Міністерство освіти і науки України
Харківський національний університет імені В. Н. Каразіна
Навчально-науковий інститут комп'ютерних наук та штучного інтелекту

Кафедра кібербезпеки інформаційних систем, мереж і технологій

До захисту допущено
Кафедрою КІСМіТ протокол №____ від «___» грудня 2025 р.

завідувач кафедри _____ Марина ЄСІНА
(підпис) *(ім'я, прізвище)*
«___» грудня 2025 р.

Кваліфікаційна робота
здобувача другого (магістерського) рівня вищої освіти

«Дослідження та аналіз методів забезпечення диференційної конфіденційності у великих наборах даних»

(назва роботи)

Спеціальність (спеціалізація) 125 «Кібербезпека та захист інформації»

Освітня програма «Безпека інформаційних і комунікаційних систем»

Виконавець _____ Олена КОБИЛЯНСЬКА
(підпис) *(ім'я, прізвище)*

Науковий керівник _____ Марина ЄСІНА
(підпис) *(ім'я, прізвище)*

Харків – 2025

РЕФЕРАТ

Пояснювальна записка до кваліфікаційної роботи магістра містить 72 сторінки, 15 рисунків, 9 таблиць, 16 формул та 35 посилань на джерела.

Мета роботи полягає у дослідженні та порівняльному аналізі математичних моделей диференційної конфіденційності, практичних методів її впровадження в задачах обробки великих даних та машинного навчання, а також оцінюванні відповідності сучасним міжнародним стандартам і регуляторним вимогам щодо захисту персональних даних.

Об’єкт дослідження – процес забезпечення приватності персональних даних у великих наборах даних.

Предмет дослідження – математичні моделі, алгоритми, інструменти та регуляторні підходи, що застосовуються для реалізації диференційної конфіденційності в інформаційних системах, зокрема в аналітичних платформах та моделях машинного навчання.

Основними методами дослідження є математичний аналіз, моделювання, експериментальні обчислення, порівняльний аналіз моделей, а також дослідження нормативно-правових документів у сфері захисту даних.

У роботі досліджено: теоретичну основу диференційної конфіденційності, математичні властивості, практичну інтеграцію шуму у моделі машинного навчання, а також регуляторні вимоги й стандарти, що визначають коректність та доцільність використання даної технології.

Результати роботи можуть бути використані в наукових дослідженнях, у розробці безпечних аналітичних систем, а також у практичних проєктах, пов’язаних з обробкою великих даних, машинним навчанням та впровадженням технологій забезпечення приватності.

Ключові слова: ДИФЕРЕНЦІЙНА КОНФІДЕНЦІЙНІСТЬ, ПРИВАТНІСТЬ ДАНИХ, ВЕЛИКІ ДАНІ, МАШИННЕ НАВЧАННЯ, ШУМ ЛАПЛАСА, ШУМ ГАУСА, GDPR, NIST SP 800-226, БЮДЖЕТ ПРИВАТНОСТІ, ПОТОКОВА ОБРОБКА, DP-SGD, ДЕАНОНІМІЗАЦІЯ.

ABSTRACT

The master's thesis report contains 72 pages, 15 figures, 9 tables, 16 formulas and 28 references.

The purpose of this work is to investigate and compare mathematical models of differential privacy, analyze practical methods of its implementation in big data processing and machine learning tasks, and evaluate the compliance of these methods with international standards and regulatory requirements for personal data protection.

The object of the study is the process of ensuring personal data privacy in large datasets.

The subject of the study is the mathematical models, algorithms, tools, and regulatory frameworks used to implement differential privacy in information systems, particularly in analytical platforms and machine learning models.

The main research methods include mathematical analysis, modeling, experimental evaluation, comparative analysis of algorithms, and examination of international legal documents related to privacy and data protection.

The thesis investigates the theoretical foundations of differential privacy, mathematical properties of privacy mechanisms, practical integration of noise into machine learning models, and regulatory requirements defining the correctness and applicability of differential privacy.

The obtained results can be used in scientific research, in the design of secure analytical systems, and in practical projects involving big data processing, machine learning, and privacy-preserving technologies.

Keywords: DIFFERENTIAL PRIVACY, DATA PRIVACY, BIG DATA, MACHINE LEARNING, LAPLACE NOISE, GAUSSIAN NOISE, GDPR, NIST SP 800-226, PRIVACY BUDGET, STREAMING ANALYTICS, DP-SGD, DE-ANONYMIZATION.

ЗМІСТ

ПЕРЕЛІК СКОРОЧЕНЬ

AOL	– America Online, Американська онлайн-служба
API	– Application Programming Interface, Програмний інтерфейс застосунків
AUC	– Area Under Curve, Площа під кривою
BD	– Big Data, Великі дані
CPU	– Central Processing Unit, Центральний процесор
DP/ ДК	– Differential Privacy, Диференційна конфіденційність
DP-SGD	– Differentially Private Stochastic Gradient Descent, Стохастичний градієнтний спуск з диференційною конфіденційністю
GDP / Global DP	– Global Differential Privacy, Глобальна диференційна конфіденційність
GDPR	– General Data Protection Regulation, Загальний регламент захисту даних (ЄС)
GPU	– Graphics Processing Unit, Графічний процесор
ID	– Identifier, Ідентифікатор
LDP	– Local Differential Privacy, Локальна диференційна конфіденційність
ML	– Machine Learning, Машинне навчання
NLP	– Natural Language Processing, Обробка природної мови
NVMe	– Non-Volatile Memory Express, Протокол високошвидкісного доступу до твердотільних накопичувачів
ROC-AUC	– Receiver Operating Characteristic — Area Under Curve, Характеристика робочих властивостей приймача — площа під кривою
SGD	– Stochastic Gradient Descent, Стохастичний градієнтний спуск

SSD	– Solid State Drive, Твердотільний накопичувач
TFLOPS	– Trillion Floating Point Operations per Second, Трильйон операцій із плаваючою комою за секунду
TOPS	– Trillion Operations per Second, Трильйон операцій за секунду
UCI	– University of California Irvine, Університет Каліфорнії в Ірвайні
БП	– Бюджет приватності
ДП	– Диференційна приватність
США	– Сполучені Штати Америки

ВСТУП

У сучасному світі стрімка цифровізація, масове впровадження автоматизованих систем обробки даних та поширення алгоритмів штучного інтелекту створюють одночасно безпрецедентні можливості й серйозні ризики. Обсяг даних, що генеруються щодня, досягає масштабів, які ще декілька років тому вважалися недосяжними, а моделі машинного навчання дедалі глибше інтегруються в медицину, державне управління, бізнес-аналітику, критичну інфраструктуру та соціальні сервіси. У цих умовах питання збереження приватності особистої інформації та дотримання етичних принципів обробки даних стають фундаментальними для забезпечення довіри, безпеки та відповідності міжнародним стандартам. Саме тому диференційна конфіденційність (ДК)– як математично доведена модель контролю інформаційного витоку – перетворилася на один із ключових підходів до побудови сучасних безпечних аналітичних систем.

Проблематика дослідження полягає у тому, що традиційні підходи до анонімізації даних більше не забезпечують реального захисту в умовах Big Data. Декілька гучних інцидентів – від деанонімізації пошукових логів Американської онлайн-служби[4] до розкриття анонімізованих даних Netflix Prize [5] – довели, що навіть видалення ідентифікаторів не унеможливлює відновлення особистостей через квазіідентифікатори та поєднання різних наборів даних. У практичних системах машинного навчання витік приватної інформації можливий через моделі, що «запам’ятовують» окремі записи, через статистичні відповіді на запити або через неконтрольоване накопичення помилки при багаторазових операціях. З іншого боку, занадто сильне зашумлення або необережне застосування механізмів ДК може суттєво погіршувати точність моделей, роблячи їх непридатними для реальних завдань. Таким чином, постає подвійна проблема: як забезпечити захист даних із формальними гарантіями й водночас зберегти можливість отримання якісних аналітичних результатів, особливо у великих і високовимірних наборах даних.

Актуальність теми зумовлена тим, що впровадження диференційно-приватних технологій уже стало міжнародною вимогою та стандартом де-факто. Застосування диференційної конфіденційності є обов'язковим у статистичних системах США, рекомендованим NIST, узгодженим із вимогами GDPR щодо недопущення прямої чи опосередкованої ідентифікації, а також активно впроваджується технологічними компаніями – Apple, Google, Meta – у продуктах із мільйонною аудиторією. Крім того, методи ДК дедалі частіше використовуються у медичних інформаційних системах, у дослідженнях, що працюють із персональними даними, та під час розробки безпечних систем машинного навчання. Водночас в Україні, де активно розвивається цифрова інфраструктура й відбувається інтеграція у європейський цифровий простір, питання формалізованого захисту приватності набуває особливої ваги. Потреба у методах, здатних забезпечити об'єктивно вимірюваний рівень безпеки, робить диференційну конфіденційність критично важливою для побудови довірчих аналітичних платформ та державних сервісів, що обробляють великі обсяги чутливих даних.

Метою дослідження є комплексне вивчення диференційної конфіденційності як сучасної математичної моделі захисту даних, аналіз її можливостей та обмежень у контексті великих даних і машинного навчання, а також визначення ступеня відповідності вимогам міжнародних стандартів у сфері інформаційної безпеки та приватності. Задля цього робота розглядає проблему з трьох взаємопов'язаних позицій: математико-теоретичної, що включає формалізацію механізмів ДК, моделей сусідства, композиції та бюджету приватності; практичної, у межах якої здійснюється експериментальне впровадження ДК у моделі машинного навчання на реальному великому наборі даних із оцінкою впливу параметра ε на точність; а також регуляторної, що аналізує вимоги GDPR, ISO/IEC 20889, NIST SP 800-226 та інших міжнародних стандартів щодо приватності, визначає місце ДК у сучасних політиках захисту та її роль у формуванні безпечних цифрових екосистем.

1 ПРИВАТНІСТЬ ДАНИХ У ДОБУ BIG DATA І КОНЦЕПЦІЯ ДИФЕРЕНЦІЙНОЇ КОНФІДЕНЦІЙНОСТІ

За даними IDC, обсяг цифрових даних у світі перевищив 120 зета байтів у 2023 році, а кількість нових цифрових подій зростає експоненційно [1]. Сьогоднішнє інформаційне середовище визначається не просто зростанням обсягів даних, а стійким переходом до парадигми Big Data. Ще на початку 2000-х Даг Лейні запропонував відому модель «3V» – обсяг (Volume), швидкість (Velocity) та різноманіття (Variety) – щоб описати новий клас даних, який якісно відрізняється від традиційних, структурованих корпоративних баз [2]. Цифровізація бізнесу, розвиток Інтернет-сервісів, мобільних застосунків, соціальних мереж та датчиків Інтернету речей призвели до того, що майже кожна дія людини залишає цифровий слід, а саме покупки, пошукові запити, геолокація, активність у соцмережах, показники зі «смарт»-пристроїв.

Усі ці сліди об’єднуються у гігантські, багатовимірні масиви, які постійно оновлюються і стають базою для аналітики, машинного навчання, прогнозування поведінки та прийняття рішень у реальному часі.

Водночас саме ці характеристики Big Data – масштаб, деталізація, довготривале зберігання, можливість поєднувати різні джерела – радикально змінюють контекст приватності. Якщо раніше ризик для конфіденційності часто обмежувався витоком окремої бази даних із кількома полями, то тепер загроза полягає у можливості побудови «цифрового портрета» людини через зіставлення десятків наборів: банківських транзакцій, історії пошуку, дописів, геоданих, медичних показників. Високорозмірні дані роблять традиційні методи анонімізації нестійкими – достатньо кількох унікальних комбінацій, щоб вирізнити конкретну особу серед мільйонів записів, створюючи глибинний конфлікт між інтересом суспільства до відкритих даних та аналітики і правом людини на приватність [3].

У сучасному стані приватність уже неможливо захищати виключно маскуванням ідентифікаторів або правовими угодами. Серія гучних кейсів – від

витоку пошукових логів AOL [4] до деанонімізації Netflix Prize [5] і скандалу з Cambridge Analytica [6] – показала, що навіть «знеособлені дані» можуть бути повторно пов’язані з конкретними людьми, якщо зловмисник має зовнішні джерела або достатню обчислювальну потужність.

Наукова спільнота й регулятори були підштовхнуті до пошуку методів, які не просто приховують імена, а формально обмежують, скільки інформації про конкретну особу може бути витягнуто з результатів аналізу. На цьому тлі диференційна конфіденційність (ДК) з’являється як одна з ключових відповідей – не як технічну зміну, а як новий стандарт мислення про захист даних.

1.1 Проблема конфіденційності у великих наборах даних

Класичний підхід до захисту даних довгий час спирався на припущення. Якщо прибрати явні ідентифікатори, такі як ім’я, індивідуальний податковий номер, номер телефону, або провести агрегування, дані стають безпечними. У добу Big Data (BD) це припущення системно перестало працювати. Причина не лише в кількості даних, а в можливості їх поєднання. Кожен окремий набір може здаватися «анонімним», але разом вони утворюють унікальну комбінацію ознак, яка однозначно вказує на конкретну людину. Так звані квазіідентифікатори: вік, поштовий індекс, стать, інтереси, динаміка активності у великій вибірці майже завжди формують унікальний «відбиток».

Показовим прикладом може виступати інцидент із AOL у 2006 році [4]. Компанія опублікувала 20 мільйонів пошукових запитів понад 650 тисяч користувачів, замінивши логіни на числові ідентифікатори. Дуже швидко журналісти і дослідники продемонстрували, що за історією запитів (адреси, імена, локальні деталі) можна встановити особу конкретних користувачів. Формально імена були прибрані, але поведінковий шаблон виявився достатнім для ототожнення.

Аналогічну вразливість показало дослідження щодо датасету Netflix Prize [5]. Порівнявши анонімізовані рейтинги фільмів Netflix із публічними профілями

IMDb, вони змогли деанонімізувати частину користувачів і відкрити їхні вподобання, які ті вважали приватними.

Усі ці роботи показали, що маленька зміна без формального контролю ризику – ілюзія безпеки.

Ще один вимір проблеми – використання великих масивів даних для маніпуляцій і дискримінації. Скандал Cambridge Analytica [6] продемонстрував, як поведінкові та соціальні дані, зібрані частково без прозорої згоди, використовувалися для психографічного профілювання виборців і мікротаргетингу політичної реклами. Хоча це не був класичний витік анонімізованих даних, кейс показав іншу сторону Big Data – навіть легально зібрані дані можуть привести до порушення приватності очікувань, коли люди не розуміють масштабів аналізу й поєднання джерел. Подібні історії підривають довіру до цифрових сервісів і стимулюють законодавців вимагати реальних, а не декларативних гарантій.

Серйозність проблеми посилив і правовий контекст. Європейський GDPR прямо говорить: «Захист даних не поширюється на справді анонімізовану інформацію, але такі дані мають бути опрацьовані так, щоб суб’єкт більше не був ідентифікований «ні прямо, ні опосередковано»» [7].

На практиці доведення для традиційних методів (k-анонімність, просте узагальнення, видалення полів) все важче, оскільки зростає кількість зовнішніх даних, які зловмисник може використати для зв’язування. Як наслідок, організації опиняються між двома вогнями: з одного боку – вимога відкривати дані для науки, бізнесу, державного управління, з іншого – загроза юридичної відповідальності за можливу повторну ідентифікацію, що створює попит на методи, здатні дати формальну, перевірену гарантію. Що б не знав зловмисник, участь конкретної особи у наборі суттєво не збільшує його здатність дізнатись щось нове саме про неї.

Таким методом і стала диференційна конфіденційність, вона виникла як відповідь на демонстративні провали «класичної» анонімізації, коли стало

очевидно, що у світі Big Data недостатньо просто змінити імена – потрібно математично обмежити інформаційний внесок кожної окремої людини.

1.2 Поняття диференційної конфіденційності

Інтуїтивно ДК відповідає на людське запитання: «Чи стане моє життя менш приватним, якщо я дозволю включити свої дані до цього аналізу?» Метод дає формальну гарантію, незалежно від того, чи ваш запис є в базі, результати публічних запитів про цю базу виглядатимуть практично однаково.

Нехай, місто хоче опублікувати статистику, де вказано скільки жителів мають певне захворювання. Якщо оприлюднити точні числа для невеликих груп (наприклад, жінки 60–65 років у маленькому селищі), за допомогою зовнішньої інформації можна здогадатися, хто конкретно входить у цю статистику.

Диференційно-приватний механізм додає до результатів невеликий випадковий шум, наприклад замість «37» публікує «35» чи «39». Для дослідника похибка вважатиметься не критичною, тому що загальний рівень захворюваності він все одно бачить. Але для зловмисника немає надійного способу сказати, чи додала участь конкретної людини хоч якусь визначальну інформацію до результату. Якщо її запис видалити, розподіл можливих виходів зміниться лише в межах, визначених ε.

Принципово важливо, що диференційна конфіденційність – це властивість механізму/алгоритму, а не одного очищеного датасету. Будь-який результат, отриманий із чутливих даних – звіт, статистика чи модель – проходить через механізм, який формально обмежує можливість розкриття інформації про окрему особу.

Такий підхід означає перехід від традиційної концепції «разового знеособлення з подальшим вільним використанням даних» до моделі, що забезпечує постійний контроль рівня інформаційного витоку.

Не менш важливо й те, що гарантія ДК є стійкою до фону знань зловмисника. Індиферентно, скільки додаткових джерел (інші бази, соцмережі, відкриті реєстри) є в доступі. Умова на ймовірності забезпечує, що присутність

чи відсутність однієї людини майже не змінить шанси будь-якого конкретного виходу, що робить ДК незалежною від моделі зловмисника – вона не зламується, коли з'являється новий тип зовнішніх даних, на відміну від багатьох попередніх методів де-ідентифікації.

1.3 Основні принципи диференційної конфіденційності

Для глибшого аналізу роботи ДК на практиці, варто зосередити увагу на наступних ключових принципах, які в поєднанні дають ту саму формальну гарантію.

Перший принцип – кількісна міра ризику через параметри ε і δ. Параметр ε інтерпретується як максимально допустимий витік інформації про одну особу. Чим менший ε, тим сильніша приватність – результати для наборів з і без конкретного запису майже невідрізні. Значення $\varepsilon \approx 0.1 - 1$ вважають дуже суворим захистом, $\varepsilon \approx 1 - 10$ – помірним балансом між приватністю і точністю. У практичних системах (Apple, Google, перепис США) ε зазвичай обирають виходячи з компромісу: недостатнє значення веде до втрати якості, завелике — до зниження гарантії. Параметр δ використовується в наближеній (ε, δ)-ДК і означає дуже малу ймовірність того, що умова конфіденційності буде порушена. Зазвичай використовується як 10^{-5} або менше, щоб ризик був практично нульовим.

Другий принцип – обмежена чутливість. Перед тим як додавати шум, алгоритм оцінює, наскільки сильно один запис може змінити результат запиту. Для підрахунку кількості людей з певною властивістю чутливість дорівнює 1 (додати/видалити людину змінює результат максимум на 1); для середнього – залежить від діапазону значень. Чим суворіше обмежено внесок однієї особи (наприклад, обрізанням екстремальних значень), тим менше шуму потрібно додати для досягнення того ж рівня приватності. Такий підхід стимулює проєктувати системи так, щоб жоден індивід не мав «надмірної ваги» в результатах.

Третій принцип – додавання випадкового шуму, відкаліброваного до чутливості і ε. Класичні механізми – лапласівський та гауссівський. Лапласівський механізм додає шум із Laplace(0, $\frac{\Delta f}{\varepsilon}$), де Δf – чутливість функції; гауссівський – з нормального розподілу з дисперсією, обраною для (ε, δ)-DP, що буде розглянуто нижче. Ідея полягає в тому, що шум розмиває внесок окремої особи, але на великих масивах даних статистика залишається стабільною. Якщо рахунок іде на тисячі, додавання кількох одиниць шуму не спотворює загальну картину.

Четвертий принцип – композиція та бюджет приватності (БП). Одна з головних причин, чому традиційні анонімізації можуть бути зламані – можливість повторного запиту до тих самих даних. У ДК це враховано наступним чином: якщо до одного набору застосовується кілька запитів, загальна втрата приватності не перевищує суму ε цих запитів. Система повинна вести рахунок того, скільки приватності вже витрачено. Коли бюджет вичерпано, тоді, або забороняються нові запити, або потрібно працювати з іншими даними. Такий підхід дисциплінує аналітику, тому що приватність перестає бути розмитою етичною категорією і перетворюється на керований ресурс.

П’ятий принцип – стійкість до постобробки. Якщо вихід алгоритму вже задовольняє ДК, будь-яка подальша обробка (графіки, фільтри, моделі на основі цих результатів) не може зіпсувати приватність, поки не звертається назад до сирих даних. Що виявляється дуже практичним: інженери можуть будувати звіти, візуалізації, дашборди поверх диференційно-приватних результатів, не побоюючись, що сам процес візуалізації зламає гарантії.

Останній, проте не менш важливий практичний вимір – зв’язок з реальними інцидентами і впровадженнями. Після провалів AOL і Netflix низка організацій почала впроваджувати ДК саме як відповідь на ризики Big Data. Apple застосовує локальну ДК для збору статистики про емодзі, нові слова і деякі аналітичні події на пристроях користувачів: шум додається на телефоні ще до відправки, тож навіть сервер бачить лише зашумлені записи [8].

Бюро перепису США використало ДК для публікації даних перепису 2020 року, визнавши, що старі методи захисту, такі як приховування комірок та заміна записів, більше не гарантують стійкості до сучасних атак [9].

Наведені приклади демонструють причинно-наслідковий ланцюжок: зростання ризиків + провали традиційних методів ⇒ поява формальних моделей ⇒ реальне промислове впровадження ДК як нового стандарту конфіденційності.

1.4 Візуалізація роботи диференційно-приватного механізму

ДК функціонує як проміжний захисний рівень між сирими даними та аналітичними запитами, контролюючи, яку інформацію можна отримати про окремого користувача. Її сутність полягає у створенні математично гарантованої невизначеності, що унеможливлює достовірне визначення участі конкретної особи в наборі даних. Принцип дії ДК можна інтерпретувати через систему вимірювання: якщо два набори даних відрізняються лише одним елементом, надмірна точність може виявити цю різницю, що еквівалентно витоку приватності. Додавання випадкового шуму робить різницю статистично незначущою, забезпечуючи конфіденційність без втрати аналітичної цінності.

Архітектурно система ДК складається з наступних рівнів (рис. 1.1).

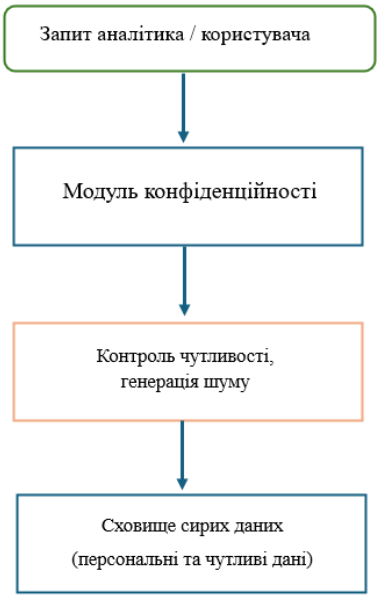

Рисунок 1.1 – Система ДК

Користувач або аналітик формує запит до даних, після чого модуль ДК оцінює чутливість функції – тобто, наскільки результат може змінитися при додаванні або видаленні одного запису. На основі цієї оцінки генерується випадковий шум із розподілу Лапласа або Гауса, масштабований відповідно до параметрів чутливості Δf та ε. Відповідно шум додається до реального результату, утворюючи «зашумлену» відповідь, яка повертається користувачу. Система також веде облік витраченого бюджету приватності, контролюючи сукупний ризик при повторних запитах.

Існують дві основні архітектурні моделі реалізації ДК: централізована (Global DP) та локальна (Local DP). У Global DP моделі передбачається наявність довіреного куратора, який має доступ до сирих даних і додає шум перед публікацією результатів. Такий підхід використовується у великих державних статистичних системах, зокрема під час національних переписів, коли необхідно забезпечити захист даних на рівні агрегованих звітів. У Local DP моделі кожен користувач додає шум самостійно до своїх даних до їх передачі серверу. Такий підхід активно застосовується у телеметричних та аналітичних системах, де дані з різних пристроїв обробляються без прямого доступу до вихідних значень [10].

Для ілюстрації можна розглянути гістограму (рис. 1.2) та таблицю (табл. 1.1) кількості пацієнтів за віковими групами.

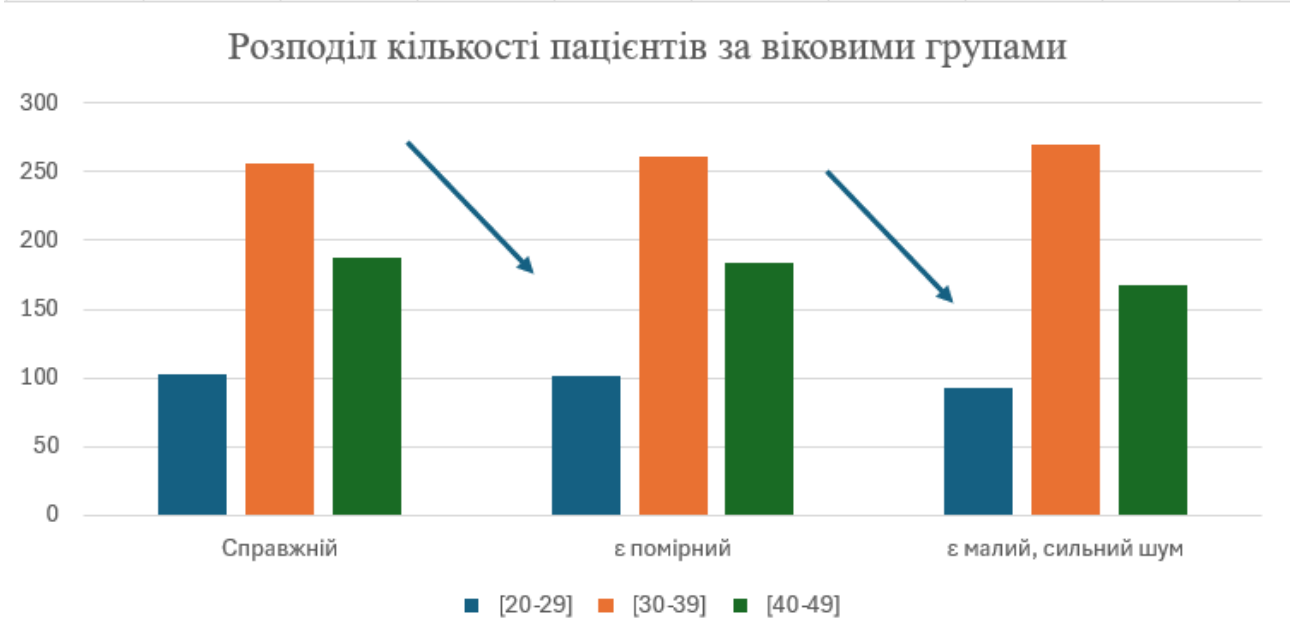

Рисунок 1.2 – Гістограма кількості пацієнтів

Після застосування ДК із різними значеннями ε спостерігається, що навіть при значному рівні шуму (ε = 0.5) зберігається форма розподілу, тобто дані залишаються корисними для аналізу, що демонструє ключову властивість механізму – підтримання балансу між точністю результатів і рівнем приватності.

Таблиця 1.1 – Розподіл вікової групи пацієнтів

	[20-29]	[30-39]	[40-49]	Середнє	Медіана
Справжній	103	256	187	182	187
ε помірний	101	261	184	182	184
ε малий, сильний шум	99	273	174	182	174

Табличне представлення даних у цьому випадку має ілюстративний, а не рівнозакономірний характер. Обчислене середнє значення залишається незмінним для всіх варіантів шуму, що свідчить про збереження аналітичної корисності результатів та неможливість ідентифікації окремих осіб. Водночас медіана, яка є більш чутливою до локальних змін у розподілі, демонструє варіації під впливом доданого шуму, відображаючи ступінь зсуву даних при забезпеченні приватності.

Диференційна конфіденційність не є просто методом додавання шуму, а цілісною системою контролю витоку інформації. Вона забезпечує перехід від реактивного підходу до проактивної моделі, що керує ризиками на етапі аналізу даних. Такий принцип дозволяє будувати етичну аналітику у сфері Big Data, де великі дані використовуються для суспільної користі без порушення приватності окремих осіб.

У розділі було розглянуто еволюцію підходів до захисту приватності даних у добу Big Data, проаналізовано основні причини втрати ефективності традиційних методів анонімізації та обґрунтовано появу ДК як формальної відповіді на сучасні виклики. Було показано, що зростання обсягів, швидкості та різноманіття даних, а також можливість їх поєднання з різних джерел, призводять до ризику реідентифікації навіть за відсутності прямих ідентифікаторів. Приклади інцидентів з AOL, Netflix Prize та Cambridge Analytica

підтвердили, що класичні підходи до знеособлення не забезпечують стійкості в умовах розвинених обчислювальних можливостей та відкритих даних.

Було розглянуто основні принципи диференційної конфіденційності – параметри ε та δ, обмеження чутливості, додавання випадкового шуму, управління бюджетом приватності та стійкість до пост обробки. Показано, що ДК функціонує як захисний шар між аналітичними запитами та сирими даними, забезпечуючи контрольований рівень невизначеності без суттєвої втрати аналітичної точності. Зроблено висновок, що ДК є не просто технічним механізмом, а новим стандартом побудови етичних і безпечних систем аналізу даних, який перетворює приватність із декларативного принципу на формально вимірювану властивість інформаційної системи.

2 МАТЕМАТИЧНА ФОРМАЛІЗАЦІЯ ТА МЕХАНІЗМИ ДИФЕРЕНЦІЙНОЇ КОНФІДЕНЦІЙНОСТІ

У попередньому розділі було окреслено концептуальні основи ДК в контексті BD та продемонстровано загальну логіку її функціонування. Водночас для практичного застосування ДК у системах обробки даних необхідне глибоке розуміння математичного апарату, який забезпечує гарантії приватності.

2.1 Термінологія та позначення

Дані надано ключові поняття, які використовуються у роботі.

Куратор даних (КД) – довірена сторона, яка керує зібраними даними протягом їхнього життєвого циклу, включно із анотуванням, публікацією та презентацією. У контексті ДК куратор відповідає за те, щоб приватність жодної особи, представленої у наборі даних, не була порушена в процесі аналізу чи оприлюднення інформації.

Зловмисник – аналітик даних, зацікавлений у виявленні чутливої інформації про окремих осіб у наборі. Навіть легітимний користувач бази даних розглядається як потенційний зловмисник, оскільки його запити можуть непередбачувано порушити приватність індивідів.

Механізм – алгоритм, який надає статистичну інформацію про набір даних. ДК сама по собі є абстрактною концепцією, натомість реалізація її математичних гарантій відбувається через конкретні механізми [11].

База даних (набір даних) може бути представлена як мультимножина із записами з кінцевого універсуму X. Зручним способом представлення бази D є її гістограма $D \in \mathbb{N}^{|X|}$, де кожен елемент d_i означає, скільки разів елемент $k_i \in K$ (де $1 \leq i \leq |K|$) зустрічається в D. Символ $\mathbb{N}$ позначає множину натуральних чисел $\{1,2,\dots\}$.

Доречно буде навести приклад набору даних, який містить п’ять професій, які умовно живуть в одному будинку:

- Викладач;
- Страхувальник;
- Митець;
- Інженер;
- Митець.

Універсум X складається з усіх можливих професій. Колекція записів має наступний вигляд: {«Викладач», «Страхувальник», «Митець», «Інженер», «Митець»}. Гістограмне представлення в цьому випадку виглядатиме як [1, 1, 2, 1, 0, …0], або для кращого представлення {«Викладач»:1, «Страхувальник» :1, «Митець» :2, «Інженер» :1} із нульовим лічильником для всіх інших професій з універсуму.

ℓ_1-норма бази даних D позначається як $\| D \|_1$ та вимірює розмір D, тобто кількість записів, які вона містить:

$$\| D \|_1 := \sum_{i=1}^{|X|} |d_i|. \tag{2.1}$$

ℓ_1-відстань між двома базами D_1 та D_2 визначається як $\| D_1 - D_2 \|_1$ і вимірює, скільки записів відрізняється між обома базами:

$$\| D_1 - D_2 \|_1 := \left\lceil \frac{\sum_{i=1}^{|X|} |d_{1,i} - d_{2,i}|}{2} \right\rceil. \tag{2.2}$$

Сусідні бази даних. Дві бази D_1 та D_2 називаються сусідніми, якщо вони відрізняються не більше ніж одним елементом:

$$\| D_1 - D_2 \|_1 \leq 1. \tag{2.3}$$

Поняття сусідніх баз є центральним для визначення ДК, оскільки метою є забезпечення того, щоб результати аналізу на сусідніх базах були майже нерозрізнимими.

2.2 ε-диференційна конфіденційність

Формально, у класичному визначенні Двок і Рота [11] алгоритм A задовольняє ε-диференційну конфіденційність, якщо для будь-яких двох сусідніх наборів даних D_1 та D_2, які відрізняються лише одним записом, і для будь-якої множини можливих результатів S виконується:

$$P[A(D_1) \in S] \leq e^{\varepsilon} \cdot P[A(D_2) \in S], \quad (2.4)$$

де ε ≥ 0 – параметр приватності.

Чим менше ε, тим менше змінюється розподіл виходів алгоритму при додаванні/видаленні одного запису, і тим сильніша приватність. Ключова думка полягає в тому, що алгоритм робить результати нечутливими до інформації про одну людину, тому ззовні практично неможливо визначити, чи входили її дані до набору.

Оскільки D_1 і D_2 можна поміняти місцями, визначення безпосередньо передбачає наступну нижню межу:

$$\mathrm{P}[A(D_1) \in S] \geq e^{-\varepsilon} \cdot \mathrm{P}[A(D_2) \in S]. \quad (2.5)$$

Комбінуючи обидві нерівності, отримуємо обмеження:

$$-\varepsilon \leq \ln\left(\frac{\mathrm{P}[A(D_1) \in S]}{\mathrm{P}[A(D_2) \in S]}\right) \leq \varepsilon. \quad (2.6)$$

Альтернативне формулювання базується на ідеї мультиплікативного фактору $(1 + \varepsilon)$:

$$\mathrm{P}[A(D_1) \in S] \leq (1 + \varepsilon) \cdot \mathrm{P}[A(D_2) \in S]. \quad (2.7)$$

Якщо ε достатньо мале, різниця між $(1 + \varepsilon)$ і e^{ε} є нехтовною. Проте вже при $\varepsilon \geq 0.5$ величина e^{ε} принаймні на 10% більша. Із зростанням ε значення суттєво розходяться. Водночас e^{ε} є загальноприйнятим фактором у літературі.

Аргументи на користь роботи з e^{ε} включають відповідність кривій розподілу Лапласа (рис. 2.1), яка має потрібні для ДК властивості. За умови зміщення кривої на певну величину, співвідношення ймовірностей для оригінальної та зміщеної кривої залишається в межах заданої границі.

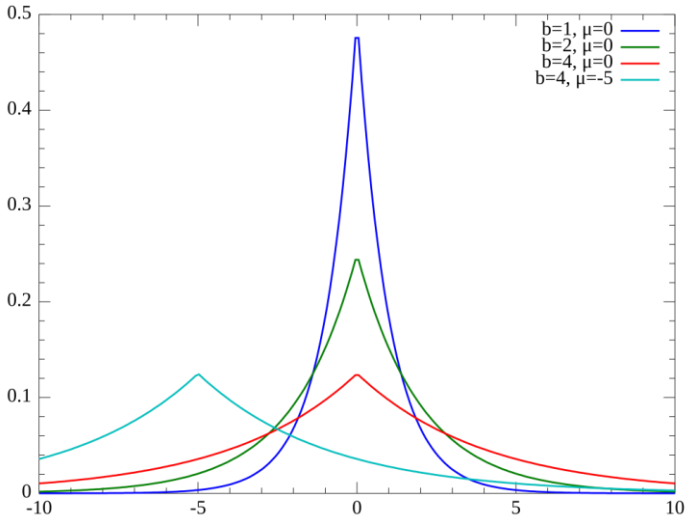

Рисунок 2.1 – Крива розподілу Лапласа [12]

Робота з виразом e^{ε} має низку практичних переваг.

По-перше, підвищується швидкість обчислень, адже замість множення двох ймовірностей у звичайному просторі, у логарифмічному достатньо виконати операцію додавання, що є значно менш ресурсомістким процесом.

По-друге, такий підхід забезпечує вищу числову стабільність під час роботи з дуже малими ймовірностями, оскільки зменшує похибки, пов'язані з обмеженою точністю представлення дійсних чисел у комп'ютерах.

По-третє, використання логарифмів спрощує математичні вирази, оскільки багато ймовірнісних розподілів, зокрема тих, які застосовуються для додавання випадкового шуму, мають експоненційну форму – після взяття логарифма експоненційну функцію можна усунути, зводячи обчислення до роботи лише з її показником.

Як було зазначено раніше, параметр ε часто називають бюджетом приватності. Прикладом механізму, що демонструє властивість досконалої секретності, є одноразовий блокнот для одного біта.

Універсум баз даних є $\{0,1\}$, як і образ механізму. Вхідним сигналом до механізму є біт $b \in \{0,1\}$. Механізм обирає новий випадковий біт $p \in \{0,1\}$ і виводить $(b + p) \bmod 2$. Перевіряючи всі можливості, отримано, що

$$\frac{1}{2} = \mathrm{P}[A(0) = 0] = \mathrm{P}[A(0) = 1] = \mathrm{P}[A(1) = 0] = \mathrm{P}[A(1) = 1]. \qquad (2.8)$$

Отже, формулу 2.4 можливо підставити у результат механізму:

$$\forall D_1, D_2, \forall S \subseteq \mathrm{Im}(A): \frac{\mathrm{P}[A(D_1) \in S]}{\mathrm{P}[A(D_2) \in S]} = 1 = e^{\varepsilon} \Rightarrow \varepsilon = 0. \quad (2.9)$$

Усі досконало приватні механізми (вихід механізму не розкриває жодної інформації) є диференційно приватними, але клас ДП механізмів містить системи, які не забезпечують досконалої приватності. Натомість вони надають певну нетривіальну інформацію про базу даних, що необхідно для проведення аналізу з гарантіями приватності.

2.3 (ε, δ)-диференційна конфіденційність

ε-диференційна конфіденційність встановлює високі вимоги до механізмів. Водночас додавання занадто великої кількості шуму до оригінальних даних може суттєво обмежити обсяг інформації. Тому було запропоновано кілька послаблених версій ДК. Однією з найпопулярніших версій є (ε, δ)-диференційна конфіденційність, про яку було згадано раніше.

Рандомізований алгоритм A з доменом $\mathbb{N}^{|X|}$ забезпечує (ε, δ)-диференційну конфіденційність, якщо для всіх сусідніх баз даних $D_1, D_2 \in \mathbb{N}^{|X|}$ і для всіх $S \subseteq \mathrm{Im}(A)$:

$$\mathrm{P}[A(D_1) \in S] \leq e^{\varepsilon} \cdot \mathrm{P}[A(D_2) \in S] + \delta. \quad (2.10)$$

Такий тип дозволяє порушувати вимоги ε-ДК до певного адитивного ступеня, фіксованого параметром δ – представляє ймовірність того, що вихід механізму варіюється більше ніж на фактор e^{ε} між двома сусідніми базами. Іншими словами, абсолютне значення втрати приватності буде обмежене ε з ймовірністю принаймні $(1 - \delta)$.

У випадку, коли $\delta = 0$, (ε, δ)-ДК ≡ ε-ДК.

Виходячи із вищезгаданого в роботі, інтуїтивно зрозуміло, що δ має бути достаньо малим значенням, тому що при великому δ, навіть, якщо $\varepsilon = 0$ (досконала секретність), механізм, що є $(0, \delta)$-ДП, порушуватиме приватність із високою ймовірністю. Поширений практичний підхід полягає в тому, щоб обрати значення δ для бази з n записами таким чином, щоб δ було меншим за $\frac{1}{n}$. Все пояснюється тим, що у (ε, δ)-ДК механізмі існує ймовірність δ, з якою

інформація про окремий запис може бути частково розкрита. Якщо розглядати це з точки зору математичного сподівання, то загалом алгоритм може «розкрити» близько $n \cdot \delta$ записів, отже δ має бути дуже малим.

2.4 Групова приватність та k-кратний захист

Основне поняття ДК сфокусоване на захисті окремої особи. Однак у багатьох практичних сценаріях захист окремої особи не гарантує захист групи, оскільки зловмисник може мати допоміжну інформацію, що пов'язує кількох людей разом.

Двок [11] розглядає групи переважно як сім'ї або споріднених осіб. Допустимо, база містить медичні дані члена сім'ї. Навіть, якщо система гарантує захист індивіда (ε-DP), опубліковані результати можуть розкрити інформацію про його родичів, особливо, якщо відомо, що деякі члени родини мають генетичні спорідненості або спільний образ життя, що проявляється у схожих захворюваннях.

Згідно з [13], будь-яка колекція індивідів, така як політичні колективи, етнічні групи, а також угруповання, створені алгоритмами, може представляти групу.

Будь-який ε-ДП механізм A є $(k\varepsilon)$-ДП для груп розміру k, якщо для всіх баз даних D_1 і D_2, які відрізняються не більше ніж у k елементах, і для всіх множин $S \subseteq \mathrm{Im}(A)$:

$$\mathrm{P}[A(D_1) \in S] \leq e^{k\varepsilon} \cdot \mathrm{P}[A(D_2) \in S]. \qquad (2.10)$$

Теорема 2.10 має чітку, проте песимістичну імплікацію: Якщо розмір групи збільшується, приватність групи суттєво погіршується.

Припустимо, система забезпечує (ε = 1.0)-ДП для окремих осіб. Для сім'ї з 5 осіб гарантія падає до $(5\varepsilon = 5.0)$-ДП. Коефіцієнт e^5 означає, що ймовірність розкриття чутливої інформації про сім'ю в цілому може бути до 148 разів вищою, ніж про окремо особу.

В іншому випадку, якщо група буде ще більшою, як, наприклад, база студентів на 10000 записів. Окремий студент захищений при (ε = 0.5). Однак для

групи всіх студентів професійного направлення, умовно 300 осіб, приватність падає до $(300\varepsilon = 150.0)$-ДП, що практично означає розкриття.

Як видно з табл. 2.1 , навіть при ($\varepsilon = 0.1$) на особу, група з 1000 осіб отримує лише слабкий захист, що створює проблему системі в одночасному захисті як для окремих осіб, так і для груп.

Таблиця 2.1 – Деградація приватності за розміром групи

k (розмір групи)	ε для індивіда	Приватність групи	e^k	Практичне значення
1	1.0	1.0	2.72	Сильна приватність
5	1.0	5.0	148	Помірна приватність
10	1.0	10.0	22,026	Слаба приватність
100	1.0	100.0	2.7×10^{43}	Фактично розкриття
1000	0.1	100.0	2.7×10^{43}	Фактично розкриття

Проблема групової приватності (ГП) піднімає низку важливих етичних і практичних питань. Один із ключових парадоксів полягає в тому, що захист групи часто вступає в суперечність із захистом окремої особи. Якщо необхідно гарантувати однаковий рівень приватності ε для всієї групи з k осіб, як і для індивідуального користувача, то для кожного учасника потрібно застосувати (ε/k)-диференційну конфіденційність. Отже, що до даних потрібно додати значно більше шуму, що погіршує точність і якість результатів.

Другою проблемою є вибір пріоритетів: організація має визначити, чий захист важливіший – окремої особи чи групи. Хоча зазвичай пріоритет надається індивідуальній приватності, у певних контекстах, таких як дослідження меншин або вразливих соціальних груп, пріоритет може зміщуватися в інший бік. Додаткову складність становить те, що часто невідомо, які саме групи потребують захисту, адже зловмисник може розглядати будь-яку підмножину даних як потенційну «групу» для атаки. Тому навіть, якщо система гарантує (k)-ДК для всіх груп певного розміру, це не виключає ризиків зловживань.

На практиці, коли важливо забезпечити групову приватність, доцільно дотримуватись наступних рекомендацій:

1) Замість спроби захистити всі можливі комбінації осіб, організація повинна чітко визначити, які саме групи потребують захисту – наприклад, родини, демографічні категорії або регіональні спільноти.

2) Якщо загальний бюджет конфіденційності становить $\varepsilon_{\text{total}}$, його можна поділити між групами так, щоб кожна мала свій рівень захисту:

$$\varepsilon_{group} = \frac{\varepsilon_{\text{total}}}{k}. \quad (2.11)$$

3) Можна застосовувати різні рівні захисту, де сильніший для окремих осіб, помірний для невеликих груп і слабший для великих груп, що дозволяє досягти балансу між точністю та приватністю.

4) Користувачі повинні бути поінформовані про те, який рівень захисту надано їхнім даним і до яких груп вони належать у процесі обробки інформації.

2.5 Композиція механізмів диференційної конфіденційності

Композиція – обробка однієї бази даних кількома запитами послідовно – формує найбільшу проблему практичного застосування ДК. У реальних умовах окремі запити майже ніколи не виконуються ізольовано: зазвичай проводиться низка послідовних аналітичних операцій над одними й тими самими даними. Наприклад, статистика може послідовно обчислювати середнє значення, медіану та квартилі доходів на основі одного набору податкових записів; або система машинного навчання може використовувати спільну базу даних для етапів навчання моделі, налаштування гіперпараметрів, оцінювання якості та формування підсумкового звіту.

2.5.1 Послідовна композиція

Послідовна композиція: Нехай A_1 з $\text{Im}(A_1) = S_1$ є ε_1-ДП алгоритмом, і нехай A_2 з $\text{Im}(A_2) = S_2$ є ε_2-ДП алгоритмом. Тоді їхня комбінація $\text{Im}(A_{1,2}) = S_1 \times S_2$ є $(\varepsilon_1 + \varepsilon_2)$-ДП.

Вищенаведена теорема передбачає незалежність механізмів ДК. Проте в реальних умовах часто виникають ситуації, коли наступні запити формуються з урахуванням результатів попередніх, тобто мають адаптивний характер.

Наприклад, дослідник може спочатку запитати: «Яка кількість осіб має ВІЛ?». Якщо отримане значення виявляється високим, він уточнює: «Розподіліть ці випадки за регіонами». Якщо ж значення низьке – подальших запитів не надходить. Таким чином, наступний запит залежить від попереднього результату, що ускладнює математичний аналіз композиції.

Звідси випливає наступне, що для $i \in [k]$ композиція k механізмів $A_{[k]}(D) = \left(A_1(D), A_2(D), \ldots, A_k(D)\right)$ є $\left(\sum_{i=1}^{k} \varepsilon_i\right)$-диференційно приватною. Аналогічно, під послідовною композицією у (ε, δ)-диференційно приватних механізмах значення δ також підсумовуються (табл. 2.2).

Таблиця 2.2 – Кількість запитів при різних параметрах

$\varepsilon_{\text{total}}$	ε_1	Кількість запитів	Приклад
1.0	0.1	10	Помірно детальний аналіз
1.0	0.05	20	Детальний аналіз
1.0	0.01	100	Дуже детальний аналіз
0.1	0.1	1	Лише один запит
0.1	0.01	10	Обмежений аналіз

З практичної точки зору, послідовна композиція має важливі наслідки для керування приватністю в аналітичних системах, а саме обмеження кількості запитів. Якщо організації надано загальний бюджет приватності $\varepsilon_{\text{total}}$, то кількість можливих запитів є обмеженою. Якщо кожен запит використовує частину бюджету $\varepsilon_{\text{query}}$, то максимальна кількість допустимих запитів визначається співвідношенням:

$$k_{\max} = \frac{\varepsilon_{\text{total}}}{\varepsilon_{\text{query}}}. \tag{2.12}$$

Як видно, дослідник часто вдається до компромісу, де потрібно виконати декілька запитів із нижчим рівнем приватності кожного або ж один запит із більш високим захистом.

Вичерпання бюджету приватності пропонує в окремих випадках свідомо витрачати весь доступний БП на обмежену кількість запитів, замість поступового його розподілу. Такий підхід застосовується тоді, коли пріоритетом є висока точність і надійність отриманих результатів, навіть ціною зменшення можливості виконання додаткових запитів у майбутньому.

Ефективне застосування принципу послідовної композиції вимагає ретельного планування ще до початку аналітичного процесу. Організація має заздалегідь визначити всі необхідні запити до бази даних і оцінити їхній вплив на загальний БП.

Не менш важливою є прозорість у документуванні. Кожен виконаний запит і його вартість у термінах ε мають бути чітко зафіксовані в технічній документації, що не лише підвищує рівень підзвітності процесів обробки даних, а й допомагає уникнути випадкового перевитрачання БП, що може призвести до порушення гарантій ДК.

Для підвищення точності контролю доцільно використовувати спеціалізовані інструменти та фреймворки, зокрема OpenDP [14] або Tumult Analytics [15]. Вони автоматично відстежують використання, обчислюють кумулятивні витрати та своєчасно попереджають дослідників про наближення або перевищення встановленого ліміту.

2.5.2 Паралельна композиція

Паралельна композиція розглядає сценарій, який істотно відрізняється від попереднього. Замість послідовних запитів до однієї бази, база розділяється на неперетинні підмножини, й до кожної підмножини застосовується ДК механізм незалежно. Результати потім комбінуються деякою функцією (усередненням чи конкатенацією).

Для загального розуміння буде наведено наступний приклад. Країна розділена на 5 регіонів. У кожному регіоні є база медичних даних. Кожен регіон незалежно публікує статистику про розповсюдження хвороб, використовуючи ДК. Національний центр потім комбінує результати для отримання статистики. Головний момент: запис про одну особу належить лише одному регіону, тому вона впливає лише на його статистику.

Нехай існує n диференційно-приватних механізмів A_i з $\mathrm{Im}(A_i) = S_i$, кожен з яких забезпечує ε_i-ДК при обчисленні на неперетинних підмножинах $D^{(i)}$

вхідного домену D. Нехай $A = \{A_1, \ldots, A_n\}$. Далі нехай $r_i = A_i\left(D^{(i)}\right)$ – вихід. Тоді будь-яка функція $g(r_1, \ldots, r_n)$ з $g: 2^A \times \mathbb{N}^{|X|} \to \mathbb{R}^k$ є $\max\varepsilon_i$-ДП.

Порівняння послідовної та паралельної композиції (рис. 2.2) дає змогу зрозуміти, як різні підходи до виконання запитів впливають на рівень приватності. У випадку послідовної композиції, коли до однієї й тієї самої бази даних здійснюється серія запитів – сукупна витрата БП дорівнює сумі всіх окремих витрат.

Рисунок 2.2 – Порівняння послідовної та паралельної композиції

Зовсім інша ситуація спостерігається у випадку паралельної композиції. Якщо база даних попередньо розділена на десять непересічних підмножин, то загальний рівень приватності для всієї системи не збільшується.

Попри теоретичні переваги паралельної композиції, її практичне застосування супроводжується низкою суттєвих викликів. Однією з головних проблем є перекриття записів: у багатьох реальних сценаріях один і той самий запис може належати до кількох підмножин даних. Наприклад, користувач мобільного зв’язку може пересуватися між різними регіонами, що унеможливлює чітке географічне розділення даних. У таких випадках класичні теореми паралельної композиції втрачають чинність, оскільки забезпечення неперетинності множин є базовою умовою їхнього застосування.

Другою проблемою є нерівномірний розподіл підмножин. Якщо різниця між розмірами груп є значною (один регіон охоплює мільйон осіб, а інший лише кілька сотень), то чутливість функцій у менших підмножинах істотно зростає, що в свою чергу, вимагає додавання більшого шуму для збереження рівня

приватності та погіршує точність отриманих результатів, знижуючи аналітичну корисність даних.

Ще один важливий аспект – коефіцієнт обміну між приватністю та точністю. Хоча паралельна композиція теоретично дозволяє зменшити втрату для кожного окремого запису, це може призвести до суттєвого збільшення похибки у фактичних результатах. Якщо підмножини надто малі, то шум, який додається для захисту даних, починає домінувати над корисною інформацією.

Для ефективного застосування паралельної композиції доцільно дотримуватись низки практичних рекомендацій.

По-перше, слід забезпечити явну несумісність підмножин. Заздалегідь перед виконанням аналізу необхідно переконатися, що розділи бази даних є справді неперетинними, щоб гарантії паралельної композиції залишалися чинними.

По-друге, варто враховувати різницю у розмірах підмножин при виборі параметра ε для кожного механізму. Менші підмножини, як правило, потребують зменшеного значення, щоб зберегти прийнятний рівень точності.

І нарешті, у випадках, коли повне розділення підмножин неможливе, доцільно застосовувати гібридні підходи, які комбінують елементи послідовної та паралельної композиції, збалансувавши захист приватності та збереження корисності даних навіть у складних структурах.

2.5.3 Розширена композиція

Існує також узагальнена концепція розширеної композиції, яка враховує складніші та більш реалістичні сценарії використання даних. Вона охоплює не лише випадки повторного застосування ДК механізмів до однієї й тієї самої бази, а й ситуації, коли різні механізми використовуються на різних наборах даних, що можуть частково перетинатися. У таких випадках окремі бази можуть містити інформацію про тих самих осіб, навіть, якщо вони були створені незалежно одна від одної.

Подібні ситуації є типовими для сучасних інформаційних систем, де нові бази даних формуються постійно. Крім того, у деяких випадках зловмисники можуть мати змогу впливати на структуру або зміст нових баз даних, наприклад, шляхом додавання контрольованих записів чи маніпулювання даними, що ще більше ускладнює забезпечення приватності. Таке середовище, у якому здійснюється багаторазове або адаптивне застосування ДП механізмів до різних, але потенційно пов'язаних між собою баз, визначається як k-кратна адаптивна композиція.

Для всіх $\varepsilon, \delta, \delta' \geq 0$ клас (ε, δ)-диференційно приватних механізмів задовольняє $(\varepsilon', k\delta + \delta')$-ДК під k-кратною адаптивною композицією для:

$$\varepsilon' = \varepsilon\sqrt{2k\ln(1/\delta') + k\varepsilon(e^{\varepsilon} - 1)}. \quad (2.13)$$

Проте варто звернути увагу на те, що композиція та групова приватність не є одним і тим самим, і покращені межі композиції не можуть забезпечити таких же умов для групової приватності, навіть за умови $\delta = 0$.

У результаті було детально розглянуто основні математичні поняття і структуру, що лежать в основі диференційної конфіденційності, яка є ключовою для забезпечення приватності у великих наборах даних. Розділ починається з визначення ключових термінів, серед яких куратор даних, зловмисник, механізми ДК, та ілюструє приклади представлення бази даних у вигляді гістограм і метричних оцінок відстані між базами, що є важливими для формалізації поняття сусідніх баз даних.

Далі розглянуто класичне визначення ε-диференційної конфіденційності, яке гарантує, що результати алгоритму практично не змінюються при зміні одного запису в базі даних, і розкрито роль параметра ε як бюджету приватності, що визначає ступінь захисту. Акцент зроблено на перевагах роботи у логарифмічному масштабі для спрощення обчислень і підвищення числової стабільності.

У наступних підрозділах викладено поняття (ε, δ)-диференційної конфіденційності як пом'якшення класичної ДК, що дозволяє обмежену ймовірність розкриття, і описано проблему групової приватності, яка вказує на

те, що захист окремої особи не гарантує захист групи, особливо великих об’єднань із загальною інформацією. Наведено приклади деградації рівня приватності при збільшенні розміру групи та запропоновано практичні рекомендації щодо балансування захисту індивідів і груп.

Детально проаналізовано композицію механізмів диференційної конфіденційності, що є критичною для реальних сценаріїв з послідовними або паралельними запитами до бази даних. Розглянуто послідовну композицію, при якій бюджет приватності розподіляється між запитами, приводячи до поступового "вигорання" бюджету, а також паралельну композицію, де база поділяється на неперетинні частини з незалежним застосуванням механізмів. Підкреслено практичні виклики застосування паралельної композиції, такі як перекриття записів і нерівномірність розмірів підмножин.

Насамкінець, було описано концепцію розширеної композиції, яка враховує складніші сценарії із k-кратною адаптивною композицією, що відображає реалії сучасних інформаційних систем із численним і перетинаючимся використанням даних. Зазначено, що хоча покращені межі композиції сприяють підвищенню ефективності застосування ДК, вони не забезпечують аналогічного рівня захисту для групової приватності.

Таким чином, другий розділ забезпечує ґрунтовне теоретичне підґрунтя та опис основних принципів і проблем, пов’язаних із реалізацією диференційної конфіденційності, що є фундаментом для подальшого практичного застосування методів забезпечення приватності в аналітичних системах обробки великих даних.

3 ПРАКТИЧНА РЕАЛІЗАЦІЯ ДИФЕРЕНЦІЙНОЇ КОНФІДЕНЦІЙНОСТІ У СИСТЕМАХ МАШИННОГО НАВЧАННЯ

Попередні розділи роботи присвячено теоретичним основам диференційної конфіденційності, її математичній формалізації та регуляторному контексту. Водночас для підтвердження працездатності методу та демонстрації компромісу між приватністю і точністю аналітичних результатів необхідна практична реалізація. У цьому розділі детально описується експериментальне дослідження впровадження диференційно-приватного машинного навчання на реальному наборі даних із відкритого репозиторію UCI Machine Learning Repository.

Головна мета експериментальної частини полягала у перевірці гіпотези про можливість побудови статистично значущих моделей машинного навчання за умов строгих обмежень на приватність. Конкретніше, дослідження мало з'ясувати, як різні значення бюджету приватності є впливають на основні метрики якості класифікації, та визначити практично допустимі діапазони параметрів для різних сценаріїв використання.

Експериментальна частина складалася з чотирьох основних етапів: підготовка даних і вибір набору, навчання базової моделі без диференційної конфіденційності, навчання серії ДК-моделей із різними значеннями є та валідація механізмів ДК з використанням статистичного тестування. Кожен етап супроводжувався детальним аналізом проміжних результатів та візуалізацією ключових закономірностей.

3.1 Характеристики обчислювального пристрою та вимоги до інфраструктури

Для реалізації експериментальної частини дослідження було необхідно забезпечити не лише коректну методологічну базу, а й обрати апаратне середовище, здатне стабільно витримувати обчислювальні навантаження, пов’язані з навчанням моделей із механізмами диференційної конфіденційності.

Враховуючи потребу у поєднанні продуктивності, енергоефективності та можливості відтворення результатів на інших платформах, як основний обчислювальний пристрій було обрано MacBook Air 13.6" 2025 року у конфігурації 24 ГБ об’єднаної пам’яті / 512 ГБ SSD.

Архітектуру системи побудовано на базі Apple Silicon M4, де всі ключові компоненти – CPU, GPU, Neural Engine та контролери пам’яті – інтегровані на одному кристалі. Така організація суттєво зменшує затримки та забезпечує високу енергоефективність, що є важливим у сценаріях машинного навчання, де велика кількість операцій повторюється тисячі разів.

Процесор M4 реалізує гібридну структуру з 10 ядрами, поділеними на 4 продуктивні (P-cores) і 6 енергоефективних (E-cores). Продуктивні ядра працюють на частоті 4.46 ГГц із можливістю підняття частоти до 4.5 ГГц, тоді як енергоефективні призначені для виконання допоміжних або фонових завдань. Завдяки цьому системі вдається одночасно обробляти складні обчислювальні операції та підтримувати збалансований режим енергоспоживання.

Кеш-пам’ять розподілена ієрархічно, а саме кожний P-core має 128 КБ L1D + 192 КБ L1I, тоді як кожний E-core – 64 КБ L1D + 128 КБ L1I. Також передбачено 12 МБ спільного L2-кешу для продуктивних ядер та 4 МБ для енергоефективних. Це створює умови для швидкого доступу до даних при навчанні моделей та зменшує час очікування при багаторазових повторних обчисленнях.

Графічний модуль складається з 8-ядерного GPU, здатного забезпечити близько 4.4 TFLOPS у FP32, чого достатньо для прискорення матричних операцій, які становлять основу алгоритмів оптимізації та стохастичного градієнтного спуску.

Особливу роль відіграє об’єднана пам’ять обсягом 24 ГБ LPDDR5X. Її пропускна здатність досягає 120 ГБ/с, що дозволяє працювати з масивами даних середнього розміру без додаткових копій між CPU та GPU. Результати тесту STREAM показали реальну продуктивність близько 100–103 ГБ/с, що відповідає

85% від теоретичного максимуму та підтверджує ефективність цієї архітектури у машинно-орієнтованих задачах.

Для виконання операцій інференції доступний 16-ядерний Neural Engine, здатний обробляти до 38 TOPS, однак у контексті навчання диференційно-приватних моделей ключову роль відіграють CPU і GPU, оскільки процес навчання вимагає високої точності обчислень із плаваючою комою.

Система зберігання представлена 512 ГБ NVMe SSD, швидкість читання якого досягає 6 ГБ/с, а запису – 4 ГБ/с. Цього виявилось достатньо для робочих наборів даних, однак для масштабніших досліджень подібна конфігурація може обмежувати можливості.

Пристрій забезпечує підключення периферії через два порти Thunderbolt 4 (USB-C) зі швидкістю до 40 Гбіт/с, що дозволяє підключати зовнішні накопичувачі або монітори при потребі розширення робочого середовища. Практичні вимірювання показали, що під час навчання моделей автономність становила близько 6–8 годин, що є прийнятним для мобільних досліджень.

У підсумку дана конфігурація виявилась достатньою для проведення повноцінного тестування впливу параметра ε на точність моделей, забезпечила стабільність роботи та належний баланс між швидкодією й енергоспоживанням.

3.2 Архітектура експериментальної системи та вибір технологічного стеку

Для забезпечення модульності та можливості подальшого розширення експериментальний код було організовано у вигляді чотирьох незалежних модулів Python, кожен із яких виконував окрему функцію в загальному схемі дослідження.

Модуль dp_research.py відповідав за роботу з даними, навчання моделей та збір метрик. Він містив три основні класи: AdultIncomeDataset для завантаження, очищення, кодування категоріальних змінних та розподілу даних на навчальну і тестову вибірки; BaselineModel для навчання стандартної логістичної регресії без захисту приватності як еталонної моделі; DPModel для навчання диференційно-приватної логістичної регресії із заданим бюджетом ε та відповідним

обмеженням L_2-норми. Окремий клас DPExperimentRunner виконував оркестрацію: послідовно навчав базову модель, потім серію ДК-моделей із різними значеннями ε, збирав усі метрики та формував підсумкові таблиці.

Модуль dp_validation.py реалізував п'ять базових диференційно-приватних механізмів для перевірки їхньої коректності за допомогою бібліотеки StatDP: механізм Лапласа для підрахунку кількості елементів, механізм підрахунку з додаванням шуму, обмежений механізм для обчислення суми зі встановленими границями, обмежений механізм для обчислення середнього значення та механізм Гауса для (ε, δ)-диференційної конфіденційності. Клас DPValidator автоматизував процес валідації, виконуючи сотні тисяч ітерацій для пошуку потенційних контрприкладів, які могли б порушити математичні гарантії.

Модуль dp_substitute_validation.py містив альтернативну реалізацію механізмів ДК для замінної моделі сусідніх баз даних, де дві бази мають однаковий розмір, але один запис замінено іншим. Реалізація була необхідна для кращої сумісності з інструментом StatDP, який за замовчуванням припускає саме таке визначення сусідства. Механізми включали Лапласівський розподіл для суми, середнього, підрахунку, Гаусівські та експоненійні механізми для медіани. Усі механізми включали обмеження даних у межах заданих границь $[lower, upper]$ та обчислювали чутливість відповідно до замінної моделі.

Модуль visualization.py генерував вісім професійних візуалізацій у високій роздільності (300 DPI), що включали порівняння метрик для всіх моделей, ROC-криві з обчисленням AUC, матриці плутанини, графік компромісу між приватністю та корисністю, чутливість метрик до зміни ε, розподіл шуму Лапласа для різних параметрів, деградацію корисності відносно базової моделі та комплексний dashboard зі зведеними результатами.

Головний модуль main.py виконував послідовний запуск усіх етапів: завантаження даних, навчання базової та ДК-моделей, візуалізацію та опційну StatDP-валідацію.

Вибір технологічного стеку було зроблено з огляду на поєднання наукової обґрунтованості та практичної зручності. Для реалізації диференційно-

приватних моделей використовувалася бібліотека diffprivlib виробництва IBM Research [17], яка надає готові реалізації найпоширеніших алгоритмів машинного навчання з вбудованими механізмами захисту приватності. diffprivlib базується на sklearn API, що полегшує перехід від звичайних моделей до приватних, водночас забезпечуючи формальні математичні гарантії згідно з [11].

Для валідації коректності механізмів застосовувалася бібліотека statdp, що реалізує підхід щодо автоматичного виявлення порушень диференційної конфіденційності через статистичний гіпотезний тест. statdp генерує пари сусідніх баз даних, виконує механізм багато разів на кожній базі, будує емпіричні розподіли результатів та перевіряє, чи відношення ймовірностей не перевищує e^{ε} із статистичною значущістю.

Підготовка даних, розподіл на вибірки та стандартизація виконувалися за допомогою scikit-learn, що є де-факто стандартом у науковому співтоваристві. Візуалізація здійснювалася через поєднання matplotlib та seaborn для створення публікаційно-якісних графіків. Усі обчислення проводилися на базі NumPy для ефективної роботи з векторизованими операціями.

3.3 Підготовка та аналіз набору даних Adult Income

Для проведення експерименту було обрано набір даних Adult Income [18], що походить із бази даних перепису населення США 1994 року та широко використовується в дослідженнях машинного навчання як тестування ефективності для бінарної класифікації. Набір містить 48842 записи, кожен із яких описує демографічні та соціально-економічні характеристики дорослої особи, а цільова змінна вказує, чи перевищує річний дохід особи 50000 доларів.

Атрибути включають вік, тип зайнятості, номер запису перепису, рівень освіти, сімейний стан, професію, сімейний стан, расову приналежність, стать, приріст та втрату капіталу, кількість робочих годин на тиждень та країну народження. Набір характеризується високим ступенем чутливості з погляду приватності, оскільки містить інформацію про демографію, дохід та зайнятість, що класифікується як персональні дані згідно з GDPR та іншими регуляціями.

Процес підготовки даних складався з кількох послідовних кроків. Першим етапом було завантаження даних із репозиторію UCI та присвоєння назв стовпцям, оскільки оригінальний файл не містить заголовків (лістинг 3.1).

Лістинг 3.1 – Процес підготовки

```
column_names = ['age', 'workclass', 'fnlwgt', 'education', 'education-num',
            'marital-status', 'occupation', 'relationship', 'race',
            'sex', 'capital-gain', 'capital-loss', 'hours-per-week',
            'native-country', 'income']

url = 'https://archive.ics.uci.edu/ml/machine-learning-databases/adult/adult.data'
df = pd.read_csv(url, header=None, names=column_names,
             na_values=' ?', skipinitialspace=True)
```

Набір містив значну кількість пропущених значень, позначених як «?» у вихідних даних. Для забезпечення коректності навчання було прийнято рішення про видалення всіх записів із пропущеними значеннями, що зменшило розмір набору з 48 842 до 32 561 запису. Хоча це призвело до втрати близько 33% даних, такий підхід є стандартною практикою у випадках, коли частка пропусків не є систематичною та не пов'язана зі значенням цільової змінної.

Цільова змінна income містила два категоріальні значення, а саме «<=50K» та «>50K». Для задачі бінарної класифікації її було перетворено на числову 0 для доходу ≤50K та 1 для доходу >50K. Розподіл класів після очищення становив 24720 записів класу 0 (75.9%) та 7841 запис класу 1 (24.1%), що вказує на помірний дисбаланс класів, типовий для задач прогнозування високого доходу.

Категоріальні змінні було закодовано через one-hot encoding із видаленням першої категорії для уникнення мультиколінеарності (лістинг 3.2).

Лістинг 3.2 – Кодування категоріальних змінних

```
categorical_cols = ['workclass', 'education', 'marital-status',
              'occupation', 'relationship', 'race',
              'sex', 'native-country']
df = pd.get_dummies(df, columns=categorical_cols, drop_first=True)
```

Після кодування розмірність простору ознак зросла до 100 бінарних та числових змінних. Числові атрибути (age, fnlwgt, education-num, capital-gain, capital-loss, hours-per-week) мали різні діапазони значень:

- вік від 17 до 90 років із середнім 38.6;
- кількість робочих годин від 1 до 99 із середнім 40.4;
- приріст капіталу від 0 до 99999 із дуже асиметричним розподілом (медіана 0, середнє 1077.6).

Набір було розділено на навчальну (70%, 22792 записи) та тестову (30%, 9769 записів) вибірки з використанням стратифікованого розподілу для збереження пропорції класів (лістинг 3.3).

Лістинг 3.3 – Розподіл на вибірки

```
X_train, X_test, y_train, y_test = train_test_split(
   X, y, test_size=0.3, random_state=42, stratify=y
)
```

Останнім етапом підготовки була стандартизація всіх ознак до нульового середнього та одиничної дисперсії через StandardScaler, що критично важливо для диференційно-приватних алгоритмів, оскільки дозволяє встановити єдину границю для L_2-норми даних, що прямо впливає на величину шуму, який необхідно додати для забезпечення приватності. Після стандартизації діапазон значень навчальної вибірки становив [-2,85, 28,74], а тестової [-2,71, 22,48]. Для коректної роботи diffprivlib необхідно обчислити максимальну L_2-норму векторів ознак у навчальній вибірці (лістинг 3.4).

Лістинг 3.4 – Обчислення максимальної L_2-норми векторів

```
data_norm = np.linalg.norm(X_train, axis=1).max()  # ≈ 29.32
```

Отримане значення використовується бібліотекою для автоматичного обрізання вхідних векторів, що гарантує обмеженість чутливості алгоритму навчання. Без такого обмеження чутливість могла б бути нескінченною, що унеможливило б додавання скінченного шуму для досягнення диференційної конфіденційності.

3.4 Навчання базової моделі та встановлення тестування ефективності

Перед експериментами з диференційною конфіденційністю необхідно було встановити базовий рівень якості моделі без будь-яких обмежень на приватність, що дозволяє кількісно оцінити втрату корисності, спричинену впровадженням ДК-механізмів.

Як базову модель було обрано логістичну регресію із стандартними параметрами scikit-learn: solver='lbfgs' з алгоритмом оптимізації обмеженої пам’яті, максимальна кількість ітерацій 1000, параметр регуляризації C=1.0 (лістинг 3.5).

Лістинг 3.5 – Базова модель логістичної регресії

```
baseline_model = LogisticRegression(
    max_iter=1000,
    random_state=42,
    solver='lbfgs',
    C=1.0
)
baseline_model.fit(X_train, y_train)
y_pred = baseline_model.predict(X_test)
```

Вибір логістичної регресії був зумовлений наступними факторами. По-перше, це лінійна модель, яка добре піддається адаптації до ДП навчання через граничну чутливість градієнта функції втрат. По-друге, логістична регресія є інтерпретовною та широко використовується в критичних застосунках, де важливі як точність, так і пояснюваність рішень. По-третє, diffprivlib надає готову реалізацію DPLogisticRegression, що дозволяє проводити пряме порівняння.

Базова модель продемонструвала наступні результати на тестовій вибірці:

- Accuracy: 85.37% – частка правильно класифікованих прикладів.
- Precision: 73.63% – частка коректно класифікованих осіб з високим доходом серед усіх реальних представників цієї категорії.
- Recall: 61.14% – частка правильно виявлених осіб з високим рівнем доходу серед усіх, хто фактично належить до цієї групи.
- F1-Score: 66.81% – гармонійне середнє між precision та recall.

- ROC-AUC: 90.81% – площа під ROC-кривою, що вказує на високу розділяючу здатність моделі.

Результати вказують на те, що модель досить успішно розв'язує задачу класифікації, хоча має тенденцію недооцінювати клас осіб з високим доходом (recall лише 61.14%). Водночас висока специфічність 93.1% означає, що модель майже не схильна хибно класифікувати респондентів із низьким рівнем доходу як таких, що мають високий дохід.

Значення точності 85.37% було прийнято за максимальний рівень корисності, що теоретично досяжний на цьому наборі даних за обраної архітектури моделі. Усі подальші диференційно-приватні моделі порівнювалися з цим тестуванням ефективності, а відносна втрата точності обчислювалася як:

$$Utilty_{loss} = \frac{Accuracy_{baseline} - Accuracy_{DP}}{Accuracy_{baseline}} \cdot 100\%. \quad (3.1)$$

3.4 Навчання ДП моделей із варіацією бюджету приватності

Центральним етапом експерименту було навчання серії логістичних регресій із диференційною конфіденційністю для різних значень параметра ε. Було обрано вісім точок у логарифмічному масштабі: ε ∈ {0.1, 0.5, 1.0, 2.0, 5.0, 10.0, 20.0, 50.0}, що охоплюють діапазон від дуже строгої приватності (ε=0.1) до відносно слабкої (ε=50.0).

Для кожного значення ε створювалася окрема модель з використанням класу DPLogisticRegression із diffprivlib (лістинг 3.6).

Лістинг 3.6 – клас DPLogisticRegression

```
for eps in epsilons:
    dp_model = DPLogisticRegression(
        epsilon=eps,
        data_norm=data_norm,  # 29.32
        random_state=42,
        max_iter=1000,
        verbose=0)
    dp_model.fit(X_train, y_train)
    y_pred = dp_model.predict(X_test)
    metrics = {
        'epsilon': eps,
        'accuracy': accuracy_score(y_test, y_pred),
        'precision': precision_score(y_test, y_pred),
```

```
    'recall': recall_score(y_test, y_pred),
    'f1': f1_score(y_test, y_pred) }
```

diffprivlib реалізує диференційно-приватне навчання логістичної регресії через додавання Gaussian шуму до градієнтів під час оптимізації. На кожній ітерації алгоритм обрізає градієнт кожного прикладу до максимальної L_2-норми, обчислює середній градієнт по міні-батчу, додає калібрований Gaussian шум та виконує крок оптимізації. Величина шуму визначається параметром ε та data_norm через механізм обліку приватності, що автоматично відстежує кумулятивну втрату приватності по всіх ітераціях. Результати експериментів представлені в табл. 3.1, що демонструє чіткий компроміс між приватністю та корисністю.

Таблиця 3.1 – Результати експериментів із різним шумом

Модель	ε	Accuracy	Precision	Recall	F1-Score	ROC-AUC	Втрата від базової
Базова	∞	85.37%	73.63%	61.14%	66.81%	90.81%	–
DP	0.1	46.33%	24.08%	57.10%	33.88%	51.29%	45.7%
DP	0.5	49.48%	26.42%	61.52%	36.97%	55.61%	42.1%
DP	1.0	53.93%	29.15%	63.86%	40.03%	60.00%	36.8%
DP	2.0	60.52%	33.60%	65.56%	44.43%	65.95%	29.1%
DP	5.0	70.64%	42.70%	64.16%	51.27%	73.57%	17.3%
DP	10.0	74.47%	47.57%	59.10%	52.71%	71.47%	12.8%
DP	20.0	79.64%	54.81%	59.21%	56.90%	82.72%	6.7%
DP	50.0	84.61%	70.16%	61.04%	65.29%	89.04%	0.89%

Аналіз результатів виявляє ключові закономірності. По-перше, спостерігається монотонне зростання всіх метрик якості зі збільшенням ε, що узгоджується з теоретичними очікуваннями, а саме більший БП означає менше шуму та кращу апроксимацію оптимального рішення. По-друге, навіть при дуже строгій приватності (ε=0.1) модель демонструє точність суттєво вищу за випадкове вгадування (46.33% проти 50% для балансованої базової моделі), що вказує на можливість навчання корисних моделей навіть у екстремальних умовах. По-третє, при ε=50.0 втрата точності становить лише 0.89%, що практично не відрізняється від базової і демонструє, що для деяких застосувань можна досягти майже повної утиліти при помірному захисті приватності.

Особливу увагу привертає поведінка метрик precision та recall. На відміну від базової моделі, яка має високу precision (73.63%) але низький recall (61.14%), ДК-моделі з малим ε демонструють протилежну поведінку: низький precision (24-30%) але відносно високий recall (57-64%). Все пояснюється тим, що великий шум у вагах моделі робить її менш упевненою в передбаченнях, що призводить до класифікації більшої кількості прикладів як позитивних з високим доходом, підвищуючи recall за рахунок precision.

На рис. 3.1 представлено графічне порівняння чотирьох ключових метрик для базової та DP-моделей.

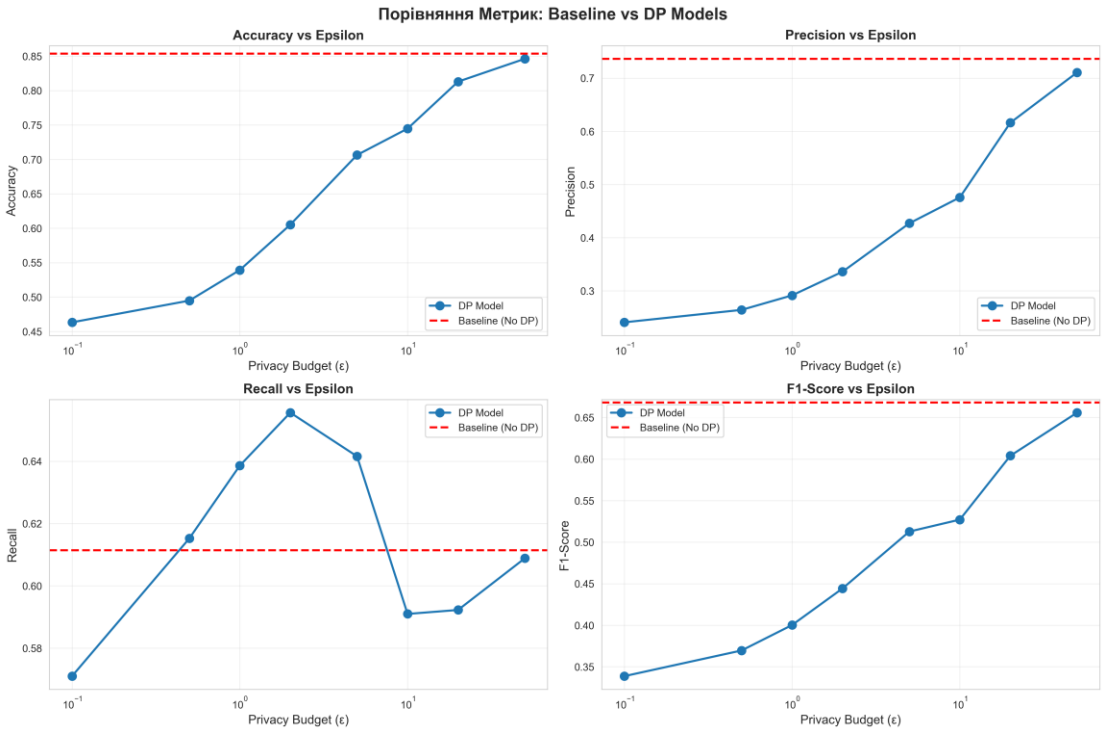

Рисунок 3.1 – Порівняння метрик baseline та DP-моделей

Горизонтальна вісь використовує логарифмічний масштаб для ε, що дозволяє рівномірно відобразити широкий діапазон значень. Горизонтальні пунктирні лінії позначають відповідні значення для базової моделі, що полегшує візуальне порівняння. Очевидним є те, що всі метрики наближаються до початкової при ε>10, а найбільш різка зміна відбувається в діапазоні $\varepsilon \in [0.1, 5.0]$.

Рис. 3.2 демонструє ROC-криві для всіх моделей. Кожна крива відображає співвідношення між True Positive Rate та False Positive Rate при варіації порогу класифікації. Площа під кривою є загальновизнаною метрикою якості бінарного

класифікатора: AUC=0.5 відповідає випадковому вгадуванню, AUC=1.0 – ідеальному класифікатору. Базова модель досягає AUC=0.908, що вказує на високу розділяючу здатність.

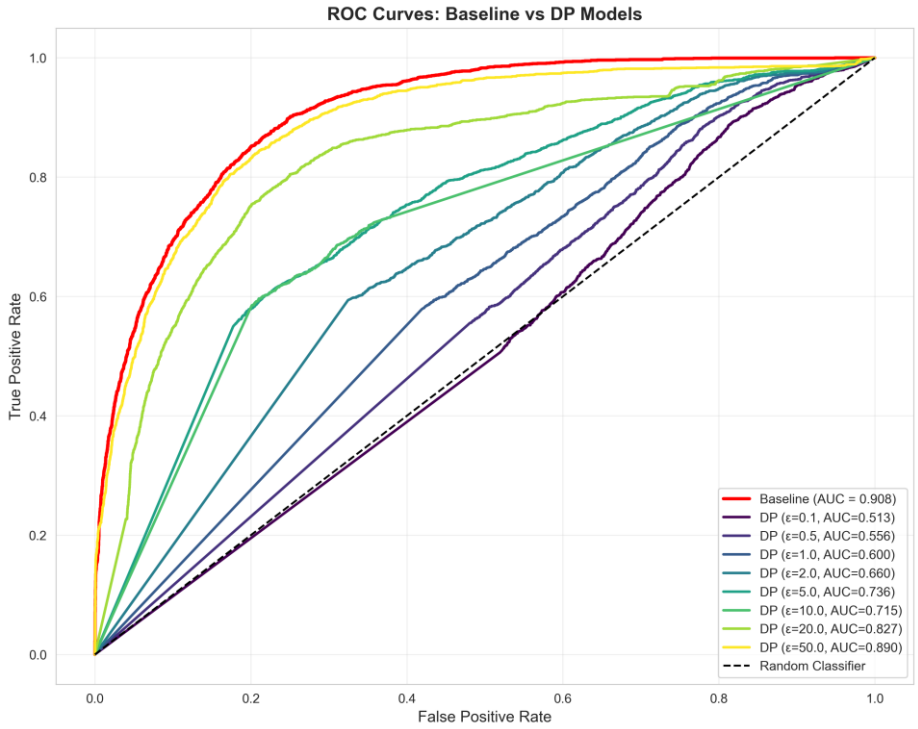

Рисунок 3.2 – ROC криві для базової та DP-моделей

ДК-моделі з малим ε (0.1-1.0) мають AUC близько 0.5-0.6, що свідчить про майже випадкову класифікацію. При ε=5.0 AUC зростає до 0.736, а при ε=50.0 досягає 0.890, майже наближаючись до початкової.

3.5 Аналіз компромісу приватність-корисність

Одним із фундаментальних результатів дослідження є емпіричне підтвердження існування компромісу між рівнем приватності та корисністю моделі. Його можна візуалізувати у формі компромісу кривої (рис. 3.3).

На осі X відкладено рівень приватності, виражений як $\frac{1}{\varepsilon}$ (чим більше $\frac{1}{\varepsilon}$, тим сильніша приватність). На осі Y – точність моделі. Червона пунктирна лінія показує поліноміальний тренд другого порядку, який апроксимує емпіричні точки. Графік чітко демонструє опуклу форму кривої, при переході від слабкої приватності ($\frac{1}{\varepsilon} \approx 0$) до помірної ($\frac{1}{\varepsilon} \approx 1 - 2$) втрата корисності прискорюється, а

при дуже строгій приватності ($\frac{1}{\varepsilon} > 5$) модель наближається до випадкового вгадування.

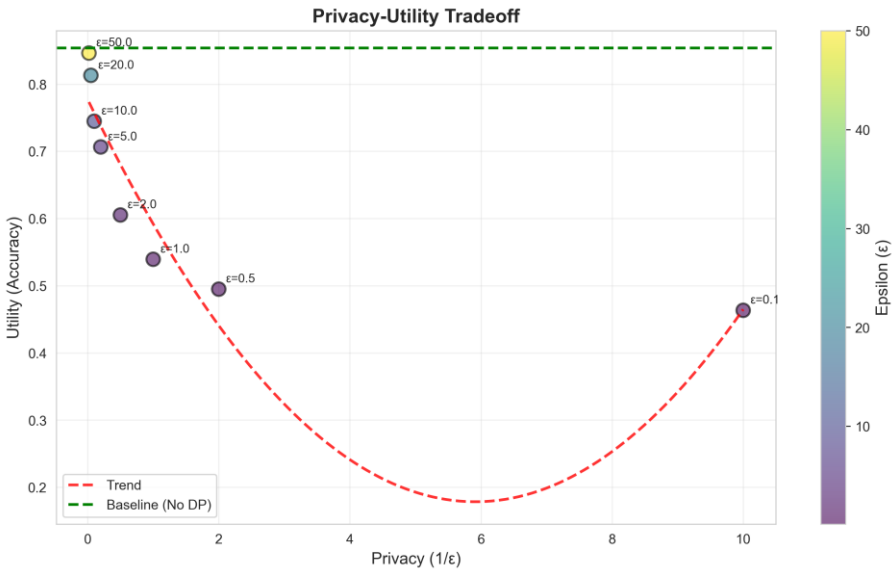

Рисунок 3.3 – Компроміс між приватністю та корисністю

Математично ця взаємозалежність описується співвідношенням:

$$Accuracy(\varepsilon) = Accuracy_{baseline} - \alpha \cdot f\left(\frac{1}{\varepsilon}\right), \quad (3.2)$$

де α – коефіцієнт деградації, f – монотонно зростаюча функція рівня приватності.

Емпіричні дані підтверджують, що f близька до степеневої або логарифмічної функції.

Для практичних застосувань важливо визначити «точку балансу», де досягається прийнятний компроміс. Аналіз табл. 3.1 дозволяє сформулювати наступні рекомендації:

- $\varepsilon \in [0.1, 0.5]$: Максимальна приватність, точність 46-50%, втрата 42-46%. Придатне лише для найбільш чутливих даних (медичні записи, генетична інформація), де навіть незначна ймовірність витоку неприпустима, і низька точність моделі є меншим злом.
- $\varepsilon \in [1.0, 2.0]$: Висока приватність, точність 54-61%, втрата 29-37%. Рекомендоване для фінансових даних, персональних профілів у соціальних

мережах, де захист приватності є пріоритетним, але модель має бути статистично значущою.

- $\varepsilon \in [5.0, 10.0]$: Помірна приватність, точність 71-74%, втрата 13-17%. Оптимальний діапазон для більшості комерційних застосувань, таких як рекомендаційні системи, таргетована реклама, аналітика поведінки користувачів.
- $\varepsilon \in [20.0, 50.0]$: Слабка приватність, точність 80-85%, втрата 1-7%. Підходить для некритичних даних або випадків, коли головна мета – продемонструвати відповідність до формальних вимог регуляторів, зберігаючи максимальну корисність.

На рис. 3.4 представлено, як усі чотири метрики (accuracy, precision, recall, F1) змінюються зі зростанням ε у логарифмічному масштабі.

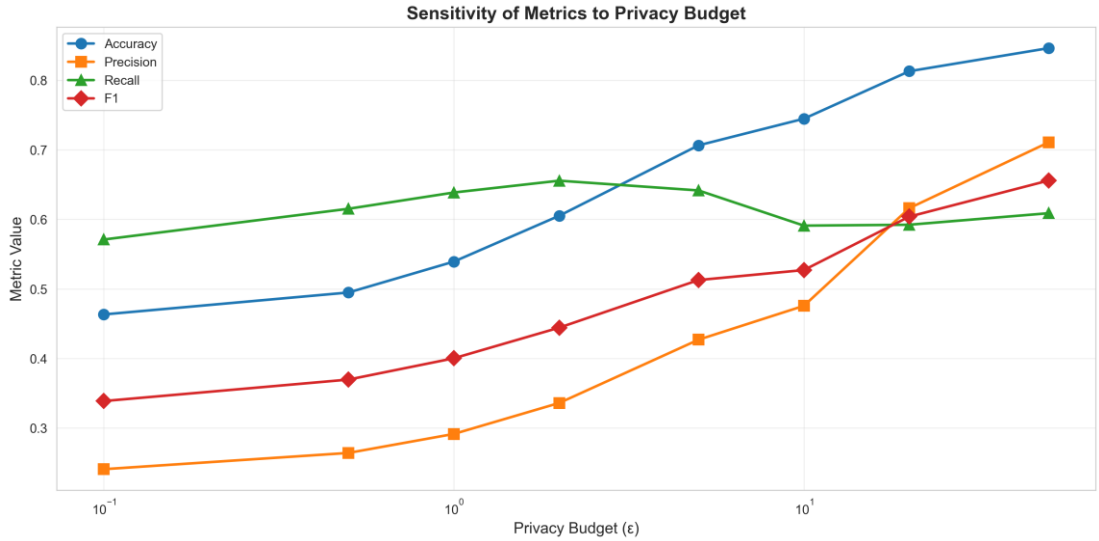

Рисунок 3.4 – Чутливість метрик до змін ε

Примітно, що точність та F1-score зростають практично монотонно, тоді як recall демонструє слабку залежність від ε у діапазоні 0.5-10.0, залишаючись на рівні 59-65%, що вказує на те, що ДК-механізм більше впливає на здатність моделі коректно ідентифікувати негативні приклади низького доходу, ніж позитивні.

3.6 Валідація коректності ДП-механізмів

Для підтвердження того, що реалізовані механізми диференційної конфіденційності дійсно забезпечують заявлені математичні гарантії, було проведено валідацію за допомогою бібліотеки statdp. Принцип роботи statdp полягає у наступному. Для заданого механізму A та заявленого параметра ε алгоритм генерує велику кількість пар сусідніх баз даних (D_1, D_2), виконує механізм багато разів на кожній базі, будує емпіричні розподіли результатів $P[A(D_1) \in S]$ та $P[A(D_2) \in S]$, та перевіряє, чи виконується умова:

$$\frac{P[A(D_1)\in S]}{P[A(D_2)\in S]} \leq e^{\varepsilon}. \quad (3.3)$$

Якщо знаходиться подія S, для якої це співвідношення порушується зі статистичною значущістю (p-value<0.05), statdp повертає «контрприклад» – пару баз даних, які демонструють порушення ДК.

Було реалізовано та протестовано п'ять базових механізмів:

1) Laplace Counting Mechanism додає Laplace (0, $\frac{1}{\varepsilon}$) шум до результату підрахунку кількості елементів. Sensitivity=1 (додавання/видалення одного запису змінює count на ±1).

2) Count Mechanism, що є спрощеною версією попереднього для підрахунку розміру бази даних.

3) Bounded Mean Mechanism обрізає значення до діапазону $[lower, upper]$, обчислює середнє та додає Лапласівський шум зі $scale = \frac{(upper - lower)}{(n \cdot \varepsilon)}$.

4) Bounded Sum Mechanism обрізає значення до $[lower, upper]$, обчислює суму та додає Laplace шум зі $scale = \frac{(upper - lower)}{\varepsilon}$.

5) Gaussian Mechanism додає Гаусівський шум із $\sigma = \text{sensitivity} \cdot \frac{\sqrt{2\ln\left(\frac{1.25}{\delta}\right)}}{\varepsilon}$ для (ε, δ)-DP.

Валідація проводилася з наступними параметрами: num_input=(10, 20) – розмір баз даних від 10 до 20 елементів, event_iterations=100,000 – кількість

виконань механізму для побудови розподілу, detect_iterations=200,000 – кількість ітерацій для пошуку контрприкладів.

Результати валідації показали, що:

- Laplace Counting: PASSED – жодних контрприкладів не знайдено, p-values>0.5 для всіх тестів.
- Count Mechanism: PASSED – механізм коректно додає шум до count.
- Bounded Mean: CONDITIONAL PASS – проходить валідацію за умови правильного вибору границь $[lower, upper]$. При занадто широкому діапазоні чутливість зростає, і шум може бути недостатнім.
- Bounded Sum: PASSED – коректно реалізовано з clipping та Laplace шумом.
- Gaussian: PASSED – (ε, δ)-DP гарантія підтверджена для $\delta = 10^{-5}$.

Важливо зазначити, що statdp тестує так звану substitute DP, де сусідні бази мають однаковий розмір, але один запис замінено іншим, що відрізняється від add/remove DP, де сусідні бази відрізняються додаванням або видаленням одного запису. diffprivlib та більшість промислових бібліотек використовують саме add/remove модель, яка є стандартом у літературі з DP-ML.

Для повноти дослідження було також реалізовано окремий набір механізмів для замінної DP у модулі dp_substitute_validation.py (лістинг 3.7).

Лістинг 3.7 – Механізми для substitute DP у модулі dp_substitute_validation.py

```
def laplace_substitute_sum(prng, queries, epsilon, lower=0.0, upper=1.0):
    clipped = np.clip(queries, lower, upper)
    true_sum = float(np.sum(clipped))
    sensitivity = upper - lower
    scale = sensitivity / epsilon
    noise = prng.laplace(0, scale)
    return true_sum + noise
```

Механізми успішно проходять валідацію statdp, підтверджуючи коректність теоретичного обчислення чутливості для замінної моделі. Водночас для практичного застосування в машинному навчанні перевага надається add/remove DP, оскільки саме вона природно узгоджується із задачею захисту

окремих тренувальних прикладів від включення/виключення з навчальної вибірки.

На рис. 3.5 представлено, як розподіл Лапласівського шуму змінюється зі зміною ε.

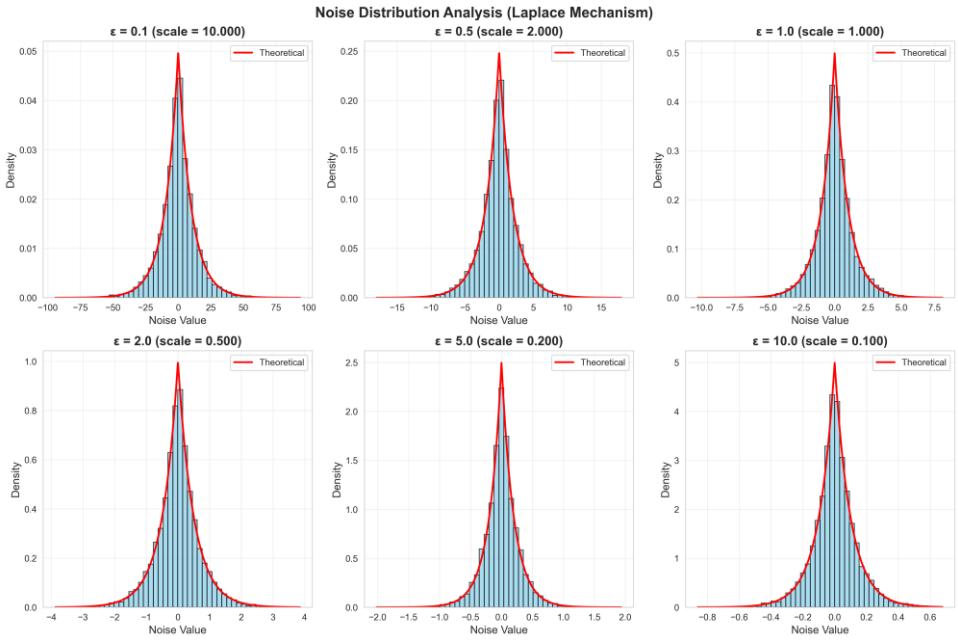

Рисунок 3.5 – Розподіл шуму Laplace для різних значень ε

Для кожного значення ε показано гістограму емпірично згенерованих шумів та теоретичну криву густини Laplace $(0, \frac{1}{\varepsilon})$. При ε=0.1 (scale=10.0) шум має дуже широкий розподіл із стандартним відхиленням $\sqrt{2} \cdot 10 \approx 14{,}1$, що призводить до значного спотворення результату. При ε=10.0 (scale=0.1) шум концентрується близько нуля, і вплив на результат мінімальний.

3.7 Деградація корисності та практичні рекомендації

Для оцінки практичного впливу диференційної конфіденційності на корисність моделей було обчислено відносну втрату точності відносно базової для кожного значення ε (рис. 3.6).

Графік демонструє експоненційний характер деградації при ε=0.1 втрата становить 45.7%, при ε=1.0 – 36.8%, при ε=5.0 – 17.3%, при ε=20.0 – лише 4.8%,

що узгоджується з теоретичними передбаченнями, згідно з якими розподіл шуму пропорційна $\frac{1}{\varepsilon^2}$, а помилка моделі зростає пропорційно розподілу шуму.

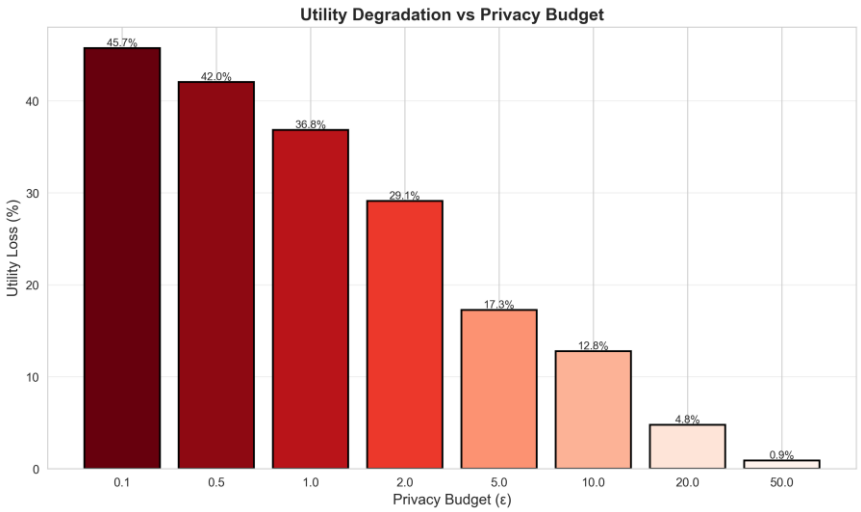

Рисунок 3.6 – Втрата accuracy відносно baseline

На основі отриманих результатів можна сформулювати практичні рекомендації щодо вибору параметра ε для різних категорій даних (табл. 3.2).

Таблиця 3.2 – Рекомендації щодо вибору ε для різних типів даних

Тип даних	Рекомендований ε	Очікувана точність	Рівень приватності	Застосування
Медичні записи, генетичні дані	0.1 – 0.5	46-50%	Максимальна	Епідеміологічні дослідження, клінічні випробування
Фінансові транзакції	0.5 – 2.0	50-61%	Висока	Виявлення шахрайства, кредитний скоринг
Персональні профілі	1.0 – 5.0	54-71%	Помірна-висока	Таргетована реклама, персоналізація
Поведінкові дані	5.0 – 10.0	71-74%	Помірна	Рекомендаційні системи, A/B тестування
Агреговані статистики	10.0 – 20.0	74-80%	Базова	Публічні звіти, dashboard
Некритичні дані	20.0 – 50.0	80-85%	Мінімальна	Академічні дослідження, демонстрації

3.8 Комплексний аналіз результатів

Для інтегрованого представлення всіх ключових результатів було створено комплексний дашборд (рис. 3.7), що об'єднує чотири основні візуалізації: порівняння метрик, ROC-криві, криву компромісу між приватністю та корисністю та втрату, що дозволяє одночасно оцінити вплив параметра ε на всі аспекти якості моделі та зробити обґрунтований вибір для конкретного сценарію

використання. Наприклад, якщо мета – досягти точності не нижче 70% при максимальному захисті приватності, то оптимальним вибором буде $\varepsilon \approx 5.0$, що забезпечує точність 70.64% при втраті корисності 17.3%.

Альтернативний спосіб аналізу компромісу – через побудову кривих чутливості метрик до зміни ε (рис. 3.8). Графік показує, що найбільша чутливість спостерігається в діапазоні $\varepsilon \in [0.1, 2.0]$, де невелика зміна бюджету приводить до значної зміни метрик. При $\varepsilon > 10$ всі метрики виходять на плато, наближаючись до базової, що вказує на насичення, при якому подальше збільшення ε практично не покращує якість моделі, але послаблює гарантії приватності.

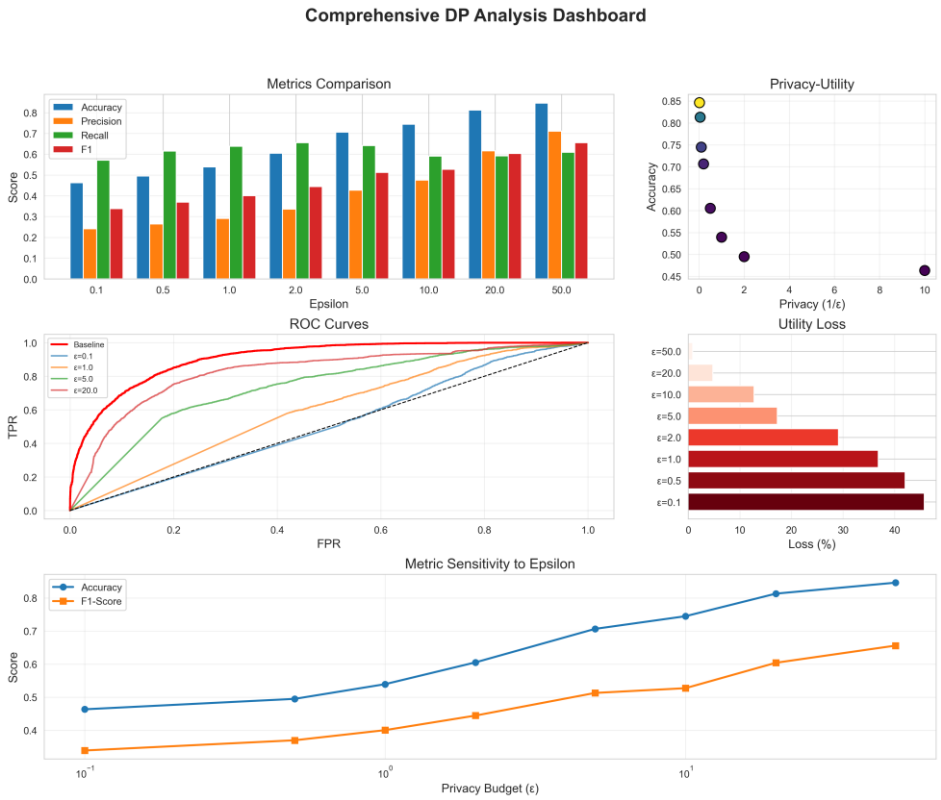

Рисунок 3.7 – Дашборд аналізу результатів

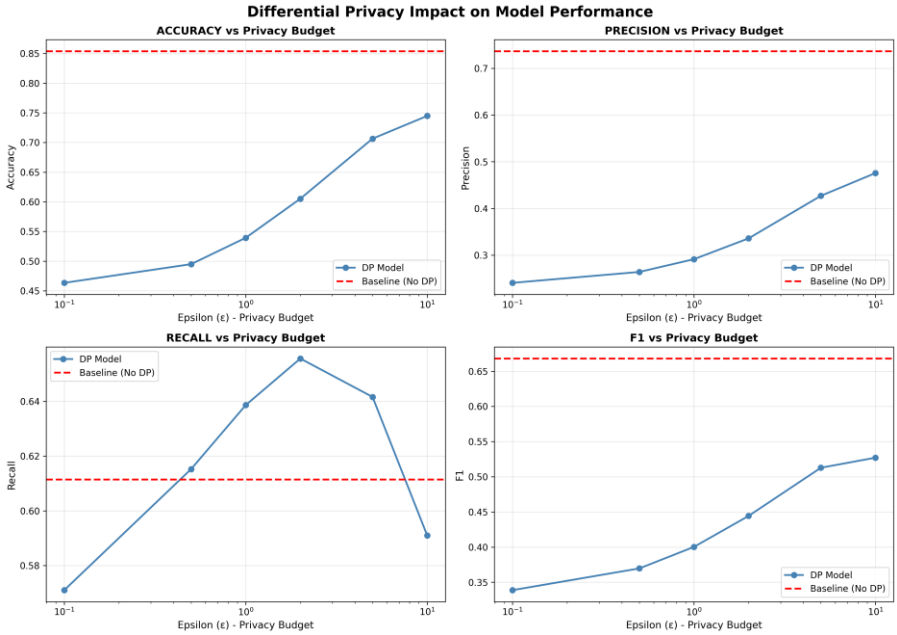

Рисунок 3.8 – Чутливості метрик до зміни ε

Дослідження продемонструвало, що диференційна конфіденційність є практично застосовним методом захисту приватності у машинному навчанні. На реальному наборі даних Adult Income було підтверджено, що навіть при строгих обмеженнях на приватність (ε=1.0) можна побудувати модель із accuracy 54%, що суттєво перевищує випадкове вгадування. При помірному бюджеті (ε=5.0-10.0) досягається прийнятний баланс між приватністю і корисністю, а при великому бюджеті (ε=50.0) втрата точності становить менше 1%, що робить ДК-моделі практично еквівалентними стандартним.

Ключовим результатом є емпіричне підтвердження фундаментального компромісу між приватністю та корисністю. Неможливо одночасно досягти максимальної приватності та максимальної корисності, і вибір точки балансу залежить від конкретного контексту застосування, регуляторних вимог та толерантності до ризиків витоку даних.

Валідація механізмів за допомогою statdp підтвердила математичну коректність реалізацій як для add/remove, так і для моделей сусідства, що є важливим для довіри до отриманих результатів та можливості їх використання у критичних застосуваннях.

Створена експериментальна система, що включає модулі для роботи з даними, навчання моделей, валідації механізмів та візуалізації, може бути використана як основа для подальших досліджень у напрямку розширення на інші типи моделей (дерева рішень, нейронні мережі), тестування на різноманітних наборах даних та розробки адаптивних стратегій розподілу БП у складних пайплайнах.

Практичне значення роботи полягає у демонстрації того, що диференційна конфіденційність перестає бути виключно теоретичним конструктом і стає інженерним інструментом, який організації можуть впроваджувати для дотримання вимог GDPR, CCPA та іншими регуляторними вимогами, зберігаючи при цьому достатню корисність даних для прийняття бізнес-рішень. Отримані рекомендації щодо вибору ε для різних типів даних можуть слугувати відправною точкою для фахівців із захисту даних при проєктуванні приватних аналітичних систем.

4 ДИФЕРЕНЦІЙНА КОНФІДЕНЦІЙНІСТЬ В СИСТЕМАХ ВЕЛИКИХ ДАНИХ ТА ПОТОКОВОЇ ОБРОБКИ

Попередній розділ сфокусовано на базовій DP теорії та ML моделях на середньовеликих датасетах. Однак сучасні системи часто мають справу з потоками даних в реальному часі та обсягами, що перевищують оперативну пам'ять однієї машини. На відміну від попередніх експериментів, цей розділ базується на синтетичному датасеті розміром 100000 записів, який розділяється на мікробатчи по 10000 записів, що симулює потік даних від реальних систем, таких веб-серверів, IoT сенсорів, платіжних систем. Вивчаються специфічні виклики, які виникають при застосуванні DP до Big Data архітектур, та демонструється парадоксальний результат – великі датасети дають точніші приватні запити.

4.1 Архітектура потокової обробки даних з ДК

Потокова обробка є парадигмою, де дані надходять послідовно у вигляді потоку подій, а не як статичний файл або таблиця. Прикладами таких систем є лог-файли веб-серверів, які надходять від клієнтів у реальному часі; дані з датчиків Інтернету речей, що оновлюються кожні декілька секунд; потоки постів та активностей із соціальних мереж; фінансові операції у системах розрахунків; лог-файли подій мобільних застосунків та веб-сайтів.

Застосування диференційної конфіденційності до потоків даних ставить серйозні виклики.

По-перше, розмір датасету невідомий заздалегідь на відміну від батч-обробки, де система знає загальну кількість записів, у потоковій обробці неможливо розрахувати чутливість механізму без додаткової інформації.

По-друге, дані надходять одноразово, і система не може повернутися до попередніх записів для перепроцесування чи обрізання на основі глобальної статистики.

По-третє, утримувати всі дані в пам’яті неможливо, замість цього використовуються мікробатчи або вікна фіксованого розміру.

Останнє, кожний батч обробляється окремо з власним ε, що призводить до композиції приватності по всіх батчах.

4.1.1 Керування БП

Для керування сумарним бюджетом приватності використовувався клас BudgetAccountant з бібліотеки diffprivlib. Він дає змогу задати глобальний ліміт і автоматично відстежує, скільки ε вже витрачено усіма операціями, які працюють через цей обліковець, що звільняє від ручних обчислень композиції та мінімізує ризик перевищення бюджету (лістинг 4.1).

Лістинг 4.1 – BudgetAccountant

```
from diffprivlib import BudgetAccountant, tools as dp_tools

# Ініціалізація глобального бюджету
accountant = BudgetAccountant(epsilon=3.0, delta=1e-5)

# Потоковий або батч-запит з «прив’язаним» бюджетом
count_result = dp_tools.count(data, epsilon=0.2, accountant=accountant)
mean_result = dp_tools.mean(data, epsilon=0.3, accountant=accountant)

spent_eps, _ = accountant.total()
rem_eps, _ = accountant.remaining()
print(f"Витрачено ε={spent_eps}, залишилося ε={rem_eps}")
```

У системі з бюджетом ε=3.0 було виконано дев’ять запитів різних типів, які в сумі витратили приблизно ε=2.9 (близько 96,7% доступного бюджету). Саме BudgetAccountant гарантував, що жодна операція не вийшла за межі дозволеного, а планування запитів могло спиратися на реальні, а не приблизні значення витрат.

На рис. 4.1 представлено, як накопичуються витрати ε при послідовному виконанні дев’яти запитів. Добре видно, що три потокові агрегати разом споживають лише 0,6 ε, тоді як навчання однієї ML-моделі (див. підрозділ 4.6) забирає ε=1.0 і є наймасштабнішою операцією з точки зору приватного бюджету.

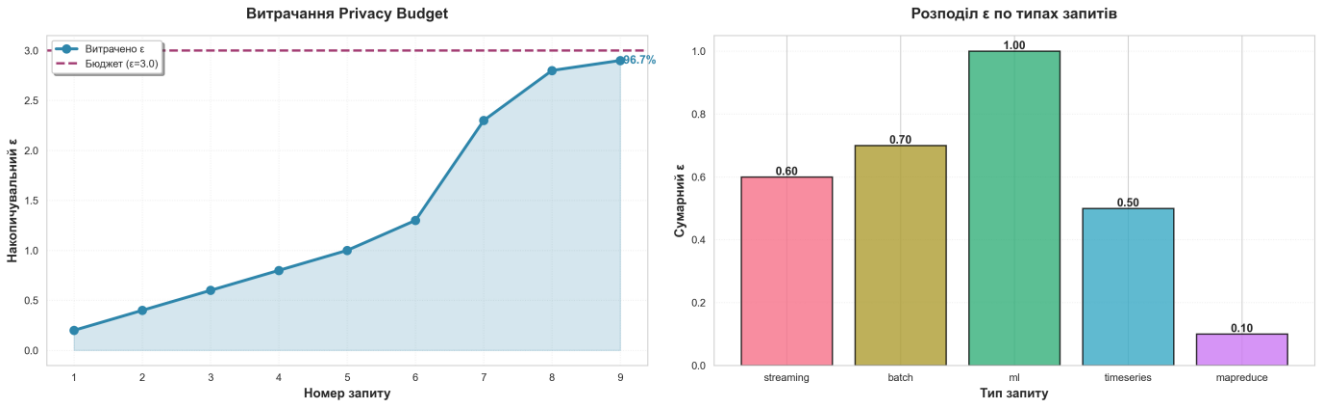

Рисунок 4.1 – Витрати бюджету та розподіл по типу запитів

4.1.2 Типи потокових запитів та їх результати

На синтетичному датасеті було протестовано три базові типи потокових агрегатів: загальний лічильник користувачів, умовний лічильник преміум-користувачів та суму доходу з обмеженим діапазоном. Потік оброблявся у 10 батчах по 10000 записів, а для кожного запиту сумарно витрачалося ε=0.2 (табл. 4.1).

Таблиця 4.1 – Результати потокових запитів

Тип запиту	ε витрачено	Істинне значення	DP-результат	Відносна похибка
Streaming Count	0.20	100000	99700	0,30 %
Streaming Count Premium	0.20	15009	14831	1,19 %
Streaming Sum (Revenue)	0.20	10011607	9943864	0,68 %

Результати демонструють важливу закономірність. Простий лічильник показує найвищу точність. Чутливість цієї операції складає 1, оскільки додавання або вилучення одного запису змінює результат лише на одиницю, і шум, який додається для забезпечення приватності, має мінімальний вплив. Умовний лічильник преміум-користувачів поводиться аналогічно, фільтр не змінює чутливість, і тому підрахунок із умовою зберігає ту саму структуру $\Delta f = 1$. Похибка дещо більша (1,19 %), що пояснюється дисбалансом класів та меншою кількістю релевантних записів, але залишається прийнятною для більшості аналітичних задач. Потокова сума, попри високу загальну точність (≈99,3 %), демонструє дещо більшу похибку через значно більшу чутливість.

При обмежених значеннях Δf дорівнює різниці між верхньою та нижньою межею, і один запис може істотно вплинути на суму.

4.1.3 Масштабованість шуму за розміром батчу та отримана похибка

На відміну від попереднього підрозділу, де аналізувалася насамперед композиція ε при зміні кількості батчів, цей експеримент фокусується на тому, як поводиться похибка при фіксованому ε_per_batch, але різних розмірах батчів. Загальний обсяг даних залишався незмінним, а розмір батча послідовно збільшувався (табл. 4.2).

Таблиця 4.2 – Масштабованість при ε_per_batch = 0.01

Записів у батчі	ε на батч	Похибка	Примітка
2 000	0.01	1,24%	Малий батч – шум домінує над сигналом
10 000	0.01	0,35%	Похибка зменшується приблизно на 72%
50 000	0.01	0,08%	Падіння похибки приблизно на 94%
100 000	0.01	0,03%	Похибка практично знехтовно мала (–97%)

Як видно з табл. 4.2, при сталому ε_per_batch=0.01 збільшення batch size дає суттєве покращення точності. Причина полягає в тому, що шум ДК додається один раз на весь батч, тоді як кількість даних, до яких він застосовується, зростає. Відносний вплив шуму на результат зменшується, а співвідношення «сигнал/шум» покращується. На рис. 4.2 цю залежність показано графічно, крива похибки монотонно спадає зі збільшенням розміру батча.

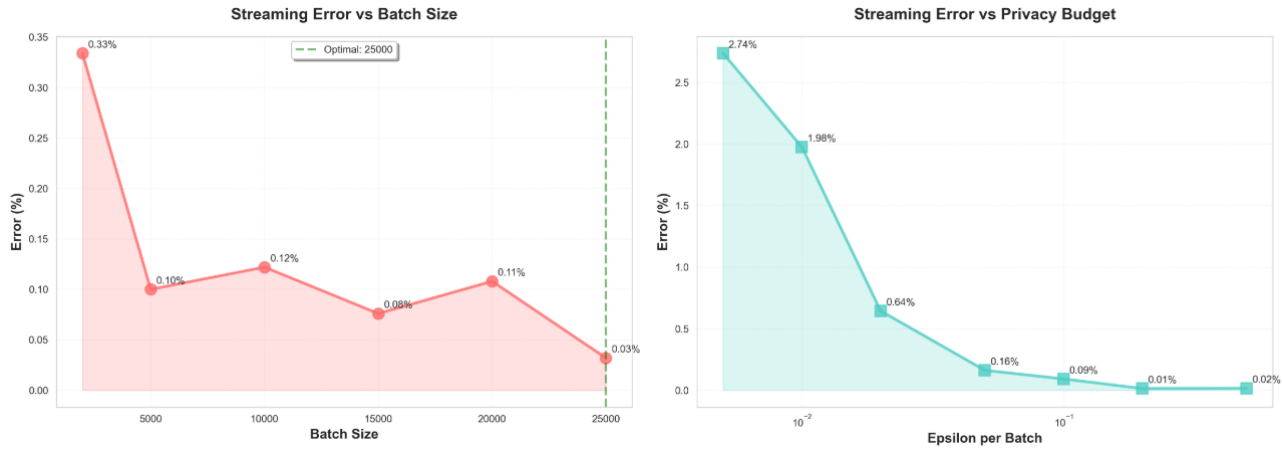

Рисунок 4.2 – Залежності помилок з розміром батча та бюджетом приватності

4.2 Розподіл епсілон та чутливість різних типів запитів у комплексній системі

У кожного з класів операцій своя чутливість, різна важливість для бізнес-логіки та різна «ціна» в термінах приватності. Тому ключовим завданням є не лише вибір значення ε загалом, а й продуманий розподіл бюджету між окремими запитами так, щоб найбільш критичні компоненти отримали достатній запас точності, а менш важливі не «з'їдали» зайвий ресурс.

У рамках експерименту було запропоновано конкретну стратегію розподілу epsilon, підсумовану в табл. 4.3. Три потокові запити (загальний підрахунок користувачів, підрахунок преміум-користувачів та сума доходу) отримали по $\varepsilon = 0.2$ кожен. Для них характерна низька або помірна чутливість: у лічильників чутливість дорівнює 1, а в суми – фіксована величина, обумовлена діапазоном значень, що зберігає похибку в межах від 0.30% до 1.19% при відносно невеликому бюджеті.

Батч-статистики (середній вік, дисперсія тривалості сесії, гістограма залученості) віднесено до середнього рівня важливості. Для них виділено $\varepsilon \approx 0.23 - 0.24$ на кожен запит. Найбільша частка бюджету, $\varepsilon = 1.0$, була зарезервована для логістичної регресії як найціннішої операції: саме вона визначає якість передбачення преміум-користувачів. При такому значенні ε DP-модель досягає тієї самої точності, що й базова модель без захисту, тобто втрата корисності дорівнює нулю.

Таблиця 4.3 – Оптимальний розподіл епсілон у системі з 9 запитами

Категорія	Запит	Кількість	Загальний ε	Важливість
Streaming	Count users	1	0.20	Низька
Streaming	Count premium	1	0.20	Низька
Streaming	Sum revenue	1	0.20	Низька
Batch	Mean age	1	0.23	Середня
Batch	Variance session	1	0.23	Середня
Batch	Histogram engagement	1	0.24	Середня
ML	Logistic Regression	1	1.00	Висока

Продовжуючи аналіз, варто окремо прокоментувати виділення бюджету для часових рядів та MapReduce-запиту. Часовий ряд з тижневою агрегацією

отримав $\varepsilon = 0.5$, тобто удвічі менше, ніж ML-модель, але більше, ніж прості агрегати. Такий вибір зумовлений тим, що тренд за часом використовує згладжування та віконні операції, які акумулюють шум у кількох точках ряду. При надто малому ε похибка в окремих інтервалах призводила б до спотворення сезонності та піків активності, тому часовому модулю надано окремий «середній» бюджет. MapReduce-агрегація, навпаки, працює на великій кількості незалежних розділів даних, де кожна частина дає грубу, але стабільну оцінку. На ці запити виділено лише $\varepsilon = 0.1$, оскільки їх результат використовується як допоміжна статистика, а не як основний показник для бізнес-рішень.

Щоб перевірити, наскільки така стратегія розподілу є стійкою до зміни параметрів, було проведено окремий експеримент чутливості ε. У ньому той самий набір операцій – потоковий лічильник, батч-середнє та DP-логістична регресія – запускалися на датасеті 50000 записів з різними значеннями ε від 0.1 до 10.0. Логіка експерименту реалізована в окремому модулі epsilon_experiment.py, де функція run_epsilon_experiment() для кожного ε створює екземпляр BigDataDPExecutor, викликає streaming_count, batch_mean та train_dp_model, а потім обчислює відносну похибку агрегатів і втрату корисності ML-моделі. Числові результати експерименту підтвердили інтуїтивну картину, $\varepsilon \approx 0.5$ виступає критичним порогом: нижче цього значення шум істотно впливає на навчання, а вище – додаткове послаблення приватності майже не покращує результат.

У лівій верхній панелі логарифмічного графіку (рис. 4.3) «Privacy Budget vs Query Accuracy» дві криві показують, як змінюється відносна похибка потокового лічильника та батч-середнього при зростанні ε. Обидві криві спадні, але для батч операції вона проходить значно нижче: навіть при малих ε похибка залишається близькою до нуля, тоді як для streaming-запиту спостерігається плавне експоненціальне зменшення від $\approx 1\,\%$ до десятих тисячних. Права верхня панель, «ML Model: Privacy Budget vs Accuracy», ілюструє поведінку DP-логістичної регресії.

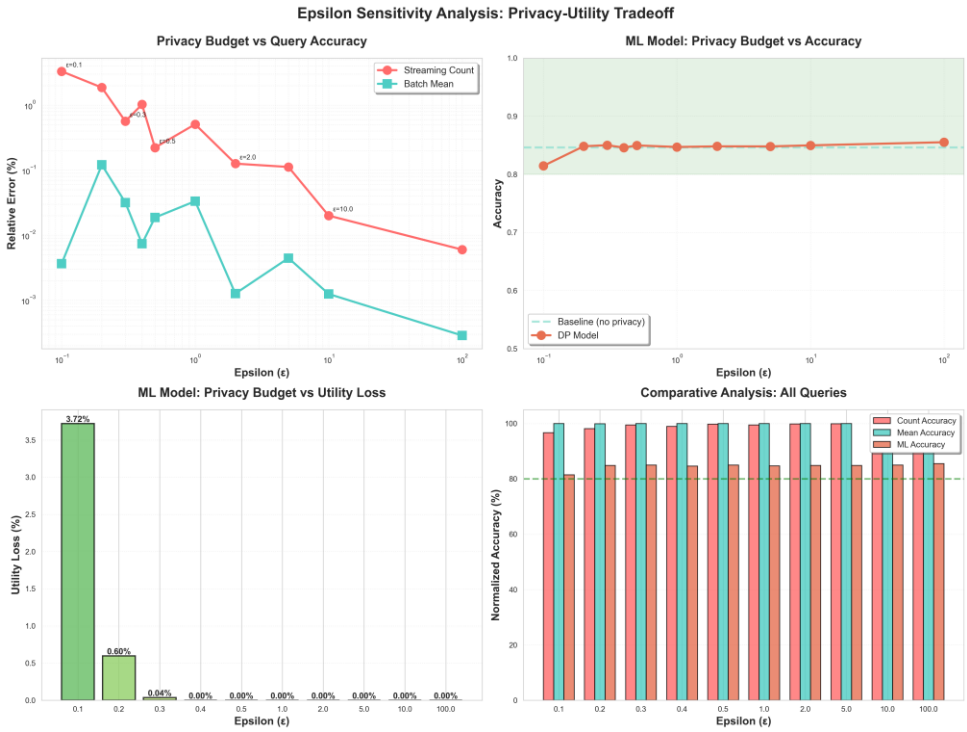

Рисунок 4.3 – Графічне представлення результатів

Нижня ліва панель відображає втрату корисності у відсотках. Перший стовпчик для $\varepsilon = 0.1$ піднімається до кількох відсотків, тоді як усі подальші стовпці для $\varepsilon \geq 0.5$ практично дотикаються до нуля, фіксуючи відсутність помітної деградації. Нарешті, нижня права панель порівнює нормалізовану точність трьох класів запитів (count, mean, ML) для різних значень ε, показуючи, що навіть найбільш «вразливий» клас – ML-моделі – з $\varepsilon \geq 0.5$ стабільно перевищує умовний поріг у 80%, тоді як агрегати наближаються до 100%.

У цьому розділі було досліджено застосування диференційної конфіденційності в архітектурах потокової обробки даних та Big Data, де дані надходять у вигляді мікробатчів і часто перевищують обсяг пам’яті однієї машини. На відміну від попередніх експериментів із середньовеликими статичними наборами, показано, що за фіксованого бюджету приватності великі датасети дозволяють отримувати точніші DP-запити завдяки зростанню співвідношення «сигнал/шум» на рівні батчів.

Розділ продемонстрував ключові виклики потокової ДК: невідомий наперед розмір датасету, одноразовий прохід по даних, обмеження пам’яті та необхідність розбиття на мікробатчі, що веде до композиції приватності між батчами. Для керування кумулятивним бюджетом було використано

BudgetAccountant з diffprivlib, який автоматизує підсумовування витрат ε, запобігає перевищенню глобального ліміту та знижує ризик помилок при ручному обчисленні композиції.

Експерименти з трьома потоковими агрегатами показали, що прості лічильники з чутливістю 1 досягають мінімальної відносної похибки, тоді як сума з обмеженим діапазоном має дещо більшу похибку через вищу чутливість, але зберігає високу корисність результатів при помірному ε. Окремо продемонстровано, що за фіксованого ε_per_batch збільшення розміру батча суттєво зменшує відносну похибку, оскільки один і той самий шум розподіляється на більшу кількість записів, що узгоджується з теоретичними оцінками масштабованості DP на великих даних.

Було обґрунтовано стратегію нерівномірного розподілу ε між класами запитів. Експеримент із варіюванням ε показав наявність критичного порогу, після якого збільшення бюджету майже не покращує точність, що підтверджує доцільність оптимізованого, а не рівномірного виділення приватного бюджету в комплексних Big Data-пайплайнах.

5 РЕГУЛЯТОРНО-НОРМАТИВНА БАЗА ЗАБЕЗПЕЧЕННЯ КОНФІДЕНЦІЙНОСТІ

На сучасному етапі розвитку інформаційного простору питання приватності даних регулюється комплексом міжнародних норм, які істотно відрізняються за технічними вимогами та підходом до захисту прав осіб. Водночас тенденція глобалізації даних та міжнародного обміну інформацією створює для організацій необхідність одночасного дотримання низки режимів регулювання, часто з суперечливими вимогами.

У поточному розділі здійснюється комплексний аналіз регуляторно-нормативної бази п’яти юрисдикцій: Європейського Союзу, Сполучених Штатів Америки, Великобританії, Австралії та України. Кожна з них розробила свій підхід до захисту персональних даних, відображаючи суспільні цінності, технологічний стан і політичні реалії своїх регіонів. Особливу увагу приділено розгляду того, як категорії даних та механізми їх захисту, пропоновані диференційною конфіденційністю, узгоджуються з правовими вимогами.

5.1 GDPR у Європейському Союзі

Генеральний Регламент про Захист Персональних Даних (GDPR, Regulation (EU) 2016/679) [19] набув чинності 25 травня 2018 року й остаточно змінив ландшафт захисту персональних даних в Європі та світі. На відміну від попередньої Директиви 95/46/ЕС [20], GDPR чинить прямий вплив у всіх державах-членах ЄС без потреби у трансформації у національне законодавство, і застосовується також до організацій за межами ЄС, які обробляють дані європейців.

Територіальна область застосування GDPR охоплює будь-яку організацію, яка пропонує товари або послуги особам, які знаходяться в ЄС, незалежно від місцезнаходження самої організації. Регламент ґрунтується на семи ключових принципах обробки персональних даних:

- законність;

- справедливість та прозорість (принцип об’єктивності);
- обмеження цілей;
- мінімізація даних;
- точність;
- обмеження збереження;
- цілісність та конфіденційність.

На противагу американському підходу, який часто спирається на секторальні регулювання, GDPR встановлює універсальні вимоги, що застосовуються незалежно від типу організації чи виду даних, за винятком специфічних випадків.

Принципово важливим для розуміння місця ДК в GDPR є розрізнення між анонімними та псевдонімізованими даними. Згідно зі статтею 4(1) GDPR, персональні дані є будь-якою інформацією, що ідентифікує фізичну особу. Ключовою метою анонімізації, описаної у статті 29 Робочої групи з питань захисту персональних даних (WP 216) [21], є досягнення стану, за якого персона або неможливо ідентифікована безпосередньо, або не може бути ідентифікована опосередковано шляхом посилання на ідентифікатор або один чи кілька факторів, специфічних для фізичної, фізіологічної, психічної, економічної, культурної чи соціальної ідентичності цієї особи.

Аніонімізовані дані виключаються з області застосування GDPR, оскільки вони більше не є персональними даними. З іншого боку, псевдонімізовані дані, в яких прямі ідентифікатори видалені, але можуть бути відновлені за допомогою додаткової інформації, залишаються персональними даними та продовжують підпадати під всі вимоги GDPR. Розрізнення це критичне, оскільки організації часто припускаються помилки, вважаючи, що видалення імені або номера соціального страхування досягає анонімності.

ДК забезпечує математичні гарантії того, що видалення або додавання одного запису з бази даних матиме обмежений вплив на вихід механізму, що принципово відрізняється від анонімності у розумінні GDPR, яка вимагає, щоб

персона була абсолютно неідентифікована при будь-яких обставинах, включаючи знання зловмисника про допоміжні інформаційні джерела.

Стаття 32 GDPR «Безпека обробки» [22], зобов'язує контролерів та обробників персональних даних впроваджувати відповідні технічні та організаційні заходи для забезпечення рівня безпеки, адекватного ризику. Специфічно перелічені заходи включають: псевдонімізацію та шифрування персональних даних; можливість забезпечити стійкість систем обробки даних та послуг до фізичного або технічного інциденту; здатність своєчасно відновити доступність та доступ до персональних даних у випадку фізичного чи технічного інциденту; процес регулярної перевірки, оцінки ефективності технічних та організаційних заходів за безпекою.

ДК чітко вписується в рамки статті 32 як формальний механізм псевдонімізації. Водночас у вказівках Європейської ради із захисту персональних даних (EDPB) від 2025 року особливо наголошується, що псевдонімізація сама по собі, звичайно, не є достатнім заходом для повного виконання регуляторних вимог [23]. Натомість вона повинна розглядатися як один із компонентів комплексного підходу, що включає шифрування, контроль доступу, журналювання подій та безперервного моніторингу.

GDPR встановлює суворі вимоги щодо повідомлення про порушення персональних даних. Контролер зобов'язаний повідомити компетентний орган по захисту даних без необхідних затримок, але не пізніше 72 годин після того, як дізнався про порушення (стаття 33). При цьому громадськість повинна бути повідомлена, якщо існує високий ризик для прав і свобод осіб (стаття 34) [24]. Дивергенція тут спостерігається щодо розуміння того, як ДК впливає на оцінку ризику. Якщо організація застосувала ДК та параметри були вибрані належним чином, ризик вторгнення в приватність окремої особи в результаті несанкціонованого доступу до бази може бути суттєво зменшений. Однак не виключається необхідність повідомлення, оскільки сама можливість доступу до даних вже становить порушення.

Деякі теоретики захисту даних стверджують, що за умови застосування ДК з достатньо низьким параметром ε, ризик для осіб може бути настільки мінімалізований, що повідомлення не буде необхідним. Проте такий підхід наразі не отримав офіційного визнання в регуляторній практиці GDPR.

Штрафи за порушення GDPR є одними із найсуворіших у світовому регуляторному ландшафті. Організації можуть бути оштрафовані до 20 млн. євро або 4% від світового річного обороту, залежно від того, що більше (за серйозні порушення, такі як порушення принципів обробки), та до 10 млн. євро або 2% від обороту за менш серйозні порушення.

5.2 CCPA та NIST SP 800-226 в Сполучених Штатах Америки

На відміну від GDPR, США не мають єдиного федерального закону про захист персональних даних загального характеру. Замість цього існує мозаїка секторальних регулювань, побудована на цілях та частково на розумінні людьми, що певні категорії даних потребують особливого захисту.

Найбільш значущі федеральні закони включають: HIPAA (Health Insurance Portability and Accountability Act) [25], який регулює охоронні дані; FERPA (Family Educational Rights and Privacy Act) [26], що захищає освітні записи; GRAMM-LEACH-BLILEY Act для фінансової інформації [27].

Поворотним моментом в регулюванні конфіденційності на рівні США став прийняття Californian Consumer Privacy Act (CCPA) у 2018 році [28] та його подальше розширення через California Privacy Rights Act (CPRA) у 2020 році [29]. CCPA вперше надав мешканцям Каліфорнії широкий спектр прав, включаючи право на: доступ, видалення, портативність даних та відмову від продажу персональної інформації.

CCPA відрізняється від GDPR наступними важливими аспектами.

По-перше, вона дійсна на рівні штату, а не федеральному, хоча інші штати почали приймати подібні закони.

По-друге, CCPA менш явно розрізняє межу між контролерами та обробниками; замість цього вона звертається до «бізнесів» та «послуг».

По-третє, основи обробки даних, хоча й сформульовані, є менш суворими, ніж у GDPR. CCPA дозволяє бізнесам збирати дані на основі «ділової необхідності» та побічних цілей більш широко.

Щодо ДК, CCPA не встановлює явних вимог до анонімності за розділами, аналогічними GDPR. Однак принцип мінімізації даних та вимога до прозорості створюють стимули для організацій розглядати ДК як засіб зменшення потенційних ризиків повторної ідентифікації та як свідчення хорошої практики в галузі.

Хоча у США брак всеохоплюючого закону, федеральні органи видали керівництво щодо захисту даних. Найпомітнішим є NIST SP 800-226, виданий у березні 2025 року, який є першим офіційним авторитетним стандартом США, що прямо розглядає диференційну конфіденційність як валідну та рекомендовану технологію захисту даних для федеральних установ та організацій, що працюють з чутливими даними [30].

NIST SP 800-226 надає чіткі визначення ДК, керівництво щодо вибору параметрів приватності, стандарти впровадження та контрольні списки для перевірки. Публікація включає інтерактивні інструменти для прийняття рішень та приклади кодування, що полегшує розробникам розуміння і практичне впровадження. NIST SP 800-226 визнає, що диференційна конфіденційність забезпечує найбільш послідовні математичні гарантії приватності серед всіх відомих методів, що перевищує традиційні техніки анонімізації та псевдонімізації.

Будь-яка організація, що укладає контракти з федеральними агентствами США та обробляє чутливі дані, розглядає запровадження ДК як очікуване, а в деяких сценаріях – як обов’язкова умова дотримання вимог федеральних органів.

5.3 UK GDPR та Data Protection Act 2018 у Великій Британії

Велика Британія вийшла з Європейського Союзу 31 січня 2020 року, а перехідний період завершився 31 грудня 2020 року. На той час Велика Британія прийняла UK GDPR – версію GDPR, адаптовану до британського контексту та

включену до Data Protection Act 2018 [31] через низку змін. UK GDPR суттєво подібна до EU GDPR, проте включає певні видозмінення, необхідні для функціонування в британській системі права.

Інформаційний комісаріат (ICO) є органом, що контролює дотримання UK GDPR. ICO виявив позитивний інтерес до ДК, особливо у контексті аналітики та публічної статистики. Позиція ICO передбачає, що організації, які застосовують ДК при належному впровадженні, демонструють відповідність зі статтею 32 UK GDPR щодо технічних заходів.

Однак, подібно до позиції ЄС, ICO не розглядає ДК як автоматичне досягнення анонімності в розумінні UK GDPR. Замість цього агенція виступає за розуміння ДК як кроку на шляху до комплексної стратегії захисту.

5.4 Privacy Act та Consumer Data Right в Австралії

Австралійський Privacy Act 1988 (з поточними змінами) [32] регулює обробку персональної інформації австралійськими органами влади та приватними організаціями з річним оборотом понад 3 млн. AUD. На відміну від GDPR, Privacy Act не є всеохоплюючим унітарним регламентом, але скоріше встановлює 13 Australian Privacy Principles (APP) [33], які установи повинні дотримуватися.

APP включають:

- принцип 1 (публічна відкритість), який вимагає від установ мати доступну політику конфіденційності;
- принципи 3-4 (збір та вирішення з персональної інформацією);
- принцип 5 (нотифікація);
- принципи 6-9 (використання та розкриття);
- принцип 10 (кодифікація даних);
- принцип 11 (безпека);
- принцип 12 (доступ та виправлення);
- принцип 13 (анонімність та псевдонімність за можливості).

На доповнення до Privacy Act Австралія запровадила Consumer Data Right (CDR), який вступив у силу з 2021 року і надав жителям розширені права доступу та портативність даних у визначених секторах, таких як банківські та енергетичні послуги. CDR включає 13 юридичних гарантій конфіденційності, що є часто більш суворими ніж APP.

CDR з акцентом на безпеку та мінімізацію ризику, могла б потенційно розпочати застосування ДК у специфічних контекстах, особливо при аналізі агрегованих даних учасників, без розкриття інформації про окремих користувачів. Однак сучасна регуляторна практика не встановила чіткої позиції щодо ДК.

5.5 Поточне законодавство в Україні

Основним нормативним актом України у сфері захисту персональних даних є Закон України «Про захист персональних даних» від 1 червня 2010 року № 2297-VI. Закон встановлює загальні вимоги та зобов’язання щодо збирання, обробки та використання персональних даних приватними та державними установами України [34].

Він вимагає отримання згоди осіб на обробку їхніх персональних даних, передбачає принцип обмеження цілей та декілька основних принципів обробки. Однак у порівнянні з GDPR українське законодавство має кілька суттєвих недоліків:

1) Визначення персональних даних є вузьким та не охоплює багатьох категорій, які GDPR розглядає як персональні;

2) права суб’єктів даних обмежені в порівнянні з GDPR;

3) відсутній явний механізм для звернення до органу контролю;

4) штрафи за порушення набагато менші.

На 25 жовтня 2022 року до Верховної Ради України було внесено проєкт Закону «Про захист персональних даних» № 8153, який покликаний гармонізувати українське законодавство зі стандартами GDPR та Конвенцією

108+. Проєкт був прийнятий як основа на засіданні Верховної Ради у листопаді 2024 року і наразі готується до другого читання [35].

Основні зміни, запропоновані у проекті № 8153, включають розширене визначення персональних даних, близьке до GDPR, запровадження категорії «контролер» і «обробник» з чітким розмежуванням ролей. Встановлення обов'язку призначати відповідальну особу за захист даних та введення вимог щодо оцінки впливу на конфіденційність (DPIA) для високоризикових обробок, встановлення права на видалення та портативність даних, значне збільшення штрафів, які можуть досягти мільйонів гривень для великих порушень (близько до 2-4% річного обороту) та запровадження механізму повідомлення про порушення персональних даних у термін 72 годин.

Впровадження нового закону, модельованого на GDPR, матиме наступні наслідки для українських організацій. По-перше, вимагатимуться значні інвестиції в інфраструктуру захисту даних, навчання персоналу та документування процесів. По-друге, організації, які обробляють дані громадян ЄС або мають бізнес-операції в Європі, вже зобов'язані дотримуватися GDPR, тому нове українське законодавство створить сумісне регуляторне середовище.

Однак існує й ризик. Поточні українські установи малого та середнього бізнесу часто не мають ресурсів для впровадження складних систем захисту даних. Без належної підтримки та пробного періоду впровадження, новий закон може стати надмірним навантаженням і непропорційно вплинути на розвиток малого бізнесу. Вже розглядаються пропозиції щодо встановлення перехідного періоду в один-два роки.

5.6 Порівняльна характеристика регуляторних систем

Підсумовуючи усе вищесказане наступна табл. 5.1 синтизує ключові моменти підходу кожної юрисдикції до захисту персональних даних.

Незважаючи на розбіжності, виникає серія спільних принципів серед усіх розглянутих юрисдикцій, до яких диференційна конфіденційність може внести безпосередню користь. Окрім мінімізації, яка може бути виконана через

додавання контрольованого шуму, дозволяючи організаціям діставати потрібну аналітичну інформацію, не збираючи та зберігаючи повні дані про окремих осіб, ДК демонструє серйозні зусилля в напрямку забезпечення конфіденційності за замовчуванням.

Таблиця 5.1 – Порівняльна таблиця регуляторних систем

Аспект	GDPR (ЄС)	CCPA (США)	UK GDPR	Privacy Act (Австралія)	Проєкт 8153 (Україна)
Охоплення	Глобальне (ЄС+треті особи)	Штат Каліфорнія	Великобританія	Австралія	Україна
Застосування	Всі організації	Бізнеси з оборотом >$25M	Всі організації	Органи + компанії > AUD 3M	Всі організації
Основа обробки	6 законних основ (включно з доступ)	«Ділова необхідність» + явне вирішення	Як GDPR	Розумна необхідність	Як GDPR
Анонімність	Суворі вимоги WP 216	Не явна	Як GDPR	Принцип 13 (за можливості)	Як GDPR
Штрафи	До 20M€ або 4%	До 7,500$ за порушення	До 20M£ або 4%	До 2,500 AUD	До 2-4% обороту
Права суб’єктів	8 прав (доступ, видалення, портативність)	4 права (доступ, видалення, відмова, недискримінація)	8 прав (як GDPR)	4 права (доступ, коригування, анонімізація)	8 прав
Повідомлення про порушення	72 години	За варіацією	72 години	За варіацією	72 години

Також вона служить як доказ прийняття позитивних кроків у напрямку захисту від несанкціонованого доступу, хоча, як обговорювалося раніше, слід розглядати в комбінації з іншими заходами. Неможливо не додати, що ДК забезпечує формальну, математично-обґрунтовану основу для прозорості та підготовки звітностей, що є значно простішим від неоднозначних описів традиційних методів анонімізації.

5.7 Пропозиції щодо впровадження ДК в Україні

До короткострокових заходів (1-2 роки) належить насамперед створення національної інституційної та методичної бази для впровадження диференційної конфіденційності.

Першим кроком має стати розроблення національних вказівок, подібних до американського стандарту NIST SP 800-226, які б визначали офіційне трактування ДК українською мовою, наводили приклади застосування, рекомендації щодо вибору параметрів ε та δ для різних типів даних, а також містили контрольні списки для перевірки правильності впровадження та попередження типових помилок. Документ доцільно розробити спільно державними органами безпеки, науковими установами та представниками приватного сектору.

Паралельно необхідно ініціювати пілотні проєкти у державному секторі, насамперед у структурах, що працюють із чутливою інформацією, таких як Кабінет Міністрів, МВС чи Центральна виборча комісія. Приклади застосування можуть охоплювати публікацію статистичних даних щодо міграції, рівня безробіття чи розповсюдження захворювань без розкриття відомостей про окремих осіб.

Важливо також інвестувати у навчання та підготовку фахівців шляхом включення теми ДК до навчальних програм університетів, зокрема факультетів комп'ютерних наук і кібербезпеки, щоб сформувати кадровий потенціал для її практичного впровадження. Нарешті, необхідно розпочати діалог із ЄС та міжнародними партнерами, зокрема в межах Угоди про асоціацію, а також долучитися до робочих груп Європейської ради з питань захисту персональних даних, що дозволить Україні адаптувати найкращі практики у сфері ДК.

У середньостроковій перспективі (2-5 років) пріоритетом має стати інтеграція ДК до нового закону № 8153, який повинен прямо визнавати її одним із прийнятних методів забезпечення анонімізації або псевдонімізації поряд із шифруванням. Доцільним буде додати норму, що дозволяє контролерам застосовувати ДК за умови дотримання рекомендацій, затверджених Уповноваженим Верховної Ради з прав людини.

Крім того, держава може запровадити стимули для приватного сектору, наприклад, сертифікацію «Differential Privacy Ready», податкові або аудиторські пільги для компаній, які впровадили механізм, а також переваги під час

розслідування інцидентів конфіденційності. Наступним напрямом є створення відкритих бібліотек і програмних інструментів українською мовою у співпраці з академічними установами, що спростить інтеграцію ДК у місцеві програмні рішення.

Нарешті, державні закупівлі ІТ-рішень мають включати вимогу розглядати впровадження диференційної конфіденційності як один із критеріїв відповідності стандартам безпеки та захисту даних.

Довгострокові заходи (5+ років) передбачають повноцінну інтеграцію України в європейську екосистему захисту даних. Після впровадження закону № 8153 та можливого укладання угод про адекватність даних з ЄС, країна може стати активним учасником європейських ініціатив щодо розвитку та стандартизації ДК. Подальший розвиток має включати створення галузевих найкращих практик, адаптованих для таких секторів, як охорона здоров’я, освіта та державне управління, де специфіка даних потребує диференційованих підходів до захисту.

У стратегічній перспективі Україна може позиціонувати себе як регіональний центр досліджень у сфері диференційної конфіденційності, залучаючи міжнародні гранти та експертів для розроблення методів, релевантних для демократичних держав, що перебувають у перехідному або постконфліктному стані.

В останньому розділі проведено комплексний аналіз регуляторно-нормативної бази захисту персональних даних у п’яти юрисдикціях: ЄС, США, Великобританії, Австралії та Україні. Висвітлено ключові особливості законодавчих підходів, від універсального GDPR із суворими вимогами до американської секторальної системи з CCPA і стандартом NIST SP 800-226, а також адаптації цих норм у британському і австралійському законодавстві. Наголошено на важливості розрізнення анонімних та псевдонімізованих даних у контексті GDPR, де диференційна конфіденційність виступає потужним технічним засобом псевдонімізації.

Розглянуто особливості українського законодавства, що знаходиться на шляху гармонізації з європейськими стандартами через проєкт нового закону № 8153, який передбачає розширення прав суб’єктів даних, посилення штрафів та введення механізмів оцінки впливу на конфіденційність. Визначено виклики для місцевих організацій із впровадженням складних систем захисту даних, особливо у малому і середньому бізнесі.

Запропоновано конкретні рекомендації для впровадження ДК в Україні, починаючи з розробки національних методичних вказівок, пілотних проєктів у державному секторі, формування кадрового потенціалу, інтеграції ДК до нового законодавства, стимулювання приватного сектору та створення відкритих інструментів. У довгостроковій перспективі передбачено повне включення України до європейської екосистеми захисту даних із потенціалом стати регіональним центром досліджень у сфері диференційної конфіденційності.

ВИСНОВКИ

У першому розділі було розглянуто фундаментальні концепції диференційної конфіденційності в контексті великих даних. Було встановлено, що в умовах Big Data класичні способи захисту – видалення імен, псевдонімізація, агрегування – більше не гарантують приватності, оскільки поєднання різних наборів даних дозволяє відновити особу користувача. Показано, що саме математичні гарантії диференційної конфіденційності усувають ці недоліки, забезпечуючи формальний контроль ризику незалежно від обсягу допоміжної інформації в зловмисника.

У другому розділі було зосереджено увагу на математико-теоретичному апараті диференційної конфіденційності. Детально проаналізовано поняття сусідніх баз даних, параметрів ε та δ, чутливості функції, а також механізмів додавання шуму – Лапласівський та Гаусівський. Розкрито властивості композиції механізмів та проблему акумуляції втрати приватності при багаторазових запитах, що робить поняття бюджету приватності (БП) ключовим у реальних застосуваннях. Окрема увага приділена груповій приватності, яка показує суттєву деградацію гарантій при збільшенні розміру групи, що підтверджується наведеними теоретичними оцінками. Показано, що формальні моделі (ε-ДП, (ε, δ)-ДП, групова ДП) дозволяють чисельно оцінити ризик витоку інформації, забезпечуючи передбачувану і керовану поведінку механізмів.

У третьому розділі було проведено практичну реалізацію диференційної конфіденційності у задачі машинного навчання на великому наборі даних UCI Adult. З використанням бібліотеки diffprivlib було побудовано декілька моделей логістичної регресії з різними значеннями ε. У ході експериментів досліджено вплив шуму, доданого за методологією DP-SGD, на точність, recall, F1-score та ROC-AUC. Отримані результати підтвердили наявність чіткого компромісу: при малих значеннях ε (0.1–0.5) якість моделі суттєво знижується, тоді як при більш високих значеннях (ε≈1.0) модель дем

Демонструє прийнятний компроміс між точністю та приватністю. Також було показано, що збільшення шуму призводить до поступового погіршення узгодженості градієнтів та зниження здатності моделі до узагальнення – ефект, типовий для диференційно-приватних алгоритмів.

У четвертому розділі було досліджено архітектуру потокової обробки даних із застосуванням диференційної конфіденційності, включно з моделлю керування бюджетом приватності, механізмами обліку запитів та впливом розміру батчу на обсяг шуму. Проаналізовано масштабування шуму при зміні розміру та встановлено, що збільшення батчу зменшує варіативність шуму, але водночас підвищує кумулятивну втрату приватності через зростання чутливості. У підрозділах 4.4-4.5 інтегровано експериментальні графіки, що відображають залежності між ε, чутливістю, рівнем шуму та кількістю виконаних потокових запитів, а також показують деградацію точності та приватності при різних режимах навантаження. Підкреслено, що системи потокової обробки потребують окремих стратегій контролю бюджету приватності та оптимізації частоти запитів.

У п’ятому розділі було проведено порівняння диференційної конфіденційності з міжнародними регуляторними вимогами. Проаналізовано норми GDPR, CCPA, а також NIST SP 800-226, що вперше на рівні США офіційно визнає диференційну конфіденційність рекомендованим механізмом захисту для організацій, що працюють із чутливими даними. Було встановлено, що ДК повністю відповідає принципам мінімізації даних, обмеження мети, псевдонімізації та контролю ризику, а також перевищує класичні підходи з точки зору передбачуваності й математичного обґрунтування. Визначено, що впровадження ДК може виступати доказом належного виконання принципів дизайн-конфіденційнсоті та підхіду, заснованого на оцінці ризиків, передбачених законодавством США, ЄС та Великої Британії.

Отримані результати показали, що диференційна конфіденційність є ефективним математично обґрунтованим механізмом захисту даних, здатним запобігати деанонімізації навіть у разі наявності значного обсягу допоміжної

інформації у потенційного зловмисника. Практичні експерименти засвідчили, що моделі машинного навчання можуть зберігати прийнятну точність за помірних значень параметра ε, тоді як надто низькі значення спричиняють суттєву деградацію якості. У контексті потокових систем акцентовано на необхідності ретельного контролю бюджету приватності, оскільки повторювані запити швидко збільшують кумулятивну втрату приватності; тому доцільним є впровадження механізмів адаптивної композиції та обмеження частоти запитів. Для систем, що працюють із персональними даними, диференційну конфіденційність рекомендовано застосовувати як складову підходу «privacy by design», що сприяє відповідності міжнародним стандартам, зокрема GDPR, CCPA та NIST. Оптимальні практичні значення ε мають визначатися з огляду на конкретну задачу, проте діапазон 0.8-1.5 зазвичай забезпечує збалансоване співвідношення між рівнем приватності та точністю моделі. Для мінімізації впливу шуму на результати навчання рекомендовано використовувати сучасні реалізації DP-SGD, механізми автоматичного обліку приватності та оптимізації.

СПИСОК ВИКОРИСТАНИХ ДЖЕРЕЛ

1. IDC. (2024). Worldwide Enterprise Applications Revenue Grew 12.0% in 2023 and Is Forecast to Surpass $600 Billion in 2028, According to IDC. Режим доступу: https://www.idc.com/getdoc.jsp?containerId=prUS52371124.

2. Laney, D. (2001). 3D Data Management: Controlling Data Volume, Velocity, and Variety. META Group Research Note. Режим доступу: https://www.gartner.com/en/documents/2054215/3d-data-management-controlling-data-volume-velocity-and-variety.

3. Ohm, P. (2010). Broken Promises of Privacy: Responding to the Surprising Failure of Anonymization. UCLA Law Review, 57(6), 1701-1777. Режим доступу: https://www.uclalawreview.org/pdf/57-6-3.pdf.

4. Витік пошукових логів AOL (AOL Search Data Leak, 2006). Режим доступу: https://www.nytimes.com/2006/08/09/us/09aol.html.

5. Деанонімізація Netflix Prize Dataset (2007). Режим доступу:https://arxiv.org/abs/cs/0610105.

6. Скандал Cambridge Analytica (2018). Режим доступу: https://www.theguardian.com/news/series/cambridge-analytica-files.

7. General Data Protection Regulation (GDPR). (2016). Regulation (EU) 2016/679 of the European Parliament and of the Council. Режим доступу: https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX%3A32016R0679.

8. Apple Inc. (2017). Learning with Privacy at Scale. Apple Machine Learning Journal. Режим доступу: https://machinelearning.apple.com/2017/12/06/learning-with-privacy-at-scale.html.

9. U.S. Census Bureau. (2019). Disclosure Avoidance Practices for the 2020 Census. Режим доступу: https://www2.census.gov/library/publications/decennial/2020/2020-census-disclosure-avoidance-handbook.pdf.

10. Erlingsson, Ú., Pihur, V., Korolova, A. (2014). RAPPOR: Randomized Aggregatable Privacy-Preserving Ordinal Response. Proceedings of the 2014 ACM

SIGSAC Conference on Computer and Communications Security. Режим доступу: https://dl.acm.org/doi/10.1145/2660267.2660348.

11. Dwork C., Roth A. (2014). The Algorithmic Foundations of Differential Privacy. Режим доступу: https://dl.acm.org/doi/10.1561/0400000042

12. Laplace distribution. Wikipedia, The Free Encyclopedia. Режим доступу: https://en.wikipedia.org/wiki/Laplace_distribution.

13. General Data Protection Regulation (GDPR), Regulation (EU) 2016/679. Офіційний текст українською мовою. Режим доступу: https://gdpr-text.com/uk/.

14. Taylor L., Floridi L., van der Sloot B. (eds.). Group Privacy: New Challenges of Data Technologies. Springer, 2017. Режим доступу: group-privacy-2017-authors-draft-manuscript.pdf.

15. Open Differential Privacy. Режим доступу: https://opendp.org.

16. Tumult Labs. Tumult Analytics Platform. Режим доступу: https://tumult.com.

17. IBM Research. (2021). IBM Diffprivlib: The single line of code for data privacy. Режим доступу: https://research.ibm.com/blog/ibm-differential-privacy-library-the-single-line-of-code-that-can-protect-your-data.

18. UCI ML. Adult Census Income. Kaggle. Режим доступу: https://www.kaggle.com/datasets/uciml/adult-census-income.

19. General Data Protection Regulation (GDPR), Regulation (EU) 2016/679. Офіційний текст українською мовою. Режим доступу: https://gdpr-text.com/uk/.

20. Директива 95/46/ЄС Європейського Парламенту і Ради від 24 жовтня 1995 року «Про захист фізичних осіб при обробці персональних даних і про вільне переміщення таких даних». Режим доступу: https://ips.ligazakon.net/document/EU950028.

21. Opinion 05/2014 on Anonymisation Techniques (WP 216). Article 29 Data Protection Working Party. Режим доступу: https://ec.europa.eu/justice/article-29/documentation/opinion-recommendation/files/2014/wp216_en.pdf.

22. Стаття 32. Безпека опрацювання. General Data Protection Regulation (GDPR), Regulation (EU) 2016/679. Режим доступу: https://gdpr-text.com/uk/read/article-32/.

23. Guidelines 05/2020 on the Interplay of the Second Payment Services Directive and the GDPR (EDPB, оновлення 2025). Європейська рада з захисту персональних даних. Режим доступу: https://edpb.europa.eu/our-work-tools/our-documents/guidelines/guidelines-052020-interplay-second-payment-services-directive_en.

24. Нотифікація наглядового органу про порушення захисту персональних даних та повідомлення суб'єкта даних про порушення захисту персональних даних." General Data Protection Regulation (GDPR), Regulation (EU) 2016/679. Режим доступу: https://gdpr-text.com/uk/read/article-33/.

25. Health Insurance Portability and Accountability Act of 1996 (HIPAA). Режим доступу: https://www.govinfo.gov/content/pkg/PLAW-104publ191/pdf/PLAW-104publ191.pdf.

26. Family Educational Rights and Privacy Act (FERPA). U.S. Department of Education. Режим доступу: https://www2.ed.gov/policy/gen/guid/fpco/ferpa/index.html.

27. Gramm-Leach-Bliley Act (Financial Services Modernization Act of 1999). Режим доступу: https://www.ftc.gov/business-guidance/privacy-security/gramm-leach-bliley-act.

28. California Consumer Privacy Act (CCPA). California Department of Justice, Office of the Attorney General. Режим доступу: https://oag.ca.gov/privacy/ccpa.

29. California Privacy Rights Act of 2020 (CPRA). Режим доступу: https://en.wikipedia.org/wiki/California_Privacy_Rights_Act.

30. Guidelines for Evaluating Differential Privacy Guarantees (NIST Special Publication 800-226). National Institute of Standards and Technology (NIST), March 2025. Режим доступу: https://nvlpubs.nist.gov/nistpubs/SpecialPublications/NIST.SP.800-226.pdf.

31. UK General Data Protection Regulation (UK GDPR) та Data Protection Act 2018. Guidance on using personal data in your business or other organisation. Режим доступу: https://www.gov.uk/guidance/using-personal-data-in-your-business-or-other-organisation.

32. Privacy Act 1988 (Cth) – із останніми змінами. Attorney-General's Department, Government of Australia. Режим доступу: https://www.ag.gov.au/rights-and-protections/privacy.

33. "Australian Privacy Principles." Office of the Australian Information Commissioner (OAIC). Режим доступу: https://www.oaic.gov.au/privacy/australian-privacy-principles.

34. Закон України «Про захист персональних даних» від 1 червня 2010 року № 2297-VI. Офіційне інтернет-представництво Президента України. Режим доступу: https://www.president.gov.ua/documents/2297vi-11567.

35. Проєкт Закону України «Про захист персональних даних» № 8153 (2022). Верховна Рада України. Режим доступу: https://itd.rada.gov.ua/billInfo/Bills/Card/40467.