In this project, we are going to use spacy for entity recognition on 200 Resume and experiment around various NLP tools for text analysis. The main purpose of this project is to help recruiters go throwing hundreds of applications within a few minutes. We have also added skills match feature so that hiring managers can follow a metric that will help them to decide whether they should move to the interview stage or not. We will be using two resume datasets; 1) the first contains resume texts from https://www.kaggle.com/datasets/snehaanbhawal/resume-dataset?select=data

2) The second is https://www.kaggle.com/code/gauravduttakiit/resume-screening-using-machine-learning/data

3) and also a data https://raw.githubusercontent.com/kingabzpro/jobzilla_ai/main/jz_skill_patterns.jsonl that contains skills that we will use to create an entity ruler.

```
import pandas as pd
import os

#spacy
import spacy
from spacy.pipeline import EntityRuler
from spacy.lang.en import English
from spacy.tokens import Doc

#gensim
import gensim
from gensim import corpora

#Visualization
from spacy import displacy
import pyLDAvis
import pyLDAvis.gensim_models as gensimvis
#pyLDAvis.enable_notebook()
import pyLDAvis.gensim_models
from wordcloud import WordCloud
import plotly.express as px
import matplotlib.pyplot as plt

#Data loading/ Data manipulation
import pandas as pd
import numpy as np
import jsonlines

#nltk
import re
import nltk
from nltk.corpus import stopwords
from nltk.stem import WordNetLemmatizer
nltk.download(['stopwords','wordnet'])

#warning
import warnings
warnings.filterwarnings('ignore')
```

```
[nltk_data] Downloading package stopwords to /root/nltk_data...
[nltk_data]   Unzipping corpora/stopwords.zip.
[nltk_data] Downloading package wordnet to /root/nltk_data...
```

```
! pip install pyLDAvis
```

```
Looking in indexes: https://pypi.org/simple, https://us-python.pkg.dev/colab-wheels/public/simple/
Collecting pyLDAvis
  Downloading pyLDAvis-3.3.1.tar.gz (1.7 MB)
     |████████████████████████████████| 1.7 MB 29.6 MB/s
  Installing build dependencies ... done
  Getting requirements to build wheel ... done
  Installing backend dependencies ... done
    Preparing wheel metadata ... done
Requirement already satisfied: jinja2 in /usr/local/lib/python3.8/dist-packages (from pyLDAvis) (2.11.3)
Requirement already satisfied: numexpr in /usr/local/lib/python3.8/dist-packages (from pyLDAvis) (2.8.4)
Requirement already satisfied: numpy>=1.20.0 in /usr/local/lib/python3.8/dist-packages (from pyLDAvis) (1.21.6)
Requirement already satisfied: scipy in /usr/local/lib/python3.8/dist-packages (from pyLDAvis) (1.7.3)
Collecting funcy
  Downloading funcy-1.17-py2.py3-none-any.whl (33 kB)
Requirement already satisfied: joblib in /usr/local/lib/python3.8/dist-packages (from pyLDAvis) (1.2.0)
Requirement already satisfied: future in /usr/local/lib/python3.8/dist-packages (from pyLDAvis) (0.16.0)
Requirement already satisfied: setuptools in /usr/local/lib/python3.8/dist-packages (from pyLDAvis) (57.4.0)
Collecting sklearn
  Downloading sklearn-0.0.post1.tar.gz (3.6 kB)
```

```
Requirement already satisfied: pandas>=1.2.0 in /usr/local/lib/python3.8/dist-packages (from pyLDAvis) (1.3.5)
Requirement already satisfied: gensim in /usr/local/lib/python3.8/dist-packages (from pyLDAvis) (3.6.0)
Requirement already satisfied: scikit-learn in /usr/local/lib/python3.8/dist-packages (from pyLDAvis) (1.0.2)
Requirement already satisfied: pytz>=2017.3 in /usr/local/lib/python3.8/dist-packages (from pandas>=1.2.0->pyLDAvis) (2022.6)
Requirement already satisfied: python-dateutil>=2.7.3 in /usr/local/lib/python3.8/dist-packages (from pandas>=1.2.0->pyLDAvis) (2.8.2)
Requirement already satisfied: six>=1.5 in /usr/local/lib/python3.8/dist-packages (from python-dateutil>=2.7.3->pandas>=1.2.0->pyLDAvis
Requirement already satisfied: smart-open>=1.2.1 in /usr/local/lib/python3.8/dist-packages (from gensim->pyLDAvis) (5.2.1)
Requirement already satisfied: MarkupSafe>=0.23 in /usr/local/lib/python3.8/dist-packages (from jinja2->pyLDAvis) (2.0.1)
Requirement already satisfied: threadpoolctl>=2.0.0 in /usr/local/lib/python3.8/dist-packages (from scikit-learn->pyLDAvis) (3.1.0)
Building wheels for collected packages: pyLDAvis, sklearn
  Building wheel for pyLDAvis (PEP 517) ... done
  Created wheel for pyLDAvis: filename=pyLDAvis-3.3.1-py2.py3-none-any.whl size=136898 sha256=841c96df90db3154634a056403114a93a6eb30204
  Stored in directory: /root/.cache/pip/wheels/90/61/ec/9dbe9efc3acf9c4e37ba70fbbcc3f3a0ebd121060aa593181a
  Building wheel for sklearn (setup.py) ... done
  Created wheel for sklearn: filename=sklearn-0.0.post1-py3-none-any.whl size=2344 sha256=f32eb0610dc7f42e7790ce71904a6f8509c213ee19ba6
  Stored in directory: /root/.cache/pip/wheels/14/25/f7/1cc0956978ae479e75140219088deb7a36f60459df242b1a72
Successfully built pyLDAvis sklearn
Installing collected packages: sklearn, funcy, pyLDAvis
Successfully installed funcy-1.17 pyLDAvis-3.3.1 sklearn-0.0.post1
```

```
! pip install jsonlines
```

```
Looking in indexes: https://pypi.org/simple, https://us-python.pkg.dev/colab-wheels/public/simple/
Collecting jsonlines
  Downloading jsonlines-3.1.0-py3-none-any.whl (8.6 kB)
Requirement already satisfied: attrs>=19.2.0 in /usr/local/lib/python3.8/dist-packages (from jsonlines) (22.1.0)
Installing collected packages: jsonlines
Successfully installed jsonlines-3.1.0
```

Dataset

Inside the CSV

The first dataset is a collection of 962 Resume Examples Resume : Contains the resume text only in string format. Category : Category of the job the resume was used to apply.

```
data_path = "/content/UpdatedResumeDataSet.csv"
```

```
resume_df0 = pd.read_csv (os.path.join(data_path))
```

```
print (resume_df0.shape)
resume_df0.tail()
```

```
(962, 2)
```

	Category	Resume
957	Testing	Computer Skills: â ¢ Proficient in MS office (...
958	Testing	â Willingness to accept the challenges. â ...
959	Testing	PERSONAL SKILLS â ¢ Quick learner, â ¢ Eagerne...
960	Testing	COMPUTER SKILLS & SOFTWARE KNOWLEDGE MS-Power ...
961	Testing	Skill Set OS Windows XP/7/8/8.1/10 Database MY...

There are total 25 unique Job categories of the resume in dataset 1

```
unique_cat = resume_df0['Category'].unique().tolist()
print ('Number of categories : ', len(unique_cat))
print ('List of categories : \n', unique_cat)
```

```
Number of categories :  25
List of categories :
 ['Data Science', 'HR', 'Advocate', 'Arts', 'Web Designing', 'Mechanical Engineer', 'Sales', 'Health and fitness', 'Civil Engineer', 'J
```

Dataset2

A collection of 2400+ Resume Examples taken from livecareer.com for categorizing a given resume into any of the labels defined in the dataset: Resume Dataset.

Inside the CSV

1) ID: Unique identifier and file name for the respective pdf.

2) Resume_str : Contains the resume text only in string format.

3) Resume_html : Contains the resume data in html format as present while web scrapping.

4) Category : Category of the job the resume was used to apply.

```
data_path1 = "/content/Resume.csv"
resume_df1 = pd.read_csv (os.path.join(data_path1))
```

```
print (resume_df1.shape)
resume_df1.tail()
```

```
(2484, 4)
```

	ID	Resume_str	Resume_html	Category
2479	99416532	RANK: SGT/E-5 NON- COMMISSIONED OFFIC...	<div class="fontsize fontface vmargins hmargin...	AVIATION
2480	24589765	GOVERNMENT RELATIONS, COMMUNICATIONS ...	<div class="fontsize fontface vmargins hmargin...	AVIATION
2481	31605080	GEEK SQUAD AGENT Professional...	<div class="fontsize fontface vmargins hmargin...	AVIATION
2482	21190805	PROGRAM DIRECTOR / OFFICE MANAGER ...	<div class="fontsize fontface vmargins hmargin...	AVIATION
2483	37473139	STOREKEEPER II Professional Sum...	<div class="fontsize fontface vmargins hmargin...	AVIATION

```
resume_df1 ['Resume_str'][0]
```

```
'         HR ADMINISTRATOR/MARKETING ASSOCIATE\n\nHR ADMINISTRATOR       Summary     Dedicated Customer Service Manager with 15+ years
of experience in Hospitality and Customer Service Management.   Respected builder and leader of customer-focused teams; strives to ins
till a shared, enthusiastic commitment to customer service.         Highlights         Focused on customer satisfaction  Team manageme
nt  Marketing savvy  Conflict resolution techniques     Training and development  Skilled multi-tasker  Client relations specialist
Accomplishments      Missouri DOT Supervisor Training Certification  Certified by IHG in Customer Loyalty and Marketing by Segment   H
ilton Worldwide General Manager Training Certification  Accomplished Trainer for cross server hospitality systems such as    Hilton On
Q  ,   Micros    Opera PMS   , Fidelio    OPERA    Reservation System (ORS) ,    Holidex    Completed courses and seminars in customer
service  sales strategies  inventory control  loss pre  '
```

There are total 24 unique Job categories of the resume in dataset 2

```
unique_cat = resume_df1['Category'].unique().tolist()
print ('Number of categories : ', len(unique_cat))
print ('List of categories : \n', unique_cat)
```

```
Number of categories :  24
List of categories :
 ['HR', 'DESIGNER', 'INFORMATION-TECHNOLOGY', 'TEACHER', 'ADVOCATE', 'BUSINESS-DEVELOPMENT', 'HEALTHCARE', 'FITNESS', 'AGRICULTURE', 'B
```

```
# Dropping un-necessary fields in resume df 2
resume_df0.columns = resume_df0.columns.str.replace('Resume', 'Resume_str')
#resume_df0.rename(columns={"Resume": "Resume_str", })
resume_df1 = resume_df1.drop (columns=['Resume_html', 'ID'])
resume_df1.tail()
```

	Resume_str	Category
2479	RANK: SGT/E-5 NON- COMMISSIONED OFFIC...	AVIATION
2480	GOVERNMENT RELATIONS, COMMUNICATIONS ...	AVIATION
2481	GEEK SQUAD AGENT Professional...	AVIATION
2482	PROGRAM DIRECTOR / OFFICE MANAGER ...	AVIATION
2483	STOREKEEPER II Professional Sum...	AVIATION

```
resume_df0.head()
```

	Category	Resume_str
0	Data Science	Skills * Programming Languages: Python (pandas...
1	Data Science	Education Details \r\nMay 2013 to May 2017 B.E...
2	Data Science	Areas of Interest Deep Learning, Control Syste...
3	Data Science	Skills â ¢ R â ¢ Python â ¢ SAP HANA â ¢ Table...
4	Data Science	Education Details \r\n MCA YMCAUST, Faridab...

```
resume_df2 = resume_df0.append(resume_df1, ignore_index=True)
resume_df2.shape
resume_df2.head()
```

	Category	Resume_str
0	Data Science	Skills * Programming Languages: Python (pandas...
1	Data Science	Education Details \r\nMay 2013 to May 2017 B.E...
2	Data Science	Areas of Interest Deep Learning, Control Syste...
3	Data Science	Skills â ¢ R â ¢ Python â ¢ SAP HANA â ¢ Table...
4	Data Science	Education Details \r\n MCA YMCAUST, Faridab...

```
unique_cat = resume_df2['Category'].unique().tolist()
print ('Number of categories : ', len(unique_cat))
print ('List of categories : \n', unique_cat)
```

```
Number of categories :  48
List of categories :
 ['Data Science', 'HR', 'Advocate', 'Arts', 'Web Designing', 'Mechanical Engineer', 'Sales', 'Health and fitness', 'Civil Engineer', 'J
```

Loading spaCy model

The jobzilla skill dataset is jsonl file containing different skills that can be used to create spaCy entity_ruler. The data set contains label and pattern-> diferent words used to descibe skills in various resume.

```
import en_core_web_sm
nlp = en_core_web_sm.load()
```

Entity Ruler

To create an entity ruler we need to add a pipeline and then load the .jsonl file containing skills into ruler. As you can see we have successfully added a new pipeline entity_ruler. Entity ruler helps us add additional rules to highlight various categories within the text, such as skills and job description in our case.

```
skill_pattern_path = "/content/jz_skill_patterns.jsonl#"
#ruler = nlp.add_pipe("entity_ruler")
ruler.from_disk(skill_pattern_path)
nlp.pipe_names
```

```
['tok2vec',
 'tagger',
 'parser',
 'attribute_ruler',
 'lemmatizer',
 'ner',
 'entity_ruler']
```

Skills Parsing

We will create two python functions to extract all the skills within a resume and create an array containing all the skills. Later we are going to apply this function to our dataset and create a new feature called skill. This will help us visualize trends and patterns within the dataset.

1) get_skills is going to extract skills from a single text.

2) unique_skills will remove duplicates.

```
def get_skills(text):
    doc = nlp(text)
    myset = []
    subset = []
    for ent in doc.ents:
        if "SKILL" in ent.label_ :
            #print ("ent.label_ ", ent.label_, "ent.text : ", ent.text)
            subset.append(ent.text)
    myset.append(subset)
    return subset


def unique_skills(x):
    return list(set(x))


def get_education(text):
    doc = nlp(text)
    myset = []
    subset = []
    for ent in doc.ents:
        if "EDUCATION" in ent.label_ :
            #print ("ent.label_ ", ent.label_, "ent.text : ", ent.text)
            subset.append(ent.text)
    myset.append(subset)
    return subset


def unique_education(x):
    return list(set(x))
```

```
nltk.download('omw-1.4')
```

```
[nltk_data] Downloading package omw-1.4 to /root/nltk_data...
True
```

Cleaning Resume Text

We are going to use nltk library to clean our dataset in a few steps:

1) We are going to use regex to remove hyperlinks, special characters, or punctuations.

2) Lowering text

3) Splitting text into array based on space

4) Lemmatizing text to its base form for normalizations

5) Removing English stopwords

6) Appending the results into an array.

```
def clean_a_text (text):
    review = re.sub(
        '(@[A-Za-z0-9]+)|([^0-9A-Za-z \t])|(\w+:\/\/\S+)|^rt|http.+?"',
        " ",
        text,
    )
    review = review.lower()
    review = review.split()
    lm = WordNetLemmatizer()
    review = [
        lm.lemmatize(word)
        for word in review
```

```
        if not word in set(stopwords.words("english"))
    ]
    review = " ".join(review)
    return review

clean = []
for i in range(resume_df2.shape[0]):
    '''
    review = re.sub(
        '(@[A-Za-z0-9]+)|([^0-9A-Za-z \t])|(\w+:\/\/\S+)|^rt|http.+?"',
        " ",
        resume_df2["Resume_str"].iloc[i],
    )
    review = review.lower()
    review = review.split()
    lm = WordNetLemmatizer()
    review = [
        lm.lemmatize(word)
        for word in review
        if not word in set(stopwords.words("english"))
    ]'''
    cleaned_text = clean_a_text (resume_df2["Resume_str"].iloc[i])
    clean.append(cleaned_text)
```

creating Clean_Resume columns and adding cleaning Resume data. creating skills columns, lowering text, and applying the get_skills function. removing duplicates from skills columns. Now we have cleaned the resume and skills columns.

```
resume_df2["Clean_Resume"] = clean
resume_df2["skills"] = resume_df2["Clean_Resume"].str.lower().apply(get_skills)
resume_df2["skills"] = resume_df2["skills"].apply(unique_skills)
resume_df2.head()
```

	Category	Resume_str	Clean_Resume	skills
0	Data Science	Skills * Programming Languages: Python (pandas...	skill programming language python panda numpy ...	[deep learning, tableau, parse, flask, cassand...
1	Data Science	Education Details \r\nMay 2013 to May 2017 B.E...	education detail may 2013 may 2017 b e uit rgp...	[python, github, machine learning, dimensional...
2	Data Science	Areas of Interest Deep Learning, Control Syste...	area interest deep learning control system des...	[deep learning, github, jupyter notebook, flas...
3	Data Science	Skills â ¢ R â ¢ Python â ¢ SAP HANA â ¢ Table...	skill r python sap hana tableau sap hana sql s...	[deep learning, tableau, time series, support,...
4	Data Science	Education Details \r\n MCA YMCAUST, Faridab...	education detail mca ymcaust faridabad haryana...	[python, data structure, java, data science, d...

```
resume_df2["education"] = resume_df2["Clean_Resume"].str.lower().apply(get_education)
```

```
resume_df2.head()
```

	Category	Resume_str	Clean_Resume	skills	education
0	Data Science	Skills * Programming Languages: Python (pandas...	skill programming language python panda numpy ...	[deep learning, tableau, parse, flask, cassand...	[]
1	Data Science	Education Details \r\nMay 2013 to May 2017 B.E...	education detail may 2013 may 2017 b e uit rgp...	[python, github, machine learning, dimensional...	[]
2	Data Science	Areas of Interest Deep Learning, Control Syste...	area interest deep learning control system des...	[deep learning, github, jupyter notebook, flas...	[]
3	Data Science	Skills â ¢ R â ¢ Python â ¢ SAP HANA â ¢ Table...	skill r python sap hana tableau sap hana sql s...	[deep learning, tableau, time series, support,...	[]
4	Data Science	Education Details \r\n MCA YMCAUST, Faridab...	education detail mca ymcaust faridabad haryana...	[python, data structure, java, data science, d...	[]

```
resume_df2["education"] = resume_df2["education"].apply(unique_education)
```

```
resume_df2.head()
```

	Category	Resume_str	Clean_Resume	skills	education
0	Data Science	Skills * Programming Languages: Python (pandas...	skill programming language python panda numpy ...	[deep learning, tableau, parse, flask, cassand...	[]
1	Data Science	Education Details \r\nMay 2013 to May 2017 B.E...	education detail may 2013 may 2017 b e uit rgp...	[python, github, machine learning, dimensional...	[]
2	Data Science	Areas of Interest Deep Learning, Control Syste...	area interest deep learning control system des...	[deep learning, github, jupyter notebook, flas...	[]
3	Data Science	Skills â ¢ R â ¢ Python â ¢ SAP HANA â ¢ Table...	skill r python sap hana tableau sap hana sql s...	[deep learning, tableau, time series, support,...	[]
4	Data Science	Education Details \r\n MCA YMCAUST, Faridab...	education detail mca ymcaust faridabad haryana...	[python, data structure, java, data science, d...	[]

```
resume_df2["Clean_Resume"][0]
```

```
'skill programming language python panda numpy scipy scikit learn matplotlib sql java javascript jquery machine learning regression sv
m na bayes knn random forest decision tree boosting technique cluster analysis word embedding sentiment analysis natural language proc
essing dimensionality reduction topic modelling lda nmf pca neural net database visualization mysql sqlserver cassandra hbase elastics
earch d3 j dc j plotly kibana matplotlib ggplot tableau others regular expression html cs angular 6 logstash kafka python flask git do
cker computer vision open cv understanding deep learning education detail data science assurance associate data science assurance asso
ciate ernst young llp skill detail javascript exprience 24 month jquery exprience 24 month python exprience 24 monthscompany detail co
mpany ernst young llp description fraud investigation dispute service assurance technology assisted review tar technology assisted rev
iew assist accelerating review process run analytics generate r...'
```

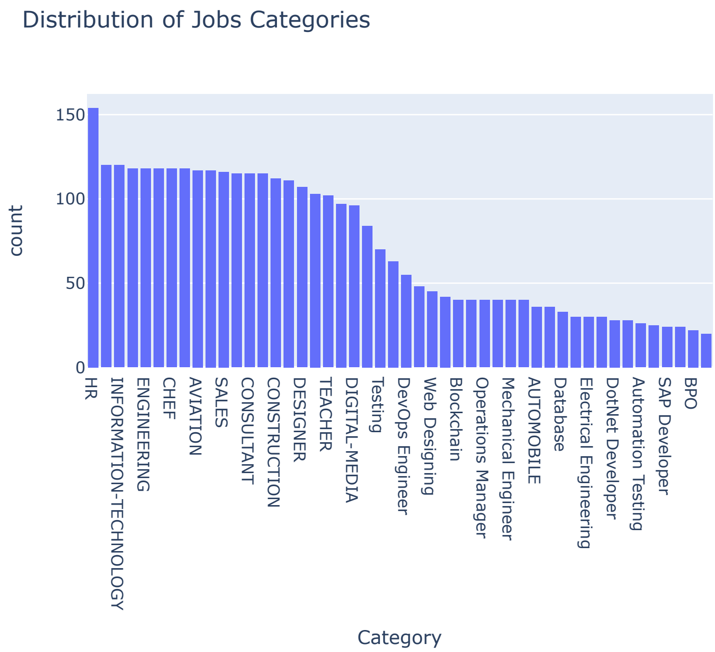
```
resume_df2["skills"][0]
```

```
['deep learning',
 'tableau',
 'parse',
 'flask',
 'cassandra',
 'hbase',
 'monitoring',
 'bot',
 'time series',
 'javascript',
 'analytics',
 'database',
 'visualization',
 'logstash',
 'natural language processing',
 'file format',
 'cluster analysis',
 'decision tree',
 'kafka',
 'sentiment analysis',
 'programming language',
 'git',
 'jquery',
 'machine learning',
 'computer vision',
 'security',
 'predictive coding',
 'plotly',
 'elasticsearch',
 'data science',
 'regular expression',
 'numpy',
 'accounting',
 'python',
 'docker',
 'dimensionality reduction',
 'random forest',
 'bootstrap',
 'mysql',
 'scikit learn']
```

Jobs Distribution

As we can see our samples contain a variety of job categories. HR, Business development, and INFORMATION-TECHNOLOGY are the top categories.

```
fig = px.histogram(
    resume_df2, x="Category", title="Distribution of Jobs Categories"
).update_xaxes(categoryorder="total descending")
fig.show()
```




```
Job_Cat = resume_df2["Category"].unique()
print (Job_Cat)
Job_Cat = np.append(Job_Cat, "ALL")
Job_Category = 'HR'
```

```
['Data Science' 'HR' 'Advocate' 'Arts' 'Web Designing'
 'Mechanical Engineer' 'Sales' 'Health and fitness' 'Civil Engineer'
 'Java Developer' 'Business Analyst' 'SAP Developer' 'Automation Testing'
 'Electrical Engineering' 'Operations Manager' 'Python Developer'
 'DevOps Engineer' 'Network Security Engineer' 'PMO' 'Database' 'Hadoop'
 'ETL Developer' 'DotNet Developer' 'Blockchain' 'Testing' 'DESIGNER'
 'INFORMATION-TECHNOLOGY' 'TEACHER' 'ADVOCATE' 'BUSINESS-DEVELOPMENT'
 'HEALTHCARE' 'FITNESS' 'AGRICULTURE' 'BPO' 'SALES' 'CONSULTANT'
 'DIGITAL-MEDIA' 'AUTOMOBILE' 'CHEF' 'FINANCE' 'APPAREL' 'ENGINEERING'
 'ACCOUNTANT' 'CONSTRUCTION' 'PUBLIC-RELATIONS' 'BANKING' 'ARTS'
 'AVIATION']
```

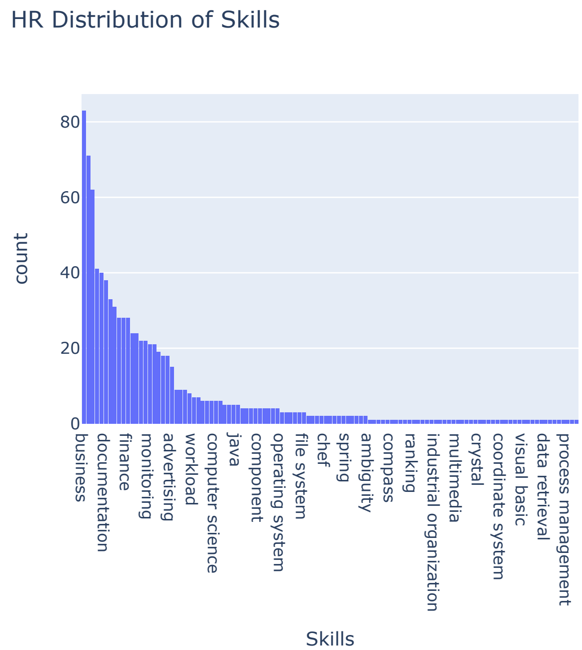
```
Total_skills = []
if Job_Category != "ALL":
    fltr = resume_df2[resume_df2["Category"] == Job_Category]["skills"]
    for x in fltr:
        for i in x:
            Total_skills.append(i)
else:
    fltr = resume_df2["skills"]
    for x in fltr:
        for i in x:
            Total_skills.append(i)
```

As we can observe HR job category skills distributions.

Top Skills are: > Business > Database > Schedule

```
fig = px.histogram(
```

```
fig = px.histogram(
    x=Total_skills,
    labels={"x": "Skills"},
    title=f"{Job_Category} Distribution of Skills",
).update_xaxes(categoryorder="total descending")
fig.show()
```



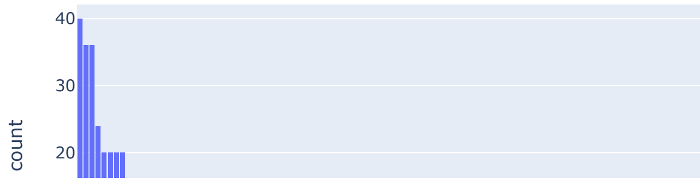
As we can observe Data Science job category skills distributions.

Top Skills are: > Python > Machine Learning > Engineering > Database

```
Job_Category= 'Data Science'
Total_skills = []
if Job_Category != "ALL":
    fltr = resume_df2[resume_df2["Category"] == Job_Category]["skills"]
    for x in fltr:
        for i in x:
            Total_skills.append(i)
else:
    fltr = resume_df2["skills"]
    for x in fltr:
        for i in x:
            Total_skills.append(i)


fig = px.histogram(
    x=Total_skills,
    labels={"x": "Skills"},
    title=f"{Job_Category} Distribution of Skills",
).update_xaxes(categoryorder="total descending")
fig.show()
```

Data Science Distribution of Skills



Most used words

In this part, we are going to display the most used words in the Resume filter by job category. In Information technology, the most words used are system, network, and database. We can also discover more patterns by exploring the word cloud below.

```
text = ""
for i in resume_df2[resume_df2["Category"] == Job_Category]["Clean_Resume"].values:
    text += i + " "

plt.figure(figsize=(8, 8))

x, y = np.ogrid[:300, :300]

mask = (x - 150) ** 2 + (y - 150) ** 2 > 130 ** 2
mask = 255 * mask.astype(int)

wc = WordCloud(
    width=800,
    height=800,
    background_color="white",
    min_font_size=6,
    repeat=True,
    mask=mask,
)
wc.generate(text)

plt.axis("off")
plt.imshow(wc, interpolation="bilinear")
plt.title(f"Most Used Words in {Job_Category} Resume", fontsize=20)
```

```
Text(0.5, 1.0, 'Most Used Words in Data Science Resume')
```

Most Used Words in Data Science Resume

Entity Recognition

We can also display various entities within our raw text by using spaCy displacy.render. I am in love with this function as it is an amazing way to look at your entire document and discover SKILL or GEP within your Resume.

```
sent = nlp(resume_df2["Resume_str"].iloc[0])
displacy.render(sent, style="ent", jupyter=True)
```

Skills * Programming Languages: Python ORG (pandas SKILL|pandas , numpy SKILL|numpy , scipy, scikit-learn, matplotlib), Sql GPE , Java PERSON , JavaScript ORG / JQuery SKILL|jquery . * Machine learning SKILL|machine-learning : Regression, SVM ORG , NaÃ¯ve Bayes ORG , KNN ORG , Random Forest ORG , Decision Trees ORG , Boosting techniques, Cluster Analysis ORG , Word Embedding PERSON , Sentiment Analysis SKILL|sentiment-analysis , Natural Language ORG processing, Dimensionality reduction SKILL|dimensionality-reduction , Topic Modelling (LDA, NMF ORG), PCA & Neural Nets ORG . * Database SKILL|database Visualizations: Mysql SKILL|mysql , SqlServer GPE , Cassandra GPE , Hbase ORG , ElasticSearch ORG D3.js SKILL|d3.js , DC.js, Plotly SKILL|plotly , kibana PERSON , matplotlib, ggplot, Tableau GPE . * Others: Regular Expression SKILL|regular-expression , HTML ORG , CSS ORG , Angular SKILL|angular 6 CARDINAL , Logstash ORG , Kafka PERSON , Python Flask GPE , Git, Docker ORG , computer vision SKILL|computer-vision - Open CV PERSON and understanding of Deep PERSON learning.Education Details

Data Science SKILL|data-science Assurance Associate

Data Science SKILL|data-science Assurance Associate - Ernst & Young LLP

Skill Details

JAVASCRIPT- Exprience PERSON - 24 months DATE

jQuery- Exprience PERSON - 24 months DATE

Python- Exprience - 24 monthsCompany Details

company - Ernst & Young LLP ORG

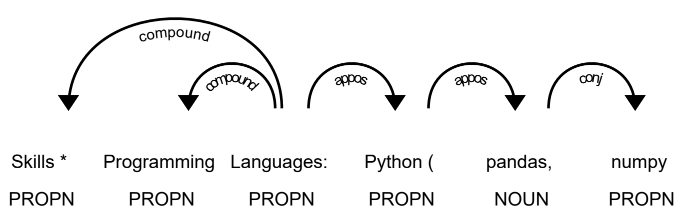
description - Fraud Investigations and Dispute Services ORG Assurance

Dependency Parsing

We can also visualize dependencies by just changing style to dep as shown below. We have also limited words to 10 which includes space too. Limiting the words will make it visualize the small chunk of data and if you want to see the dependency, you can remove the filter.

Core member of a team helped in developing automated review platform tool from scratch for assisting E

```
displacy.render(sent[0:10], style="dep", jupyter=True, options={"distance": 90})
```



Custom Entity Recognition

We have added a new entity called SKILL and is displayed in gray color. I was not impressed by colors and I also wanted to add another entity called Job Description so I started experimenting with various parameters within displace.

Adding Job-Category into entity ruler. Adding custom colors to all categories. Adding gradient colors to SKILL and Job-Category You can see the result below as the new highlighted texts look beautiful.

```
patterns = resume_df2.Category.unique()
for a in patterns:
    ruler.add_patterns([{"label": "Job-Category", "pattern": a}])
```

feedback survey data for past one year DATE . Performed sentiment (Positive, Negative & Neutral

```
# options=[{"ents": "Job-Category", "colors": "#ff3232"},{"ents": "SKILL", "colors": "#56c426"}]
colors = {
    "Job-Category": "linear-gradient(90deg, #aa9cfc, #fc9ce7)",
    "SKILL": "linear-gradient(90deg, #9BE15D, #00E3AE)",
    "ORG": "#ffd966",
    "PERSON": "#e06666",
    "GPE": "#9fc5e8",
    "DATE": "#c27ba0",
    "ORDINAL": "#674ea7",
    "PRODUCT": "#f9cb9c",
}
options = {
    "ents": [
        "Job-Category",
        "SKILL",
        "ORG",
        "PERSON",
        "GPE",
        "DATE",
        "ORDINAL",
        "PRODUCT",
    ],
    "colors": colors,
}
sent = nlp(resume_df2["Resume_str"].iloc[5])
displacy.render(sent, style="ent", jupyter=True, options=options)
```

SKILLS C Basics ORG , IOT ORG , Python GPE , MATLAB ORG , Data Science ORG , Machine Learning, HTML ORG , Microsoft Word ORG , Microsoft Excel ORG , Microsoft Powerpoint ORG . RECOGNITION Academic Secured First ORG place in B.Tech.Education Details

August 2014 DATE to May 2018 DATE B.Tech. Ghatkesar, Andhra Pradesh Aurora's Scientific and Technological Institute ORG

June 2012 DATE to May 2014 DATE Secondary Education Warangal, Telangana SR Junior College Data Science ORG

Skill Details

MS OFFICE- Exprience - Less than 1 year months

C- Exprience - Less than 1 year months

machine learning- Exprience PERSON - Less than 1 year months

data science- Exprience PERSON - Less than 1 year months

Resume Anlaysis

In this part, I am allowing users to copy&paste their resumes and see the results.

As we can see my I have added my Resume and the results are amazing. The model has successfully highlighted all the skills.

```
input_resume = "Abid Ali Awan Data Scientist I am a certified data scientist professional, who loves building machine learning models and blo
```

Custom Entity Recognition

In our case, we have added a new entity called SKILL and is displayed in Green color. I was not impressed by colors and I also wanted to add another entity called Job Description so I started experimenting with various parameters within displace.

Adding Job-Category into entity ruler. Adding custom colors to all categories. Adding gradient colors to SKILL and Job-Category You can see the result below as the new highlighted texts look beautiful.

```
sent2 = nlp(input_resume)
displacy.render(sent2, style="ent", jupyter=True, options=options)
```

Abid Ali Awan PERSON Data Scientist I am a certified data scientist professional, who loves building machine learning models and blogs about the latest AI ORG technologies. I am currently testing AI Products ORG at PEC-PITC ORG , which later gets approved for human trials. abidaliawan@tutamail.com +923456855126 Islamabad GPE , Pakistan GPE abidaliawan.me WORK EXPERIENCE Data Scientist Pakistan Innovation ORG and Testing Job-Category Center - PEC 04/2021 - Present, Islamabad GPE , Pakistan GPE Redesigned data of engineers that were mostly scattered and unavailable. Designed dashboard and data analysis report to help higher management make better decisions. Accessibility of key information has created a new culture of making data-driven decisions. Contact: Ali PERSON Raza Asif - darkslayerraza10@gmail.com Data Scientist Freelancing/Kaggle 11/2020 - Present, Islamabad GPE , Pakistan GPE Engineered a healthcare system. Used machine learning to detect some of the common decisions. The project has paved the way for others to use new techniques to get better results. Participated in Kaggle GPE machine learning competitions. Learned new techniques to get a better score and finally got to 1 percent rank. Researcher / Event Organizer CREDIT ORG 02/2017 - 07/2017, Kuala Lumpur GPE , Malaysia Marketing ORG for newly build research lab. Organized technical events and successfully invited the multiple company's CEO for talks. Reduced the gap between industries and educational institutes. Research on new development in the IoT ORG sector. Created research proposal for funding. Investigated the new communication protocol for IoT ORG devices. Contact: Dr. Tan Chye PERSON Cheah - dr.chyecheah.t@apu.edu.my EDUCATION MSc in Technology Management Staffordshire University ORG 11/2015 - 04/2017, Postgraduate with Distinction Challenges ORG in Implementing IoT-enabled Smart ORG cities in Malaysia GPE . Bachelors Electrical Telecommunication Engineering COMSATS Institute of Information Technology ORG , Islamabad 08/2010 - 01/2014 ORG , CGPA ORG : 3.09 Networking Satellite communications Programming/ Matlab Telecommunication Engineering SKILLS Designing Leadership Media/Marketing R/Python SQL Tableau ORG NLP Data Analysis Machine learning Deep learning Webapp/Cloud Feature Engineering Ensembling Time Series

```
input_skills = 'Data Science,Data Analysis,Database,SQL,Machine learning'
```

```
req_skills = input_skills.lower().split(",")
resume_skills = unique_skills(get_skills(input_resume.lower()))
score = 0
for x in req_skills:
    if x in resume_skills:
        score += 1
req_skills_len = len(req_skills)
match = round(score / req_skills_len * 100, 1)

print(f"The current Resume is {match}% matched to your requirements")
```

```
The current Resume is 60.0% matched to your requirements
```

```
print(resume_skills)
```

```
['deep learning', 'tableau', 'pytorch', 'time series', 'database', 'visualization', 'data analysis', 'communications', 'nlp', 'text pro
```

```
'''
# importing required modules
import PyPDF2

# creating a pdf file object
pdfFileObj = open('dataset/data/HR/10399912.pdf', 'rb')

# creating a pdf reader object
pdfReader = PyPDF2.PdfFileReader(pdfFileObj)

# printing number of pages in pdf file
print(pdfReader.numPages)

# creating a page object
pageObj = pdfReader.getPage(0)

# extracting text from page
print("CONTENT : " , pageObj.extractText())

# closing the pdf file object
pdfFileObj.close()
'''
```

```
'\n# importing required modules \nimport PyPDF2 \n    \n# creating a pdf file object
\npdfFileObj = open(\'dataset/data/HR/10399912.pdf\', \'rb\') \n    \n# creating a pdf
reader object \npdfReader = PyPDF2.PdfFileReader(pdfFileObj) \n    \n# printing number
of pages in pdf file \nprint(pdfReader.numPages) \n    \n# creating a page object \npa
geObj = pdfReader.getPage(0) \n    \n# extracting text from page \nprint("CONTENT : "
```

```
import slate3k as slate

#with open("dataset/data/HR/10399912.pdf",'rb') as f:
#    extracted_text = slate.PDF(f)
#print(extracted_text)
```

```
! pip install slate3k
```

```
Looking in indexes: https://pypi.org/simple, https://us-python.pkg.dev/colab-wheels/public/simple/
Collecting slate3k
  Downloading slate3k-0.5.3-py2.py3-none-any.whl (7.9 kB)
Collecting pdfminer3k
  Downloading pdfminer3k-1.3.4-py3-none-any.whl (100 kB)
     |████████████████████████████████| 100 kB 8.0 MB/s
Collecting ply
  Downloading ply-3.11-py2.py3-none-any.whl (49 kB)
     |████████████████████████████████| 49 kB 6.1 MB/s
Installing collected packages: ply, pdfminer3k, slate3k
Successfully installed pdfminer3k-1.3.4 ply-3.11 slate3k-0.5.3
```

```
def list_full_paths(directory):
    return [os.path.join(directory, file) for file in os.listdir(directory)]
```

```
!unzip /content/Data1.zip
```

```
    inflating: Data1/ENGINEERING/36517781.pdf
    inflating: Data1/ENGINEERING/37335325.pdf
    inflating: Data1/ENGINEERING/38220146.pdf
    inflating: Data1/ENGINEERING/38314236.pdf
    inflating: Data1/ENGINEERING/38535335.pdf
    inflating: Data1/ENGINEERING/39835894.pdf
    inflating: Data1/ENGINEERING/39855211.pdf
    inflating: Data1/ENGINEERING/43752620.pdf
    inflating: Data1/ENGINEERING/44624796.pdf
    inflating: Data1/ENGINEERING/47276718.pdf
    inflating: Data1/ENGINEERING/47549345.pdf
    inflating: Data1/ENGINEERING/47919212.pdf
    inflating: Data1/ENGINEERING/49127329.pdf
    inflating: Data1/ENGINEERING/50328713.pdf
    inflating: Data1/ENGINEERING/51588273.pdf
    inflating: Data1/ENGINEERING/54227873.pdf
    inflating: Data1/ENGINEERING/55595908.pdf
    inflating: Data1/ENGINEERING/55953734.pdf
    inflating: Data1/ENGINEERING/56691064.pdf
    inflating: Data1/ENGINEERING/60004873.pdf
    inflating: Data1/ENGINEERING/61579998.pdf
    inflating: Data1/ENGINEERING/62071407.pdf
    inflating: Data1/ENGINEERING/64468610.pdf
    inflating: Data1/ENGINEERING/64755882.pdf
    inflating: Data1/ENGINEERING/74236636.pdf
    inflating: Data1/ENGINEERING/77828437.pdf
    inflating: Data1/ENGINEERING/81125166.pdf
    inflating: Data1/ENGINEERING/82125182.pdf
    inflating: Data1/ENGINEERING/82246962.pdf
    inflating: Data1/ENGINEERING/86209934.pdf
    inflating: Data1/ENGINEERING/86828820.pdf
    inflating: Data1/ENGINEERING/90280583.pdf
    inflating: Data1/ENGINEERING/96029688.pdf
```

```
top_n = 3
job_category =  'ENGINEERING'
#job_category = 'HR'
test_job_folder = '/content/Data1'
test_resume_path = os.path.join (test_job_folder, job_category)
required_education = 'master, engineering, computer science, graduate, post graduate'
required_skills = 'Data Science,Data Analysis,Database,\
SQL,Machine learning,Python,tableau'
#required_skills = 'business, database, schedule'
resume_files_list = list_full_paths (test_resume_path)
print ("Total num of test resumes : ", len (resume_files_list))
```

```
Total num of test resumes :  118
```

```
resume_texts = []
match_score = []
skill_match_score = []
all_resume_skills = []
edu_match_score = []
all_resume_edu = []
for resume_file in resume_files_list:
    f = open(resume_file,'rb')
    extracted_text = slate.PDF(f)
    extracted_text = str(extracted_text)
    cleaned_text = clean_a_text (extracted_text)
    resume_texts.append (cleaned_text)
    req_skills = required_skills.lower().split(",")
    resume_skills = unique_skills(get_skills(cleaned_text.lower()))
    all_resume_skills.append (resume_skills)
    req_edu = required_education.lower().split(",")
    resume_edu = unique_skills(get_education(cleaned_text.lower()))
    all_resume_edu.append (resume_skills)

    #print (resume_skills)
    score = 0
    for x in req_skills:
        if x in resume_skills:
            score += 1
    edu_score = 0
    for x in req_edu:
        if x in resume_edu:
            edu_score += 1

    req_skills_len = len(req_skills)
    match = round(score / req_skills_len * 100, 1)
    skill_match_score.append (match)
```

```
        skill_match_score.append (match)

        req_edu_len = len(req_edu)
        edu_match = round(edu_score / req_edu_len * 100, 1)
        edu_match_score.append (edu_match)

        match_score.append (match+edu_match)
        print(f"The current-resume {resume_file} is {match}% matched to required skills and {edu_match}% with education")
```

```
The current-resume /content/Data1/ENGINEERING/32985311.pdf is 28.6% matched to required skills and 0.0% with education
The current-resume /content/Data1/ENGINEERING/17926546.pdf is 0.0% matched to required skills and 0.0% with education
The current-resume /content/Data1/ENGINEERING/11890896.pdf is 14.3% matched to required skills and 0.0% with education
The current-resume /content/Data1/ENGINEERING/47276718.pdf is 0.0% matched to required skills and 0.0% with education
The current-resume /content/Data1/ENGINEERING/15941675.pdf is 14.3% matched to required skills and 0.0% with education
The current-resume /content/Data1/ENGINEERING/10219099.pdf is 0.0% matched to required skills and 0.0% with education
The current-resume /content/Data1/ENGINEERING/31677347.pdf is 0.0% matched to required skills and 0.0% with education
The current-resume /content/Data1/ENGINEERING/21847415.pdf is 0.0% matched to required skills and 0.0% with education
The current-resume /content/Data1/ENGINEERING/15601399.pdf is 0.0% matched to required skills and 0.0% with education
The current-resume /content/Data1/ENGINEERING/19612167.pdf is 0.0% matched to required skills and 0.0% with education
The current-resume /content/Data1/ENGINEERING/90280583.pdf is 14.3% matched to required skills and 0.0% with education
The current-resume /content/Data1/ENGINEERING/30542184.pdf is 0.0% matched to required skills and 0.0% with education
The current-resume /content/Data1/ENGINEERING/38535335.pdf is 14.3% matched to required skills and 0.0% with education
The current-resume /content/Data1/ENGINEERING/28762662.pdf is 0.0% matched to required skills and 0.0% with education
The current-resume /content/Data1/ENGINEERING/20882041.pdf is 0.0% matched to required skills and 0.0% with education
The current-resume /content/Data1/ENGINEERING/64468610.pdf is 0.0% matched to required skills and 0.0% with education
The current-resume /content/Data1/ENGINEERING/81125166.pdf is 0.0% matched to required skills and 0.0% with education
The current-resume /content/Data1/ENGINEERING/16803215.pdf is 0.0% matched to required skills and 0.0% with education
The current-resume /content/Data1/ENGINEERING/23438227.pdf is 14.3% matched to required skills and 0.0% with education
The current-resume /content/Data1/ENGINEERING/17488801.pdf is 0.0% matched to required skills and 0.0% with education
The current-resume /content/Data1/ENGINEERING/21038022.pdf is 0.0% matched to required skills and 0.0% with education
The current-resume /content/Data1/ENGINEERING/61579998.pdf is 14.3% matched to required skills and 0.0% with education
The current-resume /content/Data1/ENGINEERING/19396040.pdf is 42.9% matched to required skills and 0.0% with education
The current-resume /content/Data1/ENGINEERING/21298336.pdf is 0.0% matched to required skills and 0.0% with education
The current-resume /content/Data1/ENGINEERING/11981094.pdf is 0.0% matched to required skills and 0.0% with education
The current-resume /content/Data1/ENGINEERING/64755882.pdf is 0.0% matched to required skills and 0.0% with education
The current-resume /content/Data1/ENGINEERING/12748557.pdf is 0.0% matched to required skills and 0.0% with education
The current-resume /content/Data1/ENGINEERING/31694970.pdf is 14.3% matched to required skills and 0.0% with education
The current-resume /content/Data1/ENGINEERING/38220146.pdf is 0.0% matched to required skills and 0.0% with education
The current-resume /content/Data1/ENGINEERING/44624796.pdf is 0.0% matched to required skills and 0.0% with education
The current-resume /content/Data1/ENGINEERING/28628090.pdf is 0.0% matched to required skills and 0.0% with education
The current-resume /content/Data1/ENGINEERING/12518008.pdf is 0.0% matched to required skills and 0.0% with education
The current-resume /content/Data1/ENGINEERING/14554542.pdf is 0.0% matched to required skills and 0.0% with education
The current-resume /content/Data1/ENGINEERING/12022566.pdf is 0.0% matched to required skills and 0.0% with education
The current-resume /content/Data1/ENGINEERING/35651876.pdf is 0.0% matched to required skills and 0.0% with education
The current-resume /content/Data1/ENGINEERING/50328713.pdf is 28.6% matched to required skills and 0.0% with education
The current-resume /content/Data1/ENGINEERING/14049846.pdf is 0.0% matched to required skills and 0.0% with education
The current-resume /content/Data1/ENGINEERING/77828437.pdf is 0.0% matched to required skills and 0.0% with education
The current-resume /content/Data1/ENGINEERING/82125182.pdf is 0.0% matched to required skills and 0.0% with education
The current-resume /content/Data1/ENGINEERING/14206561.pdf is 14.3% matched to required skills and 0.0% with education
The current-resume /content/Data1/ENGINEERING/28320387.pdf is 0.0% matched to required skills and 0.0% with education
The current-resume /content/Data1/ENGINEERING/25930778.pdf is 0.0% matched to required skills and 0.0% with education
The current-resume /content/Data1/ENGINEERING/33685075.pdf is 14.3% matched to required skills and 0.0% with education
The current-resume /content/Data1/ENGINEERING/18753367.pdf is 14.3% matched to required skills and 0.0% with education
The current-resume /content/Data1/ENGINEERING/24322804.pdf is 0.0% matched to required skills and 0.0% with education
The current-resume /content/Data1/ENGINEERING/10624813.pdf is 28.6% matched to required skills and 0.0% with education
The current-resume /content/Data1/ENGINEERING/25797445.pdf is 0.0% matched to required skills and 0.0% with education
The current-resume /content/Data1/ENGINEERING/27756469.pdf is 0.0% matched to required skills and 0.0% with education
The current-resume /content/Data1/ENGINEERING/32802563.pdf is 14.3% matched to required skills and 0.0% with education
The current-resume /content/Data1/ENGINEERING/21629057.pdf is 0.0% matched to required skills and 0.0% with education
The current-resume /content/Data1/ENGINEERING/47919212.pdf is 14.3% matched to required skills and 0.0% with education
The current-resume /content/Data1/ENGINEERING/32081266.pdf is 0.0% matched to required skills and 0.0% with education
The current-resume /content/Data1/ENGINEERING/28005884.pdf is 0.0% matched to required skills and 0.0% with education
The current-resume /content/Data1/ENGINEERING/25425322.pdf is 0.0% matched to required skills and 0.0% with education
The current-resume /content/Data1/ENGINEERING/54227873.pdf is 14.3% matched to required skills and 0.0% with education
The current-resume /content/Data1/ENGINEERING/19124258.pdf is 0.0% matched to required skills and 0.0% with education
The current-resume /content/Data1/ENGINEERING/28631840.pdf is 0.0% matched to required skills and 0.0% with education
The current-resume /content/Data1/ENGINEERING/10985403.pdf is 14.3% matched to required skills and 0.0% with education
```

```
    match_score = np.array(match_score)
    sort_index_match_score = np.argsort(match_score)
    #print(sort_index_match_score)

    if top_n < len (resume_files_list):
        top_n_index = list(sort_index_match_score [-top_n : ])
        top_n_index.reverse()
        for i, idx in enumerate(top_n_index):
            print ("\n\nTop %d resume is %s with match score : %f "%(i+1, resume_files_list[idx], match_score[idx] ))
            print ("\n\tSkillset of this resume is : \n\t\t", all_resume_skills[idx])
    else:
        for i, idx in enumerate(sort_index_match_score):
            print ("\n\nTop %d resume is %s with match score : %f "%(i+1, resume_files_list[idx], match_score[idx] ))
            print ("\n\tSkillset of this resume is : ", all_resume_skills[idx])
```

```
    print ("\n\tSkillset of this resume is : ", all_resume_skills[idx])
```

```
Top 1 resume is /content/Data1/ENGINEERING/12011623.pdf with match score : 71.400000

        Skillset of this resume is :
                 ['tableau', 'data mining', 'support', 'analytics', 'material', 'algorithm', 'logistic regression', 'business', 'visual


Top 2 resume is /content/Data1/ENGINEERING/19396040.pdf with match score : 42.900000

        Skillset of this resume is :
                 ['python', 'documentation', 'robot', 'engineering', 'industrial engineering', 'java', 'image quality', 'schedule', 'vi


Top 3 resume is /content/Data1/ENGINEERING/32985311.pdf with match score : 28.600000

        Skillset of this resume is :
                 ['tableau', 'python', 'testing', 'trello', 'software', 'engineering', 'business process', 'certificate', 'interaction'
```

```
sent2 = nlp(resume_texts[top_n_index[0]])
displacy.render(sent2, style="ent", jupyter=True, options=options)
```

engineering quality technician ncareer overview na highly experienced skilled graduate analytics degree

good experience sa web scraping sql predictive modelling ndata visualization excellent ability identifying

data requirement analysis data cleaning munging model building ensures organization nu effectively reach

profit growth objective comfortable data handling modeling coding appreciation nmakes sense business

standpoint six year DATE experience working researcher data analyst environmental science

ntechnology instructor experience sql data warehousing maintaining securing stabilizing data layer testing

```
docs = resume_df2["Clean_Resume"].values
dictionary = corpora.Dictionary(d.split() for d in docs)
bow = [dictionary.doc2bow(d.split()) for d in docs]
lda = gensim.models.ldamodel.LdaModel
num_topics = 4
ldamodel = lda(
    bow,
    num_topics=num_topics,
    id2word=dictionary,
    passes=50,
    minimum_probability=0
)
ldamodel.print_topics(num_topics=num_topics)
```

```
[(0,
  '0.015*"project" + 0.012*"system" + 0.008*"management" + 0.007*"state" + 0.007*"company" + 0.007*"city" + 0.007*"engineering" +
0.006*"construction" + 0.006*"design" + 0.006*"name"'),
 (1,
  '0.015*"customer" + 0.012*"company" + 0.011*"state" + 0.009*"city" + 0.009*"management" + 0.009*"service" + 0.008*"name" +
0.007*"sale" + 0.007*"financial" + 0.006*"account"'),
 (2,
  '0.013*"state" + 0.011*"city" + 0.011*"company" + 0.008*"name" + 0.008*"marketing" + 0.008*"sale" + 0.006*"business" +
0.006*"management" + 0.006*"development" + 0.006*"student"'),
 (3,
  '0.013*"project" + 0.012*"exprience" + 0.011*"month" + 0.011*"data" + 0.010*"description" + 0.010*"detail" + 0.009*"company" +
0.008*"application" + 0.008*"system" + 0.008*"database"')]
```

nprojects prepared statement monitored project schedule nidentified product defect introduced data

understand impact due defect provided valuable information nproduct shipping customer satisfaction

nmanaged multiple task accomplished goal efficiently per schedule strong work performance meet goal

ndepartment nmonitored adjusted semiconductor production process equipment improving quality

productivity achieved 10 nhigher performance rate fiscal year 2014 DATE nprovided technical support

developing building testing prototype new product process procedure provided training nand advice

engineering technician napplied database management data analysis method helped enhancing production

efficiency reduced cost ndepartment 5 every quarter n n01 2007 DATE 01 2012 DATE ncompany

name n nlecturer environmental science technology ORG effectively engaged course curriculum

✓ 4m 26s completed at 18:04