Advanced Algorithms And Parallel Programming
Theory

Christian Rossi

Academic Year 2023-2024

Abstract

Neural networks have matured into flexible and powerful non-linear data-driven models, effectively tackling complex tasks in both science and engineering. The emergence of deep learning, which utilizes neural networks to learn optimal data representations alongside their corresponding models, has significantly advanced this paradigm.

In the course, we will explore various topics in depth. We will begin with the evolution from the Perceptron to modern neural networks, focusing on the feedforward architecture. The training of neural networks through backpropagation and algorithms like Adagrad and Adam will be covered, along with best practices to prevent overfitting, including cross-validation, stopping criteria, weight decay, dropout, and data resampling techniques.

The course will also delve into specific applications such as image classification using neural networks, and we will examine recurrent neural networks and related architectures like sparse neural autoencoders. Key theoretical concepts will be discussed, including the role of neural networks as universal approximation tools, and challenges like vanishing and exploding gradients.

We will introduce the deep learning paradigm, highlighting its distinctions from traditional machine learning methods. The architecture and breakthroughs of convolutional neural networks (CNNs) will be a focal point, including their training processes and data augmentation strategies.

Furthermore, we will cover structural learning and long-short term memory (LSTM) networks, exploring their applications in text and speech processing. Topics such as autoencoders, data embedding techniques like word2vec, and variational autoencoders will also be addressed.

Finally, we will discuss transfer learning with pre-trained deep models, examine extended models such as fully convolutional CNNs for image segmentation (e.g., U-Net) and object detection methods (e.g., R-CNN, YOLO), and explore generative models like generative adversarial networks (GANs).

Contents

CHAPTER 1

Deep learning

1.1 Introduction

Definition (*Machine Learning*)**.** A computer program is considered to learn from experience E with respect to a specific class of tasks T and a performance measure P if its performance at tasks in T, as measured by P, improves with experience E.

Given a dataset $\mathcal{D} = x_1, x_2, \ldots, x_N$, Machine Learning can be broadly categorized into three types:

- *Supervised learning*: in this type of learning, the model is provided with desired outputs $t_1, t_2, \ldots, t_N$ and learns to produce the correct output for new input data. The primary tasks in supervised learning are:
  - *Classification*: the model is trained on a labeled dataset and returns a label for new data.
  - *Regression*: the model is trained on a dataset with numerical values and returns a number as the output.
- *Unsupervised learning*: here, the model identifies patterns and regularities within the dataset $\mathcal{D}$ without being provided with explicit labels. The main task in unsupervised learning is:
  - *Clustering*: the model groups similar data elements based on inherent similarities within the dataset.
- *Reinforcement learning*: in this approach, the model interacts with the environment by performing actions $a_1, a_2, \ldots, a_N$ and receives rewards $r_1, r_2, \ldots, r_N$ in return. The model learns to maximize cumulative rewards over time by adjusting its actions.

1.1.1 Deep Learning

Deep Learning, a subset of Machine Learning, focuses on utilizing large datasets and substantial computational power to automatically learn data representations. In certain cases, traditional classification may fail due to the presence of irrelevant or redundant features in the dataset.

Deep Learning addresses this issue by learning optimal features directly from the data, which are then used by Machine Learning algorithms to perform more accurate classifications. Essentially, Deep Learning involves learning how to represent data in a way that improves the performance of Machine Learning models.