

МІНІСТЕРСТВО ОСВІТИ І НАУКИ УКРАЇНИ
ДВНЗ «Ужгородський національний університет»
Факультет інформаційних технологій
Кафедра інформаційних управляючих систем та технологій

Навчально-методичний посібник до курсу

“ ТЕОРІЯ РОЗПІЗНАВАННЯ ОБРАЗІВ ”

частина I

Ужгород-2016

Копча-Горячкіна Галина Ернестівна

ТЕОРІЯ РОЗПІЗНАВАННЯ ОБРАЗІВ. Частина I: Навчально-методичний посібник для студентів факультету інформаційних технологій напрямів „Комп’ютерні науки” та „Програмна інженерія”. – Ужгород: Видавництво ДВНЗ «Ужгородського національного університету», 2016 р.

Навчально-методичний посібник містить деякі теоретичні відомості, опис предмета навчальної дисципліни у відповідності до болонського процесу, навчально-тематичний план дисципліни «Теорія розпізнавання образів», зміст лекційних тем курсу, тем практичних занять та тем самостійної та індивідуальної роботи студентів, літературу.

Друкується за рішенням кафедри інформаційних управляючих систем та технологій від . . . 2016р., протокол № .

Г.Е.Копча-Горячкіна, 2016.

Вступ

Розпізнавання образів є однією з фундаментальних проблем теорії інтелектуальних систем. З іншого боку, задача розпізнавання образів має величезне практичне значення. Оскільки практичні аспекти стосуються різних галузей сучасного життя, то спеціалістам напрямків «Комп'ютерні науки» та «Програмна інженерія» потрібно чітко розуміти основи теорії розпізнавання образів. Засвоївши базові поняття даної дисципліни, вони зможуть формулювати найпростіші прикладні задачі, створювати моделі систем розпізнавання та ефективно застосовувати потрібні методи до розв'язування практичних задач.

Мета та завдання навчальної дисципліни

Мета дисципліни "Теорія розпізнавання образів" полягає у оволодінні студентами базовими поняттями та знаннями в галузі теорії розпізнавання образів. У результаті вивчення дисципліни студент повинен оволодіти достатніми знаннями для самостійного проведення науково-дослідних проектних робіт зі створення систем розпізнавання образів, у подальшій підготовці при вивченні ряду профілюючих і спеціальних дисциплін, дипломному проектуванні та написанні магістерської роботи.

В результаті вивчення даного курсу студент повинен

знати:

- принципи побудови основних сучасних систем та процеси їх функціонування;
- основні моделі систем і методів розпізнавання образів;

вміти:

- вибирати та використовувати системи при проектуванні та експлуатації;
- мати уявлення щодо сучасного стану та перспектив розвитку інформаційних систем розпізнавання образів.

Програма навчальної дисципліни

Змістовий модуль 1.

Тема 1. Вступна лекція: основні поняття теорії розпізнавання образів, зв'язок із штучним інтелектом. Основні задачі, які виникають в теорії розпізнавання.

Тема 2. Розробка систем розпізнавання. Класифікація та огляд методів розпізнавання. Основні підходи до машинного навчання.

Тема 3. Розпізнавання з використанням гіперплощин. Перцептрони. Алгоритми навчання перцептронів.

Тема 4. Класифікація з використанням функції відстані. Алгоритм найближчого сусіда. Алгоритми кластеризації.

Змістовий модуль 2.

Тема 5. Метод потенціалів. Процедура метода потенціальних функцій.

Тема 6. Статистичні методи розпізнавання. Класифікація з використанням функцій правдоподібності. Класифікатор Байєса.

Тема 7. Розпізнавання графічних образів. Формування і представлення зображень. Методи попередньої обробки та нормалізації зображень.

Тема 8. Структурні методи в теорії розпізнавання. Синтаксичне розпізнавання образів. Навчання і граматичний вивід.

Структура навчальної дисципліни

Назви змістових модулів і тем	Кількість годин											
	денна форма						Заочна форма					
	Усьо- го	у тому числі					Усьо- го	у тому числі				
		л	п	лаб	інд	с.р.		л	п	лаб	інд	с.р.
1	2	3	4	5	6	7	8	9	10	11	12	13
Модуль 1												
Змістовий модуль 1												
Тема 1. Вступна лекція: основні поняття теорії розпізнавання образів, зв'язок із штучним інтелектом. Основні задачі, які виникають в теорії розпізнавання.	8	2				6	10	2				8
Тема 2. Розробка систем розпізнавання. Класифікація та огляд методів розпізнавання. Основні підходи до машинного навчання.	10	2	2			6	12	2				10
Тема 3. Розпізнавання використанням	12	2	2			8	10					10

гіперплощин. Перцептрони. Алгоритми навчання перцептронів.												
Тема 4. Класифікація з використанням функції відстані. Алгоритм найближчого сусіда. Алгоритми кластеризації.	12	4	2			6	10					10
Разом за змістовим модулем 1	42	10	6			26	42	4				38
Змістовий модуль 2												
Тема 5. Метод потенціалів. Процедура метода потенціальних функцій.	12	2	2			8	12	2				10
Тема 6. Статистичні методи розпізнавання. Класифікація з використанням функцій правдоподібності. Класифікатор Байєса.	12	4	2			6	16	4				12

Тема 7. Розпізнавання графічних образів. Формування і представлення зображень. Методи попередньої обробки та нормалізації зображень.	12	4	2			6	10					10
Тема 8. Структурні методи в теорії розпізнавання. Синтаксичне розпізнавання образів. Навчання і граматичний вивід.	12	2	2			8	10					10
Разом за змістовим модулем 2	48	12	8			28	48	6				42
Усього годин	90	22	14			54	90	10				80

Теми практичних занять

№ з/п	Назва теми	Кількість годин
1	Розробка систем розпізнавання. Класифікація та огляд методів розпізнавання. Основні підходи до машинного навчання.	2
2	Розпізнавання з використанням гіперплощин. Перцептрони. Алгоритми навчання перцептронів.	2
3	Класифікація з використанням функції відстані. Алгоритм найближчого сусіда. Алгоритми кластеризації.	2
4	Метод потенціалів. Процедура метода потенціальних функцій.	2
5	Статистичні методи розпізнавання. Класифікація з використанням функцій правдоподібності. Класифікатор Байєса.	2
6	Розпізнавання графічних образів. Формування і представлення зображень. Методи попередньої обробки та нормалізації зображень.	2
7	Структурні методи в теорії розпізнавання. Синтаксичне розпізнавання образів. Навчання і граматичний вивід	2
	Разом	14

Самостійна та індивідуальна робота

№ з/п	Назва теми	Кількість годин
1	Основні поняття теорії розпізнавання образів, зв'язок із штучним інтелектом. Основні задачі, які виникають в теорії розпізнавання.	6
2	Розробка систем розпізнавання. Класифікація та огляд методів розпізнавання. Основні підходи до машинного навчання.	6
3	Розпізнавання з використанням гіперплощин. Перцептрони. Алгоритми навчання перцептронів.	8
4	Класифікація з використанням функції відстані. Алгоритм найближчого сусіда. Алгоритми кластеризації.	6
5	Метод потенціалів. Процедура метода потенціальних функцій.	8
6	Статистичні методи розпізнавання. Класифікація з використанням функцій правдоподібності. Класифікатор Байєса.	6
7	Розпізнавання графічних образів. Формування і представлення зображень. Методи попередньої обробки та нормалізації зображень.	6
8	Структурні методи в теорії розпізнавання. Синтаксичне розпізнавання образів. Навчання і граматичний вивід	8
	Разом	54

Рекомендована література

Базова

1. Головкин Б.А. Машинное распознавание и линейное программирование. — М.: Советское радио. 1973. — 100 с.
2. Гренандер У. Лекции по теории образов. Т.1. Синтез образов. – М.: Мир, 1979. – 382с.
3. Патрик Э. Основы теории распознавания образов. — М.: Советское Радио, 1980. — 408 с.

Допоміжна

1. Потапов А.С. Распознавание образов и машинное восприятие: Общий подход на основе принципа минимальной длины описания. – СПб.: Политехника, 2007. – 548 с.
2. Ту Дж., Гонсалес Р. Принципы распознавания образов. – М.: Мир, 1978 – 412с.
3. Шапиро Л., Стокман Дж. Компьютерное зрение. – М.: Бинум, 2006. – 752 с.
4. Ф.Уосермен. Нейрокомп'ютерна техніка: Теорія і практика. Переклад українською І.Ю.Юрчак, 2001.

Основні поняття теорії розпізнавання образів.

Із завданням розпізнавання образів живі системи, у тому числі і людина, стикаються постійно з моменту своєї появи. Зокрема, інформація, що надходить з органів чуттів, обробляється мозком, який у свою чергу сортує інформацію, забезпечує прийняття рішення, а далі за допомогою електрохімічних імпульсів передає необхідний сигнал далі, наприклад, органам руху, які реалізують необхідні дії. Потім відбувається зміна навколишнього оточення, і вищевказані явища відбуваються заново. І якщо розібратися, то кожен етап супроводжується розпізнаванням.

З розвитком обчислювальної техніки стало можливим вирішити ряд завдань, що виникають в процесі життєдіяльності, полегшити, прискорити, підвищити якість результату. Наприклад, винайдено і працюють різні системи життєзабезпечення, роботизовані системи, відбувається взаємодія людини з комп'ютером. Хоча забезпечити задовільний результат в деяких завданнях (розпізнавання подібних об'єктів, що швидко рухаються, та ін.) ще не вдалося.

Приклади задач розпізнавання образів

1. Розпізнавання літер.
2. Розпізнавання штрих-кодів.
3. Розпізнавання автомобільних номерів.
4. Розпізнавання осіб та інших біометричних даних.
5. Розпізнавання мови.
6. Розпізнавання зображень.
7. Розпізнавання локальних ділянок земної кори, у яких знаходяться родовища корисних копалин.

Основні постановки задач розпізнавання образів.

Розпізнавання образів є однією з найфундаментальніших проблем теорії інтелектуальних систем. З іншого боку, задача розпізнавання образів має величезне практичне значення.

Замість терміна “розпізнавання” часто вживається інший – “класифікація”. Ці два терміни у багатьох випадках розглядаються як синоніми, але не є повністю взаємозамінюваними. Кожен з них має свої сфери застосування, і інтерпретація обох термінів часто залежить від специфіки конкретної задачі.

Наведемо деякі типові постановки задач розпізнавання.

1. **Задача ідентифікації**, яка полягає в тому, щоб вирізнити певний конкретний об’єкт серед його подібних (наприклад, впізнати серед інших людей свою дружину).

2. **Віднесення об’єкта до того чи іншого класу або класифікація**. Це може бути, наприклад, задача розпізнавання літер або прийняття рішення про наявність дефекту у деякій технічній деталі. Віднесення об’єкта до певного класу відображає найтипівішу проблему класифікації, і, коли говорять про розпізнавання образів, найчастіше мають на увазі саме цю проблему. Тому ми розглядатимемо її першочергово.

3. **Кластерний аналіз**, який полягає в поділі заданого набору об’єктів на класи – групи об’єктів, схожі між собою за тим чи іншим критерієм. Цю задачу часто називають **класифікацією без учителя**, оскільки, на відміну від задачі 2, класи апріорно не задані.

Розпізнавання образів (об’єктів, сигналів, ситуацій, явищ або процесів) – завдання ідентифікації об’єкта або визначення будь-яких його властивостей за його зображенням (оптичне розпізнавання) або аудіозапису (акустичне розпізнавання) та іншими характеристиками.

Одним з базових понять є поняття множини. У комп'ютері множина представляється набором неповторюваних однотипних елементів. Слово "неповторюваних" означає, що якийсь елемент у множині або є, або його там немає. Універсальна множина включає всі можливі для розв'язуваної задачі елементи, порожня не містить жодного.

Образ – класифікаційне угруповання в системі класифікації, що об'єднує (виділяє) певну групу об'єктів за деякою ознакою. Образи володіють характерною властивістю, що виявляється в тому, що ознайомлення з кінцевим числом явищ з однієї і тієї ж множини дає можливість дізнаватися про як завгодно велике число його представників. Образи мають характерні об'єктивні властивості в тому сенсі, що різні люди, що навчаються на різному матеріалі спостережень, здебільшого однаково і незалежно один від одного класифікують одні й ті ж об'єкти. У класичній постановці завдання розпізнавання універсальна множина розбивається на частини-образи. Кожне відображення якого-небудь об'єкта на сприймаючі органи системи, що розпізнає, незалежно від його положення щодо цих органів, прийнято називати зображенням об'єкта, а множини таких зображень, об'єднані якими-небудь загальними властивостями, являють собою образи.

Методика віднесення елемента до якогось образу називається вирішальним правилом. Ще одне важливе поняття – метрика, спосіб визначення відстані між елементами універсальної множини. Чим менша ця відстань, тим більш схожими є об'єкти (символи, звуки та ін.) – те, що ми розпізнаємо. Звичайно елементи задаються у вигляді набору чисел, а метрика – у вигляді функції. Від вибору представлення образів і реалізації метрики залежить ефективність програми. Один і той самий алгоритм розпізнавання з різними метриками буде помилятися з різною частотою.

Навчанням зазвичай називають процес вироблення в деякій системі тієї чи іншої реакції на групи зовнішніх ідентичних сигналів шляхом багаторазового впливу на систему зовнішньої корегування. Таке зовнішнє коригування у навчанні прийнято називати "заохоченнями" і "покараннями".

Механізм генерації цього корегування практично повністю визначає алгоритм навчання. Самонавчання відрізняється від навчання тим, що тут додаткова інформація про вірність реакції системі не повідомляється.

Адаптація – це процес зміни параметрів і структури системи, а можливо – і управляючих впливів, на основі поточної інформації з метою досягнення певного стану системи при початковій невизначеності в умовах роботи, що змінюються.

Навчання – це процес, в результаті якого система поступово набуває здатність відповідати потрібними реакціями на певні сукупності зовнішніх впливів, а адаптація – це підстроювання параметрів і структури системи з метою досягнення необхідної якості управління в умовах безперервних змін зовнішніх умов.

Проблеми розпізнавання легко вирішуються людьми, як правило, підсвідомо. Спроби ж побудувати штучні системи розпізнавання ненастільки переконливі. Основна проблема полягає в тому, що часто неможливо адекватно визначити ознаки, на основі яких слід здійснювати розпізнавання. Для задач, для яких такі ознаки вдається виокремити, штучні системи розпізнавання набули значного поширення.

Методи розпізнавання образів і технічні системи, що реалізують ці методи, широко використовуються на практиці. Наведемо деякі з них.

1. Технічна діагностика.

На виробництві часто виникає потреба автоматизувати контроль якості деталей. Задача полягає в тому, щоб виявити, чи є деталь дефектною, чи ні. Якщо ж з'ясується, що деталь має дефект, часто потрібно визначити тип цього дефекту.

2. Медична діагностика.

Системи розпізнавання часто використовуються і в медичній практиці. Найтипівіша ситуація полягає в тому, що ті чи інші захворювання діагностуються на основі аналізу кардіограм, рентгенівських знімків і т. п.

3. Розпізнавання літер.

Окрім усього іншого, ця проблема має велике значення для власне комп'ютерних технологій. Системи розпізнавання літер працюють разом зі сканерами – пристроями, які використовуються для введення до комп'ютера друкованих зображень і текстів. При введенні друкованого тексту сканер формує лише графічне зображення; для того щоб створити текстовий документ, з яким може працювати текстовий редактор, необхідно впізнати на цьому зображенні окремі літери. Аналогічно розпізнавання літер є необхідним для підтримки пристроїв рукописного введення. Цими пристроями, зовні схожими на звичайну авторучку, часто комплектуються надпортативні комп'ютери (персональні помічники). Основна мета цих пристроїв – замінити введення з клавіатури, що є незручним для багатьох користувачів.

4. *Розпізнавання мови.*

Сьогодні інтенсивно розвиваються технології, пов'язані, по перше, з голосовим керуванням комп'ютером, а по друге – з введенням текстів з голосу.

5. *Робототехніка.*

Застосування методів розпізнавання в робототехніці є абсолютно природним і необхідним, оскільки роботи повинні безпосередньо сприймати зовнішній світ і, відповідно, мати пристрої машинного зору.

6. Охоронні системи. Застосування методів розпізнавання в охоронних системах пов'язано в першу чергу з проблемою ідентифікації. Наприклад, потрібно ідентифікувати певну особу, щоб визначити, чи має вона право входити на територію, що охороняється. Розвиваються також системи, які вирішують проблему ідентифікації відбитків пальців і т. п.

Класи та їх властивості

Одне з найвдаліших формулювань ключової парадигми теорії розпізнавання таке: будь-який об'єкт у природі є унікальним, всі об'єкти є типізованими.

Зміст цієї парадигми такий. Кожний об'єкт характеризується тими чи іншими властивостями. Наявність чи відсутність таких властивостей, а також

якісні та кількісні характеристики цих властивостей розглядаються як *ознаки об'єкта*. Унікальність будь-якого об'єкта означає те, що в природі не існує двох різних об'єктів, для яких збігаються абсолютно всі ознаки, а це дозволяє, принаймні теоретично, відрізнити один об'єкт від іншого. Але деякі ознаки різних об'єктів можуть збігатися, і це дає підстави говорити про те, що ці об'єкти належать до одного типу або класу.

Фундаментальні поняття “клас” та “об'єкт” неможливо повністю формалізувати. Спробуємо навести їх неформальні визначення.

Об'єктом у теорії розпізнавання прийнято називати будь-яку сутність, що існує або могла б існувати в реальному світі, а також будь-яке явище або процес.

Це дуже широке визначення, подальші уточнення можуть бути пов'язані з тим чи іншим звуженням нашого розуміння про те, що саме слід вважати об'єктом. Так, реально існуюча Ейфелева вежа буде вважатися об'єктом практично у будь-якій інтерпретації, а “розвиток наукових досліджень про етнокультурний стан Ефіопії та Антарктиди за період з липня 1898 року по січень 1927 року” – можливо, ні. Тут усе залежить від специфіки конкретної задачі і від мети, з якою вона вирішується.

Класом у теорії розпізнавання образів прийнято називати сукупність об'єктів, які мають ті чи інші спільні ознаки.

Клас може об'єднувати фізично існуючі сутності (наприклад, людина, яблуко) або бути абстрактним поняттям (горе, економічний крах і т.п.).

Ознаки, що дають можливість відрізнити представників одного класу від іншого, прийнято називати *інформативними ознаками*.

Ознаки, спільні для всіх представників класу, називатимемо *інваріантами класу*.

Іноді вживають формальніше визначення класу: *класом* називається сукупність об'єктів, пов'язаних між собою деяким відношенням еквівалентності, або, в крайньому разі, толерантності. Нагадаємо, що відношенням еквівалентності називається відношення, яке є симетричним,

рефлексивним і транзитивним (наприклад, відношення “дорівнювати”). Для відношення ж толерантності властивість транзитивності в цілому не виконується (наприклад, відношення “бути схожим”).

Набори інформативних та інваріантних ознак можуть збігатися, але це зовсім не обов’язково.

Інколи, щоб уникнути непорозумінь, ми називатимемо об’єкти і класи Реального світу відповідно P -об’єктами і P -класами. Зрозуміло, що P -об’єкт може належати будь-якій кількості P -класів.

P -об’єкти часто називаються *реалізаціями*, або *зразками* P -класів.

Можна виокремити такі основні властивості класів:

1. Усі представники класу мають певний набір спільних ознак (впливає з визначення).

2. Змінюваність реалізацій класів. По-перше, різні об’єкти, що належать одному класу, можуть бути не схожими між собою. Вони повинні мати спільні інваріантні ознаки, але всі інші ознаки можуть як завгодно варіювати. По-друге, один і той самий об’єкт може змінюватися з часом і навіть поступово переходити від одного класу до іншого (наприклад, перетворення пуголовка на жабу). Усе це свідчить про те, що розпізнати чіткі межі класу часто неможливо.

3. Ознайомлення з деякою скінченою кількістю представників одного класу дає можливість впізнавати інших представників цього класу.

Взагалі кажучи, ця властивість може не виконуватися. Але якщо ця властивість виконується (принаймні, теоретично), це є дуже сильним і важливим твердженням. Воно по суті означає можливість навчатися на прикладах, тобто на основі спостереження певної кількості прикладів (можливо, разом з контрприкладом), сформулювати правило розпізнавання, яке дає змогу відрізнити представників даного класу від представників іншого (можливо, з певною достовірністю тобто з певним процентом помилок). У деяких випадках правилом розпізнавання може бути предикат, який залежить

від інформативних або інваріантних ознак, інколи правило розпізнавання реалізується у вигляді деякої складної процедури.

Якщо навчання на прикладах неможливе або неефективне, правило розпізнавання інколи можна задати явно. Якщо це зробити не вдається, єдиною можливістю для надійного розпізнавання залишається запам'ятовування всіх можливих представників даного класу. Цей випадок не становить інтересу з теоретичної точки зору, і його можна реалізувати лише, якщо кількість можливих представників не є надто великою.

Створення пристроїв, що виконують функції розпізнавання різних об'єктів, у багатьох випадках відкриває можливість заміни людини як елемента складної системи спеціалізованим автоматом. Така заміна дозволяє значно розширити можливості різних систем, що виконують складні інформаційно-логічні задачі. Якість робіт, що виконує людина на будь-якому робочому місці залежить від кваліфікації, досвіду, сумлінності, стану. У той же час автомат, що її замінює, діє одноманітно і забезпечує завжди однакову якість, якщо він справний.

Але не тільки зазначена заміна і звільнення людини від виконання рутинних операцій є причиною створення і пошуку шляхів створення ряду систем розпізнавання. У деяких випадках людина узагалі не в змозі вирішувати цю задачу зі швидкістю, що задається обставинами, не залежно від якостей і психологічного стану приймаючого рішення (наприклад: протиракетний маневр літака в складних метеоумовах). А автомат з такими задачами може легко справлятися.

Отже, основні цілі заміни людини в задачах розпізнавання зводяться до наступного:

- 1) Звільнення людини від одноманітних рутинних операцій для вирішення інших важливіших задач.
- 2) Підвищення якості виконуваних робіт.
- 3) Підвищення швидкості вирішення задач.

Практичні реалізації методів розпізнавання носять назву *систем розпізнавання* (СР).

Центральним завданням розпізнавання образів є побудова на основі систематичних теоретичних і експериментальних досліджень ефективних обчислювальних засобів (об'єднаних в понятті “Системи розпізнавання”) для віднесення описів з об'єктів, явищ, процесів до відповідних класів.

Таким чином, СР – складна динамічна система, яка складається в загальному випадку з колективу підготовлених фахівців і сукупності технічних засобів здобуття і переробки інформації, що забезпечують на основі спеціально сконструйованих алгоритмів рішення задачі класифікації відповідних об'єктів, явищ або процесів.

Постають такі запитання:

- які функції виконують рецептори в частині первинної обробки результатів виявлення об'єктів, явищ;
- які характеристики ліній передачі даних від рецепторів до мозку як ЦВС;
- які ознаки виділяє система обробки;
- які алгоритми використовує мозок для вирішення завдання класифікації, оптимального управління процесом розпізнавання;
- як людині вдається позбавитися від специфічності, властивої технічним СР і тому подібні.

Отже, розпізнавання образів в техніці – необхідний елемент процесу механізації і автоматизації машин, пристроїв і систем для

- заміни людини там, де використовується важка фізична праця;
- реалізації швидких реакцій в управлінні там, де немає часу на роздуми;
- заміни людини в так званих рутинних операціях, тобто, діях, що повторюються, і не вимагають розумових зусиль.

В результаті зіставлення конкретних рішень і розробок виявилось, що не дивлячись на різноманіття і особливості додатків, завдання створення систем розпізнавання мали багато загального, не залежного від вказаної специфіки.

Ось чому для вироблення методичних підходів теорії розпізнавання мало сенс виділяти загальні прийоми, що повторюються, а їх число природно має бути обмеженим і легко об'єднаним в завдання. Самі ж ці завдання повинні бути ключовими для створення будь-якої системи розпізнавання. В результаті виявилось, що знайдений методичний підхід до побудови систем розпізнавання образів інваріантний до предметної області.

На прикладах видно, що підходи до побудови систем розпізнавання практично нічим не відрізняються, не дивлячись на специфіку самих створюваних систем.

В результаті ми отримали загальні уявлення про послідовність вирішення і складові завдання створення системи розпізнавання. Не дивлячись на відмінність предметних областей підходи до побудови СР – однакові. Система розпізнавань захворювань серця будувалася так само, як і система розпізнавання літаків, але замінити її вона не може. Аналогічно СР літаків не може застосовуватися для вирішення завдань розпізнавання захворювань серця.

Системи розпізнавання об'єктів (явищ), що створюються людиною завжди вузько спеціалізовані на відміну від її власних природних можливостей.

Що ж до загального підходу до побудови будь-якої системи, то тепер, якщо у нас є деяка сукупність об'єктів або явищ, які необхідно розпізнавати (класифікувати), на основі узагальнення дій при створенні СР в двох розглянутих прикладах ми знаємо, що

послідовність вирішення відповідних завдань наступна:

-відповідно до вибраного принципу сукупність об'єктів або явищ підрозділяється на ряд класів (говорять: призначається алфавіт класів);

-розроблюється сукупність ознак (говорять: словник);

-на мові словника ознак описується кожен клас;

-вибираються і (або) створюються засоби визначення ознак;

-на обчислювальних засобах реалізується алгоритм зіставлення апостеріорних і апріорних даних і приймається рішення про результати розпізнавання.

Хоча послідовність дій визначена, залишаються запитання:

-як краще здійснювати розбиття об'єктів (літаки, захворювання і ін.) на класи;

-як накопичувати і обробляти апріорну інформацію;

-із яких міркувань вибирати ознаки;

-як описувати класи на мові ознак;

-на основі яких методів порівнювати апріорну і апостеріорну інформацію;

-коли і як з'являється вся система розпізнавання.

Отже, головні висновки:

1. Завдання, що вирішуються в процесі створення систем розпізнавання, інваріантні відносно предметної області, мають багато спільного, базуються на єдиному методологічному підході.

2. Кожна система розпізнавання індивідуальна і призначається лише для одного цілком конкретного виду об'єктів або явищ.

Якщо знайдена сфера застосування розпізнавання, то відповідна система повинна розроблятися заново з урахуванням нових специфічних властивостей об'єктів (явищ), що визначають як систему вимірів характеристик, так і словник ознак, алфавіт класів і алгоритм ухвалення рішень.

3. СР повинна створюватися методом послідовних наближень внутрішньої структури на її математичній моделі у міру накопичення необхідної інформації.

Задачі створення систем розпізнавання

Задача № 1.

Визначення повного переліку ознак (параметрів), що характеризують об'єкти чи явища, для яких дана система розробляється.

У вирішенні цієї задачі – головне знайти всі ознаки, що характеризують сутність об'єктів (явищ), що розпізнаються. Будь-які обмеження, будь-яка

неповнота приведуть до помилок чи повної неможливості правильної класифікації об'єктів (явищ).

В дійсності навіть ціла група ознак може виявитися неефективною.

Тому для вирішення 1-ої задачі створення СР необхідно знайти всі можливі ознаки, що описують об'єкти розпізнавання, для того, щоб при оцінці ефективності рішень системи не повертатися до цієї задачі, знайшовши обмеженість обраних ознак на наступних етапах розробки.

Але щоб призначати ознаки розпізнавання, необхідно, по перше, зрозуміти, що не існує способів їхньої автоматичної генерації. На сьогодні це під силу тільки людині. Тому говорять, що вибір ознак – евристична операція. По друге, вибір ознак можна здійснювати, маючи уявлення про їхні загальні властивості. З цих позицій досить прийняти, що

ознаки можуть поділятися на:

- детерміновані;
- ймовірісні;
- логічні;
- структурні.

А. Детерміновані ознаки – це такі характеристики об'єктів чи явищ, що мають конкретні і постійні числові значення.

Прикладами детермінованих ознак можуть бути, наприклад, ТТХ бомбардувальників і винищувачів.

Числові значення ознак по кожному з літаків можна інтерпретувати як координати точок, що представляють кожен літак у 11-вимірному просторі ознак.

Необхідно мати на увазі, що в задачах розпізнавання з детермінованими ознаками помилки виміру цих ознак не відіграють ніякої ролі, якщо, наприклад, точність вимірів такої ознаки, як розмах крил літака значно вище (наприклад, 1 мм), ніж розходження цієї ознаки в різних класах літаків (наприклад, 10 м).

Представити таку систему, де використовуються детерміновані ознаки, не так важко:

– розпізнавання приналежності літака, дані якого отримані розвідкою чи з відкритої преси і не прив'язані до класів (бомбардувальник – А1, винищувач – А2 і т.п.);

– розпізнавання на конвеєрі деталей про відмінності геометричних характеристик, якщо помилки вимірів істотно менші від розмірів цих деталей.

Розпізнавання здійснюється шляхом порівняння отриманих розмірів з наявними в базі даних характеристиками деталей.

Б. Ймовірнісні ознаки – це характеристики об'єкта (явища), що носять випадковий характер.

З такими ознаками в основному і мають справу в природі і техніці.

Відрізняються ці ознаки тим, що в силу випадковості відповідної величини ознака одного класу може приймати значення з області значень інших класів, кожний з яких підлягає розпізнаванню в системі.

Якщо ознака не може прийняти значення з області відповідних значень для інших класів, то, отже, маємо справу не з ймовірнісною, а з тою ж детермінованою ознакою. Це саме підкреслює, чому ймовірнісні системи є системами більш загального порядку.

Для того, щоб можна було в умовах випадковості говорити про можливість розпізнавання, варто вимагати, щоб імовірності спостереження значень ознаки у своєму класі були набагато більшими, ніж у чужих. У протилежному випадку дана ознака не дозволить побудувати СР, що використовує опис класів на його основі. Ефективність його недостатня для достовірного рішення і необхідно шукати інші ознаки, що мають більшу роздільну здатність.

Випадкова величина характеризується законом розподілу імовірностей. Тобто аналогічно законом розподілу повинна характеризуватися кожна випадкова ознака.

Приклади випадкових ознак розпізнавання:

– середнє значення потужності сигналу радіолокатора, відбитого від літака (причина – порізаність кругової діаграми розсіювання сигналу

радіолокатора літаком і електронні й атмосферні шуми в тому самому радіолокаційному діапазоні);

– розмір листа рослини (причини – відмінності в харчуванні, освітленості, вологості і т.п.);

– розмір патологічної зміни якого-небудь органа людини (причини – різні стадії захворювання при його виявленні, різні ракурси і перетини спостережень утворення і т.п.).

В. Логічні ознаки розпізнавання – це характеристики об’єкта чи явища, представлені у вигляді елементарних висловлювань про істинність («так», «ні», «істина» чи «хибність»).

Ці ознаки, як ми розуміємо, не мають кількісного вираження, тобто є якісними судженнями про наявність, або про відсутність деяких властивостей чи складових у об’єктів чи явищ.

Приклади логічних ознак:

- біль у горлі пацієнта;
- кашель;
- нежить;
- розчинність реактиву і т.п.

Тут про кожну ознаку можна сказати тільки те, що вона є, або її немає.

До логічних можна віднести також такі ознаки, у яких не важлива величина, а лише факт попадання чи непопадання її в заданий інтервал. (наприклад, крейсерська швидкість літака більше чи менше 2000 км/год).

Г. Структурні ознаки – непохідні (тобто, елементарні, не утворені з інших елементарних ознак) елементи (символи), примітиви зображення об’єкта розпізнавання.

Поява структурних ознак зобов’язана виникненню проблеми розпізнавання зображень з її специфічними особливостями і труднощами.

Приклади структурних ознак:

- а) для зображення прямокутника:
 - горизонтальний відрізок прямої;

– вертикальний відрізок прямої.

б) для будь-якого зображення на екрані дисплея:

– піксель.

Традиційно для опису зображень використовувалися розкладання його в ряди за ортогональними функціями (ряди Фур'є, поліноми Ерміта, Лежандра, Чебишева, розкладання Карунена-Лоева й ін.).

Структурний опис на відміну від розкладань:

– зрозуміліший (фізично) для людини, що вирішує задачу розпізнавання об'єкта;

– прийнятний і для комп'ютерної реалізації при розпізнаванні;

– вільний від трудомісткості обчислень і втрат інформації, властивих розкладанням.

Виявляється, що оперуючи обмеженим числом атомарних (непохідних) елементів (примітивів), можна одержати опис різноманітних об'єктів. Тобто, для об'єктів, що відрізняються, можна мати набір однакових непохідних елементів. Але для того, щоб опис можна було би виконати, поряд з визначенням непохідних елементів повинні вводитися правила комбінування, що визначають способи побудови об'єкта із згаданих непохідних елементів. У результаті два однакових непохідних елементи різних об'єктів можуть бути з'єднані один з одним за різними правилами. Це і буде їх відрізнити.

У цілому для опису якого-небудь об'єкта непохідні елементи поєднуються в ланцюжки (речення) за своїм, характерним тільки для цього об'єкта, набором правил.

У результаті зв'язків з непохідних елементів (структурних ознак) утвориться об'єкт, аналогічно до того, як речення мови будуються шляхом з'єднання слів, які в свою чергу складаються з букв. У цьому структурні методи виявляють аналогію із синтаксисом природної мови. Звідси структурні ознаки носять ще назву лінгвістичних чи синтаксичних.

(Приклад – код Фрімена).

Таким чином, розглянуто дуже докладно 1-шу задачу створення систем розпізнавання – визначення повного переліку ознак (параметрів), що характеризують об'єкти чи явища, для яких дана система розробляється. Головні висновки:

1) Вибір, призначення ознак розпізнавання – евристична операція, що залежить від творчості, винахідливості розробника.

2) Склад ознак, що обираються на цьому етапі, повинен бути якомога більш різнобічним і повним, незалежним від того, можна чи не можна ці ознаки одержати.

3) Вибір ознак повинен здійснюватися в групах детермінованих, випадкових, логічних і структурних.

Задача № 2.

Первісна класифікація об'єктів (явищ), що потрібно розпізнати, складання апіорного алфавіту класів.

Нам уже знайома на описовому рівні ця задача: необхідно вибрати (призначити) класи об'єктів (явищ) розпізнавання. Рішення її здійснюється найчастіше евристично, як і вибір ознак розпізнавання, а логіка її рішення наступна:

1-е – визначається, які рішення можуть прийматися за результатами розпізнавання або людиною, або автоматичною системою керування об'єктом (ціль розпізнавання).

2-е – на основі визначеної вище мети формулюються вимоги до системи розпізнавання, що дозволяють вибрати принцип класифікації.

3-е – складається апіорний алфавіт класів об'єктів (явищ).

Припустимо за результатами деякого методу медичної діагностики стану печінки людини необхідно приймати рішення про метод лікування (див.1-й пункт у розглянутій послідовності рішення задачі апіорної класифікації – ціль). Наскільки серйозне прийняття такого рішення з огляду на можливість хірургічного втручання.

Тоді, мабуть, вимогою до системи (див. 2-й пункт послідовності) є надійне (з високою імовірністю) діагностування кожного захворювання печінки.

Отже, в апіорний алфавіт класів (див. 3-й пункт розглянутої послідовності) необхідно включити всі можливі захворювання печінки, а їх – 11. Тобто, класів розпізнаваних захворювань печінки, діагностування деякою гіпотетичною системою розпізнавання повинне бути 11. Для більш чіткого розуміння назовемо ці класи:

- | | |
|------------------------|-----------------|
| 1.Гострий гепатит. | 7.Абсцес. |
| 2.Хронічний гепатит. | 8.Пухлина. |
| 3.Жирова інфільтрація. | 9.Метастази. |
| 4.Цироз. | 10.Гематома. |
| 5.Кіста проста. | 11.Конкременти. |
| 6.Кіста паразитарна. | |

Зауважимо, що, крім ситуації, запропонованої розглянутої задачі, можливі й інші, коли кількість класів, по яких надійно розпізнаються деякі об'єкти (явища), заздалегідь невідомо і повинне визначатися самою системою розпізнавання. Ця задача називається задачею кластеризації, у якій можна відмовитися уже від евристичного підходу. Однак рішення тут досягається при виборі деяких загальних правил кластеризації, що задає розроблювач системи.

Задача № 3.

Розробка апіорного словника ознак розпізнавання.

Вирішуючи задачу №1, ми повинні були знайти всі можливі ознаки розпізнавання заданих об'єктів чи явищ. Аналогічно при рішенні задачі №2 визначився склад класів.

Тепер, маючи в своєму розпорядженні відповідний перелік і апіорний алфавіт класів, необхідно провести аналіз можливостей вимірювання ознак або розрахунку їх за даними вимірювань, вибрати ті з них, що забезпечуються

вимірюваннями, а також у разі потреби розробити пропозиції і створити нові засоби вимірювань для забезпечення необхідної ефективності розпізнавання.

Таким чином, головний зміст розглянутої задачі побудови системи розпізнавання – створення словника, який буде забезпечуватись реально можливими вимірюваннями.

Однак, вдалий чи поганий набір ознак розпізнавання, вийшов у результаті зазначених дій розробника СР, можна зрозуміти, виконавши випробовування системи розпізнавання в цілому й оцінивши ефективність розпізнавання. Але системи розпізнавання на зазначеному етапі розробки ще не існує. У той же час, як ми помітили, з'явилася необхідність оцінки ефективності. І розглядаючи чергові задачі створення СР, ми побачимо, що розглянута задача залишається актуальною протягом усіх наступних етапів створення системи розпізнавання (опис класів, вибір алгоритму розпізнавання). Тільки методом послідовних наближень удається домогтися вибору словника ознак, що забезпечує бажану якість рішень.

Виходом з положення, що склалося, є можливість створення на даному етапі математичної моделі системи. Математичні моделі СР і використовуються для реалізації зазначених послідовних наближень, про що згадувалося на описовому рівні при розгляді задач побудови систем розпізнавання.

Задача № 4.

Опис класів апріорного алфавіту мовою апріорного словника ознак.

Апріорний опис класів – найбільш трудомістка з задач у процесі створення системи розпізнавання, що вимагає глибокого вивчення властивостей об'єктів розпізнавання, а також і найбільш творча задача.

У рамках цієї задачі необхідно кожному класу поставити у відповідність числові параметри детермінованих і випадкових ознак, значення логічних ознак і речення, складені зі структурних ознак-примітивів.

Значення цих параметрів описів можна одержати із сукупності наступних робіт і дій:

- спеціально поставлені експериментальні роботи чи експериментальні спостереження;
- результати обробки експериментальних даних;
- математичні розрахунки;
- результати математичного моделювання;
- витяги з літературних джерел.

Що ж таке опис класу мовою ознак? Розглянемо це окремо для детермінованих, випадкових, логічних і структурних ознак.

Якщо ознаки розпізнаваних об'єктів – детерміновані, то описом класу може бути точка в n -мірному просторі детермінованих ознак з апіорного словника, сума відстаней якої від точок, що представляють об'єкти даного класу, мінімальна.

Легко собі уявити такий еталон, повернувшись до розглянутих нами таблиць ТТХ літаків. Тут ми маємо справу з 11-вимірним простором ознак. Кожна координата – це одна яка-небудь характеристика, наприклад “екіпаж”. Якщо розглядати тільки одну координату “екіпаж”, то точкою еталона для винищувачів буде – 1, для бомбардувальників – 4. Це точки, суми відстаней яких від усіх винищувачів і всіх бомбардувальників, що представляють ці два класи, мінімальні.

Точно так само це можна зробити за всіма 11 координатами (тобто “стеля”, “розмах крил”, “бомбове навантаження” і т.п.), у результаті чого будемо вже мати справу з точками еталонів в 11-вимірному просторі.

Якщо ознаки розпізнавання – логічні, то для опису кожного класу необхідно насамперед мати повний набір елементарних логічних висловлень A , B , C , що входять до складу апіорного словника. Але це тільки ознаки. Для опису класів цього недостатньо. Ще необхідно встановити відповідність між набором значень вказаних ознак A , B , C і класами W_1, W_2, \dots, W_m .

Так для простоти розуміння і без домагань на медичну вірогідність візьмемо такий приклад: необхідно розпізнавати два захворювання – звичайна застуда й ангіна (W_1, W_2), а як логічні ознаки виберемо:

A – підвищена температура (A=0 – ні, A=1 – так);

B – нежить (B=0 – ні, B=1 – так);

C – нариви в горлі (C=0 – ні, C=1 – так).

Тоді так зване булеве співвідношення між класом W_1 (звичайне простудне захворювання) і значеннями ознак (а ці значення – бінарні) виглядає так

$$W_1 = ABC + \overline{A}BC + A\overline{B}C$$

Тут множення, як ви знаєте, відповідає логічному “І”, а додавання – “АБО”.

Аналогічно для другого класу захворювань одержимо наступний опис

$$W_2 = \overline{A}BC + A\overline{B}C$$

Якщо розподіл об’єктів розпізнавання, що представляються числовими значеннями їхніх ознак за областями відповідного простору випадковий, то для опису класів необхідно визначити характеристики цих розподілів. А з теорії ймовірності відомо, що це

– функції ПРВ $f_i(x_1, x_2, \dots, x_n)$, де x_1, \dots, x_n – випадкові ознаки, i – номер класу;

$P(W_i)$ – апіорна ймовірність того, що об’єкт, випадково обраний із загальної сукупності, виявиться належним класу W_i .

Як одержати ПРВ класів системи розпізнавання? У розпорядженні розроблювача СР – три способи:

- експериментальне визначення за статистичними даними;
- теоретичний висновок;
- моделювання.

Те ж стосується апіорної ймовірності класу $P(W_i)$.

Якщо ознаки розпізнавання – структурні, то описом кожного класу повинен бути набір пропозицій (ланцюжків з непохідних елементів із правилами з’єднання). Кожна з пропозицій класу – характеристика структурних особливостей об’єктів цього класу. Приклад – код Фрімена.

Задача № 5.

Вибір алгоритму класифікації, який забезпечує віднесення об'єкта чи явища, що розпізнається, до відповідного класу.

Безпосереднє вирішення задачі розпізнавання на основі використання словника ознак і алфавіту класів об'єктів чи явищ фактично полягає в розбитті простору значень ознак розпізнавання на області D_1, D_2, \dots, D_n , що відповідають класам W_1, W_2, \dots, W_n (згадуємо визначення “образа”).

Зазначене розбиття повинне бути виконане таким чином, щоб забезпечувалися мінімальні значення помилок віднесення класифікованих об'єктів чи явищ до “чужих” класів.

Результатом такої операції є віднесення об'єкта, що має набір ознак X_1, X_2, \dots, X_n (точка в n -вимірному просторі), до класу W_i , якщо зазначена точка лежить у відповідній до класу області ознак – D_i .

Розбиття простору ознак можна представляти як побудову розподіляючих функцій $f_i(x_1, x_2, \dots, x_n)$ між множинами (областями) ознак D_i , що належать різним класам.

Фізично розпізнавання ґрунтується на порівнянні значень тієї чи іншої міри близькості розпізнаваного об'єкта з кожним класом. При цьому якщо значення обраної міри близькості (подібності) L даного об'єкта w з яким-небудь класом W_g досягає екстремума щодо значень її по інших класах, то приймається рішення про належність цього об'єкта класу W_g . Якщо міра близькості не має екстремума, то ми знаходимося на границі, де не можемо віддати перевагу жодному з класів.

В алгоритмах розпізнавання, що використовують детерміновані ознаки як міру близькості, використовують середньоквадратичну відстань між даним об'єктом w і сукупністю об'єктів (w_1, w_2, \dots, w_n) , що представляють (описують) кожен клас. При цьому як методи вимірювань відстані між об'єктами $d(w, w_g)$ можуть використовуватися будь-які методи (творчий процес тут не обмежується). В алгоритмі розпізнавання, що використовує детерміновані

ознаки можна враховувати і вагу V_j (установлювати ступінь довіри чи важливості). Тоді

В алгоритмах розпізнавання, що використовують випадкові ознаки як міру близькості, використовують ризик, пов'язаний з рішенням про приналежність об'єкта до класу W_i , де i – номер класу ($i=1,2,\dots,m$).

Для алгоритмів, заснованих на логічних ознаках, поняття “міра близькості” не має змісту. Згадаємо спрощений приклад, розглянутий нами для логічних ознак захворювань (простої застуди й ангіни).

Маючи значення ознак A , B , C , досить підставити їх у булеві співвідношення між класами й ознаками, щоб відразу одержати результат як істинність чи хибність булевої функції опису того чи іншого класу.

Дійсно, нехай ознаки прийняли наступні значення:

– Підвищена температура: $A=1$

– Нежить: $B=0$

– Нариви в горлі: $C=1$

Тоді підстановка їх у булеві співвідношення дасть наступний результат:

$$W_1 = ABC + \overline{ABC} + \overline{ABC} = 0$$

$$W_2 = \overline{ABC} + \overline{ABC} = 1$$

Тобто, істинним є друге співвідношення, що відповідає розпізнаванню ангіни як діагностуючого класу з двох захворювань.

Для алгоритмів, заснованих на структурних (лінгвістичних) ознаках, поняття “міри близькості” більш специфічно.

З урахуванням того, що кожен клас описується сукупністю речень, що характеризують структурні особливості об'єктів відповідних класів, розпізнавання невідомого об'єкта здійснюється ідентифікацією речення, що описує цей об'єкт, з одним із речень у складі опису якого-небудь класу.

При цьому ідентифікація може мати на увазі найбільшу подібність речення, що описує розпізнаваний об'єкт із реченнями з наборів опису кожного класу.

Розглянувши задачу №5, ми фактично завершили розгляд кола задач створення СР. У той же час уже відзначалося, що створення СР здійснюється послідовними наближеннями в міру одержання додаткової інформації. У цьому ряді послідовних наближень головну роль відіграють ознаки розпізнавання. Від ефективності їхнього набору залежить ефективність системи в цілому. У процесі вдосконалювання системи зазначений набір поповнюється, неефективні ознаки виключаються. Тому однією з задач створення СР повинна бути і задача переходу від апріорного словника ознак до робочого. Те ж стосується й апріорного алфавіту класів.

Задача № 6.

Визначення робочого алфавіту класів і робочого словника ознак системи розпізнавання.

Справжня задача на рівні розробки, що пройшла етапи рішення задач 1 – 5, принаймні вже може бути поставлена, тому що в результаті виконання попередніх задач створена система розпізнавання першого наближення (апріорний алфавіт класів і апріорний словник ознак, обрано алгоритм розпізнавання).

Суть задачі – розробка такого (робочого) алфавіту класів і такого (робочого) словника ознак, що забезпечили б максимальне значення показника ефективності розпізнавання. Тобто, з апріорного словника ми повинні вибрати ознаки, що дозволяють при всіх наявних обмеженнях на одержання їх вимірювань, забезпечити максимум ймовірності правильної класифікації об'єктів (явищ). Такий вибір не може не припускати оцінку зазначених показників до того, як створена система.

Зазначена суть задачі змушує знову звернути увагу на можливість одержання оцінки ефективності системи розпізнавання шляхом її моделювання. Про це ми говорили при створенні апріорного словника ознак.

Що ж стосується прийомів, що забезпечують добір у процесі оптимізації систем розпізнавання, то вони є також предметом окремого розгляду.

Методи розпізнавання

Виділяють чотири групи методів розпізнавання:

1. Порівняння зі зразком.

Найбільш наочно застосування цього методу видно в розпізнаванні тексту. Якщо маємо зображення відсканованого символу і колекція зображень зразків (усіх букв абетки), і потрібно визначити, якій букві алфавіту відповідає відскановане зображення. Для цього масштабуємо зображення символу до розмірів зразків і вибираємо той, відстань до якого мінімальна.

2. Нейронні мережі.

Вибираємо вид мережі і налаштовуємо коефіцієнти. На вхід нейронної мережі подається об'єкт для розпізнавання. Група рецепторів мережі відповідає за прийом своєї характеристичної властивості.

3. Статистичні методи.

4. Структурні та синтаксичні методи.

Розбираємо об'єкт на елементи. Будуємо правило в залежності від входження окремих елементів та їх послідовностей.

Машинне навчання

1959 році Артур Самюель визначив машинне навчання як “Область науки, що надає комп’ютерам можливість навчатися у випадку, коли явно вони не є запрограмованими”.

Є три основні причини необхідності машинного навчання, коли недостатньо просто запрограмувати комп’ютер. По-перше, розробники не можуть передбачити усі ситуації, в яких може виявитись машина. Наприклад робот, розроблений для проходження лабіринтів, мусить дізнаватися план кожного нового лабіринту, який йому трапляється. По-друге, розробники не можуть передбачити усі зміни з часом: програма для прогнозування цін на фондовій біржі повинна навчатися для адаптації, коли умови на ринку змінюються. По-третє, іноді люди не мають жодного уявлення про те, як запрограмувати розв’язок. Наприклад, більшість людей добре розпізнають обличчя своїх знайомих, але навіть найкращі програмісти не можуть створити програму для виконання цієї задачі, не використовуючи алгоритми навчання.

Навчання, як і інтелект, охоплює такий широкий спектр процесів, що складно дати йому точне визначення. Словникові визначення включають такі фрази, як “отримувати знання, розуміння або вміння за допомогою вивчення, інструктування або досвіду” і “модифікація поведінкової тенденції через досвід.” Зоологи і психологи вивчають навчання людей і тварин. Існують паралелі між тваринним і машинним навчанням. Звичайно, багато технік у машинному навчанні походять від намагань психологів покращити свої теорії людського і тваринного навчання через обчислювальні моделі. Щодо машин можна сказати, що машина навчається коли змінює свою структуру, програму або дані (у відповідь на зовнішню інформацію) таким чином, що її очікувана майбутня продуктивність покращується. Машинне навчання зазвичай застосовується для змін у системах, що виконують завдання, пов’язані зі штучним інтелектом. Такі завдання охоплюють розпізнавання, діагностику, планування, контролювання роботів, прогнозування і таке інше.

“Змінами” можуть бути або покращення вже робочих систем, або синтез цілковито нових.

Розпізнавання образів (pattern recognition) — це розділ теорії штучного інтелекту (artificial intelligence), що вивчає методи класифікації об’єктів. За традицією об’єкт, що піддається класифікації, називається образом (pattern).

Образом може бути цифрова фотографія (розпізнавання зображень), буква або цифра (розпізнавання символів), запис мови (розпізнавання мови) тощо.

В межах теорії штучного інтелекту розпізнавання образів включається в більш широку наукову дисципліну – теорію машинного навчання (machine learning), метою якої є розробка методів побудови алгоритмів, що здатні навчатися.

Існують два підходи до навчання: індуктивне і дедуктивне. Індуктивне навчання, або навчання за прецедентами, засноване на виявленні загальних властивостей об’єктів на підставі неповної інформації, отриманої емпіричним шляхом. Дедуктивне навчання передбачає формалізацію знань експертів у вигляді баз знань (експертних систем тощо).

Як кожна математична дисципліна, розпізнавання образів має власний математичний апарат, який включає математичну статистику, методи оптимізації, дискретну математику, алгебру і геометрію.

Розпізнавання образів має широке застосування і використовується при створенні усіх комп’ютерних систем, на які покладаються інтелектуальні функції, тобто функції, пов’язані із прийняттям рішень замість людини: медична діагностика, криміналістична експертиза, пошук інформації та інтелектуальний аналіз даних тощо.

Прецедент – це об’єкт, приналежність якого до заданого класу визначена заздалегідь. Прецедентом може бути, наприклад, набір ознак пацієнта із відомим діагнозом, з яким слід порівнювати набір ознак людини, діагноз якої ще невідомий.

Кожний образ являє собою набір чисел, що описують його властивості і називаються ознаками (feature). Впорядкований набір ознак об'єкта називається вектором ознак (feature vector). Вектор ознак – це точка в просторі ознак (feature space).

Класифікатор, або вирішальне правило (decision rule) – це функція, яка ставить у відповідність вектору ознак образу клас, до якого він належить.

Задачу розпізнавання образів можна розділити на ряд підзадач.

1. Генерування ознак (feature generation) – вимірювання або обчислення числових ознак, що характеризують об'єкт.

2. Вибір ознак (feature selection) – визначення найбільш інформативних ознак для класифікації (в цей набір можуть входити не лише первинні ознаки, але й функції від них).

3. Побудова класифікатора (classifier construction) – конструювання вирішального правила, на підставі якого здійснюється класифікація.

4. Оцінка якості класифікації (classifier estimation) – обчислення показників правильності класифікації (точність, чутливість, специфічність, помилки першого та другого роду).

Перцептрон як метод розпізнавання образів

Ф. Розенблатт вводячи поняття про модель мозку, завдання якої полягає в тому, щоб показати, як у деякій фізичній системі, структура й функціональні властивості якої відомі, можуть виникати психологічні явища, – описав найпростіші експерименти з розпізнавання. Дані експерименти цілком стосуються і методів розпізнавання образів, але відрізняються тим, що алгоритм розв'язання не детермінований.

Найпростіший експеримент, на основі якого можна одержати психологічно значиму інформацію про деяку систему, зводиться до того, що моделі пред'являються два різних стимули й потрібно, щоб вона реагувала на них різним чином. Метою такого експерименту може бути дослідження можливості спонтанного розпізнавання стимулів системою при відсутності

втручання з боку експериментатора, або, навпаки, вивчення примусового розпізнавання, при якому експериментатор прагне навчити систему здійснювати необхідну класифікацію.

У досвіді з навчанням перцептронів зазвичай пред'являється деяка послідовність образів, у яку входять представники кожного із класів, що підлягають розпізнаванню. Відповідно до деякого правила модифікації пам'яті правильний вибір реакції підкріплюється. Потім перцептронів пред'являється контрольний стимул і визначається ймовірність одержання правильної реакції для стимулів даного класу. Залежно від того, збігається чи не збігається обраний контрольний стимул з одним з образів, які використовувалися в навчальній послідовності, отримують різні результати.

Якщо контрольний стимул не збігається з жодним із навчальних стимулів, то експеримент пов'язаний не тільки з чистим розпізнаванням, але містить у собі й елементи узагальнення.

Якщо контрольний стимул збуджує деякий набір сенсорних елементів, цілком відмінних від тих елементів, які активізувалися при впливі раніше пред'явлених стимулів того ж класу, то експеримент є дослідженням чистого узагальнення.

Перцептрони не мають здатності до чистого узагальнення, але вони цілком задовільно функціонують в експериментах із розпізнавання, особливо якщо контрольний стимул досить близько збігається з одним з образів, щодо яких перцептрон уже нагромадив певний досвід.

Перцептрони і зародження штучних нейронних мереж

Як науковий предмет штучні нейронні мережі вперше заявили про себе у 40-і роки. Прагнучи відтворити функції людського мозку, дослідники створили прості апаратні (а пізніше програмні) моделі біологічного нейрона і системи його з'єднань. Коли нейрофізіологи досягли більш глибокого розуміння нервової системи людини, ці ранні спроби стали сприйматися як досить грубі апроксимації. Проте на цьому шляху були досягнуті вражаючі результати, що

стимулювали подальші дослідження і привели до створення більш витончених мереж.

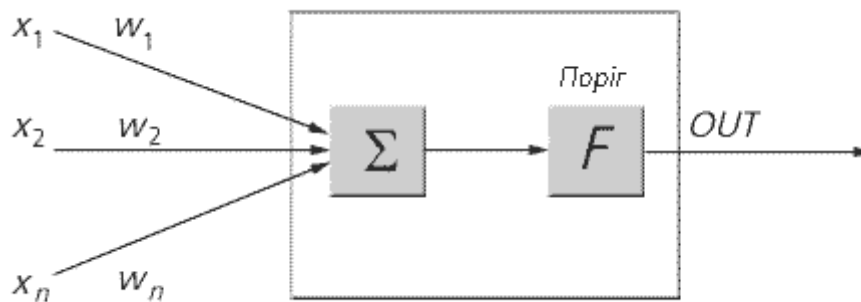


Рис.1. Перцептронний нейрон

Перше систематичне вивчення штучних нейронних мереж було зроблене Маккалоком і Пітсом в 1943 р. Пізніше вони досліджували мережеві парадигми для розпізнавання зображень, що піддаються зсувам і поворотам. Проста нейронна модель, показана на рис. 1, використовувалася в більшій частині їх роботи. Елемент Σ перемножує кожен вхід x з вагою w і підсумовує зважені входи. Якщо ця сума більше заданого порогового значення, вихід рівний одиниці, в іншому випадку нулю. Ці системи (і множина до них подібних) отримали назву *перцептронів*. Вони складаються з одного шару штучних нейронів, сполучених за допомогою вагових коефіцієнтів з множиною входів (див. рис. 2), хоч в принципі описуються і більш складні системи.

У 60-і роки перцептрони викликали великий інтерес і оптимізм. Розенблат довів чудову теорему про навчання перцептронів, що пояснюється нижче. Уидроу дав ряд переконливих демонстрацій систем перцептронного типу, і дослідники у всьому світі прагнули вивчити можливості цих систем. Первинна ейфорія змінилася розчаруванням, коли виявилось, що перцептрони не здатні навчитися рішення ряду простих задач. Мінський суворо проаналізував цю проблему і показав, що є жорсткі обмеження на те, що можуть виконувати одношарові перцептрони, і, отже, на те, чому вони можуть навчатися. Оскільки в той час методи навчання багатошарових мереж не були відомі, дослідники перейшли в більш багатообіцяючі області, і дослідження в області нейронних мереж прийшли в занепад. Недавнє відкриття методів

навчання багатшарових мереж в більшій мірі, ніж який-небудь інший чинник, вплинуло на відродження інтересу і дослідницьких зусиль.

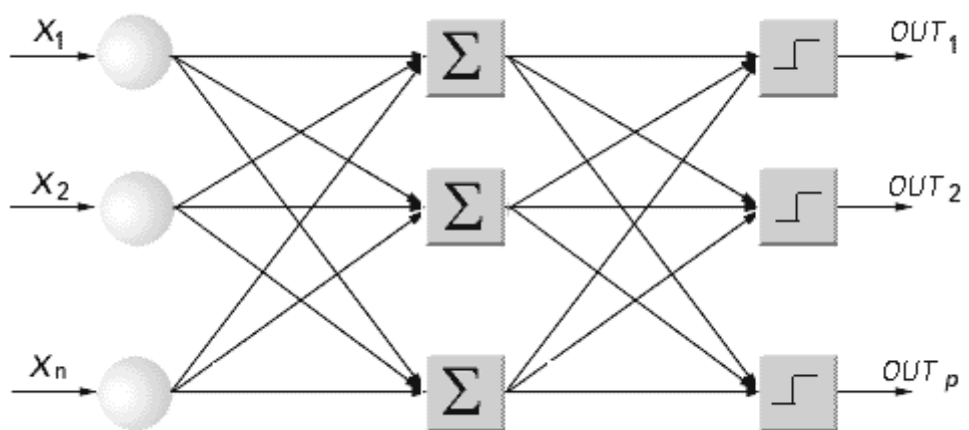


Рис. 2. Перцептрон з багатьма виходами

Робота Мінського, можливо, і охолодила запал ентузіастів перцептрону, але забезпечила час для необхідної консолідації і розвитку теорії, що лежала в основі можливості навчання перцептрону. Важливо зазначити, що аналіз Мінського не був спростований. Він залишається важливим дослідженням і повинен вивчатися, щоб помилки 60-х років не повторилися.

Незважаючи на свої обмеження перцептрони широко вивчалися (хоч не дуже широко використовувалися). Теорія перцептронів є основою для багатьох інших типів штучних нейронних мереж, і перцептрони ілюструють важливі принципи. Внаслідок цих причин вони є логічною початковою точкою для вивчення штучних нейронних мереж.

Перцептронне представлення

Доказ теореми навчання перцептрону показав, що перцептрон здатний навчитися всьому, що він здатний представляти. Важливо при цьому уміти розрізняти представлення і здатність до навчання. Поняття представлення відноситься до здатності перцептрону (або іншої мережі) моделювати певну функцію. Здатність до навчання вимагає наявності систематичної процедури налаштування ваг мережі для реалізації цієї функції.

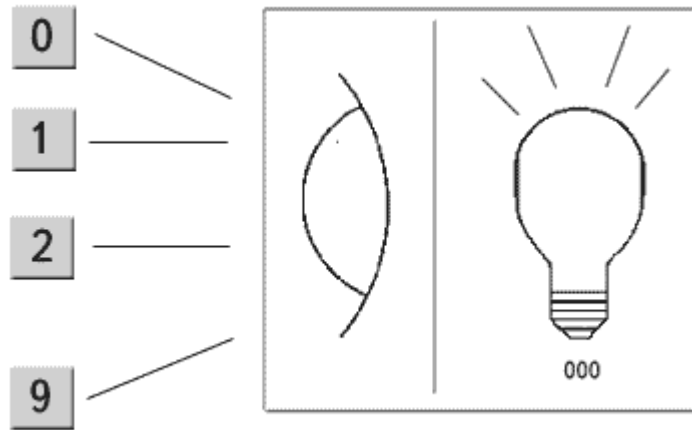


Рис. 3. Система розпізнавання зображень

Для ілюстрації проблеми представимості припустимо, що у нас є множина карт, з цифрами від 0 до 9. Припустимо також, що в нас є гіпотетична машина, що здатна відрізнити карти з непарним номером від карт з парним номером і що запалює індикатор на своїй панелі при пред'явленні карти з непарним номером (див. рис.3). Чи можна представити таку машину перцептром? Тобто, чи може бути сконструйований перцептрон і налаштовані його ваги (неважливо яким чином) так, щоб він мав таку ж роздільну здатність? Якщо це так, то кажуть, що перцептрон здатний представляти бажану машину. Ми побачимо, що можливості представлення одношарових перцептронів досить обмежені. Є багато простих машин, які не можуть бути представлені перцептром незалежно від того, як налаштовуються його ваги.

Проблема функції ВИКЛЮЧАЮЧЕ АБО

Один з самих песимістичних результатів Мінського показує, що одношаровий перцептрон не може відтворити таку просту функцію, як ВИКЛЮЧАЮЧЕ АБО. Це функція від двох аргументів, кожний з яких може бути нулем або одиницею. Вона приймає значення одиниці, коли один з аргументів рівний одиниці (але не обидва). Проблему можна проілюструвати за допомогою одношарової однеї нейронної системи з двома входами, показаною на рис. 4.

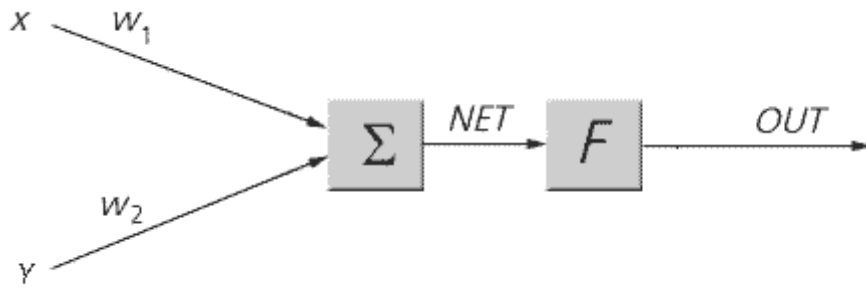


Рис.4. Однонейронна система

Визначимо один вхід через x , а інший через y , тоді всі їх можливі комбінації будуть складатися з чотирьох точок на площині x - y , як показано на рис.5. Наприклад, точка $x=0$ і $y=0$ позначена на рисунку як точка A_0 .

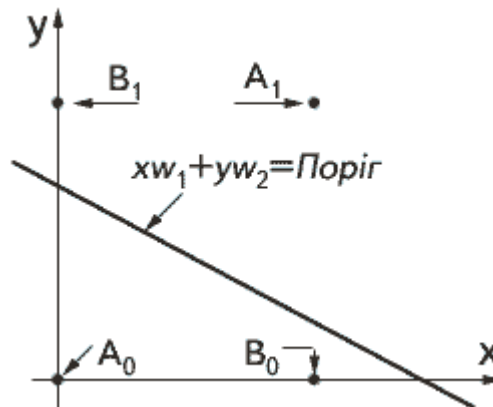


Рис. 5. Проблема ВИКЛЮЧАЮЧЕ АБО

Таблиця 2.1 показує необхідний зв'язок між входами і виходом, де вхідні комбінації, які повинні давати нульовий вихід, помічені A_0 і A_1 , одиничний вихід - B_0 і B_1 .

Таблиця 2.1. Таблиця істинності для функції ВИКЛЮЧАЮЧЕ АБО

Точки	Значення x	Значення y	Необхідний вихід
A_0	0	0	0
B_0	1	0	1
B_1	0	1	1
A_1	1	1	0

У мережі на рис.4 функція F є звичайним порогом, так що OUT приймає значення нуль, коли NET менше 0,5, і одиницю у випадку, коли NET більше або дорівнює 0,5. Нейрон виконує наступне обчислення:

$$NET = xw_1 + yw_2 \quad (2.1)$$

Жодна комбінація значень двох ваг не може дати співвідношення між входом і виходом, що задається у таблиці 2.1. Щоб зрозуміти це обмеження, зафіксуємо NET на величині порога 0,5. Мережа в цьому випадку описується рівнянням (2.2). Це рівняння лінійне щодо x і y , тобто всі значення x і y , що задовольняють цьому рівнянню, будуть лежати на деякій прямій в площині x - y .

$$xw_1 + yw_2 = 0,5 \quad (2.2)$$

Будь-які вхідні значення для x і y на цій лінії будуть давати порогове значення 0,5 для NET . Вхідні значення з одного боку прямої забезпечать значення NET більше порога, отже, $OUT=1$. Вхідні значення по іншу сторону прямої забезпечать значення NET менше порогового значення, роблячи OUT рівним 0. Зміни значень w_1 , w_2 і порога будуть міняти нахил і розташування прямої. Для того щоб мережа реалізовувала функцію ВИКЛЮЧАЮЧЕ АБО, задану у таблиці 2.1, треба розташувати пряму так, щоб точки А були з одного боку прямої, а точки В – з іншою. Спробувавши намалювати таку пряму на рис.5, переконуємося, що це неможливе. Це означає, що які б значення не приписувалися вагам і порогу, мережа нездатна відтворити співвідношення між входом і виходом, необхідне для представлення функції ВИКЛЮЧАЮЧЕ АБО.

Поглянувши на задачу з іншої точки зору, розглянемо NET як поверхню над площиною x - y . Кожна точка цієї поверхні знаходиться над відповідною точкою площини x - y на відстані, рівній значенню NET в цій точці. Можна показати, що нахил цієї NET -поверхні однаковий для всієї поверхні x - y . Всі точки, в яких значення NET дорівнює величині порога, проектується на лінію рівня площини NET (див. рис.6).

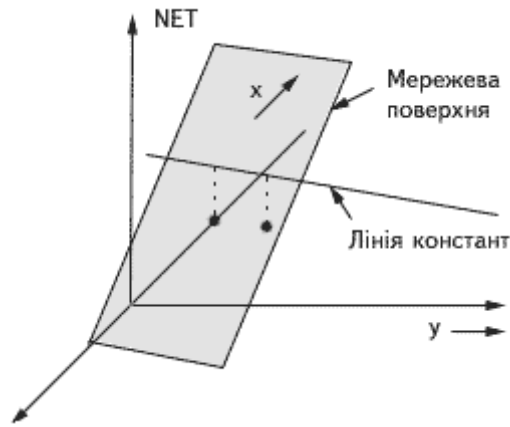


Рис. 6. Персептрона *NET*-площина

Ясно, що всі точки по один бік порогової прямої проєктуються в значення *NET*, більше за поріг, а точки по інший бік дадуть менші значення *NET*. Таким чином, порогова пряма розбиває площину *x-y* на дві області. У всіх точках по один бік порогової прямої значення *OUT* дорівнює одиниці, по інший бік – нулю.

Лінійна роздільність

Як ми бачили, неможливо намалювати пряму лінію, що розділяє площину *x-y* так, щоб реалізовувалася функція ВИКЛЮЧАЮЧЕ АБО. На жаль, цей приклад не єдиний. Є широкий клас функцій, що не реалізуються одношаровою мережею. Про ці функції кажуть, що вони є лінійно нероздільними, і вони накладають певні обмеження на можливості одношарових мереж.

Лінійна роздільність обмежує одношарові мережі задачами класифікації, в яких множини точок (відповідних входним значенням) можуть бути розділені геометрично. Для нашого випадку з двома входами роздільник є прямою лінією. У разі трьох входів розділення здійснюється площиною, що розтинає тривимірний простір. Для чотирьох або більше входів візуалізація неможлива і необхідно уявно представити *n*-вимірний простір, що розтинається "гіперплощиною" геометричним об'єктом, який розтинає простір чотирьох або більшого числа вимірів.

Оскільки лінійна роздільність обмежує можливості перцептронного представлення, то важливо знати, чи є дана функція роздільною. На жаль, не існує простого способу визначити це, якщо число змінних велике.

Нейрон з n двійковими входами може мати 2^n різних вхідних образів, що складаються з нулів і одиниць. Оскільки кожен вхідний образ може відповідати двом різним бінарним виходам (одиниця і нуль), то всього є 2^{2^n} функцій від n змінних.

Таблиця 2.2. Лінійно роздільні функції

n	2^{2^n}	Число лінійно роздільних функцій
1	4	4
2	16	14
3	256	104
4	65536	1882
5	$4,3 \times 10^9$	94572
6	$1,8 \times 10^{19}$	15 028 134

Як видно з таблиці 2.2, імовірність того, що випадково вибрана функція виявиться лінійно роздільною, досить мала навіть для помірному числа змінних. З цієї причини одношарові перцептрони на практиці обмежені простими задачами.

Подолання обмеження лінійної роздільності

До кінця 60-х років проблема лінійної роздільності була добре зрозуміла. До того ж було відомо, що це серйозне обмеження представлення одношаровими мережами можна подолати, додавши додаткові прошарки. Наприклад, двошарові мережі можна отримати каскадним з'єднанням двох одношарових мереж. Вони здатні виконувати більш загальні класифікації, відділяючи ті точки, які містяться в опуклих обмежених або необмежених областях. Область називається опуклою, якщо для будь-яких двох її точок відрізок, що з'єднує їх, повністю лежить в області. Область називається

обмеженою, якщо її можна взяти в деяке коло. Необмежену область неможливо вкласти всередину кола (наприклад, область між двома паралельними лініями). Приклади опуклих обмежених і необмежених областей наведені на рис.7.



Рис.7. Опуклі обмежені і необмежені області

Щоб уточнити вимогу опуклості, розглянемо просту двошарову мережу з двома входами, підведеними до двох нейронів першого прошарку, сполученими з єдиним нейроном в шарі 2 (див. рис.8). Нехай поріг вихідного нейрона рівний 0,75, а обидва його ваги рівні 0,5. У цьому випадку для того, щоб поріг був перевищений і на виході з'явилася одиниця, потрібно, щоб обидва нейрони першого рівня на виході мали одиницю. Таким чином, вихідний нейрон реалізує логічну функцію І. На рис.8 кожен нейрон прошарку 1 розбиває площину x - y на дві півплощини, один забезпечує одиничний вихід для входів нижче верхньої лінії, інший – для входів вище нижньої лінії. На рис. 8 показаний результат такого подвійного розділення, де вихідний сигнал нейрона другого прошарку рівний одиниці тільки всередині V-подібної області. Аналогічно у другому прошарку може бути використано три нейрони з подальшим розділенням площини і створенням області трикутної форми. Включенням достатнього числа нейронів у вхідний прошарок може бути утворений опуклий багатокутник будь-якої бажаної форми. Оскільки вони утворені за допомогою операції І над областями, що задаються лініями, то всі такі многогранники опуклі, отже, тільки опуклі області і виникають. Точки, що не складають опуклої області, не можуть бути відділені від інших точок площини двошаровою мережею.

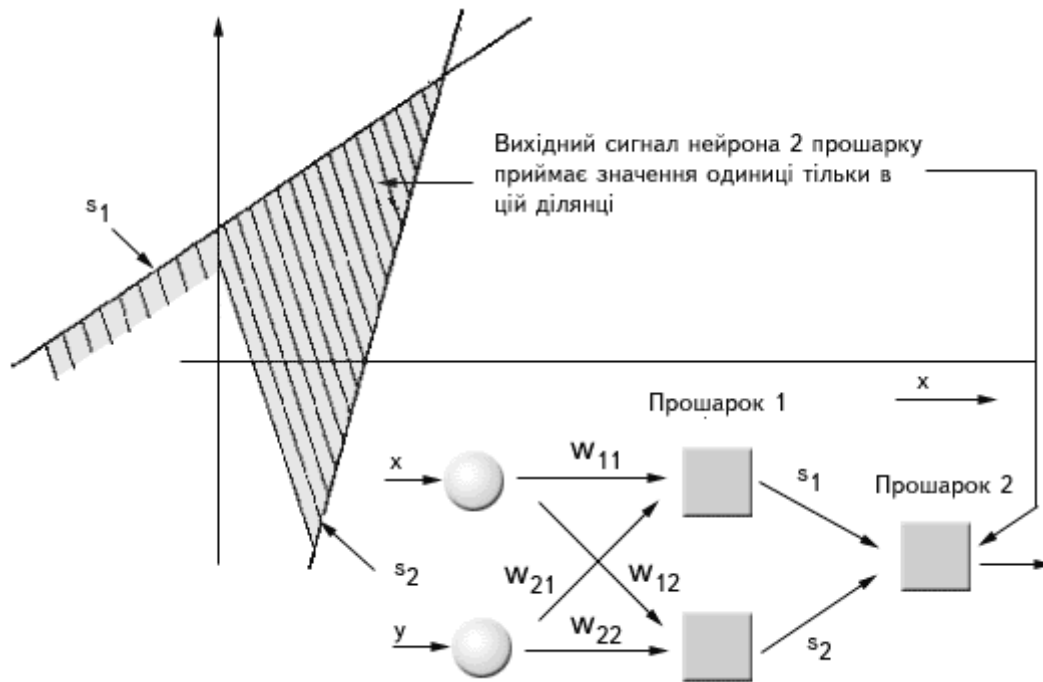


Рис. 8. Опукла область рішень, що задається двошаровою мережею

Нейрон другого прошарку не обмежений функцією I. Він може реалізовувати багато інших функцій при відповідному виборі ваг і порога. Наприклад, можна зробити так, щоб одиничний вихід будь-якого з нейронів першого прошарку призводив до появи одиниці на виході нейрона другого прошарку, реалізувавши тим самим логічне АБО. Є 16 двійкових функцій від двох змінних. Якщо вибирати відповідним чином ваги і поріг, то можна відтворити 14 з них (все, крім ВИКЛЮЧАЮЧЕ АБО і що ВИКЛЮЧАЮЧЕ НІ).

Входи не обов'язково повинні бути двійковими. Вектор безперервних входів може являти собою довільну точку на площині x - y . У цьому випадку ми маємо справу зі здатністю мережі розділяти площину на безперервні області, а не з розділенням дискретних множин точок. Для всіх цих функцій лінійна роздільність показує, що вихід нейрона другого прошарку рівний одиниці тільки в частині площини x - y , обмеженої багатокутною областю. Тому для розділення площин P і Q необхідно, щоб всі P лежали всередині опуклої багатокутної області, що не містить точок Q (або навпаки).

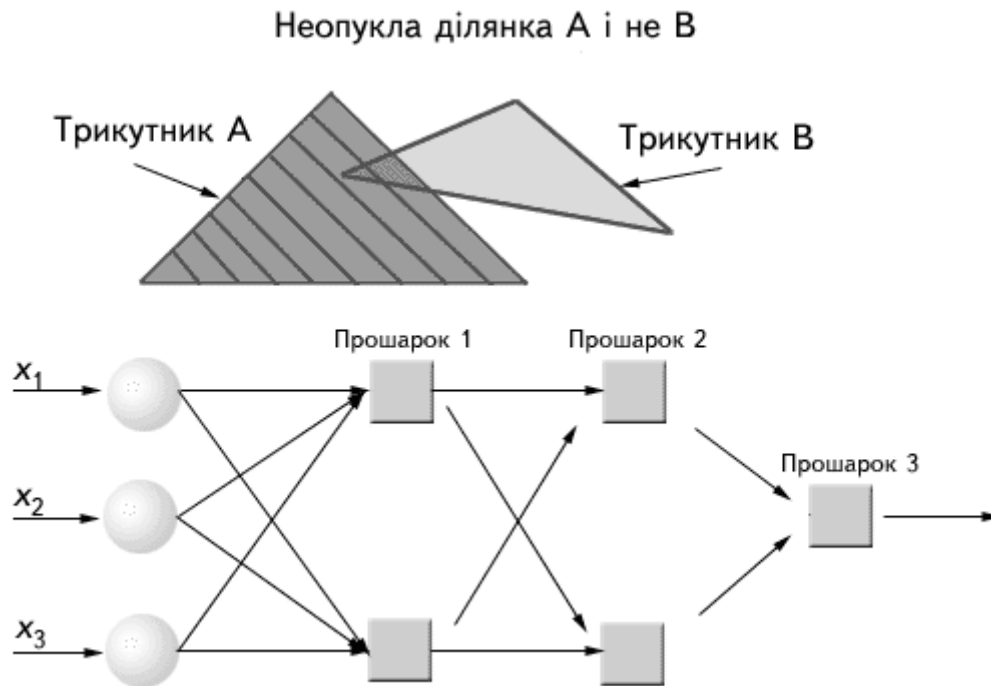


Рис. 9. "Неопукла" область рішень, що задається тришаровою мережею

Тришарова мережа є більш загальною. Її класифікуючі можливості обмежені лише числом штучних нейронів і ваг. Обмеження на опуклість відсутні. Тепер нейрон третього прошарку приймає як вхід набір опуклих багатокутників, і їх логічна комбінація може бути неопуклою. На рис.9 ілюструється випадок, коли два трикутники A і B , скомбіновані за допомогою функцій " $A \cap \bar{B}$ ", задають неопуклу область. При додаванні нейронів і ваг число сторін багатокутників може необмежено зростати. Це дозволяє апроксимувати область будь-якої форми з будь-якою точністю. Додатково не всі вихідні області другого прошарку повинні перетинатися. Можна об'єднувати різні області, опуклі і неопуклі, видаючи на виході одиницю кожен раз, коли вхідний вектор належить одній з них.

Незважаючи на те, що можливості багатшарових мереж були відомі давно, протягом багатьох років не було теоретично обґрунтованого алгоритму для налаштування їх ваг.

Ефективність запам'ятовування

Серйозним питанням є питання ефективності запам'ятовування інформації в перцептроні (або будь-яких інших нейронних мережах) в

порівнянні зі звичайною комп'ютерною пам'яттю і методами пошуку інформації в ній. Наприклад, в комп'ютерній пам'яті можна зберігати всі вхідні образи разом з класифікуючими бітами. Комп'ютер повинен знайти необхідний образ і дати його класифікацію. Для прискорення пошуку могли б бути використані різні добре відомі методи. Якщо точна відповідність не знайдена, то для відповіді може бути використане правило найближчого сусіда.

Число бітів, необхідне для зберігання цієї ж інформації у вагах перцептрону, може бути значно меншим в порівнянні з методом звичайної комп'ютерної пам'яті, якщо образи допускають економний запис. Однак Мінський розробив патологічні приклади, в яких число бітів, необхідних для представлення ваг, росте з розмірністю задачі швидше, ніж експонентне. У цих випадках вимоги до пам'яті із зростанням розмірності задачі швидко стають нездійсненними. Якщо, як він передбачив, ця ситуація не є виключенням, то перцептрони часто можуть бути обмежені тільки малими задачами. Наскільки загальними є такі неподатливі множини образів? Це залишається відкритим питанням, що відноситься до всіх нейронних мереж. Пошуки відповіді надзвичайно важливі для досліджень по нейронних мережах.

Навчання перцептрону

Здатність штучних нейронних мереж навчатися є їх найбільш інтригуючою властивістю. Подібно до біологічних систем, які вони моделюють, нейронні мережі самі моделюють себе внаслідок спроб досягнути кращої моделі поведінки.

Використовуючи критерій лінійної роздільності, можна вирішити, чи здатна одношарова нейронна мережа реалізовувати необхідну функцію. Навіть в тому випадку, коли відповідь позитивна, це принесе малу користь, якщо у нас немає способу знайти потрібні значення для ваг і порогів. Щоб мережа мала практичну цінність, потрібен систематичний метод (алгоритм) для обчислення цих значень. Розенблат зробив це в своєму алгоритмі навчання перцептрону

разом з доказом того, що перцептрон може бути навчений всьому, що він може реалізовувати.

Навчання може бути з вчителем або без нього. Для навчання з вчителем потрібен "зовнішній" вчитель, який оцінював би поведінку системи і керував її подальшими модифікаціями. При навчанні без вчителя мережа шляхом самоорганізації робить необхідні зміни. Навчання перцептрону є навчанням з вчителем.

Алгоритм навчання перцептрону може бути реалізований на цифровому комп'ютері або іншому електронному пристрої, і мережа стає в певному значенні самоорганізуючою. З цієї причини процедуру налаштування ваг переважно називають "навчанням" і кажуть, що мережа "навчається". Доказ Розенблата став основною віхою і дав могутній імпульс дослідженням в цій області. Сьогодні в тій або іншій формі елементи алгоритму навчання перцептрону зустрічаються в багатьох мережевих парадигмах.

Алгоритм навчання перцептрону

Перцептрон навчають, подаючи множину образів по одному на його вхід і налаштовуючи ваги доти, поки для всіх образів не буде досягнутий необхідний вихід. Нехай вхідні образи нанесені на демонстраційні карти. Кожна карта розбита на квадрати і від кожного квадрата на перцептрон подається вхід. Якщо в квадраті є лінія, то від нього подається одиниця, в іншому випадку – нуль. Множина квадратів на карті задає, таким чином, множину нулів і одиниць, яка і подається на входи перцептрону. Мета полягає в тому, щоб навчити перцептрон включати індикатор при подачі на нього множини входів, що задають непарне число, і не включати у разі парного.

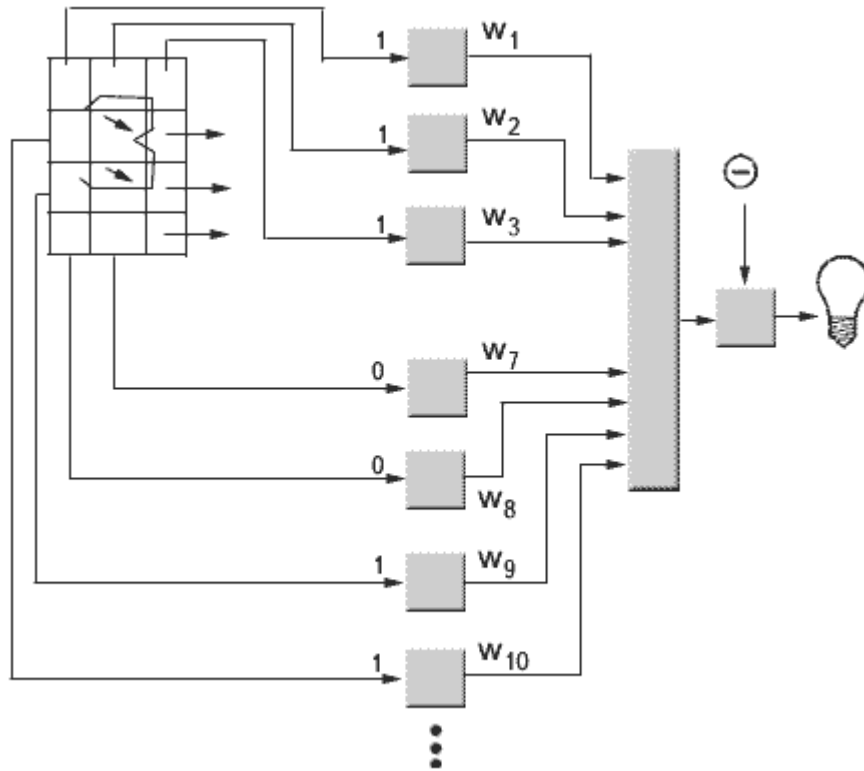


Рис. 10. Перцептронна система розпізнавання зображень

На рис.10 показана перцептронная конфігурація. Припустимо, що вектор X є образом розпізнаваної демонстраційної карти. Кожна компонента (квадрат) $X - (x_1, x_2, \dots, x_n)$ множиться на відповідну компоненту вектора ваг $W - (w_1, w_2, \dots, w_n)$. Ці добутки підсумовуються. Якщо сума перевищує поріг Θ , тоді вихід нейрона Y рівний одиниці (індикатор запалюється), в іншому випадку він – нуль. Ця операція компактно записується у векторній формі як $Y = XW$, а після неї слідує порогова операція.

Для навчання мережі образ X подається на вхід і обчислюється вихід Y . Якщо Y правильний, то нічого не міняється. Однак якщо вихід неправильний, то ваги, приєднані до входів, що підсилюють помилковий результат, модифікуються, щоб зменшити похибку.

Щоб побачити, як це здійснюється, припустимо, що демонстраційна карта з цифрою 3 подана на вхід і вихід Y рівний 1 (показуючи непарність). Оскільки це правильна відповідь, то ваги не змінюються. Якщо, однак, на вхід подається карта з номером 4 і вихід Y рівний одиниці (непарний), то ваги, приєднані до одиничних входів, повинні бути зменшені, оскільки вони прагнуть дати

невірний результат. Аналогічно, якщо карта з номером 3 дає нульовий вихід, то ваги, приєднані до одиничних входів, повинні бути збільшені, щоб скоректувати похибку.

Цей метод навчання може бути підсумований таким чином:

1. Подати вхідний образ і обчислити Y .
2.
 - Якщо вихід правильний, то перейти на крок 1;
 - Якщо вихід неправильний і рівний нулю, то додати всі входи до відповідної ним ваги; або
 - Якщо вихід неправильний і рівний одиниці, то відняти кожний вхід з відповідної йому ваги.
3. Перейти на крок 1.

За кінцеве число кроків мережа навчиться розділяти карти на парні і непарні при умові, що множина цифр лінійно роздільна. Це означає, що для всіх непарних карт вихід буде більшим за поріг, а для всіх парних менше. Зазначимо, що це навчання глобальне, тобто мережа навчається на всій множині карт. Виникає питання про те, як ця множина повинна представлятися, щоб мінімізувати час навчання. Чи повинні елементи множини представлятися послідовно один за одним або карти потрібно вибирати випадково? Нескладна теорія служить тут путівником.

Дельта-правило

Важливе узагальнення алгоритму навчання перцептрону, зване дельта-правилом, переносить цей метод на безперервні входи і виходи. Щоб зрозуміти, як воно було отримане, крок 2 алгоритму навчання перцептрону може бути сформульований в узагальненій формі за допомогою введення величини δ , яка рівна різниці між необхідним або цільовим виходом T і реальним виходом Y

$$\delta = (T - Y). \quad (2.3)$$

Випадок, коли $\delta=0$, відповідає кроку 2а, коли вихід правильний і в мережі нічого не змінюється. Крок 2б відповідає випадку $\delta > 0$, а крок 2в випадку $\delta < 0$.

У будь-якому з цих випадків перцептронний алгоритм навчання зберігається, якщо δ множиться на величину кожного входу x_i і цей добуток додається до відповідної ваги. З метою узагальнення вводиться коефіцієнт "швидкості навчання" η), який множиться на δx_i , що дозволяє управляти середньою величиною зміни ваги.

У алгебраїчній формі запису

$$\Delta_i = \eta \delta x_i, \quad (2.4)$$

$$w(n+1) = w(n) + \Delta_i, \quad (2.5)$$

де Δ_i - корекція, пов'язана з i -м входом x_i ; $w_i(n+1)$ – значення ваги i після корекції; $w_i(n)$ – значення ваги i до корекції.

Дельта-правило модифікує ваги відповідно до необхідного і дійсного значень виходу кожної полярності як для безперервних, так і для бінарних входів і виходів. Ці властивості відкрили множину нових застосувань.

Труднощі з алгоритмом навчання перцептрону

Може виявитися складним завдання визначити, чи виконується умова роздільності для конкретної навчальної множини. Крім того, в багатьох ситуаціях, що зустрічаються на практиці входи часто міняються у часі і можуть бути роздільні в один момент часу і нероздільні в інший. У доказі алгоритму навчання перцептрону нічого не говориться також про те, скільки кроків потрібно для навчання мережі. Мало втішливого в усвідомленні того, що навчання закінчиться за кінцеве число кроків, якщо необхідний для цього час порівнянний з геологічною епохою. Крім того, не доведено, що перцептронний алгоритм навчання більш швидкий в порівнянні з простим перебором всіх можливих значень ваг, і в деяких випадках цей примітивний підхід може виявитися кращим.

На ці питання ніколи не знаходилося задовільної відповіді, вони відносяться до природи навчального матеріалу. У різній формі вони виникають у випадках, де розглядаються інші мережеві парадигми. Відповіді для сучасних мереж як правило не більш задовільні, ніж для перцептронів. Ці проблеми є важливою областю сучасних досліджень.

Кластеризація

У багатьох прикладних задачах вимірювати ступінь подібності об'єктів істотно простіше, ніж формувати опис за ознаками. Наприклад, набагато легше порівняти дві фотографії і сказати, що вони належать одній людині, ніж зрозуміти, на підставі яких ознак вони схожі.

Завдання класифікації об'єктів на основі їх подібності між собою, коли належність навчальних об'єктів до будь-яких класів не задається, називається завданням *кластеризації*. В основному розглядаються статистичні, ієрархічні та графові алгоритми кластеризації. Існують методи багатовимірної шкалювання, що дозволяють відновлювати опис за ознаками об'єктів за матрицею попарних відстаней між ними.

Задача кластеризації (або навчання без вчителя) полягає в наступному. Є навчальна вибірка $X_\ell = \{x_1, \dots, x_\ell\} \subset X$ і функція відстані між об'єктами $\rho(x, x')$. Потрібно розбити вибірку на непересічні підмножини, які називаються *кластерами*, так, щоб кожен кластер складався з об'єктів, близьких за метрикою ρ , а об'єкти різних кластерів істотно відрізнялися. При цьому кожному об'єкту $x_i \in X_\ell$ приписується мітка (номер) кластера y_i . *Алгоритм кластеризації* – це функція $a: X \rightarrow Y$, яка будь-якому об'єкту $x \in X$ ставить у відповідність мітку кластера $y \in Y$.

Множина міток Y в деяких випадках відома заздалегідь, однак частіше ставиться завдання визначити оптимальну кількість кластерів з точки зору того чи іншого критерію якості кластеризації. Рішення задачі кластеризації

принципово неоднозначне, і цьому є кілька причин. По-перше, не існує однозначно найкращого критерію якості кластеризації. Відомий цілий ряд досить розумних критеріїв, а також ряд алгоритмів, які не мають чітко вираженого критерію, але здійснюють досить розумну кластеризації «з побудови». Всі вони можуть давати різні результати. По-друге, число кластерів, як правило, невідоме заздалегідь і встановлюється відповідно до деякого суб'єктивного критерію. По-третє, результат кластеризації істотно залежить від метрики ρ , вибір якої, як правило, також суб'єктивний і визначається експертом.

Кластеризація (навчання без вчителя) відрізняється від класифікації (навчання з учителем) тим, що мітки вихідних об'єктів y_i спочатку не задані, і навіть може бути невідома сама множина Y . У цьому сенсі задача кластеризації ще більшою мірою некоректно поставлена, ніж задача класифікації. Цілі кластеризації можуть бути різними в залежності від особливостей конкретної прикладної задачі:

- Зрозуміти структуру множини об'єктів X_ℓ , розділивши її на групи схожих об'єктів. Спростити подальшу обробку даних і прийняття рішень, працюючи з кожним кластером окремо (стратегія «розділяй і володарюй»).

- Скоротити обсяг збережених даних в разі надвеликої вибірки X_ℓ , залишивши по одному найбільш типовому представникові від кожного кластеру.

- Виділити нетипові об'єкти, які не підходять до жодного з кластерів. Цю задачу називають однокласовою класифікацією, виявленням нетиповості або новизни (novelty detection). У першому випадку кількість кластерів намагаються зробити меншою. У другому випадку важливіше забезпечити високу ступінь подібності об'єктів усередині кожного кластера, а кластерів може бути скільки завгодно. У третьому випадку найбільший інтерес представляють окремі об'єкти, які не вписуються в жоден з кластерів. У всіх цих випадках може застосовуватися ієрархічна кластеризація, коли великі кластери дробляться на більш дрібні, ті в свою чергу дробляться на ще

дрібніші, і т. ін. Такі завдання називаються *завданнями таксономії* (taxonomy). Результатом таксономії є не просте розбиття множини об'єктів на кластери, а деревовидна ієрархічна структура. Замість номера кластера об'єкт характеризується перерахуванням всіх кластерів, яким він належить, від великого до дрібного. Класичним прикладом таксономії на основі подібності є систематизація живих істот, запропонована Карлом Ліннеєм в середині XVIII століття. У сучасному уявленні біологічна ієрархія має близько 30 рівнів, 7 з них вважаються основними: царство, тип, клас, відділ, сімейство, рід, вид. Таксономії будуються у багатьох областях знань, щоб упорядкувати інформацію про велику кількість об'єктів. Саме алгоритми ієрархічної кластеризації дозволяють автоматизувати процес побудови таксономій.

Типи кластерних структур. Формується реєстр різних типів кластерних структур, які можуть виникати в практичних задачах. Різні алгоритми кластеризації можуть бути більш-менш успішні в цих ситуаціях. Прості алгоритми, як правило, вузько спеціалізовані і дають адекватні результати тільки в одній-двох ситуаціях. Більш складні алгоритми, такі як FOREL або агломеративна процедура Ланса-Вільямса, справляються з кількома типами ситуацій. Однак створення алгоритму, що успішно працює в усіх ситуаціях без винятку, представляється важкою і навряд чи можливою для розв'язання задачею.

Зміст

Вступ	
Мета та завдання дисципліни	
Програма навчальної дисципліни	
Структура навчальної дисципліни	
Теми практичних занять	
Теми самостійної та індивідуальної роботи	
	0
Рекомендована література	
	1
<i>Теоретичні відомості:</i>	
Основні поняття теорії розпізнавання образів	
	2

Копча-Горячкіна Галина Ернестівна – старший викладач кафедри інформаційних управляючих систем та технологій факультету інформаційних технологій ДВНЗ «Ужгородський національний університет»

ТЕОРІЯ РОЗПІЗНАВАННЯ ОБРАЗІВ

Навчально-методичний посібник. Частина I для студентів факультету інформаційних технологій напряму „Комп’ютерні науки” та „Програмна інженерія” ДВНЗ «Ужгородський національний університет»

Відповідальний за випуск: **Міца О.В.** – кандидат тех. наук, завідувач кафедри інформаційних управляючих систем та технологій.