2019

Data analytics on performance of computing students

Balogun, Oluwafemi Samson

ACM Press

Artikkelit ja abstraktit tieteellisissä konferenssijulkaisuissa
Authors
All rights reserved
http://dx.doi.org/10.1145/3364510.3366152

https://erepo.uef.fi/handle/123456789/7899
Downloaded from University of Eastern Finland's eRepository
Data analytics on performance of computing students

Oluwafemi Samson Balogun†
School of Computing
University of Eastern Finland
Kuopio, Finland
samson.balogun@uef.fi

Solomon Sunday Oyelere
School of Computing
University of Eastern Finland
Joensuu, Finland
solomon.oyelere@uef.fi

Donald Douglas Atsa’am
Department of Maths, Sta and CS
University of Agriculture
Makurdi, Nigeria
donatsaam@alumni.edu.tr

ABSTRACT
This study was conducted to determine the relationship between computing students’ initial and final academic performance to support decision making in higher education institutions. The data used in this research contains the initial GPA and CGPA of 710 computing students for a period of 5 years. Test of normality for the final graduating results and linear regression models were fitted to the data to predict the overall performance based on the initial result. This study revealed that there are strong linear relationships between the GPAs and CGPAs of computing students over the period of this study. The contribution of this work is that the result enables the student and the teacher to understand the trend of students’ academic performance for the purpose of decision-making. Keeping track of the performance of the students help provide support whenever is needed. This results aimed to decrease student dropout by means of facilitating student to predict their probability of success in computing courses after enrollment. In addition, teachers will be able to boost student performance in their courses as a result of enhanced determination of student’s abilities to learn the course and fine-tuning teaching approaches and techniques.

CCS CONCEPTS
Applied computing → Education

KEYWORDS
Computing education, Data analytics, Academic performance, Linear regression analysis, Correlation analysis

ACM Reference format:

1 Background
Numerous factors are regularly considered as key in the decision-making of university education. A portion of the variables that determine student overall performance, success, failure or dropout are favorable condition for effective learning, student-teacher interaction, teacher satisfaction, student engagement and satisfaction with the educational setting. Educational data mining methods are employed to understand patterns and trends of students’ learning. Large-scale data emerging from learning settings are collected and used to on the one hand, identify the weaknesses within the educational settings, and on the other hand, improving the performance of teachers and students [1]. In fact, one particular goal of educational data mining is to predict the performance of students using different prediction algorithms. Prediction algorithms analyze the entire dataset to generate accurate models for each potential outcome. Prediction algorithms usually provides signs about trends of students’ academic endeavor such as student performance, pedagogical activities, learning style and behaviors, teaching quality, effectiveness and efficiency, potential for success or dropout and so on [2], [3], [4].

This paper depicts the apparent association between the final graduating grade of computing students and their initial results in the university. This study examine the impact of the initial Grade Point Average (GPA) of computing students in their first year on the final graduation CGPA. This study answers the research question, to what extent can the final year CGPA of a computing student be predicted using the GPA of the first year? Correlation analysis was used to determine if there is a linear relationship between the initial GPA and the final CGPA of computing students. While linear regression analysis was used to test the normality of the final graduating results and linear regression models were fitted to the data to predict the overall performance based on the initial result [5].

2. Data source, results and interpretation
The data used in this research contains the initial GPA of computing students for a period of 5 years and their graduating CGPA from the departments of Computer Science (CS), Information and Communication Technology (ICT), and Electrical Engineering (EE). The data used for this work is a secondary data collected from Modibbo Adama University of Technology, Yola, Adamawa State, Nigeria.

2.1 Results of correlation analysis
The correlation coefficients between the GPAs of the students from the first year till the point of graduation for all the three departments are presented in Table 2.

Table 1: Correlation Analysis
The result in Table 1 show that there are positive linear relationships among the GPAs across all levels. The linear relationships are also strong except for few occasions where there exist fair relationships.

### 2.2 Results of regression Analysis

The final CGPA is made to be the dependent variable while the GPA is made to be independent variable. The regression line for ICT, CS and EE scores are:

- ICT: final CGPA = 0.648 + 0.7920 first GPA
- CS: final CGPA = 0.131 + 0.8697 first GPA
- EE: final CGPA = 0.402 + 0.8048 first GPA

Table 3 represent the test of equality of means for ICT, CS and EE scores are:

<table>
<thead>
<tr>
<th>Term</th>
<th>Coefficient</th>
<th>Std. Error Coefficient</th>
<th>t</th>
<th>p - value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Constant</td>
<td>0.648</td>
<td>0.113</td>
<td>5.75</td>
<td>0.000</td>
</tr>
<tr>
<td>First GPA</td>
<td>0.402</td>
<td>0.402</td>
<td>25.50</td>
<td>0.000</td>
</tr>
</tbody>
</table>

### 3. Implication of the study and conclusion

Inferential statistics has been implemented on the academic performance of computing students in a Nigerian university. This study revealed that there are strong linear relationships between the GPAs and CGPAs of computing students over the period of this study. All the results of the regression models significantly fit the data used and each of the regression parameters contributes to the significance of the model. Teacher could use the result of this study to understand the trends in the academic performance of the computing student to provide the needed support at each stage of the students' academic progression. Besides, the results of this study will serve as motivation for the teachers and look out for signs of trouble in each student. It will also help in keeping a track record of each students by predicting their academic performance so as to mentor, advice and nurture the students. The result of this study will encourage competition among the students and motivates students to work harder to improve their performance in order to succeed.

### REFERENCES
