

Procesamiento de resultados de APIs de IA

En esta práctica, aprenderás a manejar y analizar la respuesta JSON de 2 APIs de Inteligencia Artificial de AWS: **DetectFaces** y **DetectText**. Estas herramientas permiten identificar rostros y textos en imágenes, generando resultados estructurados que contienen información detallada sobre los objetos detectados.

La práctica se divide en dos partes:

1. **Detección de rostros:** Utilizarás los resultados del servicio **DetectFaces** para extraer y procesar información sobre las características faciales detectadas. Esto incluye datos sobre edad, emociones y detalles geométricos de las caras.
2. **Detección de texto:** Analizarás la salida de **DetectText**, que identifica y devuelve los textos encontrados en una imagen, junto con información sobre su ubicación y nivel de confianza en la detección.

Ambas partes te permitirán familiarizarte con la manipulación de datos en formato JSON, la extracción de información relevante y la visualización o interpretación de resultados.

Para resolver las prácticas tendrás que programar un script de Python (.py) diferente para cada parte, que procese cada fichero y muestre en pantalla lo que se solicita.

Parte 1: reconocimiento facial

El servicio **DetectFaces** de AWS analiza imágenes y proporciona información detallada sobre las características de los rostros detectados, como edad estimada, género, emociones, presencia de accesorios, y puntos de referencia faciales.

El objetivo principal de esta parte es que aprendas a:

1. **Cargar y explorar datos en formato JSON**, un estándar muy utilizado en la transmisión de datos en aplicaciones web y servicios de IA.
2. **Extraer y manipular información anidada** del JSON, interpretando los resultados que ofrece la API.
3. **Realizar análisis básicos y avanzados**, incluyendo el cálculo de estadísticas, identificación de patrones, y visualización de datos.
4. **Aplicar conceptos de programación en Python** relacionados con estructuras de datos y procesamiento de información en IA.

Esta práctica también permitirá familiarizarse con la **interpretación de resultados de servicios de IA**, una habilidad esencial para aquellos que trabajan en el ámbito de la inteligencia artificial y el análisis de datos.

Se te va a proporcionar un fichero .json (facedetails.json5) con la salida de la API tras el envío de una imagen que contiene una cara, tendrás que procesar ese json y mostrar en pantalla el análisis según estas tareas.

Tareas específicas que debes realizar:

1. Carga y exploración del JSON:

- Cargar el archivo JSON usando la biblioteca `json`.
- Identificar las claves principales y sus subniveles (`FaceDetails`, `Emotions`, `Landmarks`, etc.).

2. Extracción y muestra de información básica:

- Buscar y mostrar la edad estimada (rango bajo y alto).
- Mostrar si la persona lleva gafas y/o gafas de sol (y la confianza asociada).
- Determinar el género detectado y su confianza.

3. Análisis de Emociones:

- Identificar la emoción con el nivel de confianza más alto, y la de nivel más bajo.
- Calcular y mostrar un gráfico de barras con la distribución de emociones (usando `matplotlib` o `seaborn`, muestra el máximo de 100).

4. Procesamiento de Puntos de Referencia (Landmarks) [opcional]:

- Extraer las coordenadas de los puntos clave de la cara (por ejemplo, ojos, boca, nariz), luego, visualizar estos puntos en una gráfica de dispersión, simulando un mapa facial.

Parte 2: detección de textos

En esta parte de la práctica, trabajarás con la salida JSON generada por el servicio **DetectText** de AWS, mediante un fichero `.json` que se te va a proporcionar (`textdetections.json5`). Esta salida contiene información sobre los textos detectados en una imagen, sus coordenadas geométricas y el nivel de confianza de cada detección.

Objetivos:

- Extraer los textos detectados junto con su nivel de confianza.
- Analizar la estructura jerárquica de detección (líneas y palabras).
- Visualizar los textos ordenados según su posición en la imagen.
- Identificar y destacar palabras específicas según criterios (por ejemplo, nivel de confianza).

Tareas específicas que debes realizar:

1. Carga del JSON: importa el archivo JSON proporcionado, usando la librería `json` de Python, y carga los datos en un diccionario.

2. Exploración de la estructura:

- Identifica las claves principales: `TextDetections` y `TextModelVersion`.
- Observa la organización de los textos en líneas (`LINE`) y palabras (`WORD`).
- Nota la relación entre `ParentId` (vincula palabras a su línea correspondiente).

3. Extracción de textos y confianzas:

- Recorre la lista `TextDetections` para obtener:
 - El texto detectado (`DetectedText`).
 - El nivel de confianza (`Confidence`).
 - El tipo de elemento (`Type`: `LINE` o `WORD`).
- Muestra o almacena los textos junto con su nivel de confianza.

4. Organización y visualización:

- Agrupa las palabras en líneas usando la relación `ParentId`.
- Ordena las líneas según su posición vertical (`Top` de `BoundingBox`).
- Muestra las líneas de texto completas en el orden correcto.

Ejemplo de salida esperada:

Línea 1: IT'S (Confianza: 99.35%)

Línea 2: MONDAY (Confianza: 99.62%)

Línea 3: but keep (Confianza promedio: 99.59%)

Línea 4: Smiling (Confianza: 88.95%)

Texto completo detectado:

"IT'S MONDAY but keep Smiling"