Міністерство освіти і науки України
Харківський національний університет радіоелектроніки

Центр Післядипломної освіти
(повна назва)
Кафедра Штучного інтелекту
(повна назва)

КВАЛІФІКАЦІЙНА РОБОТА
Пояснювальна записка

рівень вищої освіти перший (бакалаврський)

Розробка системи генерації природномовних текстів з імітацією стилю цільового автора на основі імітаційного навчання з використанням генеративно-змагальних нейронних мереж
(тема)

Виконав:
здобувач другого року навчання,
групи ІТШІп-23-1
Віталій Галкін
(власне ім'я, прізвище)

Спеціальність 122 Комп'ютерні науки
(код і повна назва спеціальності)
Тип програми освітньо-професійна
Освітня програма Штучний інтелект
(повна назва освітньої програми)

Керівник ас. Ірина Малєєва
(посада, власне ім'я, прізвище)

Допускається до захисту

Завідувач кафедри ШІ _____ Олег ЗОЛОТУХІН
(підпис) (власне ім'я, прізвище)

2025 р.

Харківський національний університет радіоелектроніки

Центр Післядипломної освіти

Кафедра Штучного інтелекту

Рівень вищої освіти перший (бакалаврський)

Спеціальність 122 Комп'ютерні науки
(код і повна назва)

Тип програми освітньо-професійна

Освітня програма Штучний інтелект
(повна назва)

ЗАТВЕРДЖУЮ:
Зав. кафедри _____
(підпис)
«_____» _____ 20 ___ р.

ЗАВДАННЯ
НА КВАЛІФІКАЦІЙНУ РОБОТУ

здобувачеві Галкіну Віталію Вікторовичу
(прізвище, ім'я, по батькові)

1. Тема роботи Розробка системи генерації природномовних текстів з імітацією стилю цільового автора на основі імітаційного навчання з використанням генеративно-змагальних нейронних мереж
затверджена наказом університету від 19 травня 20 25 р. № 387Ст
2. Термін подання студентом роботи до екзаменаційної комісії 24 червня 20 25 р.
3. Вихідні дані до роботи Науково-технічні публікації, дані інтернет-джерел та відомих наукових проектів щодо розробки та дослідження систем з використанням штучного інтелекту

4. Перелік питань, що потрібно опрацювати в роботі
1) Аналіз предметної галузі
2) Теоретичні основи запропонованого підходу
3) Експериментальна перевірка TextGAIL

КАЛЕНДАРНИЙ ПЛАН

№	Назва етапів роботи	Строк / терміни виконання етапів роботи	Примітка
1	Отримання завдання на кваліфікаційну роботу	19.05.2025	виконано
2	Аналіз предметної галузі	08.04.2025	виконано
3	Огляд існуючих метрик оцінювання якості	11.04.2025	виконано
4	Визначення проблем при розробці системи	20.04.2025	виконано
5	Моделювання та навчання системи	25.04.2025	виконано
6	Написання пояснювальної записки	15.05.2025	виконано
7	Перевірка на академічний плагіат	27.06.2025	виконано
8	Нормоконтроль	28.06.2025	виконано
9	Підготовка презентації та доповіді	12.06.2025	виконано
10	Попередній захист	20.06.2025	виконано
11	Рецензування	20.06.2025	виконано
12	Захист перед ЕК	24.06.2025	

Дата видачі завдання 19 травня 20 25 р.

Здобувач _____
(підпис)

Керівник роботи _____ ас. Ірина Малєєва
(підпис) (посада, власне ім’я, прізвище)

РЕФЕРАТ

Пояснювальна записка: 67 с., 3 рис., 4 табл., 1 дод., 25 джерел.

ГЕНЕРАЦІЯ ТЕКСТУ, ІМІТАЦІЙНЕ НАВЧАННЯ, СТИЛІЗАЦІЯ МОВИ, ШТУЧНИЙ ІНТЕЛЕКТ, DPO, GPT-2, ROBERTA, TEXTGAIL.

Об’єкт дослідження – процес автоматичної генерації природномовного тексту із заданими стилістичними характеристиками на основі моделей машинного навчання.

Предмет дослідження – архітектури генеративних моделей, зокрема трансформерних мереж та генеративно-змагальних підходів, що дозволяють керувати стилістичною відповідністю згенерованого тексту.

Мета роботи – розробка та експериментальна оцінка системи генерації тексту з імітацією авторського стилю на основі генеративно-змагального імітаційного навчання із застосуванням методу прямої оптимізації преференцій (DPO).

Методи дослідження – генеративно-змагальні мережі, трансформерні архітектури, імітаційне навчання, метод прямої оптимізації преференцій, автоматичні метрики оцінки якості генерації тексту, експертний аналіз.

У цій кваліфікаційній роботі досліджено сучасні методи стилістично-керованої генерації текстів та визначено суттєві обмеження традиційних підходів на основі максимізації правдоподібності, зокрема проблему втрати стилістичної індивідуальності. Запропоновано та реалізовано архітектуру TextGAIL, що об'єднує генератор GPT-2 із дискримінатором RoBERTa та додатковим етапом прямої оптимізації преференцій. Експериментальна перевірка підтвердила значне покращення стилістичної відповідності та індивідуалізації текстів порівняно з традиційними підходами.

ABSTRACT

Bachelor’s thesis contains: 67 pp., 3 fig., 4 tabl., 1 ann., 25 references.

ARTIFICIAL INTELLIGENCE, DPO, GPT-2, IMITATION LEARNING, LANGUAGE STYLIZATION, ROBERTA, TEXT GENERATION, TEXTGAIL.

Object of research – the process of automatic generation of natural language text with specified stylistic characteristics using machine learning models.

Subject of research – architectures of generative models, particularly transformer networks and generative-adversarial approaches, enabling controlled stylistic text generation.

Goal of research – to develop and experimentally evaluate a text generation system capable of imitating an author’s style on the basis of generative adversarial imitation learning using the direct preference optimization (DPO) method.

Research methods – generative adversarial networks, transformer architectures, imitation learning techniques, direct preference optimization (DPO), automated metrics for evaluating text-generation quality, expert analysis.

As a result of this research, existing methods for stylistically-controlled text generation are examined, and key limitations of traditional maximum-likelihood approaches, particularly the loss of stylistic individuality, are identified. The TextGAIL architecture, combining a GPT-2-based generator, RoBERTa-based discriminator, and an additional direct preference optimization step, has been proposed and implemented. Experimental evaluation confirmed significant improvements in stylistic match and individualization compared to traditional methods.

ЗМІСТ

ПЕРЕЛІК УМОВНИХ ПОЗНАЧЕНЬ, СИМВОЛІВ, ОДИНИЦЬ, СКОРОЧЕНЬ І ТЕРМІНІВ

API – Application Programming Interface – інтерфейс прикладного програмування;

BERT – Bidirectional Encoder Representations from Transformers – двонапрямлені представлення кодувальника на основі трансформерів;

BLEU – BiLingual Evaluation Understudy – оцінка відповідності текстів еталону (метрика BLEU);

BPE – Byte-Pair Encoding – кодування пар байтів;

CommonGen – Commonsense Generation – датасет для генерації тексту на основі семантично пов’язаних концептів;

DPO – Direct Preference Optimization – пряма оптимізація преференцій;

GAIL – Generative Adversarial Imitation Learning – генеративно-змагальне імітаційне навчання;

GAN – Generative Adversarial Networks – генеративно-змагальні мережі;

GLUE – General Language Understanding Evaluation – загальна оцінка розуміння мови;

GPT – Generative Pre-trained Transformer – генеративний попередньо натренований трансформер;

GPT-J – варіант GPT з відкритим кодом, орієнтований на продуктивність і якість генерації тексту;

GPT-NeoX – масштабована архітектура GPT з відкритим кодом, орієнтована на тренування великих мовних моделей;

GPU – Graphics Processing Unit – графічний процесор;

ICML – International Conference on Machine Learning – міжнародна конференція з машинного навчання;

IL – Imitation Learning – імітаційне навчання;

LLaMA – Large Language Model Meta AI – велика мовна модель, розроблена компанією Meta;

LoRA – Low-Rank Adaptation – низькорангова адаптація;

Mistral – сімейство ефективних відкритих мовних моделей, орієнтованих на швидкість і продуктивність;

PPL – Perplexity – перплексія (метрика оцінювання мовних моделей);

PPO – Proximal Policy Optimization – проксимальна оптимізація політики;

RL – Reinforcement Learning – навчання з підкріпленням;

RLHF – Reinforcement Learning from Human Feedback – навчання з підкріпленням на основі людського зворотного зв’язку;

RoBERTa – Robustly Optimized BERT Pre-training Approach – робастно оптимізований підхід до попереднього навчання моделі BERT;

ROCStories – Story Cloze Test dataset – датасет для задачі умовно-некерованого продовження коротких історій;

SEO – Search Engine Optimization – оптимізація тексту для пошукових систем;

SFT – Supervised Fine-Tuning – кероване донавчання (за допомогою розмічених даних);

SQuAD – Stanford Question Answering Dataset – датасет Стенфордського університету для задачі відповідей на питання.

ВСТУП

Сьогодні створення природномовних текстів з використанням штучного інтелекту є однією з найбільш актуальних і перспективних задач сучасних наукових досліджень. Інтерес до генерації тексту пояснюється широким діапазоном її застосувань: від вдосконалення машинного перекладу й автогенерації інформаційного контенту до розробки діалогових систем і створення оригінальних творчих творів. Стрімкий розвиток цієї галузі протягом останніх років відбувся завдяки появі великих нейромережевих моделей на базі трансформерної архітектури, зокрема GPT-2, GPT-3 та BERT. Такі моделі дозволяють створювати тексти, які за зв'язністю й граматичною правильністю наблизилися до людського рівня.

Водночас зберігається низка суттєвих труднощів, що обмежують практичні можливості навіть найбільш досконалих генеративних моделей. Серед цих проблем можна відзначити експозиційне зміщення, при якому модель накопичує помилки через суттєву відмінність між умовами навчання та процесом безпосередньої генерації. Крім того, серйозним завданням лишається точна стилістична адаптація тексту до можливостей і манери автора. Моделі, що тренуються за принципом максимізації правдоподібності, незважаючи на високу загальну якість, часто створюють стилістично нейтральні або шаблонні тексти. Причина полягає в тому, що локальні функції втрат не враховують глобальні стилістичні характеристики, які формують індивідуальний авторський стиль. Як наслідок виникає відоме явище «стилістичного розмивання».

Для подолання цих труднощів дослідники пропонують альтернативні шляхи, такі як генеративно-змагальні нейромережі (GAN). Проте використання традиційних GAN для текстової генерації не набуло широкого впровадження через низьку стабільність навчання та складнощі з дискретною природою текстових даних.

Особливий інтерес викликає можливість створення текстів, які не відрізняються від авторських за своїми стилістичними та лексичними характеристиками. Підвищена актуальність такої проблематики обумовлена широким впровадженням великих мовних моделей у таких сферах, як журналістика, освіта та цифровий маркетинг. Наприклад, сервіси Jasper та CopyAI сьогодні активно використовуються у бізнес-середовищі для створення рекламних матеріалів. Автоматична генерація текстового контенту стала також звичною практикою для численних провідних новинних медіа та сучасних освітніх платформ, що суттєво підвищує вимоги до якості, стилістичної гнучкості та адаптивності цих технологій.

Отже, розвиток методів, здатних гнучко й точно відтворювати стилістичні особливості певного автора чи брендового стилю, є надзвичайно важливим та практично значущим завданням.

У межах цього дослідження пропонується підхід TextGAIL, що інтегрує переваги імітаційного навчання та прямої оптимізації преференцій (DPO). Для реалізації цієї ідеї використано два добре зарекомендовані нейромережеві компоненти: генератор тексту на базі GPT-2 та модель преференцій (дискримінатор) на основі RoBERTa. Поєднання цих технологій дозволяє ефективно усунути проблему експозиційного зміщення, забезпечити точну стилістичну адаптацію та значно покращити якість та правдоподібність створюваних текстів.

1 АНАЛІЗ ПРЕДМЕТНОЇ ГАЛУЗІ

Протягом останнього десятиліття автоматизована генерація природних текстів істотно трансформувалася завдяки стрімкому розвитку методів глибокого навчання. Якщо на початку 2010-х років мовні моделі спиралися переважно на статистичні методи із частотними та ймовірнісними підходами, то останнім часом вони були повністю заміщені нейронними архітектурами. На зміну рекурентним нейронним мережам (RNN), які працювали переважно послівно чи навіть посимвольно, у сферу генерації текстів прийшли сучасніші та ефективніші трансформерні моделі. Саме запровадження трансформера, яке було запропоновано Vaswani і співавторами у 2017 році, стало важливим кроком в еволюції генеративних моделій [1]. Це дозволило створювати нейронні мережі нового покоління, здатні навчатися на великих і надвеликих корпусах текстів, формуючи однорідні та змістовні генеровані результати високої якості.

Останні досягнення у створенні генеративних моделей відзначаються не лише зростанням розміру та складності архітектур, але й зміною загальної парадигми їх використання. Якщо раніше головним завданням вважалася безумовна генерація, що полягала у простому формуванні текстових послідовностей, то сьогодні інтерес дослідників і практиків зміщується до генерації контрольованої, яка враховує чітко визначені параметри й атрибути. Мова йде не тільки про логічність, граматичну та семантичну коректність тексту, але й про його відповідність заданим семантичним властивостям, стилістиці чи тематичному наповненню. Наукові групи й компанії активно досліджують методи керування тональністю тексту, його емоційним забарвленням, довжиною та здатністю до інтеграції наперед визначених ключових слів і виразів.

Серед найбільш перспективних та активно вивчених напрямів є стилістичний трансфер. Такий підхід дозволяє моделі ефективно змінювати

стилістичний реєстр тексту, переходячи, наприклад, з офіційно-ділового стилю в художній, при цьому зберігаючи його змістову складову. Водночас розвиваються й більш складні рішення, де стиль і тематична спрямованість текстів задаються не тільки зовнішніми інструкціями, але й спеціальними керуючими токенами та додатковими класифікаторами. Як наслідок, сьогодні генеративні моделі переходять від універсальних та нейтральних генераторів до більш гнучких систем, здатних адаптуватися до специфічних запитів користувача.

Вкрай суттєвим кроком вперед стало також покращення методик вибору слів під час генерації. Простий «жадібний» або випадковий вибір поступився місцем більш складним алгоритмам селективного відбору, таким як ядерна вибірка або вибірка за топ-k. Такі методики дозволяють відкидати повтори, уникати малоймовірних варіантів і формувати різноманітні, логічні та стилістично цілісні послідовності. Це наблизило якість текстів, генерованих автоматично, до такого рівня, що їх складно відрізнити від авторських матеріалів, написаних людиною.

Таким чином, сучасні генеративні моделі стали важливою інфраструктурою для потужних та гнучких рішень у сфері автоматизованої генерації. Водночас, із доступністю таких технологій зростає актуальність завдання глибокої стилістичної адаптації до вимог та особливостей конкретних користувачів і авторів. Можливості, які сьогодні відкриті завдяки інтеграції потужних попередньо навчених трансформерів і різноманітних методів керування текстом, спонукають дослідників до розробки спеціалізованих методик, спрямованих на точну та реалістичну імітацію авторського стилю. Саме на цьому завданні сфокусовано пропоноване дослідження, яке спрямоване на покращення якості генерації тексту шляхом використання гібридного підходу TextGAIL, що поєднує імітаційне навчання та пряму оптимізацію стилістичних преференцій [2].

1.1 Огляд проблеми генерації природномовних текстів

Сучасний стан розробок нейромережевих технологій для автоматичної генерації тексту характеризується суттєвим прогресом щодо зв'язності та структурної проясненості згенерованих послідовностей. Водночас залишаються відкритими і численні виклики: зокрема, проблема забезпечення довготривалої тематичної узгодженості, можливості ефективно керувати змістовими та стилістичними атрибутами, а також розроблення надійних методик для оцінювання якості генерованих текстів. Сучасні нейромережеві моделі здатні успішно освоювати великі навчальні корпуси та ефективно імітувати статистичні й синтаксичні особливості людської мови. Завдяки цьому їх результати генерації на рівні речень і навіть окремих абзаців часто не поступаються за якістю текстам, створеним людиною. Крім того, такі моделі демонструють можливість врахування початкового контексту для побудови текстів, наприклад, описуючи запропоноване зображення або продовжуючи текст, початий людиною.

В останніх наукових працях можна спостерігати численні підходи до оптимізації генерації тексту. До таких належать вдосконалення мережевих архітектур переважно заснованих на трансформерах, покращення втратних функцій (наприклад, доповнення стандартних критеріїв максимізації правдоподібності штрафами за надлишкові повторення чи спеціальними регуляризаторами) та формулювання збалансованих критеріїв навчання.

Генеративно-змагальні мережі (GAN), які активно застосовуються у задачах формування зображень, у сфері текстової генерації поки не набули переваги над традиційними підходами через специфічні особливості даних і принципи отримання градієнтів втрат [3]. Простір текстів є дискретним, і це створює несприятливі умови для стабільної роботи дискримінативних моделей. Як наслідок, моделі такого типу часто демонструють низьку результативність у порівнянні з більш стандартними методиками. Для вирішення окреслених труднощів були запропоновані модифікації GAN-

підходів, зокрема поєднання з навчанням через підкріплення (RL), або попереднє навчання окремих елементів моделей класичними методами.

Окрему увагу приділяють задачам забезпечення вищого ступеня керованості генерації. Контрольованість стилістики є часткою ширшого напряму контрольованої текстової генерації загалом. Існуючі технології адаптації стилістики моделей, як правило, здійснюють або пряме донавчання до конкретного стилю, або додатково навчають модель використовувати тексти певного автора. Попри такі методики, ситуація ускладнюється ризиками втрати загальної узгодженості створених текстів, браком достатньої кількості стилістично зразкових текстів окремого автора і тенденцією моделі до механічного запам'ятовування. У відповідь на це розробники прагнуть використовувати інтегровані рішення, поєднуючи великі попередньо треновані моделі з методами навчання через зворотний зв'язок чи з підкріпленням, що дозволяє моделі генерувати більш природний результат.

Водночас, навіть сучасні підходи не забезпечують впевненого отримання тексту, який був би гарантовано нерозрізнюваним з людським авторським текстом. Серед основних невирішених завдань – експозиційне зміщення, пов'язане з тим, що моделі, навчені за принципом попереднього показу вірних відповідей (методом учителя), не мають змоги виправляти власні помилки, що накопичуються з ростом довжини тексту і поступово призводять до втрати осмисленості. Для вирішення цієї проблеми розробляються спеціальні тренувальні стратегії, такі як запланована вибірка або використання навчання з підкріпленням для коригування генерації цілих речень.

Іншим принциповим викликом є знаходження правильного балансу між осмисленістю та різноманітністю текстів, які створює модель. Без спеціальних заходів нейромережеві моделі постійно повторюють одні й ті самі фрази та структури, втрачаючи оригінальність, тоді як спроби підвищити різноманітність нерідко погіршують логіку. З цієї причини для

оцінки розробляються спеціалізовані метрики, зокрема, вимірювання унікальності n-грам, а також самооцінки подібності текстів, створених однією і тією ж моделлю.

Для генерації тексту, який сприймається як авторський людський твір, важливим є не лише зміст, а й відповідний стиль. Стиль складається з низки глобальних ознак – формальність, емоційний тон, складність мови. Також індивідуальних авторських патернів – лексичні звички, синтаксичні конструкції. Через відсутність загальноприйнятих стилістичних метрик завдання формалізації стилю є складною проблемою, яку зараз розв'язують шляхом застосування класифікуючих моделей або людської експертизи [4]. Це пояснюється насамперед суб'єктивністю сприйняття стилістичних рис, які різняться залежно від особистого досвіду чи культурного контексту. Водночас зростання обсягів цифрового контенту зумовлює потребу у точніших та більш ефективних рішеннях щодо автоматизації таких комплексних завдань.

На відміну від задач з еталонними відповідями, творчі генеративні задачі потребують оцінювання, що відповідає людському сприйняттю, оскільки автоматичні традиційні метрики є недостатніми для точного аналізу.

Застосування стилістичної імітації вже отримує широке практичне використання у таких галузях, як автоматизована журналістика, маркетингові комунікації та освітні технології. Генерація в стилі автора сприяє посиленню автентичності медіаконтенту, спричиняє встановлення ефективніших і більш емоційних зв'язків зі споживачами в рекламі й e-commerce, а також дозволяє створювати персоналізовані та орієнтовані на студента навчальні матеріали у сфері освіти.

Сукупність згаданих чинників зумовлює високу актуальність задачі розробки та вдосконалення системи TextGAIL, яка, у межах даного дослідження, покликана забезпечити ефективне розв’язання проблем стилістичної адаптації автоматично згенерованих текстів.

1.2 Аналіз сучасних підходів до генерації текстів

Розвиток генеративних моделей для тексту пройшов шлях від класичних статистичних методів до сучасних нейронних мереж і великих мовних моделей. Попри значний прогрес у цій сфері, деякі фундаментальні питання досі залишаються відкритими. У цьому розділі розглянуто ключові підходи, що сформували на сьогодні сучасний стан методів генерації тексту. Особливу увагу приділено таким поширеним методам, як максимізація правдоподібності, генеративно-змагальні мережі, а також аналізується важлива проблема експозиційного зміщення, яке значною мірою впливає на підсумкову якість створюваних текстів.

Основу для навчання більшості сучасних мовних моделей складає принцип максимізації правдоподібності. Базова ідея цього підходу полягає в тому, що модель навчається передбачати наступне слово у послідовності, максимізуючи на кожному кроці ймовірність вибору правильного варіанту. Це дозволяє системі засвоїти як локальні, так і глобальні контекстуальні залежності, створюючи граматично й мовно узгоджені речення.

Метод максимізації правдоподібності вирізняється простотою реалізації, стабільністю навчання та добрим теоретичним обґрунтуванням. Він легко формулюється як задача оптимізації та ефективно вирішується стохастичним градієнтним методом, досягаючи низьких значень таких класичних метрик як перплексія.

Попри це, підхід має низку суттєвих обмежень. Одним із ключових є експозиційне зміщення. Воно пов'язане з тим, що в тренувальний період модель спирається лише на правильний контекст з навчальних даних. Але під час генерації модель використовує власний згенерований текст як контекст для наступних слів. В результаті виникають ситуації, яких модель не зустрічала під час навчання, а це викликає накопичення помилок, проблеми повторів або втрати логіки у згенерованому тексті.

Крім цього, модель, натренована за методом максимальної правдоподібності, оптимізує лише ймовірності окремих слів, але це не завжди відповідає високій якості фінального тексту в цілому. Вона зазвичай орієнтована на найбільш частотні, типові та ймовірні слова й речення, тому результати часто є шаблонними та нецікавими. Попри зазначені недоліки, саме максимізація правдоподібності залишається базовим методом, з якого починають розвиток більш складних підходів.

Однією з таких альтернатив є генеративно-змагальні мережі (GAN). Розроблені спочатку для генерації зображень, GAN були адаптовані й для генерації текстів. Основна ідея GAN полягає у взаємодії двох конкурентних нейромережевих компонентів – генератора, який намагається створити реалістичний текст, та дискримінатора, завданням якого визначити наскільки отриманий текст є реальним або штучним.

Для текстових задач GAN дозволяють оцінювати якість всієї послідовності цілком, а не лише окремих слів чи n-грам. Утім, застосування GAN для текстів стикається з проблемою дискретності. На відміну від зображень, текст складається з окремих дискретних символів або токенів, що ускладнює пряме передавання градієнта дискримінатора генератору. Для боротьби з цією проблемою були розроблені рішення, такі як модель SeqGAN [5], що використовує політичний градієнт для розрахунку винагороди, модель MaliGAN [6], у якій модифіковано функцію винагороди для зменшення варіативності градієнтів, модель LeakGAN [7], де генератор отримує від дискримінатора додаткову проміжну інформацію або модель RelGAN [8], що пропонує складніші структури архітектури для генерації реалістичних текстів.

Втім, незважаючи на різноманітність рішень, GAN-підходи для тексту все ще мають суттєві технічні проблеми. Навчання таких мереж залишається складним і нестабільним, сильно залежить від параметрів та архітектурних виборів, а результати демонструють значну варіативність. Попри це, GAN продовжують вважатися потенційно перспективним

напрямом досліджень, особливо в комбінації з іншими підходами, такими як навчання з підкріпленням.

Що ж стосується проблеми експозиційного зміщення, то поряд з GAN для боротьби з нею пропонували й інші методи. Одним з таких є плановий відбір, коли модель поступово переходить від справжніх слів з навчальної множини до самостійно генерованих контекстів. Ще один підхід – модель MIXER, яка поєднує класичне навчання з максимізацією правдоподібності і поступово переходить до оптимізації цілих речень за допомогою навчання з підкріпленням. Метод професорського форсування також забезпечує вирівнювання прихованих просторів моделі на тренувальних та генеративних режимах роботи.

Саме до цієї групи рішень також належить і модель TextGAIL, реалізована у даній роботі. Вона використовує GAN-архітектуру спільно з імітаційним навчанням для покращення стилістичної відповідності тексту та вирішення проблеми експозиційного зміщення.

Отже, у даний момент існує цілий спектр методів, що спрямовані на вирішення нагальних проблем в області генерації текстів – від класичних та формальних до складних та інноваційних. Від обраного підходу залежить специфіка конкретних задач, ресурсні витрати та якість згенерованих результатів. Вибір найбільш відповідних методів визначається, насамперед, конкретними задачами генерації тексту та критеріями якості отриманих результатів, що відкриває широкі можливості для подальших експериментів та розробок.

1.3 Аналіз великих попередньо навчених мовних моделей

Протягом останніх років поява потужних мовних моделей на базі трансформерної архітектури кардинально вплинула на стан задач аналізу та створення текстового контенту. Серед найбільш відомих – моделі GPT-2 [9], GPT-3 [10], BERT [11], RoBERTa [12]. Значна ефективність таких моделей

зумовлюється тренуванням на надвеликих текстових корпусах (десятки або навіть сотні гігабайтів інформації). Це дозволяє їм результативно засвоювати загальні мовні закономірності, накопичувати великий обсяг знань, а також контекстуальних взаємозв'язків. Завдяки кількості параметрів, що сягає сотень мільйонів чи десятків мільярдів, такі системи можуть без спеціального донавчання створювати природні, логічні й структуровані тексти у різних стилях та тематичних напрямках.

Однією з перших моделей, яка продемонструвала практичні результати такого підходу, стала GPT-2. Ця система показала ефективність у завданнях генерації, таких як продовження тексту, відповідь на запитання й навіть переклад – без додаткового донавчання під конкретну задачу. GPT-2 є авторегресивною трансформерною моделлю, яка генерує текст послідовно, використовуючи як контекст уже згенеровану інформацію. Її тренування відбувалось на корпусі WebText обсягом понад 40 гігабайтів. Випущена у 2019 році, GPT-2 мала від 117 мільйонів до 1,5 мільярдів параметрів, що на той час було революційним і дало їй змогу досягти небувалої різноманітності й логічності сформульованих текстів.

Подальший розвиток цього напрямку продовжила модель GPT-3, яка містить близько 175 мільярдів параметрів та здатна вирішувати завдання лише на основі кількох прикладів (few-shot learning), без додаткового налаштування параметрів. Завдяки такому масштабу, GPT-3 стала універсальною генеративною системою для текстів різної тематики та стилів. Водночас саме модель GPT-3 допомогла чітко окреслити серйозну проблему великих мовних систем – так звані «галюцинації», тобто схильність створювати факти, події чи інформацію, яка відсутня в реальних джерелах.

Паралельно із розвитком генеративних технологій прогресували і так звані дискримінативні моделі, такі як BERT і RoBERTa. На відміну від моделей родини GPT, вони не призначені для створення нових текстів напряму. Їхнім завданням є побудова високоякісних векторних

представлень слів та підслів (токенів), що необхідно для розв'язання задач класифікації, інформаційного вилучення чи надання відповідей на запитання. Прикладом таких удосконалень стала RoBERTa від компанії Facebook, що тренувалась на значно більших масивах текстів (до 160 ГБ). У цій моделі використовується задача динамічного маскування токенів, яка дала змогу істотно перевершити результати попередників (включаючи оригінальний BERT) у стандартизованих задаче-тестах типу GLUE та SQuAD.

Сучасні практичні розробки активно використовують комбінацію генеративних та дискримінативних підходів. Наприклад, запропонована модель TextGAIL успішно поєднує генеративні можливості GPT-2 з потужністю дискримінативної моделі RoBERTa. Такий підхід дозволяє успішно розв’язувати серйозну проблему ранніх моделей (наприклад, SeqGAN), у яких генератори тренувалися «з нуля» і демонстрували низьку якість формованого тексту.

Однак сучасні великі мовні моделі усе ще мають низку типових недоліків: «галюцинації», повторення однотипних сегментів тексту, логічні суперечності, труднощі з підтримкою заданих стилістичних обмежень. Для боротьби з цими проблемами активно розвиваються й удосконалюються методи «керованого» навчання: навчання за інструкціями, а також методики на основі навчання з підкріпленням та із людським експертним зворотним зв’язком.

При донавчанні великих мовних моделей відзначають два протилежних ризики – недостатня та надмірна адаптація параметрів. Недостатньо адаптована модель погано засвоює стилістичні особливості нових даних, а надмірна адаптація навпаки призводить до механічного відтворення навчальних прикладів, втрати логічності й оригінальності текстів. Для уникнення цих небажаних результатів застосовують мінімізацію швидкості навчання, а також додаткові стратегії регуляризації моделі.

Поза увагою класичного огляду моделей типу GPT та RoBERTa останнім часом було створено альтернативні, менш ресурсомісткі варіанти: LLaMA, Falcon, Mistral, GPT-J та GPT-NeoX від спільноти EleutherAI, що відзначаються доступними ресурсними вимогами завдяки відкритому коду та при цьому забезпечують гарну якість текстів. Їхнє використання відкриває перспективу для зниження апаратних ресурсів та розширення можливостей адаптації великих мовних моделей під конкретні завдання чи галузі.

Важливими напрямами досліджень вважають також інтегровані методики, здатні поєднувати сильні сторони окремих алгоритмів, і створення спеціалізованих архітектур, які матимуть знижений рівень «галюцинацій» завдяки спеціальній регуляризації вихідних даних. Такі підходи можуть підвищити практичну якість і безпеку використання великих мовних моделей у реальних задачах, що є одним із пріоритетних інтересів мовних технологій наступного покоління.

1.4 Постановка задачі дослідження

Задача автоматичної генерації тексту, який реалістично відтворює стилістичні особливості конкретного автора, є надзвичайно актуальною з теоретичної і практичної точок зору. На відміну від класичних завдань створення тексту з правильною граматичною структурою чи релевантним змістом, у цій задачі стоїть завдання глибшого моделювання індивідуального авторського стилю. Потрібно відобразити характерний синтаксис, ідіостиль, типову лексику, властиву тональність і навіть ритмічні особливості письмового мовлення автора. Саме ці елементи формують помітну унікальність авторського почерку і забезпечують впізнаваність стилю.

Традиційні підходи до побудови генеративних моделей переважно орієнтовані на максимізацію правдоподібності послідовностей слів. Хоча це

й дозволяє отримувати граматично точні й зрозумілі тексти, часто виникає явище «стилістичного розмивання». Створювані тексти мають узагальнений, нейтральний стиль, оскільки оптимізується головним чином імовірність появи конкретного слова чи короткої послідовності. Це ігнорує важливі глобальні стилістичні ознаки, властиві авторському стилю всього текстового корпусу.

Це обмеження традиційних методів зумовлює потребу у пошуку нових підходів для ефективної генерації текстів з чітко визначеними стилістичними рисами. Одним із перспективних шляхів розв'язання цієї проблеми є метод імітаційного навчання, коли модель тренується робити послідовності тексту за реальними прикладами «демонстраціями» та поступово наближає стиль власного тексту до заданого еталона. Щоб додатково покращити точність такого підходу, використовується генеративно-змагальна архітектура, де дискримінатор розрізняє справжні й згенеровані авторські тексти, а генератор поступово корегує свої параметри завдяки отримуваному зворотному зв'язку для досягнення стилістичної автентичності.

Вибір оптимальної архітектури генератора й дискримінатора суттєво впливає на підсумкову ефективність розв’язання поставленої задачі. Сучасна генеративна модель GPT-2 як трансформерна структура, що вже має узагальнені мовні уявлення, добре підходить на роль генератора. Для дискримінатора найбільш відповідними є потужні трансформерні енкодери типу RoBERTa, які демонструють високу чутливість до тонких стилістичних і контекстуальних особливостей текстів.

У рамках цього дослідження метою є розробка ефективного методу автоматичної генерації стилізованих авторських текстів з використанням підходу імітаційного навчання у поєднанні з генеративно-змагальною архітектурою [13]. Для досягнення цієї мети передбачено аналіз сучасних технологій у галузі стилістичної генерації тексту, формалізацію архітектури запропонованого методу, створення й тестування програмного прототипу, а

також порівняльний аналіз результатів із класичними підходами, зокрема зі стандартним донавчанням GPT-2.

Об’єктом дослідження є процес автоматизованої генерації тексту із заданими авторськими стилістичними ознаками.

Предметом дослідження виступають конкретні алгоритмічні рішення, методи та моделі, призначені для імітації авторського стилю засобами імітаційного навчання з генеративно-змагальним підходом і адаптації великих трансформерних моделей для цієї задачі.

Варто зазначити, що успішна реалізація запропонованого підходу може мати важливі практичні наслідки в різноманітних сферах: створення автентичного брендового контенту для маркетингу, персоналізація освітніх курсів з індивідуальним голосом викладача, стилістична реставрація історичних текстів у цифровій гуманітаристиці, а також загальне покращення якості й точності автоматичних мовних систем. Саме цим і зумовлена важливість і актуальність проведення цього дослідження.

2 ТЕОРЕТИЧНІ ОСНОВИ ЗАПРОПОНОВАНОГО ПІДХОДУ

Запропонований підхід базується на ідеях генеративно-змагального імітаційного навчання (GAIL) та методу прямої оптимізації преференцій (DPO) [14]. У своїй основі він використовує попередньо натреновані трансформер-моделі – генеративну GPT-2 та дискримінативну RoBERTa. Для забезпечення ефективності та стабільності навчання, його реалізація передбачає поетапне налаштування моделей, що описано нижче.

2.1 Архітектура системи TextGAIL

Запропонована система передбачає двокомпонентну архітектуру, яка складається з:

– генеративної нейромережевої моделі, відповідальної за безпосередню генерацію стилістично відповідних текстів;

– дискримінативної нейромережевої моделі, яка оцінює відповідність створених текстів заданому стилю.

Для реалізації цих задач залучаються трансформерні моделі GPT-2 (генератор) та RoBERTa (дискримінатор преференцій). Ефективність підходу визначається правильним вибором і початковим налаштуванням саме цих моделей, що докладно описано у подальших підрозділах.

2.1.1 Навчання базової генеративної моделі GPT-2

Як основу генеративного компоненту використано модель GPT-2 – трансформер-декодер. Вибір саме цієї моделі продиктований її високою результативністю в генерації плавних і граматично правильних текстів різних жанрів та наявністю ефективних попередньо навчених версій.

Архітектурно GPT-2 являє собою автотрансформер з авторегресійним механізмом генерації, за якого кожен наступний токен передбачається,

спираючись на усі попередні позиції. Найбільш важливим компонентом є багатоголовий механізм самоуваги, який дозволяє враховувати довгострокові зв'язки між словами. Для авторегресивної генерації у декодері GPT-2 використовується маскована самоувага, де при генерації токена модель може враховувати лише попередні токени.

Крім механізму уваги, трансформер містить нейронні мережі прямого поширення, механізми порядкового позиційного кодування, а також залишкові з'єднання і нормалізацію шарів, які пришвидшують стабільність навчання та збіжність моделі.

Для даного дослідження вибрана версія GPT-2 масштабу small з приблизно 117 мільйонами параметрів. Така конфігурація є оптимальною з погляду співвідношення продуктивності та обчислювальних ресурсів, забезпечуючи хорошу базу для експериментів без необхідності значних обчислювальних витрат та тривалого часу виконання.

Стандартне тренування GPT-2 здійснюється на корпусах типу WebText, що задає моделі потужний попередній рівень знань про загальні мовні закономірності. Проте для специфічного донавчання під конкретний авторський стиль цих загальних знань недостатньо. Таким чином, перший етап навчання є донавчанням моделі на менших, проте стилістично репрезентативних вибірках текстів, намагаючись максимально точно відтворити специфічні риси цільового автора. Процес налаштування ґрунтується на мінімізації перехресної ентропії, що відповідає стандарту для задач мовного моделювання.

У процесі підготовки та роботи з текстом використовується система токенізації на основі кодування пар байтів (BPE). Токенізатор GPT-2 має словник на приблизно 50 тисяч токенів, що дозволяє ефективно працювати зі специфічною авторською лексикою, зменшуючи кількість невідомих токенів та покращуючи стилістичну відповідність.

Безпосередній етап донавчання реалізується за допомогою бібліотеки Hugging Face Transformers, функціонал якої дозволяє зручно організовувати

та контролювати процес тренування. Донавчання є досить ресурсоємним процесом і потребує GPU з відповідним обсягом відеопам'яті.

Завершений етап донавчання забезпечує хорошу початкову адаптацію моделі GPT-2 до конкретного стилю автора та створює основу для подальшої преференційної оптимізації моделі.

2.1.2 Навчання дискримінативної моделі преференцій RoBERTa

Після формування першого компонента, другий етап присвячено створенню дискримінатора преференцій. Основною задачею цієї моделі є чітке та стабільне оцінювання генерованих GPT-2 текстів на предмет стилістичної ідентичності автентичним зразкам.

Для цієї ролі обрано RoBERTa, поліпшений варіант BERT, що також побудований на енкодерній структурі архітектури трансформера. RoBERTa вирізняється такими перевагами над BERT, як динамічне маскування в процесі тренування, більший словник BPE та відмова від менш ефективної задачі прогнозування наступних речень.

Процес тренування RoBERTa також є донавчанням, але у форматі бінарної задачі класифікації. Модель навчається розрізняти оригінальні авторські тексти і тексти, створені моделлю GPT-2 після першого етапу.

Як цільова функція використовується бінарна перехресна ентропія. Для реалізації цього навчання також застосовується бібліотека Hugging Face Transformers. Результатом цієї роботи є дискримінатор, що точно визначає відповідність текстів цільовому стилю, на базі чого можна сформувати якісний датасет для наступного етапу прямої оптимізації преференцій.

Для оцінки ефективності моделі RoBERTa використовуються стандартні метрики точності, повноти, прецизійності й F1-міри, що дають уявлення про стабільність класифікатора. Завершений другий етап формує дискримінативну основу, що використовується як оцінювач стилю на

наступних етапах роботи, включаючи фінальний процес преференційної оптимізації генеративної моделі.

2.2 Алгоритм навчання моделі TextGAIL

Після завершення навчання ключових компонентів – базової версії генератора (GPT-2 SFT) та дискримінатора стилістичних преференцій (RoBERTa) – подальшим кроком є опис алгоритму остаточного тренування генеративної моделі. Ключова мета фінального етапу це надати моделі GPT-2 здатність стабільно створювати ті варіанти текстів, що вибираються як кращі відповідно до стилістичних та якісних критеріїв, визначених дискримінатором RoBERTa. Алгоритм остаточного навчання генеративної моделі TextGAIL включає кілька важливих послідовних кроків, що представлені у вигляді схеми на рисунку 2.1.

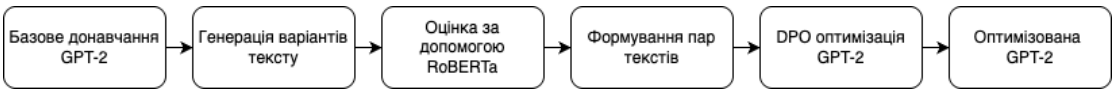

Рисунок 2.1 – Схема алгоритму навчання моделі TextGAIL

2.2.1 Формування набору даних преференцій для DPO

Для реалізації прямої преференційної оптимізації (DPO) необхідно згенерувати спеціалізований набір навчальних прикладів. Такий набір складається з трійок, що включають початковий контекст (prompt) та два згенеровані тексти: один, якому надається перевага (chosen), і один, який оцінюється як менш вдалий (rejected). Роль цих трійок – забезпечити модель GPT-2 прямими орієнтирами щодо бажаних і небажаних результатів генерації тексту.

Алгоритм побудови кожної трійки передбачає створення кількох варіантів відповідей генеративною моделлю GPT-2, попередньо налаштованою за допомогою контрольованого донавчання. Після цього всі варіанти відповіді автоматично оцінюються за допомогою моделі преференцій RoBERTa. Результати оцінювання використовуються для формування пар текстів – кращий з високою оцінкою та гірший з низькою.

У межах роботи розглянуто три основні стратегії комплектування навчального корпусу залежно від способу вибору текстів. Перший варіант передбачає створення пар, у яких еталонний «реальний» текст автора зіставляється зі згенерованим, причому «реальний» текст позначається як той, якому надається перевага. Перевагою такого підходу є чіткий і сильний навчальний сигнал, його недоліками є низька гнучкість та потреба в наявності достатньої кількості авторських текстів.

Другий підхід до комплектування полягає у виборі найкращого та найгіршого варіантів тексту з декількох, що створює модель GPT-2. Текст із найвищою оцінкою дискримінатора RoBERTa є кращим, із найнижчою – гіршим. Цей підхід гнучкіший, оскільки дозволяє формувати дані без використання еталонних текстів, але його якість сильно залежить від точності роботи дискримінативної моделі.

Третій комбінований підхід базується на спробі об'єднати переваги двох попередніх шляхом додаткової фільтрації створених пар відповідно до різниці в оцінках RoBERTa для текстів у парі. Такий метод потенційно забезпечує якісний сигнал, хоча включає додатковий етап та ускладнює формування корпусу.

Вибір конкретного способу побудови залежить від експериментального контексту та наявності текстових ресурсів, з якими дослідники працюють. У будь-якому разі отриманий корпус текстових пар повинен відповідати високим стандартам якості, оскільки неточне формування оцінок може негативно вплинути на кінцеве налаштування генеративної моделі.

Процес створення такого набору даних є ресурсоємним, оскільки потребує значного обсягу обчислень: багаторазових проходів генеративної моделі GPT-2 з наступною детальною оцінкою текстів дискримінатором RoBERTa. Підсумком роботи стає корпус попарних преференцій, який дозволяє здійснити ефективну та стійку оптимізацію генеративної моделі, що і реалізується у подальшому етапі прямої оптимізації преференцій (DPO).

Під час формування корпусу парних преференцій особливу важливість має вибір оптимального методу подальшої оптимізації генеративної моделі, оскільки створений набір буде використовуватись саме для остаточного налаштування. У цьому контексті необхідним є коротке порівняння та обґрунтування того, чому в даному дослідженні для фінального етапу навчання використано саме метод прямої оптимізації преференцій (DPO) замість класичних алгоритмів навчання з підкріпленням на основі людських оцінок (RLHF).

RLHF вимагає створення додаткової проміжної моделі винагороди, яка проходить окремий цикл налаштування й лише потім використовується для тренування основної моделі-генератора. Цей підхід є багатоступінчастим, складним та вимагає значних витрат часу й ресурсів.

У свою чергу, метод DPO не передбачає жодної проміжної моделі, а напряму використовує пари текстів, що описані вище. Це значно спрощує архітектуру системи і зменшує ресурсні затрати для остаточного донавчання. З огляду на ці практичні й теоретичні переваги, метод DPO і був обраний як остаточний для подальшого налаштування генеративної моделі TextGAIL.

Детальний опис реалізації підходу прямої оптимізації преференцій, особливостей архітектури та вибору її ключових гіперпараметрів наведено нижче.

2.2.2 DPO-донавчання генеративної моделі GPT-2

Метод прямої оптимізації преференцій (DPO), який застосовується у цій роботі, є сучасною альтернативою традиційним способам навчання з підкріпленням на основі людських оцінок (RLHF, PPO). Якщо класичні RL-методи потребують тренування додаткової нейромережі «винагороди», яка оцінює якість згенерованих текстів, то підхід DPO уникає цього складного й ресурсомісткого етапу, пропонуючи натомість прямий механізм оптимізації.

Суть ідеї DPO – у безпосередньому використанні сформованих пар текстів (обраний і відхилений) для прямого налаштування головної моделі-генератора. Метод орієнтований на оптимізацію ймовірностей генерації тих текстів, які відзначені як успішні, й одночасне зменшення ймовірностей варіантів текстів, що були позначені як невдалі. Інтуїтивно ідея полягає в тому, що модель не витрачає ресурси на створення проміжної моделі оцінювання текстів, а одразу отримує прямий сигнал, який вказує, які варіанти тексту потрібно генерувати частіше, а яких уникати. Такий безпосередній спосіб оптимізації робить процес навчання прозорішим, ефективнішим та легшим в управлінні.

На рисунку 2.2 графічно показано ключову різницю між класичним методом навчання з підкріпленням на основі людських оцінок (RLHF) і методом прямої оптимізації преференцій (DPO), обраним для реалізації в цій роботі. На відміну від RLHF, що потребує додаткового етапу тренування проміжної моделі оцінок, метод DPO передбачає пряме налаштування параметрів генератора на основі вже оцінених парних текстових варіантів. Це значно спрощує загальну архітектуру та зменшує ресурсні витрати.

Рисунок 2.2 – Порівняння схем навчання моделей із використанням RLHF та DPO

Заключний етап тренування моделі GPT-2 у цьому дослідженні полягає безпосередньо у застосуванні описаного DPO-підходу. Для реалізації цього етапу коректно сформований набір текстових пар використовується як основа процесу навчання. Архітектурно DPO-донавчання передбачає одночасну роботу двох версій моделі GPT-2:

– базова GPT-2 – основна модель, яку оптимізують і налаштовують;

– референтна GPT-2 – модель-регуляризатор із фіксованими параметрами, встановленими на початку процесу навчання і незмінними в процесі тренування.

Під час тренування базова модель GPT-2 корегує параметри таким чином, щоби збільшувати ймовірність текстів, позначених як «кращі», і зменшувати ймовірність тих, що були «відхилені». Для того, щоб модель не надто далеко відходила від початкової базової моделі, додатково вводиться регуляризаційний параметр β, який безпосередньо контролює ступінь відхилення. Саме вибір правильних значень гіперпараметрів є ключовим фактором ефективності методу DPO. Найважливішими гіперпараметрами на цьому етапі є:

– β (beta) – регуляризаційний коефіцієнт, зазвичай від 0.1 до 0.5;

– learning rate – швидкість навчання (зазвичай менше значення, ніж стандартне донавчання – порядку 1e-6 – 5e-6);

– кількість епох донавчання – зазвичай 1 – 3;

– розмір пакета, що визначається стабільністю навчання та GPU-ресурсами.

Процес такого навчання є досить ресурсоємним, оскільки тексти одночасно проходять обробку двома версіями моделі. Для контролю й моніторингу процесу використовуються спеціальні метрики, зокрема значення цільової втратної функції і середньої маржі переваг між обраними і відхиленими зразками у наборі парних порівнянь.

На рисунку 2.3 представлено алгоритмічну схему реалізації DPO-методу. Спочатку система завантажує початкові дані та моделі, перевіряє, чи є уже навчені версії GPT-2 та RoBERTa, та за потреби тренує їх і зберігає. Далі формують спеціальний набір парних преференцій для DPO. Фінально DPO-алгоритм оптимізує GPT-2. Зрештою, здійснюють генерацію текстів, оцінювання результатів і збереження отриманих метрик.

Таким чином переваги методу DPO обумовлюють вибір цього підходу як найефективнішого у контексті стилістично керованої генерації тексту. DPO поєднує простоту реалізації, прозорість алгоритмічного рішення та менші ресурсні витрати порівнянно з RLHF-методами. Як продемонструють подальші експерименти, фінально налаштована модель GPT-2 демонструє хороший рівень стилістичної точності, природності й узгодженості текстів відповідно до критеріїв, сформульованих дискримінативною моделлю RoBERTa.

Запропонований підхід, що поєднує финальну оптимізацію GPT-2 за допомогою дискримінатора RoBERTa, теоретично дозволяє досягти високого рівня стилістичної відповідності та логічної узгодженості генерованих текстів. Очікується, що створювані таким чином тексти будуть характеризуватись авторським стилем, автентичністю та природністю, що підлягає подальшій експериментальній перевірці.

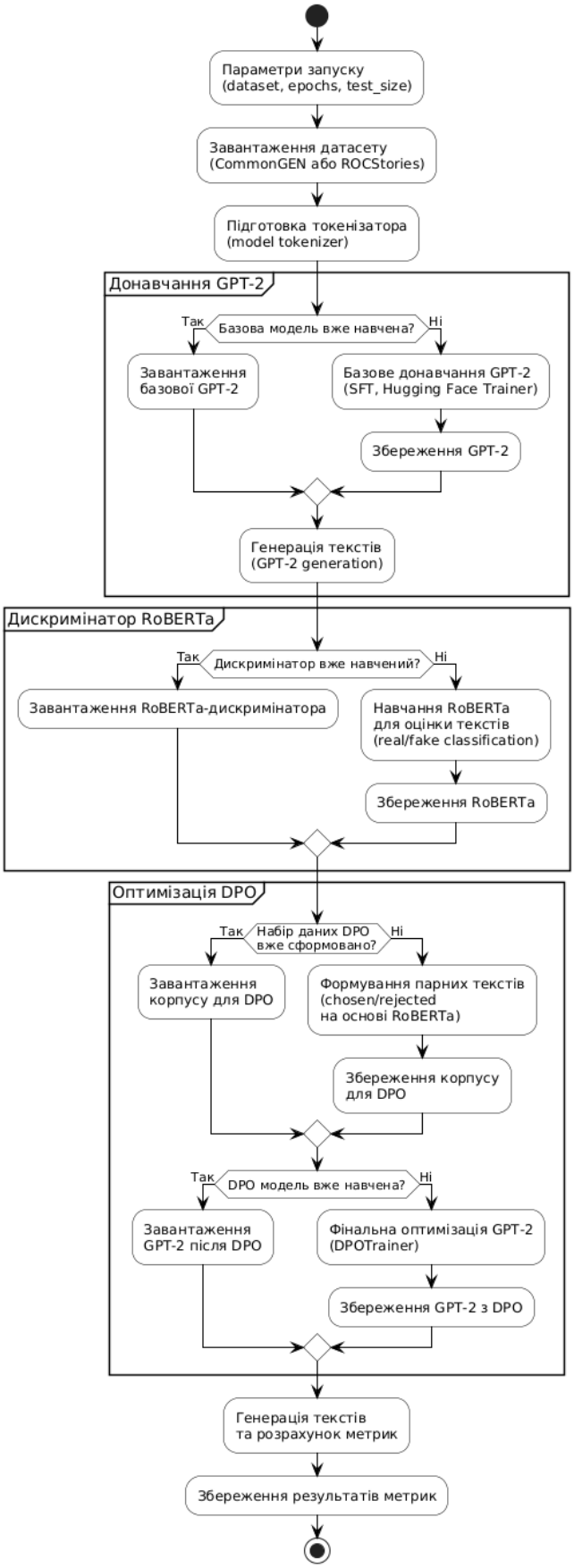

Рисунок 2.3 – Схема програмної реалізації навчального процесу TextGAIL

2.3 Порівняння TextGAIL з альтернативними підходами до стилізації тексту

Запропонована система TextGAIL має ряд істотних особливостей у порівнянні з традиційними та актуальними альтернативними технологіями стилізації й генерування текстів. Класичні генеративно-змагальні мережі GAN, а також традиційні підходи GAIL для текстових задач застосовують ітеративний процес одночасного тренування генератора й дискримінатора, що часто призводить до нестабільності процесу оптимізації. У таких умовах можливе виникнення дефектів навчання, таких як колапс режимів, розбіжність моделей або циклічне коливання без істотного покращення результатів. На відміну від цього запропонований підхід TextGAIL будується за чітко послідовною схемою, де кожен із чотирьох основних етапів навчання здійснюється автономно. Такий підхід дозволяє уникнути типових проблем стабільності класичних GAN-рішень, а завдяки прямій оптимізації преференцій (DPO) суттєво зменшується ймовірність колапсу й деградації генерації.

Що стосується порівняння з іншими варіантами генеративно-змагального імітаційного навчання (GAIL з PPO), запропонований TextGAIL має простішу й прозорішу підготовчу та навчальну стратегію. Він відмовляється від дорогого з точки зору обчислень і складного в налаштуванні алгоритму PPO [15], пропонуючи натомість прямий принцип оптимізації преференцій, що не вимагає навчання окремої моделі винагород й забезпечує значно меншу ресурсоємність навчального процесу.

Іншим розповсюдженим сучасним методом налаштування великих мовних моделей на заданий стиль є навчання з підкріпленням на основі людських оцінок (RLHF) [16], яке вже стало стандартом для багатьох великих систем, таких як Anthropic Claude або GPT-4. Порівняно із підходами RLHF, TextGAIL має спрощену архітектуру, адже не містить окремої складної моделі винагороди й уникає необхідності залучати

зовнішніх людських оцінювачів – замість цього база для оптимізації складається автоматично на етапі формування пар текстів із подальшою оцінкою дискримінатора RoBERTa. Втім, саме ця особливість водночас є певним потенційним обмеженням запропонованого методу, оскільки фінальна ефективність оптимізації прямо залежить від точності автоматичної моделі преференцій, яка не завжди може повністю відповідати реальним людським оцінкам.

У порівнянні з техніками перенесення стилю, які зазвичай використовують механізми кодування й декодування для заміни одного стилістичного компонента на інший, TextGAIL пропонує кардинально інший підхід. Тут модель генерує текст у заданому стилі безпосередньо, з поступовою оптимізацією якості на основі порівняльних парних оцінок. Через це модель має перевагу у змістовій цілісності згенерованих текстів, адже в початковій генерації використовується вже сильно налаштована GPT-2, що ефективно зберігає зміст і структуру генерованих зразків. Стандартні ж техніки трансферу стилю часто стикаються з проблемою адаптації стилю коштом втрати початкового змісту тексту.

Перевагою TextGAIL також є його гнучка здатність до масштабування. На відміну від технік стилістичного трансферу, які вимагають створення якісного латентного простору та перенавчання моделей для кожного нового стилю, використання попередньо навчених трансформерних архітектур дозволяє проводити лише помірне донавчання для кожного нового автора чи стилістичних умов без значних структурних змін, спрощуючи процес масштабування системи.

Що ж стосується порівняння із поширеними методами швидких налаштувань великих генераторів тексту через подання кількох прикладів у промпті, то TextGAIL пропонує більш фундаментальну адаптацію. Модель не потребує додавання текстових зразків стилю в промпті під час генерації, завдяки чому вільний обсяг контекстного вікна може бути повністю використаний безпосередньо для виконання основного завдання. Крім

цього, повноцінне навчання й налаштування GPT-2 через преференційні оцінки дозволяє досягти значно глибшого й тоншого наслідування стилю у порівнянні зі швидкою поверхневою адаптацією узагальнених промптів.

Тим не менш, модель TextGAIL має і певні обмеження. Для її ефективної реалізації необхідне створення репрезентативного стилістичного корпусу текстів, що є помітною трудністю у практичних умовах, особливо у порівнянні з простішими підходами, для яких достатньо кількох прикладів. Також, хоча технологічні витрати на тренування через DPO значно нижчі, ніж в RLHF, вони все ще перевищують найпростіші техніки формування стилю через промпт.

Узагальнюючи порівняльний аналіз, TextGAIL посідає унікальне місце серед сучасних підходів до стилізації тексту. Він поєднує високу результативність великих попередньо навчених трансформерних моделей із меншою ресурсомісткістю та більшою навчальною стабільністю завдяки оптимізації на базі автоматичних преференцій, забезпечуючи збалансоване рішення під задачі генерації тексту у стилях конкретних авторів.

2.4 Обмеження та потенційні напрямки вдосконалення підходу TextGAIL

Запропонований підхід TextGAIL, попри очевидні переваги, має низку суттєвих обмежень, усвідомлення яких є важливим для ефективного практичного використання методу та визначення перспективних напрямків його доопрацювання.

Першим очевидним обмеженням роботи системи є вимоги до обчислювальних ресурсів. Незважаючи на спрощення порівняно з PPO-алгоритмами, процес навчання все ще потребує значного об'єму пам'яті графічних процесорів. Це пов'язано з необхідністю одночасного збереження базової та референтної моделей у процесі преференційної

оптимізації (DPO), а також з вимогами до ресурсів під час підготовки набору преференцій, що передбачає генерацію великої кількості кандидатів.

Іншим помітним обмеженням є поточне використання моделей досить скромного масштабу GPT-2 та RoBERTa-base. Масштабування системи до більш сучасних великих нейромережевих архітектур, які мають мільярди параметрів, несе істотні складності. Таке масштабування може призвести до необхідності використання методів розподіленого навчання, ускладнення налаштування, збільшення витрат часу і ресурсів, а також погіршення продуктивності порівняно з меншими моделями.

Важливим аспектом ефективності всієї системи є точність роботи дискримінативної моделі преференцій. Якщо дискримінатор недостатньо ефективно виявляє тонкі нюанси стилю чи має суттєві стилістичні упередження, ці недоліки неминуче передаються генератору, впливаючи на кінцеву якість текстів. Крім того, використання автоматичної моделі преференцій не дозволяє цілком ідентифікувати суб'єктивні моменти стилістики, які можуть бути важливими з позиції людської оцінки. Як наслідок, навіть найретельніше навчений дискримінатор може не повністю відповідати сприйняттю стилістичної відповідності людиною, що додатково ускладнює процес автоматичної оцінки отриманих результатів.

Ще однією характерною особливістю і обмеженням усіх моделей, що навчаються імітаційно, є потенційне зниження їхньої креативності. Підхід, за якого модель навчається наслідуванню стилістичних характеристик наявних текстів, передбачає певний рівень стилістичного консерватизму, внаслідок чого в модель фактично закладається певний «стилістичний інерційний слід». Це може спричинити недостатню інноваційність та труднощі генерації на теми або в контекстах, які значно відрізняються від тих, що представлені у тренувальному наборі. Також є ризик перенавчання: якщо стиль моделі прив'язаний занадто вузько до експериментальних текстових корпусів, то окрім стилю модель може отримати небажані особливості тематичних обмежень оригінальних текстів автора.

Крім цього, сама сутність стилю є комплексним, багатошаровим і частково суб’єктивним поняттям. Відхилення від його реального людського сприйняття може бути значним навіть серед людей-експертів. Це створює труднощі об'єктивного оцінювання стилістичної точності моделі лише за допомогою автоматичних метрик, які навіть за умов коректного налаштування не завжди здатні повністю адекватно відобразити якість згенерованих текстів.

На основі зазначених обмежень можна сформулювати перспективні напрямки для подальших досліджень і вдосконалення. Так, цікавим кроком могла б бути реалізація двоступеневого підходу до оцінювання преференцій, коли автоматична модель поєднується з обмеженим залученням людського зворотного зв’язку. Також доцільним напрямком може стати розробка більш деталізованої моделі стилістичної оцінки, яка б аналізувала стиль не інтегрально, а за окремими його аспектами, такими як синтаксис, лексика, тематична спеціалізація або ритміка.

Варто звернути увагу на механізми поступового навчання, які дозволили б легко й ефективно оновлювати модель додатковими текстовими матеріалами без повторного повного навчання, що потенційно необхідно при змінах авторського стилю з часом. Перспективним напрямком є також інтеграція методів контрастивного навчання або додаткових регуляризаторів для посилення розрізнювання тонких стилістичних відмінностей авторів і підвищення точності та специфічності генерації.

Таким чином, враховуючи визначені межі використання системи та описані можливі способи поліпшення її продуктивності, подальше дослідження підходу TextGAIL матиме ціль усунути або пом'якшити його обмеження, забезпечуючи ще ширшу сферу практичного застосування та кращу відповідність реальним умовам стилізованої генерації текстів.

3 ЕКСПЕРИМЕНТАЛЬНА ПЕРЕВІРКА TEXTGAIL

3.1 Методологія експериментів

З метою ґрунтовної перевірки ефективності запропонованого у роботі методу TextGAIL, заснованого на алгоритмі прямої оптимізації преференцій (DPO), було організовано та проведено серію експериментальних досліджень. Експериментальне оцінювання здійснювалося на двох принципово відмінних наборах даних; кожен виокремлює різні класи задач генерації тексту та дозволяє проаналізувати специфічні властивості та переваги розробленого методу.

У межах дослідження було використано такі стандартні датасети:

– CommonGEN (керована генерація тексту) [17]. У задачах такого типу модель повинна створювати зв'язні, логічні й синтаксично правильні речення на основі декількох заданих понять (наприклад, «dog», «give», «sit», «teach», «treat»). Вибір цього набору є доречним, оскільки він ретельно і повноцінно дозволяє перевірити здатність застосовуваної моделі до строгої реалізації лексичних умов та логіки речень, які вона генерує;

– ROCStories (умовно-некерована генерація тексту) [18]. Особливістю цього набору є те, що модель отримує на вхід тільки перше речення короткої історії. В її завдання входить створення продовження цього текстового фрагменту таким чином, щоб воно відповідало заданому контексту стилістично, тематично та семантично. Вибір цього датасету критично важливий для перевірки здатності моделі генерувати тексти, які виглядають логічно послідовними та демонструють природний авторський стиль, властивий людині.

Для отримання статистично надійних і достовірних результатів кожен з експериментальних запусків повторювався тричі. Всі повтори виконували із незмінними гіперпараметрами й умовами, але щоразу використовували різні початкові випадкові ініціалізації ваг моделі. У кожному окремому

повторенні використовувалися абсолютно однакові 5000 прикладів тренувальних даних, 500 – валідаційних та 100 – тестових.

Варто окремо підкреслити принципову різницю між керованою (CommonGEN) та умовно-некерованою (ROCStories) генерацією. Керований підхід (CommonGEN) активно застосовують у різних практичних сценаріях, таких як:

– створення текстів для пошукової оптимізації (SEO), де необхідне точне дотримання заданих ключових слів;

– написання текстів із суворими спеціалізованими лексичними вимогами;

– генерація технічних, наукових та навчальних матеріалів, у яких важлива коректність і точність термінології.

Натомість некерований сценарій генерації (ROCStories) краще відповідає широкому класу реальних завдань, що висувають менші вимоги до наявності визначених ключових слів, проте суворо контролюють стиль та логічну послідовність генерованих текстів. Це, зокрема, такі сценарії, як:

– продовження художніх оповідань і розповідей в авторському стилі;

– створення природної стилістики діалогів із віртуальними агентами (чат-боти);

– генерація художніх текстів про ситуації або сценарії, де стиль письма займає пріоритет над конкретною лексикою.

Експериментальна процедура умовно була розбита на три основні етапи. Перший етап мав на меті базове донавчання попередньо натренованої моделі GPT-2 на вибраному текстовому корпусі (CommonGEN або ROCStories). Використовували стандартний підхід авторегресивного навчання, в основі якого лежить оптимізація з перехресною ентропією.

Другий етап охоплював підготовку дискримінатора на основі моделі RoBERTa. На цьому етапі дискримінативна модель навчалася диференціювати авторські тексти від згенерованих GPT-2 штучних варіантів. Після завершення етапу навчена RoBERTa модель

використовувалася для формування бази парних преференцій. Для кожного промпту створювалися п'ять варіантів генерації й обиралися відповідно найкращий (chosen) та найгірший (rejected) варіанти. Отакі сформовані пари і стали основою для наступного етапу застосування алгоритму DPO.

Третій етап був присвячений саме прямій оптимізації преференцій DPO: тут GPT-2 донавчалася вже на основі сформованого набору парних преференцій. Донавчання тривало три епохи, з такими параметрами: розміром пакета batch size = 4, темпом навчання 5e-6 і регуляризатором $\beta = 0.2$. Ці параметри були визначені на основі попередніх експериментальних тестів як найбільш оптимальні.

Оцінка якості генерації текстів здійснювалась за допомогою набору різноманітних кількісних метрик:

– BLEU – оцінює наближеність згенерованих текстів до авторських еталонів;

– Distinct-n та Self-BLEU – використовуються для аналізу різноманітності й варіативності текстів (високе Distinct-n і низьке Self-BLEU свідчать про хороше різноманіття текстів);

– перплексія (PPL) – оцінює впевненість моделі у власній генерації;

– Seq-Rep-n – аналізує повторюваність фрагментів тексту й дозволяє виявляти проблеми зацикленості генерації;

– оцінка RoBERTa – головна метрика, яка показує близькість тексту до людських авторських варіантів.

Для аналізу дискримінатора тексти попередньо токенізували та подавали на вхід RoBERTa. Отримувалися логіти, що конвертувалися в ймовірності від 0 до 1 – де значення, близькі до одиниці, означали текст, максимально схожий на реальний авторський.

Щоб кількісно оцінити якість, розраховували середнє значення оцінки дискримінатора для моделей TextGAIL та базової GPT-2 за всіма прикладами. Крім того, визначалась також частка випадків, коли TextGAIL отримував вищі оцінки порівняно з GPT-2.

Додатково було проведено деталізований аналіз розподілу відповідей дискримінатора на різних рівнях упевненості. Це дозволило оцінити не лише середній рівень генерації, але й зрозуміти силу дискримінатора у розмежуванні текстів, згенерованих авторські точно й автентично, та тих, що мають ознаки штучності.

Гіперпараметри – кількість епох, темп навчання, β та розмір пакета – були підібрані на підставі попередніх експериментальних тестів.

Для зручності і прозорості експериментальні результати структуровано зберігалися у спеціально каталогізованих папках. Додатково, кожний тестовий запуск виконували тричі з різними випадковими ініціалізаціями. Це дозволило перевірити загальну стабільність та якість роботи розробленого підходу TextGAIL у різних умовах.

3.2 Результати експериментів

З метою забезпечення статистичної надійності та точності оцінок експерименти для керованої генерації на наборі даних CommonGEN було виконано тричі. Результати цих запусків наведено у таблиці 3.1 нижче.

Різниця між показниками одержана при порівнянні моделі TextGAIL, натренованої за допомогою DPO, та стандартної базової моделі після донавчання. Позитивні зміни метрик BLEU, distinct-n та seq-rep-n свідчать про покращення відповідності до еталонних текстів, зростання лексичної різноманітності та зменшення повторів відповідно. Зниження метрики perplexity загалом є позитивним з погляду узгодженості тексту, хоча надмірне зниження може свідчити про потенційний дефіцит різноманітності. Проте, деяке зростання self-BLEU свідчить, що збільшення відповідності еталону може частково знижувати загальну внутрішню варіативність генерованих текстів.

Таблиця 3.1 – Порівняльні метрики для моделей на наборі даних CommonGEN

Метрика	Запуск 1	Запуск 2	Запуск 3	Середнє
BLEU (зміна)	+0.0057	+0.0075	+0.0089	+0.0074
distinct-1 (зміна)	+0.0283	-0.0032	+0.0807	+0.0353
distinct-2 (зміна)	+0.0024	+0.0015	+0.0231	+0.0090
seq-rep-3 (зміна)	+0.0300	+0.0100	+0.1600	+0.0667
perplexity (зміна)	-255.9684	+17.6667	-262.9871	-167.1296
self-bleu (зміна)	+0.0227	+0.0091	-0.0003	+0.0105
DPO перемоги	100%	58%	100%	86%
Середня оцінка (базова)	0.0176	0.5690	0.0223	0.2030
Середня оцінка (DPO)	0.6547	0.7141	0.7010	0.6899
Зміна оцінки	+0.6371	+0.1451	+0.6786	+0.4869

Під час аналізу результатів усіх трьох запусків спостерігається стабільна перевага моделі TextGAIL порівняно з базовою GPT-2, хоча масштаб цього поліпшення змінюється залежно від ініціалізації та якості першого етапу навчання.

У таблиці 3.2 наведено результати експериментів на наборі даних ROCStories, який є прикладом задачі умовно-некерованої генерації тексту.

Таблиця 3.2 – Порівняльні метрики для моделей на наборі даних ROCStories (три запуски).

Метрика	Запуск 1	Запуск 2	Запуск 3	Середнє
BLEU (зміна)	-0.0241	-0.0103	-0.0241	-0.0195
distinct-1 (зміна)	+0.1875	+0.0144	+0.1875	+0.1298
distinct-2 (зміна)	+0.0606	+0.0052	+0.0606	+0.0421
seq-rep-3 (зміна)	+0.3100	+0.0500	+0.3100	+0.2233
perplexity (зміна)	-14.2084	-13.1248	-14.2084	-13.8472
self-bleu (зміна)	-0.0479	-0.0071	-0.0397	-0.0316
DPO перемоги	92%	48%	92%	77.33%
Середня оцінка (базова)	0.3720	0.6842	0.3720	0.4761
Середня оцінка (DPO)	0.6830	0.7159	0.6830	0.6940
Зміна оцінки	+0.3110	+0.0316	+0.3111	+0.2179

Результати на ROCStories також демонструють стабільне поліпшення лексичної різноманітності та зменшення механічних повторень, що вказує на покращення творчої складової генерації. Невелике зниження показників BLEU у даному випадку є швидше ефектом більшої свободи творчості та меншої формальності при написанні продовження історій, ніж серйозною проблемою якості.

Для детальнішого порівняння якості генерації текстів моделлю GPT-2 та запропонованою нами моделлю TextGAIL DPO у таблиці 3.3 наведено додаткові приклади з експериментальних даних. Приклади було відібрано

таким чином, щоб максимально наочно відобразити відмінності у якості отриманих текстів. Були враховані такі аспекти, як наявність логічних чи семантичних помилок, здатність моделей дотримуватися початкового контексту або обов'язкове використання заданих ключових концептів.

Таблиця 3.3 – Найбільш показові приклади результатів генерації текстів моделями GPT-2 і TextGAIL DPO

Вхідні дані	Результати GPT-2	Результати TextGAIL DPO
Генерація за ключовими словами (CommonGEN)		
dog, give, sit, teach, treat	dog gives students a lesson on how to treat a dog (переплутані ролі персонажів)	dog teaches you how to give a treat to a dog (логічно і точно)
microphone, sing, stage	singer playing a piano on stage (пропущено слово microphone)	musician performs in a room with microphone on stage (точно за умовою)
beach, dig, hole, sand	deep in the sand on a beach (не всі ключові слова, не завершене речення)	dig a hole in the beach in a sandstorm (адекватно і повно)

Продовження таблиці 3.3

Вхідні дані	Результати GPT-2	Результати TextGAIL DPO
Продовження історій (ROCStories)		
Abel refused to swim in PE. No one knew it was because he couldn't swim.	He was suspended. Abel took a shot at the big time. He ran a marathon (нелогічно, втрата контексту)	He quit school (логічне завершення, відповідає контексту)
Zach was sitting at his computer and watching TV. He kept seeing Taco Bell commercials all night. He ended up driving to Taco Bell on his bike he was so hungry! He ate 15 dollars worth of food and was too full to ride back home.	He realized he was drunk and decided to go check out his new place (нелогічна зміна змісту історії)	He's disappointed he did not pay for his ticket (краще відповідає введенню, логічне завершення)
My Boss came into my office. He said they would be letting our department go. Within the week we lost our job. They didn't even give us an explanation.	My boss gave us the pay and we were offered a job (контекстне суперечення)	I was fired (коректне завершення)

Аналіз наведених прикладів показує, що генерування текстів моделлю TextGAIL DPO частіше забезпечує більш адекватні та логічні результати, на відміну від GPT-2, яка схильна порушувати контекстні умови та допускає виражені логічні помилки. Водночас навіть більш вдала модель TextGAIL DPO не позбавлена окремих семантичних артефактів та потребує подальших досліджень і коригування.

Таким чином, надані результати підтверджують, що модель TextGAIL DPO демонструє вищу стабільність у генерації логічно завершених і контекстно коректних текстів, однак перспективним залишається подальший розвиток цього підходу з метою уникнення всіх потенційних помилок.

3.3 Експертне оцінювання стилістичної відповідності

Додатково до автоматизованих метрик, представлених у попередніх розділах, було здійснено спеціалізовану процедуру експертної оцінки стилістичної відповідності згенерованих текстів. Це дозволило глибше проаналізувати тонкі стилістичні та лінгвістичні особливості отриманих текстових даних і визначити міру узгодженості автоматичних метрик з людським сприйняттям стилю текстів.

Для організації експертного оцінювання було відібрано 20 пар текстів, створених із датасетів ROCStories та CommonGen. Кожна пара містила два варіанти тексту: перший генерувався базовою моделлю GPT-2, другий – оптимізованою моделлю TextGAIL. Оцінювачами виступили 5 незалежних експертів. Вони знайомились із текстами в умовах сліпого тестування. Їм не повідомлялося, до якого методу належить той чи інший текстовий варіант.

З метою забезпечення максимальної об'єктивності тексти демонструвалися експертам у випадковому порядку. Оцінювачам була надана інформація лише про початкову умову генерації – набір ключових слів (для CommonGen) чи початкове речення історії (для ROCStories), а

також два текстові варіанти, позначені абстрактно («Текст А», «Текст В»), без зазначення моделі, яка їх створила. Учасники повинні були обрати той варіант генерації, який, на їхню думку, був найближчим до очікуваного стилю згідно з наданими початковими умовами. Додатково кожен експерт оцінював свою суб’єктивну впевненість у виборі за 5-бальною шкалою.

Отримані результати показали, що у 72% випадків експерти переважно надавали перевагу варіантам, згенерованим моделлю TextGAIL. Середній рівень суб'єктивної впевненості експертів у своїх рішеннях склав близько 4 балів, що свідчить про досить високу одностайність і визначеність у виборі. Найбільшу перевагу модель TextGAIL мала у генерації продовжень текстових історій на ROCStories: там її переважали у 84% випадків. Для керованої генерації на базі CommonGen модель теж демонструвала значні переваги, але трохи нижчі – близько 63%.

Процедура оцінювання стилю текстів проводилась за чітко визначеними та узгодженими критеріями:

– природність і легкість мови, відсутність механічних чи штучних формулювань;

– стилістична гармонійність (узгодженість стилю, тону, емоційного забарвлення);

– різноманіття лексики (лексична варіативність);

– логічність і семантична узгодженість.

Більш докладний аналіз якісних коментарів експертів підтвердив, що автоматизовані метрики текстової якості не завжди здатні повністю врахувати всі стилістичні й естетичні тонкощі текстової генерації. Зокрема, було знайдено високу позитивну кореляцію суб’єктивних оцінок стилістичної якості з високими показниками лексичної варіативності (Distinct-n) і низьким рівнем повторюваності фрагментів (Seq-rep-n). Водночас для метрики BLEU кореляція з експертними оцінками виявилася значно нижчою, що свідчить про її нижчу інформативність у завданнях стилістичної оцінки.

Підсумкові статистичні результати експертного вибору наведено у таблиці 3.4 нижче.

Таблиця 3.4 – Результати експертного вибору між GPT-2 та TextGAIL

Модель	Відсоток виборів (%)
GPT-2 (базова)	18%
TextGAIL	72%
Важко визначитися	10%

За результатами експертної оцінки було також сформульовано конкретні рекомендації для подальшого розвитку й удосконалення моделі TextGAIL. Основними важливими напрямами розвитку визнано такі:

– збільшення обсягу й різноманітності текстових корпусів для тренування, що забезпечить стабільнішу роботу моделі в різних тематичних сценаріях;

– впровадження тонших алгоритмічних механізмів для балансування між стилістичною точністю та змістовною адекватністю генерації;

– створення зручних інструментів контрольованої генерації, які дають користувачам можливість більш точно налаштовувати стилістичні параметри (формальність, емоційна насиченість, складність мовних конструкцій);

– розробка індивідуалізованих стилістичних профілів, з орієнтацією на персональні авторські особливості;

– активніше залучення людських експертних оцінок безпосередньо в цикл навчання нейромоделей у парадигмі «людина у циклі».

Таким чином, проведене експертне оцінювання дозволило не тільки додатковим чином підтвердити ефективність і раціональність архітектури TextGAIL з точки зору людського сприйняття, але й сформулювати

прикладні рекомендації щодо того, як надалі покращувати генерацію до рівня максимально наближеного до природного людського стилю письма.

3.4 Аналіз та обговорення результатів

Проведені три незалежні запуски експериментів для кожного з двох обраних датасетів дозволили зробити ґрунтовні висновки щодо ефективності запропонованого методу TextGAIL на основі прямої оптимізації преференцій (DPO).

Аналіз результатів задач керованої генерації (набір CommonGEN) показав кілька важливих закономірностей. У кожному з трьох запусків стабільно покращувалась метрика BLEU, в середньому на 0.0074. Це свідчить про те, що тексти, створювані за допомогою DPO-оптимізації, стають ближчими до людських текстів-еталонів. Крім того, середня зміна оцінки дискримінатора RoBERTa зросла на +0.4869. Цей результат додатково підтверджує суттєве підвищення якості текстів з позиції нейромережевої оцінки стилю.

Водночас було помічено цікавий ефект, якість DPO моделі сильно залежить від початкової якості базової GPT-2. У другому запуску якість початкової моделі була помітно вищою (0.5690), порівняно з першим (0.0176) та особливо третім (0.0223). Як наслідок, приріст ефективності у другому запуску був менш виражений. Ця ситуація підкреслює високу корисність розробленого методу саме у випадках низької або середньої початкової якості базової моделі.

Результати експериментів із задачі некерованої генерації (набір ROCStories) продемонстрували дещо інший характер. У всіх трьох запусках метрика BLEU, навпаки, зменшувалася в середньому на 0.0195. Але для творчих завдань це якраз є позитивним сигналом. Такі зміни означають більшу оригінальність та творчість генерованих текстів. Паралельно суттєво зросла лексична різноманітність на +0.1298, а повторюваність сегментів

тексту знизилася на 0.2233. Загалом модель DPO демонструвала виражену перевагу над базовою GPT-2 – 77.33% випадків. Втім, варто зазначити, що у другому запуску ця перевага виявилась меншою, лише 48%, що знову підкреслює вплив початкових умов ініціалізації на результати.

Під час якісного порівняння текстів від моделей легко простежується відмінність між базовою GPT-2 і модифікованою моделлю TextGAIL. Зокрема, модель GPT-2 під час вирішення задачі CommonGEN часто генерувала речення, які, попри формально правильну граматику, мали семантичні помилки, не завжди повноцінно використовували надані концепти або містили недоречні деталі. Порівняно з базовою GPT-2 тексти, згенеровані за методом DPO, були більш природними, логічними й точно відповідали висхідним умовам задачі. Проте, у деяких випадках все ж спостерігалися окремі нетипові деталі або зайві елементи, пов’язані з ймовірнісною природою генерації та обмеженим доступом до контекстуальної інформації. Ці явища пов’язані з ймовірнісною природою генерації та браком додаткового контексту.

При генерації умовного завершення історій із ROCStories базова GPT-2 здебільшого видавала довші, але частіше нелогічні, дещо заплутані тексти. Натомість модель DPO генерувала помітно коротші тексти, те ж водночас значно точніші стилістично, логічно цільніші, більш зв'язані з вихідною історією та емоційно природніші. Ці лаконічні приклади підкреслюють стилістичну перевагу моделі DPO.

Додатково звертає увагу стабільність зростання показників різноманітності (distinct-n та seq-rep-n). Це підтверджує наукове припущення, що метод оптимізації преференцій якісно покращує стилістичну гнучкість генерації. Водночас поведінка метрики BLEU суттєво залежала від контексту задачі. Для CommonGEN вона зростала, а для ROCStories показувала зниження. Це збігається з очікуваною відмінністю у характері завдань: керовані задачі орієнтовані на точність і концептуальну відповідність, некеровані – на більшу варіативність.

На окрему увагу заслуговує стабільність експериментальних результатів. Особливо помітним прикладом є близькість результатів першого й третього запусків у ROCStories. Цей факт свідчить про методологічну і технологічну міцність та надійність запропонованого методу TextGAIL за умови однакових налаштувань.

Хоча пряме порівняння з іншими актуальними підходами безпосередньо не проводилося через високу складність їхньої реалізації, можна оцінити певні потенційні переваги даного методу. Серед таких переваг можна виокремити простішу архітектуру, меншу обчислювальну складність, більшу стабільність результатів й ширший спектр можливих застосувань.

Таким чином, комплекс отриманих експериментальних результатів підтверджує перспективність, ефективність та значну прикладну цінність розробленого методу TextGAIL для завдань керованої й некерованої генерації природномовного тексту з визначеними стилістичними властивостями.

3.5 Етичні та правові аспекти генерації тексту в стилі автора

Сучасні технології автоматичного створення текстового контенту, який переконливо імітує стиль конкретних авторів, зумовлюють низку принципових викликів, які мають юридичний та етичний характер.

З юридичної точки зору авторське право у більшості країн світу захищає конкретні оригінальні твори: статті, книги чи інші тексти. Водночас сам авторський стиль – особливості письма, вираженість індивідуального ідіолекту, мовні прийоми чи риторичні фігури – наразі не є об’єктами, що юридично захищаються. Таким чином, генерування текстів у авторській стилістиці без буквального копіювання оригінальних фрагментів формально не вважається порушенням авторських прав. Однак сьогодні наразі відсутній міжнародний консенсус чи чітке законодавче регулювання

питання про допустимість використання захищених авторським правом творів під час навчання нейромережевих моделей.

Поточна позиція юридичних інституцій, зокрема Офісу з авторських прав США (рішення 2023 року), полягає в тому, що твори, створені штучним інтелектом без людського творчого внеску, не мають захисту авторських прав, адже відсутній необхідний компонент людського авторства [19]. Така ситуація створює певну юридичну невизначеність: з одного боку, текст, створений моделлю штучного інтелекту, не є формально захищеним авторським правом; з іншого – автор, чий стиль наслідується алгоритмом, втрачає будь-який юридично оформлений механізм захисту чи компенсації, якщо його стилістику активно використовують у комерційних цілях без дозволу.

Ще однією юридичною проблемою стає випадкове відтворення моделлю оригінальних фрагментів текстів, використаних для тренування. Великі мовні моделі можуть дослівно відтворювати певні сегменти навчальних корпусів за специфічних умов. Зокрема, у цій роботі модель TextGAIL має механізм запобігання таким ситуаціям: дискримінатор заохочує загальну стилістичну подібність, а не точне дослівне копіювання. Попри це, у практичному застосуванні завжди залишається необхідність додаткової перевірки згенерованого контенту на плагіат та додаткових процедур фільтрації потенційно ідентичних фрагментів.

Важливою етичною проблемою також є власне імітація стилю існуючих письменників, блогерів чи журналістів. Такі стилістично точні тексти можуть створити ефект автентичних авторських матеріалів й мотивувати поширення дезінформації. Складність визначення того, створений текст алгоритмічно чи людиною, лише підсилює ризики інформаційних фальсифікацій.

Для протидії поширенню дезінформації сьогодні активно розробляються спеціалізовані системи виявлення нейромережевого контенту, такі як Grover, GPTZero або DetectGPT [20], [21], [22]. Ці системи

розпізнають штучно створені тексти за їх статистичними особливостями. У попередньому варіанті методу TextGAIL пропонувалась спеціальна метрика – так званий показник сприйняття. У цьому дослідженні така метрика не розраховувалась безпосередньо, однак аналогічну функцію ефективно виконували дискримінатор RoBERTa та визначений показник «відсоток перемог» моделі DPO. Високі отримані значення підтверджують ефективність методу у напрямку стилістичної імітації. Однак це породжує нову суперечність між необхідністю створення стилістично правдоподібних текстів та суспільною необхідністю розрізняти машинні тексти й авторський контент.

Важливим етичним аспектом процесу генерації є також неминуче перенесення мовних упереджень, присутніх у вихідних навчальних даних. Ця проблема набуває особливої ваги з огляду на ймовірну здатність моделей навіть посилювати ненавмисні упередження через статистичні особливості генерації. Для вирішення цієї етичної дилеми важливим є вдосконалення процедур очищення даних та застосування спеціальних алгоритмів і критеріїв фільтрації текстів.

Окрема дискусія точиться навколо питання транспарентності походження автоматично створених текстів. Чітке маркування чи попередження читачів щодо створення текстів алгоритмом наразі важливе і в етичному, і в юридичному аспекті. Одним із перспективних рішень є використання цифрових водяних знаків, вбудованих безпосередньо у процес генерації на рівні розподілів ймовірностей [23]. Такі маркери мають бути непомітними для людей, але легко ідентифікуватись статистичними методами.

Ще однією проблемою є потенційні соціально-економічні наслідки генеративних технологій. Йдеться про девальвацію авторської праці та вплив на творчі професії. Зараз дискусія з цього питання є полярною: частина фахівців вважає, що такі технології підвищать продуктивність творчих професій, а інші навпаки, мають занепокоєння щодо загального

падіння якості творчого контенту через поширення поверхових стилістичних імітацій.

Таким чином, для ефективного й етично відповідального впровадження автоматичних систем стилістичного наслідування, таких як TextGAIL, необхідним є комплексний підхід до регулювання й етичної оцінки. Це передбачає контроль над авторським правом, фільтрацію даних з метою уникнення етичних упереджень та забезпечення максимальної прозорості щодо походження штучно створених текстів. Особливо прийнятним видається сценарій, коли автор свідомо використовує такі технології у власній творчості або за особистою згодою.

Не менш важливим етичним питанням є конфіденційність і захист персональних даних під час навчання моделей. Оскільки великі мовні моделі часто навчаються на текстах із відкритих джерел, потенційно можлива ситуація випадкового оприлюднення персональних чи чутливих даних, що потребує ретельної фільтрації та очищення навчальних корпусів.

Нарешті, актуальними залишаються питання відповідальності за наслідки неправомірного використання штучно створених текстів, зокрема поширення неправдивої чи небезпечної інформації. Нині відкритим питанням є те, хто саме повинен нести юридичну відповідальність у подібних ситуаціях.

Отже, безпечна й відповідальна інтеграція генеративних технологій та їх ефективний розвиток потребують не лише подальших технологічних рішень, але й глибшого осмислення правових, етичних та суспільних аспектів їх застосування. Важливу роль матиме спільна робота дослідників, правників, розробників алгоритмічних систем та творців контенту для вироблення єдиних правил і стандартів належного використання таких технологій. Зокрема, необхідно чітко регламентувати питання відповідальності за створений машиною контент, правила авторського використання стилістичних елементів та механізми прозорого маркування автоматично згенерованих текстів для читача.

Тільки за умов такого комплексного підходу буде можливим застосування розроблених у роботі технологій генерації, зокрема TextGAIL, в реальній практичній діяльності без ризику завдати шкоди суспільству чи окремим авторам. Подібний підхід дозволить максимально наблизитись до гармонійної інтеграції штучного інтелекту в людську діяльність, уникнувши водночас багатьох негативних аспектів, які він потенційно може спричинити, та отримавши максимальну користь від науки для широкого кола користувачів.

3.6 Перспективні напрями розвитку та можливості подальших удосконалень методу TextGAIL

Подальший розвиток системи TextGAIL може здійснюватися у двох основних тематичних площинах. Перша з них пов'язана із технічними й алгоритмічними вдосконаленнями підходу, друга – із практичною адаптацією та інтеграцією запропонованої методики для вирішення реальних задач.

Технічно-алгоритмічні доопрацювання включають оптимізацію використання обчислювальних та апаратних ресурсів. Важливою проблемою сучасних моделей генерації тексту залишається їх висока ресурсоємність у процесах налаштування та тренування. Вирішенням цієї задачі може бути запровадження ефективніших технологій адаптації з малою кількістю параметрів, таких як адаптери або LoRA [24]. Це дозволить суттєво економити витрати пам'яті та знизити загальні вимоги до обчислювальної інфраструктури. Перспективним також є вивчення альтернативних алгоритмів навчання з підкріпленням, які краще пристосовані для роботи у дискретних просторах інформаційних даних, ніж поточні підходи.

Інший перспективний технічний аспект – розробка моделей, зорієнтованих на генерацію великих за розміром текстів. Це вимагає

інтеграції в архітектуру додаткових механізмів довготривалої пам’яті або впровадження ієрархічних моделей нового покоління. Такі архітектури спочатку генерують верхньорівневу схему чи план текстового контенту, а потім поступово її деталізують. Це може суттєво зменшити частоту стилістичних чи логічних помилок, притаманних генерації довгих текстів.

Ще одним актуальним напрямом для розвитку є створення розширеної мультисигнальної функції винагороди для методу DPO. Такий підхід дозволяє паралельно враховувати кілька параметрів тексту: стилістичну подібність, емоційне забарвлення, тональність та формальність. Звичайно, створення такої системи винагороди потребуватиме продуманої структури балансування сигналів, отриманих від різних нейромережевих компонентів оцінювання.

Важливе алгоритмічне завдання – це вдосконалення дискримінативного компонента, який наразі базується на моделі RoBERTa. Для подальшого підвищення якості роботи дискримінатора перспективною може бути заміна поточної моделі більш сучасними нейромережевими архітектурами, наприклад DeBERTa V3. Також важливий потенціал має підхід із використанням ансамблю моделей-класифікаторів, що дозволить стабільніше й точніше оцінювати стилістичну відповідність і створювати ретельніші набори пар для преференційного навчання.

Поряд із суто технічними напрямами другою вагомою групою перспективних задач є подальша прикладна адаптація системи TextGAIL для реальних сценаріїв практичного застосування. Тут особливе значення має поглиблена інтеграція людських експертних оцінок безпосередньо у процес формування наборів даних для навчання моделей преференцій [25]. Такий підхід дозволить створити ефективнішу систему оцінки стилістичних конотацій та суб'єктивних критеріїв якості текстів.

Крім того, перспективною прикладною задачею є використання дискримінатора TextGAIL для автоматизованого фільтрування найкращих текстових варіантів із наборів, створених через потужні пропрієтарні

API-моделі (наприклад, GPT-4). Така інтеграція створить умови для додаткового підвищення якості автоматично згенерованого контенту.

Значний прикладний потенціал також може мати масштабування запропонованого методу за допомогою більш сучасних і потужних моделей: Llama, Mistral чи інших передових нейромережевих архітектур. Це дозволить суттєво підвищити абсолютну якість генерації текстового контенту та розширити спектр можливих завдань, які здатна розв'язати система.

Суттєвим практичним напрямом є інтеграція запропонованого методу із зовнішніми базами знань для зменшення фактологічних «галюцинацій». Ця інтеграція дозволить оперативніше й ефективніше перевіряти фактологічну коректність отриманого тексту з урахуванням зовнішніх авторитетних джерел інформації.

Перспективною задачею є адаптація моделі до різноманітних мультимодальних завдань, де генеровані тексти мають узгоджуватись із іншими інформаційними форматами.

Важливим напрямом для подальших досліджень є розроблення діалогових систем нового покоління. Це системи, які здатні не просто створювати тексти, а стилістично тонко адаптуватися до стилю мовлення співрозмовника чи певного заданого автора. Це значною мірою покращить користувацький досвід у взаємодії із чат-ботами й асистентами.

Перспективним напрямом є розробка багатокомпонентних стилістичних моделей оцінювання. Вони дозволять чітко аналізувати і керувати механізмами генерації за низкою ознак: лексичною, синтаксичною, прагматичною тощо. Така система стане виразно корисною у прикладних предметних контекстах.

Нарешті, важливою перспективою є адаптація розробленої методики TextGAIL до конкретних прикладних завдань із таких галузей, як журналістика, електронна комерція, освіта та право.

Важливим напрямом подальших досліджень є більш ретельне вивчення специфічних контекстів та потреб цих предметних галузей задля ефективної інтеграції запропонованого підходу в реальні бізнесові, освітні та суспільні практики. Реалізація зазначених напрямів розвитку створюватиме умови для максимального наближення результатів генерації текстів до природного людського стилю та суттєво розширить сферу їх прикладного застосування.

.

ВИСНОВКИ

У межах кваліфікаційної роботи було розроблено та експериментально перевірено модель TextGAIL – систему для генерації природномовних текстів із імітацією стилістичних особливостей заданого цільового автора. Запропонований у роботі підхід базується на поєднанні принципів імітаційного навчання та генеративно-змагальної парадигми з використанням методу прямої оптимізації преференцій (DPO).

Ключові етапи реалізації запропонованого підходу такі. На першому етапі виконувалося додаткове донавчання попередньо тренованої базової моделі GPT-2 на цільовому корпусі текстових даних. Після цього на другому етапі проводилося навчання спеціалізованого дискримінатора на основі моделі RoBERTa, завданням якого було ефективно розрізняти штучно згенеровані тексти й оригінальні авторські фрагменти. На підставі оцінок, отриманих дискримінатором, формувався набір пар текстів «обраний–відхилений», що став основою для фінального етапу налаштування генеративної моделі за допомогою алгоритму DPO.

Для перевірки розробленого у роботі методу було обрано дві принципово відмінні задачі генерації текстів, представлені наборами даних CommonGen (керована генерація за ключовими словами) та ROCStories (умовно-некерована генерація продовження художніх історій). Отримані в результаті експериментів дані переконливо підтвердили ефективність запропонованого методу TextGAIL, який продемонстрував суттєві й стабільні переваги порівняно із простим донавчанням вихідної GPT-2.

Зокрема значно покращились показники лексичної різноманітності, суттєво зменшилась повторюваність текстових фрагментів.

Більш розгорнутий якісний аналіз отриманих текстів також підтвердив перевагу методу DPO. Генеровані ним тексти характеризуються кращою семантичною й логічною узгодженістю, більш природною

стилістикою, а також помітною втратою типових шаблонів і механічності, притаманних варіантам GPT-2 без застосування додаткових процедур налаштування.

Особливо виразні переваги моделі TextGAIL спостерігаються у завданнях творчого типу, де вкрай важлива стилістично багата й природна мова без виражених повторів та логічних суперечностей. Це підтверджує, що сигнал преференцій, який отримує генератор від дискримінатора, здатен суттєво покращити гнучкість і адаптивність процесу генерації у складних стилістичних завданнях.

При цьому виявлені зміни перплексії та деякі коливання у результатах між окремими запусками експериментів вказують на складні нелінійні закономірності залежності якості фінальної моделі від її початкового стану. Запропонований метод демонструє найбільшу ефективність саме для моделей з первинно низькою чи середньою якістю генерації, оскільки саме у таких випадках приріст якості залишається найбільш помітним.

В окремому порядку було відзначено істотні переваги методу TextGAIL порівняно з класичними методами навчання з підкріпленням, зокрема PPO. Мається на увазі більш проста архітектурна реалізація, менша обчислювальна складність, зменшена кількість критично важливих гіперпараметрів та більша стабільність результатів у широкому спектрі задач – від такого типу завдань із суворими лексичними вимогами, до творчих й стилістично відкритіших завдань.

Однак, окремо ідентифіковано низку обмежень і актуальних проблем у використанні даного класу систем. До них належать високі вимоги до апаратних ресурсів, ризик втрати коректності й достовірності інформації, а також етичні чи правові питання, пов'язані з імітацією чийогось стилю. Особливими питаннями є коректність використання текстів, захищених авторським правом, при стилістичній імітації, а також ризики поширення неправдивого й маніпулятивного контенту, створеного з стилістичної імітації популярних авторів та ЗМІ.

Для вирішення зазначених проблем запропоновано перспективні напрями наступних досліджень:

– масштабування запропонованого методу TextGAIL за допомогою потужніших моделей, таких як нейромережі Llama чи Mistral, для зростання абсолютної якості генерації тексту;

– удосконалення дискримінативного компонента з переходом на сучасніші моделі класу DeBERTa V3, ансамблевими підходами чи розширенням бази навчальних текстів;

– інтеграція експертних людських оцінок безпосередньо у процес налаштування моделей у парадигмі навчання через людину у циклі;

– застосування дискримінатора, натренованого пропонованим методом, для відбору кращих варіантів текстів, отриманих сучасними моделями (наприклад, API GPT-4);

– адаптація TextGAIL для складніших типів завдань: від перефразування й стилізованих діалогів, до імітації композиційних чи наративних елементів текстів.

Отже, отримані наукові та прикладні результати підтверджують ефективність і перспективність запропонованого методу TextGAIL та відкривають шляхи для його подальшого адаптування, масштабування й інтеграції у різні прикладні сценарії стилістичної генерації текстового контенту.

ПЕРЕЛІК ДЖЕРЕЛ ПОСИЛАННЯ

1. Vaswani A., Shazeer N., Parmar N., et al. Attention Is All You Need. *Advances in Neural Information Processing Systems*. 2017. 30. URL: https://papers.nips.cc/paper_files/paper/2017/file/3f5ee243547dee91fbd053c1c4a845aa-Paper.pdf (дата звернення: 13.04.2025).

2. Wu Q., Li L., & Yu Z. TextGAIL: Generative Adversarial Imitation Learning for Text Generation. *Proceedings of the AAAI Conference on AI*. 2021. 35(11), 14067–14075. URL: https://arxiv.org/abs/2004.13796 (дата звернення: 13.04.2025).

3. Goodfellow I., Pouget-Abadie J., Mirza M., et al. Generative Adversarial Nets. *Advances in Neural Information Processing Systems*. 2014. 27. URL: https://arxiv.org/abs/1406.2661 (дата звернення: 13.04.2025).

4. Tikhonov A., Yamshchikov I. P. What is wrong with style transfer for texts? *arXiv preprint arXiv:1808.04365*. 2018. URL: https://arxiv.org/abs/1808.04365 (дата звернення: 13.04.2025).

5. Yu L., Zhang W., Wang J., Yu Y. SeqGAN: Sequence Generative Adversarial Nets with policy gradient. *Proceedings of AAAI'17*. 2017. URL: https://arxiv.org/abs/1609.05473 (дата звернення: 13.04.2025).

6. Che T., Li Y., Zhang R., Hjelm R. D., et al. Maximum-Likelihood Augmented Discrete Generative Adversarial Networks (MaliGAN). *arXiv preprint arXiv:1702.07983*. 2017. URL: https://arxiv.org/abs/1702.07983 (дата звернення: 13.04.2025).

7. Guo J., Lu S., Cai H., Zhang W., et al. Long Text Generation via Adversarial Training with Leaked Information. *Proceedings of AAAI'18*. 2018. URL: https://arxiv.org/abs/1709.08624 (дата звернення: 13.04.2025).

8. Nie, W., Narodytska, N., & Patel, A. RelGAN: Relational Generative Adversarial Networks for Text Generation. *Proceedings of the International Conference on Learning Representations*. 2019. URL: https://openreview.net/forum?id=rJedV3R5tm (дата звернення: 13.04.2025).

9. Radford A., Wu J., Child R., Luan D., et al. Language Models are Unsupervised Multitask Learners. *OpenAI*. 2019. URL: https://cdn.openai.com/better-language-models/language_models_are_unsupervised_multitask_learners.pdf (дата звернення: 13.04.2025).

10. Brown T.B., Mann B., Ryder N., et al. Language Models are Few-Shot Learners. *Advances in Neural Information Processing Systems*. 2020. 33, 1877–1901. URL: https://arxiv.org/abs/2005.14165 (дата звернення: 13.04.2025).

11. Devlin J., Chang M.-W., Lee K., Toutanova K. BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding. *NAACL*. 2019. URL: https://arxiv.org/abs/1810.04805 (дата звернення: 13.04.2025).

12. Liu Y., Ott M., Goyal N., Du J., et al. RoBERTa: A Robustly Optimized BERT Pretraining Approach. *arXiv preprint arXiv:1907.11692*. 2019. URL: https://arxiv.org/abs/1907.11692 (дата звернення: 13.04.2025).

13. Ho, J., Ermon, S. Generative Adversarial Imitation Learning. *Advances in Neural Information Processing Systems*, 29. 2016. URL: https://arxiv.org/abs/1606.03476 (дата звернення: 13.04.2025).

14. Rafailov R., Sharma A., Mitchell E., Ermon S., et al. Direct Preference Optimization: Your Language Model is Secretly a Reward Model. *arXiv preprint arXiv:2305.18290*. 2023. URL: https://arxiv.org/abs/2305.18290 (дата звернення: 13.04.2025).

15. Schulman J., Wolski F., Dhariwal P., Radford A., et al. Proximal Policy Optimization Algorithms. *arXiv preprint arXiv:1707.06347*. 2017. URL: https://arxiv.org/abs/1707.06347 (дата звернення: 13.04.2025).

16. Ouyang L., Wu J., Jiang X., Almeida D., et al. Training language models to follow instructions with human feedback (RLHF). *arXiv preprint arXiv:2203.02155*. 2022. URL: https://arxiv.org/abs/2203.02155 (дата звернення: 13.04.2025).

17. Lin B., Zhou W., Shen M., et al. CommonGen: A Constrained Text Generation Challenge. *EMNLP*. 2020. URL: https://aclanthology.org/2020.findings-emnlp.165 (дата звернення: 13.04.2025).

18. Mostafazadeh N., et al. A Corpus and Evaluation Framework for Commonsense Stories. *NAACL-HLT*. 2016. URL: https://aclanthology.org/N16-1098 (дата звернення: 13.04.2025).

19. Copyright Registration Guidance: Works Containing Material Generated by Artificial Intelligence. *U.S. Copyright Office*. 2023. URL: https://www.federalregister.gov/documents/2023/03/16/2023-05321/copyright-registration-guidance-works-containing-material-generated-by-artificial-intelligence (дата звернення: 13.04.2025).

20. Grover – Defending Against Neural Fake News. *rowanzellers.com*. URL: https://rowanzellers.com/grover/ (дата звернення: 13.04.2025).

21. GPTZero. *gptzero.me*. URL: https://gptzero.me (дата звернення: 13.04.2025).

22. DetectGPT: AI Detector You Can Trust. *detectgpt.com*. URL: https://detectgpt.com (дата звернення: 13.04.2025).

23. Kirchenbauer J., et al. A Watermark for Large Language Models. *arXiv preprint arXiv:2301.10226*. 2023. URL: https://arxiv.org/abs/2301.10226 (дата звернення: 13.04.2025).

24. Hu E., et al. LoRA: Low-Rank Adaptation of Large Language Models. *arXiv preprint arXiv:2106.09685*. 2021. URL: https://arxiv.org/abs/2106.09685 (дата звернення: 13.04.2025).

25. Christiano, P., et al. Deep reinforcement learning from human preferences. *NeurIPS*. 2017. URL: https://arxiv.org/abs/1706.03741 (дата звернення: 13.04.2025).