

Giuristi e Chat-GPT: a chi l'ultima parola?

scritto da Frank Zappa | 12/02/2024



di Frank Zappa

Una nuova [ricerca](#) dal titolo, che riprende una famosa serie tv, 'Better call GPT?' pubblicata da [AI Center of Excellence di Onit Inc.](#), Nuova Zelanda, ha analizzato un vero e proprio scontro tra giganti nella revisione dei contratti: da una parte gli avvocati junior i Legal Process Outsourcing (LPOs – fornitori esterni di servizi legali), dall'altra i 'recenti' LLM (Large Language Models come per esempio Chat GPT).

Le mansioni erano quelle classiche: analizzare contratti e trovare possibili problemi legali, un compito standard per gli avvocati. La ricerca ha confrontato sugli stessi documenti (10 contratti di appalto), presi da casi reali, i risultati prodotti da LPO, avvocati junior e diversi LLMs (Claude, GTP4, GPT3.5, Palm2).

I risultati sono stati rivisti da avvocati senior, assegnando un punteggio di precisione (quanto un problema rilevato sia categorizzato correttamente) e recupero (quanti problemi sono stati trovati sul totale) rispetto al numero di problemi rilevati nei contratti. I risultati sono sintetizzati nella tabella sottostante, ordinate per Fattore F che è una media di precisione e recupero, mentre, Loss, complemento a 1 di F, indica "l'errore" medio.

	Precision	Recall	F-score	Loss
LPO	0.933	0.823	0.874	0.126
GPT4-1106	0.835	0.910	0.871	0.129
Junior	0.876	0.845	0.860	0.140
Claude 2.0	0.743	0.907	0.817	0.183
GPT4 -32k	0.958	0.723	0.820	0.180
Claude 2.1	0.723	0.917	0.809	0.191
Palm2 text-bison	0.617	0.831	0.708	0.292
GPT3.5	0.531	0.864	0.657	0.343

Table 2. Performance ranked by F-score in determining legal issues

La tabella indica come i fornitori esterni (LPO) siano risultati migliori ma con un margine risicatissimo su GPT4-1106, che a sua volta performa meglio di un avvocato junior. E comunque altri LLM seguono con margini di errore maggiori ma non con una differenza abissale.

Quando si parla di costi e velocità, i LLM hanno spazzato via ogni concorrenza, ma questo non è esattamente una novità: come vediamo nella tabella successiva (dove è indicato il tempo medio che un avvocato senior ha impiegato per rivedere i risultati prodotti dai vari 'competitor') i vari LLM sono da 10 a 100 volte più veloci.

Role	Average Time per Document (Minutes)
Senior Lawyer	43.46
Junior Lawyer	56.17
LPO	201.00
GPT 4-1106	4.70
GPT 4-32k	2.11
GPT 3.5	1.44
Claude 2.1	2.05
Claude 2.0	1.63
Palm2 text-bison	0.73

Table 4. Time Comparison

Si ha un risultato simile, analizzando i costi per documento. Nella tabella successiva vediamo come i costi di un LLM siano 100 o addirittura 1000 volte inferiori.

Role	Average Cost per Document (\$USD)
Senior Lawyer	75.92
Junior Lawyer	74.26
LPO	36.85
GPT 4-1106	0.25
GPT 4-32k	1.24
GPT 3.5	0.05
Claude 2.1	0.02
Claude 2.0	0.02
Palm2 text-bison	0.03

Table 5. Cost Comparison

Ma cosa significano questi dati?

1. I LLM possono a breve per conquistare il mondo dei “contratti ad alto volume/basso rischio”. Non si può ignorare il vantaggio economico.
2. L’efficienza e i risparmi dei LLM sono spettacolari. I primi che adotteranno queste tecnologie in larga scala guadagneranno un vantaggio competitivo enorme, scatenando una specie di corsa all’oro tra gli studi legali.
3. I LLM attuali hanno ancora grandi margini di miglioramento e non dovremo attendere molto.
4. Purtroppo, sembra che i “giovani avvocati” possano essere le vittime di questo processo. La riduzione della domanda di posizioni legali di base è inevitabile.

Una nota ironica: fino a pochi mesi fa la sigla inglese [LLM](#) era solo sinonimo di Dottore in Giurisprudenza nel mondo anglosassone. Come cambiano velocemente le cose, nevvvero?

Foto di [Sergei Tokmakov, Esq. https://Terms.Law](#) da [Pixabay](#)