Intelligent machines for good? More focus on the context

Giovanni Delnevo, Marco Roccetti, Silvia Mirri
Department of Computer Science and Engineering
University of Bologna
Bologna, Italy
giovanni.delnevo2@unibo.it, marco.roccetti@unibo.it, silvia.mirri@unibo.it

ABSTRACT

Machine learning and modern Artificial Intelligence (AI) systems are influencing several aspects of our human lives. Many of these algorithms, based on Artificial Neural Networks (ANNs), have been empowered to make decisions and take actions, based on the well-known notions of efficiency and speed. The aura of objectivity and infallibility of such algorithms, nonetheless, have been already put into question (e.g., refer to the debate about the recent tragic car crashes that have involved self-driving cars). In this setting, our intuition identifies a key issue around the problem of AI errors and bias into the insufficient or inaccurate (human) activity of comprehension and codification of the context where the ANNs will have to operate. We present here a simple cognification ANN-based case study, in an underwater scenario, where we recovered from a situation of partial failure, by including additional contextual factors that were initially disregarded. Our final reflection is that a *nuanced* consideration of a complex context, and subsequent technical actions, should be always kept in mind before an AI-based system takes its final shape. Because machines have still no context for what they are doing, it is a human duty and responsibility to codify it.

CCS CONCEPTS

• Computing methodologies~Artificial intelligence • Computing methodologies~Machine learning

KEYWORDS

Artificial intelligence, Machine learning, Neural networks, Context formalization

ACM Reference format:

Giovanni Delnevo, Marco Roccetti, and Silvia Mirri. 2018. Intelligent machines for good? More focus on the context. In *Proceedings of ACM Goodtechs Conference (GOODTECHS'18). ACM, Bologna, Italy, 5 pages.* https://doi.org/10.1145/3284869.3284875

Goodtechs '18, November 28–30, 2018, Bologna, Italy

ACM ISBN 978-1-4503-6581-9/18/11…$15.00
https://doi.org/10.1145/3284869.3284875

1 Introduction

Artificial Intelligence (AI) has become a part of our daily lives till the point of being now present in many consumer technologies, such as smartphones, tablets, cars, and computers [1]. This is due to the performances of Artificial Neural Networks (ANNs) that are reaching and often exceeding the human ones on an increasing number of complex, yet specialized, tasks, including text translation, audio/image recognition, and strategic games for example [2-15]. Their power stems from an innate ability of crawling huge database of examples (both positive and negative), learning a behaviour from them. Unfortunately, all this is getting farther and farther from human agency. In some sense, out of the human control. If we, the designers, wish they become a *good piece of technology*, we have to multiply our effort to make the decisions taken by those systems acceptable by humans. In few words, the debate on the implications of the use of AI for people and, in general, for the society is still open. It becomes even more controversial in contexts where automatized decisions on crucial factors can have a deadly impact, like healing a patient or driving a car. Many conflicting positions are emerging. On one side, Mark Zuckerberg considers this AI as an unprecedented mean to eradicate hitherto indestructible problems that have afflicted humanity for a long time; on the other side, Elon Musk, Bill Gates and Stephen Hawking emphasize the risk that humanity is running with a high probability of becoming second-class citizens in the world. After all these premises, our modest vision is that of a people-centric AI, where humans help improve machine learning algorithms, based on a human-in-the-loop machine learning paradigm that is emerging in the HCI community. Following this line, our intuition identifies a key issue around the problem of AI errors and bias into the insufficient and inaccurate (human) activity of comprehension and codification of the context where the ANNs will have to operate [16–20]. Disregarding the context means ignoring that grey area where often human intelligence gives its best. Underestimating the role of the context is equivalent to misunderstanding the fact that often machines should be designed not to emulate humans, rather to collaborate with them. In simple words, and translating this into the machine learning world, ignoring the context can lead to a wrong training activity of ANNs and, hence, to dangerous results. Obviously, we have a typical big problem in designing and training a simple and efficient ANN able to deal with large datasets containing hundreds of thousands of pieces of information, while managing an intelligible and accurate formalization of the context where it has to operate.

In this paper, we present our experience in the cognification of an underwater scenario where ANNs were trained to find optimal routes, like for example in the case when an underwater fiber optic cable has to be installed [21]. With this running example, we demonstrate that a nuanced consideration of a complex context has to be always kept in mind before an AI system takes its final shape.

The reminder of the paper is structured as follows. Section 2 illustrates an initial ANN training scenario where many crucial factors of the context were ignored with consequent negative results. In Section 3, instead, we show how the introduction of those factors ameliorates the final results. Finally, Section 4 concludes the paper.

2 The context matters, (un)fortunately

We trained an ANN with real data coming from an underwater fiber optic cable installation case. The situation in the reality is as follows. Before a vessel installs a cable, an a-priori in-depth analysis of the seabed takes place. A vessel inspects the seabed, obtaining a grid of sampled points of which the depth and the soil typology are the main information. Using these data, a team of geologists define the optimal route. This route is the final output of the analysis process and it is provided to the crew of the vessel that will then install the cable. In this context, we trained an ANN to find an optimal route of a given African seabed. Essentially, the ANN, at each given point along the route, learns the following one, based on information concerning the soil type of each surrounding point (neighbors). Essentially, our ANN learns how to move on a square grid, like that in Table 1. No other contextual information was formalized and passed to the ANN to be learnt.

5	6	7
4	x,y	0
3	2	1

Table 1. Possible movements

2.1 When the results are unreliable

We then run our ANN and contrasted the obtained results against those provided by a team of geologists who were our advisors on this case. After several experiments, we found evidence that the routes computed by the ANN deviated from the trajectories suggested by geologists, going outside of the inspected seabed corridor, without even reaching the destination, as emphasized, for example, in Figure 1.

With the aim of ameliorating this situation, without altering the training process, we simply added some further rules to be executed after the ANN had run. In essence, we developed a backtracking mechanism. If the ANN goes out of the inspected seabed corridor, we remove the point from the route and we let the ANN re-compute an alternative path, trying to identify an admissible route.

With this a-posteriori modification, our ANN becomes always able to suggest a plausible route, from the initial point to the destination, as shown in Figure 2. Nonetheless, this good result comes with the drawback that these computed routes can be often much longer than those proposed by geologists. In Table 2, third column, we have reported some instances of the route length ratio (ANN / geologists).

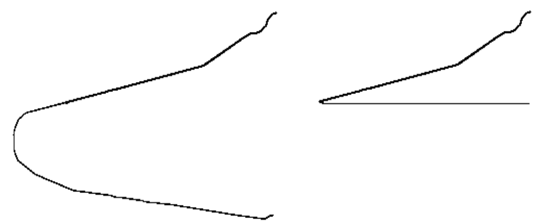

Figure 1. Route comparison. Geologists (leftmost) vs. Neural network (rightmost)

Figure 2. Route comparison II. Geologists (leftmost) vs. Neural network (rightmost)

Instance	Out	Ratio
1	No	4.98
2	No	4.53
3	No	5.71

Table 2. Route length ratio (ANN / geologists)

3 Sharpen the focus on the context

The discrepancy between the routes computed by our ANNs and those provided by the geologists, suggests us that the codification of context, so far, did not contain all the contextual factors relevant to the case (nature and characteristics of the cable, for example). Hence, we sharpened our focus on the context, looking for relevant factors, previously not considered. We identified the following four ones:

1. The direction along which a cable is pulled is relevant (e.g., pulling a cable from top to bottom or vice versa is different).

2. We cannot make a 180-degree bend in a cable, otherwise it breaks.
3. We cannot make a 90-degree bend in a cable, otherwise it breaks.
4. The cable cannot be installed outside of the inspected seabed corridor.

3.1 Formalizing (and learning) more rules

We tried to change our model of the context, including the factors previously mentioned. First of all, we focused on the direction along which the cable is pulled. The problem with our previous model emerges from the following example. Consider the following factors in Table 1. A direction from bottom to top. Rocks (dangerous soil) in point 5 and 6, sand (good soil) in point 7. The right decision to make is moving from the center of the grid to point 7. Now consider a specular situation when only the direction of the movement is changed (from top to bottom). All the rest remains unaltered. The right decision, in this case, is move to point 1. This explains simply that an explicit codification of this contextual information is needed to allow our ANN to be trained correctly. This new codification is reported in Table 3 with all the possible cases, based on the direction to follow.

5	6	7
4	↓	0
3	2	1

4	5	6
3	↙	7
2	1	0

3	4	5
2	←	6
1	0	7

2	3	4
1	↖	5
0	7	6

1	2	3
0	↑	4
7	6	5

0	1	2
7	↗	3
6	5	4

7	0	1
6	→	2
5	4	3

6	7	0
5	↘	1
4	3	2

Table 3. Possible movements: directions

Go to problem 2 now. We can not allow our ANN to go back and forth. This can be simply formalized, and then passed to our ANN to be learnt, by making not admissible those precise movements that make our ANN go backward, as per Table 4.

Think now of point 3. The cable can not make a 90-degree bend, otherwise it breaks. This can be simply formalized, and then passed to our ANN to be learnt, by making not admissible those precise movements that make a 90-degree bend. All these contextual factors, concerning the cable, can be hence formalized as per the new and final revised Table 5.

After the formalization of points 1, 2 and 3, we trained again our ANN and run it to check its behavior. Consider now the results we obtained. Figure 3 shows how many times our ANN took the (right) decision to go to points 1, 2 and 3, while ignoring all the others. The correct behavior of our ANN is confirmed by the value of 0.82, 0.79, 0.97, respectively associated to the movements towards points 1, 2 and 3, of the area under the Receiver Operating Characteristic (ROC) curves. Even the average value of those figures of merits confirmed the efficacy of our context formalization.

X	X	X
4	↓	0
3	2	1

4	X	X
3	↙	X
2	1	0

3	4	X
2	←	X
1	0	X

2	3	4
1	↖	X
0	X	X

1	2	3
0	↑	4
X	X	X

0	1	2
X	↗	3
X	X	4

X	0	1
X	→	2
X	4	3

X	X	0
X	↘	1
4	3	2

Table 4. Possible movements revised

X	X	X
X	↓	X
3	2	1

X	X	X
3	↙	X
2	1	X

3	X	X
2	←	X
1	X	X

2	3	X
1	↖	X
X	X	X

1	2	3
X	↑	X
X	X	X

X	1	2
X	↗	3
X	X	X

X	X	1
X	→	2
X	X	3

X	X	X
X	↘	1
X	3	2

Table 5. Final possible movements

Figure 4, nonetheless, reveals that we still have a problem. In fact, even if the formalization of factors 1, 2, and 3 has made our ANN tolerant to the typical cable problems that initially were ignored, it still disregards the important fact that our ANN cannot choose a route that exits the inspected seabed corridor. This is evident from an analysis of Figure 4.

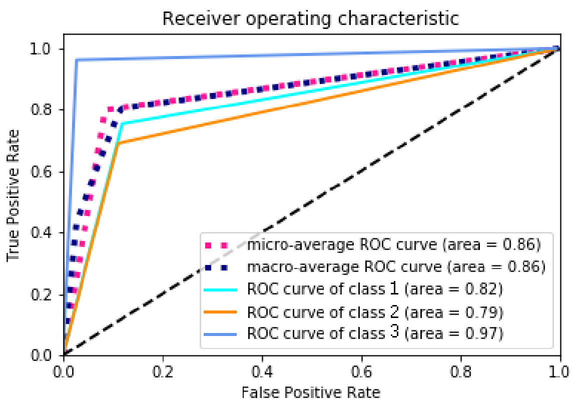

Figure 3: Movements: AUC values

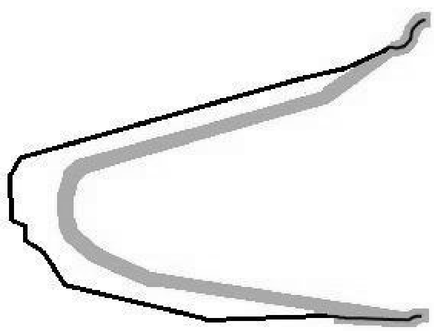

Figure 4: Route (leftmost), corridor (rightmost)

3.2 Learning through failures

A major problem here does exist. Human knowledge is limited and humans are afraid of all what is out of their knowledge. This is exactly the case. A vessel returns to us information about the soil type of a given seabed corridor. Geologists consider as acceptable only those routes that fall within that given corridor. For reasons we are not able to understand, our ANN often tends to exit that corridor, as Figure 4 confirms. We need hence a method to pass to our ANN this kind of contextual information: "Please don't get out of the corridor". To this aim, we designed and implemented a procedure to let the ANN learn this information, based on a typical learning through failures approach [22 - 24].

Our technique is as follows. We take our trained ANN and we run it. When the network predicts a point to follow, we control that the chosen point is within the corridor. If not, we force the ANN to follow an alternative point (the second best, which is within the corridor). Then, we save this information. We do like this, until a final correct route is calculated. At this point, we re-train a new ANN, based also on the errors the previous ANN have incurred in. Then we run again this ANN, and repeat the process train-run-error-codify-re-train till the point our ANN does not get out of the corridor any longer.

We have replicated this process six times with six different ANNs, with six different training configuration. Table 6 demonstrates that this process has converged for all the six cases we have examined. Each of the six ANNs converges to a correct route after a three-step long training process.

To check whether our method can be generalized, we have taken a yet new ANN and tried to train it with all the information (both positive and negative) that have stemmed from the previously described process of Table 6, where exactly 376 errors were made in total, out of 3667 route points. (Not) Surprisingly, this final ANN always remains within the seabed corridor, having learnt that getting out of it is a bad thing.

	Step 1	Step 2	Step 3
ANN 1	20	14	0
ANN 2	285	0	-
ANN 3	10	0	-
ANN 4	30	8	0
ANN 5	1	0	-
ANN 6	8	0	-

Table 6. Number of errors

To further exacerbate our analysis, we have repeated the process with other twelve different ANNs, each one learning the information summarized in Table 6. We can confirm that no one of them got out of the corridor. We conducted a kind of sensitivity analysis to understand if a smaller amount of errors can be sufficient (< 376) to let the ANN learn that it has not to get out of the corridor. This analysis has produced the graph of Figure 5, where it is shown that all the twelve ANNs are able to remain within the corridor with only a 75% of the amount of errors shown to them (i.e., 282).

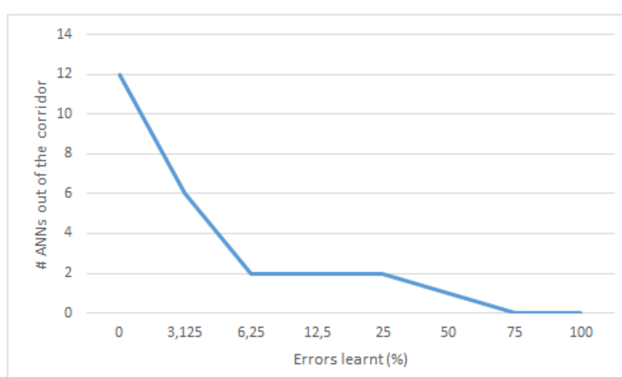

Figure 5: ANNs: remaining within the corridor

3.3 Are this kind of results really good?

To conclude, a good measure to understand if an ANN makes a correct decision, in this specific underwater cable installation case, is to verify if the length of the route computed by ANN following our training method is comparable to the length of the route provided by the geologists. The result is shown in Table 7. As evident, all the routes are in the corridor and their length is comparable to the one defined by geologists (at most 0.025% of difference).

Instance	Out	Ratio
1	No	0.998
2	No	1.010
3	No	1.021
4	No	1.000
5	No	1.017
6	No	1.019
7	No	1.013
8	No	1.019
9	No	1.025
10	No	1.018
11	No	1.005
12	No	1.019

Table 7. ANNs: route length ratio

4 Conclusion

There has been much commentary about the negative potential for ANNs to have a detrimental impact to humans and to the society. Our vision is that a comprehension and formalization of the context where the ANNs have to operate can smooth out this problem.

Obviously, a complex task is devising an efficient and speedy method to train an ANN, while managing an intelligible and accurate formalization of the context of interest.

We have presented a case study where the comprehension and codification of contextual information have improved the behaviour of an ANN, making it more *human*.

ACKNOWLEDGMENTS

The authors are grateful towards Andrea Lipparini and Maurizio Sobrero (University of Bologna) for their precious support. We are also indebted towards the company that has provided the underwater data of interest. To guarantee its privacy, we still keep it here anonymized.

REFERENCES

[1] J. Lemley, S. Bazrafkan, and P. Corcoran. "Deep Learning for Consumer Devices and Services: Pushing the limits for machine learning, artificial intelligence, and computer vision." IEEE Consumer Electronics Magazine 6.2 (2017): 48-56.

[2] Y. Wu, M. Schuster, Z. Chen, Q. V. Le, M. Norouzi, W. Macherey, M. Krikun, Y. Cao, Q. Gao, K. Macherey, J. Klingner, A. Shah, M. Johnson, X. Liu, L. Kaiser, S. Gouws, Y. Kato, T. Kudo, H. Kazawa, K. Stevens, G. Kurian, N. Patil, W. Wang, C. Young, J. Smith, J. Riesa, A. Rudnick, O. Vinyals, G. Corrado, M. Hughes, and J. Dean. 2016. Google's Neural Machine Translation System: Bridging the Gap between Human and Machine Translation. CoRR abs/1609.08144 (2016). arXiv:1609.08144 http://arxiv.org/abs/1609.08144

[3] A. Van Den Oord, S. Dieleman, H. Zen, K. Simonyan, O. Vinyals, A. Graves, N. Kalchbrenner, A. W Senior, and K. Kavukcuoglu. 2016. WaveNet: A generative model for raw audio.. In SSW. 125.

[4] J. Son Chung and A. Zisserman. 2018. Learning to lip read words by watching videos. Computer Vision and Image Understanding (2018).

[5] S. Suwajanakorn, S. M. Seitz, and I. Kemelmacher-Shlizerman. 2017. Synthesizing Obama: Learning Lip Sync from Audio. ACM Trans. Graph. 36, 4, Article 95 (July 2017), 13 pages. https://doi.org/10.1145/3072959.3073640

[6] F. Chen, B. Li, R. Dong, and P. Zhao. 2018. High-Performance OCR on Packing Boxes in Industry Based on Deep Learning. In Pacific Rim International Conference on Artificial Intelligence. Springer, 1018–1030.

[7] D. Silver, A. Huang, C. J. Maddison, A. Guez, L. Sifre, G. Van Den Driessche, J. Schrittwieser, I. Antonoglou, V. Panneershelvam, M. Lanctot, et al. 2016. Mastering the game of Go with deep neural networks and tree search. Nature 529, 7587 (2016), 484.

[8] M. Moravˇc´ık, M. Schmid, N. Burch, V. Lis´y, D. Morrill, N. Bard, et al. Deepstack: Expert-level artificial intelligence in heads-up no-limit poker. Science, 356(6337):508–513, 2017.

[9] O. AI, "The open ai dota 2 bot," 2018, [Accessed 20-August-2018]. [Online]. Available: https://blog.openai.com/more-on-dota-2/

[10] S. Ferretti, S. Mirri, C. Prandi, and P. Salomoni. 2016. Automatic web content personalization through reinforcement learning. Journal of Systems and Software 121 (2016), 157–169. https://doi.org/10.1016/j.jss.2016.02.008

[11] M. Roccetti, P. Salomoni, C. Prandi, G. Marfia, M. Montagnani, and L. Gningaye. 2016. Understanding Crohn's disease patients reaction to infliximab from facebook: A medical perspective. In Proceedings of the 2016 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining, ASONAM 2016. 1007 – 1010. https://doi.org/10.1109/ASONAM.2016.7752364

[12] M. Roccetti, C. Prandi, P. Salomoni, and G. Marfia. 2016. Unleashing the true potential of social networks: confirming infliximab medical trials through Facebook posts. Network Modeling Analysis in Health Informatics and Bioinformatics 5, 1 (2016). https://doi.org/10.1007/s13721-016-0122-9

[13] Silvia Mirri, Catia Prandi, Marco Roccetti, and Paola Salomoni. 2017. Handmade Narrations: Handling Digital Narrations on Food and Gastronomic Culture. ACM J. Comput. Cult. Herit (2017). https://doi.org/10.1145/3097569

[14] S. Ferretti, S. Mirri, C. Prandi, and P. Salomoni. 2017. On personalizing Web content through reinforcement learning. Universal Access in the Information Society (2017). https://doi.org/10.1007/s10209-016-0463-2

[15] A. Melis, S. Mirri, C. Prandi, M. Prandini, P. Salomoni, and F. Callegati. 2018. Integrating Personalized and Accessible Itineraries in MaaS Ecosystems Through Microservices. Mobile Networks and Applications (2018). https://doi.org/10.1007/s11036-017-0831-z

[16] R. Caruana, Y. Lou, J. Gehrke, P. Koch, M. Sturm, and N. Elhadad. 2015. Intelligible models for healthcare: Predicting pneumonia risk and hospital 30-day readmission. In Proceedings of the 21th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining. ACM, 1721–1730.

[17] A. Melis, S. Mirri, C. Prandi, M. Prandini, P Salomoni, and F. Callegati. 2016. CrowdSensing for smart mobility through a service-oriented architecture. In IEEE 2nd International Smart Cities Conference: Improving the Citizens Quality of Life, ISC2 2016 https://doi.org/10.1109/ISC2.2016.7580861

[18] S. Mirri, C. Prandi, P. Salomoni, and L. Monti. 2016. Social location awareness: A prototype of altruistic IoT. In 2016 8th IFIP International Conference on New Technologies, Mobility and Security, NTMS 2016. https://doi.org/10.1109/NTMS.2016.7792477

[19] J. Zou and L. Schiebinger. AI can be sexist and racist – it's time to make it fair. Nature 559, 324-326 (2018)

[20] J. Buolamwini and T. Gebru. 2018. Gender shades: Intersectional accuracy disparities in commercial gender classification. In Conference on Fairness, Accountability and Transparency. 77–91.

[21] G. Delnevo, A. Lipparini, M. Roccetti, and M. Sobrero. 2018. Human intelligence, ambidexterity and ai: Reflections on a data cognification case study in an underwater scenario. 2nd Annual Science Fiction Prototyping Conference 2018: Designing your Future with Science Fiction, SciFi-It 2018 (2018), 50–55.

[22] A. Arino and J. De La Torre. 1998. Learning from failure: Towards an evolutionary model of collaborative ventures. Organization science 9, 3 (1998), 306–325

[23] P. RP Coelho and J. E McClure. 2005. Learning from failure. American Journal of Business 20, 1 (2005), 1–1.

[24] D. A Shepherd. 2003. Learning from business failure: Propositions of grief recovery for the self-employed. Academy of management Review 28, 2 (2003), 318–328.