International Education Forum, 2025, Volume 3, Issue 2
http://ojs.bbwpublisher.com/index.php/IEF
Online ISSN: 2981-8605

Digital Intelligence Empowerment: Teaching Exploration of Mathematics Subject Data Course

Yuting Li*, Hairong Lian, Fengjie Geng

School of Science, China University of Geosciences (Beijing), Beijing 100083, China

Corresponding author: Yuting Li, 2023010017@cugb.edu.cn

Abstract: Given the rapid development of the data science field, college personnel training is facing new challenges and opportunities. This paper focuses on the "Introduction to Data Science" course, an in-depth exploration of the teaching reform of data courses. With the optimization of curriculum settings, innovation, and reform of teaching methods and other measures, the study focuses on case design, effectively promotes the deep combination of theory and practice, and integrates machine learning algorithms and Python programming language, to effectively improve students' comprehensive literacy. In addition, this paper also discusses the cross integration of mathematics and data science, to provide a valuable reference for the transportation of high-quality talents in the flourishing era of numerical intelligence.

Keywords: Data science; Mathematics; Digital intelligence empowerment; Teaching reform

Online publication: March 12, 2025

1. Introduction

With the rapid development of emerging technologies such as big data and artificial intelligence, human society has entered the era of digital intelligence [1–2]. In this wave of times, data has exploded and become a key factor driving innovation and development in various fields, and its influence in the field of education has become increasingly prominent. Innovation-driven development has always been the strategic focus of China, which strives to promote the vigorous development of new economies such as new technologies, new forms of business, new industries, and new models [3]. To occupy a dominant position in the global competition, overcome the "bottleneck" problem, and break through the core technology, the country urgently needs a large number of talents with a solid theoretical foundation and the application ability of new technology. However, the traditional training model of basic science talents has made it difficult to meet the requirements of the new era, new economic development, and new industrial revolution [4]. Therefore, the establishment of an applied science major has become an urgent need for national strategic development. As the core carrier of training students' application and practice ability, data majors are facing unprecedented

opportunities and challenges [5–6].

On the one hand, data courses are developing rapidly with the advent of the age of digital intelligence. Digital resources are very rich, online platforms for students to bring a large number of learning materials so that students can access cutting-edge knowledge. Teachers can also use big data technology to analyze students' learning trajectories, such as the mastery of knowledge points, to carry out personalized teaching more accurately and effectively complete the teaching content. The emergence of a variety of data analysis software has brought great convenience to learning data courses, such as Python, R language, etc., which allows students to quickly apply what they have learned and achieve the effect of putting what they have learned into practice.

On the other hand, the teaching of data courses has also encountered many difficulties. With the acceleration of knowledge upgrading and the endless emergence of new technologies and algorithms in the field of data science, how to keep up with the pace of the time and incorporate cutting-edge achievements in time has become a thorny problem faced by teachers. At the same time, how to build a data specialty under the background of mathematics discipline has become an increasingly prominent problem. The traditional teaching mode emphasizes theory teaching, and the practice teaching link is weak, which makes it difficult to meet the strict requirements of students' practice and innovation ability in the age of digital intelligence. At a time when data security and ethical issues are increasingly attracting social attention, how to integrate relevant educational content into the curriculum and guide students to establish correct data values cannot be ignored.

Given this, it is of great significance to deeply explore the teaching strategies of data courses in the age of digital intelligence. For the data computation and application specialty under the mathematics discipline, "Introduction to Data Science" as an introductory course, its importance is self-evident. Although mathematical theory is the cornerstone of data science, how to carry out teaching for students of this major in line with their disciplinary background, so that students can not only have a solid grasp of data science knowledge but also deeply integrate mathematical expertise, is an urgent problem to be solved, which is related to whether educators can provide the society with compound talents with profound mathematical heritage and data science application ability. This paper focuses on the construction of data courses under the background of mathematics, and discusses the teaching research and practice of "Introduction to Data Science."

2. Characteristics and status quo of data courses

2.1. Features

The era of digital intelligence is a high-level stage of information technology development, and its core essence is the deep integration of digitalization and intelligence, cutting-edge technologies such as big data, artificial intelligence, and the Internet of Things enable each other to work together to drive all-round social change. Big data is like a "digital gold mine" containing endless value, with massive, diverse, high-speed, and real data continuously pouring out from every corner. Artificial intelligence is like a "digital brain" with extraordinary wisdom, and cutting-edge algorithms such as machine learning and deep learning make it have the excellent ability to simulate human intelligence, and can accurately analyze, predict, and make decisions on massive data.

As a key component of the education system in the age of digital intelligence, data courses cover the interdisciplinary knowledge fields of statistics, computer science, mathematics, etc., as shown in **Figure 1**. From the perspective of basic theory, statistics provides the cornerstone of scientific methodology for data collection, sorting, analysis, and inference, enabling students to master core skills such as data distribution law, correlation analysis, and hypoth-

esis testing, and to mine valuable information from massive data. Computer science gives students the ability to data storage, processing, algorithm design, and programming implementation. With the help of knowledge of database management systems, data structure, algorithms, programming languages, and so on, they can effectively realize the automatic operation of data and solve complex tasks. As the foundation of logical thinking and quantitative analysis, mathematics, linear algebra, calculus, probability theory, and other branches provide powerful tools for data modeling, optimization algorithms, machine learning theoretical derivation, etc., to help students build accurate mathematical models to explain data phenomena.

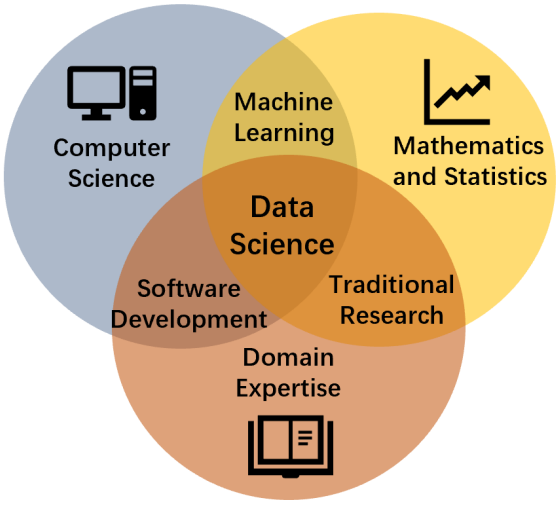

Figure 1. Data science

"Introduction to Data Science" is an important data course aimed at cultivating talents with theoretical knowledge and practical ability. On the one hand, to be able to apply the knowledge learned, students need to master the technical means of data processing, such as data acquisition, storage, processing, analysis, and visualization. At the same time, they should be proficient in data processing software SQL, Python, R language, and other tools. For structured and unstructured data, such as marketing data, social network data, image recognition, etc., students need to process and analyze. On the other hand, it is very important to cultivate students' mathematical thinking and creative consciousness to construct data courses under the mathematics discipline. Students should have the thinking habit of using data to understand the nature of problems, discover potential rules, predict development trends, and use data to drive decision-making and solve practical problems innovatively.

2.2. Status quo

The teaching and research of data courses in foreign countries started earlier and achieved a lot. Many top universities in Europe and the United States, such as Stanford University and Massachusetts Institute of Technology, rely on strong scientific research strength and rich educational resources to build a cutting-edge and systematic data science curriculum system. These courses not only cover basic theoretical knowledge such as statistics and computer science but also closely combine with the actual needs of the industry and integrate the latest research results in hot fields such as machine learning, data mining, and artificial intelligence so that what students learn is closely connected with the needs of the market.

In contrast, although the domestic data course teaching and research started a little later, in recent years, with the vigorous rise of the big data industry, the development momentum is rapid. Among institutions of higher learning in the country, key universities such as Renmin University of China and Tsinghua University started data-related courses earlier. Subsequently, two undergraduate majors, Big Data and Data Science, were officially established in 2016 and 2022 respectively [4]. Many universities have set up data science and big data technology, data analysis, and other related majors and courses, and actively explore the teaching model suitable for national conditions. In August 2017, experts from the Guiding Committee for the Teaching of Mathematics in higher education institutions of the Ministry of Education declared the major "Data Calculation and Application." In March 2018, the exploration and practice of "Data Computation and Application" of mathematics applied science major, chaired by Professor Wu Shuquan of Fudan University, was approved [4].

In the field of mathematics, the major of data computation and application, as a cutting-edge major that deeply integrates mathematical theories into data processing and application, its guiding course "Introduction to Data Science" has become more and more important. Mathematical theory, as the solid foundation of data science, provides indispensable support for data analysis, modeling, and algorithm design. However, how to carry out precise teaching based on the unique disciplinary background of students in this major, so that they can not only grasp the core knowledge system of data science but also seamlessly embed their deep mathematical expertise in data science applications, has become a key problem to be solved in the field of education. The solution to this challenge is directly related to the ability to provide a steady stream of composite high-end talents with deep mathematical literacy and excellent data science application ability for society.

3. The internal relationship between mathematics and data science

In the context of mathematics, many students who choose to major in data computation and application have a question: "Why do we need to learn mathematical theory?" Therefore, in the Introduction to Data Science course, educators should explain to students what role mathematics plays in data science, which actually provides the theoretical foundation and practical tools for understanding and processing data. For example, matrix operations in linear algebra play a key role in dimensionality reduction and image processing. Therefore, students should understand that mathematics can serve as a powerful toolbox for data scientists, enabling them to extract valuable information from complex data sets and make relatively accurate decisions based on it. Through these types of courses, students understand that a solid mathematical foundation is essential for anyone wishing to go deep into the field of data science.

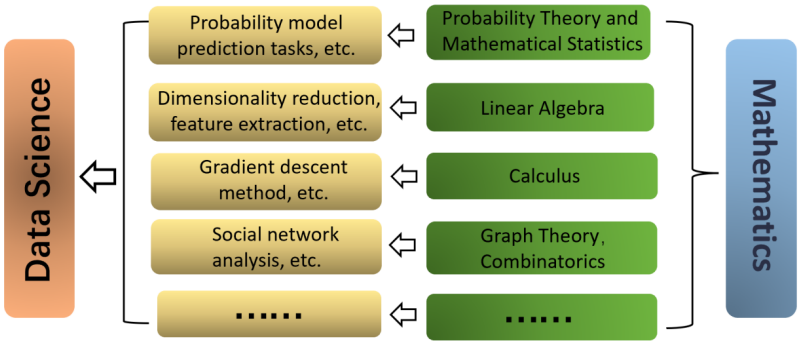

Figure 2. Connection between mathematics and data science

When using data science tools to analyze problems, students should be taught some connections between math and data science, as shown in **Figure 2**. Some examples are as follows.

Probability theory is the basis for understanding random phenomena. In data science, probability theory is used to model uncertainty and risk, such as the use of probabilistic models in machine learning algorithms to predict the likelihood of events occurring.

Linear algebra is essential for processing high-dimensional data. It involves vectors, matrices, and their operations, concepts that are widely used in data representation, feature extraction, dimensionality reduction techniques such as PCA, and many machine learning algorithms.

Calculus can be applied in the field of optimization of artificial intelligence algorithms, especially when the algorithm needs to search for the optimal value of a task. For example, the gradient descent method used in neural network training is based on calculus theory.

Graph theory, combinatorial mathematics, and other knowledge points in basic mathematics can be used to explain neural network structure, social network analysis, and graph neural networks.

Through the above analysis, educators need to let students know that data science research is based on mathematical theory, and mathematics also provides ideas and methods for data science. For data science, whether it is from theoretical research to application landing, or simple data visualization to complex artificial intelligence models, mathematics plays an indispensable role. Therefore, giving students a deep understanding and mastery of mathematics is essential to becoming a good data scientist.

4. Teaching thinking

4.1. Enhancement of mathematics and data science knowledge

For students majoring in data computation and application in mathematics, although they have a certain mathematical foundation, they still need to strengthen the mathematical knowledge closely related to data science. Explain in depth the Bayesian theory of probability theory, which has important applications in data classification, prediction, and other tasks. Strengthen the teaching of matrix decomposition and singular value decomposition in linear algebra, which is very important in data dimensionality reduction and recommendation systems. Through practical cases, such as the use of a Bayesian classifier for spam identification, students can experience the specific application of mathematical knowledge in data science.

In addition, the core body of knowledge for data science should be detailed. In terms of data mining, it explains algorithms such as association rule mining and cluster analysis, such as the Apriori algorithm for mining association rules in commodity sales data, and the K-Means clustering algorithm for classifying customer groups. In the field of machine learning, in-depth analysis of supervised learning, unsupervised learning, semi-supervised learning, and other algorithm principles, by comparing the performance of different algorithms in image recognition, speech recognition, and other tasks, so that students can understand the advantages of algorithms and suitable tasks. In the deep learning section, the structure of neural networks, backpropagation algorithms, and classical models such as convolutional neural networks and recurrent neural networks are introduced. Practical projects such as image classification and natural language processing are taken as examples to enable students to master the construction and training methods of deep learning models.

4.2. Problem-based learning

A teaching method that encourages active Learning by dealing with practical problems is called Problem-Based

learning (PBL), as shown in **Figure 3**. An important feature of the "Introduction to Data Science" course is that it is suitable for PBL teaching. Therefore, in classroom teaching, practical problems are used as examples to guide students to think and solve problems. This method encourages students to find problems, analyze problems, and find solutions by themselves in the classroom, to achieve the purpose of in-depth understanding and mastery of knowledge.

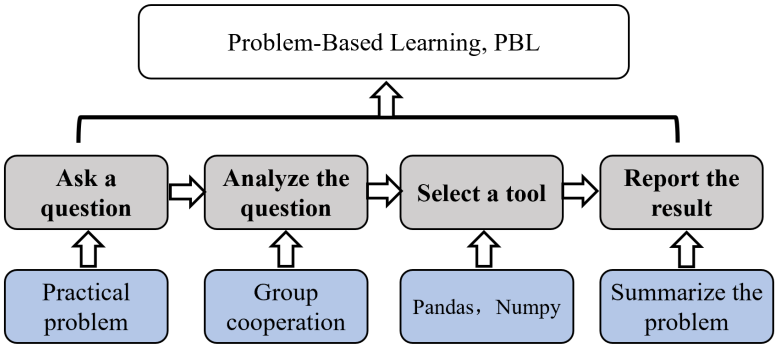

Figure 3. PBL teaching method map

The PBL teaching method is described in detail below.

Ask questions. When preparing the lesson, educators can set a problem that comes from life and integrate it with the course content to stimulate students' interest in learning. For example, when explaining the data preprocessing method, the specific case of the "Titanic sinking" incident can be introduced to analyze the relationship between gender, age, and other factors and the survival probability of passengers. The event was made famous by the movie Titanic, so students are very interested in it. Given specific problems, educators can solve the problem of "gender, age and the correlation between survivors and passengers" as a guide, and teach students relevant pretreatment methods such as data filling and data missing, which will help students understand and accept the knowledge points.

Analyze the question. After the question is raised, the analysis of the problem through data science methods can be in the form of group cooperation, for example, each group is consisted of 3–5 people, so that each student can participate in the discussion. Students can be assigned different roles within the group, such as collecting information, writing programs, and visualizing data analysis. This can not only improve work efficiency but also let students experience the importance of teamwork. For example, when talking about "exploratory data analysis", three students can be assigned to undertake professional terminology query, data collection, and data visualization. As a result, it can enhance students' classroom participation and activity, and stimulate students' interest in learning.

Select a tool. In the process of group work, teachers need to provide students with the necessary reference materials, tools, or platforms to help them gather information. For example, some commonly used data preprocessing libraries are Pandas, Numpy, etc., and visualization tools are Matplotlib, Seaborn, and so on. Teachers should regularly check the progress of each group and give feedback and guidance. At the same time, critical thinking is required in the evaluation process, guiding students to put forward open-ended questions and encouraging them to think about problems from different angles. After solving the problem, organize students to reflect and discuss what worked and what needs to be improved.

Report results. After the group debriefing, teachers need to find out how students feel about the PBL teaching style and what problems they encounter during the learning process. Finally, according to the results, the teaching methods are adjusted to improve the teaching design and classroom management. This can effectively implement the

PBL model in the classroom, not only improving students' interest in learning and engagement but also cultivating their teamwork and problem-solving skills.

4.3. Classroom thinking and politics in data science

For data courses, it is not only necessary to cultivate students' practical ability in teaching, but more importantly, to let them establish correct social values. The integration of ideological and political education in the course of "Introduction to Data Science" will not only help to cultivate well-developed scientific and technological talents but also promote the harmonious development of society.

As shown in **Figure 4**, the following examples give specific methods of integrating ideological and political education in the course of teaching "Introduction to Data Science."

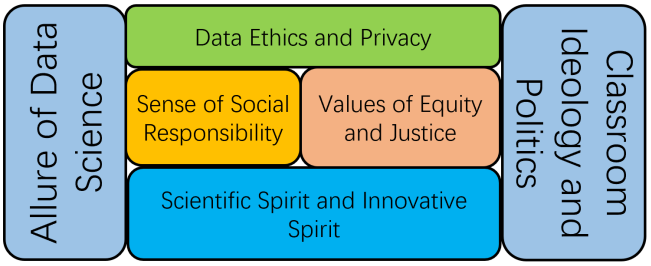

Figure 4. Classroom thinking and politics in data science

Data ethics and privacy protection. Data science cases are basically from life, so students must understand the importance of data ethics while learning knowledge. Real-world examples of data breaches and data abuse, such as the Facebook-Cambridge Analytica incident, can teach students the need for data privacy and set the bottom line for scientific research.

Enhance social responsibility. In the selection of data science projects, students can be encouraged to pay more attention to social welfare projects, such as using data science methods to deal with health care problems. By using the knowledge learned in class, students can solve social problems, enhance their sense of social responsibility, and in turn, promote students' interest in learning.

Cultivate values of fairness and fairness. Using some typical cases, such as gender discrimination in recruitment algorithms, racial bias in judicial decisions, etc., can let students know that the application of data science should consider diverse population backgrounds and establish the value of fairness and justice.

Promote the spirit of science and innovation. While learning classroom knowledge well, students' innovation ability should be cultivated, and students should be inspired to explore unknown areas through basic knowledge. For example, more students are encouraged to participate in some discipline competitions to stimulate students' sense of innovation, especially emphasizing multi-disciplinary cooperation (such as sociology, psychology, physics, etc.). In this way, students' scientific spirit, innovative spirit, and teamwork ability can be better cultivated.

Through the above methods, ideological and political education can be effectively integrated into the data science classroom to cultivate compound talents with both professional knowledge and noble character.

5. Summary and outlook

To sum up, building a data-related course teaching system that meets the needs of the digital intelligence era is a key

path to cultivating talents with a deep theoretical foundation and cutting-edge technology application ability. As a typical representative of the curriculum system, Introduction to Data Science should guide students to use knowledge flexibly and solve complex problems in reality. In particular, the close connection between mathematics and data science should be explained, so that students can understand the important value of mathematical theory. In addition, the organic integration of ideological and political education elements helps cultivate high-quality scientific and technological talents with both morality and talent and meets the needs of the national science and technology strategy and talent reserve.

Funding

This research was supported by the Fundamental Research Funds for the Central Universities (Grant No. 2652023054), the Subject Development Research Fund of China University of Geosciences (Beijing) in 2024 (Grant No. 2024XK106), the Beijing Higher Education Association project in 2023 (Grant No. MS2023276).

Disclosure statement

The authors declare no conflict of interest.

References

[1] Hu YH, 2023, Research on Talent Training Model Oriented to Adapt to Digital Transformation: Based on the Professional Perspective of Data Computing and Application. Journal of Mudanjiang College of Education, 2023(6): 7–9.

[2] Zhou JQ, Tan LJ, Li ZM, 2024, Discussion and Research on Training Model of Data Science and Big Data Professionals. Higher Education, 2024(4): 66–68.

[3] Shi YF, Liu XN, 2024, Teaching Exploration of Data Science Introduction Course under the Background of New Agriculture. Journal of Lanzhou Institute of Technology, 2024(31): 131–135.

[4] Yin ZX, Wang GQ, Fang T, 2023, Construction and Exploration of Data Computation and Application. University Mathematics, 2023(39): 20–24.

[5] Peng L, 2024, Exploration on Teaching Reform of Mathematical Analysis Course for Data Science Major. Specialty Construction and Teaching Reform, 2024(4): 57–59.

[6] Qin XP, Chen YG, Li CP, et al., 2018, “Data Science” Course Group and “Introduction to Data Science” Course Construction. Big Data, 2018(4): 19–28.