

XỬ LÝ CV BẰNG KỸ THUẬT XỬ LÝ NGÔN NGỮ TỰ NHIÊN (NLP)

Nguyễn Bảo Trí - 230201059

Tóm tắt

- Lớp: CS2205.MAR2024
- Link Github: <https://github.com/tringuyen2020/CS2205.MAR2024>
- Link YouTube video: https://youtu.be/JWrU8_GirTA
- Họ và Tên: Nguyễn Bảo Trí



Giới thiệu

- Hiện tại, các doanh nghiệp và cơ quan tuyển dụng xử lý rất nhiều CV hàng ngày. Do đó, cần có một hệ thống thông minh tự động có thể lấy ra tất cả thông tin quan trọng từ các CV không có cấu trúc và chuyển đổi tất cả sang định dạng có cấu trúc chung, sau đó có thể xếp hạng các CVs cho một vị trí công việc cụ thể.



Mục tiêu



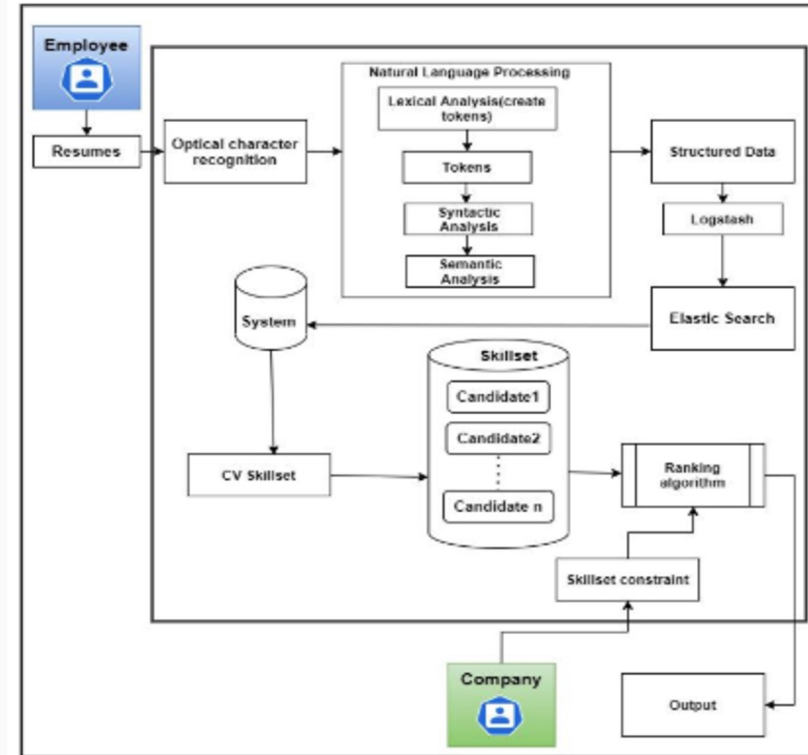
- Đơn giản hoá quy trình tuyển dụng bằng cách sử dụng Optical Character Recognition (OCR) và NLP để xử lý CV tự động nhằm cung cấp những ứng viên chất lượng cho các công ty.
- Giảm thiểu sự bất công và phân biệt đối xử trong quá trình lọc CV.
- Các CV sẽ được phân loại và xếp hạng dựa trên dữ liệu từ các mạng xã hội như LinkedIn.

Nội dung và Phương pháp

- **Nội dung:**

Trong hệ thống này, chúng tôi sử dụng NLP để phân tích CV theo các yêu cầu cụ thể. Một web portal việc làm chung cho người sử dụng lao động cũng như nhân viên được cung cấp để đăng ký và tạo công việc. Các CV nhận được sẽ được phân tích cú pháp và xếp hạng theo yêu cầu của công ty.

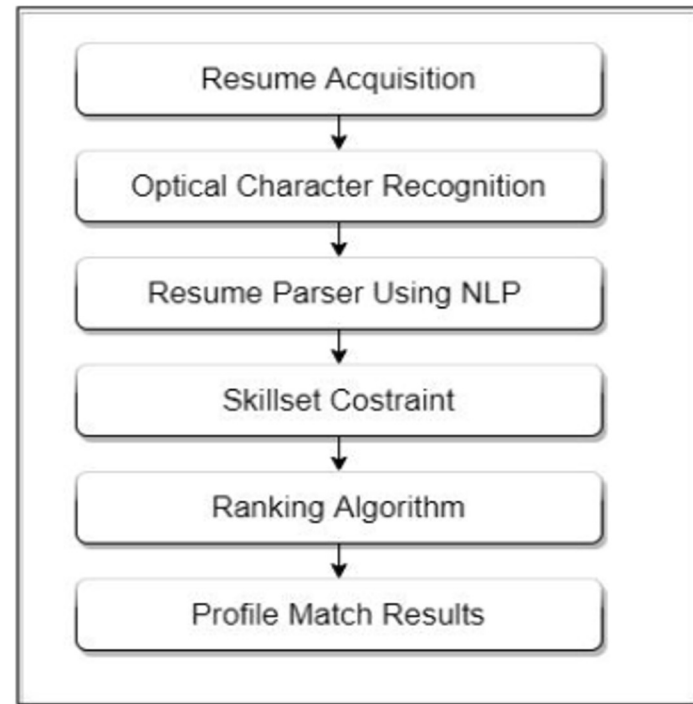
Ngoài ra, mục tiêu khác của chúng tôi là trích xuất dữ liệu từ mạng xã hội như LinkedIn để ứng tuyển, điều này sẽ giúp quá trình tuyển dụng dễ dàng nhận được các đơn đăng ký chất lượng từ nhiều khu vực khác nhau bằng cách tránh bất công và phân biệt đối xử.



Nội dung và Phương pháp

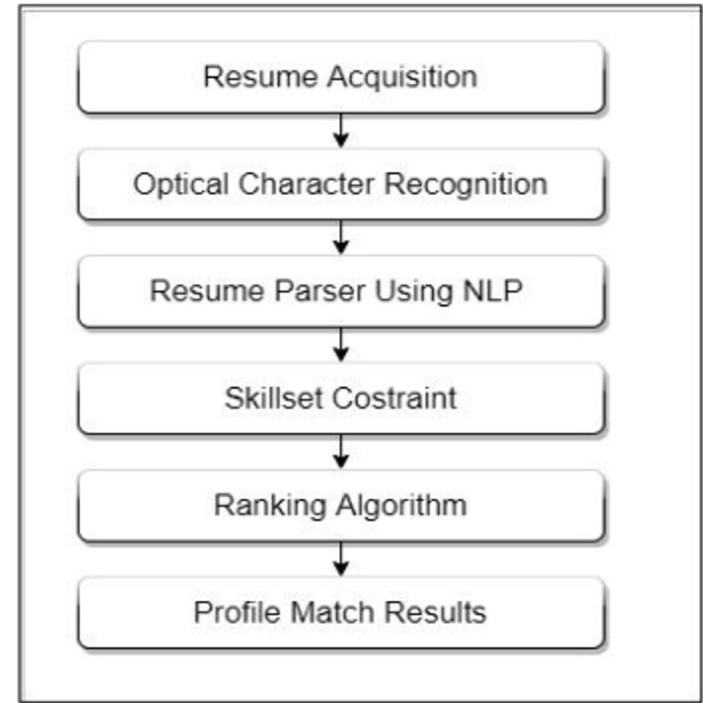
- **Phương pháp:**

1. Cung cấp web portal cho bộ phận HR
2. Ứng viên sử dụng web portal để tạo tài khoản và tải CV lên
3. Sử dụng OCR để chuyển đổi CV thành 1 dạng văn bản đơn giản
4. Sử dụng NLP để phân tích CV: Phân tích từ vựng, Phân tích cú pháp, Phân tích ngữ nghĩa, và Nhận dạng thực thể được đặt tên (Named Entity Recognition hay NER)



Nội dung và Phương pháp

5. Thể hiện dữ liệu thu được trên bảng thông tin (dashboard) chứa các biểu đồ dựa trên dữ liệu trong CV
6. Tính điểm và sắp xếp theo các ràng buộc do bộ phận HR cung cấp
7. Các đặc điểm cá nhân của ứng viên sẽ được tăng cường tương ứng dựa trên mức độ ưu tiên của đặc điểm đó
8. Điểm tích lũy sẽ được sử dụng để phân loại ứng viên. Danh sách ứng viên được sắp xếp cuối cùng sẽ được hiển thị cho bộ phận HR



Kết quả dự kiến

Việc ứng dụng NLP vào việc xử lý CV mang lại các kết quả mong đợi như sau:

- Việc xử lý CV trở nên tự động, nhanh chóng, và hiệu quả
- Tiết kiệm chi phí trong việc sàng lọc CV
- Giảm thiểu sai sót, bất công và phân biệt gây ra bởi người tuyển dụng

Tài liệu tham khảo

- [1]. Vinaya R. Kudatarkar, ManjulaRamannavar, Dr.Nandini S. Sidnal “An Unstructured Text Analytics Approach for Qualitative Evaluation of Resumes”, 2015, IJIRAE.
- [2]. Satyaki Sanyal, Neelanjan Ghosh, SouvikHazra, Soumyashree Adhikary, “Resume Parser with Natural language Processing”, 2007,IJESC.
- [3]. Papiya Das, Manjusha Pandey, Siddharth Swarup Rautaray, “A CV parser Model using Entity Extraction Process and Big Data Tools ”, 2018,IJITCS.
- [4]. Ayishathahira and Sreejith , “Combination of Neural Networks and Conditional Random Fields for Efficient Resume Parsing”,International CET Conference on Control, Communication and Computing(IC4),2018..
- [5]. Dr.Parkavi A,Pooja Pandey,Poornima J,Vaibhavi G S,Kaveri BW, “E-Recruitment System Through Resume Parsing,Psychometric Test and Social Media Analysis”, 2019,IJARBEST.
- [6]. Nirali Bhaliya, Jay Gandhi, Dheeraj Kumar Singh, “NLP based Extraction of Relevant Resume using Machine Learning ”,2020,IJITEE.