

# ELECTRONIC INVOICE PROCESSING

## DOCUMENTATION

## TEAM CASECATION

### **PROBLEM STATEMENT AND OBJECTIVES:**

Use of Image Processing models such as **\*\*Optical Character Recognition (OCR)\*\*** to digitalize the process of invoice transactions. Moving towards fully automated invoice processing and extracting data from pdf/jpeg/webcam image into scalable xlsx/csv template. UI deployment using styling scripts.

### **LINK TO DEMO VIDEO:**

<https://drive.google.com/file/d/1O3WsQxDEC859AgnvO5DiPUzSidoOYvmZ/view?usp=sharing>

### **PYTHON LIBRARIES USED:**

- Pytesseract
- Pdf2image
- OpenCv
- Xlsxwriter
- NumPy
- Pandas
- Streamlit

### **WHAT THE MODEL DOES?**

- I/O: Loading data in the form of PDF/jpeg/webcam image
- Performing feature extraction using OpenCV's existing algorithms
- Mapping the extracted features into scalable excel templates which are easy to process automatically

### **LINK TO GITHUB REPO:**

<https://github.com/Dips2705/electronic-invoice-processing>

## **INSTRUCTIONS TO RUN THE CODE:**

For Conversion from PDF to Excel Invoice:

- Open Google Colab and upload the notebook 'Jupyter Files/runner.ipynb'
- Upload the required PDF to the local machine provided by the colab system.
- Run all the cells of the notebook except the last.
- In the last cell replace 'sample\_invoice.pdf' with the actual name of the PDF file uploaded.
- Refresh the files and you will see the 'converted\_xls.xlsx' file ready to download.

For Conversion from JPEG to Excel Invoice:

- Open Google Colab and upload the notebook 'Jupyter Files/jpeg\_invoice.ipynb'
- Upload the files 'python files/app.py' and 'python files/jpeg\_invoice.py' to the local machine provided by the colab system.
- Run all the cells of the notebook.
- Go to the link generated by the second last cell of the notebook which will open on to a website developed using Streamlit library of Python.
- Click on 'Convert from JPEG' from the left sidebar and then upload the required file in the designated place (twice, if needed). A preview of the image will be generated.
- Press Execute and wait for the execution to finish.
- Click on 'Download Converted Excel File' and then click on the generated link to download the corresponding excel file.
- Once the file 'converted.xlsx' is downloaded, select "End Session".