

обчислювальних ресурсів. Мета полягає в тому, щоб зробити системи швидшими, дешевшими та надійнішими. Використання кастомізованих локальних моделей дозволить використовувати ШІ більш ефективно в різних галузях, таких як охорона здоров'я, фінанси та виробництво. Використання потужних віртуальних агентів включає розробку нових методів обробки природної мови, машинного навчання та штучного інтелекту. Мета полягає в тому, щоб зробити віртуальних агентів більш корисними для взаємодії.

Список використаних джерел

1. «The most important AI trends in 2024» [Електронний ресурс]. – Режим доступу: <https://www.ibm.com/blog/artificial-intelligence-trends/>
 2. «Discover multimodal AI, one of the most promising trends in generative AI» [Електронний ресурс]. – Режим доступу: <https://www.datacamp.com/blog/what-is-multimodal-ai>
 3. Q&A: UW researcher discusses just how much energy ChatGPT uses,” [Електронний ресурс]. University of Washington, 27 July 2023
 4. «What is reinforcement learning from human feedback (RLHF)?» [Електронний ресурс]. – Режим доступу: <https://www.ibm.com/topics/rlhf>
- Науковий керівник:** Родіонова О. В., ст. викладач.

Калінін Д. В.

аспірант

Національний технічний університет

«Харківський політехнічний інститут»

denis.kalinin.univ@gmail.com

АНОНІМІЗАЦІЯ МЕДИЧНИХ ДАНИХ НА ОСНОВІ ГЕНЕРАТИВНИХ НЕЙРОННИХ МЕРЕЖ

Наразі проблема анонімізації (деідентифікації) даних стає дедалі актуальнішою у сфері цифрової інформації через все більш активне використання особистих даних для статистичного аналізу та машинного навчання. Анонімізація – це метод обробки даних, при якому ідентифікаційна інформація видаляється або змінюється. Особливе місце ця проблема займає у медичній сфері, де анонімність даних регулюється певними правовими актами та є обов'язковою за необхідністю поширення або публікації усіх даних або деякої їх частини [1].

Особливістю проблеми анонімізації є її комплексність, яка полягає у різноманітні залежних факторів: тип та об'єм даних, однорідність даних та присутність аномалій, тип потенційного зловмисника та рівень його знань про дані тощо. Таким чином, немає єдиної універсальної формули, яка б забезпечувала стабільно високу якість анонімізації для будь-яких наборів даних. Але при цьому виділяють певні базові моделі (анонімізації), які є своєрідним фундаментом при вирішенні проблеми анонімізації: *k*-anonymity, *l*-diversity та *t*-closeness [2].

Насамперед треба визначити, що у контексті персональних даних виділяють наступні категорії атрибутів інформації: прямі ідентифікатори (direct identifier, DID), квазі-ідентифікатори (quasi-identifier, QID) та чутливі атрибути (sensitive attribute, SA). Виходячи з цього, *k*-anonymity гарантує, що як мінімум *k* осіб мають однаковий набір атрибутів у межах своєї групи (класу еквівалентності). Вирішення проблеми неоднорідності даних у свою чергу забезпечує модель *l*-diversity, яка гарантує, що кожен клас еквівалентності має *l* унікальних варіантів значень для SA. І, нарешті, *t*-closeness гарантує, що відстань між розподіленням значень SA в кожному класі еквівалентності та відповідним розподіленням в оригінальному класі не перевищує задане значення *t*.

На відміну від зазначених підходів, які оперують виключно вхідними даними, використання генеративних нейронних мереж дає змогу генерувати нові дані, які ще називають синтетичними [3]. Такі дані можуть бути використані для об'єднання з вхідними даними або навіть для їх повноцінної заміни. Звісно, що якість таких синтетичних даних залежить від об'єму, різноманіття та подібності наборів даних, на яких навчається штучний інтелект. Застосування якісних штучних даних дозволяє вирішити ряд проблем, замінюючи або доповнюючи деякі з зазначених технік, зменшуючи при цьому ризики реідентифікації конкретних осіб та покращуючи якість вихідних даних для подальшого аналізу.

Наразі розробляється інформаційна система, що дозволяє ефективно анонімізувати медичні дані, використовуючи для цього генеративні нейронні мережі. Дана система буде враховувати задані критерії корисності та рівня захисту даних.

Список використаних джерел

1. Anonymizing Health Data / Khaled El Emam, Luk Arbuckle, O'Really Media, Inc, 2013.
2. A Review of Anonymization for Healthcare Data / Iyiola E. Olatunji, Jens Rauch, Matthias Katzensteiner, Megha Khosla, Published by Oxford University Press, 2021.
3. Practical Synthetic Data Generation / Khaled El Emam, Lucy Mosquera, Richard Hoptroff, O'Really Media, Inc, 2020.

Науковий керівник: Северин В. П., д.т.н., професор.

Ковальов К.Є.

*старший науковий співробітник
Український науково-дослідний інститут
спеціальної техніки та судових експертиз
Служби безпеки України
icte@ssu.gov.ua*

ВИКОРИСТАННЯ ШТУЧНОГО ІНТЕЛЕКТУ В РОСІЙСЬКО-УКРАЇНСЬКІЙ ВІЙНІ

Війна в Україні стала поштовхом для стрімкого розвитку та удосконалення військових технологій. Наразі велика науково-технічна спільнота нашої країни активно розробляє, випробовує та втілює в життя новітні системи озброєння та військової техніки.

Однією із головних інноваційних технологій є штучний інтелект (далі – ШІ), який по суті являється сукупністю різних технологічних процесів, що об'єднують великі об'єми інформації, алгоритми та обчислювальні потужності.

Усвідомлюючи, що військово-промислова галузь росії значно перебільшує наші можливості, Україна з метою отримання військової переваги над противником вибрала стратегію розвитку та впровадження технології ШІ в різні системи життєдіяльності людини.

Перед нашими інженерами і науковцями стоїть не просте завдання з розробки та впровадження технології ШІ в різні безпекові системи та озброєння. Швидкий розвиток цього напрямку в безпековій сфері зможе дозволити нашим силам оборони суттєво стримати зовнішню агресію противника, а згодом і перемогти його.

Втілення ШІ в сучасні системи управління та координації військ дозволяє нам в режимі реального часу фіксувати місцезнаходження об'єктів противника, їх переміщення на полі бою та в тилу, нищити ворожу техніку та особовий склад.