ISSN (Online) 2394-2320

International Journal of Engineering Research in Computer Science and Engineering (IJERCSE)

Vol 10, Issue 4, April 2023

Prediction of Student Academic Performance using Linear Regression

[1] Noel A. Banca, [2] Renante A. Diamante

[1] [2] [3] Iloilo State University of Fisheries Science and Technology, Philippines
Corresponding Author Email: [1] noel.banca@gmail.com, [2] rdiamantetip@gmail.com

Abstract*— Academic assessment is based on providing instant and specific feedback after each learning step to avoid unnecessary delays in correcting students' errors. For such type of evaluation to realize its maximum benefit, it is suggested that assessment should not be a once-and-done activity. Rather, it must be a continuous act, which guides the teaching-learning process of provision for timely feedback. The Machine Learning Algorithm using linear regression was developed to Predict Student Performance probability rate to complete the course taken. It also provides students with information such as ID number, course, year and grades. The demographic data were used for personal, financial and psychological information, family background and health record. The respondents of the study are Student Affairs personnel, IT Experts, and students coming from different departments. This study used the Prototyping Model as a guide in developing the application. To evaluate the application, McCall's Quality Model was applied. The mean was used to analyze the data. In terms of system operation, system revision and system transition criteria all mean equivalents for each characteristic were described as "very effective". Based on the result of the study, the researchers concluded that the Machine Learning Algorithm to Predict Student Academic Performance is a big help to the institution, especially to the Office of Student Affairs to predict the probability rate of the student finish their studies and increase student's awareness.

Keywords: Machine Learning, Linear Regression, Student Academic Performance.

I. INTRODUCTION

For quite a long period of time, many educators have viewed assessment as a means for measuring learning final outcomes and this is mainly represented through what is known as summative assessment [1]. This is exactly what assessment for learning does as it aims to improve the quality of teaching and use assessment to modify students' learning.

Nowadays, Machine Learning (ML) is one of the most promising application areas in the field of Information Technology where its application scope is almost unlimited [2]. Machine learning is having a tremendous impact on the teaching industry. The teaching industry is adopting new technologies to predict the future of the education system. It is Machine learning which predicts the future nature of the education environment by adapting new advanced intelligent technologies. This work explores the application of Machine Learning in teaching and learning for further improvement in the learning environment in higher education [3]. In the higher education environment, tracking and analyzing student learning outcomes by subject is one approach to improving student learning outcomes in the next subjects.

Large universities using the course management system could provide stakeholders with comprehensive assessments and data to review and assess student learning outcomes. There seems to be little such research at small and medium institutions in any case. This research suggests a simple method to design a dashboard that can be used by faculty and other university stakeholders to evaluate the learning outcomes of each subject [4]. Course instructors need to assess the efficacy of their teaching methods, but experiments in education are seldom politically, administratively, or ethically feasible. Quasi-experimental tools, on the other hand, are often problematic, as they are typically too complicated to be of widespread use to educators and may suffer from selection bias occurring due to confounding variables such as students' prior knowledge [5].

This study aimed to explore the efficacy of academic assessment for learning in improving the performance of students using machine learning algorithms, specifically predictive analytics, and demographics data. The findings of this study may be used to enhance the process of modernizing the educational practice and help to facilitate the learning then, the overall academic performance of the students. It can be also used as a learning interaction on the potential failure of students to achieve learning objectives. And it may be able to make changes in learning strategies in order to facilitate student diversity in the classroom.

II. METHODOLOGY

The input-process-output model was adopted when this study was conceived as shown in Figure 1. The input variables are in two clusters. Students' demographics of the first set of input variables are made up of all the information needed for student profiling.

ISSN (Online) 2394-2320

International Journal of Engineering Research in Computer Science and Engineering (IJERCSE)

Vol 10, Issue 4, April 2023

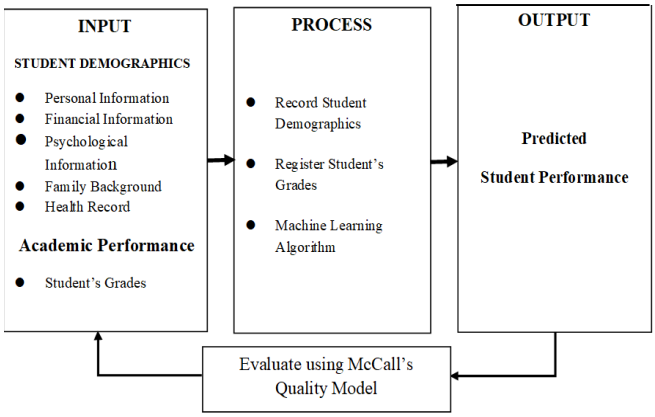

Figure 1. Schematic diagram illustrating the conceptual framework of the Machine Learning Algorithm to Predict Student Performance using the input-process-output model.

The other input group of variables refers to academic performance. Academic performances of students include all the results of their tests, tasks, outputs, assignments and projects in overall subjects.

Once the data on the students' demographics and academic performance is all collected, gathered and encoded in the system, the teachers can now assess the academic performance of students.

Development Process

The researchers used the Spiral Model as a guide to developing a proposed web system for the teacher to easily assess the academic performance of the students. The spiral model is one of the most important Software Development Life Cycle models, which provides support for Risk Handling. In its diagrammatic representation, it looks like a spiral with many loops.

Objectives determination and identify alternative solutions

In this study, student demographics are gathered. Student demographics will serve as the basis for the teachers to identify the salient points in the student's academic performance and offers an alternative solution to those students that need extra attention to on their studies.

Identify and resolve risks

During this process, all the information about the student's demographics should be evaluated by the teachers and must be classified depending on their profile. All data regarding their academic performance, including those who need extra attention are already resolved and the team is now ready for the making of the prototype.

Develop the next version of the Product

During the third quadrant, the identified features are developed and verified through testing. At the end of the third quadrant, the next version of the software is available.

Once the prototype is already done, all information about the students' demographics and academic performance must encode and ready for assessment.

Review and plan for the Next Phase

In the fourth quadrant, the Customers evaluate the so-far developed version of the software. In the end, planning for the next phase is started.

The assessment of the student's academic performance is already performed and teachers can now predict the probability of the students' passing rate. They can now also help the students adjust to their learning methods. The system is now ready to use and may start another developing phase.

III. RESULTS AND DISCUSSIONS

Machine Learning Algorithm to Predict Student Performance is designed to create a system that will help the students and teachers to evaluate the student's information and performances. They can easily know if they can pass the course that they taking up or not so that teachers and administrators can assess the students on any precautionary measures needed to improve their performance.

The development of the system on Machine Learning Algorithm to Predict Student Performance is for the students, teachers and school administration for an easier assessment of the student's performance and to address any issues or factors that can contribute to the poor academic performance of students.

Table 1. Level of Software Operation of the Machine Learning Algorithm to Predict Student Performanc

Product Operation	End Users		IT Experts		Entire Group	
	M	Desc	M	Desc	M	Desc
Correctness	4.2	VE	4.8	VE	4.8	VE
Efficiency	4.4	VE	4	VE	4.4	VE
Integrity	4.2	VE	4.4	VE	4.4	VE
Reliability	3.8	VE	4	VE	4.6	VE
Usability	4.6	VE	4.9	VE	4.6	VE
Mean Average	4.24	VE	4.42	VE	4.56	VE

As shown in Table 1, As the software quality factor Correctness, the Machine Learning Algorithm to Predict Student Performance was rated by the End Users (M=4.2) Entire Group (M=4.8) IT Experts (M=4.8) as "Very Effective".

This means that the Machine Learning Algorithm to Predict Student Performance in terms of Product Operation as the software quality factor Correctness for the entire group, the application can work accurately.

ISSN (Online) 2394-2320

International Journal of Engineering Research in Computer Science and Engineering (IJERCSE)

Vol 10, Issue 4, April 2023

As the software quality factor Efficiency, the Machine Learning Algorithm to Predict Student Performance rated by the End Users (M=4.4) and Entire Group (M=4.4) IT Experts (M=4) as "Very Effective".

This means that the Machine Learning Algorithm Predict Student Performance in terms of Software Operation as the software quality factor Efficiency for the entire group, and the application provides the correct information and result.

As the software quality factor Integrity, the Machine Learning Algorithm to Predict Student Performance rated by the End Users (M=4.2) Entire Group (M=4.4) IT Experts (M=4.4) as "Very Effective".

This means that the Machine Learning Algorithm to Predict Student Performance in terms of Product Operation as the software quality factor Integrity for the entire group, and the application is consistent throughout the system operation.

As the software quality factor Reliability, Academic Machine Learning Algorithm to Predict Student Performance rated by the End Users (M=3.8) Entire Group (M=4.6) IT Experts (M=4) as "Very Effective".

This means that the Machine Learning Algorithm to Predict Student Performance in terms of Product Operation as the as the software quality factor Reliability for the entire group, and the application operates its function without failure.

As the software quality factor Usability, Machine Learning Algorithm to Predict Student Performance rated by the End Users (M=4.6) Entire Group (M=4.6) IT Experts (M=4.9) as "Very Effective".

This means that the Machine Learning Algorithm Predict Student Performance in terms of Product Operation the as the software quality factor Usability for the entire group, and the application is easy to operate and control.

Table 2. Level of Software Revision of the Machine Learning Algorithm to Predict Student Performance

Software Revision	End Users		IT Experts		Entire Group	
	M	Descri	M	Desc	M	Desc
Maintainability	4.2	VE	4.8	VE	4.8	VE
Flexibility	4.4	VE	4.6	VE	4.8	VE
Testability	4.2	VE	4.4	VE	4.8	VE
Mean Average	4.37	VE	4.6	VE	4.8	VE

As shown in table 2, the software quality factor Maintainability Machine Learning Algorithm to Predict Student Performance, rated by the End Users (M=4.2) Entire Group (M=4.8) IT Experts (M=4.8) as "Very Effective".

This means that the Machine Learning Algorithm to Predict Student Performance in terms of Product Revision as the software quality factor Maintainability for the entire group, the application can easily maintain or manage by the administrators.

As the software quality factor Flexibility, the Machine Learning Algorithm to Predict Student Performance rated by the End Users (M=4.4) Entire Group (M=4.8) IT Experts (M=4.6) as "Effective".

This means that the Machine Learning Algorithm to Predict Student Performance in terms of Product Revision as the software quality factor Maintainability for the entire group, and the application can give satisfying interaction to the user and to its environment.

As the software quality factor Testability, the Machine Learning Algorithm to Predict Student Performance rated by the End Users (M=4.2) Entire Group (M=4.8) IT Experts (M=4.4) as "Very Effective".

This means that the Machine Learning Algorithm to Predict Student Performance in terms of Software Revision is the software quality factor Testability for the entire group, the application be tested before releasing to the user.

Table 3. Level of Software Transition of the Machine Learning Algorithm to Predict Student Performance

Software Transition	End Users		IT Experts		Entire Group	
	M	Desc	M	Desc	M	Desc
Portability	3.2	VE	4.2	VE	4.6	VE
Reusability	3.8	VE	4.4	E	4.6	VE
Interoperability	4.25	VE	4	VE	4.8	VE
Mean Average	3.75	VE	4.2	VE	4.67	VE

As shown in Table 3, as the software quality factor Portability, the Machine Learning Algorithm to Predict Student Performance rated by the End Users (M=3.2) IT Experts (M=4.2) Entire Group (M=4.6) as "Very Effective".

This means that the Machine Learning Algorithm Predict Student Performance in terms of Software Transition as the software quality factor Portability for the entire group, the application can easily be installed and uninstalled.

As the software quality factor Reusability, Machine Learning Algorithm to Predict Student Performance rated by the Entire Group (M=4.6) End Users (M=3.8) IT Experts (M=4.4) as "Effective".

This means that the Machine Learning Algorithm to Predict Student Performance in terms of Product Transition as the software quality factor Reusability for the entire group, the application can still function in several technology environments.

ISSN (Online) 2394-2320

International Journal of Engineering Research in Computer Science and Engineering (IJERCSE)

Vol 10, Issue 4, April 2023

As the software quality factor Interoperability Machine Learning Algorithm to Predict Student Performance rated by the End Users (M=4.25) IT Experts (M=4) as "Very Effective".

This means that the Machine Learning Algorithm to Predict Student Performance in terms of Software Transition is the software quality factor Interoperability for the entire group, the application can still function while the other app is in use.

Table 4. Perception of the Evaluation when taken as a whole

As a whole	End Users		IT Experts		Entire Group	
	M	Desc	M	Desc	M	Desc
Correctness	4.2	VE	4.8	VE	4.8	VE
Efficiency	4.4	VE	4	VE	4.4	VE
Integrity	4.2	VE	4.4	VE	4.4	VE
Reliability	3.8	VE	4	VE	4.6	VE
Usability	4.6	VE	4.9	VE	4.6	VE
Maintainability	4.2	VE	4.8	VE	4.8	VE
Flexibility	4.4	VE	4.6	VE	4.8	VE
Testability	4.2	VE	4.4	VE	4.8	VE
Portability	3.2	VE	4.2	VE	4.6	VE
Reusability	3.8	VE	4.4	VE	4.6	VE
Interoperability	4.25	VE	4	VE	4.8	VE

In terms of Correctness, the system rated by the End Users (M=4.2) as "Very Effective", and IT Experts (M=4.8) as "Very Effective" Entire Group (M=4.8) as "Very Effective".

In terms of Maintainability, the system rated by the End Users (M=4.2) as "Very Effective", and IT Experts (M=4.8) as "Very Effective" Entire Group (M=4.8) as "Very Effective".

In terms of Flexibility, the system rated by the End Users (M=4.4) as "Very Effective", and IT Experts (M=4.6) as "Effective" Entire Group (M=4.8) as "Very Effective".

In terms of Testability, the system rated by the End Users (M=4.2) as "Very Effective", and IT Experts (M=4.4) as "Very Effective" Entire Group (M=4.8) as "Very Effective".

In terms of Portability, the system rated by the End Users (M=3.2) as "Very Effective", and IT Experts (M=4.2) as "Very Effective" Entire Group (M=4.6) as "Very Effective".

In terms of Reusability, the system is rated by the End Users (M=4.6) as "Very Effective", and IT Experts (M=4.9) as "Very Effective" Entire Group (M=4.6). This means that the usability of the system is functioning well. It's have been rated by the End Users as "Very Effective", and IT Experts as "Effective", which means that for the entire group, the IT Experts are satisfied with the device but also expecting an improvement and more innovation in the system of the Machine Learning Algorithm to Predict Student Performance

In terms of Interoperability, the End Users rated the Machine Learning Algorithm to Predict Student Performance as "Very Effective". The End Users and IT Experts are comfortable and satisfied with the portability of the system. This means that the effectiveness of the *Machine* Learning Algorithm in Predict Student Performance towards its users is functional and good.

Table 5. Result of the Mean Average for the Mangrove Identification App

Category	Entire Group	
	Mean	Description
Software Operation	4.40	Very Effective
Software Revision	4.59	Very Effective
Software Transition	4.21	Very Effective
Mean Average	4.4	Very Effective

In terms of Product Operation, it got a mean value of 4.40 which was described as the highest rate of "Very Effective".

It indicates that the development of the Machine Learning Algorithm to Predict Student Performance conforms to the requirements standard of McCall's Software Quality Model in terms of Product Operation, Product revision and Product Transition. This outcome denotes that the developed application needs to improve or steadied a piece of extra information handling the design and the application's development and operation.

IV. CONCLUSION

In view of the study, the following conclusions were drawn:

This study was able to accomplish the objectives based on the features of the system. The system was user-friendly both to the students and administration as they were able to access, fill up, verify and view the demographics needed. The system was able to provide enough knowledge to the teachers and admins on what specific areas certain students need more effort and attention. Schools and education facilities should provide more effort into developing applications and systems that can help improve the performances of every student.

Recommendations

Based on the proceedings, the following recommendation was drawn;

The researchers recommend that the graph of linear regression should be included in the database of the system instead of manually plotting it in Excel Spreadsheet.

REFERENCES

[1] Black, P. J., & William, D. (2006). Assessment and classroom learning. *Assessment in Education: Principles, Policy, and Practice, 5* (1), 7-74.doi:10.1080/0969595980050102

[2] Kucak, D., Juricic, V., & Dambic, G. (2018. Machine learning in education – A survey of current research trends. *DAAAM International*, 1726-9679. doi: 10.2507/29th.daaam.proceedings.059

[3] Asthana, P. & Hazela, B. (2020). Applications of machine learning in improving the learning environment. *Multimedia Big Data Computing for IoT Applications.* doi: 10.1007/978-981-13-8759-3_16

[4] T hanh, T. B., Dang, T. D., Le, D. H., & Aquino, P. G. (2021). Student grades analytics dashboards in higher education in Vietnam. *Journal of Contemporary Issues in Business and Government, (27)*1. Retrieved from https://cibg.org.au./

[5] Duzhin, F. & Gustafsson, A. (2018). Machine learning-based app for self-evaluation of teacher-specific instructional style and tools. *Education Sciences, 8,* 7. doi: 10.3390/educsci 8010007