Research and Practice of Differentiated Teaching of Information Technology Public Course in Higher Vocational Colleges Based on Multi-dimensional Student Profiles

Xudong Pang

Shanghai Publishing and Printing College, Shanghai, China

pxd@sppc.edu.cn

Abstract. In this paper, practical issues in the teaching of public information technology courses in higher vocational colleges, such as significant differences in student foundations, broad coverage of majors, and difficulties in teaching evaluation are addressed. An educational big data-based student profiling system is developed. By integrating multi-dimensional features including students' background information, learning process data, and academic performances, methods such as correlation analysis, clustering analysis, and data visualization comparison are employed to achieve precise assessment and stratification of student abilities. The research results indicate that this profile system can effectively support teachers in carrying out differentiated teaching, improving teaching effectiveness, and providing data support for the research.

Keywords: Student Profiles; Education Big Data; Differentiated Teaching; Higher Vocational Education.

1 Introduction

Information technology courses are becoming increasingly important in modern society, and utilizing advanced digital technologies to empower the instructional design and ideological-political education development across various higher vocational courses is a current research priority [1].

With the development of massive enrollment expansion in higher vocational education, the student body has become increasingly diverse, and the specialized categories within these institutions are also complex and varied. The "Information Technology" series of public courses at our institution faces two main teaching challenges. Firstly, the courses cover students from all majors across the university, where in particular, new instructors often lack experience in understanding the academic proficiency levels of students from different majors, typically only gaining a clear picture by mid-term exams, which negatively impacts teaching effectiveness. Secondly, due to the combined influence of factors such as regional differences in information technology pen-

A. Y. M. A. Islam et al. (eds.), *Proceedings of the 2025 International Conference on Educational Technology and Management Information Systems (ETMIS 2025)*, Advances in Computer Science Research 129,
https://doi.org/10.2991/978-94-6239-630-2_16

etration, subject selection orientations, and admission scores, there is a significant disparity in the foundational knowledge of students upon entry. Practical teaching reveals that students' initial hands-on skill levels vary widely, and current differentiated teaching methods lack comprehensive data-driven approaches and digital estimation methods for assessing students' foundational abilities [2].

In this paper the currently popular concept of student profiling is further introduced [3], which involves filtering and secondary modeling development based on existing data from the school's digital information platform. It establishes iterative algorithms and empirical data models to ultimately form a dynamic data evaluation system [4]. The outcomes of this study will enable teachers to conduct precise analysis of student learning conditions and teaching assessments proactively and efficiently. Through big data digital empowerment, the system will guide and direct the specific implementation of differentiated teaching plans. Additionally, part of the visualized data analysis results will be presented to students, allowing them to recognize their strengths and weaknesses as well as track the development of their competency metrics [5].

2 Overview of Research Methods

The research subjects of this paper encompass students from diverse majors, with varied learning abilities and foundational knowledge. The overall methodology primarily employs quantitative statistical approaches, with qualitative research methods serving as a supplementary component. This study will conduct qualitative research to investigate the current state of the subject, verify the feasibility of data collection, and establish connections with the IT department and student affairs office. Based on the rich database collected from the school's existing digital information platform, correlation-based data filtering and secondary modeling development will be carried out. Feature information highly relevant to this course, such as student origins, ICT entrance test scores, and major-specific requirements will be extracted to establish correlation models [6]. Finally, database development and big data analysis techniques will be utilized for logical language software applications and quantitative data processing and visualization. For the research subjects, this study aims to explore a digital profiling approach for surveying and analyzing individual student data, seeking a scientifically sound multidimensional evaluation scheme.

The development of a student profiling system is generally discussed from three aspects: data organization, data platform, and application orientation. Commonly used technologies include big data development techniques, data analysis methods, evaluation algorithms, and data visualization technologies [7]. In this study, the student profiling system for the fundamental information technology course is structured across different development layers, primarily consisting of a feature data screening module, a data association modeling development module, a data analysis module, and a visual data presentation module. The feature data screening module primarily performs secondary screening of raw student data, selecting feature data relevant to the information technology course based on an empirical model which is mainly based on a foundational framework derived from data of two previous classes of students from the same major, encompassing their academic performance during the two semesters of the first year and their SCITE Level 1 exam results. The empirical model indicates that students from western provinces generally exhibit weaker foundational knowledge in information technology. Consequently, the coefficient for the region-of-origin factor in the comprehensive model requires appropriate adjustment. Given the difficulty in obtaining college entrance examination/transition examination scores, this data is mainly substituted by the basic score module from the subsequent " entrance test." For this demonstration class, there are no special addition or deduction factors related to major-specific requirements, so an average coefficient is applied. The screened data then proceeds to the data association modeling development module, where further association modeling is conducted between the categorized, layered, and structurally processed screened data and the learning outcomes or academic performance in the course, while simultaneously validating the effectiveness of the empirical model. Specific methods are employed for the data analysis processes within these modules. Developing the profiling system requires mastery of big data analysis techniques, the ability to use Java or Python for logical application development, and proficiency in data storage and processing tools. Finally, data visualization is implemented, and tools like Excel can, to some extent, meet current requirements.

3 Analysis of Teaching Data of Demonstration Classes

3.1 Partial Original Sample Data

Serial number	Comprehensive in peacetime	Midterm	End of term	General comments	Place of origin	Entrance test
1	93.4	95	68	75.9	Shanghai	100.0
2	94.0	80	83	83.5	Shanghai	100.0
3	94.2	88	43	57.1	Yunnan	90.0
4	96.0	98	85	88.7	Hebei	100.0
5	94.9	80	69	73.8	Zhejiang	100.0
6	87.5	97	89	90.5	Shanghai	90.0
7	96.3	97	88	90.6	Shanghai	100.0
8	93.9	96	79	83.9	Hebei	100.0
9	96.1	85	83	84.7	Zhejiang	100.0
10	78.1	77	52	59.6	Anhui	90.0
11	92.4	63	74	73.6	Henan	80.0
12	90.9	75	45	55.6	Guangdong	90.0
13	54.8	70	59	60.8	Shanghai	70.0
14	86.9	86	57	65.8	Shandong	100.0
15	92.7	84	65	71.6	Anhui	100.0
16	95.0	100	73	80.6	Shanghai	90.0
17	95.9	94	89	90.7	Shanghai	90.0
18	95.7	98	85	88.7	Shandong	100.0
19	91.4	100	80	85.1	Shandong	100.0
20	95.1	98	92	93.5	Shandong	100.0
21	95.8	91	93	92.9	Shanghai	100.0
22	80.6	85	76	78.3	Shanghai	100.0
23	88.0	80	43	54.9	Yunnan	100.0
24	95.9	82	79	81.3	Shanghai	100.0
25	93.9	88	80	83.0	Zhejiang	90.0
26	91.2	88	68	74.3	Shanghai	100.0
27	95.4	98	81	85.8	Henan	100.0
28	84.1	84	63	69.3	Anhui	60.0
29	95.2	93	82	85.5	Shanghai	100.0
30	96.5	98	90	92.3	Shanghai	100.0
31	79.7	90	52	62.4	Zhejiang	90.0
32	94.5	98	92	93.5	Zhejiang	100.0
33	93.5	92	59	69.0	Yunnan	90.0
34	95.7	74	89	86.7	Zhejiang	100.0
35	94.3	98	55	67.5	Shandong	100.0
36	96.4	96	87	89.7	Shanghai	100.0
37	82.4	86	75	77.9	Shanghai	100.0
38	98.4	98	92	93.8	Shanghai	100.0
39	92.4	93	55	66.3	Shandong	100.0
40	86.9	79	58	65.1	Guangdong	100.0
41	96.1	72	89	86.3	Shanghai	100.0
42	95.4	90	91	91.2	Hebei	100.0
43	95.2	95	86	88.7	Anhui	100.0
44	85.1	92	81	83.6	Shandong	100.0
45	96.1	82	83	84.1	Shanghai	100.0
46	87.6	50	58	59.4	Hebei	100.0
47	94.7	94	82	85.7	Shanghai	100.0
48	94.6	73	82	81.5	Zhejiang	100.0
49	95.0	66	78	77.3	Zhejiang	80.0
50	92.1	96	60	70.4	Shanghai	100.0

Fig. 1. Representative data samples

This study employs the International Film and Animation Program (50 students) taught by the author as a demonstration cohort. Through the student profiling system, multidimensional data analysis is conducted to investigate differences in learning capabilities among students, validate the effectiveness of differentiated instructional strategies

[5], and provide data-driven support for the development of curriculum-based ideological and political education. Representative data samples are shown in Fig. 1.

3.2 Algorithm Model Construction and Analysis

First, the Correlation analysis is studied in Table 1 where Pearson's correlation coefficient was used to analyze the correlation between the characteristics and the total score, and the results were as follows:

From the results a conclusion can be drawn that the daily attendance, online study activities, and Excel proficiency show strong positive correlations with overall scores, with the midterm score serving as the key predictor of predicting the final performance.

Table 1. The correlation between the characteristics and the total score.

Features	Correlation Coefficient	Significance (p-value)
Attendance	0.82	<0.01
Learning through activity	0.76	<0.01
Excel scores	0.68	<0.01
Midterm results	0.91	<0.01

To achieve more scientific student stratification, this study employs the K-Means clustering algorithm. Prior to modeling, key features such as "overall score," "Excel/Word practical scores," and "online-study platform activity" were standardized using Z-score normalization to eliminate scale differences. Through elbow method analysis, the inflection point of the sum of squared errors was observed at K=3, confirming that dividing students into three groups is most appropriate. The final clustering results are stable, with a silhouette score of 0.52, indicating clear separation between clusters. The specific stratification is as follows:

- Group A (15 students): Overall score ≥ 90, excellent performance in Excel/Word, high activity on the online study platform;

- Group B (20 students): Overall score between 75-89, varying practical skills, significant fluctuations in midterm scores;

- Group C (5 students): Overall score ≤70, low attendance rate, poor assignment completion.

3.3 Data Visualization and Teaching Suggestions

First, group profiles and their corresponding instructional strategies are analyzed in Table 2. Based on the characteristics of the three clustered groups identified through the previous clustering algorithm, specific instructional strategies are proposed respectively as follows:

Table 2. Group profiles and their corresponding instructional strategies.

Grouping	Proportion	Core issues	Teaching suggestions
Group A	30%	High demand for advanced content	Add elective modules like AI-assisted office skills, guide participation in competitions
Group B	40%	Performance decline after midterms	Implement "fast-slow student pairing" model, assign tiered assignments
Group C	10%	Weak foundations, lack of motivation	Develop "micro-lectures + targeted practice" learning modules, provide additional tutoring

Then, the comparative stacked line chart of the demonstration class's entrance test scores and overall final scores is shown in Fig. 2. First, it should be clarified that the entrance test scores are not included in the overall final scores, serving instead as an independent module reflecting students' foundational academic and informational competencies. The reason for using a stacked line chart here is that we aim to observe whether there is a positive correlation between the entrance test and the final overall score of Information Technology course, without comparing specific numerical values. This allows both lines to be displayed in a single chart. From the comparison, it can be seen that the orange entrance test curve aligns well with the blue final overall score curve. This indicates that students with lower entrance test scores, who had relatively weaker foundational IT skills, generally achieved lower final grades in the introductory IT course. However, there are cases where students improved through their own efforts. For instance, the student with the ending NO.28 scored 60 on the entrance test, which may have been due to forgetting to submit a corresponding image. After all, this entrance test only took 45 minutes and was relatively simple. Data from other classes also confirms that this entrance test was well-designed, engaging, and discriminative.

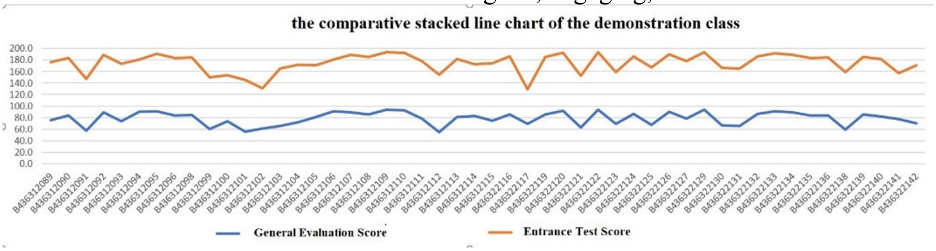

Fig. 2. Demonstration Class Entrance Test Score VS General Evaluation Score

By analyzing the relationship between students' places of origin and their academic performance in the data, the following conclusions can be drawn that students from Shanghai demonstrated the strongest overall performance, achieving the highest average score (83.1) with relatively concentrated score distribution (indicated by a small standard deviation). Students from Hebei, Shandong, and Zhejiang performed at a moderate level, though Zhejiang included some individual high scorers (93.5). Students from Anhui, Yunnan, and Guangdong scored on the lower end, particularly those from Yunnan (average score of 56.5). Differences in sample sizes should be noted: Shanghai

(29 cases) provided the most robust data, while other provinces had smaller sample sizes, meaning their statistical power is for reference only and lacks empirical guiding significance.

Through the profiling system, teachers can accurately identify the ability distribution within a class before the course begins and dynamically adjust the teaching pace (e.g., accelerating and expanding content for Group A while reinforcing foundational knowledge for Group C). This enables data-driven decision-making in instructional strategies. The implementation of tiered instruction has demonstrated significant effectiveness in differentiated teaching.

4 Conclusion

This profiling system enables teachers to gain early insights into students' admission scores, geographical backgrounds, and foundational knowledge, allowing for the implementation of differentiated, targeted, and matched instructional strategies tailored to students at varying ability levels. Furthermore, the system can be continuously refined and supplemented through preliminary teaching research, establishing a digital archive for assessing IT capabilities throughout the three-year university period. This fosters the creation of a modern, student-centered, and digitally-driven lifelong competency development record. By integrating online and offline performance as well as in-class and extracurricular activities, it deepens our understanding of each student.

Acknowledgments

This work is sponsored by the Shanghai Municipal Education Commission Young Teacher Development and Funding Program under SPPC No. Y0B-0203-25-14-04y.

References

1. Anisimova A N , Efremova Y I .Digital Transformation of Vocational Education: Challenges of Modern Society[C]. International Scientific Conference Digital Transformation of the Economy: Challenges, Trends, New Opportunities .Springer, Cham (2022).
2. Kong M .Research on Big Data Analysis and Teaching Evaluation of Big Data Accounting Major in Higher Vocational Colleges[J].Education Reform and Development, 7(1):72-79 (2025).
3. Mbodila M , Mbaleki C . The value of student profiling in enhancing academic support within the higher education sector [J]. EUREKA: Social & Humanities (2025).
4. Zaitie C .Feasibility Research and Practical Exploration on Higher Vocational Education and Differentiated Training in County-Level Cities under Students' Personalized Demands[J].Journal of Shazhou Professional Institute of Technology (2019).
5. Kokkinos, Theodore , and P. Gakis . Student teachers' differentiated teaching practices for high-achieving students. Journal of Further and Higher Education (2020).

6. Engel H .Institutional Profiling for Educational Development: Identifying Which Conditions for Student Success to Address in a Given Educational Setting—A Case Study[J].Journal of College Student Retention: Research, Theory & Practice, 2025, 26(4):1100–1122(2025).
7. Hung J L , Zhang K .Examining mobile learning trends 2003–2008: a categorical meta-trend analysis using text mining techniques[J].Journal of Computing in Higher Education, 2012, 24(1):1-17(2012).