



Programma del corso “Reti Neurali” A.A. 2016/2017 (3 cfu)

Docente: Adriano Barra

Dipartimento di Matematica e Fisica “Ennio De Giorgi”, Università del Salento

Finalità del corso:

Lo scopo ultimo del corso è condividere con lo studente i concetti salienti e, parimenti, fornire allo studente gli strumenti cardine, affinché questi possa continuare autonomamente la sua crescita culturale nell'ambito dell'Intelligenza Artificiale: tale trasferimento di informazione cercherà di essere supplito sia da un punto di vista formale/matematico, (e.g. mostrare durante il corso metodi chiari per impostare un problema di pertinenza e risolverlo opportunamente), sia logico/deduttivo, (e.g. capire cosa chiedersi -che sia ragionevole- per affrontare un dato problema, pertinente al tema del corso ovviamente). A tal fine il programma del corso è diviso in tre sezioni principali.

La prima sezione serve per assicurarci di condividere una conoscenza scientifica di base (pre-requisito ovviamente necessario per muovere insieme i primi passi nel percorso delle reti neurali verso l'Intelligenza Artificiale): a tal proposito, una conoscenza della Meccanica Statistica e dei Processi Stocastici certamente agevolerà lo studente. Nella prima sezione del corso rivedremo insieme quindi alcuni argomenti fondamentali (per questo corso in particolare) di queste due discipline.

La seconda sezione introduce invece importanti modelli di Meccanica Statistica, fondamentali per una successiva impostazione d'analisi -e relativa indagine- inerentemente il funzionamento delle reti neurali, e sviluppa gli opportuni metodi matematici necessari alla loro descrizione ed alla comprensione della loro fenomenologia.

L'ultima sezione è invece completamente dedicata alle reti neurali: dopo una succinta descrizione (sempre in termini matematici) dei meccanismi cardine inerenti il neurone e la propagazione d'informazione tra neuroni, si costruiranno “reti di neuroni” (in altre parole si spiegheranno “cosa sono” -matematicamente parlando- queste reti neurali) e se ne studieranno le proprietà emergenti (cioè non immediatamente deducibili guardando al comportamento del singolo neurone), persistendo in una prospettiva di meccanica statistica.

Nello specifico, proveremo a vedere come queste reti siano in grado di apprendere ed astrarre guardando esempi suppliti dal mondo esterno e come, successivamente, queste usino quanto appreso per rispondere opportunamente, qualora stimolate, al mondo esterno. Capiremo inoltre come queste a volte possano sbagliare, e perchè.

Idealmente alla fine del corso lo studente dovrebbe essere in grado di poter proseguire autonomamente nell'approfondimento di tale disciplina, e di fruirne di conseguenza.

Pre-requisiti per seguire il corso:

Nessuno (oltre i corsi della triennale), ed il corso è pensato tanto per Matematici quanto per Fisici. Tuttavia non si nasconde che una conoscenza di base della Meccanica Statistica aiuti significativamente. Non guastano d'altronde un'infarinatura di Processi Stocastici, Teoria della Probabilità, Sistemi Dinamici e la conoscenza di linguaggi di programmazione (e.g. C) e/o di programmi per il calcolatore (e.g. Mathematica) [questi ultimi due saranno veramente di impiego minimale, e -all'atto pratico- marginalmente necessari alla comprensione del corso nel suo insieme, quindi la loro totale ignoranza non è assolutamente grave in questo frangente].

Estremi del docente:

Dr. Adriano Barra, Stanza 455 (Dipartimento di Matematica e Fisica “Ennio De Giorgi”)

Email: adriano.barra@gmail.com, adriano.barra@unisalento.it

Web: www.adrianobarra.com

Programma di massima del corso

1. Richiami di Meccanica, Probabilità e Meccanica Statistica

- Problema del Time-Reversal con esempi di PDE di interesse in Fisica (e.g. D'Alambert & Fourier).
- Sintesi della formulazione Lagrangiana ed Hamiltoniana della meccanica classica.
- Legge dei Grandi Numeri e Teoremi del Limite Centrale.
- Il modello di Ehrenfest: analisi sia della statica che della dinamica, studio dell'entropia.
- L'approccio di Gibbs ed il metodo della “distribuzione più probabile”.
- L'approccio di Jaynes: una prospettiva del tutto inferenziale.
- Equivalenza (e non) tra le entropie di Gibbs e Shannon.
- La temperatura come rumore veloce: Dal random walk all'equazione del calore*.

2. Fenomeni critici e transizioni di fase: verso i sistemi complessi

- Modello di Ising: Approssimazione di campo medio, transizioni di fase e rotture di simmetrie.
- Proprietà degli stati puri: Fattorizzazione delle funzioni di correlazione e clustering.
- Metodo del “punto di sella”, equazioni Dobrushin-Ruelle-Lanford e disuguaglianze convesse.
- Metodo del campo di cavità, tecnica di Hamilton-Jacobi ed analisi di Fourier.
- La master equation: Hamiltoniana come funzione di Lyapunov ed il Bilancio Dettagliato.
- L'inverse-problem nel caso più semplice (un sola specie di Curie-Weiss & via log-likelihood).
- Introduzione ai sistemi complessi: Misure quenched ed annealed, overlap e repliche*.
- Ageing, rottura del time-translational-invariance e fenomenologia dei trap models*.
- Spettro di una catena di Markov semplice e frustrata: Rilassamento dei modi normali*.
- Il modello di Sherrington-Kirkpatrick: Analisi con “replica trick”, Soluzione RS e crisi entropica*.
- Analisi mediante “replica trick”: Approssimazione 1RSB e Teoria di Parisi*.

3. Elementi di Intelligenza Artificiale: le reti neurali

- Il quadro storico nel quale è nata l'AI: tanti contributi da diverse discipline.
- La “cable theory” di Hodgkin-Huxley ed il neurone “integrate & fire” di Stein.
- Il riflesso condizionato di Pavlov mediante dinamica stocastica à la Glauber.
- La memoria associativa e le reti neurali: La proposta di Hebb.
- Il modello di Mattis, trasformazioni di gauge locali e storage di un bit di informazione.
- Il modello di Hopfield a basso carico con il metodo della log-constrained entropy.
- Il modello di Hopfield ad alto carico con il replica trick: Teoria di Amit-Gutfreund-Sompolinsky*
- Il Perceptrone di Rosenblatt e la critica di Minsky&Papert.
- Apprendimento delle Boltzmann machines: scenario studente-allievo.
- Apprendimento delle Boltzmann machines: scenario non supervisionato.
- Equivalenza tra retrieval in reti di Hopfield e learning in macchine di Boltzmann.

**questi argomenti saranno solo accennati nella versione del corso da 3 cfu*

Libri di testo consigliati

[Amit] D.J. Amit, Modeling Brain Functions, Cambridge Press (1985).

[Barra] A. Barra, Appunti specifici per questo corso (2017).

[Coolen] A.C.C. Coolen, R. Kuhn, P. Sollich, Theory of Neural Information Processing Systems, Oxford Press (2005).