

ДОСЛІДЖЕННЯ ВПЛИВУ АРХІТЕКТУРНИХ ЗМІН НА ЯКІСТЬ БАГАТОШАРОВОЇ НЕЙРОННОЇ МЕРЕЖІ

Вінницький національний технічний університет

Анотація

У роботі досліджено вплив архітектурних змін багатошарової нейронної мережі на якість розв'язання задачі регресії. Проведено порівняння базової та модифікованої нейронних моделей, які відрізняються кількістю шарів та нейронів. Оцінювання виконано за середньоквадратичною похибкою та аналізом процесу навчання. Отримані результати підтверджують, що оптимізація архітектури дозволяє підвищити точність прогнозування та стабільність навчання моделі.

Ключові слова: нейронні мережі, багатошарова нейронна мережа, архітектура моделі, машинне навчання, регресія.

Abstract

The paper investigates the influence of architectural modifications of a multilayer neural network on the quality of solving a regression problem. A comparison of a basic and an improved neural model differing in the number of layers and neurons is carried out. Evaluation is performed using mean squared error and learning curve analysis. The results confirm that architectural optimization improves prediction accuracy and training stability.

Keywords: neural networks, multilayer neural network, model architecture, machine learning, regression.

Вступ

Багатошарові нейронні мережі широко застосовуються для розв'язання задач регресії, у яких необхідно встановити складні нелінійні залежності між вхідними параметрами та цільовою змінною. Якість таких моделей значною мірою залежить від вибору їх архітектури, зокрема кількості прихованих шарів, нейронів та способів їх поєднання. Невдалий вибір архітектури може призводити до недостатньої узагальнювальної здатності моделі або перенавчання, що негативно впливає на точність прогнозування.

У межах даного дослідження попередньо виконано розвідувальний аналіз даних, які містять числові та категоріальні характеристики об'єктів і відповідні значення цільової змінної. На цьому етапі було виявлено ключові ознаки, що мають найбільший вплив на формування результату, а також оцінка структури даних, їх розподілу та взаємозв'язків між окремими параметрами. Отримані результати розвідувального аналізу стали основою для подальшої побудови та навчання моделей машинного навчання.

Метою дослідження є аналіз впливу архітектурних змін багатошарової нейронної мережі на якість прогнозування. Для цього проведено порівняння базової нейронної моделі та модифікованої архітектури з розширеними можливостями апроксимації, що дозволяє оцінити доцільність використання більш складних нейронних структур для розв'язання задач регресії.

Результати дослідження

Для проведення розвідувального аналізу було обрано набір даних Auto MPG Dataset [1], який має відкритий доступ для загального використання на платформі Kaggle. Даний датасет містить інформацію про технічні характеристики автомобілів та їхні фактичні показники витрат пального, що дозволяє використовувати його для розв'язання задачі регресії. Загалом набір даних включає декілька сотень записів та містить сукупність числових і категоріальних ознак, що впливають на ефективність використання пального (рис. 1).

Основними атрибутами датасету є:

- MPG – показник витрат пального автомобіля;
- Cylinders – кількість циліндрів двигуна;
- Displacement – об'єм двигуна;

- Horsepower – потужність двигуна;
- Weight – маса автомобіля;
- Acceleration – динамічні характеристики розгону;
- Model Year – рік випуску автомобіля;
- Origin – регіон походження автомобіля;
- Car name – назва моделі автомобіля.

Перші 5 рядків датасету:

	mpg	cylinders	displacement	horsepower	weight	acceleration	model year	origin	car name
0	18.0	8	307.0	130	3504	12.0	70	1	chevrolet chevelle malibu
1	15.0	8	350.0	165	3693	11.5	70	1	buick skylark 320
2	18.0	8	318.0	150	3436	11.0	70	1	plymouth satellite
3	16.0	8	304.0	150	3433	12.0	70	1	amc rebel sst
4	17.0	8	302.0	140	3449	10.5	70	1	ford torino

Інформація про датасет:

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 398 entries, 0 to 397
Data columns (total 9 columns):
#   Column          Non-Null Count  Dtype
---  ---
0   mpg              398 non-null    float64
1   cylinders        398 non-null    int64
2   displacement     398 non-null    float64
3   horsepower       398 non-null    object
4   weight           398 non-null    int64
5   acceleration     398 non-null    float64
6   model year      398 non-null    int64
7   origin           398 non-null    int64
8   car name        398 non-null    object
dtypes: float64(3), int64(4), object(2)
memory usage: 28.1+ KB
None
```

Рис. 1. Приклад даних із датасету «Auto MPG Dataset»

У ході розвідувального аналізу було виконано попереднє очищення даних, обробку пропущених значень, а також перетворення категоріальних ознак у числовий формат (рис. 2) [2].

Колонки датасету: ['mpg', 'cylinders', 'displacement', 'horsepower', 'weight', 'acceleration', 'model year', 'origin', 'car name']

Конвертовано 'horsepower' у числовий формат.

Було видалено 6 рядків з пропущеними значеннями.

Колонку 'car name' видалено (текстова змінна).

Дані після очищення:

	mpg	cylinders	displacement	horsepower	weight	acceleration	model year	origin
0	18.0	8	307.0	130.0	3504	12.0	70	1
1	15.0	8	350.0	165.0	3693	11.5	70	1
2	18.0	8	318.0	150.0	3436	11.0	70	1
3	16.0	8	304.0	150.0	3433	12.0	70	1
4	17.0	8	302.0	140.0	3449	10.5	70	1

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 392 entries, 0 to 391
Data columns (total 8 columns):
#   Column          Non-Null Count  Dtype
---  ---
0   mpg              392 non-null    float64
1   cylinders        392 non-null    int64
2   displacement     392 non-null    float64
3   horsepower       392 non-null    float64
4   weight           392 non-null    int64
5   acceleration     392 non-null    float64
6   model year      392 non-null    int64
7   origin           392 non-null    int64
dtypes: float64(4), int64(4)
memory usage: 24.6 KB
None
```

Використовувані ознаки: ['cylinders', 'displacement', 'horsepower', 'weight', 'acceleration', 'model year', 'origin']

Рис. 2. Попередня обробка даних

Додатково проведено аналіз розподілу основних параметрів та їхнього впливу на цільову змінну, що дозволило оцінити характер взаємозв'язків між ознаками [3, 4].

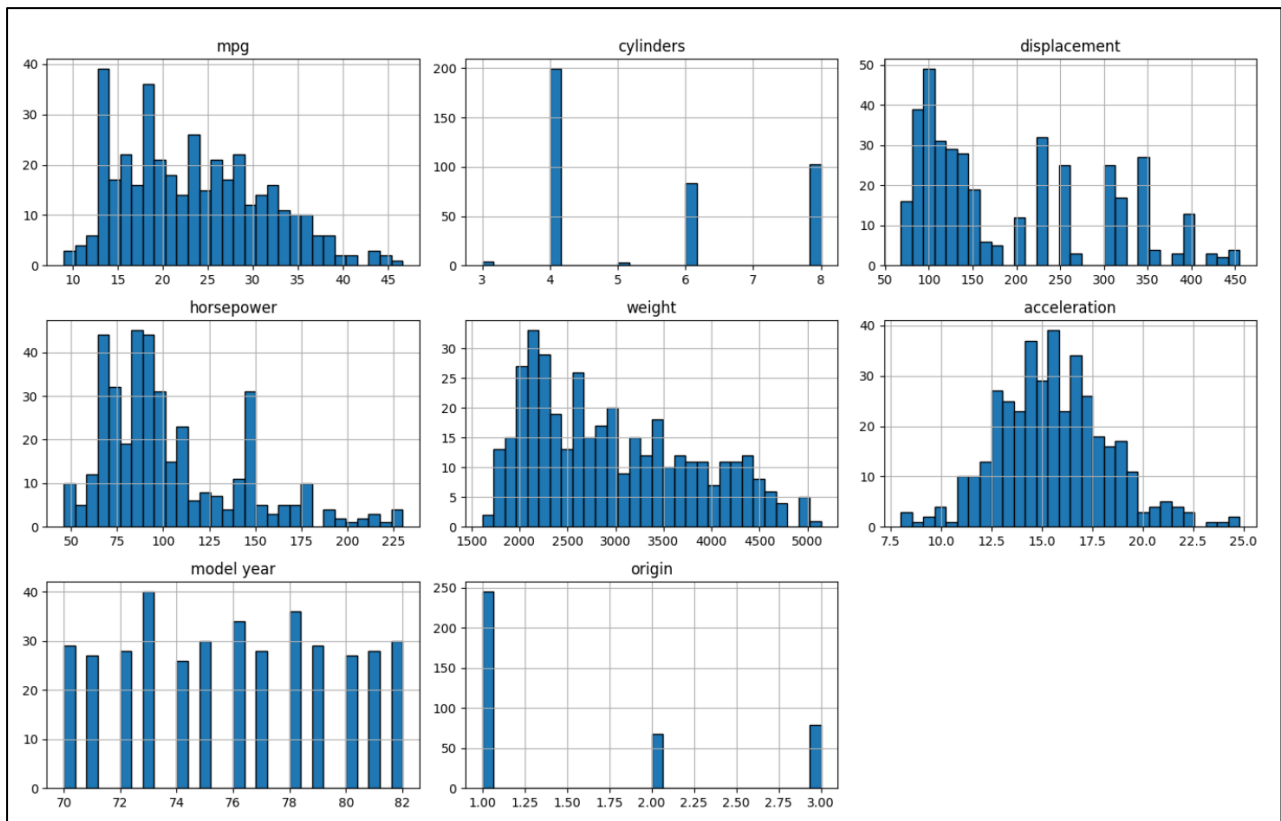


Рис. 3. Розподіл числових ознак

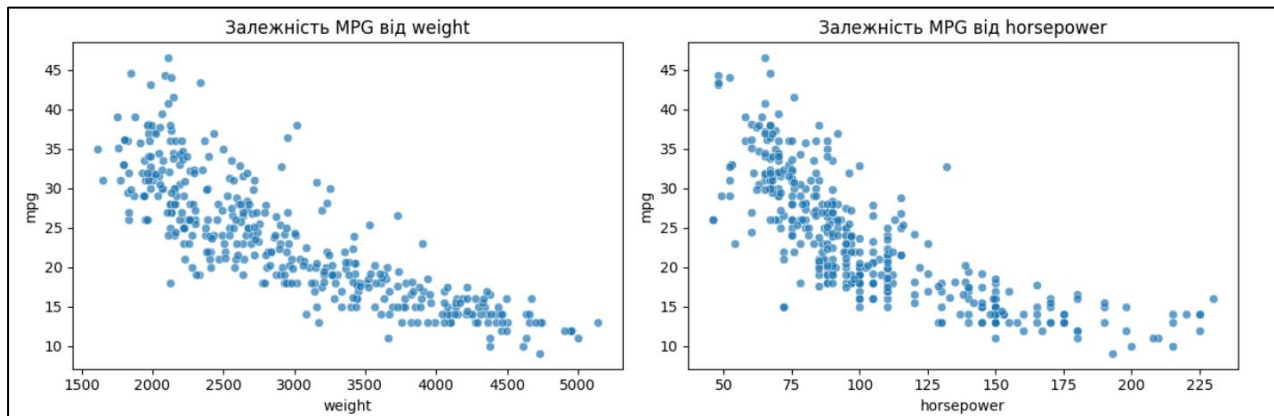


Рис. 4. Залежності цільової ознаки (MPG) від ваги та потужності двигуна машини

Результати аналізу показали наявність нелінійних залежностей між витратами пального та такими характеристиками, як маса автомобіля, потужність двигуна та кількість циліндрів. Отримані висновки підтверджують доцільність використання багатошарових нейронних мереж, здатних ефективно апроксимувати складні нелінійні функціональні залежності, а також слугують основою для подальшого дослідження впливу архітектурних змін нейронної моделі на якість прогнозування.

На основі попереднього аналізу також було побудовано кореляційну матрицю (рис. 5).

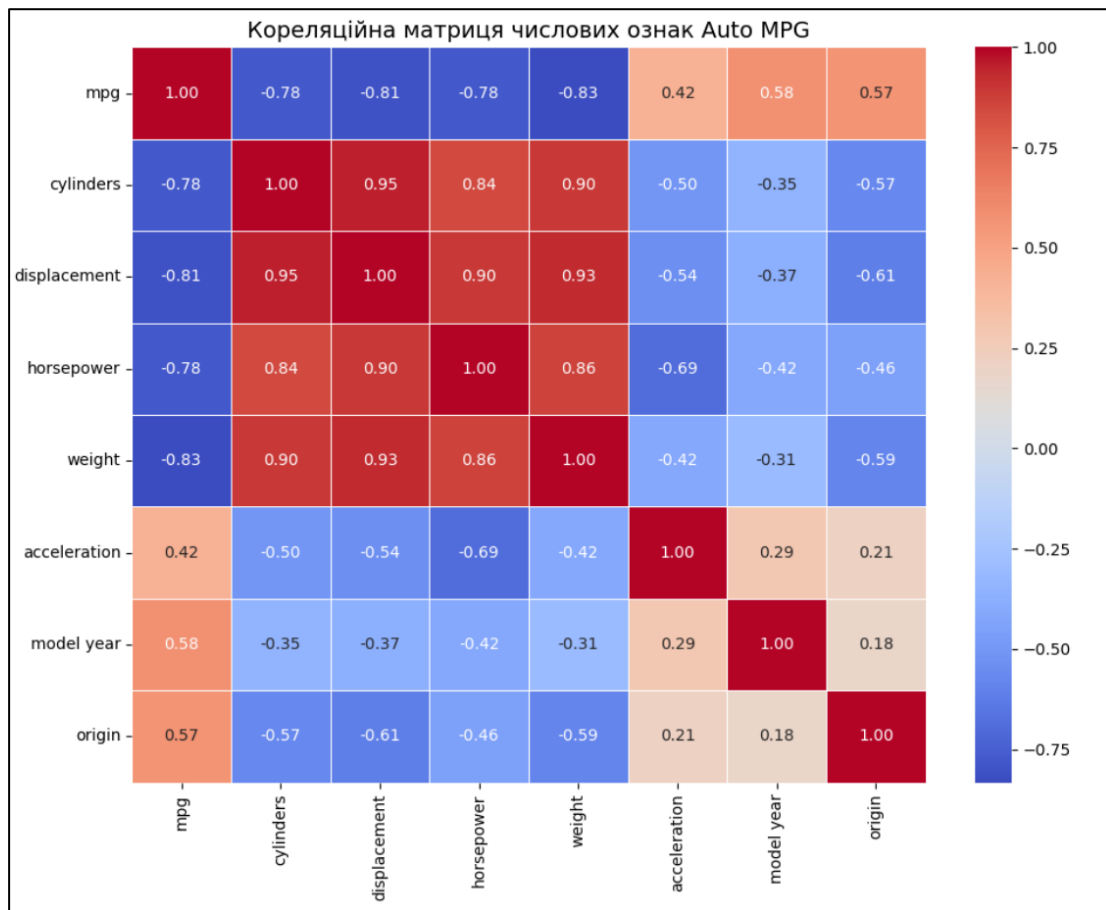


Рис. 5. Кореляційна матриця

З наведеної кореляційної матриці та отриманих залежностей можна зробити висновок, що витрати пального автомобіля (mpg) мають найбільш сильний зв'язок із такими характеристиками, як маса автомобіля, об'єм двигуна, потужність двигуна та кількість циліндрів. Це свідчить про те, що зі збільшенням габаритів і потужності транспортного засобу ефективність використання пального суттєво знижується. Водночас позитивну кореляцію з показником mpg демонструють рік випуску автомобіля та регіон походження, що може бути пов'язано з розвитком технологій і підвищенням стандартів паливної економічності у пізніші роки. Параметр прискорення має помірний зв'язок із витратами пального та іншими технічними характеристиками, що вказує на його другорядний вплив у порівнянні з масою та потужністю.

Отримані результати підтверджують наявність складних і взаємопов'язаних залежностей між технічними параметрами автомобіля, що обґрунтовує доцільність використання багатoshарових нейронних мереж для моделювання таких даних.

На основі результатів розвідувального аналізу даних було здійснено побудову простої багатoshарової нейронної моделі для розв'язання задачі регресії. Дана модель використовувалась як базовий варіант для подальшого порівняльного аналізу та оцінювання впливу архітектурних змін на якість прогнозування. Архітектура нейронної мережі складалася з вхідного шару, одного прихованого шару та вихідного нейрона, що відповідає за формування прогнозованого значення цільової змінної [5-7].

Для навчання моделі використовувався набір числових ознак, отриманих після попередньої обробки даних, включно з масштабуванням вхідних параметрів. Як функцію активації у прихованому шарі застосовано нелінійну функцію, що дозволяє моделі враховувати складні залежності між ознаками. Навчання мережі виконувалося з використанням методу зворотного поширення похибки та оптимізатора градієнтного спуску, а якість моделі оцінювалася за середньоквадратичною похибкою.

Отримані результати навчання простої моделі продемонстрували здатність нейронної мережі відтворювати загальні тенденції у даних, однак водночас виявили обмеження її представницьких

можливостей. Зокрема, для складних і нелінійних взаємозв'язків між технічними характеристиками об'єктів така архітектура може бути недостатньо гнучкою, що обґрунтовує необхідність подальшого ускладнення структури нейронної мережі. Відповідні візуалізації наведенні на рисунках 6-8

Архітектура базової моделі:
Model: "sequential_3"

Layer (type)	Output Shape	Param #
dense_10 (Dense)	(None, 64)	640
dense_11 (Dense)	(None, 32)	2,080
dense_12 (Dense)	(None, 1)	33

Total params: 2,753 (10.75 KB)
Trainable params: 2,753 (10.75 KB)
Non-trainable params: 0 (0.00 B)

Рис. 6. Вигляд архітектури базової моделі

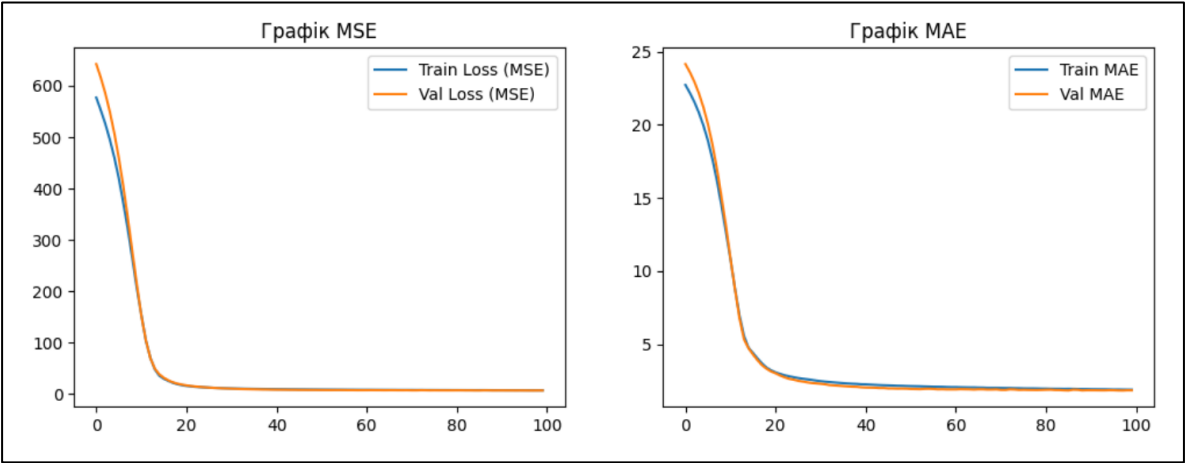


Рис. 7. Візуалізація навчання моделі

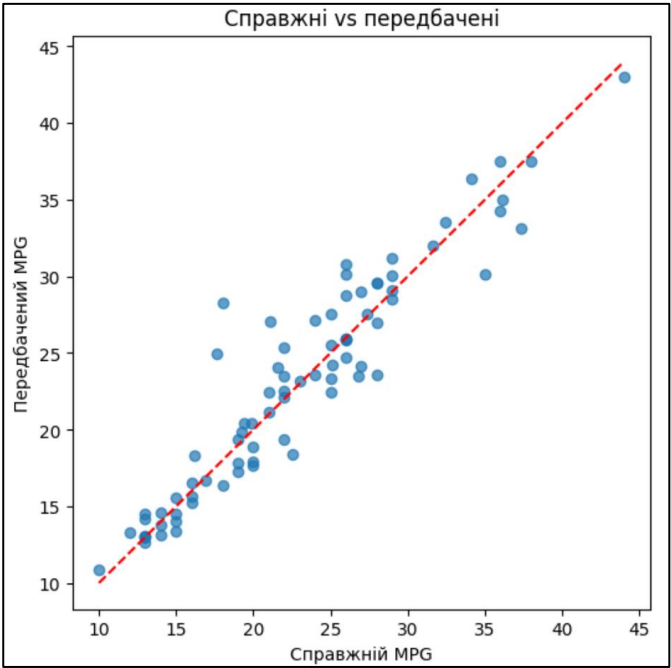


Рис. 8. Графік справжніх та передбачених значень

З метою підвищення якості прогнозування та усунення обмежень, притаманних простій архітектурі, було розроблено ширшу багатошарову нейронну модель (рис.9). Основною ідеєю побудови цієї моделі стало збільшення її представницької здатності шляхом ускладнення архітектури, зокрема за рахунок додавання прихованих шарів та збільшення кількості нейронів у кожному з них. Такий підхід дозволяє нейронній мережі ефективніше моделювати складні нелінійні залежності між вхідними ознаками та цільовою змінною.

Архітектура advanced моделі: Model: "sequential_4"		
Layer (type)	Output Shape	Param #
dense_13 (Dense)	(None, 128)	1,280
dropout_1 (Dropout)	(None, 128)	0
dense_14 (Dense)	(None, 64)	8,256
dense_15 (Dense)	(None, 32)	2,080
dense_16 (Dense)	(None, 1)	33
Total params: 11,649 (45.50 KB)		
Trainable params: 11,649 (45.50 KB)		
Non-trainable params: 0 (0.00 B)		

Рис. 9. Архітектура ширшої (покращеної) моделі

У ширшій моделі збережено ті самі вхідні ознаки та методи попередньої обробки даних, що й у базовій архітектурі, що забезпечує коректність подальшого порівняння результатів. Для прихованих шарів також використано нелінійні функції активації, а навчання здійснювалося за аналогічним алгоритмом оптимізації. Водночас збільшення кількості параметрів мережі дозволило моделі краще враховувати взаємозв'язки між технічними характеристиками та підвищити точність апроксимації.

Результати навчання ширшої нейронної мережі продемонстрували покращення показників якості у порівнянні з простою моделлю, що проявляється у зниженні значення середньоквадратичної похибки та більш стабільному процесі збіжності. Отримані результати підтверджують доцільність використання ускладненої архітектури для розв'язання задач регресії з великою кількістю взаємопов'язаних ознак. Порівняння ширшої та базової моделі зображено на рисунку 10.

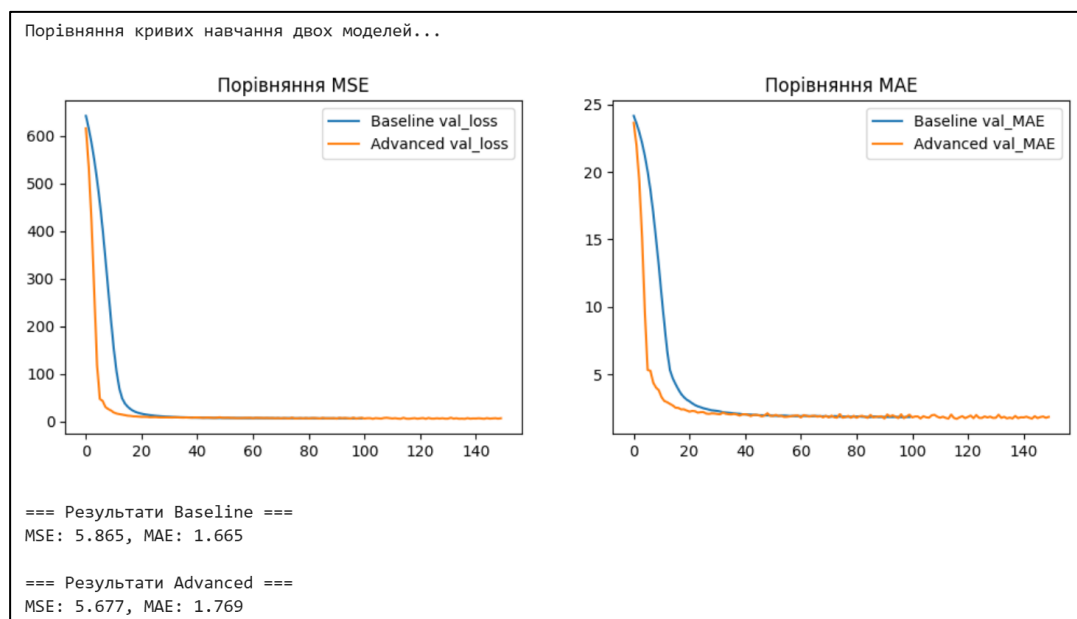


Рис. 10. Порівняння базової та покращеної моделей

Порівняльний аналіз базової та ширшої багат шарових нейронних моделей показав, що ускладнення архітектури має суттєвий вплив на якість розв'язання задачі регресії. Ширша нейронна мережа продемонструвала нижчі значення середньоквадратичної похибки у порівнянні з базовою моделлю, а також більш стабільний процес навчання, що свідчить про її вищу здатність до узагальнення даних. Це підтверджує, що додавання прихованих шарів і збільшення кількості нейронів дозволяє моделі ефективніше відображати складні нелінійні залежності між вхідними ознаками.

Водночас результати базової моделі засвідчили її обмежену представницьку здатність, яка проявляється у меншій точності прогнозування та повільнішій збіжності під час навчання. Таким чином, проведене порівняння доводить доцільність використання ширшої архітектури багат шарової нейронної мережі для задач регресії з високою складністю взаємозв'язків між параметрами, а також підкреслює важливість оптимального вибору архітектури при побудові нейронних моделей.

Висновки

У ході дослідження було проведено розвідувальний аналіз набору даних Auto MPG, що містить інформацію про технічні характеристики автомобілів та їхні показники витрат пального. На цьому етапі було досліджено розподіл основних числових ознак, що дозволило оцінити структуру даних, характер їх варіації та масштаб параметрів.

Побудовані гістограми розподілу показали суттєві відмінності між окремими характеристиками, зокрема масою автомобіля, потужністю та об'ємом двигуна, що вказує на необхідність попередньої обробки та масштабування даних перед навчанням моделей машинного навчання. Аналіз кореляційної матриці дозволив виявити найбільш значущі взаємозв'язки між технічними параметрами та цільовою змінною. Зокрема, встановлено сильний зв'язок між витратами пального та такими ознаками, як маса автомобіля, кількість циліндрів, об'єм і потужність двигуна, а також помірний вплив року випуску та регіону походження.

На основі отриманих результатів було побудовано та проаналізовано дві багат шарові нейронні моделі – базову та ширшу архітектуру. Порівняння результатів навчання показало, що ускладнення архітектури нейронної мережі сприяє зниженню похибки прогнозування та підвищенню стабільності процесу навчання. Отримані результати підтверджують доцільність використання ширших багат шарових нейронних мереж для моделювання складних нелінійних залежностей у задачах регресії та демонструють важливість обґрунтованого вибору архітектури нейронної моделі.

СПИСОК ВИКОРИСТАНОЇ ЛІТЕРАТУРИ

1. Auto MPG Dataset [Електронний ресурс]. – Режим доступу: <https://www.kaggle.com/datasets/uciml/autompg-dataset>
2. Pandas Getting started. 2026 [Електронний ресурс]. – Режим доступу: https://pandas.pydata.org/docs/getting_started/index.html
3. Matplotlib Pyplot Documentation. 2025 [Електронний ресурс]. – Режим доступу: https://matplotlib.org/3.5.3/api/as_gen/matplotlib.pyplot.html
4. Seaborn Tutorial. 2024 [Електронний ресурс]. – Режим доступу: <https://seaborn.pydata.org/tutorial.html>
5. Мокін В.Б., Дратований М.В. Наука про дані: машинне навчання та інтелектуальний аналіз даних. Навчальний посібник. м. Вінниця, 2024 р. 263 с. [Електронний ресурс]. – Режим доступу: <https://iq.vntu.edu.ua/repository/getfile.php/8163.pdf>
6. TensorFlow Documentation. 2024 [Електронний ресурс]. – Режим доступу: <https://www.tensorflow.org/learn>
7. Keras Documentation. 2024 [Електронний ресурс]. – Режим доступу: <https://keras.io>

Янковчук Михайло Ігорович – студент групи СА-236, факультет інтелектуальних інформаційних технологій та автоматизації, Вінницький національний технічний університет, Вінниця, e-mail: mykhailoyanki@gmail.com.

Войцеховська Ольга Олександрівна – PhD, доцент кафедри системного аналізу та інформаційних технологій, Вінницький національний технічний університет, м. Вінниця, e-mail: olgav1085@gmail.com

Yankovchuk Mykhailo I. – student of Faculty of Intelligent Information Technology and Automation, SA-23b, Vinnytsia National Technical University, Vinnytsia, e-mail: mykhailoyanki@gmail.com.

Voitsekhovska Olha O. – PhD, Associate Professor of the Department of System Analysis and Information Technologies, Vinnytsia National Technical University, Vinnytsia, e-mail: olgav1085@gmail.com