### **Course Title: Pretraining LLMs**

**Provider**: deeplearning.ai  
**Completion Date**: 10/03/24  
**Duration**: 1 hour  
**Key Skills Learned**:

* Learned how to build high-quality training datasets and pre-train using pre-existing datasets and online text.
* Learned how to perform data cleaning, including deduplication, quality filtering, content filtering, privacy reduction, and rule-based cleaning.
* Learned how to configure and initialize model weights, and how these decisions affect pre-training speed.
* Explored common methods and strategies for evaluating the performance of trained models.

**Relevance to the Project**:

* We will use these skills to fine-tune the language model to improve the accuracy and personalization of our mental health assessment tool. This will help us provide more accurate and personalized feedback and improve the user experience.

**Certification or Proof**:

<https://learn.deeplearning.ai/accomplishments/d71b1a76-6e0f-477e-ac93-4ec30981b4e6>

### **Course Title: Learn MongoDB**

**Provider:** Codeacademy

**Completion Date:** 10/06/24

**Duration:** 3 hours

**Progress:** 60%

**Key Skills Learned**:

* Learned how to query documents in MongoDB database and use .find() method to retrieve data efficiently.
* Learned how to insert and update documents in the database and understood CRUD operations of MongoDB database.
* Understood the structure of NoSQL database and how to use BSON and JSON formats for data storage and interaction in MongoDB.
* Learned how to create indexes (createIndex()) to improve query performance.

**Relevance to the Project**:

* MongoDB is the database used in our project to store user survey data. Mastering MongoDB query and optimizing performance is crucial to the scalability and performance optimization of the project, especially when the amount of data increases.

### **Course Title: Open Source Models with Hugging Face**

**Provider:** Codeacademy

**Completion Date:** 10/06/24

**Duration:** 1.5 hours

**Progress:** 50%

**Key Skills Learned**:

* Learned how to choose a suitable pre-trained model and use the Hugging Face library for natural language processing tasks.
* Learned how to perform multi-language translation, document summary generation, sentence embedding and other tasks.

**Relevance to the Project**:

* These skills will be used in our project’s AI-driven clustering algorithms to improve the classification of user emotional events. By using Hugging Face’s open source model, we will be able to provide more personalized feedback.