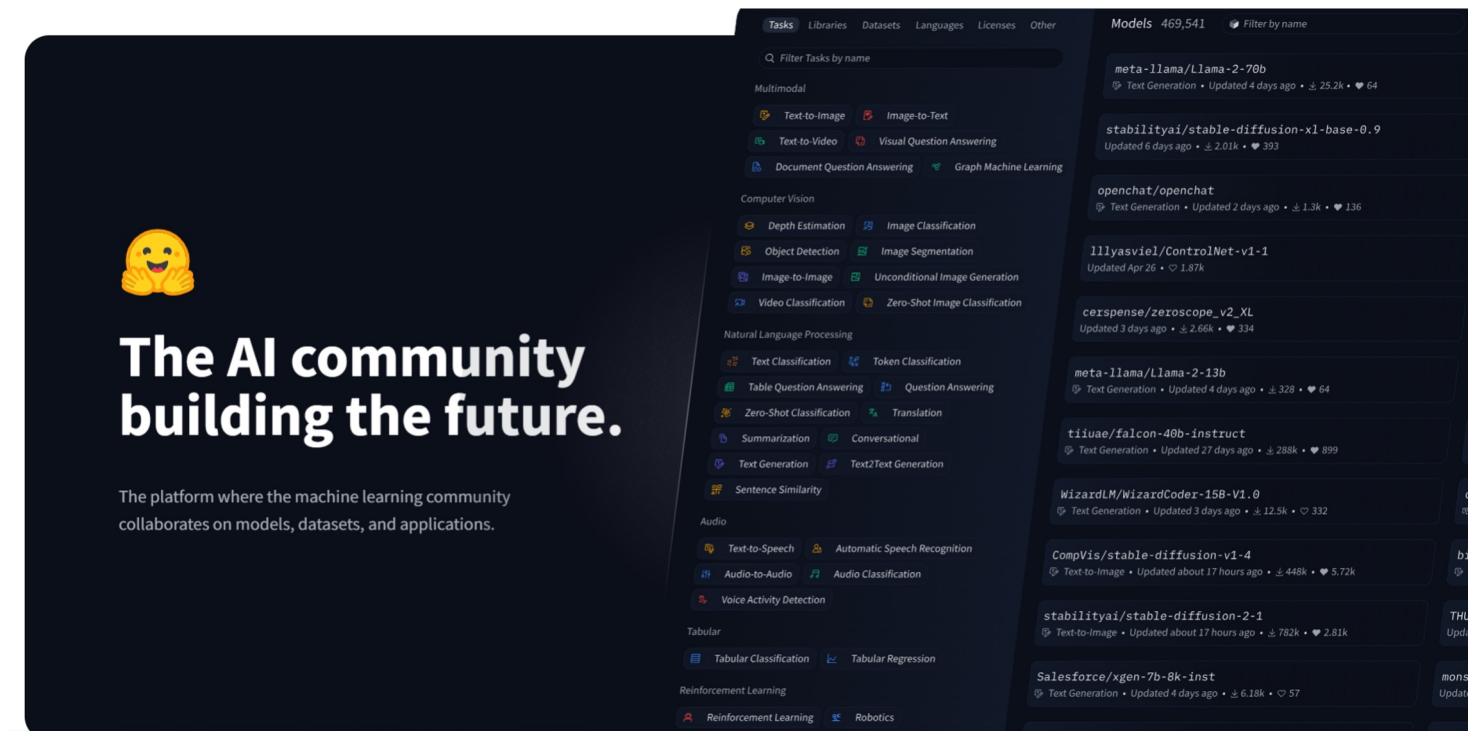


# Huggingface

Credit to TA.Cheetah

# What is Huggingface?

- HuggingFace is an AI community that promotes open source contributions. It is a hub of open source models for Natural Language Processing, computer vision, and other fields where AI plays its role. Even the tech giants like Google, Facebook, AWS, Microsoft, and others use the models, datasets, and libraries.



# Datasets in Huggingface

The screenshot displays the Huggingface Datasets interface. On the left, there are navigation tabs for 'Tasks', 'Sizes', 'Sub-tasks', 'Languages', 'Licenses', and 'Other'. Below these, a search bar 'Filter Tasks by name' is present. The main content area is divided into three sections: 'Multimodal', 'Computer Vision', and 'Natural Language Processing', each with a grid of task-specific filters. On the right, a 'Datasets' section shows a list of trending datasets, with the count '56,783' highlighted in a red box. The list includes datasets like 'allenai/dolma', 'garage-bAInd/Open-Platypus', 'BAAI/COIG-PC', 'Open-Orca/OpenOrca', 'timdettmers/openassistant-guanaco', 'OpenAssistant/oasst1', 'ehartford/dolphin', 'fka/awesome-chatgpt-prompts', 'PygmalionAI/PIPPA', 'google/dreambooth', 'b-mc2/sql-create-context', 'databricks/databricks-dolly-15k', 'togethercomputer/RedPajama-Data-1T', and 'mlabonne/guanaco-llama2-1k'. Each dataset entry shows its name, a 'Preview' or 'Viewer' link, the update time, and download/like counts.

**Tasks** Sizes Sub-tasks Languages Licenses Other

Filter Tasks by name

Multimodal

- Feature Extraction Text-to-Image
- Image-to-Text Text-to-Video
- Visual Question Answering Graph Machine Learning

Computer Vision

- Depth Estimation Image Classification
- Object Detection Image Segmentation
- Image-to-Image Unconditional Image Generation
- Video Classification Zero-Shot Image Classification

Natural Language Processing

- Text Classification Token Classification
- Table Question Answering Question Answering
- Zero-Shot Classification Translation
- Summarization Conversational
- Text Generation Text2Text Generation

**Datasets** 56,783 Filter by name new Full-text search ↑↓ Sort: Trending

- allenai/dolma**  
Preview • Updated 4 days ago • ↓ 49 • ♥ 179
- garage-bAInd/Open-Platypus**  
Viewer • Updated 8 days ago • ↓ 4.09k • ♥ 125
- BAAI/COIG-PC**  
Preview • Updated 12 days ago • ↓ 373 • ♥ 166
- Open-Orca/OpenOrca**  
Viewer • Updated 3 days ago • ↓ 20.7k • ♥ 601
- timdettmers/openassistant-guanaco**  
Viewer • Updated May 28 • ↓ 37.1k • ♥ 184
- OpenAssistant/oasst1**  
Viewer • Updated May 2 • ↓ 20.9k • ♥ 978
- ehartford/dolphin**  
Preview • Updated 23 days ago • ↓ 2.27k • ♥ 152
- fka/awesome-chatgpt-prompts**  
Viewer • Updated Mar 7 • ↓ 3.96k • ♥ 3.09k
- PygmalionAI/PIPPA**  
Preview • Updated 9 days ago • ↓ 271 • ♥ 69
- google/dreambooth**  
Viewer • Updated 8 days ago • ↓ 8 • ♥ 26
- b-mc2/sql-create-context**  
Viewer • Updated Apr 21 • ↓ 1.38k • ♥ 68
- databricks/databricks-dolly-15k**  
Viewer • Updated Jul 1 • ↓ 42.8k • ♥ 325
- togethercomputer/RedPajama-Data-1T**  
Viewer • Updated Jul 1 • ↓ 17k • ♥ 828
- mlabonne/guanaco-llama2-1k**  
Viewer • Updated 28 days ago • ↓ 11.3k • ♥ 34

Ref: <https://huggingface.co/datasets?sort=trending>

# Models in Huggingface

The screenshot displays the Huggingface Models interface. On the left, there are navigation tabs: **Tasks**, Libraries, Datasets, Languages, Licenses, and Other. Below these is a search bar 'Filter Tasks by name'. The main content area is divided into three sections: Multimodal, Computer Vision, and Natural Language Processing, each with several task-specific buttons. On the right, a 'Models' tab is selected, showing a list of 301,886 models. A red box highlights the 'Models' tab and the count '301,886'. The list of models includes:















- stabilityai/control-lora**: Text-to-Image • Updated 4 days ago • 289
- meta-llama/Llama-2-7b**: Text Generation • Updated Jul 20 • 2k
- Open-Orca/OpenOrca-Platypus2-13B**: Text Generation • Updated 2 days ago • 15.3k • 137
- diffusers/controlnet-canny-sdxl-1.0**: Text-to-Image • Updated 10 days ago • 12.1k • 298
- facebook/seamless-m4t-large**: Updated about 17 hours ago • 75
- stabilityai/stablecode-instruct-alpha-3b**: Text Generation • Updated 15 days ago • 7.02k • 229
- TheBloke/Llama-2-7B-Chat-GGML**: Text Generation • Updated 28 days ago • 8.42k • 369
- stabilityai/stable-diffusion-xl-base-1.0**: Text-to-Image • Updated 19 days ago • 856k • 2.11k
- Deci/DeciCoder-1b**: Text Generation • Updated about 17 hours ago • 2.59k • 141
- meta-llama/Llama-2-7b-chat-hf**: Text Generation • Updated 14 days ago • 474k • 865
- meta-llama/Llama-2-70b-chat-hf**: Text Generation • Updated 14 days ago • 174k • 1.1k
- garage-bAInd/Platypus2-70B-instruct**: Text Generation • Updated 3 days ago • 3.77k • 107
- runwayml/stable-diffusion-v1-5**: Text-to-Image • Updated Jul 5 • 7.77M • 9.02k
- defog/sqlcoder**: Text Generation • Updated 1 day ago • 430 • 56

Ref: <https://huggingface.co/models>

















# Examples notebook

## PyTorch Examples

### Natural Language Processing

Notebook	Description	
<a href="#">Train your tokenizer</a>	How to train and use your very own tokenizer	 <a href="#">Open in Colab</a>  <a href="#">Open in Studio Lab</a>
<a href="#">Train your language model</a>	How to easily start using transformers	 <a href="#">Open in Colab</a>  <a href="#">Open in Studio Lab</a>
<a href="#">How to fine-tune a model on text classification</a>	Show how to preprocess the data and fine-tune a pretrained model on any GLUE task.	 <a href="#">Open in Colab</a>  <a href="#">Open in Studio Lab</a>
<a href="#">How to fine-tune a model on language modeling</a>	Show how to preprocess the data and fine-tune a pretrained model on a causal or masked LM task.	 <a href="#">Open in Colab</a>  <a href="#">Open in Studio Lab</a>
<a href="#">How to fine-tune a model on token classification</a>	Show how to preprocess the data and fine-tune a pretrained model on a token classification task (NER, PoS).	 <a href="#">Open in Colab</a>  <a href="#">Open in Studio Lab</a>
<a href="#">How to fine-tune a model on question answering</a>	Show how to preprocess the data and fine-tune a pretrained model on SQUAD.	 <a href="#">Open in Colab</a>  <a href="#">Open in Studio Lab</a>
<a href="#">How to fine-tune a model on multiple choice</a>	Show how to preprocess the data and fine-tune a pretrained model on SWAG.	 <a href="#">Open in Colab</a>  <a href="#">Open in Studio Lab</a>

## Computer Vision

Notebook	Description	
<a href="#">How to fine-tune a model on image classification (Torchvision)</a>	Show how to preprocess the data using Torchvision and fine-tune any pretrained Vision model on Image Classification	 <a href="#">Open in Colab</a>  <a href="#">Open in Studio Lab</a>
<a href="#">How to fine-tune a model on image classification (Albumentations)</a>	Show how to preprocess the data using Albumentations and fine-tune any pretrained Vision model on Image Classification	 <a href="#">Open in Colab</a>  <a href="#">Open in Studio Lab</a>
<a href="#">How to fine-tune a model on image classification (Kornia)</a>	Show how to preprocess the data using Kornia and fine-tune any pretrained Vision model on Image Classification	 <a href="#">Open in Colab</a>  <a href="#">Open in Studio Lab</a>
<a href="#">How to perform zero-shot object detection with OWL-ViT</a>	Show how to perform zero-shot object detection on images with text queries	 <a href="#">Open in Colab</a>  <a href="#">Open in Studio Lab</a>
<a href="#">How to fine-tune an image captioning model</a>	Show how to fine-tune BLIP for image captioning on a custom dataset	 <a href="#">Open in Colab</a>  <a href="#">Open in Studio Lab</a>
<a href="#">How to build an image similarity system with Transformers</a>	Show how to build an image similarity system	 <a href="#">Open in Colab</a>  <a href="#">Open in Studio Lab</a>
<a href="#">How to fine-tune a SegFormer model on semantic segmentation</a>	Show how to preprocess the data and fine-tune a pretrained SegFormer model on Semantic Segmentation	 <a href="#">Open in Colab</a>  <a href="#">Open in Studio Lab</a>
<a href="#">How to fine-tune a VideoMAE model on video classification</a>	Show how to preprocess the data and fine-tune a pretrained VideoMAE model on Video Classification	 <a href="#">Open in Colab</a>  <a href="#">Open in Studio Lab</a>

Fine-tune a pretrained model: <https://huggingface.co/docs/transformers/training>

Ref: <https://huggingface.co/docs/transformers/notebooks>

# Error with tensorboard

- If tensorboard error, you will change 'Block third-party cookies' to 'Allow third-party cookies' or 'Block third-party cookies in incognito mode'.

