

<https://www.kaggle.com/code/bhupesh25/notebooka8f71da09c> This notebook has ease in running our program.

## Problem statement

In today's job market, job seekers and employers face challenges in understanding resumes. Job seekers must carefully choose content to impress employers, while employers struggle to compare resumes due to the overwhelming number of options and the difficulty in comparing different resumes. Both parties face time-consuming and time-consuming tasks.

The solution analysis a resume based on a specific focus area and highlights relevant words or phrases. It also identifies soft skills, such as leadership qualities or innovation, to aid in evaluating candidates holistically. The tool can be used to save the resume or display it on the frontend for easy visualisation.

## Implementation - Details

### Data Collection:

```
file_path = '/kaggle/input/resumes/book.csv'
df = pd.read_csv(file_path) # Adjust separator based on your CSV
```

### Labelling:

```
# Define queries and keywords for hypothetical label assignment
queries = ['Frontend', 'Backend', 'Mathematics', 'Programming', 'Leadership', 'Accounting', 'AI', 'ML']
category_keywords = {
    'Frontend': ['html', 'css', 'javascript', 'react', 'vue', 'angular'],
    'Backend': ['node.js', 'django', 'flask', 'ruby on rails', 'express'],
    'Mathematics': ['algebra', 'calculus', 'statistics', 'geometry', 'probability'],
    'Programming': ['python', 'java', 'c++', 'go', 'rust', 'kotlin'],
    'Leadership': ['management', 'team lead', 'director', 'head', 'supervisor'],
    'Accounting': ['taxes', 'accounts receivable', 'ledger', 'financial statements', 'bookkeeping'],
    'AI': ['artificial intelligence', 'neural networks', 'deep learning', 'nlp', 'computer vision'],
    'ML': ['machine learning', 'svm', 'random forest', 'k-means', 'regression']
}

def assign_query_and_label(text):
    query = np.random.choice(queries)
    keywords = category_keywords[query]
    label = 1 if any(keyword in text.lower() for keyword in keywords) else 0
    return query, label

df['Query'], df['Label'] = zip(*df['Resume_str'].apply(assign_query_and_label))
```

### Model Selection:

```
from transformers import AutoTokenizer, AutoModelForSequenceClassification, TrainingArguments, Trainer
tokenizer = AutoTokenizer.from_pretrained('bert-base-uncased')
```

### Training:

```
# Training arguments
training_args = TrainingArguments(
    output_dir='./models/',
    num_train_epochs=1,
    per_device_train_batch_size=8,
    per_device_eval_batch_size=8,
    warmup_steps=500,
    weight_decay=0.01,
    logging_dir='./logs/',
)
```

```
# Initialize Trainer
trainer = Trainer(
    model=model,
    args=training_args,
    train_dataset=train_dataset,
    eval_dataset=eval_dataset,
)

# Start training
trainer.train()
model_path = '/kaggle/working/models'
model.save_pretrained(model_path)
```

## Evaluation :

**Tokenization:** The resume text and the query are converted into tokens, which are smaller units of text used by the BERT model.

```
# Tokenize input text and query
inputs = tokenizer(query, resume_text, return_tensors='pt', padding=True, truncation=True)
```

**Embeddings:** The BERT model converts these tokens into numerical representations called embeddings. These embeddings capture the meaning of words in a high-dimensional space.

```
# Get BERT embeddings
with torch.no_grad():
    outputs = model(**inputs)
    query_embedding = outputs.last_hidden_state[:, 0, :] # Query embedding
    resume_embeddings = outputs.last_hidden_state[:, 1:, :] # Resume embeddings
```

**Cosine Similarity:** The code calculates the cosine similarity between the embedding of the query and each token in the resume text. Cosine similarity measures how similar two vectors are in direction.

```
# Compute cosine similarity between query and each token in resume text
similarities = torch.nn.functional.cosine_similarity(query_embedding, resume_embeddings, dim=-1)
```

**Keyword Extraction:** The tokens in the resume are sorted based on their similarity to the query. The top 10 most similar tokens are extracted as keywords.

**Filtering:** Any keywords containing non-alphanumeric characters are removed to ensure only meaningful words are returned.

```
# Sort tokens by similarity and extract keywords
sorted_indices = torch.argsort(similarities, descending=True, dim=-1)
keywords = [tokenizer.convert_ids_to_tokens(inputs['input_ids'][0][i].item()) for i in sorted_indices[0][:10]] # Extract top 10 keywords

# Filter out keywords containing non-alphanumeric characters
keywords = [word for word in keywords if re.match("^[a-zA-Z]+$", word)]
return keywords
```

## Main Features :-

- 1) **Accuracy & Relevance** :- We were to correctly match related words according to semantic accuracy. The semantic accuracy is about 70%. Though this is not able to match relevant non-vocabulary words as semantic matching was done with BERT.
- 2 ) **Latency** :- Our Model is able to train in 10 Minutes but can execute queries in less than 30 seconds so latency is not an issue with the model.
- 3) **Robustness & Noise Handling**: We are able to handle misspellings , abbreviations and irregularities. to some extent as BERT tries semantic analysis on the given tokens.
- 4 ) **Strength of Highlighting** :- Our model do a priority check by giving weight to the top 10 relevant points compared to less relevant ones using cosine similarity and then does a regular expression matching it with only alphabets(both lower and uppercase).

## Sample Outputs

### 1) Query :- Technical Skills

#### Link to the pdf :-

[https://drive.google.com/file/d/1wUxWtcigB3v0hFVcZETt6VBaMKoJyUvU/view?usp=drive\\_link](https://drive.google.com/file/d/1wUxWtcigB3v0hFVcZETt6VBaMKoJyUvU/view?usp=drive_link)

## Output Screenshots

### INFORMATION TECHNOLOGY TECHNICIAN I

#### Summary

Versatile Systems Administrator possessing superior troubleshooting skills for networking issues, end user problems, and network security. Experienced in server management, systems analysis, and offering in-depth understanding of IT infrastructure areas. Detail-oriented, independent, and focused on taking a systematic approach to solving complex problems. Demonstrated exceptional technical knowledge and skills while working with various teams to achieve shared goals and objectives.

#### Highlights

- Active Directory
- Group Policy Objects
- PowerShell and VBScript
- Microsoft Exchange
- VMWare experience
- New technology and product research
- Office 365 and Azure
- Storage management
- Enterprise backup management
- Disaster recovery

#### Experience

Information Technology Technician I Aug 2007 to Current

Company Name 1/4 City, State

- Migrating and managing user accounts in Microsoft Office 365 and Exchange Online.
- Creating and managing virtual machines for systems such as domain controllers and Active Directory Federation Services (ADFS) in Microsoft Windows Azure (IaaS).
- Creating and managing storage in Microsoft Windows Azure (IaaS).
- Installing and configuring StorSimple iSCSI cloud array (STaaS/BaaS).
- Installing, configuring, and testing Twinstrata iSCSI cloud array (STaaS/BaaS).
- Collaborating on project plan for Office 365 migration.
- Developing detailed specifications for the Office 365 migration, including business-case documentation, cost benefit analyses, technical diagrams, and work flow documentation.
- Received training in MVC 4 for Visual Studio using .Net Framework 4/4.5 to develop application using HTML5 and CSS3.
- Installing, configuring, and supporting Linux machines for the open Wi-Fi network project.
- Compiling and generating statistical information concerning wireless network traffic using Cacti.
- Configuring wireless LAN router networking and security access.

- Developing detailed specifications for the acquisition of an Enterprise backup system including systems design, business-case **documentation**, cost benefit analysis, technical diagrams, and work flow **documentation**.
- Reviewing, evaluating, and analyzing departmental policies, guidelines, procedures, and standards with management and staff.
- Developing test scripts for acceptance, unit, and system testing of Hyperion Phase 1 and MiamiBiz Phase 2.
- Developing Quality Assurance and testing plan for Hyperion Phase 1 and MiamiBiz Phase 2.
- Debugging and logging of errors in Hyperion and MiamiBiz using Team Foundation Server (TFS).
- Participated in various phases of the project life cycle such as: determining requirements, design conceptualization, testing, implementation, deployment, and release for the Hyperion and MiamiBiz projects.
- Collaborating on project plans for Hyperion and MiamiBiz.
- Preparing presentations and **documentation** to demonstrate Hyperion and MiamiBiz functionality or design.
- Monitoring network traffic, and compiling and generating statistical information using Solar Winds.
- Collaborating on Disaster Recovery plan and procedures.
- Researching, evaluating, and recommending new hardware and new software.
- Communicating and defining systems design and requirements for new and existing systems and applications.
- Researching, evaluating, recommending, testing, and implementing third party software/utilities.
- Planning and designing network infrastructure changes â€” adding/removing servers, appliances, network logical flow.
- Reviewing, evaluating, and analyzing existing system and application viability with management and staff.
- Administering and maintaining shares on the file servers.
- Reviewing server logs to troubleshoot issues.
- Scheduling and applying hot fixes and **security** patches on the server infrastructure which includes the operating system and application software.
- Reviewing systems reporting in SCCM (System Center Configuration Manager).
- Resolving service requests escalated by the Help Desk or other technicians.
- Troubleshooting and analyzing and system **problems** for root cause analysis.
- Giving and participating in training and education programs to explain upgrades to end users.
- Migrating users' documents from local computer storage to shares on the file servers.
- Configuring, supporting, and maintaining file shares using Distributed File System (DFS)
- Managing, implementing, and testing Enterprise backup infrastructure systems such as the Symantec Veritas Netbackup, Symantec Backup Exec System Recovery/Livestate, and VRanger backup servers.

- Installing, configuring, and supporting McAfee anti-virus software on servers.
- Migrating Exchange infrastructure from Exchange 2003 to Exchange 2007 and from Exchange 2007 to Exchange 2010.
- Supporting servers in the virtualization infrastructure using VMware vSphere.
- Installing, configuring, and testing Veeam virtual machine backup software and Virtual Desktop Infrastructure (VDI).
- Reviewing systems reporting in System Center Configuration Manager (SCCM).
- Administering and maintaining the Symantec Enterprise Vault servers.
- Managing the Active Directory Domain Controllers (DCs).
- Creating and maintaining Group Policy Objects (GPOs) in Microsoft Active Directory.
- Configuring and supporting Microsoft Exchange Active Sync on devices with Apple iOS and Android mobile operating systems.
- Configuring and supporting BlackBerry devices on the BlackBerry Enterprise Server to receive Exchange email.
- Developing, testing, designing, and implementing application scripts using languages such as command batch files, Visual Basic Script, and PowerShell.
- Creating policies and procedural **documentation**.

Information Services Liaison, T Aug 2005 to Aug 2007

Company Name 1/4 City, State

- Troubleshooting hardware and software **problems** over the telephone and through remote PC administration software.
- Installing, configuring, and supporting McAfee anti-virus software on desktops.
- Installing, configuring, and supporting BBar computer backup software.
- Developing and maintaining websites on servers running Microsoft SharePoint Server and Internet Information Services (IIS).
- Supporting Systems Management Server (SMS)
- Troubleshooting LAN, WAN, Internet, and Intranet network and **security** access.
- Troubleshooting network connectivity issues related to TCP/IP, Domain Name Service (DNS), Dynamic Host Configuration Protocol (DHCP) protocols, Internet **Security** and Acceleration (ISA) proxy server, and VPN.
- Troubleshooting web application/page issues, client browsers, and related software.
- Administering and maintaining of end user accounts, permissions, and access rights in Microsoft Active Directory.
- Administering and maintaining of NTFS **security** permissions on the file servers.
- Installing, configuring, and maintaining hardware such as: servers, workstations, laptops, printers, and scanners in a Windows Enterprise environment.
- Installing, configuring, and supporting printers on the print servers.
- Installing, configuring, and supporting Microsoft Windows Server 2000 and 2003, Microsoft Windows XP and Windows Vista, and Microsoft Office XP, 2003, and 2007.

Education

Bachelor of Science, Information Technology 2005 Florida International University 1/4 City, State, United States

- Coursework in Programming, Web Administration, Network Administration, Database Administration, and Systems Administration â€” Linux
- Programming Languages: C++, Java, JSP, HTML, CSS, VB.NET, Bash, T-SQL

Certifications

CompTIA Network+ - 2014

Skills

Active Directory, Azure, anti-virus, Backup Exec, backup, Bash, batch, Cacti, Cisco ASA, databases, DHCP, DNS, **documentation**, DataDomain, EMC, Enterprise Vault, ePO, file servers, firewall, GPO, HTML, IIS, ISA, LDAP, Linux, McAfee, Exchange, Microsoft Office, Microsoft Windows, **security**, policies, PowerShell, programming, proxy server, servers, scripts, SolarWinds, SQL, StorSimple, troubleshooting, TMG, Ubuntu, Visual Basic Script, VBS, Veritas Netbackup, VPN, VRanger, Veeam, VMWare, VDI, virtual machine, NMap, ZenMap.



## 2) Query :- leadership qualities in second

### Link to the resume :-

[https://drive.google.com/file/d/1HzRjUx3\\_1y0wpcFZq3XL9BhgKOezP2gk/view?usp=drive\\_link](https://drive.google.com/file/d/1HzRjUx3_1y0wpcFZq3XL9BhgKOezP2gk/view?usp=drive_link)

- I successfully supervised the accountability of all financial reports and statements.
- I was responsible for the completeness and accuracy of weekly, monthly, quarterly, semi-annual, and annual reports.
- Monitored errors in the General Accounting and Finance System (GAFS/BQ) and ensured corrective actions were accomplished.
- I also ensured fund balances were reconciled to the appropriate audit listings and verified reports prior to release to base activities and higher headquarters.
- I furnished accounting data to base organizations often interpreting and analyzing the data to help funds managers resolve problems and manage their programs more effectively.
- I attended Major Command (MAJCOM) and Headquarters level workshops to participate and contribute to accounting policy and system changes.
- I provided professional assistance to Data Automation relevant to processing of accounting and finance data, interpreting deficiencies in software based on output products and system related problems.
- I utilized my working knowledge of commercial and government accounting system principles and knowledge of Processing Centers (PCs) to review, verify, analyze, and evaluate accounting and finance operations.
- While serving as Chief, Account Control I ensured areas of concern were addressed, concentrating on problem areas related to the database.
- I analyzed computer output products to determine processing deficiencies.
- They included, but were not limited to, the Open Document Listing (ODL), Operating Budget Ledger (OBL), Allotment Ledger (AL), and the Accounting and Finance Workload Information Management System (A&F WIMS) Extract list.
- I provided technical assistance related to policy and procedural changes required as a result of the impending base closure.
- Analyzed/developed and recommended improved training procedures enabling better use of system procedures ensuring governing directives were followed.
- I evaluated accuracy of accounting records prior to fiscal year closeout ensuring the Accounting and Finance Officer could certify their accuracy as required by regulation.
- Examined accounting transactions and documents to ensure they conformed to established accounting policy and principles.
- Coordinated and directed fiscal year end conversion for the GAFS and Integrated Accounts Payable System (IAPS).

#### Education

Northern Maine Community College 1994 Associate : Accounting City , State , USA

#### Emphasis in Business

1994 Associates : Accounting City , State , USA GPA: GPA: 3.41

Accounting GPA: 3.41 174 Hours, Quarter

Attended Husson College, major Accounting 78 semester hours toward Bachelors degree.

Professional Military Comptroller School, 6wk, 4-98; Managerial Accounting I, 09-98; Interested-Based Bargaining Training for Management, 24hrs, 09-01; Auditing Methods and Concepts 09-98; Organizational Leadership, 32hrs, 07-03; Management Development II, 32hrs, 07-03.

#### Certifications

Certified Defense Financial Manager, CDFM, May 2005

#### Interests

American Society Of Military Comptrollers

#### Additional Information

#### Skills

Accounting, General Accounting, Accounts Payable, Program Management.

helpful Links :-

1)

<https://www.kaggle.com/code/bhupesh25/notebook8f71da09c>

2) [Hugging Face – The AI community building the future.](#)