

Tarallucci, Vino e Machine Learning
“Il paper della buonanotte”™

Come Imparano le Reti Neurali

Fabio Mardero
fabio.mardero@gmail.com
github.com/fmardero

18 aprile 2019



TVML



Capacità di Memorizzazione

Apprendimento con Dati Reali e Sintetici

Analisi tramite Information Theory



Capacità di Memorizzazione



Zhang et al. - *"Understanding Deep Learning Requires Rethinking Generalization"* - 26 Feb 2017

Contenuti Principali

- ▶ Le reti neurali hanno la capacità di approssimare qualsiasi funzione, anche illogica.
- ▶ Le capacità di generalizzazione tra training e test set della NN non sono particolarmente influenzate dalle usuali tecniche di regolarizzazione.
- ▶ Lo SGD sembra uno dei principali fattori che consentono alla NN di generalizzare.



Apprendimento con Dati Reali e Sintetici



Arpit et al. - *"A Closer Look at Memorization in Deep Networks"* - 1 Jul 2017

Contenuti Principali

- ▶ Le reti neurali non “memorizzano” i dati reali.
- ▶ Le reti neurali apprendono prima strutture semplici (frequenti nel dataset) e poi quelle complesse.
- ▶ Le usuali tecniche di regolarizzazione possono ostacolare il fenomeno “memorizzazione”, preservando le capacità di apprendimento della NN con dati reali.



Analisi tramite Information Theory

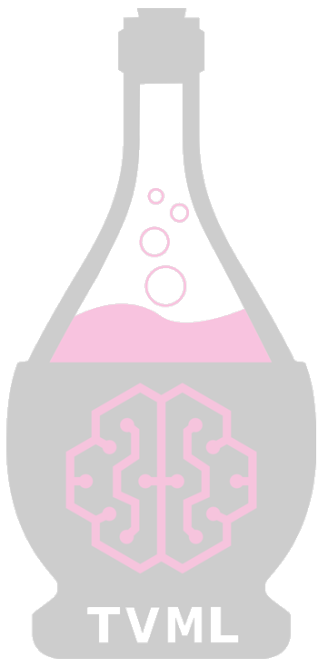


Schwartz-Ziv R. e Tishby N. - “*Opening the Black Box of Deep Neural Networks via Information*” - 29 Apr 2017

Contenuti Principali

- ▶ L'apprendimento di una rete neurale si divide in due fasi: *empirical risk minimization* (ERM) e *representation compression phase*.
- ▶ Tra le due fasi, quella di compressione è estremamente dipendente dalle dimensioni del dataset.
- ▶ Seguendo il processo di ottimizzazione tramite SGD, si individuano due fasi di ottimizzazione dell'algoritmo: *drift phase* e *diffusion phase*. Vi è uno stretto legame con quelle di ERM e compressione.

Naftali Tishby - “*Stanford Seminar - Information Theory of Deep Learning*” - 5 Apr 2018



Grazie dell'attenzione!