UNIVERSITÀ
DEGLI STUDI DI MILANO-BICOCCA

SYLLABUS DEL CORSO

Statistical Mechanics of Neural Networks

2526-2-F9102Q027

Aims

Neural Networks (NNs) are paradigmatic complex systems, characterised by an emergent collective behaviour. Statistical mechanics provides the natural theoretical framework to describe the phenomenology of complex physical systems. In view of this connection, many theoretical tools and concepts introduced in statistical physics have been later proven very useful to analyze the behaviour of NNs. The main aim of this course is to bring the students to understand some of these approaches, critically assessing their relevance in the framework of contemporary research directions. A second goal of the course is familiarize with example of applications of NN and Machine Learning to simulate dynamics of complex systems, taking biomolecules as paradigmatic examples.

Contents

First, we shall introduce some key results and models of statistical mechanics, starting from an information theory and statistical computing standpoint. Then, we shall apply these tools to the analysis of different types of NNs.
Depending on time availability, student participation and interest, the course will also cover some selected special topics (see also teaching form, below), which illustrate how NNs are being employed to investigate the dynamics of complex physical systems, using macromolecules of biological relevance as paradigmaric examples.

Detailed program

Chapter 1: From Information Theory to Statistical Mechanics

Bayes’s theorem. Prior and posterior distributions.
Shannon information content and entropy
Inference & Modelling: Maximum likelihood and Maximum Entropy approach to data modelling
Statistical Mechanics from Maximum Entropy Principle: Derivation of the Gibbs-Boltzmann’s statistical ensembles
Ising Model and its mean-field solution. Phase transitions in the Ising Model
Stochastic dynamics in complex systems. Intrinsic manifold in the data and configuration space.

Chapter 2: Neural Networks (NNs)

The perceptron
Single neuron as a classifier
Deep networks and their training. The overfitting problem
Taming the overfitting problem by the Maximum-Entropy approach
Learning as an inference process. Generalization

Chapter 3: Applications of statistical mechanics to NNs

Associative memory and Hopfield networks. Connection with the Ising Model
Capacity of a Hopfield model
Gardner volume and storage capacity
Boltzmann’s machines. The role of inner layers and analogy with auxiliary stochastic fields
Stochastic theory of Diffusive Models

Chapter 4: Learning

A case for supervised learning: interpolation
A case for unsupervised learning: clustering
Reinforcement learning.

SPECIAL TOPICS:

Machine Learning estimate of the dimensionality of the intrinsic manifold
Inferring data structure from stochastic dynamics: diffusion maps
Uncharted exploration of the intrinsic manifold
Interpolating free-energy landscapes with NNs.
Clustering in configuration space: k-means vs density-peak clustering
Reinforcement learning for enhanced sampling methods.

Prerequisites

A rudimental background in classical statistical mechanics and in particular of Boltzmann - Gibbs’ theory of equilibrium statistical ensembles is welcome but will not assumed. All the essential concepts will be provided in the course. Rudimental background in multi-dimensional calculus will be assumed.

Teaching form

The core topics of the course will be covered in conventional lectures on the black board. The selected special topics will be discussed in a mixed mode, involving both students and lecturer. Students will be divided in small groups, each presenting to the rest of the class an overview one or more special topics. The lecturer will provide the general introduction and connection with the topic discussed with rest of the course. The lecturer will also lead a discussion with the students aiming at assessing the strength and limitations of each approach.

Textbook and teaching resource

Haiping Huang, Statistical Mechanics of Neural Networks, Springer 2021. ISBN: 978-981-16-7570-6

David J.C. MacKay, Information Theory, Inference, and Learning Algorithms, Cambridge University Press 2003.

John Hertz, A. K. & Palmer, R. G. Introduction to the theory of neural computation. isbn: 9780201515602 (CRC Press, 2018).

Christopher M. Bishop: Pattern recognition and Machine Learning, Springer ISBN-10: 0-387-31073-8. ISBN-13: 978-0387-31073-2.

Katerina Gratsea, V. K. & Lewenstein, M. Storage properties of a quantum perceptron2021. https://arxiv.org/abs/2111.08414.

Semester

First semester, second year

Assessment method

The grade will be assigned on the basis of a final oral examination, also taking in consideration the student's contribution to the special topic sessions and related discussion.

Office hours

The lecturer will be available any time, previous arranging the time and date by email.

Sustainable Development Goals

INDUSTRY, INNOVATION AND INFRASTRUCTURE