

УДК 519.68

А. П. БОНДАРЧУК, канд. техн. наук, доцент,
Державний університет телекомунікацій, Київ

ДОСЛІДЖЕННЯ МЕТОДІВ САМОНАВЧАННЯ СИСТЕМ РОЗПІЗНАВАННЯ

Визначено перспективи розвитку інтелектуальних систем, що можуть навчатись, та проаналізовано сфери, де доцільно використовувати технології з самонавчанням та самоорганізацією. Досліджено методи навчання залежно від типу та кількості апріорної інформації про об'єкт. Запропоновано поділити об'єкти на класи та визначити апріорну базу ознак.

Ключові слова: методи навчання; самонавчання; алгоритм навчання; оптимальне управління; системи розпізнавання; навчальна послідовність; об'єкт інформаційної системи; оптимальне рішення.

Вступ

Сьогодні методи самонавчання систем широко використовуються насамперед у нейронних мережах та мережах із підтриманням концепції Інтернет речей (IoT).

Багато функцій вузлів мережі і деяких підсистем найближчим часом буде спрощено або автоматизовано. Стимулом цього процесу як більш ефективна заміна є застосування штучного інтелекту, що опрацьовує великий обсяг даних.

Нейронна мережа це один із способів реалізації штучного інтелекту. Нейромережа моделює роботу людської нервової системи, особливістю якої є здатність до самонавчання з урахуванням попереднього досвіду, тобто з кожним разом припускаючи все менше помилок.

Як і у разі з більшістю нейронних мереж, мета самонавчання полягає в такому навчанні мережі, щоб досягти балансу між здатністю мережі давати вірний відгук на вхідні дані, що використовувалися в процесі навчання (запам'ятовування), і здатністю видавати правильні результати у відповідь на вхідні дані, схожі, але неідентичні тим, що були використані під час навчання.

Технології, які впроваджуються з IoT, застосовують методи машинного самонавчання і автоматичної обробки даних, що надходять одночасно з кількох сенсорних пристроїв. Це дозволяє здійснювати комплексний аналіз, забезпечуючи адекватне автоматичне реагування відповідно до завдань об'єкта.

Основна частина

Система розпізнавання — складна динамічна система, сформована з технічних засобів отримання та переробки інформації для розв'язання на основі алгоритму задач розпізнавання об'єктів чи явищ.

Використання методів навчання для побудови систем розпізнавання необхідне у разі, коли відсутня повна першочергова апріорна інформація. Задано скінченну множину об'єктів, для котрих відомо до яких класів вони належать. За цих умов така множина називається *вибіркою*. Обсяг інформації

дозволяє поділити об'єкти на класи та визначити апріорну базу ознак. Проте обсяг апріорної інформації недостатній для того, щоб в багатомірному просторі ознак безпосередньою обробкою вихідних даних побудувати поверхні, які розподіляють цей простір на області, відповідні класам, або з достатньою точністю визначити апріорні ймовірності появи об'єктів різноманітних класів $P(K_i)$.

Припустимо, що вихідна апріорна інформація дозволяє скласти список об'єктів із зазначенням, до якого класу кожний із них належить. Позначимо об'єкти цього списку через $\varphi_1, \dots, \varphi_l$, а класи — через K_1, \dots, K_m , тоді вихідну інформацію можна подати у вигляді

$$\left. \begin{aligned} (\varphi_1, \varphi_2, \dots, \varphi_r) &\in K_1; \\ (\varphi_{r+1}, \dots, \varphi_q) &\in K_2; \\ \dots \dots \dots \dots & \\ (\varphi_{g+1}, \varphi_{g+2}, \dots, \varphi_l) &\in K_m. \end{aligned} \right\} \quad (1)$$

Схарактеризуємо кожний об'єкт таким значенням ознак x_1, \dots, x_N . Вважатимемо, що значення ознак у об'єкта φ_1 становлять $(x_1^{(1)}, \dots, x_N^{(1)})$, у об'єкта φ_2 дорівнюють $(x_1^{(2)}, \dots, x_N^{(2)})$ тощо. Тоді вихідний список можна подати такою навчальною послідовністю:

$$\left. \begin{aligned} [(x_1^{(1)}, x_2^{(1)}, \dots, x_N^{(1)}); \dots; (x_1^{(r)}, x_2^{(r)}, \dots, x_N^{(r)})] &\in K_1; \\ \dots \dots \dots \dots \dots \dots \dots \dots \dots \dots \dots \dots & \\ [(x_1^{(g+1)}, x_2^{(g+1)}, \dots, x_N^{(g+1)}); \dots; (x_1^{(l)}, x_2^{(l)}, \dots, x_N^{(l)})] &\in K_m. \end{aligned} \right\} \quad (2)$$

Якщо ознаки x_1, \dots, x_N статистично незалежні, а функції умовних щільностей розподілу можуть мати вигляд $f_i(x_1, \dots, x_N) = f_i(x_1)f_i(x_2)\dots f_i(x_N)$, $i = 1, \dots, m$, і якщо ознаки статистично залежні, проте вигляд функцій розподілу відомий із точністю до значення параметрів, то навчальну послідовність можна застосувати для обчислення в першому наближенні функцій $f_i(x_1, \dots, x_N)$ і $P(K_i)$.

Якщо ознаки статистично залежні і вигляд функції розподілу невідомий, то навчальною послідовністю можна скористатись для визначення в першому наближенні гіперповерхонь, які розподіляють багатомірний простір на області, відповідні класам.

© А. П. Бондарчук, 2017

Обсяг вихідної інформації не дасть змоги достатньо точно описати класи, а знайдені їхні кордони не забезпечать безпомилкового розв'язання задачі розпізнавання. Ось чому для уточнення опису класів використовується поточна апостеріорна інформація. Апостеріорна інформація відрізняється від апіорної інформації наявністю відомостей, отриманих під час проведення експерименту. Вона утворюється як наслідок функціонування сформованої системи в процесі розпізнавання невідомих об'єктів, які не належать до навчальної послідовності.

Якщо навчальна послідовність містить об'єкти, більш-менш рівномірно розташовані в областях простору ознак, відповідних класам, то така процедура зумовлює достатньо точний опис класів і, отже, можливість визначення таких меж класів, дотримуючись яких можна досягнути потенційно можливої точності роботи системи розпізнавання.

Припустимо, що всю множину об'єктів поділено на класи K_1, \dots, K_m та визначено вектор $x = \{x_1, \dots, x_N\}$, а компоненти вектора становлять апіорну базу ознак. Вектор $x = \{x_1, \dots, x_N\}$ є навчальною вибіркою фіксованої довжини. Тобто об'єкти $\varphi_1, \dots, \varphi_l$ вектора, якими вони описуються в ознаковому просторі x_1, \dots, x_l , належать до відповідних класів. Апіорні ймовірності $P(K_i)$ та умовна щільність розподілу $f_i(x_1, \dots, x_n)$ невідомі. З огляду на надання системі розпізнавання об'єктів навчальної вибірки із зазначенням класів, яким вони належать, потрібно в багатовимірному ознаковому просторі побудувати гіперповерхню, яка розподіляє цей простір на області D_i , відповідні класам $K_i, i = 1, \dots, m$. При цьому розподіл має здійснюватись якнайкраще в будь-якому сенсі.

Обмежимося класами K_1 і K_2 , тобто ситуацією, яка має назву *дихотомії* — поділу надвое. До дихотомії можна послідовно звести і загальний випадок, коли кількість класів $m > 2$. Позначимо розподільну функцію через

$$z = f(x, c), \quad (3)$$

де $c = \{c_1, \dots, c_N\}$ — невідомий вектор параметрів.

Розподільна функція, як впливає із (3), подана деякою функцією скалярної похідної векторів x і c . Знаки розподільної функції визначають області в N -вимірному ознаковому просторі D_1 і D_2 , відповідні класам K_1 і K_2 :

$$D_1 = \{x: f(x, c) < 0\}; \quad (4)$$

$$D_2 = \{x: f(x, c) > 0\}. \quad (5)$$

Наявність навчальної вибірки дає змогу отримати вказівки щодо належності об'єктів $\varphi_1, \dots, \varphi_l$ до класу K_1 чи класу K_2 :

$$y = \begin{cases} -1, & \text{якщо } \varphi_d \in K_1; \\ +1, & \text{якщо } \varphi_d \in K_2, \end{cases} \quad d = 1, \dots, l. \quad (6)$$

Ці вказівки можуть бути використані для визначення двох систем нерівностей. Якщо за допо-

могою розподільної функції $f(x, c)$ об'єкт класифікується правильно, то

$$yf(x, c) > 0, \quad (7)$$

а якщо помилково, то

$$yf(x, c) < 0. \quad (8)$$

Як міру відхилення $f(x, c)$ від y візьмемо деяку опуклу функцію від різниці y та z :

$$F(y - z) = F[y - f(x, c)]. \quad (9)$$

Подання об'єктів $\varphi_1, \dots, \varphi_l$ у процесі навчання здійснюється випадково, тому як міру, що характеризує наскільки добре вибрано розподільну функцію, доцільно взяти функціонал, який буде математичним сподіванням міри відхилення:

$$M(c) = W\{F[y - f(x, c)]\}. \quad (10)$$

Припустимо, що найкращий вибір розподільної функції зроблено за умови, коли $M(c)$ досягає мінімуму. У цьому разі задача зводиться до визначення вектора $c = c^0$, який забезпечує

$$\min_c M(c) = \min_c W\{F[y - f(x, c)]\}. \quad (11)$$

Як обмежувальна умова, що накладається на отримання оптимального рішення, розглядається вигляд розподільної функції. Отже, маємо математичну постановку задачі. Стосовно її розв'язку насамперед слід зазначити, що оскільки щільність розподілу невідома, то невідоме і математичне сподівання $W\{F\}$, яке визначається згідно з цією щільністю розподілу. За цих умов розв'язання задачі, тобто визначення вектора $c = c^0$, потребує реалізації двох етапів.

Етап 1. Скориставшись апіорною інформацією, яка міститься в навчальній послідовності, визначимо вектор $c = c^0(z)$ у першому наближенні. Це дасть змогу знайти розподілювану поверхню (у першому наближенні якнайкраще розподіляючи простір ознак на області, відповідні класам) та забезпечити процес розпізнавання нових об'єктів.

Етап 2. На цьому етапі поточна апіорна інформація, здобута внаслідок розпізнавання цих об'єктів, використовується для уточнення значення вектора $c^0(z)$ та отримання такого його значення $c = c^0$, при якому досягається мінімальна ймовірність помилкових розв'язків задачі розпізнавання.

Поділ виконання задачі на два етапи має принциповий характер, проте реалізацію кожного з етапів можна досягнути застосуванням єдиних алгоритмічних методів.

Відсутність повної апіорної інформації вимагає для розв'язання задачі визначення вектора c^0 застосовувати методи навчання, використовуючи на першому етапі апіорну інформацію, яка міститься в навчальній вибірці, а на другому етапі — поточну апостеріорну інформацію. Навчання як на основі апіорної інформації, так і на основі апостеріорної інформації можна реалізувати єдиною рекурентною процедурою. Її мета полягає в тому, щоб на кожному кроці

діставати таке значення вектора s , яке з плином часу прагне до оптимального значення s^0 .

За значенням вектора x алгоритм навчання має уможливити визначення або оцінки вектора $s[n]$ на черговому n -му кроці, якщо вектор s змінюється дискретно, або оцінки вектора $s[t]$, якщо вектор s неперервний і з плином часу прагне до оптимального вектора s^0 .

Висновки

Досліджувані технології дають змогу оптимізувати процеси в сучасних інформаційних системах, а методи самонавчання ідеально підходять для впровадження в мережі з підтримкою IoT. Навчання передбачає адаптацію системи або розвиток автономних систем збирання інформації за допомогою взаємодії елементів мережі. Слід зазначити, що інформаційні системи з навчанням мають бути інтегровані в сучасні мережі з функціями розпізнавання і логічного висновку з використанням стохастичних і статистичних методів для вирішення питань неточності інформації.

Список використаної літератури

1. *Сети следующего поколения* / [А. В. Росляков, М. Ю. Самсонов, И. В. Шибачева и др.; под ред. А. В. Рослякова].— М.: Эко-Трендз, 2008.— 424 с.

2. *Горелик, А. Л. Методы распознавания: учеб. пособие* / А. Л. Горелик, В. А. Скрипкин.— 2-е изд.— М.: Высш. шк., 1984.— 208 с.

3. *Росляков, А. Future Networks. Версия МСЭ-Т. Часть 1.— [Электронный ресурс]* / А. Росляков // ИКС No12.— 2014.— С. 68–70.— Режим доступа:

<http://www.iksmedia.ru/articles/5158228-Future-Networks-Versiya-MSET-Chast.html>.

4. *Михайлов, А. С. Проблемы и перспективы использования искусственных нейронных сетей для идентификации и диагностики технических объектов* / А. С. Михайлов, Б. А. Староверов // Вест. Ивановского гос. энерг. ун-та.— 2013.— № 3.— С. 64–68.

5. *Морозов, В. К. Основы теории информационных сетей* / В. К. Морозов, А. В. Долганов.— М.: Высш. шк., 1987.— 271 с.

6. *Каргаполов, Ю. В. Проектирование систем администрирования и управления процессами идентификации цифровых объектов в конвергентных сетях* / Ю. В. Каргаполов, А. П. Бондарчук // Телекоммуникаційні та інформаційні технології.— 2017.— № 3.— С. 36–43.

7. *Стюарт Рассел. Искусственный интеллект. Современный подход* / Стюарт Рассел, Питер Норвиг.— Издательство-Вильямс 2007.— 1408 с.

Рецензент: доктор техн. наук, доцент **В. В. Онищенко**, Державний університет телекомунікацій, Київ.

А. П. Бондарчук

ИССЛЕДОВАНИЕ МЕТОДОВ САМООБУЧЕНИЯ СИСТЕМ РАСПОЗНАВАНИЯ

Определены перспективы развития интеллектуальных систем, которые являются самообучаемыми, проанализированы те сферы, где целесообразно использовать технологии с самообучением и самоорганизацией. Исследованы методы обучения в зависимости от типа и количества априорной информации об объекте. Предложено разделить объекты на классы и определить априорную базу признаков.

Ключевые слова: методы обучения; самообучения; алгоритм обучения; оптимальное управление; системы распознавания; учебная последовательность; объект информационной системы; оптимальное решение.

A. P. Bondarchuk

STUDY OF SELF-DEVELOPMENT METHODS OF RECOGNITION SYSTEMS

The article outlines prospects for the development of intellectual systems that can study, analyzed areas where it is expedient to use technologies with self-study and self-organization. The methods of training are researched depending on the type and amount of a priori information about the object. It is proposed to divide objects into classes and determine the a priori basis of attributes.

Keywords: methods of teaching; self-learning; learning algorithm; optimal control; recognition systems; learning sequence; object of information system; optimal solution.