Це дозволить створити ефективніші системи безпеки та контролю для захисту інфраструктури та національної безпеки.

Джерела

1. Seo, Yoojeong, et. al. "Drone detection using convolutional neural networks with acoustic STFT features." 2018 15th IEEE International Conference on Advanced Video and Signal Based Surveillance (AVSS). IEEE, 2018.
2. Samadzadegan F, Dadrass Javan F, Ashtari Mahini F, Gholamshahi M. "Samadzadegan, Farhad, et al. "Detection and recognition of drones based on a deep convolutional neural network using visible imagery." *Aerospace* 9.1 (2022): 31.
3. Anwar, Zohaib M., et. al. "Machine learning inspired sound-based amateur drone detection for public safety applications." IEEE Transactions on Vehicular Technology 68.3 (2019): 2526-2534.
4. Rahman, Md Habibur, et al. "A Comprehensive Survey of Unmanned Aerial Vehicles Detection and Classification Using Machine Learning Approach: Challenges, Solutions, and Future Directions." Remote Sensing 16.5 (2024): 879.

КРИТЕРІЇ ОЦІНЮВАННЯ КВАНТИЗАЦІЇ НЕЙРОННИХ МЕРЕЖ ДЛЯ ЗАДАЧ КОМП'ЮТЕРНОГО ЗОРУ

Ткач.Н.В.

Національний університет "Києво-Могилянська академія"

04655, м. Київ, вулиця Сковороди, 2, НаУКМА, Факультет інформатики. тел.(044) 426 60 64.

E-mail: n.tkach@ukma.edu.ua, факс (044) 426 60 64

The recent advancements in the field of neural network optimization demonstrated a high potential for significant improvements in performance due to the quantization of trained models. However, there is substantial research gap in the development of criteria used for optimization algorithms, which will encapsulate a multi-faceted approach for estimating the quality of quantization methods. This research provides a practical overview of peak error-to-interval ratio as a criterion for several quantization schemes on popular neural network modules. The presented criterion establishes an easily interpretable baseline measure which is particularly helpful in the estimation of the overall performance of neural networks and their sensitivity to imposed degradation errors.

Оптимізація нейронних мереж шляхом квантизації[1] її параметрів є одним з основних сучасних методів застосування нейронних мереж на пристроях з обмеженою обчислювальною потужністю. Попри значні дослідження у сфері алгоритмів квантизації, критерії оцінювання якості квантизації є відкритою темою для нових впроваджень та ґрунтовного аналізу. Причиною цьому є деградація при конвертуванні моделей з вагами в форматі чисел з рухомою точкою(fp32) до цілочисельного формату(uint8, int16). Наявні критерії оцінювання часто не приділяють достатньо уваги до динаміки змін розподілів ваг та активацій при конвертації, а також не вказують на проміжну деградацію шарів нейронної мережі.

Однією з найпоширеніших причин вказаної деградації є розбіжність діапазонів цілочисельних типів та чисел з рухомою точкою, яка зумовлена наявністю числових значень ваг чи передавальної функції які випадають за межі кінцевих інтервалів[2]. В роботі[3] вказані статистичні дані ймовірнісних розподілів ваг та активацій і наголошено на складнощах наближення первісного розподілу внутрішніх параметрів за допомогою класичних розподілів ймовірностей. Для спрощення оцінювання якості

квантизації в даному дослідженні запропонований критерій PEIR запозичений зі обробки сигналів.

$$PEIR = \frac{\left(Max(|X - X'|)\right)}{\left(Max(X) - Min(X)\right)}$$

Де X та X' це вихідні матриці після кожного шару в оригінальній(fp32) та квантизованій моделі(uint8/int16). Для оцінювання заданого критерію було імплементовано асиметричну афінну квантизацію для 8-бітного, та динамічну фіксовану для 16-бітного діапазону. В наявному дослідженні схема квантизації застосована поканально для кожного зі згорткових шарів. Слід зазначити, що для мереж довгої та короткочасної пам'яті(ДКЧП) відповідником каналів є розмірність кожного вузла в блоці, і тому статистика діапазону отримана для кожного вектору вузла окремо. При застосуванні квантизації по каналам ми оцінюємо точку нуля відносно якої ми зміщуємо всі вхідні числа та масштабуємо їх. При 8-бітній та 16-бітній квантизації в дослідженні застосовані лінійне та експоненційне масштабування. Цю оцінку(калібрацію) отримано на основі запусків на репрезентативному датасеті. Щоб не зменшувати загальність дослідження було згенеровано датасет вхідних даних на основі роботи[4].

В описі результатів експериментів для скорочення позначень асиметрична афінна та динамічної фіксована квантизації означені як Q8 та Q16. Експерименти проводили на 100 рандомізованих вхідних матрицях. Змінний вхідний розмір необхідний для оцінки впливу масштабування на кінцеву помилку. Наведемо наступні результати для трьох поширених модулів нейронних мереж, а саме для двовимірної згортки, блоку ДКЧП, а також, двовимірної згортки в глибину. Отримані закономірності можна застосувати для покращення калібрації квантизованих моделей та їх порівняльного аналізу на кінцевих пристроях.

Тип даних	Вхідний розмір	$PEIR_{min}$	$PEIR_{median}$	$PEIR_{max}$	$PEIR_{avg}$
Q8	224x224	0.0029	0.0049	0.1685	0.01344
Q8	224x224	0.0047	0.0053	0.0817	0.086
Q16	512x512	0.0	0.0	0.1187	0.0028
Q16	512x512	0.0	0.0	0.019	0.0009

Таблиця 1. Значення критерію для згорткового двовимірного шару

Тип даних	Вхідний розмір	$PEIR_{min}$	$PEIR_{median}$	$PEIR_{max-}$	$PEIR_{avg}$
Q8	224x224	0.0065	0.0162	0.0444	0.00181
Q8	224x224	0.0616	0.0795	0.1035	0.0797
Q16	512x512	0.0	0.0001	0.0550	0.0027
Q16	512x512	0.0003	0.0003	0.0015	0.0003

Таблиця 2. Значення для блоку ДКЧП

Тип даних	Вхідний розмір	$PEIR_{min}$	$PEIR_{median}$	$PEIR_{max}$	$PEIR_{avg}$
Q8	224x224	0.0002	0.0031	0.2236	0.0194
Q8	224x224	0.0002	0.0039	0.0040	0.003
Q16	512x512	0.0	0.0	0.0957	0.0035
Q16	512x512	0.0	0.0	0.0	0.0

Таблиця 3. Значення критерію для згортки по глибині.

Перевагою даного критерію є його лінійність - він буде збільшуватись пропорційно при зміні глибини нейронної мережі. Крім цього, цей критерій є інтерпретованим, оскільки він зрозуміло пов'язаний з величиною відхилень, які впливають на вихід моделі. З недоліків, варто зазначити нечутливість до "довгих хвостів" розподілів ваг та активацій, через що неможливо в загальному випадку вказати на причину значної похибки. Більш детальний аналіз таких випадків буде проведений в наступних дослідженнях.

ДЖЕРЕЛА

1. Jacob, Benoit, et al. "Quantization and training of neural networks for efficient integer-arithmetic-only inference." *Proceedings of the IEEE conference on computer vision and pattern recognition*. 2018.
2. Nagel, Markus, et al. "A white paper on neural network quantization." *arXiv preprint arXiv:2106.08295* (2021).
3. Lee, Jun Haeng, et al. "Quantization for rapid deployment of deep neural networks." *arXiv preprint arXiv:1810.05488* (2018).
4.Nagel, Markus, et al. "Data-free quantization through weight equalization and bias correction." *Proceedings of the IEEE/CVF International Conference on Computer Vision*. 2019.

ІНСТРУМЕНТАЛЬНІ ЗАСОБИ РОЗРОБКИ ПРОГРАМНИХ СИСТЕМИ УПРАВЛІННЯ ПЕРСОНАЛОМ / INSTRUMENTAL TOOLS FOR THE DEVELOPMENT OF PERSONNEL MANAGEMENT SOFTWARE SYSTEMS

Ткаченко В.А. / Tkachenko V.A.

Київський національний університет імені Тараса Шевченка / Taras Shevchenko National University of Kyiv

01601, Київ, вул. Володимирська, 64/13, E-mail: tkachenkovitaliy@knu.ua

The article discusses software development tools for personnel management systems, with an emphasis on ensuring data security and operational efficiency. Particular attention is paid to the use of qualified electronic signatures (QES) to identify employees and guarantee the legality of operations, especially in the context of increased cybersecurity threats. The role of the Public Key Infrastructure (PKI) in ensuring information security and authentication is described. The article highlights the importance of choosing a platform for application development, especially .NET and ASP.NET Core, to create secure, high-performance web applications. The article also discusses the use of the Blazor framework for developing cross-platform solutions, including Blazor Server, Blazor WebAssembly, and Blazor Web App, which provide various rendering options and development flexibility. All these tools allow you to create reliable, scalable workforce management systems that meet security and efficiency requirements.

Системи управління персоналом є критично важливими для оптимізації процесів управління людськими ресурсами, забезпечення дотримання трудових норм і підвищення ефективності організації. У цифрову еру розробка таких систем вимагає надійних інструментів, які підтримують різноманітні функціональні можливості, включаючи керування даними, безпеку та інтеграцію із зовнішніми службами.

Одним з ключових аспектів розробки системи управління персоналом є забезпечення безпеки даних, особливо в конфіденційних областях, таких як особиста інформація співробітників, і оптимізацію зручності використання як для адміністраторів, так і для кінцевих користувачів. В умовах підвищених загроз дотримання національних стандартів цифрових підписів, таких як кваліфіковані електронні підписи (КЕП), додає додатковий рівень безпеки персональних даних.
КЕП це удосконалений електронний підпис, який створюється з використанням засобу кваліфікованого електронного підпису і базується на кваліфікованому сертифікаті відкритого ключа. Він містять інформацію про власника (працівника), а також криптографічні ключі та дані сертифікаційного центру, який їх видає. Також КЕП має таку саму юридичну силу, як і власноручний підпис.