

**How to Run it:** (Click to navigate)

-  [On Google Colab](#)
-  [On Local Machine](#)

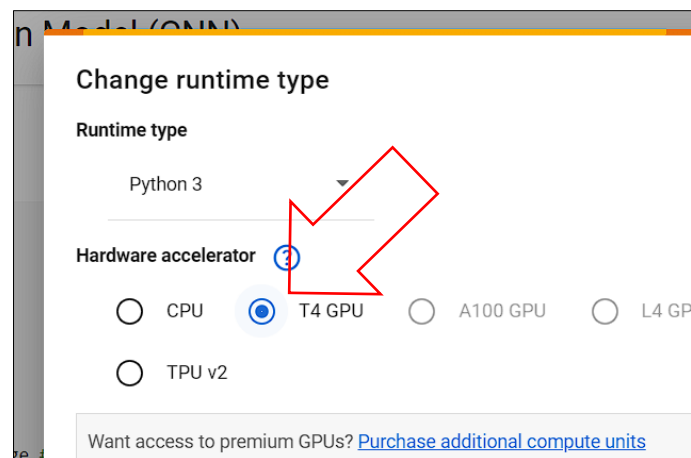
## I. Google Colab

### 1. Open the Colab link

Click on the provided Colab link to open the notebook in your browser.

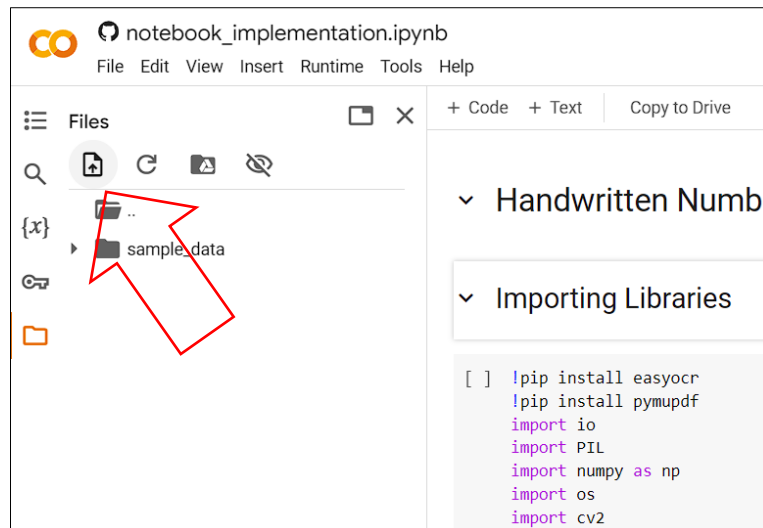
### 2. Change Runtime Type

- Go to the top menu bar.
- Click on Runtime -> Change runtime type.
- Set the **hardware accelerator** to T4 GPU.
- Then, click on **Save**.



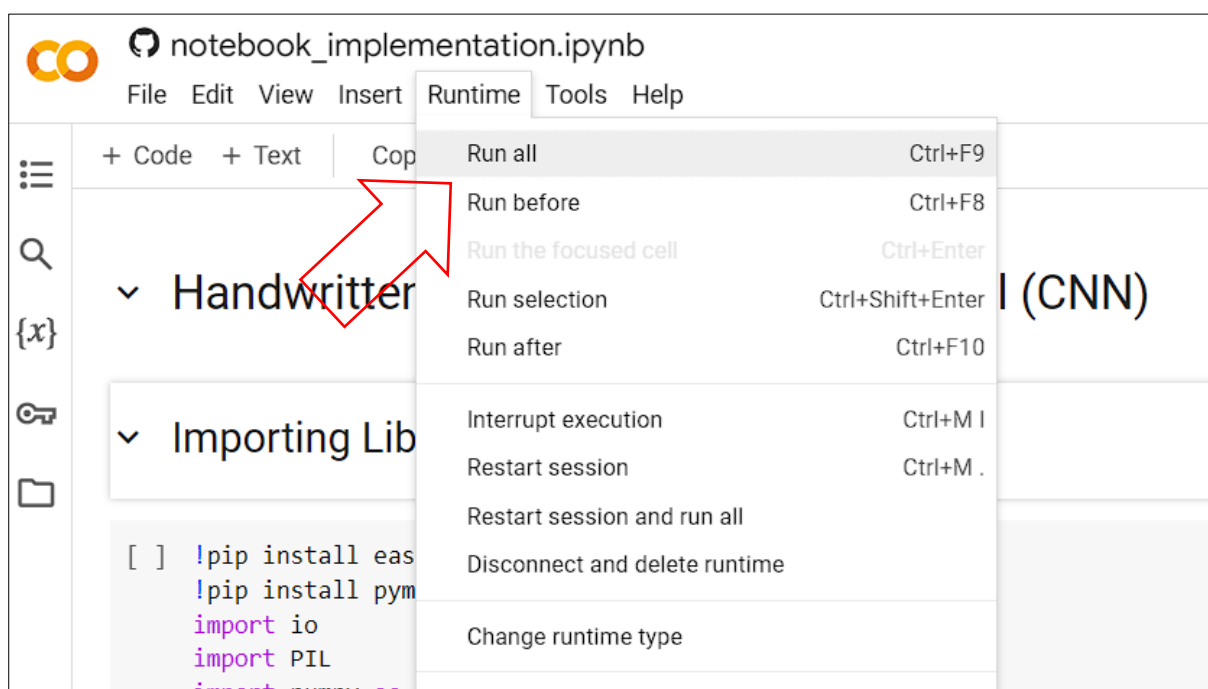
### 3. Upload PDF File

- In the code cell for file upload, click the “Run” button.
- A file uploader will appear below the cell.
- Select the PDF file you want to convert.



#### 4. Run the Code

- Click on Runtime -> Run all from the top ribbon.



#### 5. Provide Feedback

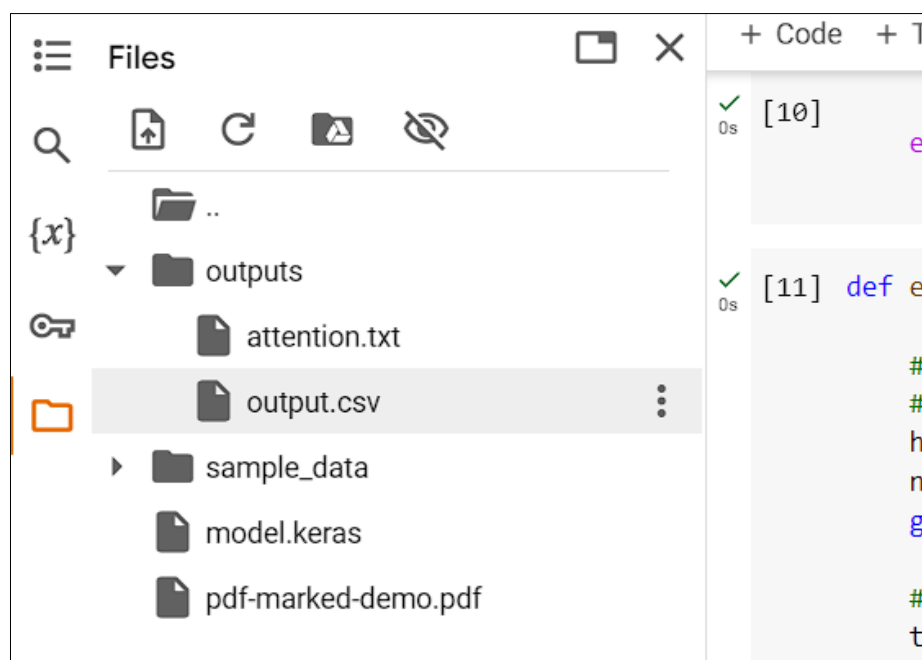
If prompted by the notebook, follow any instructions for providing feedback. This step is useful for improving the model.

```
read_pdf("pdf-marked-demo.pdf")

... 1/1 ██████████ 0s 354ms/step
0.99999785
1/1 ██████████ 0s 15ms/step
9
Predicted:9 (75.76)
Actual:
```

## 6. Check the Output

- Once the process is complete, the CSV file generated from the PDF will be available in **/outputs**
  - attention.txt contains list of cells requiring attention
  - output.csv is the generated file which may be opened with a spreadsheet software like **MS Excel**



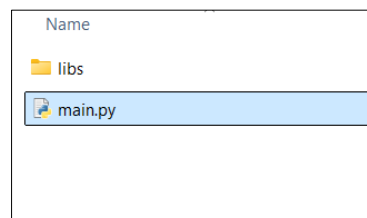
| output.csv X                 |      |                                 |             |                 |
|------------------------------|------|---------------------------------|-------------|-----------------|
| 1 to 11 of 11 entries Filter |      |                                 |             |                 |
| SNo.                         | ID   | NAME                            | DOB         | INTERVIEW MARKS |
| 1                            | 442  | Anantha N Rao (RAC, Delhi)      | 29-Nov-1999 | 39              |
| 2                            | 8319 | Deepak Kumar # (RAC, Delhi)     | 27-Mar-2000 | 88              |
| 3                            | 6220 | Jugesh Kumar (RAC, Delhi)       | 31-Dec-1997 | 83              |
| 4                            | 1465 | Krishna Nand Singh (RAC, Delhi) | 16-Mar-2000 | 88              |
| 5                            | 608  | Lilu Sahu # (RAC, Delhi)        | 04-Mar-2000 | 98              |
| 6                            | 5332 | Piyush Kumar Gupta (RAC, Delhi) | 21-Mar-2000 | 88              |
| 7                            | 691  | Pulkit Khandelwal (RAC, Delhi)  | 08-Oct-1999 | 88              |
| 8                            | 9576 | Royal Pradhan * (RAC, Delhi)    | 05-Oct-1998 | 89              |
| 9                            | 6204 | Sayan Acharya *# (RAC, Delhi)   | 16-Oct-2001 | 0               |
| 10                           | 1174 | Subhasis Biswas (RAC, Delhi)    | 26-Sep-1999 | 0               |
| 11                           | 3228 | Sujit Mal (RAC, Delhi)          | 28-Feb-1997 | 0               |

Show 25 per page

Download from runtime storage and make any required polish.

## II. On Local Machine

1. **!! Ensure python 3.10+ is installed.**
2. Navigate to **./app-implementation**



3. Run **./app-implementation/main.py**

```

C:\WINDOWS\system32\cmd.exe
2024-09-07 22:55:28.333475: I tensorflow/core/util/port.cc:113] oneDNN custom operations are on. You may see slightly different numerical results due to floating-point round-off errors from different computation orders. To turn them off, set the environment variable 'TF_ENABLE_ONEDNN_OPTS=0'.
2024-09-07 22:55:31.048431: I tensorflow/core/util/port.cc:113] oneDNN custom operations are on. You may see slightly different numerical results due to floating-point round-off errors from different computation orders. To turn them off, set the environment variable 'TF_ENABLE_ONEDNN_OPTS=0'.
2024-09-07 22:55:36.901073: I tensorflow/core/platform/cpu_feature_guard.cc:210] This TensorFlow binary is optimized to use available CPU instructions in performance-critical operations.
To enable the following instructions: AVX2 FMA, in other operations, rebuild TensorFlow with the appropriate compiler flags.
Downloading data from https://storage.googleapis.com/tensorflow/tf-keras-datasets/mnist.npz
11498434/11498434 3s 0us/step
Epoch 1/5 9s 5ms/step - accuracy: 0.8876 - loss: 0.3732 - val_accuracy: 0.9803 - val_loss: 0.0658
Epoch 2/5 8s 5ms/step - accuracy: 0.9838 - loss: 0.0587 - val_accuracy: 0.9855 - val_loss: 0.0499
Epoch 3/5 8s 5ms/step - accuracy: 0.9911 - loss: 0.0387 - val_accuracy: 0.9883 - val_loss: 0.0393
Epoch 4/5 8s 5ms/step - accuracy: 0.9923 - loss: 0.0239 - val_accuracy: 0.9864 - val_loss: 0.0479
Epoch 5/5 4s 5ms/step - accuracy: 0.9949 - loss: 0.0154

```

#### 4. Check `./output`

