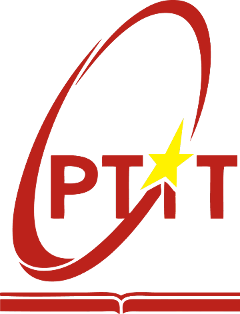
HỌC VIỆN CÔNG NGHỆ BƯU CHÍNH VIỄN THÔNG



**BÀI TẬP LỚN**

**MÔN HỌC: NHẬP MÔN KHOA HỌC DỮ LIỆU**

**Tên đề tài:  Phân tích và gợi ý việc làm trên các trang tuyển dụng**

**Sinh viên: Đỗ Công Ban**

**Mã sinh viên: B19DCCN057**

**Nhóm học phần:    03**

**Nhóm bài tập lớn: 07**

**Hà Nội, Tháng 6/2021**

**MỤC LỤC**

**Mở đầu**

**Phần I: Mô tả…………………………………………………………………….**

1. **Giới thiệu đề tài………………………………………………...………...**
2. **Phân tích yêu cầu chức năng của đề tài……………………....…………**
3. **Công nghệ sử dụng……………………………………….………………**

**Phần II: Phân tích các chức năng chính của đề tài…..……….……………….**

**Mô tả & demo chức năng…………………..……….……………………**

**Phần III: Kết luận………………………………………………………………...**

1. **Hướng phát triển………………………………….………………………**
2. **Tài liệu tham khảo……………………………….……………………….**

**Mở đầu**

Trong thời buổi hiện đại ngày nay, công nghệ thông tin cũng như những ứng

dụng của nó không ngừng phát triển, lượng thông tin và cơ sở dữ liệu được thu thập và lưu trữ cũng tích lũy ngày một nhiều lên. Con người cũng vì thế mà cần có thông tin với tốc độ nhanh nhất để đưa ra quyết định dựa trên lượng dữ liệu khổng lồ đã có. Vì thế cần phải có những phương pháp để thu thập, phân tích, khai phá dữ liệu một cách nhanh chóng và chính xác nhất.

Hiện nay vấn đề tìm việc của những sinh viên mới ra trường hay những người muốn tìm một công việc mới phù hợp hơn đang ngày càng tăng do đó bắt kịp được nhu cầu đó em quyết định thực hiện đề tài “**Phân tích và gợi ý việc làm trên các trang tuyển dụng**”. Đề tài trên là kết quả của quá trình học tập tích lũy và vận dụng những kiến thức mà em tiếp thu và tìm hiểu.

Mặc dù đã cố gắng trong quá trình xây dựng làm đề tài nhưng do còn nhiều hạn chế về thời gian và kiến thức nên sản phẩm của em không tránh khỏi những thiếu sót, những vấn đề chưa được giải quyết hoàn chỉnh. Vì vậy em rất mong nhận được những ý kiến đóng góp của thầy và những bạn trong lớp để có thể hoàn thiện và phát triển hơn.

Em xin chân thành cảm ơn!

Đỗ Công Ban

**I. Mô tả**

*1. Giới thiệu đề tài*

* Đề tài được tạo ra nhằm giải quyết vấn đề tìm việc làm của sinh viên mới ra trường hay những người muốn tìm việc mới phù hợp hơn với bản thân. Cụ thể trong đề tài này sẽ phân tích và gợi ý việc làm trên những trang tuyển dụng lớn như ITViec, Topcv, Topdev, Vietnamworks, Careerbuilder.
* Người dùng chỉ cần nhập vào chủ đề, từ khóa liên quan đến chuyên ngành mà mình muốn tìm và nhập tên thành phố muốn làm việc hệ thống sẽ thống kê số lượng việc làm trên từng trang tuyển dụng và gợi ý một vài việc làm phù hợp

*2. Phân tích yên cầu chức năng của đề tài*

Có 2 chức năng chính:

* Thống kê ra tổng số lượng việc làm ứng với keyword và nơi làm việc mà người dùng nhập vào sau đó sẽ phân tích và sẽ lên biểu đồ để người dùng có cái nhìn tổng thể nhất về việc mình đang muốn tìm từ đó sẽ chọn được công việc phù hợp
* Gợi ý ra một vài việc làm nổi bật nhất được các trang tuyển dụng ưu tiên để người dùng có thể tham khảo

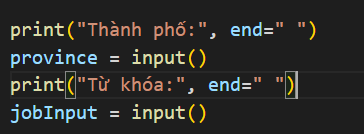
*3. Công nghệ sử dụng*

Sử dụng ngôn ngữ Python với các thư viện hỗ trợ đi kèm như: matplotlib, numpy, selenium, …

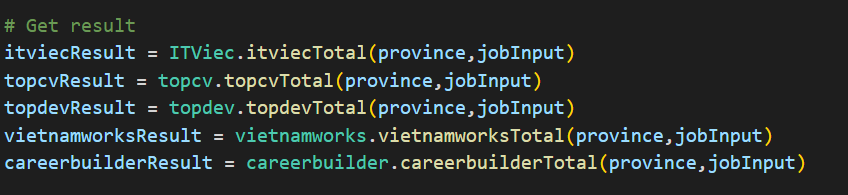
**II. Phân tích các chức năng chính**

1. *Thu nhập và thống kê*

* Hệ thống sẽ nhận 2 input đầu vào của người dùng là: thành phố làm việc và từ khóa của việc làm



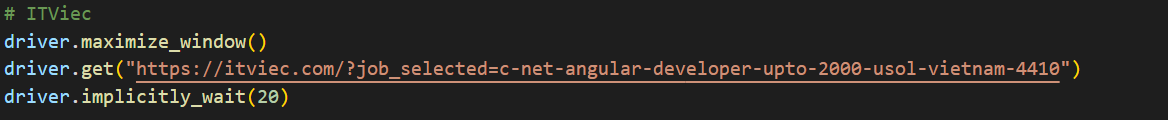
* Sau đó 2 keyword sẽ được truyền vào các hàm xử lý tương ứng với mỗi trang web tuyển dụng

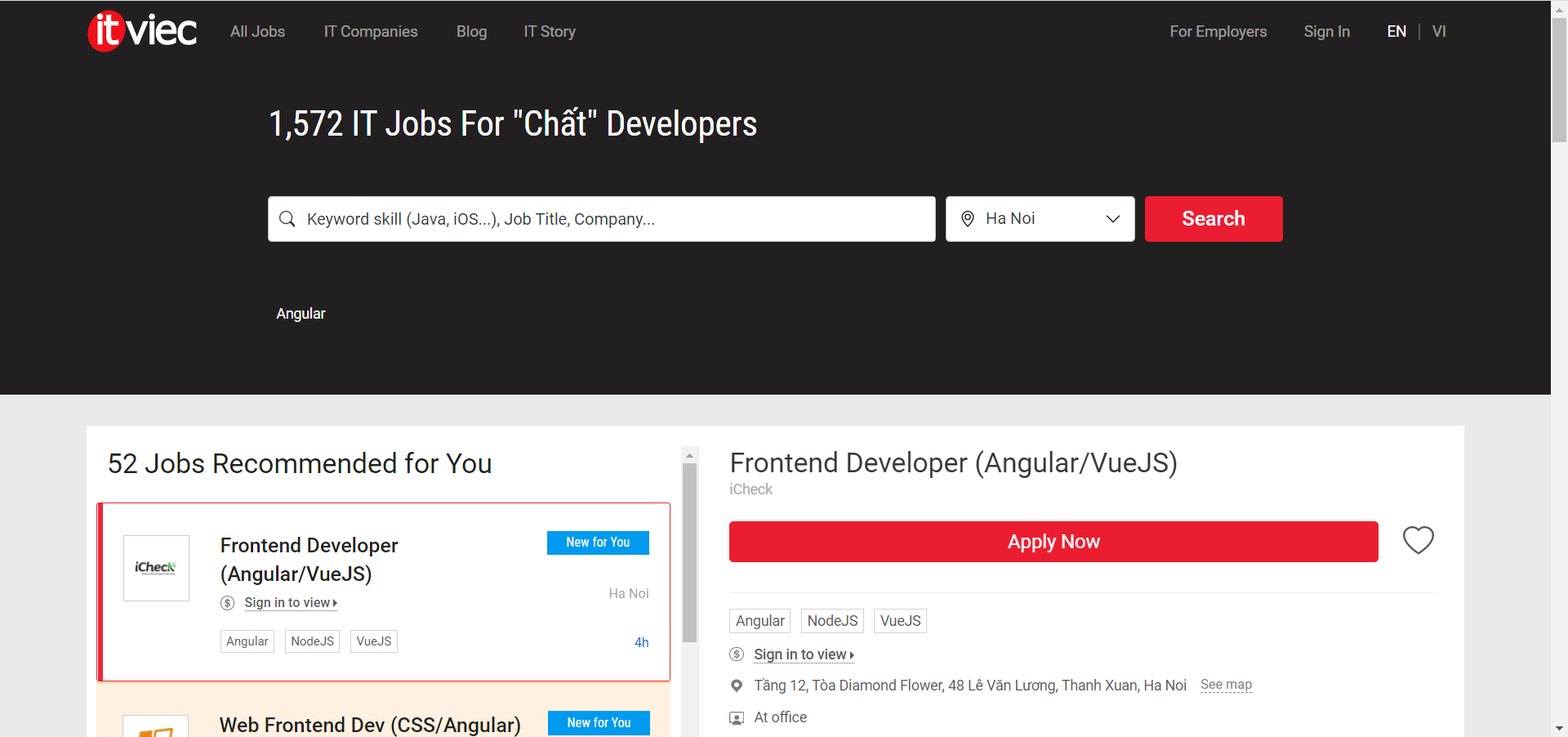


* Tại các hàm sẽ thực hiện truy cập vào trâng web, thực hiện tự động truyền các keyword và tìm kiếm

Ví dụ với trang tuyển dụng itviec.com

* Thực hiện mở trang web

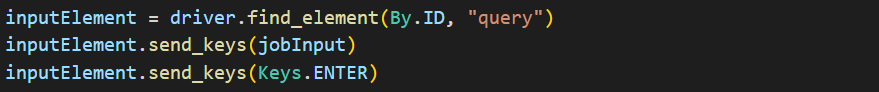


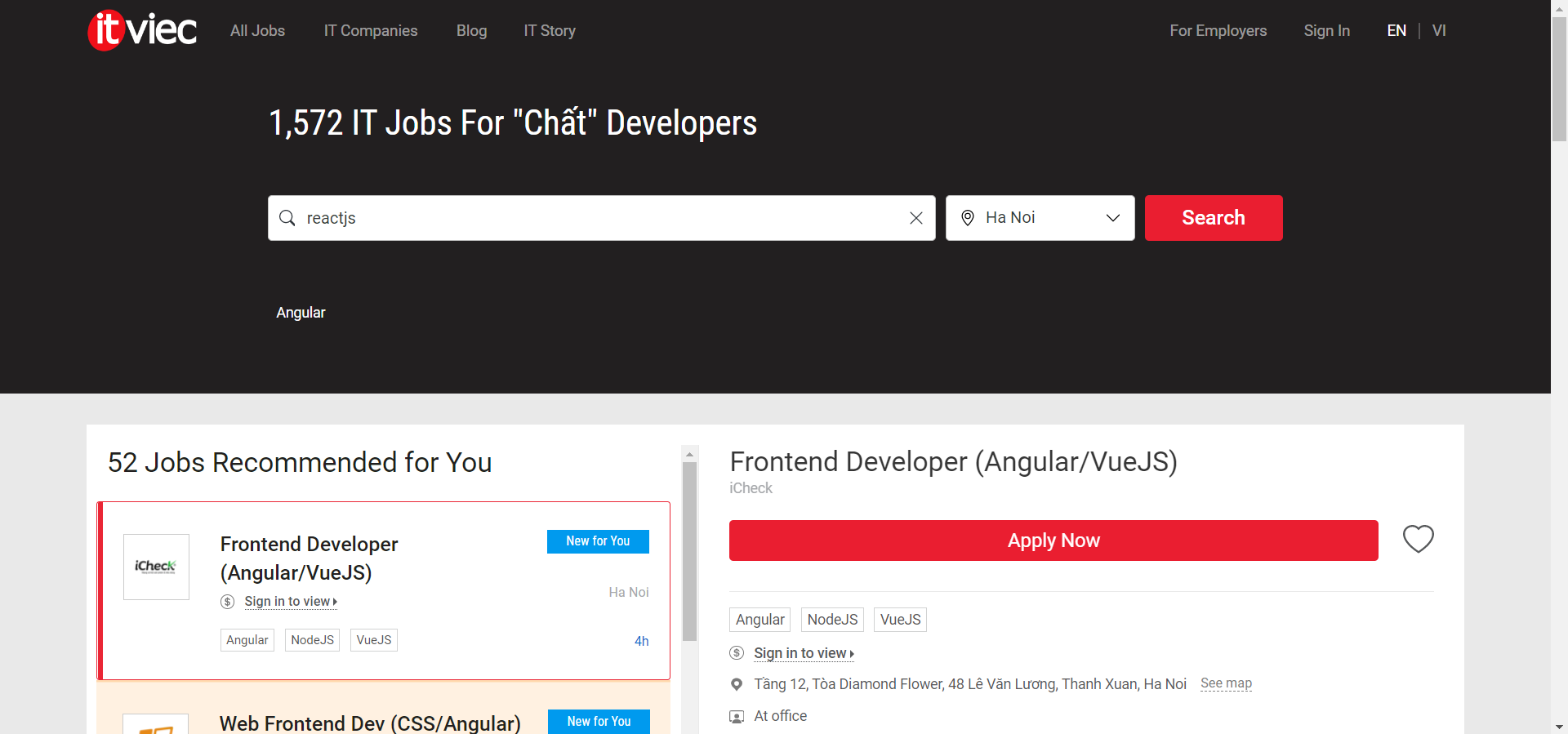


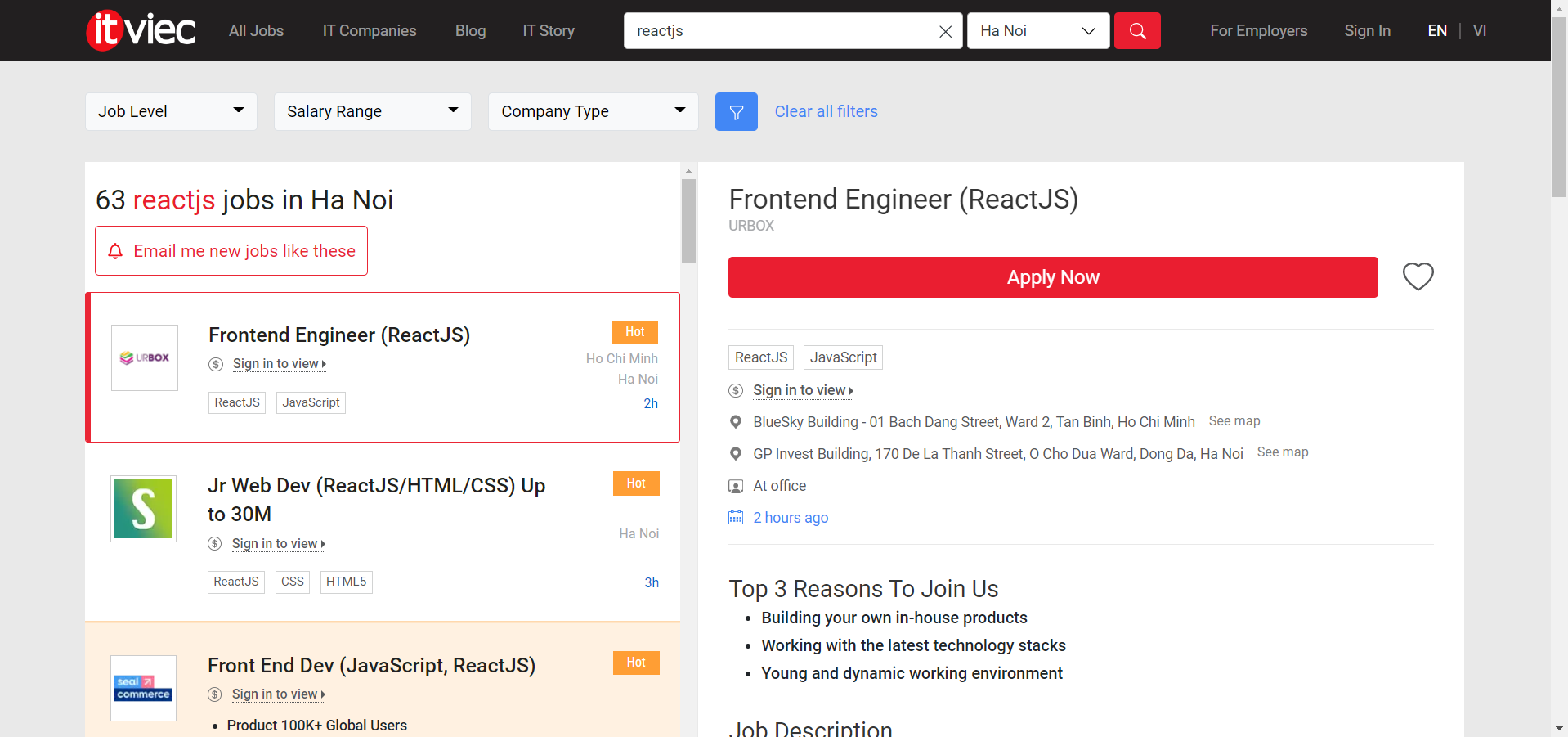
* Sau đó sẽ tự động tìm địa chỉ làm việc trong dropdownlist với địa chỉ là địa chỉ được nhập(\*Note: với thành phố mà người dùng nhập thì người dùng có thể nhập in hoa, in thường, có dấu, không dấu, hoặc viết liền không dấu thì đều được nhận. VD: người dùng muốn tìm tại thành phố Hà Nội thì có thể nhập: hà nội, Hà NộI, ha noi, hanoi, …)



* Thực hiện truyền từ khóa liên quan vào ô tìm kiếm việc làm và tự động bấm Enter để thực hiện tìm

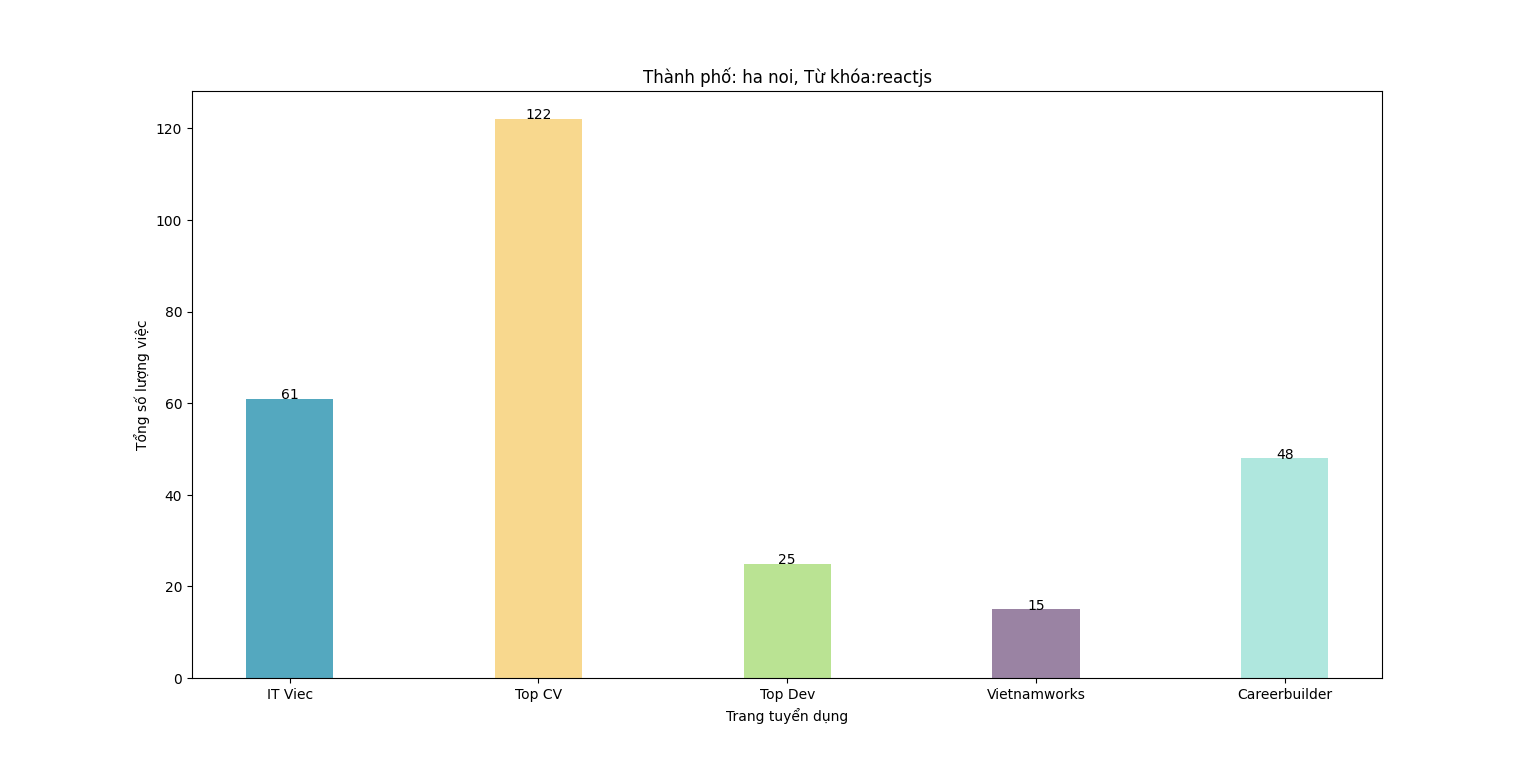






* Tổng số lượng việc làm sẽ được tách ra và phục vụ việc phân tích vẽ biểu đồ

Biểu đồ sẽ được vẽ theo dạng bar chart để người dùng có cái nhìn trực quan nhất, với hoành là từng trang tuyển dụng ứng với trục tung là tổng số lượng việc của trang đó. Trên mỗi cột sẽ hiển thị chi tiết tổng số lượng và ghi chú của biểu đồ là từ khóa và thành phố mà người dùng nhập.

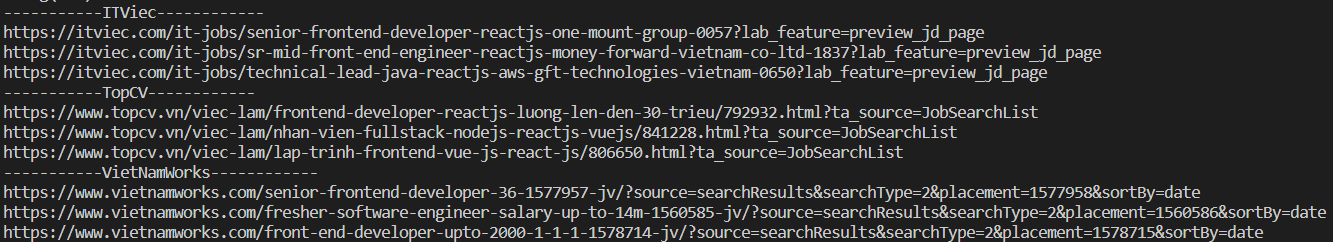


1. *Gợi ý việc làm*

Hệ thống sẽ yêu cầu nhập thêm mức lương(mức lương nhập theo USD) sau đó sẽ trả ra 3 công việc phù hợp nhất với những yêu cầu mà người dùng nhập bao gồm nơi làm việc, từ khóa công việc, mức lương.

Ví dụ:

* Nơi làm việc: Hà Nội
* Từ khóa công việc: reactjs
* Mức lương: 500 USD



**III. Kết luận**

Đề tài trên được hoàn thiện dựa trên những kiến thức em đã được tiếp thu từ trên lớp và những kiến thức thu thập từ bên ngoài. Sản phẩm còn một vài những khuyết điểm như: tốc độ phân tích và đư ra kết quả còn chậm do phụ thuộc vào tốc độ của mạng vì phải phân tích realtime trực tiếp trên các trang web. Hướng giải quyết phát triển của hệ thống sau này sẽ là tự động thu thập dữ liệu về database vào mỗi cuối ngày sau đó khi thống kê sẽ chỉ cần thống kê trên database có sẵn thì sẽ cho tốc độ cao hơn mà không phụ thuộc vào tốc độ mạng. Em rất mong sẽ nhận được những ý kiến đóng góp của thầy để đề tài của em được hoàn thiện hơn.