1. Основні поняття

Розпізнавання образів — це розділ теорії штучного інтелекту, що вивчає методи класифікації об'єктів. За традицією об'єкт, що піддається класифікації, називається *образом*. 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*, метою якої є розробка методів побудови алгоритмів, що здатні навчатися і видавати правильні відповіді з множини допустимих відповідей.

Перейдемо до опису математичного формалізму, що лежить в основі розпізнавання образів.

Озн. 1.1. *Нехай X — множина об'єктів, Y — множина допустимих відповідей, y^* — цільова функція $y^*: X \to Y$, значення якої відомі лише на скінченій підмножині об'єктів $\{x_1, x_2, ... x_m\} \subset X$. Пари (x_i, y_i) називаються **прецедентами**, а сукупність пар $T = \{(x_i, y_i)\}_{i=1}^m$ — **навчальною вибіркою**.*

Озн. 1.2. ***Задача навчання за прецедентами*** *полягає в тому, щоб за вибіркою T відновити залежність y^*, тобто побудувати **вирішальну функцію** $a: X \to Y$, яка наближає цільову функцію $y^*(x)$ не лише на об'єктах $\{x_1, x_2, ... x_m\}$, а й на всій множині X.*

Озн. 1.3. *Вирішальна функція a повинна допускати ефективну комп'ютерну реалізацію, тому її часто називають **алгоритмом**.*

Озн. 1.4. ***Ознака f об'єкта x*** *— це результат вимірювання деякої характеристики об'єкта. З формальної точки зору ознака — це відображення $f: X \to D_f$, де D_f — множина допустимих значень ознаки.*

Приклад 1.1.

$D_f = \{0,1\}$ — ***бінарна*** *ознака;*

$D_f = \{f_1, f_2, ..., f_n\}$ — ***номінальна*** *ознака;*

$D_f = \{f_{(1)}, f_{(2)}, ..., f_{(n)}\}$ — ***порядкова*** *ознака;*

$D_f = \mathbb{R}$ — ***кількісна*** *ознака.*

Якщо $D_{f_1} = D_{f_2} = ... = D_{f_n}$, *то вхідні дані називають* ***однорідними****, а інакше —* ***різнорідними****.*

Озн. 1.5. *Нехай маємо набір ознак* $f_1, f_2, ..., f_n$. *Вектор* $\left(f_1(x), f_2(x), ..., f_n(x)\right)$ *називається* ***ознаковим описом*** *об'єкта x.*

Постулат про векторну модель — об'єкт x та його ознаковий опис можна ототожнити, тобто

$$X = D_{f_1} \times D_{f_2} \times ... \times D_{f_n}.$$

Озн. 1.6. *Сукупність ознакових описів усіх об'єктів навчальної вибірки T називається* ***матрицею об'єктів–ознак****:*

$$F = \left\| f_j(x_i) \right\|_{m \times n} = \begin{pmatrix} f_1(x_1) & ... & f_n(x_1) \\ ... & ... & ... \\ f_1(x_m) & ... & f_n(x_m) \end{pmatrix}.$$

Якщо $Y = \{1, 2, ..., N\}$ — це **задача класифікації** на N диз'юнктних класів, або задача **розпізнавання образів**.

Якщо $Y = \{0,1\}^N$ — це **задача класифікації** на N класів, що перетинаються.

Якщо $Y = \mathbb{R}$ — це **задача відновлення регресії.**

Якщо вхідні дані характеризують поведінку об'єкта в минулому, а цільова функція — його поведінку у майбутньому, то задача називається **задачею прогнозування**, яка є частинним випадком задач класифікації, або відновлення регресії.

Озн. 1.7. ***Моделлю алгоритмів*** *називається параметричне сімейство відображень* $A = \left\{ g(x, \theta) \middle| \theta \in \Theta \right\}$, *де* $g : X \times \Theta \to Y$ *— деяка фіксована функція.*

Приклад 1.2. Нехай об'єкт описується n числовими ознаками

$$f_i : X \to \mathbb{R},\ j = 1, 2, ..., n.$$

Розглянемо лінійні моделі з вектором параметрів

$$\theta = (\theta_1, \theta_2, ..., \theta_n) \in \Theta = \mathbb{R}^n .$$

$g(x,\theta) = \sum_{j=1}^{n} \theta_j f_j(x), Y = \mathbb{R}$ — задача відновлення регресії,

$g(x,\theta) = sign\sum_{j=1}^{n} \theta_j f_j(x), Y = \{+1;-1\}$ — задача класифікації.

Озн. 1.8. ***Навчання алгоритма а —*** *це процес підбору оптимального параметру* θ *за навчальною вибіркою T.*

Озн. 1.9. ***Метод навчання —*** *це відображення*

$$\mu : (X \times Y)^m \to A,$$

яке довільній скінченій вибірці $T = \{(x_i, y_i)\}_{i=1}^{m}$ *ставить у відповідність деякий алгоритм **а**. Говорять, що метод* μ *будує алгоритм **а** за навчальною вибіркою T.*

Навчання зводиться до пошуку параметрів моделі, які доставляють оптимальне значення функціоналу якості.

Озн. 1.10. ***Функція втрат —*** *це невід'ємна функція* $\mathcal{L}(a,x)$, *яка характеризує величину помилки алгоритму **а** на об'єкті x. Якщо* $\mathcal{L}(a,x) = 0$, *то відповідь a(x) називається правильною.*

Озн. 1.11. ***Функціонал якості алгоритма а на вибірці T*** *має вигляд*

$$Q(a,T) = \frac{1}{m}\sum_{i=1}^{m} \mathcal{L}(a, x_i).$$

*Цей функціонал називають **функціоналом середніх втрат**, або **емпіричним ризиком**, тому що він обчислюється за емпіричними даними* $T = \{(x_i, y_i)\}_{i=1}^{m}$.

Якщо $\mathcal{L}(a,x)$ набуває лише значення 0 або 1, то вона називається ***бінарною***. Тоді якщо $\mathcal{L}(a,x) = 1$, то говорять, що алгоритм ***а*** робить помилку на об'єкті *x*, а функціонал $Q(a,T)$ дорівнює частоті помилок алгоритму ***а*** на об'єкті *x*.

Приклад 1.3. Типові функції втрат:

$\mathcal{L}(a,x)=\left[a(x)\neq y^*(x)\right]$ — індикатор помилки в задачах класифікації; $Q(a,T)$ — частота помилок.

$\mathcal{L}(a,x)=\left|a(x)-y^*(x)\right|$ — відхилення; $Q(a,T)$ — середня помилка алгоритма ***a*** на вибірці *T*.

$\mathcal{L}(a,x)=\left(a(x)-y^*(x)\right)^2$ — квадратична функція втрат; $Q(a,T)$ — середньоквадратична помилка алгоритма ***a*** на вибірці *T*.

Класичний метод навчання — мінімізація емпіричного ризику (empirical risk minimization — ERM) — полягає в тому, щоб знайти в заданій моделі *A* алгоритм ***a***, що доставляє мінімальне значення функціоналу якості *Q* на заданій навчальній виборці *T*.

$$\mu(T)=\arg\min_{a\in A} Q(a,T).$$

Література

1. Воронцов К.В. Машинное обучение. (Курс лекций). ВмиК МГУ: Москва, 2009. http://www.machinelearning.ru/wiki/index.php?title= Машинное_обучение_(курс_лекций%2C_К.В.Воронцов)
2. Дуда Р., Харт П. Распознавание образов и анализ сцен. — М.: Мир, 1976.
3. Фукунага К. Введение в статистическую теорию распознавания образов. — М.: Наука, 1979.
4. Главач В., Шлезингер М.И. Десять лекций по статистическому и структурному распознаванию образов. К.: Наукова думка, 2004. www.irtc.org.ua/image/Files/Schles/esh10_full.pdf.
5. Местецкий Л.М. Математические методы распознавания образов. (Курс лекций). ВмиК МГУ: Москва, 2004). http://www.ccas.ru/frc/papers/mestetskii04course.pdf.
6. Лепский А.Е., Броневич А.Г. Математические методы распознавания образов. (Курс лекций). Южный федеральный университет: Таганрог, 2009. http://www.lepskiy.ucoz.com/lect_Lepskiy_Bronevich_pass.pdf
7. Вапник В.Н., Червоненкис А.Я. Теория распознавания образов М.: Наука, 1974. — 416 с.