

## ФУНКЦІЇ АФФІННОСТІ ДЛЯ ПОРІВНЮВАЄМИХ ДВІЙКОВИХ ВЕКТОРІВ

*д-р техн. наук, проф. В.Д. Дмитрієнко, д-р техн. наук,  
проф. О.Ю. Заковоротний, д-р техн. наук, проф. С.Ю. Леонов,  
Національний технічний університет "Харківський політехнічний  
інститут", м. Харків*

Відстань Хеммінга  $R_X$  між довільними двійковими векторами  $J_p = (j_{p1}, j_{p2}, \dots, j_{pn})$ ,  $J_q = (j_{q1}, j_{q2}, \dots, j_{qn})$  визначається числом двійкових розрядів в яких порівнювані вектора не збігаються [1]. Це відстань часто застосовується для оцінки близькості дискретних об'єктів, які описуються за допомогою двійкових алфавітів з бінарними  $\{0, 1\}$  або біполярними  $\{-1, 1\}$  компонентами. На основі біполярних нейронів розроблена дискретна нейронна мережа Хеммінга [1], яка успішно застосовується для вирішення завдань розпізнавання та оцінки міри близькості двійкових об'єктів, компоненти яких кодуються за допомогою елементів алфавіту  $\{-1, 1\}$ . Однак, оцінка міри близькості порівнюваних об'єктів (двійкових біполярних векторів) здійснюється тільки по найбільш помітним ознакам – відрізняємим двійковим розрядам. При цьому ігноруються більш "тонкі" ознаки для порівнюваних об'єктів, які часто застосовуються для оцінки подібності бінарних векторів за допомогою функцій афінності (подібності) Рассела і Рао, Дайса, Кульчинського, Юла і т.і. [2 – 5]. Крім цього, нейронна мережа Хеммінга може зберігати в пам'яті і використовувати в якості вхідної інформації тільки біполярні вхідні вектора.

При порівнянні об'єктів з якісними ознаками, закодованими за допомогою бінарного алфавіту, для кожної пари об'єктів  $J_p = (j_{p1}, j_{p2}, \dots, j_{pn})$ ,  $J_q = (j_{q1}, j_{q2}, \dots, j_{qn})$  в роботах [2 – 5] використовуються чотири змінні:  $a, b, f, g$  (табл. 1).

Таблиця 1 – Змінні для порівняння двійкових векторів з бінарними компонентами

	$J_p$	
$J_q$	1	0
1	$a = \sum_{k=1}^n j_{qk} j_{pk}$	$g = \sum_{k=1}^n (1 - j_{qk}) j_{pk}$
0	$f = \sum_{k=1}^n (1 - j_{qk}) j_{pk}$	$b = \sum_{k=1}^n (1 - j_{qk})(1 - j_{pk})$

З аналізу табл. 1 випливає, що збільшення змінної  $a$  однозначно вказує на зростання подібності (афінності) порівнюваних векторів. Такої однозначності немає при зростанні змінної  $b$ , оскільки збільшення цієї змінної може вказувати як на збільшення схожості порівнюваних об'єктів, так і на належність їх до різних класів (якщо відсутня схожість по змінній  $a$ ). Що стосується міри близькості векторів  $J_p$  і  $J_q$ , відносно змінних  $f$  і  $g$ , то вона симетрична відносно цих змінних.

Змінні  $a$ ,  $b$ ,  $f$  і  $g$  з табл. 1 використовуються в цілому ряді відомих функцій подібності бінарних об'єктів, у яких наявність або відсутність ознак визначається компонентами бінарного алфавіту. Як приклад можна відзначити функції афінності (подібності) Сокаля і Мішера (Sokal and Michener), Жаккара і Нідмена (Jaccard and Needham) і так далі. За допомогою змінних  $f$  і  $g$  легко визначається і відстань Хеммінга між бінарними векторами  $J_p$  і  $J_q$

$$R_X(J_q, J_p) = f + g. \quad (1)$$

Змінні  $a$ ,  $b$ ,  $f$  і  $g$ , а також функції афінності Сокаля, Мішера, Жаккара, Нідмена, відстань Хеммінга обчислюються за допомогою нової дискретної нейронної мережі на основі бінарних нейронів. Недоліком цієї нейронної мережі є неможливість обробки інформації, представленій за допомогою біполярного алфавіту, а також відсутність можливості обчислення відстані Хеммінга за вищевказаними функціями афінності для бінарних об'єктів.

**Список літератури:** 1. *Дмитриенко В.Д.* Нейронные сети: архитектура, алгоритмы и использование. Учебное пособие / В.Д. Дмитриенко, А.Ю. Заковоротный, С.Ю. Леонов // Харків: НТУ "ХПИ". – 2020.– 222 с. 2. *Бабичев С.А.* Теоретичні та практичні засади інформаційної технології обробки профілів експресій генів для реконструкції генних мереж. Дисертація на здобуття наукового ступеня доктора технічних наук по спеціальності 05.13.06 – інформаційні технології. – Херсон: Херсонський національний технічний університет. – 2018. – 382 с. 3. *Дмитриенко В.Д.* Методы и алгоритмы систем искусственного интеллекта / В.Д. Дмитриенко, И.П. Хавина, А.Ю. Заковоротный, М.В. Личанский, Н.В. Мезенцев. – Киев: Кафедра, 2014. – 282 с. 4. *Фор А.* Восприятие и распознавание образов / А. Фор. – М.: Машиностроение, 1989. – 272 с. 5. *Michalski R.S.* A recent advance in data analysis: clustering jbjects into classes characterized by conjunctive concepts / R.S. Michalski, R.S. Stepp, E. Diday. – Invited chapter in the book Progress in Pattern Recognition. Vol. 1. – North-Holland Publishing Company, Amsterdam-NewYork-Oxford, 1981. – P. 33-49