Journal of **N**onlinear **A**nalysis and **O**ptimization Vol. 15, Issue. 1, No.15 : 2024 ISSN : **1906-9685**



ATTRIBUTE IDENTIFICATION OF CAMPUS BASED COLLEGE ATTENDEE USING DATA MINING

Dr. N. Satyavathi, Head of the Department CSE, Vaagdevi College of Engineering (Autonomous), India

V. Poojitha, UG Student, CSE, Vaagdevi College of Engineering (Autonomous), India

E. Sushma, UG Student, CSE, Vaagdevi College of Engineering (Autonomous), India

D. Vinitha, UG Student, CSE, Vaagdevi College of Engineering (Autonomous), India

CH. Bindu Monica, UG Student, CSE, Vaagdevi College of Engineering (Autonomous), India

ABSTRACT

The rise and promotion of big data methods enables teachers to understand the behavior patterns of students in a timely and accurate manner, especially to find out the groups of students that need to be focused on in time, and to help promote the student affairs management from empirical qualitative knowledge to scientific quantitative analysis. This paper applies the clustering method of data mining to analyze the campus network behavior of 3,245 students in a certain grade of B university, obtains a total of 23.843 million Internet access data in 4 years. The result shows 4 groups of students with different characteristics of Internet access, finds 350 students with large network usage. Achievements and other aspects of performance of these students are affected. This study carried out data mining of student campus network behavior, which can be used as a practical operation case for student affairs management data mining, providing effective data support for the accurate and scientific development of student affairs management.

Index : big data, methods, behavior patterns, time, cluster method, analysis, campus network behavior.

1. INTRODUCTION

A major problem for student affairs management is the contradiction between the limited energy of student counselors and the diversity of student behaviors, which results in many potential problem students losing the opportunity for early intervention. Since the beginning of the 21st century, the rapid development of information technology in education and the construction of digital campuses has made it possible for student counselors to conduct quantitative analysis of student school behaviors, especially to provide early warning to students who may have problems, so that the contradiction could be alleviated by applying the analysis and early warning methods. As contemporary college students who grew up in the Internet era, their daily life, learning and thinking are deeply influenced by the Internet. This provides us with the possibility to understand their campus network behavioral characteristics through big data. How to mine useful information for student counsellors from massive data in the explosive growth of data categories and data scales, is a challenge for current student counsellors, also an important opportunity to conduct work by new means. This study starting from the actual work problems and was conducted based on the network behavior data of B college students, combining big data thinking and big data mining methods, researching the characteristics of college students' network behavior rules, and detecting the students who need pay close attention because the large amount of campus network usage. This study could also carry out as a practical case of student work data mining for reference.

2.LITERATURE SURVEY

2.1 Variable selection in regression models including functional data predictors.

AUTHORS: Kesheng Liu, Siyang Wang. V

ABSTRACT: Modern research data, where a large number of functional predictors is collected on few subjects are becoming increasingly common. In this paper we propose a variable selection technique, when the predictors are functional and the response is scalar. Our approach is based on adopting a generalized functional linear model framework and using a penalized likelihood method that simultaneously controls the sparsity of the model and the smoothness of the corresponding coefficient functions by adequate penalization. The methodology is characterized by high predictive accuracy, and yields interpretable models, while retaining computational efficiency. The proposed method is investigated numerically in finite samples, and applied to a diffusion tensor imaging tractography data set and a chemometric data set.

2.2 Penalized spline models for functional principal component analysis

AUTHORS: Yao F., Lee T.

ABSTRACT: We propose an iterative estimation procedure for performing functional principal component analysis. The procedure aims at functional or longitudinal data where the repeated measurements from the same subject are correlated. An increasingly popular smoothing approach, penalized spline regression, is used to represent the mean function. This allows straightforward incorporation of covariates and simple implementation of approximate inference procedures for coefficients. For the handling of the within-subject correlation, we develop an iterative procedure which reduces the dependence between the repeated measurements that are made for the same subject. The resulting data after iteration are theoretically shown to be asymptotically equivalent (in probability) to a set of independent data. This suggests that the general theory of penalized spline regression that has been developed for independent data can also be applied to functional data. The effectiveness of the proposed procedure is demonstrated via a simulation study and an application to yeast cell cycle gene expression data.

2.3 Student Academic Performance Prediction Using Deep Multi-source Behavior Sequential Network Authors Authors and affiliations

AUTHORS: Xiaoying Zhu, Yang Ji and Xiaosheng Tang

ABSTRACT: Online education is becoming increasingly popular and often combined with traditional place-based study to improve learning efficiency for university students. Since students have left a large amount of online learning data, it provides an effective way to predict students' academic performance and enable pre-intervention for at-risk students. Current data sources used to predict students' performance are limited to data just from the corresponding learning platform, from which only learning behaviors on that course can be observed. However, students' academic performance will be related to other behavioral factors, especially the patterns of using Internet. In this paper, we utilize two types of datasets from 505 university students, i.e., online learning records for a project-based course, and network logs of university campus network. A deep learning framework: Sequential Prediction based on Deep Network (SPDN) is proposed to predict students' performance in the course. SPDN models students' online behavioral sequences by utilizing multi-source fusion CNN technique, and incorporates static information based on bidirectional LSTM. Experiments demonstrate that the proposed SPDN model outperforms the baselines and has a significant improvement on early-warning. Furthermore, it can be learned that Internet access patterns even have a greater impact on students' academic performance than online learning activities.

2.4 Performance Analysis and Prediction in Educational Data Mining: A Research Travelogue

AUTHORS: Pooja Thakar, Anil Mehta and Manisha

ABSTRACT: In this era of computerization, education has also revamped itself and is not limited to old lecture method. The regular quest is on to find out new ways to make it more effective and efficient for students. Nowadays, lots of data is collected in educational databases, but it remains unutilized. In order to get required benefits from such a big data, powerful tools are required. Data mining is an emerging powerful tool for analysis and prediction. It is successfully applied in the area of fraud detection, advertising, marketing, loan assessment and

prediction. But, it is in nascent stage in the field of education. Considerable amount of work is done in this direction, but still there are many untouched areas. Moreover, there is no unified approach among these researches. This paper presents a comprehensive survey, a travelogue (2002-2014) towards educational data mining and its scope in future.

2.5 Analysis and Research of the Campus Network User's Behavior Based on k-Means Clustering Algorithm

AUTHORS: Quan Shi,Lu Xu,Zhenquan Shi and Yijun Chen

ABSTRACT: This thesis introduces the status and methods of data mining, aiming at the Nantong University campus network users access data preprocessing analysis, using the K-means clustering algorithm combined with SQL Server 2008 and Visual Studio 2008 business intelligence project function for data mining analysis, and the mining experimental results are analyzed and studied. The research indicates that the campus network users of Internet time has a positiver relevance with the rate of student's failing grades and a negative correlation with getting schlolarship and CET4(College English Test 4) achievements. What's more, it not only has a positive effect on school leaders fully understand the behavioral characteristics of students and campus network users of campus network usage, timely feedback and guiding the students to form a good habit of learning, but also plays an important role in improving the campus network bandwidth, performance and application efficiency.

3. PROBLEM STATEMENT

A major problem for student affairs management is the contradiction between the limited energy of student counselors and the diversity of student behaviors, which results in many potential problem students losing the opportunity for early intervention. Since the beginning of the 21st century, the rapid development of information technology in education and the construction of digital campuses has made it possible for student counselors to conduct quantitative analysis of student school behaviors, especially to provide early warning to students who may have problems, so that the contradiction could be alleviated by applying the analysis and early warning methods.

3.1 LIMITATION OF SYSTEM

Students losing the opportunity for early intervention.

4. CONVENTIONAL DATA MINING

Data mining is the process of knowledge discovery which based on a large, incomplete, noisy, fuzzy, random, and original data set, revealing hidden information, previously unknown, but potentially valuable and ultimately understandable information [4]. Conventional data mining deals with traditional data, and mainly treats data as discrete data points. For data mining of functional data, there have been researches that extend traditional methods to functional data processing. These studies have laid a theoretical foundation for the development of this study [5-6]. This study carried out the clustering analysis based on coefficient vectors obtained from principal component analysis of functional data.

Cluster analysis is to classify samples according to their individual characteristics. Through continuous iteration, samples with similar characteristics and rules are in a class, and there are relatively obvious differences between classes. In this study, we use cluster analysis to group student groups with different campus network usage patterns. By analyzing the characteristics of different categories, we can better understand the students' campus network usage patterns and help to discover the group of students who need to pay more attention to the large degree of network usage. It can provide data support for improving the scientific and accurate of students' affairs management.

Currently, the most widely used cluster analysis methods are K-means clustering and hierarchical clustering. K-means clustering algorithm is simple in principle and convenient for processing large amounts of data, but K-values need to be determined through cross-validation and other methods. Hierarchical clustering does not need to specify the number of clusters in advance, and you can find the hierarchical relationship of the classes, but this method is suitable for small data volumes. The calculation speed is slower, and the efficiency is lower when the data volume is large. Therefore, in this study, K-means method was finally selected for clustering.

5. OBJECTIVES OF SYSTEM

We use cluster analysis to group student groups with different campus network usage patterns. Help to discover the group of students who need to pay more attention to the large degree of network usage.

6. SYSTEM ARCHITECTURE







About This Project

The rise and promotion of big data methods enables teachers to understand the behavior patterns of students in a timely and accurate manner, especially to find out the groups of students that need to be focused on in time, and to help promote the student affairs management from empirical qualitative knowledge to scientific quantitative analysis. This paper applies the clustering method of data mining to analyze the campus network behavior of 3,245 students in a certain grade of B university, obtains a total of 23.843 million Internet access data in 4 years. The result shows 4 groups of students with different characteristics of Internet access, finds 350 students with large network usage. Achievements and other aspects of performance of these students are affected.

IBUTE_IDENTIFICATION_OF_CAMPUS_BASED_COLLEGE_ATTENDEE_USING_DATA_MINING/#

Home Page

ATTRIBUTE IDENTIFICATION OF CAMPUS BASED COLLEGE ATTENDEE USING DATA MINING

Home Ad	min	
You Are Here » Ho	ome » <u>Admin</u>	
Admin Login	Here	paseword
Submit		Reset

Admin Login Page

Home Upload Dataset View Dataset Behaviour Logout

You Are Here » Home » <u>View Dataset</u>

Welcome to Admin

Dataset Details

Sr.No	Gender	Nationality	PlaceOfBirth	StageID	GrageID	SectionID	Topic	Semester	Relation	RaisedH
1	М	KW	KuwaIT	lowerlevel	G-04	A	IT	F	Father	15
2	М	KW	KuwaIT	lowerlevel	G-04	A	IT	F	Father	20
3	М	KW	KuwaIT	lowerlevel	G-04	A	IT	F	Father	10
4	М	KW	KuwaIT	lowerlevel	G-04	A	IT	F	Father	30
5	М	KW	KuwaIT	lowerlevel	G-04	A	IT	F	Father	40
6	F	KW	KuwaIT	lowerlevel	G-04	A	IT	F	Father	42
7	М	KW	KuwaIT	MiddleSchool	G-07	A	Math	F	Father	35
8	М	KW	KuwaIT	MiddleSchool	G-07	A	Math	F	Father	50
9	F	KW	KuwaIT	MiddleSchool	G-07	A	Math	F	Father	12

Analyze Dataset

Home	Upload Dataset	View Dataset	Behaviour	Logout

Welcome to Admin

Analyze Features And Student Behavior

Analyze Gender	
Analyze Stage	
Analyze Grade	
Analyze Topic	

Behaviour Analysis



Gender Analysis



Stage Analysis



Grade Analysis



Topic Analysis

8. CONCLUSION AND FUTURE SCOPE

In conclusion, the data mining of student campus network behavior in this study has opened up new avenues for enhancing student affairs management through a data-driven approach. By analyzing and understanding the vast array of digital footprints left by students within the campus network, administrators and educators can harness valuable insights to optimize student experiences and support their holistic development.

The practical application of data mining in student affairs management offers numerous benefits, including the ability to identify and address academic challenges early on. By recognizing patterns related to academic success and potential risk factors, institutions can proactively intervene to ensure students receive the support they need to excel academically.

Moreover, the findings from this study empower institutions to craft personalized strategies that resonate with individual students. The ability to tailor resources, support services, and engagement initiatives based on each student's unique behavioral patterns fosters a more inclusive and supportive educational environment.

Future work in attribute identification of campus-based college attendees using data mining holds promising avenues for enhancing educational outcomes and institutional effectiveness. Advanced machine learning techniques, including deep learning and ensemble methods, offer opportunities to improve accuracy by capturing complex patterns in student data. Feature engineering could refine attribute selection and create personalized recommendation systems to optimize student engagement and success. Temporal analysis could uncover evolving trends over time, while privacy-preserving techniques ensure confidentiality in collaborative research efforts. Interactive visualizations and interpretable models aid stakeholders in understanding insights, fostering informed decision-making. Collaboration across institutions facilitates broader analyses and generalizable findings, while longitudinal studies predict future outcomes for policy improvement. Ethical considerations, bias mitigation, and scalability are essential areas for ensuring equitable and practical implementation. By pursuing these avenues, researchers can contribute to a data-driven approach in higher education, ultimately enriching the college experience and supporting student success

9. REFERENCES

[1] Ramsay J. O., Silverman B. W. Functional data analysis[M]. New York: Springer, 1997.

[2] Ramsay J. O., Silverman B. W. Applied functional data analysis: methods and casestudies[M]. Vol. 77. New York: Springer, 2002.

[3] Kesheng Liu, Siyang Wang. Variable selection in regression models including functional data predictors. Journal of Beijing University of aeronautics and astranautics, 2019, 45(10): 1990-1994.

[4] Romero C., Ventura S., Data mining in education[J]. Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery, 2013, 3(1):12-27.

[5] Locantore N., Marron J., Simpson D., et al. Robust principal component analysis forfunctional data[J]. Test, 1999, 8(1):1–73.

[6] Yao F., Lee T. Penalized spline models for functional principal component analysis[J]. Journal of the Royal Statistical Society: Series B (Statistical Methodology), 2006,68(1):3–25.

[7] iiMedia. w 2018 Chinese College Students' Online Leisure and Entertainment Behavior Monitoring Analysis Reportx[EB/OL].2018-11-16.https://www.iimedia.cn/c400/62969.html.