Some Note on Artificial Intelligence

Anna Esposito[1], Marcos Faundez-Zanuy[2], Francesco Carlo Morabito[3], Eros Pasero[4]

[1]Università degli Studi della Campania "Luigi Vanvitelli, Dip.di Psicologia and IIASS, Italy
[2]Pompeu Fabra University, Spain
[3]Università degli Studi "Mediterranea" di Reggio Calabria, Italy
[4]Politecnico di Torino - Dip. Elettronica e Telecomunicazioni, Italy
{iiass.annaesp@tin.it; faundez@tecnocampus.cat; morabito@unirc.it; eros.pasero@polito.it}

Abstract. This introductory chapter discusses about some basics of artificial intelligence, and in particular on some theoretical issues concerning neural networks that are still open problems and need further investigations. This is because, independently of the large use of neural networks in several fields, using neural networks and similar artificial intelligence techniques to solve problems of non polynomial complexity is by itself a creative and intelligent problem, not rigidly tied to procedural methods and fully explainable on theoretical bases.

Keywords: Artificial Intelligence, Customer Care, Daily Life Activities, Biometric Data, Social Signal Processing, Social Behavior and Context; Complex Human-Computer Interfaces.

1. Introduction

The term Artificial Intelligence was coined by McCarthy et al. [6] in the attempt to collect funds to organize a series of seminars. In the presentation of the project McCarthy et al. wrote: "*We propose .. a ….study of Artificial Intelligence on the basis of the conjecture that every aspect of learning or any other feature of intelligence can in principle be so precisely described that a machine can be made to simulate it*" [6]. Artificial Intelligence, therefore, as it was initially defined, aimed to identify machine learning algorithms and procedures that could complement and even replace the man in the execution of tasks requiring the intervention of "intelligence". However, establishing that certain tasks requires some "intelligence" has proved to be a difficult issue, since a comprehensive understanding of the concept of "intelligence" has not yet been formulated and may remain an "ill posed problem" for the time to come. There are, however, skills or abilities which, although do not fully provide an all-inclusive rendering of the concept of intelligence, yet can be considered as important components of it. We may agree with Rhine [7] that, for example, some essential features of intelligence are:

- react flexibly to various situations:
- take advantage of fortuitous circumstances;
- recognize the relative importance of different elements in a situation;
- find similarities in situations despite of their differences;

- synthesize new concepts connecting old ones in new ways;
- produce new ideas
- …more

Nevertheless, when such features are going to be incorporate into a machine algorithm or an automatic procedure one comes across an apparent paradox. Machines, and in particular computers, adopt intrinsically rigid behaviors. However fast they may be, they represent the very essence of unawareness and therefore, how can they plan intelligent behaviors? The basic belief in artificial intelligence is that there is no contradiction in the idea of programming a computer to be intelligent, i.e., it is, in principle, perfectly conceivable, to teach a machine through precise instructions to be flexible and creative.
The fundamental characteristic that distinguishes artificial intelligence from other disciplines - such as psychology, philosophy, linguistics, logic - which investigate aspects and forms of intelligent behaviors, lies in the methodological approach, best defined as " computational approach". The computational approach requires that hypotheses and theories constituting artificial intelligence (henceforth AI) must be expressed in the form of efficient computational procedures; i.e., the description of such procedures must be so explicit and articulated that they can be translated into a computer program. The two basic concepts to which AI refers are learning and adaptation and, as previously stated, AI tries to formalize these concepts in rigorous mathematical models. A category of learning that has been easily formalized in a rigorous mathematical model is “learning from examples” [8]. Once procedures to automatically learn from examples have been formally defined, AI proposes to identify a system that can implement such learning behaviors. Among the several systems proposed by the literature, neural networks have proved to be the most successful and particularly efficient in learning by examples. Neural networks are composed of many non-linear computational elements that operate in parallel. Instead of executing a set of elementary instructions, neural networks simultaneously explore different hypotheses by exploiting the implicit parallelism present in their architecture which consists of many simple elements connected through weighted connections. The computational elements of a neural network typically perform nonlinear functions, the simplest of which is a weighted sum of the filtered input through a non-linear function called “the activation function”. A neural network is completely specified by;

a) the architecture or network topology;
b) the activation functions of its nodes in the different layers;
c) the type of learning algorithm that the network uses (supervised or unsupervised)

There are several theoretical aspects that must be accounted for the correct functioning of a neural network, the most important are:

a) Avoid the overfitting problem which is tied to the dimensioning of the network architecture and complexity of the task assigned to the net. Overfitting give rise to an undersized or oversized network with respect to the number of needed nodes for the task at the hand and affect the network performances. To overcome overfitting it is necessary not only to understand how many nodes and how many layers a network must have in order to solve the assigned task, but also ensure that the activation functions of the nodes in the various layers possess certain properties. For example, while the functions of the first layer may be strictly monotonous, those of the second layer should be chosen depending on the complexity of the task. Overfitting produces effects similar to those observed when polynomials with a degree higher than the number of available points are used to interpolate a function. The network recognizes exactly the examples seen. Nevertheless, for those not seen it responds with a behavior far from the current data distribution. This involves the loss of the network's ability to generalize, which represents one of the most important properties making neural networks appealing to solve problems of non polynomial computational complexity. How to size the network architecture is a problem still handled through trial and error processes, advocating the need of further theoretical investigations.

b) Avoid the overtraining problem. When the set of training data is presented to the network for a large number of epochs, the error committed by the network on such data (training error) decreases. However, the decrease of the training error does not imply that in turn the generalization error decreases. The network tends to specialize on the training set, losing the generalization objective. Heuristics have been suggested to avoid the overtraining problem, consisting in introducing a further set of data, the validation set, which act as a testing set for deciding when to stop the training. However, when fewer data are available this become a problem. Is there any other theoretical approach helping to avoid overtraining? More investigations are needed.

c) The input / output data encoding. Very often the input to a neural network is not in a suitable form and it is necessary to work on it with transformations. Consider, for example, a voice or a video signal. Even quantized and sampled it is practically unconceivable to consider giving in input to the network these signals in their coarse digital representation. Similarly, the output of a neural network does not always come in the form of a direct response to the problem under consideration, and may require an interpretation. Consequently, pre-processing and post-processing the network' inputs and outputs plays a fundamental role for the network's accuracy.
With deep learning algorithms [1, 3, 4, 5] it seems that the preprocessing problem has been solved. However, the advantage of using deep learning is that networks can handle large amounts of input data. More huge are these data, more work for annotating them is required, and less benefits are gained because the time required to annotate large amount of data, as for example satellites or geophysical data, reduces the advantage of eliminating the preprocessing phase. In addition, deep learning algorithms do not satisfy the European Union guidelines on developing ethical AI systems: "*AI systems should be accountable, explainable, and unbiased*" since it remains still unclear how data are processed along the deep layering of the corresponding networks' architectures [9]. More theoretical investigations are required in the future.

d) The convergence of a learning algorithm is one of the most important aspects to ensure that the network can achieve good performances in solving a given problem. There are no theoretical bases that can always guarantee the convergence, and heuristics are exploited depending on both the available computational resources and values used to initialize the network's parameters. For example the value of the learning rate must not be very small otherwise learning (and therefore convergence) is slow. However, the learning rate cannot even take very large values otherwise the network does not converge to the required solution. Theoretical procedures to infer the optimal value of the learning rate do not exist. Nevertheless, it has experimentally observed that the optimal initial learning rate value depends on dimensions and architecture of the network, learning algorithm, size of the training data, number of epochs for which the net is trained, and error committed to each training cycle. Another parameter that plays an important role in ensuring the network convergence is the way weights are initialized on the connections between one node and another. Also in this case it is necessary to take into account the architecture of the network, the learning algorithm, distinctions between recurring and non-recurring network models, and more. Studies providing accurate theoretical bases for determining the correct learning rate value or the correct way to initialize weights should be taken into consideration in the future.

2 Content of this book

The themes of this book tackled aspects of dynamics of signal exchanges either processed by natural or artificial neural networks. The volume is organized in sections, one dealing with very general neural network applications and the remaining with more specific issues associate to social exchanges and pattern recognition procedures in medicine. The contributions collected in this book were initially discussed at the International Workshop on Neural Networks (WIRN 2018) held in Vietri sul Mare, Italy, from the 13th to the 15th of June 2018. The workshop, being at its 29th edition is nowadays a historical and traditional scientific event gathering together researcher on artificial intelligence and human-machine interaction from Europe and overseas.

Section I is introductory and contain this paper [2].

Section II is dedicated to neural networks and their related applications in several research fields. It include 21 original contributes.

Section III describes the use of neural networks and pattern recognition techniques in medicine. This section includes 7 chapters discussing on advanced artificial intelligent methods for supporting health care services.

Section IV discusses themes multidisciplinary in nature, and closely connected in their final aims to identifying features from realistic dynamics of signal exchanges. Such dynamics characterize formal and informal social signals, communication modes, hearing and vision processes, and brain functionalities. The section includes 17 chapters.

3 Conclusions

Taking all of the above discussion in mind, we can conclude that using neural networks and similar artificial intelligence techniques to solve problems of non polynomial computational complexity is by itself a creative and intelligent problem, not rigidly tied to procedural methods. We expect that contributors of this book which use intelligence (natural or artificial) to solve daily problems would consider in the future investigating more on the highlighted theoretical issues.

ACKNOWLEDGEMENTS

European Commission

The research leading to these results has received funding from the European Union Horizon 2020 research and innovation programme under grant agreement N. 769872 (EMPATHIC) and N. 823907 (MENHIR) and from the project SIROBOTICS that received funding from Ministero dell'Istruzione, dell'Università, e della Ricerca (MIUR), PNR 2015-2020, Decreto Direttoriale 1735 del 13 luglio 2017.

References

1. Bengio, Y., Thibodeau-Laufer, E., Alain, G. & Yosinski, J. (2014). Deep generative stochastic networks trainable by backprop. In Proc. 31st International Conference on Machine Learning 226–234
2. Esposito A, Faundez-Zanuy M, Morabito FC, Pasero Eros (2019) Some note on artificial intelligence. This volume.
3. Graves, A, Mohamed, AR & Hinton, G. (2013) Speech recognition with deep recurrent neural networks. In Proc. International Conference on Acoustics, Speech and Signal Processing 6645–6649
4. Hinton, G. et al. Deep neural networks for acoustic modeling in speech recognition. IEEE Signal Processing Magazine 29, 82–97 (2012).
5. Krizhevsky, A, Sutskever, I. & Hinton, G. (2012). ImageNet classification with deep convolutional neural networks. In Proc. Advances in Neural Information Processing Systems 25 1090–1098
6. McCarthy, J., Minsky, M.L., Rochester, N., Shannon, C.E. (2006) A proposal for the Dartmouth summer research project on artificial intelligence. AI Mag. 27(4), 12–14
7. Rhines, W. (1985) AI: Out of the Lab and Into Business. Journal of Business Strategy, Summer 1985, 6 (1), 50-57
8. Rumelhart, D. E., Hinton, G. E. & Williams, R. J. (1986). Learning representations by back-propagating errors. Nature 323, 533–536
9. Vincent J (2019) https://www.theverge.com/2019/4/8/18300149/eu-artificial-intelligence-ai-ethical-guidelines-recommendations