

Research and Practice on the Curriculum System Design of Applied Talents Training in Colleges and Universities with Market Demand Under Big Data Technology

Xiaoling Wu, Jinxian Zhang, Liuchang Zhang

School of Information Technology and Engineering, Guangzhou College of Commerce,
Guangzhou, Guangdong, 511363, China

Abstract

Based on the operability, scientificity, applicability and other core issues of big data technology in the design of applied talents training scheme in Colleges and universities, this paper expounds the overall idea and practice of how to use big data application technology to design the curriculum system of applied talents. This method has been applied to the design practice of the course system of big data analysis and application direction of information management and information system specialty of Guangzhou Business University for two consecutive years to further verify its effectiveness and applicability.

Keywords

Applied Talents in Colleges and universities; Curriculum system design practice; Big data technology; Market oriented.

1. Research Background and Significance

The project is based on the research results of "University applied talent training curriculum system for enterprise demand under big data technology", and based on the original two major issues of "data collection easy to operate" and "practical effectiveness and applicability of the curriculum system", the research and improvement are carried out respectively.^[1]

First of all, aiming at the problem of "easy operation of data collection", the project developed and designed the online recruitment data collection program scratch, and analyzed the collected online recruitment data, including the analysis of common problems in curriculum setting such as enterprise talent demand forecast, industry development trends, industry knowledge fields, job categories, core knowledge fields, etc., analyzed the latest demand of enterprises for big data technology talents, and provided a perspective on how to timely connect the training direction of applied talents with social needs.^[2]

Then, based on the existing research results, that is, the method of corresponding conversion of each industry processing process, ability module and course module, the project has been applied to the design of the course system of business data analysis for the specialty of information management and information system for two consecutive years, and the relevant course system has been applied to teaching practice, so as to realize the test of its scientificity, rationality and applicability.

The construction and completion of this project provides an innovative idea for the discussion of talent cultivation in the field of higher education.^[3-5] The project has shown positive effects in technological innovation, methodological improvement, effectiveness verification and social impact.

2. Research Objectives

This project focuses on solving the following problems:

- (1) . Design data acquisition scheme, develop and implement data acquisition program. The original data acquisition method is adjusted and optimized, and a more convenient and efficient acquisition scheme and tool are designed.
- (2) . This paper analyzes the latest demand of enterprises for big data technology talents, and provides a perspective on how to timely connect the training direction of application-oriented talents with social needs. From the perspective of job seekers' employment concerns, this paper analyzes the enterprise talent demand in terms of enterprise demand profile, hot demand, demand trend, and important related factors in job selection.
- (3) . Based on the existing research results of "professional ability course structure derivation process", the "ability course" structure correspondence of information management and information system professional big data analysis and application direction is generated.
- (4) . The above curriculum system of "big data analysis and application direction of information management and information system specialty" is applied to teaching practice to prove its scientificity, rationality and applicability.

3. Research Contents

3.1. Design and implementation of crawling program for recruitment data collection

- (1) . Determine target data

The recruitment data of this project mainly comes from the direct recruitment of boss, a well-known domestic recruitment website. According to the project objectives, it is necessary to collect the basic information of each recruitment position on the website, including 9 items of data, including "position name", "position salary", "position city", "work experience", "education requirements", "position welfare supplement", "position label", "position content description" and "position detailed address". These information needs to be collected by the acquisition program to access the details page of each position.

- (2) . Analysis of website key parameters to be determined by the collection program

Parameter analysis of different pages of the retrieval list. This scheme needs to resolve the change of the paging parameter page in the URL and locate the location of the collected text element in the website.

Determination of the maximum number of pages in a specific retrieval condition. Considering that the collection quantity may be lower than that provided by the website, it is also necessary to obtain the information of the maximum number of pages under specific retrieval conditions. This information can be obtained through the page number text information in the pagination component at the bottom of the website.

Determine the URL links of all detail pages under each page. Using the developer tool, observe the location of elements and the commonness of link positioning of all details pages at present. It is found that they can be located through class=job card left to find the links of details pages. How to locate the information elements of the detail page in the website. Click F12 to view and locate the tag values corresponding to "position name", "position salary", "position city", "work experience", "education requirements", "position welfare supplement", "position tag", "position content description", "position detailed address" and other data.

- (3) . Overall scheme of data acquisition design

Design of data acquisition process. When collecting the website, the project uses the chromedriver driver and selenium Library in Python language to realize the automatic operation of the website.

At the same time, the chromedriver_autoinstaller library is used to realize the automatic download of the driver. By setting the selenium configuration item, the camouflage of the request state, the selection of page loading strategy, and the optimization of the collection behavior are completed to improve the stability of the collection process.

The realization of data acquisition program. After determining the implementation process and method of data collection, the final network data collection tool scratch is packaged and encapsulated into an EXE executable file.

Use scratch to collect recruitment data for data collection. Finally, the project collected 12343 recruitment data using the scratch program for the subsequent industry talent demand analysis of the project.

3.2. Analysis of current talent demand in big data industry

(1) . Workplace distribution analysis

It can be seen from Figure 1 that the job recruitment demand of big data is mainly distributed in Beijing, Shanghai, Shenzhen, Guangzhou, Chengdu, Hangzhou and other first and second tier cities, where there are more employment opportunities.

When guiding students' employment, colleges and universities should encourage them to pay attention to the job information in these areas while taking into account the cost of living and career development potential.



Figure 1. Workplace word cloud

(2) . Analysis of salary level

As shown in Figure 2, the salary level of the big data industry is about 16800 yuan per month for many jobs. On the whole, the monthly salary is mainly between 10000 and 20000 yuan. Job seekers should reasonably set salary expectations according to their own abilities and market conditions, while colleges and universities should guide students to set the expected salary to match the actual market conditions in the curriculum and career guidance.

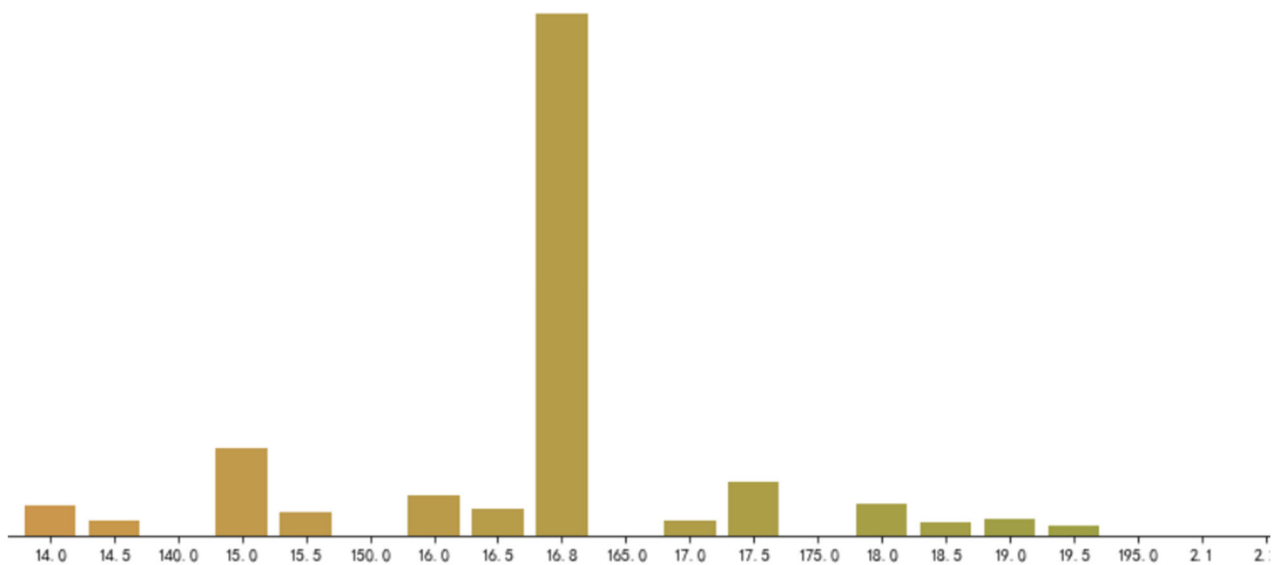


Figure 2. Salary distribution histogram

(3) . Job experience analysis

Enterprises generally pay attention to the work experience of job seekers when recruiting. Figure 3 shows that the proportion of posts without experience requirements is relatively low. About 25% of the jobs require applicants to have 1-3 years' work experience, and about 27% of the jobs require applicants to have at least 3-5 years' work experience. On the whole, about 70% of the recruitment positions require candidates to have work experience. When setting up talent training programs, colleges and universities should pay attention to the cultivation of students' practical ability, and guide students to accumulate relevant experience in many ways, such as internship, part-time work, project participation and so on.

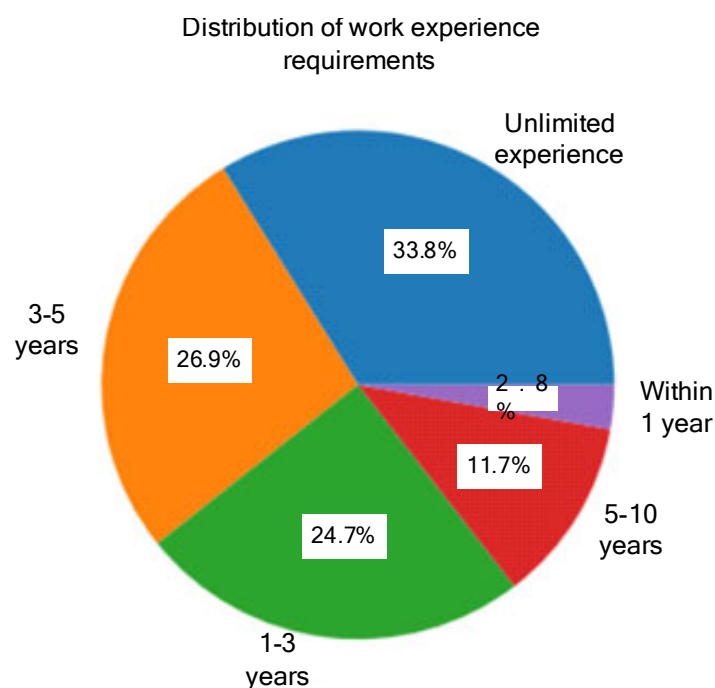


Figure 3. Analysis of work experience distribution

(4) . Analysis of academic requirements

From the bar chart 4 of the distribution of recruitment qualifications, most jobs tend to recruit candidates with bachelor's degrees, but some jobs have loose requirements for qualifications. Colleges and universities should adjust the teaching content according to the market demand, and guide students to reasonably plan the path of improving their academic qualifications. Job seekers can consider continuing education or professional certification to improve their academic background to meet job requirements.

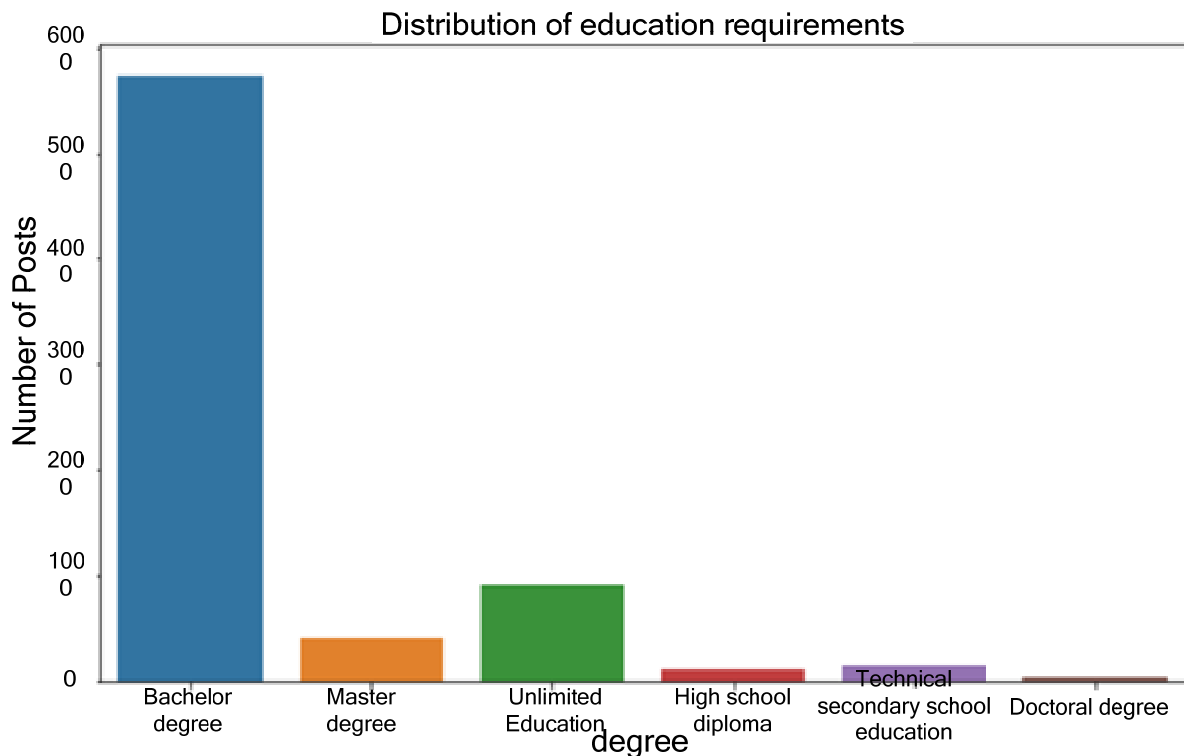


Figure 4. Bar chart of education requirements

(5) . Analysis of core skill requirements

As shown in Figure 5, the core skill of the post is the core goal of the talent training program in Colleges and universities. Analyzing the core skills of the post is to help colleges and universities clarify the key skills of cultivating talents. Mastering mainstream programming languages such as Java and Python is crucial for big data talents. In addition, data governance, data analysis, SQL and data statistics are also the core competitiveness of big data application talents. Colleges and universities should strengthen the teaching of these skills to improve students' employment competitiveness. Job seekers can also strengthen the learning and practice of these skills through online courses, workshops and other ways.

(6) . Welfare analysis

In addition to paying attention to the salary level, the welfare treatment of enterprises is an important factor to attract job seekers. According to the statistical results, it can be found that the most common forms of employee welfare are "five insurances and one fund", "paid leave", "holiday welfare" and "year-end bonus". As the legal basic welfare, the "five insurances and one fund" is the basis for enterprises to attract and retain employees. Job seekers should take the initiative to understand all the welfare items provided by the enterprise, including but not limited to "five insurances and one fund", "paid vacation", "holiday welfare" and "year-end bonus".

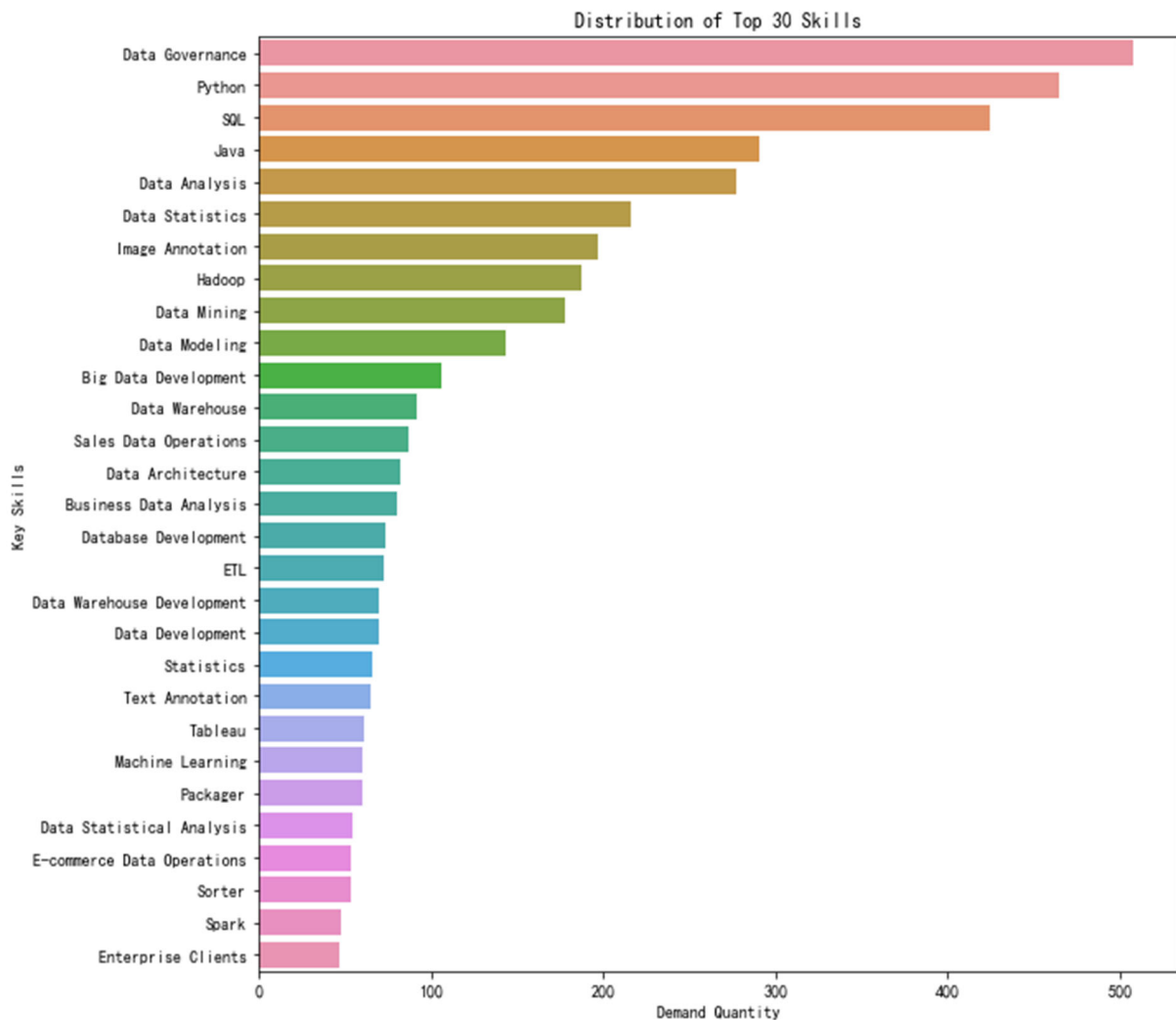


Figure 5. Bar Chart of Core Skills

(7) . Analysis of recruitment conditions most closely related to salary

Figure 6 shows the correlation coefficient between salary, experience and personnel size. The closer the correlation value is to 1, the stronger the positive correlation is, the closer the value is to -1, the stronger the negative correlation is, and the closer the value is to 0, the weaker or no correlation is.

From the correlation coefficient of the heat map, there is a certain positive correlation between salary and work experience, but the degree is low. There is little correlation between salary and personnel size. The correlation between experience and personnel size is also very weak.

It can be seen that when recruiting talents for big data application, enterprises pay little attention to the correlation between personnel size and salary, in addition to work experience, which indicates that enterprises prefer to determine salary according to job demand rather than enterprise size.

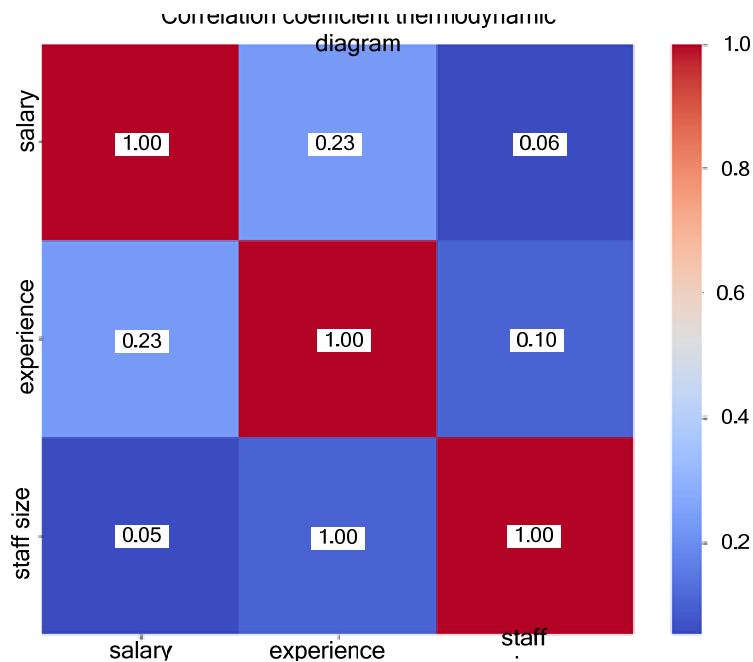


Figure 6. Correlation coefficient thermodynamic diagram

(8) . Correlation analysis of career factors

There is a certain correlation between enterprise recruitment requirements and related benefits. Analyzing these correlations can help universities and job seekers better understand the market demand and make decisions that are more in line with the market trend.

Given that the data of association mining is designed to provide decision support for the professional construction of colleges and universities and the employment of college students, considering that fresh college students usually lack work experience, tend to find jobs in economically developed regions, and have high expectations for starting salaries, this project selects rules that meet these specific requirements from the recruitment data of jobs related to big data. For example, for those job seekers seeking positions with a monthly salary ranging from 10000 to 15000 yuan, the calculation results show the following rules:

lhs	rhs	support	confidence	lift
1 {salaryRange=(10,15]}=>	{class= First tier cities }	0.215 320 48	0.7810 964	0.944 620 6
2 {salaryRange=(10,15]}=>	{Educarion= Bachelor degree }	0.211 881 19	0.768 620 0	1.053 257 5
3 {Education= Bachelor degree,				
salaryRange=(10,15]}=>	{class= First tier cities }	0.167 795 73	0.791 9331	0.957 726 0
4 {class= First tier cities,				
salaryRange=(10,15]}=>	{Education Bachelor degree }	0.167 795 73	0.779 283 6	1.067 870 1
5 {salaryRange=(10,15]}=>	{workYear=l-3S}	0.143 095 36	0.519 092 6	1.297 73 16
6 {workvear=l-3years,				
salaryRange=(10,15]}=>	{class= First tier cities }	0.115 372 59	0.806 263 7	0.975 056 7

For many fresh graduates, the monthly salary of 10000 to 15000 yuan is usually regarded as a high salary level. According to the above comprehensive analysis, if you want to obtain this salary level, job seekers may need to find job opportunities in first tier cities, and generally need to have a bachelor's degree or higher, as well as 1-3 years of relevant work experience. In addition, through similar correlation analysis, we can further determine the expectations of job seekers for the workplace, such as the trend of employment in first tier cities.

To sum up, enterprises' recruitment demand for big data talents shows a clear preference for workplace, salary level, work experience, education and core skills, and also attaches importance to the importance of welfare in attracting talents. In the recruitment process, these enterprises pay more attention to the practical skills and experience related to the position, not just the academic background.

3.3. Professional ability of "big data analysis and application" - Analysis and implementation of curriculum structure

(1) .Overview of professional competence curriculum structure transformation framework

The framework of the big data analysis and application course system, which is oriented by cultivating practical ability, is composed of three main lines, namely, business stage process chain, ability composition chain and course module chain. They constitute the three pillars of the curriculum system. The curriculum system design framework is designed according to the business process and "knowledge ability structure relationship". Each business stage corresponds to the knowledge structure requirements of the cognitive layer and the application layer. The construction of the curriculum system follows the corresponding relationship between the business process and the knowledge ability structure, as shown in Figure 7.

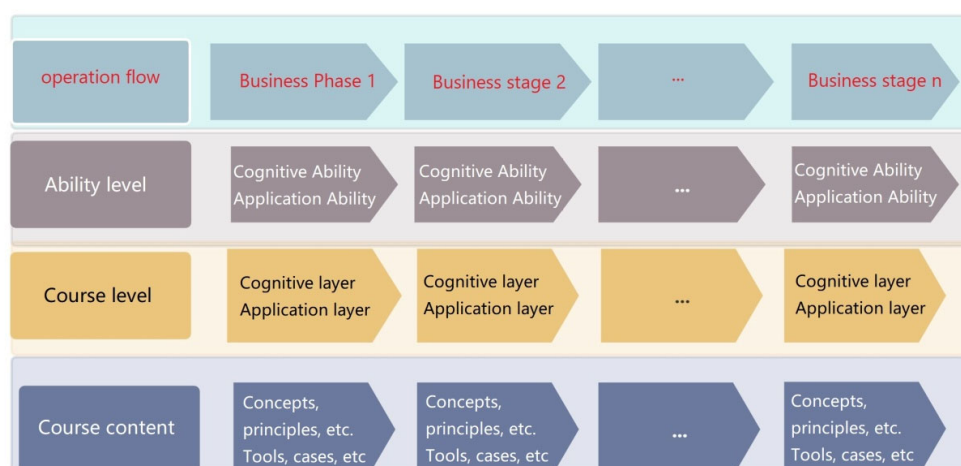


Figure 7. General idea of the framework design of big data analysis and application course system

The requirements for professional ability at each business stage include not only the cognitive ability of big data analysis and application theory, but also the practical operation ability. This requirement is reflected in the curriculum design, which is divided into two levels: cognitive level and practical level. At the cognitive level, the course aims to help students deeply understand and master the basic concepts and principles related to the big data application business stage. At the application level, the course pays more attention to the cultivation of students' practical operation skills and innovative application ability in the big data application business stage.

(1) . Introduction to the transformation process of "professional ability course structure"

The LDA topic model is used to explore the relationship between professional skills and curriculum structure based on the job description data including the professional skills information required by the position. LDA model regards each job description as a document, and each professional ability is equivalent to a topic, which is composed of a set of related vocabulary. The multiple topics calculated by the model represent the knowledge and ability required by the positions in the industry. Then, these topics are extracted and summarized

manually, and further classified into cognitive layer and application layer according to the hierarchical relationship between ability and curriculum.

To sum up, the conversion process of professional ability curriculum structure can be summarized as two main stages: "ability subject calculation" and "ability curriculum structure conversion". In the "ability subject calculation" stage, first classify the positions according to the business stage, and then use natural language processing technology to preprocess the position description of each position category, including removing numbers, punctuation, word segmentation, and stop words. Then, the LDA topic model is applied to mine the ability topic of the preprocessed text. In the process of word frequency statistics, the text related to "tools" is analyzed to identify popular technologies and provide reference for the selection of teaching tools.

In the "ability course structure transformation" stage, the LDA model is optimized through multiple parameter adjustments to obtain multiple ability topics with independent and clear topics. Then, these topics are extracted manually and classified according to cognitive ability and practical ability. Finally, according to the ability requirements of the business stage, they are divided into the cognitive layer and the application layer, and the corresponding courses are set according to the form of the course content (such as concept principles, case tools, etc.).

(3) . The transformation of talents' professional ability and curriculum structure in the direction of big data analysis and Application

(A) . Basic steps for LDA model to calculate "capability" topic

In the process of using LDA model to identify the topic of "ability", three core steps are mainly involved: the definition of job categories, the word frequency identification of key skills, and the operation of the model.

1) . Definition of job category.

As discussed earlier, when using big data technology to solve specific industry problems, the steps usually followed include setting business goals, collecting data, extracting, transforming and loading data, processing and modeling data, and data interpretation and application. Therefore, the course framework of big data analysis and application is shown in Figure 8. The "?" in it represents the part that has not been determined before using LDA model for subject calculation.

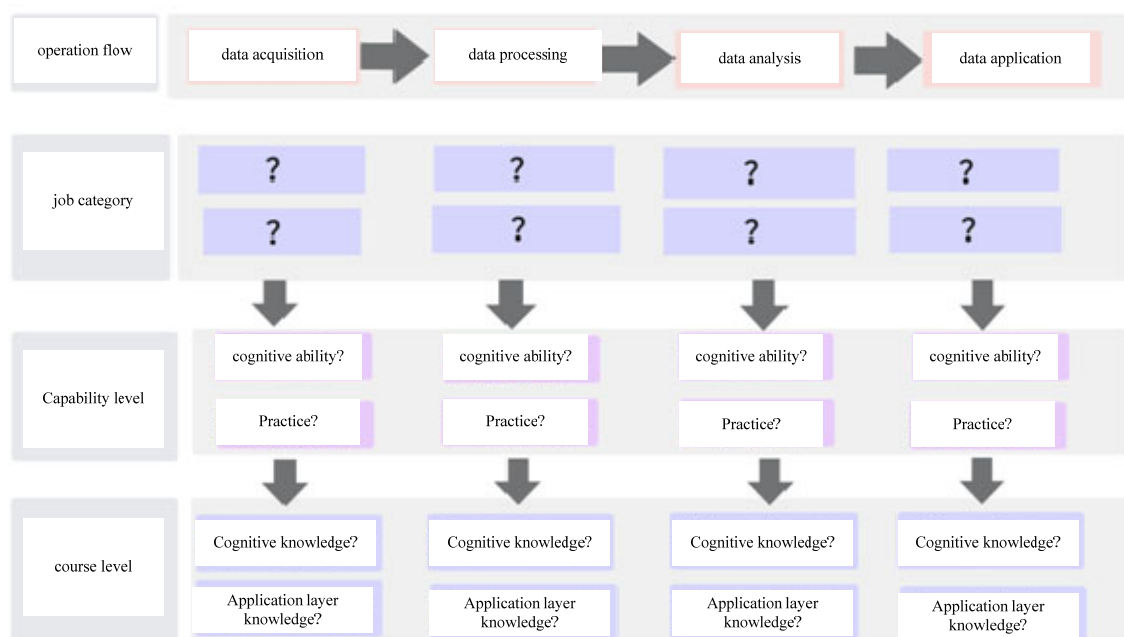


Figure 8. Course framework of big data analysis and Application

According to the data analysis process of professional ability and course structure, it is necessary to group recruitment positions according to the big data application process, and uniformly name those positions with different names but essentially the same, so as to realize the classification of positions. Through in-depth study of position name and corresponding position description, the classification method and classification details of positions can be found in Table 1.

Table 1. Job classification methods and classification results of big data

Job category	Mainly involved in data processing stage	Key words of position name	Examples of recruitment positions
Data Quality Analyst	Data cleaning, validation, quality control, metadata management	Quality, cleaning, verification, control, metadata	Data quality control engineer Data cleaning expert Data validation analyst Metadata Management Engineer Senior Data Quality Analyst Data visualization Development Engineer Senior Data Report Designer Interactive data visualization expert
Data visualization Engineer	Data visualization, report production, user interface design	Visualization, reporting, design, interface, user interaction	Data presentation architect User interface (UI) data visualization Engineer Data product manager Senior Data Product Manager Data driven product strategist Data product architect User behavior analysis product manager
Data product manager	Demand analysis, product design, iterative optimization	Product, management, strategy, user insight	

From the division results, the positions of data acquisition engineer, data platform/warehouse engineer and data algorithm engineer correspond to the data processing stages of data acquisition, data collection storage management, data modeling and data analysis respectively, while the positions of data visualization engineer and data product manager focus on the combination of data and business applications.

2) . Word frequency recognition of key skills.

Through text analysis and word segmentation of job descriptions of different job categories, key words can be identified, and the frequent occurrence of these words reveals the core skill needs of the job. The analysis of word cloud shows that the work of data Algorithm Engineers is closely related to many professional fields, including but not limited to machine learning, statistical modeling, data mining and optimization algorithms. In the use of technical tools, data Algorithm Engineers rely on a series of high-level programming languages and analysis tools, which help them build and optimize algorithm models. Through the extraction and frequency statistics of English terms of job description, it can be found that python, R, Scala and other languages, as well as tensorflow, keras, scikit learn and other machine learning frameworks are common technical tools for data algorithm engineers. They need to have a profound understanding of the data in a specific field, be able to use statistical and mathematical knowledge to analyze the data and find patterns and trends in the data. This key skill extraction method is also applicable to the analysis of core skills of other job categories.

3) . Model calculation.

According to the calculation procedure of LDA topic model, a corpus is created for each job category, and a matrix is established to associate the job description with vocabulary items. Then the LDA model is used for in-depth analysis to identify the ability topics related to a specific job category. Taking "data algorithm engineer" as an example, the subject output of the model represents the professional skill set necessary for the job category. These skill sets are defined by interrelated vocabulary groups, which need further in-depth analysis and optimization. This method helps to deepen the understanding of the professional skills and knowledge needs of various job categories.

(B) . Ability course structure transformation

1) . Theme extraction.

The calculation results of LDA topic model are further refined. The above six topics are: theoretical basis of machine learning; Master statistical knowledge; Familiar with programming language; Algorithm optimization experience; Understand model evaluation and validation methods; Ability to apply domain knowledge.

2) . Ability classification.

According to the hierarchical structure of abilities, among the five abilities, those that belong to cognitive abilities have the theoretical basis of machine learning, master statistical knowledge, and understand model evaluation and verification methods. The subjects required for practical ability are familiar with the use of programming language, algorithm optimization experience, and the application ability of knowledge in the field of data.

3) . Course design.

Combined with the divided cognitive ability and practical ability, the corresponding teaching content can be further designed to form the knowledge structure framework of the position talents. For example, after the ability course structure of the data acquisition engineer has been transformed, the results are shown in Figure 9.

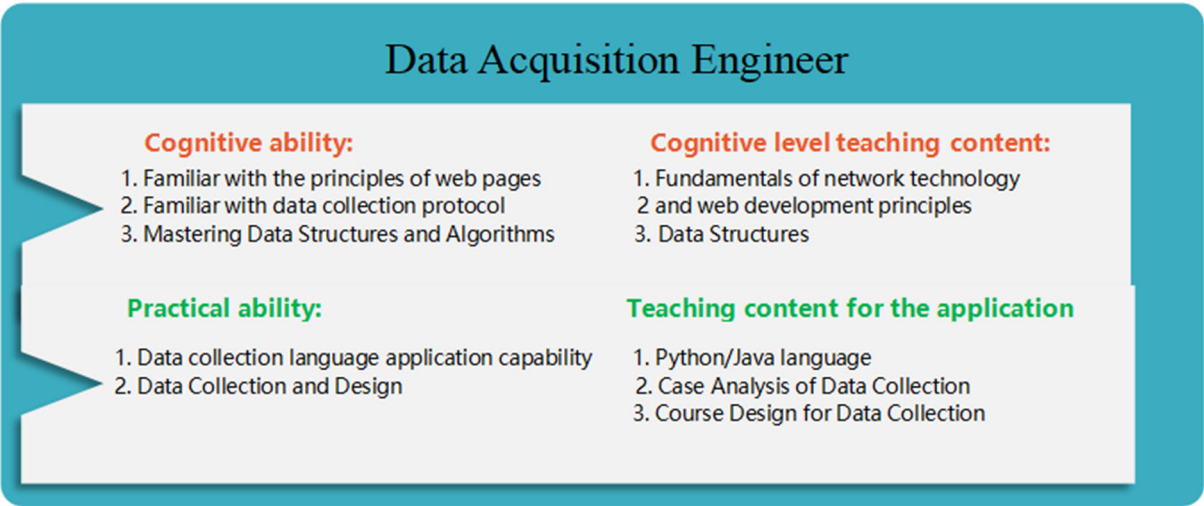


Figure 9. Data acquisition engineer "ability course structure" relationship conversion results

According to the above data acquisition engineer's ability course structure conversion method, the ability course structure information of other job categories is obtained, as shown in Table 2.

Table 2. Conversion results of "ability course structure" relationship of some big data application positions

Relevant business processes	Job category	Capability level	Course
Data display, information transmission, decision support	Data visualization Engineer	Cognitive ability - data visualization principle - Design thinking and user interface (UI) design principles - Able to understand and analyze customer business needs	Cognitive teaching content - Basic theory of data visualization - User interface design
		Practice - Ability to use data visualization tools - Capable of interactive chart development - Able to present the data analysis results in an intuitive and understandable manner	Teaching content of application layer - Data visualization technology (tableau, d3.js) - Practice of interactive data chart development - Business analysis report design
Data collection, data cleaning, data analysis, report writing, business insight	Data Analyst	Cognitive ability - Fundamentals of statistics and probability theory - Data analysis methodology - Business understanding and domain knowledge	Cognitive teaching content - statistics - Probability theory and mathematical statistics - machine learning
		Practice - Data processing and cleaning skills - Use of data analysis tools and programming languages (such as Excel, python, R) - Data visualization and report presentation	Teaching content of application layer - Fundamentals of data analysis and mining language (Python, R) - Database Principle and Application - Data visualization technology - Business analysis report design

According to the above methods, the major of information management and information system has designed the relevant curriculum system in the direction of big data analysis and application. As shown in Figure 10.

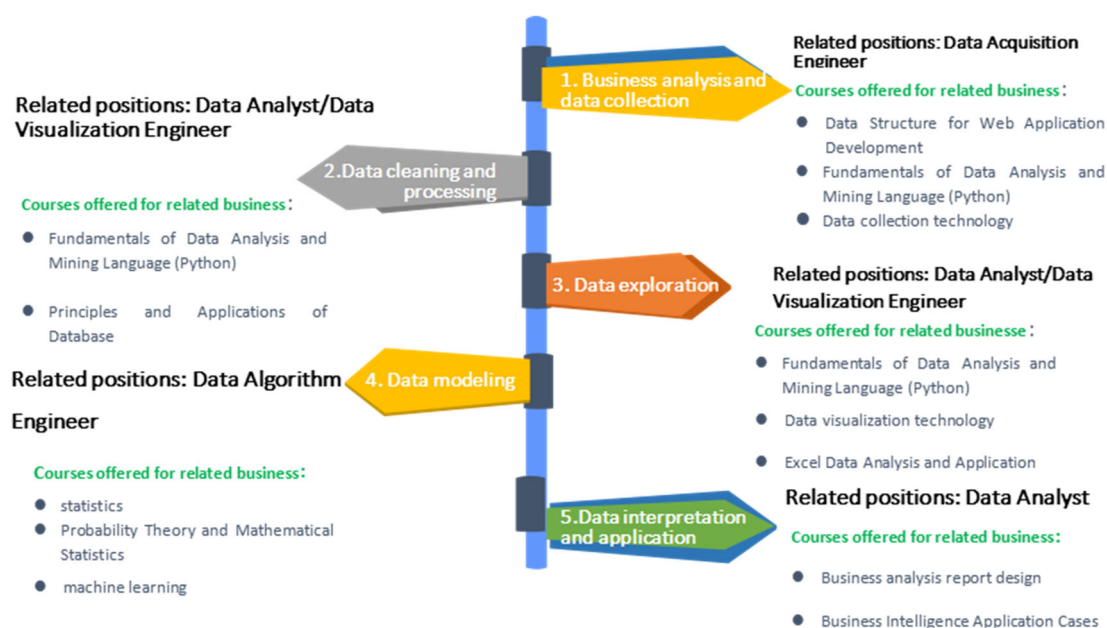


Figure 10. Curriculum system of big data analysis and Application

The professional requirements, ability training and teaching content of the relevant courses in the major direction are guided and designed with reference to the above "ability course structure" conversion results. As a result, a 2022 and 2023 level talent training program for information management and information system big data analysis and application has been formed. At the same time, during the construction of the project, the syllabus of these courses was also designed and written.

In the curriculum syllabus, the design of teaching content aims at the cultivation of cognitive ability and practical ability, and designs the corresponding content of application layer and cognitive layer. For example, in unit 1 of the course data visualization technology, the expected learning effect is to require students to understand the basic concepts and principles of data visualization, master the basic knowledge of font and color, which belongs to the cognitive level; The second unit requires to master the installation and use methods of data visualization related libraries. This is the teaching content of the application layer, and the goal is to cultivate students' practical ability in data visualization.

4. Practice Effect and Evaluation

Based on the big data technology, the subject of college curriculum system design to meet the needs of enterprises provides scientific, clear and feasible design ideas and methods for the talent training of information management and information system major in big data analysis and application direction of Guangzhou business school, effectively assists the design and implementation of the 2022 and 2023 information management and information system talent training programs of Guangzhou business school, and finally gets further verification and feedback through the design and teaching practice of the content of the relevant 12 courses.

The talent training scheme based on this topic has been discussed and adjusted by the college leaders and the teaching and research team of professional construction for many times, and has been recognized and affirmed by the top experts in the industry invited by the University for two consecutive years in 2022 and 2023.

Acknowledgements

The project number is 2021GXJK388; the project name is "Course System Design for Applied Talents Cultivation in Colleges and Universities Based on Big Data Technology for Enterprise Needs"

References

- [1] Wuxiaoling, jiangminmin, zhangliantang, et al. Curriculum system design of applied talents training in Colleges and universities for enterprise demand under big data technology [J]. computer education, 2020, (01): 86-92.doi:10.16512/j.cnki.jsjy.2020.01.021
- [2] Ouyang, Linyan, Huang Fenghua. Thinking and reform of applied talents training in data science and big data technology under the background of new engineering [J]. information system engineering, 2020, (12): 163-164
- [3] Pengguangbin, zhangxudong, dengyongsheng. Construction of talent training system for big data technology and application Majors Based on career demand analysis [J]. computer education, 2020, (04): 65-70.doi:10.16512/j.cnki.jsjy.2020.04.017
- [4] Wangjianfang. Research on the cultivation of Applied Talents in computer science and technology in the era of big data [J]. Journal of Qiqihar University (PHILOSOPHY AND SOCIAL SCIENCES EDITION), 2021, (03): 185-188.doi:10.13971/j.cnki.cn23-1435/c.2021.03.048

- [5] Tianshufang. Research on the training path of new business talents in local application oriented Universities -- a case study based on the major of big data management and application [J]. Journal of Heilongjiang Institute of Technology (Comprehensive Edition), 2022,22 (03): 17-21.doi:10.16792/j.cnki.1672-6758.2022.03.028