репрезентативні набори даних. Де можливо, розробники повинні створювати алгоритми штучного інтелекту, які будуть перевіряти та забезпечувати якість інших алгоритмів.

3. Надайте користувачам контроль. Користувачі повинні знати, коли використовуються їхні дані, чи використовується AI для прийняття рішень щодо них і чи використовуються їхні дані для створення AI. Їм також слід надати можливість дати згоду на використання таких даних.

4. Зменшити алгоритмічне зміщення. Переконайтеся, що набори даних є широкими та всеосяжними під час «навчання» AI. Алгоритмічні зміщення найчастіше становлять проблеми для жінок, меншин і груп (наприклад, осіб з порушеннями голосу, людей похилого віку), які складають лише невелику частину технологічної робочої сили.

Наш світ переживає інформаційний Великий вибух, під час якого всесвіт даних подвоюється кожні два роки, а щодня генеруються квінтильйони байтів даних.Протягом десятиліть закон Мура про подвоєння обчислювальної потужності кожні 18-24 місяці стимулював розвиток інформаційних технологій. Тепер, коли мільярди смартфонів та інших пристроїв збирають і передають дані через високошвидкісні глобальні мережі, зберігають дані у все більших центрах обробки даних і аналізують їх за допомогою все більш потужного та складного програмного забезпечення, закон Меткалфа вступає в дію. Він розглядає цінність мереж як функцію квадрата кількості вузлів, тобто мережеві ефекти експоненціально посилюють цей історичний ріст інформації. У міру розгортання мереж 5G і, зрештою, квантових обчислень, цей вибух даних зростатиме ще швидше та масштабніше.

Вплив великих даних зазвичай описують у термінах: обсяг, різноманітність і швидкість. Більше даних робить аналіз потужнішим і детальнішим. Різноманітність додає цій потужності та дозволяє робити нові та непередбачені висновки та прогнози. А швидкість полегшує аналіз, а також обмін у реальному часі.

Ймовірно, штучний інтелект прискорить цю тенденцію. Значна частина найбільш У міру розвитку штучного інтелекту він розширює можливості використання особистої інформації у спосіб, який може порушити інтереси конфіденційності, підвищуючи ефективність і швидкість аналізу особистої інформації.

Системи розпізнавання обличчя пропонують попередній перегляд проблем конфіденційності, які виникають. Завдяки багатим базам даних цифрових фотографій, доступних через соціальні мережі, веб-сайти, реєстри водійських прав, камери спостереження та багато інших джерел, машинне розпізнавання облич швидко прогресувало від нечітких зображень котів до швидкого (хоча все ще недосконалого) розпізнавання окремі люди. Системи розпізнавання облич розгортаються в містах і аеропортах Америки. Однак використання Китаєм розпізнавання облич як інструменту авторитарного контролю в Сіньцзяні та інших місцях викликало опозицію до цього розширення та закликає заборонити використання розпізнавання облич. Через занепокоєння щодо розпізнавання облич міста Окленд, Берклі та Сан-Франциско в Каліфорнії, а також Бруклін, Кембридж, Нортгемптон і Сомервіль у Массачусетсі прийняли заборону на цю технологію. У Каліфорнії, Нью-Гемпширі та Орегоні прийнято законодавство, яке забороняє використання розпізнавання обличчя за допомогою поліцейських натільних камер.

У цій аналітичній записці досліджується взаємозв’язок між AI та поточними дебатами щодо конфіденційності. Оскільки Конгрес розглядає всеосяжне законодавство про конфіденційність, щоб заповнити дедалі більші прогалини в поточній таблиці конфіденційності на федеральному рівні та штаті, йому потрібно буде розглянути питання про використання особистої інформації в системах штучного інтелекту. У цьому короткому описі я обговорюю деякі потенційні проблеми щодо штучного інтелекту та

конфіденційності, зокрема дискримінацію, етичне використання та контроль людини, а також варіанти політики, що обговорюються.

Проблеми конфіденційності в AI

Завдання Конгресу полягає в тому, щоб ухвалити законодавство про конфіденційність, яке захищатиме людей від будь-яких несприятливих наслідків використання особистої інформації в штучному інтелекті, але без необґрунтованого обмеження розвитку штучного інтелекту чи втягування законодавства про конфіденційність у складні соціальні та політичні хащі. Обговорення штучного інтелекту в контексті дебатів щодо конфіденційності часто викликає обмеження та помилки систем штучного інтелекту, таких як інтелектуальна поліція, яка може непропорційно вплинути на меншини або невдалий експеримент Amazon із алгоритмом найму, який повторює існуючу непропорційно чоловічу робочу силу компанії. Обидва вони порушують значні проблеми, але законодавство про конфіденційність досить складне навіть без урахування всіх соціальних і політичних проблем, які можуть виникнути в результаті використання інформації. Щоб оцінити вплив штучного інтелекту на конфіденційність, необхідно розрізняти проблеми з даними, які характерні для всіх видів штучного інтелекту, як-от випадки помилкових спрацьовувань і негативних результатів або надмірне пристосування до шаблонів, і ті, які характерні для використання особистої інформації.

Законодавчі пропозиції щодо конфіденційності, які стосуються цих питань, не стосуються штучного інтелекту. Швидше, вони посилаються на «автоматизовані рішення» (запозичені із законодавства ЄС про захист даних) або «алгоритмічні рішення» (використовуються в цьому обговоренні). Ця мова зміщує увагу людей від використання штучного інтелекту як такого до використання особистих даних у штучному інтелекті та до впливу, який таке використання може мати на людей. Ці дебати зосереджуються, зокрема, на алгоритмічній упередженості та потенціалі алгоритмів для незаконної або небажаної дискримінації в рішеннях, яких стосуються алгоритми. Це головне занепокоєння для громадських прав і організацій споживачів, які представляють групи населення, які зазнають неправомірної дискримінації.

Розгляд алгоритмічної дискримінації ставить основні питання щодо сфери застосування законодавства про конфіденційність. По-перше, якою мірою законодавство може або має вирішувати проблеми алгоритмічної упередженості? Дискримінація не є самоочевидною проблемою конфіденційності, оскільки вона представляє широкі соціальні проблеми, які зберігаються навіть без збору та використання особистої інформації та підпадають під дію різних законів про громадянські права. Більше того, надання цих законів для обговорення може фактично відкрити скриньку Пандори через гострі політичні питання, які вони торкаються, і численні комітети Конгресу, які мають юрисдикцію над різними такими питаннями. Незважаючи на це, дискримінація ґрунтується на особистих ознаках, таких як колір шкіри, сексуальна приналежність і національне походження. Використання особистої інформації про ці атрибути, явно або – більш імовірно й менш очевидно – через проксі-сервери, для автоматизованого прийняття рішень, що суперечить інтересам залученої особи, таким чином передбачає інтереси конфіденційності в контролі над тим, як використовується інформація.

Ця шарада згоди зробила очевидним, що помічати і вибирати втрачає сенс. Для багатьох програм штучного інтелекту це стане абсолютно неможливим.

По-друге, захист таких інтересів конфіденційності в контексті AI вимагатиме зміни парадигми регулювання конфіденційності. Більшість існуючих законів про конфіденційність, а також чинні заходи Федеральної торгової комісії проти недобросовісних і оманливих дій ґрунтуються на моделі споживчого вибору на основі «повідомлення та вибору» (також називають «повідомлення та згода»). Ця шарада згоди зробила очевидним, що помічати і вибирати втрачає сенс. Для багатьох застосувань штучного інтелекту (наприклад, інтелектуальні сигнали світлофора та інші датчики, необхідні для підтримки самокерованих автомобілів, як один із яскравих прикладів), це стане абсолютно неможливим.

Хоча майже всі законопроекти на Капітолійському пагорбі все ще певною мірою спираються на модель «повідомлення та вибір», ключові лідери Конгресу, а також зацікавлені сторони конфіденційності висловили бажання змінити цю модель, переклавши тягар захисту конфіденційності особи зі споживачів на компанії. які збирають дані. Замість вибору споживача їхня модель зосереджена на веденні бізнесу, регулюючи обробку компаніями даних – що вони збирають, як вони можуть ними користуватися та ділитися. Розгляд обробки даних, що призводить до будь-якого алгоритмічного розрізнення, може вписатися в цю модель.

Модель, орієнтована на збір і обробку даних, може впливати на AI та алгоритмічну дискримінацію декількома способами:

– Вимоги щодо управління даними, як-от обов'язки чесності чи лояльності, можуть перешкоджати використанню особистої інформації, яке є несприятливим або несправедливим по відношенню до осіб, яких ці дані стосуються.

– Правила прозорості та розкриття даних, а також права осіб на доступ до інформації, що їх стосується, можуть прояснити використання алгоритмічного прийняття рішень.

– Правила керування даними, які передбачають призначення спеціалістів із забезпечення конфіденційності, проведення оцінок впливу на конфіденційність або планування продукту через «конфіденційність за проектом», можуть виявити проблеми, пов'язані з використанням алгоритмів.

– Правила збору та обміну даними можуть зменшити агрегацію даних, що дозволяє робити висновки та прогнози, але можуть передбачати певні компроміси з перевагами великих і різноманітних наборів даних.

На додаток до цих положень загального застосування, які можуть опосередковано впливати на алгоритмічні рішення, ряд пропозицій спеціально стосуються цього предмета.

Варіанти політики AI для захисту конфіденційності

Відповіді на AI, які зараз обговорюються в законодавстві про конфіденційність, мають дві основні форми. Перший спрямований безпосередньо на дискримінацію. Група з 26 громадських правозахисних і споживчих організацій написала спільного листа, в якому закликає заборонити або контролювати використання особистої інформації з дискримінаційним впливом на «кольорових людей, жінок, релігійних меншин, членів ЛГБТК+ спільноти, людей з обмеженими можливостями, осіб, які живуть на l привабливі, іммігранти та інші вразливі верстви населення». Комітет юристів із захисту громадянських прав відповідно до закону та Free Press Action включили цей принцип до типового законодавства, спрямованого на дискримінацію даних, що впливає на економічні можливості, громадські умови або придушення виборців.

Цей підхід до алгоритмічної дискримінації передбачає дебати щодо приватних прав на дії в законодавстві про конфіденційність. Можливість такого індивідуального судового розгляду є ключовим моментом розбіжностей між демократами, які об'єднують інтереси споживачів і конфіденційності, з одного боку, і республіканцями, які об'єднують інтереси бізнесу, з іншого. Перші стверджують, що приватні позови є необхідним примножувачем сили для правоохоронних органів на федеральному рівні та штаті, тоді як другі висловлюють занепокоєння тим, що колективні позови, зокрема, обтяжують бізнес судовими процесами з тривіальних питань. У випадку багатьох видів дискримінації, перелічених у пропозиціях щодо алгоритмічної дискримінації, існуючі федеральні, державні та місцеві закони про громадянські права дозволяють особам подавати позови про дискримінацію. Будь-які федеральні переваги чи обмеження приватних прав на дії у федеральному законодавстві про конфіденційність не повинні порушувати ці закони.

Другий підхід розглядає ризик більш побіжно, із заходами підзвітності, призначеними для виявлення дискримінації під час обробки персональних даних. Численні організації та компанії, а також деякі законодавці пропонують таку підзвітність. Їхні пропозиції мають різні форми:

– Прозорість: це стосується розкриття інформації щодо використання алгоритмічного прийняття рішень. Хоча довгі детальні політики конфіденційності не є корисними для більшості споживачів, вони надають регуляторам та іншим органам захисту конфіденційності орієнтир, за яким можна перевірити обробку даних компанії та притягнути її до відповідальності. Заміна поточної політики конфіденційності на «розкриття конфіденційності», яка вимагає повного опису того, які та як дані збираються, використовуються та захищаються, покращить цю еталонну функцію. У свою чергу, вимога, щоб ці розкриття ідентифікували значні випадки використання особистої інформації для прийняття алгоритмічних рішень, допомогло б спостерігачам і споживачам знати, де слід остерігатися несприятливих результатів.

– Пояснюваність: у той час як прозорість забезпечує завчасне повідомлення про прийняття алгоритмічних рішень, пояснюваність передбачає ретроактивну інформацію про використання алгоритмів у конкретних рішеннях. Це основний підхід, прийнятий у Загальному регламенті захисту даних Європейського Союзу (GDPR). GDPR вимагає, щоб для будь-якого автоматизованого рішення з «юридичними чи подібними значними наслідками», як-от працевлаштування, кредит або страхове покриття, особа, на яку це впливає, мала звернутися до людини, яка може переглянути рішення та пояснити його логіку. Це включає компонент «людини в циклі» та елемент належного процесу, який забезпечує перевірку аномальних або несправедливих результатів.

Почуття справедливості припускає, що такий запобіжний клапан має бути доступним для алгоритмічних рішень, які мають суттєвий вплив на життя людей. Пояснення вимагає (1) ідентифікації алгоритмічних рішень, (2) деконструкції конкретних рішень і (3) встановлення каналу, за яким людина може шукати пояснення. Алгоритми зворотного проектування, засновані на машинному навчанні, можуть бути складними та навіть неможливими, і ця складність зростає, коли машинне навчання стає складнішим. Таким чином, пояснюваність тягне за собою значне нормативне навантаження та обмеження на використання алгоритмічного прийняття рішень і, у цьому світлі, повинна бути зосереджена на його застосуванні, як це зробив ЄС (принаймні в принципі) з його «правовими наслідками або подібними значними наслідками» поріг. У міру того як зростає розуміння порівняльних переваг людських і машинних можливостей, наявність «людини в курсі» рішень, які впливають на життя людей, дає спосіб поєднати потужність машин із людським судженням і співчуттям.

– Оцінка ризику: в Законі про конфіденційність 1974 року оцінки ризиків спочатку були розроблені як «оцінка впливу на конфіденційність» у рамках федерального уряду. З тих пір вони перетворилися на широко використовувані інструменти керування конфіденційністю, щоб заздалегідь оцінювати та зменшувати ризики конфіденційності, і вони вимагаються GDPR для нових технологій або використання даних із високим ризиком.

– Перевірки: перевірки оцінюють практику конфіденційності ретроспективно. Більшість законодавчих пропозицій містять деякі загальні вимоги до підзвітності, щоб переконатися, що компанії дотримуються своїх програм конфіденційності, а деякі включають самоконтроль або аудит третьої сторони. У поєднанні з проактивною оцінкою ризиків аудит результатів алгоритмічного прийняття рішень може допомогти поєднати передбачення з заднім числом; хоча, як і пояснюваність, аудит процедур машинного навчання є складним і все ще розвивається.

Через труднощі передбачення результатів машинного навчання, а також алгоритмічних рішень зворотного проектування, жодна окрема міра не може бути повністю ефективною для уникнення шкідливих ефектів. Таким чином, якщо алгоритмічні рішення є послідовними, має сенс комбінувати заходи для спільної роботи. Попередні заходи, такі як прозорість і оцінка ризиків, у поєднанні з ретроспективними перевірками аудитів і людським переглядом рішень, можуть допомогти виявити та усунути несправедливі результати. Комбінація цих показників може доповнювати один одного і давати більше, ніж сума частин. Оцінка ризиків, прозорість, пояснюваність і перевірки також посилили б існуючі засоби

правового захисту від дискримінації, яка є причиною покарання, шляхом надання документальних доказів, які можна було б використовувати в судовому процесі. Однак не всі алгоритми прийняття рішень є послідовними, тому ці вимоги мають змінюватися відповідно до об'єктивного ризику. Штучний інтелект (AI) вийшов за межі наукової фантастики та став сучасним технологічним рішенням, яке сьогодні використовують багато компаній. Його швидка інтеграція в різні сектори, від охорони здоров'я до фінансів, змінює те, як ми взаємодіємо з даними та приймаємо рішення. У дослідницькому звіті Currents за 2023 рік, в якому опитувалися засновники, керівники та співробітники технологічних компаній, було виявлено, що 49% респондентів використовують інструменти штучного інтелекту та машинного навчання для бізнесу. Однак вагання щодо цих технологій залишаються. На запитання, що заважає їхнім організаціям більше використовувати інструменти AI/ML, 29% згадали про етичні та юридичні проблеми, тоді як 34% відзначили проблеми безпеки.

З цією інновацією виникає нагальна проблема: конфіденційність AI. Оскільки системи AI обробляють величезні обсяги особистої інформації, межа між корисністю та вторгненням стає дедалі розмитішою. Компанії, які використовують бізнес-інструменти штучного інтелекту або розробляють власні, повинні ретельно поєднувати захист конфіденційної інформації з максимальним використанням можливостей технології.

Що таке конфіденційність AI?

Конфіденційність штучного інтелекту – це набір практик і проблем, зосереджених навколо етичного збору, зберігання та використання особистої інформації системами штучного інтелекту. Він відповідає критичній потребі захисту прав особи на дані та збереження конфіденційності, оскільки алгоритми штучного інтелекту обробляють величезну кількість особистих даних і вивчають їх. Забезпечення конфіденційності штучного інтелекту передбачає налагодження балансу між технологічними інноваціями та збереженням особистої конфіденційності в епоху, коли дані є дуже цінним товаром.

Методи збору даних AI та конфіденційність

Системи штучного інтелекту покладаються на велику кількість даних, щоб покращити свої алгоритми та результати, використовуючи ряд методів збору, які можуть становити значні ризики для конфіденційності. Методи, які використовуються для збору цих даних, часто невидимі для осіб (наприклад, клієнтів), від яких збираються дані, що може призвести до порушень конфіденційності, які важко виявити або контролювати.

Ось кілька методів збору даних штучного інтелекту, які мають наслідки для конфіденційності:

– Веб-скрейпінг. AI може накопичувати величезні обсяги інформації, автоматично збираючи дані з веб-сайтів. Хоча деякі з цих даних є загальнодоступними, веб-збирання також може отримувати особисті дані, потенційно без згоди користувача.

– Біометричні дані. Системи штучного інтелекту, які використовують розпізнавання обличчя, зйомку відбитків пальців та інші біометричні технології, можуть втручатися в особисту конфіденційність, збираючи конфіденційні дані, які є унікальними для окремих людей і, якщо їх зламано, незамінні.

– Пристрої IoT. Пристрої, підключені до Інтернету речей (IoT), надають системам AI дані в реальному часі з наших домівок, робочих місць і громадських місць. Ці дані можуть розкривати інтимні подробиці нашого повсякденного життя, створюючи безперервний потік інформації про наші звички та поведінку.

– Моніторинг соціальних медіа. Алгоритми AI можуть аналізувати активність у соціальних мережах, фіксуючи демографічну інформацію, уподобання та навіть емоційний стан, часто без явного відома або згоди користувача.

Наслідки цих методів для конфіденційності є далекосяжними. Вони можуть призвести до несанкціонованого стеження, крадіжки особистих даних і втрати анонімності. Оскільки технології штучного інтелекту стають все більш інтегрованими в повсякденне життя, забезпечення того, щоб збір даних був прозорим і безпечним і щоб люди зберегли контроль над своєю особистою інформацією, стає все більш критичним.

Унікальні виклики конфіденційності AI

Згідно з даними Crunchbase, у 2023 році понад 25% інвестицій в американські стартапи було спрямовано на компанії, що спеціалізуються на AI. Ця хвиля штучного інтелекту розкрила безпрецедентні можливості в обробці даних, аналізі та прогнозному моделюванні. Однак штучний інтелект створює складні та багатогранні проблеми з конфіденційністю, які відрізняються від тих, які створює традиційна обробка даних:

– Обсяг і різноманітність даних. Системи штучного інтелекту можуть переробляти та аналізувати експоненціально більше даних, ніж традиційні системи, що підвищує ризик розкриття особистих даних.

– Прогностична аналітика. Завдяки розпізнаванню образів і прогнозному моделюванню штучний інтелект може визначати особисту поведінку та вподобання, часто без відома чи згоди людини.

– Непрозоре прийняття рішень. Алгоритми штучного інтелекту можуть приймати рішення, що впливають на життя людей, без прозорих міркувань, що ускладнює відстеження або заперечення вторгнень у конфіденційність.

– Безпека даних. Великі набори даних, необхідні AI для ефективного функціонування, є привабливими цілями для кіберзагроз, що збільшує ризик злому, який може поставити під загрозу особисту конфіденційність.

– Вбудоване упередження. Без ретельного нагляду штучний інтелект може зберегти існуючі упередження в даних, які він передає, що призведе до дискримінаційних результатів і порушень конфіденційності.

Ці проблеми підкреслюють необхідність надійних заходів захисту конфіденційності в AI. Збалансування переваг штучного інтелекту з правом на конфіденційність вимагає ретельного проектування, впровадження та управління, щоб запобігти неправомірному використанню персональних даних.

Ключові проблеми конфіденційності AI для компаній

Оскільки компанії все більше інтегрують штучний інтелект у свою діяльність або створюють системи штучного інтелекту для використання своїми клієнтами, вони стикаються з багатьма проблемами конфіденційності, які слід вирішувати завчасно. Ці занепокоєння формують довіру клієнтів і мають значні юридичні та етичні наслідки, до яких компанії повинні обережно орієнтуватися.

Відсутність прозорості в алгоритмах AI

Природа «чорної скриньки» систем AI означає, що їхні процеси прийняття рішень часто непрозорі. Ця невідомість викликає занепокоєння у компаній, користувачів і регуляторів, оскільки вони часто не можуть бачити або розуміти, як алгоритми AI приходять до певних висновків або дій. Відсутність алгоритмічної прозорості також може приховати упередження або недоліки в системах AI, що призведе до результатів, які можуть ненавмисно завдати шкоди певним групам або окремим особам. Без такої прозорості підприємства ризикують підірвати довіру клієнтів і потенційно порушити нормативні вимоги.

Несанкціоноване використання персональних даних

Включення персональних даних у моделі штучного інтелекту без явної згоди створює значні ризики, зокрема юридичні наслідки відповідно до законів про захист даних, як-от GDPR, і потенційні порушення етичних стандартів. Несанкціоноване використання цих даних може призвести до порушення конфіденційності, значних штрафів і шкоди репутації компанії. З етичної точки зору такі дії ставлять під сумнів цілісність бізнесу та підривають довіру клієнтів.

Дискримінаційні результати застосування AI

Упередженість AI, що виникає через спотворені навчальні дані або помилкові алгоритми, може призвести до дискримінаційних результатів. Ці упередження можуть увічнити і навіть посилити існуючу соціальну нерівність, впливаючи на групи, засновані на расі, статі чи соціально-економічному статусі. Наслідки для конфіденційності є серйозними,

оскільки особи можуть бути несправедливо створені та піддані необґрунтованому контролю або виключенню. Для компаній це підриває чесну практику та може призвести до втрати довіри та юридичних наслідків.

Проблеми з авторським правом та інтелектуальною власністю з AI

Системи AI часто вимагають великих наборів даних для навчання, що може призвести до використання захищених авторським правом матеріалів без дозволу. Це порушує закони про авторське право та викликає занепокоєння щодо конфіденційності, коли вміст містить особисті дані. Компанії повинні обережно орієнтуватися в цих викликах, щоб уникнути судових розглядів і потенційних наслідків використання інтелектуальної власності третіх сторін без згоди.

Використання біометричної інформації

Використання біометричних даних у системах штучного інтелекту, таких як технології розпізнавання обличчя, викликає серйозні занепокоєння щодо конфіденційності. Біометрична інформація є особливо чутливою, оскільки вона за своєю суттю особиста і, у більшості випадків, незмінна. Несанкціонований збір, зберігання або використання цих даних може призвести до значного порушення конфіденційності та можливого зловживання. Підприємства, які використовують біометричний штучний інтелект, повинні забезпечити надійний захист конфіденційності, щоб зберегти довіру користувачів і дотримуватися суворих правових стандартів щодо біометричних даних.

Стратегії пом’якшення ризиків конфіденційності AI

Дослідження Deloitte у 2023 році показує, що 56% учасників опитування або не знають, або не впевнені щодо існування етичних принципів щодо генеративного використання AI в їхніх організаціях. Щоб захиститися від інвазивного потенціалу штучного інтелекту, компанії повинні активно приймати стратегії, які гарантують, що конфіденційність не буде порушена. Пом’якшення ризиків конфіденційності штучного інтелекту передбачає поєднання технічних рішень, етичних принципів і надійної політики управління даними.

Вбудуйте конфіденційність у дизайн AI

Щоб зменшити ризики конфіденційності штучного інтелекту, інтегруйте питання конфіденційності на початкових етапах розробки системи AI. Це передбачає прийняття принципів «конфіденційності за проектом», гарантуючи, що захист даних не є запізнілою думкою, а основним компонентом технології, яку створює ваша команда. Завдяки цьому моделі штучного інтелекту побудовані з необхідними засобами захисту, щоб обмежити непотрібний доступ до даних і забезпечити надійний захист із самого початку. Шифрування має бути стандартним для захисту даних у стані спокою та під час передачі, тоді як регулярні перевірки можуть забезпечити постійне дотримання політики конфіденційності.

Анонімізація та зведення даних

Використання методів анонімізації може захистити індивідуальні особи шляхом видалення ідентифікаційної інформації з наборів даних, які використовують системи AI. Цей процес передбачає зміну, шифрування або видалення особистих ідентифікаторів, що гарантує, що дані неможливо відстежити до особи. У поєднанні з анонімізацією агрегація даних об’єднує окремі точки даних у більші набори даних, які можна аналізувати без розкриття особистих даних. Ці стратегії зменшують ризик порушення конфіденційності, запобігаючи асоціації даних із конкретними особами під час аналізу AI.

Обмежте час зберігання даних

Впровадження суворої політики збереження даних мінімізує ризики конфіденційності, пов’язані зі штучним інтелектом. Встановлення чітких обмежень на тривалість зберігання даних запобігає непотрібному тривалому накопиченню особистої інформації, зменшуючи ймовірність її розголошення під час порушення. Ці політики змушують організації регулярно переглядати та видаляти застарілі або нерелевантні дані, оптимізуючи бази даних і мінімізуючи кількість даних, які піддаються ризику.

Збільште прозорість і контроль користувача

Підвищення прозорості в системах штучного інтелекту створює довіру та відповідальність користувачів. Компанії повинні повідомляти, які типи даних збираються, як алгоритми AI їх обробляють і для яких цілей. Надання користувачам контролю над їхніми даними, як-от можливість переглядати, редагувати чи видаляти їхню інформацію, розширює можливості людей і створює відчуття волі над їхнім цифровим слідом. Вони відповідають етичним стандартам і забезпечують відповідність нормам захисту даних, що розвиваються, які все більше вимагають згоди користувачів і управління.

Зрозумійте вплив нормативних актів

Розуміння наслідків GDPR та подібних нормативних актів має важливе значення для зменшення ризиків конфіденційності AI, оскільки ці закони встановлюють суворі стандарти захисту даних і надають особам значний контроль над своєю особистою інформацією. Ці правила зобов'язують організації бути прозорими щодо своєї діяльності з обробки AI та забезпечувати дотримання прав окремих осіб на дані, включаючи право на пояснення алгоритмічних рішень.

Компанії повинні вживати заходів, які гарантують точність, справедливість і підзвітність їхніх систем штучного інтелекту, особливо коли рішення мають юридичний або значний вплив на окремих осіб. Недотримання таких нормативних стандартів може призвести до значних штрафів.

Ось кілька правил і вказівок, на які варто звернути увагу:

– Закон ЄС про штучний інтелект.

– Запропонований Канадою законопроект C-27 (включає Закон про штучний інтелект і дані).

– Рекомендації Федеральної торгової комісії США (FTC) щодо використання штучного інтелекту та алгоритмів.

Розвивайте культуру етичного використання AI

Щоб зменшити ризики конфіденційності AI, установіть етичні принципи використання AI, які надають пріоритет захисту даних і дотриманню прав інтелектуальної власності. Організації повинні проводити регулярне навчання, щоб гарантувати, що всі співробітники розуміють ці вказівки та важливість їх дотримання у своїй повсякденній роботі з технологіями AI. Майте прозорі політики, які регулюють збір, зберігання та використання надзвичайно особистої та конфіденційної інформації. Нарешті, сприяння створенню середовища, де етичні проблеми можна відкрито обговорювати та вирішувати, допоможе підтримувати пильну позицію щодо потенційних порушень конфіденційності.

Майбутнє залежатиме від спільного підходу, коли безперервний діалог між технологами, компаніями, регуляторами та громадськістю формує розробку штучного інтелекту для захисту прав на конфіденційність, одночасно сприяючи технологічному прогресу.

Розробка структурної схеми

Система аналізу криптографічних алгоритмів складається із двох підсистем: підсистеми криптографічного захисту інформації з використанням представників різних класів шифрів і підсистеми криптоаналізу. Структурна схема системи аналізу криптографічних алгоритмів представлена на рисунку 1.

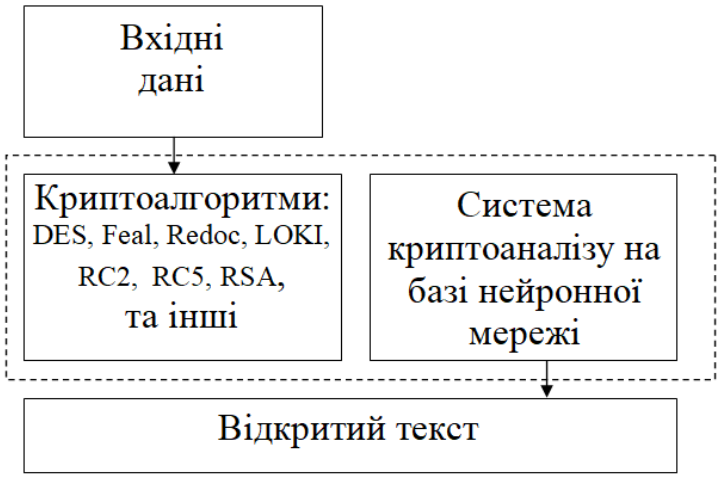

Рисунок 1 – Структурна схема системи криптоаналізу

З нього ми бачимо, що існують чотири структурні блоки які взаємодіють між собою:

– вхідні данні;

– криптоалгоритми, які потребують криптоаналізу;

– власне сама система криптоаналізу, яка складається з використання нейромережі на основі шарів Кохонена та Гроссберга;

– відкритий текст, у якому даються дані про можливість криптоаналізу, того або іншого криптоалгоритму.

Користувач вибирає файл який буде вхідним текстом та відкриває його у програмі.

Потім обирає криптоалгоритм, який хоче протестувати.

Файл, обраний користувачем, шифрується вибраним крптоалгоритмом.

Зашифрований файл підлягає криптоаналізу на базі нейронної мережі та на основі його результатів визначається стійкійкість вибраного алгоритму шифрування.

Висновки. У статті наведені теоретичні узагальнення й рішення наукового завдання дослідження методів визначення рівня стійкості сервісів забезпечення конфіденційності на основі методів AI. Рішення даного завдання полягало у вирішенні наступних задач: Був проведений огляд існуючих систем визначення рівня стійкості сервісів забезпечення конфіденційності на основі методів AI. Досліджена система визначення рівня стійкості сервісів забезпечення конфіденційності на основі методів AI. На основі отриманих результатів досліджень створена програмна реалізація системи визначення рівня стійкості сервісів забезпечення конфіденційності на основі методів AI. Розроблені під час виконання випускної кваліфікаційної роботи за другим (магістерським) рівнем вищої освіти алгоритми дозволяють успішно вирішувати завдання визначення рівня стійкості сервісів забезпечення конфіденційності на основі методів AI. Проведено аналіз предметної галузі в ході якого були виявлені об'єкти, взаємодія яких носить істотний характер для функціональної діяльності предметної галузі, і їхні основні характеристики; побудована алгоритм і вибраний середовище розробки.

Список літератури

1. Smirnov O., Smirnova T., Anas M. Al-Oraiqat, Drieiev O., Polishchuk L., Sheroz Khan, Yassin M. Y. Hasan, Aladdein M. Amro, Hazim S. AlRawashdeh «Method for Determining Treated Metal Surface Quality Using Computer Vision Technology». Sensors (Basel, Switzerland) Volume 22, Issue 16, 6223, 2022. (Scopus).

2. Smirnov, O., Lakhno, V., Akhmetov, B., Chubaievskyi, V., Khorolska, K., Bebeshko, B. «Selection of a Rational Composition of İnformation Protection Means Using a Genetic Algorithm». In: Rajakumar, G., Du, KL., Vuppalapati, C., Beligiannis, G.N. (eds) Intelligent Communication Technologies and Virtual Mobile Networks. Lecture Notes on Data Engineering and Communications Technologies, vol 131. 2023. Springer, Singapore. pp. 21-34. (Scopus).

3. Smirnov O.A., Al-Oraiqat A.M., Ulichev O.S., Meleshko Ye.V., Al-Rawashdeh H.S., Polishchuk L.I. «Modeling strategies for information influence dissemination in social networks». Journal of Ambient Intelligence and Humanized Computing Volume 13, Issue 5. Springer, Cham. 2022, pp. 2463-2477. (Scopus).
4. Smirnov O., Kuznetsov A., Kryvinska N., Kiian A., Kuznetsova K. «Full Non-Binary Constant-Weight Codes». SN Computer Science, Vol 2, 337, 2021. https://doi.org/10.1007/s42979-021-00739-w (Scopus).
5. Smirnov O., Kovalenko O., Kovalenko A., Kavun S. «Quantitative Risk Assessment Method Development in the Context of the SDLC-model». 2021 IEEE 8th International Conference on Problems of Infocommunications, Science and Technology (PIC S&T), 2021, pp. 203-208, doi: 10.1109/PICST54195.2021.9772143 (Scopus).
6. Smirnov O., Neskorodieva T., Fedorov E., Rymar P. «Neural Network Modeling Method of Transformations Data of Audit Production with Returnable Waste». CEUR Workshop Proceedings Volume 3101, 2021, Pages 192-207. (Scopus).
7. Smirnov O., Kuznetsov A., Kiian A., Kuznetsova K. «Data hiding scheme based on spread sequence addressing». CEUR Workshop Proceedings Volume 2805, 2020, Pages 44-58. (Scopus).
8. Smirnov, O., Kuznetsov, A., Kolovanova, I., Kuznetsova, T., «Noise immunity of the algebraic geometric codes». International Journal of Computing; 2019, Volume 18, Issue 4 – Research Institute for Intelligent Computer Systems – 2019. – P. 393-407. (Scopus).
9. Smirnov, O., Ulichev, O., Meleshko, Y., Khokh, V., Goncharenko, I. «Method of Choosing Objects for Informational Influence in Social Networks during Information Campaign Based on the Analytic Hierarchy Process». CEUR Workshop Proceedings, Vol 2588, P. 215-227, 2019. (Scopus).
10. Smirnov, O., Krasnobayev, V., Yanko, A., Kuznetsova, T. «Methods of nulling numbers in the system of residual classes». CEUR Workshop Proceedings, Vol 2588, P. 90-106, 2019. (Scopus).
11. Smirnov, O., Kuznetsov, A., Kovalchuk, D., Pastukhov, M., Kuznetsova, K., Prokopovych-Tkachenko, D., «Discrete Signals with Special Correlation Properties», CEUR Workshop Proceedings Volume 2353, CEUR Workshop Proceedings 2019, Pages 618-629. (Scopus).
12. Smirnov, O., Kuznetsov, A., Kiian, A., Kuznetsova, K., Ivko, T., Prokopovych-Tkachenko, D., «Soft Decoding Based on Ordered Subsets of Verification Equations of Turbo-Productive Codes», CEUR Workshop Proceedings Volume 2353, CEUR Workshop Proceedings 2019, Pages 873-884. (Scopus).
13. Smirnov, O., Kuznetsov, A., Prokopovych-Tkachenko, D. «Hiding Data in Images Using a Pseudo-Random Sequence». ISCI'2020: Information Security in Critical Infrastructures. Collective monograph. Edited by Ivan D. Gorbenko, Victor A. Krasnobayev and Alexandr A. Kuznetsov. ASC Academic Publishing, USA, 2020. pp. 46-59. – ISBN: 978-1-7362833-0-1 (Hardback), ISBN: 978-1-7362833-1-8 (Ebook).
14. Smirnov, O., Kuznetsov, A., Shekhanin, K., Chepurko, I. Detecting Hidden Information in FAT. Монографія: In.: ISCI'2019: Information Security in Critical Infrastructures. Collective monograph. Edited by Ivan D. Gorbenko and Alexandr A. Kuznetsov, ASC Academic Publishing, USA, 2019, pp. 412-429. – ISBN: 978-0-9989826-8-7 (Hardback), ISBN: 978-0-9989826-9-4 (Ebook).
15. Smirnov, O., Kuznetsov, A., Kuznetsova., K. Synthesis of Discrete Signals with Improved Correlation Properties. Монографія: In.: ISCI'2019: Information Security in Critical Infrastructures. Collective monograph. Edited by Ivan D. Gorbenko and Alexandr A. Kuznetsov, ASC Academic Publishing, USA, 2019, pp. 281-299. – ISBN: 978-0-9989826-8-7 (Hardback), ISBN: 978-0-9989826-9-4 (Ebook).
16. О.А. Смірнов, П.С. Усік, «дослідження перспектив використання технологічних рішень в мережах 5g» у Кібербезпека та інформаційні технології: монографія. – Х. : ТОВ «ДІСА ПЛЮС», 2020.С. 122-135.
17. Смірнов О.А., Дрєєва Г.М., «Метод генерування фрактального трафіку за допомогою моделі генератора на графі» у Інформаційна безпека та інформаційні технології: монографія / за заг. ред. В. С. Пономаренка. – Х. : Вид. Рожко С.Г. 2019. С. 123-139.
18. Smirnov, O., Drieieva, H., Drieiev, O., Simakhin, V., Bondar, S., Odarchenko, R. «Managing multifractal properties of the binary sequence generated with the Markov chains», CEUR Workshop Proceedings Volume 2608, 2020, Pages 633-645. (Scopus).
19. Smirnov, O., Kuznetsov, A., Gorbacheva, L., Babenko, V., «Hiding data in images using a pseudo-random sequence», CEUR Workshop Proceedings Volume 2608, 2020, Pages 646-660., (Scopus).
20. Smirnov, O., Shekhanin, K., Kuznetsov, A., Krasnobayev, V. «Detecting Hidden Information in FAT». International Journal of Computer Network and Information Security (IJCNIS). Vol. 12, No. 3, 2020. PP.33-43. (Scopus).