STUDENT SCHOLARSHIP PREDICTION Using MACHINE LEARNING

Mr. G.RAMA RAO, Assistant Professor, Department of CSE

Huda Kousar, D. Sowmya, A. Ashok, K. Sai Kiran (UG Students, Department of CSE)
itsmerams@gmail.com,hudakousarmohammad@gmail.com, damerasowmya99599@gmail.com, anugulaashok@gmail.com, 22681a0550@cjits.com

Christu Jyothi Institute of Technology and Science, Telangana, India

Abstract

This paper consists of literature survey to prediction of scholarship by using Machine Learning and Data Mining technique. Along with this it contains a small description of ML/DM which are used by the researchers. It also describes data sets as very important in ML/DM methods. Machine Learning becomes most popular in the field of IT industry. Nowadays Machine Learning and Data Mining turn as a powerful technique which applicable for various fields such as IT, Education sector and also in business sector too. The different types of ML/DM algorithms are addressed by using all this technique. The algorithms which give more accuracy results in detection of continuity of every student's scholarship such as Naïve Bayes, Decision Tree and k-NN. Finally, the proposed model will provide a list of candidates, who deserve to have a scholarship and also discussion has been made on accuracy of each techniques which was used to get a result.

Keywords: Machine Learning (ML),Data Mining (DM), Student Performance Analysis, Decision Tree, k-Nearest Neighbors (k-NN), Predictive Modeling, Data-Driven Decision Making, Academic Analytics

1.Introduction

This paper talks about the different Machine Learning and Data Mining Techniques used to prediction of scholarship and out coming results. Some of the Machine Learning and Data Mining methods are explained and also the problems faced. The merits and demerits of ML and

DM methods are discussed. This also concluded the best method used for scholarship prediction. A scholarship gives a financial help for a student to continue with further education. Usually, scholarships are provided by the governments or non-governments organizations. When students are recognized for their accomplishments. It gives them the confidence to purpose other goals. There are some of the common types of scholarships such as merit based, student specific and carrier specific. Different schemes of scholarships are provided for the students based on the different eligibility criteria. Academic scholarships probably use a Grade point average (GPA) to select awardees. Athletic scholarships are generally awarded based on many criteria such as performance in an all subject, club participation, community services etc. This paper discussed about ML/DM method-based scholarship prediction to provide scholarships to students where eligible students are automatically predicted based on the given constraints such as percentage, grade point average, marks, annual income, communication skills etc.

2.Literature Survey

A) Helping Students Get More than Their Money from an Engineering Scholarship Program

Author: Thalia Anagnos

- ELPS program supports engineering students with mentoring and industry exposure.
- Improved graduation rates, leadership skills, and professional development.
- Post-scholarship survey shows strong impact of mentoring.

B) Characteristics of Engineering Faculty Engaged in the Scholarship of Teaching and Learning

Author: Angela R. Bielefeldt

- Data mining used to analyze engineering faculty involved in SoTL research.
- SoTL faculty more likely to be assistant professors, women, and from less research-intensive institutions.
- Highlights cultural and demographic differences across education-focused faculty groups.

C) Data Mining: Prediction for Performance Improvement of Graduate Students Using Classification

Authors: Kamal Bunkar et al.

- Used decision tree classification on student data from Vikram University.
- Identified key attributes influencing course performance.
- Developed a predictive system to estimate student grades.

D) Artificial Intelligence-Based Scholarship and Credit Pre-Assessment System
Authors: Sercan Saatçi et al.

- AI and ML used to automate and objectify scholarship eligibility assessment.
- System evaluates student data using criteria from public scholarship programs.
- Aims to reduce bias and time in scholarship decision-making.

E) A Comparative Analysis of Techniques for Predicting Academic Performance
Authors: Nguyen Thai Nghe et al.

- Compared decision tree and Bayesian network on datasets from CTU (Vietnam) and AIT (Thailand).
- Decision trees were slightly more accurate (up to 12%) than Bayesian networks.
- Useful for predicting scholarship-eligible students and supporting at-risk ones.

3.Methodology

The proposed system incorporate the following components:

- **Data Collection:** Gather student data including academics, financial status, and extracurricular.
- **Data Cleaning:** Handle missing or inconsistent data to ensure accuracy.
- **Model Selection:** Choose algorithms like Decision Trees, Logistic Regression, and K-Means Clustering.
- **Model Training:** Apply supervised or unsupervised learning based on data availability.
- **Model Evaluation:** Assess model performance using accuracy, precision, recall, and F1-score.
- **Deployment:** Integrate the model into a system for real-time scholarship prediction.

4.System Architecture and Design

The system architecture for the scholarship prediction model is structured to handle data collection, processing, machine learning-based prediction, and result presentation. It integrates various components such as data input, preprocessing, model training, evaluation, and a user interface. The Unified Modeling Language (UML) is used to visually represent the system's structure and behavior. It helps in understanding, designing, and documenting different aspects of the system.

- ***Use Case Diagram***

The use case diagram identifies system actors (e.g., admin, student) and their interactions with the system, such as data input, prediction, and viewing results.

- ***Class Diagram***

The class diagram shows the static structure of the system by illustrating classes, their attributes, methods, and relationships.

- ***Sequence Diagram***

The sequence diagram describes the interaction between system components over time. It highlights the sequence of operations from data submission to scholarship prediction.

- ***Activity Diagram***

The activity diagram represents the workflow of scholarship prediction, showing the control flow between steps such as data input, preprocessing, prediction, and result display.

5.Implementation Details

The system uses a Python-based backend with machine learning models—Decision Tree, K-Nearest Neighbors (k-NN), and Naive Bayes—to predict student scholarship eligibility. It begins by loading and preprocessing student data from a CSV file, then splits the data into training and testing sets. Each model is trained using this data and evaluated using accuracy, precision, recall, and F1-score. The results help identify which model best predicts scholarship eligibility, and this logic is integrated into a Django web application for user interaction.

6.Testing and Results

System testing ensures the entire scholarship prediction system works as expected by evaluating each module and their integration. Various testing methods were used:

- **Unit Testing**: Verified individual components (like user login, registration, prediction logic) produce correct outputs for valid/invalid inputs.
- **Integration Testing**: Checked whether different modules (data input, processing, and output) interact correctly.
- **Functional Testing**: Ensured the application functions (form submissions, predictions, model execution) behave according to requirements.
- **System Testing**: Validated the complete application meets user needs and functional specs.
- **White Box Testing**: Internals like data flow and logic paths were examined.
- **Black Box Testing**: User interactions were tested without seeing the internal code.

- **Acceptance Testing**: Final confirmation by end users that the system fulfills all requirements.

Result: All test cases passed successfully with no defects encountered.

7.Future Scope

In today's world, education has become one of the most essential and expensive assets. Most high school pass outs look for financial assistance to pursue the career of their dreams that requires several years of education. In such a scenario, scholarships come up as the most desired form of financial aid. This is because, unlike educational loans, you don't have to repay scholarships. If you are deserving, you can get scholarships on merit basis. This makes scholarships prestigious and also pays for your degree. You should definitely apply for scholarships in the college of your choice. After all, a scholarship comes with numerous benefits.

8.Conclusion

This paper discloses the survey of Machine Learning and Data Mining techniques used for scholarship prediction process. By using ML and DM techniques the prediction of scholarship much easier and effective also. While predicting the data we need to look upon many parameters. ML and DM techniques unable to work without training data. For collecting of datasets makes more difficult. We cannot be able to collect the best algorithm suitable for prediction. Because prediction is based on training data values: By using the ML and DM methods, the instructions can easily track on the scholarships and the problems faced by the recipients in their studies.

9.References

[1] Thalia Anagnos, Eva Schiorring, "Helping Students Get More than Teir Money from an Engineering Scholarship Program" IEEE, 2018.

[2] Angela R. Bielefeldt, "Characteristics of Engineering Faculty Engaged in the Scholarship of Teaching and Learning" IEEE, 2015.

[3] K. Bunkar, "Data Mining: Prediction for Performance Improvement of Graduate Students using," IEEE, 2012.

[4] Sercan, Saatci, Hande cansiz, gulsha aslan,erkan ozhan “Artificial intelligence-based Scholarship DQG Credit pre-assessment system” IEEE, 2017.

[5] G. Upton, "The Effects of Merit-Based Scholarships on Educational Outcomes," Journal of Labor Research, vol. 37, no. 2, pp. 235-261.

[6] Okfalisa, Ratik a fitriani, YelliVitriana, “ The Comparison of Linear Regression Method and k- Nearest Neighbors in Scholarship Recipient” IEEE, 2018.

[7] N. T. Nghe, "A Comparative Analysis of Techniques for Predicting Academic Performance," IEEE, 2007.

[8] Jonalyn Joy B.Labayne, Lester L. Mercado, Jheanel Espiritu Estrada, “Model Development Of Students Scholarship Status at First at Asia Institute oftechnology And Humanities(FAITH)” IEEE,2018.

[9] Suma, V., and Shavige Malleshwara Hills. "Data