Slide 1: Titolo

Slide 2: Cosa è una rete neurale

Slide 3: Storia con McCulloch, Pitts e Rosenblatt

Slide 4: Applicazioni

Slide 5: Sistema nervoso umano

Slide 6: Modello formale di un neurone artificiale

Slide 7: Attivazione e varie funzioni di attivazione

Slide 8: Architettura

Slide 9: Addestramento e paradigmi

Slide 10: Processi

Slide 11: Attività di apprendimento

Slide 12: Cosa sono le reti di Hopfield

Slide 13: Architettura ed aggiornamento

Slide 14: Stabilità e funzione di energia

Slide 15: Caso discreto

Slide 16: Caso continuo

**Slide 6: Modello formale di un neurone artificiale**

Questa breve illustrazione del sistema neurale umano suggerisce uno schema per delineare l’organizzazione delle reti neurali artiﬁciali. Queste ultime sono formate da un elevato numero di unità computazionali, che possono essere equiparate ai neuroni umani, capaci di eseguire una somma pesata. Tra gli elementi base che formano queste unità troviamo:

• un insieme di connessioni;

• un sommatore;

• una funzione di attivazione;

• un valore di soglia

Le unità sono collegate tra loro attraverso delle connessioni, così come le sinapsi collegano i neuroni nella rete umana. Consideriamo una generica unità costituita da canali di ingresso . Gli input provenienti da strati precedenti o dall’esterno entrano nel neurone tramite tali canali. Sulle connessioni sono presenti dei pesi sinaptici wi, numeri reali che denotano l’eﬃcacia sinaptica, ovvero la forza della connessione. Se wi >0 il canale è detto eccitatorio, se wi<0 il canale è inibitorio. I segnali in entrata sono convogliati nel soma del neurone artiﬁciale all’interno del quale vengono sommati producendo una combinazione lineare indicata con la parola net ed il segnale con cui il neurone trasmette la sua attività all’esterno è calcolato applicando una funzione di attivazione ϕ che limita l’ampiezza dell’output. Il modello neuronale include anche un valore soglia che ha l’eﬀetto, a seconda della sua positività o negatività, di aumentare o diminuire il valore in ingresso alla funzione di attivazione. Interpretando la soglia come il peso associato ad un ulteriore canale di ingresso x0, e quindi w0 = θ, l’output ﬁnale sarà:

**Slide 7: Attivazione e funzioni di attivazione**

La funzione di attivazione determina il tipo di risposta che un neurone è in grado di emettere. Deﬁnisce, quindi, l’uscita di un neurone in funzione del livello di attivazione. L’uscita può essere un numero reale, un numero reale appartenente ad un intervallo, oppure un numero appartenente ad un insieme discreto.

Nelle reti neurali si può avere una attivazione asincrona in cui i neuroni si attivano uno per volta ed un’attivazione sincrona in cui tutti i neuroni si attivano contemporaneamente.

Gli step per l’attivazione di un neurone sono:

1. i valori di input arrivano al neurone e vengono combinati con i relativi pesi;

2. i prodotti risultanti vengono sommati e tale somma viene confrontata con un valore di soglia che dipende dalla funzione di attivazione del neurone;

3. se la somma supera il valore di soglia, il neurone si attiva inviando output sulle connessioni pesate in uscita, altrimenti viene inibito.

Esistono varie tipologie di funzioni di attivazione, le principali sono: funzioni a soglia, funzioni lineari, funzioni lineari a tratti, funzione sigmoide.

* Iniziamo con la funzione a soglia che è tra le più utilizzate. Imponendo

il valore di uscita è ed assume questa forma (figura)

* Per quanto riguarda la funzione lineare e la funzione lineare a trattiimponendo sempre otteniamo come valore di uscita y = ϕ(a) = a per quanto riguarda la lineare mentre un esempio di output in presenza di un lineare a tratti è

y = ϕ(a)=

e rappresentandole si hanno questi grafici

* Infine per quanto concerne la funzione sigmoide, questa risulta essere appartenente alla famiglia di funzioni continue non lineari ed insieme alla funzione a soglia risulta essere tra le più utilizzate. In particolare riportiamo una funzione appartenente a questa classe che prende il nome di funzione logistica ed è definita secondo questa forma y = ϕ(a) =

mentre la sua rappresentazione è quella in figura (figura)

**Slide 8: Architettura**

L’architettura di una rete neurale artiﬁciale è caratterizzata da:

• numero di strati di sinapsi;

• numero di neuroni presenti nell’input layer ovvero lo strato di ingresso;

• numero di neuroni nell’output layer, lo strato di uscita.

Le reti neurali si suddividono principalmente in tre grandi classi: le reti completamente connesse, le reti feedforward e le reti feedback o ricorrenti.

* Reti completamente connesse:

Nelle reti completamente connesse ogni neurone è connesso con tutti gli altri. Le connessioni tra i neuroni sono bidirezionali e possono essere rappresentate per mezzo di una matrice quadrata W di dimensione pari al numero di neuroni. Un suo generico elemento wi,j rappresenta il peso della connessione tra il neurone i ed il neurone j.

* Reti feedforward

In queste reti i segnali viaggiano dallo strato di ingresso verso lo strato di uscita e non esistono connessioni tra i neuroni all’interno di uno stesso strato, nè tra neuroni di strati non limitroﬁ. Ogni neurone in un generico strato è connesso con tutti quelli dello strato successivo ed i neuroni dello strato di ingresso hanno come unico compito quello di trasmettere i segnali ricevuti allo strato successivo, all’interno di essi non avviene alcuna computazione. Le reti feedforward a strati si distinguono in base al numero di strati che presentano, numero che dipende dallo speciﬁco problema che si intende risolvere.

- Reti feedforward ad uno strato che è la forma più semplice.

- Reti feedforward a più strati dette anche reti multilivello , in inglese Multi-Layer Perceptron, MLP. Tra l’input layer e l’output layer sono presenti uno o piu` strati di neuroni nascosti, si parla quindi di hidden layers. Nelle MLP ogni strato ha connessioni entranti dal precedente strato e uscenti in quello successivo, quindi la propagazione del segnale avviene in avanti in modo aciclico e senza connessioni trasversali.

- Reti ricorrenti che si distinguono dalle precedenti nel fatto che sono cicliche. Il sistema di tali reti è dinamico, dato un determinato stimolo la risposta della rete non viene dettata soltanto dai caratteri strutturali della rete stessa, ma varia in funzione del precedente contesto in cui si è manifestato lo stimolo. L’output non è determinato solo dall’input, ma anche da una cronologia di input che fornisce una forma di memoria a breve termine.

**Slide 9: Addestramento e paradigmi**

Per costruire una rete neurale eﬃciente un passo fondamentale è individuare un insieme di apprendimento detto training set ovvero una collezione di esempi dal quale la rete attingere per raggiungere il suo scopo ed un algoritmo di apprendimento ovvero un procedimento che permetta di prelevare le informazioni dal training set e di ﬁssare dei parametri che vengano poi modiﬁcati attraverso operazioni iterative interfaccianti con l’ambiente. L’apprendimento avviene sempre grazie ad un certo numero di esempi prelevati dal mondo reale ed opera in due fasi distinte:

• Fase di apprendimento o addestramento nota come learning durante la quale si cerca di far imparare alla rete tutte le informazioni contenute nel training set ottenendo un modello che verrà poi utilizzato nella fase di generalizzazione per analizzare nuovi ingressi

• Fase di generalizzazione nota come recall dove per generalizzazione di una rete si intende la sua capacità di fornire le risposte appropriate a pattern di input che non sono mai stati incontrati e quindi in questi termini si dice che una rete generalizza bene se produce mappature input/output corrette anche se l’input `e lievemente diﬀerente dagli esempi usati in fase di training.

Le reti neurali si ispirano al tratto caratteristico del sistema nervoso, ovvero la capacità di acquisire esperienza da esempi del mondo reale: per questo, oltre che di apprendimento, si parla di addestramento delle reti neurali attraverso dei paradigmi di apprendimento. I paradigmi di apprendimento si suddividono in:

• apprendimento supervisionato che presuppone un training set nel quale sono presenti coppie del tipo (xk,ydk) dove la prima variabile indica il k-esimo ingresso e la seconda la k-esima uscita desiderata. Con yk si indica l’uscita reale e la si confronta con l’uscita desiderata: l’obiettivo è modiﬁcare i pesi aﬃnchè si minimizzi la diﬀerenza tra le due uscite. Il training set iniziale viene proposto ripetutamente ﬁnch´e yk ≈ ydk, ovvero l’uscita reale sia il piu` simile possibile a quella desiderata, il tutto modiﬁcando i pesi in base alla legge di apprendimento scelta.

• apprendimento non supervisionato durante il quale la rete modiﬁca i pesi autonomamente quindi si auto-organizza. Viene fornito solo il training set senza precisare le uscite.