

СПЕЦІАЛЬНЕ КОДУВАННЯ ДЛЯ СИСТЕМ МАШИННОГО НАВЧАННЯ

Національний авіаційний університет

Gamayun@nau.edu.ua

Вступ

Машинне навчання є необхідною складовою застосування комп'ютерів на сучасному етапі розвитку прикладної інформатики. Великі масиви даних для яких треба виконати аналіз, упорядкування, різні процедури оцінювання та широкий спектр досліджень по підвищенню продуктивності комп'ютерних засобів є одними з головних у Computer Science. Актуальним залишається розв'язання питання: який принцип є найбільш продуктивним для реального часу, за якими положеннями слід будувати структури пристроїв та вузлів спецпроцесорів та процесорів універсального призначення, кластерів та систем, мереж обчислювачів для підтримки систем машинного навчання.

Постановка проблеми

Для вибору напрямку досліджень було вибрано фундаментальне положення про аналіз арифметико-алгоритмічних базисів для побудови комп'ютерних засобів.

Наприклад для цифрової обробки сигналів [1-2] реалізується загальна схема перетворення даних- обчислення великого об'єму даних та виконання альтернативних процедур визначення максимальних та мінімальних значень по масиву даних-результатів. При цьому використовують асоціативно-логічні алгоритми та реалізують індуктивні методи.

Такий комплекс робіт включає

- накопичення масивів даних з початковою обробкою;
- виконання обчислювальних процедур типу швидкого перетворення Фур'є;

- вибір максимальних або мінімальних значень у результатах;
- прийняття рішень по результатах завдання.

Головні витрати часу полягають при виконанні обчислень та вибору альтернативних значень. Вимоги застосування у реальному часі обумовили ряд досліджень [2] в яких отримано позитивний результат при застосуванні розрядно-логарифмічного (РЛ) кодування.

Основні положення розрядно-логарифмічного кодування даних полягають у застосуванні спеціальних кодів для позначення розрядів цифрових операндів, які не є нульовими, тобто значущих розрядів:

Операнд $A = A_{рл} = \{ N_1 N_2 N_3 N_4 \dots, N_Q \}$. Кожен код значущого $N_i = (a_i p_i \neq 0)$ обчислюється як логарифм від кількості, яка визначається цим розрядом: $N_i = \log_2 a_i p_i = i$. Таке кодування є однозначним та виконується без функціонального перетворення.

Основні етапи виконання процедур з навчанням полягають у наступному:

- перетворення у РЛ форму кількох старших розрядів кожного з операндів масиву, що перетворюються у задачі;
- реалізація алгоритмів задачі над отриманими РЛ кодами;
- оцінка та виконання альтернативних процедур;
- перевірка результатів по повним розрядним сіткам;
- розробка моделі по скороченим кодам операндів;
- робота по моделі с періодичною перевіркою.

Основні етапи виконання моделі по 100 каналах цифрової обробки з визначенням найбільших значень наведені у табл. 1. *Таблиця 1.*

Номер каналу перетворення	Значення старших розрядів результату
1	35 34 33 28 26
4	35 34 33 29 27
5	35 34 33 32 29
65	35 34 29 27 26
79	35 34 33 32 31
99	35 34 33 31 29
31	35 33 29 28 24
59	35 33 28 27 24
75	35 33 30 29 27
78	35 33 30 36 24
83	35 33 30 29 28
90	35 33 30 28 27

Відповідно [2] об'єм обчислень (для розв'язання всієї задачі) зменшується приблизно у $2^6 - 2^7$ разів. Аналіз виконувався по критерію значення об'єму обчислень всіх процедур для задачі цифрової обробки. Таким чином, при застосуванні розрядно-логарифмічного кодування для розв'язань з оціночними обчисленнями отримаємо зменшення кроків визначення результату та відповідне підвищення продуктивності.

Для систем машинного навчання властиві оціночні процедури, які були розглянуті у [2, 3] можуть бути виконані на базі розрядно-логарифмічного кодування також- що і є метою досліджень.

Розрядно-логарифмічне кодування має значний потенціал для застосування у різних прикладних задачах. Розглянемо інше представлення двійкових операндів.

Основна форма РЛ операнда може бути представлена як масив у формі набору кодів різниці між сусідньому розрядами (РЛ кодами)

$$A_{рл} = \{ N_1 N_2 N_3 N_4 \dots N_Q \} = \{ N_1 - N_2, N_2 - N_3, N_3 - N_4 \dots, N_{Q-1} - N_Q \}.$$

$$A_{рл} = 35\ 33\ 30\ 28\ 27 = A^*_{рл} = (-2, -3, -2, -1).$$

Відповідно $A^*_{рл}$ має менше значення $Q^* = Q - 1$.

Іншим варіантом може бути представлення як масив у формі набору кодів різниці між першим (старшим) кодом та

іншими кодами масиву- розрядами (РЛ кодами)

$$A_{рл} = \{ N_1 N_2 N_3 N_4 \dots N_Q \} = \{ N_1 - N_2, N_1 - N_3, N_1 - N_4 \dots, N_1 - N_Q \}.$$

$$A_{рл} = 35\ 33\ 30\ 28\ 27 = A^{**}_{рл} = (-2, -5, -7, -8).$$

Відповідно $A^{**}_{рл}$ має менше значення $Q^{**} = Q - 1$.

У табл. 2 представленні нові форми операндів $A^*_{рл}$ та $A^{**}_{рл}$

Таблиця 2.

Номер каналу перетворення	Значення старших розрядів результату
1	35 34 33 28 26 -1-1-5-2 -1-2-7-9
4	35 34 33 29 27 -1-1-4 -2 -1-2-6-8
5	35 34 33 32 29 -1-1-1-3 -1-2-3-6
65	35 34 29 27 26 -1-5-2-1 -1-6-8-9
79	35 34 33 32 31 -1-1-1-1 -1-2-3-4
99	35 34 33 31 29 -1-1-2-2 -1-2-4-6
31	35 33 29 28 24 -2-4-1-4 -2-6-7-11
59	35 33 28 27 24 -2-5-1-3 -2-7-8-11
75	35 33 30 29 27 -2-3-1-3 -2-5-6-8
78	35 33 30 26 24 -2-3-4-2 -2-5-7--8
83	35 33 30 29 28 -2-3-1-1 -2-5-6-7
90	35 33 30 28 27 -2-3-2-1 -2-5-7-8

Кодування з використанням кодів різниці між елементами масивів представлення ще не має досліджень як основи операційного базису, але переваги у процедурах аналізу, порівняння та упорядкування та інших асоціативно-логічних алгоритмах дають значне зменшення кроків розв'язання в цілому.

Моделювання доводить, що використання двох етапної схеми – виконання обчислювальних процедур з використанням оціночних підходів та запропонованого кодування розрядно-логарифмічного представлення різниці кодів значущих одиниць (РЛКРЗО) прискорює розв'язання в комплексі.

Розглянемо табл. 2 як варіант застосування згаданої двоетапної схеми.

Перший рядок по кожному каналу отримано як результат обчислень по найбільш вагомим розрядам [1]. Найбільші значення отриманні по каналах

Номер каналу перетворення	Значення старших розрядів результату
1	35 34 33 28 26 -1-1-5-2 -1-2-7-9
4	35 34 33 29 27 -1-1-4 -2 -1-2-6-8
5	35 34 33 32 29 -1-1-1-3 -1-2-3-6
79	35 34 33 32 31 -1-1-1-1 -1-2-3-4
99	35 34 33 31 29 -1-1-2-2 -1-2-4-6

По 79 (табл. 2) каналу перетворення маємо найбільше значення, бо такі розряди представляють повну розрядну сітку. Такий висновок є результатом порівняльного аналізу з результатами по іншим каналам.

Моделювання показує що продуктивність зростає на один-два порядки при застосуванні РЛКРЗО порівняно з впровадженням РЛ кодування.

Машинне навчання (Machine Learning) – підрозділ штучного інтелекту,

що вивчає методи побудови алгоритмів, здатних навчати. серед методів навчання

- аналітичного;
- синтетичного;
- аналітично-синтетичного;
- індуктивного;
- дедуктивного;
- традуктивного

На сучасному етапі розвитку цієї галузі широко впроваджуються нейронні мережі, від складності яких розрізняють глибинне машинне навчання. Саме запропоновані підходи на базисі розрядно-логарифмічного кодування дозволяють виконувати завдання з високою продуктивністю у актуальній галузі розвитку Computer Science.

Висновки

Запропоновано будувати процедури машинного навчання на базисі розрядно-логарифмічного представлення даних. В масивах цифрових операндів, що підлягають аналізу та упорядкуванню, вибору максимальних або мінімальних значень обчислюється додаткова різниця між кодами представлення значущих одиниць. Таке перетворення забезпечує інформацією, що дозволяє підвищити продуктивність на один-два порядки при виконанні завдання в цілому. Запропоновані підходи розширюють теоретичну базу моделювання та реалізації нейронних мереж для машинного навчання Machine Learning.

Література

1. Гамаюн В.П. Розрядно-логарифмічна арифметика. Методи та алгоритми. Монографія. – К.: Книжкове вид-во НАУ, 2007. – 272 с.

2. Гамаюн В.П., Шалаш Л.А. Высокопроизводительная обработка на основе оценочных вычислений // Нові комп'ютерні засоби, обчислювальні машини та мережі: Зб. наук. пр. Ін-т кібернетики ім. Глушкова НАНУ. – К., 2001. – № 1. – С. 104-108.

3. Гупал А.М., Пашко С.В. Эффективность байесовских процедур распознавания // Кибернетика и системный анализ. – К., 1995. – № 4. – С. 76-89.

Гамаюн В.П., Андрєєв О.В., Андрєєв В.І.

СПЕЦІАЛЬНЕ КОДУВАННЯ ДЛЯ СИСТЕМ МАШИННОГО НАВЧАННЯ

Розглянуто напрямки розвитку комп'ютерних засобів для систем машинного навчання на рівні вдосконалення арифметико-логічного базису, а саме спеціального кодування даних та відповідного операційного базису.

Запропоновано будувати процедури машинного навчання на базисі розрядно-логарифмічного представлення даних. Такий базис має всі властивості системи числення та є розвитком двійкової системи числення. В масивах цифрових операндів, що підлягають аналізу та упорядкуванню, вибору максимальних або мінімальних значень обчислюється додаткова різниця між кодами представлення значущих одиниць. Таке перетворення забезпечує інформацією, що дозволяє підвищити продуктивність на один-два порядки при виконанні завдання в цілому.

Результат роботи моделі показує рівні продуктивності, які можливо досягнути при застосуванні спеціального кодування з орієнтацією на нейроні мережі. Реалізація моделі залишається відкритою, бо дозволяє приєднувати нові методи обробки та нові архітектури для дослідження. Запропоновані підходи розширюють теоретичну базу моделювання та реалізації нейронних мереж для машинного навчання Machine Learning

Ключові слова: розрядно-логарифмічне кодування, нейроні мережі, машинне навчання.

Gamayun V.P., Andreev O.V., Andreev V.I.

SPECIAL CODING FOR SYSTEMS OF MACHINE LEARNING

The direction of development computer structure for system of machine learning on level arithmetic logical base with data of special coding operational system is considered.

The procedure development of machine learning on base of bit-logarithmic system is suggested. This system has all ability of digital system and is improving of bit system. Distance between nonzero bits is resolving. This result is data for increasing productivity for whole result.

The model result defines level of productivity this proposition for neuron system. Model realization is open system. This proposition is improving for neuron system for machine learning.

Keywords: machine learning, neuron system, special coding bit-logarithmic system.