Machine Learning and Deep Learning Revolutionize Artificial Intelligence

Adil KORCHI, Fayçal MESSAOUDI, Lahcen Oughdir

Abstract— The concept of Artificial Intelligence (AI) is present in many aspects of our professional life. From automatic translation to predictive maintenance, to infused intelligence in the applications you use every day, it's not about the future, but about the company present. The fields of application and potential uses of Artificial Intelligence are more and more diverse: understanding of natural language, visual recognition, robotics, autonomous system, Machine Learning.

In this manuscript, we will show what is the Machine Learning concept and the Deep Learning as well as their position in artificial intelligence, their strengths and their flaws. We will present some algorithms that the Learning machine uses. We will also discuss the statistics used in these algorithms to adapt the links between the Artificial Neural Networks to strengthen or destroy the links in order to have a good approximation of the input data.

Index Terms— Artificial intelligence - Machine Learning - Deep Learning - Algorithms - Statistics - Artificial Neural Networks - Cloud Computing - Big Data

1 Introduction

This manuscript explains the power of artificial intelligence, machine learning and deep learning as well as the relationship between them. It also discusses statistical curves that approximate data and make it easy to generalize and cre- ate algorithms that can receive a lot of data to make a decision and create a model to automate tasks according to situations.

It also talks about Deep Learning and explains how it is inspired by our brain (with networks of neurons) to push the analysis further and know how to extract the data itself.

2 Definitions

2.1 Artificial Intelligence

Artificial intelligence can be defined as "the set of theories and techniques used to create machines capable of simulating intelligence (Krishnamoorthy, 2018). Either computers or programs capable of performance usually associated with human intelligence, and amplified by technology, namely:

1. Ability to reason
2. Ability to process large amounts of data
3. Ability to discern patterns and patterns undetectable by a human
4. Ability to understand and analyze these models
5. Ability to interact with humans
6. Faculty to learn gradually
7. And continually improve its performance

"Artificial intelligence" thus covers a vast subject, in perpetual mutation. And the dazzling progress since 1950, the founding year of AI.

2.2 Machine Learning (ML)

It is a concept that is more and more talked about in the world of computing, and that relates to the field of artificial intelligence. Still called "statistical learning", this term refers to a process of development, analysis and implementation leading to the establishment of systematic processes. To put it simply, it is a kind of program that allows a computer or a machine an automated learning, so that it can perform a number of very complex operations. The aim is to make the machine or computer capable of providing solutions to complicated problems by processing an astronomical amount of information. This offers an opportunity to analyze and highlight the correlations that exist between two or more given situations, and to predict their different implications.

2.3 Deep Learning (DL)

Deep Learning is a subdomain of Machine Learning. It is a form of artificial intelligence, derived from Machine Learning. To understand what Deep Learning is, it is important to understand what machine learning is. Deep learning usually has two phases. The first is to estimate a model from data, called observations, that are available and in finite numbers, during the design phase of the system (Goodfellow, 2016). Model estimation involves solving a practical task, such as translating a speech, estimating a probability density, recognizing the presence of a cat in a photograph, or participating in the driving of an autonomous vehicle. This so-called "learning" or "training" phase is generally performed prior to the practical use of the model. The second phase corresponds to the setting in production: the model being determined, new data can then be submitted in order to obtain the result corresponding to the desired task. In practice, some systems can continue their learning once in production, provided they have a way to get a return on the quality of the results produced.

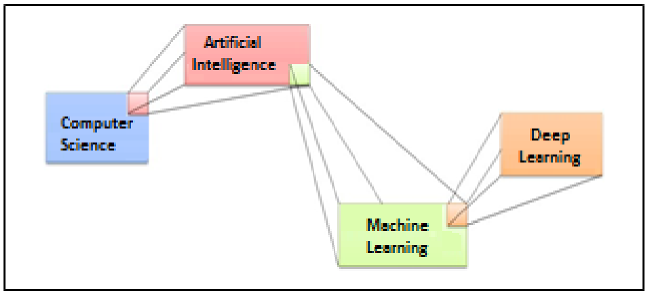

Fig. 1. Place of Deep Learning in the world of computing

Deep learning has yielded impressive results in many different fields:

1. Recognition of image, text, voice, face ...
2. Segmentation in the medical field, understanding a scene, a text ...
3. Generation of image, text, voice, artwork, human face ... Autonomous car, autonomous robot ...
4. Road monitoring, pedestrian ...
5. The victory in the GO game against the world champion of this discipline.

We could still mention many more original examples than others, but what must be remembered is that deep learning makes it possible to teach a computer a precise task by observing a large number of examples.

3 Sections Machine Learning and Deep Learning are Artificial Intelligence

3.1 Machine Learning

Artificial Intelligence as we know it is a weak Artificial Intelligence, as opposed to Strong Artificial Intelligence, which does not exist yet (Hamming, 2017). Today, machines are able to reproduce human behavior, but without awareness. Later, their abilities could grow to the point of becoming machines with consciousness, sensitivity and spirit.
If Machine Learning and Deep Learning are Artificial Intelligences, the opposite is not true. For example, knowledge graphs or rule engines are Artificial Intelligences but are not Machine Learning or Deep Learning. Deep Learning is a branch of Machine Learning.
Artificial Intelligence has evolved a lot thanks to the emergence of Cloud Computing and Big Data (Beam, 2018), a low-cost computing power and accessibility to a large amount of data. Thus, the machines are no longer programmed; they learn.
Machine Learning, or machine learning, is able to reproduce behavior thanks to algorithms, themselves powered by a large amount of data. Faced with many situations, the algorithm learns the decision to adopt and creates a model. The machine can automate tasks according to the situations.
To better understand this principle, we give the following example:
Before 1980, programs were made to calculate the price of apartments according to their area, for example, "if the area is less than 30m ^{2}, the price is worth € 65 000, if it is between 30m ^{2} and 50m ^{2}, the price is worth € 85,000, etc ... ", or maybe" price = area * € 3,500 "
These approximations are not satisfactory, and it would suffice to note the price of the apartment knowing its true area to estimate the price of a new flat of non-referenced. By adopting this reasoning, we give birth to machine learning. So we can create statistical curves that approximate the data and make it easy to generalize and create algorithms that can receive a lot of data.
Machine learning is a broad field, which includes many algorithms. Among the most famous, we find:

- Regressions (linear, multivariate, polynomial, regularized, logistic ...): these are curves that approximate the data (see diagram above)

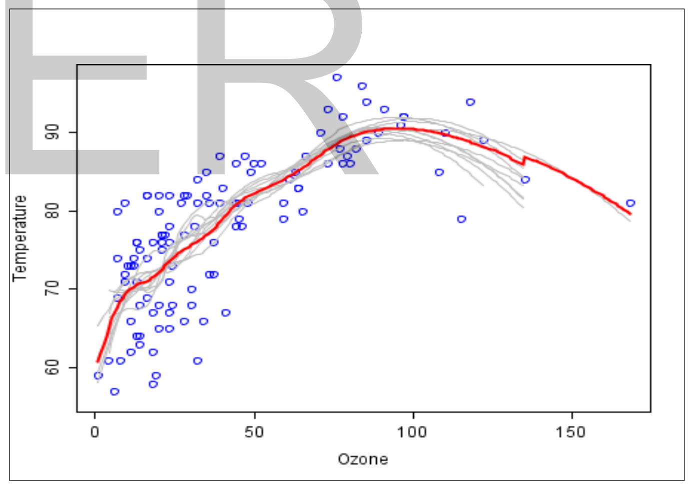

Fig. 2. Data approximation curves

- The Naïve Bayes algorithm: the algorithm gives the probability of the prediction, knowing the previous events (El Kourdi, 2004). For example, what is the most likely price of an 87.6m² apartment.
- Clustering: always thanks to mathematics, we will group the data in packets so that in each packet the data are as close as possible to each other. It is used especially for recommendations of movies "close" movies you have already seen!
- Decision trees: by answering a certain number of questions and following the branches of the tree that carry these answers, we arrive at a result (with a

probability score)

- More sophisticated algorithms based on several statistical techniques: Random Forest (a forest of decision trees that vote), Gradient Boosting, Support Vector Machine ...

```
TRAINMULTINOMIALNB(C, D)
  V ← EXTRACTVOCABULARY(D)
  N ← COUNTDOCS(D)
  for each c ∈ C
  do N_c ← COUNTDOCSINCLASS(D, c)
     prior[c] ← N_c / N
     text_c ← CONCATENATETEXTOFALLDOCSINCLASS(D, c)
     for each t ∈ V
     do T_ct ← COUNTTOKENSOFTERM(text_c, t)
     for each t ∈ V
     do condprob[t][c] ← (T_ct + 1) / Σ_t'(T_ct' + 1)
  return V, prior, condprob

APPLYMULTINOMIALNB(C, V, prior, condprob, d)
  W ← EXTRACTTOKENSFROMDOC(V, d)
  for each c ∈ C
  do score[c] ← log prior[c]
     for each t ∈ W
     do score[c] += log condprob[t][c]
  return arg max_{c∈C} score[c]
```

Figure 3: Naive Bayes algorithm (multinomial model): Training and testing.

3.2 Deep Learning

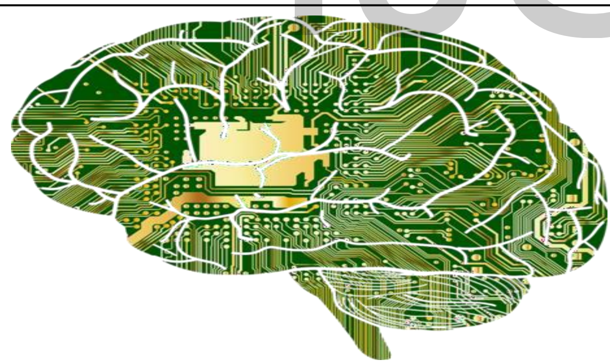

Fig 4: Artificial Neural Networks (Deep Learning)

Despite its power, Pure Machine Learning has many flaws. The first is that a human expert must sort the data to remove any unnecessary information and could clutter or penalize the system.
For example, for the sale of an apartment, if you think that the owner's age does not affect the price, there is no point in giving this information to the algorithm because if you give the system too much, it could see relationships where there are none.
Then, the second (which follows from the first): how to recognize a face? You could give the algorithm a lot of information about the person (gap between the eyes, forehead height, etc ...), but it would not be very adaptive or precise.

Thus was born the idea of Deep Learning towards the year 2010: to take inspiration from how our brain works (with networks of neurons) to push the analysis further and know how to extract the data itself!
Deep learning is therefore based on what are called artificial neural networks (deep), that is to say a set of neurons (they are small calculators that perform a mathematical operation) that send themselves numbers in function of their links, up to output neurons (Hertz, 2018). Thanks to this architecture, the Deep Learning is able to recognize faces, to synthesize texts or to drive an autonomous car!

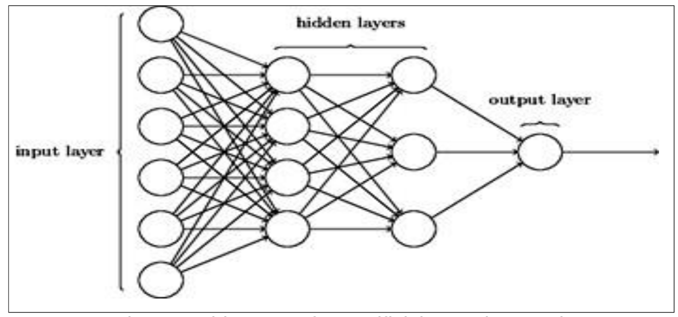

Fig. 5. Architecture of an artificial neural network

The next step is to use the statistics for the algorithm to adapt the links between its neurons to reinforce or destroy them, to ensure that at the output we have a good approximation of the input data. For example, in a network that has learned to predict the price of an apartment, if we give in input "30", then out we will have a number very close to "65,000".

We can say that :
DL = ML + NEURONS NETWORKS
where DL extracts automatically relevant data information and adapts its algorithm.

3.3 Examples of Deep Learning algorithms:

Artificial neural networks (ANN): these are the simplest and are often used in addition because they sort the information well

- Convolutional Neural Networks (CNNs): specialized in image processing, they apply filters to data to bring out new information (for example, highlighting outlines in an image can help find where the face)
- Recurrent neural networks (RNN): the most well known are the LSTMs, which have the advantage of retaining information and reusing it soon after (Snoek, 2012). They are used for the analysis of text (NLP), since each word depends on the few previous words (so that the grammar is correct)
- More advanced versions, such as auto-encoders, Boltzmann machines, self-organizing maps (SOM) ...

The figure below shows where machine learning and deep learning are in artificial intelligence.

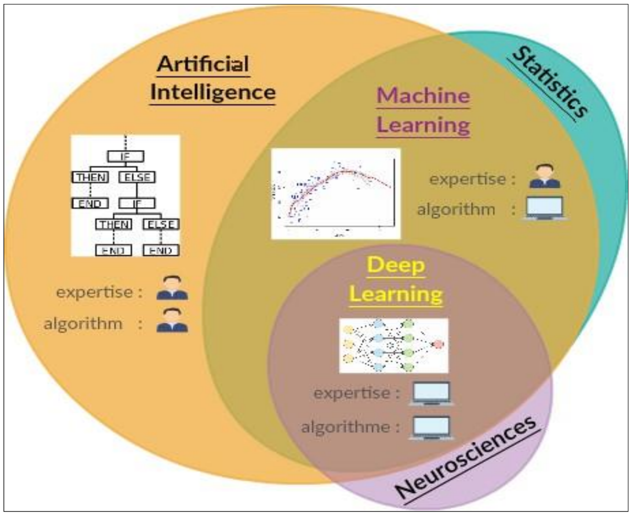

Fig. 6 : DL and ML used with neural networks using Artificial intelligence

4 Conclusion

Machine learning and deep learning are very powerful tools that make it possible to perform multiple actions such as classifying data, teaching a program from experiments, or creating an evolutionary program that is constantly improving. Thus, even with a little provided sample and data influenced by the subjectivity of the one whom measures them, these two tools remain relatively precise despite some shortcomings.

Nevertheless, the machine and the deep learning do not have only qualities, they must be constantly adapted to the problems that they try to solve. In fact, the programmer must first obtain the most representative sample possible. Then he will have to choose the function most faithful to the sample. Finally, they should be used as tools because not all problems require a complex program in machine or deep learning.

Deep learning and machine learning are going to be generalized "in all decision-making electronics", as in cars or planes, they will make it possible to do without a human expert to sort the data, since the algorithm will find its own correlations. To take the example of facial recognition, the DL algorithm will determine for itself whether to take into account the difference between the eyes (between the pixels) or if this information is not decisive enough compared to others (and this is indeed the case).

References

[1] BEAM, Andrew L. et KOHANE, Isaac S. Big data and machine learning in health care. Jama, 2018, vol. 319, no 13, p. 1317-1318.

[2] BIAMONTE, Jacob, WITTEK, Peter, PANCOTTI, Nicola, et al. Quantum machine learning. Nature, 2017, vol. 549, no 7671, p. 195.

[3] BOTTOU, Léon, CURTIS, Frank E., et NOCEDAL, Jorge. Optimization methods for large-scale machine learning. Siam Review, 2018, vol. 60, no 2, p. 223-311.

[4] BRADLEY, Andrew P. The use of the area under the ROC curve in the evaluation of machine learning algorithms. Pattern recognition, 1997, vol. 30, no 7, p. 1145-1159.

[5] DENG, Li, YU, Dong, et al. Deep learning: methods and applications. Foundations and Trends® in Signal Processing, 2014, vol. 7, no 3–4, p. 197-387.

[6] DIETTERICH, Thomas G. Ensemble methods in machine learning. In : International workshop on multiple classifier systems. Springer, Berlin, Heidelberg, 2000. p. 1-15.

[7] DONG, Chao, LOY, Chen Change, HE, Kaiming, et al. Learning a deep convolutional network for image super-resolution. In : European conference on computer vision. Springer, Cham, 2014. p. 184-199.

[8] EL KOURDI, Mohamed, BENSAID, Amine, et RACHIDI, Tajje-eddine. Automatic Arabic document categorization based on the Naïve Bayes algorithm. In : proceedings of the Workshop on Computational Approaches to Arabic Script-based Languages. Association for Computational Linguistics, 2004. p. 51-58.

[9] FERBER, Jacques et WEISS, Gerhard. Multi-agent systems: an introduction to distributed artificial intelligence. Reading : Addison-Wesley, 1999.

[10] GOLDBERG, David E. et HOLLAND, John H. Genetic algorithms and machine learning. Machine learning, 1988, vol. 3, no 2, p. 95-99.

[11] GOODFELLOW, Ian, BENGIO, Yoshua, et COURVILLE, Aaron. Deep learning. MIT press, 2016.

[12] HAMMING, Richard W. et FEIGENBAUM, Edward A. Problem solving methods in artificial intelligence. 2017.

[13] HERTZ, John A. Introduction to the theory of neural computation. CRC Press, 2018.

[14] KRISHNAMOORTHY, C. S. et RAJEEV, S. Artificial intelligence and expert systems for engineers. CRC press, 2018.

[15] LU, Huimin, LI, Yujie, CHEN, Min, et al. Brain intelligence: go beyond artificial intelligence. Mobile Networks and Applications, 2018, vol. 23, no 2, p. 368-375.

[16] MÜLLER, Vincent C. et BOSTROM, Nick. Future progress in artificial intelligence: A survey of expert opinion. In : Fundamental issues of artificial intelligence. Springer, Cham, 2016. p. 555-572.

[17] PAN, Yunhe. Heading toward artificial intelligence 2.0. Engineering, 2016, vol. 2, no 4, p. 409-413.

[18] SEBASTIANI, Fabrizio. Machine learning in automated text categorization. ACM computing surveys (CSUR), 2002, vol. 34, no 1, p. 1-47.

[19] SNOEK, Jasper, LAROCHELLE, Hugo, et ADAMS, Ryan P. Practical bayesian optimization of machine learning algorithms. In : Advances in neural information processing systems. 2012. p. 2951-2959.

[20] Outputs, with Relationships to Statistical Pattern Recognition," *Neurocomputing – Algorithms, Architectures and Applications,* F. Fogelman-Soulie and J.