

Learning Spettrale nelle Reti Neurali

Relatore: Duccio Fanelli - duccio.fanelli@unifi.it

Candidato: Lorenzo Giambagli - lorenzo.giambagli@stud.unifi.it

Le reti neurali (Neural Network) sono di solito addestrate aggiustando opportunamente le proprie connessioni (pesi) in un'operazione chiamata learning supervisionato.

In questa tesi abbiamo concentrato l'attenzione sulle reti neurali feedforward, in cui gruppi di neuroni collegano, in cascata, i neuroni (nodi) di ingresso con quelli di riconoscimento e come banco di prova abbiamo usato la discriminazione delle cifre scritte a mano da 0 a 9 (database MNIST).

Nella fattispecie ci siamo concentrati su un'analisi del funzionamento delle reti neurali mediante lo sviluppo di una tecnica di addestramento che si basa sullo spazio reciproco della rete. Questo viene opportunamente strutturato in modo da favorire il trasferimento di informazione dai neuroni di ingresso della rete verso i nodi di riconoscimento. Aniché modificare le connessioni tra i vari nodi, si vanno a cambiare gli ingressi dei vettori della base del suddetto spazio così che, attraverso un progressivo embedding, sia possibile discriminare le varie cifre.

I vettori di base dello spazio reciproco sono stati identificati con gli autovettori di un operatore di trasferimento. Mediante l'addestramento dei suoi autovalori si ottiene un classificatore lineare che, di fatto, costituisce la matrice di adiacenza di una rete. Concatenando questi operatori si possono ottenere classificatori lineari e non. Alternando, infatti, la loro applicazione a delle funzioni non lineari si riesce ad estendere il ragionamento riuscendo a costruire quello che è l'analogo di una rete neurale feedforward.

Speculiamo che il passaggio di informazione consentito dai vettori di base non ortogonali dello spazio reciproco sia il paradigma del funzionamento di reti neurali generiche.

L'accuracy ottenuta sul dataset MNIST mediante l'addestramento dello spazio reciproco della rete è stata confrontata con quella ottenuta con un addestramento classico, nello spazio diretto, di una rete analoga. I livelli di accuracy raggiunti sono pressoché identici, a fronte di un numero di parametri utilizzato circa 10 volte inferiore. Si avanza, inoltre, l'ipotesi che un addestramento solo degli autovalori degli operatori di trasferimento possa essere un'efficiente (visto il numero estremamente ridotto di parametri liberi) tecnica di pre-training per una qualsiasi rete neurale.

Nella tesi si esplora inoltre una tecnica di ricostruzione delle features dei dati che parte dall'inversione del processo di discriminazione. Questa tecnica, ottenuta a partire dagli operatori di trasferimento addestrati, è stata sfruttata per l'analisi delle porzioni di immagine considerate dalla rete per un'efficiente discriminazione.

Spectral Learning in Neural Networks

Deep neural networks are usually trained in the space of the nodes, by adjusting the weights of existing links via suitable optimization protocols.

In this thesis we have focused our attention on a class of network called “feedforward”, characterized by cascades of parallel neurons organized in layers. Inside each layer there are no connections between neurons, which are, on the contrary, fully connected with the ones of the previous and following layer.

Here we propose a radically new approach which anchors the learning process to reciprocal space. Specifically, the training acts on the spectral domain and seeks to modify the eigenvectors and eigenvalues of transfer operators in direct space. The proposed method is ductile and can be tailored to return either linear or non linear classifiers.

The performance are competitive with standard schemes, while allowing for a significant reduction of the learning parameter space.

As a reference we have used the discrimination of handwritten digits from 0 to 9, collected in the MNIST database.

Spectral learning restricted to eigenvalues could be also employed for pre-training of the deep neural network, in conjunction with conventional machine-learning schemes. Further, it is surmised that the nested indentation of eigenvectors that defines the core idea of spectral learning could help understanding why deep networks work as well as they do.

In the present work a features reconstruction technique, based on the trained transfer operators, is explored.

The aforementioned technique has been exploited in order to understand how each portion of the images is used for an efficient discrimination process.