

УДК 004.93

Р. Ковальчук, М. Петрик, докт. фіз.-мат. наук; проф.

(Тернопільський національний технічний університет імені Івана Пулюя, Україна)

ВИКОРИСТАННЯ ЗАСОБІВ МАШИННОГО НАВЧАННЯ ПРИ РОЗРОБЦІ IOS ДОДАТКУ ДЛЯ РОЗПІЗНАВАННЯ ТЕКСТУ

UDC 004.93

R. Kovalchuk, M. Petryk, Dr.; Prof.

USE OF MACHINE LEARNING TOOLS IN THE DEVELOPMENT OF A MOBILE APPLICATION FOR TEXT RECOGNITION

Машинне навчання – це підгалузь штучного інтелекту в галузі інформатики, яка часто застосовує статистичні прийоми для надання комп'ютерам здатності «навчатися» (тобто, поступово покращувати продуктивність у певній задачі) з даних, без того, щоби бути програмованими явно [1]. Дана технологія є прикладом навчання на ходу. Чим більше кваліфікованих даних мають програми ML, тим точнішим стає алгоритм ML.

Щоб побудувати модель, яка розкриває зв'язки, машинне навчання використовує такі три алгоритми:

- Контрольоване навчання, коли алгоритм вивчає дані прикладу та відповідні відповіді цілі. Ці дані можуть включати числові значення або мітки рядків, наприклад класи чи теги. Пізніше, коли подаються нові приклади, ML може передбачити правильну відповідь.

- Навчання без нагляду. ML навчається на прикладах без жодних відповідей. Таким чином, алгоритм самостійно визначає шаблони даних.

- Навчання з підкріпленням. Розробники навчають алгоритми машинного навчання приймати конкретні рішення з середовища. Таким чином, машина отримує найкращі знання, щоб приймати точні рішення.

Дана технологія набула великої популярності через широкий спектр застосування, зокрема для програмування під мобільні платформи.

Яскравими прикладами використання засобів машинного навчання на мобільних платформах є програми Shazam, Google Lens, Snapchat, Spotify. Дані застосунки використовують натреновані моделі машинного навчання для розпізнавання пісень, аналізу зображень, обробки відео у реальному часі.

Компанія Apple надає широкий спектр технологій для розробника, що дозволяють йому впровадити алгоритми комп'ютерного зору (OCR) в свій проєкт (Vision framework) та полегшити класифікацію об'єктів (Core ML) тим самим створивши систему оптичного розпізнавання. Системи оптичного розпізнавання це один з небагатьох інструментів, що мають змогу надати людям з вадами зору можливість отримати доступ до печатної інформації. Загалом такі системи складаються з трьох основних частин:

- сканування;
- оптичне розпізнавання символів (англ. OCR);
- зачитування тексту за допомогою синтезованої мови.

Оптичне розпізнавання символів (англ. Optical character recognition) – це електронна трансляція зображень рукописного, машинописного або друкованого тексту в текстові дані.

Точне розпізнавання символів в печатному тексті можливе лише якщо наявні чіткі зображення. Точність в такому випадку може перевищувати 99%. Це досягається

сегментацією зображення, обробкою та подальшою класифікацією. Ще однією з широко досліджувальних тем є – розпізнавання рукописного тексту. Але в даний час точність сильно нижча ніж для “печатного” тексту. Форми окремих рукописних символів іноді можуть не містити достатньої кількості інформації для того, щоб розпізнати весь текст.

Література.

1. Машинне навчання [Електронний ресурс] – Режим доступу: https://uk.wikipedia.org/wiki/%D0%9C%D0%B0%D1%88%D0%B8%D0%BD%D0%BD%D0%B5_%D0%BD%D0%B0%D0%B2%D1%87%D0%B0%D0%BD%D0%BD%D1%8F.
2. Оптичне розпізнавання символів [Електронний ресурс] – Режим доступу: https://uk.wikipedia.org/wiki/%D0%9E%D0%BF%D1%82%D0%B8%D1%87%D0%BD%D0%B5_%D1%80%D0%BE%D0%B7%D0%BF%D1%96%D0%B7%D0%BD%D0%B0%D0%B2%D0%B0%D0%BD%D0%BD%D1%8F_%D1%81%D0%B8%D0%BC%D0%B2%D0%BE%D0%BB%D1%96%D0%B2.