6.2 Інтелектуальні засоби ідентифікації

Інтелектуальний підхід до моделювання передбачає в першу чергу використання таких способів та особливостей побудови моделі, які підсвідомо використовує людина при створенні своєї уяви про навколишній світ. Серед них слід відзначити:

–поступове покращення моделі (навчання);

–використання відносно простих математичних операцій;

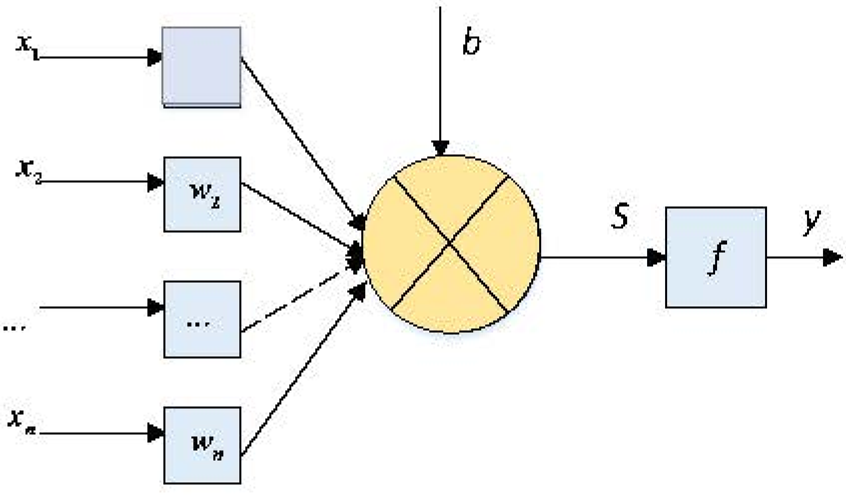
–образне уявлення про навколишній світ.

На сьогоднішньому етапі повністю відтворити людську свідомість неможлива, але деякі засоби мають з нею певну схожість. До таких засобів можна віднести нейронні мережі, нечіткі моделі, моделі на основі продукційних правил тощо.

6.2.1 Ідентифікація шляхом навчання нейронних мереж

Нейронною мережею називають структуру, що складається зі зв'язаних між собою нейронів.

Нейрон (рис. 6.3) – це складова частина нейронної мережі. Він складається з елементів трьох типів. Математично елементи нейрона: множники (синапси), суматор і нелінійний перетворювач. *Синапси* здійснюють зв’язок між нейронами, множать вхідний сигнал на число, що характеризує силу зв'язку, – вагу синапсу. *Суматор* виконує додавання сигналів, що надходять по синоптичних зв'язках від інших нейронів, і зовнішніх вхідних сигналів. *Нелінійний перетворювач* реалізує нелінійну функцію одного аргументу – виходу суматора. Ця функція називається "*функція активації*" (рис. 6.4).



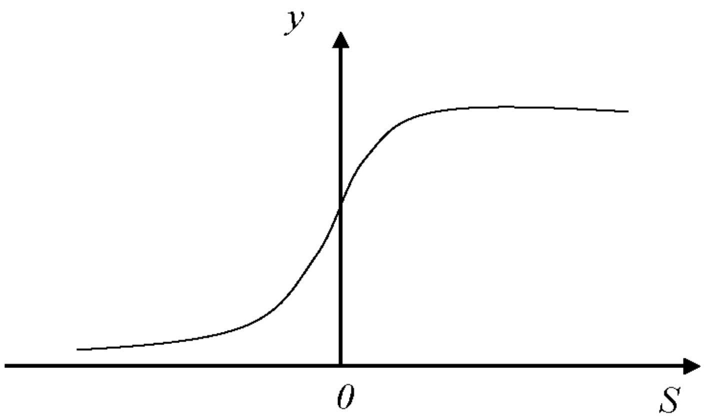


Рисунок 6.4 - Сигмоїдна функція активації

Нейрон у цілому реалізує скалярну функцію векторного аргументу. Математична модель нейрона:

$$\begin{cases} S = \sum\limits_{i=1}^{n} w_i x_i + b \\ y = f(S) \end{cases}, \qquad (6.37)$$

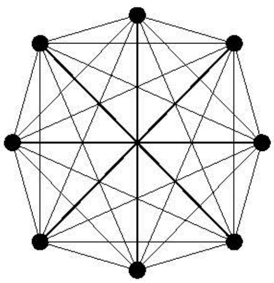
де w_i – вага синапсу, ($i=1,2...n$); b – значення зсуву; S – результат підсумовування; x_i – компонента вхідного вектора (вхідний сигнал), ($i=1,2,...n$); y – вихідний сигнал нейрона; n – число входів нейрона; f – нелінійне перетворення (функція активації).

У загальному випадку вхідний сигнал, вагові коефіцієнти і значення зсуву можуть приймати дійсні значення. Вихід (y) визначається видом функції активації і може бути як дійсним, так і цілим (при функції активації релейного типу). У багатьох практичних задачах входи, ваги і зсуви можуть приймати лише деякі фіксовані значення.

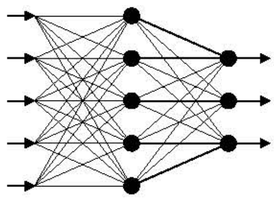
Синаптичні зв'язки з додатними вагами називають *збудливими,* з від’ємними вагами – *гальмівними*.

Нейронні мережі можуть мати різні архітектури. Можна виділити три основних типи нейтронних мереж:

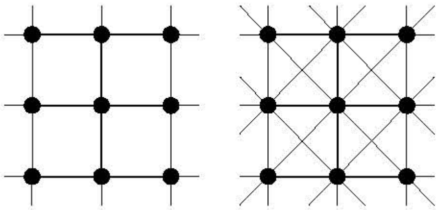
- повнозв’язні мережі (рис. 6.5, а),
- багатошарові мережі (рис. 6.5,б),
- слабкозв’язні мережі (мережі з локальними зв'язками, рис. 6.5, в).



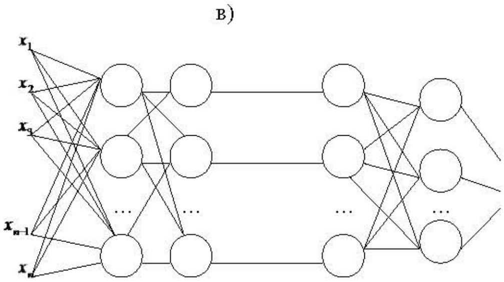
а)



б)



в)



У *слабкозв'язних* мережах нейрони розташовуються у вузлах прямокутних ґрат. Кожен нейрон зв'язаний з чотирма чи вісьмома своїми найближчими сусідами.

У *повнозв'язній* мережі кожен нейрон зв'язаний з усіма іншими (на входи кожного нейрона подаються вихідні сигнали інших нейронів).

У *багатошарових* мережах нейрони об'єднуються в шари. *Шар* – це сукупність нейронів зі спільним вхідним сигналом. Зовнішні вхідні сигнали подаються на входи нейронів першого шару, а виходами мережі є вихідні сигнали останнього шару. Крім вхідного і вихідного шарів у багатошарової нейронної мережі є один чи декілька проміжних (прихованих) шарів. Зв'язки від виходів нейронів деякого шару m до входів нейронів наступного шару ($m+1$) називаються *послідовними*.

Якщо нейрони кожного шару мережі мають однакову функцію активації, то таку нейронну мережу називають *однорідною*.

Нейронні мережі з локальними зв'язками. Нейрони в таких мережах розташовуються у вузлах прямокутних ґрат. Кожен нейрон зв'язаний з невеликим числом (4 чи 8) своїх сусідів.

Неструктуровані нейронні мережі. До цієї групи відносять всі моделі нейронних мереж, які не можна віднести ні до однієї з попередніх груп.

Кожна група нейронних мереж може бути використана для розв’язання лише деякого обмеженого класу практичних задач. Так багатошарові і повнозв’язні нейронні мережі з сигмоїдними функціями використовуються для розпізнавання образів і адаптивного керування; нейронні мережі з локальними зв'язками – для обробки зображень і деяких інших задач. Для розв’язання задач лінійної алгебри використовуються багатошарові мережі.

Найбільшою мірою теоретично обґрунтовані тришарові нейронні мережі з сигмоїдними функціями. На основі теореми Колмогорова-Арнольда доведено, що такі мережі можуть реалізовувати будь-які відображення вхідного сигналу у вихідний.

Апаратна реалізація нейронної мережі може бути виконана на нейрочіпах – мікросхемах, що містять фрагменти нейронних мереж, або на ПЛІС – програмованих логічних інтегральних схемах.

Нейронна мережа може бути реалізована не тільки апаратно, але й програмно, у вигляді відповідного алгоритму. Для побудови такої мережі, орієнтованої на розв’язання конкретної задачі, використовуються процедури формування нейронних мереж. Ці процедури забезпечують введення характеристик моделей нейронів і структур нейронних мереж. Як правило, у кожній окремій програмі реалізована лише частина з описаних моделей нейронів і нейронних мереж.

Вибір структури нейронної мережі еквівалентний процедурі структурної ідентифікації.

В процесі параметричної ідентифікації необхідно провести налаштування параметрів мережі – *навчання нейронної мережі.* Налаштування здійснюється за навчальною вибіркою, що складається з пар (<вхід>, <бажаний вихід>) – навчальних прикладів.

Якщо в ненавчену нейронну мережу ввести вхідний сигнал одного з прикладів навчальної вибірки, то вихідний сигнал мережі буде істотно відрізнятися від бажаного вихідного сигналу, визначеного в навчальній вибірці. Функція помилки чисельно визначає подібність всіх поточних вихідних сигналів мережі і відповідних бажаних вихідних сигналів навчальної вибірки. Найбільш розповсюдженою функцією помилки є середньоквадратичне відхилення.

Мета навчання – мінімізувати функцію помилки, тобто знайти такі значення параметрів мережі, при яких поточні вихідні сигнали мережі мінімально відрізняються від відповідних бажаних вихідних сигналів, заданих навчальною вибіркою. Отже, задача навчання є задачею оптимізації.

Навчання – це ітераційна процедура, яка при реалізації на звичайних комп’ютерах вимагає значного часу. Алгоритми навчання істотно розрізняються за швидкістю збіжності. Однією з найважливіших характеристик нейронних мереж є швидкість збіжності алгоритмів навчання, що реалізовані в програмі.

Найпоширенішим способом навчання нейронної мережі є застосування алгоритму зворотного поширення помилки.

Алгоритм зворотного поширення помилки застосовується для багатошарового перцептрона. У такій мережі є множина входів x_1 ,..., x_n, множина виходів *Outputs* і множина внутрішніх вузлів. Перенумеруємо всі вузли (включаючи входи і виходи) числами від 1 до N (наскрізна нумерація, незалежно від топології шарів). Позначимо через $w_{i,\,j}$ вагу, що стоїть на ребрі, яке сполучає і-й і j-й вузли, а через o_i – вихід і-го вузла. Якщо нам відомий навчальний приклад (правильні відповіді мережі t_k, $k \in Outputs$), то функція помилки, отримана за методом найменших квадратів, матиме вигляд

$$E\left(\left\{w_{i,j}\right\}\right)=\frac{1}{2}\sum_{k\in Outputs}\left(t_k-o_k\right)^2 \tag{6.38}$$

Для налаштовування ваги нейронів коректуються після кожного навчального прикладу і, таким чином, відбувається рух у багатовимірному просторі ваг. Щоб дістатися до мінімуму помилки потрібно рухатися у бік, протилежний градієнту, тобто на підставі кожного прикладу додавати до кожної ваги

$$\Delta w_{i,j}=-\eta\frac{\partial E}{\partial w_{i,j}}, \tag{6.39}$$

де 0 <η <1 – множник, що задає швидкість руху.

Похідна підраховується так. Нехай спочатку $j\in Outputs$, тобто, відповідна вага входить в нейрон останнього рівня. Вага $w_{i,j}$ впливає на вихід мережі тільки як частина суми

$$S_j=\sum_i w_{i,j}x_i\,, \tag{6.40}$$

де сума береться на входах j-го вузла. Тому

$$\frac{\partial E}{\partial w_{i,j}}=\frac{\partial E}{\partial S_j}\cdot\frac{\partial S_j}{\partial w_{i,j}}=x_i\frac{\partial E}{\partial S_j}. \tag{6.41}$$

Аналогічно, S_j впливає на загальну помилку тільки виходу j-го вузла o_j. Тому

$$\frac{\partial E}{\partial S_j}=\frac{\partial E}{\partial o_j}\cdot\frac{\partial o_j}{\partial S_j}=\left\{\frac{\partial}{\partial o_j}\left[\frac{1}{2}\sum_{k\in Outputs}\left(t_k-o_k\right)^2\right]\right\}\cdot\left[\frac{\partial o_j}{\partial S_j}\right]=\left[\frac{1}{2}\cdot\frac{\partial}{\partial o_j}\left(t_j-o_j\right)^2\right]\cdot\left[o_j\left(1-o_j\right)\right]=$$

$$=-o_j\left(1-o_j\right)\left(t_j-o_j\right). \tag{6.42}$$

Якщо ж j-й вузол не на останньому рівні, то у нього є виходи; позначимо їх через *Children(j)*. У цьому випадку

$$\frac{\partial E}{\partial S_j} = \sum_{k \in Children(j)} \frac{\partial E}{\partial S_k} \frac{\partial S_k}{\partial S_j}, \tag{6.43}$$

і

$$\frac{\partial S_k}{\partial S_j} = \frac{\partial S_k}{\partial o_j} \frac{\partial o_j}{\partial S_j} = \omega_{i,j} \frac{\partial o_j}{\partial S_j} = \omega_{i,j} o_j \left(1 - o_j\right). \tag{6.44}$$

Похідна $\partial = \dfrac{\partial E}{\partial S_k}$ обчислюється аналогічно для вузла наступного рівня (від Δ_k вона відрізняється відсутністю множника $-\eta x_{i,j}$).

Саме через особливість обчислення поправок алгоритм називається алгоритмом зворотного поширення помилки (backpropagation):

– для вузла останнього рівня (для внутрішнього вузла мережі) $\delta_j = -o_j\left(1 - o_j\right) \sum\limits_{k \in Outputs(j)} \delta_k w_{j,k}$;

– для всіх інших вузлів $\Delta w_{i,j} = -\eta \delta_j x_i$.

На вхід алгоритму, крім зазначених параметрів, потрібно задавати структуру мережі. На практиці дуже хороші результати показують мережі досить простої структури, що складаються з двох рівнів нейронів – прихованого рівня (*hidden units*) і нейронів-виходів (*output units*); кожен вхід мережі з'єднаний з усіма прихованими нейронами, а результат роботи кожного прихованого нейрона подається на вхід кожному з нейронів-виходів. У такому випадку достатньо подавати на вхід кількість нейронів прихованого рівня.

Алгоритм: *BackPropagation (η, α,* $\left\{x_i^d, t^d\right\}_{i=1,d=1}^{n,m}$, *NUMBER_OF_STEPS)*, де α – коефіцієнт інерціальності для згладжування різких стрибків при переміщенні по поверхні цільової функції:

1. Ініціалізувати (w_{ij}) маленькими випадковими значеннями, (Δw_{ij}), $i, j = 0$.
2. Повторити NUMBER_OF_STEPS разів:

   Для всіх d від 1 до m:

   1. Подати $\left\{x_i^d\right\}$ на вхід мережі і підрахувати виходи
   2. $_i$ кожного вузла.
   3. Для всіх $k \in Outputs$ $\delta_k = o_k(1 - o_k) \cdot (t_k - o_k)$.

   3. Для кожного рівня l, починаючи з передостаннього:

   Для кожного вузла j рівня l обчислити

$$\delta_j = o_j(1 - o_j) \sum_{k \in Children(j)} \delta_k w_{j,k}.$$

4. Для кожного ребра мережі (i, j)

$$\Delta w_{i,j} = \alpha \Delta w_{i,j} + (1 - \alpha)\, \eta \delta_j o_i.$$

$$w_{i,j} = w_{i,j} + \Delta w_{i,j}.$$

3. Видати значення w_{ij}.

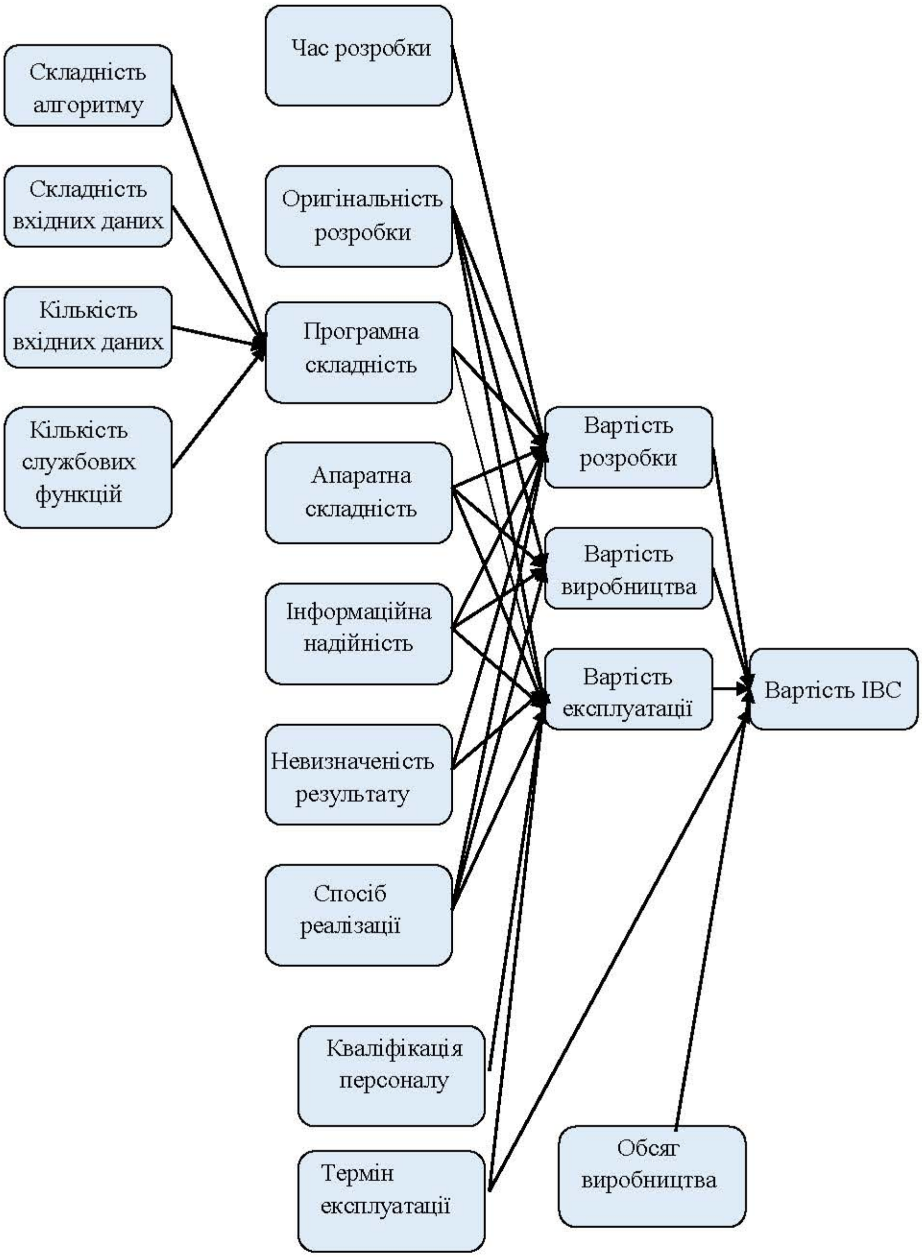
Алгоритм зворотного розповсюдження помилки достатньо простий, але має багато недоліків. По-перше, це алгоритм багатокритеріальної оптимізації (критеріями є помилки для кожного набору даних навчальної вибірки), причому критерії застосовуються послідовно – отже, немає гарантії, що навчання колись закінчиться. По-друге, навчання нейронної мережі до прийнятної точності може вимагати дуже великої вибірки і багато часу. По-третє, якість навчання суттєво залежить від вибору навчальних прикладів.

6.2.2 Нечітка ідентифікація

У багатьох випадках принаймні частина даних, за якими необхідно ідентифікувати модель, задається експертами у лінгвістичній формі «великий», «малий», «середній» тощо. У таких випадках може бути побудована нечітка регресійна модель.

Ще одним інтелектуальним засобом є *ідентифікація* математичної моделі у вигляді *нечіткої бази знань*. Нечітка база знань будується на основі висновків експерта і містить продукційні правила «якщо-то» (див. п. 4.2). Частина «якщо» містить нечіткі вхідні змінні моделі, частина «то» містить результати. Нечіткі вхідні змінні можуть приймати значення з невеликої множини термів, а можливість використання у даному правилі іншого терму з тим самим результатом характеризується функцією належності (див. п. 5.2. Набір правил разом з відповідними функціями належності і складають модель системи у вигляді нечіткої бази знань.

Процес ідентифікації починається зі створення структурно-логічної схеми залежності вихідних змінних від вхідних. Приклад структурно-логічної схеми нечіткої моделі залежності вартості інформаційно-вимірювальної системи (ІВС) від різноманітних факторів показано на рис. 6.6.



Структурно-логічна схема розробляється у напрямку від результату до вхідних даних. Спочатку експерт визначає, які головні фактори визначають результат (у наведеному прикладі вартість ІВС залежить від вартості розробки, вартості виробництва, вартості експлуатації, терміну експлуатації і обсягу виробництва. Далі на схемі визначають другорядні фактори, які впливають на зазначені головні (у наведеному прикладі вартість розробки ІВС залежить від часу розробки, оригінальності розробки, програмної складності, апаратної складності, інформаційної надійності, невизначеності результатів (точності ІВС) і способу реалізації) і так далі, доки усі впливові фактори не будуть такими, що експерт зможе їх безпосередньо оцінити.

На основі структурно-логічної схеми будується структура бази знань у вигляді сукупності таблиць. Кожна таблиця відображає окрему частину структурно-логічної схеми «фактори-результат».

Далі база знань наповнюється правилами. До набору правил у базі знань висуваються певні вимоги:

1. Кожен терм кожної нечіткої змінної повинен зустрітися у базі знань не менше одного разу;

2. У базі знань не повинні зустрічатися правила, які при однакових вхідних даних дають різний результат.

Мінімальна кількість правил, які задовольняють дані умови, $P_{\min} = \max_i n_{T\,i}$, де $n_{T\,i}$ – кількість термів i-ї нечіткої змінної, i=1…m, m – кількість нечітких змінних. Максимальна кількість правил $P_{\max} = \prod_{i=1}^{m} n_{T\,i}$. Чим більше правил, тим точніший результат, але тим більше часу необхідно для ідентифікації і обробки даних.

Звичайно кількість правил обирають так. Якщо розглядати кожну комбінацію значень нечітких змінних як точку у m-вимірному просторі, а з урахуванням функцій належності, які мають певний розкид, як m-вимірні кубики, то мінімально необхідна кількість правил повинна своїми m-вимірними кубиками покривати увесь простір значень нечітких змінних.

Експертне оцінювання усіх елементів бази знань повинно здійснюватися за певними методологічними нормами.

Метод експертних оцінок включає в себе три складові.

1. Інтуїтивно-логічний аналіз задачі. Будується на логічному мисленні та інтуїції експертів, які ґрунтуються на їх знаннях і досвіді. Цим пояснюється високий рівень вимог, що висуваються до експертів.

2. Розв’язання та видача кількісних або якісних оцінок. Ця процедура є завершальною частиною роботи експерта.

3. Обробка результатів розв’язання. Отримані від експертів оцінки повинні бути оброблені з метою одержання підсумкової оцінки проблеми. Залежно від поставленої задачі змінюється кількість виконуваних на цьому етапі розрахункових і логічних процедур.

Етапи підготовки і проведення експертизи

Якість одержуваних експертних оцінок значною мірою визначається підготовкою експертизи, а також застосовуваними методами обробки інформації, що отримується від експертів.

Можна виділити основні етапи підготовки та проведення експертизи. До них належать:

- формулювання мети експертного аналізу;
- формування групи організаторів експертизи;
- розробка процедур проведення експертного оцінювання;
- підбір експертів;
- отримання експертних оцінок;
- обробка результатів опитування та аналіз отриманих даних;
- встановлення ступеня досягнення мети експертизи.

Отримання експертних оцінок. Поняття шкали. Типи шкал

Раціональне використання інформації, що отримується від експертів, можливе за умови перетворення її у форму, зручну для подальшого аналізу. Одна з головних труднощів при оцінюванні полягає в тому, що крім явищ, об’єктів, факторів, стан яких може бути виражено кількісно (в гривнях, Кг, км, % і т. п.), доводиться оцінювати якісні фактори, рівень яких не можна точно визначити. Частину інформації, що не піддається кількісному вимірюванню, необхідно подати у вигляді непрямих оцінок.

Якщо експерт здатний порівняти і оцінити будь-які об’єкти, явища, чинники, варіанти дій, приписавши кожному з них деяке число, то кажуть, що він володіє певною системою переваг. Залежно від того, за якою шкалою задані ці переваги, експертні оцінки містять більший або менший обсяг інформації та мають різну придатність до математичної формалізації.

Шкала – це інструмент (прийнята система правил) оцінювання (вимірювання) будь-яких об’єктів чи явищ.

Розрізняють чотири типи шкал:

1. *Номінальна (класифікаційна) шкала*. Реалізує найпростіший тип вимірювання. У цьому випадку проводиться порівняння властивостей об’єкта (явища) з ознакою-еталоном, результатом є впорядкування за двоелементною шкалою, де кожному з об’єктів (явищ) присвоюється бал, що дорівнює нулю або одиниці.

Прикладом вимірювання за номінальною шкалою може служити проведення заліку. У цьому випадку експерт-викладач оцінює рівень знань студентів і виносить рішення: залік (об’єкту-студенту присвоюється бал, який дорівнює одиниці) або незалік (об’єкту-студенту присвоюється бал, який дорівнює нулю).

2. *Порядкова (рангова) шкала*. Мета полягає в упорядкуванні об’єктів (явищ), а точніше, у виявленні за допомогою експертів прихованої впорядкованості, яка, за припущенням, властива множині об’єктів. Результатом оцінювання є рішення про те, що який-небудь об’єкт (явище) переважає інший за якимось критерієм.

Прикладом може служити визначення переможців і призерів будь-якого конкурсу. Тут експерти повинні вирішити, що учасник, який посів перше місце, виявився кращим (з точки зору цілей конкурсу) за учасника, що посів друге місце. Учасник, що посів друге місце, у свою чергу, визнається кращим відносно третього і т. д.

3. *Інтервальна (різницева) шкала.* Оцінка за даною шкалою дозволяє не тільки визначити, що один об’єкт (явище) переважає інший, але також визначити: на скільки переважає. Нульова точка і одиниця вимірювання вибираються при цьому довільно.

Яскравим прикладом оцінювання за інтервальної шкалою є проведення іспиту. Тут експерт-викладач, оцінюючи рівень знань студентів, повинен не тільки вирішити, що один студент знає матеріал краще іншого, але сказати: наскільки краще. Вимірювання фактично здійснюється за шкалою з чотирьох балів ("незадовільно", "задовільно", "добре", "відмінно").

4. *Абсолютна шкала (шкала відносин).* У даному випадку передбачається, що відомо абсолютне значення властивостей об’єкта, тобто відома справжня нульова точка. Шкала використовується для тих факторів, які можуть бути подані кількісно.

Наприклад, за допомогою такої шкали експерти можуть оцінити розмір прибутку, що може бути отриманий в результаті реалізації проекту.

Залежно від сутності досліджуваних об’єктів для їх оцінювання можуть бути використані різні шкали.

Способи вимірювання об’єктів

Розглянемо в першу чергу способи вимірювання, що дозволяють розташувати об’єкти за порядковою або інтервальною шкалою, оскільки саме такий тип оцінок найчастіше використовується при проведенні експертизи. Це пояснюється тим, що оцінка за номінальною шкалою припускає лише два варіанти відповідей – ТАК, НІ. За шкалою відносини вимірюються фактори, що мають кількісний характер. Значення цих факторів часто можна отримати розрахунковим шляхом без використання експертних оцінок.

Виділимо способи вимірювання об’єктів, які найчастіше застосовуються при оцінюванні за порядковою або інтервальною шкалою: ранжування, парне порівняння, безпосередня оцінка.

1. *Ранжування* – це розташування об’єктів у порядку зростання або зменшення будь-якої притаманної їм властивості. Ранжування дозволяє вибрати з досліджуваної сукупності факторів найбільш істотний.

Якщо є n об’єктів, то в результаті їх ранжування j-им експертом кожен об’єкт отримує оцінку x_{ij} – ранг, який приписується i-му об’єкту j-им експертом.

Значення x_{ij} знаходяться в інтервалі від 1 до n. Ранг найкращого об’єкта (найважливішого чинника) дорівнює одиниці, найменш значущого – n.

Ранжуванням, що здійснене j-тим експертом, називається послідовність рангів $x_{1j}, x_{2j}, ..., x_{nj}$.

Перевагою методу є його простота, а недоліком – обмежені можливості використання. При оцінюванні великої кількості об’єктів експертам дуже важко будувати ряд, оскільки доводиться враховувати безліч складних зв’язків.

2. *Парне порівняння* – це встановлення переваги об’єктів при порівнянні всіх можливих пар. Тут не потрібно, як при ранжуванні, впорядковувати всі об’єкти – необхідно в кожній з пар виявити більш значущий об’єкт або встановити їх рівність.

Парне порівняння можна проводити при великій кількості об’єктів, а також у тих випадках, коли відмінність між об’єктами настільки незначна, що практично неможливо виконати їх ранжування.

При використанні методу найчастіше складається матриця розміром $n \times n$, де n – кількість порівнюваних об’єктів. Загальний вигляд матриці парних порівнянь показаний у таблиці 6.1.

Таблиця 6.1 – Загальний вигляд матриці парних порівнянь

Об’єкти	**1**	**2**	**...**	**j**	**...**	**n**	$\sum$
1	1	0		1		0	
2	2	1				2	
...							
n	2	0					

При порівнянні об’єктів матриця заповнюється елементами a_{ij} так:

$$a_{ij} = \begin{cases} 2, \text{ якщо об'єкт } i \text{ переважає об'єкт } j, \\ 1, \text{ якщо об'єкти рівнозначні}, \\ 0, \text{ якщо об'єкт } j \text{ переважає об'єкт } i. \end{cases}$$

Сума $\sum_{j=1}^{n} a_{ij}$ (по рядку) дозволяє оцінити відносну значимість об’єктів. Той об’єкт, для якого сума виявиться найбільшою, може бути визнаний найбільш важливим (значущим).

3. *Безпосереднє оцінювання.* Часто буває бажаним не тільки впорядкувати (ранжувати об’єкти аналізу), а й визначити, наскільки один фактор більш значущий, ніж інші.

У цьому випадку діапазон зміни характеристик об’єкта розбивається на окремі інтервали, кожному з яких приписується певна оцінка (бал), наприклад, від 0 до 10. Тому метод безпосереднього оцінювання іноді називають також бальних методом.

Число інтервалів, на які розбивається діапазон зміни властивості, може бути різним для різних експертів. Крім того, метод дозволяє давати одну і ту ж оцінку (тобто поміщати в один і той же інтервал) різним об'єктам.

Обробка результатів опитування експертів

Залежно від цілей експертизи при обробці оцінок можуть вирішуватися такі проблеми:

- формування узагальненої оцінки;
- визначення відносних ваг об'єктів;
- встановлення ступеня узгодженості думок експертів і ін.

Формування узагальненої оцінки

Нехай група експертів оцінила об'єкт, тоді x_j – оцінка j-го експерта, j=1,…m, де m – число експертів.

Для формування узагальненої оцінки групи експертів найчастіше використовуються середні величини. Наприклад, медіана M_E, за яку приймається така оцінка, відносно якої число більших оцінок дорівнює числу менших.

Може використовуватися також точкова оцінка для групи експертів, що обчислюється як середнє арифметичне:

$$\overline{x_e} = \frac{\sum_{j=1}^{m} x_j}{m}.$$

Визначення відносних ваг об'єктів

Іноді потрібно визначити, наскільки той чи інший фактор (об'єкт) важливий (істотний) з точки зору будь-якого критерію. У цьому випадку говорять, що потрібно визначити вагу кожного фактора.

Один з методів визначення ваги полягає в такому. Нехай x_{ij} – оцінка фактора i, яка надана j-м експертом, i=1,…n, j=1,…m, n – кількість факторів (об'єктів), m – кількість експертів. Тоді вага i-го об'єкта, підрахована за оцінками усіх експертів,

$$w_i = \frac{\sum_{j=1}^{m} w_{ij}}{m}, \quad i = \overline{1,n},$$

де w_{ij} – вага i-го об'єкта, підрахована за оцінками j-го експерта

$$w_{ij} = \frac{x_{ij}}{\sum_{i=1}^{n} x_{ij}}, \quad i=1,\ldots n, j=1,\ldots m.$$

Встановлення ступеня узгодженості думок експертів

У разі участі в опитуванні декількох експертів розбіжності в їхніх оцінках неминучі, проте величина цієї розбіжності має важливе значення. Групова оцінка може вважатися достатньо надійною тільки за умови достатності узгодженості відповідей окремих фахівців.

Для аналізу розкиду і узгодженості оцінок застосовуються міри розкиду:

– варіаційний розкид

$$R = x_{\max} - x_{\min},$$

де $x_{\max}$ – максимальна оцінка об'єкта; $x_{\min}$ – мінімальна оцінка об'єкта;

– середнє квадратичне відхилення

$$\sigma = \sqrt{\frac{\sum_{j=1}^{m}(x_j - \overline{x_э})^2}{m-1}},$$

де x_j – оцінка, надана j-м експертом;

– коефіцієнт варіації

$$V = \frac{\sigma}{x_e} \cdot 100\%.$$

Узгодженість між ранжуваннями двох експертів можна оцінити коефіцієнтом Спірмена:

$$\rho = 1 - \frac{6\sum_{i=1}^{n}(x_{ij} - x_{ik})^2}{n(n^2-1)} = 1 - \frac{6\sum_{i=1}^{n} d_i^2}{n(n^2-1)},$$

де x_{ij} – ранг, присвоєний i-му об'єкту j-м експертом; x_{ik} – ранг, присвоєний i-му об'єкту k-м експертом; d_i – різниця між рангами, присвоєними i-му об'єкту.

Величина ρ може змінюватися в діапазоні від -1 до +1. При повному збігу оцінок коефіцієнт дорівнює 1. Рівність коефіцієнта -1 спостерігається при найбільшій розбіжності у думках експертів.

Коли необхідно визначити узгодженість у ранжуваннях великої (більше двох) кількості експертів, розраховується коефіцієнт конкордації

$$W = \frac{12 \cdot S}{m^2(n^3 - n)},$$

де $S = \sum_{i=1}^{n}\left(\sum_{j=1}^{m} x_{ij} - \frac{1}{2}m(n+1)\right)^2$.

Коефіцієнт W змінюється в діапазоні від 0 до 1. Його рівність одиниці означає, що всі експерти присвоїли об'єктам однакові ранги. Чим ближче значення коефіцієнта до нуля, тим менш узгодженими є оцінки експертів.

Для забезпечення спільної роботи бази правил з механізмом вводу-виводу використаємо систему, структура якої наведено на рис. 6.7:

– фазифікатор вирішує задачу перетворення чітких значень x, у нечіткі визначення вхідних величин $\tilde{x}$;

– машина нечіткого логічного висновку, що на основі правил бази знань визначає значення вихідної змінної у вигляді нечіткої множини $\tilde{x}$, яка відповідає нечітким значенням вхідних змінних $\tilde{x}$;

– композиція. Всі нечіткі висновки всіх правил поєднуються для того, щоб сформувати одну нечітку множину для виходів всіх правил;

– дефазифікатор перетворює вихідну нечітку множину $\tilde{y}$ в конкретне (чітке) число y.

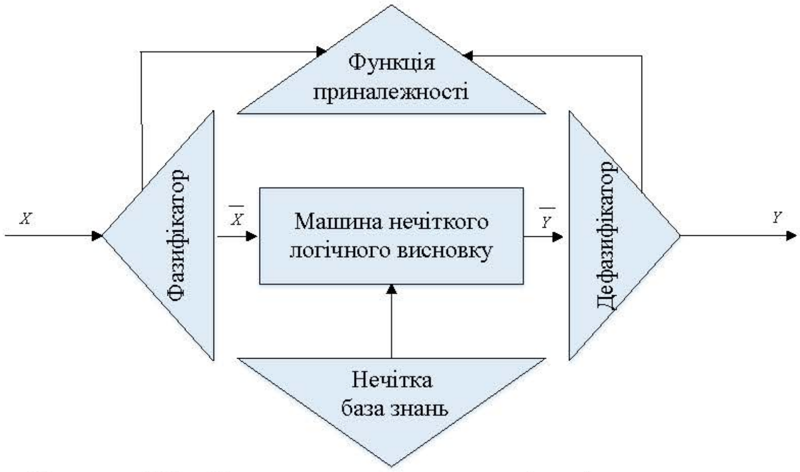


Рисунок 6.7 – Типова структура моделі нечіткого виведення

Якщо база знань сформована і функції належності усіх нечітких даних визначені, то знаходження виходу при заданих значеннях входів здійснюється за допомогою макси-мінного правила нечіткого висновку (5.42). Існують різні способи отримання нечіткого висновку: Мамдані, Цукамото, Ларсена, Сугено. Найпоширенішим є висновок Мамдані. Для цього з *бази нечітких даних* вибираються рядки з однаковими значеннями результату і з них утворюється *диз'юнктивна нормальна форма*

$$\forall y_j \in Y : \mu_y\left(y = y_j\right) = \bigcup_{j: y = y_j} \bigcap_{i=1}^{5} \mu_{x_{ij}} . \tag{6.45}$$

Операція $\bigcap$ виконується шляхом пошуку мінімуму, а операція $\bigcap$ – шляхом пошуку максимуму.

Нехай, наприклад, дані про незалежні нечіткі аргументи оцінені експертами і занесені до таблиці 6.4. Якісні фактори можуть бути оцінені термами {*“відсутній”, “малий”, “нижче середнього”, “середній”, “вище середнього”, “великий”*}, яким для зручності ставиться у відповідність шкала у діапазоні (1, 6).

Припустимо, що всі експерти, які надавали наведені дані, мають однаковий *ступінь впевненості* в кожному з них, який можна подати у вигляді трикутної функції належності

$$\mu_i(x) = \begin{cases} \dfrac{D/2 - x_i + x}{D/2}, & \max\left(1, x_i - D/2\right) \le x \le x_i \\ \dfrac{D/2 + x_i + x}{D/2}, & x_i \le x \le \min\left(D, x_i + D/2\right) \end{cases},$$

де D – діапазон шкали (у нашому випадку D=6).

Таблиця 6.2 – База знань

Номер правила	**Вхідні дані**					**Результат**
j	x_1	x_2	x_3	x_4	x_5	**Y**
1	5	4	4	5	5	4
2	2	3	1	5	4	2
3	3	3	2	4	5	3
4	5	5	5	5	4	4
5	2	3	1	5	3	1
6	2	3	2	4	3	2
7	4	4	4	6	2	6
8	4	3	4	5	3	3
9	4	4	3	3	2	5
10	3	3	5	3	4	3

Нехай тепер задана інша *експертна оцінка* аргументів

Таблиця 6.3 – Значення аргументів

x_{10}	x_{20}	x_{30}	X_{40}	x_{50}	**y**
3	4	2	2	5	

Визначимо для цього набору аргументів *функцію належності* результату.

Користуючись даними таблиці 6.4, знаходимо $\mu_y(y=1...6)$:

$$\mu_y(y=1)=\max_{j:y_j=1}\left[\min_{i=1}^{5}\mu_{x_{ij}}(x_{i0})\right]=\max_{j=5}[\min(\mu_{x_{15}=2}(x_{10}=3),\mu_{x_{25}=3}(x_{20}=4),$$

$$\mu_{x_{35}=1}(x_{30}=2),\mu_{x_{45}=5}(x_{40}=2),\mu_{x_{55}=3}(x_{50}=5))]=\max_{j=5}\left[\min(0.7,0.7,0.7,0,1)\right]=$$

$$=\max[0]=0;$$

$$\mu_y(y=1)=\max_{j:y_j=1}\left[\min_{i=1}^{5}\mu_{x_{ij}}(x_{i0})\right]=\max_{j=5}[\min(\mu_{x_{15}=2}(x_{10}=3),\mu_{x_{25}=3}(x_{20}=4),$$

$$\mu_{x_{35}=1}(x_{30}=2),\mu_{x_{45}=5}(x_{40}=2),\mu_{x_{55}=3}(x_{50}=5))]=\max_{j=5}\left[\min(0.7,0.7,0.7,0,1)\right]=$$

$$=\max[0]=0;$$

$$\mu_y(y=2)=\max_{j:y_j=2}\left[\min_{i=1}^{5}\mu_{x_{ij}}(x_{i0})\right]=\max_{j=2,6}\{\min[\mu_{x_{12}=2}(x_{10}=3),\mu_{x_{22}=3}(x_{20}=4),$$

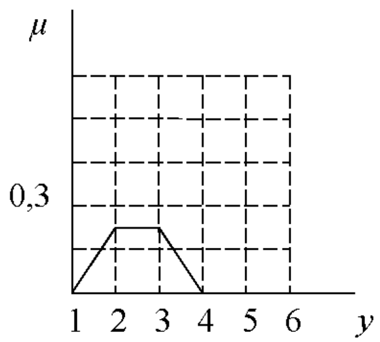
$$\mu_{x_{32}=1}(x_{30}=2),\mu_{x_{42}=5}(x_{40}=2),\mu_{x_{52}=4}(x_{50}=5)],\min[\mu_{x_{16}=2}(x_{10}=3),\mu_{x_{26}=3}(x_{20}=4),$$

$$\mu_{x_{36}=2}(x_{30}=2),\mu_{x_{46}=4}(x_{40}=2),\mu_{x_{56}=3}(x_{50}=5)]\}=$$

$$=\max_{j=2,6}\left[\min_{j=2}(0.7,0.7,0.7,0,0.7),\min_{j=6}(0.7,0.7,1,0.3,0.3)\right]=\max[0,0.3]=0.3.$$

Аналогічно для $y=3\ ...\ 6$.

Таким чином, функція належності показаний на рис. 6.8.



З графіка видно, що найможливіший рівень значення функції для даного прикладу є y=2,5, тобто між "малим" і "нижчим середнього". Конкретний результат моделювання отримується за допомогою операції *дефазифікації*. Ця операція може здійснюватися різними способами і давати відповідно різні результати:

– центроїдний метод (Center of Area):

для неперервної вихідної змінної

$$y = \frac{\int\limits_{\Omega} y\mu_y(y)dy}{\int\limits_{\Omega} \mu_y(y)dy}, \tag{6.46}$$

для дискретної змінної

$$y = \frac{\sum\limits_{i=1}^{n} y_i\mu_y(y_i)}{\sum\limits_{i=1}^{n} \mu_y(y_i)}; \tag{6.47}$$

– максимум функції належності

$$y: \max_y \mu_y(y).$$

Висновок на основі бази знань може здійснюватися також з урахуванням ваги правил. Вага правила встановлюється коефіцієнтом в інтервалі [0, 1] і обирається залежно від достовірності правила (враховуючи джерело його походження). Тоді база знань для наведеного прикладу матиме вигляд

Таблиця 6.4 – Форма бази знань з вагами правил

Номер правила	**Вхідні дані**					**Результат**	**Вага правила**
j	**x_1**	**x_2**	**x_3**	**x_4**	**x_5**	**y**	**d**
…	…	…	…	…	…	…	…

Висновок здійснюватиметься за правилом

$$\forall y_j \in Y : \mu_y\left(y = y_j\right) = \bigcup_{j:y=y_j} \left[d_j \cdot \bigcap_{i=1}^{5} \mu_{x_{ij}} \right], \tag{6.48}$$

тобто, перед знаходженням максимуму належності серед результатів застосування окремих правил ці результати (тобто мінімальні значення належності у рядку) множаться на вагу відповідного правила.

6.2.3 Узгодження моделей

Побудова формальних моделей зазвичай здійснюється шляхом висунення гіпотез про характер процесів і навколишнього середовища для зведення вихідної задачі на природній мові до однієї з типових схем моделювання. Зазвичай виконання всіх цих припущень перевіряється для кожного конкретного випадку застосування моделі. При цьому, природно, відсутні як точне виконання цих гіпотез на практиці, так і відповідні кількісні характеристики ступеня застосовності моделі для кожної конкретної ситуації, тому в модель об'єкта вони не входять. Тим часом такі характеристики дуже важливі, особливо за наявності декількох моделей і необхідності їх узгодження та координації.

В даний час як основа мови, що дозволяє провести узгодження в умовах невизначеності, як вже зазначалося вище, широке застосування отримала теорія нечітких множин. Застосування цієї теорії тим більш виправдано, якщо є різні вихідні припущення про характер невизначеності параметрів процесу.

Нехай розв'язок для кожної з m моделей задано нечіткою множиною D_i на множині розв'язків X, тоді узгоджений розв'язок D може бути поданий як перетин окремих нечітких розв'язків:

$$D = D_1 \cap D_2 \cap \ldots \cap D_m.$$

Для кожного нечіткого розв'язку D_i задається функція належності $\mu_{D_i}(x)$, що отримана в результаті дослідження застосовності моделі або вихідного припущення для процесу.

Факт виконання k припущень про характер процесів враховується введенням показника $L_i = L_{i1}, L_{i2}, \ldots, L_{ik}$, де $0 \le L_{ij} \le 1$. При цьому, якщо припущення j про процес для моделі i ідеально виконується, – $L_{ij} = 1$, якщо це припущення не виконується, $L_{ij} = 0$, в інших випадках $0 < L_{ij} < 1$. Тоді підсумкова функція належності виглядатиме [5]:

$$\mu_D(x) = \left[\mu_{D_1}(x)\right]^{L_1} \wedge \ldots \wedge \left[\mu_{D_m}(x)\right]^{L_m}.$$

Параметри L_i, що характеризують нечіткість моделей при порушенні умов їх існування, можуть бути визначені з урахуванням дисперсії моделей, яка обчислюється в результаті ідентифікації моделей для реальних об'єктів. У випадку, якщо показник L_i неможливо визначити об'єктивно, він може бути визначений на основі експертних оцінок і повинен характеризувати ступінь здійснимості припущень, точність і застосовність моделі в даній конкретній ситуації.