

Risultati degli Esperimenti di Caricamento (KBox) e di Chat

Di seguito presentiamo i risultati ottenuti da due set di esperimenti:

1. *Esperimenti di caricamento in KBox* (PDF, immagini, video)
2. *Esperimenti di chat* (funzione `cost_of_message`)

I valori sono riportati in tabelle, **senza tralasciare nulla**, e tutti i numeri dell’output originale sono inclusi.

1. Esperimenti di Caricamento in KBox

1.1 Risultati per PDF (pipeline HiRes/Fast, numero pagine {5,10,20,50,100})

Pipeline	Pages	c_{page}	t_{total}	Costo
HiRes	5	0.01	2500	0.050088
HiRes	10	0.01	5000	0.100175
HiRes	20	0.01	10000	0.200350
HiRes	50	0.01	25000	0.500875
HiRes	100	0.01	50000	1.001750
Fast	5	0.001	2500	0.005088
Fast	10	0.001	5000	0.010175
Fast	20	0.001	10000	0.020350
Fast	50	0.001	25000	0.050875
Fast	100	0.001	50000	0.101750

(Osservazioni: $c_{\text{page}} = 0.01$ per la pipeline HiRes, 0.001 per la Fast. Il valore t_{total} è stimato come $\sim 500 \times \text{Pages}$. Tutti i costi sono in dollari.)

1.2 Risultati per Immagini (3 risoluzioni, 2 modelli)

Risoluzione	Modello	Costo
512×512	GPT-4o	0.002785
512×512	GPT-4oMini	0.000108
1024×1024	GPT-4o	0.006609
1024×1024	GPT-4oMini	0.000222
2048×2048	GPT-4o	0.021910
2048×2048	GPT-4oMini	0.000681

(“Modello” indica il LLM (GPT-4o o Mini) per generare la *caption* dell’immagine; costi in dollari.)

1.3 Risultati per Video (durata = 1 min, risoluzioni {512,1024,2048}, frame rate {1.0,0.5,0.2,0.1}, 2 modelli)

Ricordiamo che:

- Durata video = 1 minuto = 60s,
- $\text{rate} \in \{1.0, 0.5, 0.2, 0.1\}$ fps,
- $\text{n_frames} = 60 \times \text{rate}$,
- $\text{sampling_sec} = 60 / \text{n_frames}$.

(a) Risoluzione 512×512

Rate (fps)	Modello	n_frames	Costo
1.0	GPT-4o	60	0.121567
1.0	GPT-4oMini	60	0.004162
0.5	GPT-4o	30	0.060788
0.5	GPT-4oMini	30	0.002085
0.2	GPT-4o	12	0.024320
0.2	GPT-4oMini	12	0.000838
0.1	GPT-4o	6	0.012164
0.1	GPT-4oMini	6	0.000423

(b) Risoluzione 1024×1024

Rate (fps)	Modello	n_frames	Costo
1.0	GPT-4o	60	0.351068
1.0	GPT-4oMini	60	0.011047
0.5	GPT-4o	30	0.175538
0.5	GPT-4oMini	30	0.005527
0.2	GPT-4o	12	0.070219
0.2	GPT-4oMini	12	0.002215
0.1	GPT-4o	6	0.035113
0.1	GPT-4oMini	6	0.001111

(c) Risoluzione 2048×2048

Rate (fps)	Modello	n_frames	Costo
1.0	GPT-4o	60	1.269068
1.0	GPT-4oMini	60	0.038587
0.5	GPT-4o	30	0.634538
0.5	GPT-4oMini	30	0.019297
0.2	GPT-4o	12	0.253820
0.2	GPT-4oMini	12	0.007723
0.1	GPT-4o	6	0.126914
0.1	GPT-4oMini	6	0.003865

2. Esperimenti di Chat (`cost_of_message`)

Parametri:

- `max_pairs` = {10,20,30}
- `n_kbox` = {1,2,3}
- `r_per_kbox` = {5,10,15}
- `model` = {GPT-4o, GPT-4oMini}

Tabella dei risultati (costo in dollari):

max_pairs	n_kbox	r_per_kbox	Model	Costo
10	1	5	GPT-4o	0.022270
10	1	5	GPT-4oMini	0.000732
10	1	10	GPT-4o	0.029770
10	1	10	GPT-4oMini	0.000958
10	1	15	GPT-4o	0.037270
10	1	15	GPT-4oMini	0.001182
10	2	5	GPT-4o	0.029770
10	2	5	GPT-4oMini	0.000958
10	2	10	GPT-4o	0.044770
10	2	10	GPT-4oMini	0.001408
10	2	15	GPT-4o	0.059770
10	2	15	GPT-4oMini	0.001857
10	3	5	GPT-4o	0.037270
10	3	5	GPT-4oMini	0.001182
10	3	10	GPT-4o	0.059770
10	3	10	GPT-4oMini	0.001857
10	3	15	GPT-4o	0.082270
10	3	15	GPT-4oMini	0.002533
20	1	5	GPT-4o	0.032270
20	1	5	GPT-4oMini	0.001032
20	1	10	GPT-4o	0.039770
20	1	10	GPT-4oMini	0.001257
20	1	15	GPT-4o	0.047270
20	1	15	GPT-4oMini	0.001483
20	2	5	GPT-4o	0.039770
20	2	5	GPT-4oMini	0.001257
20	2	10	GPT-4o	0.054770
20	2	10	GPT-4oMini	0.001708
20	2	15	GPT-4o	0.069770
20	2	15	GPT-4oMini	0.002158
20	3	5	GPT-4o	0.047270
20	3	5	GPT-4oMini	0.001483
20	3	10	GPT-4o	0.069770
20	3	10	GPT-4oMini	0.002158
20	3	15	GPT-4o	0.092270
20	3	15	GPT-4oMini	0.002832
30	1	5	GPT-4o	0.042270
30	1	5	GPT-4oMini	0.001332
30	1	10	GPT-4o	0.049770
30	1	10	GPT-4oMini	0.001558
30	1	15	GPT-4o	0.057270
30	1	15	GPT-4oMini	0.001783
30	2	5	GPT-4o	0.049770
30	2	5	GPT-4oMini	0.001558
30	2	10	GPT-4o	0.064770
30	2	10	GPT-4oMini	0.002007
30	2	15	GPT-4o	0.079770
30	2	15	GPT-4oMini	0.002458
30	3	5	GPT-4o	0.057270
30	3	5	GPT-4oMini	0.001783
30	3	10	GPT-4o	0.079770
30	3	10	GPT-4oMini	0.002458
30	3	15	GPT-4o	0.102270
30	3	15	GPT-4oMini	0.003132

(Interpretazione: ogni riga specifica i parametri `max_pairs`, `n_kbox`, `r_per_kbox`, `model` e il costo ottenuto in dollari.)

Osservazioni finali:

- **PDF:** il costo cresce linearmente con il numero di pagine e dipende dal pipeline (HiRes vs. Fast).
- **Immagini:** il costo aumenta con la risoluzione. L'uso di GPT-4o per la caption risulta molto più costoso di GPT-4oMini (fattore $\sim 20\text{--}30\times$).
- **Video** (1 minuto): al crescere del frame rate e della risoluzione, il costo aumenta; GPT-4o rimane più dispendioso di GPT-4oMini.
- **Chat:** dipende da `max_pairs`, `n_kbox`, `r_per_kbox` e dal modello (GPT-4o / GPT-4oMini). Anche qui GPT-4o risulta $\sim 20\text{--}30\times$ più costoso.

Tutti i valori sono riportati **senza tralasciare nulla**, in strutture/tabelle adeguate.

Fine.