

Notes on Hugging Face

1. Data preprocessing

- Text inputs usually require *tokenization* (breaking text into tokens).
- Image inputs often need *resizing* and *normalization* before being passed to a model.

2. Multimodal models

- Models that take in different types of data (e.g., text + images, or audio + text) are called *multimodal* models.

3. Available tasks

- Hugging Face provides many pretrained models for specific tasks. Examples:
 - Image → Text (image captioning)
 - Text → Audio (text-to-speech)
 - Translation, summarization, classification, etc.

4. How to use a pretrained model

- Go to the **Models** page on Hugging Face.
- Select a **Task** (e.g., text classification, image segmentation, speech recognition).
- Use **Sort** to order models by criteria such as downloads, likes, or recent activity.
- Click on a model that you prefer.
- On the right-hand panel, Hugging Face shows the **Transformers code snippet** to load and use that model directly.

5. Open LLM Leaderboard

- Hugging Face hosts the **Open LLM Leaderboard**, which is used for **tracking, ranking, and evaluating** large language models and chatbots based on benchmark performance, openness, and usability.

6. **Pricing** → **Task** page

- The **Pricing** → **Task** section on Hugging Face provides explanations of different machine learning tasks and how they relate to available models and usage costs.
- When you click on each task, you can explore:
 - **Suggested models** relevant to the task
 - **Datasets** that can be used for training or evaluation
 - **Interactive demos** you can try out directly in the browser