

# Застосування нейронних мереж та бібліотеки Keras для розпізнавання цифр на зображеннях

Вінницький національний технічний університет

## Анотація

У статті розглянуто застосування нейронних мереж та бібліотеки Keras для розпізнавання рукописних цифр на основі даних MNIST. Основна мета полягає у демонстрації побудови моделі, процесу її тренування та оцінки точності. Було використано нейронну мережу з трьома прихованими шарами, причому вихідний шар використовує функцію активації Softmax для класифікації цифр. Дані були нормалізовані для стабільнішого та швидшого навчання моделі. Тренування моделі проводилось на тренувальній та валідаційній вибірках, а оцінка точності — на тестовій вибірці. Модель досягла високої точності розпізнавання на рівні 97%. Обговорено основні фактори, що вплинули на успіх моделі, зокрема архітектура мережі, нормалізація даних та використання оптимізатора Adam. Запропоновані напрями подальших досліджень включають використання більш складних архітектур нейронних мереж та експерименти з різними гіперпараметрами.

**Ключові слова:** ШІ, Нейронні мережі, Keras, оптимізація, розпізнавання.

## Abstract

The article discusses the application of neural networks and the Keras library for recognizing handwritten digits based on the MNIST dataset. The primary goal is to demonstrate the construction of the model, its training process, and the evaluation of accuracy. A neural network with three hidden layers was used, with the output layer employing the Softmax activation function for digit classification. The data was normalized for more stable and faster model training. The model was trained on training and validation sets and evaluated on a test set. The model achieved a high recognition accuracy of 97%. The main factors contributing to the model's success were discussed, including network architecture, data normalization, and the use of the Adam optimizer. Proposed future research directions include using more complex neural network architectures and experimenting with different hyperparameters.

**Keywords:** AI, Neural Networks, Keras, optimization, recognition.

## Вступ

Розпізнавання рукописних цифр є важливим завданням у сфері комп'ютерного зору та машинного навчання, яке знаходить застосування у багатьох практичних завданнях, таких як автоматична обробка чеків, розпізнавання номерних знаків автомобілів, ідентифікація поштових індексів тощо. Одним із найбільш відомих наборів даних для тренування та тестування алгоритмів розпізнавання рукописних цифр є MNIST (Modified National Institute of Standards and Technology database), який містить 70 000 зображень розміром 28x28 пікселів, кожне з яких представляє одну з цифр від 0 до 9. У сучасних дослідженнях нейронні мережі, зокрема глибокі нейронні мережі, показали високу ефективність у завданні розпізнавання зображень. Бібліотека Keras, яка є інтерфейсом високого рівня для TensorFlow, надає зручний та потужний інструментарій для побудови та тренування нейронних мереж. Метою цієї статті є опис та демонстрація застосування нейронних мереж та бібліотеки Keras для розпізнавання рукописних цифр на основі даних MNIST. Ми розглянемо побудову моделі, процес тренування та оцінки точності моделі.

## Постановка задачі та методологія

Задача полягає у розпізнаванні рукописних цифр на зображеннях з використанням нейронної мережі. Ми будемо використовувати бібліотеку Keras для побудови моделі та набору даних MNIST для тренування та тестування.

1. Завантаження та підготовка даних: Спочатку завантажимо набір даних MNIST та розділимо його на тренувальну, валідаційну та тестову вибірки. Нормалізуємо зображення, щоб значення пікселів знаходились в діапазоні [0, 1].

2. Побудова моделі: Створимо нейронну мережу з трьома прихованими шарами. Вхідний шар буде Flatten для перетворення зображення розміром 28x28 в один вектор. Перший прихований шар буде

містити 30 нейронів з активацією ReLU, другий прихований шар - 10 нейронів з активацією ReLU, і вихідний шар - 10 нейронів з активацією Softmax для класифікації цифр. (1)

```
import sys
assert sys.version_info >= (3, 7)
from packaging import version
import sklearn
assert version.parse(sklearn.__version__) >= version.parse("1.0.1")
import tensorflow as tf
assert version.parse(tf.__version__) >= version.parse("2.8.0")
import matplotlib.pyplot as plt
tf.random.set_seed(42)
model = tf.keras.Sequential()
model.add(tf.keras.layers.InputLayer(input_shape=[28, 28]))
model.add(tf.keras.layers.Flatten())
model.add(tf.keras.layers.Dense(30, activation="relu"))
model.add(tf.keras.layers.Dense(10, activation="relu"))
model.add(tf.keras.layers.Dense(10, activation="softmax"))
```

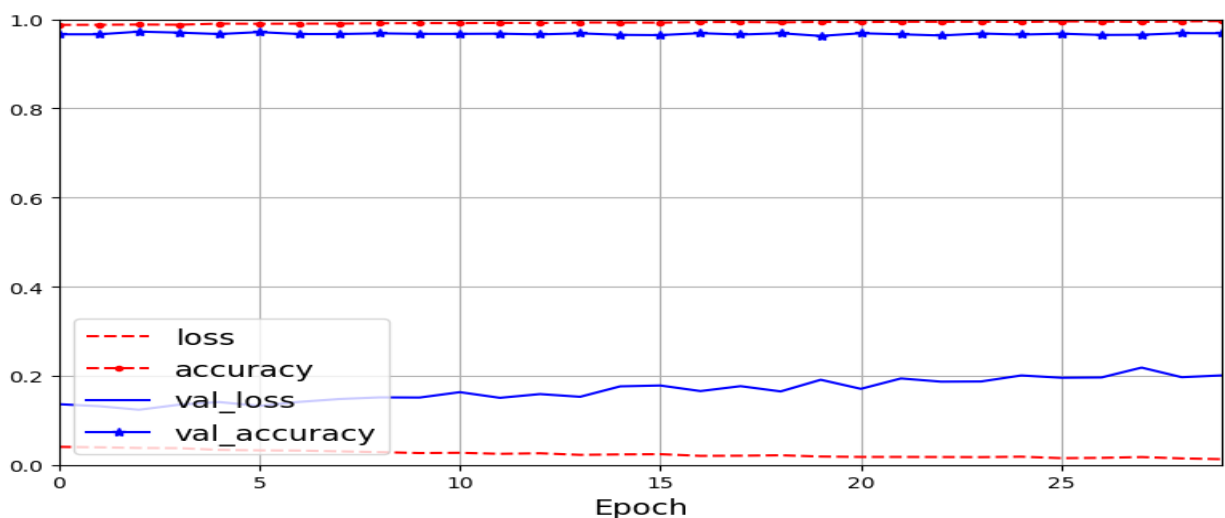
 (1)

3. Тренування моделі: В процесі тренування моделі будемо використовувати функцію втрат `sparse_categorical_crossentropy`, оптимізатор Adam та метрику точності. (2)

```
model.compile(loss="sparse_categorical_crossentropy",
              optimizer="adam",
              metrics=["accuracy"])
history = model.fit(X_train, y_train, epochs=30,
                  validation_data=(X_valid, y_valid))
```

 (2)

4. Оцінка моделі: Після тренування моделі на тренувальній та валідаційній вибірках, проведемо оцінку точності на тестовій вибірці. (рис. 3)



```
model.evaluate(X_test, y_test)
313/313 [=====] - 1s 2ms/step - loss: 0.2351 - accuracy: 0.9646
[0.23507927358150482, 0.964600266075134]
```

Рис. 3. Діаграма оцінки моделі по епохам

## Аналіз результатів

Результати тренування моделі демонструють високу точність розпізнавання на валідаційних та тестових даних. Після 10 епох тренування, модель досягає точності приблизно 97% на тестовій вибірці. Графік навчання (рис. 3) показує, що модель швидко навчається, знижуючи значення втрат та підвищуючи точність на кожній епосі. Ці результати свідчать про те, що модель не переобучається, оскільки точність на валідаційній вибірці продовжує зростати разом з точністю на тренувальній вибірці. Побудована модель показала високу ефективність у розпізнаванні рукописних цифр. Основні фактори, що вплинули на успіх моделі:

- 1) Архітектура мережі: Використання двох прихованих шарів з активацією ReLU дозволило моделі ефективно навчитися розпізнавати складні патерни у зображеннях.
- 2) Нормалізація даних: Нормалізація значень пікселів до діапазону  $[0, 1]$  сприяла стабільнішому та швидшому навчанню моделі.
- 3) Оптимізатор Adam: Цей оптимізатор є одним із найпопулярніших у сучасних дослідженнях завдяки його здатності швидко знаходити оптимальні значення ваг мережі.

## Висновки

У статті розглянуто застосування нейронних мереж та бібліотеки Keras для розпізнавання рукописних цифр на основі даних MNIST. Побудована модель продемонструвала високу точність (близько 97%) на тестовій вибірці. Використання бібліотеки Keras дозволило швидко та ефективно створити та налаштувати модель, що робить її зручним інструментом для задач комп'ютерного зору. Подальші дослідження можуть включати експерименти з більш складними архітектурами нейронних мереж, такими як згорткові нейронні мережі (Convolutional Neural Networks, CNN), для досягнення ще вищої точності. Також варто дослідити вплив різних гіперпараметрів на результати моделі.

## СПИСОК ВИКОРИСТАНОЇ ЛІТЕРАТУРИ

1. F. Chollet, "Deep Learning with Python," 2nd ed. Greenwich, CT, USA: Manning Publications Co., 2021.
2. L. Deng, "The MNIST Database of Handwritten Digit Images for Machine Learning Research," IEEE Signal Processing Magazine, vol. 29, no. 6, pp. 141-142, Nov. 2012.
3. M. Abadi et al., "TensorFlow: A System for Large-Scale Machine Learning," in 12th USENIX Symposium on Operating Systems Design and Implementation (OSDI 16), 2016, pp. 265-283.
4. A. Geron, "Hands-On Machine Learning with Scikit-Learn, Keras & TensorFlow", O'Reilly, third edition, 2023.
5. A. Glassner, "Deep Learning A Visual Approach", No Starch Press, Inc., San Francisco, 2021.

**Козловський Олексій Андрійович** – магістрант кафедри комп'ютерних наук, факультет інтелектуальних інформаційних технологій та автоматизації, Вінницький національний технічний університет, м. Вінниця, e-mail: ak@vin.ua

Науковий керівник: **Іванчук Ярослав Володимирович** – д.т.н., проф., професор кафедри комп'ютерних наук, факультет інтелектуальних інформаційних технологій та автоматизації, Вінницький національний технічний університет, м. Вінниця, e-mail: ivnchuck@vntu.edu.ua

**Oleksii Kozlovskyi** – Bachelor, Department of Computer Science, Faculty of Intelligent Information Technologies and Automation, Vinnytsia National Technical University, Vinnytsia, e-mail: ak@vin.ua

Supervisor: **Yaroslav Ivanchuk** – Doctor of Technical Sciences, Professor, Department of Computer Science, Faculty of Intelligent Information Technologies and Automation, Vinnytsia National Technical University, Vinnytsia, e-mail: ivnchuck@vntu.edu.ua