```python
!pip install docx2txt
import docx2txt
```

```
Collecting docx2txt
  Downloading docx2txt-0.8.tar.gz (2.8 kB)
  Preparing metadata (setup.py) ... done
Building wheels for collected packages: docx2txt
  Building wheel for docx2txt (setup.py) ... done
  Created wheel for docx2txt: filename=docx2txt-0.8-py3-none-any.whl size=3
  Stored in directory: /root/.cache/pip/wheels/22/58/cf/093d0a6c3ecfdfc5f6d
Successfully built docx2txt
Installing collected packages: docx2txt
Successfully installed docx2txt-0.8
```

```python
job_description = docx2txt.process('/content/artificial-intelligence-v1.docx')
resume = docx2txt.process('/content/Resume_suvro2.docx')
```

```python
print(resume)
```

```
Final Year Project: Decentralised Machine Learning Operations - Digital Cop

2. M.Sc. in Pure Mathematics, Specialization in Algebraic aspects of Crypto

University of Calcutta, Ballygunge Science College  | CGPA: 4.314/6.00

Dissertation & Thesis Works: "Wannacry Ransomware Cyberattack" and "Non-ori


EXPERIENCE

1. Deep Learning-Computer Vision, NLP and ML Internship at Mentorness, IT S

Developed CNN model for food item detection with 92% accuracy using ImageNe

Implemented FaceNet architecture for face detection with >90% accuracy, enh

Analysed customer feedback for banking systems, improving service quality b


2. Data Science Internship at Veridical Technologies, Digital Marketing and

Leveraged Credit Card Fraud Detection system development, achieving 94% acc

Implemented deep neural network model with 90% accuracy in classifying emot

Engineered Road Lane Line Detection system with 85% accurate lane line iden
```

3. Machine Learning Research Internship at Feynn Labs Services, Business In

Enhanced business analytics and market research, achieving 20% boost in dat

Orchestrated market segmentation, resulting in 18% categorization precision

Contributed to 30% increase in data-driven decision-making potency and 25%

PROJECTS

Leveraged sentiment analysis on english news paper report, achieving >80% a

Conducted graph neural network implementation for edge connectivity represe

Implemented real-time machine translation module applying LSTM and transfor

Predicted gold stock prices with finely tuned RNN and bollinger bands tradi

Detected vehicle fastag fraud with Random Forest, Logistic Regression, and

Spearheaded market segmentation for McDonald's case study and Indian EV mar

Developed recommender engine for small-scale tourism industry, increasing c

```python
content = [job_description, resume]
```

```python
from sklearn.feature_extraction.text import CountVectorizer
cv = CountVectorizer()
matrix = cv.fit_transform(content)
from sklearn.metrics.pairwise import cosine_similarity
similarity_matrix = cosine_similarity(matrix)
print(similarity_matrix)
```

```
[[1.        0.5378286]
 [0.5378286 1.        ]]
```

```python
print('Resume matches by : '+ str(round(similarity_matrix[1][0]* 100))+ '%')
```

```
Resume matches by : 54%
```

```python
CV = docx2txt.process('/content/CV_suvro.docx')
```

```python
content2 = [job_description, CV]
from sklearn.feature_extraction.text import CountVectorizer
cv = CountVectorizer()
```

```
cv = CountVectorizer()
matrix = cv.fit_transform(content2)
from sklearn.metrics.pairwise import cosine_similarity
similarity_matrix = cosine_similarity(matrix)
print(similarity_matrix)
print('CV matches by : '+ str(round(similarity_matrix[1][0]* 100))+ '%')
```

```
[[1.         0.61855003]
 [0.61855003 1.        ]]
CV matches by : 62%
```

```
!pip install -U spacy
!pip install spacy_transformers
```

```
Requirement already satisfied: tqdm<5.0.0,>=4.38.0 in /usr/local/lib/python
Requirement already satisfied: requests<3.0.0,>=2.13.0 in /usr/local/lib/py
Requirement already satisfied: pydantic!=1.8,!=1.8.1,<3.0.0,>=1.7.4 in /usr
Requirement already satisfied: jinja2 in /usr/local/lib/python3.10/dist-pac
Requirement already satisfied: setuptools in /usr/local/lib/python3.10/dist
Requirement already satisfied: packaging>=20.0 in /usr/local/lib/python3.10
Requirement already satisfied: langcodes<4.0.0,>=3.2.0 in /usr/local/lib/py
Requirement already satisfied: filelock in /usr/local/lib/python3.10/dist-p
Requirement already satisfied: typing-extensions>=4.8.0 in /usr/local/lib/p
Requirement already satisfied: sympy in /usr/local/lib/python3.10/dist-pack
Requirement already satisfied: networkx in /usr/local/lib/python3.10/dist-p
Requirement already satisfied: fsspec in /usr/local/lib/python3.10/dist-pac
Collecting nvidia-cuda-nvrtc-cu12==12.1.105 (from torch>=1.8.0->spacy_trans
  Using cached nvidia_cuda_nvrtc_cu12-12.1.105-py3-none-manylinux1_x86_64.w
Collecting nvidia-cuda-runtime-cu12==12.1.105 (from torch>=1.8.0->spacy_tra
  Using cached nvidia_cuda_runtime_cu12-12.1.105-py3-none-manylinux1_x86_64
Collecting nvidia-cuda-cupti-cu12==12.1.105 (from torch>=1.8.0->spacy_trans
  Using cached nvidia_cuda_cupti_cu12-12.1.105-py3-none-manylinux1_x86_64.w
Collecting nvidia-cudnn-cu12==8.9.2.26 (from torch>=1.8.0->spacy_transforme
  Using cached nvidia_cudnn_cu12-8.9.2.26-py3-none-manylinux1_x86_64.whl (7
Collecting nvidia-cublas-cu12==12.1.3.1 (from torch>=1.8.0->spacy_transform
  Using cached nvidia_cublas_cu12-12.1.3.1-py3-none-manylinux1_x86_64.whl (
Collecting nvidia-cufft-cu12==11.0.2.54 (from torch>=1.8.0->spacy_transform
  Using cached nvidia_cufft_cu12-11.0.2.54-py3-none-manylinux1_x86_64.whl (
Collecting nvidia-curand-cu12==10.3.2.106 (from torch>=1.8.0->spacy_transfo
  Using cached nvidia_curand_cu12-10.3.2.106-py3-none-manylinux1_x86_64.whl
Collecting nvidia-cusolver-cu12==11.4.5.107 (from torch>=1.8.0->spacy_trans
  Using cached nvidia_cusolver_cu12-11.4.5.107-py3-none-manylinux1_x86_64.w
Collecting nvidia-cusparse-cu12==12.1.0.106 (from torch>=1.8.0->spacy_trans
  Using cached nvidia_cusparse_cu12-12.1.0.106-py3-none-manylinux1_x86_64.w
Collecting nvidia-nccl-cu12==2.19.3 (from torch>=1.8.0->spacy_transformers)
  Using cached nvidia_nccl_cu12-2.19.3-py3-none-manylinux1_x86_64.whl (166.
Collecting nvidia-nvtx-cu12==12.1.105 (from torch>=1.8.0->spacy_transformer
  Using cached nvidia_nvtx_cu12-12.1.105-py3-none-manylinux1_x86_64.whl (99
Requirement already satisfied: triton==2.2.0 in /usr/local/lib/python3.10/d
Collecting nvidia-nvjitlink-cu12 (from nvidia-cusolver-cu12==11.4.5.107->to
  Using cached nvidia_nvjitlink_cu12-12.4.127-py3-none-manylinux2014_x86_64
Requirement already satisfied: huggingface-hub<1.0,>=0.19.3 in /usr/local/l
Requirement already satisfied: pyyaml>=5.1 in /usr/local/lib/python3.10/dis
Requirement already satisfied: regex!=2019.12.17 in /usr/local/lib/python3.
Requirement already satisfied: tokenizers<0.19,>=0.14 in /usr/local/lib/pyt
```

```
Requirement already satisfied: safetensors>=0.3.1 in /usr/local/lib/python3
Requirement already satisfied: annotated-types>=0.4.0 in /usr/local/lib/pyt
Requirement already satisfied: pydantic-core==2.16.3 in /usr/local/lib/pyth
Requirement already satisfied: charset-normalizer<4,>=2 in /usr/local/lib/p
Requirement already satisfied: idna<4,>=2.5 in /usr/local/lib/python3.10/di
Requirement already satisfied: urllib3<3,>=1.21.1 in /usr/local/lib/python3
Requirement already satisfied: certifi>=2017.4.17 in /usr/local/lib/python3
Requirement already satisfied: blis<0.8.0,>=0.7.8 in /usr/local/lib/python3
Requirement already satisfied: confection<1.0.0,>=0.0.1 in /usr/local/lib/p
Requirement already satisfied: click<9.0.0,>=7.1.1 in /usr/local/lib/python
Requirement already satisfied: cloudpathlib<0.17.0,>=0.7.0 in /usr/local/li
Requirement already satisfied: MarkupSafe>=2.0 in /usr/local/lib/python3.10
Requirement already satisfied: mpmath>=0.19 in /usr/local/lib/python3.10/di
Installing collected packages: spacy-alignments, nvidia-nvtx-cu12, nvidia-n
  Attempting uninstall: transformers
    Found existing installation: transformers 4.38.2
    Uninstalling transformers-4.38.2:
      Successfully uninstalled transformers-4.38.2
Successfully installed nvidia-cublas-cu12-12.1.3.1 nvidia-cuda-cupti-cu12-1
```

```python
import spacy
from spacy.tokens import DocBin
from tqdm import tqdm
import json
```

```python
cv_data = json.load(open('/content/train_data.json', 'r'))
```

```python
len(cv_data)
```

```
200
```

```python
cv_data[0]
```

```
['Govardhana K Senior Software Engineer  Bengaluru, Karnataka, Karnataka –
Email me on Indeed: indeed.com/r/Govardhana-K/ b2de315d95905b68  Total IT
experience 5 Years 6 Months Cloud Lending Solutions INC 4 Month •
Salesforce Developer Oracle 5 Years 2 Month • Core Java Developer
Languages Core Java, Go Lang Oracle PL-SQL programming, Sales Force
Developer with APEX.  Designations & Promotions  Willing to relocate:
```

Anywhere WORK EXPERIENCE Senior Software Engineer Cloud Lending
Solutions — Bangalore, Karnataka — January 2018 to Present Present
Senior Consultant Oracle — Bangalore, Karnataka — November 2016 to
December 2017 Staff Consultant Oracle — Bangalore, Karnataka — January
2014 to October 2016 Associate Consultant Oracle — Bangalore, Karnataka
— November 2012 to December 2013 EDUCATION B.E in Computer Science
Engineering Adithya Institute of Technology — Tamil Nadu September 2008
to June 2012 https://www.indeed.com/r/Govardhana-K/b2de315d95905b68?
isid=rex-download&ikw=download-top&co=IN
https://www.indeed.com/r/Govardhana-K/b2de315d95905b68?isid=rex-
download&ikw=download-top&co=IN SKILLS APEX. (Less than 1 year), Data
Structures (3 years), FLEXCUBE (5 years), Oracle (5 years), Algorithms (3
years) LINKS https://www.linkedin.com/in/govardhana-k-61024944/
ADDITIONAL INFORMATION Technical Proficiency: Languages: Core Java, Go
Lang, Data Structures & Algorithms, Oracle PL-SQL programming, Sales Force
with APEX. Tools: RADTool, Jdeveloper, NetBeans, Eclipse, SQL developer,
PL/SQL Developer, WinSCP, Putty Web Technologies: JavaScript, XML, HTML,
Webservice Operating Systems: Linux, Windows Version control system SVN &
Git-Hub Databases: Oracle Middleware: Web logic, OC4J Product FLEXCUBE:
Oracle FLEXCUBE Versions 10.x, 11.x and 12.x
https://www.linkedin.com/in/govardhana-k-61024944/',

{'entities': [[1749, 1755, 'Companies worked at'],
   [1696, 1702, 'Companies worked at'],
   [1417, 1423, 'Companies worked at'],
   [1356, 1793, 'Skills'],
   [1209, 1215, 'Companies worked at'],
   [1136, 1247, 'Skills'],
   [928, 932, 'Graduation Year'],
   [858, 889, 'College Name'],
   [821, 856, 'Degree'],
   [787, 791, 'Graduation Year'],
   [744, 750, 'Companies worked at'],
   [722, 742, 'Designation'],
   [658, 664, 'Companies worked at'],
   [640, 656, 'Designation'],
   [574, 580, 'Companies worked at'],
   [555, 572, 'Designation'],
   [470, 493, 'Companies worked at'],
   [444, 468, 'Designation'],
   [308, 314, 'Companies worked at'],
   [234, 240, 'Companies worked at'],
   [175, 198, 'Companies worked at'],
   [93, 136, 'Email Address'],
   [39, 48, 'Location'],
   [13, 37, 'Designation'],
   [0, 12, 'Name']]}]

Start coding or generate with AI.