ISSN: 2637-6067 DOI: 10.6918/IJOSSER.202505_8(5).0006

Research on Teaching Reform of "Big Data Analysis and Modeling" Course Based on POPBL

Shimin Liang

School of Computer Science, Guangdong University of Science and Technology, Dongguan, China

Abstract

As a core elective course in Data Science and Big Data Technology, "Big Data Analysis and Modeling" is designed for students who are interested in delving into data analysis and mining techniques, with the aim of sending data analysts to companies. In the past, the teaching of Big Data Analysis and Modeling is faced with such challenges as inaccurate orientation of teaching objectives, unitary traditional teaching methods and unscientific teaching evaluation methods. It is urgent to carry out teaching reform based on POPBL mode. Set teaching objectives and organize teaching content based on job requirements, integrate Problem-Based Learning and Project-Based Learning for teaching, build a reasonable course evaluation method, in order to provide reference for improving the teaching quality and student learning effectiveness of "Big Data Analysis and Modeling".

Keywords

POPBL; Big data analysis and modeling; Problem-Based Learning; Project-Based Learning.

1. Introduction

POPBL (Problem-Oriented Project-Based Learning) is a problem oriented, project based teaching model proposed by Aalborg University [1], which integrates problem-based and project-based PBL for teaching, encourages interdisciplinary practice, and emphasizes group cooperation to achieve the learning process [2]. Based on the POPBL model, Beijing Union University has carried out teaching reforms in the course of Microbiology Experiment, and formulated reform strategies such as problem-based heuristic teaching, interdisciplinary and cross course teaching content design, and the construction of a more scientific comprehensive evaluation method for the course. The teaching effectiveness has been significantly improved. Zhejiang University has introduced the POPBL model into the course of "Marketing Terminal Management", which has stimulated students' interest in learning and conducted in-depth analysis on key issues such as selecting appropriate projects, scientifically creating problem scenarios, and teacher role transformation. Based on the core architecture of POPBL, Beijing University of Electronic Information Technology has carried out teaching reforms in the course of "Software Project Management", achieving a significant improvement in students' engineering practice ability and learning enthusiasm.

2. Analysis of the Teaching Status of "Big Data Analysis and Modeling" Course

As a core elective course in Data Science and Big Data Technology, "Big Data Analysis and Modeling" is designed for students who are interested in delving deeper into data mining and data analysis techniques, with the aim of delivering data analysts to enterprises. This course is built upon a series of solid foundational courses, including Python Fundamentals and Applications, Machine Learning, Data Visualization, and more. After completing this course,

ISSN: 2637-6067 DOI: 10.6918/IJOSSER.202505_8(5).0006

students will also embark on a series of closely related practical learning activities, including "Graduation Internship", "Graduation Thesis (Design)", etc. Big Data Analysis and Modeling combines practicality, applicability, and comprehensiveness. However, in previous teaching, the following issues urgently need to be addressed in this course:

2.1. The positioning of teaching objectives is not accurate enough

This course aims to deliver data analysts to enterprises, but the shortcomings of previous teaching, which emphasized theory over practice, have led to a focus on achieving knowledge and skill goals, rather than process and method goals. This is reflected in the emphasis on explaining functions and algorithms, and the neglect of project programming implementation. This results in students having relatively weak abilities in flexibly applying knowledge, writing data analysis reports, and conducting project presentations. The inaccurate positioning of teaching objectives hinders students from meeting the job requirements of enterprises, making it difficult for them to seamlessly transition from campus to the workplace after graduation.

2.2. Traditional and singular teaching methods

In the past, teaching focused more on teachers' teaching and less on students' learning, using a single teaching method mainly based on teachers' lectures and demonstrations. This mode often leads to students being in a passive state of receiving knowledge, with more passive listening in the classroom rather than active exploration, lacking effective interaction, and thus greatly reducing students' participation. With the vigorous development of the Internet, this old teaching method has been difficult to stimulate the learning interest and enthusiasm of contemporary college students, and can not meet the learning needs of college students in the new era.

2.3. The teaching evaluation method is not scientific

In the past, teaching used a summative evaluation method, which only evaluated students' final course grades based on their final exam scores. This evaluation method not only lacked comprehensiveness and scientific nature, but also invisibly fostered students' laziness in their daily learning. In previous teaching, final exams often took the form of closed book written tests, and the assessment focused on memorization and memorization of knowledge points, which led students to tend towards mechanical memorization rather than deep understanding, thereby hindering their ability to internalize knowledge and apply it in practice.

3. Teaching Reform Strategies for the Course of Big Data Analysis and Modeling Based on POPBL

3.1. Set teaching objectives and organize teaching content based on job requirements

With the development of the economy, the demand for data analysis talents in society continues to increase, while also placing higher demands on them. The research team closely follows the pulse of social demand, comprehensively considers the development goals of the school's "building a high-level applied innovative university" and the talent training plan for data science and big data technology majors, and formulates teaching objectives that meet both social needs and professional characteristics. Aiming at the corresponding position of the course "Big Data Analysis and Modeling", the research team has conducted in-depth enterprise research and summarized the job requirements for data analysts in the corresponding position as follows: (1) comprehensively master the complete data analysis process of "requirement analysis--data acquisition--data exploration--data preprocessing--data analysis and modeling--data visualization--data analysis report"; (2) Proficient in mainstream data analysis tools such as Python, R, and SPSS, able to independently complete data analysis reports and make

ISSN: 2637-6067 DOI: 10.6918/IJOSSER.202505_8(5).0006

presentations, with strong logical thinking and outstanding communication skills; (3) Capable of establishing and maintaining machine learning models based on actual needs, conducting data analysis and prediction. Based on job requirements, the research team has set the following teaching objectives for the course "Big Data Analysis and Modeling": (1) Familiarize oneself with the basic concepts, principles, and implementation processes of big data analysis and modeling; (2) Master the use of Pandas, Matplotlib, scikit-learn and other libraries to complete key tasks such as data preprocessing, data visualization, and data analysis; (3) Master machine learning models such as decision trees, random forests, and neural networks, and conduct data analysis according to actual needs; (4) Master report writing and project reporting skills, enhance logical thinking and language expression abilities.

Based on the established teaching objectives, the research group has developed a teaching content that covers two main parts: basic accumulation and practical application. When delving into the refinement of organizational teaching content, the research team realized that although textbooks serve as important carriers of knowledge transmission and follow the principles of systematicity and comprehensiveness in their development, it is not reasonable to rely solely on textbook content for teaching in the highly practical and comprehensive the course "Big Data Analysis and Modeling". Therefore, in order to cultivate students' comprehensive ability to proficiently use machine learning libraries to conduct in-depth analysis and prediction of real data based on actual needs, the research team organizes teaching content according to a complete data analysis process, which covers key links such as data exploration, data analysis, and modeling. In terms of grasping the difficulty of teaching content, the research group combined the theory of the zone of proximal development to construct an interdisciplinary business project library that students can complete through hard work, including comprehensive analysis of catering enterprises, segmentation of mall customers, prediction of telecom customer churn, housing price prediction and other data analysis projects. The teaching content is moderately difficult and closely related to students' daily lives, stimulating their interest and enthusiasm for learning.

3.2. Integrating Problem-Based Learning and Project-Based Learning for teaching

In the teaching of the core chapter "Fundamentals of Data Mining Algorithms" in the basic part, the comprehensive analysis project of catering enterprises runs through the entire process, which includes five sub projects: customer churn prediction, sales prediction, customer value analysis, dish correlation analysis, and dish recommendation. In the classroom introduction stage, after introducing each sub item, students are asked questions through the class exercises of Learning Pass: What types of tasks do the five sub items of customer churn prediction, sales prediction, customer value analysis, dish correlation analysis, and dish recommendation belong to? How to implement it using Python? After a brief introduction to five types of data mining algorithms: classification, regression, clustering, association rules, and intelligent recommendation, return to the opening question and guide students to map knowledge points to the problem. After students think about it, provide options. In the process of hands-on practice in the comprehensive analysis project of catering enterprises, students master the principles and Python implementation of classification algorithms such as logistic regression, decision tree, support vector machine by completing customer churn prediction; Master the principles and Python implementation of regression algorithms such as linear regression, neural networks, and random forests by completing sales forecasting; Master the principles and Python implementation of hierarchical clustering, K-Means and other clustering algorithms by completing customer value analysis; Master the principle and Python implementation of Apriori association rule algorithm by completing dish association analysis; Master the principles and Python implementation of recommendation algorithms such as user based

ISSN: 2637-6067 DOI: 10.6918/IJOSSER.202505_8(5).0006

collaborative filtering and item based collaborative filtering by completing dish recommendations.

The competition is similar in content and form to the working conditions for students after graduation, providing them with a practical "training ground" that can meet their actual needs, and is an important way of practical combat. In the practical part, three interdisciplinary business projects from Kaggle competition are introduced, including mall customer segmentation, telecom customer churn prediction, and house price prediction. Each project covers the complete data analysis process of "requirement analysis--data acquisition--data exploration--data preprocessing--data analysis and modeling--data visualization--data analysis report". Taking the customer segmentation project in a shopping mall as an example, after conducting requirement analysis, students are asked a thought-provoking question through the "Learning Pass" quiz: What algorithm can achieve customer segmentation in shopping malls? How to implement it using Python? After sorting out their thoughts, students implement projects, write reports, and report on projects in groups to improve their ability to apply their knowledge to solve practical problems, cultivate their teamwork spirit, and exercise their language expression skills. In the practical learning process, students' subjectivity is fully utilized, and teachers act as guides and organizers.

3.3. Construct reasonable curriculum evaluation methods

The evaluation of Big Data Analysis and Modeling includes two parts: forward assessment and reverse assessment. The forward assessment evaluates students' learning effectiveness by taking into account their regular grades and final grades. Students are graded based on classroom performance, project reports and project reports; The final grade is determined by the final exam grade. The final exam takes the form of open-book test, with 50% of the practical exercises, focusing on the ability of students to apply the knowledge to solve practical problems. The way of forward assessment was announced to students in the first class, and students no longer need to memorize, but pay attention to the understanding and flexible use of knowledge, and finally obtained gratifying learning results. In the process of reverse evaluation, students' satisfaction with teaching is investigated, and curriculum evaluation methods and teaching strategies are improved based on feedback, so as to continuously improve teaching quality and students' learning results.

4. Conclusions

Based on the course "Big Data Analysis and Modeling", POPBL-based teaching reform is implemented, problem-oriented and project-oriented are promoted, students are guided to cross-disciplinary practice, and the reform path of learning process is realized through group cooperation, so as to cultivate students' teamwork ability and knowledge application ability. At the beginning of the teaching reform, based on the social post needs, set teaching goals, according to the complete data analysis process to organize teaching content in line with students' recent development area, improve students' learning interest and enthusiasm; Secondly, it integrates Problem-Based Learning and Project-Based Learning to improve students' participation and learning initiative. Finally, it evaluates the curriculum from the perspective of forward assessment and reverse assessment, adopts diversified evaluation methods, and constantly improves the curriculum evaluation methods and curriculum teaching strategies to meet the needs of society.

Acknowledgements

Project source: 2024 Guangdong University of Science and Technology "Quality Engineering" Higher Education Teaching Reform Project

ISSN: 2637-6067 DOI: 10.6918/IJOSSER.202505_8(5).0006

Project Name: Teaching Reform and Practice of "Big Data Analysis and Modeling" Course Based on POPBL (Project Number: GKZLGC2024182)

References

[1] Lehmann M, Christensen P, Du X, et al. Problem-oriented and project-based learning (POPBL) as an innovative learning strategy for sustainable development in engineering education[J]. European journal of engineering education, 2008, 33(3): 283-295.

[2] Kolmos A. Problem-based and project-based learning[J]. University science and mathematics education in transition, 2009: 261-280.

[3] Yasin R M, Rahman S. Problem oriented project based learning (POPBL) in promoting education for sustainable development[J]. Procedia-Social and Behavioral Sciences, 2011, 15(1): 289-293.

[4] Asim H M, Vaz A, Ahmed A, et al. A Review on Outcome Based Education and Factors That Impact Student Learning Outcomes in Tertiary Education System[J]. International Education Studies, 2021, 14(2): 1-11.

[5] Harland T. Vygotsky's zone of proximal development and problem-based learning: Linking a theoretical concept with practice through action research[J]. Teaching in higher education, 2003, 8(2): 263-272.

[6] Bojer C S, Meldgaard J P. Kaggle forecasting competitions: An overlooked learning opportunity[J]. International Journal of Forecasting, 2021, 37(2): 587-603.