

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

Wayuan Xiao<sup>1,\*</sup>

<sup>1</sup>Guangdong 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.**

	type of jobs	name	salary	work experience	Academic requirements	Place
0	Education and training	Primary school teachers of various subjects	6-11K	No experience limit	College	Shenzhen·Guangming District·Gongming
1	Electronics/Electrical/Communications	Verification Engineer	35-65K 15 salary	3-5 years	master	Shanghai·Jing'an District·Pengpu Town
2	Electronics/Electrical/Communications	Solutions Engineer-DE	20-40K 15 salary	No experience limit	Undergraduate	Shanghai·Pudong New Area·Zhangjiang
3	Internet/AI	Python senior development engineer	12-24K 13 salary	3-5 years	Undergraduate	Shanghai·Pudong New Area·World Expo
4	Electronics/Electrical/Communications	SI/PI engineer	35-60K	5-10 years	master	Shanghai·Pudong New Area·Zhangjiang
...	...	...	...	...	...	...
2783	medical health	Urgently hiring nurses	5-6K	1-3 years	Technical secondary school/technical school	Xiamen·Huli District·Jinshang
2784	Internet/AI	Java	12-20K	5-10 years	Undergraduate	Chongqing·Yubei District·Renhe
2785	Internet/AI	Software Development Engineer-Backend Direction	6-10K	No experience limit	Undergraduate	Shanghai
2786	Hotel/Travel	Hotel front desk receptionist	4-6K	No experience limit	Technical secondary school/technical school	Guangzhou·Yuexiu District·Zhongshan 6th Road
2787	Electronics/Electrical/Communications	Power supply qualification engineer	4-7K	1-3 years	Undergraduate	Beijing·Changping District·Shahe

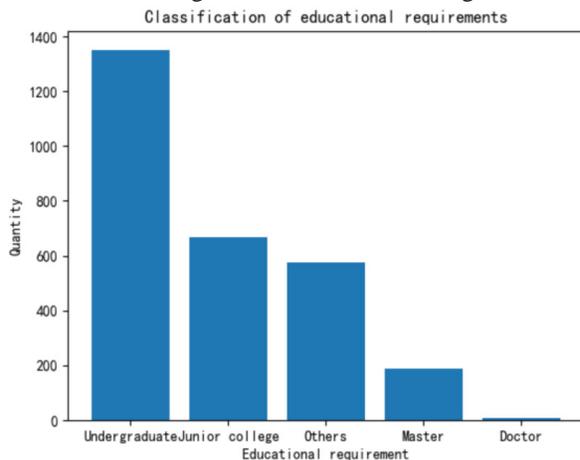
2788 rows × 6 columns

### 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.



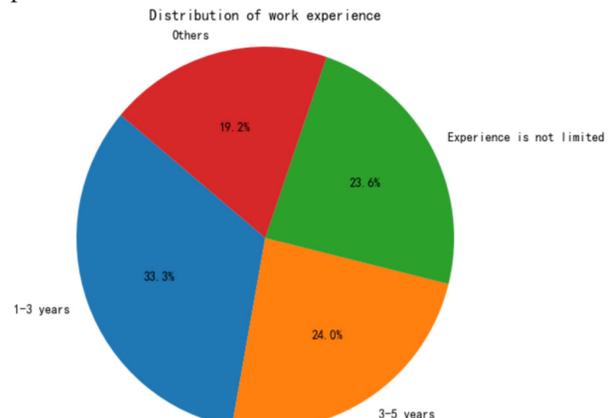
**Fig 1. Classification of educational requirements**

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.

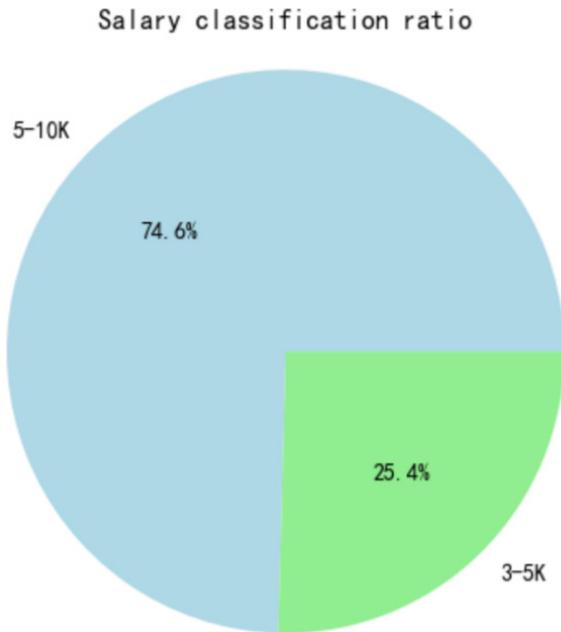


**Fig. 2 Work experience classification.**

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.

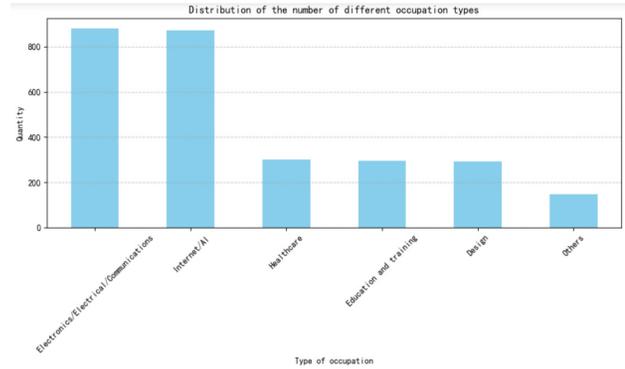


**Fig 3. Salary classification.**

As can be seen from Figure 3, more than half of the enterprises offer a monthly salary ranging from 5,000 to 10,000 yuan. According to the report "Employment Trends and Prospects for College Graduates in China in 2023," the average monthly salary for graduates in 2023 will be 10,342 yuan. However, as reflected in current online trends, only a small percentage of graduates are able to earn high salaries upon graduation. This means that, in addition to paying attention to the salary offered by the enterprise, college students also need to consider the impact of the industry market and market demand on salary levels. At the same time, job seekers also need to improve their work ability, so that their abilities align with the salary level of the enterprise.

### 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 4. Occupational type 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 preprocessing 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.

```
Index(['type of jobs', 'name', 'salary', 'work experience',  
      'Academic requirements', 'Place'],  
      dtype='object')  
Mean Squared Error: 26.712801682037927
```

### 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

- [1] Wenjie Song. Data Visualization of Employment Guidance for College Students Based on Python in Big Data Environment[J]. *Microcomputer application*, 2023, 39 (12): 127-129+134.
- [2] Jining Guo. Analysis on the Problems and Countermeasures of College Students' Employment in the New Era[J]. *Journal of Jiamusi Vocational College*, 2024, 40 (03): 112-114.
- [3] Hui Liu. Research on Influencing Factors of College Students' Employment[J]. *Cooperative economy and technology*, 2024, (17): 54-56.
- [4] Mingliang Fu, Gang Li. Research on Matching the needs of Job Seekers with the needs of Enterprises based on Big Data Analysis of Recruitment Websites[J]. *Statistics and management*, 2021, 36 (08): 91-97.
- [5] Jin Guo. Research on Recruitment Data Crawling and Data Visualization Analysis Based on Python[J]. *Light industry technology*, 2024, 40 (02): 94-96+99.
- [6] Wang, Lili and Chen, Wei and Wang, Haoliang. Data Visualization Analysis for Job Recruitment Based on Big Data. 2023. Association for Computing Machinery.
- [7] Yuanyuan Chen, Ruijie Pan, "Research on Data Analysis and Visualization of Recruitment Positions Based on Text Mining", *Advances in Multimedia*, vol.2022, Article ID 9047202, 11 pages, 2022.
- [8] Chenxiang Zhang, Lu Ning, Mingge Feng. Design and Implementation of University Employment Data Visualization Management System[J]. *Computer programming skills and maintenance*, 2022, (09): 74-77.
- [9] Feihong Tang. Visual Analysis of Recruitment Information Data Based on Python crawler[J]. *Software*, 2023, 44 (01): 176-179.
- [10] Jingjing Shen, Shuyan Yu. Visual Analysis of Big Data Related Job Recruitment Information Based on 51job. *Proceedings of the 3<sup>rd</sup> International Conference on Internet, Education and Information Technology*, 2023, 10.2991/978-94-6463-230-9\_27.
- [11] Oleksii Romanko, Mart O'Mahony. The Use of Online Job Sites for Measuring Skills and Labour Market Trends: A Review. ESCoE Technical Report No. 2022-19, May 2022.
- [12] Malik Nabeel Ahmed Awan, Sharifullah Khan, Khalid Latif and Asad Masood Khattak, A New Approach to Information Extraction in User-Centric E-Recruitment Systems. *Section Computing and Artificial Intelligence, Appl. Sci.* 2019, 9(14), 2852.
- [13] Ishraq Hatif Abd Almajed, Ghalia Nassreddine and Joumana Younis. New Recruitment Approach Based on Logistic Regression Model. *Mesopotamian journal of Computer Science*, 2024, Vol.2024, pp.15-27.
- [14] D.F.Mujtaba and N.R.Mahapatra. "Mining and Analyzing Occupational Characteristics from Job Postings". 2020 International Conference on Computational Science and Computational Intelligence (CSCI), Las Vegas, NV, USA, 2020, pp.683-685.
- [15] Jingting Wan, Fuqing Wang, Zhaoling Li and Hao Zhang. Big Data Crawling and Mining Based on Internet Recruitment Website. *International Journal of Frontiers in Engineering Technology*, 2022, ISSN 2706-655X Vol.2, Issue 3: 36-47.
- [16] Wenlong Ning, Hongxia Mao. 51job Website Content Crawling Based on Python Crawler Technology. *Information and Computers*, 2021, 33(04):180-182.
- [17] Zhimin Gu, Mingzhu Wu. Design and Implementation of Web Crawler Based on Python[J]. *Computer programming technology and maintenance*, 2023(09): 163-166.
- [18] Kaili Yand, Meijuan Shan. Python Based Data Visualization[J]. *Modern information technology*, 2019, 3(05): 30-31+34.
- [19] Haixin Zhang, Yifang Zhang, Zhilan Xie, et al. Construction of a Risk Assessment Model for Mild Cognitive Impairment in Elderly Diabetic Patients Based on Machine Learning[J]. *Chinese Journal of Disease Control*, 2024, 28(03): 284+309.
- [20] Jin Zhang, Yanyan Fu, Xuewei He, et al. Research on the Application of Machine Learning Random Forest Model in Fiscal Revenue Forecasting[J]. *Western finance and Accounting*, 2024(01): 7-10.
- [21] Chuang Zhou, Xueqin Jin, Zhengli Guo, et al. Comparison of Cognitive Frailty Risk Prediction Models of Elderly People in Community Based on Different Machine Learning Algorithms[J]. *Journal of nursing*, 2023, 38(19): 1-5+11.