

Matching CV_or_Resume with Job Description_production_ready_Format

```
In [1]: import pandas as pd
from sklearn.feature_extraction.text import CountVectorizer
from sklearn.metrics.pairwise import cosine_similarity
import pdfplumber
import PyPDF2

def read_pdf(file_path):
    """Read and extract text from a PDF file."""
    with open(file_path, 'rb') as pdf_file:
        script = []
        with pdfplumber.open(pdf_file) as pdf:
            for page in pdf.pages:
                text = page.extract_text()
                script.append(text)
        return ''.join(script).replace("\n", "")

def calculate_similarity(cv_text, req_text):
    """Calculate cosine similarity between two text documents."""
    match_test = [cv_text, req_text]
    cv = CountVectorizer()
    count_matrix = cv.fit_transform(match_test)
    similarity = cosine_similarity(count_matrix)[0][1]
    match_percentage = round(similarity * 100, 2)
    return match_percentage

def main():
    cv_path = "C:\\Users\\venka\\OneDrive\\Documents\\Venkatesh_DS_Resumes\\venkatesh_Data_Scientist.pdf"
    req_path = "C:\\Users\\venka\\OneDrive\\Documents\\Venkatesh_DS_Resumes\\JD.pdf"

    try:
        cv_text = read_pdf(cv_path)
        req_text = read_pdf(req_path)

        print("CV Text:")
        print(cv_text)
        print("\nRequirement Text:")
        print(req_text)

        match_percentage = calculate_similarity(cv_text, req_text)
        print(f"\nMatch Percentage is: {match_percentage}% to Requirement")

    except Exception as e:
        print(f"An error occurred: {e}")
```

```
if __name__ == "__main__":  
    main()
```

CV Text:

MUNGI VENKATESH Junior Data Scientist +91vvcxvzbcbvbE-mail: venkatesh.mungi.datasien(cid:2236)st@gmail.com || Blogs: h(cid:130)ps://bhaarathi-ai.blogspot.com/ || Address: KPHB Colony, Hyderabad, Telangana LinkedIn: h(cid:130)ps://www.linkedin.com/in/venkatesh-mungi-320233258 || GitHub: h (cid:130) p s : / / g i t h u b . c o m / v e n k a t e s h m u n g i ? t a b = r e p o s i t o r i e s || Phone: +91 9676079392

TECHNICAL SKILLS SUMMARY

Program ming Language & Tools Highly skilled and results-driven Data Scientist with over 2 years of experience in data acquisition, predictive modelling, machine learning, statistical analysis, Natural Language processing and Deep Python- Procedural, OOPs, SciPy, Learning. Proven track record of delivering actionable insights and driving business growth through NumPy, Pandas, Statistics data-driven decisions. Specialized in natural language processing and deep learning, consistently achieving high success rates in complex projects. Adept at applying a wide range of algorithms to Data Analysis & Visualization extract valuable information from data and develop innovative solutions. Strong analytical and Python, Seaborn, Matplotlib problem-solving abilities, coupled with a passion for staying current with the latest developments in Tableau, SQL-SMSS the field. Looking to contribute my expertise to a forward-thinking organization seeking to leverage Machine Learning Tools data for strategic advantage. Scikit-Learn, PyCaret, TensorFlow, Keras, PyTorch, SciPy, Pandas, WORK EXPERIENCE

Stats model Junior Data Scientist Machine Learning Algorithms Social Tek AI ML Business solutions Pvt Ltd, Hyderabad. Sep 2021 – Oct 2023

Linear Regression, Random Forest, Logistic Regression, Decision Tree, SVM, KNN, K-Means, H-Means, Successfully collected and extracted data from diverse sources using a range of scraping Naïve Bayes, XG-Bosst, LGBM tools, ensuring data quality and integrity. DL Tools & Frame Works Conducted thorough data cleaning and wrangling processes to transform raw data into a TensorFlow, Keras, PyTorch, NLTK structured and usable format. Spacy, PyCaret, Gensim, Textblob Employed advanced techniques for feature extraction and engineering, enhancing the Deep Learning Expertise predictive power of machine learning models. OpenCV, Computer Vision, YOLO, Utilized statistical analysis to uncover patterns, correlations, and insights within complex Natural Language Processing, ANN, datasets, driving informed decision-making. CNN, LSTM, Transformers, FRCNN, Developed and implemented machine learning models for various business applications, BE RT, Autoencoders, GANs optimizing performance and accuracy. Web scrapping tools Fine-tuned machine learning models by adjusting hyperparameters, conducting cross-BeautifulSoup, Selenium validation, and employing ensemble methods. CUDA-GPU Deep Learning Expert Applied natural language processing techniques to extract valuable information from Skilled in training deep learning models unstructured text data, enabling sentiment analysis, topic modeling, and text classification. with CUDA for faster GPU processing, Built and fine-tuned deep learning models, including neural networks, convolutional neural improving efficiency and performance networks (CNNs), and recurrent neural networks (RNNs), for tasks such as image recognition in both training and using models. and natural language understanding. Collaborated with cross-functional teams, including data engineers and business SOFT SKILLS stakeholders, to deliver data-driven solutions that aligned with organizational goals. Analytical Thinking, Problem Solving, Stayed updated with the latest advancements in data science, attending conferences, participating in online forums, and pursuing continuous learning opportunities. Adaptability, Collaboration, Communication, Provided technical guidance and mentorship to data science interns, fostering a Time management, Attention to detail, collaborative and knowledge-sharing environment. Creativity, Self-motivation, Continuous PROJECTS

Learning Resume Screening using PyTorch BERT – NLP EDUCATION Description: Developed a resume screening system using PyTorch BERT with an accuracy of 82%. Collected and Master of Science, Andhra University cleaned data to create a high-quality dataset. Implemented data preprocessing and tokenization techniques to B.S c. Computer Science, Andhra University prepare the data for modelling. Successfully fine-tuned the PyTorch BERT model for resume classification. Contributed to automating the resume screening process, reducing manual

effort. Technologies used: Python, PyTorch, BERT, data scraping libraries, Visual Studio Code

CERTIFICATIONS

House Price Prediction – Machine Learning – Predictive Modelling TS SET (Telangana State Eligibility Test) Qualified, [2018]. Description: Collected, Cleaned and prepared the dataset by performing EDA and Feature Engineering to implement the model, and optimize its performance. Successfully Built an efficient house price prediction system using Python's K-Nearest Neighbours (KNN) algorithm with R2 score of 0.85 and RMSE of 7.9.

Technologies used: Certificate of completion on "Coding for Python, Scikit-learn, data preprocessing techniques, data visualization every body in Python 3x" powered by AttainU (A Unit of Leetu Educational Services Pvt. Traffic signs identification using CNN Ltd.)

Description: The information was gathered by our data collection team and given to us afterwards. Worked with pre-processed data. Applied convolution neural network's sequential model to obtain best accuracy to PUBLICATIONS identify the given traffic signs. CNN Sequential model yields 0.71 % accuracy.

Technologies used: Pandas, Keras, <https://medium.com/@venkateshmungi1247NumPy>, Visual Studio Code, Convolution Neural Networks

Requirement Text:

About the job We are actively searching for a talented and experienced Machine Learning (ML) Engineer to join our team. As a Machine Learning Engineer, you will play a crucial role in the development and implementation of cutting-edge artificial intelligence products. Your responsibilities will involve designing and constructing sophisticated machine learning models, as well as refining and updating existing systems. In order to thrive in this position, you must possess exceptional skills in statistics and programming, as well as a deep understanding of data science and software engineering principles. Your ultimate objective will be to create highly efficient self-learning applications that can adapt and evolve over time, pushing the boundaries of AI technology. Join us and be at the forefront of innovation in the field of machine learning.

Responsibilities

- Study and transform data science prototypes
- Design machine learning systems
- Research and implement appropriate ML algorithms and tools
- Develop machine learning applications according to requirements
- Select appropriate datasets and data representation methods
- Run machine learning tests and experiments
- Perform statistical analysis and fine-tuning using test results
- Train and retrain systems when necessary
- Extend existing ML libraries and frameworks
- Keep abreast of developments in the field

Requirements and skills

- Proven experience as a Machine Learning Engineer or similar role
- Understanding of data structures, data modeling and software architecture
- Deep knowledge of math, probability, statistics and algorithms
- Ability to write robust code in Python, Java and R
- Familiarity with machine learning frameworks (like Keras or PyTorch) and libraries (like scikit-learn)
- Excellent communication skills
- Ability to work in a team
- Outstanding analytical and problem-solving skills

Match Percentage is: 67.65% to Requirement