Happimind: Easing Academic Pressure with GPT-Powered Virtual Companions

CONTROL SYSTEMS AND INSTRUMENTATION ENGINEERING PROGRAM

Pavaris Sonpradit , Wachita Bunsit , Advisor : Dr.Teema Leangarun , Co-advisor : Dr. Kajornvut Ounjai

Abstract

Academic pressure can arise from family expectations, a desire for perfection, or social pressure, which leads students to be stressed, mentally fatigued, or non-productive. While talking to peers can serve as a stress-relief mechanism, some students may find it difficult to openly discuss their issues with others. Generative AI technology has the potential to enhance mental well-being by providing a non-judgmental listener. We have developed "Happimind" for operating generative pre-trained transformer (GPT) technology. Our research focuses on creating prompt conversations with virtual companions, mimicking the experience of chatting with real-life friends. The goal is to stimulate positive emotions in students during the learning process. Various conversation techniques, including zero-shot and few-shot, were experimentally evaluated, along with corresponding parameter adjustments. The results showed that creating your own prompt successfully generated conversations resembling those with friends. Additionally, create your own prompts that demonstrate adaptability across diverse conversation topics without the need for pre-training on specific datasets, reducing the time required for conversation model training preparation.

Introduction

Undergraduate students face high stress due to academic pressure and career demands. Existing chatbots designed for Thai communication have limitations in responding to complex and unexpected queries. This research proposed that “Happimind” involves the customization of GPT [1] models to generate and modify sentences, creating an environment that closely resembles interactions with virtual companions. This allows undergraduate students facing emotional stress to freely express themselves.

Proposed Method

This research focuses on customizing GPT to create virtual friends who can listen to and understand the situations of undergraduate students dealing with stress caused by self-imposed study pressure (Fig. 1).

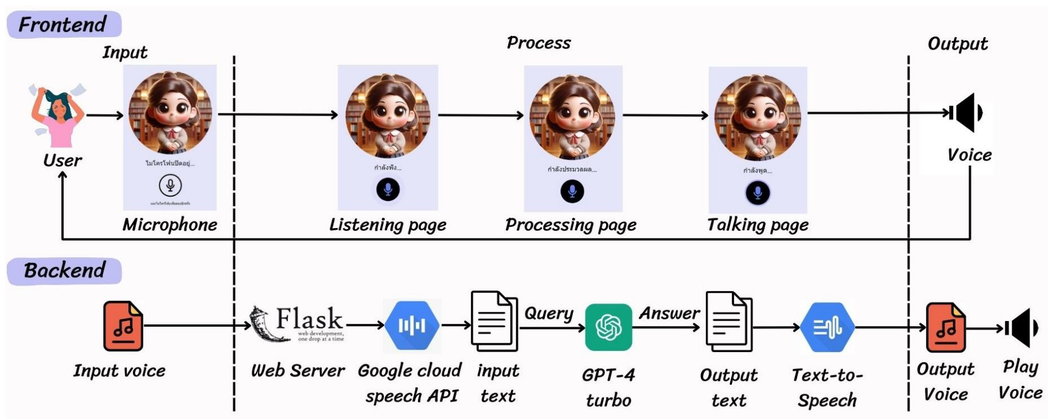

Fig. 1. System Overview

A) Adjustable parameters

Parameters are crucial settings that influence how models, which when adjusted, can significantly change the output of the GPT model[2].

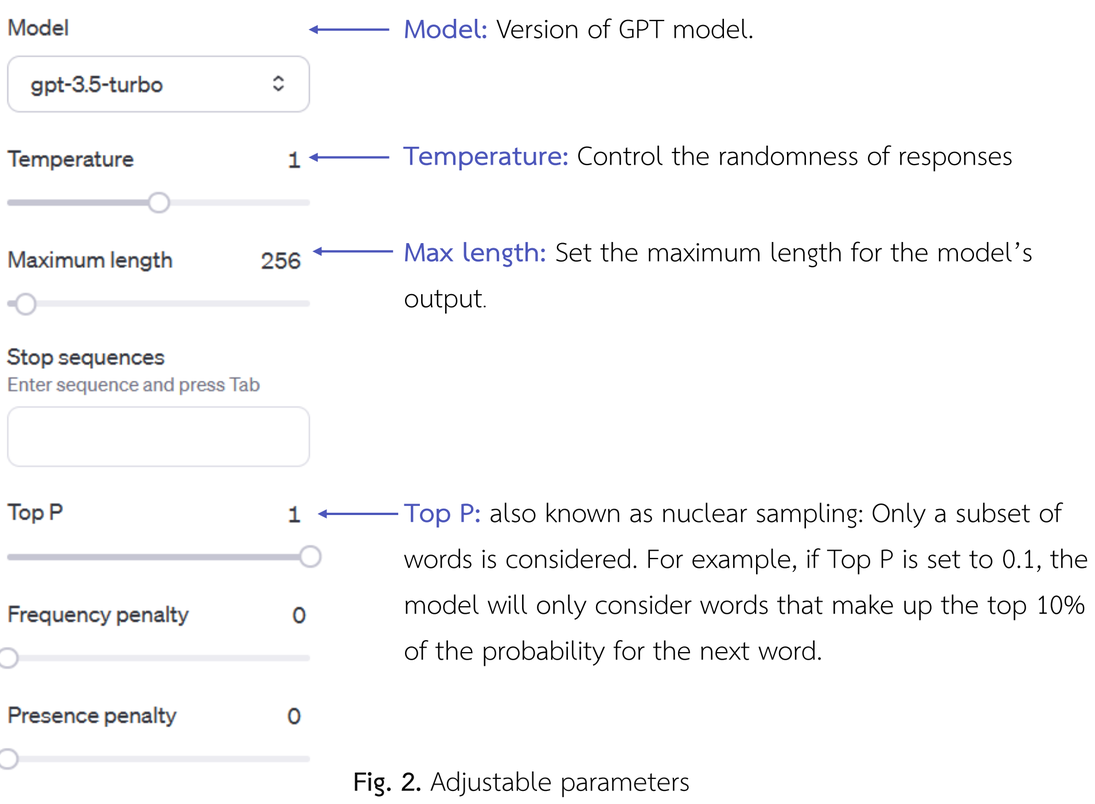

Fig. 2. Adjustable parameters

B) Two prompting techniques

Prompts provide a human language framework for the GPT model to help the model learn and act according to its intended purpose. We used two different techniques: zero-shot and few-shot.

1) Zero-shot prompting technique provides guidelines and frameworks without example sentences, specifying what the GPT model should do and how it should play its role.

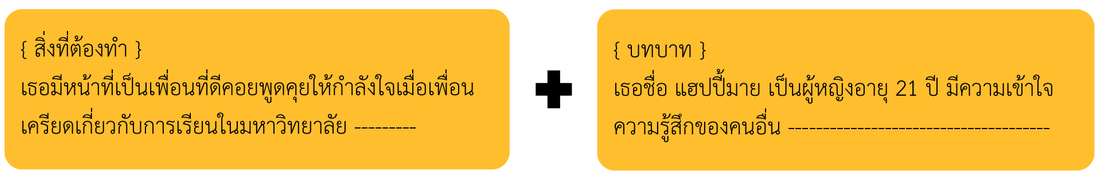

Fig. 3. Example prompt for zero-shot technique

2) Few-shot prompting technique sets guidelines and frameworks with examples in more than one sentence. In terms of what the GPT model wants to follow and the GPT model's role-playing guidelines, it is the same as the zero-shot technique, but there is a difference in the example of a conversation that involves giving encouragement. The GPT model takes groups of words about encouragement to create response sentences.

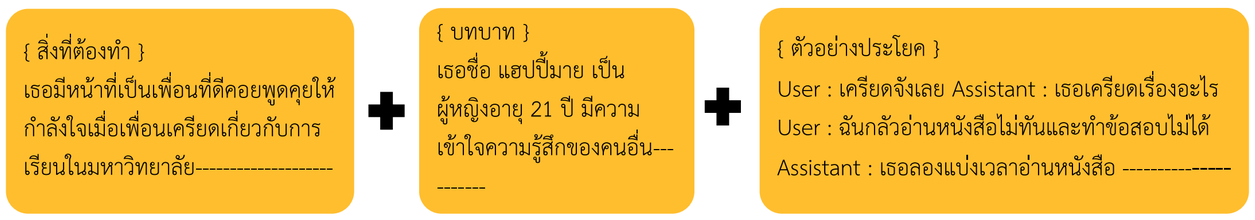

Fig. 4. Example prompt for few-shot technique

Result

We designed an experiment that explored two prompting approaches for GPT conversations: zero-shot and few-shot. Zero-shot generated discussions about student stress without a large dataset, but the conversations lacked naturalness. Few-shot learning excelled, acting like a mini-training session with additional samples. This resulted in GPT generating more natural, coherent, and concise sentences, which is particularly beneficial for friendly interactions. All techniques benefit from parameter tuning and input prompts to achieve near-natural virtual friends (see Fig. 3 for parameter tuning examples.)

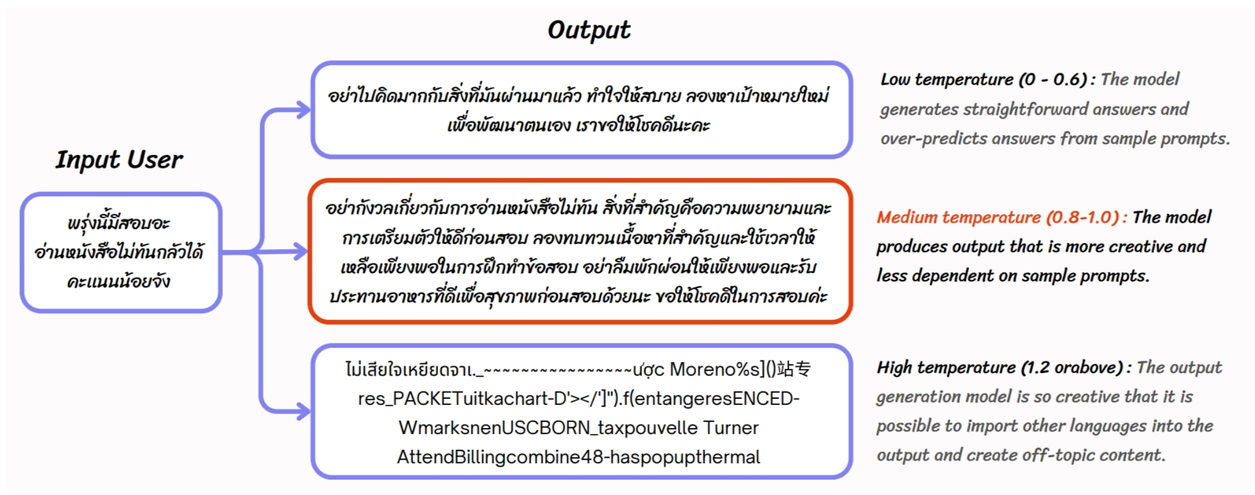

Fig. 5. An example of parameter tuning with a different temperature range

Our research examined how the number of tokens (units for processing language in GPT) affects cost and response time. We experimented with two techniques using similar input token lengths (around 31 tokens). The result is in Table 1. While the few-shot technique used more prompt input tokens than the zero-shot technique, the model's cost-effectiveness is reduced by a factor of 4 based on the rate of results per prompt. (output/prompt ratio). Even though it uses a lot number of input tokens, the few-shot technique provides the fewest output tokens.

Prompting technique	Input tokens	Prompt tokens	Output tokens	Output/Prompt tokens ratio
Zero-shot	31 ± 6	306	258 ± 32	76% ± 32
Few-shot	31 ± 6	925	223 ± 25	23% ± 24

Table. 1. Number of tokens used for the zero-shot technique and few-shot technique.

Conclusion

This research aims to create a virtual listening companion for students struggling with study stress. By leveraging GPT, a powerful AI technology can understand complex conversations and learn from user interactions. Our experiment compared two techniques, zero-shot and few-shot, along with parameter adjustments. The results show that: GPT can generate natural and interactive conversations similar to those with human peers, responding appropriately to topics without requiring massive datasets.

References

[1] Openai, “Openai,” Available: OpenAI, 23 August 2023.

[2] Codecademy Team, ”Setting Parameters in OpenAI” Available: https://www.codecademy.com/article/setting-parameters- in-open-ai, 23 August 2023.