

Problemi di Matematica delle Reti Neurali

Montanari Andrea
Università di Stanford

Gli ultimi quindici anni hanno testimoniato lo sviluppo di nuove tecnologie di apprendimento automatico (deep learning) con applicazioni che spaziano dalla robotica alla generazione automatica di testo. Questi sistemi sono basati su architetture (reti neurali) ampiamente overparametrizzate: il numero di parametri è dello stesso ordine o più grande della taglia del campione di dati. Di conseguenza, queste reti sono spesso capaci di interpolare i dati usati per l'apprendimento, anche se questi dati vengono rimpiazzati (parzialmente o interamente) da dati puramente aleatori. Nonostante questo, i modelli appresi in questo modo appaiono robusti rispetto alla frazione di dati aleatoria, e generalizzano bene a dati non visti. Queste osservazioni sollevano numerosi problemi sia dal punto di vista statistico che da quello dell'ottimizzazione. Cercherò di dare un riassunto di quello che è stabilito dal punto di vista della teoria matematica e di elencare qualche problema aperto.