[Fecha]

YOSHUA CALAHORRANO Y JOHN GUERRA

uNIVERSIDAD cENTRAL DEL ECUADOR

CURSO: SIS8-001

iDENTIFICACION DE VEHICULOS USANDO INTELIGENCIA ARTIFICIAL GENERATIVA CON GEMINI 2.0 FLASH

Contenido

[1. **Introducción** 2](#_Toc203215984)

[2. **Objetivo General** 2](#_Toc203215985)

[3. **Objetivos específicos** 2](#_Toc203215986)

[4. **Arquitectura del sistema** 2](#_Toc203215987)

[5. **Detalles de implementación** 2](#_Toc203215988)

[6. **Costos estimados** 3](#_Toc203215989)

[7. **Ventajas de la propuesta** 3](#_Toc203215990)

[8. **Conclusión** 3](#_Toc203215991)

1. **Introducción**

La identificación de vehículos a partir de imágenes ha sido de manera tradicional tomada en cuenta por medio de modelos de clasificación por redes neuronales convolucionales, pero, con el surgimiento de modelos generativos multimodales como la IA Generativa de Google, puede ser posible formular el problema integrando capacidades de visión por computadora y comprender de manera contextual.  
Este proyecto propone desarrollar una aplicación móvil que logre identificar vehículos mediante captura de imágenes y obtener la información requerida como: marca, modelo, ano aproximado, precio referencial con respecto a Ecuador y una breve reseña del este, simplificando así la necesidad de entrenamiento de modelos locales.

1. **Objetivo General**

Desarrollar una aplicación móvil que permita identificar vehículos a partir de imágenes mediante el uso de la API Gemini 2.0 Flash, con técnicas de detección de objetos con YOLO, para lograr otorgar información específica y clara.

1. **Objetivos específicos**

* Implementar un sistema de detección de vehículos en tiempo real utilizando YOLO.
* Integrar la API Gemini 2.0 Flash para generar descripciones del vehículo deseado.
* Diseñar y desarrollar una aplicación móvil con interfaz intuitiva para la captura e identificación.
* Calcular el costo estimado del uso de la API y optimizar su consumo mediante prompts y catching.

1. **Arquitectura del sistema**

**Entrada:** Imagen capturada por la cámara del dispositivo.

**Módulo de detección:** YOLO (You Only Look Once) para aislar el vehículo.

**Módulo de inferencia:** Gemini 2.0 Flash para generar información textual.

**Salida:** JSON estructurado con marca, modelo, año, país precio y reseña.

**Interfaz:** Aplicación móvil construida con React Native.

1. **Detalles de implementación**
2. **Gemini 2.0 Flash:** Modelo multimodal de Google AI Studio, constando con precios por token: $0.10 (input), $0.40 (output) por 1M tokens. Permite procesamiento de imagen y generación de texto en tiempo real.
3. **YOLO:** Framework ligero para detectar objetos para localizar el vehículo en la imagen.
4. **Prompt:** Comando integrado en el código de programación para que interactúe la aplicación con el API y otorgue la información específica.
5. **Costos estimados**

* **Uso promedio por imagen:** 500 tokens input + 1000 tokens output = 1500 tokens.
* **Costo estimado por imagen:**
* **Costo por 1000 imágenes:** $0.45 USD

1. **Ventajas de la propuesta**

* Elimina la necesidad de data sets extensos y entrenamiento local y continuo para situaciones de no existir la información del vehículo en el entrenamiento.
* Mantiene una precisión alta y contexto enriquecido por medio de Gemini.
* Se mantiene costos bajos por consulta y escalabilidad.
* Portabilidad a través de la aplicación móvil intuitiva.

1. **Conclusión**

El uso de modelos generativos como Gemini 2.0 nos ayuda a replantear soluciones tradicionales de clasificación de imágenes. Este proyecto integra herramientas modernas de detección y generación de información contextual en contexto vehicular para ofrecer una solución eficiente y con potencial de impacto real.