

**ОСНОВНІ ПРИНЦИПИ АЛГОРИТМІВ ШТУЧНОГО ІНТЕЛЕКТУ, ЯКІ
ВИКОРИСТОВУЮТЬСЯ ДЛЯ СТВОРЕННЯ КОНТЕНТУ, ЗОКРЕМА ГЕНЕРАТИВНІ
МОВНІ МОДЕЛІ**
**THE FUNDAMENTAL PRINCIPLES OF
ARTIFICIAL INTELLIGENCE ALGORITHMS USED FOR CONTENT CREATION,
INCLUDING GENERATIVE LANGUAGE MODELS**

Технології сучасності постійно змінюють повсякденне життя людей. За останні десятиліття людство створило винаходи, які було важко уявити навіть найсвітлішим розумам минулого. Сучасні технології штучного інтелекту (ШІ) все більше проникають у різні сфери життя, починаючи від обробки зображень до відео, звуку і до автоматизації повсякденних завдань. Особливо відчутним є досягнення у сфері обробки природної, або ж розмовної мови (Natural Language Processing, NLP), саме завдяки якій алгоритми ШІ вчаться формувати текстів контент високої якості. Штучний інтелект може створювати контент у різних форматах: аналітичні статті, рекламні тексти, новини для різного типу ресурсів, а також використовуватися для персоналізованої комунікації та чат-ботів. Це все стало можливе завдяки створенню генеративних мовних моделей, таких як GPT, BERT, T5 та інші. Ці моделі базуються на провідних алгоритмах машинного навчання та глибоких нейронних мережах, зокрема трансформерах, і використовують значні обсяги даних для навчання й удосконалення, що дозволяє їм розуміти й відтворювати текст, максимально схожий до людського. Тобто, чим більше запитів надходить до генеративної моделі, тим кращі відповіді вона формує. Саме таким чином відбувається її навчання.

Генеративні мовні моделі – це спеціалізовані архітектури, розроблені для генерування тексту на основі навчання з великих об'ємів текстових даних. На сьогодні основою значної кількості сучасних моделей є архітектура трансформерів, яка була представлена командою Google у 2017 році [1] і стала справжньою революцією у сфері NLP. Трансформери дозволяють моделі паралельно обробляти дані, використовуючи механізм само уваги (self-attention), який дозволяє алгоритму ефективно вивчати та зберігати контекст на різних рівнях тексту. Тобто всі частини тексту будуть взаємоз'язані і логічними.

Трансформер складається з двох основних частин: енкодера та декодера. Енкодер опрацьовує вхідні дані, виділяючи найважливішу інформацію, сюди можна віднести як текстову, так і графічну, тоді як декодер генерує вихідний текст на основі отриманого контексту. Генеративні мовні моделі, такі як GPT (Generative Pre-trained Transformer), використовують лише декодерну частину, що дає змогу сконцентруватися на створенні послідовного та зв'язного тексту. Механізм само уваги дозволяє кожному слову в реченні «звертати увагу» на інші слова, що робить моделі надзвичайно чутливими до контексту. Завдяки цьому GPT-3, GPT-4 та інші генеративні моделі мають здатність писати тексти, схожі на ті, що створюють люди, при цьому розуміючи зв'язок між словами та пропозиціями.

Модель GPT, розроблена компанією OpenAI, є однією з найпопулярніших моделей для створення контенту. Перша версія системи (GPT-2) мала 1.5 мільярда параметрів, а GPT-3 вже 175 мільярдів, що дозволило відчутно підвищити якість текстів. Завдяки великому обсягу даних, на яких модель була навчена, GPT-3 може генерувати контент різного типу: від звичайних текстів до складних технічних документів [2].

Хоча BERT не є генеративною моделлю, його використовують для аналізу та розуміння тексту, що є важливим аспектом при створенні контенту. BERT може розуміти значення слів

у контексті обох сторін, тобто зліва і справа, що робить його ідеальним для завдань, де потрібна якісна семантична обробка, як-от пошукові системи. Окрім того, ця генеративна система може обробити набагато більший обсяг інформації в порівнянні з іншими [3].

T5 – це універсальна модель, яка підходить як для генерації, так і для перекладу тексту, класифікації та іншої роботи. Вона перетворює будь-яке завдання в текстове, що дозволяє значно збільшити спектр застосування, включно з маркетинговим і журналістським контентом, який має свої специфічні тематики [4].

На сьогодні ми можемо виокремити наступні переваги генеративних мовних моделей:

1. Автоматизація повсякденних завдань – тобто генерувати великі обсяги з щонайменшою витратою часу.

2. Персоналізація – можливість створювати персоналізоване наповнення для різних цільових аудиторій. Задавши потрібні алгоритми штучний інтелект сформує різно жанрові тексти.

3. Пристосування до контексту – моделі можуть формувати тексти, що враховують контекст і стиль попередніх даних.

Та попри значні переваги генеративних систем з'являються і виклики, серед яких можна виокремити наступні:

1. Етична складова – є ризик зловживання моделями для створення маніпулятивного або шкідливого контенту, який буде використовуватися в протизаконних цілях.

2. Точність і правдивість інформації – моделі можуть генерувати неправдиву або помилкову інформацію, яка невідповідає дійсності, що є серйозним викликом для інформаційної безпеки. Тобто модель може сформувати текст на запит, який абсолютно не відповідає дійсності.

3. Обмеження в розумінні культурних та емоційних аспектів – інструменти штучного інтелекту не завжди можуть зрозуміти тонкі нюанси культури та емоцій, що може призвести до створення текстів, які не відповідають очікуванням, адже моделі нездатні обробляти емоційну складову. Також генеративні моделі можуть не працювати з аналогіями, порівняннями та метафорами.

Вищеперераховані мовні моделі на основі трансформерів відкрили нові можливості для створення текстового контенту, що спростило процес формування текстів та персоналізації комунікації з користувачами. Моделі GPT, BERT, T5 та інші, завдяки своїм особливостям, дозволяють створювати тексти, що відповідають всім стандартам якості, про те їх використання потребує контролю та редагування безпосередньо людиною, особливо в питаннях правдивості та етики. Розуміння принципів їхньої роботи, обмежень і викликів, є необхідним кроком для належного використання цих технологій.

Список використаних джерел:

1. Vaswani,A., Shazeer,N., Parmar,N., Uszkoreit,J., Jones,L., Gomez,A.N., Kaiser,L., & Polosukhin ,I. (2017). Attention Is All You Need. Advances in Neural Information Processing Systems,30. URL: <https://arxiv.org/abs/1706.03762>
2. Brown,T.B.,Mann,B.,Ryder,N.,Subbiah,M.,Kaplan,J., Dhariwal,P.,Neelakantan,A.,Shyam, P.,Sastry,G.,Aspell,A.,Agarwal,S.,Herbert-Voss,A., Krueger, G., Henighan,T., Child,R., Ramesh,A.,Ziegler,D.M.,Wu,J., Winter,C.,...&Amodei,D.(2020).LanguageModelsareFew-ShotLearners.arXivpreprintarXiv:2005.14165. URL:<https://arxiv.org/abs/2005.14165>
3. Соколенко В.А.Сучасні тенденції розвитку маркетингу / В.А.Соколенко / /НТУ «ХПІ». 2012 [Електроннийресурс].URL: <https://arxiv.org/abs/2005.14165>
4. Devlin,J.,Chang,M.W.,Lee,K.,&Toutanova,K.(2018).BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding ar Xivpreprintar Xiv:1810.04805. URL:<https://arxiv.org/abs/1810.04805>
5. Raffel,C.,Shazeer,N.,Roberts,A.,Lee,K.,Narang,S.,Matena,M.,Zhou,Y.,Li,W.,&Liu,P.J.(2020).Exploring the Limits of Transfer Learning with a Unified Text-to-Text Transformer. Journal of Machine Learning Research, 21(140),1-67. URL:<https://jmlr.org/papers/v21/20-074.html>