REVIEW ARTICLE

Towards a Mathematical Understanding of Neural Network-Based Machine Learning: What We Know and What We Don't

Weinan E[1,2,3], Chao Ma[4], Lei Wu[2] and Stephan Wojtowytsch[2,i]

[1] *Department of Mathematics, Princeton University, Princeton, NJ 08544, USA.*
[2] *Program in Applied and Computational Mathematics, Princeton University, Princeton, NJ 08544, USA.*
[3] *Beijing Institute of Big Data Research, Beijing 100871, China.*
[4] *Department of Mathematics, Stanford University, Stanford, CA 94305, USA.*

Abstract. The purpose of this article is to review the achievements made in the last few years towards the understanding of the reasons behind the success and subtleties of neural network-based machine learning. In the tradition of good old applied mathematics, we will not only give attention to rigorous mathematical results, but also the insight we have gained from careful numerical experiments as well as the analysis of simplified models. Along the way, we also list the open problems which we believe to be the most important topics for further study. This is not a complete overview over this quickly moving field, but we hope to provide a perspective which may be helpful especially to new researchers in the area.

1 Introduction

Neural network-based machine learning is both very powerful and very fragile. On the one hand, it can be used to approximate functions in very high dimensions with the

---

[i]Corresponding author. *Email addresses:* `weinan@math.princeton.edu` (W. E), `chaoma@stanford.edu` (C. Ma), `leiwu@princeton.edu` (L. Wu), `stephanw@princeton.edu` (S. Wojtowytsch)

efficiency and accuracy never possible before. This has opened up brand new possibilities in a wide spectrum of different disciplines. On the other hand, it has got the reputation of being somewhat of a "black magic": Its success depends on lots of tricks, and parameter tuning can be quite an art. The main objective for a mathematical study of machine learning is to

1. explain the reasons behind the success and the subtleties, and

2. propose new models that are equally successful but much less fragile.

We are still quite far from completely achieving these goals but it is fair to say that a reasonable big picture is emerging.

The purpose of this article is to review the main achievements towards the first goal and discuss the main remaining puzzles. In the tradition of good old applied mathematics, we will not only give attention to rigorous mathematical results, but also discuss the insight we have gained from careful numerical experiments as well as the analysis of simplified models.

At the moment much less attention has been given to the second goal. One proposal that we should mention is the continuous formulation advocated in [33]. The idea there is to first formulate "well-posed" continuous models of machine learning problems and then discretize to get concrete algorithms. What makes this proposal attractive is the following:

- many existing machine learning models and algorithms can be recovered in this way in a *scaled* form;

- there is evidence suggesting that indeed machine learning models obtained this way is more robust with respect to the choice of hyper-parameters than conventional ones (see for example Fig. 5 below);

- new models and algorithms are borne out naturally in this way. One particularly interesting example is the maximum principle-based training algorithm for ResNet-like models [58].

However, at this stage one cannot yet make the claim that the continuous formulation is the way to go. For this reason we will postpone a full discussion of this issue to future work.

1.1 The setup of supervised learning

The model problem of supervised learning which we focus on in this article can be formulated as follows: given a dataset $S = \{(x_i, y_i = f^*(x_i)), i \in [n]\}$, approximate f^* as accurately as we can. If f^* takes continuous values, this is called a regression problem. If f^* takes discrete values, this is called a classification problem.