

## СИСТЕМА РОЗПІЗНАВАННЯ МОВИ НА ОСНОВІ НЕЙРОМЕРЕЖІ

Нейрон – елементарний елемент, що виконує зважене підсумовування й нелінійне перетворення вхідного сигналу. Сигнал на виході нейрона  $Y$  залежить від величини порушення нейрона, що формується суматором  $S$  і перетворюється функцією активації  $f$ .

$$S = \sum_{j=1}^n w_j x_j, \quad Y = f\left(\sum_{j=1}^n w_j x_j\right),$$

де  $W = (w_1, w_2, \dots, w_n)$  – вектор вагових коефіцієнтів,  $\bar{X} = (x_1, x_2, \dots, x_n)$  – вектор вхідних сигналів.

Таким чином, кожний нейрон характеризується вектором коефіцієнтів  $W$  і параметрами перетворюючої функції  $f$ . Найчастіше використовується гранична функція активації:

$$F(S) = \begin{cases} 1, & S > 0 \\ 0, & S \leq 0 \end{cases}$$

Для аналітичних досліджень вибирають сигмоїдну функцію, тобто функція називається сигмоїдною, якщо вона обмежена по максимальному й мініимальному значеннях і має скрізь позитивну похідну. І крім того швидко сходиться до верхньої й нижньої меж при  $S \rightarrow +\infty$  й  $S \rightarrow -\infty$ . Графік сигмоїдної функції близький до зображення передатної характеристики біологічного нейрона.

Нейрон виконує функцію класифікації на два класи, якщо в процесі його навчання сформований вектор вагових коефіцієнтів  $W$ , значення якого зберігаються в пам'яті.

Нейромережа – набір елементів (нейронів), що виконують операції над вхідними даними, які мають властивості класифікації. В нейромережі пам'ять не локалізована, а розподілена по всій структурі. Одна з переваг подібної архітектури - "життєздатність", тобто вихід з ладу  $n$  нейронів не порушить функціонування системи. Нейромережі вирішують завдання: генерації, навчання й тестування нейромереж, визначення значимості вхідних параметрів, контрастування нейромережі, пошук мініимального вирішального набору вхідних даних. Розглянуто мережу Хопфилда та мережі, що само організуються, Кохонена.

На рисунку 1 наведена архітектура системи (Массачусетський технологічний інститут) розпізнавання мови, у якій використовують нейромережу для класифікації ділянок фонем.

Основні етапи роботи системи розпізнавання мови:

1. Сигнал оцифровується з деякою частотою дискретизації.
2. Отриманий набір відліків розбивається на вікна по 10 мс.
3. Для кожного вікна обчислюються його спектральні характеристики: 12 мел-частотних кепстральних коефіцієнтів (коефіцієнти MFCC); 12 коефіцієнтів дельта від середнього значення MFCC, які вказують ступінь спектрального відхилення; одну енергетичну характеристику, і одна дельта характеристики енергії (від середнього). Усього: для кожного вікна - 26 параметрів.
4. Щоб урахувати мінливість мови, стосовно досліджуваного вікна, розглядаються характеристики чотирьох сусідніх – що знаходяться на відстані: -60 мс, -30 мс, 30 мс і 60 мс.
5. На вхід нейромережі – багатопарового перцептрона – надходить у підсумку 130 параметрів (5 вікон по 26 характеристик у кожному). На виході нейромережі – класифікація кожного вхідного вікна, зважена в ймовірностях груп фонем (фонемні ймовірнісні характеристики).
6. Подаючи на вхід нейромережі всі контекстні вікна, формується матриця ймовірностей фонем протягом деякого часу. На рис. 2.16 вихідні дані нейромережі фіксують розпізнаване слово – "two" [tu:]. Більше темні області в "t", "t<u" і "u" указують більші ймовірності появи цих класів у цей момент часу.
7. Використовуючи алгоритм Витерби вибирається словник і знаходиться слово.
8. Слово, що відповідає кращому шляху в алгоритмі Витерби (має найбільшу суму після проходження шляху) є рішенням.

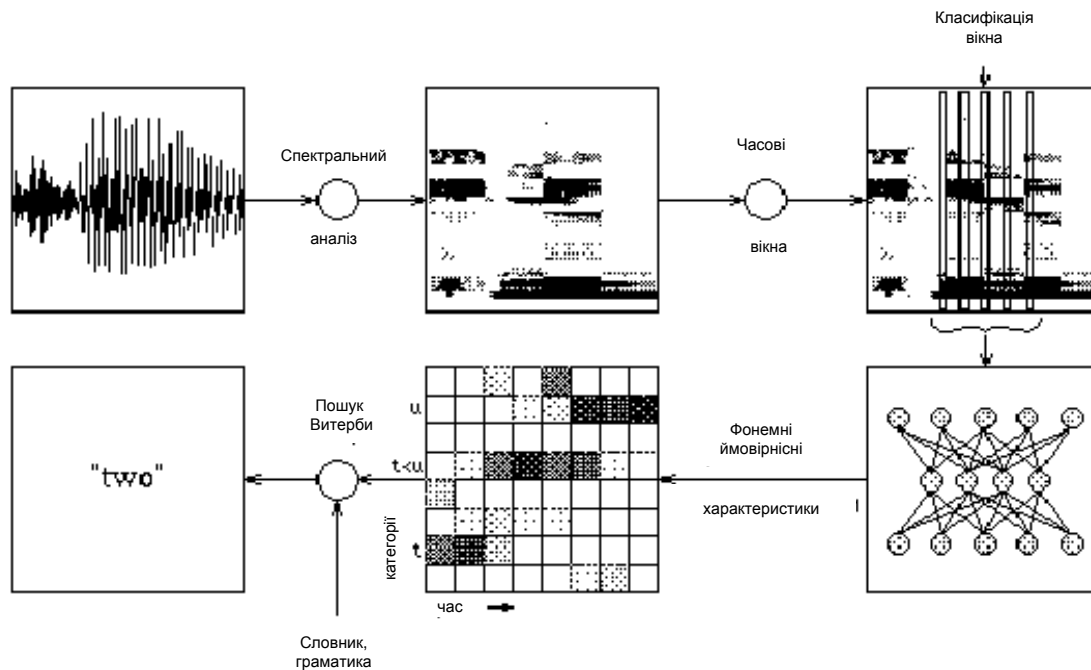


Рис. 1. Узагальнена схема системи розпізнавання мови

Великий набір різних функцій для роботи з нейромережами включений у пакет Neural Network Toolbox системи Matlab, що дозволяє не тільки задати архітектуру мережі, але й вибрати алгоритм навчання. Приведемо деякі з них:

#### 1. вибір архітектури мережі

**newhop(T)** - створення мережі Хофилда, де T - цільовий вектор,

**newp(pr,s,tf,lf)** – створення перцептрона, де pr – матриця мінімальних і максимальних значень вхідних елементів; s – число нейронів; tf – передатна функція; lf – навчальна функція,

**newff( pr , [S1 S2 S3 ] ,{'logsig' 'logsig' 'logsig'}, 'traingdx' )** - створення односпрямованої мережі, де pr – мінімальні й максимальні значення елементів; S1, S2, S3 – кількість нейронів в 1, 2 і 3 шарі; logsig – функція передачі для i-го шару; traingdx – функція навчання.

#### 2. Навчання

**initlay(Net)** – ініціалізація мережі, де Net – нейромережа,

**learngd(w,p,z,n,a,t,e,gw,ga,d,lp,ls)** - функція градієнтного спуска, де w – вагарня матриця; p – вхідний вектор; z – зважений вхідний вектор; n – мережний вхідний вектор; a – вихідне вектор; t – вектор цілей шару; e – вектор помилок шару; gw – градієнт; ga – вихідний градієнт; d - відстані нейронів; lp - параметри навчання; ls - стан навчання

#### 3. Моделювання роботи мережі

**sim(net,P)**, де net – нейромережа; P – приклади для розпізнавання

Параметри мережі:

**net.trainParam.goal** – припустима середньоквадратична погрішність

**net.trainParam.epochs** – максимальне число циклів навчання.

Основні переваги нейромереж: здатність до навчання; внутрішній паралелізм; висока відмовостійкість і завадостійкість мережі (можливість вирішувати завдання при неповних даних); асоціативна пам'ять (коефіцієнти вагової матриці утворюють асоціації й можуть змінюватися в процесі навчання).