

Resumé Parsing using NER

ASHISH PATIL 46

ARPIT BURLEY 9

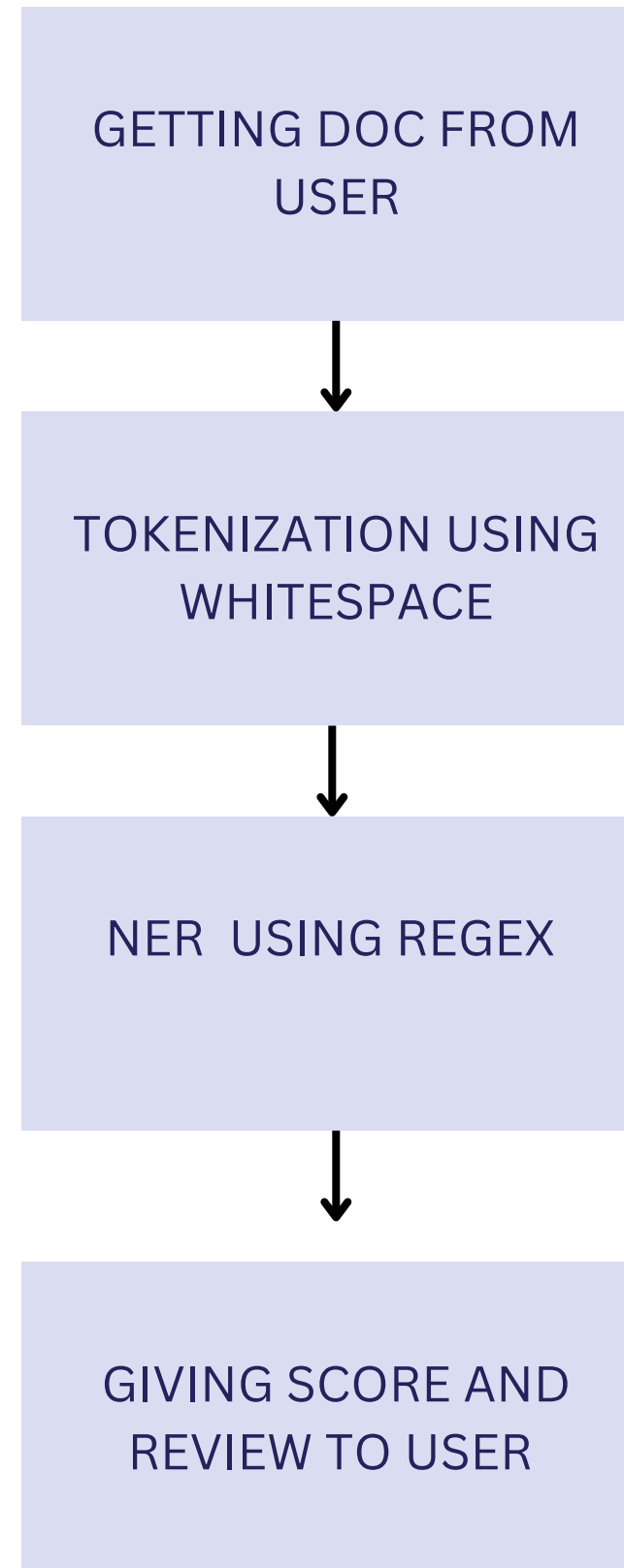
ATHARVA BAHETI 3

introduction

Without CV parsing, recruiters often spend hours reading multiple resumes to figure out the most suitable candidate. Parsing software makes this process more manageable, faster, and more cost-effective. By using algorithms and comparing candidates, parsing provides recruiters access to the best talent for the role.



Methodology



extracting names using regex

```
# extract name (assume name is the first capitalized word in the resume)
def extract_name(text):
    name_match = re.match(r"^[A-Z][a-z]+\s[A-Z][a-z]+", text) # Assumes name is the first thing in the resume
    return name_match.group() if name_match else "Not found"
```

email

```
# Extract contact info (email, phone)
def extract_contact_info(text):
    email = re.findall(r"[a-zA-Z0-9+_.-]+@[a-zA-Z0-9.-]+", text)
    phone = re.findall(r"(\(?\+?\d{1,3}?\)?-?\s?\d{1,4}-?\s?\d{1,4}-?\s?\d{4})", text)
    return {
        "Email": email[0] if email else None,
        "Phone": phone[0] if phone else None
    }
```

job titles, dates, organizations

```
# Extract organizations (companies/universities) based on common patterns
def extract_organizations(text):
    org_pattern = r"\b(?:Inc|Corporation|Corp|Ltd|LLC|University|College|Institute)\b"
    orgs = re.findall(fr"[A-Z][a-zA-Z\s]*(?:{org_pattern})", text)
    return orgs if orgs else ["Not found"]

# Extract dates (experience or education periods)
def extract_dates(text):
    date_pattern = r"\b(?:\d{2}/\d{4}|\d{4})\b"
    dates = re.findall(date_pattern, text)
    return dates if dates else ["Not found"]

# Extract job titles (look for capitalized words that follow 'at' or 'in')
def extract_job_titles(text):
    titles = re.findall(r"(?:\bat\b|\bin\b)\s([A-Z][a-zA-Z\s]+)", text)
    return titles if titles else ["Not found"]
```


extracting skills

```
# Extract skills using a predefined list
def extract_skills(text):
    skills_list = ['Python', 'Java', 'C++', 'JavaScript', 'SQL', 'HTML', 'CSS', 'Machine Learning',
                  'Data Analysis', 'React', 'Django', 'Flask', 'TensorFlow', 'Keras', 'AWS',
                  'Docker', 'Kubernetes', 'Leadership', 'Communication', 'Problem-solving']

    skills_found = [skill for skill in skills_list if skill.lower() in text.lower()]
    return skills_found if skills_found else ["No skills found"]

# Give a resume score based on presence of key elements
def calculate_resume_score(name, contact_info, organizations, dates, job_titles, skills):
    score = 0
    if name != "Not found": score += 10
    if contact_info['Email']: score += 10
    if contact_info['Phone']: score += 10
    if organizations != ["Not found"]: score += 20
    if dates != ["Not found"]: score += 20
    if job_titles != ["Not found"]: score += 20
    if skills != ["No skills found"]: score += 10
    return score
```

Enhanced Resume Parser with Scoring, Suggestions, and Visualization

Upload your resume (.pdf, .doc, .docx, .txt)



Drag and drop file here

Limit 200MB per file • PDF, DOC, DOCX, TXT

Browse files



CEC-Poor-vs-Good-Resume-2.pdf 133.5KB



Uploaded Resume Text

Caroline Jones Poor Quality Resume

72 Main Street Apt. 2

Boston, MA 02115

635-7310/ partygirl@hotmail.com

O

bjective: I am seeking a position where my education and experience will be utilize
more experience in the field of Art Administration.

E

ducation

09/[Year] - Prese nt Bachelors of Arts in Art Administration, Simmons

Honors: GPA: 2.9

Joined student clubs including, Student Government Assoc.

With Study Abroad program, traveled throughout Europe and

visited countries including England, F rance, Germany, Italy

M

ay [Year] Received diploma from City High School Boston Class of

Extracted Information

Name: Caroline Jones

Email: partygirl@hotmail.com

Phone: 635-7310

Organizations (Companies/Universities): EDUCATION Simmons University

Dates (Experience/Periods): 7310, 7310

Job Titles: Not found

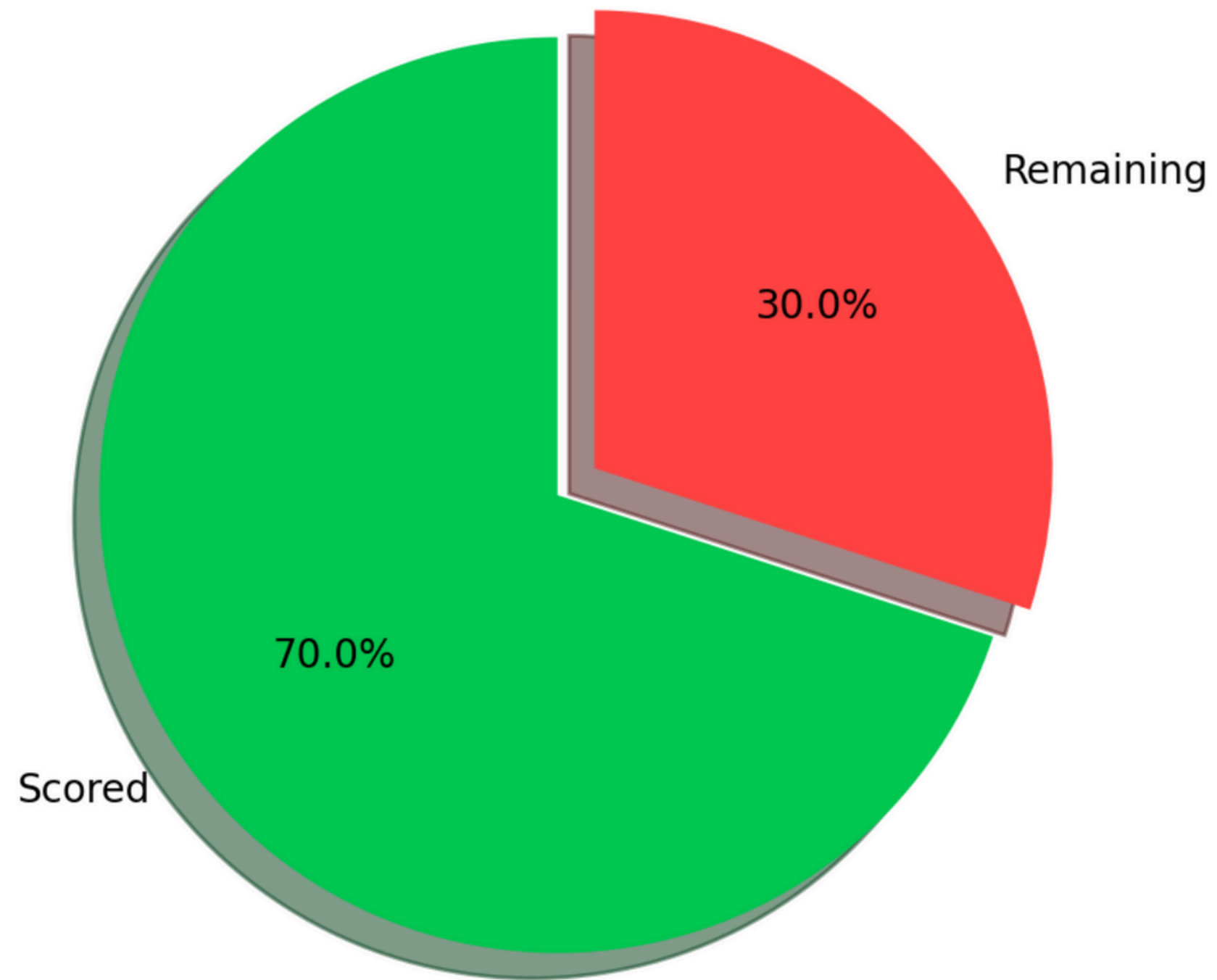
Skills: No skills found

Resume Score

Your resume score is: 70/100

Resume Score

Your resume score is: 70/100



scored



Suggestions for Improvement

- Job titles not found. Add your job titles and roles in previous companies.
- Skills not found. List technical and soft skills relevant to the job.