

інтелект може масштабувати виробництво якісного контенту навіть для невеликих команд, забезпечуючи конкурентоспроможність у digital-просторі [4].

Крім того, JPMorgan Chase досяг рекордного зростання CTR на 450% завдяки використанню Persado — платформи генеративного ШІ для створення та оптимізації рекламних текстів. У 2019 році банк уклав 5-річний договір з Persado, що дозволило переписувати існуючі маркетингові матеріали та створювати персоналізовані повідомлення для конкретних аудиторій. Це показує, наскільки ефективним може бути ШІ у створенні змісту, який привертає увагу користувачів [4].

Хоча інструменти автогенерації текстів мають значний потенціал у маркетингу, їх обмеження – залежність від даних, проблеми точності мовлення та «галюцинації» – потребують ретельного людського контролю і комбінованого підходу. Це дозволяє забезпечити високу якість контенту, зберегти емоційну привабливість і підвищити довіру аудиторії.

Незважаючи на наведені виклики, тренд на використання ШІ-генераторів у маркетингу буде лише посилюватися, тому подальші дослідження повинні зосередитися на удосконаленні алгоритмів для покращення точності й емоційної насиченості генерованого контенту, що стане ключовим кроком до більш ефективного використання ШІ у маркетингових стратегіях.

Список використаних джерел

1. Paris, M. (2024). The Top 8 AI Text Generator Tools To Improve Writing. *Forbes* 101. URL: <https://www.forbes.com/sites/technology/article/ai-text-generators/>
2. Sembly AI. (2024). Top 10 AI Text Generators in 2025. Sembly AI. URL: <https://www.sembly.ai/blog/top-ai-text-generators/>
3. Aravind, B. (2025). AI-Generated Content vs. Human-Written Content: What the KPIs Don't Tell You (Until You Live It). URL: <https://www.linkedin.com/pulse/ai-generated-content-vs-human-written-what-kpis-dont-you-aravind-jf0ac>
4. Van der Wilt, J. (2025). 11 Best AI Advertising Examples of 2025. *DataFeed Watch Blog*. URL: <https://www.datafeedwatch.com/blog/best-ai-advertising-examples>

3. МОДЕЛІ ГЛИБОКОГО НАВЧАННЯ В ЗАДАЧАХ РОЗПІЗНАВАННЯ ЗОБРАЖЕНЬ ТА ОБРОБКИ МОВИ

Єрмоленко С.В.,
здобувач третього (наукового) рівня вищої освіти,
ПВНЗ «Європейський університет»

РОЗПІЗНАВАННЯ ТЕКСТОВИХ ТА ВІЗУАЛЬНИХ ОБ'ЄКТІВ ДЛЯ МУЛЬТИМОДАЛЬНОГО АНАЛІЗУ КОНТЕНТУ

Сучасний цифровий світ характеризується швидким зростанням обсягів мультимедійного контенту, що включає текстову, візуальну та відео інформацію. Цей інформаційний вибух створює значні виклики для ефективного аналізу та вилучення знань. Традиційні методи аналізу контенту, які часто зосереджуються лише на одній модальності (текст або зображення), демонструють обмежену ефективність у контексті багатоваріантного, складного контенту, де семантичне значення часто формується на перетині різних типів даних. Недостатність однотипного аналізу стає все більш очевидною, що зумовлює нагальну потребу в розробці та впровадженні автоматизованих систем, здатних до глибокого, цілісного аналізу контенту. Такі системи є затребуваними у широкому спектрі галузей, включаючи інформаційний пошук, моніторинг засобів масової інформації, забезпечення безпеки, маркетинг, медицину та освіту.

Ключовою проблемою в цій галузі є розробка ефективних комп'ютерних систем машинного навчання, які б могли одночасно розпізнавати, інтерпретувати та інтегрувати інформацію з текстових та візуальних об'єктів. Це вимагає не лише створення досконалих моделей для обробки природної мови (Natural Language Processing, NLP) та комп'ютерного зору (Computer Vision, CV) окремо, але й розробки синергетичних архітектур, де різні модальності взаємодоповнюють та збагачують одна одну для всебічного аналізу контенту. Таким чином, дослідження повинні бути спрямовані на створення систем, здатних до повного розуміння мультимодального контенту, що є актуальним науково-технічним завданням.

Порядок проведення досліджень в цьому напрямку наступний: формулювання ключової ідеї інтегрованого розпізнавання текстових та візуальних об'єктів для аналізу контенту, представлення огляду базових принципів розпізнавання в кожній модальності, обґрунтування переваг мультимодального підходу та сучасна реалізація системи. Викладені положення закладають концептуальну основу для подальшого дослідження шляхів підвищення ефективності систем розпізнавання.

Розпізнавання текстових об'єктів – це процес ідентифікації та категоризації значущих елементів у текстових даних. Це включає виявлення іменованих сутностей (імен людей, назв організацій, географічних локацій), визначення ключових слів та фраз, тематичне моделювання, аналіз тональності (сентимент-аналіз) та встановлення семантичних взаємозв'язків між поняттями. Фундаментальною дисципліною для розв'язання цих задач є обробка природної мови (NLP).

Історично, перші підходи до аналізу тексту базувалися на правилах та статистичних моделях, таких як Наївний Байєс (Naive Bayes), який і досі знаходить застосування в задачах класифікації тексту, наприклад, для фільтрації спаму. Однак, справжній прорив у NLP пов'язаний з розвитком методів машинного навчання, зокрема глибокого навчання. Рекурентні нейронні мережі (Recurrent Neural Networks, RNN), включаючи їх більш складні варіанти, такі як мережі з довгою короткостроковою пам'яттю (Long Short-Term Memory, LSTM), продемонстрували здатність ефективно обробляти послідовності текстових даних, враховуючи попередній контекст.

Сучасний етап розвитку NLP характеризується домінуванням моделей на основі архітектури Трансформер (Transformer), таких як BERT (Bidirectional Encoder Representations from Transformers) та GPT (Generative Pre-trained Transformer). Ці моделі, завдяки механізмам уваги (attention mechanisms), здатні вловлювати довгострокові залежності та складні семантичні зв'язки в тексті, що значно підвищило якість виконання різноманітних NLP-задач. Еволюція від простих статистичних моделей до трансформерних архітектур кардинально змінила можливості автоматизованого аналізу тексту.

Основні задачі, що вирішуються за допомогою розпізнавання текстових об'єктів, включають:

- Класифікація тексту – віднесення документа до однієї чи кількох категорій.
- Визначення тематики (Topic Modeling) – автоматичне виявлення прихованих тем у великих колекціях документів.
- Аналіз тональності (Sentiment Analysis) – визначення емоційного забарвлення тексту (позитивне, негативне, нейтральне).
- Вилучення іменованих сутностей (Named Entity Recognition, NER) – ідентифікація та класифікація іменованих сутностей у тексті.
- Машинний переклад.

Список використаних джерел

1. Ramachandram, D., Taylor, G. W. (2017). Deep multimodal learning: A survey on recent advances and trends. *IEEE Signal Processing Magazine*, 34(6), 96-108.
2. Baltrusaitis, T., Ahuja, C., & Morency, L. P. (2018). Multimodal machine learning: A survey and taxonomy. *IEEE transactions on pattern analysis and machine intelligence*, 41(2), 423-443.
3. Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., Polosukhin, I. (2017). Attention is all you need. *Advances in neural information processing systems*, 30.

Замлинна Я.В.,

здобувач другого (магістерського) рівня вищої освіти,

Катуніна О.С.,

кандидат економічних наук, доцент,

Київський національний економічний Університет імені Вадима Гетьмана

ПРИКЛАДНІ ЗАДАЧІ ВИКОРИСТАННЯ ДИФУЗІЙНИХ МОДЕЛЕЙ В ЕКОНОМІЦІ ТА УПРАВЛІННІ

До основних інструментів генеративного моделювання належать варіаційні автоенкодера (Variational Autoencoder, VAE), генеративно-змагальні мережі (Generative Adversarial Networks, GAN) [1] та дифузійні моделі (Diffusion

Models). Варіаційні автокодери (VAE) використовують варіаційний байєсівський підхід у латентному просторі та здебільшого застосовуються для відтворення даних і дослідження їхньої внутрішньої структури. GAN реалізують змагальну взаємодію генератора і дискримінатора, що дозволяє досягати високої якості зображень, хоча часто супроводжується нестабільністю навчання. Дифузійні моделі, засновані на зворотному стохастичному процесі, демонструють стабільність і здатність до генерації реалістичних зображень. У таблиці 1 наведено узагальнене порівняння зазначених підходів:

Таблиця 1. Порівняння основних підходів у генеративному моделюванні

Характеристика	VAE (Variational Autoencoder)	GAN (Generative Adversarial Network)	Diffusion Models (DM)
Принцип роботи	Байєсівський латентний простір, реконструкція	Змагання генератора та дискримінатора	Зворотний процес денойзингу (відновлення з шуму)
Якість генерації	Помірна, часто змазана	Висока, але нестабільна	Дуже висока, фотореалістична
Стабільність навчання	Висока	Часто нестабільна (mode collapse)	Висока, масштабована
Контрольованість	Через латентний простір	Обмежена, потребує кастомізації	Висока (текстові промпти, умови)
Швидкість генерації	Швидка	Швидка	Повільна (багато кроків)
Вимоги до обчислень	Помірні	Середні–високі	Високі, добре масштабується на GPU/TPU
Основні застосування	Реконструкція, аномалії, генерація латентного простору	Deepfakes, креативний контент, перетворення стилів	Текст-до-зображення, медицина, дизайн, симуляція
Приклади моделей	β -VAE, VQ-VAE	DCGAN, StyleGAN, CycleGAN	DDPM, Stable Diffusion, DALL·E 2, Imagen

Дифузійні моделі набули популярності як інструмент ефективного створення мультимедійного контенту — зображень, аудіо й відео на

базі зворотного дифузійного процесу [2]. Дифузійні моделі - це клас генеративних моделей, які вчаться поступово знімати шуми в даних шляхом реверсування процесу дифузії. Основна ідея полягає в тому, щоб почати з чистого шуму та ітеративно вдосконалити його до високоякісної вибірки з цільового розподілу [2]. В основі методу лежить поступове перетворення шуму в остаточне зображення за допомогою серії дифузійних кроків. Кожен крок має на меті наблизити шум до реального розподілу даних, що сприятиме видаленню шуму з зображення [4].

Створення реалістичних зображень на базі дифузійних моделей відбувається при використанні нейронної мережі типу U-Net, яка побудована за допомогою двовимірних CNN. Зображення, яке попадає в дифузійну модель, зменшується в розмірах, переходячи в так званий latent space, після чого до нього додається шум певною кількістю кроків. Маючи текстову вказівку, або запит, з блоку «Conditioning» шум видалається, формуючи матрицю, яка потім, виходячи з latent space, перетворюється в повноцінне зображення [4].

Відомими дифузійними алгоритмами є Stable Diffusion, DALL-E та Midjourne, які поєднують як текстове представлення інформації, що притаманне великим мовним моделям (LLMs), так і здатність згорткових шарів знаходити та локалізувати деталі на зображеннях [4]. Перевагами дифузійних моделей є такі:

Дифузійні моделі здатні генерувати дуже реалістичні та деталізовані зображення. У багатьох випадках вони перевершують інші методи, зокрема GAN, за візуальною якістю [5].

На відміну від GAN, де генератор і дискримінатор змагаються між собою, що часто призводить до нестабільного навчання, дифузійні моделі не мають цієї змагальної динаміки, завдяки чому навчання проходить стабільніше [5].

Обчислення в дифузійних моделях можна легко розпаралелювати, особливо на сучасних графічних процесорах (GPU — *Graphics Processing Unit*) або тензорних процесорах (TPU — *Tensor Processing Unit*). Це дозволяє тренувати великі моделі швидше [6].

Дифузійні моделі можна побудувати на основі різних нейромережових архітектур, наприклад, U-Net, що дозволяє адаптувати дифузійні моделі до різних типів даних: зображення, аудіо, текст тощо [6].

У економіці та управлінні дифузійні моделі демонструють високу ефективність (Таблиця 2) у низці прикладних задач, зокрема: автоматизованій генерації візуального контенту для бізнес-презентацій і звітності; створенні рекламних матеріалів і маркетингових креативів на основі текстових запитів; візуалізації складних фінансових або логістичних даних у доступній та естетичній формі; симуляції сценаріїв ринкової поведінки з метою стратегічного планування; генерації synthetic data для навчання або тестування економетричних моделей; а також побудові інтерактивних середовищ для освіти, аналітики та віртуальної співпраці.

Таблиця 2. Прикладні напрями використання дифузійних моделей в економіці та управлінні

Завдання / Галузь	Застосування дифузійної моделі	Очікуваний ефект / Результат
Бізнес-аналітика	Візуалізація аналітичних звітів за текстовими описами	Швидке створення інфографіки, підвищення сприйняття даних
Маркетинг і реклама	Генерація банерів, постерів і візуальних креативів за описом кампанії	Економія часу і витрат на дизайн, A/B тестування образів
Фінансове планування	Симуляція ринкових сценаріїв у візуальній формі	Краще розуміння ризиків і трендів
Електронне навчання та EdTech	Створення навчальних візуалізацій за короткими описами тем	Персоналізація навчання, зростання залученості
Управління персоналом	Візуалізація корпоративних структур або сценаріїв співпраці	Оптимізація управлінських процесів, комунікація змін
Розробка продукту	Швидке створення концептуальних зображень продуктів за текстовим описом	Прототипування ідей без участі дизайнерів
Економетричне моделювання	Генерація синтетичних даних для тренування моделей	Більше даних для точніших прогнозів

Проведене практичне дослідження включало генерацію зображень на основі текстових запитів за допомогою моделі Stable Diffusion v1.5, яка є представником дифузійних моделей нового покоління. Модель була завантажена

з відкритого репозиторію Hugging Face та розгорнута у середовищі Python із використанням бібліотеки *diffusers*. У якості прикладного сценарію було задано текстову інструкцію (*prompt*): «*A futuristic city at sunset, flying cars, neon lights, ultra detailed, 4K*». Робочий процес включав такі етапи:

- ініціалізацію моделі та підключення авторизаційного токена;
- вибір відповідного обчислювального пристрою (CPU або GPU);
- генерацію зображення з подальшим його візуальним аналізом і збереженням.

Генерація здійснювалася за допомогою латентного дифузійного процесу: випадковий шум поступово трансформувався у семантично змістовне зображення, згідно з умовою текстового запиту. Ключовим компонентом архітектури моделі виступала нейромережа на базі U-Net, яка реалізує зворотну трансформацію з латентного простору шляхом ітеративного усунення шуму.

Результатом серії експериментів стала низка альтернативних варіантів зображень, що якісно та деталізовано відтворюють зміст текстового запиту. На Рис. 1 представлено одне з синтезованих зображень з ознаками футуризму, неоновим підсвічуванням та виразною атмосферністю, що демонструє здатність дифузійної моделі ефективно інтерпретувати високорівневі текстові дескриптори та відображати їх у візуальній формі.



Рисунок 1. Згенероване зображення промту «*A futuristic city at sunset, flying cars, neon lights, ultra detailed, 4K*»

Джерело: розроблено автором

Крім основної генерації, було додатково протестовано можливість варіювання параметрів генерації, зокрема `num_inference_steps` та `guidance_scale`, що безпосередньо впливають на чіткість, деталізацію та відповідність зображення текстовому опису. Це дало змогу оцінити гнучкість моделі у адаптації під різні цілі: від швидкого чернеткового ескізу до високоякісного художнього рендеру. Збільшення кількості кроків покращувало деталізацію, але вимагало більше обчислювальних ресурсів, що підтверджує необхідність балансування між якістю результату та продуктивністю системи.

У дослідженні продемонстровано, що зміна навіть одного ключового слова у запиті («futuristic» на «cyberpunk», або додавання стилістичних маркерів як-от «digital art», «isometric», «concept art») кардинально змінює композицію сцени. Це відкриває широкі можливості для інтерактивного дизайну: користувач може поступово коригувати промт і миттєво отримувати візуальний фідбек, фактично створюючи адаптивне середовище для візуального прототипування — без потреби у залученні художника. Такий підхід може знайти застосування у швидкій розробці контенту та ідей для економічних застосувань, а також створенні ігрових рівнів, інтерфейсів, у галузі освіти та маркетингу тощо.

Практичне застосування дифузійних моделей засвідчує їх високу ефективність у завданнях генерації зображень на основі текстових інструкцій. Висока якість, стабільність навчання та універсальність дифузійних моделей забезпечують їхню значущість у сучасних дослідницьких і прикладних проєктах, особливо в креативних індустріях, дизайні, освіті та мультимедійній продукції. Очікується, що дифузійні моделі й надалі розширюватимуть сфери свого використання, стаючи невід'ємною частиною повсякденної реальності.[3].

На основі проведеного дослідження можна зробити висновок про широкі перспективи використання дифузійних моделей у таких галузях, як:

генерація візуального контенту у сфері дизайну, ігор та кіноіндустрії;

покращення медичних зображень (наприклад, відновлення томограм);
автоматизоване створення навчальних матеріалів;
створення нових засобів взаємодії у віртуальній та доповненій реальності.

Таким чином, дифузійні моделі є не лише інструментом генерації зображень, а й ефективною технологічною платформою для побудови інтелектуальних систем для широкого використання в економіці та управлінні.

Список використаних джерел

1. Gupta, R., Tiwari, S., & Chaudhary, P. (2025). *Generative AI: Techniques, Models and Applications*. Springer Nature Switzerland. <https://doi.org/10.1007/978-3-031-82062-5>
2. Unite.AI. (n.d.). *Розуміння дифузійних моделей: глибоке занурення в генеративний ШІ*. Retrieved May 16, 2025, from <https://www.unite.ai/uk/understanding-diffusion-models-a-deep-dive-into-generative-ai/>
3. Unite.AI. (n.d.). *Дифузійні моделі в ШІ – усе, що вам потрібно знати*. Retrieved May 16, 2025, from <https://www.unite.ai/uk/diffusion-models-in-ai-everything-you-need-to-know/>
4. Малиш, О. (n.d.). *Використовуємо CNN для обробки зображень. Частина перша*. DOU. Retrieved May 16, 2025, from <https://dou.ua/forums/topic/48368/>
5. Kaddoura, S. (2023). *A Primer on Generative Adversarial Networks*. Springer International Publishing. <https://doi.org/10.1007/978-3-031-32661-5>
6. Raut, R., et al. (2023). *Generative Adversarial Networks and Deep Learning*. Chapman and Hall/CRC. <https://doi.org/10.1201/9781003203964>