

자카드 유사도 기법을 이용한 채용 매칭 멘토링 시스템

장승훈^o, 최봉준^{*}

^o동서대학교 소프트웨어학과,

^{*}동서대학교 소프트웨어학과

e-mail: tmdgns5647@gmail.com^o, bongjun.choi@dongseo.ac.kr^{*}

Recruitment matching mentoring system using Jaccard Similarity

Seunghun Jang^o, Bong-Jun Choi^{*}

^oDept. of Software, Dongseo University,

^{*}Dept. of Software, Dongseo University

● 요약 ●

최근 국내 기업에서는 블라인드 테스트나 포트폴리오와 같은 자료를 활용하여 채용하는 추세이다. 지원자마다 개인의 역량이 다를 뿐만 아니라 기업에서 요구하는 기술/경험, 지원 자격, 특정 기술에 대한 경험을 요구한다. 따라서 본 논문에서는 국내 기업의 채용 공고에 기재된 지원 자격, 우대 기술, 우대 사항 등의 데이터와 지원자의 개인 역량(기술 스택, 전공 역량, 진행 프로젝트 등) 데이터를 활용하여 키워드를 추출한다. 지원자와 기업이 입력한 데이터를 통해 추출한 키워드들을 두 개의 집합으로 나눈 뒤 각각의 키워드를 할당한다. 할당받은 집합들을 비교하여 지원자의 정보가 기업의 채용 조건에 얼마나 부합하는지 계산한 후, 해당 확률을 지원자에게 제공하는 방식의 시스템이다.

키워드: 매칭(Matching), 멘토링(Mentoring), 자카드 유사도(Jaccard Similarity), 키워드(Keyword)

I. Introduction

최근 국내 기업의 취업률에 대한 문제는 뉴스와 같은 매체에서 주로 다루는 주제이다.

졸업 미루고 쉬는 학생 3년째 증가... “일자리 경험 늘려야”

Fig. 1. 졸업 미루고 쉬는 학생 3년째 증가... “일자리 경험 늘려야”, 동아일보

[그림1]에서 이들은 졸업 학점만 수료하고 졸업은 하지 않았거나 휴학한 상태에서 다른 활동도 하지 않고 있는 학생들이다. 한 취업 포털에서 취업에 실패한 지원자 589명을 대상으로 ‘스스로 생각하는 취업 실패 요인’에 대해 조사한 결과, 가장 많은 응답률을 차지한 24.6%가 ‘미비한 업무경력’을 원인으로 꼽았다. 이러한 뉴스 기사와 같이 최근 기업에서는 지원자의 역량을 먼저 확인하는 추세이다. 이 말은 신입 사원으로 지원한 경우에도 지원자마다 개별적인 역량이 다르며 기업마다 요구하는 기술과 우대 사항이 다르기 때문에, 타 지원자보다 경험이 부족한 상황에서도 우대하는 기술을 보유하고 있다면 합격한 사례가 있는 등, 기업의 우대 사항에 따른 지원자의

합격률은 천차만별이다. [1]

따라서 본 논문에서는 기업의 채용 공고에서 요구 기술, 지원 자격, 우대 사항과 같은 조건들을 키워드 형식으로 추출한 후, 추출된 키워드를 기업, 지원자 두 개의 집합으로 분할하고, 각각의 키워드를 할당하여 매칭한다. 이때, 조건에 부합 정도를 자카드 유사도(Jaccard Similarity) 기법을 활용하여 확률로 계산한다. 본 논문에서는 자카드 유사도 기법을 활용하여 매칭된 결과를 수치화하고, 이를 멘토링 방식의 시스템에 활용한다. [2]

II. Preliminaries

1. Related works

기업의 채용 조건과 사용자의 정보를 키워드 형식으로 추출한 뒤 비교하여 채용 조건에 부합하는 정도를 확률을 이용하여 수치화한다. 기업별 채용 조건을 [Table 1]과 같이 디셔너리 형태로 저장한다. 요구 기술, 지원 자격, 우대 사항과 같은 조건을 key 값으로 저장하고,

해당 조건의 기술 또는 경험 등을 value 값으로 저장한다. 기존의 키워드를 추출하는 알고리즘인 TextRank, TF-IDF 등은 키워드를 우선순위에 따라 할당하는 방식이지만, 본 논문에서는 지원자가 입력한 정보를 기업의 채용 조건과 비교하여 수치화하는 방식을 사용한다. [3]

Table 1. Company Posting

| key | Value |
|-----------------|------------------------|
| Required skills | ["java", "Spring", ..] |
| Preferential | ["MySQL", "Kafka", ..] |
| Certificate | ["정보처리기사", ..] |

덕서너리에 저장된 기업의 채용 조건을 지원자가 입력한 정보와 비교하여 확률을 계산하기 위해서 자카드 유사도 기법을 활용한다.

$$J(A, B) = \frac{|A \cap B|}{|A \cup B|} = \frac{|A \cap B|}{|A| + |B| - |A \cap B|}$$

Fig. 2. 자카드 유사도 공식

자카드 유사도는 위 [그림 2]와 같이 교집합을 합집합으로 나눈 것이며, 여기서 합집합은 기업의 채용 조건이고, 교집합은 기업과 지원자의 일치하는 키워드이다. 위 공식을 활용하여 두 집합을 비교하였을 때 도출되는 확률(%)로 계산한다.

III. The Proposed Scheme

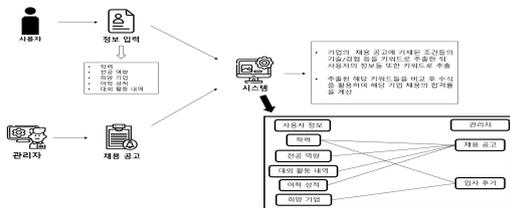


Fig. 3. 채용 매칭 시스템 개념도

기업의 채용 공고를 조건(지원 자격/요구 기술, 우대 사항 등)을 저장하고 지원자는 본 시스템 사용 시 정보(학력, 역량, 기술 스택 등)를 입력하면 시스템은 기업의 채용 공고 조건, 사용자의 정보를 키워드 형식으로 추출하고 각각 두 개의 집합으로 할당하여 비교 후 사용자의 정보가 기업의 채용 조건에 부합하는 정도를 자카드 유사도 기법을 이용하여 수치화한 확률을 지원자에게 제공한다.

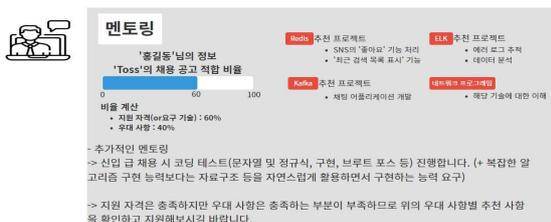


Fig. 4. 지원자와 기업 채용 조건의 적합 비율

IV. Conclusions

본 논문에서는 자카드 유사도 기법을 이용하여 지원자가 개인 정보와 기업의 채용 조건을 키워드로 추출하고 두 집합으로 할당 후 비교하여 해당 지원자의 정보가 해당 기업의 채용 확률을 결정값으로 생성되는 내용에 대해 기술하였다. 본 시스템을 통해 지원자는 기업이 원하는 인재상을 보다 쉽게 분석할 수 있으며, 해당 기업의 채용 조건에 미흡한 부분을 보완하여 채용 확률을 높일 수 있을 것으로 기대된다.

ACKNOWLEDGEMENT

본 연구는 2023년 과학기술정보통신부 및 정보통신기획평가원의 SW중심대학사업의 연구결과로 수행되었음 (2019-0-01817)

REFERENCES

- [1] Lee Mige, "The number of students taking a break after postponing graduation has increased for the third consecutive year..." "I need to increase my job experience" DongA, <https://www.donga.com/news/Society/article/all/20>
- [2] Kyoung-Min Kim, Joong-Jo Park, Buhm Lee, Young-Jin Go, Soon-Won Jung. "Efficient 1:N Fingerprint Matching Algorithm using Matching Score Distribution" Journal of Institute of Control 2018(3):208-217
- [3] Lee Han-Dong, "A Keyword Extraciton Algorithm Considering Association Between Words" Soongsil University Soft Specialization Graduate School, 2018