

УДК 004.8

Мустаєв Т.В., Мазурець О.В., Молчанова М.О.

Хмельницький національний університет

АНАЛІЗ ВПЛИВУ ПАРАМЕТРІВ АЛГОРИТМУ НАВЧАННЯ НА ПОКАЗНИКИ ДИНАМІКИ РОБОТИ НЕЙРОННОЇ МЕРЕЖІ

Робота присвячена аналізу впливу параметрів алгоритму навчання на показники динаміки роботи нейронної мережі. Була виконана розробка дослідницької програмної системи та її використання для дослідження впливу параметрів алгоритму навчання на показники динаміки роботи нейронної мережі перцептрон у задачах розпізнавання образів. Встановлено, що вплив коефіцієнту навчання на показники динаміки роботи нейронної мережі перцептрон є відносно низьким як з точки зору ефективності, так і з часу виконання процесів, проте при значеннях, що близькі до 1, точність розпізнавань образів дещо знижується; що стосується кількості навчальних ітерацій, то при збільшенні їх значення точність роботи мережі хоч і не завжди, але в загальному зростає, час навчання моделі перцептрон також збільшується. В загальному, навіть враховуючи зміни параметрів навчання, ефективність роботи мережі перцептрон залишалась високою: від 80 до 100%.

Research is devoted to analysis of influence of parameters of training algorithm on indicators of neural network dynamics. The research software system was developed and used to study the influence of the parameters of the training algorithm on the indicators of the dynamics of the perceptron neural network in pattern recognition tasks. It was established that the influence of the training coefficient on the indicators of the dynamics of the perceptron neural network is relatively low both in terms of efficiency and in terms of the execution time of the processes, however, at values close to 1, the accuracy of pattern recognition decreases somewhat; as for the number of training iterations, when their value increases, the accuracy of the network, although not always, but in general, increases, the training time of the perceptron model also increases. In general, even taking into account changes in the training parameters, the efficiency of the perceptron network remained high: from 80 to 100%.

Станом на сьогодні, комп'ютерні технології стрімко розвиваються та є однією з визначальних складових існування людського суспільства. Не останню роль у цьому відіграє така галузь, як штучний інтелект (ШІ), завдяки якій ряд важливих задач було полегшено, автоматизовано, виведено на новий рівень тощо. Наприклад, ШІ допомагає у розв'язанні різних завдань з розпізнавання та аналізу зображень, прийняття рішень, обробленні текстової інформації тощо.

Імітація людського мозку та його структури є основним підходом ШІ, з розвитком якого виникли нейронні мережі. Однією з найперших та найперших була

і залишається мережа під назвою «перцептрон» – комп'ютерна чи математична модель сприйняття мозком інформації [1]. Ця нейромережа не є надто комплексною у своїй архітектурі у порівнянні з наслідниками, але, незважаючи на це, може навчатися і виконувати ряд складних задач [2].

Перцептрон використовують для розв'язання різного спектру завдань, але однією з основних галузей є розпізнавання образів [3]. Використання саме цієї нейронної мережі у згаданій сфері є досить хорошим рішенням, якщо враховувати відношення складності архітектури та її реалізації до отриманих результатів та можливостей використання, вдосконалення [4, 5].

Після аналізу перелічених вище відомостей, актуальною темою для дослідження є вплив параметрів алгоритму навчання на показники динаміки роботи нейронної мережі перцептрон. Проведене дослідження дозволить краще охарактеризувати динаміку та принципи роботи згаданої мережі, а також з певною імовірністю знайти межі нормального функціонування.

Метою роботи є розробка дослідницької програмної системи та її використання для дослідження впливу параметрів алгоритму навчання на показники динаміки роботи нейронної мережі перцептрон у задачах розпізнавання образів. Проведене дослідження допоможе краще зрозуміти нейромережу та її будову, надасть можливість для збільшення ефективності шляхом можливих змін, модифікацій.

Для дослідження впливу параметрів алгоритму навчання нейронної мережі перцептрон на показники динаміки роботи нейронної мережі перцептрон проведено ряд тестувань з використанням програмного модулю автоматичного тестування розробленої інформаційної системи.



Рисунок 1 – Сформований набір застосованих зображень цифри «сім»

Щоб дослідження було більш об'єктивним, прийнято рішення використовувати фіксовану кількість зображень у наборах для навчання та розпізнавання – 20 (фактично це ще й кількість розпізнавань), та самі зображення також будуть фіксованими. Для цього було завантажено датасет, що містить зображення цифр з двох інших популярних дата сетів: chars74k та mnist [34]. Точніше, будуть використані зображення цифр 7, 8, та 9 у стилях комп'ютерних шрифтів для навчання та цифри 7 для розпізнавання (рисунки 1, 2, 3).



Рисунок 2 – Сформований набір застосованих зображень цифри «вісім»



Рисунок 3 – Сформований набір застосованих зображень цифри «дев'ять»

Спочатку було проведено кілька тестувань із сталою кількістю ітерацій – 10000, щоб дослідити вплив саме коефіцієнту навчання на показники динаміки роботи перцептронну: спершу, було обрано низький коефіцієнт навчання – 0.2. Результати тестування наведено на рисунку 4.

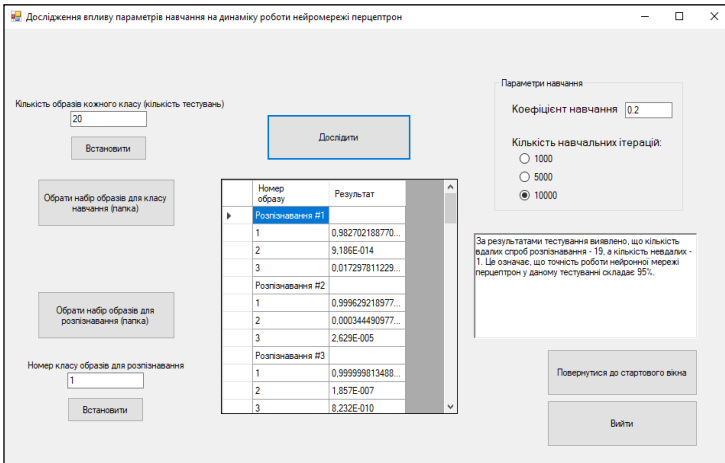


Рисунок 4 – Результати тестування з коефіцієнтом навчання 0.2

Далі коефіцієнт було збільшено до 0.5. Відповідні результати можна побачити на рисунку 5.

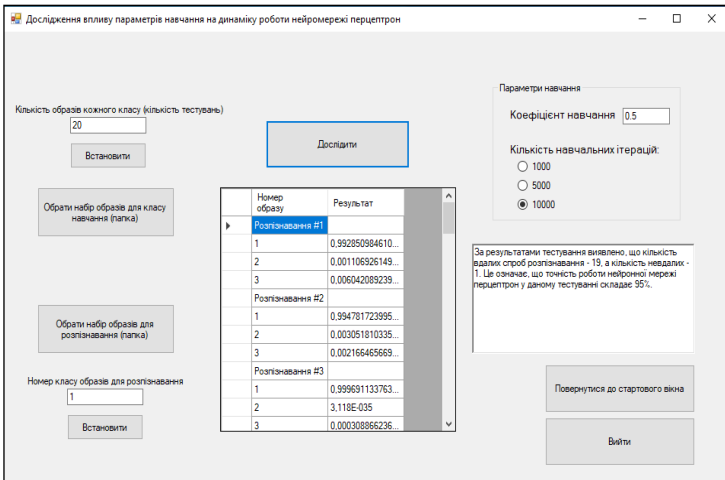


Рисунок 5 – Результати тестування з коефіцієнтом навчання 0.5

Наостанок, значення коефіцієнту встановлено за 0.8. Результати даного тестування наведено на рисунку 6.

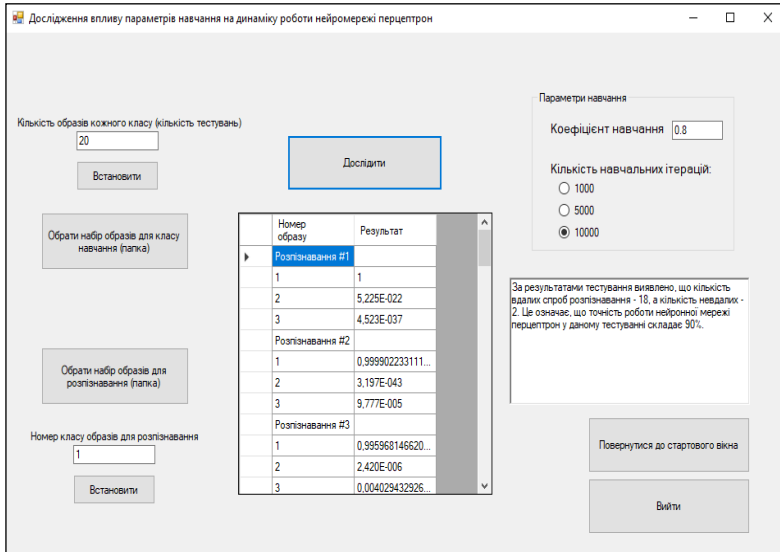


Рисунок 6 – Результати тестування з коефіцієнтом навчання 0.8

Для підсумовування результатів сформовано відповідну таблицю, що містить дані про проведені тестування з різними коефіцієнтами навчання (таблиця 1).

Таблиця 1 – підсумки тестувань з коефіцієнтами навчання

Номер тестування	Коефіцієнт навчання	Точність роботи нейромережі (%)
1	0.2	95
2	0.5	95
3	0.8	90

У підсумку можна побачити, що в цілому коефіцієнт навчання має досить низький вплив на роботу мережі перцептрон, проте при значеннях, що наближаються до 1, ефективність розпізнавань може трохи знизитись. Також варто зазначити, що час виконання для кожного тестування був приблизно однаковим – від 10 до 11 хвилин, що також є аргументом до тези низького впливу розглядуваного параметру навчання на показники динаміки роботи мережі.

Наступним кроком є дослідження впливу кількості ітерацій на показники динаміки роботи моделі нейромережі перцептрон – для цього у всіх тестуваннях обрано низький коефіцієнт навчання – 0.2. Для першого тесту встановлено кількість ітерацій рівну 1000. Його результати наведено нижче (рисунок 7).

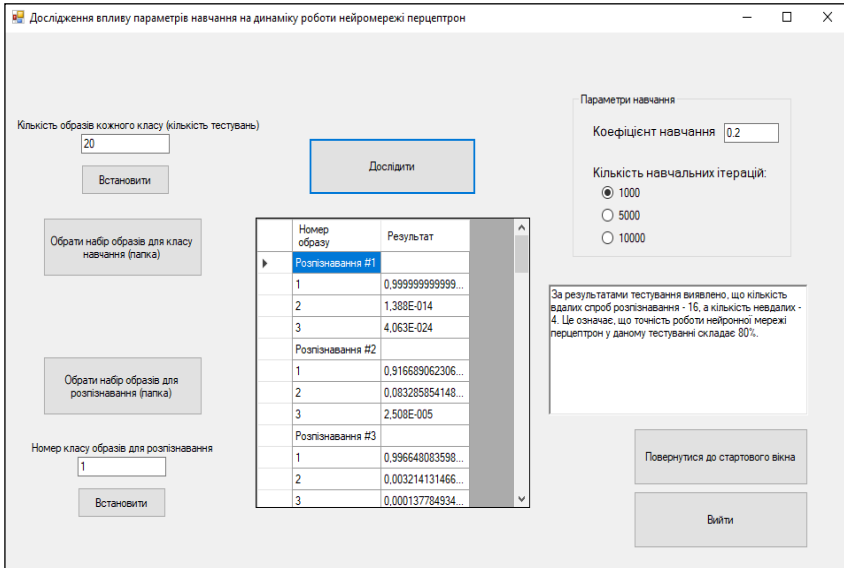


Рисунок 7 – Результати тестування з кількістю ітерацій 1000

У другому тесті кількість ітерацій збільшено до 5000. Кінцеві дані можна побачити на рисунку 8.

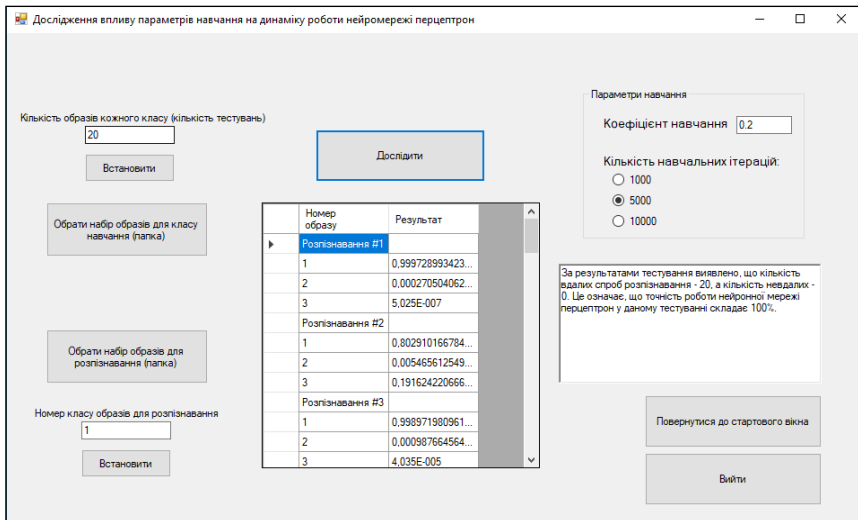


Рисунок 8 – Результати тестування з кількістю ітерацій 5000

Для останнього тесту кількість ітерацій – 10000. Результати відображено на рисунку 9.

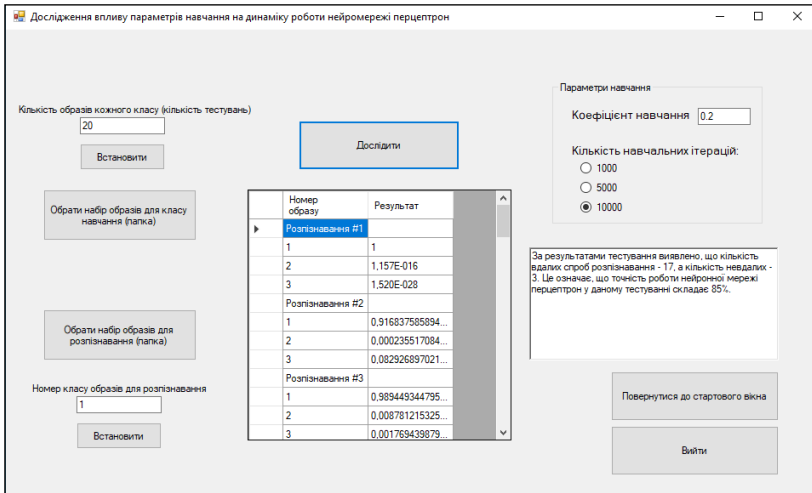


Рисунок 9 – Результати тестування з кількістю ітерацій 10000

З метою підбиття підсумків сформовано таблицю, що містить всі дані про проведені тестування з різною кількістю ітерацій (таблиця 2).

Таблиця 2 – Підсумки тестувань з кількостями ітерацій

Номер тестування	Кількість навчальних ітерацій	Точність роботи нейромережі (%)	Час виконання (у хвиликах)
1	1000	80	1.5
2	5000	100	5.5
3	10000	85	10.5

Таким чином можна спостерігати, що час виконання прямопропорційно залежить від кількості ітерацій, крім того, при збільшенні кількості ітерацій ефективність роботи мережі перцептрон збільшилась: хоча з таблиці видно, що за 5000 ітерацій точність більша, ніж за 10000, проте зважаючи на тестування, проведені з коефіцієнтом навчання, де для всіх випадків кількість ітерацій була 10000, а саме на те, що точність розпізнавань перебувала на рівні 90% і більше, то у даному випадку це можна віднести до виключення із загальної тенденції.

Отже, з проведених досліджень можна зробити висновки, що вплив коефіцієнту навчання на показники динаміки роботи нейронної мережі перцептрон є відносно низьким як з точки зору ефективності, так і з часу виконання процесів, проте при значеннях, що близькі до 1, точність розпізнавань образів дещо знижується; що стосується кількості навчальних ітерацій, то при збільшенні їх значення точність роботи мережі хоч і не завжди, але в загальному зростає, час навчання моделі перцептрону також збільшується. В загальному, навіть враховуючи зміни параметрів навчання, ефективність роботи мережі перцептрон залишалась високою: від 80 до 100%.

Перелік посилань

1. Молчанова М. О. Метод квазіоптимального підбору параметрів навчання нейронної мережі неокогнітрон у задачах класифікації плоских стаціонарних об'єктів із великою кількістю бінарних ознак. Збірник наукових праць за матеріалами другої науково-технічної конференції студентів, магістрів та аспірантів «Інформатика, керування та штучний інтелект» – Харків: ХПІ, 2015.
2. Kharysh I., Sobko O., Mazurets O. Designing CNN Neural Network Model for Detecting Fractures of Lower Extremities by X-ray Images. The Impact of Scientific Research on the Development of the Modern World. Proceedings of the XLIV International scientific and practical conference. Dubrovnik, Croatia. 2024. Pp. 91-96. .
3. Ткачук Б.О., Мазурець О.В., Молчанова М.О., Собко О.В. Метод автоматизованого прийняття рішень щодо керуванням ігровим персонажем з використанням штучної нейронної мережі перцептрон. Матеріали III Всеукраїнської науково-технічної конференції «Комп'ютерні ігри та мультимедіа як інноваційний підхід до комунікації - 2023». Одеса. 28-29 вересня 2023. С. 223-225.
4. Молчанова М. О. Розробка системи розпізнавання психоемоційного стану людини на основі класифікатора нейронної мережі типу багат шаровий перцептрон. Збірник наукових праць за матеріалами восьмої міжнародної науково-технічної конференції «Актуальні проблеми комп'ютерних технологій 2014» – Хмельницький: ХНУ, 2014. – С.257–262.
5. Молчанова М. О. Дослідження розпізнавання плоских стаціонарних об'єктів в умовах розв'язання задачі класифікації із великою кількістю схожих бінарних ознак. Збірник наукових праць за матеріалами дев'ятої міжнародної науково-технічної конференції «Актуальні проблеми комп'ютерних технологій 2015» – Хмельницький: ХНУ, 2015. – С.189 – 194.