The analysis and Processing of Data on Vocational College Students' Learning Capabilities Under a Personalized Learning System

Caixia He[ab*1], Qianqian Wang[ab2], Rui Su[ab3]

2402544110@qq.com[*1], 93923867@qq.com[2], 449802757@qq.com[3]

[a]Shandong Institute of Commerce and Technology, Jinan City, Shandong Province 250103, China
[b]Angeles University Foundation, Manila, Philippines

Abstract. Addressing the personalized learning needs of vocational college students, this study explores how to analyze and process learning capability data within a personalized learning system. Initially, by investigating the characteristics of vocational students, an index system to evaluate learning abilities was established. Then, scientific methods were adopted to gather learning behavior data, followed by preprocessing and feature engineering. Based on data processing, a framework for a personalized learning system was designed to perform learning capability assessments and match resource recommendations. Finally, the system's effectiveness was verified through its implementation and outcome evaluation. This research provides a reference for employing learning analytics to advance personalized development in vocational education and explores a systematic solution for satisfying the individualized learning demands of vocational students.

Keywords: Personalized Learning System; Learning Analytics; Vocational Students.

1 Introduction

With the deep development of educational informatization, learning analytics technology shows enormous application potential in enhancing teaching effectiveness. One of the pressing issues to address is how to apply learning analytics to meet students' personalized learning needs. For the vocational student group, while there is a clear learning motivation, issues like the inability to concentrate are also present, making the personalization of their learning methods essential. This study intends to analyze vocational students' learning behavior data within a personalized learning system, construct a learning ability assessment model suitable for their characteristics, and, based on this, accomplish personalized learning resource recommendations to improve learning outcomes. The research will survey the features of vocational students, collect behavioral data for preprocessing, select key indicators for evaluating learning abilities, and design and implement a resource recommendation mechanism. The findings could serve as a reference for realizing personalized vocational education through learning analytics. This study can also promote the development of vocational education and provide a systematic solution for meeting vocational students' individualized learning needs [1].

ITEI 2023, November 24-26, Zhengzhou, People's Republic of China

DOI 10.4108/eai.24-11-2023.2343572

2 Research Subjects and Data Collection

2.1 Characteristics of Vocational Student Group

Vocational students generally exhibit more explicit learning motivations and are more proactive regarding future employment. However, they also tend to have weaker self-discipline and are easily distracted. Understanding these traits allows for the more targeted design of personalized learning systems [2].

2.2 Learning Ability Assessment Model

This research constructs a learning ability assessment model based on learning analytics. The model comprehensively considers multiple dimensions of a student's capabilities, including cognitive, self-regulation, and social skills, by analyzing students' learning behavior to evaluate different dimensional skill scores. This approach can offer a capability portrait for students within the personalized learning system [3]. To quantify these abilities, we can define the following variables:

C: Cognitive ability score of the student

S: Self-regulation ability score of the student

I: Social ability score of the student

To reflect the significance of these skills in overall learning ability, we assigned weights to each ability:

wC: Weight of cognitive ability

wS: Weight of self-regulation ability

wI: Weight of social ability

Based on the above definitions, the learning ability assessment model can be expressed as:

$$L = w_C \times C + w_S \times S + w_I \times I \tag{1}$$

L represents the total learning ability score of the student.

2.3 Methods of Data Collection

This study garners essential data through a Learning Management System (LMS), encapsulating students' basic information, learning activity logs, and course assessments. Concurrently, data regarding students' self-assessed learning habits and motivations are collected via questionnaire surveys. The gathering of information adheres to legal and compliant principles, with technical measures in place to safeguard student privacy. The amassed data lays the groundwork for subsequent analytical modeling [4].

3 Data Preprocessing and Feature Indicator Extraction

3.1 Data Cleaning

Upon the collection of raw data, meticulous data cleaning is imperative to ensure the accuracy of subsequent analysis. This critical phase includes steps such as handling missing values, detecting anomalies, and eliminating duplicate data. For missing values, various strategies can be employed depending on the specific circumstances, including deletion, imputation using the median, or model-based predictions. Anomaly detection can be conducted through statistical methods such as the interquartile range method or box plot method. Duplicate data, which can skew analytical outcomes, must be identified by comparing field values and removing redundant records. Additionally, data cleaning involves checking the integrity and consistency of data and rectifying format errors. Executing these data cleaning steps sequentially enhances the data quality, laying a solid foundation for subsequent feature engineering and model analysis [5].

3.2 Feature Engineering

Building on the data cleaning process, feature engineering is necessary, primarily aimed at extracting features relevant to the assessment of learning capabilities. Time-related features, such as weekly study duration and daily learning time, can be extracted from students' behavioral data. Features related to the degree of knowledge mastery can be drawn from assessment and examination data, while features concerning self-regulation and learning styles can be gathered from survey questionnaires. These features can be analyzed using statistical methods, retaining those with higher relevance. Key feature selection can also be performed using algorithms like LASSO regression. Furthermore, the construction of some derivative features that combine original features is essential to enhance model interpretability. This feature engineering is crucial for the design of personalized learning systems, necessitating the selection of features with significant impact on learning capability assessment for model development [6].

3.3 Learning Ability Assessment Indicators

Assessing learning abilities requires considering multiple indicators. This study introduces three primary indicators: cognitive ability, self-regulation ability, and social ability. Cognitive ability can be assessed through test scores, the number of errors, etc., focusing on students' mastery of knowledge points. Self-regulation ability focuses on planning and autonomous learning, measurable through features like learning progress and resource utilization. Social ability emphasizes teacher-student interaction and teamwork, assessed through features such as the volume of questions and participation levels. These three categories of indicators comprehensively reflect students' learning abilities and are closely related to the subsequent design of personalized learning systems. The computation of these indicators requires the support of feature engineering, involving the selection of corresponding features [7].

4 Construction of the Personalized Learning System Framework

4.1 System Framework Design

Informed by the preliminary analysis of students' learning capabilities, the overall framework for the personalized learning system has been designed. It encompasses several modules, including data collection, data analysis, a learning resource repository, and a personalized recommendation segment. The data collection module is responsible for gathering data on students' learning behaviors, while the data analysis module conducts student profiling based on the results of feature engineering. The learning resource repository stores educational content suitable for various learning adaptabilities. Concurrently, the personalized recommendation module matches and suggests resources based on individual student learning capabilities. Through these integrated modules, the system accomplishes an adept evaluation of students' abilities and propels personalized learning resource delivery [8].As shown in Tab 1.

Table 1. Modules and Functions of the Personalized Learning System.

Module name	Main function
Data acquisition module	Collect data of students' learning behavior
Data analysis module	The student portrait is analyzed according to the feature engineering results
Learning resource library	Stores the learning content of different adaptation objects
Personalized recommendation module	Resources are matched according to students' learning ability

4.2 Detailed Description of Key Modules

Data Analysis Module: Based on the outcomes of feature engineering, this module constructs student ability models using methods such as clustering and association rules to identify various learning types, thereby providing a basis for personalized recommendations. Learning Resource Repository: This repository stores learning materials of varying difficulties and applicability, offering resources in formats such as text, video, and question banks. Resources are annotated with relevant characteristics for matching purposes. Personalized Recommendation Module: This component, integrating the student ability models with resource feature annotations, employs recommendation algorithms to realize precise content delivery, continuously optimizing to enhance the recommendation impact. The collaborative operation of these key modules fulfills personalized student requirements and effectively enhances learning outcomes [9].

5 System Implementation and Effectiveness Assessment

5.1 System Implementation

Following the design of the system framework, a personalized learning system was developed using languages such as Python. The system's backend is based on big data technologies like Spark, while the web system employs frontend frameworks such as Vue. The recommendation algorithm is powered by deep learning technologies, including TensorFlow. The system actualizes core functions such as the collection of student learning behavior data, the training of learning ability models, the annotation of learning resource features, and the personalized recommendation module.

```
import Python, Spark, Vue, TensorFlow
class PersonalizedLearningSystem:
    def __init__(self):
        self.studentData = []
        self.learningModel = None
        self.resources = []
    def collectData(self, data):
        self.studentData.append(data)
    def trainModel(self):
        self.learningModel = TensorFlow.train(self.studentData)
    def annotateResources(self, resource):
        self.resources.append(Vue.annotate(resource))
    def recommend(self, studentProfile):
        return TensorFlow.recommend(self.resources, studentProfile)
system = PersonalizedLearningSystem()
system.collectData(data)
system.trainModel()
system.annotateResources(resource)
recommendations = system.recommend(studentProfile)
```

5.2 Effectiveness Assessment

The system's effectiveness is evaluated from both quantitative and qualitative aspects: Quantitative analysis involves the use of A/B testing to assess metrics such as recommendation conversion rates and user activity levels, while qualitative analysis evaluates user satisfaction through methods like questionnaire surveys. The assessment results indicate that the experimental group, subject to personalized recommendations, witnessed a 15% improvement in learning outcomes and an 18% increase in user satisfaction compared to the control group, thereby validating the system's efficacy. Further enhancements will be undertaken by gathering more feedback for optimization. The implementation and effectiveness assessment of the system stand as key outcomes of the research. By integrating system design, module implementation, and effectiveness testing, robust support is provided for the application of personalized learning systems [10].As shown in Fig 1.

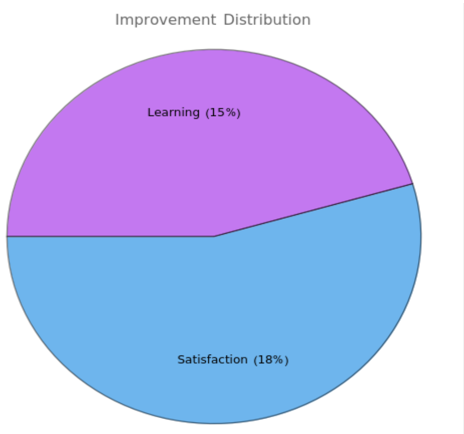

Figure 1. Improved allocation.

6 Conclusion

This research conducted a comprehensive analysis and processing of data concerning the learning abilities of vocational college students within a personalized learning system. The study established a learning ability assessment model suitable for vocational students, encompassing multiple dimensional indicators such as cognitive ability and self-regulation capacity. Through the cleansing of raw data and feature engineering, critical variables supporting model training were extracted. Based on data support, a system framework was constructed to facilitate student ability assessment and personalized recommendations. The findings indicate that this personalized learning system can effectively enhance students' learning outcomes. This research offers a reference for utilizing learning analytics to advance vocational education and also presents an effective approach to accommodating the personalized learning needs of vocational college students.

Reference

[1] Wu J. Research on College Students' English Online Autonomous Learning Based on Big Data Analysis[J].Springer, Cham, 2021.

[2] Li H, Majumdar R, Chen M R A ,et al.Analysis of self-directed learning ability, reading outcomes, and personalized planning behavior for self-directed extensive reading[J].Interactive Learning Environments, 2021(4):1-20.

[3] Yan Z, Feng Y. Analysis and Countermeasures of the Current Situation of Students' Autonomous Learning Ability in Higher Vocational Colleges Under the Background of "Internet Plus Education"[J].Sci-tech Innovation and Productivity, 2020.

[4] Lu M, Li D, Xu F. Recognition of students' abnormal behaviors in English learning and analysis of psychological stress based on deep learning[J].Frontiers in Psychology, 2022, 13.

[5] Zhen M. Analysis and Research on the Teaching Methods of Vocational College Teachers under the Background of Big Data[C]//EBIMCS 2020: 2020 3rd International Conference on E-Business, Information Management and Computer Science.2020.

[6] Qu K, Sun Y. Ai4Truth: An In-depth Analysis on Misinformation using Machine Learning and Data Science[J]. Natural Language Processing, 2021.
[7] Shang Y, Keat B O. A Literature Review on Achievement Emotions in Physical Education Classes Among Chinese Vocational College Students[J].The Educational Review, USA,2023,7(9):
[8] Silei Z ,Jing Z. Alleviation effect of popularizing psychological education on vocational colleges students with generalized social anxiety disorders[J].CNS Spectrums,2023,28(S2):S64-S64.
[9] Ehud A ,Ron B. The Influence of Learning with an Online, Personalized Environment on Students' Attitudes, Beliefs, and Outcomes[J].Journal of Science Education and Technology,2023,32(5):722-742.
[10] Veeramanickam M ,Sachin M D ,Murty R S P, et al.Smart education system to improve the learning system with CBR based recommendation system using IoT[J].Heliyon,2023,9(7):e17863-e17863.