

ДОСЛІДЖЕННЯ ШВИДКОСТІ НАВЧАННЯ НЕЙРОННИХ МЕРЕЖ В ЗАЛЕЖНОСТІ ВІД ФУНКЦІЇ АКТИВАЦІЇ

Кравченко А. В., Романов А.О.

Навчально-науковий інститут телекомунікаційних систем

КПІ ім. Ігоря Сікорського, Україна;

Національний Авіаційний Університет, Україна

E-mail: kravchenko.andriy@lkl.kpi.ua, anton3329@gmail.com

STUDY OF THE LEARNING SPEED OF NEURAL NETWORKS DEPENDING ON THE ACTIVATION FUNCTION

The paper studies the learning speed of a neural network depending on the activation function used. Three types of activation functions were considered: the Heaviside function, the Sigmoid activation function, and the Hyperbolic Tangent activation function. It was determined that the activation function has a significant impact on the quality of learning. In addition, the coupling coefficients of the neural network and the range of input data values have a great impact on the speed and quality of solving the problem.

В роботі проведено дослідження швидкості навчання нейронної мережі в залежності від функції активації, що використовується. Було розглянуто три види функцій активації: функція Хевісайду, функція активації Сигмоїда та функція активації Гіперболічний тангенс. Було визначено, що функція активація має значний вплив якості навчання. Крім того, на швидкість та якість розв'язання задачі мають великий вплив коефіцієнти зв'язку нейронної мережі та діапазон вхідних значень даних.

Метою дослідження є аналіз впливу різних функцій активації на швидкість навчання нейронної мережі. Дослідити залежність швидкості навчання від типу функції активації, коефіцієнтів зв'язку.

Сучасні телекомунікаційні мережі постійно розширюються. Це в свою чергу вимагає ефективного аналізу трафіку, стану телекомунікацій. Нейронні мережі дозволяють оптимізувати цей процес. Шляхом визначення оптимальної функції активації можна покращити швидкість навчання та краще адаптуватися до нових умов [1,2].

В дослідженні використовувався одношаровий перцептрон (рис.1). Він працюють по принципу:

- на входи мережі подаються значення, які перемножуються на коефіцієнти
- далі значення підсумовуються та до них застосовується функція активації.

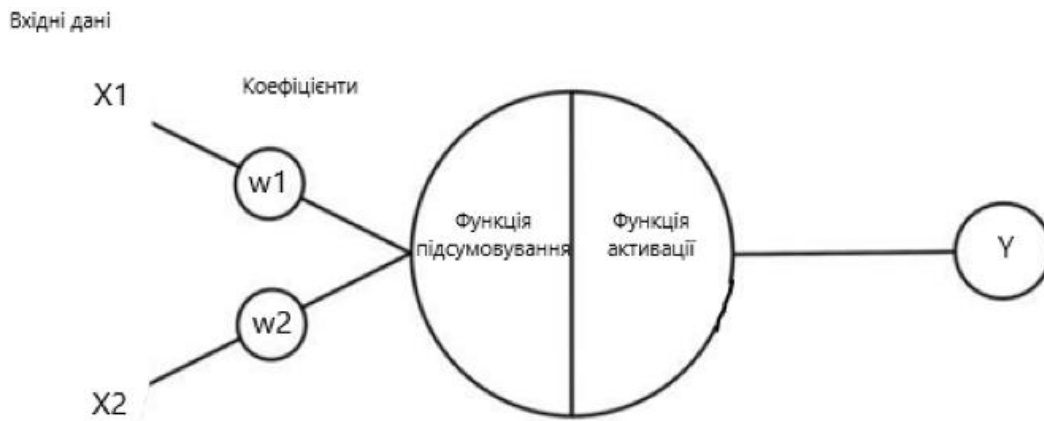


Рис. 1. Схематична модель одношарового перцептрону.

На рис. 1 позначено: X_1 , X_2 – вхідні параметри; Y – вихідний результат; w_1 , w_2 – коефіцієнти нейронної мережі.

Даними, на яких навчався перцептрон слугували значення, які приймає логічна функція імплікації:

$$y = (x_1 \rightarrow x_2) \quad (1)$$

де: x_1 , x_2 – використовуються як вхідні значення в нейронну мережу; y – вихідні значення; $\rightarrow$ - оператор логічної імплікації.

Мережа навчалась за алгоритмом градієнтного спуску. Метрикою, за допомогою якої визначалась швидкість навчання має наступний вид: N_{epoch} – кількість епох навчання. Епоха – один повний цикл навчання мережі на наборі даних, який використовується для тренування моделі. Початкові коефіцієнти w_1 , w_2 встановлювались рівними одному значенню w [3, 4, 5].

Результати експериментів показали, що при навчанні з кожною функцією активації є значення коефіцієнтів при яких швидкість навчання є найбільшою.

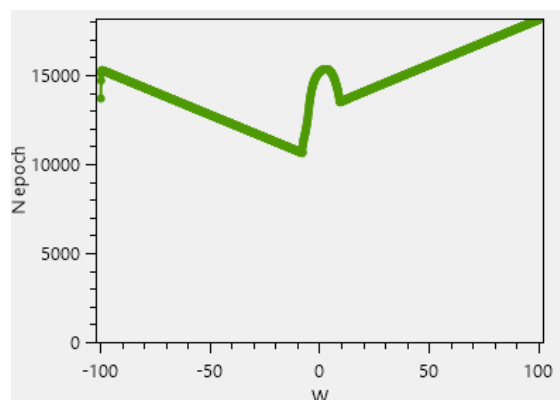


Рис 2. Результати навчання одношарового перцептрону з функцією активації Сигмоїда.

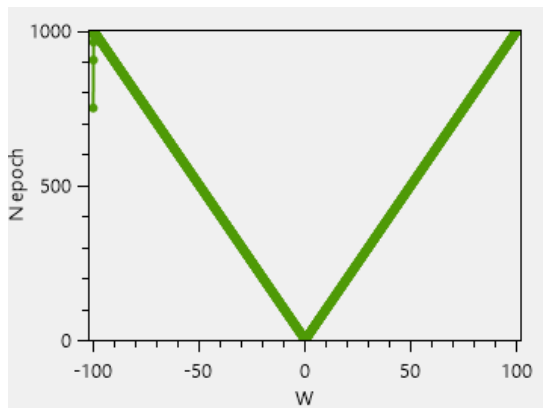


Рис. 3. Результати навчання одношарового перцептрону з функцією активації Хевісайда.

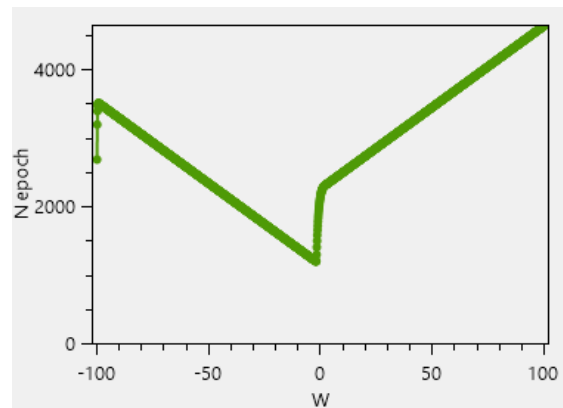


Рис. 4. Результати навчання одношарового перцептрону з функцією активації Гіперболічний тангенс.

На рис. 2, 3, 4 позначено: N_{epoch} – кількість епох навчання; W – значення, які набувають коефіцієнти перцептрону w_1, w_2 .

Висновки: Результати експериментів показують, що перцептрон з різними функціями активації має різний час навчання при однакових коефіцієнтах. Було встановлено, що при використанні функції Хевісайда і діапазоні вхідних значень $x_1, x_2 \in [0;1]$ швидкість навчання мережі є найбільшою. Перцептрон з даною функцією найшвидше навчається при величині коефіцієнтів, рівних нулю, що свідчить про те, що при ініціалізації коефіцієнтів випадковими числами краще брати значення, близькі до нуля.

Література

1. Mahmoud Abbasi, Amin Shahraki, Amir Deep Learning for Network Traffic Monitoring and Analysis (NTMA): A Survey / Taherkordi 2021 <https://www.sciencedirect.com/science/article/pii/S0140366421000426>
2. Abbasi, M., Shahraki, A., & Taherkordi, A. (2021). Deep learning for Network Traffic Monitoring and Analysis (NTMA): A survey. Computer Communications, 170, 19–41. <https://doi.org/10.1016/j.comcom.2021.01.021>
3. Annette Lopez Davila, Professor Chi-Kwong Li Neural Networks: The Backpropagation Algorithm / 2020 <https://cklixx.people.wm.edu/teaching/math400/Annette-paper.pdf>.
4. Gal Vardi, Gilad Yehudai 2021 Learning a Single Neuron with Bias Using Gradient Descent https://papers.neurips.cc/paper_files/paper/2021/file/f0f6cc51dacebe556699ccb45e2d43a8-Paper.pdf.
5. Romanov, O.I., Fediushyna, D.M., Dong, T.T. Model and method of Li-Fi network calculation with multipath light signals/(2018) 2018 International Conference on Information and Telecommunication Technologies and Radio Electronics, UkrMiCo 2018. DOI: 10.1109/UkrMiCo43733.2018.9047550