Problem: Resume screening is still the most time-consuming part of recruiting. When a job opening receives 250 resumes on average 70% of them are unqualified, therefore we need a tool that Screens the most appropriate resumes from that long list on the basis of job description. That Ultimately saves a lot of human efforts and essential hours.

Approach Taken:

Please find below steps followed for resume screening parser.

- As a first step kept all my data in google drive to access it from COLAB.
- Import required modules like **pdfminer.six, nltk, numpy, pandas subprocess, re etc**
- Developed functions to get name, phone number and email address
    - Read resume pdf using pdfminer python module and converted it into text
    - Used nltk to extract name
    - Developed regex to extract email and phone number.
- Developed functions to get Technical and Non-Technical skills from text
    - read csv files for technical and Non-technical skill to create list.
    - Removed the stop words from text and apply filter on Technical and Non-technical DB list to get vice-versa skills

Interpretation of Results:

- Got exact Email & Phone Number
- Extracted Name partially
- Able to get Technical skills & Non-technical fields

Storyboarding –

Based on the results candidate have below technical skills which is required for the JOB role

```
'Php', 'Java', 'MySQL', 'Github', 'CSS', 'newspaper', 'Framework',
'framework', 'C', 'github', 'Visual', 'PostgreSQL'
```

Also we can tell more on like, we should contact candidate or not if JOB profile and role is provided.

Scope Of Improvement:

Can use NER rather than NLTK to get Name properly, Also technical skills filter parser may improve a bit.