European IT Certification Curriculum
Self-Learning Preparatory Materials

EITC/AI/GVAPI

Google Vision API

This document constitutes European IT Certification curriculum self-learning preparatory material for the EITC/AI/GVAPI Google Vision API programme.

This self-learning preparatory material covers requirements of the corresponding EITC certification programme examination. It is intended to facilitate certification programme's participant learning and preparation towards the EITC/AI/GVAPI Google Vision API programme examination. The knowledge contained within the material is sufficient to pass the corresponding EITC certification examination in regard to relevant curriculum parts. The document specifies the knowledge and skills that participants of the EITC/AI/GVAPI Google Vision API certification programme should have in order to attain the corresponding EITC certificate.

Disclaimer

This document has been automatically generated and published based on the most recent updates of the EITC/AI/GVAPI Google Vision API certification programme curriculum as published on its relevant webpage, accessible at:

https://eitca.org/certification/eitc-ai-gvapi-google-vision-api/

As such, despite every effort to make it complete and corresponding with the current EITC curriculum it may contain inaccuracies and incomplete sections, subject to ongoing updates and corrections directly on the EITC webpage. No warranty is given by EITCI as a publisher in regard to completeness of the information contained within the document and neither shall EITCI be responsible or liable for any errors, omissions, inaccuracies, losses or damages whatsoever arising by virtue of such information or any instructions or advice contained within this publication. Changes in the document may be made by EITCI at its own discretion and at any time without notice, to maintain relevance of the self-learning material with the most current EITC curriculum. The self-learning preparatory material is provided by EITCI free of charge and does not constitute the paid certification service, the costs of which cover examination, certification and verification procedures, as well as related infrastructures.

EITC/AI/GVAPI GOOGLE VISION API DIDACTIC MATERIALS
LESSON: UNDERSTANDING WEB VISUAL DATA
TOPIC: DETECTING WEB ENTITIES AND PAGES

INTRODUCTION

Artificial Intelligence - Google Vision API - Understanding web visual data - Detecting web entities and pages

Artificial Intelligence (AI) has revolutionized various industries, and one area where it has made significant advancements is in computer vision. Google Vision API is a powerful tool that utilizes AI to analyze and understand visual data. In this didactic material, we will consider the topic of understanding web visual data using the Google Vision API, specifically focusing on detecting web entities and pages.

Web visual data refers to images and videos that are available on the internet. With the vast amount of visual information present online, it becomes important to extract meaningful insights from this data. Google Vision API provides a comprehensive set of features that enable developers to analyze and interpret web visual data effectively.

One of the key functionalities of the Google Vision API is the ability to detect web entities. Web entities are objects, landmarks, or other visual elements present in an image or video. By leveraging machine learning models, the API can accurately identify and classify these entities. This capability opens up possibilities for various applications, such as content moderation, image recognition, and visual search.

To detect web entities using the Google Vision API, developers can make use of the `webDetection` feature. This feature provides information about web entities found in an image or video, including their descriptions, relevance scores, and full matching images. By analyzing these results, developers can gain insights into the visual content and extract valuable information.

In addition to detecting web entities, the Google Vision API also allows for the identification of web pages associated with an image or video. This feature, known as `webPageDetection`, provides information about web pages that contain similar or visually related content. By analyzing the web page information, developers can gain a deeper understanding of the context in which the visual data is being used.

The process of detecting web entities and pages using the Google Vision API involves several steps. First, the API receives an image or video as input. It then analyzes the visual content using its machine learning models to identify relevant entities and web pages. The API provides a structured response containing the detected entities and associated web page information.

Developers can integrate the Google Vision API into their applications using various programming languages and frameworks. The API provides a simple and well-documented interface, allowing developers to easily make requests and process the responses. By leveraging the power of the Google Cloud Platform, developers can scale their applications to handle large volumes of web visual data efficiently.

The Google Vision API is a powerful tool for understanding web visual data. By utilizing its capabilities to detect web entities and pages, developers can extract valuable insights from images and videos available on the internet. This opens up possibilities for a wide range of applications, including content moderation, image recognition, and visual search.

DETAILED DIDACTIC MATERIAL

The Google Vision API provides a powerful tool for detecting web entities and pages based on visual data. This feature allows users to search for similar images or images with 4-pixel matching, making it useful for tasks such as identifying stolen photos or generating tags for uploaded images.

To begin, users can upload an image to a website or the web and utilize the Web Detection feature to search for identical or visually similar images. This can be particularly helpful in cases where someone has stolen a photo and the user wants to find other instances of the image on the web.

Additionally, the Web Detection feature can be used to automatically generate tags for images uploaded to platforms like Flickr. This eliminates the need for manual tagging and keyword guessing, as the API can recommend relevant tags based on the visual content of the image.

To demonstrate the usage of the Google Vision API for detecting web entities and pages, we will use Python. First, we define the file names of the images we want to analyze. In this example, we have three images: a cat, a popular web image, and a photo taken in Japan.

Next, we open each image and pass the byte data to a content object. This content object is then used to construct an image object using the vision module. Once the image object is created, we can make use of the web_detection method from the client instance to obtain the response.

The response object contains a list of results, including web entities, visually similar images, and best-guess labels. Web entities provide recommended tags that can be associated with the image, while visually similar images are images found on the web that closely resemble the uploaded image. The best-guess label represents the label name that can be used for the image.

To extract the web detection response, we create a web detection object and reference the web detection attribute from the response object. This object provides several methods to access different elements of the detection, such as best-guess labels, full match images, partial matching images, and pages with matching images.

By calling these methods, we can retrieve the corresponding attributes of each element. For example, the best-guess label attribute returns the recommended label associated with the image, while the full match images attribute provides a list of images that are identical to the uploaded image. The partial matching images attribute returns a list of images that partially match the uploaded image.

Furthermore, the API allows us to search for visually similar images. By clicking on the links provided in the response, we can view these visually similar images in the Chrome browser. This can help users explore different images recommended by Google based on their similarity to the uploaded image.

The Google Vision API provides a convenient way to detect web entities and pages based on visual data. By utilizing the Web Detection feature, users can search for similar images, generate tags for uploaded images, and explore visually similar images recommended by the API.

When using the Google Vision API for web detection and web entities, the results obtained are presented in the form of a list. To extract the relevant information, it is necessary to iterate through each element of the list. This can be achieved by implementing a loop. In this loop, the description and confidence score of each entity can be printed. The entity ID can be ignored for now.

By running this loop, a list of recommended tags suggested by the Vision API will be displayed. It is important to note that the script provided encompasses the entire process.

To demonstrate the functionality using a different image, a new image can be selected and the process can be repeated. In the example, an image downloaded from the web is utilized. It is essential to ensure that the file extension is removed before running the script. Once this is done, a greater number of results will be obtained based on the uploaded image.

The Vision API is capable of identifying various details about the image, such as the location where it was taken and the device used. In addition to this, the API also provides a method to obtain matching images. These matching images are identical to the uploaded image. By opening one of the provided links, it is possible to view the image and access the webpage associated with it. It is important to note that not all links will be functional, as some may result in a "page not found" error.

Furthermore, the Vision API allows for the identification of visually similar images from different websites. By selecting a different link, it is possible to access the same image from a different website. This information can be useful when searching for pages that contain matching images.

The API also provides additional information, such as the webpage URL, title, and the image link from the

website. It is worth mentioning that there are other details available, such as partial image matching and visually similar image links. These can be explored further on an individual basis.

This material on the Vision API for Python covers the process of understanding and utilizing web visual data. It demonstrates how to detect web entities and pages using the API.

EITC/AI/GVAPI GOOGLE VISION API - UNDERSTANDING WEB VISUAL DATA - DETECTING WEB ENTITIES AND PAGES - REVIEW QUESTIONS:

WHAT IS THE PURPOSE OF THE WEB DETECTION FEATURE IN THE GOOGLE VISION API?

The Web Detection feature in the Google Vision API serves a important role in understanding web visual data by enabling the detection of web entities and pages. This powerful tool allows developers and researchers to extract valuable information from images and videos found on the internet, expanding the capabilities of computer vision systems.

The primary purpose of the Web Detection feature is to identify and analyze objects, landmarks, and text present in web images and videos. By leveraging Google's vast knowledge graph and web index, the API can recognize a wide range of entities, including popular landmarks, famous people, animals, plants, and more. This functionality empowers developers to build applications that can automatically recognize and classify objects in web visual data.

Furthermore, the Web Detection feature provides insights into the context and relevance of web images and videos. It can detect and analyze the presence of logos, text, and other visual elements associated with specific web pages or websites. This information can be leveraged to enhance content moderation, brand monitoring, and search engine optimization strategies. For instance, a company could use the API to automatically detect and categorize user-generated content containing their logo, enabling them to track brand exposure across various online platforms.

To achieve these capabilities, the Google Vision API utilizes state-of-the-art machine learning algorithms. These algorithms are trained on vast amounts of labeled data, including images and videos from the web. By learning from this diverse dataset, the API can accurately identify and classify objects and entities in web visual data.

The didactic value of the Web Detection feature lies in its ability to provide developers and researchers with a comprehensive understanding of web visual data. By leveraging this feature, they can extract valuable insights and knowledge from the vast amount of information available on the internet. This knowledge can be used to enhance a wide range of applications, including content analysis, recommendation systems, augmented reality, and more.

The Web Detection feature in the Google Vision API plays a pivotal role in understanding web visual data by enabling the detection of web entities and pages. Its ability to recognize objects, landmarks, text, and contextual elements empowers developers and researchers to extract valuable insights and knowledge from web images and videos. By leveraging this feature, they can build powerful applications that enhance various domains, including content moderation, brand monitoring, and search engine optimization.

HOW CAN THE WEB DETECTION FEATURE BE HELPFUL IN IDENTIFYING STOLEN PHOTOS?

The Web Detection feature offered by the Google Vision API can be immensely helpful in identifying stolen photos by leveraging the power of artificial intelligence and machine learning. This feature allows users to analyze images and retrieve information about similar images found on the web. By comparing the uploaded image with a vast database of images available on the internet, the API can detect if the image has been stolen or used without proper authorization.

One way the Web Detection feature can be useful in identifying stolen photos is through its ability to find visually similar images. When a photo is stolen, it is often reposted on various websites or social media platforms. The API can scan these platforms and identify instances where the stolen photo has been shared. By providing a list of visually similar images, it becomes easier to track down the unauthorized use of the original photo.

Moreover, the Web Detection feature can also identify the web entities associated with an image. These entities include objects, landmarks, logos, and text that appear in the image. By analyzing the web entities, the API can provide additional contextual information about the image and help determine if it has been stolen. For

example, if a stolen photo of a famous landmark is detected, the API can reveal if the image is being used without proper permission.

Additionally, the API can detect the presence of visually similar pages on the web. This feature is particularly useful when dealing with images that have been modified or altered in some way. Even if the stolen photo has been edited or manipulated, the API can still identify similar pages where the modified version of the image is being used.

To illustrate the practical application of the Web Detection feature, consider a scenario where a photographer suspects that their photo has been stolen and used without permission. By uploading the photo to the Google Vision API and utilizing the Web Detection feature, the photographer can obtain a list of visually similar images found on the web. This list can then be analyzed to identify websites or social media profiles where the stolen photo is being used. This information can be important in taking appropriate legal action or requesting the removal of the unauthorized usage.

The Web Detection feature provided by the Google Vision API is an invaluable tool for identifying stolen photos. By leveraging artificial intelligence and machine learning, this feature can detect visually similar images, identify associated web entities, and locate visually similar pages on the web. This comprehensive analysis empowers users to take action against unauthorized usage of their photos and protect their intellectual property rights.

HOW DOES THE WEB DETECTION FEATURE ASSIST IN GENERATING TAGS FOR UPLOADED IMAGES?

The Web Detection feature in the Google Vision API plays a important role in assisting the generation of tags for uploaded images. By leveraging advanced artificial intelligence techniques, this feature enables the identification and extraction of relevant web entities and pages associated with an image. This process involves a comprehensive analysis of the visual content, which is then compared to a vast database of web images and information.

When an image is uploaded, the Web Detection feature utilizes powerful algorithms to analyze the visual elements within the image. It identifies various objects, landmarks, logos, and text present in the image, and then proceeds to match these visual cues with similar images and information available on the web. This matching process is performed by comparing the extracted features of the uploaded image against a large dataset of web images.

The Web Detection feature provides valuable insights by associating the uploaded image with relevant web entities. These entities can include famous landmarks, popular products, well-known brands, and even specific web pages that feature the same or similar visual content. By generating tags based on these associations, the feature facilitates a more comprehensive understanding of the image's content and context.

For example, consider a user uploading an image of the Eiffel Tower. The Web Detection feature would not only identify the Eiffel Tower as the main object in the image but also associate it with web entities such as famous landmarks, travel websites, or historical information about the Eiffel Tower. Consequently, tags generated by the feature may include "Eiffel Tower," "Paris," "landmark," "travel," and "architecture," among others.

Moreover, the Web Detection feature can assist in generating tags for images that contain less prominent objects or elements. For instance, if an image contains a lesser-known product or logo, the feature can identify and associate it with relevant web entities such as e-commerce websites, product reviews, or brand information. This enables more accurate and descriptive tags, enhancing the image's discoverability and categorization.

The Web Detection feature in the Google Vision API significantly aids in generating tags for uploaded images. By leveraging advanced AI algorithms, it identifies and associates the visual content of an image with relevant web entities and pages. This process enhances the understanding of the image's content and context, resulting in more accurate and descriptive tags.

WHAT ARE THE DIFFERENT ELEMENTS PROVIDED IN THE RESPONSE OBJECT OF THE GOOGLE VISION API'S WEB DETECTION FEATURE?

The response object of the Google Vision API's web detection feature contains several elements that provide valuable information about the web entities and pages detected in an image. These elements include web entities, full matching images, partial matching images, pages with matching images, visually similar images, and visually similar pages.

1. Web entities: The web entities element represents the entities that are related to the image based on the web. These entities can be objects, landmarks, logos, or other visual elements. Each web entity is accompanied by a description and a score that indicates the relevance of the entity to the image.

Example:

– Web entity: "Eiffel Tower"

Description: "The Eiffel Tower is a wrought-iron lattice tower located on the Champ de Mars in Paris, France."

Score: 0.9876

2. Full matching images: This element provides a list of images found on the web that are visually identical or highly similar to the input image. Each entry includes the URL of the image and the page where it was found.

Example:

– Image URL: "https://example.com/image1.jpg"

Page URL: "https://example.com/page1.html"

3. Partial matching images: Similar to full matching images, this element provides a list of images that are partially similar to the input image. These images may contain some common visual elements but are not exact matches.

Example:

– Image URL: "https://example.com/image2.jpg"

Page URL: "https://example.com/page2.html"

4. Pages with matching images: This element provides a list of web pages that contain images that are visually similar to the input image. Each entry includes the URL of the page and the URL of the matching image.

Example:

– Page URL: "https://example.com/page3.html"

Image URL: "https://example.com/image3.jpg"

5. Visually similar images: This element provides a list of images that are visually similar to the input image but may not be exact matches. These images may share common visual elements or have similar color schemes.

Example:

– Image URL: "https://example.com/image4.jpg"

6. Visually similar pages: Similar to visually similar images, this element provides a list of web pages that are visually similar to the input image. These pages may contain images with common visual elements or similar color schemes.

Example:

– Page URL: "https://example.com/page4.html"

These elements in the response object of the Google Vision API's web detection feature allow developers to gain insights into the web entities and pages associated with an image. This information can be used for various applications, such as content moderation, image search, and recommendation systems.

HOW CAN USERS EXPLORE VISUALLY SIMILAR IMAGES RECOMMENDED BY THE API?

To explore visually similar images recommended by the Google Vision API, users can leverage the powerful capabilities of the API's web entity detection feature. This feature allows users to detect and understand web entities and pages, providing them with a comprehensive understanding of the visual data present on the web.

When utilizing the Google Vision API's web entity detection, users can obtain a list of visually similar images by following these steps:

1. Send an image to the API: Users need to send an image to the API for analysis. This can be done by making a POST request to the API's web entity detection endpoint, providing the image file or URL as input.

2. Retrieve the web entities: Once the API processes the image, it returns a response containing a list of web entities detected in the image. These web entities represent objects, scenes, or concepts that are visually present in the image.

3. Extract relevant information: From the list of web entities, users can extract the relevant information for each entity. This information includes the entity's description, score, and relevance to the image.

4. Explore visually similar images: To explore visually similar images, users can utilize the information obtained from the web entities. They can perform a web search using the entity's description or use the entity's relevance score to find similar images on the web. By analyzing these visually similar images, users can gain further insights and understanding of the visual context and related content.

For example, if a user submits an image of a cat to the Google Vision API, the API may detect web entities such as "cat," "animal," and "pet." The user can then use the description "cat" to perform a web search and find visually similar images of cats. Additionally, by exploring the visually similar images, the user may discover related concepts such as "kitten," "feline," or "domestic animal."

By following these steps, users can effectively explore visually similar images recommended by the Google Vision API. This functionality can be utilized in various applications such as image search, content recommendation, and visual data analysis.

Users can explore visually similar images recommended by the Google Vision API by sending an image for analysis, retrieving the web entities detected in the image, extracting relevant information, and utilizing this information to perform a web search or find visually similar images. This process allows users to gain a deeper understanding of the visual context and related content present in the web visual data.