Algerian Journal of Sciences-A, Special issue (NCAIIT'2021 conference) 139-143

REVUE ALGERIENNE DES SCIENCES
Section -A-
Sciences de la Nature et de la Vie
Sciences techniques
http://univ-eltarf.dz/fr/

An overview of machine learning and deep learning

Karima Saidi*, Ouassila Hioual
Abbes Laghror University, Khenchela, 40004, ALGERIA

Informations

Keywords:
Artificial Intelligence
Machine learning
Deep learning

***Corresponding author:**
kariming2008@gmail.com

Abstract

Artificial Intelligence, machine learning, and deep learning have now become commonplace in our daily lives. However, the words are sometimes used interchangeably. In this paper, we learn how to tell them apart. We will explain the field of machine learning and its concepts. Furthermore, the field of deep learning known as automatic learning is extremely important. Finally, we will present the main differences between them.

1. Introduction

How to organize to make decisions, that is, to process information, is the main problem in organizational structure study in any system. If the problem of organizational structure is the effective and efficient processing of information, then it is worth exploring the role of technology to enable automatic information processing such as using machine learning(ML) and deep learning (DL) in the current time [1]. Decision-making is the method of choosing an element from the available ones based on single or multiple criteria. Therefore, ML or DL is,increasingly,used for making the decision and development of forecasting systems. In this paper, we provide a brief overview of ML and DL by explaining the main difference among them.
In the next section, we present an overview of Artificial Intelligent with its sub-fields Machine learning & Deep learning. Then, section 3 gives a comparison between conventional Machine Learning and Deep Learning. Finally, section 4 concludes the paper.

2. Artificial Intelligence (AI)

AIhas been a topic of extreme world attention in recent years. It is a general field covering machine learning and deep learning. These latter are the subfields of AI as shown in Fig.1.AI could be defined as a system capable of interacting with its environment.

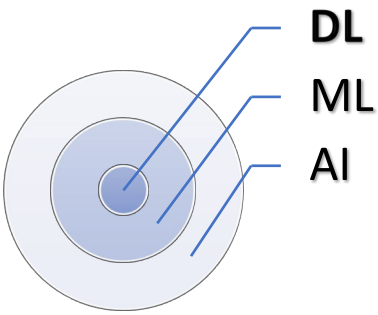


Fig.1: Relationship between AI, ML, and DL

2.1. Machine Learning

Machine learning is a tool used in the processing of large-scale data. Complex datasets with huge numbers of variables and features are well suited for this. These very large datasets use machine-learning techniques to make sense. We can say that machine learning is the brain of an AI system.
Machine learning can quickly and automatically produce models that can analyze larger, more complex data.It delivers faster, more accurate results on a very large scale. In addition,it enables the construction of accurate models [2]. It has recently emerged as a revolutionary technology to balance the difficulty of computation and the performance gap to solve NP-hard issues.

A- Different machine learning algorithms

Machine learning techniques (algorithms) can be divided into three major classes[1]:

- Supervised learning: Supervised learning algorithms are applied to learn to map between inputs and output, typically referred to like features (X) and targets (Y).

We distinguish two tasks in this class: Classification and regression

- Unsupervised learning: unsupervised learning algorithms learn patterns within the feature inputs (X) in the absence of a particular target variable. It has two tasks dimensionality reduction and clustering.
- Reinforcement learning: Their algorithms rely on determining actions that enable it to attain its goals, by interacting with the environment.

ML algorithms are the best choice when we have several criteria in one task. As shown in Fig. 2, there are many used algorithms in ML such asclassification, Support Vector Machine (SVM), Decision Tree, Logistic Regression, Naïve Bayes, Neural network, and K-nearest Neighbor (KNN) [3].

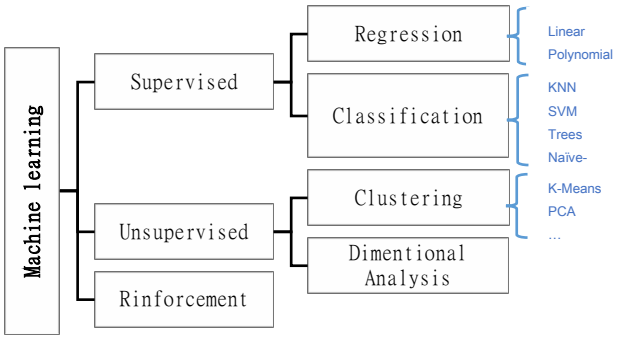


Fig.2: Machine Learning algorithms

B- Special concepts used in ML

Before discussing ML and DL architectures, we first describe the basics of some keywords, which are the foundation for such algorithms.

a) Neural network

We describe the fundamentals of **neural networks** on which such algorithms are conceptually based. As mentioned in[4],a neural network can be described as a mathematical model for information processing. Learning in a neural network means finding a set of values for the weights of all layers in a network. As shown in Fig.3, to compute the weighted sum of the input features, a neuron uses input features x_1 ; x_2 ; . . .; $x_n$and weight parameters w_0 ; w_1 ; . . .; w_n. The most well-known and very simple architecture to understand neural networkis the feed-forward multilayer network with an input layer, one or many hidden layers, and a single output layer. It can represent any function, and, generally, trained by a learning algorithm called backpropagation learning. This latter is used to reduce error in a neural network, we change the weights progressively as our algorithm finds the best solution.

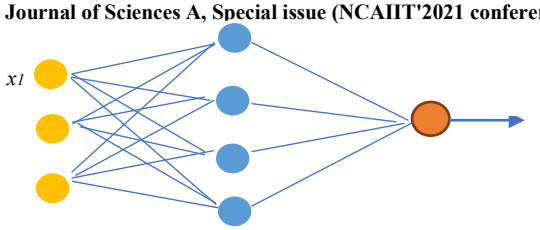


Fig. 3: Single hiddenlayer Neural network.

b) Perceptron

Perceptron is the linear model used for binary classification with simple input-output. An artificial neural is the perceptron in the field of neural network.

- *The single-layer Perceptron*

As the name implies, a single-layer artificial neural network, also known as a single-layer, has a single layer of nodes. Each node in the single layer corresponds to an output variable and is directly connected to an input variable.

- *The multilayer perceptron (MLP)*

A node, also known as a neuron or perceptron, is a computational unit that consists of one or more weighted input connections, a transfer function that incorporates the inputs, and an output connection. After that, nodes are organized into layers to form a network. The layers in a MLP can be summarized as follows[4]:

- Input Layer: Input variables, also known as the visible layer.
- Hidden Layers: Node layers that exist between the input and output layers. One or more of these layers can exist.
- Output Layer: A layer of nodes responsible for producing output variables.

In addition, we have someterms, which are used to define the shape and capabilities of a neural network, among which we can cite:

- Size: The number of nodes in the model. The number of nodes in each layer is defined as an integer, from the input layer to the output layer.
- Width: The number of nodes in a given layer.
- Depth: The number of layers in a neural network is defined as its depth.
- Capacity: The form or arrangement of functions that can be learned by a network configuration. Often referred to as "representative capacity."
- Architecture: The basic structure of the network's layers and nodes.

c) Regularization

Regularization is used when a network is overfitting.

d) Generalization

Generalization refers to the capacity to perform well on previously unobserved inputs.

e) Underfitting and overfitting

Underfitting happens when the model is unable to achieve a low enough error value on the training set. When the difference between the training and test errors is too high, overfitting occurs.

f) Transfer Learning (TL)

TL is a technique for improving a learner from one domain by transferring information from a related domain [5]. Moreover,Transfer learning attempts to transfer algorithmic operations to related tasks from a focal task. Since the algorithm does not require transfer learning with pre-trained models to learn model parameters from scratch, fine-tuning of models is not only easier but also helps avoid overfitting when small datasets are available.When target training data are scarce, transfer learning becomes essential. This may be due to the data being scarce, costly to obtain, or unavailable. Thus, in such cases, it would be beneficial if classification information of another domain could be transferred tothe desired domain.

Conventional machine learning techniques attempt to learn each task from the scratch, while transfer learning techniques attempt to transfer knowledge from previous tasks to a new task if the target has less high-quality target data[6]. This is the main difference between conventional machine learning and transfer learning as we can show in Fig.4.

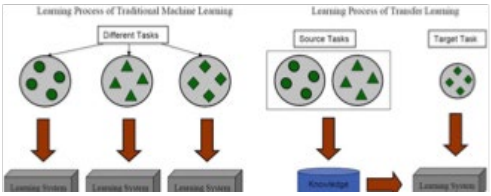


Fig.4: The learning processes between conventional machine learning and transfer learning [6]

d) Multi-task learning

The multi-task learning framework is a closely related learning strategy for transferring learning. Learning several tasks at the same time allows for the transfer of knowledge between similar tasks and increased efficiency as opposed to learning each task separately[7]. Multi-task learning seeks to increase learning efficiency and prediction accuracy by learning multiple goals from a shared representation[8]. Additionally, Multi-task learning has been used successfully across all applications of machine learning and deep learning. Multi-task learning with multiple regression and classification goals benefits multiple deep learning applications [9].

In a distributed computing environment, to make learning multiple tasks with unknown task relationships possible [10],we first derive a general dual form for a family of regularized multitask relationship learning methods.

In [11], the author partitioned the MTL architectures into four groups: architectures for a particular task domain, multi-modal architectures, learned architectures, and conditional architectures

Table 1: Architectures of MTL

Architectures	Explanation
Architectures for a particular task domain	This group of architecture is used in particular domains such as: • Computer vision, • Natural language processing • Reinforcement learning
Multi-modal architectures	Used in visual question answering with both a visual and a language component
Learned architectures	Learned architectures are fixed between steps of architecture learning. So, the same computation is performed for each input from the same task
Conditional architectures	This architecture is used for a given piece of data is dependent on the data itself.

2.2. Deep Learning

DL is a subfield of machine learning which is capable of learning the appropriate features by itself, requiring little steering by the user[12]. It is not a fixed algorithm that does not change, but rather a model or a system; we can change according to our information processing. DL is not always the best method for a given area. This is because maybe there are not enough data to allow deep learning, and or maybe it exists another algorithm that solves the problem better than it. Otherwise, its techniques are so effective as they learn how to better represent the problem while learning how to solve it. The Deep represents how many layers contribute to a model of the data. We call it the depth of the model.

2.2.1. Fundamentals steps for building deep networks

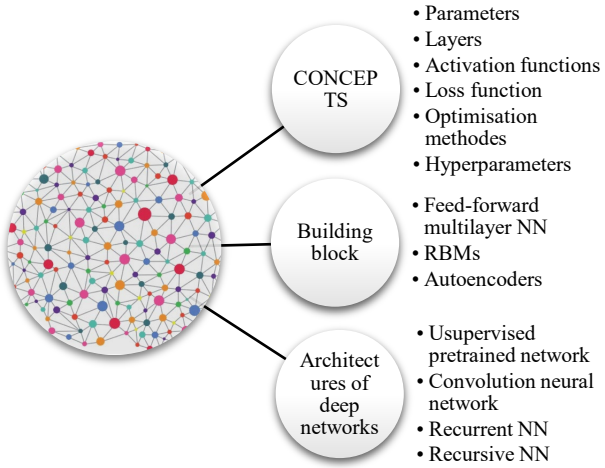


Fig. 5: Fundamentals steps for building deep networks

To build a model in deep learning, we need to grasp the fundamentals steps as we can have showed in Fig. 5. The following are some of the concepts to consider:

Parameters equivalent to the weight on the connections in the network

Layer input, hidden layers, and output layer define feedforward neural networks

Activation functions

- Output layer for binary classification. In this case, we'd use a sigmoid output layer with a single neuron to give us a real value in the range of 0.0 to 1.0 (excluding those values) for the single class.
- Output layer for multiclass classification. If we have a multiclass modeling problem yet we only care about the best score across these classes, we'd use a SoftMax output layer with an arg- max () to achieve the highest score for all the classes. The SoftMax output layer gives us a probability distribution over all the classes.

Loss Functions for classification

- Hing loss
- Logistic loss
- Negative log like a likelihood

Hyperparameters: Hyperparameters are divided into several categories as mentioned bellow:

- Layer size refers to the number of neurons in each layer.
- Magnitude (momentum, learning rate): such as the Learning rate is considered as one of the key hyperparameters in neural networks. It is how fast we change the weights.
- Regularization (dropout, drop connect, L1, L2): Regularization is a measure taken against overfitting,
- Dropout is a mechanism used to improve the training of neural networks by omitting a hidden unit.
- Activations (and activation function families)
- Weight initialization strategy
- Loss functions
- Settings for epochs during training (mini-batch size): a group or batch of vectors to be trained in the learning system. This allows using of hardware and resources more efficiently at the computer architecture level
- Normalization scheme for input data (vectorization)

3. Comparison between Conventional Machine Learning and Deep Learning

	Conventional Machine Learning	Deep Learning
Feature Learning	Using the feature extraction method with human intervention	Automatic feature extraction
Model Construction	Applying extracted features to construct a data-driven model, with shallow structures	An end-to-end high hierarchical model structure with nonlinear combining of multi-layers
Model training	Training each module step-by-step	Jointly training the parameters
Data requirement	Requiring large data	Lesser data are trained
Training Time	Long time	Less time
Transfer learning	Useful	Very useful
Multi-task learning	successful	Very successful

Conventional machine learning methods include algorithms such as logistic and linear regression, k-means clustering, decision trees, support vector machine (SVM), random forest, and Bayes learning... etc.

Among the benefits of machine learning techniques, and particularly deep learning is that when used on large datasets, they perform best, thus improving their analytical and predictive power. One of the benefits of DL over conventional machine learning approaches is that feature learning is performed automatically, so there is no need to develop a separate algorithm to do this part [13]. The main differences between conventional machine learning and deep learning are summarized in several points according to [14]. So, we show in Table 2 the conventional (traditional) ML versus DL algorithm comparison.

Table 2: The main differences between conventional machine learning and deep learning

Keys to a successful ML/DL project

As we have shown, it is best to be very precise about the issue that we are attempting to solve. That way, we will know exactly what it means when the model works and behaves correctly.
If we do plan to build our own, we will need to have a lot of data to allow for analysis and building a model. This could be data we have within your business, or sometimes publicly available data sets could be

combined and used to train a model (using transfer learning).
Once we do have a tool implemented, it is good to plan out how we'll continue to improve it using information about its results (such as multi-task learning).
Even if we do start off with all the right steps, some applications are still technically difficult and may not provide the results we're looking for. But that is always an open challenge to research.

4. Conclusion

In this paper, wehave introduced the concepts of AI with its sub-fields that we intend to use for solving the NP-hard problem. Moreover, we gave the main differences between conventional Machine learning and Deep learning.

Reference

[1] Y. R. Shrestha, V. Krishna, and G. von Krogh, "Augmenting organizational decision-making with deep learning algorithms: Principles, promises, and challenges," *Journal of Business Research*, vol. 123, pp. 588–603, Feb. 2021, doi: 10.1016/j.jbusres.2020.09.068.
[2] J. Lin, Y. Dai, X. Chen, and Y. Wu, "Resource Allocation of Cloud Application Through Machine Learning: A Case Study," in *2017 International Conference on Green Informatics (ICGI)*, 2017, pp. 263–268.
[3] I. Goodfellow, Y. Bengio, and A. Courville, *Deep Learning*. MIT Press, 2016.
[4] J. Patterson and A. Gibson, *Deep Learning: A Practitioner's Approach*. O'Reilly Media, Inc., 2017.
[5] K. Weiss, T. M. Khoshgoftaar, and D. Wang, "A survey of transfer learning," *Journal of Big Data*, vol. 3, no. 1, p. 9, May 2016, doi: 10.1186/s40537-016-0043-6.
[6] S. Panigrahi, A. Nanda, and T. Swarnkar, "A Survey on Transfer Learning," in *Intelligent and Cloud Computing*, Singapore, 2021, pp. 781–789, doi: 10.1007/978-981-15-5971-6_83.
[7] J. Wang, M. Kolar, and N. Srerbo, "Distributed Multi-Task Learning," in *Artificial Intelligence and Statistics*, May 2016, pp. 751–760, Accessed: Apr. 05, 2021. [Online]. Available: http://proceedings.mlr.press/v51/wang16d.html.
[8] A. Kendall, Y. Gal, and R. Cipolla, "Multi-Task Learning Using Uncertainty to Weigh Losses for Scene Geometry and Semantics," 2018, pp. 7482–7491, Accessed: Apr. 05, 2021. [Online]. Available: https://openaccess.thecvf.com/content_cvpr_2018/html/Kendall_Multi-Task_Learning_Using_CVPR_2018_paper.html.
[9] S. Ruder, "An Overview of Multi-Task Learning in Deep Neural Networks," *arXiv:1706.05098 [cs, stat]*, Jun. 2017, Accessed: Apr. 05, 2021. [Online]. Available: http://arxiv.org/abs/1706.05098.
[10] S. Liu, S. J. Pan, and Q. Ho, "Distributed Multi-Task Relationship Learning," in *Proceedings of the 23rd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, New York, NY, USA, Aug. 2017, pp. 937–946, doi: 10.1145/3097983.3098136.
[11] M. Crawshaw, "Multi-Task Learning with Deep Neural Networks: A Survey," *arXiv:2009.09796 [cs, stat]*, Sep. 2020, Accessed: Apr. 07, 2021. [Online]. Available: http://arxiv.org/abs/2009.09796.
[12] M. A. Wani, F. A. Bhat, S. Afzal, and A. I. Khan, *Advances in Deep Learning*, vol. 57. Singapore: Springer Singapore, 2020.
[13] R. A. Khalil, N. Saeed, Y. M. Fard, T. Y. Al-Naffouri, and M.-S. Alouini, "Deep Learning in Industrial Internet of Things: Potentials, Challenges, and Emerging Applications," *arXiv:2008.06701 [eess]*, Aug. 2020, Accessed: Aug. 22, 2020. [Online]. Available: http://arxiv.org/abs/2008.06701.
[14] S. Shadroo, A. M. Rahmani, and A. Rezaee, "The two-phase scheduling based on deep learning in the Internet of Things," *Computer Networks*, p. 107684, Nov. 2020, doi: 10.1016/j.comnet.2020.107684.