

Downloadable Pre-requisites

For the "Modern NLP: Pre-training, Fine-tuning, Prompt Engineering, and Human Feedback" workshop at ODSC East 2023

Python Environment

You will need a functioning `python` environment and your choice of a code editor. If you aren't sure where to start, you can follow this [VS Code guide](#). You will NOT need to install anaconda or any specific package/ environment manager for this workshop.

Dependencies

Install the following dependencies via `pip install` or similar.

- [LangChain](#)
- [chromadb](#)
- [predictionguard](#)

Embedding/ Vector Store

Although we won't be running large language models locally (on our laptops) during the workshop, it will be helpful to download a model that can generate embeddings for us. After you install `chromadb` as mentioned above, open up a `python` interpreter or script and run the following to make sure you have the `sentence transformers` model downloaded:

```
import chromadb

client = chromadb.Client()
collection = client.create_collection("sample_collection")

collection.add(
    documents=["This is a document", "This is another document"],
    metadatas=[{"source": "my_source"}, {"source": "other_source"}],
    ids=["doc1", "doc2"]
)
```

When you run this, you should see the `sentence transformers` model downloaded automatically and cached locally for future use.