

Student Academic Performance Prediction System by Using Decision Tree

Norfatin Firzana Mohd Salehuddin¹, Siti Rohaidah Ahmad^{2,*}, Wan Yusmawati Wan Yusoff³, Yuhanim Hani Yahaya⁴, and Siti Hajar Zainal Rashid⁵

¹ Norfatin Firzana Binti Mohd Salehuddin; 2200254@alfateh.upnm.edu.my

² Siti Rohaidah Binti Ahmad; sitiroidah@upnm.edu.my;  ORCID ID: 0000-0001-5045-9803

³ Wan Yusmawati Binti Wan Yusoff; yusmawati@upnm.edu.my;  ORCID ID: 0000-0003-0407-3031

⁴ Yuhanim Hani Binti Yahaya; yuhanim@upnm.edu.my;  ORCID ID: 0000-0001-9616-8649

⁵ Siti Hajar Binti Zainal Rashid; sitiroidah@upnm.edu.my;  ORCID ID: 0000-0002-0017-7063

* Correspondence: sitiroidah@upnm.edu.my; +60182737020.

Abstract: The main goal of the education system is not only to provide good quality education but to ensure that students can graduate with excellent academic performance. The main challenge of educational institutions is the drop-in rates for the success of students in academics, some even decide to quit studying. To overcome such problems, early forecasting, and analysis related to student failure or factors that may affect academic performance students need to be done. This method can help educational institutions identify students who need special attention to reduce the failure rate and take appropriate action. Storage of information related to student exam results is often collected and kept by the management department of the faculty as a record only. Thus, Student Performance Prediction System is developed using tree classification techniques ID3 results where IF-THEN rules are generated through ORANGE software and applied into the system to be developed. Student information such as results examination, gender, student status, state of origin is collected to help generate tips. Data storage such as user information and information students are stored through a MySQL database. In addition, HTML, CSS, and PHP as well used to design the system interface. Overall, the development of this Student Academic Performance Prediction System by Using Decision Tree is intended to help educational institutions increase the rate of excellence and reduce student dropout or failure rate.

Keywords: Decision Tree; Orange Data Mining; Student Academic Performance Prediction System.



Copyright: © 2023 by the authors. Submitted for open access publication under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. INTRODUCTION

Students are the main asset for educational institutions, especially for university-level. These subjects play an important role in producing quality graduates with outstanding academic achievements and commendable personalities. In order to measure a student's level of performance in academic, tests such as exams, evaluations and student involvement are taken into account. The final results of each test will then be recorded and stored in the student management system possibly in the form of a file, excel, document or other type of format. However, the academic performance of each student might be different due to certain factors such as the way of learning, background of each individual or environmental factors. The increasing number of students in an educational institution causes the amount of student data become enormous and process of analysing the data becoming difficult if it be done by traditional statistical techniques. Therefore, a system is needed to help the

academic institution analyse student performance as well as weak areas for improvement. Thus, the proposal to develop a Student Academic Performance Prediction System is made to ensure the activity of analysing student performance in academics becoming easier and could help the academic institution successfully achieve the aim for a higher success rate in the future.

2. METHOD & MATERIAL

The methodology used for the system development is the Agile methodology. There are several phases involves in this methodology such as information gathering, analysis, design, development, testing, and feedback. Agile methodology then be modified into information gathering, analysis, design, development, testing and feedback to meet the activities be carried throughout developing the system. Figure 1 shows the details about the activities carried out in each phase.

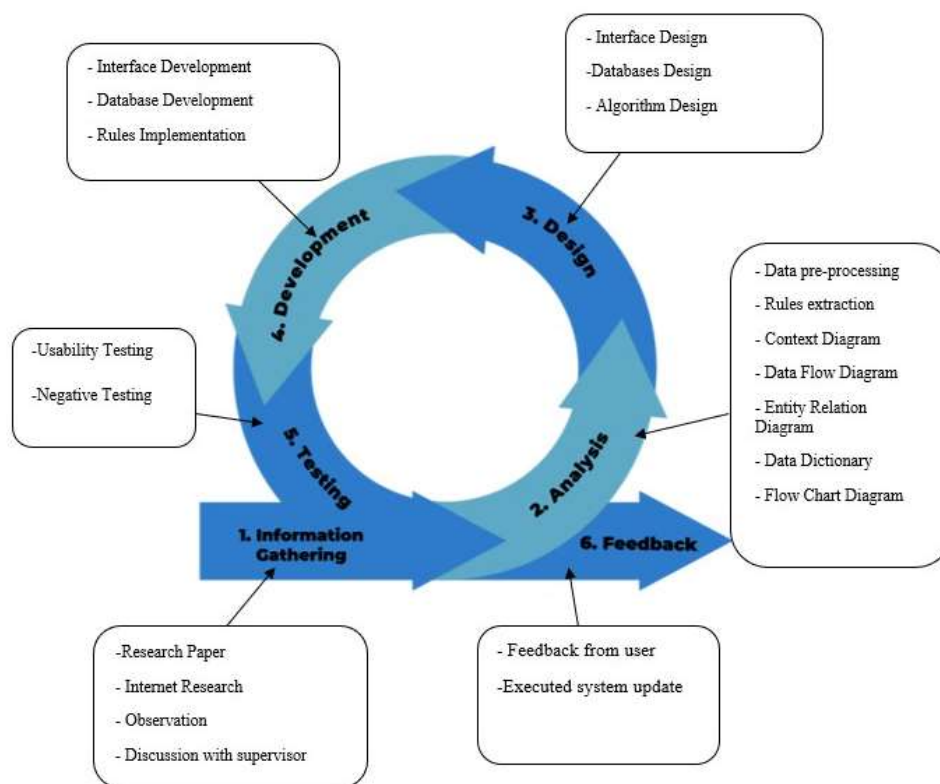


Figure 1. Agile Methodology for Student Academic Performance Prediction System

There are several materials are involved throughout the development of the project. These materials are needed to ensure the project meet the objectives successfully. The materials are as follow:

1. Users

User is a person who uses or operates something. In this project, there are two main users which are administrator and lecturer. Lecturer used the system to predict and view the result of the prediction after insert all the data need by the system while administrator used the system to monitor the performance of the students.

2. Software and Programming Language

Along the development of the system there are few of tools need to ensure the successful of the project. These are the following software and programming language used throughout the development:

- i. Sublime software used to write the code of PHP language, HTML, CSS and MySQL
- ii. Xampp used as the webserver
- iii. Google Chrome platform to access the databases and displays the system.
- iv. Orange Data Mining Software used to generate the Decision Tree.
- v. MySQL used for databases.
- vi. PHP language used to develop the system.

3. FINDINGS

Concept of data mining is applied in build up the system to enable the prediction activity to be carried out. The system be trained to produce student performance results based on previous data. The pass data need to go through data cleaning process in order to produce a useful dataset and rules for the system to be used. The IF-THEN rules are generated from the decision tree build by the Orange software. These rules will be implemented into the system in IF-ELSE condition using PHP language. In order to predict the performance of a student, there are some information are needed by the system such as student ID, student's state, student status, gender of student and grade for every subject on a semester. These data will be stored in MySQL databases for analysis and produced the result.

3.1 Interface of The System

Figure 2 illustrates an interface of the system that required lecturer to insert all information needed by the system.

Figure 2. Student's Information Interface

These data then will be stored in databases and displayed in the system on the student list page as shown in Figure 3.

3.2 Design of System Algorithm

Algorithm design is a logical approach that describes the procedure or step by step of a system. In this system, the algorithm design is described for each existing process. Figure 5 until Figure 9 display the algorithm design for the system.

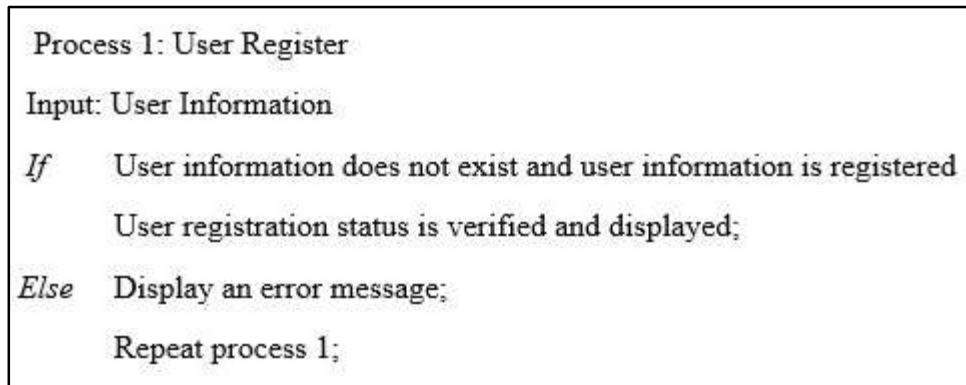


Figure 5. Algorithm Design for Process 1

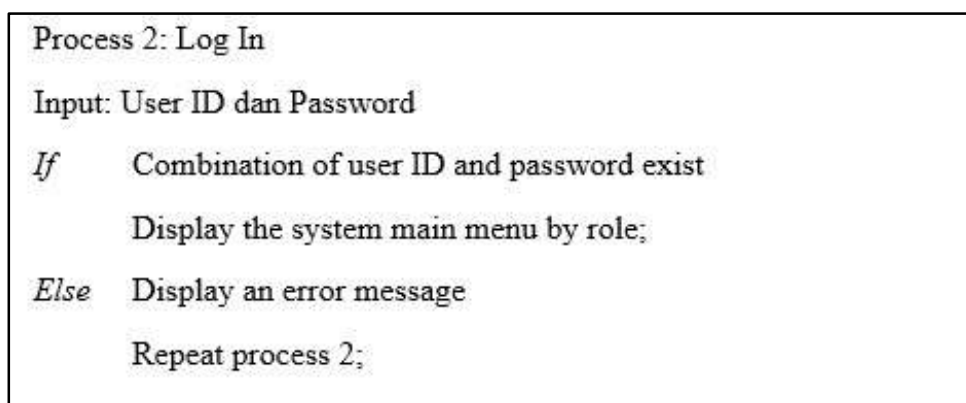


Figure 6. Algorithm Design for Process 2

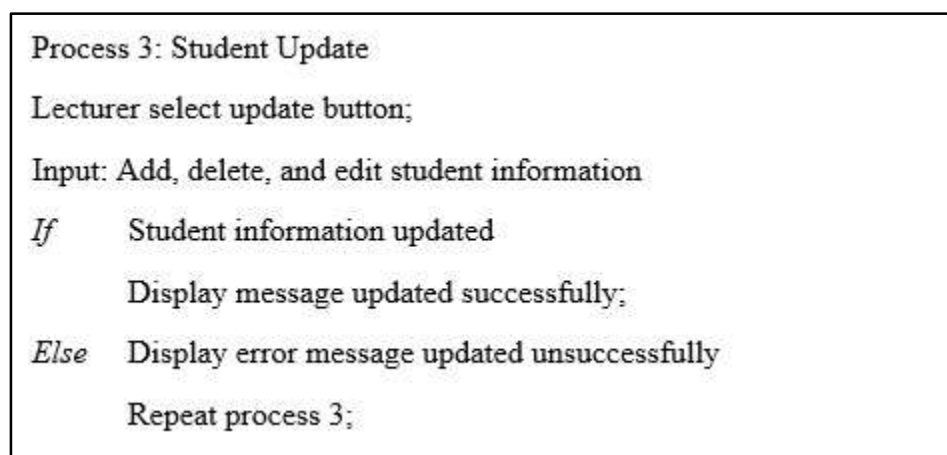


Figure 7. Algorithm Design for Process 3

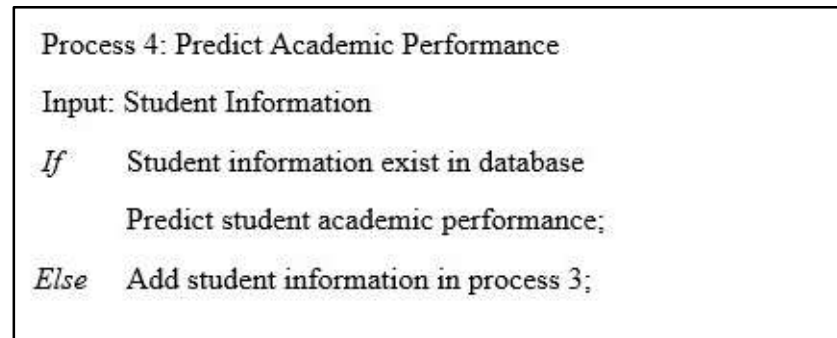


Figure 8. Algorithm Design for Process 4

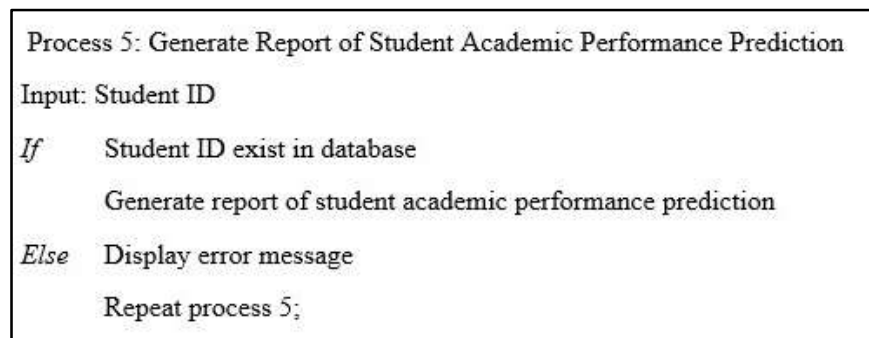


Figure 9. Algorithm Design for Process 5

4. DISCUSSION

This Student Academic Performance Prediction System was developed to provide facilities for lecturers to predict a student's academic performance. Thus, the development of this system can be measured through several objectives that had been focusing to. The advantages of this system are produced depending on the needs that have been targeted and the problem that arise. The advantages of the system are as follows:

- i. A simple and use-friendly interface system. All buttons can work well according to the needs of the user.
- ii. Allow the lecturers to predict the potential of students for the upcoming semester.
- iii. Lecturers can set appropriate targets and plans to help students achieve excellent performance for the next semester.
- iv. The system has a database that stores student information to be used for the prediction safely and systematically.

Although the system can run well, it still has some flaws that need to be improved for ease in the future. The weakness that might occurred in this system are as follows:

- i. This system only produces results according to the category "Excellent, Very Good, Good, Fail" without any Average Grade being specified.
- ii. The predictions made by the system are limited by issuing predictions for one student only without displaying overall statistics such as comparisons between courses, states, student types or subjects.
- iii. The system is only focused on one course only.
- iv. The administrator can only see the results produced by the lecturer without being able to give any comments for improvement.

5. CONCLUSION

The conclusion from the overall development of the system can be stated that the Student Academic Performance Prediction System has achieved the objectives and scope that have been set. All processes and work can be carried out well even though there are time constraints throughout the development process. In order to ensure the improvement of the excellent academic performance achieve, this system can be used as the guide or benchmark for system improvement in the future. Thus, excellent student can be produced to develop a stable economic development of country's society.

Acknowledgments: Special thanks to the Faculty of Defense Science and Technology, National Defence University of Malaysia for supporting this particular project and also the Student Affairs Department for the sponsorship to participate i3DC 2023.

References

- Abana, E. C. (2019). A decision tree approach for predicting student grades in research project using Weka. *International Journal of Advanced Computer Science and Applications*, 10(7). <https://doi.org/10.14569/ijacsa.2019.0100739>
- Bioinformatics Laboratory, U.of L. (no date) *Data mining, orange data mining - Data mining*. Retrieved from: <https://orangedatamining.com/>
- Du, H. (2022). *Data Mining techniques and applications: An introduction* (1st ed.). Cengage, India.
- Hamsa, H., Indiradevi, S., & Kizhakkethottam, J. J. (2016). Student academic performance prediction model using Decision Tree and Fuzzy Genetic Algorithm. *Procedia Technology*, 25, 326–332. <https://doi.org/10.1016/j.protcy.2016.08.114>
- Hasan, R., Palaniappan, S., Raziff, A. R. A., Mahmood, S., & Sarker, K. U. (2018). Student academic performance prediction by using Decision Tree Algorithm. *4th International Conference on Computer and Information Sciences (ICCOINS), Kuala Lumpur, Malaysia*, 1-5. <https://doi.org/10.1109/iccoins.2018.8510600>
- Intelligent mining and decision support system (inminds)* (n.d.). Retrieved from: https://ir.unimas.my/id/eprint/5713/1/AINC_INMINDS.pdf
- Jackson, M. B. (2019, June 13). Chapter 3: Data, information and meaning. Retrieved from: https://www.academia.edu/39544504/Chapter_3_Data_Information_and_Meaning
- Malaysia Communications and Multimedia Commission. (2019). *Internet Users Survey 2020*. <https://www.mcmc.gov.my/skmmgovmy/media/General/pdf/IUS-2020-Infographic.pdf>
- Mantri, N. (2021) *Using ID3 algorithm to build a decision tree to predict the weather*. OpenGenus IQ: Computing Expertise & Legacy. Retrieved from: <https://iq.opengenus.org/id3-algorithm>
- Pandey, M., & Kumar Sharma, V. (2013). A Decision Tree Algorithm pertaining to the student performance analysis and prediction. *International Journal of Computer Applications*, 61(13), 1–5. <https://doi.org/10.5120/9985-4822>
- Russell, S. J., & Norvig, P. (2022). *Artificial intelligence: A modern approach*. 2nd ed. Prentice Hall.
- Sa, C. L., Abang Ibrahim, D. H. B., Dahliana Hossain, E., & Hossin, M. (2014). Student performance analysis system (SPAS). *The 5th International Conference on Information and Communication Technology for the Muslim World (ICT4M)*. <https://doi.org/10.1109/ict4m.2014.7020662>
- Sewwandi, P. (2021, December 14). How to connect MySQL databases in Weka using MySQL jdbc connector? Medium. Retrieved from: <https://medium.com/@piyumika546/how-to-connect-mysql-databases-in-weka-using-mysql-jdbc-connector-4dbb60149577>
- Student Performance Analysis System (SPAS)*. (2015). ResearchGate.
- The Art of Agile Development*. O'Reilly Media.

Universiti Pertahanan Nasional Malaysia. (2017). *Buku Panduan Akademik Sarjana Muda Sesi Akademik 2017/2018* (2017th ed.).

Zoeken.(n.d.). *A Faculty Support System for student academic performance analysis*. Retrieved from: https://www.researchgate.net/publication/298646821_