Short Course on Scientific Deep Learning

Prof. Tan Bui-Thanh
Oden Institute for Computational Engineering and Sciences
University of Texas at Austin

11-13 June 2024

Abstract

This course discusses various aspects of an important class of scientific machine learning, namely scientific deep learning. To that end, the course will cover the following topics (with certain depths including proofs of relevant theorems): 1) statistical machine learning framework with deep learning, 2) universal approximation theorem for ReLU network 3) Back-propagation for gradient computation. If time permits, we will cover the gradient vanishing problem in deep learning and ResNet as an approach to overcome these problems.

Outline of the course

This short course limits itself on a few important topics in foundational machine learning. We shall focus on supervised learning with deep learning. The following are the topics of the course:

1. **Statistical machine learning framework with deep learning**. In the first part, we start with the fundamental supervised problem of learning from data. We shall use linear least square problem as a prototype example and then immediately generalize it to an abstract setting of risk minimization versus empirical risk minimization. This setting will allow us to introduce and discuss several concepts of supervised learning including hypothesis spaces, (stochastic) gradient descent methods, and bias-variance trade-off. Prerequisite: Audiences are familiar with basics of least square method, optimization, and probability.

2. **Why does deep learning work**. The second part discuss the hypothesis space generated by deep neural network. We shall limit ourselves to the universality of ReLU network, by far the most popular one, and show why such a network can approximate any function/task to any desired accuracy. Prerequisite: Audiences are familiar with basics of function spaces (space of continuous function with uniform norm), convergence, and interpolation.

3. **Backpropagation and gradient vanishing**. The third part discusses what backpropagation means and its equivalence to the adjoint/Lagrangian approach as an efficient way to compute the gradient of deep neural networks. Prerequisite: Audiences are familiar with basics of inequality-constrained optimization using Lagrangian/adjoint approach (we will quickly review it), and directional derivatives (with respect to vectors/matrices).

Course Schedule

Guest Lecturer: **Prof. Tan Bui-Thanh**
Welcome and opening notes: **Prof. Laura Mainini**.

- **Lecture 1** Tuesday 11 June 2024 | 1000-1300 | DIMEAS Aula Ferrari
- **Lecture 2** Tuesday 11 June 2024 | 1400-1700 | DIMEAS Aula Ferrari
- **Lecture 3** Thursday 13 June 2024 | 1000-1300 | DIMEAS Aula Ferrari
- **Lecture 4** Thursday 13 June 2024 | 1400-1700 | DIMEAS Aula Ferrari

Lecture Notes

Typed and more comprehensive lecture notes will be provided before the class.