
Emozioni ed Apprendimento: Intelligenza artificiale per la formazione

Matteo Avalle • fluentify.com • Linux Day Ivrea • 2019

Il problema

Lezioni di inglese in videoconferenza

- Facilmente registrabili
- Possibilita' di identificare i punti "critici" su cui concentrarsi
- Solitamente, sono i momenti di maggior stress per lo studente
- Ottimo spunto per suggerire argomenti di ripasso

Però

E' veramente possibile analizzare manualmente tutte le registrazioni?

Automatizziamo l'analisi!

- E' sufficiente la traccia audio?
 - I risultati saranno realistici?
-

La soluzione

(forse)

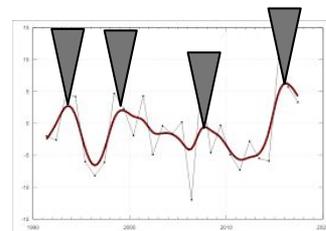
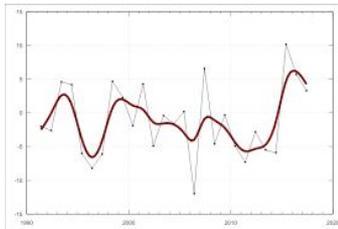
Le risposte in letteratura

SI PUO' FARE

- Provata correlazione tra l'intonazione della voce ed il livello di stress
- Estrarre l'intonazione da una traccia audio è relativamente semplice

NON SI PUO' FARE

- Soggetti diversi mostrano variazioni diverse
- Ci sono fattori esterni che causano variazioni, difficile essere precisi



Videolezione

Registrazione
Audio

Grafico
Intonazione

Analisi
statistica

Aggiunta
marker



$$\frac{\partial}{\partial a} \ln f_{a, \sigma^2}(\xi_1) = \frac{(\xi_1 - a)}{\sigma^2} f_{a, \sigma^2}(\xi_1) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left\{-\frac{(\xi_1 - a)^2}{2\sigma^2}\right\}$$
$$\int \mathcal{T}(x) \cdot \frac{\partial}{\partial \theta} f(x, \theta) dx = \mathbb{M}\left(\mathcal{T}(\xi) \cdot \frac{\partial}{\partial \theta} \ln L(\xi, \theta)\right) \int \frac{\partial}{\partial \theta} f(x, \theta) dx$$
$$\int \mathcal{T}(x) \left(\frac{\partial}{\partial \theta} \ln L(x, \theta) \right) \cdot f(x, \theta) dx = \int \mathcal{T}(x) \left(\frac{\partial}{\partial \theta} f(x, \theta) \right) dx$$
$$\frac{\partial}{\partial \theta} \mathbb{M}(\mathcal{T}) = \frac{\partial}{\partial \theta} \int \mathcal{T}(x) f(x, \theta) dx = \int \frac{\partial}{\partial \theta} \mathcal{T}(x) f(x, \theta) dx$$

Risultati dei primi test



- Persone diverse hanno comportamenti diversi
 - Impossibile trovare una regola matematica per identificare i punti chiave
 - Risultati non “credibili”
 - Informazioni aggiuntive non facilmente collegabili
-

Scenario ideale per IA

<https://medium.com/@erikhallstrm/hello-world-rnn-83cd7105b767>

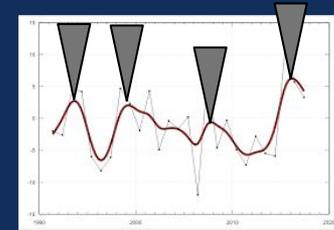
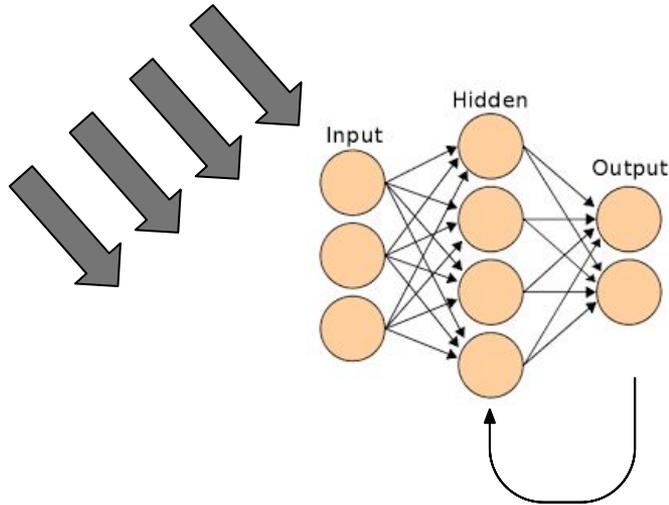
<http://karpathy.github.io/2015/05/21/rnn-effectiveness/>

- Abbiamo tanti possibili input diversi, ma non sappiamo bene come combinarli tra loro
 - Sembra evidente ci sia presenza di una correlazione tra le variazioni di tono ed il livello di stress, ma la correlazione e' difficile da "spiegare"
 - E' molto semplice produrre manualmente dei dati "di verifica": basta guardare delle registrazioni
-

Perche' rivolgersi all'IA?

Il nuovo prototipo

- **Intonazione**
- **Volume**
- (Audio completo)
- (Analisi in frequenza)



**Identificazione
stress marker**

- Risultati validi con i dati di training
 - Risultati “credibili” anche su registrazioni inedite
 - Ampi margini di miglioramento (nuovi input, ottimizzazione parametri, etc)
-

Risultati dei nuovi test



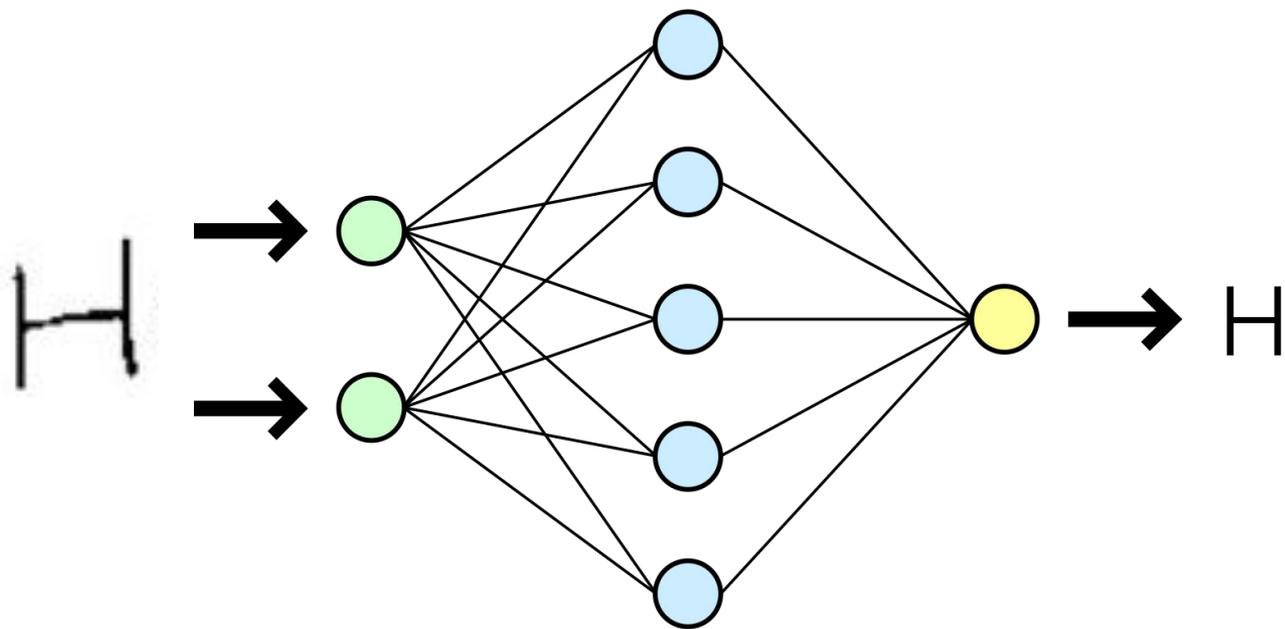
Ok, ma... perché?

Tre tecnologie adottate

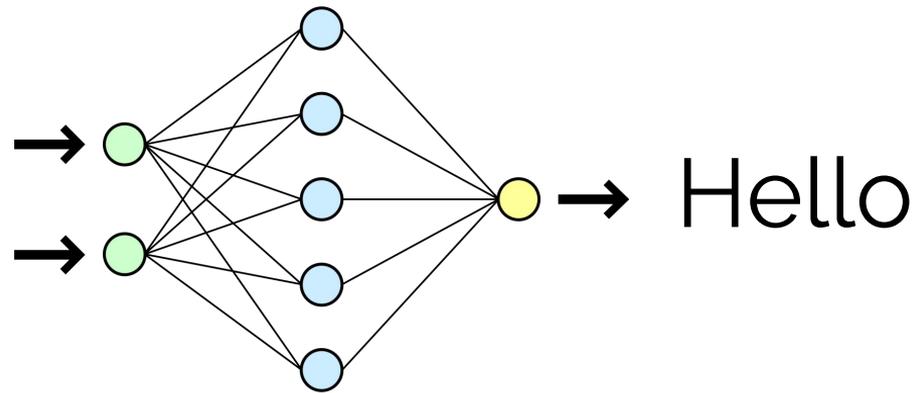
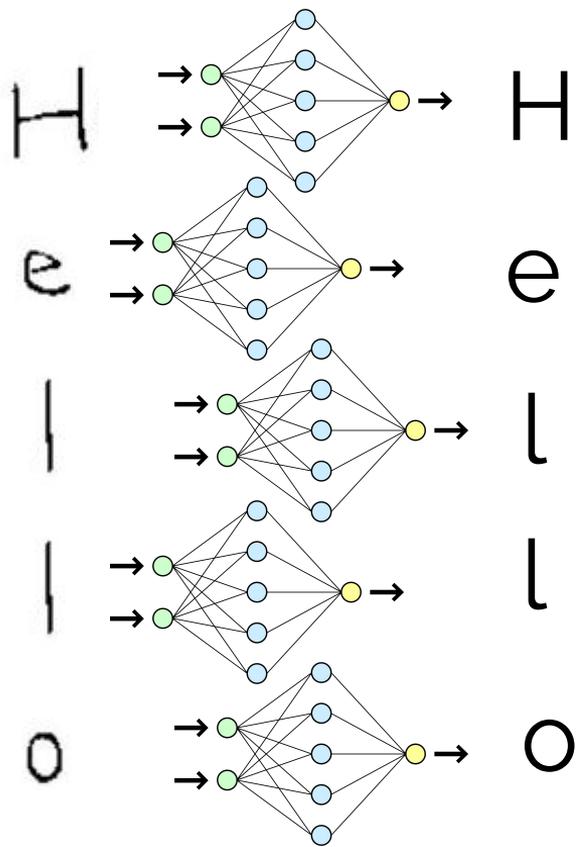
- Neural Network
- Deep Neural Network
- Recurrent Neural Network

**Cosa
significano?**

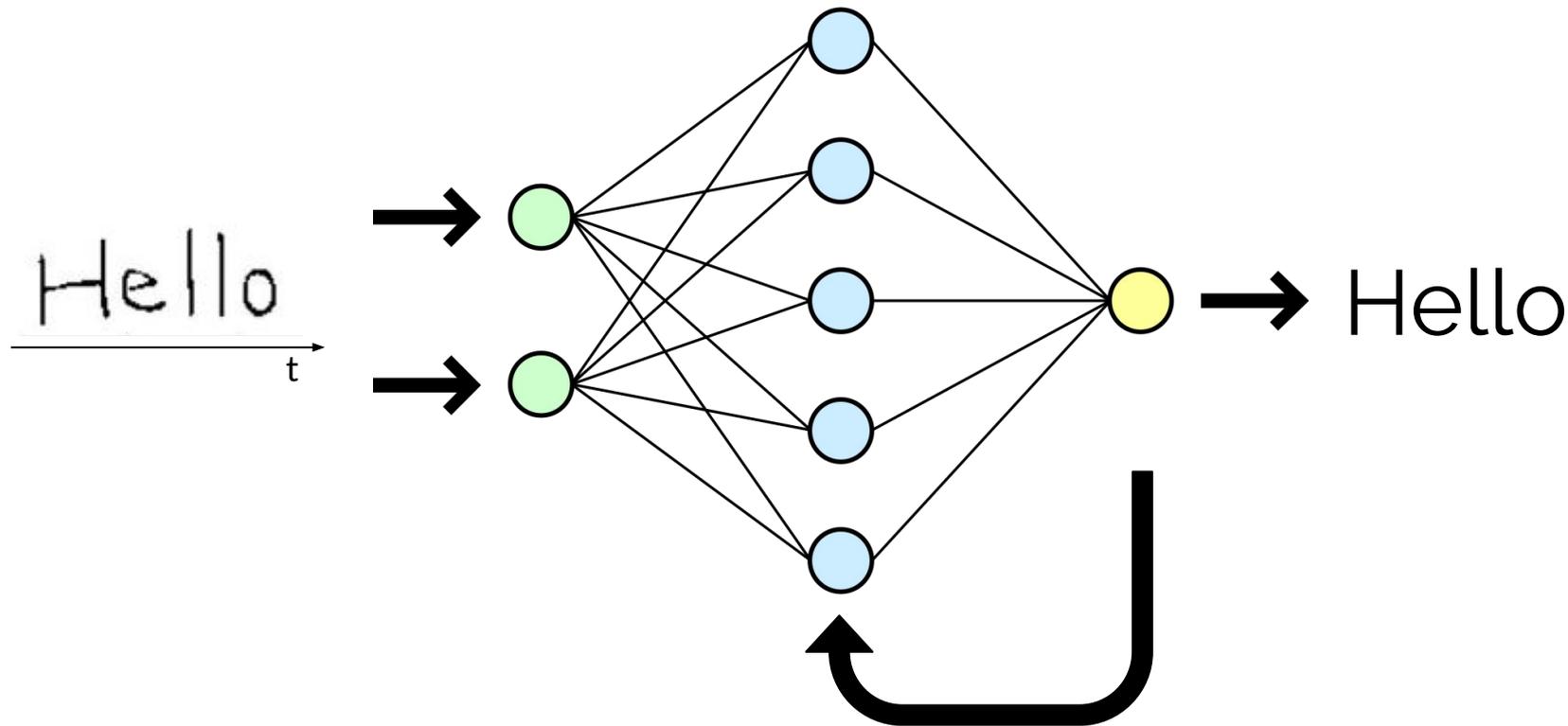
Rete neurale



Rete neurale "Profonda": livelli di astrazione



Rete neurale "ricorrente": la memoria



Conclusioni

Funziona

- Il team di sviluppo non era assolutamente all'altezza di risolvere il problema: minime conoscenze di psicologia e statistica, assolutamente inadeguate data la problematica
- Le conoscenze di IA sono altrettanto minime
- L'IA ha permesso comunque di ottenere una soluzione, non ottimale ma comunque sufficiente a risolvere il problema
- Le estreme complessità di una IA in grado di ragionare, ricordarsi concetti a più livelli di astrazione, sono mascherate da piattaforme open

Grazie
