

ЗАСТОСУВАННЯ НЕЙРОННИХ МЕРЕЖ В РОЗПІЗНАВАННІ РУКОПИСНОГО ТЕКСТУ

На сьогоднішній день великої популярності в світі набула така галузь штучного інтелекту як нейронні мережі. Актуальність розробок в галузі нейронних мереж обумовлена перш за все тим, що застосування даної моделі широко використовуються в найрізноманітніших областях. За допомогою вирішення задач на основі нейронних мереж функціонування будь-якої системи стає ефективнішим.

Сьогодні відома велика кількість галузей застосування штучних нейронних мереж. Найбільш розповсюдженими серед них є: фінанси, економіка, медицина, наукові дослідження, інформаційні технології, штучний інтелект та ін.

Необхідно зазначити що існує велика кількість програмного забезпечення, що використовує можливості технологій штучних нейронних мереж. Прикладом можуть бути універсальні програми, що вирішують задачі від розпізнавання рукописного тексту до задач прогнозування.

Проте розглядаючи розпізнавання рукописного тексту, ми маємо деякі його властивості які в подальшому викликають проблеми при розпізнаванні.

Так, однією з проблем яка викликає труднощі при розпізнаванні є те, що всі рукописні символи мають статичні і динамічні властивості. Статичні можуть полягати в розмірі або формі символу, а відмінності в динаміці можуть бути в кількості штрихів та їх порядку.

Проблеми розпізнавання символів можуть виникати, наприклад, і через великий алфавіт мови. Варто зазначити що найбільшою проблемою розпізнавання рукописного тексту була і буде неоднозначність при читанні. Іноді у людей виникають труднощі при спробі прочитати навіть власний почерк.

Ряд проблем виникає через той факт, що одні і ті ж самі символи можуть бути написані по-різному. Також рідко можна зустріти двох людей з однаковим почерком. Ця задача пов'язана з різницею шрифтів в класичній задачі розпізнавання тексту.

На відміну від шрифтів, кожен символ в рукописному тексті може мати зовсім інший стиль і вигляд в залежності від контексту, в якому здійснюється написання букв і багатьох інших факторів.

Найбільш універсальний підхід до вирішення завдання про розпізнавання рукописного тексту є нейронні мережі. Основні переваги нейронних мереж полягають в здатності навчатися самостійно і автоматично на основі вибірок, бути продуктивними на зашумлених або нечітких даних, мати можливість паралельної реалізації і бути ефективними інструментами для обробки великих баз даних.

Найпопулярнішою нейронною мережею, що набула широкого використання, є багатошаровий перцептрон – Multi-Layer Perceptron (MLP). Ця структура навчається за допомогою зворотного розповсюдження помилки, є однією з найбільш популярних і універсальних форм нейронних мереж-класифікаторів. Її часто використовують для розпізнавання рукописного тексту. Однак такі системи мають і суттєві недоліки. Насамперед, це те, що мережа може бути ненавчена за вказаний час.

При навчанні такої мережі вхідна множина сигналів розглядається як вектор. Як вектор для розпізнавання рукописного тексту обираються закодовані значення пікселів. Навчання здійснюється шляхом послідовного представлення вхідних векторів з одночасним налаштуванням ваг відповідно визначеній процедурі. Мережа має вхідний, прихований та вихідний шари. В процесі навчання кожен вхідний вектор трансформується у вихідний вектор. Маючи вихідний вектор необхідно визначити який з нейронів вихідного шару має найвище значення активації – це і є результатом роботи нейронної мережі.

Нейронна мережа використовує навчаючу вибірку для автоматичного виводу правил щоб розпізнати рукописний текст. Чим більше навчаючих прикладів, тим більше нейронна мережа може дізнатися про рукописний текст і в результаті це підвищить точність розпізнавання.

В нашій роботі ми використовуємо багатошаровий перцептрон (MLP) для розпізнавання рукописного тексту та множини зображень бази MNIST. Під час тестування такої мережі було виявлено, що на 10 000 тестових зображеннях точність розпізнавання досягає приблизно 92%. Проте, чим більша кількість прихованих шарів в мережі, тим більша точність розпізнавання. Але необхідно зважати, що велика кількість прихованих шарів впливає на швидкість навчання тому в нашій мережі емпіричним методом було обрано оптимальну кількість прихованих шарів в мережі – 2.