RESUME SHORTLIST USING EXCEL FILE

SOURCE CODE:

```python
import os
import nltk
import pandas as pd
from nltk.corpus import stopwords
from nltk.tokenize import word_tokenize
# Configure paths
folder_path = "D:/Saketh/Python Programming/sampleprograms/NLP/NLPLab"
files = [f for f in os.listdir(folder_path) if f.endswith('.txt')]
# Download necessary NLTK resources
nltk.download('punkt', quiet=True)
nltk.download('stopwords', quiet=True)
stop_words = set(stopwords.words('english'))
keywords = ["python", "java", "javascript", "developer", "engineer"]
# Initialize lists to store shortlisted and non-shortlisted data
shortlisted_data = []
non_shortlisted_data = []
# Function to extract resume information
def extract_info(lines):
    info = {
        "Name": "",
        "Position Applied": "",
        "Skills": "",
        "Experience": "",
        "Contact": ""
    }
    for line in lines:
        if line.startswith('Name: '):
            info["Name"] = line.replace("Name: ", "").strip()
        elif line.startswith('Position Applied: '):
            info["Position Applied"] = line.replace("Position Applied: ", "").strip()
        elif line.startswith('Skills: '):
            info["Skills"] = line.replace("Skills: ", "").strip()
        elif line.startswith('Experience: '):
            info["Experience"] = line.replace("Experience: ", "").strip()
        elif line.startswith('Contact: '):
            info["Contact"] = line.replace("Contact: ", "").strip()
    # Convert skills string into a list
    skills = info["Skills"].split(',') if info["Skills"] else []
    info["Skills"] = [skill.strip() for skill in skills]
    return info
# Iterate over all files
for file in files:
    file_path = os.path.join(folder_path, file)
    with open(file_path, 'r') as f:
        content = f.read()
    lines = content.split('\n')
    # Extract information from resume
    resume_info = extract_info(lines)
    # Extract keywords from skills for shortlisting
```

```python
        resume_content = ' '.join([line for line in lines if not line.startswith("Name: ") and not line.startswith("Contact: ")])
        tokens = word_tokenize(resume_content.lower())
        filtered_tokens = [word for word in tokens if word not in stop_words]
        matches = [word for word in keywords if word in filtered_tokens]
        # Check if the resume is shortlisted
        if len(matches) >= 2:
            resume_info["Shortlisted"] = "Yes"
            shortlisted_data.append(resume_info)
        else:
            resume_info["Shortlisted"] = "No"
            non_shortlisted_data.append(resume_info)
# Convert lists to DataFrames
df_shortlisted = pd.DataFrame(shortlisted_data)
df_non_shortlisted = pd.DataFrame(non_shortlisted_data)
# Save to an Excel file with two sheets
output_file = "shortlisted_candidates.xlsx"
with pd.ExcelWriter(output_file, engine='xlsxwriter') as writer:
    df_shortlisted.to_excel(writer, sheet_name="Shortlisted", index=False)
    df_non_shortlisted.to_excel(writer, sheet_name="Non-Shortlisted", index=False)
print(f"Data has been saved to {output_file} with two sheets: 'Shortlisted' and 'Non-Shortlisted'")
```

OUTPUT:

Data has been saved to shortlisted_candidates.xlsx with two sheets: 'Shortlisted' and 'Non-Shortlisted'

SHEET 1: (Shortlisted)

	Name	Position Applied	Skills	Experience	Contact	Shortlisted
2	Vamsi Krishna Bavirisetty	Software Developer	['Python', 'Ruby', 'React', 'SQL']	3 years at	vamsib0609@gmail.com	Yes
3	Vivek Bavandlapally	Software Engineer	['Java', 'Python', 'C++', 'SQL']	5 years at	Vivek.bavandlapally@gmail.com	Yes
4	Sai Kushal	Frontend Developer	['HTML', 'CSS', 'JavaScript', 'React'	4 years at	bachukushal28@gmail.com	Yes
5	Sreedhar Reddy	Backend Developer	['Python', 'Django', 'SQL', 'AWS']	5 years at	sreedharreddybollup@gmail.cor	Yes
6	Raghu Sai Varun Oleti	Full Stack Developer	['Python', 'JavaScript', 'React', 'No	4 years at	raghusaivaruno@gmail.com	Yes
7	Saketh Birdouzee	Machine Learning Engineer	['Python', 'TensorFlow', 'Keras', 'C	3 years at	machinelearner94@gmail.com	Yes
8	Puneeth Reddy	Web Developer	['HTML', 'CSS', 'PHP', 'JavaScript']	2 years at	puneethreddypuneethreddy658	Yes
9						

SHEET 2: (Non-Shortlisted)

	Name	Position Applied	Skills	Experience	Contact	Shortlisted
2	Praveen Kumar Meda	Data Scientist	['Python',	2 years at	praveenmeda143@gmail.com	No
3	Marcus John	DevOps Engineer	['Linux', 'C	3 years at	marcusjohn22@gmail.com	No
4	Vishnu Vardhan Reddy Mukka	Data Analyst	['SQL', 'Exc	2 years at	vishnuvardhanreddy.mukka@gr	No

DEPENDENCIES TO BE INSTALLED:

pip install xlsxwriter
pip install openpyxl