

## ЗАСТОСУВАННЯ МЕРЕЖІ ГЛУБОКОЇ ДОВІРИ В ЗАДАЧІ РОЗПІЗНАВАННЯ ЗОБРАЖЕНЬ

На сьогодні нейронні мережі різних топологій використовуються для вирішення різноманітних задач, наприклад прогнозування, розпізнавання образів, зокрема обробка сигналів і зображень, аналіз рукописного тексту, кластеризація образів.

Універсальної топологією, що дозволяє вирішувати всі подібні завдання, є багатошаровий перцептрон. Багатошаровий перцептрон є мережею прямого поширення і навчається методом зворотного поширення помилки, який належить до групи навчання з учителем. Незважаючи на відносну простоту алгоритму, у нього є недолік - ризик «застрягти» в локальному мінімумі поверхні помилок, тоді алгоритм не прямуватиме до бажаного мінімуму і процес навчання не може бути закінчений.

Зі збільшенням розмірності задач, ускладнюється поверхня помилок. В умовах, коли образи описуються десятками або сотнями параметрів, навчання перцептрона методом зворотного поширення помилки дозволяє досягти лише певної точності, яка буде варіюватися в області 60-80%. Крім того, при великій розмірності потрібен перехід до даних меншої розмірності. Це викликає необхідність нарощувати шари перцептрона, а значить з кожен новим шаром, помилка навчання буде все менше впливати на налаштування параметрів перших шарів, і тим довше буде процес навчання.

Для того, щоб вирішити поставлене завдання, замінюємо перцептрон мережею глибинної довіри. Мережа глибинної довіри складається з декількох обмежених машин Больцмана (англійською Restricted Boltzmann Machine, скорочено – RBM), з'єднаних так, щоб невидимий шар однієї машини виконував роль видимого шару для наступної.

Навчання перцептрона зводиться до задачі навчання обмежених машин Больцмана і використанні отриманих вагових коефіцієнтів для роботи алгоритму зворотнього розповсюдження помилки.

Навчання мережі глибинної довіри відбувається за алгоритмом layer-wise learning. В ході роботи даного алгоритму обмежені машини Больцмана навчаються послідовно, отже основою попереднього навчання є саме алгоритм навчання обмежених машин Больцмана.

Кожна з машин Больцмана навчається за алгоритмом незалежного розходження (англійською contrastive divergence або CD-k), який запропонував Дж.Хінтон. Критерій зупину алгоритму наведений нижче.

$$\frac{\partial \ln P(v^{(k)})}{\partial w_{ij}} = M [v_i^{(k)} h_j]_{data} - M [v_i h_j]_{model} \quad (1)$$

Дж. Хінтон запропонував замінити математичні очікування даними наступним чином.

$$M [v_i^{(k)} h_j]_{data} \approx \frac{\partial E(v_i^{(k)} h_j)}{\partial w_{ij}},$$
$$M [\bar{v}_i \bar{h}_j]_{model} \approx \frac{\partial E(\bar{v}_i \bar{h}_j)}{\partial w_{ij}} \quad (2)$$

де  $E$  – енергія машини Больцмана,  $w_{ij}$  - вагові коефіцієнти,  $v$  – нейрон видимого шару,  $h$  – нейрон невидимого шару

Був змодельований процес навчання нейронної мережі на тестовій вибірці MNIST. Завдання включало в себе навчання на наборі зображень рукописних цифр від 0 до 9 і розпізнавання відповідних образів на інших наборах.

Як видно з таблиці 1, використання глибинного навчання значно підвищує точність роботи нейромережі. Разом із тим варто відмітити, що використання лише однієї ітерації алгоритму CD-k дозволяє досягти значних покращень точності.

Таблиця 1. Результати навчання

	Без глибинного навчання	CD-10	CD-1
Точність навчання	63,4%	97%	90,2%

Застосування методів глибинного навчання для вирішення задач великої розмірності дозволяє суттєво підвищити точність навчання, що було перевірено при комп'ютерному моделюванні. Запропонований підхід дозволяє вирішувати великий спектр задач, таких як прогнозування, аналіз рукописного тексту, розпізнавання голосу, тощо.