Resume Screening App using AI

Bulla Rajesh, Mrinal S

Atria Institute of Technology, Bengaluru, India

Abstract—The flood of resumes in today's talent acquisition landscape makes it difficult for hiring managers and recruiters to find qualified applicants quickly. This study presents an innovative method to integrate artificial intelligence techniques into a custom application to expedite the resume screening process. The program increases the effectiveness of recruitment workflows by a webapp that can be created to take inputs in form of pdf or csv format from the user and screens the resume. The development of a prototype application is the first step in the research methodology. It is then thoroughly verified and analyzed.

I. Introduction

The initial phase in the hiring process is resume screening, during which recruiters or hiring managers go through an extensive collection of resumes in looking for qualified applicants for a position. It entails reviewing each resume to ascertain whether the candidate has the education, training, and work history needed for the role.

They play a vital role in the hiring process due to increased volume of the received applications. Further subject expertise is needed for the screening process in order to comprehend the requirements and eligibility for the employment function.

Resume screening is done using zero shot classification from open-source Hugging Face Bart large mnli model. zero shot classification enables the app to classify the resumes accurately according to the classes provided.

Principles of Resume Screening App

The resume screening app is primarily dependent upon 3 factors

- Labels: For any resume screening process mentioning the labels is important as it defines the goal or the objective of the classification task

- File type: The file type of the resume submitted by the user must be pdf or csv.

- Data Cleaning: Before classification process the resume undergoes a data cleaning process to eliminate special characters, unwanted spaces etc.

Alternative Work Going On In The Field And TheirDrawback

Contract or Freelance Work: Although projects involving flexibility and a wide range of skills might be demonstrated, resume screening applications could find it difficult to assess the value or applicability of such experiences. Applicants who have chosen alternative career paths may be at a disadvantage as they may give preference to applicants with traditional job history over those on contract.

Non-Traditional Education and Skills Development: Resume screening tools that give preference to degrees from traditional educational institutions may provide difficulties for candidates who have followed nontraditional educational pathways, such as online courses, boot camps, or self-directed study. Similar to this, these applications might not be able to adequately record or assess abilities learned through unconventional methods.

What our system proposes: User uploads resume and zero shot classification from open-source Hugging Face Bart large mnli model is applied to the resume. By specifying labels, the zero-shot classification model predicts the suitable job.

II. Data Flow Model
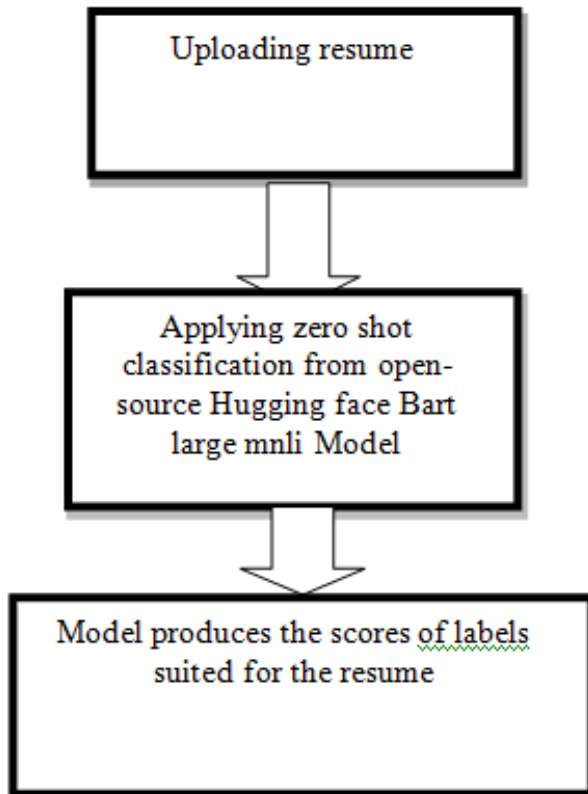
Development of Model

- <u>Inputs:</u> resume, labels.

- <u>Output</u>: scores of the labels suited for the resume.

Learning Code

This particular problem comes under transfer learning which refers to using a model trained for one task in a different application than what it was generally trained for.

In zero shot classification we provide the model with a prompt and a sequence of text that describes what we want our model to do, in natural language.

Loading the model

Facebook/Bart-Large-mnli model is loaded with zero shot classification task from Hugging Face.

zero-shot-classification can improve the effectiveness and productivity of a resume screening software through enabling it to classify resume applications without requiring explicit training on every category or skill. It is an effective solution for automating and optimizing the resume screening process because of its scalability and versatility.

```
# Load zero-shot classification model

classifier = pipeline("zero-shot-classification", model="facebook/bart-large-mnli")
```

III. Data Cleaning

By importing Regular expression library, we can use it for Data Cleaning by replacing special characters, non-ASCII characters, unwanted spaces, occurrences of RT and cc and unwanted string values with single whitespace character

```python
def clean_text(text):
    # Implement your data cleaning steps here
    cleaned_text = re.sub('http\S+\s', ' ', text)
    cleaned_text = re.sub('RT|cc', ' ', cleaned_text)
    cleaned_text = re.sub('#\S+\s', ' ',cleaned_text)
    cleaned_text = re.sub('@\S+', ' ', cleaned_text)
    cleaned_text = re.sub('[%s]' % re.escape("""!"#$%&'()*+,-./:;<=>?@[\]^_`{|}~"""), ' ',
    cleaned_text = re.sub(r'[^\x00-\x7f]', ' ', cleaned_text)
    cleaned_text = re.sub('\s+', ' ', cleaned_text)
    return cleaned_text
```

StreamLit WebApp

StreamLit is an open-source python library that facilitates the development and dissemination of stunning, unique online applications for data science and machine learning. It takes only a few minutes to create and implement robust data applications.

For the resume screening project, the web application can come very useful as users may submit resumes straight into the program with StreamLit' s user friendly interface.

Just by submitting the 2 inputs which are labels and resume we can predict the scores of labels suited for the resume.

Scenario 1: When input is a data science resume and input labels provided are machine learning engineer and software developer. The output scores of labels suited for the job is given below.

```
labels" : [
  0 : "Machine learning engineer"
  1 : "Software Developer"

scores" : [
  0 : 0.8814599514007568
  1 : 0.11854010075330734
```

We can see that machine learning engineer is a highly suitable job compared to the rest of the labels.

Scenario 2: When input is a software developer resume and input labels are software developer, IoT developer and web developer. The output scores of labels suited for the job is given below.

```
"labels" : [
  0 : "Software developer"
  1 : " web developer"
  2 : " Iot developer"

"scores" : [
  0 : 0.5966267585754395
  1 : 0.2822297513484955
  2 : 0.12114345282316208
```

We can see that software developer is a highly suitable job compared to the rest of the labels.

IV. Conclusion

Hence, we can see that zero-shot classification can definitely help us in screening of resumes. After integration of this model with web or app development it can turn into a large-scale application which can ease the burden of the recruiters as they will not have to manually view a flood of resumes.

References

1. Jie Chen, Chunxia Zhang and Zhendong Niu, "A Two- Step Resume Information Extraction Algorithm", Mathematical Problems in Engineering, vol. 2018, pp. 8, 2018.
2. Kajal Kumari, Resume Screening with Natural Language Processing in Python, June 2021.
3. T Schmitt, P Caillou and M Sebag, "Matching Jobs and Resumes: a Deep Collaborative Filtering Task", Proc. of the 2nd Global Conf. on Artificial Intelligence, pp. 1-14, 2016.
4. C. Tapsai, "Information Processing and Retrieval from CSV File by Natural Language", 2018 IEEE 3rd International Conference on Communication and Information Systems (ICCIS), pp. 212-216, 2018.
5. Abdul Wahab and M Nachappa, "Resume Parser with Natural Language Processing", International Research Journal of Engineering and Technology, vol. 09, no. 03, 2022.
6. Tim Zimmermann, Leo Kotschenreuther and Karsten Schmidt, "Data-driven HR Resume Analysis Based on Natural Language Processing and Machine Learning", arXiv, 2016.
7. J Chen, Z Niu and H Fu, "A Novel Knowledge Extraction Framework for Resumes Based on Text Classifier", Proceedings of the International Conference on Web-Age Information Management, pp. 540-543, 2015
8. A. Zaroor, M. Maree and M. Sabha, "A Hybrid Approach to Conceptual Classification and Ranking of Resumes and Their Corresponding Job Posts" in IDT 2017. Smart Innovation Systems and Technologies, Cham:Springer, vol. 72, 2017
9. Arvind Sinha, Amir Khusru and Ashwani Kumar, "Resume Screening Using Natural Language Processing and Machine Learning: A Systematic Review", Machine Learning and Information Processing Advances in Intelligent Systems and Computing, vol. 1311, 2021.
10. Yasunobu Kinoa, Hiroshi Kurokia, Tomomi Machidab, Norio Furuyab and Kanako Takano, "Text Analysis for Job Matching Quality Improvement", International Conference on Knowledge Based and Intelligent Information and Engineering Systems Procedia Computer Science, vol. 112, pp. 1523-1530, 2017.
11. Pradeep Roy, Sarabjeet Chowdhary and Rocky Bhatia, "A Machine Learning approach for automation of Resume Recommendation system", Procedia Computer Science, vol. 167, pp. 2318-2327, 2020.
12. P. K. Pushp and M. M. Srivastava, "Train once test anywhere: Zero-shot learning for text classification", CoRR, 2017.
13. A. Eriguchi, M. Johnson, O. Firat, Hideto Kazawa and Wolfgang Macherey, "Zero-Shot Cross-lingual Classification Using Multilingual Neural Machine Translation", CoRR, 2018.
14. J. Zhang, P. Lertvittayakumjorn and Y. Guo, "Integrating semantic knowledge to tackle zero-shot text classification", NAACL-HLT (2019), pp. 1031-1040, 2019.
15. S. Antol, C. Zitnick and D. Parikh, "Zero-shot learning via visual abstraction", Proceedings of ECCV, 2014.