

Exploration of Classification Rules for Hugging Face Model Tree

I . Background

There are numerous models on Hugging Face, each with its corresponding model tree. The model tree generally includes three types of labels as well as a merge flag (indicating whether the model is merged). For example:

Model tree for Qwen/Qwen-Image-2512 ⓘ	
Adapters	22 models
Finetunes	12 models
Quantizations	7 models

This study explores the classification relationships of the aforementioned labels (flags) on the platform.

II . Implementation Approach

Obtain the HF statistical information database compiled officially by Hugging Face for models:

<https://huggingface.co/datasets/cfahlgren1/hub-stats>

Python

```
1 # 获取所有模型数据
2 import pandas as pd
3
4 df = pd.read_parquet(
5     "hf://datasets/cfahlgren1/hub-stats/models.parquet"
6 )
7
8 df.to_parquet(
9     "models.parquet",
10    engine="pyarrow",
11    index=False
12 )
```

A `models.parquet` file will be generated locally.

名称	修改日期	类型	大小
hf.duckdb	2026/1/13 9:50	DUCKDB 文件	12 KB
hf.duckdb.wal	2026/1/13 9:52	WAL 文件	2 KB
hf_models.py	2026/1/12 10:15	Python 源文件	1 KB
hf_models_search.py	2026/1/12 21:36	Python 源文件	4 KB
models.parquet	2026/1/12 11:25	PARQUET 文件	913,370 KB

Visualize and review the dataDownload database software:

① Install DuckDB

- <https://duckdb.org/docs/installation/>

② Install DBeaver Community

- <https://dbeaver.io/>

The data can be viewed as follows:

Beaver 25.3.2 - models

文件(F) 编辑(E) 导航(N) 搜索(A) SQL 编辑器 数据库(D) 窗口(W) 帮助(H)

连接到数据库 > 项目

命名的连接选择器

hf.duckdb

main

表

视图

索引

序列

数据类型

属性

数据

图

模型名: models

表描述:

表名: models

类型: VIEW

Catalog 名称: main

Schema 名称: main

列名 | 数据类型

AZ _id | VARCHAR

AZ id | VARCHAR

AZ author | VARCHAR

AZ cardData | VARCHAR

AZ inferenceProviderMapping | STRUCT<id VARCHAR, adapter VARCHAR, adapterWeightsPath VARCHAR, features STRUCT<structuredOutput BOOLEAN, toolCalling BOOLEAN>>

AZ lastModified | TIMESTAMP_NANOSECOND

AZ likes | BIGINT

AZ trendingScore | DOUBLE

AZ config | VARCHAR

AZ downloads | BIGINT

AZ downloadsAllTime | BIGINT

AZ tags | VARCHAR[]

AZ pipeline_tag | VARCHAR

AZ library_name | VARCHAR

AZ siblings | STRUCT<filename VARCHAR[]>

AZ createdAt | TIMESTAMP_NANOSECOND

AZ safetensors | STRUCT<parameters STRUCT<BF16 DOUBLE, BOOL DOUBLE, F16 DOUBLE, F32 DOUBLE, F64 DOUBLE, F8_E5M2 DOUBLE, F8_E5M2 DOUBLE>, I16 DOUBLE>>

AZ transformersInfo | STRUCT<auto_model VARCHAR, custom_class VARCHAR, pipeline_tag VARCHAR, processor VARCHAR>

AZ baseModels | STRUCT<models STRUCT<id VARCHAR, id VARCHAR[]>, relation VARCHAR>

AZ gguf | VARCHAR

Show SQL | 输入一个SQL表达式来过滤结果(使用Ctrl+Space)

	AZ custom_class	AZ pipeline_tag	AZ processor	models
▶ 235	[NULL]	[NULL]	[NULL]	[NULL]
▶ 236	[NULL]	[NULL]	[NULL]	[NULL]
▶ 237_LM	[NULL]	text-generation	AutoTokenizer	[{"_id": "664dc156dba1a2ae958dc90, id=mistralai/Mistral-7B-v0.3"}]
▶ 238	[NULL]	[NULL]	[NULL]	[NULL]
▶ 239	[NULL]	[NULL]	[NULL]	[NULL]
▶ 240_LM	[NULL]	text-generation	AutoTokenizer	[{"_id": "66e6cfeb6998c3d86c14e163, id=Qwen/Qwen2.5-0.5B"}]
▶ 241_LM	[NULL]	text-generation	AutoTokenizer	[{"_id": "66e81d2494457ccb513ef9c, id=Qwen/Qwen2.5-Coder-7B"}]
▶ 242	[NULL]	[NULL]	[NULL]	[{"_id": "66e81cbc683a3f4e5291bbf, id=Qwen/Qwen2.5-7B-Instruct"}]
▶ 243	[NULL]	[NULL]	[NULL]	[NULL]
▶ 244_LM	[NULL]	text-generation	AutoTokenizer	[{"_id": "67b79cc230e38c400f496b93, id=google/gemma-3-1b-pt"}]
▶ 245	[NULL]	[NULL]	[NULL]	[NULL]
▶ 246	[NULL]	[NULL]	[NULL]	[NULL]
▶ 247_LM	[NULL]	text-generation	AutoTokenizer	[NULL]
▶ 248	[NULL]	[NULL]	[NULL]	[NULL]
▶ 249	[NULL]	[NULL]	[NULL]	[NULL]
▶ 250	[NULL]	[NULL]	[NULL]	[NULL]
▶ 251_Seq	[NULL]	image-to-text	AutoProcessor	[{"_id": "6795ffcd88cd7c0294702a72, id=Qwen/Qwen2.5-VL-7B-Instruct"}]
▶ 252	[NULL]	[NULL]	[NULL]	[NULL]
▶ 253_Seq	[NULL]	image-to-text	AutoProcessor	[NULL]
▶ 254_Seq	[NULL]	image-to-text	AutoProcessor	[NULL]
▶ 255_qLM	[NULL]	text2text-generation	AutoProcessor	[NULL]
▶ 256_qLM	[NULL]	text2text-generation	AutoProcessor	[NULL]
▶ 257_LM	[NULL]	text-generation	AutoTokenizer	[{"_id": "69045311eb7b630406985187, id=Nanbeige/Nanbeige4-3B-Base"}]
▶ 258	[NULL]	[NULL]	[NULL]	[{"_id": "66aaa908fc35e079a941470d, id=black-forest-labs/FLUX.1-dev"}]
▶ 259_qLM	[NULL]	text2text-generation	AutoProcessor	[NULL]
▶ 260	modelling_maincoder.MaincoderForCausalLM	feature-extraction	[NULL]	[NULL]
▶ 261	[NULL]	[NULL]	[NULL]	[NULL]
▶ 262	[NULL]	[NULL]	[NULL]	[{"_id": "69533484e2ff1a5f54e4226f, id=allura-forge/Llama-3.3-8B-Instruct"}]
▶ 263	[NULL]	[NULL]	[NULL]	[{"_id": "695338edbd4a0b8d93bb997, id=Qwen/Qwen-Image-2512"}]
▶ 264	[NULL]	[NULL]	[NULL]	[{"_id": "680a289582ee1640eea64f70, id=ResembleAI/chatterbox"}]
▶ 265	[NULL]	[NULL]	[NULL]	[{"_id": "694e77986d03127b5892d03d, id=LGAI-EXAONE/K-EXAONE-236B-A23B"}]
▶ 266	[NULL]	[NULL]	[NULL]	[{"_id": "6953f0cbdefedb21ff0ffa, id=IQuestLab/IQuest-Coder-V1-40B-Loop-Instruct"}]
▶ 267_LM	[NULL]	text-generation	AutoTokenizer	[{"_id": "694637e104ef97f5fb8e495, id=MinimaxAI/Minimax-M2.1"}]
▶ 268	[NULL]	[NULL]	[NULL]	[{"_id": "6925c6bcc2b55c92698bf47c, id=Tongyi-MAI/Z-Image-Turbo"}]
▶ 269	[NULL]	[NULL]	[NULL]	[{"_id": "6958ee2ba50a975be8b20d03, id=Lightricks/LTX-2"}]

III. Exploring Correlative Relationships



Changelog

NEW Changes July 25th

- added `baseModels` field to models which shows the models that the user tagged as base models for that model

Example:

```
{  
  "models": [  
    {  
      "_id": "687de260234339fed21e768a",  
      "id": "Qwen/Qwen3-235B-A22B-Instruct-2507"  
    }  
  ],  
  "relation": "quantized"  
}
```

NEW Changes July 9th

- Fixed issue with `gguf` column with integer overflow causing import pipeline to be broken over a few weeks

3.1 baseModels

According to the logs, `baseModels` is newly added data starting from July 25, 2025, which is filled in by users voluntarily. After reviewing the data, it was found that a small amount of `baseModels` data also had values before this date, which were derived from the `cardData` field. For example:

1. Models without the `baseModel` field

`cardData`:

C/C++

```
1 {"license": "apache-2.0", "tags": ["generated_from_trainer"], "model_index": [{"name": "PENGMEGJIE-finetuned-emotion", "results": [{"task": {"name": "Text Classification", "type": "text-classification"}}]}]}
```

Its corresponding `baseModel`

SELECT m.* FROM hf.main.models AS m ORDER BY m.created_at DESC

	baseModels	gguf		
class	pipeline_tag	processor	models	relation
▶ 170	text-generation	AutoTokenizer	[NULL]	[NULL]
▶ 171	[NULL]	[NULL]	[NULL]	[NULL]
▶ 172	text-generation	AutoTokenizer	[NULL]	[NULL]
▶ 173	fill-mask	AutoTokenizer	[NULL]	[NULL]
▶ 174	text-generation	AutoTokenizer	[NULL]	[NULL]
▶ 175	[NULL]	[NULL]	[NULL]	[NULL]
▶ 176	[NULL]	[NULL]	[NULL]	[NULL]
▶ 177	automatic-speech-recog	AutoProcessor	[NULL]	[NULL]
▶ 178	automatic-speech-recog	AutoProcessor	[NULL]	[NULL]
▶ 179	[NULL]	[NULL]	[NULL]	[NULL]
▶ 180	automatic-speech-recog	AutoProcessor	[NULL]	[NULL]
▶ 181	fill-mask	AutoTokenizer	[NULL]	[NULL]
▶ 182	fill-mask	AutoTokenizer	[NULL]	[NULL]
▶ 183	[NULL]	[NULL]	[NULL]	[NULL]
▶ 184	[NULL]	[NULL]	[NULL]	[NULL]
▶ 185	[NULL]	[NULL]	[NULL]	[NULL]
▶ 186	text2text-generation	AutoTokenizer	[NULL]	[NULL]
▶ 187	text2text-generation	AutoTokenizer	[{"_id": "621ffdc136468d709f17b862", "id": "gsarti/it5-base"}]	finetune
▶ 188	text2text-generation	AutoTokenizer	[{"_id": "621ffdc136468d709f17b862", "id": "gsarti/it5-base"}]	finetune
▶ 189	text2text-generation	AutoTokenizer	[{"_id": "621ffdc136468d709f17b862", "id": "gsarti/it5-base"}]	finetune
▶ 190	text2text-generation	AutoTokenizer	[{"_id": "621ffdc136468d709f17ae06", "id": "facebook/mbart-large-cc25"}]	finetune
▶ 191	text2text-generation	AutoTokenizer	[{"_id": "621ffdc136468d709f17ae06", "id": "facebook/mbart-large-cc25"}]	finetune
▶ 192	text2text-generation	AutoTokenizer	[{"_id": "621ffdc136468d709f17ae06", "id": "facebook/mbart-large-cc25"}]	finetune
▶ 193	text-classification	AutoTokenizer	[NULL]	[NULL]
▶ 194	[NULL]	[NULL]	[NULL]	[NULL]
▶ 195	[NULL]	[NULL]	[NULL]	[NULL]
▶ 196	text-classification	AutoTokenizer	[NULL]	[NULL]
▶ 197	text-classification	AutoTokenizer	[NULL]	[NULL]

Models with the `baseModel` field

cardData:

C/C++

```
1 {"language": ["it"], "tags": ["summarization"], "datasets": ["ARTELab/mlsum-it"], "metrics": ["rouge"], "base_model": "facebook/mbart-large-cc25", "model-index": [{"name": "summarization_mbart_mlsum", "results": []}]}]
```

Its corresponding `baseModel`:

属性 | 数据 | 图

SELECT m.* FROM hf.main.models AS m ORDER BY m.created_at DESC

ID	AZ pipeline_tag	AZ processor	baseModels		AZ relation	AZ gguf
			models			
170	text-generation	AutoTokenizer	[NULL]		[NULL]	[NULL]
171	[NULL]	[NULL]	[NULL]		[NULL]	[NULL]
172	text-generation	AutoTokenizer	[NULL]		[NULL]	[NULL]
173	fill-mask	AutoTokenizer	[NULL]		[NULL]	[NULL]
174	text-generation	AutoTokenizer	[NULL]		[NULL]	[NULL]
175	[NULL]	[NULL]	[NULL]		[NULL]	[NULL]
176	[NULL]	[NULL]	[NULL]		[NULL]	[NULL]
177	automatic-speech-recog	AutoProcessor	[NULL]		[NULL]	[NULL]
178	automatic-speech-recog	AutoProcessor	[NULL]		[NULL]	[NULL]
179	[NULL]	[NULL]	[NULL]		[NULL]	[NULL]
180	automatic-speech-recog	AutoProcessor	[NULL]		[NULL]	[NULL]
181	fill-mask	AutoTokenizer	[NULL]		[NULL]	[NULL]
182	fill-mask	AutoTokenizer	[NULL]		[NULL]	[NULL]
183	[NULL]	[NULL]	[NULL]		[NULL]	[NULL]
184	[NULL]	[NULL]	[NULL]		[NULL]	[NULL]
185	[NULL]	[NULL]	[NULL]		[NULL]	[NULL]
186	text2text-generation	AutoTokenizer	[NULL]		[NULL]	[NULL]
187	text2text-generation	AutoTokenizer	[{"id": "621ffdc136468d709f17b862", "id": "gsarti/it5-base"}]		finetune	[NULL]
188	text2text-generation	AutoTokenizer	[{"id": "621ffdc136468d709f17b862", "id": "gsarti/it5-base"}]		finetune	[NULL]
189	text2text-generation	AutoTokenizer	[{"id": "621ffdc136468d709f17b862", "id": "gsarti/it5-base"}]		finetune	[NULL]
190	text2text-generation	AutoTokenizer	[{"id": "621ffdc136468d709f17ae06", "id": "facebook/mbart-large-cc25"}]		finetune	[NULL]
191	text2text-generation	AutoTokenizer	[{"id": "621ffdc136468d709f17ae06", "id": "facebook/mbart-large-cc25"}]		finetune	[NULL]
192	text2text-generation	AutoTokenizer	[{"id": "621ffdc136468d709f17ae06", "id": "facebook/mbart-large-cc25"}]		finetune	[NULL]
193	text-classification	AutoTokenizer	[NULL]		[NULL]	[NULL]

3.2 relation

finetune、adapter、quantized

Take ten entries for each of the three types of labels:

 **finetune.csv**
文件大小: 65.7 KB

 **quantized.csv**
文件大小: 86.21 KB

 **adapter.csv**
文件大小: 16.64 KB

① finetune:

`cardData.base_model` must exist.

`tags` almost certainly contain:

base_model:xxx

base_model:finetune:xxx

The complete weights (not adapters) are in the repo.

Reverse inference rules:

💡 if:

- `cardData.base_model` exists
- The model has complete weights (not LoRA / not adapters)

then:

`relation = finetune`

② adapter:

`tags` contain:

lora

adapter

peft

In the repo files:

No complete `model.safetensors` exists

Only adapter weights (e.g., `adapter_model.safetensors`) are present

`baseModels.relation = adapter`

Reverse inference rules:

💡 if:

- `cardData.base_model` exists
- The `tags` / repo structure indicates it is PEFT / LoRA

then:

`relation = adapter`

③ **quantized:**

`tags` contain:

quantized

gguf

int8, int4, 4bit, 8bit

Strong characteristic features in filenames:

*.gguf

int8.safetensors

baseModels.relation = quantized

Reverse inference rules:



if:

- `cardData.base_model` exists
- The `tags` or filenames indicate it is a quantized version

then:

`relation = quantized`

IV. Verify the aforementioned hypotheses

4.1 Whether `finetune` truly represents "fine-tune"

SQL

```
1 SELECT
2   baseModels.relation,
3   COUNT(*) AS cnt,
4   SUM(
5     CASE
6       WHEN list_contains(tags, 'lora')
7         OR list_contains(tags, 'adapter')
8         OR list_contains(tags, 'peft')
9       THEN 1 ELSE 0
10    END
11  ) AS suspicious_adapter_like
12 FROM models
13 WHERE baseModels.relation = 'finetune'
14 GROUP BY baseModels.relation;
15
```

The screenshot shows a DuckDB interface with the following details:

SQL Editor:

```
SELECT
    baseModels.relation,
    COUNT(*) AS cnt,
    SUM(
        CASE
            WHEN list_contains(tags, 'lora')
                OR list_contains(tags, 'adapter')
                OR list_contains(tags, 'peft')
            THEN 1 ELSE 0
        END
    ) AS suspicious_adapter_like
FROM models
WHERE baseModels.relation = 'finetune'
GROUP BY baseModels.relation;
```

Results:

relation	cnt	suspicious_adapter_like
finetune	244,723	1,130

- The `cnt` value is significantly large
- `suspicious_adapter_like` ≈ 0 or extremely small

Verification result: Normal

4.2 Whether `adapter` truly represents "adapter"

SQL

```
1 SELECT
2   COUNT(*) AS total_adapter,
3   SUM(
4     CASE
5       WHEN
6         list_contains(tags, 'lora')
7         OR list_contains(tags, 'adapter')
8         OR list_contains(tags, 'peft')
9       THEN 1 ELSE 0
10    END
11  ) AS has_adapter_tag
12 FROM models
13 WHERE baseModels.relation = 'adapter';
14
```

The screenshot shows a DuckDB interface with the following details:

SQL Query:

```
SELECT
    COUNT(*) AS total_adapter,
    SUM(
        CASE
            WHEN
                list_contains(tags, 'lora')
                OR list_contains(tags, 'adapter')
                OR list_contains(tags, 'peft')
            THEN 1 ELSE 0
        END
    ) AS has_adapter_tag
FROM models
WHERE baseModels.relation = 'adapter';
```

Results:

	total_adapter	has_adapter_tag
1	266,041	262,423

has_adapter_tag / total_adapter $\approx 98\%$



Verification result: Normal

MaLi888/test

Token Classification, BERTopic, biology, License: apache-2.0

Model card, Files and versions, Community, Settings, Use this model

Getting started with your model

Complete model information
Add metadata and complete your model card to make your model more discoverable.

Edit Model Card

Push your model files
Upload your model weights to this repository.

CLI Python, Git, HTTPS, SSH

```
# Install the Hugging Face CLI
powershell -ExecutionPolicy ByPass -c "irm https://hf.co/cli/install.ps1 | iex"
# (optional) Login with your Hugging Face credentials
hf auth login
# Push your model files
hf upload MaLi888/test .
```

You can also upload directly from the website using the File Uploader.

Downloads last month
Downloads are not tracked for this model. [How to track](#)

Inference Providers NEW
Token Classification
This model isn't deployed by any Inference Provider.
Ask for provider support

Model tree for MaLi888/test

Base model
Quantized, Finetuned (1)

zai-org/GLM-4.7-GGUF, bartowski/zai-org,GLM-4.7-GGUF, this model

<hfduckdb> Script, main, models, *<hfduckdb> hf.main.models

```
SELECT *
FROM hf.main.models
where id = 'bdsqlsz/qinglong_DetailedEyes_Z-Image';
```

结果 1

SELECT * FROM hf.main.models where id = 'bdsqlsz/qingl' | 输入一个 SQL 表达式来过滤结果 (使用 Ctrl+Space)

id	elAuthor	performance	AZ provider	AZ providerDetails	AZ providerId	AZ status	AZ task
1	[NULL]	fal-ai	[NULL]	[2]	fal-ai/z-image/turbo/lora wavespeed-ai/z-image/turbo-lora	[2]	live
1.1							text-to-image
1.2							text-to-image
1.3							
1.4							
1.5							

Define **tags** for self-uploaded models:

MaLi888/test

Image-to-Image, lora, arxiv:1910.09700, License: apache-2.0

Model card, Files and versions (selected), Community, Settings

main / test / README.md

MaLi888 Update README.md 5552ae3 VERIFIED

Preview, Code, raw, Copy download link, history, blame, edit, delete

```
metadata
license: apache-2.0
base_model:
- tencent/HY-MT1.5-1.8B-GGUF
- zai-org/GLM-4.7
- bartowski/zai-org_GLM-4.7-GGUF
tags:
- lora
pipeline_tag: image-to-image
```

Also review the existing models with the `adapter` tag:

bdsqsz/qinglong_DetailedEyes_Z-Image

Text-to-Image, Diffusers, lora, License: apache-2.0

Model card, Files and versions (selected), Community 3

main / qinglong_DetailedEyes_Z-Image / README.md

bdsqsz Update README.md 0a91266 VERIFIED

Preview, Code, raw, Copy download link, history, blame, contribute, delete, Safe

```
metadata
license: apache-2.0
base_model:
- Tongyi-MAI/Z-Image-Turbo
pipeline_tag: text-to-image
library_name: diffusers
tags:
- lora
```

4.3 Whether `quantized` truly represents "quantized"

SQL

```
1 SELECT
2   COUNT(*) AS total_quantized,
3   SUM(
4     CASE
5       WHEN
6         list_contains(tags, 'quantized')
7         OR list_contains(tags, 'gguf')
8         OR list_contains(tags, 'int8')
9         OR list_contains(tags, 'int4')
10        OR list_contains(tags, '4bit')
11        OR list_contains(tags, '8bit')
12      THEN 1 ELSE 0
13    END
14  ) AS has_quant_tag
15 FROM models
16 WHERE baseModels.relation = 'quantized';
17
```

```
ht.duckdb ht <hf.duckdb> Script main models *<hf.duckdb> hf.main.models X

SELECT
    COUNT(*) AS total_quantized,
    SUM(
        CASE
            WHEN
                list_contains(tags, 'quantized')
                OR list_contains(tags, 'gguf')
                OR list_contains(tags, 'int8')
                OR list_contains(tags, 'int4')
                OR list_contains(tags, '4bit')
                OR list_contains(tags, '8bit')
            THEN 1 ELSE 0
        END
    ) AS has_quant_tag
FROM models
WHERE baseModels.relation = 'quantized';
```



结果 1 X

	total_quantized	has_quant_tag
1	125,673	101,205



The probability is already greater than 80%, indicating that the hypotheses are basically correct.

4.4 Source of the `merge` tag

`merge` indicates that the model is the `baseModel` of other models, and the other models have more than one `baseModel`.

The verification process is as follows:

The screenshot shows a 'Model tree' interface with the following categories and counts:

- Adapters: 136 models
- Finetunes: 378 models
- Merges: 13 models (highlighted with a red box)
- Quantizations: 177 models

The Qianwen model (abbreviation) has the `Merges` tag, but the model is found to have no `baseModel` upon database query.

The screenshot shows a database query interface with the following details:

- Query: `SELECT * FROM hf.main.models WHERE id = 'Qwen/Qwen3-4B-Instruct-2507';`
- Results:

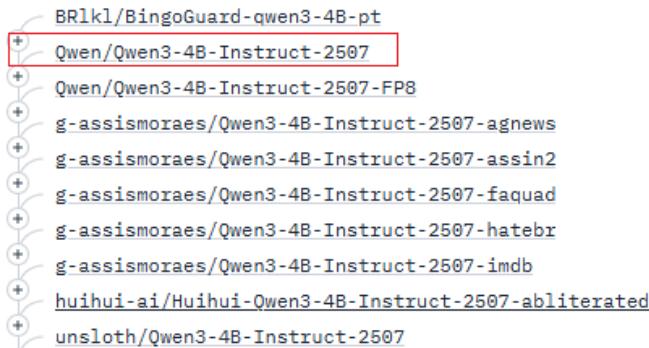
id	baseModels
Qwen/Qwen3-4B-Instruct-2507	[NULL]

Click on the model `ParrotRouter/Qwen3-4B-Instruct-2507-20250808-233922-0` in the `merges` section:

This model isn't deployed by any Inference Provider.

 Ask for provider support

Model tree for ParrotRouter/Qwen3-4B-Instruct-2507-20250808-233922-0 ⓘ



Merge model

this model

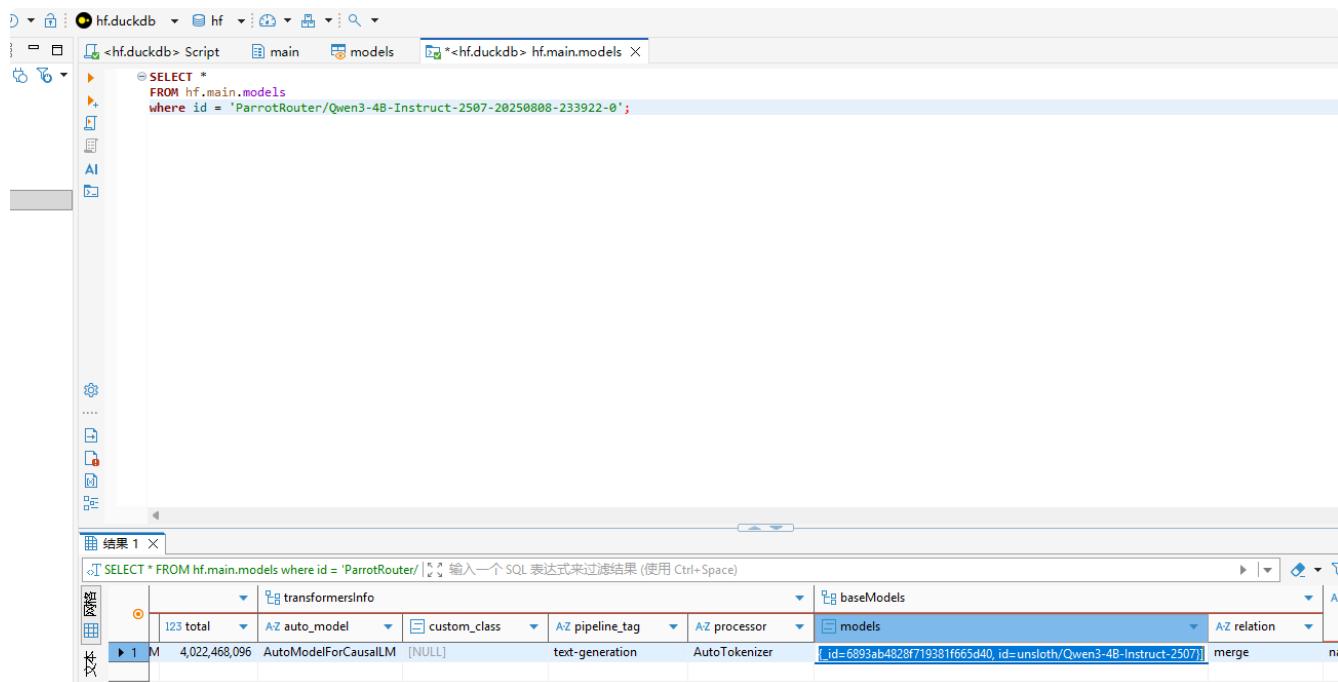
Merges

2 models

Quantizations

2 models

It is found that the corresponding Qianwen model exists in its `model tree`, and the database query results are as follows:



```
SELECT *  
FROM hf.main.models  
WHERE id = 'ParrotRouter/Qwen3-4B-Instruct-2507-20250808-233922-0';
```

transformersInfo	baseModels
123 total	AutoModelForCausalLM [NULL]
4,022,468,096	text-generation
	AutoTokenizer
	_id=6893ab4828f719381f665d40, id=unslloth/Qwen3-4B-Instruct-2507

baseModels of this model

[{_id=68956805ea6b891bf6a55a78, id=BRlkl/BingoGuard-qwen3-4B-pt},
{_id=6891e3bb084ce75acffb033d, id=Qwen/Qwen3-4B-Instruct-2507},

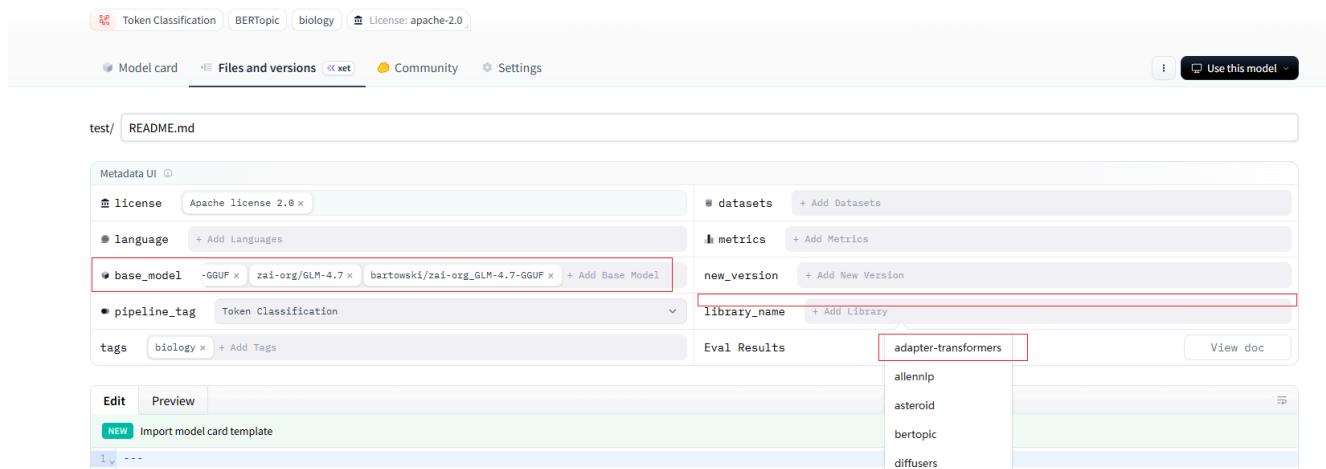
```
{_id=6892fb7f7e589749f520da30, id=Qwen/Qwen3-4B-Instruct-2507-FP8},  
{_id=6893bbd0a3c083bbfc22b0a2, id=g-assismoraes/Qwen3-4B-Instruct-2507-  
agnews}, {_id=6893aa104f29d79eef316513, id=g-assismoraes/Qwen3-4B-  
Instruct-2507-assin2}, {_id=6893b1529e0760254790bfa1, id=g-  
assismoraes/Qwen3-4B-Instruct-2507-faquad},  
{_id=6893b5756a284e3c9f0a984d, id=g-assismoraes/Qwen3-4B-Instruct-2507-  
hatebr}, {_id=68939105ae69d218f46a573c, id=g-assismoraes/Qwen3-4B-Instruct-  
2507-imdb}, {_id=6894967025b5b79c17de8b27, id=huihui-ai/Huihui-Qwen3-4B-  
Instruct-2507-abliterated}, {_id=6893ab4828f719381f665d40, id=unsloth/Qwen3-  
4B-Instruct-2507}]
```

It is verified that its `base models` are consistent with the `model tree`.

V. Upload models to verify `cardData` and the source of `tags`.

Introduction to Model Cards:

<https://huggingface.co/docs/hub/model-cards#evaluation-results>



Focus on the `base model` and `library_name` options.

The `base model` option specifies which models can be selected as the base models for the uploaded model, while the `library_name` option determines whether the model is a `finetune` or an `adapter`.

Unique `tags` can also be edited in `cardData`, and the background logical reasoning mechanism will identify the tags in `tags` and then determine whether it is a `finetune` or an `adapter`.

VI. Identify models where the `baseModel` and `adapter` are released simultaneously

case	Meaning
Base + Adapter in the same repository	A standalone delivery unit
Adapter in a separate repository	Clearly depends on an external base model
Base model in a separate repository	A standard base model

6.1 Judgment Criteria

In the file list (siblings) of the same model repository, complete base model weight files and adapter-specific files appear simultaneously.

SQL

```
1 WITH repo_flags AS (
2     SELECT
3         id,
4         bool_or(
5             lower(r.rfilename) LIKE '%model.safetensors%'
6             OR lower(r.rfilename) LIKE '%pytorch_model.bin%'
7         ) AS has_base,
8         bool_or(
9             lower(r.rfilename) LIKE '%adapter_model%'
10            OR lower(r.rfilename) LIKE '%adapter_config%'
11        ) AS has_adapter
12     FROM models,
13     UNNESTsiblings) AS s(r)
14     GROUP BY id
15 )
16     SELECT *
17     FROM repo_flags
18     WHERE has_base = TRUE
19     AND has_adapter = TRUE
20
```

hf.duckdb > Script main models *<hf.duckdb> hf.main.models X

```

WITH repo_flags AS (
    SELECT
        id,
        bool_or(
            lower(r.rfilename) LIKE '%model.safetensors%'
            OR lower(r.rfilename) LIKE '%pytorch_model.bin%'
        ) AS has_base,
        bool_or(
            lower(r.rfilename) LIKE '%adapter_model%'
            OR lower(r.rfilename) LIKE '%adapter_config%'
        ) AS has_adapter
    FROM models,
    UNNEST(siblings) AS s(r)
    GROUP BY id
)
SELECT *
FROM repo_flags
WHERE has_base = TRUE
    AND has_adapter = TRUE

```

结果 1 X

WITH repo_flags AS (SELECT id, bool_or(lower(r.rfilename)) AS has_base, bool_or(lower(r.rfilename)) AS has_adapter FROM models, UNNEST(siblings) AS s(r) GROUP BY id) SELECT * FROM repo_flags WHERE has_base = TRUE AND has_adapter = TRUE

	AZ id	has_base	has_adapter
1	charvabannur/OpenAssistant-Phi2-QLoRA	[v]	[v]
2	Keerthana98/peft-starcoder-lora-a100-v3	[v]	[v]
3	alifzl/SQLChef-1.8B	[v]	[v]
4	Chat-Error/Llama-3-Kimiko-LoRA	[v]	[v]
5	andmev/lora_model	[v]	[v]
6	SALUTEASD/Qwen-Qwen1.5-0.5B-1726079652	[v]	[v]
7	xuejj/Qwen-Qwen1.5-1.8B-1726665376	[v]	[v]
8	rejjieQ/Qwen-Qwen1.5-0.5B-1727158860	[v]	[v]
9	downtown1/Qwen-Qwen1.5-0.5B-1727658753	[v]	[v]
10	jbreuch/cs329-dpo	[v]	[v]
11	eeeebbb2/50cb5edb-0039-45a3-a254-b8681a16d6ba	[v]	[v]
12	dimasik87/15fe4780-5642-4be8-a12e-d53036d8db64	[v]	[v]
13	Raj-hf/Llama3.2-Vision-Radiology	[v]	[v]
14	morturr/Mistral-7B-v0.1-PAIR_headlines_one_liners-COMB-one_liners-comb-3-seed-28-2025-02-04	[v]	[v]
15	clembench-playpen/meta_llama_KTO_KTO_Wordle_ExperimentAborted_ErrorsInAbortedOnly	[v]	[v]
16	nwbawwa/DeepSeek-R1-Distill-Llama-8B-unsloth-bnb-4bit-Medical-COT	[v]	[v]
17	chukwubuikemstephen/llama3-8b-Medical-COT	[v]	[v]
18	dzanbek/c67f49c6-3e32-471e-923b-be5eb67f1936	[v]	[v]
19	fevohh/GenExtract-3B-v0	[v]	[v]
20	gauravparajuli/florence2_4_r16	[v]	[v]
21	aamijar/Llama-2-7b-hf-lora-r8-mrpc-portlora-epochs1	[v]	[v]
22	darturi/Llama-3.1-8B-Instruct_es_RFA_ADAPTER-orthogonal-vaccine2	[v]	[v]
23	laniqo/WMT25-EuroLLM-9B-CPO	[v]	[v]
24	SalimBou5/dpo_model_sft_mmlu_1	[v]	[v]
25	sidvash/famus_multi_instance_v3_unsloth_llama-3-8b-Instruct-bnb-4bit-3-epochs	[v]	[v]

刷新 保存 取消 导出数... 200 400+ ... 400 行已获取 - 9.689s (0.001s 耗取), 2024-02-01 14:45:21

6.2 Verification

Do base + adapter files exist, yet the semantic meaning is not "adapter" (meeting the above characteristics but not labeled as `adapter` on HF)?

SQL

```
1 SELECT
2   m.id,
3   m.tags
4 FROM models m,
5     UNNEST(m.siblings) AS s(unnest)
6 GROUP BY m.id, m.tags
7 HAVING
8   bool_or(lower(s.unnest.rfilename) LIKE '%adapter_model%')
9   AND
10  bool_or(lower(s.unnest.rfilename) LIKE '%model.safetensors%')
11  AND NOT (
12    array_to_string(m.tags, ',') ILIKE '%adapter%'
13    OR array_to_string(m.tags, ',') ILIKE '%lora%'
14  )
15 LIMIT 10;
16
```

The screenshot shows a DuckDB interface with a query editor and a results table. The query in the editor is:

```

SELECT
    m.id,
    m.tags
FROM models m,
    UNNEST(m.siblings) AS s(unnest)
GROUP BY m.id, m.tags
HAVING
    bool_or(lower(s.unnest.rfilename) LIKE '%adapter_model%')
    AND
    bool_or(lower(s.unnest.rfilename) LIKE '%model.safetensors%')
    AND NOT (
        array_to_string(m.tags, ',') ILIKE '%adapter%'
        OR array_to_string(m.tags, ',') ILIKE '%lora%'
    )
LIMIT 10;

```

The results table has two columns: AZ id and tags. The AZ id column lists model names, and the tags column lists model tags. A red box highlights the entry 'blake41/lora_model' in the AZ id column.

AZ id	tags
1 yuri-no/gemma-argos	▶ transformers [+4]
2 Katyc/llama-3-8b-bnb-4bit	▶ transformers [+11]
3 aengusl/25jun_r2d2_lat_eps1pt5_lr2e-5_layers8-16_ckpt240	▶ transformers [+4]
4 blake41/lora_model	[12] <ul style="list-style-type: none"> ▶ transformers safetensors text-generation-inference unsloth llama trl en base_model:unsloth/llama-3-8b-Instruct-bnb-4t base_model:finetune:unsloth/llama-3-8b-Instruc license:apache-2.0 endpoints_compatible region:us
5 amit-chhabria/llama-3-8b-chat-doctor	▶ transformers [+8]
6 dalaug627/text-to-sql-finetuned-mistral-small-2409-lora	▶ transformers [+4]
7 nitish-11/friends Joey_trained_Llama-3-8B	▶ transformers [+9]

It is also found that some models with LoRA are not defined as `adapter`, indicating that there is a discrepancy between semantics and physical structure. Can such data be defined as `ambiguous`?

VII. Summary

- A tag relationship exists **only if** a `baseModel` is present;
- `finetune` means a `baseModel` has been selected;

- The display of `adapter` depends on either the `library_name` or the inclusion of keywords such as "lora" in the repository;
- `merge` indicates the presence of multiple `baseModels` ;
- For `relation` , `merge` is prioritized for display, followed by `adapter` and then `finetune` .