



Universitat Autònoma de Barcelona

Degree Thesis

---

# ALTERNATIVES TO THE STEEPEST GRADIENT DESCEND IN NEURAL NETWORK TRAINING

---

Author:  
Gerard Lahuerta Martín

Supervisor:  
Dr. Lluís Alsedà Soler

A THESIS SUBMITTED IN FULFILMENT OF THE  
REQUIREMENTS FOR THE DEGREE OF COMPUTATIONAL  
MATHEMATICS AND DATA ANALYTICS IN THE

SCIENCE FACULTY

June 2024



## Declaration of Authorship

## Abstract

## Acknowledgements

I would like to thank all my family that gave me support while doing this thesis, the degree and all my life.

I am especially grateful to my mother, Maria Montserrat for encouraging and advising me to face the challenge of doing this thesis and my career.

I also appreciate the cheering of my colleagues as well as my girlfriend, who supports me too.

Finally, I want to give a special mention to Dr. Lluís Alsedà for guiding me while doing the thesis as well as for being an inspiration.

Thanks for all.

# Contents

<b>Declaration of Authorship</b>	<b>2</b>
<b>Abstract</b>	<b>3</b>
<b>Acknowledgements</b>	<b>4</b>
<b>Preface</b>	<b>6</b>
<b>Introducción</b>	<b>7</b>
<b>1 Neural Networks</b>	<b>8</b>
1.1 Operation of Neural Networks . . . . .	9
1.1.1 kjjkgj . . . . .	10
1.2 Cost Function . . . . .	11
<b>2 Optimizers</b>	<b>12</b>
2.1 Types of Optimizers . . . . .	13
2.2 Deterministic Optimizers . . . . .	14
2.3 Heuristic Optimizers . . . . .	15
2.4 Exploid and Explore philosophy . . . . .	16
<b>3</b>	<b>16</b>

## Preface

Since in the early starts of my degree in *Matemàtiques Computacionals i Analítica de Dades*, when one of my professors explains the Gradient Descend method to optimise functions, in every subject about Artificial Intelligence or Neural Networks the method used to obtain the best parameters is always the Steepest Gradient Descend.

This made me think about the nonexistence of an alternative method to optimise Neural Network's parameters, and this thought made me uncomfortable, causing me to reflect on the subject.

No professors could give me an answer to why this is the best way and why it will give the best configuration of the Neural Network's weights, they just told me that this is the fastest way to arrive to the minimum of the *cost function*.

At first thought, the idea that a method whose only purpose is to minimise a value that represents how bad the model predicts seems a good proposition.

Nevertheless, there is no evidence that the results given by the method are the best ones.

I was feeling like the main idea was optimising a function without any information about it, with one algorithm that may or may not give the best result, only a relatively good one; and this way of thinking does not match with me.

This is why I am trying to change the ideology that the Steepest Gradient Descend is the best way to train Neural Network to obtain the best prediction; researching and testing different alternatives of optimization algorithms to train Neural Networks.

The final objective to this thesis is to obtain a better way to arrive to the best configuration; in other words, obtain an algorithm whose results have less error rate in Neural Network predictions (than the one obtained by Steepest Gradient Descend).

## Introducción



## 1 Neural Networks

## 1.1 Operation of Neural Networks

### 1.1.1 kjjkgj

## 1.2 Cost Function

## 2 Optimizers

## 2.1 Types of Optimizers

## 2.2 Deterministic Optimizers

## 2.3 Heuristic Optimizers



## **2.4 Exploid and Explore philosophy**

### **3**