

## The Practice Teaching System of Big Data Talent Training in Universities

Allam MAALLA<sup>a,\*</sup> and Shu-Qing LIN<sup>b,\*</sup>

Research Institute of Open Data Commercial Application, School of Information Technology and Engineering, Guangzhou College of Commerce, Guangzhou, China, 511363

<sup>a</sup>allammaalla@yahoo.com <sup>b</sup>975871783@qq.com

**Keywords:** Financial Big Data Talents, Curriculum Framework Design, Big Data

**Abstract.** This paper goal is to build a practical teaching system for financial big data talent training in business colleges, accelerate the training of qualified and excellent financial big data engineering talents, meet the talent needs of the industry, and expand the professional reputation of the colleges. The big data experiment platform with advanced big data integrated experimental system and training management system provides a virtual simulation environment for practical teaching. The analysis and design theory and methods of financial information systems have the ability to analyze, design, implement and maintain financial information systems.

### Introduction

Benefiting from the rapid expansion of the big data market, the demand for related IT support has exploded. Among them, enterprises that provide big data infrastructure, big data software technology services, and industry big data content consulting services have brought unprecedented Customer group. DC predicts that by 2020, the company's expenditure based on big data computing and analysis platform will exceed 500 billion US dollars, and the compound growth rate will reach 34.1% in the next 5 years; in the next 3 to 5 years, China needs 1.8 million data talents, but currently only About 300,000 people. At the same time, China's colleges and universities in cloud computing, data science and other majors are still in their infancy, and the talents cultivated each year are far from meeting the needs of the industry. Therefore, it is imperative to open a big data major and accelerate the cultivation of talents.

The financial industry is the industry that relies most on data and is the easiest to realize data. In recent years, emerging financial institutions such as consumer loans and P2P are the products of the combination of big data technology and finance. At present, the demand for big data talents in finance is extremely strong in China. Only Internet finance is one year, and the growth rate is 3-5 times per year. It is generally believed that there will be a gap of 1 million talents in Internet finance, and the most lacking is big data risk control talents, including data mining and statistical modeling talents from primary to advanced.

### Financial Big Data Talents

The financial industry is the industry that relies most on data and is the easiest to realize data. In recent years, emerging financial institutions such as consumer loans and P2P are the products of the combination of big data technology and finance. At present, the demand for big data talents in finance is extremely strong in China. Only Internet finance is one year, and

the growth rate is 3-5 times per year. It is generally believed that there will be a gap of 1 million talents in Internet finance, and the most lacking is big data risk control talents, including data mining and statistical modeling talents from primary to advanced.

4. Cultivate qualified and excellent financial big data engineering talents, explore new talent training mechanisms for the development of business education, and expand the professional reputation of colleges and universities.

Accelerate the training of qualified and excellent financial big data application talents to meet the needs of social talents and expand the professional reputation of colleges and universities. At the same time, the goal of the three-dimensional practical teaching system of business talents has broken through the research perspective of traditional economic management professional practice teaching system, focusing on solving the bottleneck problem of market demand and talent supply contradiction, and exploring new talent training mechanism for the development of business education. According to the characteristics of business talent training, relying on the advantages of economic management disciplines, it is fully achievable to build a practical teaching system for business talent training with the goal of cultivating students' innovative practical ability.

### **Constructing a Practical Teaching System**

Constructing a practical teaching system integrating the big data engineering practice capability framework, experimental environment, teacher strength and teaching cases is a necessary condition for cultivating financial big data application technology talents. At present, the training of financial big data talents has a certain foundation for construction in the course system design, teacher strength, and teaching cases. Using big data technology to analyze the online recruitment information that can best reflect the talent needs of enterprises, and extract and extract the information about the talent knowledge and core competence of big data application talents. We designed a financial system that is connected with the needs of enterprises and has a complete knowledge chain. Data course system. The School of Information Engineering and Technology has a certain number of big data technology faculty members, including middle and senior professional teachers with practical experience in big data technology related projects.

With the advent of the era of big data, the surge in data volume has brought unprecedented opportunities and challenges to data mining. The complexity of financial big data has prompted data mining talents to have stronger financial data collation, analysis capabilities and higher overall quality. The construction of the financial big data talent training practice system with the characteristics of commercial colleges and universities has a good professional construction demonstration significance for the application of talent training in our school to break through the bottleneck between market demand and talent supply.

### **Key Issues to Be Solved**

According to the above analysis, under the guidance of the principles of overall planning and system coordination, how to break through and implement in place is the key issue in the construction of the entire practical teaching system. These problems and corresponding solutions are as follows:

a) How to design a financial big data technology talent training course to better meet the needs of the society?

Fully researched to form a complete, hierarchical, and cutting-edge curriculum framework design. The professional direction determines the rationality and tries to assist decision-making in methods, forms, and data. The survey method adopted is to use the big data technology to analyze the network recruitment information, and apply the knowledge tree principle to design the course group from the big data application process. School-enterprise professionally builds joint talent training and injects project engineers. The school-enterprise professional joint training of talents adopts the 2.5 + 0.5 + 1 model. That is: the 1-4th semester (2 academic year) course is completed in the school, mainly taught by the school teachers; the practice session sends enterprise teachers to the school to teach. These links include: entrance to school, new students' enrollment education, and primary school curriculum design training. The fifth semester (0.5 school year) is completed at the school and is mainly taught by school teachers. For individual technical courses, The sixth semester (0.5 academic year) course is completed at the training base. This part of the course is a professional course and is linked to the needs and requirements of the employer. In the 7th and 8th semester (1 academic year), students are also instructed to complete the graduation project.

A pyramid-like curriculum teaching case library that serves the teaching needs of different levels and objectives. The introduction of the big data comprehensive experimental system and platform has a case database of independent development and design of the enterprise, which is used to realize practical training. Cases are available in different scales and difficulty levels, and can be used in all aspects of teaching, including curriculum design, primary school, training, and graduation design. The design of the case corresponds to the curriculum of the university, such as demand analysis, software modeling, object-oriented programming, etc., so the system can be used in conjunction with the course to ensure that students can combine theory with practice and improve learning.

b) How to improve the technical ability, project experience, teaching and research level of the teachers in the field of big data?

Through the way that enterprise engineers lead the way, help the teachers in the school to understand and master the relevant financial big data analysis technology in practice teaching. Cooperate with the enterprise to train talents and implement the practical training of students, and promote the technical ability level and project experience of the teachers in the field of big data. The project plan and the school-enterprise professional jointly build joint talent training. The joint talent training adopts the 2.5 + 0.5 + 1 model. That is: the 1-4th semester (2 academic years) course is completed in the school, mainly taught by the school teachers; the practice session sends enterprise teachers to the school to teach. The fifth semester (0.5 academic years) is completed at the school and is mainly taught by school teachers. For individual technical courses. At the same time, teachers from the school are encouraged to participate in these classrooms.

c) What is the principle of student knowledge internalization? How to design efficient and scientific methods to improve students' interest and ability in learning? First, the overall cognitive teaching is the talent training medium. For the whole profession, from the introduction course of the whole profession to each course and every class, we insist on combing the knowledge framework according to the overall cognition, but do not fall into the pure theoretical introduction of the concept, but draw the relevant background knowledge by case practice. Second, a systematic case library, supplemented by understanding digestion knowledge. Third, comprehensively enrich the case library to make up for the shortcomings

of the case. Cases are often designed for individual practical problems. The content coverage is narrow and not systematic. For hidden knowledge that is not within the system, it needs to be expanded by other cases to supplement the implementation. Therefore, in addition to the single case within the textbook Teaching, also need to expand through extracurricular activities, extracurricular auxiliary learning materials and other means to ensure the acquisition of hidden knowledge within the system.

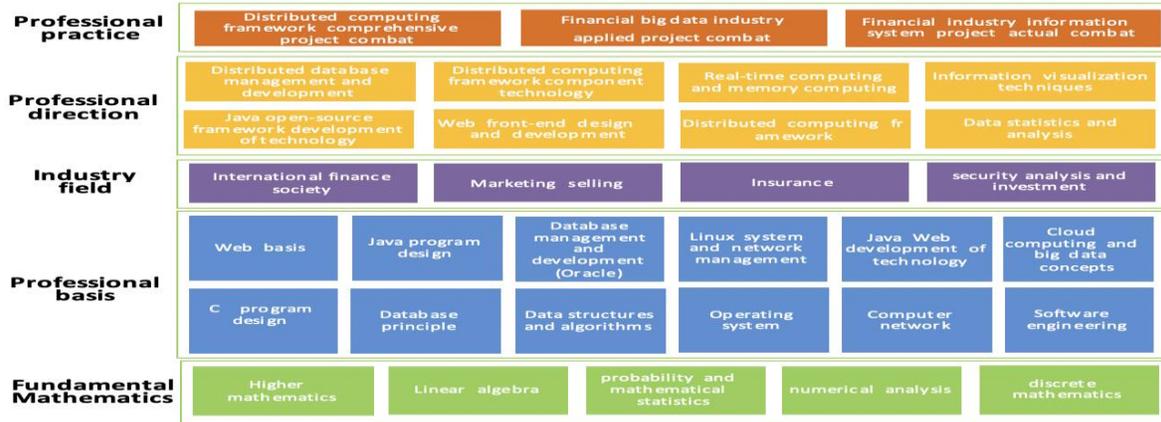
### **Curriculum Framework Design**

(1) Cloud Computing and Introduction to Big Data. As an introductory course in the direction of this major, this course introduces students to the concepts, technologies and applications related to cloud computing and big data, so that students can establish a preliminary understanding of the relevant knowledge, technology and development prospects of the profession, as a guide for the follow-up course.

(2) Distributed computing framework foundation. As the foundation and core technology course of this major, this course introduces students to the basic concepts, installation and configuration of the Hadoop distributed computing system, distributed programming model (Map/Reduce), distributed file system (HDFS), and related scheduling, monitoring and maintenance tools enable students to build a basic understanding of distributed computing systems, master the primary distributed application design and implementation methods, and lay the theoretical and practical foundation for subsequent in-depth courses.

(3) Distributed database management and development. As a core technical course in this major, this course introduces students to the basic concepts of distributed databases, installation and configuration, management and maintenance, data access and development. The course focuses on NoSQL databases such as HBase, MongoDB, Redis, etc., and describes their use and development in a distributed environment. To enable students to establish a basic understanding of distributed databases, master the primary distributed database application system design and development methods, and lay the theoretical and practical foundation for the subsequent in-depth courses.

(4) Distributed Computing Framework component technology. As a core advanced course in this major, this course introduces students to mainstream components on the Hadoop distributed computing platform, including Hive, Pig, Sqoop, Flume, Kafka, Zookeeper and more. Enable students to have a complete Hadoop ecosystem-based design and implementation of big data applications.



**Figure 1.** Curriculum Framework Design

### Building a Big Data and Business Intelligence Experiment System

This experimental system is designed to provide students with a complete set of Hadoop and its environment, design, development, monitoring, maintenance tools, software and services. With this experimental system, the experimental and training environment requirements of the core technology courses of this major can be met. This experimental system is divided into two major components:(1)A virtual lab environment for students to learn big data. The environment is carried out by means of the aforementioned virtualized desktop teaching system, and the network administrator configures the big data learning virtual machine in advance for the students to use.

(2) The real environment for research or large-scale case presentations. This environment is carried out through several servers.

### System Architecture

#### Main Functions

**Basic platform:** The basic platform for big data storage and processing, which can realize the storage and management of massive data, support common components of platforms such as Hive, Impala, Pig, Spark, and Yarn, and provide support for data analysis services on the platform. These common components increase the ease of use of platform data, making data manipulation and data analysis easier to use, saving labor and reducing labor time.

**Data integration:** support the unified storage of massive structured data, semi-structured data, and unstructured data, deepen the expansion of enterprise intelligence and service capabilities, and improve the decision-making level of enterprises. We can use enterprise-level data ETL tools or open source ETL tools. For example, Flume, Sqoop, Kafka, etc.. Through this platform, faster and more convenient data integration can be achieved.

**Data analysis and processing:** Provide different data analysis and processing modules for different industry applications. Through data analysis and processing, enterprises can clarify the scope of their intelligence and service capabilities, as well as the requirements and solutions implied by the data. Solutions and the ability to meet the deep data mining and analysis needs of different industries.

Platform operation and maintenance: real-time and uninterrupted monitoring and management of the services and data resources involved in the platform through the visual management terminal, reasonable setting of early warning thresholds, ensuring timely notification and display of abnormal resource status, and timely notification via SMS and email Managers discover and deal with problems. In the case that the cluster is uncontrollable, you can use the platform to perform data backup after the disaster, version upgrade of the cluster, or rollback of the version.

The system is a comprehensive product that includes data storage, data integration, data analysis and processing, and application API services. Supports enterprise-level data integration and processing, and integrates multiple open source ETL and data processing and analysis tools; it includes both in-memory data analysis (Spark) engine and Ad Hoc. Query provides both a real-time search engine and deep analysis and mining tools. The composition is shown as below:

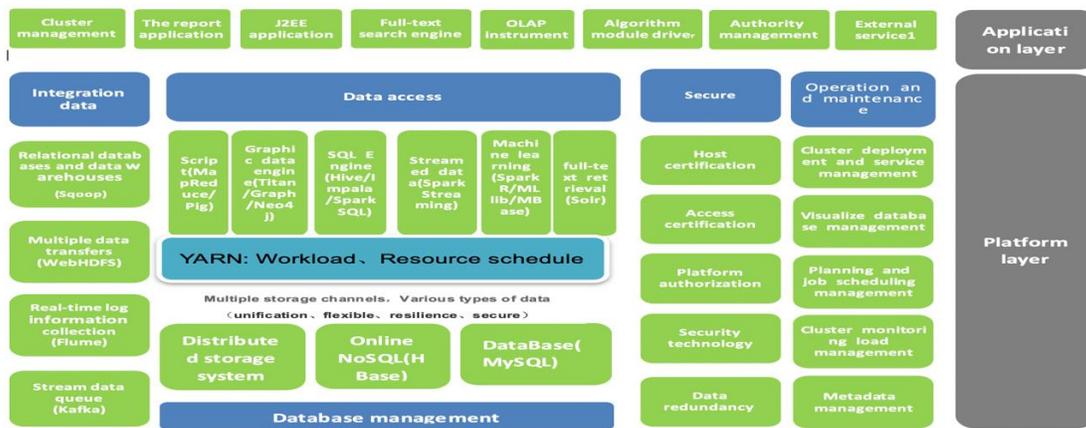


Figure 2. System architecture diagram 'Cloud training and learning platform'

### Summary

Curriculum design is an important carrier for talent training, with the core of cultivating students' professional application ability, information application ability and social adaptability, the proportion of public basic courses, professional courses and practical teaching courses should be reasonably set to guide students to make full use of network teaching resources, such as Excellent courses and teaching videos, etc., enable students to extensively access a large number of course information, cultivate "cross-border" thinking and awareness, and continuously improve the ability to identify information, analyze information and comprehensively apply information. In the process of cultivating financial talents, we should focus on the teaching methods that focus on stimulating students' interest in learning and the reform of examination methods based on the application of students' ability to ensure students' knowledge and use of knowledge. Teachers should not only pay attention to students' understanding of theoretical knowledge and the operation of practical skills, but also to the flexible use of layered teaching, peer interaction, teamwork and other means to create high-performance teaching modes, in many complicated Construct a three-dimensional, diverse and complete knowledge system in information and problems.

## **Acknowledgement**

This research was financially supported by the 2017 Higher Education Teaching Research and Reform Project of Guangdong Province, "The Big Data Talent Cultivation Practice Teaching System of Business Schools" (Project No. 2017SJXGG01), and Characteristic Key Subject of E-commerce Construction Project of Guangzhou College of Commerce Foundation, (Project No. TSZDXK201601).

## **References**

- [1] De Nisi, A.S. and R.W. Griffin, "Human Resource Management". 2015, New York: Houghton Mifflin Company.
- [2] A TP Track Research Report Talent Management: "A State of the Art". 2015, Tower Perrin HR Services.
- [3] Maalla A, Zhou C, Guang-Yu W U. The Construction of Heterogeneous Platform of Unified Service System Based on Cloud Computing [J].2017(cmee). DOI: 10.12783/dtce/cmee2017/20056.
- [4] Tso, G.K.F. and K.K.W. Yau, "Predicting electricity energy consumption : A comparison of regression analysis, decision tree and neural networks". *Energy*, 2016. 32: p. 1761 - 1768.
- [5] Han, J. and M. Kamber, "Data Mining: Concepts and Techniques". 2014, San Francisco: Morgan Kaufmann Publisher.
- [6] Chien, C.F. and L.F. Chen, "Data mining to improve personnel selection and enhance human capital: A case study in high-technology industry". *Expert Systems and Applications*, 2015. 34(1): p. 380-290.