Are Data Structures and Algorithms obsolete in the Age of Machine Learning

Sidharth Ramesh

***Abstract*— The field of Machine Learning deals with the development of systems that can automatically learn and improve from experiences. Machine Learning & Artificial Intelligence have been rapidly developing over the past decade, having its effect on many diverse aspects of the society. This evolution in computer science technology has led to innovative and unexpected solutions, particularly because the traditional algorithmic approach relies on human understanding of a particular problem, whereas Machine Learning deals with deriving solutions from data. In this paper, we compare and contrast the benefits of Machine Learning approach with the traditional Data Structures and Algorithms approach towards a problem solution.**

I. INTRODUCTION

In Computer Science, an algorithm[1] is any well-defined computational procedure that takes some value, or set of values, as input and produces some value, or set of values, as output. An algorithm is thus a sequence of computational steps that transform the input into the output. A data structure is an arrangement of data in a computer's memory or even disk storage. An example of several common data structures are arrays, linked lists, queues, stacks, binary trees, and hash tables. Algorithms are used to manipulate the data contained in these data structures as in searching and sorting.

However, in the above scenario of algorithmic approach, the computer can be thought of a naïve child – a very fast child nonetheless, but a child who knows only to follow step by step instructions provided to him. The concept of learning remained alien to computers, as humans themselves could not directly tell a computer how to learn something since we could not describe it in an algorithmic manner. This is where Machine Learning[2] showed its significance. Instead of scripting out a list of instructions on how to a particular task, we make use of labelled or classified data to enable the machine figure out the logic to map the input to the output. To take the example of a system that identifies handwritten alphabets, we provide the system with labelled images of handwritten characters Thus, the area of Machine Learning deals with the design of programs that can learn rules from data, adapt to changes, and improve performance with experience.

II. COMPARISON OF MACHINE LEARNING WITH DATA STRUCTURES AND ALGORITHMS

A. Algorithmic approach to developing solution to a problem

An algorithm is a blueprint for solving a problem. There is no one fixed way of developing an algorithm. Some algorithms are very informal in nature, whereas some may have a more mathematical approach.[3] For example, the algorithm to buy a pencil from a store would be very simple, compared to an algorithm sort n numbers.

In any scenario, the general approach to develop an algorithm would involve the following steps:

1) Obtain a description of the problem.
2) Analyze the problem.
3) Develop a high-level algorithm.
4) Refine the algorithm by adding more detail.
5) Review the algorithm.

Algorithms have an associated complexity associated with them – time complexity and compute complexity[4]. For an algorithm to a problem with a given complexity, the time and accuracy of the result remains the same.

Due to this reason, algorithms are more suited for scenarios where human beings have an accurate solution to the problem.[5][6] Knowing the solution enables us to prescribe a finite set of steps that need to be taken to solve it – which is what an algorithm is.

B. Machine Learning Approach to solving a problem

Here, we tackle the problem with a different approach. Consider an example of a Neural Network that is used to detect a handwritten digits[7][8]. This is done using training of the neural network using labelled data. So for a handwritten character recognition, images of handwritten characters categorized and labelled as a digit is used as training data. The Neural Network then learns the relationship between the input - the set of pixels of the captured image and the output – the recognized digit[9][10].

With Machine Learning, the focus is not on coming up with a solution for the problem ourselves, but enabling the system to be able to come up with it.[11]. Machine Learning Theory, also known as Computational Learning Theory[2] aims understand the fundamental principles of learning as a computational process. Since in this approach we don't have to be concerned about devising a sequence of steps that equates to the solution, it is perfectly suitable for scenarios where we don't know how. Examples for this are object recognition, facial recognition where we ourselves know precisely how to do it, but we cannot directly tell a computer how to do it[9][12].

Unlike algorithmic approach, the time and accuracy of the result from a system using machine learning could vary from one system to another – as there is huge dependency on the data used for training.[13] The training data set also causes the system to run into problems of classification bias and

over-fitting[14]. However, the case of unsupervised learning[15][16][17] brings a whole new approach to Machine Learning, wherein the system does not need any human intervention or supervision to learn. However, unsupervised learning[18][19][20] relies heavily on algorithms designed to enable the system to perform the learning process in an unsupervised manner, such as in the cases of k-means algorithm and clustering.

III. CONCLUSIONS

The advent of Machine Learning started a huge revolution in the field of Artificial Intelligence and Robotics, whereby machines gain the capability to learn and improve from feedback data. Though the possibilities of such systems are enormous, they are far from eliminating the need of traditional data structures and algorithms. The core area of application for the two are almost mutually exclusive. Algorithmic approach can be applied in areas where human beings have a clear understanding of the solution to be taken for a problem, whereas Machine Learning can be applied in scenarios where the input to output relationship is oblivious to us. This does not necessarily eliminate the possibility of both approaches working for solving the same problem. In fact, the application of algorithms are now taking a different turn – in implementing Machine Learning training methods. Even in the most autonomous system, we make use of algorithms for its auxiliary subsystems or sub-components. The day when algorithms are completely obsolete for human beings may only come at a time of technological singularity – when machines are self-aware and capable of complete independence in decision making.

REFERENCES

[1] Ronald L Thomas H. Cormen Charles E. Leiserson. *Introduction to Algorithms, Second Edition*.

[2] S. Shalev-Shwartz and S. Ben-David. "Understanding Machine Learning: From Theory to Algorithms". In: *Cambridge University Press* (2014).

[3] A. Ben-Tal and A. Nemirovski. "Lectures on modern convex optimization: analysis, algorithms, and engineering applications". In: *Society for Industrial and Applied Mathematics (SIAM)* (2001).

[4] H. Samet and R.E. Webber. "Hierarchical Data Structures and Algorithms for Computer Graphics, Part I: Fundamentals". In: *CGe.4* (1988).

[5] H. Samet. "The Quadtree and Related Heirarchical Data Structures". In: *ACM Computing Surveys* (1984).

[6] Amir A and Samet H. *Effecient regular data structures and algorithms for location and proximity problems*.

[7] William T. Freeman and Edward H. Adelson. "The design and use of steerable filters". In: *IEEE Transactions on Pattern Analysis and Machine Intelligence* (1991).

[8] H. Greenspan and C. Anderson. "Overcomplete steerable pyramid filters and rotation invariance". In: *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition* (1994).

[9] K. Sung and T. Poggio. "Example-based learning for viewbased face detection". In: *IEEE Patt. Anal. Mach. Intell.* (1998).

[10] A. Kassim P. Dias and V. Srinivasan. "A neural network based corner detection method". In: *IEEE International Conference on Neural Networks* (1995).

[11] Lee H and Ranganath R. "Convolutional deep belief networks for scalable unsupervised learning of hierarchical representations". In: *ICML* (2009).

[12] Pinto N and J. J DiCarlo. "Why is real-world visual object recognition hard?" In: *PLoS Computational Biology* (2008).

[13] A. Krizhevsky. "Learning multiple layers of features from tiny images". In: *Technical report, University of Toronto* (2009).

[14] J.M Buhmann. "Data clustering and learning". In: *the Hand-book of Brain Theory and Neural Networks, M. A. Arbib, Ed., MIT Press, Boston* (1995).

[15] D Silver. "Mastering the game of Go without human knowledge". In: *Nature 550* (2017).

[16] B. Mulgrew S. Chen and P. M. Grant. "A clustering technique for digital communications channel equalization using radial basis function networks". In: *IEEE Trans. Neural Networks* (1993).

[17] Ang C.H and Shaffer C. A. "A new region expansion for quadtrees". In: *IEEE Transactions on Pattern Analysis and Machine Intelligence* (1990).

[18] A. Coates and H Lee. "An analysis of singlelayer networks in unsupervised feature learning". In: *AISTATS* (2011).

[19] Osindero S Hinton G. E and Teh Y. W. "A fast learning algorithm for deep belief nets". In: *Neural Computation* (2006).

[20] Raina R and Lee H. "Self-taught learning: Transfer learning from unlabelled data". In: *ICML* (2007).