Hugging Face

Hugging Face

Study Notes — API, Models & Pipelines

1. Hugging Face Access Token (API Key)

An access token is a secure string of characters used to authenticate and access Hugging Face services and resources. The API key and access token are the same thing.

1.1 When You NEED a Token

- **Private / gated models:** e.g., Meta's LLaMA requires authentication
- **Hugging Face Inference API:** to make API calls
- **Uploading to the Hub:** uploading models, datasets, or Spaces

1.2 When You DON'T Need a Token

- **Public models:** e.g., GPT-2 is freely downloadable
- **Using the Transformers library locally:** most open-source models don't require a key

1.3 How to Create a Token

- **Step 1:** Go to huggingface.co/join and create an account
- **Step 2:** Verify your email
- **Step 3:** Click your profile → Access Tokens → Create New Token
- **Step 4:** Add a token name, click Create Token, copy and save it — you won't see it again
- **Note:** You can edit permissions or delete the token anytime. Never share your token.

2. Download a Dataset from Hugging Face

A dataset is a collection of structured data used for training, evaluating, or testing ML models. Hugging Face hosts thousands of datasets for various NLP and ML tasks.

2.1 How to Download

Use the load_dataset() function from the datasets library. It can download datasets from the Hugging Face Hub or load from local files.

Function	`load_dataset('dataset_name')`
Example	`load_dataset('imdb')`

2.2 IMDB Dataset Example

- **Purpose:** Movie reviews labeled positive or negative (sentiment classification)
- **Train split:** 25,000 rows — features: text (review), label (sentiment)
- **Test split:** 25,000 rows — same features, for evaluation
- **Unsupervised split:** 50,000 rows — no labels, used for pre-training or semi-supervised learning

3. Download a Model from Hugging Face

Use the from_pretrained() method from the transformers library to download model weights, configuration, and tokenizer from the Hugging Face Hub.

3.1 BERT-base-uncased Output Example

For input "hello hugging face", the last hidden state output shape is (1, 7, 768):

- **1:** Batch size — one input sentence
- **7:** Sequence length — tokenized input including special tokens
- **768:** Hidden size — each token represented as a 768-dimensional vector (standard for BERT base)

4. Sentiment Analysis

Sentiment analysis determines the sentiment expressed in a piece of text — positive, negative, or neutral.

4.1 Types of Sentiment Analysis

- **Polarity Detection:** "I love this product" → Positive | "The service is terrible" → Negative | "The package arrived on time" → Neutral
- **Emotion Detection:** Anger: "This is so frustrating" | Joy: "I'm thrilled about the results"
- **Aspect-Based:** "The food was great but the service was slow" → food: positive, service: negative
- **Intent Analysis:** "Where can I buy this product?" → Purchase intent

4.2 How to Perform Sentiment Analysis

- **Library:** Use the pipeline() function from the transformers library
- **Input:** Pass a list of strings to analyze multiple texts at once
- **Output:** List of dictionaries with label (POSITIVE/NEGATIVE) and score (0 to 1)

4.3 Understanding the Output Score

- **Score close to 1:** Model is very confident in its prediction

- **Score close to 0:** Model is less confident
- **High scores occur because:** Model fine-tuned on large datasets; text contains strong, unambiguous language

4.4 Performance Tip

For complex or large-scale projects on Google Colab, change the runtime to T4 GPU or V2-8 TPU for faster inference.

5. Text Classification

Text classification categorizes text into predefined classes. It is a broader task than sentiment analysis.

5.1 Use Cases

- **Spam Detection:** Classify emails as spam or not spam
- **News/Document Categorization:** Sports, Technology, Politics, etc.
- **Intent Detection:** Cancel an order, book a flight, etc.

5.2 Sentiment Analysis vs. Text Classification

Feature	Sentiment Analysis	Text Classification
Scope	Narrow — specific to sentiment	Broad — task-dependent labels
Labels	Positive, Negative, Neutral	Spam/Not Spam, Sports/Tech/...

5.3 Label Mapping for Spam Detection

- **NEGATIVE:** → Spam
- **NEUTRAL / POSITIVE:** → Not Spam

Note: Low confidence scores may indicate the model was fine-tuned for sentiment, not spam specifically. Use a dedicated spam-detection model for better accuracy.

6. Text Summarization

Summarization condenses long content into shorter, meaningful output. Real-world uses include summarizing articles, research papers, chatbot responses, and key points from large datasets.

6.1 Key Parameters

max_length	`Maximum number of tokens in the summary`
min_length	`Minimum number of tokens in the summary`

length_penalty	`Value of 2 encourages longer summaries`
num_beams	`Beam search width; value of 4 = 4 beams (higher = better quality, slower)`

6.2 Process

- **Step 1:** Install transformers and PyTorch
- **Step 2:** Load model with AutoModel (seq2seq LM) and tokenizer with AutoTokenizer
- **Step 3:** Tokenize the input text
- **Step 4:** Call generate() with the parameters above
- **Step 5:** Decode output tokens to text

7. Text-to-Text Translation

Translation is part of text-to-text generation. This class of models takes an input sequence and generates an output sequence — covering translation, summarization, paraphrasing, and QA.

7.1 Text Generation vs. Text-to-Text Generation

Feature	Text Generation	Text-to-Text Generation
Type	Autoregressive, one token at a time	Sequence-to-sequence (input → output)
Use Cases	Dialogue systems, text completions	Translation, summarization, paraphrasing

7.2 T5-Small Model

- **Publicly available on Hugging Face:**
- **Tasks:** Summarization, translation, question answering
- **Usage:** Prefix input with a task prompt, e.g., "translate English to Spanish: ..."
- **Parameters:** max_length and num_beams for customizing generation

8. Question Answering

Question answering extracts an answer from a given context. The model needs both a context paragraph and a question.

8.1 Process

- **Step 1:** Load a pre-trained QA model and tokenizer
- **Step 2:** Tokenize both the context and question
- **Step 3:** Pass tokenized input to the model

- **Step 4:** Model extracts start and end scores to find the answer span
- **Step 5:** Convert token IDs back to words to display the answer

Example: Given a context about a person, the question "Where is this person based?" will extract the location directly from the context.

9. Text to Image Generation

9.1 Key Library: Diffusers

The diffusers library is an open-source Python library by Hugging Face focused on diffusion models for generating images, audio, and other data.

9.2 Stable Diffusion

- **Type:** Latent diffusion model for high-quality image generation
- **Input:** Text prompt
- **Output:** Generated image (PNG file)
- **Example prompt:** "Flying cars soar over a futuristic cityscape at sunset"

9.3 Process

- **Step 1:** Install required libraries (diffusers, etc.)
- **Step 2:** Load the Stable Diffusion pipeline
- **Step 3:** Pass a text prompt to the pipeline
- **Step 4:** Save the output as a PNG file

10. Text to Video Synthesis

Text-to-video synthesis generates a sequence of video frames based on a textual description. It is a complex task combining NLP models with generative/diffusion models.

10.1 Popular Frameworks

- **Runway ML:** Tools for video generation and editing
- **Pictory:** AI-generated videos
- **DeepMind's Perceiver IO:** Handles multimodal inputs — text, images, videos
- **PyTorch / TensorFlow:** Used to build custom video generation pipelines

10.2 Process Using Hugging Face

- **Step 1:** Use the diffusers library to load a pre-trained text-to-image model (Stable Diffusion)

- **Step 2:** Set a text prompt, e.g., "A futuristic cityscape at night with flying cars"
- **Step 3:** Generate individual frames using a for-in loop (e.g., 10 frames)
- **Step 4:** Use OpenCV library to stitch frames into a video
- **Step 5:** Use NumPy arrays in the process

10.3 Output

- **Frame files:** frame_0.png through frame_9.png (individual images)
- **Final video:** output_video.mp4 (all stitched frames)

LINKS

GOOGLE COLAB —-> Hugging Face.ipynb

Video Link —-> Hugging Face Crash Course | Learn Hugging Face in 1 hour | Am…

hugging face (bangla short)

১. হাগিং ফেস পরিচিতি

হাগিং ফেস কী?

- একটি বিখ্যাত কোম্পানি ও ওপেন-সোর্স কমিউনিটি
- মূল ফোকাস → NLP (ন্যাচারাল ল্যাঙ্গুয়েজ প্রসেসিং) ও AI

তিনটি প্রধান লাইব্রেরি:

- Transformers
- Datasets
- Tokenizers

মূল বৈশিষ্ট্যসমূহ:

- Model Hub → প্রি-ট্রেইনড মডেল শেয়ার ও ডাউনলোড
- Datasets Library → বিভিন্ন NLP ডেটাসেট
- Spaces → ML ডেমো ও অ্যাপ হোস্টিং প্ল্যাটফর্ম
- প্রোডাকশনে সহজ মডেল ডিপ্লয়মেন্ট
- শক্তিশালী ডেভেলপার কমিউনিটি

জনপ্রিয় মডেলসমূহ:

মডেল	পূর্ণরূপ	বিশেষত্ব
BERT	Bidirectional Encoder Representations from Transformers	দ্বিমুখী পদ্ধতিতে শব্দের প্রসঙ্গ বোঝে
GPT	Generative Pre-trained Transformer	টেক্সট জেনারেশন
RoBERTa	Robustly Optimized BERT Approach	BERT-এর উন্নত সংস্করণ, ডায়নামিক মাস্কিং

⚠️ গুরুত্বপূর্ণ: RoBERTa-র মূল লক্ষ্য হলো BERT-এর সীমাবদ্ধতা সমাধান করে পারফরম্যান্স উন্নত করা।

২. হাগিং ফেসের ব্যবহারের ক্ষেত্র

প্রধান ব্যবহারের ক্ষেত্রসমূহ:

🤖 কনভার্সেশনাল **AI** (চ্যাটবট)

- GPT, Blenderbot মডেল ব্যবহার করে
- কাস্টমার সাপোর্ট, শিক্ষামূলক অ্যাসিস্ট্যান্ট, থেরাপি বট

📊 সেন্টিমেন্ট অ্যানালাইসিস

- গ্রাহকের ফিডব্যাক ও সোশ্যাল মিডিয়া বিশ্লেষণ

- পজিটিভ / নেগেটিভ / নিউট্রাল নির্ধারণ

📝 টেক্সট সামারাইজেশন

- দীর্ঘ ডকুমেন্ট সংক্ষিপ্তকরণ
- মিটিং নোট তৈরি

🏷️ নেমড এন্টিটি রিকগনিশন (**NER**)

- রেজুমে থেকে তথ্য বের করা
- স্বাস্থ্যসেবা ও ফিনান্স ডোমেইনে ব্যবহার

🌐 মেশিন ট্রান্সলেশন

- এক ভাষা থেকে অন্য ভাষায় অনুবাদ
- কম সম্পদযুক্ত ভাষায়ও প্রযোজ্য

❓ কোয়েশ্চেন আন্সারিং

- কাস্টমার সাপোর্ট, FAQ, শিক্ষা ক্ষেত্রে

🎤 স্পিচ রিকগনিশন ও সিনথেসিস

- স্পিচ-টু-টেক্সট ও টেক্সট-টু-স্পিচ
- রিয়েল-টাইম ক্যাপশনিং, ভিজ্যুয়াল QA

🎯 অন্যান্য উল্লেখযোগ্য ক্ষেত্র:

- রেকমেন্ডেশন সিস্টেম
- জালিয়াতি শনাক্তকরণ
- মানসিক স্বাস্থ্য পর্যবেক্ষণ
- হেট স্পিচ মডারেশন
- শেয়ার বাজার পূর্বাভাস
- মাল্টিমোডাল অ্যাপ্লিকেশন (টেক্সট + ভিডিও + অডিও)
- AR অ্যাপ্লিকেশন
- মডেল ফাইন-টিউনিং ও তুলনা

৩. ট্রান্সফর্মার্স লাইব্রেরি

কী? → প্রি-ট্রেইনড মডেল ও পাইপলাইনের মূল ওপেন-সোর্স পাইথন লাইব্রেরি

কেন ব্যবহার করবেন?

- জটিল NLP মডেল সহজে ব্যবহার করা যায়
- অত্যাধুনিক মডেলে অ্যাক্সেস
- বড় সক্রিয় কমিউনিটি
- কাস্টমাইজেশন ও ফাইন-টিউনিং সমর্থন
- অন্যান্য টুলের সাথে ইন্টিগ্রেশন সম্ভব

ব্যবহারের ক্ষেত্র:

- টেক্সট ক্লাসিফিকেশন (স্প্যাম শনাক্তকরণ)
- NER (নাম, তারিখ, অবস্থান চিহ্নিতকরণ)
- টেক্সট ট্রান্সলেশন
- টেক্সট জেনারেশন (GPT)
- কোয়েশ্চেন আন্সারিং

ইনস্টলেশন:

```
pip install transformers        ← সাধারণ
!pip install transformers       ← Google Colab
```

৪. ডেটাসেটস লাইব্রেরি

কী? → NLP ও ML কাজের জন্য বিভিন্ন ডেটাসেটে সহজ অ্যাক্সেসের পাইথন লাইব্রেরি

কেন ব্যবহার করবেন?

- লেজি লোডিং ও স্ট্রিমিং → বড় ডেটাসেটে দক্ষতা
- ইউনিফাইড API
- ট্রান্সফর্মার্স ও ML ফ্রেমওয়ার্কের সাথে ইন্টিগ্রেশন
- ৩৫০K+ ডেটাসেট উপলব্ধ (huggingface.co/datasets)
- কাস্টম ডেটাসেট সমর্থন

ব্যবহারের ক্ষেত্র:

- টেক্সট ক্লাসিফিকেশন, সেন্টিমেন্ট অ্যানালাইসিস
- কোয়েশ্চেন আন্সারিং, ট্রান্সলেশন
- NER, কাস্টম ডেটাসেট প্রি-প্রসেসিং

ইনস্টলেশন:

```
pip install datasets            ← সাধারণ
!pip install datasets           ← Google Colab
git+github.com/huggingface/datasets  ← GitHub থেকে
```

৫. টোকেনাইজার্স লাইব্রেরি

কী? → টেক্সট ডেটা টোকেনাইজ করার দ্রুত, দক্ষ ও নমনীয় লাইব্রেরি

টোকেনাইজেশন কী?

টেক্সটকে ছোট ইউনিটে (শব্দ/সাব-ওয়ার্ড/অক্ষর) বিভক্ত করে সংখ্যাসূচক রূপে রূপান্তর করার প্রক্রিয়া

কেন ব্যবহার করবেন?

- বড় ডেটাসেটেও দ্রুত টোকেনাইজেশন
- কাস্টম টোকেনাইজার ও একাধিক অ্যালগরিদম সমর্থন
- ট্রান্সফর্মার্সের সাথে ইন্টিগ্রেশন
- সহজ API (টোকেনাইজিং, ডিকোডিং, ভোকাবুলারি)

- প্রি-ট্রেইনড টোকেনাইজার অ্যাক্সেস

ব্যবহারের ক্ষেত্র:

- স্প্যাম শনাক্তকরণ, সেন্টিমেন্ট অ্যানালাইসিস
- মেশিন ট্রান্সলেশন, টেক্সট জেনারেশন
- ডোমেইন-স্পেসিফিক টোকেনাইজার তৈরি

ইনস্টলেশন:

```
pip install tokenizers          ← সাধারণ
!pip install tokenizers         ← Google Colab
pip install git+[GitHub পাথ]    ← GitHub থেকে
```

৬. হাগিং ফেস অ্যাক্সেস টোকেন (API কী)

অ্যাক্সেস টোকেন কী?

> হাগিং ফেস সার্ভিস ও রিসোর্স অ্যাক্সেস করার জন্য ব্যবহৃত একটি সুরক্ষিত অক্ষর-স্ট্রিং। API কী এবং অ্যাক্সেস টোকেন একই জিনিস।

কখন টোকেন প্রয়োজন? ✅

- প্রাইভেট/গেটেড মডেল ব্যবহারে (যেমন: Meta LLaMA)
- Inference API ব্যবহারে
- হাগিং ফেস হাবে মডেল/ডেটাসেট/Spaces আপলোড করতে

কখন টোকেন প্রয়োজন নেই? ❌

- পাবলিক মডেল ব্যবহারে (যেমন: GPT-2)
- ট্রান্সফর্মার্স লাইব্রেরির মাধ্যমে পাবলিক মডেল ব্যবহারে
- ফ্রি ওপেন-সোর্স মডেল ব্যবহারে

টোকেন তৈরির ধাপ:

1. huggingface.co/join → অ্যাকাউন্ট তৈরি
2. ইমেইল যাচাই করুন
3. প্রোফাইল → Access Tokens → Create New Token
4. টোকেনের নাম দিন → Create Token
5. কপি করে নিরাপদে সংরক্ষণ করুন

> ⚠️ সতর্কতা: মডেল বন্ধ করলে আর দেখা যাবে না। কারো সাথে শেয়ার করবেন না।

৭. হাগিং ফেস থেকে ডেটাসেট ডাউনলোড

ডেটাসেট কী?

> মেশিন লার্নিং মডেল ট্রেনিং, মূল্যায়ন বা পরীক্ষার জন্য কাঠামোবদ্ধ ডেটার সংগ্রহ

ডাউনলোড পদ্ধতি:

```
from datasets import load_dataset
dataset = load_dataset("imdb")
```

IMDB ডেটাসেটের গঠন:

বিভাজন	সারি সংখ্যা	বিষয়বস্তু
Train	২৫K	মুভি রিভিউ + লেবেল (পজিটিভ/নেগেটিভ)
Test	২৫K	মুভি রিভিউ + লেবেল
Unsupervised	৫০K	শুধু টেক্সট, কোনো লেবেল নেই

📌 Unsupervised স্প্লিট প্রি-ট্রেনিং বা সেমি-সুপারভাইজড লার্নিংয়ে ব্যবহৃত হয়।

৮. হাগিং ফেস থেকে মডেল ডাউনলোড

ডাউনলোড পদ্ধতি:

```
from transformers import BertModel, BertTokenizer
model = BertModel.from_pretrained("bert-base-uncased")
```

from_pretrained মেথড কী করে?

- মডেল ওয়েটস ডাউনলোড
- কনফিগারেশন ডাউনলোড
- টোকেনাইজার ডাউনলোড

BERT আউটপুট শেপ ব্যাখ্যা: "hello hugging face" ইনপুটের জন্য আউটপুট → **(1, 7, 768)**

মান	অর্থ
1	ব্যাচ সাইজ (১টি বাক্য)
7	সিকোয়েন্স লেংথ (বিশেষ টোকেনসহ)
768	হিডেন সাইজ (BERT বেসের আদর্শ মান)

৯. সেন্টিমেন্ট অ্যানালাইসিস

সেন্টিমেন্ট অ্যানালাইসিস কী?

টেক্সটে প্রকাশিত সেন্টিমেন্ট নির্ধারণ করা (পজিটিভ/নেগেটিভ/নিউট্রাল)

সেন্টিমেন্ট অ্যানালাইসিসের ধরন:

ধরন	উদাহরণ
পোলারিটি ডিটেকশন	"পণ্যটি ভালো" → পজিটিভ
ইমোশন ডিটেকশন	"এত হতাশাজনক" → রাগ
অ্যাসপেক্ট-বেসড	"খাবার ভালো, সার্ভিস খারাপ" → মিশ্র
ইন্টেন্ট অ্যানালাইসিস	"কোথায় কিনব?" → কেনার ইচ্ছা

কোড প্রক্রিয়া:

```
from transformers import pipeline
analyzer = pipeline("sentiment-analysis")
result = analyzer(["আমি ক্রিকেট পছন্দ করি"])
```

আউটপুট বোঝার নিয়ম:

স্কোর	অর্থ
১-এর কাছাকাছি	মডেল অত্যন্ত আত্মবিশ্বাসী
০-এর কাছাকাছি	মডেল কম আত্মবিশ্বাসী

📌 **Google Colab** টিপস: জটিল প্রজেক্টে T4 GPU বা V2-8 TPU ব্যবহার করুন।

১০. টেক্সট ক্লাসিফিকেশন

কী? → টেক্সটকে নির্দিষ্ট ক্যাটাগরিতে শ্রেণিবদ্ধ করা

সেন্টিমেন্ট অ্যানালাইসিস বনাম টেক্সট ক্লাসিফিকেশন:

বিষয়	সেন্টিমেন্ট অ্যানালাইসিস	টেক্সট ক্লাসিফিকেশন
পরিধি	সংকীর্ণ (আবেগ-নির্দিষ্ট)	বিস্তৃত (কাজ-নির্ভর)
লেবেল	পজিটিভ/নেগেটিভ/নিউট্রাল	স্প্যাম/নট স্প্যাম, টপিক ক্যাটাগরি

স্প্যাম শনাক্তকরণের লেবেল ম্যাপিং:

- Negative → স্প্যাম
- Neutral → স্প্যাম নয়
- Positive → স্প্যাম নয়

⚠️ **গুরুত্বপূর্ণ:** কম কনফিডেন্স স্কোর মানে মডেল অনিশ্চিত। সেন্টিমেন্ট মডেল স্প্যামের জন্য সর্বোত্তম নয় — নির্দিষ্ট কাজের জন্য নির্দিষ্ট মডেল বেছে নিন।

১১. টেক্সট সামারাইজেশন

বাস্তব ব্যবহার:

- দীর্ঘ আর্টিকেল সংক্ষিপ্তকরণ
- গবেষণাপত্র সারসংক্ষেপ
- চ্যাটবটে দ্রুত প্রতিক্রিয়া
- বড় ডেটাসেট থেকে মূল পয়েন্ট বের করা

মূল প্যারামিটারসমূহ:

প্যারামিটার	কাজ
max_length	সর্বোচ্চ টোকেন সংখ্যা
min_length	সর্বনিম্ন টোকেন সংখ্যা
length_penalty	দীর্ঘ (২) বা সংক্ষিপ্ত সারসংক্ষেপ
num_beams	বিম সার্চের প্রস্থ (৪ = উন্নত মান, ধীর গতি)

প্রক্রিয়া:

1. AutoModel ও AutoTokenizer লোড করুন
2. ইনপুট টোকেনাইজ করুন
3. generate() মেথড দিয়ে সারসংক্ষেপ তৈরি করুন
4. আউটপুট টোকেন ডিকোড করে প্রিন্ট করুন

১২. টেক্সট-টু-টেক্সট (অনুবাদ)

টেক্সট জেনারেশন বনাম টেক্সট-টু-টেক্সট:

বিষয়	টেক্সট জেনারেশন	টেক্সট-টু-টেক্সট
পদ্ধতি	অটোরিগ্রেসিভ (একটি একটি টোকেন)	সিকোয়েন্স-টু-সিকোয়েন্স
ব্যবহার	ডায়ালগ, টেক্সট কমপ্লিশন	অনুবাদ, সামারাইজেশন, প্যারাফ্রেজিং

T5-small মডেল:

- T5 (Text-to-Text Transfer Transformer)-এর ছোট সংস্করণ
- সামারাইজেশন, অনুবাদ, QA-তে ব্যবহারযোগ্য
- টাস্ক প্রিফিক্স প্রয়োজন → যেমন: "translate English to Spanish:"

প্রক্রিয়া:

1. T5 মডেল ও টোকেনাইজার লোড
2. "translate English to Spanish: [টেক্সট]" ফরম্যাটে ইনপুট তৈরি
3. generate() দিয়ে আউটপুট তৈরি
4. ডিকোড করে ফলাফল প্রিন্ট

১৩. কোয়েশ্চেন আন্সারিং

প্রয়োজনীয় উপাদান:

- কনটেক্সট → যে প্যারাগ্রাফ থেকে উত্তর খোঁজা হবে
- প্রশ্ন → যা জানতে চাওয়া হচ্ছে

প্রক্রিয়া:

1. প্রি-ট্রেইনড QA মডেল ও টোকেনাইজার লোড
2. কনটেক্সট ও প্রশ্ন একসাথে টোকেনাইজ
3. মডেলে পাস করুন
4. মডেল স্টার্ট ও এন্ড স্কোর বের করে উত্তরের অবস্থান নির্ধারণ করে
5. টোকেন আইডি → শব্দে রূপান্তর করে উত্তর প্রদর্শন

১৪. টেক্সট-টু-ইমেজ

ডিফিউজার্স লাইব্রেরি:

- হাগিং ফেস কর্তৃক তৈরি ওপেন-সোর্স লাইব্রেরি
- ইমেজ, অডিও ও অন্যান্য ডেটা জেনারেশনের জন্য
- ডিফিউশন মডেল ব্যবহার করে

স্টেবল ডিফিউশন:

- একটি লেটেন্ট ডিফিউশন মডেল
- টেক্সট প্রম্পট থেকে উচ্চমানের ইমেজ তৈরি করে
- সবচেয়ে জনপ্রিয় জেনারেটিভ মডেলগুলোর একটি

প্রক্রিয়া:

1. Stable Diffusion Pipeline লোড করুন
2. টেক্সট প্রম্পট পাস করুন
3. ইমেজ PNG ফাইল হিসেবে সংরক্ষিত হয়

উদাহরণ প্রম্পট:

> "Flying cars soar over a futuristic cityscape at sunset"

১৫. টেক্সট-টু-ভিডিও সিনথেসিস

কী?

টেক্সচুয়াল বিবরণের উপর ভিত্তি করে ভিডিও ফ্রেমের সিকোয়েন্স তৈরি করা

জটিলতার কারণ:

- NLP মডেল + জেনারেটিভ/ডিফিউশন মডেল একত্রে প্রয়োজন

জনপ্রিয় ফ্রেমওয়ার্ক:

ফ্রেমওয়ার্ক	বিশেষত্ব
Runway ML	ভিডিও জেনারেশন ও এডিটিং
Pictory	AI-জেনারেটেড ভিডিও
DeepMind Perceiver IO	মাল্টিমোডাল ইনপুট (টেক্সট+ইমেজ+ভিডিও)
PyTorch / TensorFlow	ভিডিও ফ্রেম পাইপলাইন তৈরি

প্রক্রিয়া:

1. Stable Diffusion মডেল লোড (ডিফিউজার্স লাইব্রেরি)
2. টেক্সট প্রম্পট সেট করুন
3. For-in লুপে ১০টি ফ্রেম তৈরি করুন
4. OpenCV দিয়ে ফ্রেম একত্রিত করুন
5. NumPy অ্যারে ব্যবহার করুন

আউটপুট ফাইলসমূহ:

frame_0.png → frame_9.png (১০টি ফ্রেম)
output_video.mp4 (চূড়ান্ত ভিডিও)

Video content in bangla

হাগিং ফেস - পরিচিতি এবং বৈশিষ্ট্য:

হাগিং ফেস একটি বিখ্যাত কোম্পানি এবং ওপেন-সোর্স কমিউনিটি যা NLP অর্থাৎ ন্যাচারাল ল্যাঙ্গুয়েজ প্রসেসিং এবং আর্টিফিশিয়াল ইন্টেলিজেন্সের উপর কাজ করে। হাগিং ফেস তার ট্রান্সফর্মার্স লাইব্রেরির জন্য সবচেয়ে বেশি পরিচিত, যা টুলস এবং প্রি-ট্রেইনড মডেল সরবরাহ করে যাতে সেন্টিমেন্ট অ্যানালাইসিস, মেশিন ট্রান্সলেশন এবং টেক্সট সামারাইজেশনের মতো বিস্তৃত NLP কাজগুলো করা যায়। সবচেয়ে বেশি ব্যবহৃত হাগিং ফেস লাইব্রেরিগুলো হলো ট্রান্সফর্মার্স, ডেটাসেটস এবং টোকেনাইজার্স।

এর মধ্যে অনেক লাইব্রেরি রয়েছে। একটি মূল লাইব্রেরি হলো ট্রান্সফর্মার্স, যেখানে BERT-এর মতো প্রি-ট্রেইনড মডেল রয়েছে। হাগিং ফেসে একটি মডেল হাব রয়েছে, যা একটি প্ল্যাটফর্ম যেখানে ব্যবহারকারীরা প্রি-ট্রেইনড মডেল শেয়ার এবং ডাউনলোড করতে পারেন। এর পাশাপাশি ব্যবহারকারীরা ডেটাসেট এবং অন্যান্য রিসোর্সও ডাউনলোড করতে পারেন। হাগিং ফেসে বিভিন্ন ধরনের ডেটাসেটের জন্য একটি লাইব্রেরি রয়েছে যার নাম ডেটাসেটস লাইব্রেরি, যা NLP কাজে ব্যবহৃত হয়। হাগিং ফেসে মেশিন লার্নিং ডেমো এবং অ্যাপ্লিকেশন হোস্ট ও শেয়ার করার একটি প্ল্যাটফর্ম রয়েছে যার নাম Spaces। এর মাধ্যমে প্রোডাকশন পরিবেশে মডেল সহজেই ডিপ্লয় এবং ব্যবহার করা যায়। হাগিং ফেসে একটি শক্তিশালী কমিউনিটি রয়েছে, যেখানে ডেভেলপার এবং AI প্রেমীরা এই ইকোসিস্টেমে অবদান রাখেন।

হাগিং ফেসের কিছু জনপ্রিয় মডেলের মধ্যে রয়েছে বহুল ব্যবহৃত BERT — এটি একটি বাক্যে শব্দের প্রসঙ্গ বোঝার জন্য ব্যবহৃত হয়। এর পূর্ণরূপ হলো Bidirectional Encoder Representations from Transformers। এটি Google দ্বারা NLP-এর জন্য তৈরি একটি শক্তিশালী ওপেন-সোর্স মেশিন লার্নিং ফ্রেমওয়ার্ক। এটি দ্বিমুখী পদ্ধতিতে শব্দ ও বাক্যের মধ্যে সম্পর্ক বিশ্লেষণ করে প্রসঙ্গ বোঝার ক্ষেত্রে উৎকৃষ্ট, যা কম্পিউটারকে অস্পষ্ট ভাষার অর্থ আরও ভালোভাবে বুঝতে সাহায্য করে। এছাড়াও রয়েছে GPT অর্থাৎ Generative Pre-trained Transformer এবং টেক্সট-টু-টেক্সট ট্রান্সফর্মার। সেইসাথে RoBERTa অর্থাৎ Robustly Optimized BERT Approach। RoBERTa একটি ট্রান্সফর্মার-ভিত্তিক ল্যাঙ্গুয়েজ মডেল যা ইনপুট সিকোয়েন্স বিশ্লেষণ করতে সেলফ-অ্যাটেনশন ব্যবহার করে। RoBERTa ডায়নামিক মাস্কিং প্রয়োগ করে যেখানে মাস্কিং প্যাটার্ন পরিবর্তন করা হয়। এটি বিভিন্ন NLP কাজে উন্নত পারফরম্যান্স প্রদান করে। RoBERTa মডেলের মূল লক্ষ্য হলো BERT মডেলের সীমাবদ্ধতাগুলো সমাধান করে তার পারফরম্যান্স উন্নত করা।

হাগিং ফেস - ব্যবহারের ক্ষেত্র:

হাগিং ফেস NLP, কম্পিউটার ভিশন এবং মাল্টিমোডাল অ্যাপ্লিকেশনে বিস্তৃত ব্যবহারের ক্ষেত্র সমর্থন করে।

শুরুতে রয়েছে কনভার্সেশনাল AI অর্থাৎ চ্যাটবট। GPT, Blenderbot এবং অন্যান্য মডেল ব্যবহার করে বুদ্ধিমান চ্যাটবট তৈরি করা যায়। এই চ্যাটবটগুলো ভার্চুয়াল অ্যাসিস্ট্যান্টের মতো কাস্টমার সাপোর্টে ব্যবহার করা যায়। এর পাশাপাশি ইন্টারেক্টিভ ডায়ালগ সিস্টেম তৈরি করা যায়। এগুলো শিক্ষামূলক অ্যাসিস্ট্যান্ট এবং থেরাপি বট হিসেবেও ব্যবহার করা যায়।

এরপর আসে সেন্টিমেন্ট অ্যানালাইসিস। গ্রাহকের ফিডব্যাক, সোশ্যাল মিডিয়া পোস্ট বা সার্ভিসের প্রতিক্রিয়া সহজেই বিশ্লেষণ করা যায় যাতে সেন্টিমেন্ট নির্ধারণ করা যায় — অর্থাৎ এটি পজিটিভ, নেগেটিভ নাকি নিউট্রাল। এর পাশাপাশি GPT-এর মতো মডেল ব্যবহার করে সহজেই টেক্সট, আর্টিকেল, ব্লগ এমনকি কবিতা জেনারেট করা যায়। যেকোনো প্রোগ্রামিং ভাষায় কোড স্নিপেটও সহজেই জেনারেট করা যায়। এর সাথে পণ্যের বিবরণ, রিভিউ, মার্কেটিং প্ল্যানসহ বিভিন্ন কন্টেন্ট তৈরি করা যায়।

এরপর আসে টেক্সট সামারাইজেশন। দীর্ঘ ডকুমেন্ট সহজেই গুরুত্বপূর্ণ পয়েন্টে সংক্ষিপ্ত করা যায়। মিটিং নোটও লেখা যায়।

এরপর আসে নেমড এন্টিটি রিকগনিশন। রেজুমে থেকে নাম, দক্ষতা এবং অভিজ্ঞতা সহজেই বের করা যায়। স্বাস্থ্যসেবায় রোগ নির্ণয়, রোগীর নাম এবং চিকিৎসা পরিভাষা চিহ্নিত করতেও এটি উপকারী। এছাড়াও আর্থিক প্রতিবেদন থেকে কোম্পানির নাম এবং তাদের কার্যক্রম বের করা যায়। তাই এটি ফিনান্স ডোমেইনেও ব্যবহৃত হয়।

মেশিন ট্রান্সলেশন — ওয়েবসাইট, অ্যাপ এবং ডকুমেন্ট এক ভাষা থেকে অন্য ভাষায় অনুবাদ করা যায়, যেমন ইংরেজি থেকে স্প্যানিশে। এটি কম সম্পদযুক্ত ভাষাগুলোতেও ব্যবহার করা যায়।

কোয়েশ্চেন আন্সারিং ব্যবহারের ক্ষেত্র কাস্টমার সাপোর্টের জন্য সবচেয়ে বেশি উপকারী। এর পাশাপাশি নির্দিষ্ট পাঠ্যপুস্তক বা নোটের উপর ভিত্তি করে শিক্ষার্থীদের প্রশ্নের উত্তর সহজেই দেওয়া যায়। FAQ সহজেই উত্তর দেওয়া যায়। বড় ডকুমেন্ট বা ডেটাবেজ থেকে উত্তর সহজেই খুঁজে পাওয়া যায়।

স্পিচ রিকগনিশন এবং সিনথেসিসেও এটি ব্যবহার করা যায়। স্পিচ থেকে টেক্সটে রূপান্তর করা যায়। স্পিচ-টু-টেক্সট এবং টেক্সট-টু-স্পিচ মডেল ব্যবহার করে ভয়েস-কন্ট্রোল্ড অ্যাপ্লিকেশন তৈরি করা যায়। রিয়েল-টাইম ক্যাপশনিং, ছবির বিবরণও তৈরি করা যায়। ভিজ্যুয়াল কোয়েশ্চেন আন্সারিংও করা যায়।

এরপর আসে রেকমেন্ডেশন সিস্টেম। ব্যবহারকারীরা কী পছন্দ করছেন তার উপর ভিত্তি করে সিনেমা বা ওয়েব সিরিজ সহজেই রেকমেন্ড করা যায়। ব্যবহারকারীর উদ্দেশ্য এবং প্রসঙ্গ বুঝে সার্চ রেজাল্ট উন্নত করা যায়। জালিয়াতি সহজেই শনাক্ত করা যায়। স্প্যাম ইমেইল সহজেই শনাক্ত ও ফিল্টার করা যায়। মডেল ব্যবহার করে মানসিক স্বাস্থ্যও পর্যবেক্ষণ করা যায়। টেক্সট বা স্পিচ বিশ্লেষণ করে স্ট্রেস, উদ্বেগ বা বিষণ্ণতার মতো আবেগ শনাক্ত করা যায়। সাপোর্ট কল বা চ্যাটে গ্রাহকের আবেগও বোঝা যায়। টেক্সট-টু-স্পিচ, স্পিচ-টু-টেক্সট এবং রিয়েল-টাইম ট্রান্সলেশনও করা যায়।

মাল্টিমোডাল অ্যাপ্লিকেশন — শুধু টেক্সট নয়, ভিডিও এবং অডিও কন্টেন্টও সহজেই বিশ্লেষণ করা যায়। অগমেন্টেড রিয়েলিটি অ্যাপ্লিকেশনও তৈরি করা যায়। মেশিন লার্নিং মডেল ট্রেনিংয়ের জন্য সিনথেটিক টেক্সট ডেটা তৈরি করা যায়। টেক্সট প্যারাফ্রেজ করা যায়। এন্টিটির মধ্যে সম্পর্ক চিহ্নিত করা যায়। শিক্ষার্থীদের রচনা বা অ্যাসাইনমেন্ট গ্রেড করা যায়। ব্যাকরণ সংশোধন, শব্দভান্ডার পুনর্বিন্যাস এবং কন্টেন্টের জন্য টুল তৈরি করা যায়। স্বাস্থ্যসেবায় রোগীর মেডিকেল রেকর্ড থেকে তথ্য বের করা যায়, আইনি ডকুমেন্ট থেকে মূল ধারা, বাধ্যবাধকতা বা ঝুঁকি বের করা যায়। গেমের জন্য AI-চালিত ন্যারেটিভ তৈরি করা যায়। সোশ্যাল মিডিয়া বিশ্লেষণ করে ট্রেন্ডিং টপিক বা হ্যাশট্যাগ চিহ্নিত করা যায়। ইনফ্লুয়েন্সারদের পোস্টের প্রভাব বিশ্লেষণ করা যায়। একাধিক ভাষায় সার্চ সক্ষম করতে মাল্টিলিঙ্গুয়াল অ্যাপ্লিকেশন তৈরি করা যায়। হেট স্পিচ এবং ক্ষতিকর কন্টেন্ট সহজেই শনাক্ত ও মডারেট করা যায়। নিউজ আর্টিকেল, সোশ্যাল মিডিয়া সেন্টিমেন্ট, টুইটার পোস্ট বিশ্লেষণ করে শেয়ার বাজারের প্রবণতা পূর্বাভাস দেওয়া যায়। পণ্য লঞ্চের মতো ইভেন্ট পূর্বাভাস দেওয়া যায়। মার্কেটিং ক্যাম্পেইনের জন্য ইমেইল কন্টেন্ট ব্যক্তিগতকৃত করা যায়। ব্যবহারকারীর পছন্দ এবং আচরণের উপর ভিত্তি করে ওয়েবসাইট বা অ্যাপ কন্টেন্ট কাস্টমাইজ করা যায়। নির্দিষ্ট কাজের জন্য প্রি-ট্রেইনড মডেল ফাইন-টিউন করা যায়। কাস্টম ডেটাসেটে বিভিন্ন মডেলের পারফরম্যান্স তুলনাও করা যায়।

হাগিং ফেসের ট্রান্সফর্মার্স লাইব্রেরি:

ট্রান্সফর্মার্স লাইব্রেরি হলো প্রি-ট্রেইনড মডেল এবং পাইপলাইনের মূল লাইব্রেরি। এটি একটি ওপেন-সোর্স পাইথন লাইব্রেরি। হাগিং ফেস এই ট্রান্সফর্মার্স লাইব্রেরি তৈরি করেছে এবং এটি মডুলার এবং এক্সটেনসিবল। এতে ট্রান্সলেশন, টেক্সট সামারাইজেশন, টেক্সট ক্লাসিফিকেশনসহ বিস্তৃত NLP কাজের জন্য হাজার হাজার প্রি-ট্রেইনড মডেল রয়েছে।

ট্রান্সফর্মার্স লাইব্রেরি ব্যাপকভাবে ব্যবহৃত হয় কারণ এটি জটিল NLP মডেলের সাথে ব্যবহার করা বেশ সহজ। এটি অত্যাধুনিক মডেলে অ্যাক্সেস দেয়। এর পেছনে একটি বড় এবং সক্রিয় কমিউনিটি রয়েছে। এটি কাস্টমাইজেশন এবং ফাইন-টিউনিং সমর্থন করে। এর পাশাপাশি ট্রান্সফর্মার্স লাইব্রেরি অন্যান্য টুলের সাথে ইন্টিগ্রেট করা যায়।

ট্রান্সফর্মার্স লাইব্রেরির ব্যবহারের ক্ষেত্রগুলোর মধ্যে রয়েছে: ইমেইলে স্প্যাম শনাক্তকরণের ক্ষেত্রে টেক্সট ক্লাসিফিকেশনের মতো ক্যাটাগরিতে টেক্সট শ্রেণিবদ্ধ করা। নেমড এন্টিটি রিকগনিশনের মাধ্যমে টেক্সটে নাম, তারিখ এবং অবস্থান চিহ্নিত করা। ইংরেজি থেকে জার্মানের মতো বিভিন্ন ভাষার মধ্যে টেক্সট অনুবাদ করা। GPT-এর মতো মডেল ব্যবহার করে টেক্সট জেনারেট করা। প্রদত্ত প্রসঙ্গের উপর ভিত্তি করে কোয়েশ্চেন আন্সারিং করা।

ট্রান্সফর্মার্স লাইব্রেরি ইনস্টল করতে পাইপ ব্যবহার করুন — পাইপ হলো পাইথন প্যাকেজ এবং লাইব্রেরি ডাউনলোড, ইনস্টল এবং পরিচালনার একটি প্যাকেজ ম্যানেজার — কমান্ড হলো: pip install transformers। গুগল কোলাবে পিপ কমান্ডের আগে একটি বিস্ময় চিহ্ন যোগ করুন। হাগিং ফেস গিটহাব রিপোজিটরি থেকেও সরাসরি ট্রান্সফর্মার্স লাইব্রেরি ইনস্টল করা যায়।

হাগিং ফেসের ডেটাসেটস লাইব্রেরি:

ডেটাসেটস লাইব্রেরি NLP এবং অন্যান্য মেশিন লার্নিং কাজের জন্য বিস্তৃত বিভিন্ন ডেটাসেটে সহজ অ্যাক্সেস প্রদান করে। এটি হাগিং ফেস দ্বারা তৈরি একটি পাইথন লাইব্রেরি। এটি ডেভেলপার এবং গবেষকদের মডেল ট্রেনিং এবং মূল্যায়নের জন্য ডেটা নিয়ে কাজ করা সহজ করে তোলে।

ডেটাসেটস লাইব্রেরি ব্যাপকভাবে ব্যবহৃত হয় কারণ এর দক্ষতা — লেজি লোডিং এবং স্ট্রিমিং বড় ডেটাসেটের সাথে কাজ করা সহজ করে তোলে। এতে ডেটাসেট প্রসেস করার জন্য একটি ইউনিফাইড API রয়েছে। ট্রান্সফর্মার্স লাইব্রেরি এবং অন্যান্য ML ফ্রেমওয়ার্কের সাথেও ডেটাসেটস লাইব্রেরি ব্যবহার করা যায়, তাই ইন্টিগ্রেশন এবং ইন্টারঅপারেবিলিটি সম্ভব। হাগিং ফেস হাজার হাজার ডেটাসেট সরবরাহ করে। এটি কাস্টম ডেটাসেট এবং প্রি-প্রসেসিং পাইপলাইন সমর্থন করে।

huggingface.co/datasets-এ হাগিং ফেসের ডেটাসেটস পেজে ৩৫০K-এরও বেশি ডেটাসেট পাওয়া যায়। যেকোনো একটিতে ক্লিক করলে সব বিস্তারিত তথ্য পাওয়া যায়।

ডেটাসেটস লাইব্রেরির ব্যবহারের ক্ষেত্রগুলোর মধ্যে রয়েছে: টেক্সট ক্লাসিফিকেশনের অধীনে স্প্যাম শনাক্তকরণের মতো কাজের জন্য ডেটাসেট সহজেই লোড এবং প্রি-প্রসেস করা। সেন্টিমেন্ট অ্যানালাইসিস এবং কোয়েশ্চেন আন্সারিং নিয়েও কাজ করা যায়। ট্রান্সলেশনের জন্যও কিছু ডেটাসেট পাওয়া যায়। নেমড এন্টিটি রিকগনিশনের কাজও হাগিং ফেস দ্বারা ইতিমধ্যে প্রদত্ত কিছু ডেটাসেট দিয়ে করা যায়। ডেটাসেটস লাইব্রেরি ব্যবহার করে কাস্টম ডেটাসেট লোড ও প্রি-প্রসেস করা যায়।

ডেটাসেটস লাইব্রেরি ইনস্টল করতে কমান্ড ব্যবহার করুন: pip install datasets। গুগল কোলাবে পিপ কমান্ডের আগে বিস্ময় চিহ্ন যোগ করুন। হাগিং ফেস গিটহাব রিপোজিটরি থেকেও সরাসরি ডাউনলোড করা যায় — git+github.com/huggingface/datasets পিপকে গিটহাবের হাগিং ফেস ডেটাসেটস রিপোজিটরি থেকে প্যাকেজ ইনস্টল করতে বলে।

হাগিং ফেসের টোকেনাইজার্স লাইব্রেরি:

টোকেনাইজার্স লাইব্রেরি হলো টেক্সট ডেটা টোকেনাইজ করার জন্য একটি দ্রুত, দক্ষ এবং নমনীয় লাইব্রেরি, যা ন্যাচারাল ল্যাঙ্গুয়েজ প্রসেসিংয়ের একটি গুরুত্বপূর্ণ ধাপ। টোকেনাইজেশনে টেক্সটকে শব্দ, সাব-ওয়ার্ড বা অক্ষরের মতো ছোট ছোট ইউনিটে বিভক্ত করা হয়। এরপর সেগুলোকে সংখ্যাসূচক প্রতিনিধিত্বে রূপান্তর করা হয় যা ML মডেল প্রসেস করতে পারে।

টোকেনাইজার্স লাইব্রেরি ব্যাপকভাবে ব্যবহৃত হয় কারণ এটি বড় ডেটাসেটেও দ্রুত টোকেনাইজেশনের জন্য অপ্টিমাইজড। এটি কাস্টম টোকেনাইজার এবং একাধিক টোকেনাইজেশন অ্যালগরিদম সমর্থন করে। ইন্টিগ্রেশন সম্ভব — ট্রান্সফর্মার্সের মতো অন্যান্য হাগিং ফেস লাইব্রেরির সাথে ব্যবহার করা যায়। টোকেনাইজিং, ডিকোডিং এবং ভোকাবুলারি পরিচালনার জন্য এটিতে একটি সহজ API রয়েছে। এর পাশাপাশি প্রি-ট্রেইনড টোকেনাইজারগুলো সহজেই অ্যাক্সেস করা যায়।

ব্যবহারের ক্ষেত্রগুলোর মধ্যে রয়েছে: ইমেইলে স্প্যাম শনাক্তকরণের মতো টেক্সট ক্লাসিফিকেশনের জন্য টেক্সচুয়াল ডেটা সহজেই টোকেনাইজ করা। সেন্টিমেন্ট অ্যানালাইসিস করা, এন্টিটি লেবেলের সাথে টোকেন সারিবদ্ধ করা। মেশিন ট্রান্সলেশন, টেক্সট জেনারেশন এবং কোয়েশ্চেন আন্সারিংয়েও ব্যবহৃত হয়। ডোমেইন-স্পেসিফিক ডেটাসেটের জন্য টোকেনাইজার ট্রেন ও ব্যবহার করা যায়।

টোকেনাইজার্স লাইব্রেরি ইনস্টল করতে কমান্ড ব্যবহার করুন: pip install tokenizers। গুগল কোলাবে পিপ কমান্ডের আগে বিস্ময় চিহ্ন যোগ করুন। গিটহাব রিপোজিটরি থেকেও সরাসরি ইনস্টল করা যায়: pip install git+ এরপর গিটহাব পাথ।

হাগিং ফেস অ্যাক্সেস টোকেন (API কী) এবং কীভাবে তৈরি করবেন:

অ্যাক্সেস টোকেন হলো একটি সুরক্ষিত অক্ষর-স্ট্রিং যা মূলত হাগিং ফেস সার্ভিস এবং রিসোর্স অ্যাক্সেস করতে ব্যবহৃত হয়। হাগিং ফেস API কী এবং হাগিং ফেস অ্যাক্সেস টোকেন একই জিনিস।

হাগিং ফেস অ্যাক্সেস টোকেন প্রয়োজন হয় যখন: আপনি একটি প্রাইভেট বা গেটেড মডেল বা ইনফারেন্স **API** ব্যবহার করছেন — যেমন Meta-র LLaMA একটি প্রাইভেট মডেল এবং এটি অ্যাক্সেস করতে অথেনটিকেট করতে হবে। হাগিং ফেস ইনফারেন্স API ব্যবহার করলে API কল করতে একটি অ্যাক্সেস টোকেন প্রয়োজন। হাগিং ফেস হাবে মডেল, ডেটাসেট বা Spaces আপলোড করলেও একটি অ্যাক্সেস টোকেন প্রয়োজন।

হাগিং ফেস অ্যাক্সেস টোকেন প্রয়োজন নেই যখন: আপনি GPT-2-এর মতো পাবলিক মডেলগুলো অ্যাক্সেস করছেন যা ডাউনলোড এবং ব্যবহারের জন্য সর্বজনীনভাবে উপলব্ধ। হাগিং ফেসের ট্রান্সফর্মার্স লাইব্রেরির মাধ্যমে মডেল ব্যবহার করলেও অ্যাক্সেস টোকেন প্রয়োজন নেই। অনেক ওপেন-সোর্স মডেল বিনামূল্যে পাওয়া যায় এবং API কী ছাড়াই ব্যবহার করা যায়।

হাগিং ফেস অ্যাক্সেস টোকেন তৈরি করতে: huggingface.co/join-এ যান এবং আপনার ইমেইল ঠিকানা ব্যবহার করে একটি অ্যাকাউন্ট তৈরি করুন। ইমেইল যাচাই করার পরে প্রোফাইলে ক্লিক করুন, Access Tokens-এ যান এবং create new token-এ ক্লিক করুন। একটি টোকেনের নাম যোগ করুন, create token-এ ক্লিক করুন, তারপর এটি কপি করে নিরাপদ কোথাও সংরক্ষণ করুন কারণ মডেল বন্ধ করার পরে আর দেখতে পাবেন না। যেকোনো সময় টোকেনের অনুমতি সম্পাদনা বা মুছে ফেলা যায়। মনে রাখবেন, কারো সাথে আপনার অ্যাক্সেস টোকেন শেয়ার করবেন না।

হাগিং ফেস থেকে ডেটাসেট ডাউনলোড:

একটি ডেটাসেট হলো কাঠামোবদ্ধ ডেটার একটি সংগ্রহ যা মেশিন লার্নিং মডেল ট্রেনিং, মূল্যায়ন বা পরীক্ষার জন্য ব্যবহার করা যায়। হাগিং ফেসের প্ল্যাটফর্মে অনেক ডেটাসেট রয়েছে যা NLP-এর মতো বিভিন্ন ব্যবহারের ক্ষেত্রে ব্যবহার করা যায়।

ডেটাসেট ডাউনলোড করতে, ডেটাসেটস লাইব্রেরি থেকে load_dataset ফাংশন ব্যবহার করুন। এই ফাংশন হাগিং ফেস হাব থেকে ডেটাসেট ডাউনলোড করতে বা লোকাল ফাইল থেকে লোড করতে পারে।

উদাহরণস্বরূপ, IMDB ডেটাসেট লোড করা — IMDB ডেটাসেটে সেন্টিমেন্ট ক্লাসিফিকেশনের জন্য পজিটিভ বা নেগেটিভ হিসেবে লেবেলযুক্ত মুভি রিভিউ রয়েছে। এটি চালালে স্বয়ংক্রিয়ভাবে ডেটাসেট ডাউনলোড এবং প্রসেস হবে। ডেটাসেটটি ট্রেন এবং টেস্টে বিভক্ত। ট্রেনে ২৫K সারি রয়েছে যেখানে ফিচার হিসেবে আছে মুভি রিভিউর জন্য টেক্সট এবং সেন্টিমেন্টের জন্য লেবেল যেমন পজিটিভ বা নেগেটিভ। টেস্টে একই ফিচার সহ পরীক্ষার উদ্দেশ্যে ২৫K সারি রয়েছে। এতে আরো ৫০K সারি রয়েছে যেটিতে সাধারণত সেন্টিমেন্ট অ্যানালাইসিসের জন্য লেবেল থাকে না — এটি প্রায়ই প্রি-ট্রেনিং বা সেমি-সুপারভাইজড লার্নিংয়ের মতো কাজে ব্যবহৃত হয়।

হাগিং ফেস থেকে মডেল ডাউনলোড:

হাগিং ফেস থেকে মডেল ডাউনলোড করতে, ট্রান্সফর্মার্স লাইব্রেরি ব্যবহার করুন। হাগিং ফেস হাব থেকেও সরাসরি ডাউনলোড করা যায়।

মডেল ডাউনলোড করতে from_pretrained মেথড ব্যবহার করুন। এই মেথড হাগিং ফেস হাব থেকে মডেল ওয়েটস, কনফিগারেশন এবং টোকেনাইজার ডাউনলোড করে।

উদাহরণস্বরূপ, BERT-base-uncased মডেল ডাউনলোড করে "hello hugging face" ইনপুট দিলে, শেষ হিডেন স্টেট আউটপুট শেপ টেনসর ডাইমেনশন উপস্থাপন করে। শেপটি গঠিত হয়: ১ — ব্যাচ সাইজ যেহেতু একটি ইনপুট বাক্য রয়েছে। ৭ — সিকোয়েন্স লেংথ, যা বিশেষ টোকেনসহ "hello hugging face"-এর টোকেনাইজড সংস্করণের সাথে সঙ্গতিপূর্ণ। ৭৬৮ — হিডেন সাইজ, যেখানে প্রতিটি টোকেন ৭৬৮-ডাইমেনশনাল ভেক্টর হিসেবে উপস্থাপিত হয়, যা BERT-এর বেস আর্কিটেকচারের জন্য আদর্শ।

হাগিং ফেস দিয়ে সেন্টিমেন্ট অ্যানালাইসিস:

সেন্টিমেন্ট অ্যানালাইসিসে একটি টেক্সটে প্রকাশিত সেন্টিমেন্ট নির্ধারণ করা হয় — যেমন পজিটিভ, নেগেটিভ বা নিউট্রাল।

সেন্টিমেন্ট অ্যানালাইসিসের ধরনগুলো:

পোলারিটি ডিটেকশন — অর্থাৎ পজিটিভ, নেগেটিভ বা নিউট্রাল। যেমন, "আমি এই পণ্যটি পছন্দ করি" পজিটিভ, "সার্ভিসটি ভয়ানক" নেগেটিভ এবং "প্যাকেজটি সময়মতো এসেছে" নিউট্রাল।

ইমোশন ডিটেকশন — রাগের মতো আবেগ অন্তর্ভুক্ত করে, যেমন "এটি এত হতাশাজনক," এবং আনন্দ, যেমন "আমি ফলাফলে রোমাঞ্চিত।" ইমোশন ডিটেকশনে সুখ, হতাশা এবং অন্যান্য আবেগ অন্তর্ভুক্ত।

অ্যাসপেক্ট-বেসড সেন্টিমেন্ট অ্যানালাইসিস — একটি পণ্য বা সার্ভিসের নির্দিষ্ট দিকের প্রতি সেন্টিমেন্ট। যেমন, "খাবার দারুণ ছিল কিন্তু সার্ভিস ধীর ছিল" — খাবারে পজিটিভ সেন্টিমেন্ট কিন্তু সার্ভিসে নেগেটিভ সেন্টিমেন্ট।

ইন্টেন্ট অ্যানালাইসিস — কিছু কেনার বা অভিযোগ করার ইচ্ছা। যেমন, "আমি এটি কোথায় কিনতে পারি" একটি পার্চেজ ইন্টেন্ট নির্দেশ করে।

হাগিং ফেস ব্যবহার করে সেন্টিমেন্ট অ্যানালাইসিস করতে, ট্রান্সফর্মার্স লাইব্রেরি থেকে পাইপলাইন ফাংশন ব্যবহার করুন। পাইপলাইন ফাংশন সেন্টিমেন্ট অ্যানালাইসিসসহ বিভিন্ন NLP কাজ সম্পাদনের একটি সহজ উপায় প্রদান করে। একটি প্রি-ট্রেইনড সেন্টিমেন্ট অ্যানালাইসিস মডেল লোড করুন এবং একসাথে একাধিক টেক্সট বিশ্লেষণ করতে স্ট্রিংয়ের একটি লিস্ট পাস করুন।

আউটপুট হলো ডিকশনারির একটি লিস্ট। প্রতিটি ডিকশনারিতে সেন্টিমেন্ট লেবেল এবং কনফিডেন্স স্কোর রয়েছে। স্কোর ০ থেকে ১-এর মধ্যে একটি মান। ১-এর কাছাকাছি স্কোর মানে মডেল তার পূর্বাভাসে খুব আত্মবিশ্বাসী। ০-এর কাছাকাছি স্কোর মানে মডেল কম আত্মবিশ্বাসী। পজিটিভ লেবেল নির্দেশ করে যে মডেল টেক্সটের সেন্টিমেন্ট পজিটিভ বলে পূর্বাভাস দিচ্ছে। নেগেটিভ

লেবেল এর বিপরীত। ১-এর কাছাকাছি উচ্চ স্কোর হয় কারণ মডেলটি একটি বড় ডেটাসেটে ফাইন-টিউন করা হয়েছে এবং সেন্টিমেন্ট অ্যানালাইসিস কাজে অত্যন্ত নির্ভুল। ইনপুট টেক্সটে শক্তিশালী, স্পষ্ট ভাষা থাকলে মডেলের পক্ষে উচ্চ আত্মবিশ্বাসের সাথে সেন্টিমেন্ট পূর্বাভাস দেওয়া সহজ হয় — যেমন "ঘৃণা" মানে নেগেটিভ এবং "ভালোবাসা" মানে পজিটিভ।

দক্ষতার জন্য, গুগল কোলাবে রানটাইম টাইপ পরিবর্তন করা যায়। জটিল বা বড় মাপের প্রজেক্টের জন্য T4 GPU বা V2-8 TPU বেছে নেওয়া যায়।

হাগিং ফেস দিয়ে টেক্সট ক্লাসিফিকেশন:

টেক্সট ক্লাসিফিকেশন স্প্যাম শনাক্তকরণে ব্যবহার করা যায় — ইমেইলকে স্প্যাম বা স্প্যাম নয় হিসেবে শ্রেণিবদ্ধ করা। নিউজ আর্টিকেল বা ডকুমেন্টও শ্রেণিবদ্ধ করা যায় — যেমন স্পোর্টস ক্যাটাগরির অধীনে স্পোর্টস আর্টিকেল, প্রযুক্তি ক্যাটাগরির অধীনে টেক-সম্পর্কিত আর্টিকেল। ইন্টেন্ট ডিটেকশনের ক্ষেত্রেও এটি ব্যবহার করা হয়, যেমন অর্ডার বাতিল করা, ফ্লাইট বুক করা ইত্যাদি।

সেন্টিমেন্ট অ্যানালাইসিস এবং টেক্সট ক্লাসিফিকেশনের মধ্যে পার্থক্য: সেন্টিমেন্ট অ্যানালাইসিস সংকীর্ণ — অর্থাৎ পজিটিভ, নেগেটিভ বা নিউট্রালের মতো সেন্টিমেন্ট-নির্দিষ্ট। টেক্সট ক্লাসিফিকেশনের লেবেলগুলো কাজের উপর নির্ভর করে — যেমন স্প্যাম বা স্প্যাম নয়, বা বিভিন্ন টপিক ক্যাটাগরি।

হাগিং ফেস ব্যবহার করে স্প্যাম শনাক্তকরণের জন্য টেক্সট ক্লাসিফিকেশন করতে, ট্রান্সফর্মার্স এবং টর্চ লাইব্রেরি ইনস্টল করুন, পাইপলাইন মডিউল ইম্পোর্ট করুন এবং একটি প্রি-ট্রেইনড স্প্যাম শনাক্তকরণ মডেল লোড করুন। একাধিক টেক্সট সেট করুন এবং পাইপলাইনে স্ট্রিংয়ের একটি লিস্ট পাস করে একসাথে শ্রেণিবদ্ধ করুন। লেবেলগুলো স্প্যাম এবং স্প্যাম নয়-তে ম্যাপ করুন — নেগেটিভ মানে স্প্যাম, নিউট্রাল মানে স্প্যাম নয়, পজিটিভ মানে স্প্যাম নয়। লেবেল এবং কনফিডেন্স স্কোরসহ ফলাফল প্রদর্শন করতে ফর-ইন লুপ ব্যবহার করুন।

কম কনফিডেন্স স্কোর নির্দেশ করে যে মডেল তার পূর্বাভাস সম্পর্কে অনিশ্চিত। যদি একটি মডেল সেন্টিমেন্ট অ্যানালাইসিসের জন্য ফাইন-টিউন করা হয় কিন্তু বিশেষভাবে স্প্যাম শনাক্তকরণের জন্য নয়, তাহলে কনফিডেন্স স্কোর কম হতে পারে। আরও ভালো ফলাফলের জন্য হাগিং ফেস থেকে একটি ভিন্ন, আরও নির্দিষ্ট মডেল বেছে নেওয়া যায়।

হাগিং ফেস দিয়ে টেক্সট সামারাইজেশন:

সামারাইজেশন অনেক বাস্তব-জীবনের অ্যাপ্লিকেশনে ব্যবহৃত হয় — দীর্ঘ আর্টিকেলকে সংক্ষিপ্ত স্নিপেটে রূপান্তর করা, ডকুমেন্ট বা গবেষণাপত্র সংক্ষিপ্ত করা, চ্যাটবটে দ্রুত সংক্ষিপ্ত প্রতিক্রিয়া প্রদান করা এবং বড় ডেটাসেট থেকে মূল পয়েন্ট বের করা।

হাগিং ফেস ব্যবহার করে টেক্সট সামারাইজেশন করতে, ট্রান্সফর্মার্স এবং পাইটর্চ লাইব্রেরি ইনস্টল করুন। মডেল এবং টোকেনাইজার সরাসরি লোড করতে sequence-to-sequence LM-এর জন্য AutoModel এবং AutoTokenizer ব্যবহার করুন। একটি প্রি-ট্রেইনড সামারাইজেশন মডেল লোড করুন, তারপর ইনপুট টেক্সট টোকেনাইজ করুন।

সারসংক্ষেপের দৈর্ঘ্য এবং মান নিয়ন্ত্রণের মূল প্যারামিটারগুলো হলো: max_length — সারসংক্ষেপে সর্বোচ্চ টোকেন সংখ্যা। min_length — সারসংক্ষেপে সর্বনিম্ন টোকেন সংখ্যা। length_penalty — দীর্ঘ বা সংক্ষিপ্ত সারসংক্ষেপ উৎসাহিত করে, যেখানে ২ মান মানে দীর্ঘ সারসংক্ষেপ। num_beams — বিম সার্চের প্রস্থ নিয়ন্ত্রণ করে, যেখানে উচ্চতর মান মান উন্নত করে কিন্তু ইনফারেন্স ধীর করে, এবং ৪ মান মানে ডিকোডিংয়ের জন্য ৪টি বিম। সারসংক্ষেপ তৈরি করতে generate মেথড ব্যবহার করুন, তারপর আউটপুট টোকেন টেক্সটে ডিকোড করুন এবং সারসংক্ষেপ প্রিন্ট করুন।

হাগিং ফেস দিয়ে টেক্সট-টু-টেক্সট (অনুবাদ):

অনুবাদ হলো টেক্সট-টু-টেক্সট মডেলের একটি অংশ যার জন্য কাজের ধরন নির্দিষ্ট করতে একটি টাস্ক প্রিফিক্স প্রয়োজন — যেমন ট্রান্সলেশন, সামারাইজেশন ইত্যাদি। টেক্সট-টু-টেক্সট জেনারেশনে শুধু অনুবাদ নয়, সামারাইজেশন, প্যারাফ্রেজিং, কোয়েশ্চেন আন্সারিং এবং এমনকি সেন্টিমেন্ট ক্লাসিফিকেশনও অন্তর্ভুক্ত।

টেক্সট-টু-টেক্সট এবং টেক্সট জেনারেশনের মধ্যে পার্থক্য: টেক্সট জেনারেশন অটোরিগ্রেসিভ টেক্সট জেনারেশনের জন্য ব্যবহৃত হয় যেখানে মডেল একসময়ে একটি টোকেন ক্রমানুসারে টেক্সট তৈরি করে — যেমন ডায়ালগ সিস্টেম, টেক্সট কমপ্লিশন ইত্যাদি। টেক্সট-টু-টেক্সট জেনারেশন ক্লাস সিকোয়েন্স-টু-সিকোয়েন্স কাজের জন্য ব্যবহৃত হয় যেখানে মডেল একটি ইনপুট সিকোয়েন্স নেয় এবং একটি আউটপুট সিকোয়েন্স তৈরি করে — যেমন টেক্সট সামারাইজেশন, প্যারাফ্রেজিং এবং অনুবাদ।

হাগিং ফেস ব্যবহার করে অনুবাদ করতে, T5-small মডেল ব্যবহার করুন, যা হাগিং ফেসে সর্বজনীনভাবে উপলব্ধ। এই মডেলটি T5 মডেলের একটি ছোট সংস্করণ এবং সামারাইজেশন, অনুবাদ এবং এমনকি কোয়েশ্চেন আন্সারিংয়ের মতো কাজে ব্যবহার করা যায়। T5 মডেল একটি বহুমুখী টেক্সট-টু-টেক্সট মডেল যা ইনপুটে একটি টাস্ক-নির্দিষ্ট প্রম্পট প্রিফিক্স করে অনুবাদ পরিচালনা করতে পারে — যেমন "translate English to Spanish।" ইনপুট টেক্সটকে মডেল প্রসেস করতে পারে এমন ইনপুট আইডিতে টোকেনাইজ করুন। অনুবাদিত টেক্সট তৈরি করতে মডেল ব্যবহার করুন। max_length এবং num_beams-এর মতো প্যারামিটার দিয়ে জেনারেশন প্রক্রিয়া কাস্টমাইজ করা যায়। আউটপুট টোকেনগুলো টেক্সটে ডিকোড করা হয় এবং অনুবাদিত ফলাফল হিসেবে প্রিন্ট করা হয়।

হাগিং ফেস দিয়ে কোয়েশ্চেন আন্সারিং:

হাগিং ফেস ব্যবহার করে কোয়েশ্চেন আন্সারিং করতে, ট্রান্সফর্মার্স লাইব্রেরি ব্যবহার করুন। কোয়েশ্চেন আন্সারিং কাজের জন্য একটি কনটেক্সট এবং একটি প্রশ্ন প্রয়োজন। কনটেক্সট হলো একটি প্যারাগ্রাফ বা টেক্সট যেখানে উত্তর পাওয়া যেতে পারে।

একটি প্রি-ট্রেইনড QA মডেল এবং টোকেনাইজার লোড করুন। টোকেনাইজার ব্যবহার করে কনটেক্সট এবং প্রশ্ন উভয়ই টোকেনাইজ করুন। উত্তর পেতে মডেলে টোকেনাইজড ইনপুট পাস করুন। মডেল উত্তরের সবচেয়ে সম্ভাব্য শুরু এবং শেষ অবস্থান পেতে স্টার্ট এবং এন্ড স্কোর বের করবে। তারপর উত্তর প্রদর্শন করতে টোকেন আইডিগুলোকে শব্দে রূপান্তর করুন।

উদাহরণস্বরূপ, একজন ব্যক্তির সম্পর্কে একটি কনটেক্সট এবং "এই ব্যক্তি কোথায় থাকেন?" প্রশ্ন দিলে মডেল সরাসরি কনটেক্সট থেকে উত্তর বের করবে।

হাগিং ফেস দিয়ে টেক্সট-টু-ইমেজ:

ডিফিউজার্স লাইব্রেরি হলো ইমেজ, অডিও এবং অন্যান্য ধরনের ডেটা তৈরির জন্য ডিফিউশন মডেলের উপর দৃষ্টি নিবদ্ধ করা একটি ওপেন-সোর্স পাইথন লাইব্রেরি। এগুলো হাগিং ফেস দ্বারা তৈরি জেনারেটিভ মডেলের একটি শ্রেণি।

স্টেবল ডিফিউশন হলো উচ্চ-মানের ইমেজ জেনারেশনের জন্য ডিজাইন করা একটি লেটেন্ট ডিফিউশন মডেল। এই মডেল ব্যবহার করে টেক্সট প্রম্পট থেকে ইমেজ তৈরি করা যায়। এটি সবচেয়ে জনপ্রিয় জেনারেটিভ মডেলগুলোর একটি।

হাগিং ফেস ব্যবহার করে টেক্সট থেকে ইমেজ তৈরি করতে, প্রয়োজনীয় লাইব্রেরি ইনস্টল করুন এবং স্টেবল ডিফিউশন পাইপলাইন লোড করুন। ডিফিউজার্স লাইব্রেরি একটি স্টেবল ডিফিউশন পাইপলাইন সরবরাহ করে যা টেক্সট প্রম্পট থেকে ইমেজ তৈরি করা সহজ করে। পাইপলাইনে একটি টেক্সট প্রম্পট পাস করুন ইমেজ তৈরি করতে — যেমন "Flying cars soar over a futuristic cityscape at sunset।" ইমেজটি একটি PNG ফাইল হিসেবে সংরক্ষিত হয় যা পরে ডাউনলোড করা যায়।

হাগিং ফেস দিয়ে টেক্সট-টু-ভিডিও সিনথেসিস:

টেক্সট-টু-ভিডিও সিনথেসিসে একটি টেক্সচুয়াল বিবরণের উপর ভিত্তি করে ভিডিও ফ্রেমের একটি সিকোয়েন্স তৈরি করা হয়। যেহেতু এটি একটি জটিল কাজ, এতে বিভিন্ন NLP মডেলকে জেনারেটিভ মডেল বা ডিফিউশন মডেলের সাথে একত্রিত করতে হয়।

কিছু জনপ্রিয় ভিডিও জেনারেশন ফ্রেমওয়ার্ক হলো: Runway ML — এটি ভিডিও জেনারেশন এবং এডিটিংয়ের জন্য টুলস অফার করে। Pictory — বেশিরভাগ AI-জেনারেটেড ভিডিওর জন্য। DeepMind-এর Perceiver IO — টেক্সট, ইমেজ এবং ভিডিওসহ মাল্টিমোডাল ইনপুট পরিচালনা করতে পারে। ভিডিও ফ্রেম জেনারেশনের জন্য পাইপলাইন তৈরিতে পাইটর্চ বা টেনসরফ্লো-র মতো লাইব্রেরি ব্যবহার করা হয়।

হাগিং ফেস ব্যবহার করে টেক্সট থেকে ভিডিও তৈরি করতে, স্টেবল ডিফিউশনের মতো একটি প্রি-ট্রেইনড টেক্সট-টু-ইমেজ মডেল লোড করতে ডিফিউজার্স লাইব্রেরি ব্যবহার করুন। একটি টেক্সট প্রম্পট সেট করুন — যেমন "A futuristic cityscape at night with flying cars।" ফর-ইন লুপ ব্যবহার করে টেক্সট বিবরণের উপর ভিত্তি করে পৃথক ফ্রেম তৈরি করুন — যেমন ১০টি ফ্রেম তৈরি করা। ফ্রেমগুলো একটি ভিডিওতে একত্রিত করতে OpenCV লাইব্রেরি ব্যবহার করুন। প্রক্রিয়ায় NumPy লাইব্রেরি এবং NumPy অ্যারে ব্যবহার করুন।

আউটপুটে অন্তর্ভুক্ত: frame_0.png থেকে frame_9.png নামে পৃথক ফ্রেম ইমেজ ফাইল এবং সমস্ত একত্রিত ফ্রেম সমন্বিত output_video.mp4 নামে একটি চূড়ান্ত ভিডিও ফাইল।