Analysis of the Combination of Enterprise Recruitment and College Student Employment Based on Data Visualization

Wayuan Xiao¹,*

¹Guangdong University of Technology School of International Education, Guangzhou City, Guangdong Province, 510000, China
*Corresponding author: mengwei@ldy.edu.rs

Abstract:
The purpose of this paper is to explore how to combine the recruitment needs of enterprises with the employment of college students through data visualization technology, and to provide new insights for addressing the employment matching challenge. First of all, it will introduce the current situation and challenges of enterprise recruitment and college student employment, and analyze the problems and reasons for this. Secondly, a method based on data visualization will be proposed. By collecting and analyzing relevant data from the Boss direct employment platform, including educational requirements, salary, work experience, occupation type, and other information. Visualization tools will be used to present data, and these tools will also demonstrate the recruitment needs of enterprises. Then, it will explore how to make targeted recommendations and provide career guidance based on the results of data analysis to promote a better match between enterprises and college students. Finally, the research results will be summarized, and the future research direction and development trend will be prospected.

Keywords: Data visualization; online recruitment; college student employment.

1. Introduction
Every graduation season, college students are faced with the problem of employment. On this issue, big data plays a significant role in the analysis of college students’ employment status and employment guidance, as well as the analysis of the current situation and trend of enterprise recruitment [1]. It also plays a crucial role in addressing the nexus between enterprise recruitment and college students’ employment. This paper relies on the Internet platform, uses big data technology and machine learning models, takes Boss recruitment website as an example, analyzes the information on the recruitment website, and provides guidance and help for college students’ employment.

2. Analysis of Current Situation of Enterprise Recruitment and College Students Employment

2.1 Enterprise Recruitment Demand Analysis
Online recruitment has become an efficient recruitment channel, particularly suitable given the current thriving context of Internet development. In the face of a growing number of college students, constantly changing employment demands, and increasingly stringent recruitment requirements from enterprises, online recruitment from enterprises, online recruitment holds great significance.

2.2 Analysis of Employment Status of College Students
According to the National Conference on Employment and Entrepreneurship for College Graduates, the number of college graduates in 2024 has increased by 210,000 compared to the previous year, and is expected to reach 11.79 million [2]. With the increasing number of graduates, the difficulty of finding jobs for them has also increased. The challenging employment situation for college graduates has become an undisputed problem [3].

2.3 Matching Problem Analysis
Currently, the number of graduates seeking jobs is increasing year by year, and yet the demand for suitable candidates among enterprises remains strong. Despite this, there are still graduates who struggle to find suitable employment, while enterprises find it challenging to attract qualified candidates [4]. This reflects the problem of poor information and poor matching between enterprise recruitment and college students’ employment [4].

3. The Application of Data Visualization in it

3.1 Data Acquisition and Processing
This paper aims to utilize web crawler technology to gath-
er recruitment data from Boss direct recruitment network, analyze the collected data from enterprises, and subsequently use these insights to assist and advise college students in their job search and employment process. The ultimate goal is to address the mismatch between corporate recruitment needs and college students’ employment expectations.

Firstly, basic information such as work location, occupation type, salary, educational requirements, and work experience was extracted from the Boss direct employment web page, and the extracted data was then saved into a CSV file. Table 1 below illustrates the obtained results.

### Table 1. Extract data result graph.

<table>
<thead>
<tr>
<th>type of jobs</th>
<th>name</th>
<th>salary</th>
<th>work experience</th>
<th>Academic requirements</th>
<th>Place</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>Education and training</td>
<td></td>
<td></td>
<td></td>
<td>Shenzhen Guangming District-Gongming</td>
</tr>
<tr>
<td>1</td>
<td>Electronics/Electrical/Communications</td>
<td>Verification Engineer</td>
<td>35-40K</td>
<td>3-5 years</td>
<td>master</td>
</tr>
<tr>
<td>2</td>
<td>Electronics/Electrical/Communications</td>
<td>Solutions Engineer DE</td>
<td>20-40K</td>
<td>3-5 years</td>
<td>Undergraduate</td>
</tr>
<tr>
<td>3</td>
<td>InternetAI</td>
<td>Python senior development engineer</td>
<td>12-24K</td>
<td>3-5 years</td>
<td>Undergraduate</td>
</tr>
<tr>
<td>4</td>
<td>Electronics/Electrical/Communications</td>
<td>S/PI Engineer</td>
<td>35-60K</td>
<td>5-10 years</td>
<td>master</td>
</tr>
<tr>
<td>2783</td>
<td>medical health</td>
<td>Urgently hiring nurses</td>
<td>5-8K</td>
<td>1-3 years</td>
<td>Technical secondary school/technical school</td>
</tr>
<tr>
<td>2784</td>
<td>InternetAI</td>
<td>Java</td>
<td>12-20K</td>
<td>5-10 years</td>
<td>Undergraduate</td>
</tr>
<tr>
<td>2785</td>
<td>InternetAI</td>
<td>Software Development Engineer-Backend Direction</td>
<td>6-10K</td>
<td>3-5 years</td>
<td>Undergraduate</td>
</tr>
<tr>
<td>2786</td>
<td>Hotel/Travel</td>
<td>Hotel front desk receptionist</td>
<td>4-6K</td>
<td>No experience limit</td>
<td>Technical secondary school/technical school</td>
</tr>
<tr>
<td>2787</td>
<td>Electronics/Electrical/Communications</td>
<td>Power supply qualification engineer</td>
<td>4-7K</td>
<td>1-3 years</td>
<td>Undergraduate</td>
</tr>
</tbody>
</table>

3.2 Visualization Results and Analysis

In this experiment, pandas and matplotlib visualization tools were utilized to visualize the extracted CSV files.

#### 3.2.1 Classification of educational requirements

By classifying the educational requirements in the dataset, a bar chart can be generated, as shown in Figure 1.

![Fig 1. Classification of educational requirements](image)

As can be seen from Figure 1, most enterprises require a bachelor’s degree for job applicants. On the one hand, this reflects the increase in the college entrance rate. On the other hand, it also indicates the ample employment opportunities available for graduates who do not intend to pursue further studies.

#### 3.2.2 Work experience classification

Work experience is also a requirement that enterprises pay attention to when recruiting talents. Figure 2 illustrates the classification of work experience requirements for enterprise recruitment in the extracted dataset.

![Fig. 2 Work experience classification](image)
As can be seen from Figure 2, enterprises have diverse requirements regarding job seekers’ work experience, with those requiring 1-3 years of work experience being the most common. Therefore, college students need to be attentive to whether the recruiter has a demand for work experience and its specific requirements when submitting their resumes. Moreover, it also reminds college students that they can participate in campus recruitment or engage in part-time jobs to enrich their work experience and resume, which will be beneficial for job hunting in the future.

3.2.3 Salary classification

Salary is not only an important factor for enterprises to consider during recruitment, but it is also a crucial consideration for contemporary college students when choosing a job. Figure 3 illustrates the result of categorizing the salaries offered by the firms in this dataset.

3.2.4 Occupational type classification

The hiring needs for different job types vary. By classifying the different occupational types in the data collected in this experiment, the bar chart shown in Figure 4 can be obtained.

Fig 3. Salary classification.

With the rapid development of science and technology, the demand for positions in the Internet, electronic, and electrical communication industries is very strong. These two industries also continue to release more employment opportunities. Therefore, they have important guiding significance for college students’ major selection or ability cultivation and improvement.

3.3 Model Analysis

In this experiment, the data from the dataset was obtained, with “salary” serving as the target variable. Five machine learning models, namely the random forest model, linear regression model, support vector machine model, logistic regression model, and decision tree model, were utilized to analyze and process the obtained data.

3.3.1 Random Forest Model

Figure 5 depicts the results of data analysis utilizing the random forest model.

Mean Squared Error: 26.712801682037927

Fig 5. The result of the random forest model.

Compared to the other four models, the performance of the random forest model is better as it is capable of capturing complex relationships in the data. Additionally, the model processes the data with high accuracy and can effectively handle issues such as nonlinear relationships, missing data, and outliers. Furthermore, during the feature engineering stage, this experiment conducted various processing steps on the dataset. These included converting the target variable column of “salary” into numerical data and eliminating rows containing missing values. These preprocessing steps aimed to provide improved training data for the models. The following is a preliminary analysis of the dataset structure: Various factors such as occu-
pational types, job positions, work experience, education level, cost of living, and economic development in different regions can influence salary levels. Figure 6 depicts the results of the feature projection.

```python
Index(['type of jobs', 'name', 'salary', 'work experience', 'academic requirements', 'place'], dtype='object')
Mean Squared Error: 26.71280169820307027
```

Fig 6. The result of the feature projection.

Through the interpretation of the model, a feature importance analysis was conducted using the random forest model. Subsequently, the degree of influence of different feature objects on the target variable, “salary,” was obtained. The results indicate that “Internet server development engineer” and “analog integrated circuit expert” have a significant impact on salary levels. This reflects the broad employment prospects within the Internet and electronic and electrical communication industries. Simultaneously, the findings obtained from the feature importance analysis can guide further data analysis and feature engineering, thereby contributing to the enhancement and optimization of prediction models.

3.3.2 Other machine learning models

Compared to the random forest model, the linear regression model, support vector machine model, logistic regression model, and decision tree model performed poorly in this experiment. This indicates the presence of numerous outliers and incomplete situations within the dataset, necessitating further improvement and updates in subsequent experiments.

4. Recruitment and Employment Matching Analysis Based on Data Visualization

4.1 Recruitment Needs Analysis

The data visualization results obtained from the experiment can reflect the varying recruitment needs of enterprises on the online recruitment platform, enabling job seekers to choose their favorite positions based on their needs and deliver resumes accordingly. Taking the Boss direct employment platform as an example, an online recruitment platform offers various options such as “location,” “occupation type,” “salary,” “work experience,” and “educational requirements” to intuitively demonstrate the recruitment needs of job seekers. Based on the above experimental results, most positions require applicants to possess a bachelor’s degree or higher and have some work experience.

4.2 Analysis of Employment Hot Spots and Trends

Based on the data visualization analysis results of this experiment, it is evident that the current employment hot spots are concentrated in the internet industry and the electronic and electrical communication industry. These two industries exhibit a larger and growing demand for jobs compared to other industries. This indicates that college students need to prioritize their own skill development to equip themselves with the necessary skills and knowledge to adapt to and keep up with the latest industry trends and technologies. In addition, college students should actively participate in internships and practical opportunities to enrich their work experience, thereby increasing their chances of having more opportunities to choose from when it comes to job positions.

5. Conclusion

5.1 Summary of Research Results

This paper employs web crawler technology to retrieve a subset of recruitment data from the Boss direct employment website. It then utilizes data visualization techniques and machine learning models to analyze the obtained dataset, aiming to derive insights into different occupational types, salaries, work experience requirements, and educational qualifications found on the online recruitment platform. Furthermore, it investigates the influence of each factor on salary levels. The purpose of this paper is to visually display and analyze the job demands of enterprises on an online recruitment website through the use of web crawler technology. By doing so, we aim to gain a comprehensive understanding of the current situation and trends in enterprise recruitment. This information can provide valuable guidance and assistance to college students seeking employment, helping them to address the issue of information asymmetry in the job search process. Additionally, we hope to facilitate a more effective matching between enterprise recruitment needs and the employment aspirations of college students [5].

5.2 Outlook

There are still many shortcomings in this experiment. For example, the amount of data in the dataset is not large enough and not comprehensive enough to make good use of machine learning models such as decision trees to obtain better results. The experimental results would be more convincing if the amount of extracted data were increased and a more comprehensive analysis was conducted. In addition to analyzing online recruitment data, it can also collect information on the employment needs of college students, thus better achieving the goal of matching the recruitment needs of enterprises with the employment aspirations of job-seeking college students, and providing greater assistance to the future recruitment and job market.
References