ISSN 2079-0031 (Print) ISSN 2411-0558 (Online)

УДК 004.94:004.22 **DOI: 10.20998/2411-0558.2025.01.01**

Л. І. КОРОТКА, канд. техн. наук, доц., Український державний університет науки і технологій,
В. С. МАКАРЧЕНКО, асп. Український державний університет науки і технологій

ЗАСТОСУВАННЯ ГЕНЕРАТИВНО-ЗМАГАЛЬНИХ МЕРЕЖ ПРИ МОДЕЛЮВАННІ ПЛАЗМОХІМІЧНИХ ПРОЦЕСІВ ОТРИМАННЯ НАНОСИСТЕМ

Для генерації додаткових даних при моделюванні плазмохімічних процесів отримання наносистем пропонується застосувати генеративно-змагальні мережі. Розглянуто проблемні аспекти процесу одержання таких навчальних даних, зокрема: якість синтетичних даних, збереження функціональних залежностей між вхідними та вихідними даними, стабільне навчання генератора та дискримінатора. Розібрано налаштування параметрів навчання нейронних мереж: значення шуму в генераторі, розмір латентного простору, функцій активації, відслідковування границь синтетичних даних після генерації. Спроєктовано та реалізовано відповідну архітектуру генеративно-змагальної мережі. Отримана синтетична вибірка використовується для навчання неглибоких нейронних мереж, а реальний набір даних використовується для перевірки якості навчання та тестування. Іл.: 3. Бібл. 11 назв..

Ключові слова: генеративно-змагальні мережі; моделювання плазмохімічних процесів; отримання наносистем; синтетичні дані; навчання нейронних мереж.

Вступ. Зацікавленість нанотехнологіями постійно зростає через їх широке застосування, наприклад, у біомедицині, енергетиці, електроніці тощо. Моделювання плазмохімічних процесів є складним особливо через їх багатофакторну природу. Фізичне моделювання є затратним і непростим, тому комп’ютерне моделювання буває раціональною альтернативою. Побудова адекватних моделей з наявними даними представляє трудомісткий процес та потребує якісної оцінки, як самих даних так і моделей, що описують плазмохімічний процес отримання наносистем. Останні можуть мати обмеження щодо їх точності та якості.

Підходи, які використовують класичні аналітичні методи, а саме: побудови, наприклад, регресійних моделей не завжди задовольняють їх розробників, і особливо ця проблема постає, коли кінцевими

користувачами створених моделей не є фахівці в області математичного чи комп'ютерного моделювання.

Для вирішення таких завдань необхідно мати достатню кількість експериментальних даних, зробити їх аналіз (і статистичний у тому числі), перевірити якість вибірки та ін. Оскільки процес моделювання плазмохімічних процесів отримання наносистем носить складний нелінійний характер, із залежностями між параметрами, які важко формалізувати класичними методами, то і побудова моделей, які його описують теж представляється не тривіальним завданням. Використання напрямів штучного інтелекту, в тому числі, нейронних мереж є перспективним та доцільним [1 – 2].

Актуальність проблеми. Проведення натурних експериментів потребує додаткових фінансових витрат та пов'язано зі складністю плазмохімічного процесу, тому побудова адекватних моделей і їх реалізація не потребує зайвих пояснень. Зауважимо, що маємо обмежену кількість визначених експериментальних даних, тоді використання нейромережевих технологій при побудові моделей розглядуваного плазмохімічного процесу має суттєве значення. При умові достатньої кількості навчальної та тренувальної вибірок можна застосувати навчені моделі нейронних мереж в якості альтернативи аналітичним методам.

Метою дослідження є отримання синтетичних даних за допомогою генеративних змагальних мереж та використання їх для навчання шарової нейронної мережі (НМ) в якості навчального масиву даних. Протестувати її на реальних експериментальних даних та довести, що такий підхід дозволяє отримати адекватну модель плазмохімічного процесу отримання наносистем у вигляді нейронної мережі.

Постановка завдання дослідження полягає в тому, що маємо деякі результати проведення натурних експериментів з отримання наносистем у фізико-хімічних процесах, які можна використовувати як множину даних (вхідних та вихідних). Для побудови моделей при симуляції плазмохімічних процесів отримання наночасток формалізуються вхідні та вихідні дані. На етапі попереднього аналізу вхідних змінних розглядалися: концентрація прекурсора $AgNO_3$ (x_1, г/л); співвідношення прекурсора та стабілізатора (x_2, г/л); тривалість обробки (x_3, с. та x_4, ум. од.); довжина хвилі (x_5, нм та x_6, в долях); вид стабілізатора (x_7); сила струму (x_8, мВ та

x_9, в долях); тиск газової фази (x_{10}, мПа); напруження (x_{11}, В). Вихідними параметрами є: розмір отриманих наночасток (y_1, нм); індекс полідисперсності (*IPD*) (y_2); *Z*-потенціал (y_3). Інтервали варіювання усіх змінних та їх експериментальні значення є відомими, але, очевидно, що залежність F між ними невідома та її необхідно встановити. У такому варіанті маємо відобразити одинадцяти вимірний простір у тривимірний за допомогою відображення F: $R^{11} \rightarrow R^3$.

Після аналізу даних вирішено залишити п’ять вхідних змінних та три вихідних: $R^5 \rightarrow R^3$. Стосовно вхідних змінних було залишено: співвідношення прекурсора та стабілізатора ($x_2 \in [0{,}10; 0{,}30]$); тривалість обробки ($x_3 \in [10; 300]$ с.); довжина хвилі ($x_5 \in [360; 412]$ нм); вид стабілізатора ($x_7 \in [0{,}86; 0{,}98]$); сила струму($x_8 \in [0{,}545; 1{,}00]$. Вихідні дані лишаються незмінними. Стосовно останньої вхідної змінної (сила струму), то тут можна її розглянути у декількох дискретних значеннях, а значить працювати і без неї, але враховувати її як параметр протікання самого плазмохімічного процесу. В такому випадку необхідно рознести дані на відповідні класи та виконати генерацію даних для кожного з них. Зауважимо, що без обмеження суджень у роботі було прийнято рішення не розбивати експериментальні дані на окремі підмножини відносно цієї змінної.

Як зазначалося, у [3] побудова аналітичних моделей можлива, але представляє собою кропіткий процес. Крім того, користувачами програмного забезпечення, яке буде реалізовувати моделювання, є хіміки-технологи. Тому інформаційна система, яка реалізується, повинна бути зрозумілою та не потребувати від користувачів особливих знань щодо моделювання.

Основна частина. Використання елементів штучного інтелекту (і нейронних мереж у тому числі) у плазмохімії є перспективним, оскільки сам досліджуваний складний процес має: велику кількість змінних, які взаємодіють не лінійно; високий рівень турбулентності та нестабільності у плазмі, що в свою чергу ускладнює прогнозування результатів класичними методами. Тому застосування некласичних підходів є далекосяжним напрямом [4 – 5]. Як відомо, штучні нейронні мережі представляють собою універсальні апроксимуючі системи, то і плазмохімія з їх використанням не є виключенням [6, 7]. Для врахування

складних залежностей між параметрами, проблема вибору відповідної архітектури НМ є окремим завданням для проєктувальників таких систем. Навіть якщо проблема з архітектурою та топологією мережі вирішена, то далі постає питання розробки алгоритмів навчання НМ, які забезпечують стабільну роботу моделі навіть у разі неповних або зашумлених даних [4, 8]. Якщо все ж таки не вдається отримати якісну навчену мережу, то тоді необхідно звернути увагу безпосередньо на дані, з яким вона працює. Існує достатня кількість формул розрахунків залежності між об'ємом вибірки та її похибкою навчання. Як правило, така залежність носить обернено пропорційний характер, тобто для отримання більш точної моделі необхідно відповідно більший об'єм даних. І ці зауваження стосуються моментів, які враховують похибки вимірювань даних та похибки обчислень/округлень при комп'ютерному моделюванні.

У роботі пропонується для прогнозування вихідних даних плазмохімічного процесу отримання наночасток використовувати шарові нейронні мережі, але проблема їх навчання є тому, що експериментальних даних обмаль. Для роботи з недостатньою кількістю навчальних масивів існують різні підходи в залежності від того які завдання розв'язуються. Якщо, наприклад, відбувається робота із зображеннями, то можна створити додаткові зразки шляхом їхнього повороту, зсуву, розтягнення тощо [9]. Таким чином існують деякі підходи вирішення проблеми подібного роду, але їх вибір залежить безпосередньо від розробника такої системи [10, 11].

У випадку необхідності роботи із числовими масивами, на думку авторів доцільно використовувати методи генерації синтетичних даних, наприклад, SMOTE (Synthetic Minority Over-sampling Technique), автоенкодер та інші. Для генерації табличних даних, що імітують розподіл оригінальних невеликих табличних наборів, корисно використовувати генеративно-змагальні мережі (ГЗМ або GANs). Такий підхід дозволяє зберегти статистичну схожість синтетичних даних з реальними. Якість навчання нейронної мережі значною мірою буде залежати від того, наскільки добре синтетичні дані, які отримано з використанням GANs, відображають оригінальний розподіл у наборі.

Технологія GANs поєднує в собі, як відомо, міждисциплінарні підходи, а саме: двоосібну гру з нульовою сумою, оптимізацію при навчанні НМ, принцип максимізації правдоподібності та інші. В даному

випадку, генератор $G(z, P_G)$ використовується для створення синтетичних даних. Тут z – латентний простір, на основі якого формуються нові фейкові дані подібні до реальних, P_G – параметри (ваги та зміщення), які використовуються для опису архітектури та функціонування генератора $G(z, P_G)$. У контексті нейронних мереж вказані параметри визначають зв'язки між нейронами, які навчаються під час тренування. Завдання дискримінатору $D(x, P_D)$ полягає в тому, щоб відрізнити/розпізнати генеровані дані z від реальних x. Відповідно P_D – є параметрами $D(x, P_D)$, оскільки дискримінатор представляє собою теж нейронну мережу. Функція виграшу з нульовою сумою формулюється як:

$$U(D,G) = \min_G \max_D V(D, G) = \mathscr{E}_{x\sim Pdata}[lnD(x)] + \mathscr{E}_{z\sim Pz(z)}[ln(1-D(G(z)))], \quad (1)$$

тут $G(z)$ – дані, що створені генератором G, який бере випадковий вектор z із латентного простору й перетворює його у згенеровані дані, які схожі на реальні; $D(G(z))$ – це ймовірність, яку дискримінатор D присвоює синтетичним даним, вважаючи їх реальними (результат знаходиться в діапазоні [0, 1], де одиниця – дані схожі на реальні та нуль – дані виглядають підробленими).

Роль у функції втрат полягає в тому, що частина втрат відповідає за покарання генератора G, якщо дискримінатор D добре виконує свою роботу і правильно класифікує G, як підробку. В той же час генератор G намагається мінімізувати вираз, змушуючи дискримінатор оцінювати D: $D(G(z))\rightarrow 1$ та вважати згенеровані дані реальними. Таким чином, у формулі (1) першої частини дискримінатор намагається максимізувати ймовірність правильного класифікування реальних даних; а у другій її частині дискримінатор також намагається максимізувати ймовірність правильної класифікації згенерованих даних як підроблених. Щодо використання натурального логарифму у формулі (1), то з огляду статистичної інтерпретації він зазвичай використовується у задачах максимізації правдоподібності, а з математичної – спрощення процесу обчислення градієнтів при оптимізації, оскільки має прості похідні.

Для проєктувальника будь-якої нейронної мережі постають питання щодо архітектури та вибору їх параметрів P_G, P_D. На рисунку 1 приведено

ISSN 2079-0031 (Print) ISSN 2411-0558 (Online)

одну з множини мереж, з якими проводилися експерименти. Архітектуру дискримінатора та генератора містить рис. 2.

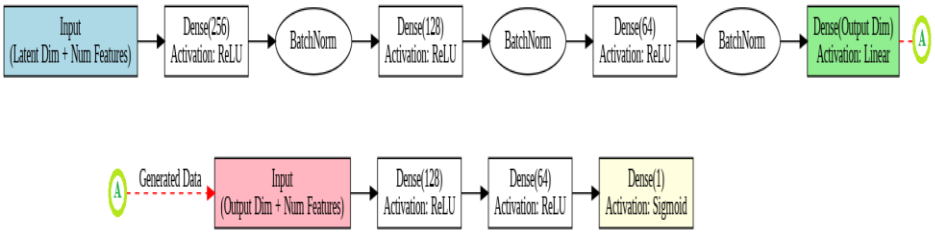


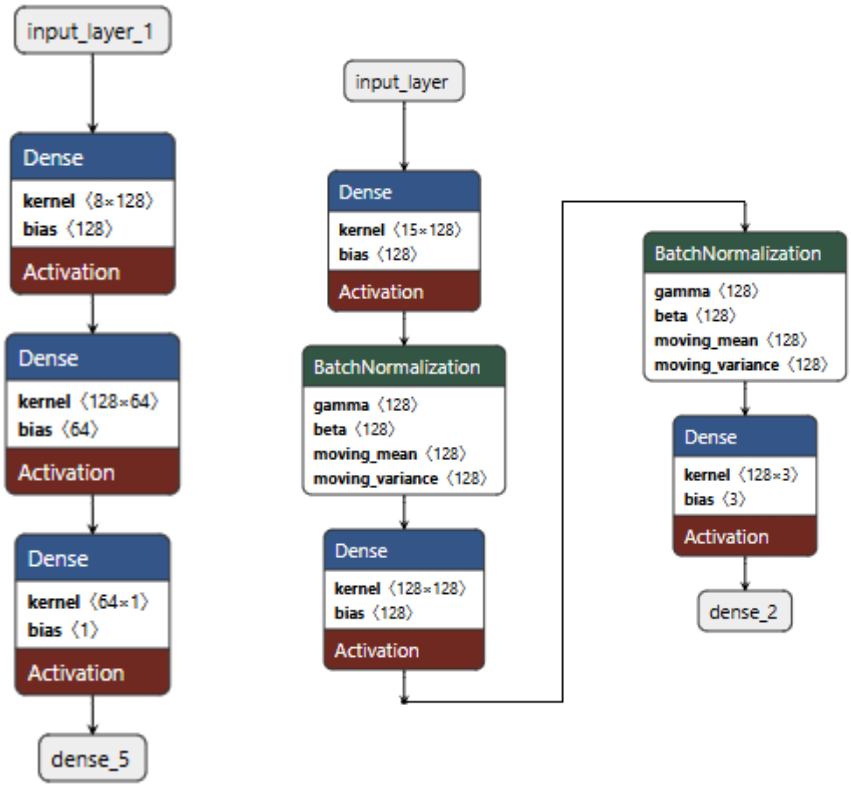
Рис. 1. Архітектура мережі GANs

У процесі проведення досліджень з мережами було виявлено декілька фактів, на яких автори вважають за доцільне зупинитися та проаналізувати. Оскільки, синтетичні дані повинні мати однакові з реальними границі їх змінення, то цей факт є очевидним та не потребує додаткових пояснень. При навчанні мережі на кожній епосі можна прослідкувати за кроками оновлення для генератора *G* і дискримінатора *D*. Проміжні результати, які виводяться під час тренування GANs, відображають показники втрат (loss) для генератора (*G* loss) та дискримінатора (*D* loss).

Втрати дискримінатора (*D* loss) показують наскільки добре дискримінатор здатен відрізнити справжні дані від синтетичних (генерованих). Якщо зразки класифікуються успішно, то і значення втрат (*D* loss) є близьким до нуля. В противному випадку, якщо значення близьке до одиниці або є більшим неї, то це може свідчити про труднощі дискримінатора з точним розрізненням справжніх та фейкових даних.

Величина показника втрат генератора (*G* loss) показує наскільки добре генератор "обманює" дискримінатор, створюючи синтетичні дані, які дискримінатор приймає за справжні. Тому вважається, що для GANs ідеальним сценарієм є збалансованість між розглянутими величинами (*D* loss) та (*G* loss). Очевидно, що генератор прагне "обманювати" дискримінатор як найкраще, а дискримінатор, у свою чергу, прагне достатньо добре відрізняти синтетичні дані від справжніх [4].

ISSN 2079-0031 (Print) ISSN 2411-0558 (Online)



а) дискримінатор б) генератор

Рис. 2. Моделі дискримінатора та генератора

Якщо з цією проблемою вдається впоратися, то може виникнути наступна. При прагненні генератора ввести в оманну дискримінатор та при аналізі отриманих фейкових даних, може скластися ситуація, коли ці дані концентруються в околі якоїсь точки багатовимірного простору, де генератор навчився найкраще "обманювати" дискримінатор. Таким чином, отримуються неякісні синтетичні дані. Виходить, що генератор концентрується на моделюванні обмеженого підпростору вихідних даних. Причин цьому може бути декілька, наприклад: модель недостатньо складна, тоді генератор може не навчитися представляти всю варіативність даних; дані не рівномірно розподілені, тобто сильно згруповані і модель відтворює подібні структури; розмір або структура латентного простору можуть не відповідати складності даних.

При проведенні експериментів з мережами авторам прийшлося варіювати розмір останнього. Зазвичай на практиці розмір латентного

простору рекомендовано збільшувати, але приблизно не більше п’ятдесяти. Результати чисельних експериментів довели, що при моделюванні плазмохімічних процесів для реальних даних його значення залишили рівним десяти. Розширення латентного простору не надало суттєвих зрушень щодо розширення границь діапазону змінення згенерованих даних. Очевидно, що розмір латентного простору залежить від завдання, яке розв’язується. Одним із покращень роботи генератору стало додавання штрафу у функцію втрат генератора, який змушує генерувати дані рівномірно на всьому інтервалі.

Паралельно з процесом побудови та налаштування параметрів GANs зберігається питання щодо якості синтетичної вибірки. Проводилися експерименти з архітектурами, а саме: збільшення кількості шарів та вузлів у генератора. Наведемо деякі параметри моделі та її архітектури. Було застосовано відносно простий підхід побудови моделі генератору, зокрема, використано стек шарів без розгалужень чи злиттів. Звісно, що існують інші підходи такі як, складні архітектури з розгалуженнями, злиттями, кількома входами чи виходами, тощо. У роботі використано в якості першого шару повнозв’язний шар (Dense з 256 нейронами), який має кількість параметрів: сума розміру латентного простору, властивостей помножена на кількість нейронів у шарі, включаючи ваги та зсуви. Функції активації, що застосовано у роботі, який містить рисунок 2. Зауважити слід, що проводилися експерименти з різними функціями, але очікувано практично найкраще себе зарекомендувала ReLu для прихованих шарів.

Для стабілізації та прискорення навчання використано техніку нормалізації виходів цього шару під час тренування. Далі комбінація повнозв’язного шару та нормалізація застосовується двічі, але з різною кількістю нейронів (128 та 64 відповідно у шарах Dense). Вихідний шар містить 64 нейрони. У процесі чисельних експериментів з архітектурами генеративно-змагальних мереж та, виходячи зі значень втрат у процесі навчання, було спрощено архітектуру GANs. Прибрано один повнозв’язний шар Dense з нормалізвцією. Крім того зменшення кількості нейронів спростило структуру мережі генератора (рис. 2б). Кількість усього параметрів 19971; параметрів 19459, які оптимізуються під час навчання та 512, які не змінюються. Щодо моделі дискримінатору, то вона значно простіша, а саме: містить два повнозв’язних шари (Dense) відповідно з кількістю нейронів 128 та 64. Вихідний шар має один нейрон.

ISSN 2079-0031 (Print) ISSN 2411-0558 (Online)

Усього модель має 9473 параметри, з них ті ж 9473, які оптимізуються під час навчання (рис. 2а).

Як вже зазначалося у роботі, що у моделях мереж GANs слід дотримуватися балансу між складністю генератора та дискримінатора, що і продемонстрували експерименти з різними моделями при навчанні. Очевидно, що відношення кількості параметрів і складності архітектури залежить від задачі, але можна виділити, що ключовим моментом для успішного тренування є співвідношення складності генератора та дискримінатора і кілька загальних принципів. Оскільки генератор має завданням створити дані такі, щоб вдало "обманювати" дискримінатор, то генератор повинен бути достатньо складним для генерації реалістичних фейкових даних. Якщо генератор буде занадто простим, то він не зможе ефективно створювати синтетичні дані.

В свою чергу дискримінатор теж не може бути занадто простим, бо не зможе відрізняти та правильно класифікувати реальні і генеровані дані, але він не повинен бути надмірно складним, бо може швидко навчитися розрізняти справжні дані від підроблених, що призводить до «домінування» дискримінатора та поганого тренування. Тому баланс між моделями повинен зберігатися і тому важливо експериментувати з архітектурами мереж та кількістю їх параметрів. Отже дискримінатор скоріш за все буде менш складним, але достатньо потужним, щоб виявляти найдрібніші відмінності між справжніми та згенерованими даними. Використання меншої кількості параметрів у дискримінаторі може запобігти його надмірному навчанню.

Як показують експерименти з моделями генератора та дискримінатора, підтримка балансу складності досить важлива. При використанні симетричного підходу, тобто моделі мають однакову кількість параметрів, на простих задачах є доцільним. Застосування дисбалансу у сторону дискримінатору може бути корисним на початкових етапах навчання та при складних розподілах даних. Використаний у роботі датасет, дозволяє зробити висновки, що показники плазмохімічного процесу не мають занадто складного розподілу, тому невеликий дисбаланс в сторону генератора є доцільним. Можна стверджувати, що знаходження балансу/дисбалансу та зміна пропорцій було вирішено шляхом емпіричного налаштування, а от регуляризація дискримінатора задля уникнення його переваги над генератором не дала очікуваного результату.

ISSN 2079-0031 (Print) ISSN 2411-0558 (Online)

Якщо вважати вирішеною задачу балансу моделей та те, що дані генеруються не в околі деякої точки/точок, то це не означає, що синтетичні дані будуть отримані саме ті, які влаштують дослідників. Автори стикнулися з наступною проблемою, що границі змінення генерованих даних повністю відповідають реальним, але при подальшому їх застосуванню виявляється, що отриманий датасет не дуже якісний. При детальному аналізі синтетичних даних виявилося, що порушується вплив та залежність факторів, який був наявний у реальних даних.

З'ясувалося, що (рис. 3) мульти залежність вихідних змінних (розмір наночасток y_1, полідисперсність y_2, Z-потенціал y_3) від вхідних даних (співвідношення прекурсора та стабілізатора; тривалість обробки; довжина хвилі; вид стабілізатора; сила струму) при генерації синтетичних даних порушується. Саме ця залежність була попередньо встановлена у відсотках для реальних даних при застосуванні різних підходів. Приведено результати із використанням алгоритму ансамблевого навчання, який реалізує метод випадкових лісів (RandomForest) (рис. 3). Як відомо, він базується на комбінації кількох дерев рішень (decision trees) і забезпечує високу точність прогнозування, стійкість до перенавчання та гнучкість для роботи зі складними нелінійними даними. Не важко помітити, що для вихідних змінних, таких як: розмір наночасток та полідисперсність суттєвий вплив має тривалість обробки, потім співвідношення прекурсора та стабілізатора для Z-потенціалу, і тільки потім присутній вплив усіх інших вхідних змінних. Очевидно, що це необхідно враховувати при генерації даних, що і було далі зроблено у роботі.

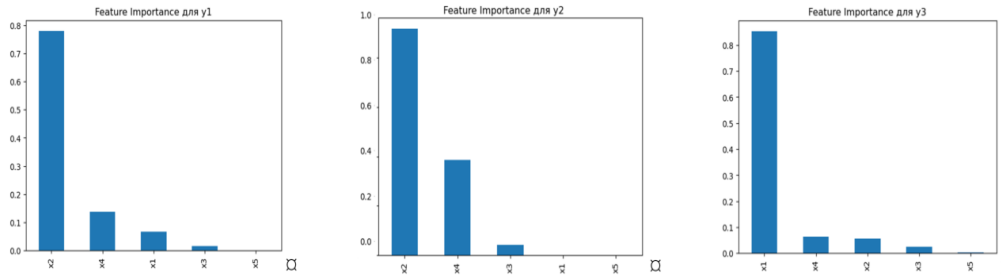


Рис. 3. Залежність між вхідними та вихідними змінними реальних даних

ISSN 2079-0031 (Print) ISSN 2411-0558 (Online)

Таким чином, маємо збалансовані моделі змагальних мереж та рівномірно розподілені у певних границях синтетичні дані, але порушується зв'язок між змінними, який необхідно враховувати. Повстала нова проблема відносно генерації синтетичних даних. Оскільки залежність реальних вихідних та вхідних змінних відома, то для збереження цих зв'язків було прийнято рішення щодо регуляризації для фейкових даних.

З метою мінімізації відхилення від заданої залежності було введено обчислення спеціальної функції втрат (loss function), яка враховує залежності між вхідними та генерованими даними. Функція втрат додається до загальної функції втрат нейронної мережі для того, щоб змусити модель дотримуватись певних апріорних залежностей між вхідними та вихідними змінними. У результаті модель навчається не лише мінімізувати загальну похибку, а й враховувати ці зв'язки, що робить її прогноз більш інтерпретованим та реалістичним.

Нарешті було отримано навчальну синтетичну вибірку, яка відповідала усім вимогам та була застосована для навчання неглибокої нейронної мережі алгоритмом зворотного поширення похибки. Фактично названу мережу було використано задля перевірки якості синтетичних даних. Оскільки такі мережі є достатньо відомими, то автори вважають не зупинятися детально на їх навчанні та налаштуванні їх параметрів (функції активації швидкість навчання, тощо) [10]. Тестування цієї мережі відбувалося на реальних даних. Зауважимо, що відносна сумарна похибка мережі по кожній вихідній змінній склала не більше 5-8%, що є достатнім для тестової вибірки. Отримані результати чисельних досліджень задовільнили експертів хіміків-технологів, які мають чітке уявлення щодо складності моделювання плазмохімічного процесу отримання наночасток.

Висновки. Використання нейромереж дає можливість враховувати нелінійність та складні взаємозв'язки між параметрами технологічного процесу, який є багатофакторним, із залежностями між параметрами, які важко формалізувати класичними методами.

Для навчання нейромереж необхідна стратегія доповнення даних, наприклад, за рахунок генерації синтетичних даних. Доведено, що відсутність достатньої кількості експериментальних даних через високу

вартість та складність проведення експериментів, можна компенсувати фейковими генерованими даними.

Проведено аналіз проблемних аспектів, які можуть виникати при генерації даних. Обґрунтовано необхідність адаптації нейромережевих підходів під специфіку процесу. Розглянуто пошук оптимальної архітектури генеративно-змагальних мереж та налаштування параметрів.

Продемонстровано доцільність застосування генеративно-змагальних мереж для отримання синтетичних даних, які у подальшому використано для навчання неглибокої нейронної мережі при моделюванні параметрів плазмохімічних процесів отримання наночасток.

Список літератури

1. *Zhengwei Wang, Qi She and Tomas, E. Ward* “Generative Adversarial Networks in Computer Vision: A Survey and Taxonomy”. ACM Computing Surveys, 2020. pp.1-41.

2. *Goodfellow J., Pouget-Abadie M., Mirza B., Xu D., Warde-Farley S., Ozair A., Courville and Y. Bengio* “Generative adversarial nets”, Advances in Neural Information Processing Systems, 2014. pp. 2672–2680.

3. *Макарченко В.С., Коротка Л.І.* «Аналіз та підготовка даних при моделюванні плазмохімічних процесів отримання наносистем», Системні технології, 5(154), 2024. С. 30-39.

4. *Korotka L., Klevzhyts D. and Shvydko D.* “Use of generative-adversarial networks when creating content”, Artificial intelligence. National Academy of Sciences of Ukraine Institute of Artificial Intelligence Problems MES of Ukraine and NAS of Ukraine, 2024, № 2 (99). pp. 32-47. (https://doi.org/10.15407/jai2024.02.032)

5. *M.-Y. Liu and O. Tuzel* “Coupled generative adversarial networks”, Advances in neural information processing systems, 2016. pp. 469–477.

6. Verma, S., Chugh, S., Ghosh, S., Rahman, B.M. (2022) Artificial Neural Network Modelling for Optimizing the Optical Parameters of Plasmonic Paired Nanostructures. *Nanomaterials*, 12, 2022. p. 170. (https://doi.org/10.3390/nano12010170)

7. *Skiba M. I., Vorobyova V. I.* Synthesis of silver nanoparticles using orange peel extract prepared by plasmochemical extraction method and degradation of methylene blue under solar irradiation. Advances in Materials Science and Engineering, 2019. P.1-8.

8. Aurelien Geron (2020) Hands-On Machine Learning with Scikit-Learn and TensorFlow: Concept, Tools, and Techniques to Build Intelligent Systems. *Київ: Комп’ютерне видавництво «Діалектика»* (пер. з англ. Ю.Н. Артеменко), 2020. 688 р.

9. *Дмитрієнко В., Леонов С., Мезенцев М.* Нейронна мережа для розпізнавання та класифікації зображень на кордонах декількох класів. Вісник Національного технічного університету «Харківський політехнічний інститут». 2024, Том 1 № 1-2 (11-12). С. 87-95.

10. *Коротка Л.І.* Застосування нейро-нечітких мереж для визначення раціональних параметрів чисельного розв’язку систем диференціальних рівнянь. Вісник

Кременчуцького національного університету імені Михайла Остроградського. – Кременчук: КрНУ, 2021. Випуск 5(130). С. 55-61.

11. Ian Goodfellow, Yoshua Bengio, Aaron Courville. (2016) Deep Learning (Adaptive Computation and Machine Learning series). *The MIT Press*. 2016. 800p.

References:

1. Zhengwei Wang, Qi She and Tomas, E. Ward (2020), "Generative Adversarial Networks in Computer Vision: A Survey and Taxonomy", *ACM Computing Surveys*, pp.1-41.

2. Goodfellow J., Pouget-Abadie M., Mirza B., Xu D., Warde-Farley S., Ozair A., Courville and Y. Bengio (2014), "Generative adversarial nets", *Advances in Neural Information Processing Systems*, pp. 2672–2680.

3. Makarchenko V.S., Korotka L.I. (2024), «Analiz ta pidhotovka danykh pry modeliuvanni plazmokhimichnykh protsesiv otrymannia nanosystem», *Systemni tekhnolohii*, 5(154), S. 30-39. (DOI:https://doi.org/10.34185/1562-9945-5-154-2024-04)

4. Korotka L., Klevzhyts D. and Shvydko D. (2024), "Use of generative-adversarial networks when creating content", *Artificial intelligence. National Academy of Sciences of Ukraine Institute of Artificial Intelligence Problems MES of Ukraine and NAS of Ukraine*, № 2 (99). pp. 32-47. (https://doi.org/10.15407/jai2024.02.032)

5. M.-Y. Liu and O. Tuzel, (2016) "Coupled generative adversarial networks", *Advances in neural information processing systems*, pp. 469–477.

6. Verma S., Chugh S., Ghosh S., Rahman B.M.A. (2022), Artificial Neural Network Modelling for Optimizing the Optical Parameters of Plasmonic Paired Nanostructures. *Nanomaterials*, 12, p. 170. (https://doi.org/10.3390/nano12010170)

7. Skiba M. I., Vorobyova V. I. (2019), Synthesis of silver nanoparticles using orange peel extract prepared by plasmochemical extraction method and degradation of methylene blue under solar irradiation. *Advances in Materials Science and Engineering*, P.1-8.

8. Aurelien Geron (2020), Hands-On Machine Learning with Scikit-Learn and TensorFlow: Concept, Tools, and Techniques to Build Intelligent Systems. *Kyiv: Kompiuterne vydavnytstvo «Dialektyka»* (per. z anhl. Yu.N. Artemenko), 688 p.

9. Dmytriienko V., Leonov S., Mezentsev M. (2024), Neironna merezha dlia rozpiznavannia ta klasyfikatsii zobrazhen na kordonakh dekilkokh klasiv. *Visnyk Natsionalnoho tekhnichnoho universytetu "Kharkivskyi politekhnichnyi instytut"*. Tom 1 № 1-2 (11-12). S. 87-95.

10. Korotka L.I., (2021) Zastosuvannia neiro-nechitkykh merezh dlia vyznachennia ratsionalnykh parametriv chyselnoho rozviazku system dyferentsialnykh rivnian. *Visnyk Kremenchutskoho natsionalnoho universytetu imeni Mykhaila Ostrohradskoho*. – Kremenchuk: KrNU, Vypusk 5(130). C. 55-61.

11. Ian Goodfellow, Yoshua Bengio, Aaron Courville (2016), Deep Learning (Adaptive Computation and Machine Learning series). *The MIT Press.* 800p.

Статтю представив д.т.н., проф. Національного технічного університету "Харківський політехнічний інститут" В.І. Носков.

Надійшла (received) 21.02.2025

Коротка Лариса Іванівна, канд. техн. наук, доцент
Український університет науки і технологій ННІ УДХТУ
просп. Науки, 8, м. Дніпро, 49005
Korotka Larysa Ivanivna, Candidate of Technical Sciences
Ukrainian State University of Science and Technology ESI USUCT
Nauky Ave., 8, Dnipro, 49005
e-mail: larysakorotka@gmail.com
ORCID ID: 0000-0003-0780-7571

Макарченко Віктор Сергійович, аспірант
Український університет науки і технологій ННІ УДХТУ
просп. Науки, 8, м. Дніпро, 49005
Makarchenko Viktor Serhiyovych, Postgraduate
Ukrainian State University of Science and Technology ESI USUCT
Nauky Ave., 8, Dnipro, 49005
e-mail: viktor_makarchenko@udhtu.edu.ua

УДК 004.94:004.22

Застосування генеративно-змагальних мереж при моделюванні плазмохімічних процесів отримання наносистем / Коротка Л.І., Макарченко В.С. // Вісник НТУ "ХПІ". Серія: Інформатика та моделювання. – Харків: НТУ "ХПІ". – 2025. – № 1 (13). – С. 7 – 21.

Для генерації додаткових даних при моделюванні плазмохімічних процесів отримання наносистем пропонується застосувати генеративно-змагальні мережі. Розглянуто проблемні аспекти процесу одержання таких навчальних даних, зокрема: якість синтетичних даних, збереження функціональних залежностей між вхідними та вихідними даними, стабільне навчання генератора та дискримінатора. Розібрано налаштування параметрів навчання нейронних мереж: значення шуму в генераторі, розмір латентного простору, функцій активації, відслідковування границь синтетичних даних після генерації. Спроєктовано та реалізовано відповідну архітектуру генеративно-змагальної мережі. Отримана синтетична вибірка використовується для навчання неглибоких нейронних мереж, а реальний набір даних використовується для перевірки якості навчання та тестування. Іл.: 3. Бібл. 11.

Ключові слова: генеративно-змагальні мережі; моделювання плазмохімічних процесів; отримання наносистем; синтетичні дані; навчання нейронних мереж.

UDC 004.94:004.22

Application of generative-adversarial networks in modeling plasma-chemical processes for obtaining nanosystems / Korotka L.I., Makarchenko V.S. // Herald of the National Technical University "KhPI". Series of "Informatics and Modeling". – Kharkov: NTU "KhPI". – 2025. – № 1 (13). – P. 7 – 21.

The use of neural networks makes it possible to take into account the nonlinearity and complex relationships between the parameters of the technological process, which is multifactorial, with dependencies between the parameters that are difficult to formalize by classical methods. To generate additional data in modeling physiochemical processes for the production of nanosystems, it is proposed that generative adversarial networks (GANs) be used.

The problematic aspects of the process of obtaining such training arrays are considered, in particular: the quality of synthetic data, preservation of functional dependencies between input and output data, and stable training of the generator and discriminator. The settings of neural network training parameters are analyzed: the value of noise in the generator, the size of the latent space, activation functions, and tracking the boundaries of synthetic data after generation. The necessity of adapting neural network approaches to the specifics of the process is substantiated. The search for the optimal architecture of the generative-adversarial network and parameter settings are considered.

For example, a data augmentation strategy is required to train no-deep neural networks by generating synthetic data. It is proved that the lack of sufficient experimental data due to the high cost and complexity of conducting experiments can be compensated by fake generated data.

In this paper, we design and implement an appropriate GAN architecture. The resulting synthetic sample is used to train no-deep neural networks, and the real data set is used to test. The expediency of using generative-adversarial networks to obtain synthetic data, which were subsequently used to train a no-deep neural network for modeling the parameters of plasma chemical processes of nanoparticle production, was demonstrated.

Keywords: generative-adversarial networks; modeling of plasma-chemical processes; obtaining a nanosystem; synthetic data; training neural networks.