

УДК 004.032.26

MSC 37D45, 68T07

INFLUENCE OF CHAOS ON ACTIVATION FUNCTIONS IN HOPFIELD NETWORKS

O. S. MAISTRENKO, D. A. KLYUSHIN

Faculty of Computer Science and Cybernetics, Taras Shevchenko National University of Kyiv,
Kyiv, Ukraine, E-mail: o.maistrenko@knu.ua, dmytroklyushin@knu.ua

ВПЛИВ ХАОСУ У ФУНКЦІЯХ АКТИВАЦІЇ МОДЕЛЕЙ ХОПФІЛДА

О. С. МАЙСТРЕНКО, Д. А. КЛЮШИН

Факультет комп'ютерних наук та кібернетики, Київський національний університет імені
Тараса Шевченка, Київ, Україна, E-mail: o.maistrenko@knu.ua, dmytroklyushin@knu.ua

ABSTRACT. Hopfield networks are known for their ability to store and recall patterns. Recently, there has been interest in new types of activation functions and how they can be used in these networks. This paper looks at how chaotic activation functions can be used in Hopfield networks and what effects they have on how the networks work.

KEYWORDS: Hopfield network, activation function, chaotic neural networks, breast cancer.

АНОТАЦІЯ. Мережі Хопфілда, відомі своїми можливостями асоціативної пам'яті, активно вивчалися з моменту їх створення. Новітні досягнення в функціях активації викликали інтерес до вивчення їх застосовності в рамках цих мереж. У цій статті досліджується використання функцій хаотичної активації в мережі Хопфілда та їх вплив на динаміку та продуктивність мережі.

КЛЮЧОВІ СЛОВА: Мережа Хопфілда, функція активації, хаотичні нейронні мережі, рак молочної залози.

ВСТУП

В глибинному навчанні, функції активації відіграють ключову роль у передачі сигналів між шарами нейронів. Хоча функції, такі як сигмоїди або гіперболічні тангенси, були популярні раніше, останні дослідження підкреслюють переваги використання функцій активації, які базуються на хаосі [1]. Функції, які використовують хаос виявляють властивості, які сприяють швидшій та більш ефективній збіжності під час навчання нейронних мереж. Вони дозволяють збільшити область пошуку оптимальних параметрів моделі, сприяючи зменшенню проблеми застрягання в локальних мінімумах та забезпечуючи більш стабільний процес навчання [2, 3].

Це може призвести до покращення якості прогнозування та зменшення часу, необхідного для навчання глибинних моделей. Більше того, використання хаосу у функціях активації може допомогти розв'язати проблему відтворюваності навчання, оскільки ці функції демонструють меншу чутливість до початкових умов [4]. Однак, необхідно враховувати, що використання функцій активації на основі хаосу може потребувати більш складних методів оптимізації та збільшувати вимоги до обчислювальних ресурсів.

Однією з математичних моделей теорії нейронних мереж є модель Хопфілда. Нейронна мережа Хопфілда — це тип рекурентної, повнозв'язної, штучної нейронної мережі з симетричною матрицею зв'язків. У процесі роботи динаміка таких мереж збігається (конвергує) до одного з положень рівноваги. Ці положення рівноваги є локальними мінімумами функціоналу, що називається енергія мережі (у найпростішому випадку — локальними мінімумами негативно визначеної квадратичної форми на n -вимірному кубі). Така мережа може бути використана як автоасоціативна пам'ять, як фільтр, а також для розв'язання деяких завдань оптимізації. На відміну від багатьох нейронних мереж, що працюють до отримання відповіді через певну кількість тактів, мережі Хопфілда працюють до досягнення рівноваги, коли наступний стан мережі дорівнює попередньому. Ця модель описує процес навчання та подальшого розпізнавання образів.

У моделях Хопфілда активаційні функції використовуються для моделювання поведінки нейронів у мозку. Хоча зазвичай використовуються лінійні або сигмоїдальні функції активації, деякі дослідження демонструють, що застосування хаотичних функцій активації може призвести до цікавих результатів [5]. Наприклад, хаотичні функції можуть додати стохастичний елемент до процесу вирішення задач, що може бути корисним для деяких класів задач, таких як оптимізація або розпізнавання образів, де різноманітність та динаміка має важливе значення [6].

Використання хаотичних функцій активації також може поліпшити роботу моделей Хопфілда у вирішенні проблем з великим числом локальних мінімумів. Хаотичні функції можуть допомогти у знаходженні глобальних оптимумів через їхню складнішу та випадкову природу [7]. Але використання хаотичних функцій активації також може ускладнити навчання та збільшити обчислювальні витрати, оскільки вони вимагають більш складних методів оптимізації та обробки.

Додатковою перевагою використання хаотичних функцій активації у моделях Хопфілда є їхня здатність до створення більш складних динамічних залежностей між нейронами. Це може призвести до покращення здатності мережі до адаптації до змін у вхідних даних або до різних умов навколишнього середовища. Такі хаотичні залежності можуть допомогти моделі відтворювати складні динамічні процеси, такі як пам'ять або навчання зі зміщенням [8, 9].

Важливо враховувати, що використання хаотичних функцій активації потребує додаткового уважного налаштування параметрів моделі та підбору підходящих методів навчання. Невірне налаштування може призвести

до непередбачуваної поведінки мережі або до збоїв у процесі навчання. Також слід враховувати, що хаотичні функції активації можуть бути менш ефективними в задачах, де потрібна висока точність або стабільність, так як їхня природа може призвести до більшої варіативності результатів.

Необхідність глибокого дослідження використання хаотичних функцій активації в мережах Хопфілда набуває все більшої актуальності у зв'язку зі зростанням інтересу до нейромереж та їх потенційних застосувань у складних задачах обробки інформації.

Хоча деякі дослідження вже показали перспективність такого підходу, існує недостатнє розкриття цієї теми в науковій літературі. Важливо провести більш глибокі експериментальні та теоретичні дослідження з метою вивчення потенційних переваг та обмежень використання хаотичних функцій активації в мережах Хопфілда [10, 11].

Розуміння динаміки таких мереж з хаотичними функціями активації може привести до важливих відкриттів у сферах зберігання та відновлення інформації, а також у роботі з великими наборами даних.

1. ХАОС У МОДЕЛЯХ ХОПФІЛДА

Позначимо змінною $x(t)$ активність нейрона у момент часу t . Стан $x(t) = 1$ відповідає стану «збудження», а стан $x(t) = -1$ — стану «гальмування».

Нехай $h(t)$ — зовнішній вплив, що чиниться на нейрон, наприклад, з боку інших нейронів. Для зручності вважатимемо час дискретним ($t = 0, 1, 2, \dots$). У [12] було запропоновано, що еволюція нейрона визначається динамічною системою, що представляє собою кусково-лінійне, кусково-лінійне одновимірне відображення

$$x(t + 1) = F(x(t), h(t)) \tag{1}$$

Це сімейство відображень визначається функцією $K(h)$ кута нахилу. Потрібно, щоб ця залежність задовольняла умовам: 1) парність $K(-h) = K(h)$; 2) монотонне зменшення при $h > 0$ (і, отже, монотонне зростання при $h < 0$); 3) $K(0) = 2$. Всі ці вимоги виконані для функції

$$K(h) = \frac{2}{1 + |h|/\mu} \tag{2}$$

Динаміка формального нейрону визначається співвідношенням

$$x(t + 1) = \begin{cases} 1, & h(t) > 0 \\ -1, & h(t) < 0 \end{cases} \tag{3}$$

Тобто без будь-якого перехідного процесу $x(t)$ приймає значення $+1$ або -1 незалежно стану нейрону в попередній момент часу.

Відомо, що реальний нейрон є надзвичайно складною системою інерції [13–16]. Процеси, які у ньому, мають характерні часи від одиниць до сотень (і більше) мілісекунд [17–20]. У [16] було припущено, що при великому значенні зовнішнього впливу h нейрон «швидко» переходить у стан

збудження, тоді як при менших значеннях швидкість переходу в стан збудження зменшується. Це і враховується в наведеній вище хаотичній функції активації Іжикевича у (1), тому що при зменшенні h збільшується $K(h)$ і збільшується час перехідного процесу у (2), хоча єдина особлива точка $x = 1$ все ще є стійкою. Коли ж $K(h)$ переходить через 1, $x = 1$ втрачає стійкість і динаміка нейрона стає хаотичною.

Вище наведений сценарій переходу від порядку до хаосу у (3) спостерігається в багатьох розривних відображеннях або відображеннях з гострою вершиною і називається гомоклінічним вибухом [21]. Як показали чисельні дослідження, результати слабко залежать від конкретного виду функції $K(h)$, якщо цей сценарій виникнення хаосу. Найважливішим параметром для відображення є лише значення $h = p$, у якому $K(h) = 1$, тобто при якому відбувається перехід до хаосу.

2. ПОРІВНЯННЯ ФУНКЦІЙ АКТИВАЦІЇ

Диференційовні функції активації в нейромережах мають ряд вагомих переваг порівняно з їх недиференційованими аналогами. Перш за все, вони дозволяють ефективно використовувати алгоритми оптимізації, такі як зворотне розповсюдження помилки, для навчання мережі. Гладка природа диференційовних функцій дозволяє знаходити оптимальні параметри шляхом пошуку градієнтів, що спрощує процес навчання та покращує швидкість збіжності.

Важливо враховувати, що диференційовні функції активації дозволяють використовувати методи регуляризації, які зменшують перенавчання мережі та покращують її загальну здатність до узагальнення. Наприклад, можна застосовувати $L1$ або $L2$ регуляризацію до ваг моделі, що допомагає контролювати їхню складність та уникнути перенавчання.

Крім того, диференційовні функції активації забезпечують гладку та неперервну залежність між вхідними та вихідними даними мережі, що сприяє плавному переходу між значеннями та полегшує оптимізаційні процеси. Вони також дозволяють використовувати більш потужні алгоритми оптимізації, такі як адаптивні методи градієнтного спуску, що сприяє швидшій та ефективнішій збіжності мережі.

Хаотична функція активації Іжикевича, описана у попередньому розділі, має недолік, а саме те, що вона не є диференційовною при малих значеннях $h(t)$. Цей недолік може бути вирішений шляхом заміни лінійного відображення параметричною функцією Безье (рис. 1). Така заміна дозволяє забезпечити диференційованість функцій активації шляхом виконання умови про біективність рівняння $x(p)$, де x є вектором вхідних значень, а p — параметрами функції Безье.

Для порівняння активаційних функцій в мережі Хопфілда — сигнуму, хаотичної функції активації Іжикевича та її диференційовної модифікації — було проведено наступний експеримент. Вхідними даними були вибірки з нормальних розподілів $N(\mu, \sigma^2)$ двох класів. Параметри μ та σ^2 були підібрані відповідно до того щоб розглянути різноманітні випадки перетину

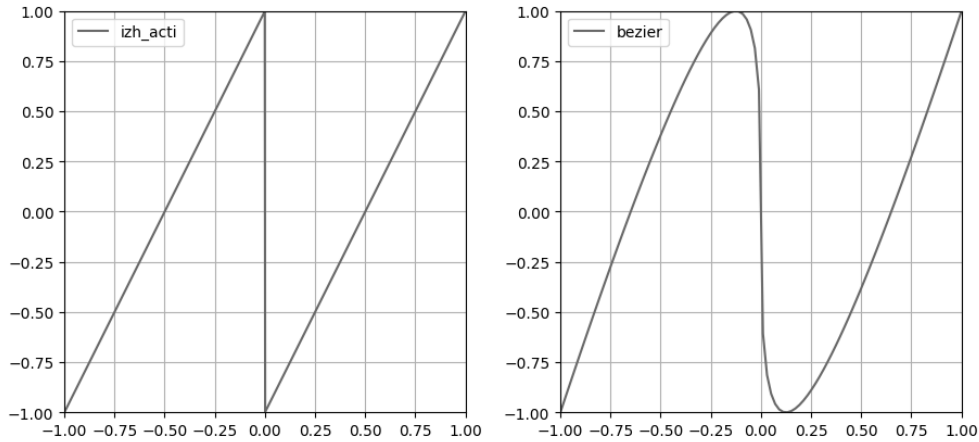


Рис. 1. Хаотична функція активації Іжикевича та її модифікована версія з параметричною кривою Безьє.

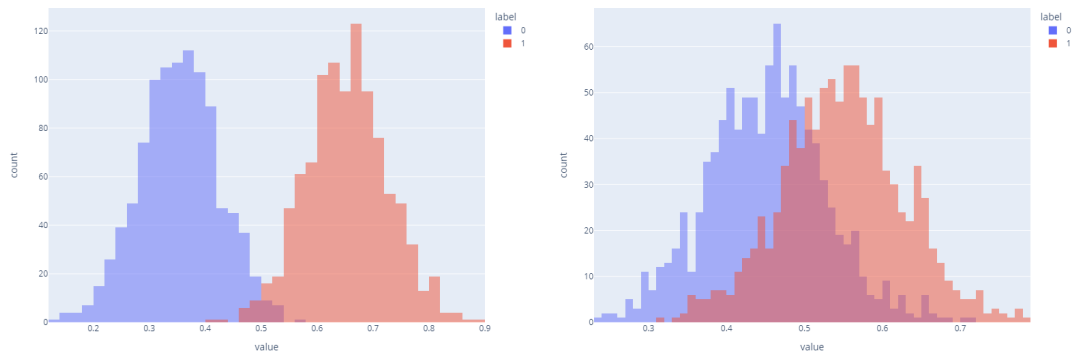


Рис. 2. Вибірки для бінарної класифікації з нормальних розподілів з однаковою дисперсією та різним математичним сподіванням.

розподілів. На тренувальних даних будувалися тренувальні образи відповідно до кожної категорії бінарної класифікації, якість моделей оцінювалися на тестових даних за допомогою Leave One Out – валідації, цільовою метрикою була F1-міра.

Гіпотеза нуль полягає у тому, що функції активації не впливають на результати моделі. Альтернативна гіпотеза полягає у тому, що хаотичні функції активації дають кращі результати.

Розглянемо випадок, коли розподіли перетинаються, $\mu_1 = 0.33$, $\mu_2 = 0.66$, $\sigma_1^2 = \sigma_2^2 = 0.05$. Після 100 експериментів використовуючи сигнум як функцію активації, отримуємо середнє значення F1-міри рівне 0.79. Використовуючи статистичну потужність, можна отримати необхідну кількість експериментів для значущості приросту до середнього значення F1-міри.

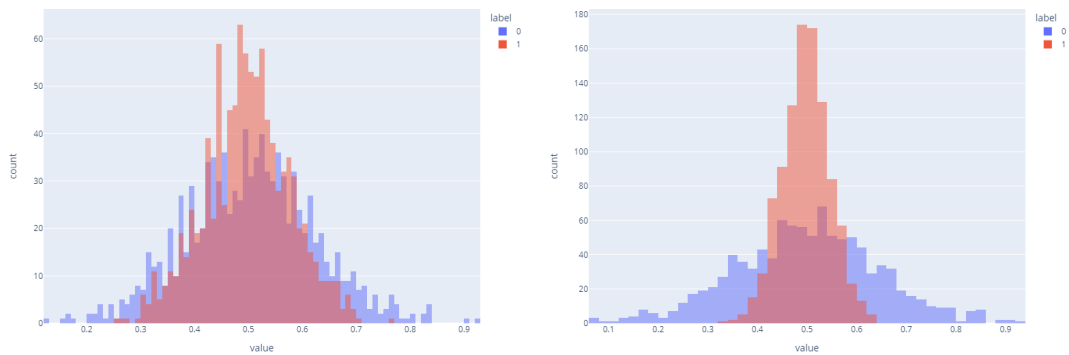


Рис. 3. Вибірки для бінарної класифікації з нормальних розподілів з однаковим математичним сподіванням та різною дисперсією.

№	Назва	z-статистика	p-значення
1	sign_vs_izh	-4.8378	0.0001
2	sign_vs_bezier	-5.0268	0.0001
3	izh_vs_bezier	-1.5855	0.1129

ТАБЛ. 1. Результати z-тестів для експериментів з функціями активацій.

Відповідно, необхідно щонайменше 25580 експерименти для кожної групи для значущості z-тесту про приріст не менш ніж на 1%.

Було проведено 26000 експериментів для кожної з порівнюваних функцій активацій (сигнуму, хаотичної функції активації Іжикевича та її диференційовної модифікації), обчислено F1-міри для кожного експеримента. Середнє значення F1-міри для експериментів з сигнумом склало 0.86, з хаотичною функцією активації Іжикевича — 0.87, з диференційовною модифікацією — 0.873. Після цього було проведено z-тести для кожної пари експериментів (табл. 1).

Обидві вибірки з хаотичними функціями статистично значуще відрізняються від вибірки експериментів з сигнумом, але не мають статистично значущої різниці між собою, хоча модифікована функція дала кращі результати, що можна побачити зі значень z-статистик.

Аналогічно проведено експерименти з іншими параметрами нормальних розподілів, обчислено F1-міри для кожної вибірки та проведено попарні z-тести (табл. 2).

Для експериментів з різним математичним сподіванням збереглася попередня тенденція — хаотичні функції показують статистично значуще покращення у порівнянні з сигнумом, проте диференційовна модифікація демонструє кращі результати.

В експериментах з однаковим математичним сподіванням усі функції продемонстрували схожі результати і не мають статистично значущої різниці.

№	Назва	Параметри	z-стат	p-знач
1	sign_vs_izh	$\mu_1 = 0.33, \mu_2 = 0.69, \sigma_1^2 = \sigma_2^2 = 0.05$	-5.6564	0.0001
2	sign_vs_bezier	$\mu_1 = 0.33, \mu_2 = 0.69, \sigma_1^2 = \sigma_2^2 = 0.05$	-5.4406	0.0001
3	izh_vs_bezier	$\mu_1 = 0.33, \mu_2 = 0.69, \sigma_1^2 = \sigma_2^2 = 0.05$	-1.4139	0.1574
4	sign_vs_izh	$\mu_1 = 0.47, \mu_2 = 0.58, \sigma_1^2 = \sigma_2^2 = 0.05$	-3.0115	0.0026
5	sign_vs_bezier	$\mu_1 = 0.47, \mu_2 = 0.58, \sigma_1^2 = \sigma_2^2 = 0.05$	-3.4758	0.0005
6	izh_vs_bezier	$\mu_1 = 0.47, \mu_2 = 0.58, \sigma_1^2 = \sigma_2^2 = 0.05$	-0.3991	0.6898
7	sign_vs_izh	$\mu_1 = \mu_2 = 0.5, \sigma_1^2 = 0.05, \sigma_2^2 = 0.1$	-1.6944	0.0902
8	sign_vs_bezier	$\mu_1 = \mu_2 = 0.5, \sigma_1^2 = 0.05, \sigma_2^2 = 0.1$	-1.3373	0.1811
9	izh_vs_bezier	$\mu_1 = \mu_2 = 0.5, \sigma_1^2 = 0.05, \sigma_2^2 = 0.1$	0.9461	0.3441
10	sign_vs_izh	$\mu_1 = \mu_2 = 0.5, \sigma_1^2 = 0.05, \sigma_2^2 = 0.067$	-1.2368	0.2162
11	sign_vs_bezier	$\mu_1 = \mu_2 = 0.5, \sigma_1^2 = 0.05, \sigma_2^2 = 0.067$	-1.5615	0.1184
12	izh_vs_bezier	$\mu_1 = \mu_2 = 0.5, \sigma_1^2 = 0.05, \sigma_2^2 = 0.067$	-0.4817	0.6300

ТАБЛ. 2. Результати z-тестів для експериментів з функціями активацій.

3. ЕКСПЕРИМЕНТ З ФРАКТАЛЬНОГО АНАЛІЗУ ЯДЕР БУКАЛЬНОГО ЕПІТЕЛІУ

Тепер розглянемо ефективність хаотичних функцій активації в експерименті з реальними даними.

Одні з перших повідомлень про злякисні зміни з'явилися в 1960-х роках, коли широко вивчався вміст Х-хроматину в соматичних клітинах і стала очевидною його нестабільність при різних функціональних змінах в організмі і загальній патології соматичних клітин.

При наявності пухлин в організмі спостерігаються очевидні зміни вмісту Х-хроматину в букальному епітелії та нейтрофілах периферичної крові.

Також було показано, що зміни кількості клітин з Х-хромосомами викликані дефектом функціонального стану гетерозиготної Х-хромосоми [22].

Значний інтерес викликають дослідження, що демонструють зміни в клітинах букального епітелію у пацієнтів з пухлинами. У 1960-х роках Н. Nieburgs та співавтори [23] повідомили про характерний перерозподіл маси хроматину в соматичних клітинах у 77% хворих на рак і назвали ці зміни пухлинними. Ці зміни характеризувалися збільшенням розмірів ядер епітеліальних клітин і збільшенням розмірів «обмежених» ділянок хроматину, оточених світлими ділянками [24]. Такі ж зміни спостерігалися в клітинах печінки, нирок та інших органів. У хворих на рак молочної залози виявлено збільшення вмісту ДНК у букальному епітелії та розміру інтерфазних ядер.

Однак деякі автори не виявили достовірної різниці за цим показником між хворими і практично здоровими чоловіками, коли вміст ДНК клітин букального епітелію у чоловіків з бронхіальною епітеліомою вимірювали методом клітинної спектрофотометрії [25].

У дослідженні [26] було обстежено контрольну групу (29 пацієнток), групу пацієнток з раком молочної залози II стадії (68 пацієнток) та групу пацієнток з фіброаденоматозом (33 пацієнтки). Всі діагнози були підтверджені

гістологічно. Морфологічний набір даних складався з 20256 зображень інтерфазних ядер буккального епітелію (6752 ядра були відскановані в трьох варіантах — без фільтрації, з жовтим фільтром і фіолетовим фільтром).

Морфологічним матеріалом слугували мазки епітеліальних клітин слизової оболонки порожнини рота в середній глибині шипуватого шару. В середньому кожен препарат складався з 52 клітин. Вміст ДНК-фуксину в ядрах епітеліальних клітин розраховували як добуток оптичної щільності на площу. На першому етапі аналізу отримували зображення розподілу хроматину у вигляді матриці 128 x 128 пікселів [26].

Щоб відобразити фрактальну природу розподілу хроматину та забезпечити інваріантність до обертання зображення, вздовж кожного пікселя зображення було створено просторову криву заповнення, а значення кольорів RGB пікселів зображення зчитувалися послідовно, а не по рядках. В результаті матрицю пікселів можна зіставити з трьома векторами, що відповідають трьом каналам колірної моделі RGB. Як криві заповнення простору використано криві Серпінського [26].

Для того, щоб застосувати метод фрактального аналізу зображень, зображення необхідно попередньо обробити. Для цього в роботі [26] було застосовано метод Отсу. Цей метод використовується для порогової бінаризації напівтонових зображень. Алгоритм припускає, що на зображенні є два класи пікселів (основний і фоновий), і шукає оптимальне порогове значення, яке розділяє пікселі на два класи так, щоб внутрішньокласова дисперсія була мінімальною.

Існує кілька методів обчислення фрактальної розмірності зображення. У [26] було обрано показник Херста, тому що він дуже підходить для послідовного аналізу. Показник Херста пов'язаний з фрактальною розмірністю D формулою $H = 2 - D$.

Вхідними даними для моделі стали трьохканальні (RGB) вибірки фрактальних розмірностей ядер для кожного пацієнта. Вибірki істотно відрізнялися за кількістю елементів. Тому при підготовці даних перед навчанням нейронної мережі для кожної вибірки було обчислено n квантилів, де n — кількість елементів у найменшій з вибірок (рис. 4).

З вхідних трьохканальних RGB даних були створені додаткові набори даних. В експериментах були використані наступні вибірки: RGB (три канали одночасно), R (тільки червоний), G (тільки зелений), B (тільки синій), Grey (сірий канал обчислений за формулою $0.299 * R + 0.587 * G + 0.114 * B$), Mean (середнє арифметичне трьох каналів).

Для побудови тренувальних образів для мережі Хопфілда було застосовано два способи агрегації квантилів — середнє арифметичне та медіана.

На цьому кроці тренувальні образи представляли з себе послідовності з дійсних чисел.

Нейрони мережі Хопфілда мають два значення: 1 або -1. Тому до образів було застосовано додаткову бінаризацію: дійсні значення округлювалися до дев'ятого знаку після коми, перераховувалися в двійкову систему числення і конкатенувалися в фінальний бінарний портрет (рис. 5).

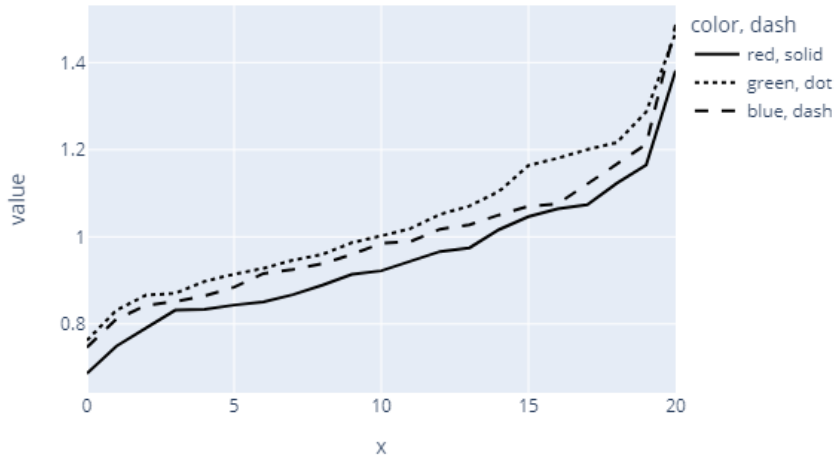


Рис. 4. Приклад узагальнення навчальних вибірок через квантилі.

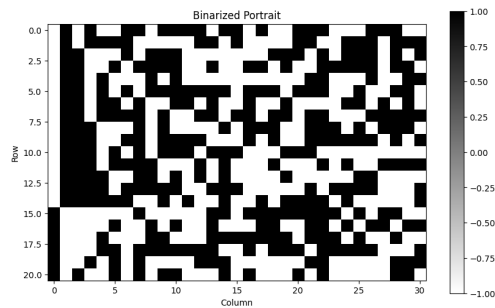


Рис. 5. Приклад підготовленого тренувального образу.

Усього було поставлено 144 експериментів, вони поділялися на 4 категорії: хворі на рак проти здорових; хворі на рак проти здорових та хворих на фіброаденоматоз; хворі на рак проти хворих на фіброаденоматоз; хворі на фіброаденоматоз проти здорових.

На тренувальних даних будувалися тренувальні образи відповідно до кожної категорії бінарної класифікації, якість моделей оцінювалися на тестових даних за допомогою Leave One Out – валідації, цільовими метриками були влучність (precision), повнота (recall) та F1-міра.

В порівнянні «хворі на рак проти хворих на фіброаденоматоз» диференційовна модифікація показала кращі результати в експериментах з найкращими результатами, медіана виявилася кращою агрегаційною функцією за середнє арифметичне. Найбільшу точність дали вхідні дані у форматі RGB та одноканальні аналоги R, G та B (табл. 3).

№	Вхідні дані	Агрегація	Ф. актив	PRECISION	RECALL	F1
1	B	avg	bezier	0.6389	0.3382	0.4423
2	B	avg	izh	0.6316	0.3529	0.4528
3	B	avg	sign	0.6750	0.3971	0.5000
4	B	median	bezier	0.8077	0.6176	0.7000
5	B	median	izh	0.8269	0.6324	0.7167
6	B	median	sign	0.8077	0.6176	0.7000
7	G	avg	bezier	0.6538	0.5000	0.5667
8	G	avg	izh	0.6400	0.4706	0.5424
9	G	avg	sign	0.6538	0.5000	0.5667
10	G	median	bezier	0.8727	0.7059	0.7805
11	G	median	izh	0.8679	0.6765	0.7603
12	G	median	sign	0.8727	0.7059	0.7805
13	gray	avg	bezier	0.7292	0.5147	0.6034
14	gray	avg	izh	0.7347	0.5294	0.6154
15	gray	avg	sign	0.7292	0.5147	0.6034
16	gray	median	bezier	0.7347	0.5294	0.6154
17	gray	median	izh	0.7059	0.5294	0.6050
18	gray	median	sign	0.7347	0.5294	0.6154
19	mean	avg	bezier	0.6667	0.4706	0.5517
20	mean	avg	izh	0.6600	0.4853	0.5593
21	mean	avg	sign	0.6735	0.4853	0.5641
22	mean	median	bezier	0.7143	0.5882	0.6452
23	mean	median	izh	0.7193	0.6029	0.6560
24	mean	median	sign	0.7143	0.5882	0.6452
25	R	avg	bezier	0.6800	0.5000	0.5763
26	R	avg	izh	0.6596	0.4559	0.5391
27	R	avg	sign	0.5429	0.2794	0.3689
28	R	median	bezier	0.8636	0.8382	0.8507
29	R	median	izh	0.8636	0.8382	0.8507
30	R	median	sign	0.8636	0.8382	0.8507
31	RGB	avg	bezier	0.7333	0.4853	0.5841
32	RGB	avg	izh	0.6923	0.3971	0.5047
33	RGB	avg	sign	0.6667	0.3529	0.4615
34	RGB	median	bezier	0.9474	0.7941	0.8640
35	RGB	median	izh	0.9310	0.7941	0.8571
36	RGB	median	sign	0.9474	0.7941	0.8640

ТАБЛ. 3. Результати порівняльного експерименту з фрактального аналізу ядер букального епітелію (хворі на рак проти хворих на фіброаденоматоз).

В порівнянні «хворі на рак проти здорових» диференційовна модифікація продемонструвала схожі результати з хаотичною функцією Іжикевича. Обидві хаотичні функції дали кращі результати за сігнум в більшості

ХАОС У МОДЕЛЯХ ХОПФІЛДА

№	Вхідні дані	Агрегація	Ф. актив	PRECISION	RECALL	F1
1	B	avg	bezier	0.8980	0.6471	0.7521
2	B	avg	izh	0.9000	0.6618	0.7627
3	B	avg	sign	0.8980	0.6471	0.7521
4	B	median	bezier	0.8364	0.6765	0.7480
5	B	median	izh	0.8364	0.6765	0.7480
6	B	median	sign	0.8364	0.6765	0.7480
7	G	avg	bezier	0.6591	0.4265	0.5179
8	G	avg	izh	0.6744	0.4265	0.5225
9	G	avg	sign	0.6591	0.4265	0.5179
10	G	median	bezier	0.7347	0.5294	0.6154
11	G	median	izh	0.7174	0.4853	0.5789
12	G	median	sign	0.7347	0.5294	0.6154
13	gray	avg	bezier	0.7500	0.5735	0.6500
14	gray	avg	izh	0.7692	0.5882	0.6667
15	gray	avg	sign	0.7500	0.5735	0.6500
16	gray	median	bezier	0.7059	0.5294	0.6050
17	gray	median	izh	0.7000	0.5147	0.5932
18	gray	median	sign	0.7059	0.5294	0.6050
19	mean	avg	bezier	0.7556	0.5000	0.6018
20	mean	avg	izh	0.7660	0.5294	0.6261
21	mean	avg	sign	0.7556	0.5000	0.6018
22	mean	median	bezier	0.7917	0.5588	0.6552
23	mean	median	izh	0.7917	0.5588	0.6552
24	mean	median	sign	0.7917	0.5588	0.6552
25	R	avg	bezier	0.7347	0.5294	0.6154
26	R	avg	izh	0.7347	0.5294	0.6154
27	R	avg	sign	0.7347	0.5294	0.6154
28	R	median	bezier	0.7143	0.5147	0.5983
29	R	median	izh	0.7143	0.5147	0.5983
30	R	median	sign	0.7143	0.5147	0.5983
31	RGB	avg	bezier	0.8889	0.5882	0.7080
32	RGB	avg	izh	0.8636	0.5588	0.6786
33	RGB	avg	sign	0.8696	0.5882	0.7018
34	RGB	median	bezier	0.7959	0.5735	0.6667
35	RGB	median	izh	0.7959	0.5735	0.6667
36	RGB	median	sign	0.7959	0.5735	0.6667

ТАБЛ. 4. Результати порівняльного експерименту з фрактального аналізу ядер букального епітелію (хворі на рак проти здорових).

експериментів. Медіана та середнє арифметичне дали схожі результати в аналогічних експериментах. Найбільшу точність дали вхідні дані у форматі B та RGB (табл. 4).

№	Вхідні дані	Агрегація	Ф. актив	PRECISION	RECALL	F1
1	B	avg	bezier	0.6207	0.5294	0.5714
2	B	avg	izh	0.6000	0.5294	0.5625
3	B	avg	sign	0.6316	0.5294	0.5760
4	B	median	bezier	0.5857	0.6029	0.5942
5	B	median	izh	0.5857	0.6029	0.5942
6	B	median	sign	0.5857	0.6029	0.5942
7	G	avg	bezier	0.5507	0.5588	0.5547
8	G	avg	izh	0.5606	0.5441	0.5522
9	G	avg	sign	0.5507	0.5588	0.5547
10	G	median	bezier	0.5254	0.4559	0.4882
11	G	median	izh	0.5345	0.4559	0.4921
12	G	median	sign	0.5254	0.4559	0.4882
13	gray	avg	bezier	0.5588	0.5588	0.5588
14	gray	avg	izh	0.5571	0.5735	0.5652
15	gray	avg	sign	0.5588	0.5588	0.5588
16	gray	median	bezier	0.6034	0.5147	0.5556
17	gray	median	izh	0.6034	0.5147	0.5556
18	gray	median	sign	0.6034	0.5147	0.5556
19	mean	avg	bezier	0.5333	0.4706	0.5000
20	mean	avg	izh	0.5469	0.5147	0.5303
21	mean	avg	sign	0.5333	0.4706	0.5000
22	mean	median	bezier	0.6071	0.5000	0.5484
23	mean	median	izh	0.6140	0.5147	0.5600
24	mean	median	sign	0.6071	0.5000	0.5484
25	R	avg	bezier	0.5294	0.5294	0.5294
26	R	avg	izh	0.5441	0.5441	0.5441
27	R	avg	sign	0.5294	0.5294	0.5294
28	R	median	bezier	0.5741	0.4559	0.5082
29	R	median	izh	0.5741	0.4559	0.5082
30	R	median	sign	0.5741	0.4559	0.5082
31	RGB	avg	bezier	0.5667	0.5000	0.5313
32	RGB	avg	izh	0.5738	0.5147	0.5426
33	RGB	avg	sign	0.5667	0.5000	0.5313
34	RGB	median	bezier	0.6000	0.5294	0.5625
35	RGB	median	izh	0.6066	0.5441	0.5736
36	RGB	median	sign	0.6000	0.5294	0.5625

ТАБЛ. 5. Результати порівняльного експерименту з фрактального аналізу ядер букального епітелію (хворі на рак проти здорових та хворих на фіброаденоматоз).

В порівнянні «хворі на рак проти здорових та хворих на фіброаденоматоз» усі три функції активації продемонстрували схожі результати. Так

ХАОС У МОДЕЛЯХ ХОПФІЛДА

№	Вхідні дані	Агрегація	Ф. актив	PRECISION	RECALL	F1
1	B	avg	bezier	0.6429	0.5625	0.6000
2	B	avg	izh	0.5926	0.5000	0.5424
3	B	avg	sign	0.6429	0.5625	0.6000
4	B	median	bezier	0.7353	0.7813	0.7576
5	B	median	izh	0.7429	0.8125	0.7761
6	B	median	sign	0.7353	0.7813	0.7576
7	G	avg	bezier	0.4583	0.3438	0.3929
8	G	avg	izh	0.4231	0.3438	0.3793
9	G	avg	sign	0.4400	0.3438	0.3860
10	G	median	bezier	0.4688	0.4688	0.4688
11	G	median	izh	0.4516	0.4375	0.4444
12	G	median	sign	0.4688	0.4688	0.4688
13	gray	avg	bezier	0.4800	0.3750	0.4211
14	gray	avg	izh	0.4815	0.4063	0.4407
15	gray	avg	sign	0.4615	0.3750	0.4138
16	gray	median	bezier	0.4688	0.4688	0.4688
17	gray	median	izh	0.4333	0.4063	0.4194
18	gray	median	sign	0.4688	0.4688	0.4688
19	mean	avg	bezier	0.6429	0.5625	0.6000
20	mean	avg	izh	0.6552	0.5938	0.6230
21	mean	avg	sign	0.6000	0.5625	0.5806
22	mean	median	bezier	0.4444	0.5000	0.4706
23	mean	median	izh	0.4444	0.5000	0.4706
24	mean	median	sign	0.4444	0.5000	0.4706
25	R	avg	bezier	0.6667	0.5625	0.6102
26	R	avg	izh	0.6000	0.5625	0.5806
27	R	avg	sign	0.5000	0.5625	0.5294
28	R	median	bezier	0.5200	0.4063	0.4561
29	R	median	izh	0.4815	0.4063	0.4407
30	R	median	sign	0.5200	0.4063	0.4561
31	RGB	avg	bezier	0.7083	0.5313	0.6071
32	RGB	avg	izh	0.6538	0.5313	0.5862
33	RGB	avg	sign	0.5862	0.5313	0.5574
34	RGB	median	bezier	0.6774	0.6563	0.6667
35	RGB	median	izh	0.7000	0.6563	0.6774
36	RGB	median	sign	0.6774	0.6563	0.6667

ТАБЛ. 6. Результати порівняльного експерименту з фрактального аналізу ядер букального епітелію (хворі на фіброаденоматоз проти здорових).

само нема переважаючих альтернатив серед агрегацій та формату вхідних даних (табл. 5).

В порівнянні «хворі на фіброаденоматоз проти здорових» хаотичні функції показали кращі результати за сигнум в більшості експериментів. Медіана була ефективнішою за середнє арифметичне в аналогічних експериментах.

Найбільшу точність дали вхідні дані у форматі RGB та B (табл. 6).

Найкращі результати були отримані у експериментах «Хворі на рак проти хворих на фіброаденоматоз» — середнє значення міри F1 5 найкращих експериментів дорівнює 0.84, найгірші результати було отримано у групі експериментів «Хворі на рак проти здорових та хворих на фіброаденоматоз» — середнє значення міри F1 5 найкращих експериментів дорівнює 0.57. Це можна пояснити розрізненістю другого класу — здорові та хворі на фіброаденоматоз — які при агрегації генерували образ, який модель важко відрізняла від хворих на рак.

Також медіана виявилась кращою агрегацією для тренувальних образів — значення F1-міри в середньому більше у більшості експериментів.

Найкращі формати вхідних вибірок — RGB і B — вони присутні у найкращій п'ятірці для кожної групи експериментів.

ВИСНОВКИ

У глибинному навчанні, особливо у контексті штучних нейронних мереж, використання хаосу у функціях активації відіграє значну роль у забезпеченні ефективності та гнучкості моделей.

Хаотичні функції активації дозволяють збагатити простір розподілу признаков, що дозволяє моделі адаптуватися до різноманітних умов та зберігати стійкість до шуму в даних.

Такий підхід сприяє збільшенню робастності моделей до змін в навчальних даних і може покращити їх загальну здатність до узагальнення на нові дані. Використання хаотичних функцій активації також може мати важливе значення у випадках, коли модель повинна адаптуватися до непередбачуваних чи нелінійних зв'язків у даних, що робить її більш універсальною та потужною у вирішенні складних завдань машинного навчання.

Також функції активації, які можна диференціювати у нейромережах, мають кілька важливих переваг над їх недиференційованими відповідниками. В першу чергу, вони дозволяють ефективно використовувати оптимізаційні алгоритми, такі як зворотне розповсюдження помилок, для тренування мережі. Гладка природа диференційованих функцій дозволяє знаходити оптимальні параметри за допомогою градієнтного спуску, що спрощує процес тренування та покращує швидкість збіжності.

У статті обгрунтовано та продемонстровано переваги хаотичних функцій активації моделі Хопфілда на прикладі експериментів зі штучними вибірками з нормальних розподілів та на реальних даних в прогнозуванні раку молочної залози.

Цей напрямок ще потребує подальших досліджень та експериментів, розробка та впровадження нових хаотичних функцій активації може стати перспективним шляхом удосконалення глибинного навчання не тільки в

моделях Хопфілда, а й в альтернативних алгоритмах та нейромережєвих архітектурах та покращення результатів у різноманітних задачах.

Автори висловлюють вдячність колегам Голубєвой К.М. та Бородай Н.В. за надані дані для навчання моделей.

ЛІТЕРАТУРА

1. Dubey S. R., Singh S. K., Chaudhuri B. B. Activation Functions in Deep Learning: A comprehensive Survey and Benchmark. *Neurocomputing* 2022. Vol. 503. P. 24–27.
2. Kilicarslan S., Adem K., Celik M. An overview of the activation functions used in deep learning algorithms. *Journal of New Results in Science* 2021. Vol. 10. P. 75–88.
3. Wang X., Ren H., Wang A. Smish: A Novel Activation Function for Deep Learning Methods. *Electronics* 2022. Vol. 11. P. 540–545.
4. Chen X., Wang Y. A Chaotic Neuron and its Ability to Prevent Overfitting. *Frontiers in Computing and Intelligent Systems* 2023. Vol. 5. P. 53–61.
5. Deng Q., Wong C., Lin H. Memristive Hopfield neural network dynamics with heterogeneous activation functions and its application. *Chaos Solitons and Fractals* 2024. Vol. 178. P. 101–124.
6. Deng Q., Wong C., Lin H. Chaotic dynamical system of Hopfield neural network influenced by neuron activation threshold and its image encryption. *Nonlinear Dynamics* 2024. Vol. 112. P. 1–18.
7. Magallon D., Garcia J., Huerta G., Jaimes R. Real-time neural identification using a recurrent wavelet first-order neural network of a chaotic system implemented in an FPAA. *Integration* 2024. Vol. 96. P. 384–393.
8. Krotov D. A new frontier for Hopfield networks. *Nature Reviews Physics* 2023. Vol. 5. P. 52–70.
9. Cursino C., Dias L. Hybrid Hopfield Neural Network. *SN Computer Science* 2024. Vol. 5. P. 25–74.
10. Kashyap S., Dsouza N., Shi L., Beymer D. Hopfield Encoding Networks. 2024.
11. Liang X., Yang Y., Wang R., Chen J. Synchronization of delayed stochastic reaction-diffusion Hopfield neural networks via sliding mode control. *Nonlinear Analysis Modelling and Control* 2024. Vol. 29. P. 1–19.
12. Lieberman-Aiden E. Comprehensive mapping of long-range interactions reveals folding principles of the human Genome *Science* 2009. Vol. 326. P. 289–193.
13. Babloyantz A., Nicolis S., Salazar M. Evidence of chaotic dynamics of brain activity during the sleep cycle. *Physics Letters A*. 1985. Vol. 111. P. 152–156.
14. Freeman W. Simulation of Chaotic EEG Patterns with a Dynamic Model of the Olfactory System. *Biological Cybernetics* 1987. Vol. 56. P. 139–150.
15. Hirsch M. Convergent Activation Dynamics in Continuous Time Neural Networks. *Neural Networks* 1989. Vol. 2. P. 331–351.
16. Hopfield J. Neural networks and physical systems with emergent collective computational abilities *Proc Natl Acad Sci U S A*. 1982. Vol. 79(Apr). P. 2554–2558.
17. Shepherd G. The synaptic organization of the brain. *Oxford University Press* 1990.
18. Skarda C., Freeman W. How brains make chaos in order to make sense of the world. *Behavioral and Brain Sciences* 1987. Vol. 10. P. 161–195.

-
19. Tsuda I. Dynamic link of memory - Chaotic memory map in nonequilibrium neural networks. *Neural Networks* 1992. Vol. 5. P. 857–893.
 20. Yao Y., Freeman W. Model of Biological Pattern Recognition with Spatially Chaotic Dynamics. *Neural Networks* 1990. Vol. 3. P. 153–170.
 21. Sparrow A. The Lorenz Equations. Bifurcations, Chaos, and Strange Attractors. *Applied Mathematical Sciences* 1982. Vol. 41. P. 26–50.
 22. Radak M., Lafta H., Fallahi H. Machine learning and deep learning techniques for breast cancer diagnosis and classification: a comprehensive review of medical imaging studies. *Journal of Cancer Research and Clinical Oncology* 2023. Vol. 49. P. 10473–10491.
 23. Nieburgs H. Recent progress in the interpretation of malignancy associated changes (MAC). *Acta Cytologica* 1968. Vol. 12. P. 445–453.
 24. Boroday N., Chekhun V., Golubeva E., Klyushin D. In vitro and in vivo densitometric analysis of DNA content and chromatin texture in tumor cell nuclei under the influence of a nano composite and magnetic field. *Advances in Cancer Research & Treatment* 2016. Vol. 2016. P. 1–11.
 25. Ogden G., Cowpe J., Green M. The effect of distant malignancy upon quantitative cytologic assessment of normal oral mucosa. *Cancer* 1990. Vol. 65. P. 477–480.
 26. Klyushin D., Golubeva K., Boroday N., Shervarly D. Breast Cancer Diagnosis Using Machine Learning and Fractal Analysis of Malignancy-Associated Changes in Buccal Epithelium. *Artificial Intelligence, Machine Learning, and Data Science Technologies Future Impact and Well-Being for Society* 2021. Vol. 5. P. 1–18.

Надійшла: 12.07.2024 / Прийнята: 10.11.2024