1 КОРОТКІ ТЕОРЕТИЧНІ ВІДОМОСТІ

Нервова система людини складається з клітин, які називаються нейронами, і має надзвичайну складність: близько 1011 нейронів беруть участь у близько 1015 передавальних зв'язках, що мають довжину метр і більше. Кожен нейрон має багато якостей, спільних з іншими клітинами, але його унікальною здатністю є прийом, обробка і передача електрохімічних сигналів по нервовим шляхам, що утворюють комунікаційну систему мозку.

На рис. 1 показана структура біологічного нейрона. Дендрити йдуть від тіла нервової клітини до інших нейронів, де вони приймають сигнали в точках з'єднання, які називаються **синапсами**.

Рисунок 1 – Структура біологічного нейрона

Прийняті синапсом вхідні сигнали підводяться до тіла нейрона. Тут вони сумуються, причому одні входи прагнуть збудити нейрон, інші – перешкодити його збудженню. Коли сумарне збудження в тілі нейрона перевищує деякий поріг, нейрон збуджується, посилаючи по аксону сигнал іншим нейронам. У цієї основної функціональної схеми багато ускладнень і виключень, проте більшість штучних нейронних мереж (НМ) моделюють лише ці прості властивості.

Отже, нейрон (формальний нейрон, нейроподібний елемент) являє собою примітивний обчислювальний пристрій, що має кілька входів і один вихід, і є основним обчислювальним елементом НМ.

Одношаровий персептрон є одним з найпростіших варіантів НМ (рис. 2) і містить лише один нейрон. Будучи самостійною моделлю НМ з одного боку, одношаровий персептрон (формальний нейрон) є основним конструкційним елементом для більшості моделей НМ, з іншого боку.

На вхід одношарового персептрона (формального нейрона) надходить набір вхідних сигналів *x1, x2, ..., xN* або вхідний вектор *x*. Кожен вхідний сигнал помножується на відповідну вагу зв'язку *w1, w2,..., wN* – аналог ефективності синапсу (міжнейронного контакту).

Рисунок 2 – Одношаровий персептрон (формальний нейрон)

Вага зв'язку є скалярною величиною, позитивною для збудливих і негативною для гальмуючих зв'язків. Зважені вагами зв'язків вхідні сигнали надходять на блок сумації, що відповідає тілу клітини, де здійснюється їхня алгебраїчна сумація та визначається рівень збудження нейроподібного елемента *net*:

.

Вихідний сигнал нейрона *y* визначається шляхом пропущення рівня збудження *net* через нелінійну функцію активації *ψ*:

*y = ψ (net – θ),*

де *θ* – деякий постійний зсув (аналог порога нейрона). Звичайно використовуються найпростіші нелінійні функції:

бінарна (порогова): 

чи сигмоїдна: *ψ(x*) = 1 / (1 *+ e–x*)*.*

У залежності від типу функції активації розрізняють дискретні персептрони, що використовують порогову функцію активації, і дійсні – що використовують дійсні функції активації, наприклад, сигмоїдну функцію.

Кожен нейрон має невелику пам'ять, що реалізовується ваговими коефіцієнтами і порогом нейрона. Тому нейрони можна розглядати як запам'ятовуючі пристрої. У той же час нейрони можуть розглядатися як примітивні процесори, що здійснюють обчислення значення функції активації на основі різниці зваженої суми вхідних сигналів і порога.

**Алгоритм навчання одношарового дискретного персептрона** має вигляд.

*Крок 1.* Вагам *wi*(0) (*i*= 1, ..., *N*) і порогові *θ*(0) присвоюються випадкові значення (через *wi*(*t*) позначений ваговий коефіцієнт *i-*го входу персептрона в момент часу *t*, через *θ*(*t*) позначена величина зсуву (порога) нейрона в момент часу *t*).

*Крок 2*. Пред'являються черговий вхідний вектор *х* = {*х1, ..., хN*}т з навчальної множини і бажаний вихід *y*\*(*t*) (*y*\*(*t*) = 1, якщо *х*(*t*) відноситься до класу *A*, *y*\*(*t*) = 0, якщо *х*(*t*) відноситься до класу *В*).

*Крок 3*. Обчислюється значення на виході персептрона за формулами:

,

**.

*Крок 4*. Корегуються ваги персептрона відповідно до рівностей:

*wi*(*t*+ 1) = *wi*(*t*) + *η*(*y*\*(*t*) - *y*(*t*))*xi*(*t*), *i* = 1, 2, ..., *N*,

*θ*(*t* + 1) = *θ*(*t*) + *η*(*y*\*(*t*) – *y*(*t*)),

де *η* – позитивний коригувальний приріст.

*Крок 5*. Якщо досягнута збіжність, то процедура навчання закінчується; у протилежному випадку – перехід до кроку 2.

Відповідно до даного алгоритму спочатку виробляється ініціалізація параметрів персептрона випадковими значеннями. Потім по черзі пред'являються екземпляри з відомою класифікацією, обрані з навчальної множини, і корегуються ваги відповідно до формул кроків 3 і 4. Величина корекції визначається позитивним коригувальним приростом *η*, конкретне значення якого вибирається досить великим, щоб швидше вироблялася корекція ваг, і в той же час досить малим, щоб не допустити надмірного зростання значень ваг.

Процедура навчання продовжується до тих пір, поки не буде досягнута збіжність, тобто поки не будуть отримані ваги, що забезпечують правильну класифікацію для всіх образів з навчальної множини.

У тому випадку, коли навчальні вибірки розділити гіперплощиною неможливо для навчання персептрона можна використовувати **алгоритм** **Уідроу-Хоффа**, який мінімізує середньоквадратичну помилку між бажаними і реальними виходами мережі для навчальних даних. Цей алгоритм також можна застосовувати для навчання одношарового дійсного персептрона. Алгоритм Уідроу-Хоффа можна записати в тому ж вигляді, що і вищеописаний алгоритм, припускаючи, що у вузлах персептрона нелінійні елементи відсутні, а коригувальний приріст *η* у процесі ітерацій поступово зменшується до нуля.

Якщо для рішення задачі розпізнавання образів використовується дискретний персептрон, то вирішальне правило відносить вхідний образ до класу *A*, якщо вихід персептрона дорівнює 1, і до класу *В* – у протилежному випадку.

У випадку, якщо для рішення задач розпізнавання образів використовується дійсний персептрон, то вирішальне правило відносить вхідний образ до класу *А*, якщо вихід мережі більше 0.5, і до класу *В* – у протилежному випадку.