

## ВПЛИВ ЗМЕНШЕННЯ РОЗМІРІВ НЕЙРОННОЇ МЕРЕЖІ НА ЇЇ ЗДАТНІСТЬ ДО УЗАГАЛЬНЕННЯ

*Анотація:* У даній науковій роботі досліджено вплив розмірів нейромережі та кількості вхідних ознак на її здатність до узагальнення. Застосовуючи аналіз на семи різних задачах класифікації з використанням різноманітних конфігурацій моделей, ми визначили, що розширення розмірів нейромережі за мінімальний обсяг, необхідний для запам'ятовування навчальної вибірки, призводить до покращення її здатності до узагальнення. Однак вплив кількості вхідних ознак виявився більш складним та залежним від конкретної задачі. Результати наголошують на важливості обережного вибору розмірів нейромережі та кількості ознак у контексті навчання за традиційними алгоритмами, а також вказують на можливість оптимізації обраної конфігурації моделі для кращого балансу між ефективністю та здатністю до узагальнення.

*Ключові слова:* нейромережа, узагальнення, розмір нейромережі, ознаки, класифікація, перенавчання.

### Вступ

При використанні алгоритмів навчання для штучних нейронних мереж, які не спрямовані явно на максимізацію загальних здатностей мережі, зменшення її розмірів, таке як зменшення кількості нейронів у прихованому шарі або кількості вхідних сигналів, часто не призводить до покращення її здатності до узагальнення. Цей результат може здатися суперечливим загальноприйнятій думці, що виключення шумових або неінформативних ознак, а також зайвих нейронів, є обов'язковим та корисним практичним кроком.

Виключення шумових ознак та зайвих нейронів зазвичай розглядається як спосіб підвищити продуктивність штучної нейронної мережі та зробити її більш ефективною. Проте в деяких випадках, коли алгоритми навчання не оптимізовані для врахування таких змін, зменшення розмірів мережі може навіть погіршити її здатність до узагальнення. Це може бути пов'язано з тим, що занадто радикальне скорочення розмірів мережі може призвести до втрати важливої інформації або зниження її виразних можливостей.

Отже, важливо враховувати конкретні умови та цілі навчання штучної нейронної мережі при вирішенні питання, чи слід зменшувати її розміри, і якщо так, то наскільки радикально. Це рішення слід приймати з урахуванням алгоритмів навчання та характеру даних, з якими працює нейронна мережа, для досягнення найкращого балансу між ефективністю та здатністю мережі до узагальнення.

### **Постановка мети та задач**

Мета дослідження - аналіз семи задач класифікації, використовуючи набори даних з репозиторію UCI KDD Database Repository, з метою визначення впливу розміру нейронної мережі та кількості незалежних ознак на помилку узагальнення.

Завдання дослідження включають наступне:

1. Визначити вплив розміру нейронної мережі на здатність моделі до узагальнення шляхом порівняння результатів для мереж різного розміру.

2. Дослідити, чи підвищення розміру мережі, яке перевищує мінімальний обсяг, необхідний для запам'ятовування навчального набору даних, покращує здатність до узагальнення.

3. Визначити, чи може надмірний розмір мережі діяти як регуляризатор моделі.

Паралельно дослідити вплив кількості незалежних ознак на помилку узагальнення, звертаючи увагу на різні варіації цього впливу в різних задачах класифікації.

Загальні висновки мають підкреслити важливість вибору оптимального розміру мережі та кількості ознак при навчанні класифікаторів, особливо для традиційних алгоритмів, які мінімізують помилку апроксимації навчальних даних, та наголосити на значенні оцінки інформативності ознак перед навчанням мережі.

### **Стратегії управління розмірами нейронних мереж для запобігання перенавчанню**

Дослідження [1,2] показали, що однією з основних причин перенавчання нейронних мереж часто є традиційні алгоритми навчання, особливо ті, які мають високу швидкість збіжності, такі як навчання за загальним градієнтом і його види, включаючи метод спряжених градієнтів. Ці алгоритми можуть бути занадто агресивними в процесі навчання, що призводить до того, що нейронна мережа починає запам'ятовувати дані навчального набору, замість того, щоб узагальнювати їх.

Для боротьби з цим ефектом часто використовують стратегію зменшення розмірів нейронної мережі до досягнення необхідної точності на навчальному наборі даних. Для зменшення розмірів мережі застосовують два основних підходи: зменшення кількості незалежних змінних, включаючи видалення неінформативних, зайвих та шумових ознак, а також зменшення кількості нейронів у прихованих шарах мережі.

Зменшення кількості незалежних змінних дозволяє мережі сконцентруватися на найважливіших ознаках даних, що сприяє кращій здатності до узагальнення. Зменшення кількості нейронів у прихованих шарах також може зменшити інформаційну потужність мережі, що допомагає уникнути перенавчання, оскільки менша кількість нейронів ефективніше фільтрує шум і викиди в даних.

Отже, керування розмірами нейронної мережі є важливою стратегією для забезпечення її здатності до узагальнення та запобігання перенавчанню, особливо при використанні швидких алгоритмів навчання.

### **Оптимізація розмірів нейронних мереж для запобігання перенавчанню та поліпшення загальної здатності**

У цій роботі ми досліджуємо фундаментальне припущення, пов'язане зі зменшенням надлишковості нейронних мереж та його впливом на загальну здатність моделей до узагальнення. Ці дослідження спрямовані на визначення моменту, коли зменшення розмірів мережі призводить до покращення здатності моделі до узагальнення, і коли обмеження розміру починає заважати досягненню бажаної помилки на навчальних даних.

Конкретніше, наша робота ставить перед собою завдання визначити, чи відповідає мінімум помилки узагальнення мінімально необхідній кількості ознак для даної задачі, чи мінімум помилки узагальнення досягається при значно більшій кількості ознак, ніж потрібно для меморизації даних. Якщо перше припущення вірне, то ми можемо оптимізувати процес зменшення кількості ознак, проводячи перевірки і аналізуючи рівень помилки узагальнення після видалення невеликої кількості ознак.

У випадку другого припущення, коли мінімум помилки узагальнення досягається при більшій кількості ознак, ми можемо автоматично виключити всі зайві ознаки для даної задачі без додаткових перевірок.

Крім того, ми розглядаємо аналогічні питання щодо кількості нейронів у прихованих шарах нейронної мережі. Це особливо актуально для багатозадачного навчання (MTL), враховуючи зауваження [4] щодо можливості перенавчання на різних етапах навчання для кожної з вирішуваних задач.

Нашою метою є проведення дослідження, яке дозволить точніше визначити оптимальні розміри нейронних мереж і ознак для різних задач машинного навчання та глибокого навчання, враховуючи їх вплив на здатність моделі до узагальнення та запобігання перенавчанню.

### **Аналіз розмірів нейронних мереж та кількості ознак у завданнях класифікації з надмірними даними**

У семи реальних задач класифікації із репозиторію UCI KDD Database Repository було виявлено надмірну кількість незалежних ознак, і це надає можливість перевірити помилки узагальнення в різних точках по вісі кількості ознак - для повного набору ознак, мінімального набору та кількох проміжних значень. Крім того, були перевірені різні значення кількості нейронів у прихованому шарі нейронної мережі - від 10 до 35 нейронів з кроком у 5 нейронів (10 нейронів виявилось достатнім для 100% правильної класифікації навчального набору даних в кожній задачі, отже, розмір мережі починався з мінімального розміру або незначно більшого).

Були використані такі набори даних як Musk (168 незалежних ознак), OptDigits (64 ознаки) та п'ять варіацій задачі Multiple features: Zer (47), Kar (64), Fou (76), Fac (216) та Pix (240 ознак). Далі на графіках горизонтально відображувалася кількість нейронів, вертикально - помилка узагальнення (відсоток неправильних відповідей на тестовому наборі даних, чим менше тим краще), отримана під час навчання мережі на 4/5 набору даних і тестування на залишку 1/5 частини.

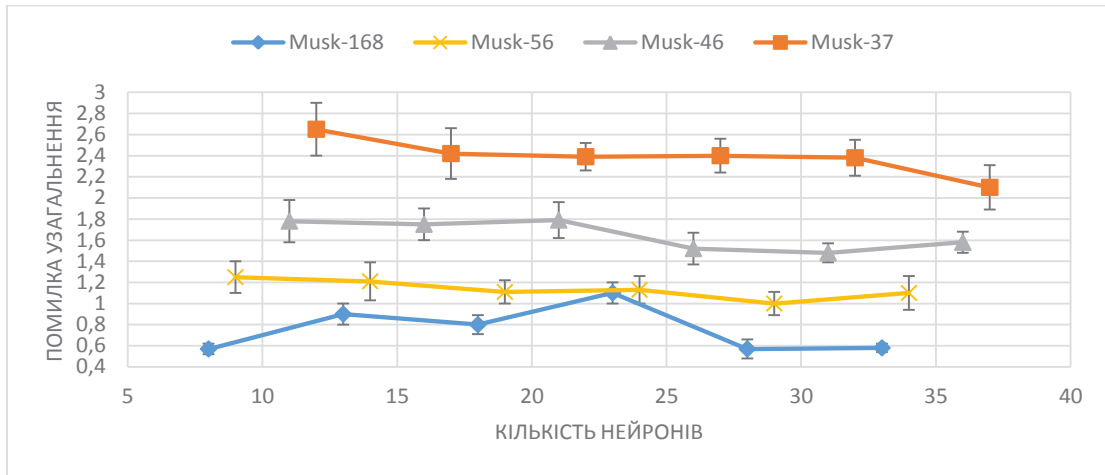


Рис. 1. Графік Musk

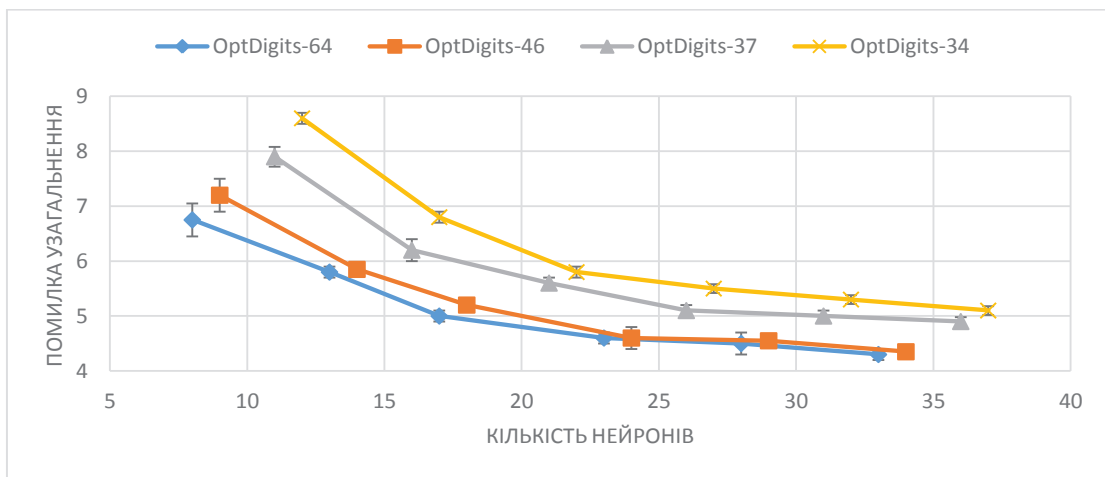


Рис. 2. Графік OptDigits

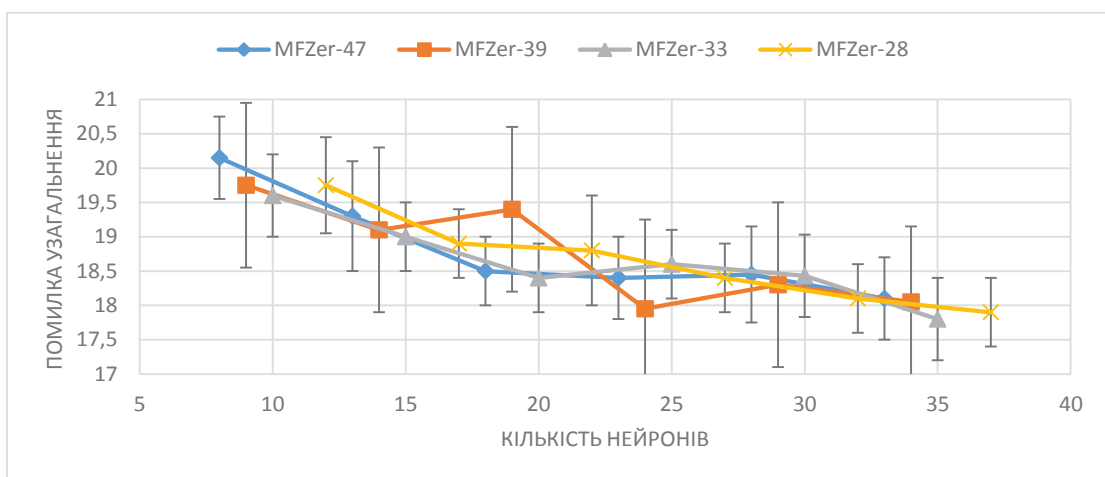


Рис. 3. Графік MFZer

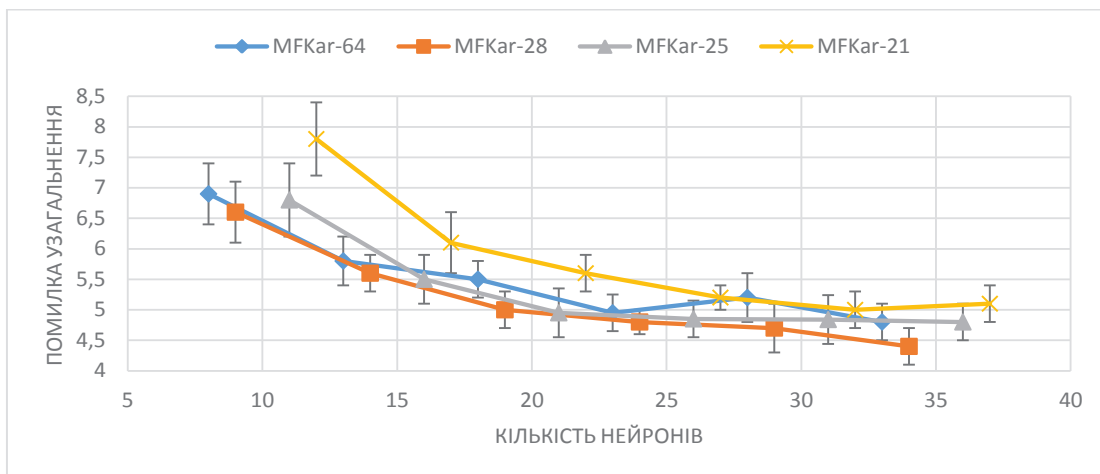


Рис. 4. Графік MFKar

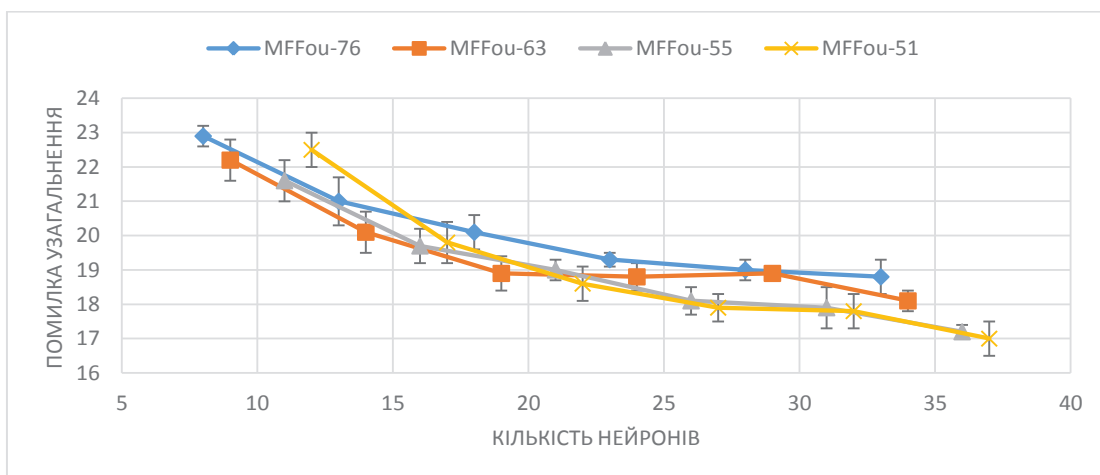


Рис. 5. Графік MFFou

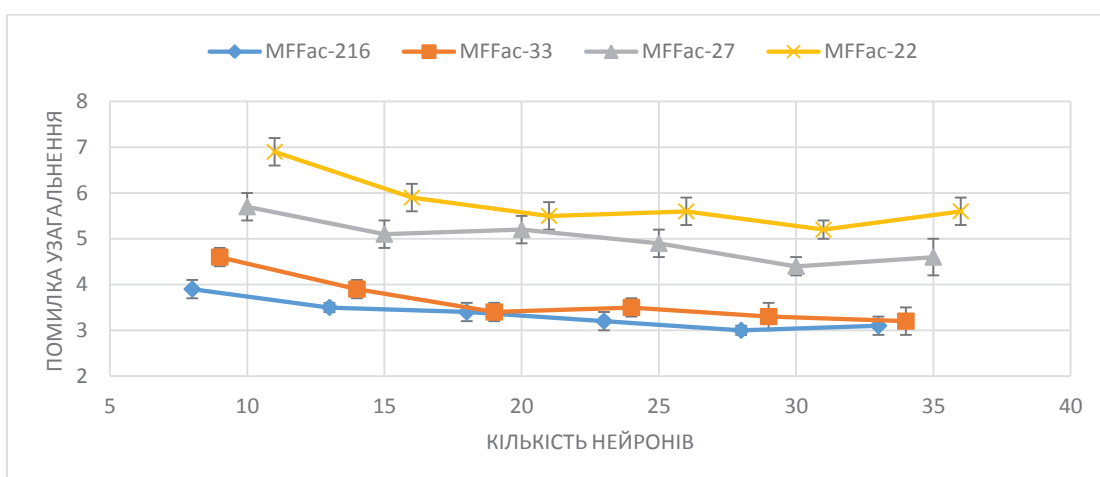


Рис. 6. Графік MFFac

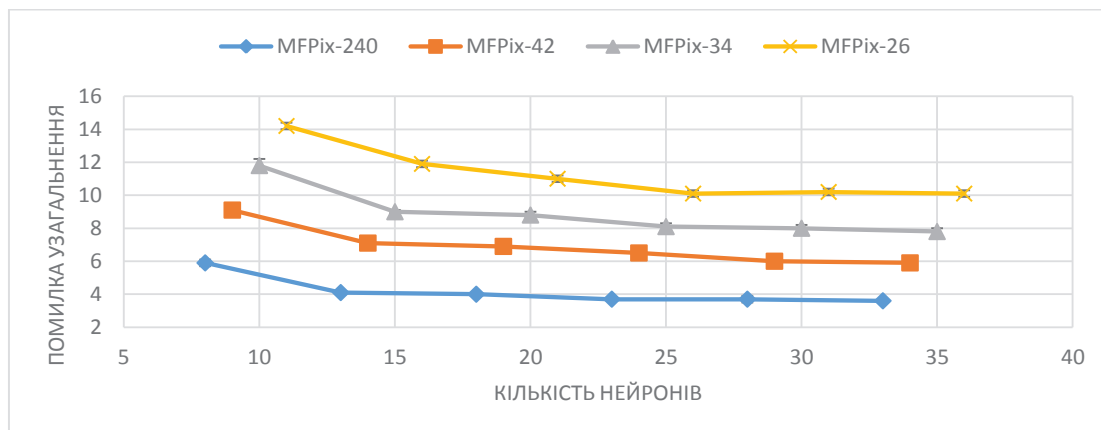


Рис. 7. Графік MFPIx

Для кожної задачі, кожного розміру мережі та кількості вхідних ознак було навчено 25 нейромереж з різними розбиттями набору даних на навчальну та тестову вибірки, і на основі цих 25 результатів обчислювалися середні значення помилок узагальнення та довірчі інтервали для них (довірчі інтервали відображалися у вигляді "вусів" навколо кожної точки графіка). У рядках підписані кількості ознак для кожного графіка.

### Аналіз результатів

У нашому дослідженні ми проаналізували сім задач класифікації, використовуючи набори даних із репозиторію бази даних UCI KDD. Основним питанням було визначення впливу розміру нейронної мережі та кількості незалежних ознак на помилку узагальнення.

Ми звернули увагу, що збільшення розміру мережі, що перевищує мінімальний обсяг, необхідний для запам'ятовування навчальної вибірки, призводить до покращення здатності моделі до узагальнення. Цей результат відповідає спостереженням, зробленим у [3]. Ефект зменшення помилки узагальнення при збільшенні розміру мережі спостерігався при будь-якій кількості вхідних сигналів, що свідчить про те, що надмірний розмір мережі діє як деякий регуляризатор моделі.

Проте ситуація з кількістю незалежних ознак є складнішою. Для різних завдань ми спостерігали різний вплив зміни кількості ознак на помилку узагальнення. У деяких завданнях навіть збільшення кількості ознак призводило до погіршення помилки узагальнення, особливо коли кількість ознак перевищувала певне критичне значення. Це явище спостерігалось у завданнях Musk і Multiple features - Pix. У інших завданнях, таких як OptDigits і Multiple features - Fac, погіршення узагальнення починалося при близькій до мінімальної кількості ознак.

В цілому, наші результати підкреслюють важливість ретельного вибору розміру нейронної мережі та кількості ознак в контексті навчання традиційними алгоритмами,

які мінімізують помилку апроксимації навчальної вибірки, але не емпіричну помилку. У деяких випадках зменшення розміру мережі або кількості ознак може погіршити емпіричну помилку, тому не завжди варто прагнути до мінімізації розмірів моделі. Оцінка інформативності ознак перед навчанням мережі може бути більш перевагоюним методом, ніж операції з нейромережею.

### СПИСОК ЛІТЕРАТУРИ

1. Fukumizu K. Effect of batch learning in multilayer neural networks // Proceedings of the 5th International Conference on Neural Information Processing (ICONIP'1998). - 1998. - С. 67-70.
2. Lawrence S., Giles C.L. Overfitting and neural networks: conjugate gradient and backpropagation // Proceedings of the International Joint Conference on Neural Networks (IJCNN'2000), Como, Italy. - 2000. - С. 114-119.
3. Caruana R.A., de Sa V.R. Benefitting from the variables that variable selection discards // Journal of Machine Learning Research. - 2003. - Том 3. - С. 1245-1264.
4. Caruana R.A. Multitask learning // Machine Learning. - 1997. - Том 28. - С. 41-75.