**Integrante 4 - Nicolás Di Domenico:**

##### APIs de imagen y video.

* + [Vision AI](https://cloud.google.com/vision) (Cloud Vision)
  + [Video AI](https://cloud.google.com/video-intelligence)

##### APIs de documentos y datos.

* + [Document AI](https://cloud.google.com/document-ai)

**Que desarrollar en cada servicio escogido:**

* Como se llama.
* Cómo funciona.
* Para qué sirve.
* Ventajas.
* Desventajas.
* Beneficios vs costos (opcional).

**Google Cloud ofrece** APIs de imagen y video como **Vision AI** y **Video AI**. Y tambíen APIs de documentos y datos como **Document AI** que están basados en inteligencia artificial y machine learning.

Comenzando con…

**Vision AI**

También conocido como Cloud Vision API, es parte de los servicios de inteligencia artificial que ofrece Google Cloud y se enfoca en la visión artificial para el análisis y procesamiento de imágenes, videos y documentos.

### **¿Cómo funciona?**

**Vision AI** se basa en un concepto importante dentro de la inteligencia artificial, que es el de **visión artificial**. Este es un campo de la inteligencia artificial que se encarga de enseñarle a las computadoras a, por así decirlo,"ver" y entender imágenes o videos, de manera similar a como lo hace un ser humano. Pero para las computadoras, "ver" no es lo mismo que para nosotros. Mientras que los humanos tenemos ojos y un cerebro que interpreta lo que vemos gracias a años y años de evolución, las computadoras necesitan de sensores y modelos matemáticos para analizar los datos de manera visual y así poder comprenderlos y mas tarde compararlos.

Entonces para una máquina, una imagen va a ser un montón de números que representan colores y píxeles. **La visión artificial** justamente se va a encargar de convertir esos números en información útil, cómo identificar un objeto, reconocer un rostro, o leer el texto en una imagen. Y Bueno esto se va a lograr a través de algoritmos avanzados que encuentran patrones y características que se van a ir repitiendo en diferentes imágenes.

**Y es acá donde la visión artificial** **se apoya en machine learning haciendo posible que la computadora aprenda a identificar esos patrones y características sin necesidad de programar explícitamente cada objeto.** Entonces Vision AI va autilizar modelos de **machine learning** que ya han sido entrenados por Google, a través de millones de ejemplos que le fueron tirando, o sea se utilizaron grandes cantidades de imágenes para enseñarle al modelo a distinguir diferentes objetos. Y bueno este entrenamiento es lo que le va a permitir a la máquina a "aprender a ver".

Una vez que los modelos han aprendido a identificar diferentes cosas en las imágenes, Vision AI pone esas capacidades a disposición de cualquier persona a través de una API y asi permitiendo que los desarrolladores envíen imágenes a Vision AI, y reciban la info sobre lo que hay en estas, sin necesidad de tener conocimientos avanzados de inteligencia artificial.

### **¿Para qué sirve?**

**Vision AI** tiene múltiples aplicaciones prácticas, cómo por ejemplo:

* **Detectar texto en archivos sin procesar y resumirlos automáticamente:** donde básicamente subis un documento al **Cloud Storage de google** y automáticamente se activa un proceso que extrae el texto del PDF, lo resume, y luego lo almacena en una base de datos donde se puede consultar.
* **Búsqueda de productos basados en imágenes para comercio electrónico:** Imaginen que tienen una foto de la remera o pantalón que les gusta ya sea porque sacaste de internet o porque le sacaron una foto en algún local, entonces a través de la Product Search API de vision AI, solo necesitas subir esa imagen. Ya que los comercios tienen fotos de sus productos almacenadas, entonces cuando al subis, **Vision AI** analiza y la compara con el catálogo del comercio. Haciendo que te muestra una lista de productos que son visualmente similares o justamente similar al que estaba buscando.

### **Crea un flujo de procesamiento de imágenes:** donde vision AI, utiliza modelos ya entrenados para revisar cada imagen que los usuarios suben, y luego genera anotaciones o etiquetas (como "perro", "árbol", etc.). También puede ayudar a filtrar contenido inapropiado o digitalizar texto de documentos. Esta herramienta se las voy a mostrar a lo último.

* **Consigue descripciones de imágenes automatizadas con la IA generativa**: por ejemplo si tengo un catálogo de productos para mi negocio y necesito una descripción rápida de cada uno. Vision AI puede ver la imagen del catalogo y describir cada uno de los productos que quiero vender
* **Extrae texto y estadísticas de documentos con IA generativa:** Document AI, una parte de Vision AI, puede leer estos documentos, extraer la información clave (como nombres, fechas, cifras), y convertirla en datos organizados que puedes usar.
* **Inspección visual de alta precisión:** Visual Inspection AI puede revisar productos en una línea de producción para detectar defectos automáticamente, como si estuviera haciendo una inspección de calidad. Entonces se entrena al modelo para asi reconocer lo que se considera "defectuoso" y luego se encarga de encontrar esos defectos en tiempo real.

### **Ventajas**

* **Automatización de Procesos Complejos**: Vision AI permite automatizar tareas que tradicionalmente requerirían intervención humana, como la lectura y resumen de documentos, inspección de calidad en fábricas, y clasificación de imágenes. Esto ahorra tiempo y recursos.
* **Facilidad de Integración con API**: Gracias a su API sencilla, los desarrolladores pueden integrar Vision AI en sus aplicaciones con poco esfuerzo, sin necesidad de tener conocimientos profundos en inteligencia artificial o visión artificial.
* **Escalabilidad y Flexibilidad**: Al funcionar en la nube, permite escalar capaz de escalar verticalmente sin necesidad de gestionar servidores o infraestructura local siendo mas caro no. Además, se puede adaptar a diversas necesidades comerciales, desde la búsqueda de productos hasta la moderación de contenido.
* **Modelos Pre Entrenados**: Al utilizar modelos ya entrenados por Google, se evita el costo y el tiempo necesario para desarrollar modelos desde cero. Igualmente recordemos que si se necesita de modelos más concretos que resuelva necesidades mas específicas de negocio lo podemos hacer a través de **AutoML** que ya bien explicó el compañero Estefano.
* **Reducción de Errores Humanos**: Al automatizar tareas repetitivas que están sujetas al error humano, se mejora la precisión de los procesos. Esto es importante en áreas críticas como por ejemplo en la producción industrial, donde un error puede afectar la calidad del producto, o en la gestión de datos de un hospital, donde la precisión es justamente vital para el cuidado del paciente.

### **Desventajas**

* **Limitaciones en Personalización**: Como ya mencioné en las ventajas, Visual IA proporciona modelos pre-enetrenados, eso puede llevarnos a una desventaja que es justamente que al estar adecuado para tipos especificos de dominios de negocio vamos a tener que recurrir a otros modelos que a través de autoML le enseñemos que tipos de respuesta queremos según nuestras necesidades particulares.
* **Privacidad y Seguridad de los Datos**: Enviar datos visuales a la nube para ser procesados puede representar un riesgo para empresas que manejen información sensible. Aunque Google implementa medidas de seguridad avanzadas, algunos usuarios podrían tener inquietudes sobre el almacenamiento y el acceso a sus datos en la nube.

**Ejemplo de uso de Vision AI en “Try the API”:** <https://cloud.google.com/vision?hl=es#how-it-works>

**Video AI**

…va a realizar un análisis preciso de videos, reconociendo una gran cantidad de objetos, lugares y acciones. ¿Peor como se fiderencia de vision AI si dije que tambièn analiza videos?, bueno….Vision **AI** se centra principalmente en el análisis de imágenes estáticas (analizando imágenes individuales extraídas de un video), **Video AI** está diseñado específicamente para trabajar con los videos completos, tanto los que están almacenamiento local o en la nube como una transmisión en vivo, y siendo capaz de procesar enteramente los videos, analizando las escenas, identificando objetos en movimientos tareas específicas de un objeto persona, y esto lo hace desde el inicio hasta el fin del flujo del video.

Bueno…**el funcionamiento** de **Video AI** es similar al de **Vision AI** y se realiza a través de la **API de Video Intelligence**, que emplea modelos de **machine learning** previamente entrenados, entonces le enviamos un video a esta api y nos devuelve **metadatos detallados** sobre el contenido del video como por ejemplo etiquetas de objetos como auto, persona, perro, con una marca de tiempo que indica cuándo aparece un objeto. entre otras funciones que después vamos a ver.

las ventajas y desventajas van a ser muy similares a la de vision AI, pero orientado al procesamiento de información de los videos.

### **¿Para qué sirve?**

###### Anuncios contextuales: Puedes identificar las secciones de los videos más adecuadas para insertar anuncios pertinentes para el contexto según el contenido del video. Para ello, basta con asociar las etiquetas de tu video, relacionadas con marcas de tiempo específicas, con el contenido de tus anuncios.

* **Procesa videos en streaming**: Si estás trabajando con videos en tiempo real, Vision AI puede analizar lo que está ocurriendo en el video a medida que se reproduce. Los datos analizados se almacenan para que puedas hacer búsquedas sobre lo que ocurrió en el video, como "¿Cuántas veces apareció el auto rojo?").
* En la pagina de Google cloud mencionaba 2 ejemplos concretos como:
* detectar pozos en las calles de la ciudad en tiempo real haciendo que cuando va pasando un camión que los detecta y rellena a partir de este servicio de GCP.
* Había tambien un caso de un padre que grababa todos los sucesos o eventos familiares y los subia al cloud storage, entonces un momento la hija a través de esta API creó un album familiar que mostraba imagenes o secciones de video a partir de las grabaciones del padre, con la posibilidad de tener una barra de búsqueda para determinar como como se queria mostar todo eso, por ejemplo se filtraba cuando habia un cumpleaños, o fotos de nacimiento de la hija, o cuando la la mujer estab embarazada. O sea la api recorta fragmentos del video segun el criterio de entarda

### **Ventajas:**

Las ventaja van a ser la Flexibilidad, la escalabilidad, la facilidad de integracion, **Reducción de Errores Humanos,**  etc.

### **Desventajas:**

Las desventajas vamos a seguir teniendo los **problemas de Privacidad** con los datos sensibles cuando se analizan los videos.

**Limitaciones en Tiempo Real**: Aunque es posible procesar videos en streaming, hay ciertos casos en los que la latencia puede ser un problema si se necesita una respuesta inmediata, especialmente cuando se analiza contenido crítico o situaciones de seguridad.

Y bueno otra desventtaja al igual que Vision AI es que **los** modelos preentrenados de video AI pueden no cubrir necesidades específicas de ciertos negocios. Para estos casos, se necesita usar **AutoML Video**, que permite entrenar los modelos no.

**Document AI**

**Document AI** de Google Cloud es un servicio de inteligencia artificial diseñado para automatizar la extracción y comprensión de datos a partir de documentos, ya sean estructurados o no estructurados. Si bien Vision tambien puede leer documentos extrayendolos como imagenes y detectando etiquetas en el texto, Docuemn AI busca obtener un entendicmiento completo logrando **analizar, clasificar y estructurar la información extraída.**

**¿Cómo funciona?** Utiliza **modelos de machine learning** y **IA generativa** para digitalizar documentos, clasificar información, y extraer datos relevantes de forma precisa.. Document AI emplea **procesadores de documentos** que extraen texto y datos estructurados de documentos escaneados usando **OCR**, permitiendo que se almacenen, clasifiquen y analicen de manera automatizada. Los modelos preentrenados se ajustan a necesidades específicas mediante la personalización con Document AI Workbench, facilitando su uso para diferentes tipos de documentos.

*“****Document AI*** *funciona usando* ***inteligencia artificial*** *para digitalizar, clasificar y extraer datos importantes de documentos. Primero, los documentos pasan por un* ***preprocesamiento*** *donde se mejora su calidad, eliminando ruido o imperfecciones que puedan dificultar la lectura. Luego, se* ***clasifican y separan*** *las diferentes partes del documento, identificando las páginas más importantes.*

*Después, se realiza la* ***extracción de datos*** *con tecnologías como* ***OCR (Reconocimiento Óptico de Caracteres)****, que convierten las palabras e información del documento en datos digitales y estructurados. Estos datos, finalmente, pasan por una* ***validación****, usando reglas específicas para asegurar que todo se ha extraído correctamente y de forma precisa.*

*Por último, si la empresa tiene necesidades específicas, se puede* ***personalizar el modelo*** *usando* ***Document AI Workbench****.*

*Por ejemplo: imagina una tienda de ropa que recibe facturas de sus proveedores.* ***Document AI*** *puede leer estas facturas y extraer datos básicos como el nombre del proveedor y el monto total. Sin embargo, si la tienda necesita extraer detalles más específicos, como el número de lote de los productos o las instrucciones de envío, puede personalizar el modelo usando* ***Document AI Workbench*** *para que también capture esa información. Esto ayuda a la tienda a tener un control más detallado sobre su inventario y pedidos, automatizando el proceso y haciendo que todo sea más eficiente.”*

**¿Para qué sirve?** Document AI se usa para automatizar procesos empresariales, clasificar y analizar formularios, digitalizar documentos antiguos, y generar estadísticas que ayudan en la toma de decisiones. También es útil para casos como el análisis de formularios y la extracción de datos para alimentar sistemas empresariales.

**Ventajas** incluyen rapidez en la obtención de valor, alta precisión, integración con otras herramientas de Google Cloud y la capacidad de tomar decisiones mejor informadas al estructurar la información.

**Desventajas**: Por otro lado, existen preocupaciones sobre la privacidad al procesar datos sensibles en la nube, limitaciones en la personalización de los modelos preentrenados para necesidades específicas, y la dependencia de una conexión a internet estable.

**Ejemplo de uso de Document AI en “Try the API”:** <https://cloud.google.com/document-ai?hl=es_419>

(opcional)