4. Aslan M. F., Sabanci K., Aslan B. Artificial intelligence techniques in crop yield estimation based on Sentinel-2 data: a comprehensive survey. Sustainability. 2024. Vol. 16, No. 18. Article 8277. DOI: https://doi.org/10.3390/su16188277.

5. Estimates of crop yield anomalies for 2022 in Ukraine based on Copernicus Sentinel-1, Sentinel-3 satellite data, and ERA-5 agrometeorological indicators / E. Panek-Chwastyk et al. Sensors. 2024. Vol. 24, No. 7. Article 2257. DOI: https://doi.org/10.3390/s24072257.

6. Machine learning in agriculture: a comprehensive updated review / L. Benos et al. Sensors. 2021. Vol. 21, No. 11. Article 3758. DOI: https://doi.org/10.3390/s21113758.

7. Site-specific sunflower yield forecasting based on spatial analysis and machine learning / V. H. Hnatiienko et al. Reports of the National Academy of Sciences of Ukraine. 2025. No. 4. P. 17-26. DOI: https://doi.org/10.15407/dopovidi2025.04.017.

УДК 004.8

ПОРІВНЯННЯ ЛІНІЙНОГО ТА КОСИНУСНОГО ГРАФІКІВ ШУМУ ДЛЯ ВИБОРУ КІЛЬКОСТІ КРОКІВ СЕМПЛЮВАННЯ В ІНФОРМАЦІЙНІЙ СИСТЕМІ ГЕНЕРАЦІЇ ЗОБРАЖЕНЬ НА ОСНОВІ ДИФУЗІЙНИХ МОДЕЛЕЙ

Дмитро Мокан

Миколаївський національний аграрний університет, м. Миколаїв, Україна

Науковий керівник: В. В. Співак,
асистент Миколаївський національний аграрний університет

Дифузійні моделі [1, 4] стали основою сучасних інформаційних систем генерації зображень. Час відповіді такої системи майже пропорційний кількості кроків семплювання, тому цей параметр є ключовим під час проєктування. Графік шуму визначає, як відношення сигнал/шум (ВСШ) розподіляється вздовж процесу, а отже, скільки кроків припадає на область, де зображення майже повністю втрачене [2, 5].

Принцип дифузійних моделей полягає в поступовому додаванні гауссового шуму до зображення в прямому процесі та навчанні нейромережі відновлювати зображення в зворотному процесі [1]. У вихідній постановці DDPM для генерації одного зображення потрібно близько тисячі послідовних викликів мережі, а детерміністичні семплери типу DDIM [3] дозволяють істотно скоротити їх кількість без повторного навчання. Саме тому питання, які кроки зберегти під час скорочення, має практичну вагу для проєктувальника системи.

Типова інформаційна система генерації зображень містить інтерфейс введення запиту, модуль кодування тексту, модуль семплювання, що багатократно викликає нейромережу-денойзер, та декодер; у моделях латентної дифузії [4] семплювання виконується в стисненому латентному просторі, проте кількість викликів денойзера залишається головним чинником затримки. Саме тому модуль налаштування параметрів семплера потребує обґрунтованих критеріїв вибору кількості кроків.

Мета дослідження – кількісно порівняти лінійний [1] і косинусний [2] графіки шуму за розподілом ВСШ уздовж $T = 1000$ кроків та оцінити частку кроків семплювання DDIM [3], що припадає на майже чистий шум.

Дослідження обчислювальне й не передбачає навчання моделі: використано лише аналітичні означення графіків із [1, 2]. У прямому процесі зашумлення зображення x_0 до кроку t має вигляд $x_t = \sqrt{\bar{\alpha}_t}\cdot x_0 + \sqrt{(1-\bar{\alpha}_t)}\cdot \varepsilon$, де ε – гауссів шум, а

$$\bar{\alpha}_t = \prod_{s=1}^{t}(1 - \beta_s) \qquad (1)$$

Для лінійного графіка β_t зростає від 10^{-4} до 0,02 [1]. Для косинусного $\bar{\alpha}_t = f(t)/f(0)$, $f(t) = \cos^2[(t/T + s)/(1 + s)\cdot\pi/2]$, $s = 0{,}008$, $\beta_t = \min(1 - \bar{\alpha}_t/\bar{\alpha}_{t-1}; 0{,}999)$ [2]. ВСШ на кроці t:

$$SNR(t) = \frac{\bar{\alpha}_t}{1-\bar{\alpha}_t} \qquad (2)$$

Крок вважали майже шумовим, якщо ВСШ < 0,01, тобто частка сигналу в x_t не перевищує приблизно 1 % (для даних одиничної дисперсії); цей поріг є умовним і задається дослідником. Для DDIM [3] розглянуто S = 50 рівномірних кроків t = 20, 40, …, 1000. Обчислення виконано в Python (NumPy).

Вибір ВСШ як показника зумовлений тим, що з виразу для x_t відношення дисперсії сигналу до дисперсії шуму дорівнює $\bar{\alpha}_t/(1-\bar{\alpha}_t)$, тобто воно напряму характеризує, скільки інформації про вихідне зображення зберігається на кроці t. Для наочності на рис. 1 ВСШ подано в децибелах ($10\cdot\lg$ SNR). Такий підхід не потребує навчання нейромережі, тому його можна застосувати на етапі проєктування системи ще до запуску дорогих експериментів.

Основні показники подано в табл. 1, залежність ВСШ від t – на рис. 1. За лінійного графіка ВСШ опускається нижче одиниці вже при t* = 260, а 326 кроків із 1000 (32,6 %) мають ВСШ < 0,01. За косинусного графіка t* = 497, а таких кроків лише 64 (6,4 %). Для DDIM із S = 50 це відповідає 17 кроків (34 %) проти 4 (8 %).

Різницю між графіками добре ілюструють проміжні значення. На початку процесу (t = 100) ВСШ становить 8,71 (9,4 дБ) для лінійного графіка і 34,8 (15,4 дБ) для косинусного, тобто косинусний графік зашумлює зображення повільніше. Наприкінці процесу (t = 750) ВСШ дорівнює 0,0034 (−24,7 дБ) для лінійного та 0,169 (−7,7 дБ) для косинусного графіка. Таким чином, за лінійного графіка зображення втрачається значно раніше, і пізні кроки несуть мало корисного сигналу.

Таблиця 1. Показники розподілу ВСШ для T = 1000

Показник	Лінійний	Косинусний
$\bar{\alpha}T$ (за t = 1000)	$4{,}04\cdot10^{-5}$	$2{,}43\cdot10^{-9}$
ВСШ при t = 500, дБ	−10,7	−0,1
Перший крок t*, де ВСШ < 1	260	497
Кроків з ВСШ < 0,01 із 1000	326 (32,6 %)	64 (6,4 %)
Кроків з ВСШ < 0,01 із 50 (DDIM)	17 (34 %)	4 (8 %)

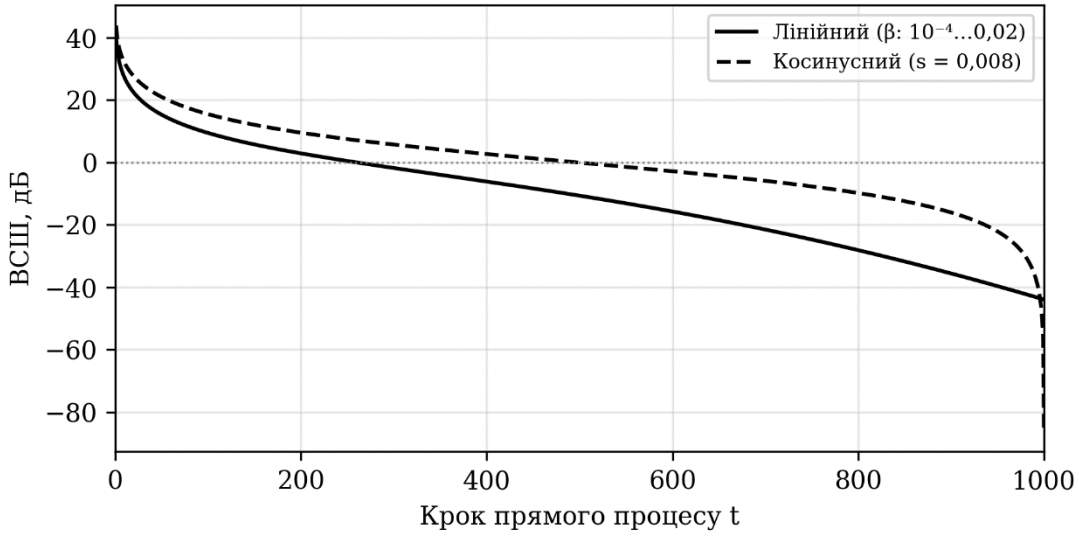


Рис. 1. Залежність ВСШ (дБ) від кроку прямого процесу t для лінійного та косинусного графіків шуму

Лінійний графік «спалює» значну частину кроків в області, де x_t практично не містить інформації про зображення; за малого S це означає, що близько третини обчислювального бюджету системи

витрачається на цю область. Косинусний графік розподіляє кроки рівномірніше за ВСШ. Крім того, жоден із графіків не забезпечує нульового кінцевого ВСШ: $\bar{\alpha}_T = 4{,}04 \cdot 10^{-5}$ (лінійний) і $2{,}43 \cdot 10^{-9}$ (косинусний), тобто вхід моделі при навчанні не є чистим шумом, який подається при інференсі [5]. Виправлення [5] полягає в масштабуванні

$$\sqrt{\bar{\alpha}'_t} = \frac{(\sqrt{\bar{\alpha}_t} - \sqrt{\bar{\alpha}_T}) \cdot \sqrt{\bar{\alpha}_1}}{\sqrt{\bar{\alpha}_1} - \sqrt{\bar{\alpha}_T}} \qquad (3)$$

що дає $\bar{\alpha}'_T = 0$ за збереження $\bar{\alpha}'_1 = \bar{\alpha}_1$ (перевірено розрахунком для лінійного графіка).

Наочно наслідки видно з арифметики кроків DDIM: із 50 рівномірно вибраних кроків в області ВСШ ≥ 0,01 за лінійного графіка залишається 33 кроки, а за косинусного – 46. Отже, за однакових затримок косинусний графік надає семплеру більше кроків у діапазоні, де відбувається відновлення структури зображення. Водночас це не означає, що лінійний графік не можна застосовувати: моделі, навчені з ним, зокрема в латентній дифузії [4], працюють коректно, а нерівномірний розподіл кроків можна частково компенсувати вибором нерівномірної сітки кроків семплера. Цей напрям потребує окремої перевірки і в роботі не досліджувався.

Зазначимо особливість косинусного графіка на рис. 1: різке падіння ВСШ на останніх кроках зумовлене обмеженням $\beta_t \leq 0{,}999$, яке запроваджене в [2] для уникнення сингулярності, тому значення ВСШ у кінцевій точці (−86,1 дБ) слід розглядати як наслідок цього обмеження, а не як властивість основної частини кривої. Це не впливає на висновок про частку майже шумових кроків, оскільки він спирається на основну область графіка.

Розробнику інформаційної системи доцільно до вибору графіка та кількості кроків S перевіряти розподіл ВСШ за формулою (2) і частку кроків із ВСШ < 0,01, а також контролювати нульове кінцеве ВСШ. Обмеження: не оцінювалась якість згенерованих зображень (наприклад, за FID), тому результат стосується лише розподілу ВСШ і не доводить переваги одного графіка за якістю; поріг 0,01 умовний.

У межах інформаційної системи це можна реалізувати як окремий модуль налаштування семплера: під час завантаження моделі він обчислює $\bar{\alpha}_t$ за конфігурацією графіка, будує розподіл ВСШ і повідомляє адміністратору частку майже шумових кроків та значення кінцевого ВСШ. На основі цього звіту обирається кількість кроків S, яка відповідає вимогам до часу відповіді. Кінцевий вибір доцільно підтверджувати оцінюванням якості на контрольному наборі запитів, оскільки аналіз розподілу ВСШ лише звужує коло розумних варіантів. Подальші дослідження можуть охопити вибір нерівномірної сітки кроків та зіставлення розподілу ВСШ з метриками якості генерації.

Усі наведені значення можна відтворити за формулами (1)–(3) та параметрами графіків із [1, 2]: розрахунок не містить випадкових величин, не залежить від апаратного забезпечення й виконується за частки секунди на звичайному комп'ютері. Це дозволяє вбудувати перевірку розподілу ВСШ безпосередньо в процедуру розгортання моделі та повторювати її для кожної нової конфігурації графіка без додаткових витрат на навчання. Водночас остаточна оцінка придатності конфігурації має спиратися також на експертне або метричне оцінювання якості зображень.

Висновки. 1. За лінійного графіка ВСШ < 1 настає при $t^* = 260$, за косинусного – при $t^* = 497$. 2. Частка майже шумових кроків становить 32,6 % для лінійного та 6,4 % для косинусного графіка (34 % і 8 % для DDIM із S = 50). 3. Обидва графіки мають ненульове кінцеве ВСШ, що слід враховувати в системі, зокрема застосувавши масштабування (3).

Список використаних джерел

1. Ho J., Jain A., Abbeel P. Denoising diffusion probabilistic models. Advances in Neural Information Processing Systems. 2020. Vol. 33. P. 6840–6851. URL: https://arxiv.org/abs/2006.11239 (дата звернення: 12.10.2026).

2. Nichol A., Dhariwal P. Improved denoising diffusion probabilistic models. Proceedings of Machine Learning Research. 2021. Vol. 139. P. 8162–8171. URL: https://arxiv.org/abs/2102.09672 (дата звернення: 12.10.2026).

3. Song J., Meng C., Ermon S. Denoising diffusion implicit models. International Conference on Learning Representations (ICLR 2021). URL: https://arxiv.org/abs/2010.02502 (дата звернення: 12.10.2026).

4. Rombach R., Blattmann A., Lorenz D., Esser P., Ommer B. High-resolution image synthesis with latent diffusion models. Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). 2022. P. 10684–10695. URL: https://arxiv.org/abs/2112.10752 (дата звернення: 12.10.2026).

5. Lin S., Liu B., Li J., Yang X. Common diffusion noise schedules and sample steps are flawed. IEEE/CVF Winter Conference on Applications of Computer Vision (WACV 2024). URL: https://arxiv.org/abs/2305.08891 (дата звернення: 12.10.2026).

УДК 004.85

ПОРІВНЯЛЬНИЙ АНАЛІЗ ПРОСТИХ МЕТОДІВ КЛАСИФІКАЦІЇ ЗА ФІЗИКО-ХІМІЧНИМИ ПОКАЗНИКАМИ ЗРАЗКІВ ВИНА

Ренат Бондар

Миколаївський національний аграрний університет, м. Миколаїв, Україна

Науковий керівник: О. Є. Богатєнкова,
асистент Миколаївський національний аграрний університет

Оцінювання та класифікація виноробної продукції за лабораторними показниками є типовим завданням підтримки рішень у виноробстві: фізико-хімічні тести доступні на етапі сертифікації, а моделі інтелектуального аналізу даних можуть доповнювати роботу експертів [1]. Для невеликих вибірок практично важливо обрати простий, інтерпретований класифікатор і з’ясувати, наскільки його результат залежить від попередньої обробки ознак [2].

Мета дослідження – порівняти за точністю та макро-F1 чотири прості методи класифікації (метод k найближчих сусідів, наївний байєсівський класифікатор, логістичну регресію, дерево рішень) на відкритих даних хімічного аналізу вина та оцінити вплив стандартизації ознак на метод k найближчих сусідів.

Використано відкритий набір Wine з репозиторію UCI [3]: 178 зразків вина з однієї області Італії, 13 неперервних показників хімічного аналізу (вміст спирту, яблучної кислоти, золи, фенолів, проліну тощо) та три класи за сортом винограду (59, 71 і 48 зразків). Слід зазначити, що цільовою змінною є сорт, а не органолептична оцінка якості, яку моделюють у [1] на значно більших вибірках вина Vinho Verde; тому отримані результати стосуються порівняння методів на лабораторних показниках вина, а не прогнозування експертної оцінки якості. Моделі реалізовано в бібліотеці scikit-learn [4]. Для методу k найближчих сусідів (k = 5) ознаки стандартизовано за формулою (1), відстань між зразками обчислювали за формулою (2):

$$z_{ij} = \frac{x_{ij} - \mu_j}{\sigma_j} \qquad (1)$$

$$d(x, x') = \sqrt{\sum_{j=1}^{13} (x_j - x'_j)^2} \qquad (2)$$

де x_{ij} – значення j-ї ознаки i-го зразка; μ_j і σ_j – середнє та стандартне відхилення j-ї ознаки на навчальній вибірці; $d(x, x')$ – евклідова відстань між зразками. Також застосовано гаусів наївний Байєс, логістичну регресію з L2-регуляризацією (C = 1) та дерево рішень глибиною 3; гіперпараметри не підбирали. Якість оцінювали п’ятикратною стратифікованою перехресною перевіркою, повтореною 10 разів (50 розбиттів, початкове значення генератора 42); стандартизацію виконували лише на навчальних блоках, щоб уникнути витоку інформації. Показники – частка правильних відповідей (accuracy) і макро-F1.