Deep learning: concepts and implementation tools

Eddy Sánchez-DelaCruz and David Lara-Alabazares

Postgraduate Department, Technological Institute of Misantla, Veracruz Mexico.
{eddsacx, dlaraalab}@gmail.com

Abstract. In this paper, the concepts and tool available to use Deep learning in scholar projects are given. We carry out experiments by combining meta-classifiers with a deep artificial neural network in four binary datasets. The results show optimal percentages of correct classification in some cases. The sample criteria that prevalence in this study was a representative sample over traditional criteria.

Keywords: Artificial Intelligence · Deep Learning · Artificial Neural Network · Meta-classifiers

1 Introduction

At the beginning of the 21st century, Artificial Intelligence (AI) in its various disciplines that integrate it, has started in a surprising way, to emulate faithfully human behavior and reasoning. As results of this, remarkable progress have emerged in different fields such as the computer-assisted medical diagnosis, classification of DNA sequences, data mining, artificial vision, voice recognition, analysis of written language, virtual games, robotics, and any others where the reasoning is the main element [9].

Among the different disciplines of AI, the *deep learning* is a novel alternative (from a few decades ago) that has as main objective to make that an intelligent agent can be capable to make its own decisions, something that currently is only possible in science fiction. In this sense, different approaches of Artificial Neural Networks (ANN) are used in deep learning, having as goal provides to an agent of personality comparable to a human. Among these approaches, we have Deep Neural Networks, Convolutional Neural Networks, and Deep Belief Networks. Therefore, below is exposed briefly the topic of ANN that will be helpful to understand how deep learning works.

In this study, we described basic concepts of ANN, we mentioned some ANN tools and we experimented with various datasets.

The rest of article is divided as follow: in the section 2 the ANN, deep learning concepts are described and the tools for implement deep learning algorithms are mentioned, in the sections 3 and 4 the experiments and results are showed, finally, the conclusions are described in section 5.

2 Artificial Neural Network

ANN is inspired by the way the biological nervous system works. One of the pioneering work can be found in [8] where Warren McCulloch and Walter Pitts proposed the first mathematical model for a biological neuron.

Formally, an ANN can be defined as a processing element that receives a set of input elements, denoted as $X = x_1, x_2, \ldots x_n$, which are modified respectively for a series of weights $W = w_1, w_2, \ldots, w_n$. The different values that are modified by the weights are added together in what is called the net input (The net input is the result of the addition of the products of each input value by its corresponding weight adding a bias value or threshold of the neuron, denoted by b, which is determined when it is activated). The activation of the neuron depends on the activation function that acts on the net input which also regulates the output of the neuron. As it can be observed in the figure 1, that both the sum and the activation function represent the cell body of the neuron, in this place are realized the corresponding mathematical calculations and the results are transmitted to the output y [5]. In general, an ANN is divided into three layers: the input, hidden and output (see Figure 1), where the hidden layer has three neurons for this example.

2.1 Deep Learning

Deep Learning is a discipline for the search of patterns through deep abstractions that are achieved with multiple hidden layers of an ANN [2,1]. To achieve abstraction, in a hidden layer, for example, the border of the area of interest in an image is selected (Figure 2); depth is achieved by repeating abstraction in as many hidden layers as desired. Both the number of hidden layers and that of neurons range from 1 to n, since there is no metric that establishes how many to use, rather this tries to resolve the agreement to the problem, the dimensions and properties of the data set, and based on the experience of who implements the ANN. The output of our ANN will be the answer that we are looking for, which corresponds to a classification that tells us that in an input image such as the one in figure 2, there will be a representation of a "giraffe". Also, in the same figure, we can see how the input and output images would be represented for Deep learning.

2.2 Tools for implementing deep learning

There is a wide variety of tools in which Deep learning can be implemented. Table 1 lists the ones that to the authors' knowledge are the most representative. These tools can be found in the web as open source and are oriented to any passionate person with not too much experience in programming or to self-taught persons with a little patience, in such way that they can explore the Deep learning with some tool of their choice.

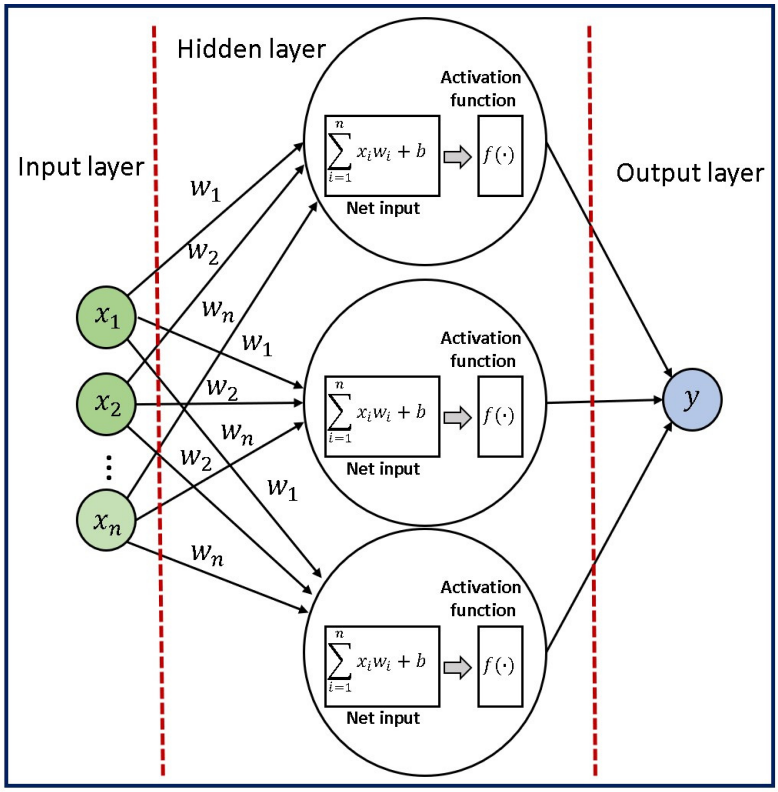
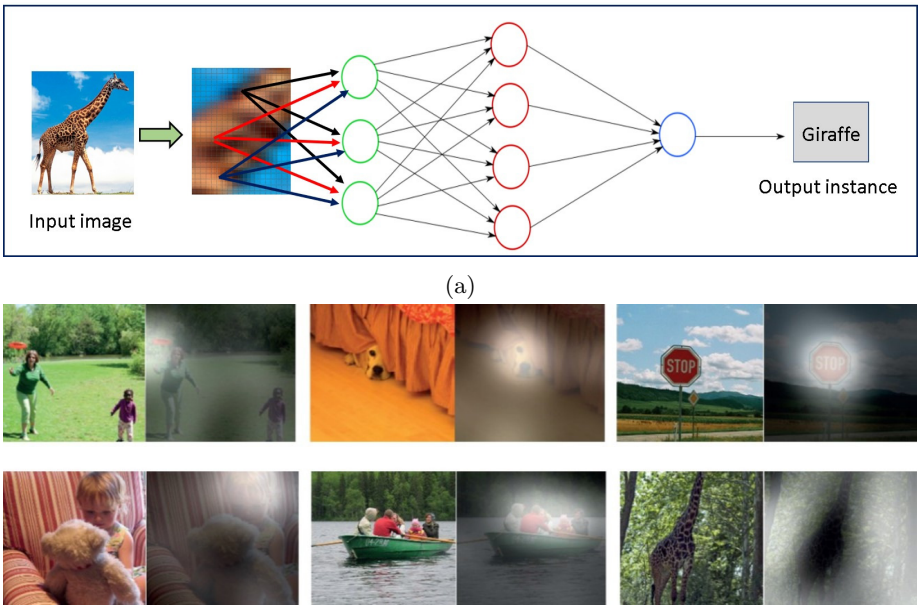


Fig. 1: Scheme example of an ANN.

3 Applying Deep learning

We carry out four experiments using datasets from public access, which are listed in Table 2.

- *Car-evaluation* dataset was created from a simple hierarchical decision model, originally developed for DEX demonstration, which is a system based on a methodology that combines multi-attribute decision making with expert systems [3].
- *Daily-deals* is a set of information about products supply and demand in an establishment (mall), the dataset is available, for download using RapidMiner, a data mining tool.
- *Pap-smear* dataset corresponds to microscopic images of pap-smears, which are categorized in seven states where cervical cancer disease could be [6].
- *Blood-transfusion* dataset consists of blood information of subjects to know if they are eligible (1) or not (0) to be donors [14].



(a)

(b)

Fig. 2: a) Abstractions using the Deep learning approach, b) input and output images before and after processing with Deep learning [7]

Each dataset was used with two classes; datasets with more than two classes were adapted to binary classes with the One-vs-All approach [10].

4 Results and analysis

The experiments were carried out using the WEKA v.3.8 tool (Waikato Enviroment Knowledge Analysis) [13,11], using a Lenovo G470 laptop, having an Intel Celeron @ 1.50GHz processor, with 2.00 GB RAM and 32 Bit professional windows 7 OS. The WEKA tool has several categories of algorithms, each category contains a variety of algorithms that can be implemented separately or by combining them. In this case, for the tests of each of the experiments, each classifier of the Meta-classifiers (assembled classifier) category was combined with the deep learning algorithm Dl4JMlp Classifier, that is, a classifier assembled with a deep ANN. The implementations were made using three sampling criteria: cross validation, $2/3 - 1/3$ and representative sample (RS). In Table 3 we can observe the highest results.

These results are notable due to the percentage of correctly classified instances, which in some cases have reached the optimum value of 100%. However, it worth to mention that binary classification is the easiest to solve as compared

Table 1: Open source tools available to implement Deep learning.

Software	**Developer**	**website**
Computational Network Toolkit CNTK	Microsoft Research	cntk.ai
Deeplearning4j	Adam Gibson	deeplaerning4j.org
Caffe	Berkeley Vision and Learning Center	caffe.berkeleyvision.org
Theano	University of Montreal's LISA Group	deeplearning.net/software/theano
Thorch	Ronan Collobert, Kuray Kavukcuoglu and Clement Farabet	torch.ch
Tensor Flow	Google Brain Team	tensorflow.org
WEKA	Machine Learning group at the University of Waikato	cs.waikato.ac.nz/ml/weka/

with the treatment of multi-class attributes. Then, a future study will be to observe the behaviour of multi-class classification applying Deep Learning with these datasets.

It is important to highlight that both the cross-validation and the $2/3 - 1/3$ sampling criteria are well established in AI for the classification of large volumes of data. Therefore, being these four datasets of reasonable dimensions, the prevalent criteria was statistical for RS, which was justified in reason that it represents the equilibrium point for a given frequency or the data concentration gave a distribution [4].

According to [12], in many applications, it can be advantageous to combine the best of two approaches, in this case: assembled classifiers and deep learning. Assembled classifiers is an approach based on successive results refinements obtained with standard classification algorithms and Dl4JMlpClassifier is an algorithm based on advantages of ANN multilayer perceptron.

5 Conclusion and future works

In these previous results can be appreciated that the binary classification by combining assembled algorithms with a Deep Learning algorithm give satisfactory results. Then, it can be speculated that the results in multi-class classifica-

Table 2: Datasets used in the experiments and their characteristics.

Dataset	**instances**	**attributes**	**classes**
car-evaluation	1728	6	4
Daily-days	1500	4	2
pap-smear	917	20	7
Blood-transfusion	748	5	2

Table 3: Highest results for each dataset.

Dataset	Classifier	Sampling	%
Car-evaluation	RandommCommitee+Dl4JMlpClassifier	RS	95.54
Daily-days	RandomCommitee+Dl4JMlpClassifier	2/3-1/3	100
	RandomizableFilteredClassifier + Dl4JMlpClassifier	2/3-1/3	100
	AdaBoost+ Dl4JMlpClassifier	RS	100
	Bagging + Dl4JMlpClassifier	RS	100
	RandommCommitee+Dl4JMlpClassifier	RS	100
Pap-Smear	RandommSubSpace+Dl4JMlpClassifier	RS	95.54
Blood-transfusion	Multischeme + Dl4JMlpClassifier	RS	81.57

tions may be acceptable, i.e., in a range not less than 80% of correctly classified instances. As future work we propose:

– To make experiments with the algorithm combinations of Table 3, but with multi-class datasets.

Any person, with basic knowledge of programming and computer skills, can take their first steps and experiment in Deep Learning with the data processing tools such as those listed in Table 1; there is also a variety of public domain datasets to carry out experiments such as those presented in the previous section; In addition to the above, there is information in books, scientific and dissemination journals, tutorials, and videos that indicate step by step how to implement a deep ANN.

Finally, as an additional comment: we would like to mention that someone can also experience with Deep Learning from a digital artistic expression perspective and make surrealist image designs such as those shown in Figure 3, because as long as AI continues on its way to make an intelligent agent able to think and feel like a human, a variety of algorithms have been developed to explore and exploit the benefits offered by Deep Learning, for example, generate artistic representations in digital format. A tool that allows you to create this type of

art is “Deep Dream Generator”, available as shareware for free on the website: https://deepdreamgenerator.com/

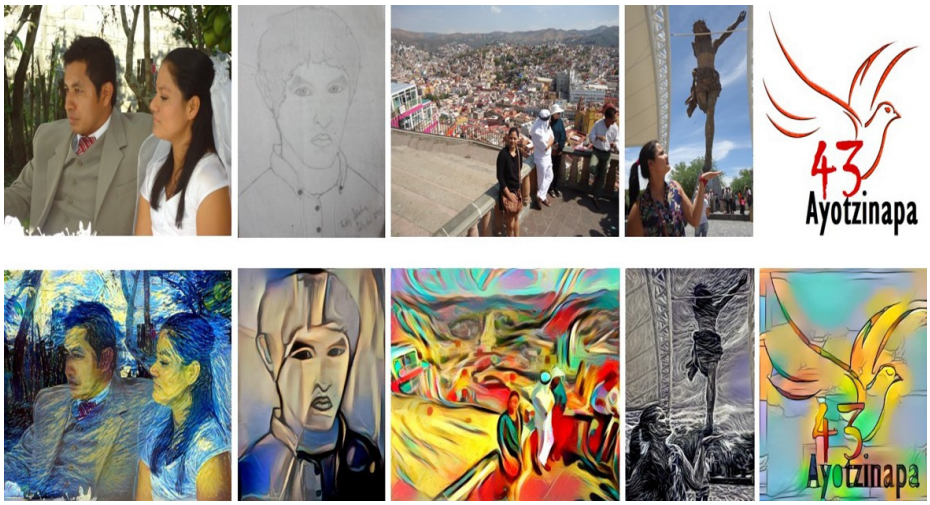


Fig. 3: Input and output images after processing with deep learning algorithms.

Acknowledgement

Authors acknowledge to the National Technological of Mexico, TECNM, and to the council of science and techonology, CONACYT, for the support given to realise this research. Authors acknowledge Editors and Reviewers for their constructive comments to produce the final version of this manuscript.

References

1. Martín Abadi, Ashish Agarwal, Paul Barham, Eugene Brevdo, Zhifeng Chen, Craig Citro, Greg S Corrado, Andy Davis, Jeffrey Dean, Matthieu Devin, et al. Tensorflow: Large-scale machine learning on heterogeneous distributed systems. *arXiv preprint arXiv:1603.04467*, 2016.
2. Itamar Arel, Derek C Rose, Thomas P Karnowski, et al. Deep machine learning-a new frontier in artificial intelligence research. *IEEE computational intelligence magazine*, 5(4):13–18, 2010.
3. Arthur Asuncion and David Newman. Uci machine learning repository, 2007.
4. DC Baird. Experimentación. una introducción a la teoría de mediciones y al diseño de experimentos, 1998.
5. Alejandro Cárdenas Cardona. *Inteligencia artificial, métodos bio-inspirados: un enfoque funcional para las ciencias de la computación.* PhD thesis, Universidad Tecnológica de Pereira. Facultad de Ingenierías Eléctrica , 2012.

6. Jan Jantzen and George Dounias. Analysis of pap-smear image data. In *Nature-Inspired Smart Information Systems 2nd Annual Symposium*. NiSIS, 2006.
7. Yann LeCun, Yoshua Bengio, and Geoffrey Hinton. Deep learning. *nature*, 521(7553):436, 2015.
8. Warren S McCulloch and Walter Pitts. A logical calculus of the ideas immanent in nervous activity. *The bulletin of mathematical biophysics*, 5(4):115–133, 1943.
9. Nils J Nilsson. *Principles of artificial intelligence*. Morgan Kaufmann, 2014.
10. Anderson Rocha and Siome Klein Goldenstein. Multiclass from binary: Expanding one-versus-all, one-versus-one and ecoc-based approaches. *IEEE Transactions on Neural Networks and Learning Systems*, 25(2):289–302, 2013.
11. Ingrid Russell and Zdravko Markov. An introduction to the weka data mining system. In *Proceedings of the 2017 ACM SIGCSE Technical Symposium on Computer Science Education*, pages 742–742. ACM, 2017.
12. Jürgen Schmidhuber. Deep learning in neural networks: An overview. *Neural networks*, 61:85–117, 2015.
13. Ian H Witten, Eibe Frank, Mark A Hall, and Christopher J Pal. *Data Mining: Practical machine learning tools and techniques*. Morgan Kaufmann, 2016.
14. I-Cheng Yeh, King-Jang Yang, and Tao-Ming Ting. Knowledge discovery on rfm model using bernoulli sequence. *Expert Systems with Applications*, 36(3):5866–5871, 2009.