



Hugging Face

Hugging Face 勉強会

この資料は下記GitHubの一部を再利用して作っています。

<https://github.com/nlp-with-transformers/notebooks>

是非ともこの本を手にとって勉強してください

目次

- ・ 自己紹介
- ・ 勉強会の目的
- ・ Hugging Faceを使おうと思ったわけ
- ・ Hugging Faceとは
- ・ Hugging Faceサイトの構成と見方
- ・ デモンストレーション
 - ・ デモンストレーションの前に
 - ・ Google Colabで使ってみる
- ・ UseCaseを考えてみよう
- ・ まとめ

本日の想定レベル：☆☆○○○

Python : ちょっとわかる

LLM : ちょっとわかる

英語 : ほとんどわかんない

自己紹介

- ・ザ・JTC企業で製造現場の効率化の支援
(システム導入/開発/AI活用)
- ・1974年生まれの50歳



Community of
Deep Learning Evangelists
CDLE

勉強会の私の目的

Hugging Faceを使う仲間作り

- なかなか日本語の資料がない
- 日本語の書籍が少ない
- Udemyにも講座がない
- あわよくばGENIACみたいなのに参加したい

[GENIAC 採択プロジェクト] 多様な日本語能力の向上を目指した公開の基盤モデル開発 | 東京大学松尾研究室 - Matsuo Lab (u-tokyo.ac.jp)



'23企業協働メンバーが二人参加！！(カワイイ)

Hugging Faceを使おうと思ったわけ

- Transformerの理解をしたくて買った本
→ 大半がTransformersの本だった。
- いろんなモデルがここ一箇所ですべて探せる/試せるというより、
メモリーに乗り切らない学習データを扱えるところ

メモリーマッピング

- メモリーに学習データを読み込むのではなく、ディスク上のファイルを直接読み込むことでメモリーの節約を行う

ストリーミング

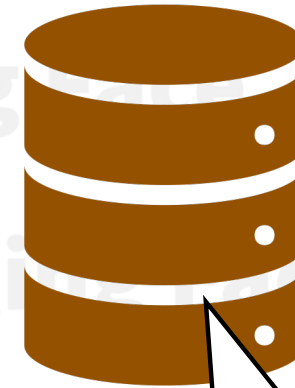
- 大量の学習データをローカルPCに置くのではなく、その時に必要な分だけをダウンロードしながら学習を進めていく



Hugging Faceとは 1



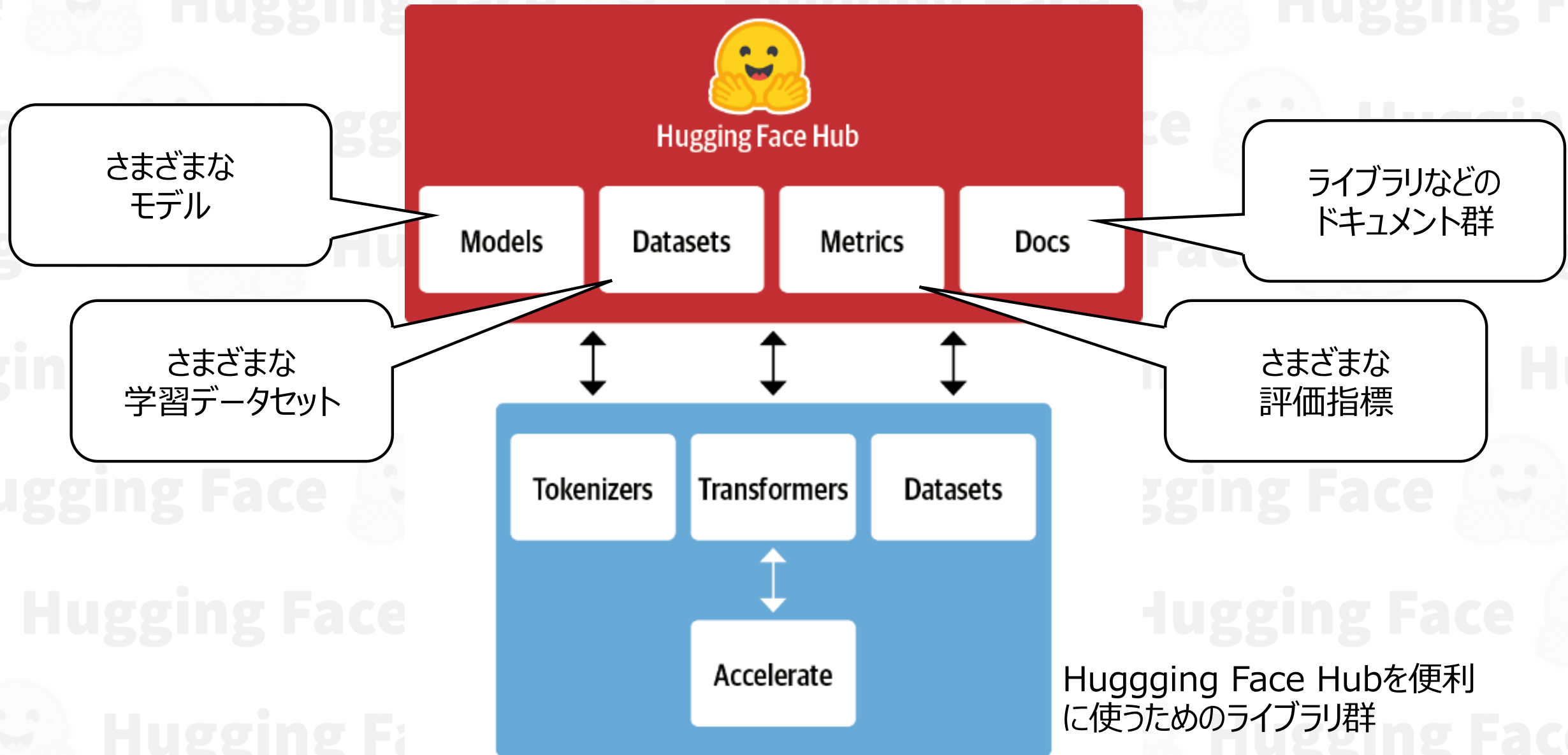
さまざまな
モデル



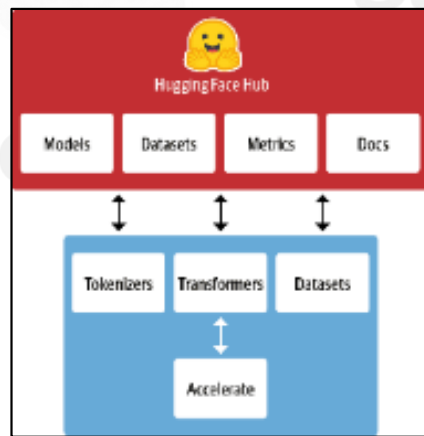
さまざまな
データセット

モデルやデータセット、使い方などを共有し、容易に活用
(AIモデルのGitHub、辞書みたいなもの)

Hugging Faceとは 2



Hugging Faceサイトの構成と見方



Organizations

- Create New

Resources

- Hub guide
- Transformers doc
- Forum
- Tasks
- Learn

Dark theme

Navigation Bar (B)

- Models
- Datasets
- Spaces
- Posts
- Docs
- Solutions
- Pricing

Left Sidebar (A)

- New
- ikedachin
 - Profile
 - Inbox (0)
 - Settings
 - Get Pro
- Organizations
 - Create New
- Resources
 - Hub guide
 - Transformers doc
 - Forum
 - Tasks
 - Learn
- Dark theme

Main Content Area

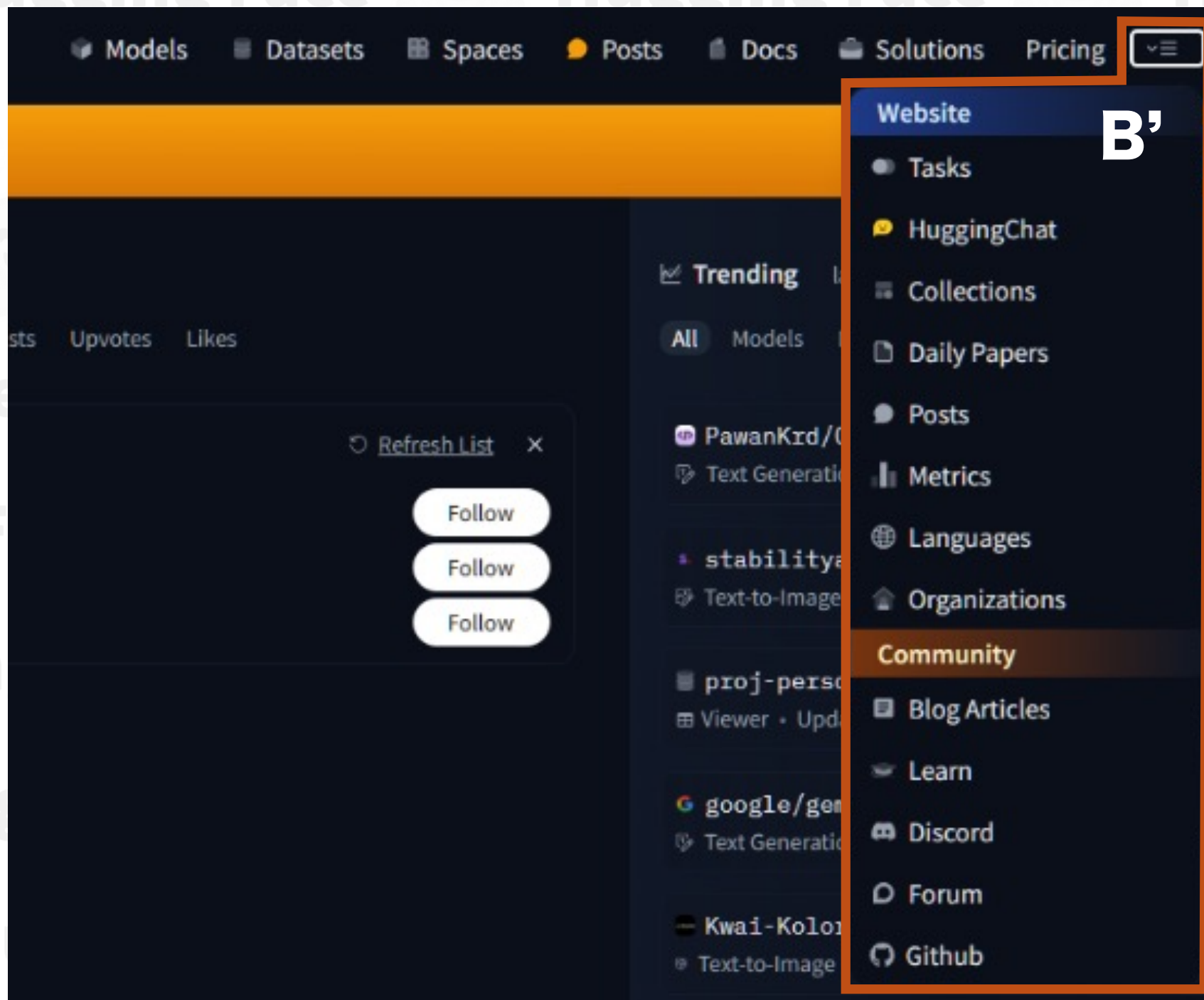
Following

Follow your favorite AI creators

- ilyasvick - Creator of ControlNet and other image gen gems
- ylecun - The godfather of deep learning
- bwang0911 - Research on sentence embedding models

Trending

- PawanKrd/CosmosRP-8k
- stabilityai/stable-diffusion-3-medium
- proj-persona/PersonaHub
- google/gemma-2-9b
- meta-llama/Meta-Llama-3-8B
- BAAI/Infinity-Instruct
- Salesforce/xlam-function-calling-60k



B'

デモンストレーション

- 結局、使ってみた方が理解が早い！
- 「Textbooks Are All You Need」の論文で有名になったPhi-1の派生
“**Phi-1.5** (Microsoft)” を使ってみよう
- [論文はこちら](#)
- [HuggingfaceのPhi-1.5のページ](#)
- [Google Colaboratory](#)

必要なもの

- Googleアカウント（Google Colaboratoryを使う場合）
- HuggingFaceアカウント（必須ではないと思うけど・・・）

デモンストレーションの前に 1

Phi-1.5のページを見てみる

The image shows a screenshot of the Hugging Face website for the model `microsoft/phi-1_5`. The interface is dark-themed. At the top, there's a navigation bar with 'Model card', 'Files and versions', and 'Community 87'. Below this, a banner says 'Hugging Face is way more fun with friends and colleagues!'. The model name 'microsoft/phi-1_5' is displayed with '1.3k' likes. A row of tags includes 'Text Generation', 'Transformers', 'Safetensors', 'English', 'phi', 'nlp', 'code', 'Inference Endpoints', 'text-generation-inference', 'arxiv:2309.05463', and 'License: mit'. Three orange boxes with white numbers highlight specific areas: Box 1 points to the 'Model card' tab; Box 2 points to the 'Train', 'Deploy', and 'Use this model' buttons; Box 3 points to the 'Model Summary' section. The 'Model Summary' text describes the model as a Transformer with 1.3 billion parameters, trained on synthetic texts, and notes that it was not fine-tuned for instruction following. The right sidebar shows 'Downloads last month: 85,902', 'Safetensors' download options, and an 'Inference API' section with a text input field containing 'My name is Lewis and I like to'.

Model card Files and versions Community 87

Hugging Face Search models, datasets, users... Models Datasets Spaces Posts Docs Solutions Pricing

Hugging Face is way more fun with friends and colleagues! Join an organization

microsoft/phi-1_5 like 1.3k

Text Generation Transformers Safetensors English phi nlp code Inference Endpoints text-generation-inference arxiv:2309.05463 License: mit

Model card Files and versions Community 87

Model Summary

The language model Phi-1.5 is a Transformer with 1.3 billion parameters. It was trained using the same data sources as phi-1, augmented with a new data source that consists of various NLP synthetic texts. When assessed against benchmarks testing common sense, language understanding, and logical reasoning, Phi-1.5 demonstrates a nearly state-of-the-art performance among models with less than 10 billion parameters.

We did not fine-tune Phi-1.5 either for instruction following or through reinforcement learning

Downloads last month 85,902

Safetensors Model size 1.42B params Tensor type FP16

Inference API

Text Generation Examples

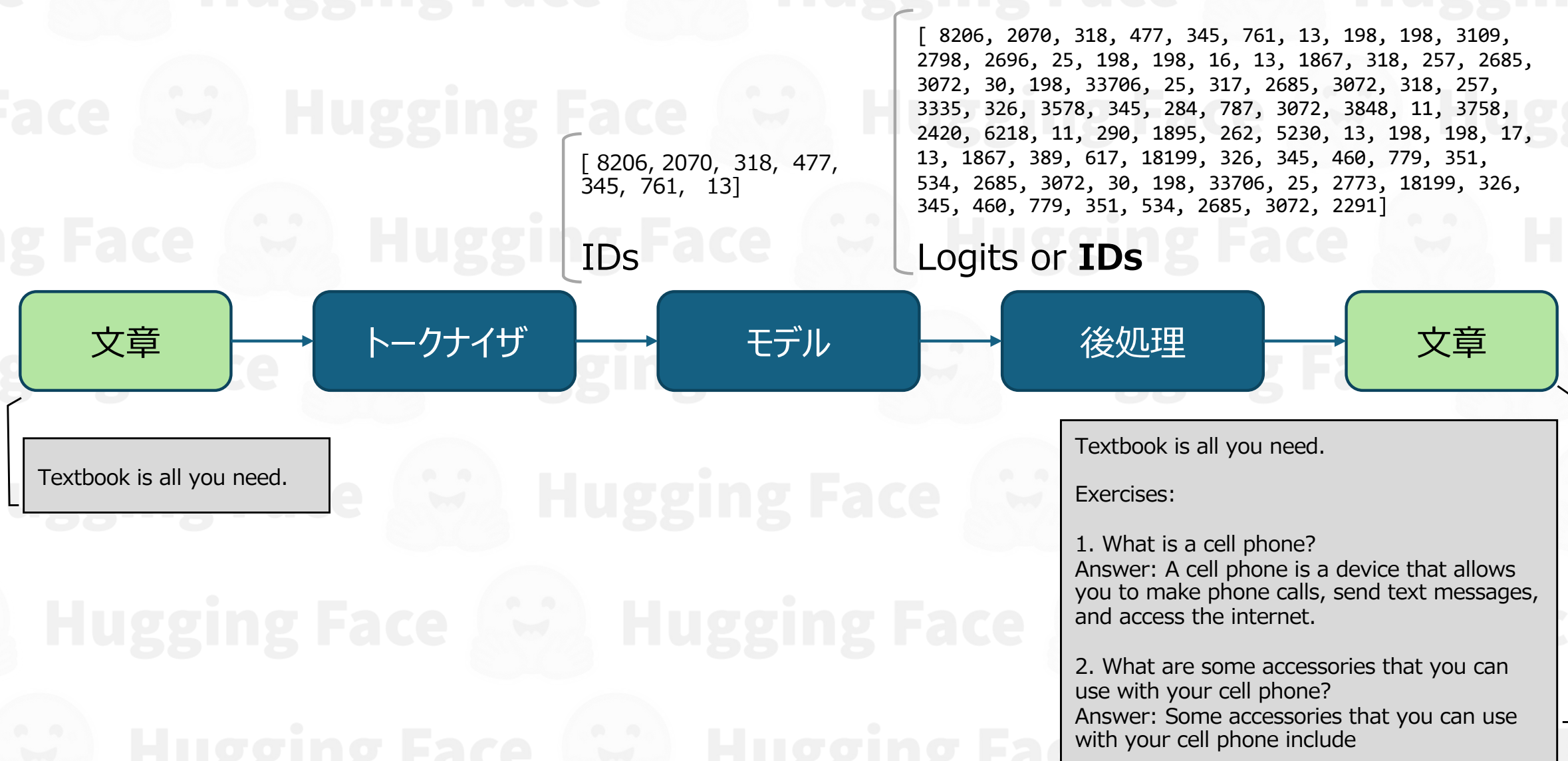
My name is Lewis and I like to

Train Deploy Use this model

https://huggingface.co/microsoft/phi-1_5

デモンストレーションの前に 2

LLMの処理をイメージする



GoogleColabで使ってみよう！

[Google Colab](#) ← click

[HuggingfaceのPhi-1.5のページ](#) ← click

UseCaseを考えてみよう

- ローカルLLMに活路？
 - 社内特化？特定の目的特化？

- 一般的でない専門用語

視差呼称、はつり、デレッキ、うま

逃げ、刺す、芋掘り、三味線...



まとめ

- ちょっと難しそうだけど、なんやら便利そう！
- 日本語の資料が少ないから、みんなで頑張ろうぜ！

みんなでHuggingFaceを使ってみよう！



補足

Datasetsの便利なところ

- 強力なmap関数
Pythonのmap関数より遥かに強力
- 使いやすいDataset形式
ほぼ、辞書形式
- stream
必要なデータだけを取得しながら使える

Input IDs

Text 1	101	23	74	2	67	102	0	0	0	0	0	0	0
Text 2	101	14	66	53	7	87	14	37	31	17	9	21	102
Text 3	101	91	20	15	98	36	81	85	23	102	0	0	0

Attention masks

1	1	1	1	1	1	0	0	0	0	0	0	0	0
1	1	1	1	1	1	1	1	1	1	1	1	1	1
1	1	1	1	1	1	1	1	1	1	1	0	0	0



Stack vectors to input matrices



Input batch

101	23	74	2	67	102	0	0	0	0	0	0	0
101	14	66	53	7	87	14	37	31	17	9	21	102
101	91	20	15	98	36	81	85	23	102	0	0	0

1	1	1	1	1	1	0	0	0	0	0	0	0	0
1	1	1	1	1	1	1	1	1	1	1	1	1	1
1	1	1	1	1	1	1	1	1	1	0	0	0	0

Login

Make sure to have access to it at <https://huggingface.co/google/gemma-2-9b>.

401 Client Error. (Request ID: Root=1-66873ee3-45a5c76834cae61a35b65299;8994a406-1c2a-4867-9a54-7e105b88f969)

Cannot access gated repo for url <https://huggingface.co/google/gemma-2-9b/resolve/main/config.json>.

Access to model google/gemma-2-9b is restricted. You must be authenticated to access it.

Output is truncated. View as a [scrollable element](#) or open in a [text editor](#). Adjust cell output [settings...](#)

Windows固有の設定

開発者向け

これらの設定は開発目的での使用のみを意図しています。

[詳細情報](#)

開発者モード

圧縮されていないファイルも含め、任意のソースからのアプリをインストールします。

☒ オン