Intelligenza artificiale libera (autore: Vittorio Albertoni)

Premessa

Se chiediamo a ChatGPT di scrivere, utilizzando il non facile linguaggio C++, un software che chieda all'utente il suo nome per rivolgergli un saluto, in meno di un secondo ci troviamo di fronte il listato del codice scritto in maniera perfetta.

Ho chiesto a ChatGPT:
«Come mai riesci a scrivere un programma in linguaggio C++ ?»

e la risposta è stata:
«Posso "scrivere" programmi in linguaggio C++ solo nel senso che sono in grado di generare codice sorgente basato su modelli o esempi. La mia capacità di generare codice è basata sulle informazioni che ho acquisito durante il mio addestramento, che include una vasta gamma di conoscenze in diversi campi, compresa la programmazione.
Tuttavia, è importante sottolineare che la mia generazione di codice non è equivalente a scrivere un programma in modo interattivo come farebbe un programmatore umano. Non ho la capacità di interagire direttamente con ambienti di sviluppo, compilatori o eseguire effettivamente il codice. La mia capacità deriva dalla comprensione di modelli linguistici e di codice che emergono dai dati di addestramento.»

Allo stesso modo, se chiediamo a ChatGPT di descriverci la storia di Cremona lo fa con linguaggio appropriato e con dovizia di particolari.

Ho allora chiesto a ChatGPT:
«Come mai conosci la storia di Cremona ?»

e la risposta è stata:
«Conosco informazioni generali sulla storia di Cremona e di molti altri argomenti perché durante il mio addestramento ho analizzato e assimilato una vasta gamma di testi e documenti provenienti da fonti pubbliche disponibili su Internet. La mia capacità di fornire informazioni deriva dalla comprensione dei modelli linguistici e delle relazioni concettuali presenti in questi dati.
Tuttavia, è importante notare che non ho accesso a informazioni in tempo reale o a fonti riservate. Le informazioni che fornisco si basano sulla conoscenza disponibile fino alla mia data di ultimo aggiornamento.»

Cioè la così detta intelligenza artificiale generativa, che, su comando, produce qualcosa che apparentemente non c'era prima, si realizza attraverso l'elaborazione di grandi masse di dati con un software tendente ad individuare in maniera sempre più approfondita le relazioni esistenti tra questi dati e ad addestrare un macchina a generare, in modo sempre più coerente, altri dati relativi a specifiche situazioni in base all'assimilazione delle relazioni esistenti tra i dati di partenza.

Un'idea - pur vaga, data la difficoltà della materia - di come si possa produrre questo software possiamo averla leggendo il Capitolo 7 del mio testo «Tutto Python per principianti», disponibile in formato PDF in questa stessa rubrica[1] .

Bene ricordare che l'intelligenza artificiale generativa, di cui ChatGPT è forse il più noto esempio, non si limita alla generazione di scritti ma può riguardare anche generazione di immagini e suoni.

In questa piccola rassegna presento alcune realizzazioni che si fondano su software open source.

Il tutto destinato ad un pubblico di inesperti affinché abbiano la possibilità di fare qualche esperimento.

[1] A proposito di Python e di ChatGPT, rammento che ChatGPT è stato programmato con il linguaggio Python.

Indice

1 Suono

Nel settembre 2018 ho pubblicato sul mio blog all'indirizzo *www.vittal.it* l'articolo «Intelligenza artificiale per l'improvvisazione musicale» presentando nell'allegato «impro_visor» un software in grado di aiutarci a scrivere musica jazzistica imitando lo stile dei più noti artisti del genere.

Quel software applica la stessa tecnica applicata dai nuovi software generativi partendo, però, da basi di dati molto ridotte, nel caso specifico una serie di rappresentazioni in codice MIDI di fraseggi tipici utilizzati da vari jazzisti, aiutandoci a scrivere codice MIDI che ne riprenda lo stile.

Anche quella è intelligenza artificiale ma fa meno scena di ciò che abbiamo a disposizione oggi.

1.1 Stable audio

Stable audio genera suono secondo istruzioni ricevute in formato testo da un prompt[2] .

Forte dell'addestramento avuto sulla base dati costituita dall'archivio AudioSparx, contenente 800.000 file audio per quasi 20.000 ore di registrazione di svariati rumori, suoni e brani musicali di vario stile genera nuovi suoni tenendo conto di quanto indicato nel prompt.

Rilasciato nel settembre 2023, il software è stato prodotto con linguaggio Python da Stability AI e possiamo vederlo su GitHub.

Per usufruirne andiamo all'indirizzo *https://www.stableaudio.com/* dove abbiamo un campione di brano musicale prodotto secondo un prompt di una certa complessità, a dimostrazione della potenza del software.
Cliccando sul pulsante TRY IT OUT veniamo invitati ad accedere con le credenziali Google oppure creando un account gratuito ed acquisiamo così il diritto a produrre fino a 20 brani, non utilizzabili commercialmente, di durata fino a 45 secondi ogni mese, scaricabili in formato MP3 (possiamo diventare utente professional con € 11,99 al mese, avendo diritto a 500 brani, utilizzabili commercialmente, di durata fino a 90 secondi ogni mese, scaricabili in formato MP3 o in formato WAV e possiamo diventare utente enterprise trattando le condizioni).

La difficoltà sta nello scrivere il prompt, che va scritto in lingua inglese e deve descrivere bene ciò che vogliamo.

Se ci accontentiamo di produrre un file audio che ci faccia sentire il canto dell'usignolo possiamo semplicemente scrivere il prompt `nightingale`, a dimostrazione del fatto che possiamo produrre non solo musica ma anche suoni e rumori di altro tipo.

Per ottenere brani musicali dobbiamo essere più creativi nello scrivere il prompt, sapendo che Stable audio comprende l'indicazione degli strumenti da utilizzare, la descrizione dell'atmosfera che intendiamo creare, il tempo indicato in BPM (battiti al minuto) e il tipo di musica per grande categoria di appartenenza (jazz, country, ecc.).

Indicazioni e dimostrazioni al riguardo si trovano nella User Guide accessibile scorrendo fino in fondo la pagina di presentazione di Stable Audio all'indirizzo *https://www.stableaudio.com/*.

Francamente devo dire che, per i miei gusti musicali orientati al classico, al jazz d'autore e alla strumentazione tradizionale, posso benissimo fare a meno di Stable Audio.

2 Immagini

Quello dell'elaborazione di immagini è forse il filone più antico, se così si può dire, dell'intelligenza artificiale generativa.

[2] In tutte le forme di intelligenza artificiale generativa abbiamo la formulazione di una richiesta da parte dell'utente e una risposta da parte del computer, a ripetizione dell'interattività tipica dell'informatica un tempo praticata da terminale, dove il prompt indicava la riga di terminale in cui l'utente doveva scrivere le sue richieste.

La tecnica è sempre la solita: enormi database, questa volta di immagini, sui quali la macchina si esercita in maniera approfondita per riuscire a produrre nuove immagini aventi caratteristiche predeterminate.

2.1 ThisPersonDoesNotExist

E' un prodotto della rete generativa StyleGAN, creata a fine 2018 da alcuni ricercatori della NVIDIA.

Il codice è stato distribuito open source su GitHub a partire dal 4 febbraio 2019 su licenza libera Creative Commons.

Se andiamo all'indirizzo *https://this-person-does-not-exist.com/it* possiamo divertirci a generare visi assolutamente credibili di persone che non esistono.

Abbiamo a disposizione tre finestrelle in cui scegliere il sesso, la fascia di età e l'etnia della persona non esistente di cui produrre il viso cliccando sul pulsante AGGIORNA IMMAGINE.

Indicando come sesso maschile, come fascia di età 35-50 e come etnia nero ho ottenuto questo

2.2 Stable Diffusion

Rilasciato nell'agosto 2022 dalla Stability AI, la stessa di Stable Audio, è programmato in linguaggio Python e il codice è disponibile su GitHub con licenza libera.

Qui possiamo generare qualsiasi figura e non solo volti immaginari: come per Stable Audio dobbiamo essere bravi a compilare il prompt in cui descrivere in lingua inglese le caratteristiche della figura che vogliamo sia generata.

All'indirizzo *https://stablediffusionweb.com/* clicchiamo FREE AI IMAGE GENERATOR e inseriamo il prompt nella relativa finestrella. Se siamo grafici esperti possiamo scegliere uno stile aprendo il menu STYLES, ma noi possiamo accontentarci di ciò che viene proposto per default. Cliccando sul pulsante GENERATE avviamo il processo di creazione dell'immagine, processo che avrà una certa durata.

Al prompt `moonrise` ho ottenuto questa bella immagine

Al prompt `cat with his friend dog` ho ottenuto quest'altra

e qui mi è venuto il sospetto che l'intelligenza artificiale abbia preso una piccola allucinazione, perché il cane mi sembra avere un muso piuttosto da gatto.

3 Testo

Un computer che risponde a una nostra domanda, anche complessa, e lo fa in modo pertinente fa impressione: fino a che produce un brano musicale o una immagine, abituati come siamo a trattare musica e immagini con il computer, ci sembra quasi normale. Ma un computer che ti sforna tutto un ragionamento che ben si combina con quanto gli hai chiesto è un'altra cosa, perché sembra che il computer non solo capisca ciò che gli chiediamo ma pensi, come fa il nostro cervello prima di rispondere a una domanda, e poi risponda correttamente.

Pertanto, nonostante ormai da anni esistano applicazioni di intelligenza artificiale in svariati contesti, quando ci siamo trovati di fronte a ChatGPT abbiamo in massa scoperto l'intelligenza artificiale.

Di ChatGPT non parlo in questa rassegna dedicata all'intelligenza artificiale libera perché, nonostante il relativo software sia stato prodotto da una struttura che si chiama Open AI, di open non ha nulla in quanto si tratta di software proprietario e non si conoscono esattamente le origini dei dati di addestramento.

Peraltro, anche se quella che è diventata la più popolare manifestazione di intelligenza artificiale si deve a software proprietario (occorre comunque riconoscere che si merita la fama perché funziona bene), il più grande grande fermento di iniziative in materia di LLM (Large Language Models) lo ritroviamo nel mondo del software libero.

Gran parte dei modelli prodotti resta sconosciuta ai comuni mortali, essendo materia per addetti ai lavori.

Ogni tanto, però, c'è qualche addetto ai lavori che fa in modo che anche gente qualsiasi abbia accesso a modelli in forma di chat, proprio come ha fatto Open AI con i vari modelli GPT che alimentano ChatGPT.

Cito un paio di esempi.

3.1 Llama 2

LLaMA sta per Large Language Model Meta AI ed è stato prodotto da Meta AI, laboratorio di intelligenza artificiale fondato da Mark Zuckerberg.

Possiamo chattare con questo modello all'indirizzo

https://www.llama2.ai/.

3.2 Falcon 180 B

Prodotto dalla Huggin Face, nata grazie a investimenti da parte di Google, Amazon, Nvidia, Salesforce, AMD, Intel, IBM e Qualcomm.

Possiamo chattare con questo modello all'indirizzo

https://huggingface.co/spaces/tiiuae/falcon-180b-demo

nel momento in cui scrivo molto affollato di richieste e praticamente inaccessibile. Speriamo venga potenziato.

* * *

Molto spesso fanno la loro comparsa chatbot di cui si perdono presto le tracce.

Il caso più recente è quello di Open Assistant, progetto della LAION (Large-scale Artificial Intelligence Open Network), la cui chatbot è comparsa nell'aprile 2023 ed è scomparsa a fine novembre 2023 in quanto il progetto è stato chiuso.

Frettolosamente interpretato come la risposta del software libero a ChatGPT, in quanto funzionava molto bene, appariva strano tutto un meccanismo che vi si ricollegava di valutazione delle risposte e di richieste di contributi agli utenti che venivano premiati con punteggi per la loro collaborazione.

Ora scopriamo che era tutto un meccanismo per testare dataset e software ed ha rappresentato un ottimo esempio di software nato dalla collaborazione di un mondo.

All'indirizzo *https://projects.laion.ai/Open-Assistant/* troviamo tutta la spiega.

Su GitHub è stato rilasciato il software per far girare il quale non basta un semplice computer da consumer electronics.

Pertanto noi consumatori dovremo aspettare che qualcuno utilizzi questo software per offrirci una chatbot.

* * *

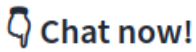
In tutto questo fermento merita di essere ricordata una lodevole iniziativa della LMSYS Org (Large Model Systems Organization), struttura creata da un gruppo di studenti dell'Università di Berkeley, che mette in grado gente comune come noi, con risorse hardware contenute, di sperimentare almeno una ventina di modelli Large Language, open source e proprietari, potendo anche metterli a confronto.

3.3 Chatbot Arena

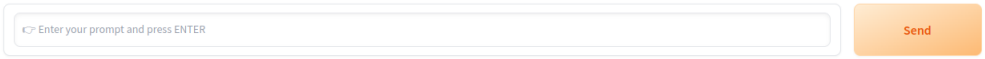
Andiamo all'indirizzo

https://chat.lmsys.org/

e si apre per noi la pagina per la chat, la cui zona inizia dove troviamo l'indicazione



e la zona del prompt

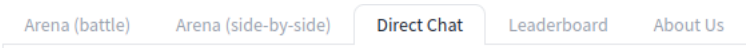


la troviamo verso il fondo della pagina, sotto la zona dedicata alle risposte.

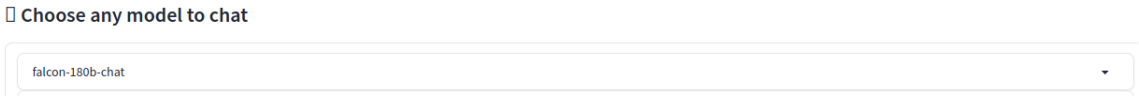
Nel momento in cui scrivo (dicembre 2023) Chatbot Arena può utilizzare una ventina di modelli. La risposta al prompt avviene nella lingua utilizzata per il prompt stesso.

Abbiamo tre modalità per divertirci e scegliamo la modalità agendo sulla prima riga della pagina che abbiamo di fronte.

La modalità più semplice è la DIRECT CHAT



attraverso la quale possiamo scegliere il modello con cui chattare e lo facciamo aprendo il menu che troviamo appena prima della zona delle risposte



Nell'illustrazione è stato scelto il modello `Falcon 180 B` ma cliccando sul triangolino in fondo a destra nella finestrella possiamo scegliere un altro modello tra quelli disponibili.

Trattandosi di un Chatbot Arena possiamo far gareggiare a due a due i modelli e indicare quale sia stata la risposta migliore.

Con la modalità ARENA (BATTLE) mettiamo a confronto due modelli scelti a caso dal sistema, senza conoscerne il nome.

Con la modalità ARENA (SIDE-BY-SYDE) mettiamo a confronto due modelli scelti da noi.

Aprendo la pagina LEADERBOARD possiamo vedere la classifica aggiornata sulla base dei giudizi espressi da chi si è divertito a far gareggiare i modelli e aprendo la pagina ABOUT US veniamo a conoscere i nomi di chi ci ha regalato questo simpatico strumento.

Il testo generato da questi strumenti può essere un testo letterario, come una poesia, un racconto o una barzelletta composti secondo indicazioni di contenuto e stile fornite nel prompt. Può essere il listato di un programma per computer scritto con un linguaggio e per fare le cose indicate nel prompt. Può riguardare la soluzione di un problema, matematico o di altra natura, indicato nel prompt. Può fornire conoscenza su un qualsiasi argomento indicato nel prompt.

Nel confezionare le risposte al prompt il modello può soffrire di allucinazioni e creare testi insignificanti, come barzellette con non fanno ridere nessuno, o scorretti.

Chattando qua e là ho, per esempio, scoperto che il film «Un dollaro d'onore» è la versione italiana del film «The man who shot Liberty Valance» e che il Torrazzo di Cremona è stato fatto costruire dai Gonzaga. Pertanto, almeno per conoscenze di questo tipo, è meglio che ci rivolgiamo a Wikipedia.

4 Applicazioni e variazioni sul tema

Ciò che ho presentato nei capitoli precedenti è alla portata di tutti su normalissimi computer, persino su smartphone.

Si tratta infatti di applicazioni che interfacciano attraverso un browser web il lavoro che viene svolto sui potenti server che stanno dietro e il browser rende possibile trasmettere al server la nostra richiesta e al server recapitarci il risultato dell'elaborazione, ma l'elaborazione non avviene sul nostro computer.

A volte questo servizio viene assoggettato a compenso oltre certi livelli di prestazione.
Nel caso di Stable Audio per poter produrre più di 20 brani al mese allungandone la durata, nel caso di Stable Diffusion per produrre le immagini più in fretta, per utilizzare GP4 con l'applicazione ChatGPT Plus, ecc.

Se il software è opensource ciascuno può procurarselo, adattarlo alle proprie esigenze e farlo girare su proprio hardware della potenza necessaria.

Se il software è proprietario occorre pagare i diritti per il suo utilizzo a chi lo ha prodotto.

Per quanto riguarda i modelli utilizzabili attraverso Chatbot Arena, quando apriamo la pagina della classifica scegliendo la linguetta LEADERBOARD possiamo vedere se si tratta di modelli proprietari (proprietary) o liberi (tutti gli altri).

Circa l'utilizzo che viene fatto di questo software possiamo distinguere tra Chatbot di conversazione, che sono quelli di testo aperti a qualsiasi argomento che ho presentato nel Capitolo 3, per i quali il lavoro di machine learning è molto pesante e complesso, e Chatbot di transazione, che sono quelli addestrati su argomenti specifici con machine learning molto ridotto, del tipo di quelli usati per l'assistenza clienti, per rispondere a faq, ecc.

Spesso, soprattutto i chatbot di transazione, sono vocali.

A volte sono vocali anche chatbot di conversazione, come Alexa della Amazon.

La vocalizzazione avviene applicando tecniche di STT (speech to text) a prompt vocali e TTS (text to speech) alle risposte testuali, ma il lavorio che sta sotto è sempre lo stesso.
Per un'informativa sulle cennate tecniche suggerisco la lettura del Capitolo 4 del mio testo «Tutto Python per principianti», disponibile in formato PDF in questa stessa rubrica.

5 Conclusione

Spero che leggendo questo manualetto anche chi non aveva dimestichezza con i nuovi strumenti che ci mette a disposizione quella che chiamiamo intelligenza artificiale abbia imparato qualche cosa.

Comunque la si voglia vedere, siamo in presenza di stravolgimenti sul modo di fare molte cose e come tutti gli stravolgimenti anche questo porterà cose positive.

Me lo fa sperare soprattutto il fatto che quanto è stato realizzato sia il prodotto di forti collaborazioni tra i colossi dell'informatica, con grande ricorso al software libero, garanzia di trasparenza e di progresso per tutti.