# Assignment: Fine-Tuning a Language Model (LLM) for a Specific Topic

## Introduction:

This document serves as technical documentation detailing the process of automatically constructing a fine-tuning model based on user instructions.

## Usecase:

Our selected use case is to fine-tune a Large Language Model (LLM) capable of answering any questions from the [Tracified documentation](https://help.tracified.com/docs/intro).

## High-Level overview of the System

Input: "Develop a fine-tuned LLM capable of responding to queries sourced from Tracified documentation."

Output: Fine-Tuned LLM System

The process entails the following steps:

Dataset Generation: Automatically generating the dataset.

System Message Generation: The proposed system will create an effective system prompt for your model.

Fine-Tuning: Upon dataset generation, the system will automatically fine-tune a model, and prepare it for inference.

Dataset Generation

In this step, we automatically generate the dataset to fine-tune the model. We initialize the following parameters:

1. Temperature: We opt for a lower value to ensure more precise question answers from the documentation.

2. Number of examples: We leave this parameter empty as the number of examples generated depends on the given content.

3. Reference document: For our selected use case, we provide a reference document to the model for generating data samples. If you don't need to provide reference data for your use case, which means the model itself needs to generate examples, you can leave the reference document field empty. In our case, we automatically scrape the content from the corresponding web documentation and input the extracted content into the model to generate data samples. You can find the web-scraped data [here](link\_here).

Next, we define the data generation chain, which comprises the following components:

1. Prompt: We establish a clear prompt with input variables.

2. Model: We utilize the latest 'gpt-4-1106-preview' model.

3. Parser: We specify a parser to ensure that the output is structured and consistent at all times.

The output of this chain will consist of the generated data samples as outlined in the parser.

System Message Generation:

In this stage, we provide a clear and precise prompt to generate system messages for use in inference. We then utilize the output parser to structure the output more consistently and systematically.

Fine-Tuning:

Prior to fine-tuning the model, it is necessary to format the dataset in the following manner For the OpenAI API, the data must be stored in jsonl format.:

Top of Form