ДОСЛІДЖЕННЯ ЕЛЕМЕНТАРНОГО ПЕРЦЕПТРОНА РОЗЕНБЛАТТА

Перцептрони – клас моделей мозку

Перцептрони або персептрони (від *perceptio* – сприйняття) були першими штучними нейронними мережами, що з'явилися у результаті багаторічних досліджень мозку тварин та людини. Автор першого перцептрона – американський вчений Френк Розенблатт, що вперше опублікував свої дослідження в цій галузі у 1957 році. На думку Ф. Розенблатта, перцептрони, насамперед, є класом моделей мозку, що пояснюють деякі його характерні функції. Зокрема, перцептрони, нехай і в самій елементарній формі, пояснюють деякі проблеми організації пам'яті біологічних систем, демонструють механізм отримання знань "систем, що пізнають (*cognitive*)", про навколишній світ та показують, що ці знання залежать як від когнітивної системи, так й від навколишнього середовища. За Розенблаттом, для різних видів тварин найпростіше подання про анатомічну структуру нервової системи може бути отримане за допомогою схеми, наведеної на рис. 2.1.

Кожний з п'яти видів інформації про зовнішнє середовище сприймається своїми спеціалізованими сенсорними нейронами й передається по своїх окремих сенсорних трактах у центральну нервову систему. Через моторні нейрони центральна нервова система пов'язана із м'язами та залозами організму. У своїх перших роботах Розенблатт розглядав модель тільки зорової системи. У найбільш простому вигляді ця модель містить у собі три послідовно з'єднані безлічі нейронів: чутливі (*S*-елементи); нейрони, що асоціюють (*A*-елементи), та реагуючі (*R*-елементи). *S*-елементам у нервовій системі тварини або людини відповідають сенсорні або рецепторні нейрони, які генерують сигнали на зовнішні подразнення, що надходять (зображення), та передають їх *A*-нейронам. *A*-елементи у нервовій системі живого організму аналогічні нейронам, що утворюють локальний спеціалізований зоровий центр у корі головного мозку та єднають рецепторні нейрони із моторними. *R*-елементам у нервовій системі відповідають ефекторні (моторні) нейрони, які упорядковані в обмежені топологічні структури і які пересилають сигнали керування центральної нервової системи до м'язів й залоз організму.

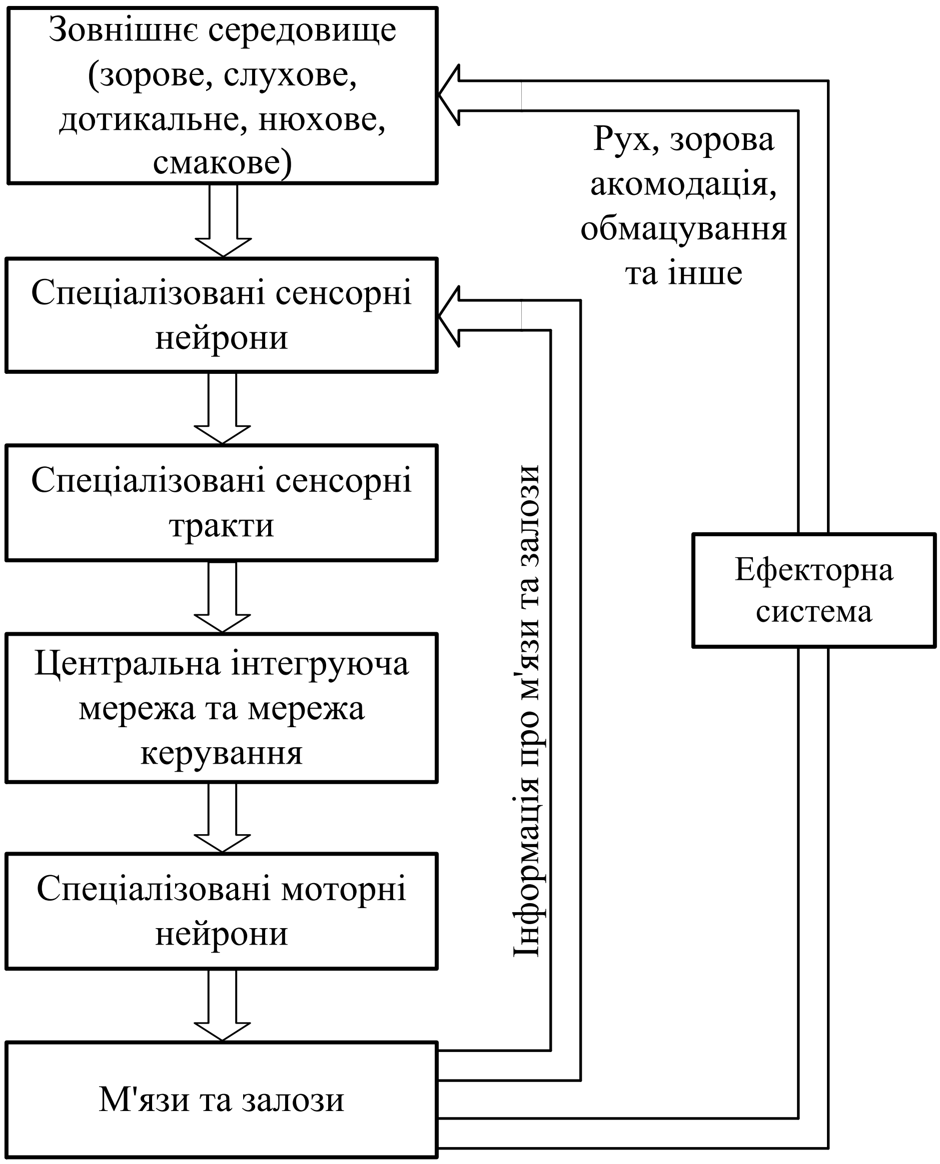

Рис. 2.1. Структура нервової системи

Визначення 2.1. *S*-елемент називається простим, якщо він видає одиничний вихідний сигнал при вхідному сигналі, що перевищує деякий заданий поріг θ, та нульовий сигнал – у противному випадку.

Визначення 2.2. Простим асоціативним елементом називається *A*-елемент, що видає одиничний вихідний сигнал, якщо алгебраїчна сума його вхідних сигналів перевищує деякий заданий поріг $\theta > 0$, інакше – вихідний сигнал асоціативного нейрона дорівнює нулю.

Визначення 2.3. Простим біполярним (бінарним) реагуючим елементом називається *R*-елемент, що видає одиничний вихідний сигнал, якщо алгебраїчна сума його вхідних сигналів більша за граничне значення або дорівнює йому, та від’ємний одиничний (нульовий) сигнал, якщо сума його вхідних сигналів менша від заданого порога.

Чутливі *S*-елементи живого організму (рис. 2.2) збуджуються від впливу енергії світла, якщо величини їх вхідних сигналів перевищують деякий поріг θ_i. Рецепторні нейрони випадковим чином пов'язані із *A*-елементами, вихідні сигнали яких відмінні від нуля тільки у тому випадку, коли збуджена досить велика кількість сенсорних нейронів, які впливають на входи одного елемента, що асоціює. Простий *A*-елемент, аналогічно простому *S*-елементу, є активним та видає одиничний вихідний сигнал, якщо алгебраїчна сума сигналів на його вході перевищує задану порогову величину, у противному випадку нейрон перебуває не у збудженому стані. Коефіцієнти (ваги) зв'язків між *S*- та *A*-елементами постійні.

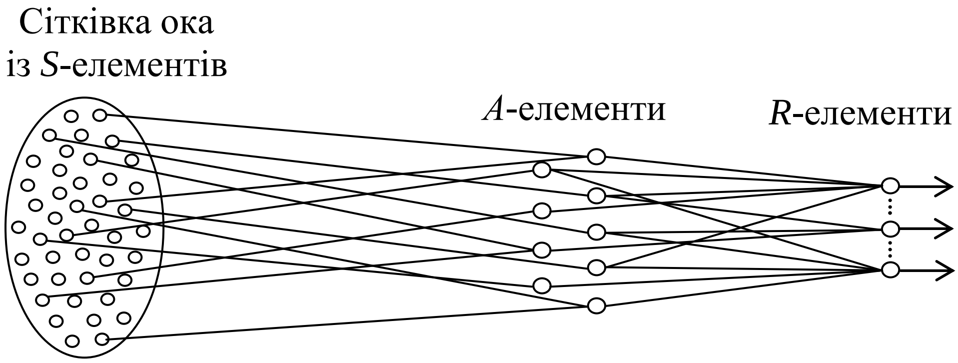

Рис. 2.2. Структура моделі зорової системи

Комбінація виходів всіх *A*-елементів являє собою реакцію двох перших шарів перцептрона на пред'явлене вхідне зображення, що за допомогою вихідного шару нейронів перетвориться в необхідну комбінацію вихідних сигналів системи. Часто вимагають, щоб кожному класу вхідних зображень відповідав тільки один певний активний *R*-нейрон. Необхідних комбінацій вихідних сигналів на кожний клас зображень домагаються на етапі навчання або адаптації перцептрона за рахунок зміни змінних ваг зв'язків між *A*- та *R*-елементами.

Поділ множини *G* зображень на два класи G_1 та G_2 можна виконати за допомогою одного вихідного елемента. У цьому випадку зображенням першого класу може відповідати додатний вихідний сигнал (+1) *R*-елемента, а другого класу – від'ємний (–1). На прикладі найпростішого (елементарного) перцептрона розглянемо різні способи навчання цих нейромереж, уперше запропонованих та досліджених Розенблаттом.

Визначення 2.4. Простим перцептроном називається нейронна мережа, що складається із *S*-, *A*- й *R*-елементів та задовольняє такі п'ять умов:

1. У мережі є тільки один *R*-нейрон, що з'єднаний зв'язками зі змінними вагами з усіма *A*-нейронами.
2. У мережі є тільки послідовні зв'язки від *S*- до *A*-елементів та від *A*-елементів до *R*-елемента.
3. Ваги зв'язків між *S*- та *A*-елементами є фіксованими.
4. Час передачі сигналів кожним окремим зв'язком дорівнює нулю (або фіксованій постійній величині).
5. Вихідні сигнали всіх нейронів мережі формуються у такому вигляді:

$$U_{\text{вих.}} = f\left(\sum_i U_{\text{вх.}i}(t)\right),$$

де $\sum_i U_{\text{вх.}i}(t)$ – алгебраїчна сума сигналів, що надходять одночасно на вхід нейрона.

Визначення 2.5. Елементарним перцептроном називається простий перцептрон із простими *A*- та *R*-елементами й передатними функціями зв'язків вигляду

$$C_{ij}(t) = w_{ij}(t) U_{\text{вих.}i}(t - \tau_{ij}),$$

де $w_{ij}(t)$ – вага зв'язку між *i*-м та *j*-м нейронами в момент часу *t*; $U_{\text{вих.}i}(t - \tau_{ij})$ – вихідний сигнал *i*-го нейрона в момент часу $(t - \tau_{ij})$; τ_{ij} – час передачі сигналу $U_{\text{вих.}i}(t - \tau_{ij})$ з виходу *i*-го нейрона на вхід *j*-го елемента.

Елементарний перцептрон навчається або настроюється на розпізнавання двох класів зображень G_1, G_2 шляхом пред'явлення йому деяких послідовностей зображень із цих класів. Учитель (людина або обчислювальна машина), що спостерігає реакцію перцептрона на кожне вхідне зображення, при наявності помилкових рішень мережі повинен корегувати ваги зв'язків між *R*- та *A*-елементами відповідно до деякої системи правил.

Визначення 2.6. Матрицею взаємодії перцептрона називається матриця, елементами якої є ваги зв'язків w_{ij} для всіх пар нейронів U_i, U_j мережі.

Якщо зв'язок між нейронами U_i, U_j відсутній (наприклад, у простому перцептроні немає зв'язків між *R*- та *S*-нейронами), то припускають, що $w_{ij} = 0$.

Матриця взаємодії фактично відображає стан пам'яті перцептрона. Множина всіх можливих станів пам'яті мережі утворить фазовий простір мережі, що може бути представлений у вигляді області в *n*-мірному евклідовому просторі, кожна координатна вісь якого відповідає одному зв'язку мережі.

2.2.2. Навчання перцептронів α- та γ-системи підкріплень

Визначення 2.7. Системою підкріплень нейронної мережі називається будь-який набір правил, за допомогою яких можна змінювати у часі стан пам'яті мережі (або матрицю взаємодії).

Визначення 2.8. Додатним (від’ємним) підкріпленням називається такий процес корекції ваг зв'язків, при якому вага зв'язку $w_{ij}(t)$, що починається на виході активного *i*-го елемента й закінчується на вході *j*-го елемента, змінюється на величину $\Delta w_{ij}(t)$, знак якої збігається зі знаком вихідного сигналу *j*-го нейрона (знак якої протилежний знаку вихідного сигналу *j*-го нейрона).

Існує велика кількість різних систем підкріплень, більша частина з яких становить лише історичний інтерес. Тому зупинимося тільки на системі підкріплень з корекцією похибок, що є основною на цей час.

У системі підкріплень з корекцією похибок насамперед необхідно визначити, чи є реакція перцептрона правильною. Доти, поки вихідний сигнал *R*-елемента набуває бажаного значення, величина сигналу підкріплення η дорівнює нулю. З появою неправильної реакції перцептрона використовується підкріплення, величина та знак якого в загальному випадку визначається монотонно зростаючою функцією f

$$\eta = f(R^* - R), \tag{2.1}$$

де R^* – бажана реакція; R – отримана реакція; $f(0) = 0$.

Таким чином, з появою помилки для корекції ваг зв'язків використовується сигнал, знак якого протилежний знаку вихідного сигналу R-елемента. У зв'язку із цим розглянутий метод корекції ваг одержав назву системи з від’ємним підкріпленням.

Конкретним прикладом системи підкріплень з корекцією похибок є альфа-система підкріплень. У цій системі при наявності похибок ваги всіх активних зв'язків, які скінчуються на R-елементі, змінюються на однакову величину η, а ваги всіх неактивних зв'язків залишаються без змін. Перцептрони, у яких застосовується альфа-система підкріплення, називаються альфа-перцептронами.

При використанні альфа-системи підкріплень сума ваг всіх зв'язків між R- та A-нейронами може зростати (або убувати) від кроку до кроку, що повинно приводити до небажаних ситуацій, коли багато зв'язків мають максимальні (або мінімальні) ваги й не можуть використовуватися надалі у процесі навчання нейронної мережі. Для усунення цього недоліку α-системи підкріплень була запропонована гамма-система підкріплень, що має властивість консервативності щодо суми Σ_1 ваг всіх зв'язків між нейронами, тобто сума Σ_1 залишається постійною в процесі навчання перцептрона. Це досягається за рахунок того, що при наявності помилкової реакції перцептрона спочатку ваги всіх активних зв'язків змінюються на однакове значення η, а услід за цим з ваг всіх активних та пасивних зв'язків віднімається величина, яка дорівнює відношенню суми зміни ваг всіх активних зв'язків до числа всіх зв'язків. Зміна ваг окремих зв'язків при цьому визначається співвідношенням

$$\Delta w_{ij} = \begin{cases} \eta - \dfrac{N_{\text{ак}}\eta}{N}, & \text{якщо } i\text{-й } A\text{-нейрон збуджений,} \\ -\dfrac{N_{\text{ак}}\eta}{N}, & \text{якщо } i\text{-й } A\text{-нейрон загальмований,} \end{cases} \tag{2.2}$$

де Δw_{ij} – у загальному випадку приріст ваги зв'язку між i-м A-нейроном та j-м R-нейроном, для елементарного перцептрона j = const = 1; η – величина сигналу підкріплення; $N_{\text{ак}}$ – число активних зв'язків; N – число зв'язків, що закінчуються на вході j-го елемента.

При такій системі корекції ваг зв'язків виконується рівність

$$\eta N_{\text{ак}} - \frac{N_{\text{ак}}\eta}{N} \cdot N = 0\,,$$

з якої й виходить консервативність гамма-системи підкріплень відносно суми всіх ваг зв'язків, що навчаються.

Зауваження 2.1. Відзначимо, що співвідношення (2.2) у неявній формі передбачає, що ваги w_{ij} зв'язків, що корегуються, досить далекі від своїх граничних значень $w_{ij\,\min}= 0$ й $w_{ij\,\max} = 1$, тобто

$$w_{ij\,\min} \le w_{ij} + \Delta w_{ij} \le w_{ij\,\max}\,. \tag{2.3}$$

Якщо нерівності (2.3) порушуються, а вимога консервативності щодо суми Σ_1 ваг зв'язків залишається незмінною, то співвідношення (2.2) необхідно уточнити. Нехай, наприклад, серед активних зв'язків $N_{\text{а гр.}}$ зв'язків мають граничні значення ваг $w_{ij\,\max}$ або $w_{ij\,\min}$ й для них виконуються умови

$$w_{ij\,\max} + \Delta w_{ij} > w_{ij\,\max} \quad \text{або} \quad w_{ij\,\min} + \Delta w_{ij} < w_{ij\,\min}\,. \tag{2.4}$$

Нехай також $N_{\text{а бгр.}}$ активних зв'язків мають ваги, близькі до граничних, для яких справедливі такі нерівності:

$$w_{ij} + \Delta w_{ij} > w_{ij\,\max} \quad \text{або} \quad w_{ij} + \Delta w_{ij} < w_{ij\,\min}\,. \tag{2.5}$$

У цьому випадку загальна сума $S_{\text{а}}$ первісних змін ваг активних зв'язків буде становитиме

$$S_{\text{а}} = (N_{\text{ак}} - N_{\text{а гр.}} - N_{\text{а бгр.}})\eta + \text{sign}(\eta) \sum_{k=1}^{N_{\text{а бгр.}}} \left| w_{kj}^{\text{гр.}} - w_{kj} \right|, \quad (2.6)$$
$$\left| w_{kj}^{\text{гр.}} - w_{kj} \right| < \left| \Delta w_{kj} \right|,$$

де $w_{kj}^{\text{гр.}}$ – граничне значення ваги зв'язку між k-м та j-м нейронами, $w_{kj}^{\text{гр.}} \in \{0,1\}$; Δw_{kj} – зміна ваги зв'язку, яка обчислюється за співвідношенням (2.2) без врахування наявності множини $\{w_{ij\,\min},\ w_{ij\,\max}\} \equiv \{0,1\}$; $\text{sign}(\cdot)$ – знакова функція.

Зміна ваг окремих зв'язків при цьому може бути визначена співвідношенням

$$\Delta w_{ij} = \begin{cases} \eta_i - \dfrac{S_{\text{а}}}{N}, & \text{якщо } i\text{-й } A\text{-нейрон збуджений,} \\ -\dfrac{S_{\text{а}}}{N}, & \text{якщо } i\text{-й } A\text{-нейрон загальмований.} \end{cases}$$

Прикладом ще одного загального способу навчання перцептронів є метод корекції похибок випадковими збурюваннями. Він передбачає, як і у альфа-системи підкріплень, з появою помилок – корекцію ваг активних зв'язків, але знак та величина корекції для кожного зв'язку вибирається випадково відповідно до деякого заданого розподілу ймовірностей.

Приклад 2.1. Навчимо елементарний перцептрон з бінарними *S*- та *A*-нейронами й біполярним *R*-нейроном (рис. 2.3) розпізнавати зображення букв Н і П (рис. 2.4, *а*, *б*) на рецепторному полі з дев'яти елементів (рис. 2.4, *в*).

При цьому зажадаємо, щоб при пред'явленні зображення букви Н на виході *R*-елемента був сигнал “–1”, з появою другого зображення – сигнал “+1”.

Задамо у таблицях 2.1 та 2.2 ваги зв'язків w_{ij}^{1} ($i = \overline{1,9}$, $j = \overline{1,6}$), w_{k}^{2} ($k = \overline{1,6}$) відповідно між бінарними *S*- та *A*-нейронами й між *A*-нейронами та біполярним нейроном *R* за допомогою генератора випадкових чисел, що генерує їх із скінченої множини $\{0,1;\ 0,2;\ \ldots;\ 0,9\}$.

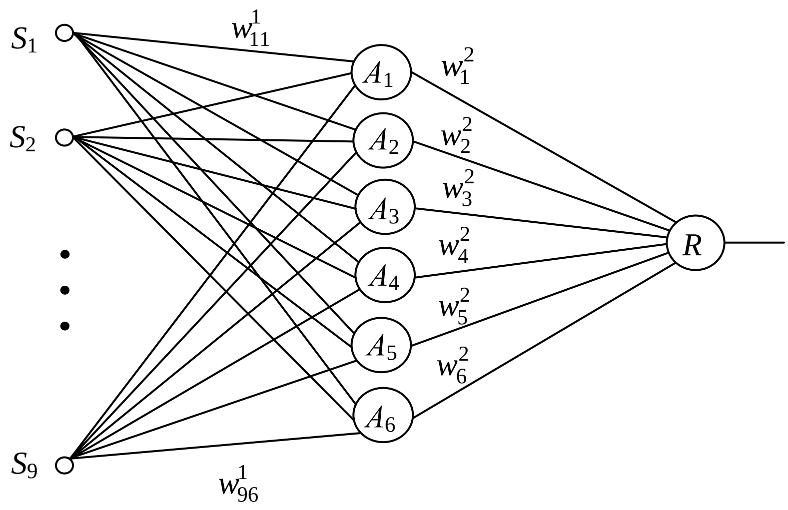

Рис. 2.3. Елементарний перцептрон

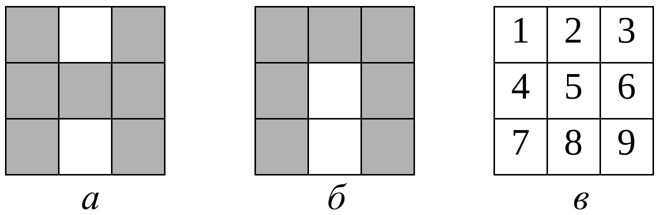

Рис 2.4. Зображення букв Н та П

Таблиця 2.1 – Ваги w_{ij}^1 зв'язків перцептрона між *S*- та *A*-елементами

w_{ij}^1	S_1	S_2	S_3	S_4	S_5	S_6	S_7	S_8	S_9
A_1	0,3	0,2	0,1	0,6	0,5	0,4	0,9	0,6	0,7
A_2	0,2	0,1	0,3	0,4	0,5	0,6	0,7	0,1	0,9
A_3	0,3	0,5	0,1	0,6	0,5	0,4	0,9	0,4	0,7
A_4	0,4	0,3	0,2	0,1	0,8	0,7	0,6	0,6	0,9
A_5	0,5	0,3	0,3	0,6	0,1	0,2	0,9	0,2	0,7
A_6	0,6	0,5	0,4	0,1	0,2	0,3	0,8	0,5	0,8

Таблиця 2.2 – Ваги w_k^2 зв'язків перцептрона між *R*- та *A*-елементами

w_k^2	A_1	A_2	A_3	A_4	A_5	A_6
R	0,2	0,8	0,6	0,9	0,8	0,1

Подамо на вхід перцептрона зображення букви Н (рис 2.4, *а*). Це зображення збуджує всі *S*-нейрони, крім другого та восьмого. Одиничні сигнали з виходів збуджених бінарних *S*-нейронів через зв'язки, вагові коефіцієнти яких задані табл. 2.1, надходять на входи A-нейронів. Сумарний вхідний сигнал на вході i-го A-елемента визначається співвідношенням

$$U_{\text{вх.}\,Ai} = \sum_{j=1}^{9} U_{\text{вих.}\,Sj} w_{ji}^{1}, \;\; i = \overline{1,6}, \tag{2.7}$$

де $U_{\text{вх.}Ai}$ – сигнал на вході i-го A-нейрона; $U_{\text{вых.}Sj}$ – сигнал на виході j-го S-нейрона; w_{ji} – вага зв'язку між j-м S-нейроном та i-м A-елементом.

Для першого A-нейрона маємо

$$U_{\text{вх.}A1} = \sum_{j=1}^{9} U_{\text{вих.}Sj} w_{j1}^{1} = 1 \cdot 0{,}3 + 0 \cdot 0{,}2 + 1 \cdot 0{,}1 + 1 \cdot 0{,}6 + 1 \cdot 0{,}5 + 1 \cdot 0{,}4 +$$
$$+ 1 \cdot 0{,}9 + 0 \cdot 0{,}6 + 1 \cdot 0{,}7 = 3{,}5.$$

Аналогічно обчислюються сигнали на входах інших A-елементів. Результати цих обчислень наведено в другому рядку табл. 2.3. У третьому рядку цієї таблиці – результати розрахунків сигналів на входах A-елементів при пред'явленні перцептрону зображення букви П.

Таблиця 2.3 – Величини сигналів на входах A-елементів

Зображення	Сигнали на входах A-елементів					
	$U_{\text{вх.}A1}$	$U_{\text{вх.}A2}$	$U_{\text{вх.}A3}$	$U_{\text{вх.}A41}$	$U_{\text{вх.}A5}$	$U_{\text{вх.}A6}$
Буква Н	3,5	3,6	3,5	3,7	3,3	3,2
Буква П	3,2	3,2	3,4	3,2	3,6	3,5

Для спрощення розрахунків припустимо, що пороги θ_i, $i = \overline{1,6}$ всіх A-нейронів однакові:

$$\theta_1 = \theta_2 = \ldots = \theta_6 = \theta.$$

Якщо величина порога θ обрана менше 3,2, то при пред'явленні будь-якого зображення будуть збуджені всі A-нейрони, а якщо вибрати $\theta > 3{,}7$, то на виходах всіх нейронів будуть нульові сигнали. В обох цих випадках перцептрон не може виконувати розпізнавання пропонованих зображень.

Очевидно, що для забезпечення працездатності нейронної мережі поріг θ необхідно вибрати між 3,2 та 3,7 й таким чином, щоб при пред'явленні різних зображень збуджувалися різні множини M_1, M_2 A-елементів, причому бажано, щоб ці множини не перетиналися, тобто

$$M_1 \cap M_2 = 0. \tag{2.8}$$

Нехай вихідний сигнал A-елементів визначається співвідношенням

$$U_{\text{вих}.A} = \begin{cases} 1, & \text{якщо}\, U_{\text{вх}.A} \geq \theta, \\ 0, & \text{якщо}\, U_{\text{вх}.A} < \theta, \end{cases}$$

тоді умова (2.8) виконується при $\theta = 3{,}5$ й при пред'явленні зображення букви Н будуть збуджені елементи A_1, A_2, A_3 й A_4, а при пред'явленні букви П – нейрони A_5 й A_6. Розрахуємо з урахуванням даних табл. 2.2 сигнали $U_{\text{вх}.R\boldsymbol{Н}}$, $U_{\text{вх}.R\boldsymbol{П}}$ на вході R-нейрона при пред'явленні зображень букв Н и П:

$$U_{\text{вх}.R\boldsymbol{Н}} = \sum_{i=1}^{6} U_{\text{вих}.Ai} w_i^2 = 1 \cdot 0{,}2 + 1 \cdot 0{,}8 + 1 \cdot 0{,}6 + 1 \cdot 0{,}9 + 0 \cdot 0{,}8 + 0 \cdot 0{,}1 = 2{,}5,$$

$$U_{\text{вх}.R\boldsymbol{П}} = \sum_{i=1}^{6} U_{\text{вих}.Ai} w_i^2 = 0 \cdot 0{,}2 + 0 \cdot 0{,}8 + 0 \cdot 0{,}6 + 0 \cdot 0{,}9 + 1 \cdot 0{,}8 + 1 \cdot 0{,}1 = 0{,}9.$$

При величині порога R-елемента $\theta_R = 1{,}7$ й пред'явленні зображення букви Н на виході перцептрона буде сигнал “+1”, а при пред'явленні другого зображення – сигнал “–1”, що не відповідає вихідним вимогам до вихідних сигналів нейронної мережі. Використовуємо для настроювання перцептрона α-систему підкріплень при величині сигналу підкріплення η, що дорівнює 0,1, й при пред'явленні послідовності зображень Н, П, Н, П, ... у моменти часу t_1, t_2, t_3, ... Процес адаптації ваг зв'язків між R- та A-нейронами ілюструється в табл. 2.4.

У другому стовпці цієї таблиці при $t = t_0$ наведено значення вихідних ваг зв'язків та величини сигналів $U_{\text{вх}.R\text{Н}}$, $U_{\text{вх}.R\text{П}}$ на вході R-елемента при пред'явленні відповідно зображень букв Н та П. При першому пред'явленні зображення букви Н у момент часу t_1 (позначено $t_1^{\text{Н}}$) у силу наявності помилкового сигналу на виході перцептрона корегуються ваги активних зв'язків $w_1^2,\dots,w_4^2$ на величину η = 0,1. Ця корекція зменшує сумарний вхідний сигнал $U_{\text{вх.}R\text{Н}}$ до величини 2,1.

Нехай функціонування *R*-елемента описується співвідношенням

$$U_{\text{вих.}R} = \begin{cases} +1, & \text{якщо } U_{\text{вх.}R} \ge \theta_R, \\ -1, & \text{якщо } U_{\text{вх.}R} < \theta_R, \end{cases}$$

де θ_R – поріг *R*-елемента, тоді для досягнення правильної реакції *R*-елемента на зображення букви Н необхідні дві повторні корекції ваг зв'язків $w_1^2,\dots,w_4^2$. Результати цих корекцій наведено в табл. 2.4 відповідно в п'ятому й сьомому стовпцях при $t = t_3^{\text{Н}}$ й $t = t_5^{\text{Н}}$.

Таблиця 2.4 – Адаптація ваг зв'язків перцептрона за допомогою α-системи підкріплень

Вагові коефіцієнти та вхідні сигнали	Моменти часу									
	t_0	$t_1^{\text{Н}}$	$t_2^{\text{П}}$	$t_3^{\text{Н}}$	$t_4^{\text{П}}$	$t_5^{\text{Н}}$	$t_6^{\text{П}}$	$t_7^{\text{П}}$	$t_8^{\text{П}}$	$t_9^{\text{П}}$
w_1^2	0,2	0,1	0	0	0	0	0	0	0	0
w_2^2	0,8	0,7	0	0,6	0	0,5	0	0	0	0
w_3^2	0,6	0,5	0	0,4	0	0,3	0	0	0	0
w_4^2	0,9	0,8	0	0,7	0	0,6	0	0	0	0
w_5^2	0,8	0	0,9	0	1	0	1	1	1	1
w_6^2	0,1	0	0,2	0	0,3	0	0,4	0,5	0,6	0,7
$U_{\text{вх.}R\text{Н}}$	2,5	2,1	–	1,7	–	1,4	–	–	–	–
$U_{\text{вх.}R\text{П}}$	0,9	–	1,1	–	1,3	–	1,4	1,5	1,6	1,7

Після моменту часу $t = t_5^{\text{Н}}$, через те що виконується співвідношення (2.8), із вхідної послідовності можуть бути виключені зображення букви Н та можуть пред'являтися тільки зображення букви П. Результати корекції ваг зв'язків w_5^2, w_6^2, що визначають сигнал на вході *R*-елемента при пред'явленні зображення букви П, наведено в останньому рядку табл. 2.4. Оскільки в розглянутому прикладі корекція вхідного сигналу *R*-елемента при пред'явленні зображення букви П здійснюється тільки за допомогою ваг двох зв'язків, причому, після другої корекції вага зв'язку w_5^2 набуває максимального значення та надалі зростати не може, то процес навчання нейронної мережі правильної реакції на друге зображення більш тривалий та закінчується тільки при $t = t_9^{\text{П}}$.

У табл. 2.5 наведено результати настроювання елементарного перцептрона при тих же вихідних даних, але за допомогою γ-системи підкріплень.

Таблиця 2.5 – Адаптація ваг зв'язків перцептрона за допомогою γ-системи підкріплень

Вагові коефіцієнти та вхідні сигнали	Моменти часу									
	t_0	$t_1^{\text{Н}}$	$t_2^{\text{П}}$	$t_3^{\text{Н}}$	$t_4^{\text{П}}$	$t_5^{\text{Н}}$	$t_6^{\text{П}}$	$t_7^{\text{Н}}$	$t_8^{\text{П}}$	$t_9^{\text{Н}}$
w_1^2	0,2	0,1666	0,1333	0,1000	0,0800	0,0600	0,0400	0,0200	0,0000	0,0000
w_2^2	0,8	0,7666	0,7333	0,7000	0,6800	0,6600	0,6400	0,6200	0,6000	0,5750
w_3^2	0,6	0,5666	0,5333	0,5000	0,4800	0,4600	0,4400	0,4200	0,4000	0,3750
w_4^2	0,9	0,8666	0,8333	0,8000	0,7800	0,7600	0,7400	0,7200	0,7000	0,6750
w_5^2	0,8	0,8667	0,9341	1,0000	1,0000	1,0000	1,0000	1,0000	1,0000	1,0000
w_6^2	0,1	0,1667	0,2341	0,3004	0,1380	0,4604	0,5404	0,6204	0,7004	0,7754
$U_{\text{вх.}R\text{Н}}$	2,5	2,3664	–	2,1000	–	1,9400	–	1,7800	–	1,6250
$U_{\text{вх.}R\text{П}}$	0,9	–	1,1682	–	1,3800	–	1,5404	–	1,7004	1,7754
Σ_1	3,4	3,3998	3,4004	3,4004	3,4000	3,4004	3,4004	3,4004	3,4004	3,4004

У другому стовпці табл. 2.5 при $t = t_0$ наведено вихідні ваги зв'язків та величини сигналів на вході R-елемента при пред'явленні зображень букв Н та П, а також сума Σ_1 ваг всіх зв'язків між R- та A-нейронами. Значення ваг зв'язків у третьому стовпці таблиці отримано після пред'явлення зображення букви Н у момент часу $t = t_1^{\text{Н}}$. Оскільки $\eta = 0{,}1$, $N_{\text{ак}} = 4$ й $N = 6$, то, використовуючи співвідношення (2.2) для розрахунку збільшення ваг активних зв'язків, одержимо

$$\Delta w_1 = \Delta w_2 = \Delta w_3 = \Delta w_4 = \eta - N_{\text{ак}}\eta / N = 0{,}1 - 4 \cdot 0{,}1 / 0{,}6 = 0{,}0334, \qquad (2.9)$$

а для зміни ваг пасивних зв'язків маємо

$$\Delta w_5 = \Delta w_6 = -N_{\text{ак}}\eta / N = -4 \cdot 0{,}1 / 0{,}6 = 0{,}0667. \qquad (2.10)$$

Знаючи збільшення ваг зв'язків та використовуючи співвідношення

$$w_i^2(t_1^{\text{Н}}) = w_i^2(t_0) - \Delta w_i\ , \quad i = \overline{1{,}6},$$

неважко одержати й чисельні значення, наведені у третьому стовпці табл. 2.5.

При пред'явленні зображення букви П у момент часу $t = t_2^{\text{П}}$ активними є тільки нейрони A_5 та A_6, тому $N_{\text{ак}}$ = 2 і співвідношення (2.2) дає наступні чисельні значення збільшень ваг зв'язків:

$$\Delta w_5 = \Delta w_6 = \eta - N_{\text{ак}}\eta / N = 0{,}1 - 2 \cdot 0{,}1 / 0{,}6 = 0{,}0667, \qquad (2.11)$$

$$\Delta w_1 = \ldots = \Delta w_4 = -N_{\text{ак}}\eta / N = -2 \cdot 0{,}1 / 0{,}6 = 0{,}0333. \qquad (2.12)$$

Знаючи збільшення Δw_i $(i = \overline{1{,}6})$ й використовуючи вираз

$$w_i^2(t_2^{\text{П}}) = w_i^2(t_1^{\text{Н}}) + \Delta w_i\ , \quad i = \overline{1{,}6}\,,$$

нескладно одержати дані четвертого стовпця табл. 2.5.

При розрахунку збільшень ваг зв'язків при $t = t_3^{\text{Н}}$ за співвідношеннями (2.9), (2.10) виявляється, що збільшення Δw_5 більше, ніж можлива зміна ваги зв'язку w_5^2:

$$\Delta w_5 = -N_{\text{ак}}\eta / N = -4 \cdot 0{,}1 / 0{,}6 = 0{,}0667 > 1 - w_5^2(t_2^{\text{П}}) = 1 - 0{,}9341 = 0{,}0659.$$

Тому для виконання умови консервативності щодо суми Σ_1 ваг зв'язків необхідно величину різниці

$$\Delta_1 w_5 = \Delta w_5 - \left|1 - w_5^2(t_2^{\text{П}})\right| = 0{,}0667 - 0{,}0659 = 0{,}0008$$

використовувати для зміни ваг зв'язків, які не набули граничних значень. Один з можливих способів використання різниці $\Delta_1 w_5$ – змінити кожний з таких (N – 1) активних та пасивних зв'язків на величину

$$\eta_1 = \frac{\Delta_1 w_5}{N-1}.$$

Чисельні значення ваг зв'язків при $t = t_3^{\text{Н}}$ наведено у п'ятому стовпці табл. 2.5.

При розрахунку ваг зв'язків за допомогою виразів (2.11), (2.12) для $t = t_4^{\text{П}}$ необхідно враховувати, що змінюватися може вага тільки одного активного зв'язку:

$$\Delta w_5 = 0, \tag{2.13}$$

$$\Delta w_6 = \eta - \frac{N_{\text{ак}} - 1}{(N-1)}\eta = 0{,}1 - 1 \cdot 0{,}1/5 = 0{,}0800, \tag{2.14}$$

$$\Delta w_1 = \ldots = \Delta w_4 = (N_{\text{ак}} \text{ -}1)\eta/(N-1) = -1 \cdot 0{,}1/5 = -0{,}0200. \tag{2.15}$$

Аналогічно при $t = t_5^{\text{Н}}$ маємо

$$\Delta w_1 = \ldots = \Delta w_4 = -\eta + N_{\text{ак}}\eta/(N-1) = -0{,}1 + 4 \cdot 0{,}1/5 = -0{,}0200, \tag{2.16}$$

$$\Delta w_5 = 0, \tag{2.17}$$

$$\Delta w_6 = N_{\text{ак}}\eta/(N-1) = 4 \cdot 0{,}1/5 = 0{,}0800. \tag{2.18}$$

Використовуючи співвідношення (2.13) – (2.18), аналогічним образом розраховуються ваги зв'язків при $t = t_6^{\text{П}}, t_7^{\text{Н}}, t_8^{\text{П}}$. Коли $t = t_9^{\text{Н}}$ число активних *A*-нейронів при пред'явленні зображення букви Н буде дорівнювати чотирьом, але у виразі (2.16) необхідно використовувати $N_{\text{ак}}$ = 3, оскільки ваговий коефіцієнт w_1^2 = 0. Зменшується до чотирьох й число ваг зв'язків, які використовуються для забезпечення постійності суми Σ_1 ваг всіх змінюваних зв'язків перцептрона.

У результаті одержуємо

$$\Delta w_1 = \Delta w_5 = 0,$$
$$\Delta w_2 = \Delta w_3 = \Delta w_4 = -\eta + (N_{\text{ак}} - 1)\eta/(N-2) = -0{,}1 + 3\cdot 0{,}1/4 = -0{,}025,$$
$$\Delta w_6 = (N_{\text{ак}} - 1)\eta/(N-2) = 3\cdot 0{,}1/(6-2) = 0{,}075,$$
$$U_{\text{вх}.R\text{Н}} = 1{,}6250 < \theta_R = 1{,}7, \;\; U_{\text{вх}.R\text{П}} = 1{,}7754 > \theta_R = 1{,}7.$$

Таким чином, при пред'явленні букви Н на вході R-елемента сигнал менший від величини порога θ_R й, отже, на виході R-нейрона буде необхідний сигнал “–1”, а при пред'явленні зображення букви П на вихідному нейроні з'явиться заданий сигнал “+1”.

Виходячи з співвідношення (2.2) в γ-системі підкріплень при тій самій величині сигналу підкріплення η, що й у α-системі, збільшення Δw_{ij} ваги активного зв'язку, що корегується, менше, ніж в α-системі. У зв'язку з цим можливо очікувати, що у загальному випадку процес настроювання нейронної мережі другим методом вимагає більшого числа ітерацій. Однак аналіз даних табл. 2.4 та 2.5 розглянутого прикладу показує, що число ітерацій при використанні γ-системи підкріплень не більше, ніж при застосуванні α-системи. Це пояснюється наступним. Якщо ваги всіх зв'язків далекі від граничних значень, то загальна сума S_α зміни ваг в α-системі пов'язана тільки з активними зв'язками:

$$S_\alpha = \eta N_{\text{ак}}.$$

В γ-системі аналогічна сума S_γ може дорівнювати, бути менше або більше S_α, тому що на кожній ітерації коректуються ваги всіх зв'язків:

$$S_\gamma = (\eta - \frac{N_{\text{ак}}\eta}{N})N_{\text{ак}} + \frac{N_{\text{ак}}\eta}{N}(N - N_{\text{ак}}).$$

За допомогою першого доданка в цьому виразі підраховується сума зміни ваг активних зв'язків, а за допомогою другого – сума зміни ваг пасивних зв'язків. Після перетворень маємо

$$S_{\gamma} = N_{\text{ак}}\eta + (N_{\text{ак}} - \frac{2N_{\text{ак}}^2}{N})\eta. \quad (2.19)$$

З аналізу виразу (2.19) виходить, що залежно від співвідношення величин $N_{\text{ак}}$ та N можливі три вирази:

$$S_{\gamma} = S_{\alpha}, \text{ якщо } N_{\text{ак}} = 0{,}5N, \quad (2.20)$$

$$S_{\gamma} > S_{\alpha}, \text{ якщо } N_{\text{ак}} < 0{,}5N, \quad (2.21)$$

$$S_{\gamma} < S_{\alpha}, \text{ якщо } N_{\text{ак}} > 0{,}5N. \quad (2.22)$$

Зі співвідношень (2.20) – (2.22) можна зробити висновок, що в загальному випадку за числом ітерацій у процесі навчання елементарного перцептрона жодна з розглянутих систем підкріплень не має помітної переваги.

2.2.3. Теореми Розенблатта про елементарний перцептрон

Перша теорема Розенблатта доводить існування елементарного перцептрона, здатного виконати будь-яку класифікацію заданої множини чорно-білих зображень, тобто вона показує, що перцептрон є універсальним пристроєм для вирішення будь-якого завдання класифікації зображень.

Теорема 2.1. Нехай дана множина $W = \{W_1, \ldots, W_n\}$ чорно-білих зображень на деякій сітківці $S = \{S_1, \ldots, S_m\}$, тоді для будь-якої класифікації $C(W)$ множини W чорно-білих зображень на дві підмножини W^1, W^2 існує не порожній клас елементарних перцептронів із сітківкою S, здатних виконати цю класифікацію.

Доведення. Для доведення досить показати існування хоча б одного елементарного перцептрона, здатного виконати довільну класифікацію $C(W)$. Розглянемо перцептрон, кожному зображенню W_k $(k = \overline{1,n})$ на сітківці S якого відповідає один A-елемент – нейрон A_k, функціонування якого визначається виразом

$$U_{\text{вих}.Ak} = \begin{cases} 1, \text{ якщо } U_{\text{вх}.Ak} \ge \theta, \\ 0, \text{ якщо } U_{\text{вх}.Ak} < \theta, \end{cases} \tag{2.23}$$

де $U_{\text{вих}.Ak}$ – вихідний сигнал нейрона A_k; $U_{\text{вх}.Ak}$ – сигнал на вході нейрона A_k,

$$U_{\text{вх}.Ak} = \sum_{j=1}^{m} U_{\text{вих}.Sj} w_{jk}, \tag{2.24}$$

$U_{\text{вих}.Sj}$ – вихідний сигнал j-го S-елемента,

$$U_{\text{вих}.Sj} = \begin{cases} 1, \text{ якщо } S\text{-елемент збуджений,} \\ -1, \text{ якщо } S\text{-елемент загальмований,} \end{cases} \tag{2.25}$$

w_{jk} – вага зв'язку між j-м S-елементом та k-м A-нейроном; θ – поріг спрацьовування k-го A-елемента; $\theta = m$.

Для кожного зображення W_k задамо ваги w_{jk} $(j = \overline{1, m},\ k = \overline{1, n})$ співвідношенням

$$w_{jk} = \begin{cases} 1, \text{ якщо елемент } S_j \text{ збуджений } k\text{-м зображенням,} \\ -1, \text{ якщо елемент } S_j \text{ загальмований } k\text{-м зображенням.} \end{cases} \tag{2.26}$$

При пред'явленні будь-якого зображення W_k $(k = \overline{1, n})$ перцептрону, що задовольняє співвідношенням (2.23) – (2.26), тільки на вході одного k-го A-нейрона буде сигнал, що дорівнює, відповідно до співвідношення (2.24), числу m, й тільки на виході цього нейрона відповідно до виразу (2.23) буде одиничний вихідний сигнал. Тепер для правильного виконання поділу вихідної множини W на дві підмножини W^1, W^2 за допомогою елементарного перцептрона необхідно тільки всім вагам зв'язків між R- та A-нейронами, які відповідають A-елементам, збудженим зображеннями підмножини W^1, додати додатні значення, а всім вагам зв'язків A-нейронів, які збуджуються зображеннями підмножини W^2, – від'ємні значення, а потім задати вихідний сигнал R-нейрона $U_{вих.R}$ виразом вигляду

$$U_{\text{вих.}R} = \begin{cases} 1, \ \text{якщо}\, U_{\text{вх.}R} \geq 0\,, \\ -1, \ \text{якщо}\, U_{\text{вх.}R} < 0\,, \end{cases}$$

де $U_{\text{вх.}R}$ – вхідний сигнал *R*-елемента.

У цьому випадку всі зображення підмножини W^1 будуть кодуватися додатними одиничним вихідним сигналом нейронної мережі, а підмножина W^2 – від'ємними, тобто буде правильно виконуватися класифікація *C*(*W*) вихідної множини *W* зображень.

Хоча побудована в такий спосіб нейронна мережа не має істотного практичного значення, проте її наявність показує, що елементарний перцептрон є універсальним пристроєм розподілу будь-якої множини зображень на два класи. У тому випадку, коли число зображень множини *W* перевищує число *A*-нейронів, елементарний перцептрон втрачає свою універсальну здатність класифікувати зображення.

Теорема 2.2. Якщо число *n* зображень у множині *W* більше від числа *A*-елементів елементарного перцептрона, то існує деяка класифікація *C*(*W*) множини *W* чорно-білих зображень на дві підмножини W^1, W^2, що не може бути виконана перцептроном.

Теорема 2.3. Для будь-якого елементарного перцептрона із скінченним числом *A*-нейронів ймовірність виконання класифікації *C*(*W*), що вибирається з рівномірного розподілу за всіма можливими розподілами множини $W = \{W_1, \ldots, W_n\}$ зображень на два класи, прямує до нуля при *n*, що прямує до нескінченності.

Теорема 2.1 доводить існування елементарного перцептрона, здатного виконувати будь-який заданий розподіл *C*(*W*) зображень деякої множини *W* на два класи, однак вона не вказує алгоритмів досягнення цієї здатності в процесі навчання нейронної мережі. Розенблаттом була доведена важлива для теорії елементарних перцептронів теорема про наявність таких алгоритмів для α-перцептронів.

Розглянемо деяку класифікацію *C*(*W*) множини *W* зображень на дві підмножини W^1, W^2, що може здійснюватися перцептроном з *R*-елементом, вихідний сигнал якого задовольняє умови

$$U_{\text{вих.}R} = \begin{cases} 1, \ \text{якщо}\, W_k \in W^1, \\ -1, \ \text{якщо}\, W_k \in W^2, \end{cases} \quad k = \overline{1,n}. \tag{2.27}$$

Припустимо також, що

$$\rho_k = \begin{cases} 1, \text{ якщо } W_k \in W^1, \\ -1, \text{ якщо } W_k \in W^2, \end{cases} \quad k = \overline{1,n}. \tag{2.28}$$

Визначення 2.9. Метод корекції похибок без квантування – це метод системи підкріплень з корекцією похибок, коли при помилковій реакції *R*-елемента з порогом θ на деяке зображення $W_k \in W$ до ваги кожної зі зв'язків, що з'єднують активні *А*-нейрони з *R*-елементом, додається величина $\eta = \rho_k R_k$, де коефіцієнт R_k вибирається з умови, що після корекції ваг зв'язків виконується співвідношення (2.27), тобто перцептрон правильно класифікує пред'явлене зображення. У методі корекції похибок з квантуванням застосовується це ж правило корекції ваг зв'язків, але величина R_k в загальному випадку набагато менша й правильний сигнал на виході *R*-нейрона, як правило, досягається не за одну ітерацію.

Теорема 2.4. Нехай є елементарний α-перцептрон, множина чорно-білих зображень $W = \{W_1, \ldots, W_n\}$, деяка класифікація *C*(*W*) цих зображень на дві підмножини, що може бути виконана α-перцептроном. Зображення $W_1, \ldots, W_n \in W$ подаються на вхід перцептрона в довільній послідовності, у якій кожне з них з'являється неодноразово, через скінченне число пред'явлень інших зображень. Тоді процес навчання перцептрона методом корекції похибок (із квантуванням або без квантування підкріплень) незалежно від початкових значень ваг зв'язків між *R*- та *A*-елементами завжди приводить за скінченне число ітерацій до множини ваг зв'язків, за допомогою яких α-перцептрон може виконати задану класифікацію зображень.

Теорема 2.4 доводить наявність збіжного детермінованого методу навчання з корекцією помилок для елементарного перцептрона із квантуванням або без квантування підкріплень. Наступна теорема доводить, що навчання елементарного перцептрона може бути виконане й при менш жорстких вимогах до виду корекції – методом корекції похибок з випадковим законом підкріплень, коли з появою похибки сигнал підкріплення формується як в α-системі, але знак підкріплення з імовірністю 0,5 може бути додатним або від’ємним.

Теорема 2.5. Нехай є елементарний α-перцептрон із скінченним числом значень ваг зв'язків між *R*- та *А*-нейронами, множина чорно-білих зображень $W = \{W_1, \ldots, W_n\}$, деяка класифікація $C(W)$ цих зображень на дві підмножини, що може бути виконана α-перцептроном при деякому наборі ваг зв'язків між *R*- та *A*-нейронами, зображення $W_1, \ldots, W_n \in W$, що подаються на вхід перцептрона в довільній послідовності, у якій кожне з них з'являється неодноразово через скінченне число пред'явлень інших зображень. Тоді процес навчання перцептрона, величина сигналів підкріплень якого формується, як у α-системі із квантуванням підкріплень, а знак підкріплення вибирається з імовірністю 0,5 додатним або від'ємним, може бути виконаний за скінченний час із ймовірністю, що дорівнює одиниці, незалежно від початкових значень ваг зв'язків між *R*- та *A*-нейронами.

Природно, що метод з випадковим знаком підкріплення вимагає більшого обсягу обчислень при навчанні перцептрона, ніж пряма корекція похибок із квантуванням або без квантування підкріплень. Ще більшого обсягу обчислень вимагає метод, у якому виробляється випадковий вибір не тільки знака, але й величини підкріплення. Розенблаттом доведена теорема про те, що з імовірністю, яка дорівнює одиниці, навчання перцептрона може бути виконане за скінченний час і методом корекції випадковими збурюваннями, коли підкріплення формується, як у α-системі, але при цьому величина η та знак підкріплень для кожної ваги зв'язку вибираються окремо й незалежно відповідно до деякого заданого закону розподілу ймовірностей.

Менш універсальною системою підкріплень, ніж α-система й системи з випадковим формуванням підкріплень, є γ-система. Це доводить наступна теорема Розенблатта.

Теорема 2.6. Нехай є елементарний γ-перцептрон, підмножина $W = \{W_1, \ldots, W_n\}$ чорно-білих зображень та деяка класифікація $C(W)$ цих зображень на два класи W^1, W^2. Тоді для виконання класифікації $C(W)$ може існувати набір ваг зв'язків, недосяжний для γ-системи підкріплень.

Доведення. Нехай функціонування *A*-нейронів визначається виразом

$$U_{\text{вих}.Ak} = \begin{cases} 1, \ \text{якщо}\, U_{\text{вх}.Ak} \geq \theta\,, \\ 0, \ \text{якщо}\, U_{\text{вх}.Ak} < \theta\,, \end{cases} \quad k = \overline{1,n},$$

де $U_{\text{вих}.Ak}, U_{\text{вх}.Ak}$ – відповідно вихідний й вхідний сигнали *k*-го *A*-нейрона; θ – поріг спрацьовування *A*-нейронів.

Нехай також кожний *A*-нейрон збуджується тільки одним зображенням з множини *W*, а класифікація *C*(*W*) здійснюється за допомогою *R*-елемента, функціонування якого описується співвідношенням

$$U_{\text{вих}.R} = \begin{cases} 1, \ \text{якщо}\, U_{\text{вх}.R} > 0\,, \\ -1, \ \text{якщо}\, U_{\text{вх}.R} < 0\,, \end{cases}$$

де $U_{\text{вих}.R}, U_{\text{вх}.R}$ – відповідно вихідний й вхідний сигнали *R*-нейрона. Виберемо класифікацію *C*(*W*), що відносить всі зображення до класу $W^1\, (U_{\text{вих}.R},\, U_{\text{вх}.R} > 0)$ або до класу $W^2\, (U_{\text{вих}.R},\, U_{\text{вх}.R} < 0)$. Очевидно, що у першому випадку рішення існує тільки тоді, коли ваги всіх зв'язків між *R*- та *A*-елементами додатні (або від'ємні, якщо всі зображення відносяться до класу W^2). Якщо для першого випадку початкові ваги всіх зв'язків між *R*- та *A*-нейронами від'ємні, а для другого – додатні, то в силу властивості консервативності γ-системи підкріплень щодо суми всіх ваг зв'язків між *R*- та *A*-нейронами, вона не зможе виконати правильне настроювання ваг зв'язків перцептрона для розглянутої класифікації *C*(*W*).

2.3. Індивідуальні завдання

1. Розробити структуру елементарного перцептрона, здатного розпізнавати перші дві букви вашого імені та вашого прізвища. При цьому необхідно обґрунтувати вибір:

- числа рецепторних нейронів (число *n* *S*-елементів перцептрона повинне бути в межах $12 \le n \le 30$);
- числа нейронів схованого шару;
- величину кроку в алгоритмі навчання перцептрона;
- види функцій активації нейронів кожного шару;
- величини порогів нейронів кожного шару.

2. Навчити нейронну мережу методом α- та γ-підкріплень. Порівняти роботу обох методів.