MATRIX AND COMPUTATION SYSTEM SUPPORTING AI DEVELOPMENT

NABAJYOTI DAS
Research Scholar
MATHEMATICS
Asian International University Imphal, Manipu

ABSTRACT

Artificial intelligence (AI) aims to simulate human-like intelligence in machines, covering perceptual, cognitive, and decision-making abilities. This paper explores the development of AI, focusing on matrix-based algorithms and computational systems supporting AI applications. It discusses the integration of computational intelligence (CI) techniques, such as artificial neural networks (ANN) and evolutionary computation (EC), with deep learning (DL) for real-world applications. Additionally, it highlights the role of numerical linear algebra in extracting meaningful information from data and enabling efficient computations, with a focus on matrix operations. The paper also addresses challenges in dense matrix multiplication and emphasizes the importance of addressing computational complexities and memory constraints. Overall, it underscores the interdisciplinary nature of AI research and its profound impact on various domains, necessitating continued innovation and collaboration to advance AI technologies.

Keywords: *Artificial intelligence, Computational intelligence, Matrix-based algorithms,Numerical linear algebra, Dense matrix multiplication*.

1. INTRODUCTION

Artificial intelligence (simulated intelligence) is the capacity of PCs or different frameworks to copy human intelligence. A machine that can see, reason, learn, plan, conjecture, and participate in other human-like ways of behaving is the sacred goal of artificial intelligence. One of the most striking contrasts among people and different creatures is our degree of intelligence. A developing assortment of machines are bit by bit dislodging people in a wide range of occupations because of the perpetual pattern of modern upsets; the following extraordinary impediment will be to keep machine intelligence from in the end overriding human specialists. The wide and various group of man-made intelligence research is an immediate consequence of the numerous researchers who are giving their professions to the subject. Search calculations, master frameworks, information diagrams, regular language handling, advancement calculations, ML, DL, and a lot more areas of artificial intelligence study are all essential for the field.

Different types of intelligence, including visual, mental, and direction, are engaged with man-made intelligence improvement. At the point when a machine has perceptual intelligence, it implies it can see, hear, feel, and do other fundamental errands that individuals do. An elevated degree of mental intelligence is portrayed by predominant capacities with regards to thinking, enlistment, and information obtaining. The objective is to give PCs human-level thinking and mental capacities, drawing motivation from fields like mental science, neuroscience, and mind like intelligence. It is regularly accepted that machines would have the option to make the best decisions for individuals, manufacturing plants, and the climate once they foster the limit with regards to discernment and perception, similar as people. For choice intelligence to work, information science should be extended using sociology, choice hypothesis, the executives science, and applied information science. The point of perceptual, mental, and dynamic intelligence requires a man-made intelligence foundation layer that is upheld by information, capacity, handling limit, ML calculations, and artificial intelligence structures. From that point onward, it can comprehend the information's innate standards to help and carry out artificial intelligence applications through model preparation. Our work and ways of life are overall essentially impacted by the rising broadness and profundity of computer based intelligence's application layer, which is being converged with fields like key sciences, modern assembling, human existence, social administration, and the internet.

2. MATRIX-BASED ALGORITHMS IN AI

The three fundamental fields of computational intelligence (CI) — artificial brain organizations (ANNs), rationale induction, and developmental calculation — share numerous likenesses with artificial intelligence (simulated intelligence). CI is spurred by both science and language. Truly, man-made intelligence and the essential artificial intelligence research are both covered under the IEEE CI Society (CIS). With regards to artificial intelligence (simulated intelligence), the three grade ways of thinking — Imagery, Connectionism, and Evolutionism — to a great extent reflect the three significant ways of thinking inside CI: rationale, artificial brain organizations (ANN), and transformative calculation (EC). Profound learning (DL) and profound brain organizations (DNN) have made astounding progress in various true applications, including shrewd city improvement, compound union preparation, and discourse acknowledgment, exhibiting the quick headway of CI and simulated intelligence as of late. The outcome of profound learning (DL) and profound brain organizations (DNN) is intently attached to the assets of huge information (BD) and superior execution figuring (HPC), which incredibly grow the abilities of exemplary artificial brain organizations (ANN) to DNN and their applications. Specifically, BD assets consider great DNN preparing, and HPC assets, for example, distributed computing, supercomputing, and realistic handling units (GPUs), accommodate sensible DNN execution times. To tackle complex issues in true applications like industry, government, climate, economy, finance, medication, schooling, the executives, and nature, new CI strategies (like DNN) are created through the reconciliation of Calculations (like ANN), Large Information assets, and Processing assets (like HPC).

Ongoing years have seen quick development in the EC people group because of the far reaching utilization of EC calculations in fields like PC organizations, electromagnetic designing, blockchain applications, versatile figuring, and data security, where customary numerical techniques have demonstrated incapable in tackling improvement issues. In reasonable settings, these accomplishments matter significantly for issues of a little or medium scale. Nowadays, by the by, improvement issues can be huge scope or super-enormous scope since the quantity of decision factors could arrive at many thousands, in the event that not thousands, because of the developing intricacy of useful circumstances. While handling muddled issues, it's feasible to set the populace size very huge to increment variety. Since they are iterative and populace based, EC calculations can cause gigantic computational requests from either enormous scope choice factors or large populaces. Until further notice, this is the principal justification for why EC calculations aren't more famous. Huge Information and Superior Execution Figuring (HPC) are the groundworks of enormous scope FS and DNN, which actually tackle perplexing difficulties.

3. COMPUTATIONAL SYSTEMS FOR MATRIX OPERATIONS

This document does not intend to cover all of the many possible uses for such data. In various fields, matrices are used, although they are most commonly associated with spatial and time-series data analysis. We consider a matrix to be a representation of the data here:

$$A = \begin{bmatrix} — & a_1^T & — \\ & \vdots & \\ — & a_n^T & — \end{bmatrix} \in \mathbb{R}^{n,m}$$

concerning the underlying data, along the dimensions m.n..The feature vectors ai ∈ ℘m are seen as the rows of A in this context, where m is the dimension of the feature space. The amount of data points, or dimension n, could be extremely enormous. Assumption n > m is common. Another possible form that the data can take is a d-order tensor.

$$A \in \mathbb{R}^{n_1 \times n_2 \times \ldots \times n_d}$$

with dimensions $n_i \in \mathbb{N} \quad \forall i.$

The challenges of deriving useful insights from the gathered data differ across domains of application. We frequently come into tasks like supervised learning, unsupervised learning, and semi-supervised learning while we extract knowledge from data. Training data availability and utilization is a good way to explain the differences between supervised and unsupervised learning. In supervised learning, all of the data is available, but in semi-supervised learning, just a fraction of the data is available. A specific example that is fundamental to computer science, statistics, numerical linear algebra, and data science would like to be highlighted before we begin the thorough discussion.

Basically, the goal is to fit a linear model to labelled data, and this is known as the least squares problem. For the sake of argument, let's say we have data points ∈ℝm and usually get a value of either -1 or 1 when solving a classification problem. Next, we want to discover a weight vector ∈ℝm that minimizes the 2-norm of the residual. A linear regression issue would inevitably result from this, as is widely recognized.

$$\min_{w} \|Aw - y\|_2^2 + \lambda \|w\|_2^2$$

same way as before. Also introduced here is the regularization term 2 2's ridge parameter, which can have the value > 0. There are a number of possible norms for measuring the regularization term, including the 1-norm and the total variation norm. Next, we find the answer to this problem by:

$$w_* = \left(A^\top A + \lambda I\right)^{-1} A^T y,$$

an example of a standard linear equation system containing a positive definite matrix that is symmetrical with respect to the value of < 0. All of these fields are interdependent, as this example shows. Data mining that uses numerical linear algebra techniques to make evaluations robust and efficient runs into problems. Studying problems with real data, on the other hand, drives the development of complex numerical approaches. This survey aims to demonstrate how numerical linear algebraic approaches enable calculations to disclose hidden information through information extraction from data, data modeling, and pattern detection. In this way, the paper is organized. The application of conventional factorizations like QR and SDV is demonstrated initially, followed by the introduction of interpretable factorizations like NMF and CUR decomposition. The literature on kernel approaches, particularly the graph Laplacian, is also covered. Up next, we'll talk about randomization and matrices' functions. Finally, we touch on some new findings on deep learning and high-dimensional problems. Before we dive in, we should warn you that numerical linear algebra is a huge subject and that data analysis using its methods is nothing new. We regret that we were unable to include the authors whose results were omitted from this poll in our efforts to highlight current trends. Eldén and Strang have written lovely books that serve as introductions to linear algebra for data science applications. We would like to highlight these works.

4. THE DENSE MATRIX MULTIPLICATION

Numerous computational issues involve dense matrix multiplication. The reduction of its run-time complexity is hindered by its intrinsic data dependencies and arithmetic complexity.

Reducing the execution time of dense matrix multiplication can be achieved in three main ways. 1) Increasing the number of additions and decreasing the number of scalar multiplications, which are both time-consuming. When applied to computers with a single CPU and main memory, these methods performed admirably, and their theoretical arithmetic complexity was excellent. But they were inefficient to use on hierarchical memory systems because of complicated data access patterns. In addition, for matrices of size N × N, recursive matrix multiplication necessitates additional memory, as demonstrated by Cray's implementation of Strassen's technique, which necessitated 2.34 ∏ N2 of extra space. (2) The four types of schedules that can be used for parallelizing matrix multiplication are (i) Broadcast-Compute Shift, (ii) All-Shift-Compute, (iii) Broadcast Compute-Roll, and (iv) Compute-Roll-All, also known as Orbital. The second type is defined by features like consistent and localized data transfer, maximum data reuse without replication, the ability to compute all matrix data simultaneously (retina I/O), etc. (3) A mix of the two methods, where the intricacy of the hardware makes the data transportation anomalies more noticeable. It is worth noting that there appears to be a decline in research on recursive (and recursive-parallel) matrix multiplication. The most recent study that we are aware of dates back to 2006. Keep in mind that sparse matrix algebra primarily aimed to conserve memory by implementing sophisticated data access patterns and indexing structures to store information about non-zero members. But now the fundamental building block is a dense block, and regular data access makes up for the zero multiplications.

5. CONCLUSION

The goal of artificial intelligence (AI) is to imitate human-like abilities in machines through a variety of research domains, such as perceptual, cognitive, and decision-making intelligence. Deep learning (DL) and deep neural networks (DNN) are two examples of how the integration of computational intelligence (CI) approaches, such as artificial neural networks (ANN), logic inference, and evolutionary computation (EC), has significantly advanced AI. Access to big data and high-performance computing (HPC) resources is critical to the success of these methods. Furthermore, although there are computing difficulties in large-scale settings, the use of EC algorithms has demonstrated promise in the resolution of challenging optimization problems. AI development heavily relies on computational systems that support matrix operations for tasks like data analysis, unsupervised, semi-supervised, and supervised learning. Numerical linear algebraic techniques, including randomization, interpretable factorizations, kernel methods, and classical factorizations, are essential for gleaning meaningful information from data and facilitating speedy calculations. However, managing memory limitations, data dependencies, and computational complexity in dense matrix multiplication continues to be a challenge, requiring creative solutions to maximize run-time complexity and resource use.

REFERENCES

[1] D. Silver et al., "A general reinforcement learning algorithm that masters chess, shogi, and go through self-play," Science, vol. 362, no. 6419, pp. 1140–1144, 2018.

[2] D. Silver et al., "Mastering the game of go with deep neural networks and tree search," Nature, vol. 529, no. 7587, pp. 484–489, 2016.

[3] J. C. Bezdek, "(Computational) intelligence: What's in a name?," IEEE Syst., Man, Cybern. Mag., vol. 2, no. 2, pp. 4–14, Apr. 2016.

[4] M. H. S. Segler, M. Preuss, and M. P. Waller, "Planning chemical syntheses with deep neural networks and symbolic AI," Nature, vol. 555, pp. 604–610, 2018.

[5] Q. Chen et al., "A survey on an emerging area: Deep learning for smart city data," IEEE Trans. Emerg. Top. Comput. Intell., vol. 3, no. 5, pp. 392–410, Oct. 2019.

[6] S. M. Nekooei and G. Chen, "Cooperative coevolution design of multilevel fuzzy logic controllers for media access control in wireless body area networks," IEEE Trans. Emerg. Top. Comput. Intell., vol. 4, no. 3, pp. 336–350, Jun. 2020.

[7] Y. Guo, J. Y. Li, and Z. H. Zhan, "Efficient hyperparameter optimization for convolution neural networks in deep learning: A distributed particle swarm optimization approach," Cybern. Syst., to be published, doi: 10.1080/01969722.2020.1827797.

[8] Y. LeCun, Y. Bengio, and G. Hinton, "Deep learning," Nature, vol. 521, no. 7553, pp. 436–444, 2015.

[9] Y. Ong and A. Gupta, "AIR5: Five pillars of artificial intelligence research," IEEE Trans. Emerg. Top. Comput. Intell., vol. 3, no. 5, pp. 411–415, Oct. 2019.

[10] Y. Sun, G. G. Yen, and Y. Zhang, "Evolving unsupervised deep neural networks for learning meaningful representations," IEEE Trans. Evol. Comput., vol. 23, no. 1, pp. 89–103, Feb. 2019.

*****