

L'esplosione dell'intelligenza artificiale dal punto di vista di Piero Savastano

[LINK AL VIDEO DI TIKTOK](#)

dal 2012 in poi vivendolo da dentro da esperto di nicchia e poi fino a fine è diventata una cosa mondiale.

Vi ricordo che nel 2012 che l'anno in cui AlexNet ha vinto la competizione scientifica ImageNet, se ne parlava soprattutto nel mondo del computer vision e dell'intelligenza artificiale basata su simulazioni del tessuto nervoso.

[Era la nicchia della nicchia] AlexNet e il team che ha vinto questa competizione in cui Devi costruire un algoritmo in grado di classificare immagini.

Data un'immagine ti dico cosa c'è dentro una tra migliaia di categorie, quindi un compito di classificazione difficile.

il training set, l'immagine d'addestramento stava nell'ordine di grandezza dei milioni e quindi arriva questo team AlexNet che dice OK scaliamo sulle reti neurali, facciamo le profondissime hanno rispolverato le convolutioni di Yann LeCun che adesso è il capo della ricerca di Facebook e hanno creato questa rete a convolutioni profondissima che ha sbaragliato ImageNet puntando sulla scala.

Quindi la rivoluzione del deep learning che ricalca su scoperte scientifiche di decenni prima si basa sulla scala, ed è interessante che il team di AlexNet era composto da Geoffrey Hinton che è il padre del deep learning di base, quello lì che ha inventato la *back propagation* e si deve a lui praticamente il grosso di questo settore.

A mio avviso, sarà considerato nella storia dell'umanità la strega di un alla stregua di un Darwin e o di un Freud e non sto esagerando perché Geoffrey Hinton ci ha dato la conferma che una macchina può fare le stesse cose che facciamo noi simulando il tessuto nervoso e per questo lo ritengo all'altezza di un Freud che ti ha detto che in realtà pensiamo sempre alla sfera sessuale e lo mette all'altezza di un Darwin che ci ha detto che siamo delle scimmie, cioè sono tutte tutti i pensatori che ci hanno un po' declassato che però a lungo da me ci hanno dato più libertà di pensiero e di sogno.

tornando ad AlexNet c'era Geoffrey Hinton, c'era Ilya Sutskever che al momento è il capo della ricerca di OpenAi (era un ragazzo poi ha fatto strada). AlexNet Vince la competizione da lì in poi, che cosa è successo?

che ci sono state diverse ondate di computer vision che si alternava modelli per il linguaggio, (modelli generativi per mettere in fila le parole).

Mi ricordo che la computer vision in realtà da AlexNet in poi è passata nel tempo di 2 3 anni direttamente all'*object-detection* quindi non solo classifico l'immagine ma dico anche quale oggetto si trova dove all'interno di un'immagine e più oggetti per immagine.

Questa roba è arrivata molto velocemente, invece il linguaggio *laggava*, c'erano delle cose bellissime, nel 2015/16 17 sono ero di BERT transformer, erano Belle architetture, scalavano il linguaggio era sempre più figo però non ci potevi fare affidamento, era un linguaggio strampalato, sconclusionato e quindi questi oggetti venivano utilizzati soprattutto per fare sempre classificatori linguistici embedding, per proiettare il testo nello spazio geometrico, per fare task molto limitati.

Il ruolo di OpenAi in questo è soprattutto legato al linguaggio più che alla computer vision e si è visto tanto ruolo di deep mind acquisita Google e il genio di Demis Hassabis che invece hanno un ruolo rilevante in un in un aspetto di questo mondo che ancora non è scalato ma che sicuramente è ancora più dirompente di quello che abbiamo visto finora che è quello

dell'apprendimento per rinforzo.

Prima che cominciassero a uscire i vari gpt2 e poi gpt3 che hanno rivoluzionato.

Deep mind aveva rivoluzionato eh il mondo degli agenti intelligenti con una rete neurale che imparava a giocare a non so quanti giochi dell'Atari guardando i pixel dello schermo e ricevendo come unico segnale i punteggi del gioco (se morivi oppure no.)

Quindi senza **l'apprendimento supervisionato** per cui tu Sai qual è la risposta e la indica alla macchina, la macchina poi si ottimizza i pesi no con **l'apprendimento del rinforzo**. Tu indirettamente dai dei segnali scalari (es. fuoco, fuochino, acqua, acquazzone) e la rete deriva da sola alla *policy* quello che deve fare.

Deepmind con questa tecnologia qui non solo ha vinto l'atari, ma poi ha vinto anche a scacchi, ha vinto anche a go alfa go e quindi però questa cosa qui non è passata tanto al mondo delle applicazioni, più di tanto perché scala poco.

apprendimento per rinforzo è molto più potente della semplice *back propagation*, anche perché nell'apprendimento per rinforzo puoi incapsulare la back propagation che è il deep learning classico, però scala molto meno, ci vogliono molti più dati, ci vogliono molte più risorse di calcolo per far convergere questi agenti con la rete dentro che tu addestri per rinforzo e quindi siamo ancora nell'era dell'apprendimento supervisionato. ma tenetevi forte perché il meglio deve ancora arrivare e il meglio sta dentro Google.

anche OpenAi inizialmente faceva prendere un rinforzo, si sono accorti che scalava poco.

[Elon Mask era già uscito da OpenAi che nasce come no profit.]

Siamo ancora nel periodo in cui è la computer vision a essere industrializzata con l'object detection, ma il linguaggio ancora insomma è interessante ma ancora un Po' ancora poco di rompenza.

Si arriva g. p. t. 2 che non aveva niente di particolare se non di essere l'ennesimo transformer architettura coniata da Google, simile a tanti altri sistemi di quel tipo, non aveva nulla di speciale in realtà.

poi nel 2020 arriva il fulmine, c'è il sereno che è OpenAi i articolo del 2020, "*language models are few shot learners*", e qui si dice sistemi; quindi, abbiamo preso un oggetto molto simile in realtà a g. p. t. 2, l'abbiamo fatto **scalare** 10x per se tu ti leggi l'abstract di quell'articolo, lo vado in giro a dire è la scala che conta, è la scala che ha fatto la differenza: trovo il classico ciarlatano che mi dice no perché ci stanno i transformer, perché ci sta questo... [nota]

la dirompenza di g. p. t. 3 e dei Language model la trovi scritta nell'abstract del loro articolo in cui dicono abbiamo scalato questi sistemi, quindi non si sono inventati in realtà niente di più, c'erano tutte le **premesse** teoriche, ma le premesse teoriche le stiamo preparando dai dagli anni 80, quindi **back propagation**, reti multistrato, reti profonde, il ramo della computer vision, poi sono arrivati i transformer per le serie multivariate. per il linguaggio l'ultimo ingrediente, in realtà ingrediente che era già stato anticipato da AlexNet nella computer vision era la scala.

il linguaggio ha bisogno di scalare molto di più della visione. Perché si usano un sacco di parametri dentro questa rete e forse perché il linguaggio è molto più complesso. forse (forse non è detto) ma comunque esce questo g. p. t. 3 nell'abstract ti dicono noi abbiamo semplicemente scalato e abbiamo trovato proprietà emergenti pazzesche, tra cui il fatto che non Devi più fare *fine-tuning* delle reti come si faceva con quelle precedenti, incluso g. p. t. 2 quando ti serviva declinato. su una cosa Devi andare a modificare la rete, Devi fare il fare *fine-tuning*.

Openai ci ha detto nel 2020 questa cosa, se tu gli metti nel prompt, nel testo di input, quello che deve fare è in grado di seguire le istruzioni. e lì proprio boom.

Gente che faceva ricerca in **natural language process** in da vent'anni, che piava la valigetta e faceva raga ciao (significa che se ne andava data la scoperta rivoluzionaria).

Una rivoluzione totale nel mondo scientifico, quindi un botto molto più grande di quello di Alex. net in computer vision. nel frattempo, c'erano state varie ondate e tutti quanti stavamo lì che dicevamo OK deep learning, è arrivata a saturazione, è arrivata a saturazione. no non era arrivata a saturazione e forse ancora non c'è arrivato.

arriva il g. p. t. 3, mondo scientifico in subbuglio totale, tempo un paio d'anni quando esce chat g. p. t. nel frattempo già si facevano i chatbot basati su g. p. t. 3, già si cominciava a vedere roba, però niente ha fatto il botto a livello poi di **consumer** di Massa come chat g. p. t. e quindi dal mio punto di Vista ho vissuto questo crescendo in cui sembrava sempre che stesse per finire, se stesse per finire, invece continuava, continuava, continuava.

Proprio un andamento esponenziale e da parte Mia l'interessante è stato vedere che eh ci facciamo le rete in casa essenzialmente e piano piano negli anni inesorabilmente si è allontanato sempre di più questa cosa di farti le rete in casa, studiarsi pay torch e tenso fluo è diventato sempre meno sensato.

sono nate librerie di più alto livello, tipo i transformers di Hugging face per il linguaggio e yolo v. x per la computer vision, con cui avevi a che fare con questi oggetti senza sapere nulla della rete neurale che c'era sotto di base e e si è allontanato sempre di più questo discorso di farteli in casa, è diventato fatto il *fine tuning* fino al 2018 19. Poi pure il *fine tuning* è passato totalmente in secondo piano in favore di questi modelli di fondazione che vengono, eh batterie incluse per aiutarti in quello che vuoi, perché glielo dici direttamente nel prompt e sono felice di essere vivo in questo periodo storico.