

REALIZADO POR:
Pedro Palisa.

IDENTIDAD GRAL.

ALVEAR

COMISION: 76165

Resumen

El objetivo de este proyecto es desarrollar un sistema que genere automáticamente un emblema para un club de fútbol basado en una breve descripción sobre el mismo y que representa.

Objetivos

- Generar una descripción textual del emblema basado en una breve descripción sobre el mismo y que representa.
- Convertir la descripción textual en una imagen.

Introducción

En un pueblo localizado en la provincia de Buenos Aires, se encuentra un club de fútbol llamado Club Gral. Alvear el cual acaba de ascender a Primera C. Debido a esto, los socios del club desean diseñar el emblema del club.

Este proyecto desea abordar la problemática puntualmente desarrollando un sistema que genere automáticamente un emblema para el club utilizando la API de Gemini para ***gemini-1.5-flash*** para generar una descripción textual del emblema, a partir de la descripción de los socios sobre el club y que representa el mismo para ellos. Luego se empleará DALL-E para transformar esa descripción en un diseño visual del emblema.

Justificación de la viabilidad del proyecto

Este proyecto es factible debido a que al ser un club de barrio, facilita que gran parte del trabajo creativo sea realizado por las herramientas de IA, que a su vez las mismas aceleran significativamente el proceso de diseño comparado con los métodos tradicionales.

Cálculo de costos

En cuanto a los costos del mismo, la utilización de **gemini-1.5-flash** en estos momentos es gratuita; mientras que los costos de la utilización del modelo DALL-E-3 según la información de OpenAI tiene los siguientes costos:

- **Input (tokens de entrada):** \$0.60 USD por 1 millón de tokens

Dado que el prompt de salida generado tiene un límite de 1000 tokens, podemos calcular el costo aproximado de la siguiente manera:

Costo por tokens de entrada:

- $1000 \text{ tokens} / 1,000,000 \text{ tokens} = 0.001$
- $0.001 * \$0.60 = \0.0006 USD

Y al día de hoy el costo de generación de imagen con calidad estándar y tamaño 1024x1024 es de \$0.04 USD, por lo tanto el costo total de la operación sería:

- $0.0006 + 0.04 = \$0.0406 \text{ USD}$

Metodología

- Enviar un prompt describiendo brevemente el club y lo que representa.
- Generación de texto: Utilizar Generative AI de Google para generar una descripción de un emblema mediante la técnica de prompting conocida como "In-context Learning"
- Generación de imagen: Utilizar DALL-E para convertir la descripción en una imagen.

Herramientas y tecnologías

- Python
- Generative AI gemini-1.5-flash API
- OpenAI DALL-E API

Implementación

Código utilizado para la generación de prompt texto a texto

```
import google.generativeai as genai
import openai
import requests
from io import BytesIO
from PIL import Image

genai.configure(api_key="AQUI-VA-API-Key-GenAI")
model = genai.GenerativeModel("gemini-1.5-flash")

generation_config = genai.GenerationConfig(
    max_output_tokens=1000,
    temperature=0.1,
)

# Instrucción del sistema
system_instruction = "Eres un diseñador gráfico experto que describe imágenes de forma simple y concisa. Tu objetivo es realizar una descripción de un emblema heráldico del club Gral. Alvear que pueda ser introducida como prompt en un modelo texto a imagen basado en una descripción del club. Los colores del club son azul marino, plateado y dorado"

user_prompt = input("Ingresa tu prompt: ")

prompt = system_instruction + "\n" + user_prompt

# Generar la descripción con Google Generative AI
response = model.generate_content(
    prompt,
    generation_config=generation_config,
)

description = response.text.strip()
print("Descripción generada:", description)
```

Donde se utiliza ***system_instruction*** para darle contexto a la IA sobre cómo debe responder y ***generation_config*** para darle un límite de tokens de salida, y se almacena la respuesta en una variable para luego generar la imagen.

Código utilizado para la generación de prompt texto a imagen

```
# Configurando la API KEY de OpenAI
openai.api_key = "AQUI-VA-OpenAI-API-Key"

# Generar la imagen con OpenAI DALL-E usando la descripción generada
response_image = openai.Image.create(
    model="dall-e-3",
    prompt=description,
    size="1024x1024"
)

# Obtener el link de la imagen generada
url_image = response_image['data'][0]['url']
print("URL de la imagen generada:", url_image)

# Guardar la imagen
imagen = Image.open(BytesIO(requests.get(url_image).content))
imagen.save('imagen.png')
print("La imagen ha sido guardada como 'imagen.png'.")
```

Donde se utiliza la respuesta generada de texto a texto y se genera una imagen con la descripción.

Link de colab donde se encuentra el código completo:

https://colab.research.google.com/drive/1ZWPeONrNZbRj90DyznASy_Euk349xEWD#scrollTo=TA0p0oPipXCG

Ejemplo y resultados

Ingresa tu prompt : El club fue fundado en 1960 en un pueblo llamado General Alvear. El club es nuestro orgullo, un lugar de encuentro donde se tienen valores como el esfuerzo, la perseverancia y el compañerismo, donde se puede respirar la pasión por el fútbol y el orgullo de pertenecer y representar a este pueblo hermoso.

print(salida): Un escudo heráldico azul marino con un borde dorado. En el centro, una estrella plateada de cinco puntas sobre una pelota de fútbol dorada. Debajo, la inscripción "Gral. Alvear 1960" en dorado.

Imagen generada:



Conclusión:

Al desarrollar el proyecto, se puede concluir que se obtuvieron resultados positivos utilizando la IA ya que logramos el objetivo planteado a un bajo costo y de forma rápida sin la necesidad de métodos tradicionales.