UNIVERSITÀ
DEGLI STUDI DI MILANO-BICOCCA

COURSE SYLLABUS

Statistical Mechanics of Neural Networks

2324-2-F9102Q027

Obiettivi

Da un lato, le Reti Neurali (RN) sono sistemi complessi paradigmatici, caratterizzati da un comportamento collettivo emergente. Dall'altro, la meccanica statistica fornisce il quadro teorico naturale per descrivere la fenomenologia dei sistemi fisici complessi. Alla luce di questo collegamento, molti strumenti e concetti teorici introdotti nella fisica statistica sono stati successivamente applicati per analizzare il comportamento delle RN. Lo scopo di questo corso è quello di portare gli studenti a comprendere alcuni di questi approcci, valutandone criticamente la rilevanza in il quadro delle direzioni di ricerca contemporanee.

Contenuti sintetici

In primo luogo, verrsnno introdotti alcuni modelli paradigmatici e risultati chiave della meccanica statistica, a partire da un punto di vista della teoria dell'informazione e del calcolo statistico. Quindi, applicheremo questi strumenti all'analisi di diversi tipi di RN.
A seconda della disponibilità di tempo, della partecipazione e dell'interesse degli studenti, il corso coprirà anche alcuni temi speciali selezionati (vedi anche relativo modulo didattico), che illustrano come le RN possono essere impiegate per studiare la dinamica di sistemi fisici complessi.

Programma esteso

Capitolo 1: Dalla teoria dell'informazione alla meccanica statistica

Il teorema di Bayes. Distribuzioni a priori e a posteriori.
Informazione di Shannon ed entropia

Inferenza e modellizzazione: l'approccio di maximum likelyhood e maximum entropy alla modellazione dei dati
Derivazione della meccanica statistica dal principio di massima entropia: derivazione degli insiemi statistici di Gibbs-Boltzmann. Modello di Ising e sua risoluzione a livello di campo medio. Transizioni di fase nel modello di Ising. Dinamica stocastica nei sistemi complessi. Intrinsic manifold nello spazio dei dati e delle configurazioni.

Capitolo 2: Reti neurali (RN)

Il percettrone
Singolo neurone come classificatore
Reti profonde e loro allenamento. Il problema dell'overfitting
Affrontare il problema dell'overfitting con l'approccio Maximum-Entropy
Apprendimento come processo di inferenza. Generalizzazione

Capitolo 3: Applicazioni della meccanica statistica alle RN

Memoria associativa e reti di Hopfield. Collegamento con il modello di Ising
Capacità di un modello Hopfield
Volume Gardner e capacità di storing
Le macchine di Boltzmann

Capitolo 4: Apprendimento

Un'applicazione di apprendimento supervisionato: l'interpolazione
Un'applicazione di apprendimento non supervisionato: il clustering
Reinforcement learning

TEMI SPECIALI:

Stima di Machine Learning della dimensionalità della varietà intrinseca di un sistema multidimensionale
Inferire la struttura dei dati dalla dinamica stocastica: mappe di diffusione
L' Uncharted exploration problem
Interpolazione di paesaggi a energia libera con RN.
Clustering nello spazio di configurazione: k-means vs density peak clustering
Enhanced sampling da reinforcement learning.

Prerequisiti

Un background rudimentale nella meccanica statistica classica e in particolare nella teoria di Boltzmann-Gibbs degli insiemi statistici all'equilibrio è ben accetto ma non sarà assunto. Tutti i concetti essenziali verranno forniti durante il corso. Sono richieste competenze di base di calcolo a molte variabili

Modalità didattica

Gli argomenti centrali del corso saranno trattati in lezioni frontali alla lavagna. Gli argomenti speciali selezionati saranno discussi in modalità mista, coinvolgendo sia gli studenti che il docente. Gli studenti saranno divisi in piccoli gruppi, ognuno dei quali presenterà al resto della classe una panoramica su uno o più argomenti speciali. Il docente fornirà l'introduzione generale e il collegamento con l'argomento trattato con il resto del corso. Il docente condurrà inoltre una discussione con gli studenti volta a valutare i punti di forza ei limiti di ciascun approccio.

Materiale didattico

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.

Periodo di erogazione dell'insegnamento

Primo semestre, secondo anno

Modalità di verifica del profitto e valutazione

Il voto verrà assegnato sulla base di una prova orale finale, tenuto anche conto del contributo dello studente alle sessioni tematiche e relative discussioni.

Orario di ricevimento

In qualsiasi momento, previo accordo con il docente via email

Sustainable Development Goals

IMPRESE, INNOVAZIONE E INFRASTRUTTURE