As per UGC guidelines an electronic bar code is provided to seure your paper

International Journal for Modern Trends in Science and Technology, 8(06): 192-194, 2022

ISSN: 2455-3778 online
DOI: https://doi.org/10.46501/IJMTST0806030
Available online at: http://www.ijmtst.com/vol8issue06.html

Academic Performance Prediction Based on Multi Sorce Multi Feature Behavioural Data

K.Venkateswara Rao[1], Y.Dhana Lakshmi[2], N.Roshitha[2], P.Kavya Sree[2], T.Laya[2]

[1]Assistant Professor, Department of Computer Science & Engineering, Sri Vasavi Institute of Engineering & Technology, Pedana, A.P, India
[2]Department of Computer Science & Engineering, Sri Vasavi Institute of Engineering & Technology, Pedana, A.P, India

To Cite this Article
K.Venkateswara Rao, Y.Dhana Lakshmi, N.Roshitha, P.Kavya Sree and T.Laya. Academic Performance Prediction Based on Multi Sorce Multi Feature Behavioural Data. International Journal for Modern Trends in Science and Technology 2022, 8(06), pp. 192-194. https://doi.org/10.46501/IJMTST0806030

Article Info
Received: 09 May 2022; Accepted: 02 June 2022; Published: 07 June 2022.

ABSTRACT

Digital data trails from disparate sources covering different aspects of student life are stored daily in most modern university campuses .However, it remains challenging to Combine these data to obtain a holistic view of a student ,Use these data to accurately predict academic performance , and Use such predictions to promote positive student engagement with the university. To initially alleviate this problem, in this article, a model named Augmented Education (Augment ED) is proposed .In our study, first, an experiment is conducted based on a real-world campus dataset of college students N , that aggregates multisource behavioral data covering not only online and offline learning but also behaviors inside and outside of the classroom. Specifically, to gain in-depth insight into the features leading to excellent or poor performance, metrics measuring the linear and nonlinear behavioral changes (e.g., regularity and stability) of campus lifestyles are estimated; furthermore, features representing dynamic changes in temporal lifestyle patterns are extracted by the means of long short-term memory (LSTM) Second, machine learning-based classification algorithms are developed to predict academic performance. Finally, visualized feedback enabling students (especially at-risk students) to potentially optimize their interactions with the university and achieve a study-life balance is designed. The experiments show that the Augment ED model can predict students' academic performance with high accuracy.

1. INTRODUCTION

As an important step to achieving personalized education, academic performance prediction is a key issue in the education data mining field. It has been extensively demonstrated that academic performance can be profoundly affected by the following factors:
Students' Personality (e.g., neuroticism, extraversion, and agreeableness) Personal Status (e.g., gender, age, height, weight, physical fitness, cardio respiratory fitness, aerobic fitness, stress, mood, mental health, intelligence, and executive functions)
Lifestyle Behaviors (e.g., eating, physical activity, sleep patterns, social tie, and time management)
Learning Behaviors (e.g., class attendance, study duration, library entry, and online learning)

LITERATURE SURVEY

- Systematic literature review is a literature review method that aims to answer research questions by identifying, assessing, evaluating and interpreting all findings related to the research topic.
- Systematic literature review has proven to be an effective research method to provide an overview of trends in certain research topics, both results,

methodology and coverage of previous research fields.

- The systematic literature review method is carried out by following certain protocols.
- The stages of implementing a systematic literature review begin with formulating research questions and making a systematic literature review protocol.
- The literature review protocol in question is to determine what keywords to use in searching for articles in the database and what criteria to cover (inclusion criteria) and what criteria for articles that are not used (exclusion criteria).
- After the articles through the protocol are collected, the next step is to identify the appropriate literature, select from the primary study, extract data and assess the quality of the research results obtained.
- In the final stage, a synthesis of various research results is carried out and the stage of writing a systematic literature review article can be carried out.

PROPOSED METHOD:

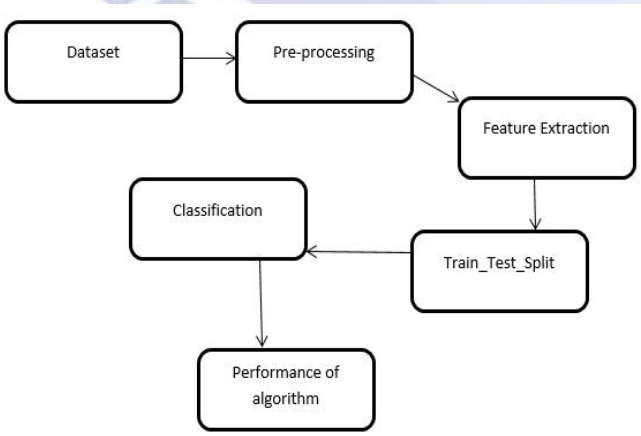
The main objective of our project is to predict student academic prediction based on multi source and multi-features.

ADVANTAGES:

- Develops a robust academic performance prediction.
- Potentially help students to optimize their interactions with the university

.

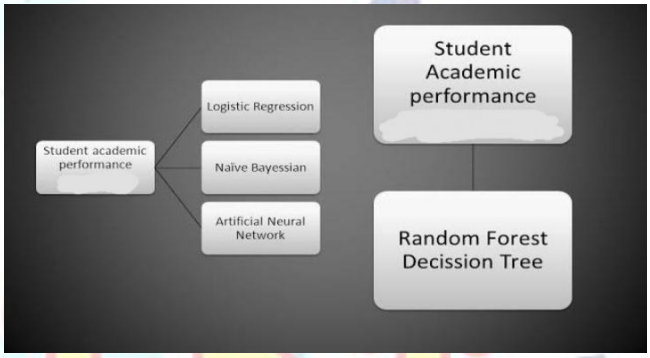
ARCHITECTURE:



DATA SET:

No	Variable Name	Information	Value
1	Marital Status	Marital status when take magister college	0 = not married 1 = married
2	Gender	Gender of magister student	0 = woman 1 = man
3	Scholar University	Rating of university with scale from 1-10 from survey of Webomatrics	10 = 35 big first rank 9 = 35 big second rank, etc
4	Period of Study	Time period for study	Nominal (2-4 years)
5	Work Status	Work status when take magister college	0 = not work 1 = work
6	Scholar GPA	GPA value at scholar	Nominal (0-4)
7	Age (new student)	Age when take magister college	Nominal

IMPLEMENTATION:



WORKING ALGORITHMS:

Artificial Neural Network:

ANN is useful for application in several areas including pattern recognition, classification, forecasting, process control etc. Robust for noisy dataset.

Limitations:

- ANN's do not have parametric statistical properties
- ANN may converge to local instead of global minima, there by providing non optimal data fits.

Logistic Regression:

LR is able to provide information about significance value of predicator. There are no assumption about normality of dataset.

Limitation:

- Only able to work with binary criterion variable.

Naïve Bayessian:

Naïve Bayessain requires data training fewer than other classification method.

Limitation:

- Dataset should satisfy independent assumption.

Random Forest Decision Tree:

Random forest runs efficiently on large databases. Random forest can handle thousands Input variables without variable deletion.

Random forest gives estimates of what variables are important in the classification. Random forest has an effective method for estimating missing data and maintains accuracy when a large proportion of the data are missing.

Random forest able to do classification, clustering and outlier detection.

Limitations:

- Random forest have been observed to overfit for some dataset with noisy classification/regression tasks.
- Unlike decision trees, the classifications made by random forest are difficult for humans to interpret.

IMPLEMENTATION RESULT:

	Multilayer Perceptron		
	Using Training Set	Cross Validation	Percentage Split
Correctly Classified Instances	84.62%	73.08%	71.43%
Incorrectly Classified Instances	15.38%	26.92%	28.57%
Kappa Statistic	0.5649	0.2877	0.3243
Mean Absolute Error	0.2701	0.3486	0.3971
Root Mean Absolute Error	0.354	0.4764	0.5047
Relative Absolute Error	65.53%	84.51%	90.96%
	Logistic Regression		
	Using Training Set	Cross Validation	Percentage Split
Correctly Classified Instances	77.88%	75.96%	65.71%
Incorrectly Classified Instances	22.12%	24.04%	34.29%
Kappa Statistic	0.3197	0.2786	0.1286
Mean Absolute Error	0.3343	0.3658	0.3988
Root Mean Absolute Error	0.408	0.4447	0.452
Relative Absolute Error	81.10%	88.68%	91.35%

	Naïve Bayes		
	Using Training Set	Cross Validation	Percentage Split
Correctly Classified Instances	76.92%	75.96%	68.57%
Incorrectly Classified Instances	23.08%	24.04%	31.43%
Kappa Statistic	0.2811	0.2415	0.2159
Mean Absolute Error	0.2585	0.2737	0.32
Root Mean Absolute Error	0.4504	0.4611	0.4952
Relative Absolute Error	62.73%	66.35%	73.28%
	Random Forest		
	Using Training Set	Cross Validation	Percentage Split
Correctly Classified Instances	**100%**	**68.27%**	**68.57%**
Incorrectly Classified Instances	0%	42.31%	31.43%
Kappa Statistic	1	-0.0732	0.2159
Mean Absolute Error	0.0933	0.3952	0.3686
Root Mean Absolute Error	0.1431	0.5162	0.5235
Relative Absolute Error	22.63%	95.82%	84.41%

CONCLUSION:

As an important issue in the education data mining field, academic performance prediction has been studied by many researchers. However, due to lack of richness and diversity in both data sources and features, there still exist a lot of challenges in prediction accuracy and interpretability. To initially alleviate this problem, our study aims at developing a robust academic performance prediction model, to gain an in- depth insight into student behavioral patterns and potentially help students to optimize their interactions with the university. In our study, a model named Augment ED is proposed to predict the academic performance of college students. Our contributions in this study are related to three sources. First, regarding data fusion, to the best of our knowledge, this work is the first to capture, analyze and use multisource data covering not only online and offline learning but also campus-life behaviors inside and outside of the classroom for academic performance prediction. Based on these multisource data, a rich profile of a student is obtained. Second, regarding the feature evaluation, behavioral change is evaluated by linear, nonlinear, and deep learning (LSTM) methods.

Conflict of interest statement

Authors declare that they do not have any conflict of interest.

REFERENCES

[1] A. Furnham, and J. Monsen, "Personality traits and intelligence predict academic school grades," Learning and Individual Differences, vol. 19, no. 1, pp. 0-33, 2009.

[2] M. A. Conard, "Aptitude is not enough: How personality and behavior predict academic performance," Journal of Research in Personality, vol. 40, no. 3, pp. 339-346, 2006.

[3] A. Jones, and K. Issroff, "Learning technologies: Affective and social issues in computer-supported collaborative learning," Computers & Education, vol. 44, no. 4, pp. 395-408, 2005.

[4] L. Qiu, Q. Lei, and Z. Zhang, "Advanced sentiment classification of Tibetan microblogs on smart campuses based on multi-feature fusion," IEEE Access, vol. 6, pp. 17896-17904, 2018.