

МІНІСТЕРСТВО ОСВІТИ І НАУКИ УКРАЇНИ
КРИВОРІЗЬКИЙ ДЕРЖАВНИЙ ПЕДАГОГІЧНИЙ УНІВЕРСИТЕТ
Фізико-математичний факультет
Кафедра інформатики та прикладної математики

«Допущено до захисту»

Завідувач кафедри

_____ Моїсеєнко Н. В.

Реєстраційний № _____

«_____» _____ 2024 р.

«_____» _____ 2024 р.

**ОЦІНКА ЕФЕКТИВНОСТІ РІЗНИХ МОДЕЛЕЙ
НАВЧАННЯ ЧАТ-БОТІВ НА ДІАЛОГОВИХ
НАБОРАХ ДАНИХ**

Кваліфікаційна робота студента групи І-20
ступінь вищої освіти «бакалавр»
спеціальності 014.09 Середня освіта (Інформатика)
Ляшенка Романа Олеговича

Керівник: доктор педагогічних наук, професор,
старший дослідник
Семеріков Сергій Олексійович

Оцінка:

Національна шкала _____

Шкала ECTS _____ Кількість балів _____

Голова ФК _____

Члени ФК _____

ЗАПЕВНЕННЯ

Я, Ляшенко Роман Олегович, розумію і підтримую політику Криворізького державного педагогічного університету з академічної доброчесності. Запевняю, що ця кваліфікаційна робота виконана самостійно, не містить академічного плагіату, фабрикації, фальсифікації. Я не надавав і не одержував недозволену допомогу під час підготовки цієї роботи. Використання ідей, результатів і текстів інших авторів мають покликання на відповідне джерело.

Із чинним Положенням про запобігання та виявлення академічного плагіату в роботах здобувачів вищої освіти Криворізького державного педагогічного університету ознайомлений. Чітко усвідомлюю, що в разі виявлення у кваліфікаційній роботі порушення академічної доброчесності робота не допускається до захисту або оцінюється незадовільно.



ЗМІСТ

ВСТУП	3
1. БІБЛІОМЕТРИЧНИЙ АНАЛІЗ	6
1.1. Обґрунтування доцільності	6
1.2. Результати	7
Висновки до 1 розділу	16
2. МОДЕЛІ НАВЧАННЯ ЧАТ-БОТІВ	20
2.1. Навчання з учителем	20
2.1.1. Seq2Seq моделі на основі рекурентних нейронних мереж	21
2.1.2. Трансформерні архітектури	22
2.2. Навчання з підкріпленням	25
2.3. Трансферне навчання	27
2.4. Оцінка ефективності навчання чат-ботів	30
Висновки до 2 розділу	31
3. ЕКСПЕРИМЕНТИ	33
3.1. Створення наборів даних для навчання чат-ботів	33
3.2. Вибір моделей для донавчання	34
3.3. Процес донавчання моделі чат-боту на базі GPT2-uk	35
Висновки до 3 розділу	41
ВИСНОВКИ	42
СПИСОК ВИКОРИСТАНИХ ДЖЕРЕЛ	43
ДОДАТКИ	55
А. Результати застосування генеративних чатботів для ви- значення імен кластерів	55

ВСТУП

Актуальність теми. Розвиток штучного інтелекту та технологій опрацювання природної мови зумовлює стрімке зростання інтересу до створення програмних агентів, здатних вести природномовний діалог – так званих чат-ботів. Провідні ІТ-компанії, такі як Google [1–5; 11; 16; 17; 22; 30; 34; 37; 41; 46; 53; 55; 56; 58; 62; 68; 69; 71; 72; 74; 76; 82; 83; 85; 87; 93], Microsoft [4; 9; 28; 39; 40; 52; 59; 68; 72; 75; 89], Meta (Facebook) [7; 13; 50; 54; 70; 73; 80; 81], OpenAI [6; 14; 25; 38; 43; 49; 53; 56; 57; 60; 65; 66; 77; 79; 86] активно працюють над розробкою чат-ботів. Успішні проекти на кшталт ChatGPT від OpenAI [63] демонструють значний потенціал застосування таких систем у різноманітних сферах людської діяльності. У вказаних дослідження можна виділити дві основні групи задач: 1) навчання (викладання та учіння) студентів; 2) побудова чат-ботів – систем типу “питання-відповідь”, які навчають власні моделі за даними, отриманими від користувача. Водночас постає низка важливих наукових та практичних завдань щодо ефективного навчання таких систем, забезпечення їхньої надійності, безпеки та етичності.

Чат-боти як природномовні системи штучного інтелекту мають величезний потенціал для підвищення ефективності та якості різноманітних видів людської діяльності шляхом автоматизації рутинних процесів, надання інтелектуальної підтримки користувачам, забезпечення персоналізованого навчання тощо. Успішні розробки у цій галузі здатні у перспективі докорінно змінити характер взаємодії людини з комп’ютерними системами, підвищити продуктивність праці та навчання. Разом з тим, вирішення етичних та безпекових проблем є критично важливим для впровадження цих технологій в інтересах суспільства.

Об’єкт дослідження – навчання чат-ботів.

Предмет дослідження – моделі навчання чат-ботів на діалогових наборах даних.

Мета дослідження – оцінка ефективності різних моделей навчання чат-ботів на діалогових наборах даних.

Відповідно до мети визначено такі основні **завдання дослідження**:

- 1) провести бібліометричний аналіз джерел з проблеми навчання чат-

ботів для визначення ключових понять дослідження;

- 2) визначити моделі навчання чат-ботів та критерії оцінки їх ефективності;
- 3) створити два набори даних та виконати донавчання доцільно обраних великих мовних моделей на них.

Методи дослідження: *бібліометричний аналіз* результатів пошуку для а) виявлення хронологічної межі, починаючи з якої спостерігається стійке зростання кількості публікацій за темою дослідження; б) для створення карти зв'язків ключових слів; в) для групування ключових слів у кластери; г) для визначення ключових понять дослідження; *великі мовні моделі Claude 2, Google Bard, GPT 3.5, GPT 4.0* для генерації припущень щодо назв кластерів ключевих слів; *методи інженерії програмного забезпечення* для проєктування, кодування, тестування та випробування.

Новизна результатів дослідження:

- виявлено нижню хронологічну межу (2018 рік), починаючи з якої спостерігається стійке зростання кількості публікацій із навчанням чат-ботів;
- визначено 4 кластери досліджень із навчання чат-ботів: 1) опрацювання природної мови; 2) застосування технологій опрацювання природної мови у суспільстві; 3) застосування машинного навчання для опрацювання природної мови; 4) чат-боти у навчанні та сфері послуг.

Апробація кваліфікаційної роботи відбувалась шляхом:

- участі у наукових конференціях:
 1. 6th Workshop for Young Scientists in Computer Science & Software Engineering (February 2, 2024, Kryvyi Rih, Ukraine) [24];
 2. VII International Scientific and Practical Conference “Information Technology for Education, Science, and Technics” (May 23-24, 2024, Cherkasy, Ukraine) [94];
- публікацій за темою дослідження:

1. *Ляшенко Р. О., Семеріков С. О.* Бібліометричний аналіз навчання чат-ботів // Тези доповідей VII Міжнародної науково-практичної конференції «Інформаційні технології в освіті, науці і техніці» (ІТОНТ-2024), (Черкаси, 23-24 травня 2024 р.) Черкаси : ЧДТУ, 2024. С. 314–316. URL: <https://itest.chdtu.edu.ua/#materials>
2. *Ляшенко Р. О., Семеріков С. О.* Бібліометричний аналіз досліджень з навчання чат-ботів: ключові поняття та тенденції // Інформаційні технології і засоби навчання. 2024. Т. 101, № 3. (у процесі публікації)

Структура та обсяг роботи. Кваліфікаційна робота складається зі вступу, двох розділів, висновків, списку використаних джерел (95 найменувань), 1 додаток. Робота містить 4 таблиці та 12 рисунків. Загальний обсяг роботи – 60 сторінок.

РОЗДІЛ 1

БІБЛІОМЕТРИЧНИЙ АНАЛІЗ

1.1. Обґрунтування доцільності

У галузі створення та навчання чат-ботів активно працюють як науковці, так і розробники ІТ-компаній. Окремі аспекти цієї проблематики висвітлюються у низці публікацій. Зокрема, результати пошуку у блозі DeepLearning.AI [27] за 2019-2023 рр. надають можливість виокремити наступні практично розв'язані задачі:

- з 2020 року чатботи (Generative BST від Facebook та Google Meena) можна використовувати для *коротких діалогів на загальні теми* [10; 12];
- з 2023 року Microsoft надає можливість використання чатботів у Office 365 та Windows для *покращення продуктивності* [19];
- з 2020 року чатботи активно *застосовуються у бізнесі* для обслуговування клієнтів, продажів тощо [20];
- з 2023 року Google та Microsoft вводять чатботи у *пошукові системи* [78].

Статті у блозі DeepLearning.AI вказують також на прогалини знань та відкритів проблеми, що існують у даній галузі:

- медичні чатботи можуть ставити помилкові діагнози [18];
- такі чатботи, як BlenderBot 3 та Galactica, можуть виявляти упередженість, токсичність та перекручувати факти [48];
- зависока вартість розгортання великих мовних моделей [23];
- недостатня прозорість моделей чатботів може підіймати етичні питання [92].

Аналіз останніх досліджень і публікацій виявляє необхідність консолідації зусиль для визначення актуальних напрямів і пріоритетів досліджень з метою ефективного навчання чат-ботів.

1.2. Результати

Для бібліометричного аналізу було обрано наукометричну базу даних Scopus видавництва Elsevier як одну з найбільших реферативних баз даних, що індексує більше 42 тис. періодичних видань та містить переважно якісні наукові джерела з технічних, природничих, медичних та суспільних наук, що робить її репрезентативною для аналізу публікацій з проблематики навчання чат-ботів на стику кількох наукових галузей. Збалансоване покриття різних наукових дисциплін у Scopus забезпечує необхідну повноту для виявлення ключових тенденцій та напрямків досліджень в обраній галузі.

Для визначення, в яких роботах зустрічаються ключові поняття дослідження “навчання” та “чатботи”, 18 листопада 2023 року було виконано пошук у базі Scopus [32]. Відбирались ті документи, у назвах, анотаціях та ключових словах яких одночасно зустрічались слова “chatbot” та “training” (рис. 1.1).

The screenshot shows the Scopus search interface. At the top, there is a search bar with the query "chatbot AND training" and a search button. Below the search bar, there are options to "Save search" and "Set search alert". The search results are displayed in a table with columns for Document title, Authors, Source, Year, and Citations. Two results are visible:

Document title	Authors	Source	Year	Citations
1 Exploring the design and utility of an integrated web-based chatbot for young adults to support healthy eating: a qualitative study	Ashton, L.M., Adam, M.T., Whatnall, M., ...Hansen, V., Collins, C.E.	International Journal of Behavioral Nutrition and Physical Activity, 20(1), 119	2023	0
2 Investigating ChatGPT and cybersecurity: A perspective on topic modeling and sentiment analysis	Okey, O.D., Udo, E.U., Rosa, R.L., Rodriguez, D.Z., Kleinschmidt, J.H.	Computers and Security, 135, 103476	2023	0

Рис. 1.1. Результати пошуку у базі Scopus за словами “chatbot”, “training”.

Загальна кількість документів – 549; розподіл документів за роками подано на рис. 1.2.

Перегляд ключових слів показує, що найбільш часто серед них зустрічаються такі: Chatbot (244 рази), Chatbots (221), Artificial Intelligence (126), Natural Language Processing Systems (90), Natural Language Processing

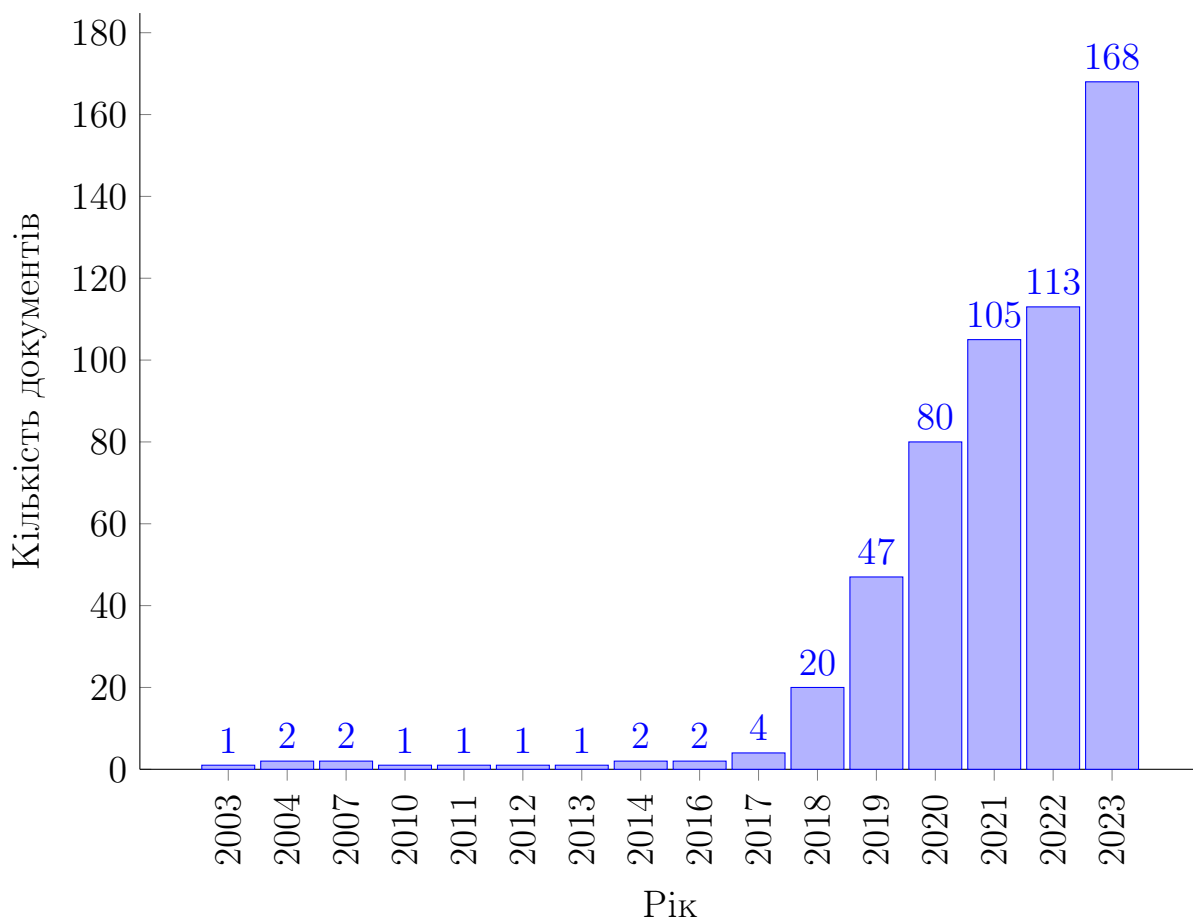


Рис. 1.2. Розподіл результатів пошуку за роками.

(90). Впорядкування ключових слів за алфавітом надає можливість визначити, які з них відрізняються лише за числом (у однині та у множині), та привести всі до однини:

- Chatbots є множиною від Chatbot;
- Conversational Agents є множиною від Conversational Agent;
- Convolutional Neural Networks є множиною від Convolutional Neural Network;
- Customer Services є множиною від Customer Service (Customer-service);
- Dialogue Generations є множиною від Dialogue Generation;
- Humans є множиною від Human;
- Large Language Models є множиною від Large Language Model;
- LSTM є аббревіатурою Long Short-term Memory;

- Machine-learning є синонімом Machine Learning;
- Mobile Applications є множиною від Mobile Application;
- NLP є аббревіатурою від Natural Language Processing (Natural Language Processing);
- NLU є аббревіатурою від Natural Language Understanding;
- Reinforcement Learnings є множиною від Reinforcement Learning;
- Virtual Assistants є множиною від Virtual Assistant.

Знайдені документи були експортовані у форматах CSV [91, с. 30] та BibTeX. У файлі CSV були виконані заміни ключових слів у множині та аббревіатур на ключові слова в однині.

Для виконання бібліометричного аналізу був використаний VOSviewer 1.6.20 [15]: *Create...* → *Create a map based on bibliographic data* → *Read*

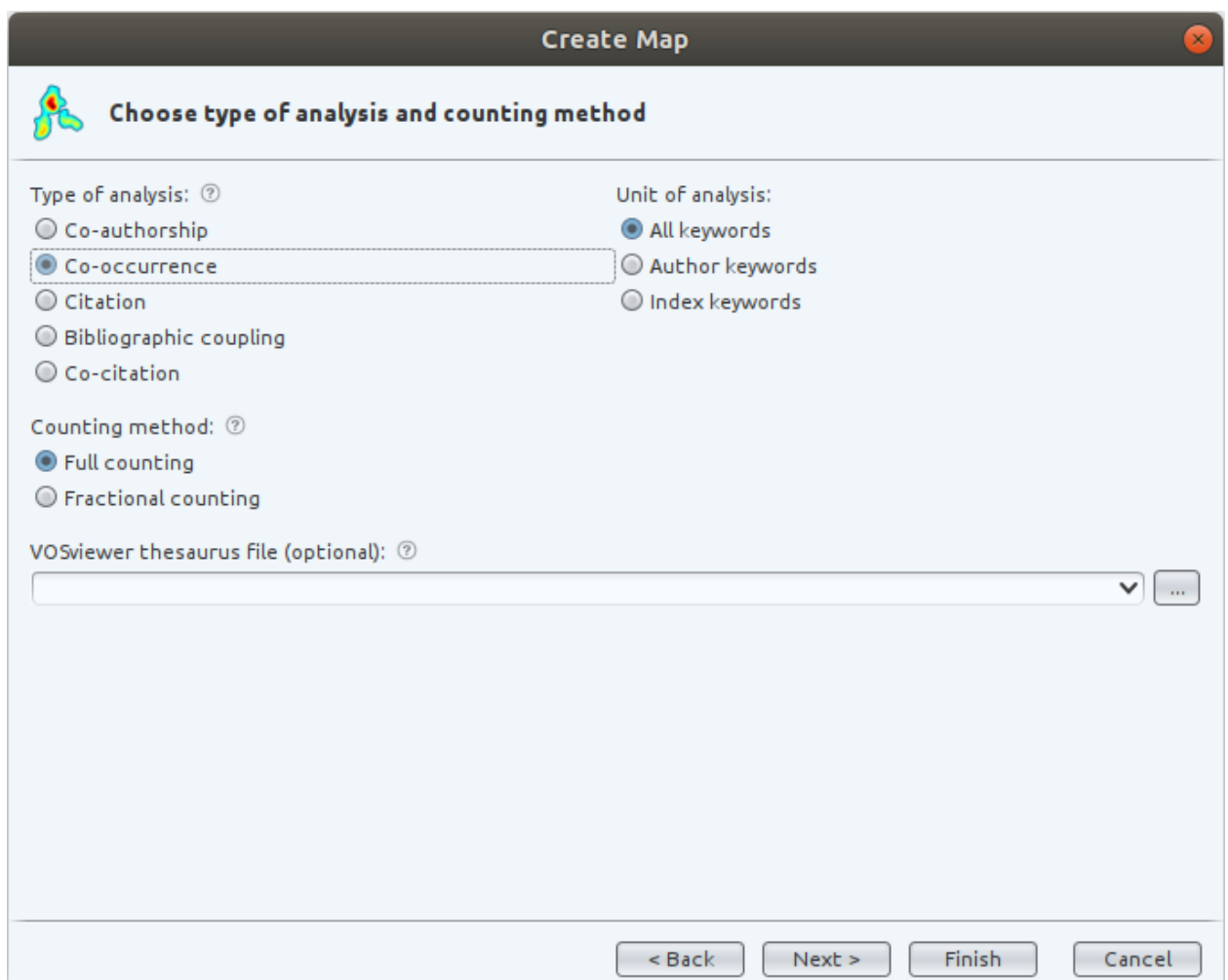


Рис. 1.3. Створення карти.

data from bibliographic database file → *Scopus* (рис. 1.3).

Карта створюється із такими параметрами: тип аналізу – за одночасною появою терміну (co-occurrence); одиниця аналізу – всі ключові слова; метод підрахунку – повний підрахунок (full counting).

Загальна кількість ключових слів – 3705 – може бути зменшена до 58 відкиданням тих з них, що зустрічаються менше ніж 12 разів. Із відібраних слів додатково вилучено слів “article” як таке, що не є специфічним для пошукового запиту.

Побудовану карту (табл. 1.1) подано на рис. 1.4.

Таблиця 1.1

Розподіл ключових слів за кластерами.

Ключове слово	Кластер	W_{Links}	$W_{Total\ link\ strength}$	$W_{Occurrences}$	$S_{Avg.\ pub.\ year}$	$S_{Avg.\ citations}$	$S_{Avg.\ norm.\ citations}$	Значущість
classification (of information)	1	37	122	24	2021.2083	3.7917	0.5908	3
training data	1	36	120	27	2021.1111	12.8519	1.8358	3
speech processing	1	36	102	26	2020.4615	9.4231	1.1619	1
performance	1	36	88	24	2021.5833	6.2917	1.02	2
semantics	1	34	90	18	2020.7222	8.8333	0.9379	
dialogue systems	1	33	96	23	2020.4348	8.4348	1.0096	
natural language understanding	1	32	123	32	2020.9375	6.0938	0.8055	3
computational linguistics	1	32	114	33	2020.1515	16.1515	2.0909	6
text processing	1	31	97	16	2021.25	4.4375	0.6438	
long short-term memory	1	30	113	24	2020.7917	4.8333	0.6383	
state of the art	1	29	54	14	2021.0714	4.0714	0.5852	
language model	1	28	61	14	2021.9286	3.5	0.6501	2
question answering	1	28	60	12	2021.1667	1.4167	0.2476	
information retrieval	1	27	51	12	2020.4167	9.3333	0.9633	
embeddings	1	24	53	13	2020.8462	7.5385	0.6933	
reinforcement learning	1	23	52	15	2020.0667	9.0667	1.1396	

Продовження на наступній сторінці

Табл. 1.1 – продовження з попередньої сторінки

Ключове слово	Кластер	W_{Links}	$W_{Total\ link\ strength}$	$W_{Occurrences}$	$S_{Avg.\ pub.\ year}$	$S_{Avg.\ citations}$	$S_{Avg.\ norm.\ citations}$	Значущість
response generation	1	17	37	12	2020.75	11.5833	1.343	
artificial intelligence	2	52	452	126	2021.3333	9.8571	1.8428	6
conversational agent	2	47	203	50	2020.18	18.68	1.9112	1
human	2	37	266	60	2022.1167	15.7667	2.7326	2
adult	2	29	117	17	2021.9412	17	2.3017	
training	2	29	84	22	2021.7727	16.7273	3.5069	2
male	2	27	94	13	2021.4615	23.3846	3.0896	1
medical education	2	27	66	12	2022	3.5833	0.557	
female	2	26	101	16	2021.375	22.6875	2.6942	
controlled study	2	26	76	15	2022.2	9.2	1.6989	1
health care	2	26	67	15	2021.4	3.6	0.51	
mental health	2	25	56	13	2021.3846	22.6154	2.2811	
chatgpt	2	24	124	35	2023	4.8	1.4989	2
education	2	24	77	19	2021.1053	12.4211	3.3792	1
covid-19	2	23	47	13	2021.9231	3.5385	0.884	
review	2	19	70	12	2021.9167	34.25	3.3067	2
large language model	2	19	50	12	2023	2.9167	0.9108	2
natural language processing	3	54	505	100	2021.28	6.29	1.1837	6
natural language processing systems	3	52	484	90	2020.7667	5.5667	0.6711	3
machine learning	3	50	276	51	2021.4314	8.9412	1.0081	
deep learning	3	46	265	56	2021.1429	5.6607	0.6799	
natural languages	3	42	235	40	2021.5	4.1	0.6042	
language processing	3	37	165	25	2022.56	1.6	0.5094	2
learning algorithms	3	37	153	26	2021.0769	3.4231	0.4283	
virtual assistant	3	35	67	15	2021.0667	13.3333	1.5093	3
user interfaces	3	29	61	14	2020.8571	8.9286	1.7491	2
convolutional neural network	3	26	53	12	2020.9167	6.4167	0.7242	

Продовження на наступній сторінці

Табл. 1.1 – продовження з попередньої сторінки

Ключове слово	Кластер	W_{Links}	$W_{Total\ link\ strength}$	$W_{Occurrences}$	$S_{Avg.\ pub.\ year}$	$S_{Avg.\ citations}$	$S_{Avg.\ norm.\ citations}$	Значущість
query processing	3	24	82	14	2021.1429	1.4286	0.3373	
diagnosis	3	24	53	13	2021.5385	9.8462	1.3824	2
chatbot	4	56	1065	365	2021.137	6.3014	0.901	6
learning systems	4	46	223	43	2020.2326	7.186	0.4899	3
students	4	42	176	38	2020.9737	8.2105	0.9839	
e-learning	4	35	128	37	2020.8378	5.0541	0.6268	
learn+	4	27	53	13	2021.7692	3.9231	0.6622	2
sales	4	26	64	17	2021	5.5882	0.6926	
human computer interaction	4	26	53	14	2019.8571	6.0714	0.5181	
knowledge based systems	4	26	51	13	2019.5385	15.5385	0.7796	2
personnel training	4	25	62	18	2020.8889	6.5556	1.4229	1
curricula	4	24	58	12	2021	9.5833	2.41	3
customer service	4	24	49	15	2021.2	4.2667	0.4819	1
engineering education	4	24	46	12	2020.75	4.5833	0.4115	

У табл. 1.1 використовуються такі позначення [91, с. 6, 38]:

- ключове слово – визначений автором або системою індексування термін, характерний для публікації;
- кластер – номер групи ключових слів;
- W_{Links} – кількість зв'язків слова з іншими словами;
- $W_{Total\ link\ strength}$ – загальна сила зв'язків слова з іншими словами (тут кількість публікацій, в яких два терміни зустрічаються разом);
- $W_{Occurrences}$ – кількість документів, в яких зустрічається ключове слово;

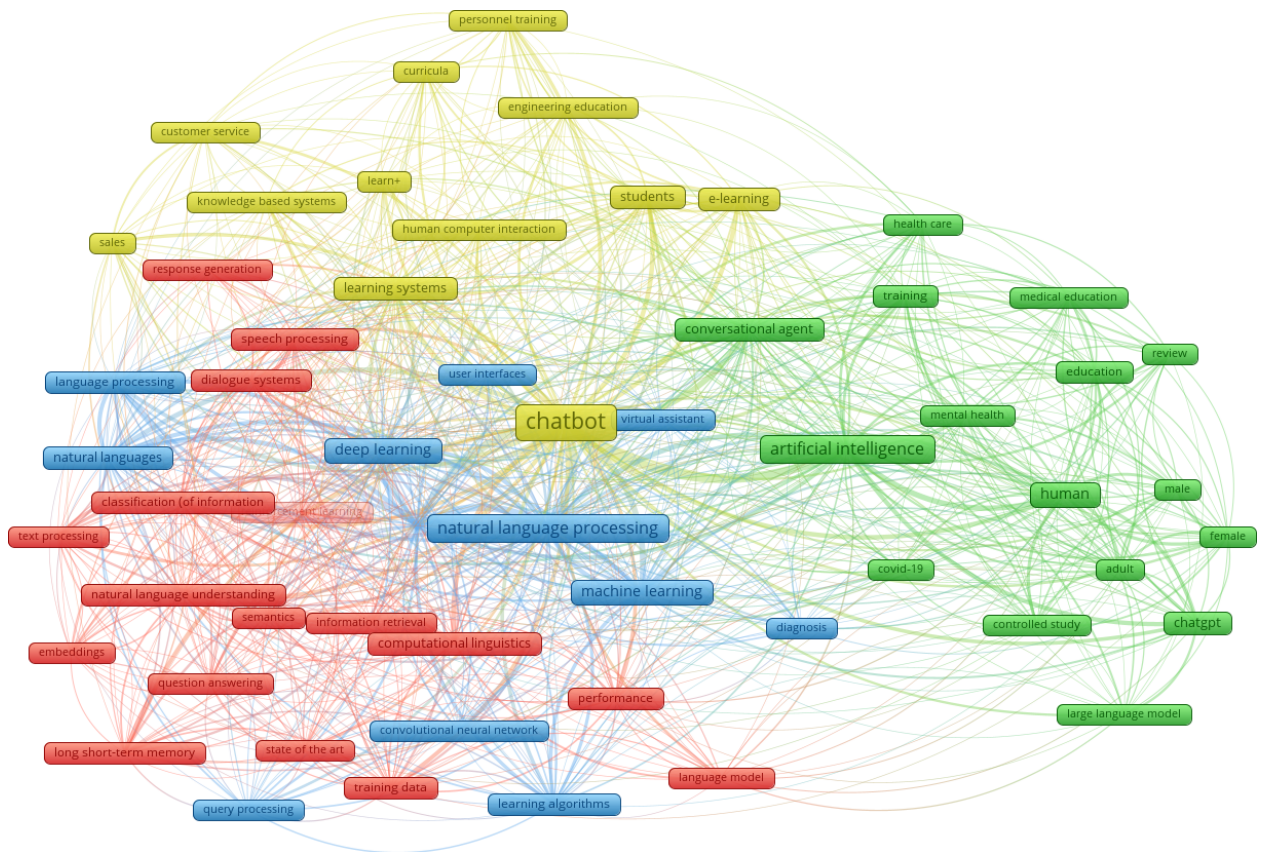


Рис. 1.4. Карта зв'язків ключових слів.

- $S_{Avg. pub. year}$ – середній рік публікації документів, в яких зустрічається ключове слово;
- $S_{Avg. citations}$ – середня кількість цитувань документів, в яких зустрічається ключове слово;
- $S_{Avg. norm. citations}$ – середня нормалізована кількість цитувань документів, в яких зустрічається ключове слово;
- значущість – обчислюється як сума, де для кожного ключового слова обирається 2, якщо воно має найбільше значення за колонкою, та 1, якщо воно має значення, що перевищує найбільшому.

У додатку 3.3 подано результати застосування генеративних чатботів ChatGPT (GPT 3.5), Bing (GPT 4.0, 3 режими), Claude 2 та Google Bard для визначення імен кластерів (табл. 1.2).

Залучені великі мовні моделі генерували декілька варіантів назв та описів кластерів на основі аналізу ключових слів, які входять до їх складу. Для кожного кластеру наведені всі запропоновані моделями варіанти з позначеннями А.1, А.2 і т.ін.

Імена кластерів ключових слів.

№	Пропозиції від чатботів	Обране ім'я
1	Advanced Language Processing and AI Applications (A.1); Language Processing and Information Retrieval (A.2); Natural Language Processing (NLP) (A.3); Natural Language Understanding and Generation (A.4); Natural Language Processing Applications (A.5); Technical NLP (A.6)	Опрацювання природної мови (A.3)
2	AI in Healthcare and Social Context (A.1); AI in Healthcare and Education (A.2); Artificial Intelligence (A.3); Artificial Intelligence and Health Care (A.4); Conversational AI and Impacts (A.5); Applications of NLP (A.6)	Застосування технологій опрацювання природної мови у суспільстві (A.6, A.5, A.1)
3	Core Technologies in Natural Language Processing and Machine Learning (A.1); Machine Learning and User Interaction (A.2); Machine Learning (A.3); Natural Language Processing and Machine Learning: (A.4); Core AI and ML Techniques (A.5); Machine Learning and Deep Learning for NLP (A.6)	Застосування машинного навчання для опрацювання природної мови (A.6, A.4)
4	Educational and Service-oriented AI Applications (A.1); AI in Education and Customer Service (A.2); Learning Systems (LS) (A.3); Chatbot and Education (A.4); AI for Training and Customer Service (A.5); Chatbots and Learning Systems (A.6)	Чат-боти у навчанні та сфері послуг (A.1, A.2, A.4, A.5)

Вибір остаточних назв кластерів, представлених у стовпці “Обране ім'я” табл. 1.2 як українською, так і англійською мовами, здійснювався шляхом ретельного аналізу та критичної оцінки авторами усіх згенерованих різними моделями пропозицій. Ключовими критеріями відбору стали змістовність, чіткість, лаконічність назв із одночасним максимально повним охопленням ключових термінів у складі кластерів.

Обрані назви кластерів не лише узагальнюють їх зміст, а й логічно

відображають виявлені за результатами бібліометричного аналізу основні напрями досліджень у галузі навчання чат-ботів: 1) базові методи опрацювання природної мови; 2) застосування відповідних технологій у різних сферах суспільної діяльності; 3) використання машинного навчання як провідного інструментарію розробки систем опрацювання природномовних даних; 4) спектр ключових застосувань чат-ботів, зокрема у освітній галузі та сфері послуг.

Таким чином, остаточний вибір назв кластерів став результатом поєднання можливостей великих мовних моделей із експертною думкою авторів і ґрунтується на чітких обґрунтованих позиціях, спрямованих на уникнення невизначеностей у трактуванні отриманих результатів.

На рис. 1.5 подано карту зв'язків ключових слів із значущістю не менше 1, тобто найбільш значущих ключових слів, що належать до четвірки виділених кластерів. Дана карта дозволяє проаналізувати взаємозв'язки між центральними поняттями дослідження та візуалізувати провідні напрямки у межах кожного кластера.

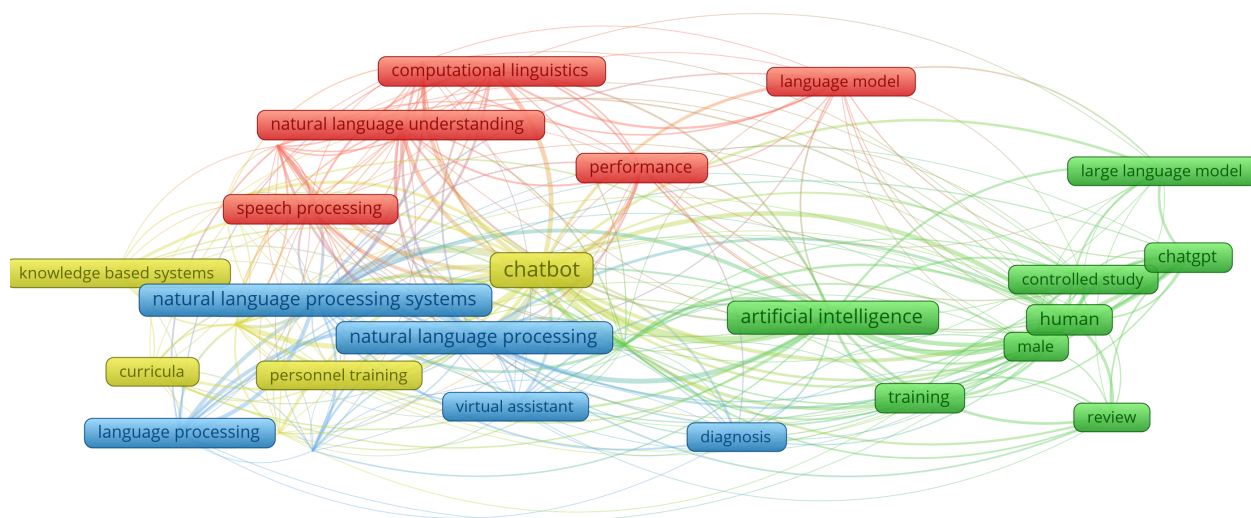


Рис. 1.5. Карта зв'язків найбільш значущих ключових слів.

Побудова цієї фокусної карти здійснювалась з метою концентрації уваги на найбільш важливих ключових словах, відібраних за принципом максимальних та передмаксимальних значень низки показників (кількість зв'язків, загальна вага зв'язків, кількість згадувань у публікаціях, середні значення року публікації, цитованості тощо). Візуалізація найбільш значущих термінів та їх взаємозв'язків дає змогу ідентифікувати пріоритетні напрямки у кожній з досліджуваних сфер для окреслення перспективних

напрямків майбутніх наукових пошуків.

Такий підхід дозволяє деталізувати і поглибити висновки, зроблені на основі загальної карти ключових слів, концентруючись безпосередньо на центральних поняттях навчання чат-ботів як провідних агентів штучного інтелекту.

Висновки до 1 розділу

1. Бібліометричний аналіз 549 джерел із бази Scopus з проблеми навчання чат-ботів надав можливість:
 - а) виявити нижню хронологічну межу (2018 рік), починаючи з якої спостерігається стійке зростання кількості публікацій із навчання чат-ботів;
 - б) висловити припущення про те, що причиною суттєвого зростання кількості робіт у 2023 році (на 55 порівняно із 2022 роком) є поява у відкритому доступі ChatGPT та пов'язаний з цим сплеск інтересу наукової спільноти до тематики великих мовних моделей та чат-ботів: для підтвердження чи спростування цього припущення можна проаналізувати контекст ключових слів у публікаціях 2023 року на предмет безпосередніх згадувань ChatGPT або близьких до нього систем, а також розглянути динаміку цитувань праць, присвячених ChatGPT, в інших роботах за цей період;
 - в) згрупувати у 4 кластери авторські та індексні ключові слова джерел із навчання чат-ботів: 1) опрацювання природної мови; 2) застосування технологій опрацювання природної мови у суспільстві; 3) застосування машинного навчання для опрацювання природної мови; 4) чат-боти у навчанні та сфері послуг.
2. Для визначення найбільш значущих ключових понять дослідження у кожному кластері біли визначені ключові слова, що мають найбільше та наступне за ним значення за такими показниками: кількість зв'язків, загальна сила зв'язків, кількість документів із заданими ключовим словом та середні: рік публікації, кількість цитувань,

нормалізована кількість цитувань. Для відібраних ключових слів була обчислена їх значущість від 1 до 6:

- а) у першому кластері (опрацювання природної мови) найбільш значущими є 7 ключових слів (41%): computational linguistics (6), natural language understanding (3), performance (2), language model (2), speech processing (1);
 - б) у другому кластері (застосування технологій опрацювання природної мови у суспільстві) найбільш значущими є 10 ключових слів (63%): artificial intelligence (6), large language model (2), chatgpt (2), human (2), training (2), review (2), conversational agent (1), controlled study (1), education (1), male (1);
 - в) у третьому кластері (застосування машинного навчання для опрацювання природної мови) найбільш значущими є 6 ключових слів (50%): natural language processing (6), natural language processing systems (3), virtual assistant (3), language processing (2), user interfaces (2), diagnosis (2);
 - г) у четвертому кластері (чат-боти у навчанні та сфері послуг) найбільш значущими є 7 ключових слів (58%): chatbot (6), learning systems (3), curricula (3), learn+ (2), knowledge based systems (2), personnel training (1), customer service (1).
3. Під час аналізу карти ключових слів було виявлено, що третій кластер (застосування машинного навчання для опрацювання природної мови) має певні перетини зі змістом першого (опрацювання природної мови) та другого (застосування технологій опрацювання природної мови у суспільстві) кластерів. Це пояснюється тим, що методи машинного навчання, зокрема глибинного навчання, є фундаментальними інструментами для розробки систем опрацювання природної мови та їх застосувань. Тому деякі ключові терміни третього кластера, такі як опрацювання природної мови, віртуальні асистенти, нейронні мережі тощо, тісно пов'язані з концепціями інших кластерів. Це свідчить про тісну інтеграцію різних напрямків досліджень у єдиному науковому полі створення інтелектуальних розмовних агентів. Подальший

аналіз перетинів між кластерами може стати перспективним напрямком майбутніх наукових пошуків для більш чіткого розмежування предметних областей.

4. Аналіз карти зв'язків найбільш значущих ключових слів надав можливість визначити провідні напрями досліджень:
 - а) у опрацюванні природної мови засобами обчислювальної лінгвістики – розуміння природної мови, побудова мовних моделей та розпізнавання мовлення;
 - б) у застосуванні технологій штучного інтелекту для опрацювання природної мови – контрольоване використання великих мовних моделей та чат-ботів (зокрема, ChatGPT) у освіті;
 - в) у застосуванні машинного навчання для опрацювання природної мови – використання віртуальних асистентів, природномовних інтерфейсів користувача та інших систем опрацювання природної мови, зокрема, для діагностики;
 - г) у застосуванні чат-ботів у навчанні та сфері послуг – використання чат-ботів, систем навчання та систем управління знаннями для доповненого та адаптивного навчання.

Визначені напрями можуть бути корисними для організацій при розробці стратегій використання штучного інтелекту та інтеграції розмовних агентів у виробничу діяльність.

5. Результати проведеного бібліометричного аналізу можуть знайти застосування:
 - в освіті:
 - а) визначені пріоритетні напрями досліджень у галузі навчання чат-ботів можуть бути використані для формування тематики курсів, програм підготовки фахівців з опрацювання природної мови та розробки систем штучного інтелекту;
 - б) виявлені ключові поняття здатні слугувати основою для розробки навчальних матеріалів, покликаних висвітлити

найважливіші концепції та технології створення розмовних агентів;

- в) огляд застосувань чат-ботів в освітньому процесі окреслює перспективні шляхи їх інтеграції в системи адаптивного і персоналізованого навчання;
- у наукових дослідженнях:
 - а) карта зв'язків ключових понять може слугувати базою для формування наукових гіпотез та побудови концептуальних моделей під час проєктування досліджень;
 - б) результати аналізу відкривають перспективи для подальших бібліометричних студій динаміки та взаємозв'язків у предметному полі;
 - в) виявлені пріоритетні напрямки визначають актуальні вектори майбутніх наукових розвідок у галузі штучного інтелекту та його використання у наукових дослідженнях у галузі соціальних наук.

РОЗДІЛ 2

МОДЕЛІ НАВЧАННЯ ЧАТ-БОТІВ

2.1. Навчання з учителем

Навчання з учителем (supervised learning), або кероване навчання є одним із основних підходів до машинного навчання, який широко використовується для побудови чат-ботів та інших діалогових систем [90, с. 3940] (рис. 2.1). Цей підхід передбачає навчання моделі на розмічених даних, де кожен приклад містить вхідну інформацію (запит користувача) та відповідну цільову відповідь (репліку чат-бота).

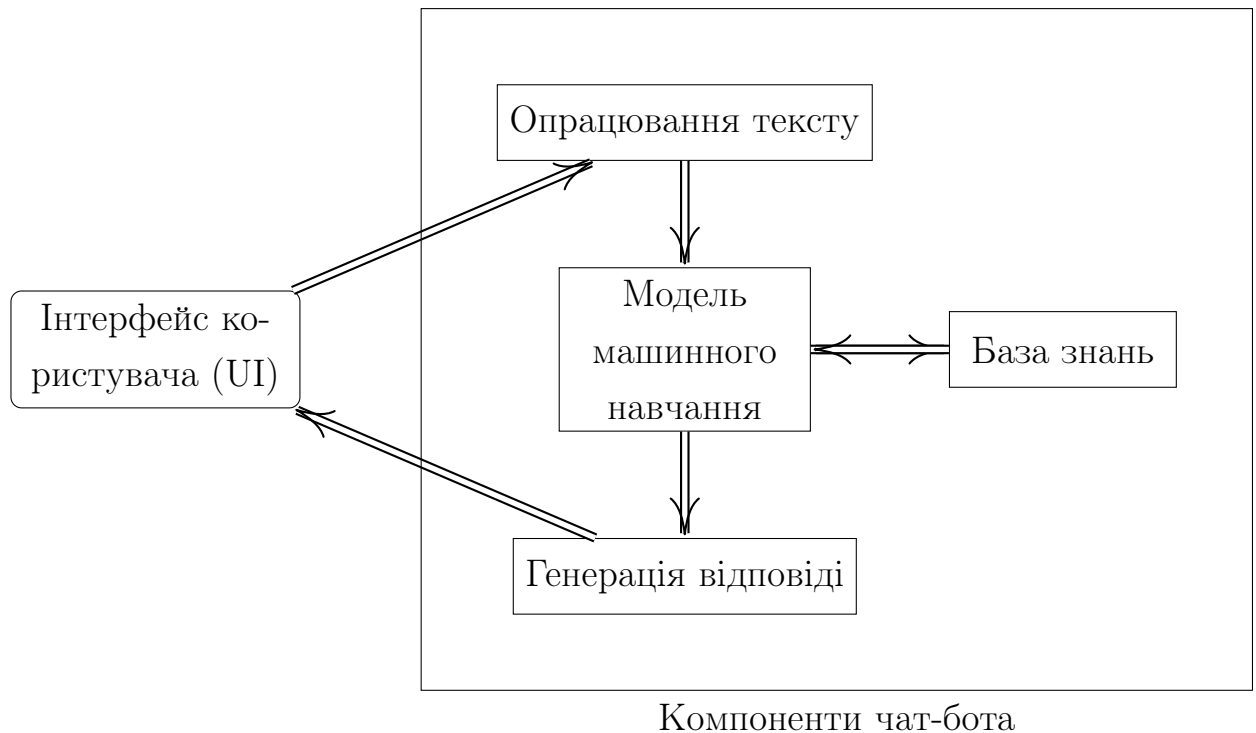


Рис. 2.1. Загальна схема роботи чат-бота (за [67, с. 2]).

Patil, Mudaliar, Kamat [67] виокремлюють наступні компоненти чат-боту [67, с. 2]):

- *опрацювання тексту* – векторні подання слів (word embeddings) є векторними зображеннями слів у межах певного словника, що дозволяють краще реалізовувати та використовувати статистичні моделі машинного навчання;

- *модель машинного навчання* – концепція штучної нейронної мережі, яка широко використовується для опрацювання входу, класифікації та генерації найбільш відповідної відповіді на вхідний запит;
- *база знань* – набір даних, що використовується для навчання моделі, який може бути відкритим або закритим за предметною областю (доменом): чат-боти відкритого домену поступаються релевантністю та точністю відповідей, а чат-боти закритого домену працюють добре завдяки обмеженим, але чітко визначеним рамкам набору даних;
- *генерація відповіді*: відповідь, що повертається на вхідний запит, або обирається з бази, або генерується залежно від особливостей вхідних векторів, словника та навченого класифікатора.

Процес навчання з учителем для чат-ботів включає наступні етапи:

1. Вибір архітектури моделі.
2. Збір та анотування даних для навчання.
3. Попереднє опрацювання даних.
4. Навчання моделі.
5. Оцінка якості навченої моделі.

Серед проблем навчання з учителем можна виділити потребу у великому обсязі якісних розмічених даних, збір та анотування яких може бути трудомістким та затратним процесом [90, с. 3940]. Одним із способів її розв’язання є використання напівкерованого навчання, яке надає можливість навчати модель як на розмічених, так й на нерозмічених даних [90, с. 3943].

2.1.1. Seq2Seq моделі на основі рекурентних нейронних мереж

Seq2Seq (Sequence-to-Sequence) моделі на основі рекурентних нейронних мереж із застосуванням модулів LSTM (Long Short-Term Memory) є одним із підходів до побудови діалогових систем з використанням методів машинного навчання [44, с. 230].

Архітектура Seq2Seq складається з двох основних компонентів: кодувальник та декодувальник. Кодувальник обробляє вхідну послідовність (запит користувача) і генерує її векторне подання у прихованому просторі. Декодувальник, у свою чергу, використовує це подання для генерації вихідної послідовності (відповіді чат-бота) [67, с. 9].

Ключовою особливістю Seq2Seq моделей є використання рекурентних шарів, зокрема LSTM, для обробки послідовних даних [67, с. 4]. LSTM модулі дозволяють ефективно моделювати довготривалі залежності у послідовностях, що є критично важливим для генерації зв'язних та контекстно-релевантних відповідей у діалозі [67, с. 5].

Seq2Seq моделі на основі LSTM мають певні обмеження. Зокрема, вони можуть страждати від проблеми “затухаючих градієнтів” (vanishing gradients) при опрацюванні довгих послідовностей [67, с. 4].

Тим не менш, Seq2Seq архітектури на базі LSTM досі залишаються важливим будівельним блоком сучасних діалогових систем. Вони часто використовуються як базові моделі, які надалі можуть бути вдосконалені та розширені за рахунок інших методів, таких як модулі уваги (attention) або ієрархічні архітектури [90, с. 3941-3942].

У контексті розробки чат-ботів Seq2Seq моделі на основі LSTM зарекомендували себе як ефективний інструмент для генерації відповідей на запити користувачів у різноманітних предметних областях [42].

2.1.2. Трансформерні архітектури

Трансформерні архітектури, зокрема моделі сімейства GPT (Generative Pre-trained Transformer), являють собою передовий підхід до побудови мовних моделей та діалогових систем. Ці архітектури набули надзвичайної популярності завдяки тому, що вони послідовно перевершують інші мовні моделі, такі як рекурентні нейронні мережі. Ці моделі базуються на механізмі самоуваги (self-attention) і здатні ефективно обробляти та генерувати послідовності довільної довжини, зберігати довші історії розмов, що веде до послідовної, контекстної та покращеної бесіди [29, с. 2].

Архітектура трансформера складається з кількох блоків кодувальника та декодувальника, кожен з яких містить шари самоуваги та прямо-

го поширення (рис. 2.2). Завдяки цьому трансформери здатні паралельно обробляти всі елементи вхідної послідовності [8, с. 2], що значно пришвидшує процес навчання порівняно з рекурентними моделями [8, с. 10].

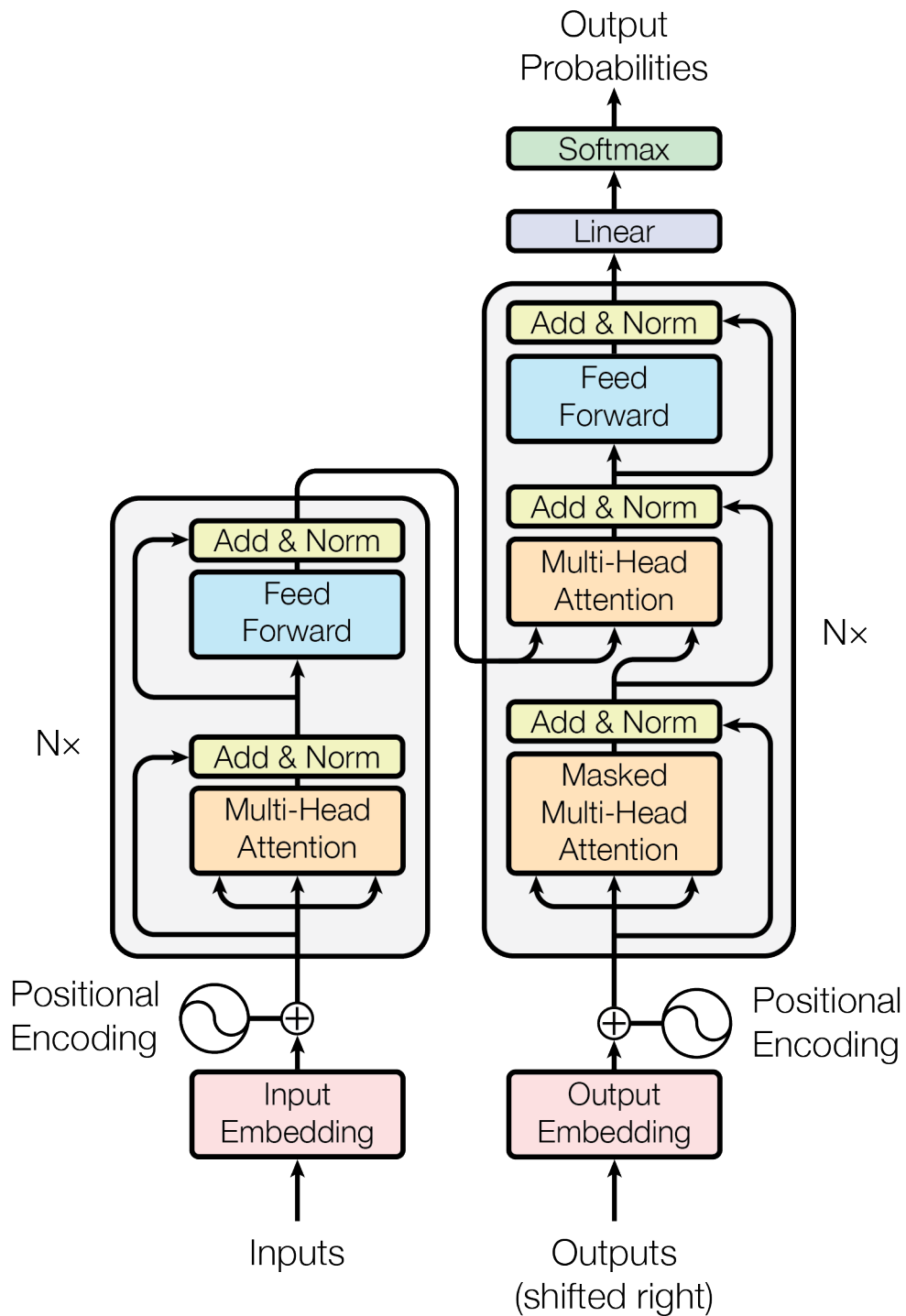


Рис. 2.2. Архітектура трансформера [8, с. 3].

Дослідники також продемонстрували, як продуктивність великих мовних моделей, навчених на великому корпусі даних, може бути покращена при донавчанні (fine-tuning) на конкретних завданнях. Це можна чітко

спостерігати, коли ми дивимося на серію GPT від OpenAI (рис. 2.3): GPT, GPT-2, GPT-3 та GPT-4, яка здатна виконувати завдання типу відповіді на запитання, розуміння прочитаного, узагальнення тексту, генерація тексту чи моделювання розмови [29, с. 2].

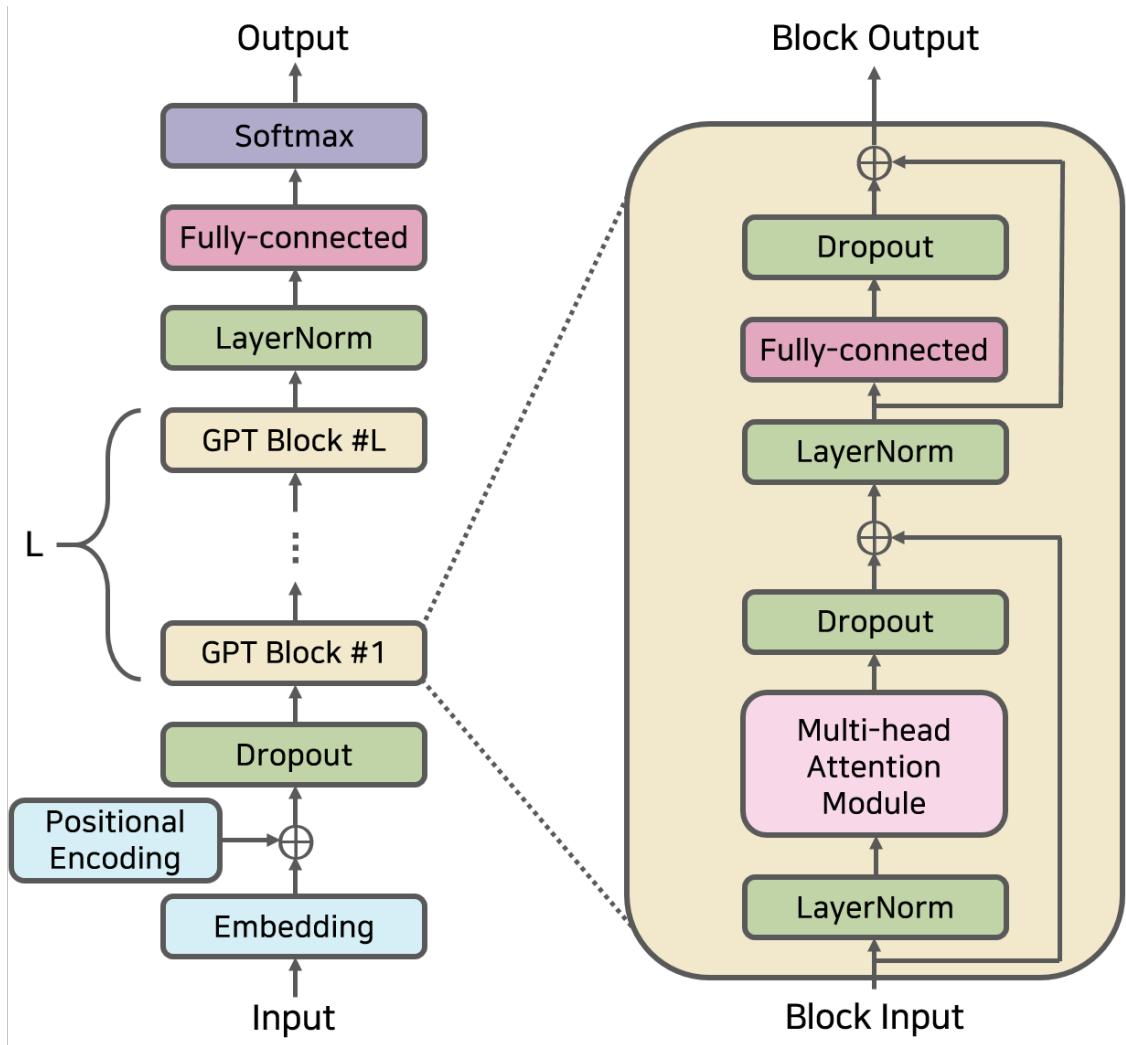


Рис. 2.3. Conceptual architecture of a GPT model [51, с. 4].

Незважаючи на вражаючу продуктивність моделей GPT, відомо, що вони демонструють явище, яке називається галюцинацією, при якому вони генерують результати, які є контекстуально неправдоподібними або несумісними з реальним світом [51, с. 1]. Незважаючи на це, трансформерні архітектури, і зокрема моделі GPT, стали де-факто стандартом для побудови сучасних діалогових систем та чат-ботів. Їхня здатність обробляти контекст, генерувати людиноподібні відповіді та адаптуватися до різних предметних областей робить їх потужним інструментом для створення інтелектуальних асистентів та віртуальних співрозмовників.

2.2. Навчання з підкріпленням

Навчання з підкріпленням (reinforcement learning) є одним із підходів до машинного навчання, який набуває все більшої популярності у галузі побудови чат-ботів та інших діалогових систем. На відміну від навчання з учителем, де модель навчається на розмічених парах “запит-відповідь”, навчання з підкріпленням дозволяє моделі навчатися через взаємодію з середовищем та отримання зворотного зв'язку у вигляді винагород на свої дії.

Для випадку, коли середовищем є користувач, процес навчання з підкріпленням для чат-ботів можна представити у такий спосіб:

1. Агент (чат-бот) взаємодіє з користувачем у дискретні моменти часу.
2. На кожному кроці агент отримує інформацію про поточний стан діалогу та генерує відповідь.
3. Користувач надає зворотний зв'язок у вигляді винагороди, яка відображає якість згенерованої відповіді у контексті поточного діалогу. Функція винагороди визначає, яким чином обчислюється винагорода за кожну дію агента, та може враховувати різні аспекти, такі як релевантність відповіді, її граматична правильність, задоволеність користувача тощо.
4. Агент оновлює свою стратегію таким чином, щоб максимізувати функція цінності – сумарну очікувану винагороду протягом всього діалогу.

Перевагою чат-ботів на основі навчання з підкріпленням є їхня здатність адаптуватися до різних сценаріїв взаємодії з користувачем та оптимізувати свою поведінку для досягнення бажаних результатів. Крім того, такі агенти можуть навчатися на основі неявного зворотного зв'язку від користувача, що дозволяє уникнути необхідності в явному розмічуванні даних.

Навчання з підкріпленням на основі зворотного зв'язку від користувача (reinforcement learning from human feedback, RLHF (рис. 2.4)) зазвичай виконується ітераційно [45, с. 92]:

1. Модель винагороди спочатку навчається на наборі даних людських суджень.
2. Модель стратегії навчена максимізувати винагороду за допомогою поточної моделі винагороди.
3. Потім люди оцінюють результати моделі стратегії для створення нового набору даних.
4. Цей новий набір даних використовується для оновлення моделі винагороди, що робить її більш точною.
5. Модель стратегії знову налаштовується за допомогою оновленої моделі винагороди.

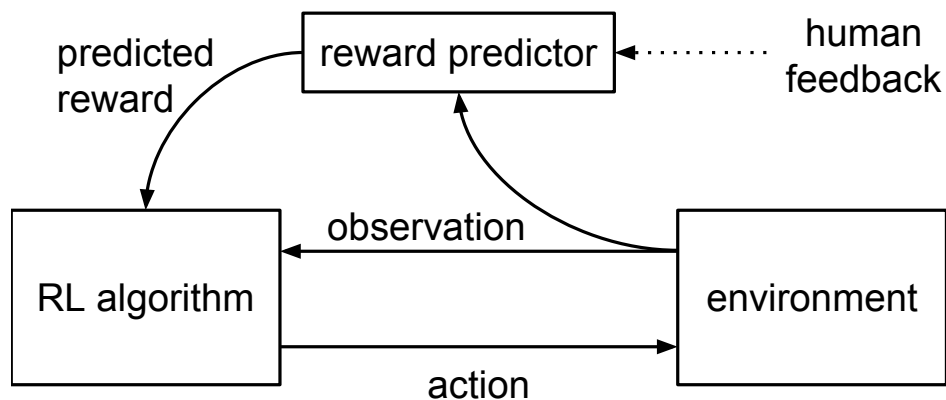


Рис. 2.4. Навчання з підкріпленням на основі зворотного зв'язку від користувача [26, с. 2].

Генеративно-змагальні мережі (Generative Adversarial Networks, GAN) є одним із найбільш інноваційних підходів до керування діалогом у чат-ботах. GAN складаються з двох нейронних мереж – генератора та дискримінатора, які змагаються одна з одною у процесі навчання. Мета генератора – створити відповіді, які будуть настільки реалістичними, що дискримінатор не зможе відрізнити їх від справжніх відповідей людини. Мета дискримінатора – навчитися розрізняти згенеровані відповіді та справжні відповіді.

Chou, Hsueh [21] запропонувано метод створення чат-бота за допомогою моделі, яка генерує послідовні речення на основі генеративної змагальної мережі. Архітектура моделі містить генератор, який генерує

різноманітне речення, і дискримінатор, який оцінює речення. У генераторі поєднані модель уваги, яка реагує на відстеження станів розмови, з моделлю Seq2Seq, за допомогою LSTM для отримання інформації про речення. Для дискримінатора розраховуються два типи винагород, щоб призначити низькі винагороди за повторювані речення та високі винагороди за різноманітні речення. За такого підходу середовищем виступає сама модель.

Запропонована Tran, Le, Huynh [88] модель поєднує навчання з підкріпленням та генеративно-змагальні мережі для генерації як точних, так і подібних до людських відповідей (рис. 2.5).

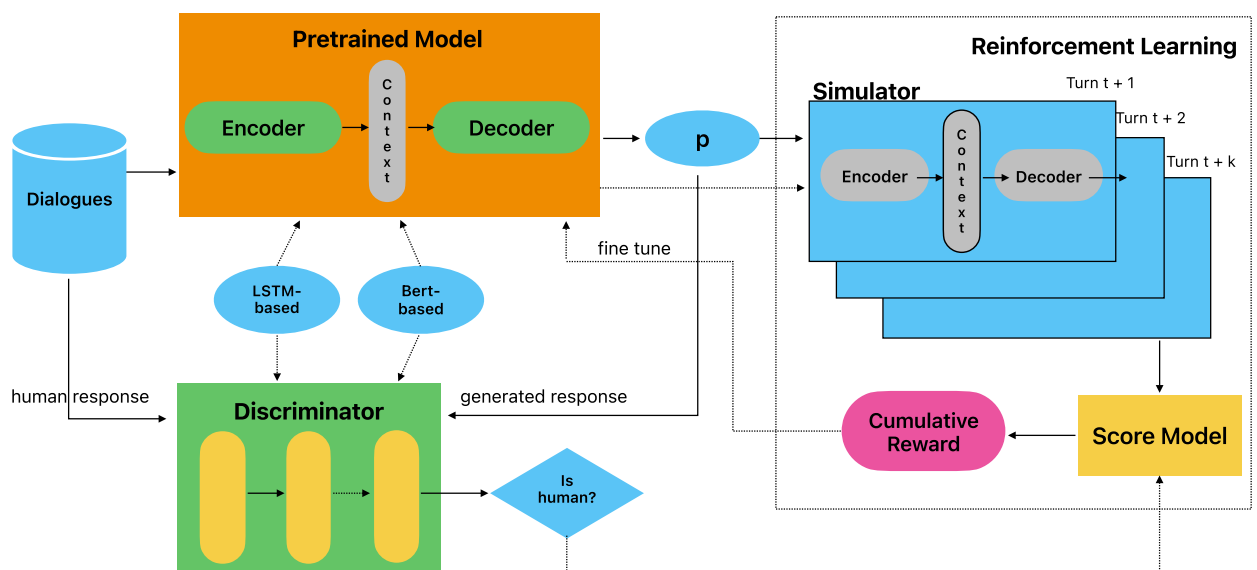


Рис. 2.5. Гібридна модель, що поєднує RLHF та GAN для навчання чат-ботів (за [88, с. 75962]).

2.3. Трансферне навчання

Трансферне навчання (transfer learning), або навчання із пересенням, є підходом до машинного навчання, який дозволяє використовувати знання, отримані моделлю при вирішенні однієї задачі, для покращення її ефективності при вирішенні іншої, схожої задачі [45, с. 81].

Ідея трансферного навчання полягає у тому, щоб спочатку навчити модель на великому обсязі даних для виконання загального завдання (наприклад, передбачення наступного слова у реченні), а потім донавчити її

на меншій кількості даних для виконання специфічного завдання (наприклад, генерації відповідей чат-бота у певній предметній області).

Адаптація моделей під конкретну предметну область (домен) є важливим напрямком трансферного навчання, який дозволяє покращити якість роботи чат-ботів у специфічних сферах застосування. Автори [33, с. 4116] розрізняють два випадки, коли два домени перекриваються та коли один домен є розширенням іншого, й вказують на переваги використання трансферного навчання для навчання спеціалізованих чат-ботів:

- навчання чат-ботів на меншій кількості даних: у середовищах з обмеженням даних моделі, навчені за допомогою трансферного навчання, досягають кращих результатів навчання та тестування, ніж моделі, навчені незалежно;
- краща продуктивність чат-боту: використання трансферного навчання має значний позитивний вплив на продуктивність, навіть коли всі дані з цільового домену доступні.

Тонке налаштування (файнтюнінг, від fine-tuning) існуючих великих мовних моделей замість створення їх з нуля часто є більш практичним і ефективним підходом з таких причин [45, с. 80-82]:

- ефективність ресурсів – файнтюнінг використовує попередньо навчені моделі, що дозволяє досягти високої продуктивності без значних інвестицій у обчислювальні ресурси;
- ефективність даних – файнтюнінг дозволяє скористатися великим обсягом даних, на яких були попередньо навчені моделі, потребуючи лише меншого, специфічного для задачі набору даних для адаптації;
- перенесення навчання – попередньо навчені моделі мають загальне розуміння мови, контексту і певні предметні знання, які переносяться на конкретну задачу при файнтюнінгу;
- висока продуктивність – файнтюнінг дозволяє використовувати сучасні архітектури, ретельно оптимізовані та протестовані експертами галузі;

- нижчий бар'єр входу – файнтюнінг є більш доступним для організацій та осіб, які не мають достатньої інфраструктури для навчання моделей з нуля;
- постійне навчання – попередньо навчені моделі можна постійно оновлювати та налаштовувати для різних завдань, що робить їх універсальними та адаптивними;
- широке застосування – одну попередньо навчену модель можна налаштувати для багатьох доменів і завдань.

Процес тонкого налаштування включає такі кроки (рис. 2.6) [45, с. 82-84]:

1. Завантаження параметрів попередньо навченої моделі.
2. Підготовка специфічного для завдання набору даних.
3. Вилучення ознак з використанням попередньо навчених шарів моделі.
4. Регулювання параметрів моделі за допомогою зворотного поширення помилки та градієнтного спуску з меншою швидкістю навчання.
5. Оновлення градієнтів, застосування регуляризації для запобігання перенавчання.
6. Вибір стратегії файнтюнінгу: повне або часткове налаштування моделі.
7. Оцінка ефективності та оптимізація гіперпараметрів моделі.

Під час тонкого налаштування нейронної мережі відбуваються зміни на рівні архітектури та окремих нейронів [45, с. 85-88]:

- коригування ваг і зміна виходів функцій активації;
- налаштування верхніх шарів моделі та заморожування нижніх шарів;
- регулювання простору ознак для адаптації до нового завдання;
- заміна останнього шару для відповідності новому завданню;

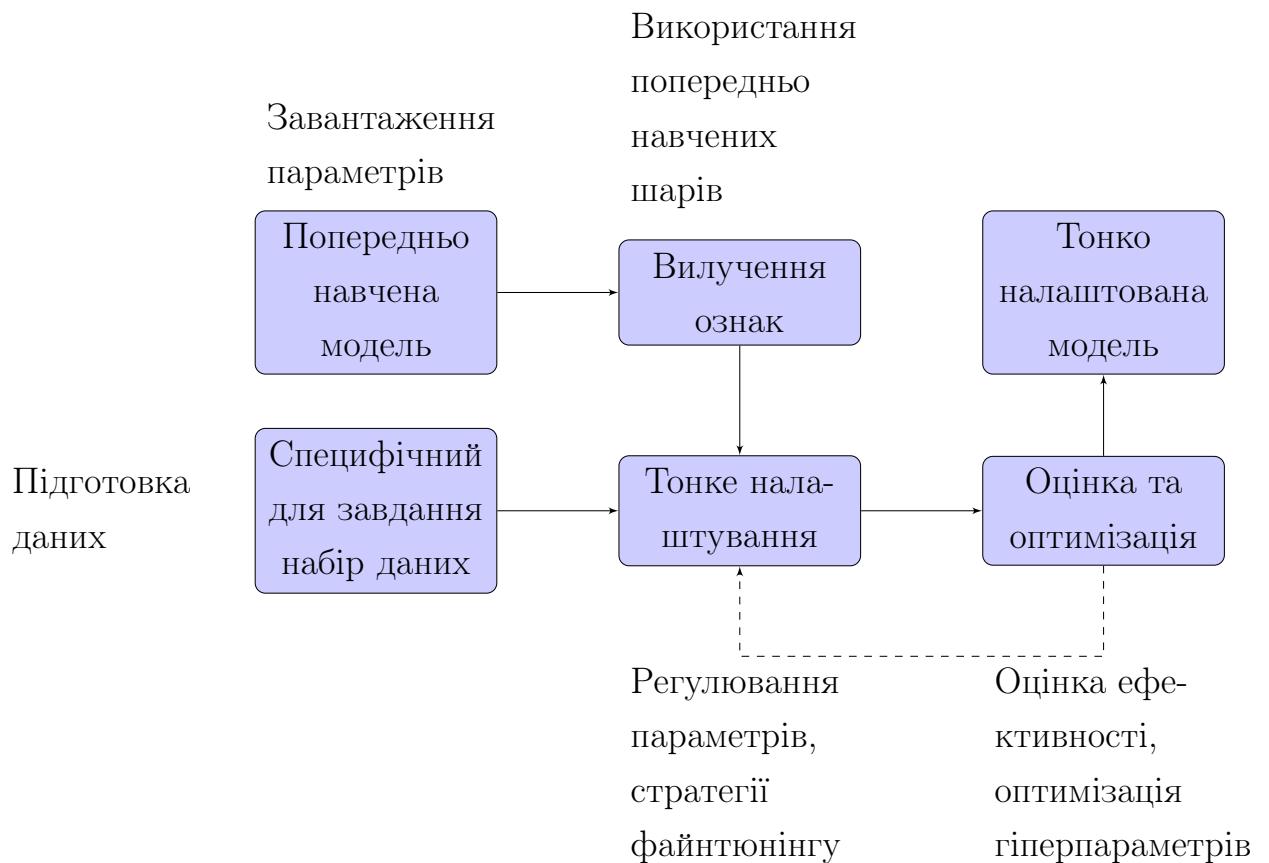


Рис. 2.6. Схема процесу тонкого налаштування.

- оновлення параметрів нормалізації пакетів даних.

Тонке налаштування дозволяє ефективно адаптувати попередньо навчені моделі до конкретних завдань, зберігаючи їхнє “чуття” та оптимізуючи для специфіки нової задачі.

2.4. Оцінка ефективності навчання чат-ботів

Taulli [84] розглядає ряд метрик ефективності великих мовних моделей (табл. 2.1).

Human evaluation (оцінка людьми, або людська оцінка) є групою методів, які передбачають безпосередню оцінку якості роботи чат-ботів реальними користувачами або експертами. Ці методи дозволяють отримати більш цілісну та комплексну оцінку ефективності моделі, враховуючи такі аспекти як релевантність, зв’язність, природність та корисність згенерованих відповідей.

Для людської оцінки якості тонко налаштованої моделі необхідно ви-

Метрики для оцінки моделей генерації тексту (за [84, с. 36-37]).

Метрика	Опис
BERTScore	оцінює шляхом порівняння згенерованого тексту з еталонним текстом з використанням вкладень BERT
Перплексивність	оцінює, наскільки добре розподіл ймовірностей, передбачений моделлю, відповідає фактичному розподілу даних
BLEU	обчислює оцінки точності n-грам подібності між згенерованим і еталонним текстом
ROUGE	обчислює перетин n-грам між згенерованим і еталонним текстом

конати такі кроки:

1. Використати тонко налаштовану модель для генерації нових текстів.
2. Залучити людей (експертів у відповідній області або звичайних користувачів) для оцінки якості згенерованих текстів.
3. Розробити систему оцінювання, яка враховує такі аспекти, як граматики, зв'язність, релевантність, стиль і загальна якість.
4. Зібрати відгуки та оцінки від кількох людей та усереднити результати, щоб отримати загальну оцінку ефективності моделі.

Висновки до 2 розділу

1. Навчання з учителем є одним з основних підходів, який передбачає навчання моделі на розмічених парах “запит-відповідь”. Для цього використовуються архітектури на основі рекурентних нейронних мереж (наприклад, Seq2Seq з LSTM) та трансформерів (наприклад, GPT). Ці моделі здатні генерувати контекстно-релевантні відповіді, але потребують великих обсягів якісних розмічених даних.
2. Навчання з підкріпленням дозволяє моделі навчатися через взаємодію з середовищем (користувачем) та отримання зворотного зв'язку у вигляді винагород. Цей підхід реалізується за допомогою

генеративно-змагальних мереж (GAN) та ітеративного процесу навчання з підкріпленням на основі зворотного зв'язку від користувача (RLHF).

3. Трансферне навчання полягає у використанні знань, отриманих моделлю при вирішенні однієї задачі, для покращення її ефективності при вирішенні іншої, схожої задачі. Найпоширенішим підходом є тонке налаштування (файнтюнінг) попередньо навченої моделі на специфічному наборі даних для адаптації до конкретної предметної області.
4. Для оцінки ефективності навчання чат-ботів використовуються як автоматичні метрики (BERTScore, перплексивність, BLEU, ROUGE), так і методи людської оцінки, які дозволяють врахувати релевантність, зв'язність та природність згенерованих відповідей.

РОЗДІЛ 3

ЕКСПЕРИМЕНТИ

3.1. Створення наборів даних для навчання чат-ботів

Для проведення експериментів з навчання чат-ботів було створено два набори даних, що містять тексти наукових публікацій з галузі інформаційних технологій.

Набір даних із публікацій “CEUR Workshop Proceedings” – це видання, яке публікує матеріали наукових конференцій та семінарів, що охоплюють широкий спектр досліджень у галузі комп’ютерних наук та інженерії. Для формування набору даних було виконано наступні кроки:

- завантаження сайту <https://ceur-ws.org/> за допомогою утиліти `wget`;
- видобування текстів із 68791 PDF файлів, що відповідають томам 1-3583 за 1995-2024 роки;
- створення єдиного текстового файлу `ceur-ws.txt` розміром 1917 Мб (2009797694 байт).

Відносний розподіл публікацій за мовами статей згідно Scopus: англійська – 94.807%, російська – 1.368%, німецька – 1.101%, іспанська – 0.772%, португальська – 0.691%, турецька – 0.596%, французька – 0.340%, українська – 0.160%, італійська – 0.131%, чеська – 0.015%, інші мови разом – 0.019%.

Набір даних із публікацій журналу “Інформаційні технології та засоби навчання” – це видання, яке публікує статті з теоретичних та прикладних аспектів використання інформаційно-комунікаційних технологій в освіті. Процедура створення цього набору включала:

- завантаження сайту журналу <https://journal.iitta.gov.ua/index.php/itlt> утилітою `wget`;
- видобування текстів із 1732 PDF файлів, що відповідають томам 1-100 журналу за 2006-2024 роки;
- формування текстового файлу `journal.txt` об’ємом 107 Мб (112245281 байт).

Специфікою цього набору є вужча предметна спеціалізація текстів, що зосереджуються на питаннях цифрової трансформації освіти, розробки та використання електронних освітніх ресурсів, технологіях дистанційного та змішаного навчання тощо.

Орієнтований відносний розподіл публікацій за мовами статей згідно Web of Science: українська – 52.54%, російська – 26.73%, англійська – 20.73%.

Сформовані набори даних відрізняються за обсягом, часовим охопленням, тематичною спрямованістю текстів та розподілом мов написання текстів, що дозволяє використовувати їх для порівняльного аналізу ефективності моделей чат-ботів, навчених на різномірних текстових корпусах.

3.2. Вибір моделей для донавчання

Ураховуючи поширеність моделей сімейства GPT (Generative Pre-trained Transformer), були вивчені можливості доступу як до їх сучасних версій (GPT 3.5, 4.0), так й до альтернативних моделей (Gemini 1.0, Claude 3):

- OpenAI надає програмний та візуальний [64] інтерфейс для донавчання моделей gpt-3.5-turbo-0125, gpt-3.5-turbo-1106, gpt-3.5-turbo-0613, babbage-002, davinci-002 та gpt-4-0613. Жодна з цих моделей не є безоплатною для донавчання – найдешевша babbage-002 потребує оплати 0.40\$ за 1 млн. лексем, у той час як досконаліші моделі дорожчі у 20 та більше разів. Крім оплати за донавчання, є оплата за використання – так, вартість вхідних та вихідних повідомлень для моделі davinci-002 складає 12.00\$ за 1 млн. лексем.
- Google також надає програмний та візуальний [35] інтерфейс для донавчання моделей gemini-1.0-pro, gemini-1.5-pro, gemini-1.5-flash, gemini-1.0-pro, але останній в Україні недоступний [36].
- Anthropic станом на 1 квартал 2024 року не надає можливості донавчання моделей сімейства Claude (крім відібраних груп експертів) [31].

Ураховуючи непрогнозованість бюджету на донавчання та використання донавчених моделей, замість сучасних версій GPT була обрана

історична модель GPT-2 [47], навчена на внутрішньому наборі даних фірми OpenAI – WebText (40 Гб). Доцільність використання цієї моделі як основи для чат-боту показана у роботі [61].

Hugging Face надає доступ до всіх базових версій моделі GPT-2 за посиланням <https://huggingface.co/openai-community>: GPT2, GPT2-Medium, GPT2-Large та GPT2-XL, навчених на англійських текстах. Ураховуючи сумірність набору даних із публікацій “CEUR Workshop Proceedings” (2 Гб) та набору даних WebText (40 Гб), а також переважну мову (англійську), для донавчання доцільним був вибір моделі GPT2-XL.

Hugging Face містить велику (більше 11 тисяч) кількість донавчених моделей GPT-2 (<https://huggingface.co/models?sort=trending&search=gpt2>), серед яких менше 10 були донавчені на україномовних текстах (табл. 3.1).

Ураховуючи малий обсяг набору даних із публікацій журналу “Інформаційні технології та засоби навчання” та переважання у ньому української мови, доцільним є вибір для донавчання на ньому моделі, яка вже була донавчена на україномовному наборі текстів. Серед представлених у табл. 3.1 моделей для донавчання була обрана `malteos/gpt2-uk` через її невеликий розмір (124М параметрів) та неспецифічність наборів даних OSCAR та Wikimedia порівняно з наборами, використаними для донавчання інших моделей – художніх текстів, новин, творів для ЗНО.

3.3. Процес донавчання моделі чат-боту на базі GPT2-uk

Для донавчання моделі чат-боту на базі GPT2-uk на україномовному текстовому корпусі у середовищі Google Colab був реалізований наступний код на мові Python з використанням бібліотеки `transformers`:

```
from transformers import TextDataset
from transformers import DataCollatorForLanguageModeling
from transformers import GPT2Tokenizer, GPT2LMHeadModel
from transformers import Trainer, TrainingArguments
```

Спочатку імпортуються необхідні класи з бібліотеки `transformers`

Україномовні моделі GPT-2 на Hugging Face (станом на 11.06.2024).

Назва	Версія	Набір даних
Tereveni-AI/gpt2-124M-uk-fiction	GPT2	Model was trained on corpus of 4040 fiction books, 2.77 GiB in total
kyryl0s/gpt2-uk-xxs	GPT2	GPT2 being trained on Ukrainian news.
benjamin/gpt2-large-wechsel-ukrainian	GPT2-Large	gpt2-large transferred to Ukrainian using the method from the NAACL2022 paper WECHSEL: Effective initialization of subword embeddings for cross-lingual transfer of monolingual language models.
kyryl0s/gpt2-uk-zno-edition	GPT2?	GPT2 trained to generate ЗНО (Ukrainian exam SAT type of thing) essays
malteos/gpt2-uk	GPT2	A generative language model for the Ukrainian language follows the GPT-2 architecture (124M parameters). Training data: OSCAR, Wikimedia dumps
nvidia/GPT-2B-001	GPT2-XL	The model was trained on 1.1T tokens obtained from publicly available data sources. The dataset comprises 53 languages and code.

для роботи з текстовими даними, розбивачем на лексеми (токенізатором) та моделлю GPT2, а також для налаштування процесу навчання.

Далі були визначені допоміжні функції:

```
def load_dataset(file_path, tokenizer, block_size = 128):
    dataset = TextDataset(
        tokenizer = tokenizer,
        file_path = file_path,
        block_size = block_size,
    )
    return dataset
```

```

def load_data_collator(tokenizer, mlm = False):
    data_collator = DataCollatorForLanguageModeling(
        tokenizer=tokenizer,
        mlm=mlm,
    )
    return data_collator

```

Функція `load_dataset` завантажує текстовий набір даних, виконуючи його токенизацію із заданим розміром блоку `block_size`. Функція `load_data_collator` створює об'єкт, що підготує пакети даних для навчання моделі (з можливістю маскованого моделювання мови).

Основна функція `train` виконує безпосереднє донавчання моделі:

```

def train(train_file_path,model_name,
          output_dir,
          overwrite_output_dir,
          per_device_train_batch_size,
          num_train_epochs,
          save_steps, resume_from_checkpoint):
    tokenizer = GPT2Tokenizer.from_pretrained(model_name)
    train_dataset = load_dataset(train_file_path, tokenizer)
    data_collator = load_data_collator(tokenizer)

    tokenizer.save_pretrained(output_dir)

    model = GPT2LMHeadModel.from_pretrained(model_name)

    model.save_pretrained(output_dir)

    training_args = TrainingArguments(
        output_dir=output_dir,
        overwrite_output_dir=overwrite_output_dir,
        per_device_train_batch_size=per_device_train_batch_size,
        num_train_epochs=num_train_epochs,
    )

```

```

trainer = Trainer(
    model=model,
    args=training_args,
    data_collator=data_collator,
    train_dataset=train_dataset,
)

trainer.train(resume_from_checkpoint=resume_from_checkpoint)
trainer.save_model()

```

Ця функція виконує наступні кроки:

1. Завантаження та збереження токенизатора і моделі, ініціалізованої з вказаної точки `model_name`.
2. Підготовка набору даних для навчання та об'єкту `data_collator`.
3. Визначення параметрів навчання `training_args`: каталогу для збереження моделі, розмір пакету, кількість ітерацій.
4. Створення об'єкту `trainer` з моделлю, параметрами навчання та навчальними даними.
5. Запуск навчання моделі методом `train()` з можливістю продовження з контрольної точки.
6. Збереження донавченої моделі.

Параметри донавчання задаються наступними змінними:

```

#оригінальна модель та точки збереження
train_directory =
    "/content/drive/MyDrive/SemerikovProject/q_and_a"
train_file_path =
    "/content/drive/MyDrive/SemerikovProject/journal.txt"
model_name = train_directory
#донавчена модель та точки збереження

```

```
output_dir =
    "/content/drive/MyDrive/SemerikovProject/custom_full_text"
overwrite_output_dir = False
per_device_train_batch_size = 8
num_train_epochs = 50
save_steps = 50000
```

Змінна `train_file_path` вказує шлях до текстового файлу з корпусом для донавчання, `model_name` – каталог із файлами початкової моделі (або контрольною точкою продовження донавчання), `output_dir` – вихідний каталог для збереження донавчених файлів моделі та контрольних точок.

Початкова модель `malteos/gpt2-uk` завантажується за допомогою методів бібліотеки `transformers` з хабу Hugging Face:

```
from transformers import AutoTokenizer, AutoModelForCausalLM

tokenizer = AutoTokenizer.from_pretrained("malteos/gpt2-uk")
model = AutoModelForCausalLM.from_pretrained("malteos/gpt2-uk")
tokenizer.save_pretrained(train_directory)
model.save_pretrained(train_directory)
```

Завантажені файли моделі зберігаються у вказану директорію `train_directory`.

Запуск процесу донавчання виконується викликом функції `train` із вказаними параметрами:

```
train(train_file_path=train_file_path,
      model_name=model_name,
      output_dir=output_dir,
      overwrite_output_dir=overwrite_output_dir,
      per_device_train_batch_size=per_device_train_batch_size,
      num_train_epochs=num_train_epochs,
      save_steps=save_steps,
      # False для першого разу,
      # True - для відновлення з точки зупинки
      resume_from_checkpoint=False)
```


Тривалість навчання може бути зменшена шляхом використання GPU на відділеному сервері або локальному комп'ютері. Останнє потребувало встановлення набору драйверів CUDA 12.5 для підтримки відеокарти GeForce RTX 3080 (10 Gb) та спеціалізованої версії PyTorch, а також оновлення бібліотек accelerate та transformers.

Для фінального експерименту загальний час навчання склав 37 годин (середня тривалість однієї ітерації навчання – 43 хвилини).

Для тестування донавченої на наборі даних моделі було створено інтерфейс користувача з використанням бібліотеки gradio (рис. 3.1):

```
from transformers import pipeline
import gradio as gr

model = pipeline("text-generation",
                 model="/content/drive/MyDrive/SemerikovProject/lastversion")

def predict(prompt):
    completion = model(prompt, max_length=50)[0]["generated_text"]
```

The image shows a web-based chatbot interface. At the top, there is a text input field labeled 'prompt' containing the word 'Рашевська'. Below this input field are two buttons: a grey 'Clear' button and an orange 'Submit' button. Below the buttons is an output area labeled 'output' which displays the generated text: 'Рашевська, С.О. Семеріков, Ю.В. Триус, А.М. Стрюк та ін. Питання оцінювання результатів навчання з використанням хмарних сервісів висвітлені'.

Рис. 3.1. Інтерфейс прототипу чат-боту для моделі GPT-2, донавченої на наборі із публікацій журналу “Інформаційні технології та засоби навчання”.

```
return completion
```

```
gr.Interface(fn=predict, inputs="text", outputs="text").launch()
```

Висновки до 3 розділу

1. Сформовано два набори даних для навчання чат-ботів: набір із публікацій “CEUR Workshop Proceedings” (широкий спектр досліджень у галузі комп’ютерних наук англійською мовою) та набір із публікацій журналу “Інформаційні технології та засоби навчання” (предметно-орієнтований корпус переважно українською мовою). Створені набори даних відрізняються за обсягом, часовим охопленням, тематичною спрямованістю та розподілом мов написання текстів, що забезпечує умови для порівняльного аналізу ефективності моделей чат-ботів, навчених на різномірних текстових корпусах.
2. Обґрунтовано вибір моделей для донавчання: базової багатомовної моделі GPT2-XL для першого набору даних та попередньо донавченої на україномовних текстах моделі gpt2-uk для другого. Вибір моделей базувався на співставленні характеристик наборів текстових даних (обсягу, мовного складу) та доступних моделей сімейства GPT з урахуванням їх розміру та попереднього досвіду навчання.
3. Реалізовано процес донавчання обраних моделей на сформованих текстових корпусах з використанням можливостей бібліотеки **transformers** від Hugging Face. Розроблений програмний код дозволяє виконувати завантаження попередньо навчених моделей, їх донавчання на текстових даних, що надаються користувачем, та збереження донавчених моделей для подальшого використання.
4. Протестовано роботу донавчених моделей для генерації відповідей чат-боту на тематично споріднені запити користувача. Для зручності взаємодії користувача з чат-ботом розроблено графічний інтерфейс з використанням бібліотеки **gradio**.

ВИСНОВКИ

1. Бібліометричний аналіз 549 джерел із бази Scopus з проблеми навчання чат-ботів надав можливість:
 - а) виявити нижню хронологічну межу (2018 рік), починаючи з якої спостерігається стійке зростання кількості публікацій із навчання чат-ботів;
 - б) згрупувати у 4 кластери авторські та індексні ключові слова джерел із навчання чат-ботів: 1) опрацювання природної мови; 2) застосування технологій опрацювання природної мови у суспільстві; 3) застосування машинного навчання для опрацювання природної мови; 4) чат-боти у навчанні та сфері послуг.
2. Визначено основні моделі навчання чат-ботів: навчання з учителем на основі рекурентних нейронних мереж та трансформерів; навчання з підкріпленням, включаючи використання генеративно-змагальних мереж; трансферне навчання шляхом тонкого налаштування попередньо навчених великих мовних моделей. Для оцінки ефективності навчання чат-ботів використовуються як автоматичні метрики (BERTScore, перплексивність, BLEU, ROUGE), так і методи людської оцінки, які дозволяють врахувати релевантність, зв'язність та природність згенерованих відповідей.
3. Сформовано два набори даних для навчання чат-ботів: набір із публікацій “CEUR Workshop Proceedings” та набір із публікацій журналу “Інформаційні технології та засоби навчання”. Обґрунтовано вибір базових моделей GPT2-XL та gpt2-uk для донавчання на створених текстових корпусах. Реалізовано процес донавчання обраних моделей на сформованих текстових корпусах та протестовано роботу донавчених моделей для генерації відповідей чат-боту на тематично споріднені запити користувача. Для зручності взаємодії розроблено графічний інтерфейс користувача на основі бібліотеки gradio.

Отримані результати можуть бути використані для подальших досліджень у напрямку створення ефективних та спеціалізованих чат-ботів із застосуванням сучасних підходів до навчання великих мовних моделей.

СПИСОК ВИКОРИСТАНИХ ДЖЕРЕЛ

1. “I wouldn’t say offensive but...”: Disability-Centered Perspectives on Large Language Models / V. Gadiraju [та ін.] // ACM International Conference Proceeding Series. — Association for Computing Machinery, 2023. — С. 205—216. — DOI: 10.1145/3593013.3593989.
2. A chatbot proposal for tele orientation on breastfeeding; [Uma proposta de chatbot para tele orientação sobre aleitamento materno] / J. B. dos Santos Junior [та ін.] // RISTI - Revista Iberica de Sistemas e Tecnologias de Informacao. — 2021. — Т. 2021, E42. — С. 357—363.
3. A virtual counseling application using artificial intelligence for communication skills training in nursing education: Development study / S. Shorey [та ін.] // Journal of Medical Internet Research. — 2019. — Т. 21, № 10. — DOI: 10.2196/14658.
4. *Alaswad S., Kalganova T., Awad W.* Using ChatGPT and other LLMs in Professional Environments // Information Sciences Letters. — 2023. — Т. 12, № 9. — С. 2097—2108. — DOI: 10.18576/isl/120916.
5. An overview of the features of chatbots in mental health: A scoping review / A. A. Abd-alrazaq [та ін.] // International Journal of Medical Informatics. — 2019. — Т. 132. — DOI: 10.1016/j.ijmedinf.2019.103978.
6. Artificial intelligence chatbot performance in triage of ophthalmic conditions / R. J. Lyons [та ін.] // Canadian Journal of Ophthalmology. — 2023. — DOI: 10.1016/j.jcjo.2023.07.016.
7. Artificial Neural Network Based University Chatbot System / N. Bhartiya [та ін.] // 2019 IEEE Bombay Section Signature Conference, IBSSC 2019. 2019January. — Institute of Electrical, Electronics Engineers Inc., 2019. — DOI: 10.1109/IBSSC47189.2019.8973095.
8. Attention Is All You Need / A. Vaswani [та ін.]. — 2023. — arXiv: 1706.03762 [cs.CL].
9. Attitude detection for one-round conversation: Jointly extracting target-polarity pairs / Z. Zeng [та ін.] // Journal of Information Processing. — 2019. — Т. 27. — С. 742—751. — DOI: 10.2197/IPSJJIP.27.742.

10. Big Bot Makes Small Talk: A research summary of Facebook's Generative BST chatbot. — 05.2020. — URL: <https://www.deeplearning.ai/the-batch/big-bot-makes-small-talk/>.
11. BilinBot: A Bilingual Chatbot using Deep Learning / S. S. Jennifer [та ін.] // HORA 2023 - 2023 5th International Congress on Human-Computer Interaction, Optimization and Robotic Applications, Proceedings. — Institute of Electrical, Electronics Engineers Inc., 2023. — DOI: 10.1109/HORA58378.2023.10156681.
12. Bot Comic: How Google's Meena chatbot developed a sense of humor. — 02.2020. — URL: <https://www.deeplearning.ai/the-batch/bot-comic/>.
13. *Bozzon A.* Enterprise crowd computing for human aided chatbots // Proceedings - International Conference on Software Engineering. — IEEE Computer Society, 2018. — C. 29—30. — DOI: 10.1145/3195555.3195566.
14. Can large language models write reflectively / Y. Li [та ін.] // Computers and Education: Artificial Intelligence. — 2023. — Т. 4. — DOI: 10.1016/j.caeai.2023.100140.
15. *Centre for Science and Technology Studies, Leiden University, The Netherlands.* VOSviewer - Visualizing scientific landscapes. — 2023. — URL: <https://www.vosviewer.com/>.
16. Chatbot design method using hybrid word vector expression model based on real telemarketing data / J. Zhang [та ін.] // KSII Transactions on Internet and Information Systems. — 2020. — Т. 14, № 4. — C. 1400—1418. — DOI: 10.3837/TIIS.2020.04.001.
17. Chatbot Integrated with Machine Learning Deployed in the Cloud and Performance Evaluation / G. R. Gunnam [та ін.] // IS and T International Symposium on Electronic Imaging Science and Technology. Т. 34. — Society for Imaging Science, Technology, 2022. — DOI: 10.2352/EI.2022.34.3.MOBMU-205.

18. Chatbots Disagree on Covid-19: Medical chatbots offered conflicting Covid advice. — 04.2020. — URL: <https://www.deeplearning.ai/the-batch/chatbots-disagree-on-covid-19/>.
19. Chatbots for Productivity: Microsoft extends Copilot to 365 and Windows. — 09.2023. — URL: <https://www.deeplearning.ai/the-batch/microsoft-extends-copilot-365-windows/>.
20. China Chases Chatbots: Chinese tech companies race to cash in on ChatGPT fever. — 03.2023. — URL: <https://www.deeplearning.ai/the-batch/chinese-tech-companies-race-to-cash-in-on-chatgpt-fever/>.
21. *Chou T.-L., Hsueh Y.-L.* A Task-oriented Chatbot Based on LSTM and Reinforcement Learning // Proceedings of the 2019 3rd International Conference on Natural Language Processing and Information Retrieval. — Tokushima, Japan : Association for Computing Machinery, 2019. — C. 87—91. — (NLP/IR '19). — DOI: 10.1145/3342827.3342844.
22. Comparative Performance of ChatGPT and Bard in a Text-Based Radiology Knowledge Assessment / N. S. Patil [та ін.] // Canadian Association of Radiologists Journal. — 2023. — DOI: 10.1177/08465371231193716.
23. Cost Containment for Generative AI: Microsoft's quest to reduce the size and cost of language models. — 10.2023. — URL: <https://www.deeplearning.ai/the-batch/microsofts-quest-to-reduce-the-size-and-cost-of-language-models/>.
24. CS&SE@SW 2023: 6th Workshop for Young Scientists in Computer Science & Software Engineering. — 02.2024. — URL: <https://cssesw.ccjournals.eu/cssesw2023/>.
25. *Day M.-Y., Shaw S.-R.* AI Customer Service System with Pre-trained Language and Response Ranking Models for University Admissions // Proceedings - 2021 IEEE 22nd International Conference on Information Reuse and Integration for Data Science, IRI 2021. — Institute of Electrical, Electronics Engineers Inc., 2021. — C. 395—401. — DOI: 10.1109/IRI51335.2021.00062.

26. Deep reinforcement learning from human preferences / P. Christiano [та ил.]. — 2023. — arXiv: 1706.03741 [stat.ML].
27. *DeepLearning.AI*. Search | The Batch | AI News & Insights. — 12.2023. — URL: <https://www.deeplearning.ai/search/>.
28. Detecting Entities of Works for Chinese Chatbot / C. Wu [та ил.] // ACM Transactions on Asian and Low-Resource Language Information Processing. — 2020. — Т. 19, № 6. — DOI: 10.1145/3414901.
29. *Dharrao D., Gite S.* TherapyBot: a chatbot for mental well-being using transformers // International Journal of Advances in Applied Sciences. — 2024. — Бep. — Т. 13, № 1. — С. 1–12. — ISSN 2252-8814. — DOI: 10.11591/ijaas.v13.i1.pp1-12.
30. *Doshi J.* Chatbot User Interface for Customer Relationship Management using NLP models // Proceedings - 2021 1st IEEE International Conference on Artificial Intelligence and Machine Vision, AIMV 2021 / за ред. P. S., B. S.K., G. R.K. — Institute of Electrical, Electronics Engineers Inc., 2021. — DOI: 10.1109/AIMV53313.2021.9670914.
31. *Elaprolu S.* Introducing the AWS Generative AI Innovation Center’s Custom Model Program for Anthropic Claude. — 11.2023. — URL: <https://aws.amazon.com/blogs/machine-learning/introducing-the-aws-generative-ai-innovation-centers-custom-model-program-for-anthropic-claude/>.
32. *Elsevier B.V.* Scopus - Document search | Signed in. — 2023. — URL: <https://www.scopus.com/search/form.uri?display=basic#basic>.
33. Goal-Oriented Chatbot Dialog Management Bootstrapping with Transfer Learning / V. Ilievski [та ил.] // Proceedings of the Twenty-Seventh International Joint Conference on Artificial Intelligence, IJCAI-18. — International Joint Conferences on Artificial Intelligence Organization, 07.2018. — С. 4115–4121. — DOI: 10.24963/ijcai.2018/572.
34. *Gonda D. E., Chu B.* Chatbot as a learning resource? Creating conversational bots as a supplement for teaching assistant training course // TALE 2019 - 2019 IEEE International Conference on Engineering, Technology

- and Education. — Institute of Electrical, Electronics Engineers Inc., 2019. — DOI: 10.1109/TALE48000.2019.9225974.
35. *Google*. Google AI Studio. — 2024. — URL: <https://aistudio.google.com/>.
 36. *Google*. Available regions for Google AI Studio and Gemini API. — 2024. — URL: <https://ai.google.dev/gemini-api/docs/available-regions>.
 37. Healthcare ex Machina: Are conversational agents ready for prime time in oncology? / J.-E. Bibault [та ил.] // *Clinical and Translational Radiation Oncology*. — 2019. — Т. 16. — С. 55—59. — DOI: 10.1016/j.ctro.2019.04.002.
 38. *Henrickson L*. Chatting with the dead: The hermeneutics of thanabots // *Media, Culture and Society*. — 2023. — Т. 45, № 5. — С. 949—966. — DOI: 10.1177/01634437221147626.
 39. Implementation of an Undergraduate Admission Chatbot Using Microsoft Azure’s Question Answering and Bot Framework / A. G. Usigan [та ил.] // *ACM International Conference Proceeding Series*. — Association for Computing Machinery, 2022. — С. 240—245. — DOI: 10.1145/3582099.3582135.
 40. Improving Matching Models with Hierarchical Contextualized Representations for Multi-turn Response Selection / C. Tao [та ил.] // *SIGIR 2020 - Proceedings of the 43rd International ACM SIGIR Conference on Research and Development in Information Retrieval*. — Association for Computing Machinery, Inc, 2020. — С. 1865—1868. — DOI: 10.1145/3397271.3401290.
 41. Insights and Reflections of the Impact of ChatGPT on Intelligent Knowledge Services in Libraries / R. Zhao [та ил.] // *Journal of Library and Information Science in Agriculture*. — 2023. — Т. 35, № 1. — С. 29—38. — DOI: 10.13998/j.cnki.issn1002-1248.23-0116.
 42. Intelligent Chatbot Adapted from Question and Answer System Using RNN-LSTM Model / P. Anki [та ил.] // *Journal of Physics: Conference*

- Series. — 2021. — Бep. — Т. 1844, № 1. — С. 012001. — DOI: 10.1088/1742-6596/1844/1/012001.
43. Investigating ChatGPT and cybersecurity: A perspective on topic modeling and sentiment analysis / O. D. Okey [та иH.] // Computers and Security. — 2023. — Т. 135. — DOI: 10.1016/j.cose.2023.103476.
 44. Ji Z. A Multi-modal Seq2seq Chatbot Framework // Proceeding of 2021 International Conference on Wireless Communications, Networking and Applications / за ред. Z. Qian, M. Jabbar, X. Li. — Singapore : Springer Nature, 2022. — С. 225–233. — DOI: 10.1007/978-981-19-2456-9_24.
 45. Kansal A. Finetuning: The Theory // Building Generative AI-Powered Apps: A Hands-on Guide for Developers. — Berkeley, CA : Apress, 2024. — С. 77–100. — DOI: 10.1007/979-8-8688-0205-8_5.
 46. Kim B., Seo J., Koo M.-W. Randomly Wired Network Based on RoBERTa and Dialog History Attention for Response Selection // IEEE/ACM Transactions on Audio Speech and Language Processing. — 2021. — Т. 29. — С. 2437–2442. — DOI: 10.1109/TASLP.2021.3077119.
 47. Language Models are Unsupervised Multitask Learners / A. Radford [та иH.]. — 02.2019. — URL: https://cdn.openai.com/better-language-models/language_models_are_unsupervised_multitask_learners.pdf.
 48. Language Models, Extended: Large language models grew more reliable and less biased in 2022. — 12.2022. — URL: <https://www.deeplearning.ai/the-batch/language-models-grew-more-reliable-and-less-biased-in-2022/>.
 49. Large language models in health care: Development, applications, and challenges / R. Yang [та иH.] // Health Care Science. — 2023. — Т. 2, № 4. — С. 255–263. — DOI: 10.1002/hcs2.61.
 50. Learning from dialogue after deployment: Feed yourself, Chatbot! / B. Hancock [та иH.] // ACL 2019 - 57th Annual Meeting of the Association for Computational Linguistics, Proceedings of the Conference. — Association for Computational Linguistics (ACL), 2020. — С. 3667–3684.

51. *Lee M.* A Mathematical Investigation of Hallucination and Creativity in GPT Models // *Mathematics*. — 2023. — T. 11, № 10. — C. 2320. — DOI: 10.3390/math11102320.
52. *Lee P., Bubeck S., Petro J.* Benefits, Limits, and Risks of GPT-4 as an AI Chatbot for Medicine. // *New England Journal of Medicine*. — 2023. — T. 388, № 13. — C. 1233—1239. — DOI: 10.1056/NEJMsr2214184.
53. *Lee S.-W., Choi W.-J.* Utilizing ChatGPT in clinical research related to anesthesiology: a comprehensive review of opportunities and limitations // *Anesthesia and Pain Medicine*. — 2023. — T. 18, № 3. — C. 244—251. — DOI: 10.17085/apm.23056.
54. *Lin L., D'Haro L. F., Banchs R. E.* A web-based platform for collection of human-chatbot interactions // *HAI 2016 - Proceedings of the 4th International Conference on Human Agent Interaction*. — Association for Computing Machinery, Inc, 2016. — C. 363—366. — DOI: 10.1145/2974804.2980500.
55. *Lin Y., Yu Z.* A bibliometric analysis of artificial intelligence chatbots in educational contexts // *Interactive Technology and Smart Education*. — 2023. — DOI: 10.1108/ITSE-12-2022-0165.
56. *Lo C. K.* What Is the Impact of ChatGPT on Education? A Rapid Review of the Literature // *Education Sciences*. — 2023. — T. 13, № 4. — DOI: 10.3390/educsci13040410.
57. *Lum Z. C.* Can Artificial Intelligence Pass the American Board of Orthopaedic Surgery Examination? Orthopaedic Residents Versus ChatGPT // *Clinical Orthopaedics and Related Research*. — 2023. — T. 481, № 8. — C. 1623—1630. — DOI: 10.1097/CORR.0000000000002704.
58. *Luthier G., Popescu-Belis A.* Chat or learn: A Data-driven robust question-answering system // *LREC 2020 - 12th International Conference on Language Resources and Evaluation, Conference Proceedings*. — European Language Resources Association (ELRA), 2020. — C. 5474—5480.

59. Measuring and Mitigating Bias in AI-Chatbots / H. Beattie [та ил.] // Proceeding - 2022 IEEE International Conference on Assured Autonomy, ICAA 2022. — Institute of Electrical, Electronics Engineers Inc., 2022. — C. 117—123. — DOI: 10.1109/ICAA52185.2022.00023.
60. *Nielsen J. P. S., Buchwald C. von, Grønhøj C.* Validity of the large language model ChatGPT (GPT4) as a patient information source in otolaryngology by a variety of doctors in a tertiary otorhinolaryngology department // *Acta Oto-Laryngologica*. — 2023. — T. 143, № 9. — C. 779—782. — DOI: 10.1080/00016489.2023.2254809.
61. *Ohlheiser A. W., Hao K.* An AI is training counselors to deal with teens in crisis. — 02.2021. — URL: <https://www.technologyreview.com/2021/02/26/1020010/trevor-project-ai-suicide-hotline-training/>.
62. Open domain chatbot based on attentive end-to-end Seq2Seq mechanism / S. S. Abdullahi [та ил.] // ACM International Conference Proceeding Series. — Association for Computing Machinery, 2019. — C. 339—344. — DOI: 10.1145/3377713.3377773.
63. *OpenAI.* Introducing ChatGPT. — 11.2022. — URL: <https://openai.com/blog/chatgpt>.
64. *OpenAI.* Fine-tuning - OpenAI API. — 2024. — URL: <https://platform.openai.com/finetune>.
65. *Pack A., Maloney J.* Using Generative Artificial Intelligence for Language Education Research: Insights from Using OpenAI’s ChatGPT // *TESOL Quarterly*. — 2023. — T. 57, № 4. — C. 1571—1582. — DOI: 10.1002/tesq.3253.
66. *Park I., Joshi A. S., Javan R.* Potential role of ChatGPT in clinical otolaryngology explained by ChatGPT // *American Journal of Otolaryngology - Head and Neck Medicine and Surgery*. — 2023. — T. 44, № 4. — DOI: 10.1016/j.amjoto.2023.103873.
67. *Patil S., Mudaliar V., Kamat P.* LSTM based Ensemble Network to enhance the learning of Long-term Dependencies in Chatbot // *International Journal of Automation and Smart Technology*. — 2022. — Січ. — T. 12, № 1. — C. 2286—2286. — DOI: 10.5875/ausmt.v12i1.2286.

68. *Peyton K., Unnikrishnan S.* A comparison of chatbot platforms with the state-of-the-art sentence BERT for answering online student FAQs // Results in Engineering. — 2023. — T. 17. — DOI: 10.1016/j.rineng.2022.100856.
69. *Povinsky M., Melicherik M., Siladi V.* A Chatbot based on Deep Neural Network and Public Cloud Services with TJBOT Interface // INFORMATICS 2019 - IEEE 15th International Scientific Conference on Informatics, Proceedings. — Institute of Electrical, Electronics Engineers Inc., 2019. — C. 101–106. — DOI: 10.1109/Informatics47936.2019.9119304.
70. *Prasomphan S.* Using chatbot in trading system for Small and Medium Enterprise (SMEs) by convolution neural network technique // ACM International Conference Proceeding Series. — 2019. — C. 93–98. — DOI: 10.1145/3341069.3341092.
71. *Raundale P., Sawale A.* Dialog prediction in institute admission: A deep learning way // 2021 2nd International Conference for Emerging Technology, INCET 2021. — Institute of Electrical, Electronics Engineers Inc., 2021. — DOI: 10.1109/INCET51464.2021.9456159.
72. *Ray A., Mathew R.* Review of Cloud-Based Natural Language Processing Services and Tools for Chatbots // Lecture Notes on Data Engineering and Communications Technologies. — 2020. — T. 31. — C. 156–162. — DOI: 10.1007/978-3-030-24643-3_18.
73. Recipes for building an open-domain chatbot / S. Roller [та ін.] // EACL 2021 - 16th Conference of the European Chapter of the Association for Computational Linguistics, Proceedings of the Conference. — Association for Computational Linguistics (ACL), 2021. — C. 300–325.
74. *Rodriguez S., Mune C.* Uncoding library chatbots: deploying a new virtual reference tool at the San Jose State University library // Reference Services Review. — 2022. — T. 50, № 3/4. — C. 392–405. — DOI: 10.1108/RSR-05-2022-0020.
75. *Ruane E., Young R., Ventresque A.* Training a chatbot with microsoft LUIS: Effect of intent imbalance on prediction accuracy // International Conference on Intelligent User Interfaces, Proceedings IUI. — Association

- for Computing Machinery, 2020. — C. 63–64. — DOI: 10.1145/3379336.3381494.
76. *Salas-Pilco S. Z., Xiao K., Hu X.* Correction to: Artificial Intelligence and Learning Analytics in Teacher Education: A Systematic Review (Education Sciences, (2022), 12, 8, (569), 10.3390/educsci12080569) // Education Sciences. — 2023. — T. 13, № 9. — DOI: 10.3390/educsci13090897.
77. *Sarma G., Kashyap H., Medhi P. P.* ChatGPT in Head and Neck Oncology-Opportunities and Challenges // Indian Journal of Otolaryngology and Head and Neck Surgery. — 2023. — DOI: 10.1007/s12070-023-04201-6.
78. Search War! Google and Microsoft both announce AI-Powered search. — 02.2023. — URL: <https://www.deeplearning.ai/the-batch/google-and-microsoft-both-announce-ai-powered-search/>.
79. *Singh S., Djalilian A., Ali M. J.* ChatGPT and Ophthalmology: Exploring Its Potential with Discharge Summaries and Operative Notes // Seminars in Ophthalmology. — 2023. — T. 38, № 5. — C. 503–507. — DOI: 10.1080/08820538.2023.2209166.
80. Smart and intelligent chatbot assistance for future industry 4.0 / H. Khatter [та ит.]. — Springer International Publishing, 2021. — C. 153–168. — DOI: 10.1007/978-3-030-77070-9_10.
81. *Sreedhar M. N., Ni K., Reddy S.* Learning improvised chatbots from adversarial modifications of natural language feedback // Findings of the Association for Computational Linguistics Findings of ACL: EMNLP 2020. — 2020. — C. 2445–2453.
82. Study on emotion recognition and companion Chatbot using deep neural network / M.-C. Lee [та ит.] // Multimedia Tools and Applications. — 2020. — T. 79, № 27/28. — C. 19629–19657. — DOI: 10.1007/s11042-020-08841-6.
83. *Syzdykbayeva A., Baikulova A., Kerimbayeva R.* Introduction of Artificial Intelligence as the Basis of Modern Online Education on the Example of Higher Education // SIST 2021 - 2021 IEEE International Conference on Smart Information Systems and Technologies. — Institute of Electrical,

- Electronics Engineers Inc., 2021. — DOI: 10.1109/SIST50301.2021.9465974.
84. *Taulli T.* AI-Assisted Programming: Better Planning, Coding, Testing, and Deployment. — Sebastopol, CA : O'Reilly Media, Inc., 2024. — URL: <https://www.oreilly.com/library/view/ai-assisted-programming/9781098164553/>.
85. The neural network conversation model enables the commonly asked student query agents / N. Muangnak [та ил.] // International Journal of Advanced Computer Science and Applications. — 2020. — T. 11, № 4. — С. 154—164. — DOI: 10.14569/IJACSA.2020.0110421.
86. The promise and peril of using a large language model to obtain clinical information: ChatGPT performs strongly as a fertility counseling tool with limitations / J. Chervenak [та ил.] // Fertility and Sterility. — 2023. — T. 120, № 3. — С. 575—583. — DOI: 10.1016/j.fertnstert.2023.05.151.
87. Towards coherent and engaging spoken dialog response generation using automatic conversation evaluators / S. Yi [та ил.] // INLG 2019 - 12th International Conference on Natural Language Generation, Proceedings of the Conference. — Association for Computational Linguistics (ACL), 2019. — С. 65—75.
88. *Tran Q.-D. L., Le A.-C., Huynh V.-N.* Enhancing Conversational Model With Deep Reinforcement Learning and Adversarial Learning // IEEE Access. — 2023. — T. 11. — С. 75955—75970. — DOI: 10.1109/ACCESS.2023.3297652.
89. Unsupervised cross-domain adaptation for response selection using self-supervised and adversarial training / J. Li [та ил.] // WSDM 2022 - Proceedings of the 15th ACM International Conference on Web Search and Data Mining. — Association for Computing Machinery, Inc, 2022. — С. 562—570. — DOI: 10.1145/3488560.3498404.
90. *Uprety S. P., Jeong S. R.* The Impact of Semi-Supervised Learning on the Performance of Intelligent Chatbot System // Computers, Materials & Continua. — 2022. — T. 71, № 2. — С. 3937—3952. — DOI: 10.32604/cmc.2022.023127.

91. *Van Eck N. J., Waltman L.* VOSviewer Manual. — Universiteit Leiden, 2023. — URL: https://www.vosviewer.com/documentation/Manual_VOSviewer_1.6.20.pdf.
92. What We Know — and Don't Know — About Foundation Models: A new Stanford index to assess the transparency of leading AI models. — 11.2023. — URL: <https://www.deeplearning.ai/the-batch/a-new-stanford-index-to-assess-the-transparency-of-leading-ai-models/>.
93. *Yoo J., Cho Y.* ICSA: Intelligent chatbot security assistant using Text-CNN and multi-phase real-time defense against SNS phishing attacks // Expert Systems with Applications. — 2022. — Т. 207. — DOI: 10.1016/j.eswa.2022.117893.
94. *Ляшенко Р. О., Семеріков С. О.* Бібліометричний аналіз навчання чат-ботів // Тези доповідей VII Міжнародної науково-практичної конференції «Інформаційні технології в освіті, науці і техніці» (ІТОНТ-2024), (Черкаси, 23-24 травня 2024 р.) — Черкаси : ЧДТУ, 2024. — С. 314—316. — URL: <https://itest.chdtu.edu.ua/#materials>.
95. *Ляшенко Р. О., Семеріков С. О.* Бібліометричний аналіз досліджень з навчання чат-ботів: ключові поняття та тенденції // Інформаційні технології і засоби навчання. — 2024. — Т. 101, № 3. — (у процесі публікації).

Додаток А

Результати застосування генеративних чатботів для визначення імен кластерів

Запит був створений англійською мовою 30.11.2023. У запиті містились відомості про кількість кластерів та ключові слова, що відносились до кожного кластеру. Від чатботу очікувались пропозиції щодо імен та змісту кластерів. Запит мав наступний вигляд:

“There are 4 clusters of keywords related to chatbot training research:

Cluster 1: classification (of information), training data, speech processing, performance, semantics, dialogue systems, natural language understanding, computational linguistics, text processing, long short-term memory, state of the art, language model, question answering, information retrieval, embeddings, reinforcement learning, response generation

Cluster 2: artificial intelligence, conversational agent, human, adult, training, male, medical education, female, controlled study, health care, mental health, chatgpt, education, covid-19, review, large language model

Cluster 3: natural language processing, natural language processing systems, machine learning, deep learning, natural languages, language processing, learning algorithms, virtual assistant, user interfaces, convolutional neural network, query processing, diagnosis

Cluster 4: chatbot, learning systems, students, e-learning, learn+, sales, human computer interaction, knowledge based systems, personnel training, curricula, customer service, engineering education

Please suggest cluster names and brief descriptions for each of the 4 clusters based on the keywords.”

ChatGPT (GPT 3.5)

1. *Advanced Language Processing and AI Applications*: This cluster focuses on advanced language processing techniques and various AI applications related to text and speech. It encompasses topics like natural language understanding, computational linguistics, information retrieval, dialogue

systems, and the utilization of cutting-edge technologies such as long short-term memory (LSTM), embeddings, and reinforcement learning in language models and question answering systems.

2. *AI in Healthcare and Social Context*: This cluster centers around the applications of artificial intelligence in healthcare, mental health, and education with a focus on human interaction. It explores the use of conversational agents, particularly in medical education, mental health care, and addressing issues like COVID-19. Additionally, it touches on the gender aspect in AI studies by mentioning controlled studies with adults and gender variations.
3. *Core Technologies in Natural Language Processing and Machine Learning*: This cluster revolves around foundational technologies crucial in natural language processing and machine learning. It includes topics such as deep learning, machine learning algorithms, language processing, neural networks (like convolutional neural networks), query processing, and the development of virtual assistants and user interfaces.
4. *Educational and Service-oriented AI Applications*: This cluster focuses on AI applications in education, learning systems, and service-oriented domains. It highlights topics like e-learning, knowledge-based systems, human-computer interaction, customer service, and sales. It also touches upon the use of AI in student learning, engineering education, and personnel training, emphasizing knowledge acquisition and application in various educational and service-based contexts.

Bing, precise mode (GPT 4.0)

1. *Language Processing and Information Retrieval*: This cluster focuses on the technical aspects of natural language understanding and computational linguistics. It includes topics such as speech processing, semantics, dialogue systems, text processing, language modeling, question answering, and reinforcement learning for response generation.
2. *AI in Healthcare and Education*: This cluster revolves around the application of artificial intelligence, particularly conversational agents like

ChatGPT, in fields such as medical education, healthcare, and mental health. It also touches on demographic factors (male, female, adult) and current issues like COVID-19.

3. *Machine Learning and User Interaction*: This cluster is centered on machine learning, deep learning, and natural language processing systems. It also includes user interfaces and virtual assistants, indicating a focus on how users interact with these AI systems.
4. *AI in Education and Customer Service*: This cluster pertains to the use of AI systems like chatbots in educational settings (e-learning, engineering education, curricula) and customer service. It also includes topics related to human-computer interaction and knowledge-based systems.

Bing, balanced mode (GPT 4.0)

1. *Natural Language Processing (NLP)*: This cluster includes topics related to the processing of human language by computers. It covers a wide range of subfields such as computational linguistics, text processing, information retrieval, and question answering. The cluster also includes techniques such as embeddings, long short-term memory, and reinforcement learning.:
2. *Artificial Intelligence (AI)*: This cluster includes topics related to the development of intelligent agents that can perform tasks that typically require human intelligence. It covers a wide range of subfields such as conversational agents, large language models, and chatbots. The cluster also includes topics related to health care, education, and COVID-19.:
3. *Machine Learning (ML)*: This cluster includes topics related to the development of algorithms that can learn from data and make predictions or decisions. It covers a wide range of subfields such as deep learning, learning algorithms, and convolutional neural networks. The cluster also includes topics related to virtual assistants, user interfaces, and diagnosis.:
4. *Learning Systems (LS)*: This cluster includes topics related to the development of systems that can learn from data and improve over time.

It covers a wide range of subfields such as e-learning, personnel training, and customer service. The cluster also includes topics related to sales, human-computer interaction, and engineering education.:

Bing, creative mode (GPT 4.0)

1. *Natural Language Understanding and Generation*: This cluster contains keywords related to the tasks and methods of natural language understanding and generation, such as classification, dialogue systems, question answering, and response generation. It also includes keywords that indicate the performance and state of the art of these tasks, such as semantics, long short-term memory, embeddings, and reinforcement learning.
2. *Artificial Intelligence and Health Care*: This cluster contains keywords related to the applications and implications of artificial intelligence in health care, such as conversational agent, medical education, mental health, and covid-19. It also includes keywords that indicate the characteristics and challenges of these applications, such as human, adult, male, female, controlled study, review, and large language model.
3. *Natural Language Processing and Machine Learning*: This cluster contains keywords related to the general concepts and techniques of natural language processing and machine learning, such as natural language processing systems, machine learning, deep learning, learning algorithms, and convolutional neural network. It also includes keywords that indicate the interfaces and domains of these techniques, such as virtual assistant, user interfaces, query processing, and diagnosis.
4. *Chatbot and Education*: This cluster contains keywords related to the design and evaluation of chatbot systems for education, such as chatbot, learning systems, students, e-learning, learn+, and curricula. It also includes keywords that indicate the contexts and outcomes of these systems, such as sales, human computer interaction, knowledge based systems, personnel training, customer service, and engineering education.

Claude 2

1. *Natural Language Processing Applications*: This cluster focuses on various NLP applications like classification, speech processing, dialogue systems, question answering, etc. Central themes include semantics, computational linguistics, LSTMs, performance, and building practical natural language systems.
2. *Conversational AI and Impacts*: This cluster deals with conversational AI, its use in healthcare/education, and analyzing its impacts. Key themes look at chatbots/assistants, how they interact with and impact humans, with analysis around ethics, proper training, and responsible AI.
3. *Core AI and ML Techniques*: This cluster covers fundamental AI and ML techniques like natural language processing, machine learning, deep learning, algorithms, etc. It focuses on the key technical methods and capabilities that power many AI systems.
4. *AI for Training and Customer Service*: The fourth cluster deals with leveraging AI for use cases like education/training and customer service. Central ideas look at e-learning, knowledge systems, training curriculums, human-computer interaction, and providing automation for repetitive customer service queries.

Google Bard

1. *Technical NLP*: This cluster focuses on the technical aspects of natural language processing (NLP), including classification, training data, speech processing, semantics, dialogue systems, and language models.
2. *Applications of NLP*: This cluster focuses on the applications of NLP in various domains, including conversational agents, healthcare, education, and customer service.
3. *Machine Learning and Deep Learning for NLP*: This cluster focuses on the machine learning and deep learning techniques used in NLP.

4. *Chatbots and Learning Systems*: This cluster focuses on the development and evaluation of chatbots and other learning systems.