





 **Spaces**

 **mteb/leaderboard** 

 like 3.67k

● Running on **CPU UPGRADE**



Massive Text Embedding Benchmark (MTEB) Leaderboard. To submit, refer to the [MTEB GitHub repository](#) 🤖 Refer to the [MTEB paper](#) for details on metrics, tasks and models. Also check out [MTEB Arena](#) ✂️

Search Bar (separate multiple queries with `;`)

🔍

Search for a model and press enter

Model types

☐ Open

☐ Proprietary

☐ Sentence Transformers

☐ Cross-Encoders

☐ Bi-Encoders

☐ Uses Instructions

☐ No Instructions

Model sizes (in number of parameters)

☐ <100M

☐ 100M to 250M

☐ 250M to 500M

☐ 500M to 1B

☐ >1B

Overall

Bitext Mining

Classification

Clustering

Pair Classification

Reranking

Retrieval

STS

Summarization

Retrieval w/Instructions

Retrieval is the task of finding relevant documents for a query.

English

Chinese

French

Polish

Russian

Law

LongEmbed

RAR-b

BRIGHT

BRIGHT Long

CoIR

Retrieval French leaderboard 🇫🇷

◦ Metric: Normalized Discounted Cumulative Gain @ 10 (nDCG@10)

◦ Languages: French

◦ Credits: [Lyon-NLP](#): [Gabriel Sequeira](#), [Imene Kerboua](#), [Wissam Siblini](#), [Mathieu Ciancone](#), [Marion Schaeffer](#)

Rank ▲	Model ▲	Model Size (Million Parameters) ▲	Memory Usage (GB, fp32) ▲
1	bge-multilingual-gemma2	9242	34.4
2	gte-Qwen2-7B-instruct	7613	28.3
3	voyage-multilingual-2		
4	voyage-law-2		
5	gte-multilingual-base	305	1.14

6	<u>intention</u>	306	1.14
7	<u>gte-Qwen2-1.5B-instruct</u>	1776	6.62
8	<u>gte-Qwen2-1.5B-instruct-Q4 K</u>		
9	<u>bilingual-embedding-large</u>	560	2.09
10	bilingual-embedding-large	560	2.09

- Total Datasets: 199
- Total Languages: 113
- Total Scores: 54992