## **演習環境についてのご案内**

LLM講座のDay2からは講義に加えて演習も実施する予定です。

演習の時間においてはクラウドのGPU環境を用いて演習を行います。

演習環境はDay1でアンケート等の出欠確認に利用しているOmnicampusを利用します。

ご自身のアカウントでログインしてご利用ください。

(提供方法や内容に関しては変更になる可能性がございます。予めご理解をお願いいたします。)

### Omnicampus GPU仮想環境の基本仕様

Omnicanpus上でクラウドの仮想環境を立ち上げ、LLMの学習に適したGPU環境を提供します。

* 講座を通して1人当たり1GPU換算で合計50時間分の仮想環境を利用することが出来ます。
  + 毎回の演習と、講義の最終課題(LLM開発コンペ(予定))をこの合計50時間分のGPUリソースで行うことを想定しています。
* GPUはNVIDIA L4 のカード枚数を1~4枚まで選ぶことが出来ます。
* 1回の起動につき2時間まで利用する事が出来ます。2時間経過後にシャットダウンします。
* 2GPU以上使った場合は2倍、4GPUの場合は4倍分の速度で消費されますので、ご注意ください。
  + ただし有効に使う分には学習時間を短縮する事が期待できます。
* 使用時間例
  + 毎回1GPUで利用した場合、講座通して合計50時間利用可能　(50/1=50)
  + 毎回4GPUで利用した場合、講座通して合計12.5時間利用可能 (50/4=12.5)
* 注意: 上記の事から仮想環境の利用が終わったら、必ずインスタンスを停止してください。  
  (不必要に利用可能時間が消費されます。)

***Tips****: 通常の演習はGoogle Colabやローカル環境等で行い、計算力が必要な演習や課題などはOmnicampusを使う等する事で有効活用できる場合があります。(その他環境は各自でご用意ください。)*

(9/11 追記 解消後削除予定)

**先週発生した環境環境のトラブルについて**

先週Day2の演習時間において、Omnicampus側で受講者からアクセスが集中した為、大量のリクエストを処理仕切れず、正常に演習環境を提供する事ができなかった。

障害後にシステムの増強を行なった為、現在は正常に使用できている様に見受けられる。

**本日の演習時間について**

万全を期して本日のDay3は講師によるデモンストレーションで実施いたします。

受講者の皆様は環境を起動せず、講義終了後に各自起動、実行いただく様お願いいたします。

ご協力のほどお願いいたします。

**演習環境の実施形態について**

もし今後、同様に演習環境が立ち上がらなくなったりエラーなどが出た場合、

LLM講座Slackの [#04\_運営へのご意見\_質問](https://llmsummer2023.slack.com/archives/C05FE159UV8)へエラーメッセージや、事象をご報告の上、講座運営からのご連絡をお待ちください。

それまでの間は、(操作ができれば)演習環境を停止し、一旦時間を置いて、後日再度起動を試してください。

**次回以降の演習の時間について**

Day4 (9/15) 開催当日までに実施の方針を確定させ、皆様へご連絡させて頂きます。

### 演習環境の準備

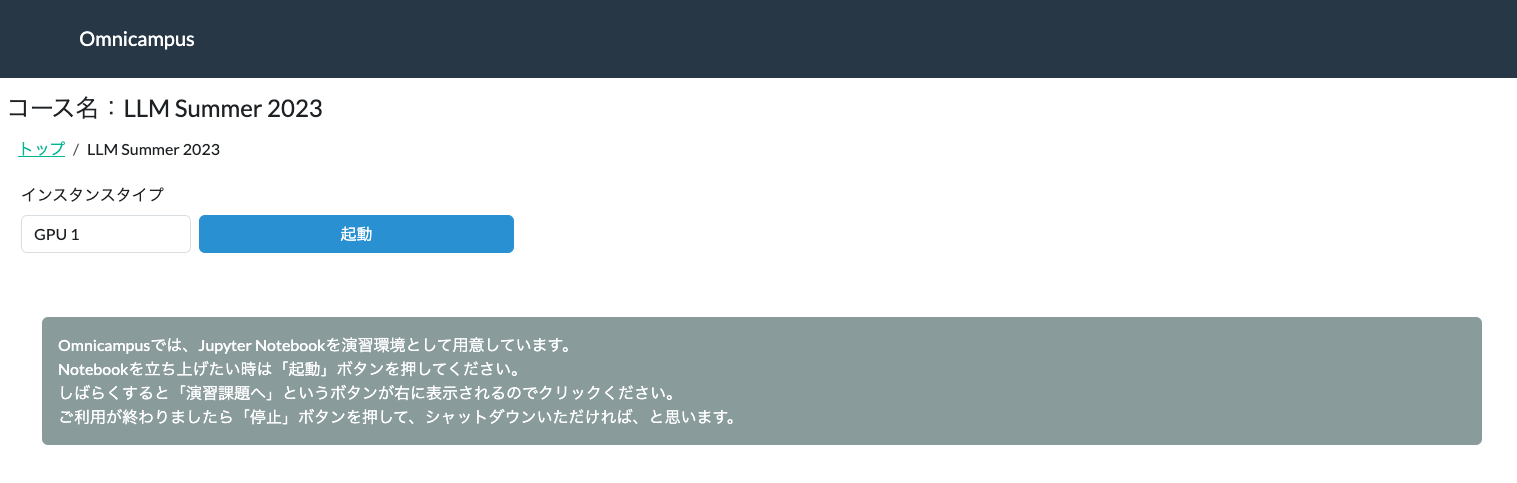
#### Omnicampusからインスタンスを立ち上げる

下記ページにアクセスします。

<https://edu.omnicamp.us/courses/34/notebook/>

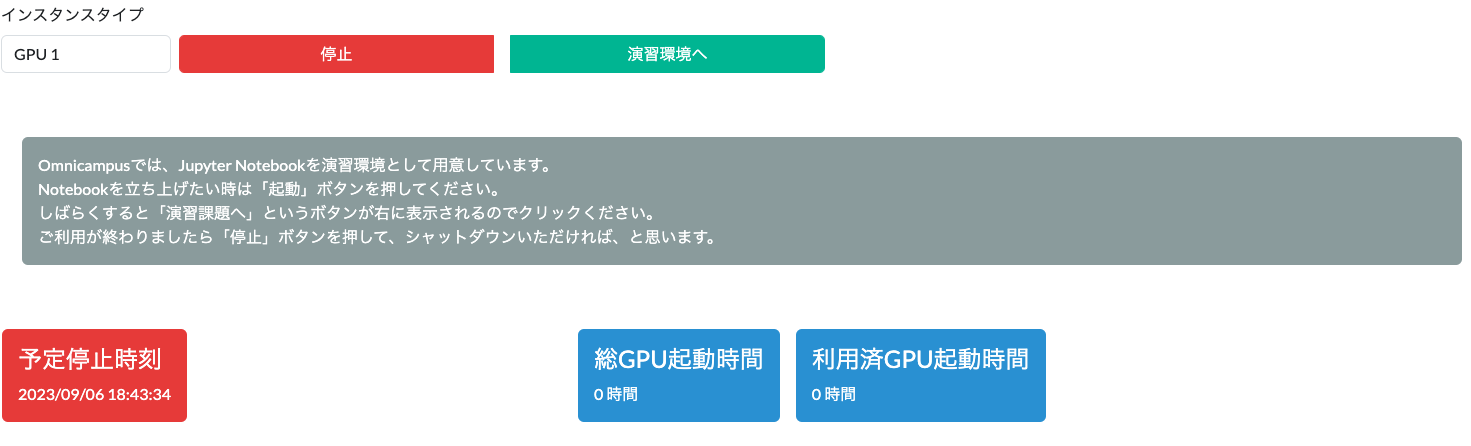
ご利用になられたい任意のインスタンスタイプ(GPU1~4)を選択し起動してください。  
起動ボタンを押下後に外部のクラウドサービス上にGPUの仮想環境を立ち上げます。  
その為、起動には平均1~2分程度時間がかかる見込みです。  
(クラウド側のアクセス状況によってはさらに時間が掛かる可能性もあります。)

起動するまでそのままお待ちください。



起動完了後、「演習環境へ」を押下しJupyter Notebookの環境を起動してください。  
(別タブで開きます。)

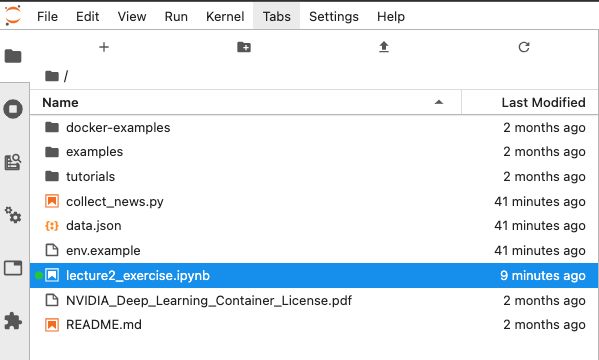
使用を終えたら「停止」を押下して環境をシャットダウンしてください。



#### 演習資料をローカルにダウンロード、その後Omnicampus上のJupyter環境へアップロードし、notebookを開いて実行してください。

演習資料

<https://drive.google.com/drive/u/1/folders/1X5T-INm1QC0pn4QEhd5MVF0I5pBX1xTm>



### 演習環境のライブラリ等 (Google Colab等外部の環境で実行する際にご利用下さい。)

注意: Google Colabだと大きなモデルサイズが動かない可能性があるので、GPU memoryに合うようにサイズを小さくする必要があるかもしれません。

```

pip install wandb datasets tqdm tiktoken transformers deepspeed openai PyYAML accelerate datasets einops evaluate peft protobuf scikit-learn scipy sentencepiece fire mpi4py

# for JGLUE

pip install bs4 zenhan mecab-python3 pyknp

# Day2

pip install langchain sentence\_transformers faiss-gpu python-dotenv

# Day4

pip install unidic-lite

# Day5

pip install bitsandbytes trl huggingface\_hub emoji neologdn ipywidgets

# Day6

pip install git+https://github.com/CarperAI/trlx

```

## 参考リンク

LLM講座 GPU演習環境 (Omnicampus)

<https://edu.omnicamp.us/>

LLM講座 演習資料格納先 (Google ドライブ)

<https://drive.google.com/drive/u/1/folders/16Mq_Q3cdmohw20P57GMq_Eqhsh6Zwahb>

LLM講座 受講の手引き(Google スライド)　<https://docs.google.com/presentation/d/1C78_vO1kVRb3I197DzjyS7OGZWWQCe1c/edit?usp=drive_link&ouid=111635753884204102602&rtpof=true&sd=true>

LLM講座 各資料URL一覧表 (Google スプレッドシート)

<https://docs.google.com/spreadsheets/d/1hKQaQDx0arTMtCWhjQX8fZBy1sHPeP9kwTzlaGipMgE/edit?usp=drive_link>