1. from docx import Document
2. import os
3. import re
4. #should have docx library installed
5. #the resumes should be in word file
6. folder\_path = "/content/sample\_docx\_path"
7. tfp = "/content/tfp.txt"
8. files=[]
9. def find\_q(fp,tfp): # fp = folder\_path, tfp = target file path (to save output data)
10. result=open(tfp,"w")
11. for f in os.listdir(fp):
12. files.append(os.path.join(fp,f))
13. #files.pop(0) # remove this if you use this anywhere outside google colab
14. for doc in files:
15. document = Document(doc)
16. for line in document.paragraphs:
17. text = line.text
18. if re.search("PYTHON|python|Python", text): and re.search("Java|JAVA|java", text): and re.search("javascript|JAVASCRIPT|Javascript", text):
19. result.write(doc)
20. result.close()
21. find\_q(folder\_path,tfp)
22. f=open(tfp,'r')
23. txt=f.read()
24. print(txt)