

UNIVERSITÀ DEGLI STUDI DI TRIESTE
FACOLTÀ DI SCIENZE MATEMATICHE, FISICHE E NATURALI
CORSO DI LAUREA IN FISICA

Introduzione alla Teoria delle Reti Neurali (Codice: SM307 - 6 CFU)
(Marco Budinich, A.A. 2007 - 2008)

Cenni alle basi biologiche - il cervello, aree corticali, microcolonne, neuroni, dendriti e assoni, sinapsi. Membrana e potenziali del neurone, cenni allo sviluppo del segnale e analogia circuitale. Codifica dell'informazione nei segnali, plasticità delle sinapsi e regola di Hebb.

Modelli neuronali - il neurone di McCulloch e Pitts: similarità e differenze con i neuroni biologici. Apprendimento e sua interpretazione come minimizzazione del numero di errori. Il Perceptron: enunciazione del problema nello spazio degli input, problemi linearmente separabili, l'algoritmo di apprendimento ispirato alla regola di Hebb. Teorema del Perceptron e sua dimostrazione. Perceptron con funzione di trasferimento continua e confronto con il caso binario, esempio della tangente iperbolica.

Caratteristiche e limiti del neurone di McCulloch e Pitts - Growth function e dimensione di Vapnik Cernonenkis d_{VC} (VC-dimension). Caso del Perceptron: calcolo del numero delle dicotomie implementabili con un discriminante lineare; punti in posizione generale, definizione della capacità, relazione con d_{VC} . Approssimazione Gaussiana per il calcolo della capacità.

Reti "feed-forward" - definizione di reti feed-forward e teorema di approssimazione universale. Mancanza di un valido algoritmo di apprendimento e metodo costruttivo per la "grandmother cell representation". Back-propagation e ambiguità intrinseca nella rappresentazione nel livello intermedio, caso dell'XOR, condizione necessaria della rappresentazione fedele ed esempi basati sull'XOR. Algoritmi di apprendimento basati sulla back-propagation: vantaggi dell'uso della tangente iperbolica. Descrizione dettagliata dell'implementazione dell'algoritmo: parte forward, parte backward, criteri di stop, esecuzione batch e incrementale. Generalizzazione e sue caratteristiche, definizione del funzionale $g(f)$ e limite di Vapnik e Cervonenkis. Esempi di applicazione del limite per calcolare il numero di esempi nel Perceptron e in una rete feed-forward. Funzione d'errore alternativa per il Perceptron basata sulla teoria dell'informazione, capacità per uno strato di Perceptron binari e suo limite superiore, ruolo dell'architettura e dell'apprendimento, principi di apprendimento basati sull'informazione.

Apprendimento non supervisionato - neurone di Oja e sue proprietà con relative dimostrazioni. Analisi delle componenti principali e metodi di proiezione: proprietà e difetti. Cenni alla rete di Sanger per il calcolo delle componenti principali dei dati. Altra possibilità tramite una rete parzialmente supervisionata con gli ingressi (n-h-n). Rete WTA (pigliatutto), tassellazione di Voronoi, funzione minimizzata nell'apprendimento discussione del metodo. Reti non supervisionate con distanza definita anche nello spazio dei neuroni: mappe che preservano le distanze, reti di Kohonen e loro ispirazione biologica. Proprietà generali delle reti di Kohonen: conservazione delle distanze, riduzione dimensionale, ordinamento e adattamento, funzione minimizzata dall'apprendimento.

Memoria associativa - memoria associativa e memoria tradizionale dei computer: punti fissi e bacini di attrazione. Rete di neuroni di McCulloch e Pitts completamente interconnessi:

memorizzazione di un solo pattern. Random walk unidimensionale. Memorizzazione di più pattern, ipotesi sulla casualità dei bit di un pattern, stabilità del singolo bit. Capacità della memoria associativa calcolata con diverse definizioni della probabilità di errore sulla stabilità dei pattern memorizzati. Limite della stabilità dinamica a $m/n = 0.138$. Numero di bit memorizzati e numero di bit necessari per memorizzarli. Caso pratico: algoritmo per costruire una memoria associativa. Funzione di Lyapunov per la memoria associativa e necessità di avere pesi simmetrici.

Meccanica statistica e reti neurali - Cenni di meccanica statistica: distribuzione di Boltzmann, calcolo del valor medio di una quantità, funzione di partizione ed energia libera, identità formale fra entropia e informazione. Cenni ai sistemi di Ising con dinamica di Glauber; caso del singolo spin e di semplici sistemi paramagnetici e ferromagnetici con interazione costante. Relazione fra sistemi di Ising e memorie associative; discussione qualitativa del ruolo della temperatura, discussione qualitativa dei diversi tipi di soluzioni a $T = 0$ e $T > 0$; cenni ai vetri di spin (spin-glass). Cenni di teoria dell'informazione: quantità di informazione e entropia di una sorgente, esempi particolari, caso dell'entropia massima (distribuzione uniforme), cenni alla mutua informazione. La macchina di Boltzmann come generalizzazione delle varie reti studiate finora, introduzione di unità nascoste, regola di apprendimento derivata dalla minimizzazione della distanza delle distribuzioni e sua implementazione pratica. Simulazione dei sistemi con dinamica di Glauber e algoritmo di Metropolis. L'algoritmo di Simulated Annealing e sua generalità.

Visione - Cenni alle diverse componenti biologiche del sistema visivo umano e alle relative funzionalità. Ipotesi sull'interpretazione dei campi visivi nella corteccia: trasformate spaziali e modelli computazionali.

Programmi per informatici (per i quali non è in programma la parte di Meccanica Statistica)

1. Verifica della dipendenza da m (numero di esempi) della frazione di problemi casuali risolvibili dal Perceptron per 10 e 1000 ingressi.
2. Back-propagation su una rete feed-forward con 2 insiemi di esempi: uno per l'apprendimento e uno per il test. Apprendimento sull'insieme di esempi con criterio di stop basato sulla minimizzazione dell'errore sull'insieme di test. Uso di un numero di esempi sufficiente per ottenere una ragionevole probabilità di generalizzazione.
3. Verifica del numero di esempi necessari per una buona generalizzazione per un Perceptron basato sul limite inferiore alla generalizzazione di Vapnik e Cervonenkis.
4. Rete di Sanger per il calcolo delle componenti principali di una serie di dati.
5. Rete di Kohonen unidimensionale con dati in 1, 2 e 3 dimensioni.
6. Rete a due strati: il primo implementa una rete di Kohonen, il secondo una rete feed-forward con back-propagation.
7. Perceptron con funzione d'errore basata sulla distanza di Kullback-Leibler fra output e output desiderati.
8. Semplice memoria associativa con 10.000 neuroni.
9. Rete a due strati con un primo strato per il calcolo delle componenti principali (rete di Sanger) ed un secondo strato di memoria associativa che memorizza gli insiemi trasformati (per due persone).
10. Calcolare analiticamente un algoritmo di apprendimento per uno strato di h Perceptron binari basato sulla massimizzazione della mutua informazione fra input e output (per chi ama la teoria).
11. Memoria associativa con la dinamica di Glauber.
12. Programmazione del Simulated Annealing per un problema combinatorio, per esempio Perceptron con pesi discreti o con problemi irrisolvibili come l'XOR.
13. Altri programmi possono essere concordati con il docente.

Testi consigliati

Collocazione biblioteca Dipartimento di Fisica:

Griffith J.S., Mathematical Neurobiology, Academic Press, 1971, pp. x-162; I-224

Hertz J., Krogh A., Palmer R.G., Introduction to the Theory of Neural Computation, Addison-Wesley, 1991, pp. xxii-328; I-295

Muller B., Reinhardt J., Neural Networks - An Introduction, Springer Verlag, 1991, pp. xiv-266; I-296

Palmer S.E., Vision science: photons to phenomenology, The MIT Press, 2002, pp. xxii-810;

Biblioteca Psicologia: 01./2342

Sessioni ufficiali d'esame di: Introduzione alla Teoria delle Reti Neurali

11 gennaio 2008	12 giugno 2008	11 settembre 2008
18 febbraio 2008	10 luglio 2008	
20 gennaio 2009		

Tutti gli esami si svolgono alle ore 9:30 al Dipartimento di Fisica