The program uses REST api to send requests, the request and response bodies are according to the GOOGLE API VISION Platform.

PDF text detection has been done using the function files:asyncBatchAnnotate functioN

The program uses google-cloud python library

The program takes the credentials from the JSON file which is present into the root folder.

The program takes the pdf link which is present in the G-Cloud storage bucket.

The request JSON BODY:
Input config:
In the request json we have provided the URI to the PDF
the mime type has been specified to application/PDF

Output config:
In the output config we've mentioned the destination uri, which is present in the cloud bucket,
The program automatically takes the input uri and creates a new destination based on the file name.

RESPONSE:
A successful response will return operation ID.

After parsing the data, multiple json files with pages names 1-2.json, 2-3.json are created.

The program parses the data from PDF and converts the response into the json file.

The program uses the JSON library to parse the data from the blobs.

There's a for loop which iterates over the blobs and converts the json file response into strings which are therefore feeded into an html file using (with open (write)) function available in the python.

The name of the html file is automatically from the input URI and the uri is formatted in a way that only the filename.pdf is left and this string is used as the name for HTML file.

During the data feeding, html tags(<HTML><HEAD> <BODY>) have been added,
 tags have been added to provide spacing in the output file.

Once the iteration is completed the html file will be automatically loaded into the browser and you can see the output. I've used the webbrowser module to open the html file.