

Les collections

Notation {taille_des_chunks_en_caractères}_{nombre de documents}

Collection 256_1000

- Nombre de chunk: 199116
- Temps d'ingest: 7:24
- La collection contient les documents spécifiés par Q&A pour répondre aux questions

Collection 256_950

- Nombre de chunk: 191366
- Temps d'ingest: 6:36
- La collection ne contient pas les documents spécifiés par Q&A pour répondre aux questions

Collection 512_1000

- Nombre de chunk: 86111
- Temps d'ingest: 5:53
- La collection contient les documents spécifiés par Q&A pour répondre aux questions

Collection 512_950

- Nombre de chunk: 82566
- Temps d'ingest: 5:00
- La collection ne contient pas les documents spécifiés par Q&A pour répondre aux questions

Collection 768_1000

- Nombre de chunk: 56208
- Temps d'ingest: 4:59
- La collection contient les documents spécifiés par Q&A pour répondre aux questions

Collection 768_950

- Nombre de chunk: 53865
- Temps d'ingest: 4:09
- La collection ne contient pas les documents spécifiés par Q&A pour répondre aux questions

Les résultats

Les LLM seuls

Test:

Modèle: mistral-7b-instruct-v0.2.Q8_0 en GGUF local Prompt: aucun, uniquement la question de Q&A Temps d'inférence: 1:00:00 Scores:

Métrique	Rouge-1	Rouge-2	Rouge-L	Bleu	Avg Precision (Bleu)	Avg Precision	Avg Recall	Avg F1
Valeur	0.1814	0.0425	0.1216	0.0128	0.0348	0.4667	0.6041	0.5241

Test:

Modèle: gemma-7b-it de Groq Prompt: aucun, uniquement la question de Q&A Temps d'inférence: 1:24 Scores:

Métrique	Rouge-1	Rouge-2	Rouge-L	Bleu	Avg Precision (Bleu)	Avg Precision	Avg Recall	Avg F1
Valeur	0.1364	0.0437	0.1037	0.0143	0.0251	0.4811	0.6347	0.5458

Les SimpleRAG

Ne contient pas les documents spécifiés par Q&A

Test:

Modèle: SimpleRAG avec gemma-7b-it de Groq Prompt: {instruction}{contexte}{question} avec 5 retrieved documents Collection: 256_950 Temps d'inférence: / Scores:

Métrique	Rouge-1	Rouge-2	Rouge-L	Bleu	Avg Precision (Bleu)	Avg Precision	Avg Recall	Avg F1
Valeur	0.244	0.0693	0.1745	0.0377	0.0721	0.5294	0.5993	0.5603

Test:

Modèle: SimpleRAG avec gemma-7b-it de Groq Prompt: {instruction}{contexte}{question} avec 5 retrieved documents Collection: 512_950 Temps d'inférence: / Scores:

Métrique	Rouge-1	Rouge-2	Rouge-L	Bleu	Avg Precision (Bleu)	Avg Precision	Avg Recall	Avg F1
Valeur	0.2341	0.0621	0.1691	0.0317	0.0668	0.519	0.5895	0.55

Test:

Modèle: SimpleRAG avec gemma-7b-it de Groq Prompt: {instruction}{contexte}{question} avec 5 retrieved documents Collection: 768_950 Temps d'inférence: / Scores:

Métrique	Rouge-1	Rouge-2	Rouge-L	Bleu	Avg Precision (Bleu)	Avg Precision	Avg Recall	Avg F1
----------	---------	---------	---------	------	----------------------	---------------	------------	--------

Métrique	Rouge-1	Rouge-2	Rouge-L	Bleu	Avg Precision (Bleu)	Avg Precision	Avg Recall	Avg F1
Valeur	0.213	0.0594	0.152	0.0304	0.0584	0.5075	0.5792	0.5395

Contient les documents spécifiés par Q&A

Test:

Modèle: SimpleRAG avec gemma-7b-it de Groq Prompt: {instruction}{contexte}{question} avec 5 retrieved documents Collection: 256_1000 Temps d'inférence: / Scores:

Métrique	Rouge-1	Rouge-2	Rouge-L	Bleu	Avg Precision (Bleu)	Avg Precision	Avg Recall	Avg F1
Valeur	0.3167	0.1215	0.2253	0.0654	0.0972	0.5867	0.6715	0.6245

Test:

Modèle: SimpleRAG avec gemma-7b-it de Groq Prompt: {instruction}{contexte}{question} avec 5 retrieved documents Collection: 512_1000 Temps d'inférence: / Scores:

Métrique	Rouge-1	Rouge-2	Rouge-L	Bleu	Avg Precision (Bleu)	Avg Precision	Avg Recall	Avg F1
Valeur	0.3098	0.1252	0.2221	0.0697	0.0986	0.5788	0.6703	0.6199

Test:

Modèle: SimpleRAG avec gemma-7b-it de Groq Prompt: {instruction}{contexte}{question} avec 5 retrieved documents Collection: 768_1000 Temps d'inférence: / Scores:

Métrique	Rouge-1	Rouge-2	Rouge-L	Bleu	Avg Precision (Bleu)	Avg Precision	Avg Recall	Avg F1
Valeur	0.3071	0.1298	0.2287	0.0715	0.0975	0.5775	0.6815	0.6231

Les MetadataRAG

Ne contient pas les documents spécifiés par Q&A

Test:

Modèle: MetadataRAG avec gemma-7b-it de Groq Prompt: {instruction}{contexte}{question} avec 5 retrieved documents Collection: 256_950 Temps d'inférence: / Scores:

Métrique	Rouge-1	Rouge-2	Rouge-L	Bleu	Avg Precision (Bleu)	Avg Precision	Avg Recall	Avg F1
Valeur	0.1481	0.0285	0.1081	0.0096	0.0317	0.4633	0.4968	0.4759

Test:

Modèle: MetadataRAG avec gemma-7b-it de Groq Prompt: {instruction}{contexte}{question} avec 5 retrieved documents Collection: 512_950 Temps d'inférence: / Scores:

Métrique	Rouge-1	Rouge-2	Rouge-L	Bleu	Avg Precision (Bleu)	Avg Precision	Avg Recall	Avg F1
Valeur	0.1654	0.0395	0.1232	0.0169	0.0399	0.4767	0.4975	0.483

Test:

Modèle: MetadataRAG avec gemma-7b-it de Groq Prompt: {instruction}{contexte}{question} avec 5 retrieved documents Collection: 768_950 Temps d'inférence: / Scores:

Métrique	Rouge-1	Rouge-2	Rouge-L	Bleu	Avg Precision (Bleu)	Avg Precision	Avg Recall	Avg F1
Valeur	0.1163	0.0161	0.0825	0.0051	0.027	0.4421	0.4361	0.436

Contient les documents spécifiés par Q&A

Test:

Modèle: MetadataRAG avec gemma-7b-it de Groq Prompt: {instruction}{contexte}{question} avec 5 retrieved documents Collection: 256_1000 Temps d'inférence: / Scores:

Métrique	Rouge-1	Rouge-2	Rouge-L	Bleu	Avg Precision (Bleu)	Avg Precision	Avg Recall	Avg F1
Valeur	0.3138	0.1823	0.2584	0.1042	0.1143	0.5845	0.7292	0.6462

Test:

Modèle: MetadataRAG avec gemma-7b-it de Groq Prompt: {instruction}{contexte}{question} avec 5 retrieved documents Collection: 512_1000 Temps d'inférence: / Scores:

Métrique	Rouge-1	Rouge-2	Rouge-L	Bleu	Avg Precision (Bleu)	Avg Precision	Avg Recall	Avg F1
Valeur	0.2956	0.1782	0.2447	0.0965	0.105	0.5771	0.7294	0.6412

Test:

Modèle: MetadataRAG avec gemma-7b-it de Groq Prompt: {instruction}{contexte}{question} avec 5 retrieved documents Collection: 768_1000 Temps d'inférence: / Scores:

Métrique	Rouge-1	Rouge-2	Rouge-L	Bleu	Avg Precision (Bleu)	Avg Precision	Avg Recall	Avg F1

Métrique	Rouge-1	Rouge-2	Rouge-L	Bleu	Avg Precision (Bleu)	Avg Precision	Avg Recall	Avg F1
Valeur	0.278	0.1575	0.2204	0.075	0.0849	0.5723	0.7343	0.6408