

Валько В.В.*студент,**Національний технічний університет України**«Київський політехнічний інститут імені Ігоря Сікорського»*

АЛГОРИТМ НАВЧАННЯ БАГАТОШАРОВОЇ НЕЙРОННОЇ МЕРЕЖІ МЕТОДОМ ЗВОРОТНОГО РОЗПОВСЮДЖЕННЯ ПОМИЛКИ

Метою даної статті було проаналізувати алгоритм навчання багатошарової нейронної мережі методом зворотного розповсюдження помилки. Для того щоб провести аналіз даної нейронної мережі потрібно провести її навчання таким чином, щоб досягнути збалансованих здібностей мережі давати правильний відгук на вхідні дані, які використовувались під час навчання (дана здібність називається запам'ятовування), і здатністю видавати правильні результати у відповідь на вхідні дані, схожі, але не ідентичні тим, що були використані під час навчання (реалізація принципу узагальнення).

Навчання мережі методом зворотного розповсюдження помилки включає в себе три етапи: задання на вхід навчальних даних, з подальшим розповсюдженням даних в напрямку виходів, обчислення і зворотній розподіл відповідної помилки і корегування параметрів прийняття рішень. Після навчання передбачається лише подання на вхід мережі даних і розповсюдження їх в напрямку виходів. При цьому, якщо навчання мережі може бути досить тривалим процесом, то безпосереднє обчислення результатів навченою мережею відбувається дуже швидко. Крім того існують деякі варіації методу зворотного розповсюдження помилки, які забезпечують суттєве покращення швидкості процесу навчання мережі.

Нейрони, які представляють собою виходи мережі (позначені як Y), і приховані нейрони можуть мати зміщення. Зміщення, яке відповідає виходу Y_k позначено як w_{0k} , а прихованому елементу Z_j – v_{0j} . Ці зміщення слугують в якості ваги на зв'язках, які виходять з нейронів, на виході яких завжди з'являється 1 (показано на рис. 1). В процесі навчання сигнали розповсюджуються в зворотному напрямку.

Алгоритм, який представлено нижче, застосовується до нейронної мережі з одним прихованим шаром, що підходить для більшості ситуацій на практиці. Як вже було описано вище, навчання мережі включає в собі три етапи: подача на входи мережі навчальних даних, зворотній розподіл помилок та корекція ваги нейронів. В ході першого етапу кожен вхідний нейрон x_i отримує сигнал і транслює його кожному з прихованих нейронів Z_1, Z_2, \dots, Z_p . Кожен прихований нейрон обчислює результат його мережевої функції і розсилає свій сигнал z_j всім вихідним нейронам. Кожен вихідний нейрон Y_k , в свою чергу, обчислює результат своєї мережевої функції, який являє собою вихідний сигнал даного нейрону для відповідних вхідних даних. В процесі навчання, кожен нейрон на виході мережі порівнює обчислене значення Y_k , з наданим цільовим значенням, визначаючи відповідне значення помилки для даного вхідного шаблону. На

основі даної помилки обчислюється $\sigma_k (k = 1, 2, \dots, m)$, σ_k використовується при розповсюдженні помилки від Y_k до всіх елементів попереднього шару, а також пізніше при зміні ваги зв'язків між прихованими та вихідними нейронами. Аналогічно обчислюється $\sigma_j (j = 1, 2, \dots, p)$ для кожного прихованого нейрона Z_j [2]. Незважаючи на те, що розповсюджувати помилку до вхідного шару немає необхідності, σ_j використовується для зміни ваги зв'язків між нейронами прихованого шару та вхідними нейронами. Після того, як всі σ було визначено, відбувається одночасна корекція ваги всіх зв'язків.

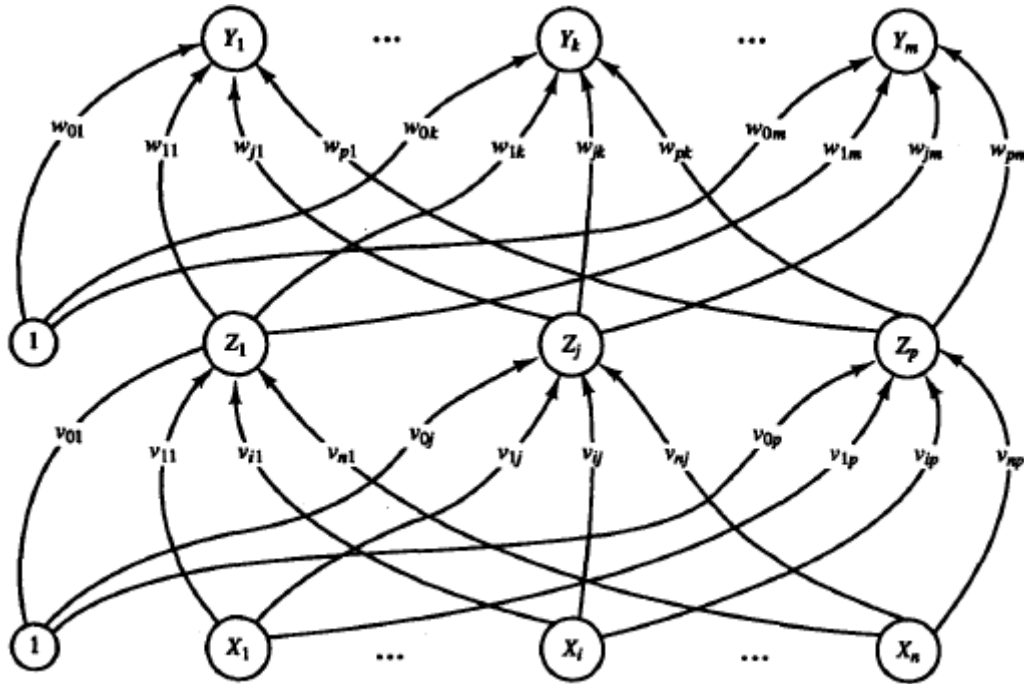


Рис. 1. Мережа з зворотнім розповсюдженням помилки з одним прихованим шаром

Джерело: [1]

Вибір початкової ваги впливає на те, чи зможе мережа досягнути глобального(або тільки локального) мінімуму помилки, і наскільки швидко цей процес буде відбуватися. Зміна ваги між двома нейронами, яка зв'язана з активаційною функцією нейрона з наступного шару і активаційної функції нейрона з попереднього шару. В зв'язку з цим, важливо уникати вибору такої початкової ваги, які обнуляють активаційну функцію або її похідну. Також початкові ваги повинні бути не надто великими, щоб вхідний сигнал для кожного прихованого або вихідного нейрона не потрапив в регіон дуже малих значень сигмоїда. З іншого боку, якщо початкові ваги будуть занадто малими, то вхідний сигнал на приховані або вихідні нейрони буде близьким до нуля, що призведе до дуже низької швидкості навчання. Стандартна процедура ініціалізації ваги полягає в присвоєнні їм випадкових значень в інтервалі $(0,5; 0,5)$ Значення можуть бути і додатними і від'ємними, так як кінцева вага, яка буде отримана після навчання, може бути і додатною і від'ємною.

Нижче наведено модифікацію процедури ініціалізації за алгоритмом Nguyen-Widrow [3], яка дає можливість швидшого навчання.

Введемо позначення:

n – кількість вхідних нейронів

p – кількість прихованих нейронів

β – фактор масштабування: $\beta = 0.7 * (p)^{\frac{1}{n}}$

Процедура складається з наступних кроків:

1. Для кожного прихованого нейрону ($Z_j, j = 1, 2, \dots, p$) ініціалізувати його вектор зв'язків з вхідними нейронами: $v_{ij}(old) = rand(-0.5, 0.5)$
2. Обчислити $v_{ij} = (V_{1j}(old)^2 + V_{2j}(old)^2 + \dots + V_{nj}(old)^2)^{\frac{1}{2}}$
3. Переініціалізувати зв'язки: $v_{ij} = \frac{\beta * v_{ij}(old)}{v_{ij}(old)}$
4. Задати значення зміщення: $m_{0j} = rand(-\beta, \beta)$.

Перевагами даної нейронної мережі є досить швидкий процес обчислення результатів навченою мережею, крім того ця мережа не має проблеми з обмеженням в шаблонах даних, які можуть бути використані для її навчання, на відміну від одношарових нейронних мереж. Завдяки швидкості обчислення результатів нейронна мережа з зворотнім розповсюдженням помилки може бути застосована для реальних задач.

Список використаних джерел:

1. Хайкин С. Нейронные сети: Полный курс // Пер. с англ. Н. Н. Куссуль, А. Ю. Шелестова. 2-е изд., испр. – М.: Издательский дом Вильямс, 2008, 1103 с.
2. Laurene V. Fausett Fundamentals of Neural Networks: Architectures, Algorithms And Applications 1st Edition // Pearson, 2004, 461 p.
3. Martin T Hagan, Neural Network Design, 2nd Edition // Martin Hagan 2014, 800 p.

Глущенко Д.К.

студент;

Черняк Л.П.

доктор технічних наук, професор,

Національний технічний університет України

«Київський політехнічний інститут імені Ігоря Сікорського»

АНАЛІЗ МАГМАТИЧНИХ ПОРІД ЗАХІДНОЇ УКРАЇНИ

Розширення і ефективне використання сировинної бази є передумовою розвитку виробництва силікатних матеріалів різного призначення [1]. При цьому значна увага приділяється магматичним породам, які складають превалюючу частку літосфери, в тому числі в Україні [2-5]. Для визначення напрямків застосування і оптимізації технологічних параметрів силікатних виробництв при застосуванні магматичних порід як вихідної сировини слід