DOI: 10.20396/revpibic2720193040

Fundamentals of Deep Neural Networks – Application Based in Convolutional Networks.

Rafael Gonçalves*, Romis Attux.

Abstract
Deep neural networks are becoming the state-of-the-art in the field of machine learning, especially because of their notable performance in a variety of current problems. Among these algorithms we can cite convolutional neural networks as one of the most powerful models for image classification problems. This research aimed to perform a theoretical study on deep neural networks, as well as to implement a convolutional neural network applied to the problem of handwritten digits classification.

Key words:
Deep Neural Networks, Machine Learning, Computational Inteligence.

Introduction

Artificial neural networks (ANN) are non-linear and adaptative computing systems originally inspired in biological neural networks present in animal brains. Especially with the advent of deep neural networks, this became an important paradigm in the field of machine learning (ML) and is used to solve a variety of current problems.

In this context, this research has two objectives, of which the first was to study the theory of deep neural networks and the second was to implement one of these algorithms – namely, a convolutional neural network, CNN – applied to a problem of handwritten digits classification (MNIST dataset[2]).

Results and Discussion

The study of deep learning networks was carried out with a discussion about the initial chapters of the book “Deep Learning”[1]. This stage was fulfilled with reading and discussion and with the implementation of a shallow network - multilyer perceptron (MLP) - and a learning algorithm - stochastic gradient descent (SGD) with error backpropagation (BP) for computation of the derivatives.

The second part of the research was done with the implementation of ANN’s for image classification applied to the dataset[2]. The hyperparameters for the models were chosen with the use of cross validation (CV). Both implemented architectures, MLP and CNN, had final models with accuracy achieving rates superior to 94% as seen in Chart 1. The ANN’s were implemented in Python programming language with the use of the PyTorch library and its source code is avaiable in the code repository[3].

Chart 1. Comparison between models.

Model	Accuracy
MLP	0.9490
CNN	0.9726

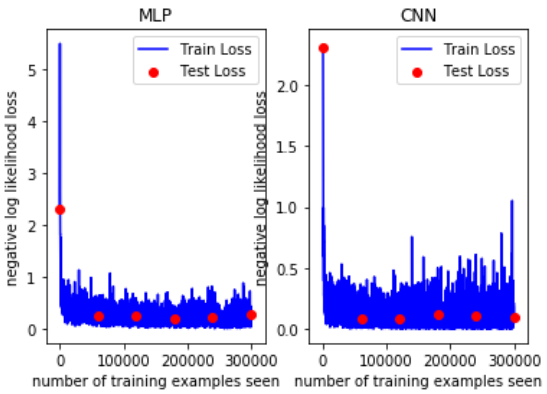

Image 1. Learning curve for each model (MLP on the left and CNN on the right).

Conclusions

The student was able to study the fundamentals of deep neural networks (basic concepts of linear algebra, probability and information theory, numerical computation and machine learning) and is able to understand and implement simple deep neural networks. With the implementation of the ANN’s the student was able to develop important algorithms of deep learning (such as BP and SGD) and apply studied practices (such as the use of CV for choosing hyperparameters and the use of a learning curve to analyse visually if there is underfit or overfit in the training).

Acknowledgement

The student wishes to express his gratitude to Prof. Romis Attux for all the encouragement and support and also thanks PIBIC/CNPq/Unicamp for the financial support.

[1] Goodfellow, I.; Bengio, Y.; Courville, A. *Deep Learning.* **2016**.

[2] LeCun, Y. *The MNIST Database of Handwritten Digits.* Available at: http://yann.lecun.com/exdb/mnist/. Last acessed in: 07/07/2019.

[3] Gonçalves, R. *MNIST nn.* Avaiable at: https://github.com/RafaelGoncalves8/mnist_nn/. Last acessed in: 07/07/2019.