

УДК 004.05:004.85

*Кучер М.О., студент 2 курсу
магістратури, спеціальність 122
«Комп'ютерні науки»
Бабаков Р.М. к.т.н., доцент, доцент
кафедри інформаційних технологій*

МЕТОДИ ОЦІНКИ ЯКОСТІ НАВЧАННЯ НЕЙРОННОЇ МЕРЕЖІ

Донецький національний університет імені Василя Стуса, м. Вінниця

Оскільки потреба в навчанні нових нейронних мереж зростає з кожним роком, актуальним є оцінка якості їх навчання, оскільки від вірної оцінки залежить як якість моделі, так і витрати на її створення.

Окрім специфічних для конкретної задачі метрик є більш універсальні метрики, які дослідники можуть використовувати для значної частини моделей [1].

Розглянемо набір даних «Ships in Satellite Imagery»[2] та проаналізуємо методи оцінки якості навчання навченої на цих даних моделі.

Наведемо короткий опис даних в даному наборі, датасет представляє собою набір з сцен, що є супутниковими знімками поверхні землі з різними об'єктами, а також набір відсортованих зображень маленького розміру з прикладами кораблів і зображення всього іншого, що не можна вважати кораблем: вода, будівлі, шматки кораблів, уламки.

Всього датасет містить майже тисячу зображень кораблів і три тисячі інших зображень.

Обробивши датасет усунувши дисбаланс класів і навчивши нейронну мережу можемо оцінити якість цього навчання.

Одною з базових метрик оцінки якості навчання є середньоквадратична помилка, або «MSE», що є середнім квадратичної похибки між передбаченим значенням та істинним [3].

Часто представлена метрика використовується для оцінки якості моделі безпосередньо в процесі навчання для вчасної зупинки навчання і запобігання перенавчанню моделі, скористаємось аналогічним чином зберігши той стан моделі, який буде краще за інші беручи за початкові ваги випадкові числа і зберігаючи нові кожен раз коли похибка буде менше ніж для збережених раніше чисел.

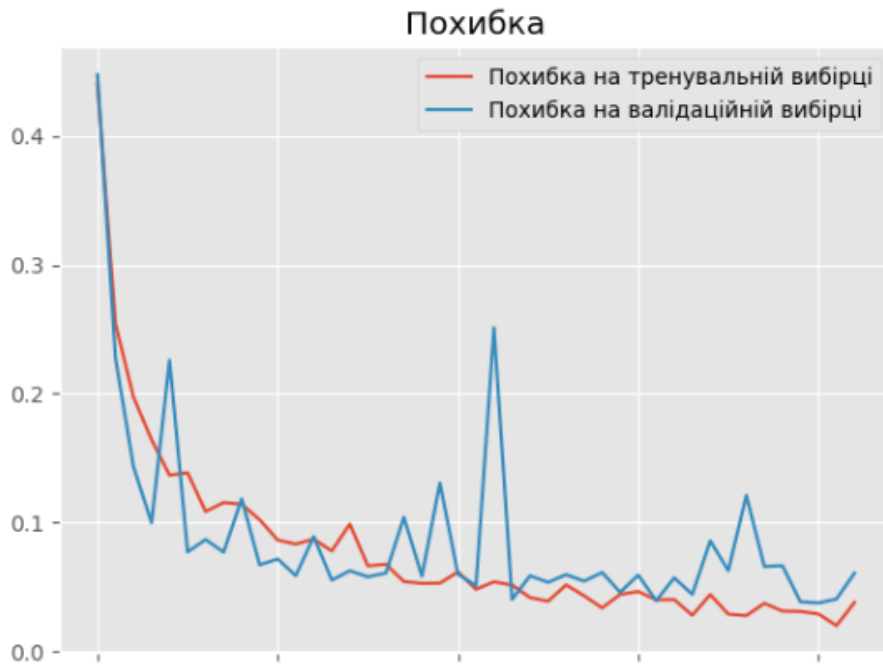


Рисунок 1 – Результат виведення середньоквадратичної помилки

На продемонстрованому рисунку видно поступове зменшення похибки як на валідаційній, так і на тренувальній виборках. Важливим є тренування на тренувальній вибірці, а оцінка якості саме на валідаційній, таким чином менше шансів на завчання мережею прикладів.

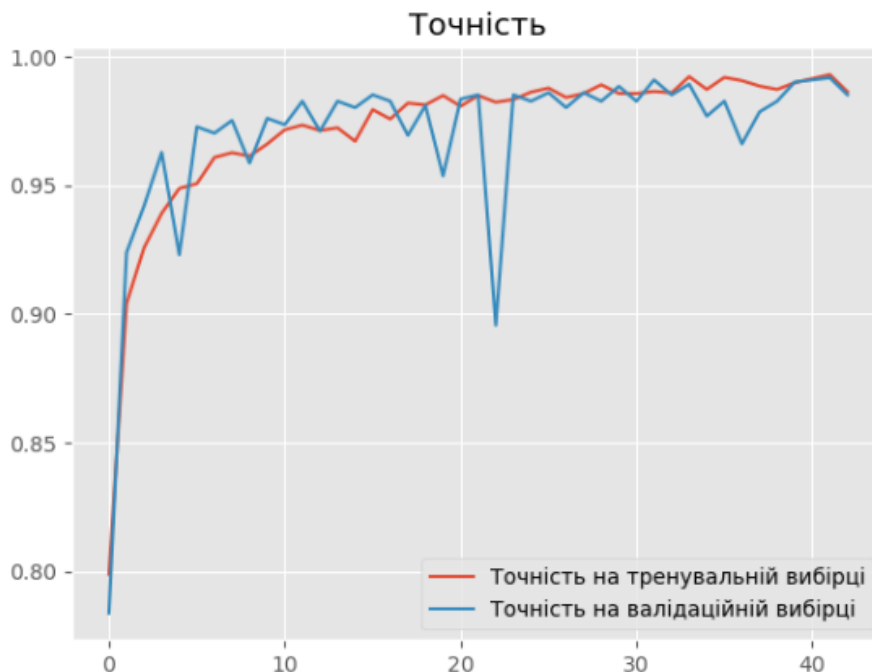


Рисунок 2 – Результат виведення точності

Аналогічно до середньоквадратичної помилки використовують ще й метрику точності «ассигасу», яка продемонстрована на рисунку 2.

Навчена модель, показує хороші результати якості навчання згідно з приведеними метриками, які застосовувались під час навчання, для подальшого

аналізу використана матриця заплутаності, яка зроблена за результатами перевірки роботи моделі з вагами, які були в моделі на момент найточнішого результату з точки зору квадратичної помилки і точності.

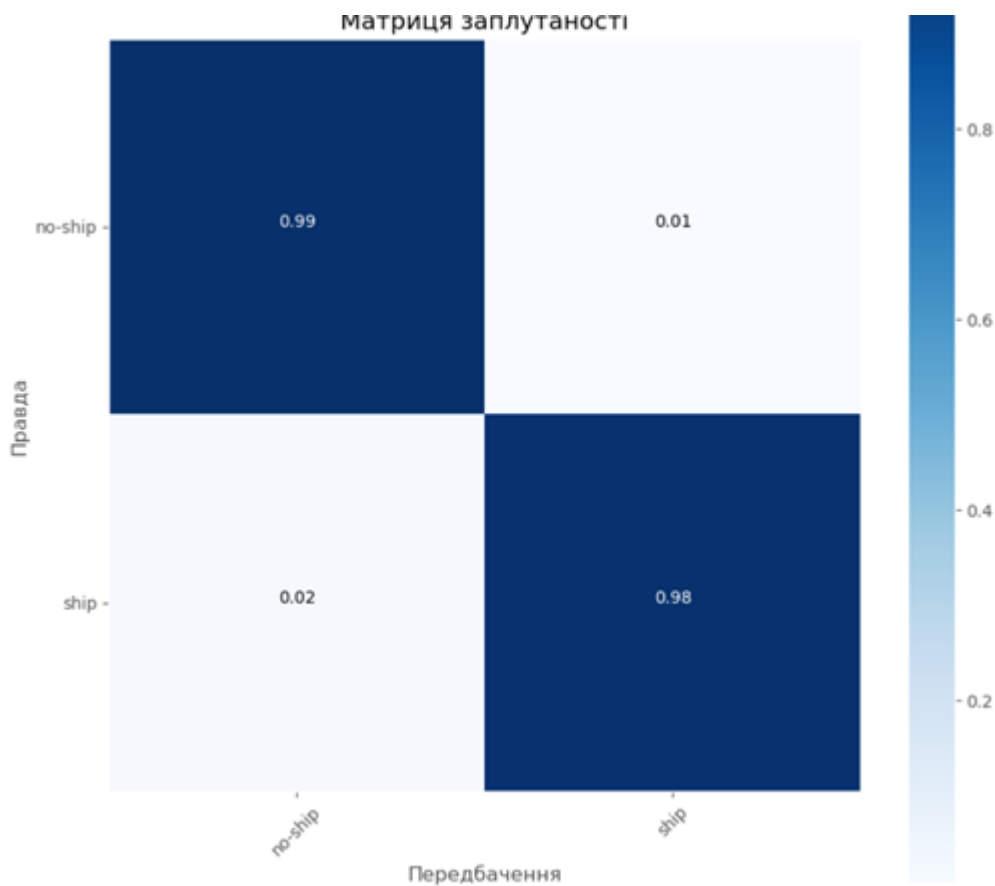


Рисунок 3 – Результат виведення матриці заплутаності

На рисунку видно, що матриця заплутаності представляє собою матрицю 2 на 2 в якій видно відношення реальності/правди, до нашого передбачення. Тобто матриця має 4 клітинки, які є комбінацією правди і вірного вгадування, правди і невірного вгадування, неправди і вірного вгадування, а також неправди в невірного вгадування. Дана матриця дає змогу оцінити не абстрактну точність чи похибку, а більш детально проаналізувати які помилки частіше виникають, що полегшує їх виправлення.

Отже, продемонстровані методи оцінки якості навчання нейромережі дозволяють дослідникам краще розуміти процес і наслідки навчання їх моделей, що було продемонстровано на прикладі обраного датасету.

Список використаних джерел

1. *Neural network-based non-intrusive speech quality assessment using attention pooling function* [Електронний ресурс]. Режим доступу до ресурсу: <https://asmp-urasipjournals.springeropen.com/articles/10.1186/s13636-021-00209-4>
2. *Ships in Satellite Imagery* [Електронний ресурс] // *Online Journal for Research and Education*. – 2021. – Режим доступу до ресурсу: <https://www.kaggle.com/datasets/rhammell/ships-in-satellite-imagery>
3. *Mean Squared Error* [Електронний ресурс]. Режим доступу до ресурсу: <https://congyuzhou.medium.com/mse-2adead8c9e35>