

Hugging Face 란 무엇인가?

Hugging Face 는 인공지능(AI), 특히 자연어 처리(NLP) 분야에서 널리 사용되는 오픈소스 플랫폼이자 회사입니다.

1. Hugging Face 의 정체

- 원래는 챗봇 스타트업이었지만, 지금은 AI 모델과 데이터셋을 공유하고 협업할 수 있는 대표적인 허브 역할을 합니다.
- GitHub 처럼 AI 연구자와 개발자들이 만든 모델을 업로드하고 다른 사람들이 자유롭게 다운로드하여 활용할 수 있습니다.

2. 주요 서비스

1) 😊 Transformers 라이브러리

- BERT, GPT, T5, RoBERTa 등 최신 딥러닝 모델을 쉽게 불러와 사용할 수 있는 파이썬 라이브러리.

예시 코드:

```
from transformers import AutoTokenizer, AutoModel
```

```
tokenizer = AutoTokenizer.from_pretrained('bert-base-uncased')
model = AutoModel.from_pretrained('bert-base-uncased')
```

2) Datasets 라이브러리

- 대규모 공개 데이터셋을 손쉽게 다운로드하고 전처리할 수 있습니다. (예: IMDB, SQuAD 등)

3) Model Hub (모델 저장소)

- 전 세계 사람들이 업로드한 수십만 개의 모델을 저장하고, 바로 불러와 활용할 수 있습니다.

4) Spaces (AI 앱 배포 플랫폼)

- Gradio, Streamlit 같은 툴로 만든 AI 데모를 쉽게 공유할 수 있습니다.

Hugging Face 사이트 구조

🌐 공식 사이트: <https://huggingface.co>

1. Models (모델 허브)

- 전 세계 연구자/기업이 업로드한 수십만 개의 AI 모델을 모아둔 공간.
- NLP, Computer Vision, Audio, Multimodal 등 다양한 분야별 모델 존재.
- 예: BERT, GPT-2, Stable Diffusion, LLaMA 등.
- 원하는 모델을 검색 후 Transformers 라이브러리로 바로 불러와 사용 가능.

2. Datasets (데이터셋 허브)

- 학습 및 평가용 데이터셋을 모아둔 저장소.
- 공개 데이터셋뿐 아니라 커뮤니티 업로드 데이터셋도 포함.
- 예: IMDB, SQuAD, CommonVoice 등.

3. Spaces (AI 앱 데모)

- AI 모델을 활용한 웹 애플리케이션을 공유할 수 있는 공간.
- Gradio 또는 Streamlit 기반으로 앱 제작 가능.
- 누구나 쉽게 데모 페이지를 만들어 배포할 수 있음.

4. Docs (문서/가이드)

- Hugging Face 공식 문서.
- Transformers, Datasets, Accelerate, Diffusers 등 라이브러리 사용법 제공.
- 초보자 튜토리얼부터 고급 개발자용 가이드까지 포함.

5. Community (커뮤니티)

- 포럼, 블로그, 이벤트를 통해 연구자와 개발자가 교류하는 공간.
- 협업 프로젝트, 논문 구현, 오픈소스 기여가 활발히 이루어짐.