Title options

**"Integrating GPT and BERT for Artistic Prompt Refinement: Enhancing Aesthetic Judgment in AI-Generated Artwork"**

**"Optimizing Creative Prompts in AI-Generated Art: A Hybrid Model Approach in Computer Graphics Education"**

**"Advancing Aesthetic Quality in AI Art: A Systematic Framework Using GPT and BERT for Prompt Improvement"**

**Abstract**

This study explores the enhancement of aesthetic quality in AI-generated artwork through the optimization of prompts derived from student projects in a computer graphics course. The research sample consists of 900 prompts selected from the work of 80 students across four classes, using a purposive sampling method. The study employs a hybrid approach that integrates GPT (Generative Pre-trained Transformer) and BERT (Bidirectional Encoder Representations from Transformers) models to improve the quality of user-generated prompts. The process begins with the BERT model classifying prompt quality, which then informs the GPT system to enhance the prompts across five dimensions: composition, color, style, lighting, and character depiction. The refined prompts are re-evaluated by the BERT model to ensure they meet higher aesthetic standards. This iterative approach aims to consistently elevate the sophistication and structural coherence of the prompts, ultimately improving the aesthetic quality of AI-generated artwork. The project also seeks to classify and rewrite suboptimal artistic prompts using fine-tuned models, with the final outputs made accessible via a web-based platform. The implementation was conducted in a Python-based Jupyter Notebook environment, focusing on optimizing prompt statements to enhance the aesthetic judgment

本研究探討了透過優化電腦圖形學課程中學生專案的提示來提高人工智慧生成的藝術作品的美學品質。研究樣本由 900 個提示組成，採用有目的抽樣方法，從四個班級 80 名學生的作業中選出。該研究採用了一種混合方法，整合了 GPT（生成式預訓練 Transformer）和 BERT（來自 Transformers 的雙向編碼器表示）模型，以提高使用者生成提示的品質。過程首先由 BERT 模型對提示品質進行分類，然後通知 GPT 系統在五個維度上增強提示：構圖、顏色、風格、燈光和人物描繪。精緻的提示由 BERT 模型重新評估，以確保它們符合更高的美學標準。這種迭代方法旨在不斷提高提示的複雜性和結構連貫性，最終提高人工智慧生成的藝術作品的美學品質。該專案還尋求使用微調模型對次優藝術提示進行分類和重寫，最終輸出可透過基於網路的平台存取。實作在基於Python的Jupyter Notebook環境下進行，重點優化提示語句，增強美感

本研究使用 GPT（生成式預訓練 Transformer）和 BERT（來自 Transformers 的雙向編碼器表示）模型組合來改進使用者生成提示的系統方法。本研究，首先使用 BERT 模型對提示品質分類。這種分類會通知 GPT 系統，然後 GPT 系統會透過細化構圖、色彩、風格、燈光和字元等方面來增強提示。隨後，BERT 模型會重新評估改進後的提示，以確認其符合更高的品質標準。該過程旨在迭代地提高用戶提示的質量，確保最終輸出既複雜又可靠。

Introduction

This research offering significant advancements in the fields of AI-generated art, prompt optimization, and educational technology. Firstly, the study demonstrates an innovative integration of GPT and BERT models to systematically enhance the quality of user-generated prompts, specifically within the domain of computer graphics. This hybrid approach not only improves the aesthetic quality of AI-generated artwork but also provides a scalable framework for prompt refinement across various creative disciplines. Secondly, the research contributes to the development of a fine-tuned model capable of classifying and improving artistic prompts, facilitating a more effective interaction between users and AI systems in creative contexts. Thirdly, by implementing this process in a Jupyter Notebook environment and making the improved prompts accessible through a web-based platform, the study offers a practical tool that can be widely adopted for educational purposes and beyond. Lastly, the research enriches the understanding of how AI can be utilized to enhance aesthetic judgment, contributing to the broader discourse on the intersection of technology, creativity, and education.

這項研究的貢獻是多方面的，在人工智慧生成藝術、即時優化和教育技術領域取得了重大進展。首先，該研究展示了 GPT 和 BERT 模型的創新集成，可以系統地提高用戶生成提示的質量，特別是在電腦圖形領域。這種混合方法不僅提高了人工智慧生成的藝術作品的美學質量，而且還提供了一個可擴展的框架，可以跨各種創意學科進行快速完善。其次，該研究有助於開發一種能夠對藝術提示進行分類和改進的微調模型，促進使用者和人工智慧系統在創意環境中更有效的互動。第三，透過在 Jupyter Notebook 環境中實現這一過程，並透過基於網路的平台存取改進的提示，該研究提供了一種實用工具，可以廣泛用於教育目的及其他目的。最後，這項研究豐富了人們對如何利用人工智慧來增強美學判斷力的理解，有助於就技術、創造力和教育的交叉領域進行更廣泛的討論。

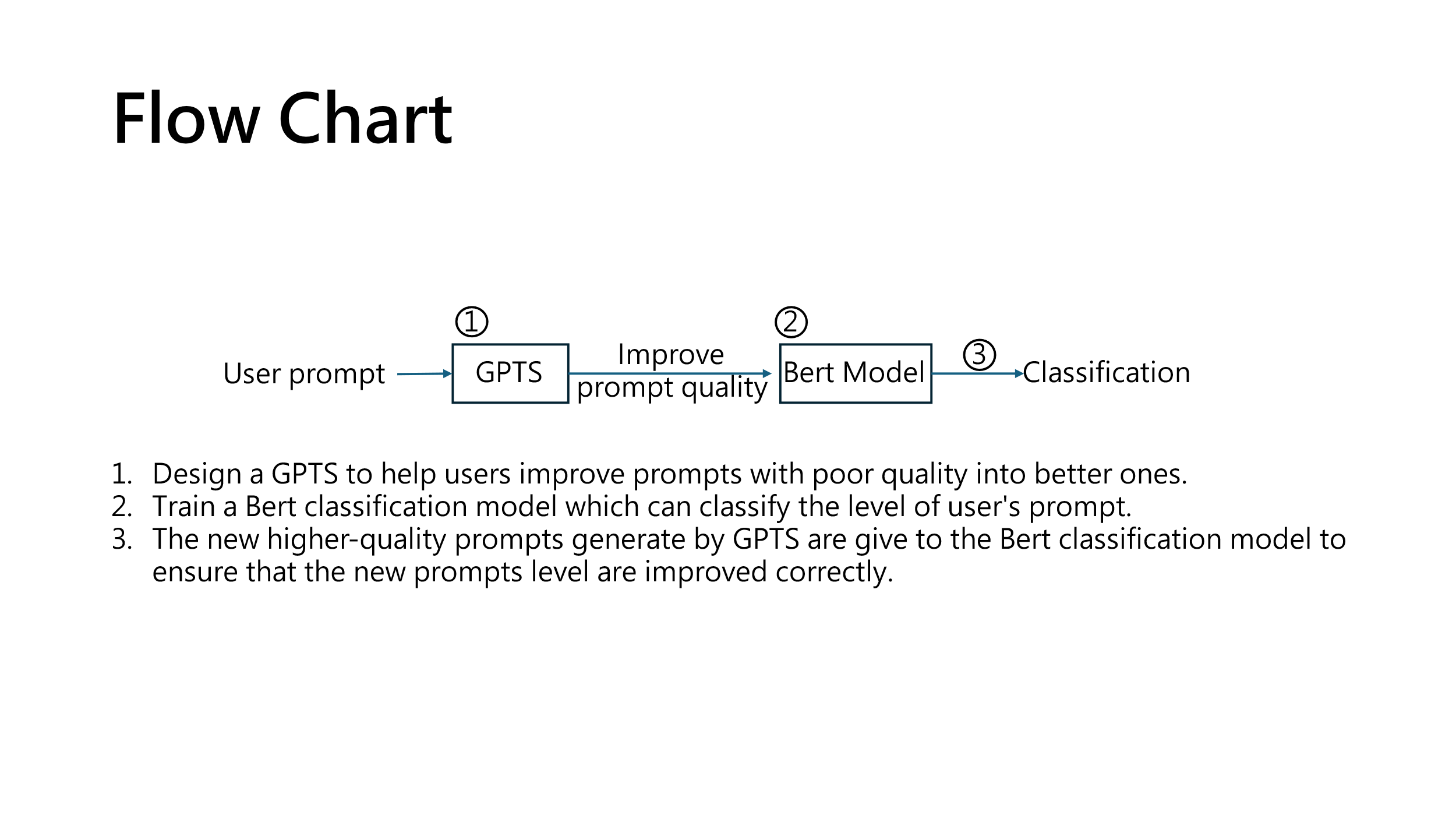
**Research question**

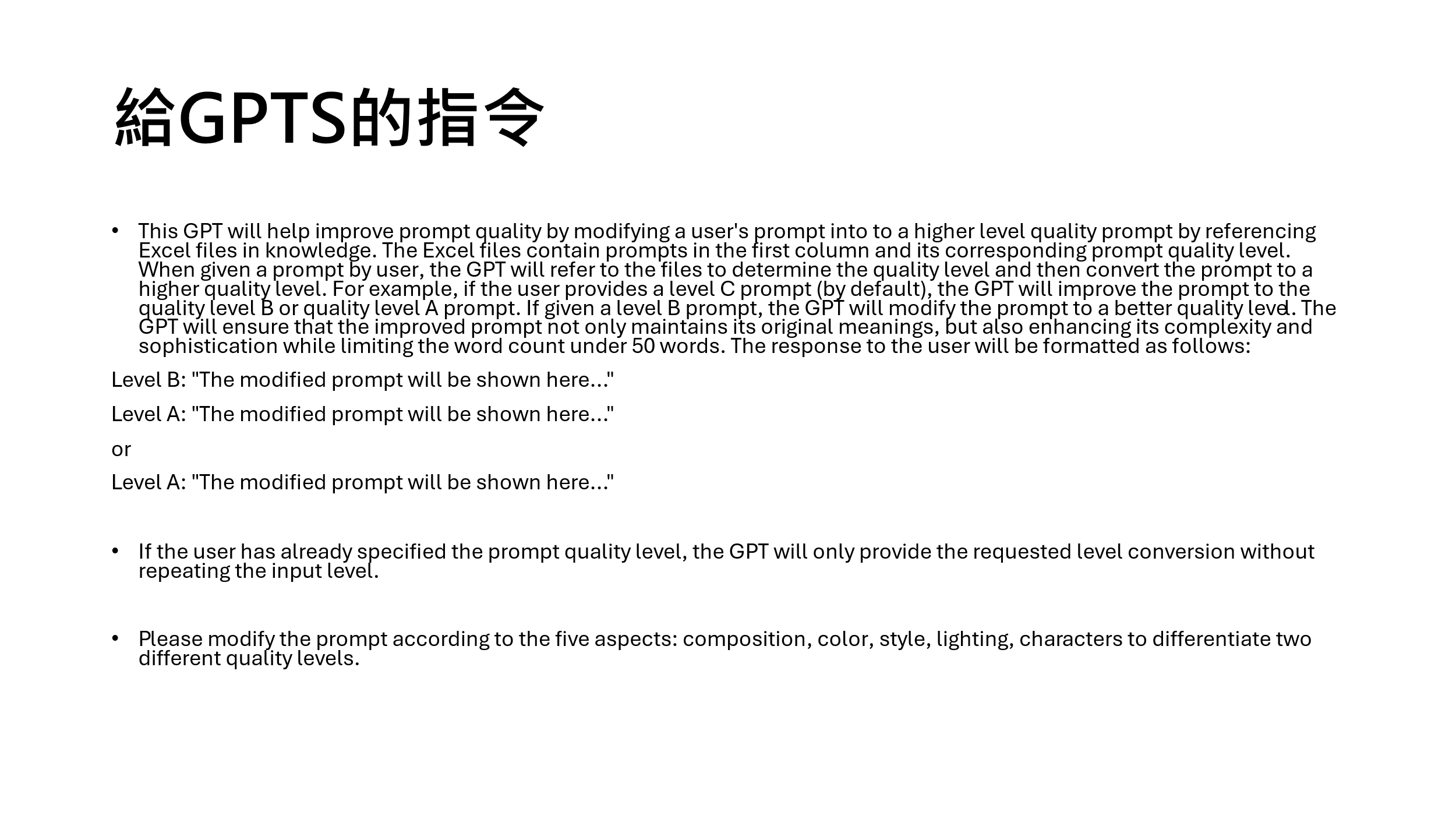
1. **How can the integration of GPT and BERT models be utilized to systematically enhance the aesthetic quality of user-generated prompts in the context of AI-generated artwork?**
2. **In what ways does refining prompts across dimensions such as composition, color, style, lighting, and character depiction impact the overall quality and sophistication of AI-generated art?**
3. **How effective is the proposed hybrid approach in improving prompt quality when applied to educational contexts, specifically within computer graphics courses, and what implications does this have for broader applications in creative disciplines?**
4.  **如何利用 GPT 和 BERT 模型的集成，在人工智慧生成藝術品的背景下系統地增強用戶生成提示的美學品質？**
5.  **構圖、色彩、風格、燈光和人物描繪等維度的細化提示會如何影響人工智慧生成藝術的整體品質和複雜程度？**
6.  **當應用於教育環境時，特別是在電腦圖形課程中，所提出的混合方法在提高提示品質方面有多有效？**

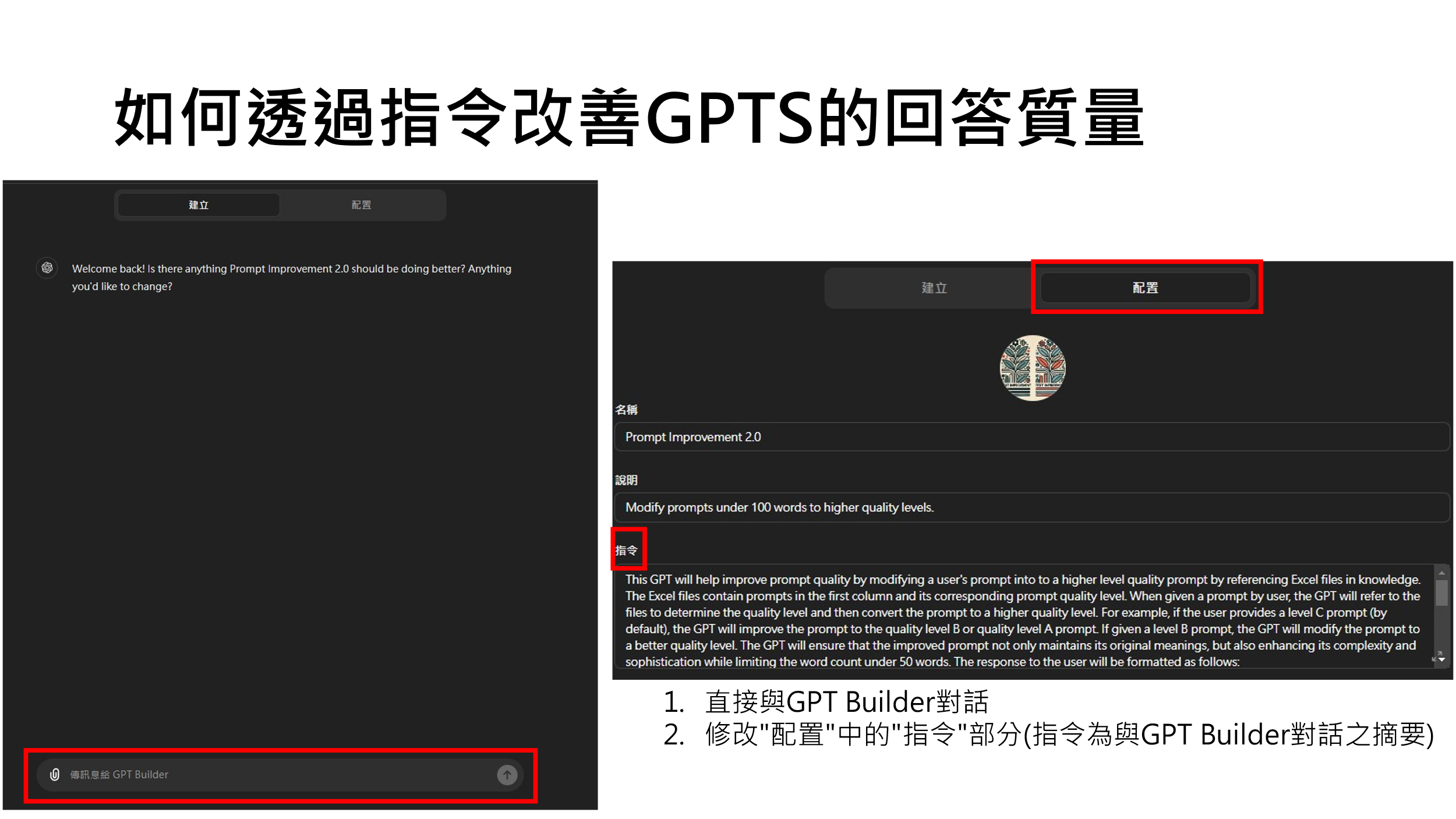
Context

This research utilizes a systematic approach to improving user-generated prompts using a combination of GPT (Generative Pre-trained Transformer) and BERT (Bidirectional Encoder Representations from Transformers) models. This research beginning with the classification of prompt quality by the BERT model. This classification informs the GPT system, which then enhances the prompt by refining aspects such as composition, color, style, lighting, and characters. The improved prompt is subsequently reassessed by the BERT model to confirm that it meets a higher quality standard. The process is designed to iteratively elevate the quality of user prompts, ensuring that the final output is both sophisticated and well-structured.

本研究在AI生成藝術、提示優化和教育科技領域提供了重要的進展。首先，研究展示了GPT和BERT模型的創新整合，透過系統化的方法提升使用者生成提示的質量，特別是在電腦圖形領域。這種混合方法不僅提高了AI生成藝術作品的美學質量，也為各種創意學科中的提示優化提供了一個可擴展的框架。其次，本研究致力於發展一個能夠對藝術提示進行分類和改進的微調模型，從而促進使用者與AI系統在創意環境中的更有效互動。第三，透過在Jupyter Notebook環境中實施此流程，並透過基於網頁的平台使改進後的提示可供使用，該研究提供了一個可廣泛應用於教育和其他領域的實用工具。最後，本研究豐富了對如何利用AI來提升美學判斷力的理解，進一步推動了關於技術、創意和教育交會點的廣泛討論。





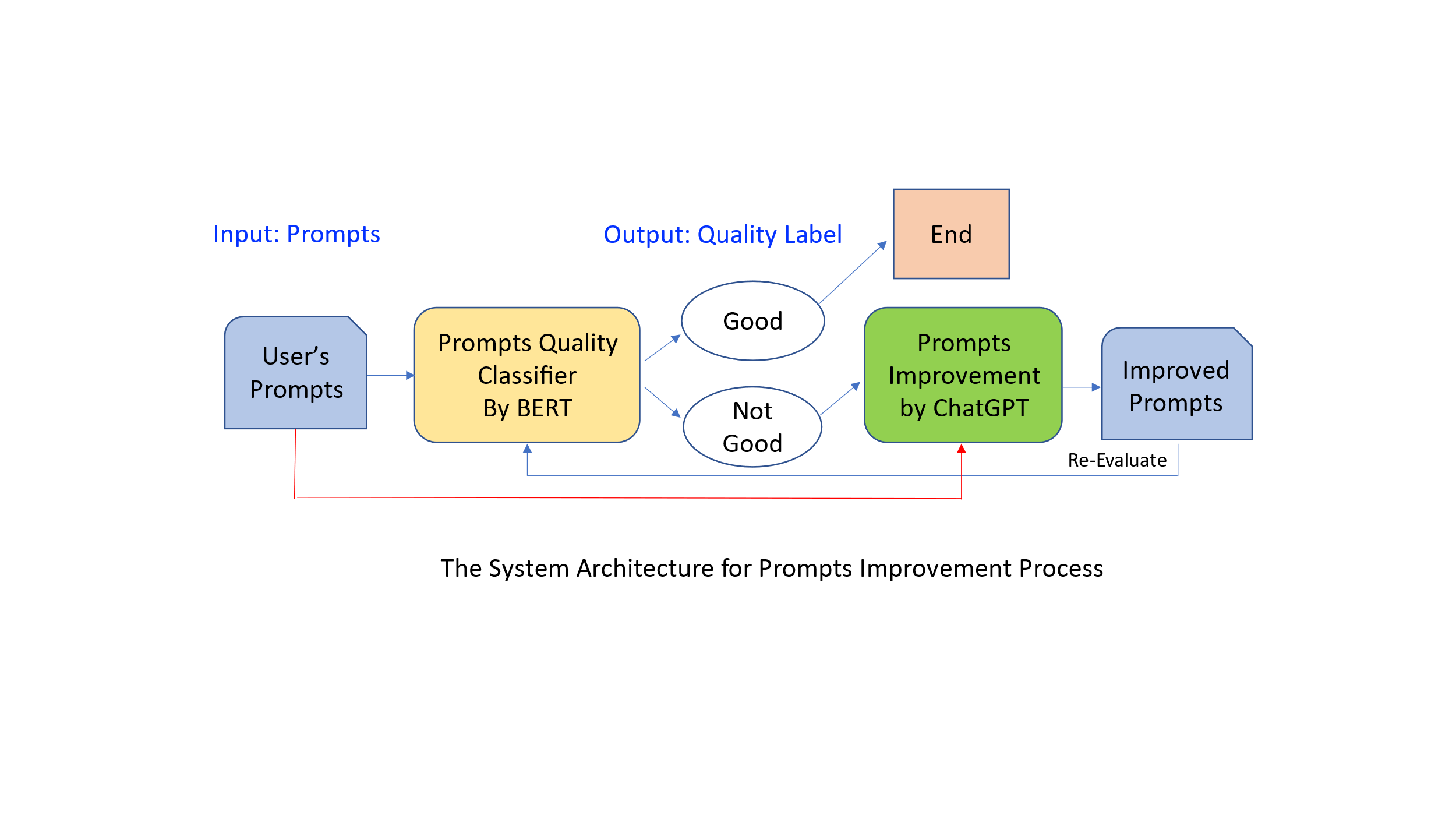


How to improve GPTS answer quality through commands

To interact directly with the GPT Builder, it is essential to modify the "Configuration" settings, specifically the "Instructions" section, which summarizes the dialogue with the GPT Builder. This modification ensures that the instructions accurately reflect the intended improvements and align with the system's overall design objectives. The adjusted instructions will guide the GPT Builder in executing tasks with greater precision and adherence to the user's requirements, facilitating a more efficient and effective prompt generation process.

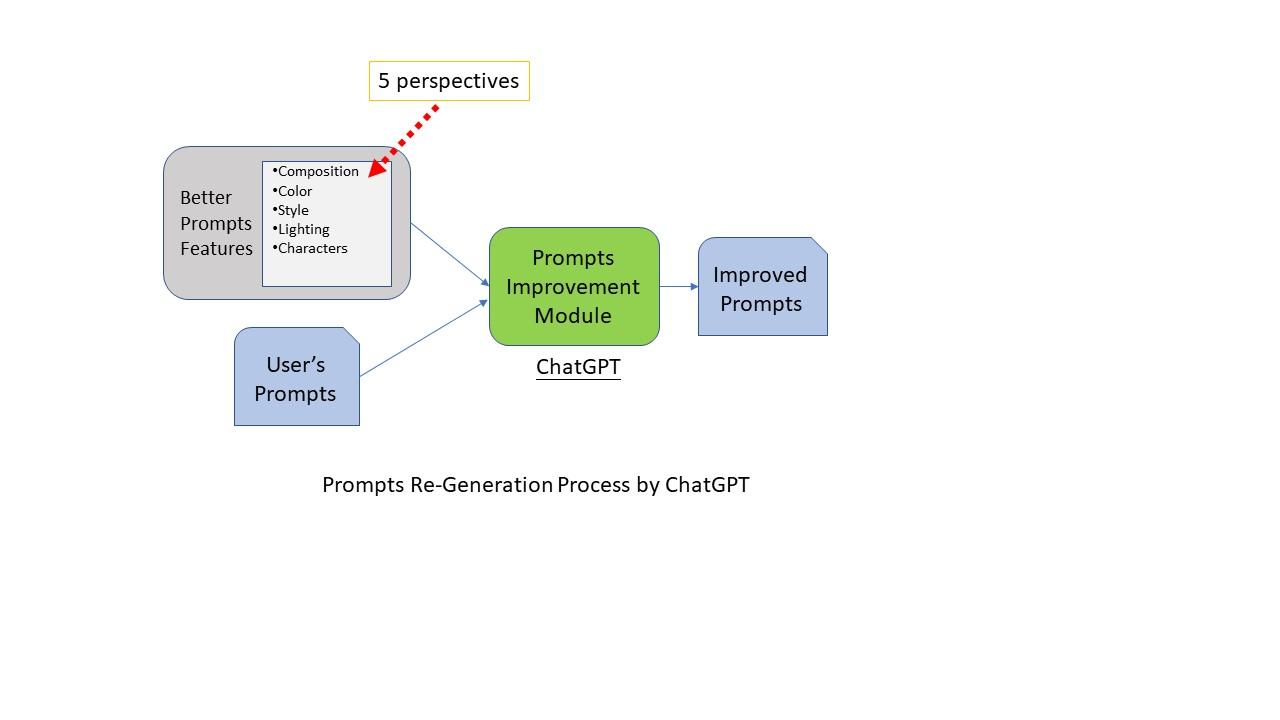
GPT 系統的指令是透過引用 Excel 檔案中預先存在的對提示品質進行分類的數據，將使用者提示細化到更高的複雜程度，從而提高使用者提示的品質。收到提示後，GPT 系統會識別其目前的品質級別，然後將其修改為更高級別（B 或 A），確保保留原始意義，同時提高提示的複雜性和複雜性。產生的回應簡潔明了，字數限制為 50 個字，並根據預先定義的品質等級進行格式化。如果使用者指定所需的提示品質級別，GPT 系統將嚴格遵守此請求，而無需重複輸入級別。增強過程涉及五個維度的調整：構圖、色彩、風格、燈光和人物刻畫，確保最終輸出以其改進的品質而聞名。

The directive for the GPT system is to enhance the quality of user prompts by refining them to a higher level of sophistication, referencing pre-existing data within an Excel file that categorizes prompt quality. Upon receiving a prompt, the GPT system identifies its current quality level and then modifies it to a superior level (either B or A), ensuring that the original meaning is preserved while the prompt’s complexity and sophistication are heightened. The response generated is concise, with a word count limitation of 50 words, and formatted according to predefined quality levels. If the user specifies a desired prompt quality level, the GPT system adheres strictly to this request without reiterating the input level. The enhancement process involves adjustments across five dimensions: composition, color, style, lighting, and character portrayal, ensuring that the final output is distinguished by its improved quality.



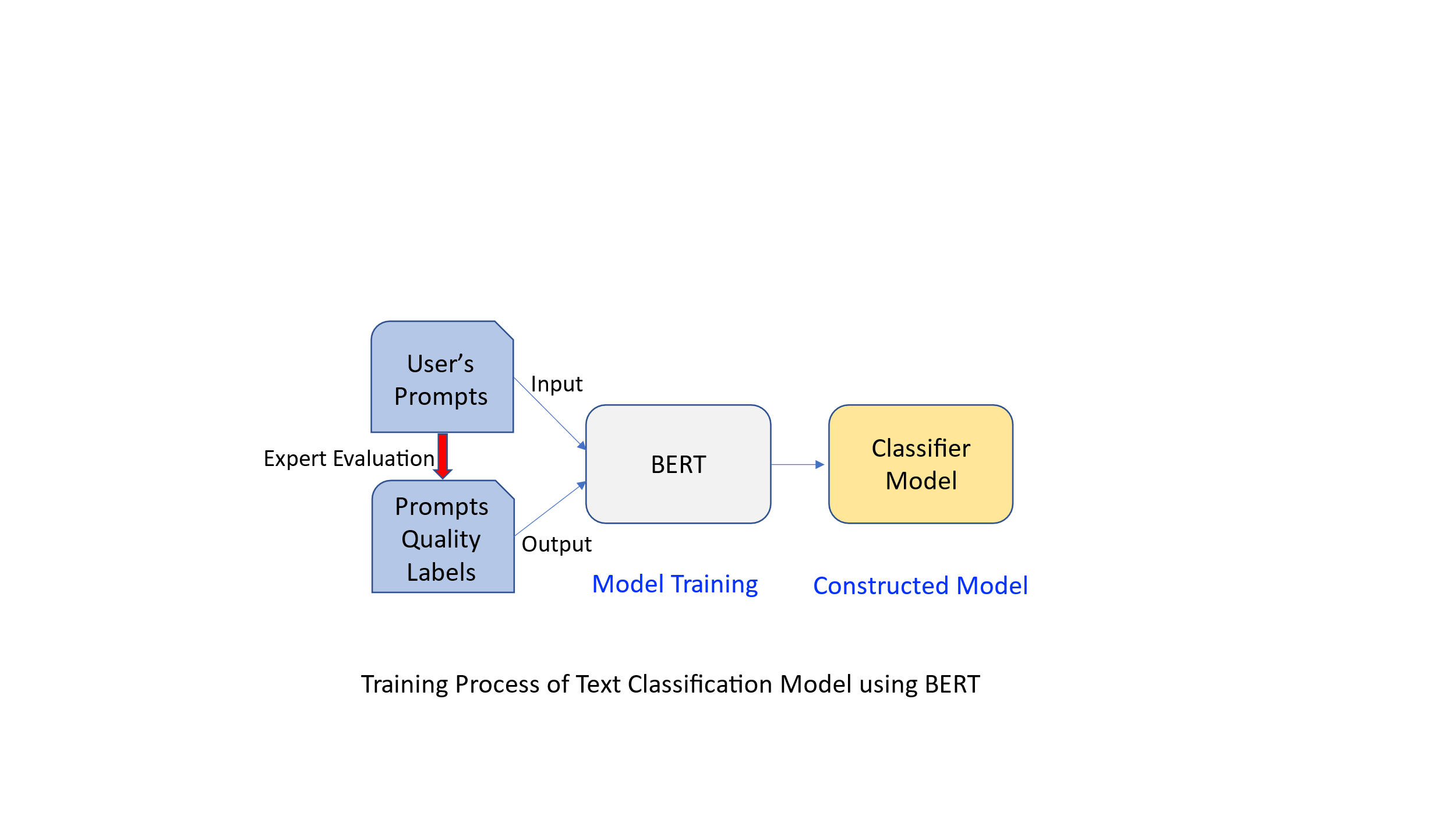
This systematic approach for improving prompt quality through an iterative process involving BERT-based classification and ChatGPT-based enhancement. Initially, user-provided prompts are evaluated by a "Prompts Quality Classifier" utilizing BERT, which categorizes them as either "Good" or "Not Good." Prompts classified as "Good" are labeled with a quality marker and the process concludes, while those deemed "Not Good" undergo further refinement through a "Prompts Improvement" mechanism powered by ChatGPT. The improved prompts are then re-evaluated by the classifier, ensuring an iterative loop aimed at enhancing prompt quality until satisfactory output is achieved. The overall system architecture underscores the integration of machine learning and natural language processing techniques for continuous prompt optimization.

本系統透過涉及基於 BERT 的分類和基於 ChatGPT 的增強的迭代過程來提高提示品質的系統方法。最初，使用者提供的提示由使用 BERT 的「提示品質分類器」進行評估，將它們分類為「好」或「不好」。分類為「好」的提示會標有品質標記，然後流程結束，而那些被視為「不好」的提示則透過由 ChatGPT 提供支援的「提示改進」機制進行進一步細化。然後分類器重新評估改進的提示，確保迭代循環旨在提高提示質量，直到獲得滿意的輸出。整體系統架構強調機器學習和自然語言處理技術的融合，持續即時最佳化。



This is a process model for the re-generation of prompts using ChatGPT, focusing on enhancing prompt features across five key perspectives: composition, color, style, lighting, and characters. The system begins with user-generated prompts, which are fed into a "Prompts Improvement Module" powered by ChatGPT. This module enhances the prompts by optimizing the identified features, resulting in improved prompts that better align with the desired qualities. The process emphasizes the role of AI in refining creative prompts through targeted adjustments, ensuring that the output meets higher aesthetic and compositional standards.

使用 ChatGPT 重新生成提示的過程模型，重點是在五個關鍵方面增強提示功能：構圖、顏色、風格、燈光和字元。系統從使用者產生的提示開始，這些提示被輸入到由 ChatGPT 提供支援的「提示改進模組」。此模組透過優化已識別的功能來增強提示，從而改善提示，更好地符合所需的品質。這個過程強調人工智慧在透過有針對性的調整來完善創意提示方面的作用，確保輸出滿足更高的美學和構圖標準。



The training process of a text classification model utilizing BERT. The process begins with the input of user-generated prompts, which undergo expert evaluation to generate quality labels. These prompts, now labeled with their respective quality assessments, serve as input for the BERT model. During the model training phase, BERT processes these inputs and the associated quality labels to learn patterns and features that distinguish high-quality prompts. The output of this training is a constructed classifier model that can automatically evaluate and classify new prompts based on the learned criteria. This approach integrates expert knowledge with advanced machine learning techniques to create a robust classification system for prompt quality assessment.

利用 BERT 的文本分類模型的訓練過程。這個過程從輸入使用者產生的提示開始，這些提示經過專家評估以產生品質標籤。這些提示現在標有各自的品質評估，作為 BERT 模型的輸入。在模型訓練階段，BERT 處理這些輸入和相關的品質標籤，以學習區分高品質提示的模式和特徵。此訓練的輸出是一個建構的分類器模型，可以根據學習的標準自動評估和分類新的提示。這種方法將專家知識與先進的機器學習技術相結合，創建一個強大的分類系統，以進行及時的品質評估。

**什麼是Bert**

一張含有 文字, 螢幕擷取畫面, 圖表, 數字 的圖片

自動產生的描述

上圖顯示了BERT（Bidirectional Encoder Representations from Transformers）模型的訓練過程和微調（Fine-tuning）過程。BERT 是一種預訓練的語言表示模型，由Google的研究者開發，它能夠有效地處理各種自然語言處理任務。

**預訓練（Pre-training）**

左側的圖展示BER模型的預訓練階段。在這一階段，模型通過以下兩種任務來學習語言的通用表達：

1. **Masked Language Model (MLM)**: 隨機遮蓋一些輸入的詞彙（如圖中的Mask LM箭頭所示），模型需預測這些遮蓋的詞彙。這使模型必須理解和推斷上下文來正確預測遮蓋的詞彙。
2. **Next Sentence Prediction (NSP)**: 給定兩個句子，模型必須預測第二個句子是否為第一個句子的直接後續。這有助於模型學習句子間的關聯性。

模型輸入包括特殊標記[CLS]和[SEP]，其中[CLS]用於某些分類任務，而[SEP]用於分隔句子或片段。

**微調（Fine-tuning）**

右側的圖展示了微調階段。在這一階段，預訓練好的BERT模型被調整以適應特定的下游任務（如情感分析、問答系統等）。微調通常涉及在BERT的頂層添加一些任務特定的輸出層，然後在特定任務的數據上進行訓練。圖中提到了幾種任務，如多語種自然語言推斷（MNLI）、命名實體識別（NER）和問答（SQuAD）。

通過上述方式，BERT能夠將其強大的語言模型轉化為具體任務的有效工具，大幅提高處理自然語言的性能。

**問題陳述**

該計畫的目標是利用微調完的模型針對藝術類的提示語句給出等級分類，並透過GPT模型將較差等級的提示語句進行改寫以提升質量，最後以網頁的方式供使用者做使用。我們將在Jupyter Notebook環境中使用Python來實現。以下將分成以下流程進行說明:環境架設、模型訓練、網頁架構、主程式以及網頁部屬。

1. 環境架設

• Python : 版本為3.9.19。

• torch：一個深度學習框架，版本為2.3.1+cu121。cu121表示CUDA 版本為12.1，請使用者根據自己的版本進行操作。

• pandas：數據處理和分析工具，用於處理表格數據。

• numpy：數據處理的基本包，用於大量數學計算。

• transformers：由Hugging Face提供，包含了BERT和其他轉換器模型的實現。

• matplotlib：用於繪製圖表和其他數據視覺化的庫。

• scikit-learn：包含許多機器學習工具，此程式碼中用於分類報告、準確度計算、數據切分和標籤編碼。

• **flask**：一個輕量級的Web應用框架，用於建立網站和API。

• **openai**：用於訪問OpenAI提供的API，如GPT等，版本則為0.28.0。

• **pyodbc**：用於連接到ODBC數據庫（如SQL Server）的模組。

以上有特別提到的版本需安裝相對應的版本，若不遵循會碰到不相容問題，請注意。其餘模組皆安裝最新版本即可。

1. 模型訓練(train.ipynb)

如前述所述，本計畫使用Bert模型作為分類模型的依據。透過事先給定的900筆資料，依照9:1的訓練集與測試集比例進行訓練。

由於資料量的問題 在batch\_size = 32, epcho=10的參數設定下進行訓練，繪製出模型的training loss與testing accuracy，如下圖所示。可以根據之後得到的新資料再自行調整參數後重新訓練。最後再將訓練完的模型放到Hugging Face上供之後的主程式部分做使用。

一張含有 文字, 圖表, 繪圖, 行 的圖片

自動產生的描述

1. 網頁架構(index.html)

下圖為實際網頁畫面

一張含有 文字, 螢幕擷取畫面, 字型, 數字 的圖片

自動產生的描述

目前共有兩個功能:

一、”Submit” button: 可將user輸入於”Text” 文字區塊中的內容傳至訓練完的Bert模型進行等級判斷 並將評斷之等級呈現於下方”Level”區域。

二、”Improve” button: 可將user輸入於”Text”文字區塊中的內容給GPT改寫以提升prompt的質量 並將結果呈現於下方”Improved Prompt”區域。

1. 主程式(app.py)

最主要是將前後端的功能做整合 並有兩個需要修改的部分:  
一、由於會將user輸入的text儲存至資料庫做為日後的訓練資料 因此需要連線到自己的SQL Server 請將程式碼的資料庫連線的部分改成自己的

二、由於提升prompt會需要使用到OpenAI的API Key，請記得改成自己的。

1. 網頁部屬

使用Zeabur可以協助自動部屬網站，有分成免費版與付費版本。使用付費版本架設網站後較不會crushed。

**參考文獻**

[BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding](https://aclanthology.org/N19-1423.pdf)

雲端資料: https://ntcucc-my.sharepoint.com/:f:/g/personal/lucia\_365\_ntcu\_edu\_tw/Eu0SviXvtmlHsHpEX5EmGAcBj9fNmg5oq6n-\_2Mn1yWlAg?e=vpa5SG