International Journal Of Health, Engineering And Technology (IJHET) *E-ISSN 2829 - 8683*
Volume 3, Number 1, May 2024, *Page. 294 - 300*
Email : editorijhet@gmail.com

Analysis of Factors Causing Students' Failure to Complete Their Thesis on Time Using the Random Forest Algorithm

Riszki Fadillah[1)], Intan Nur Fitriyani [2)], Nur Indah Nasution [3)], Rahadatul 'Aisy Riadi [4)], Dinda Salsabila Ritonga [5)]
[1)] Information Technology, Faculty of Computer Science, Ika Bina Institute of Technology and Health
[2,5)] Information Systems, Faculty of Computer Science, Ika Bina Institute of Technology and Health
[3)]Health Administration, Faculty of Health Sciences, Ika Bina Institute of Technology and Health
[4)] Software Engineering, Faculty of Computer Science, Dumai University

*Coresponding Author
Email : fadillahriszki@gmail.com

Abstract
This research aims to analyze the factors that influence students' delays in completing final assignments using the Random Forest algorithm. The data used includes variables such as GPA, number of credits, employment status, frequency of guidance, organizational activities, and personal motivation. These variables were analyzed to determine their effect on students' ability to complete their final assignments on time. The Random Forest model is applied to predict whether students complete their final assignments on time or not. The model results show an accuracy of 63.33%, with the frequency of guidance and personal motivation being the most influential factors in completing the final assignment on time. Followed by the number of credits and GPA, which also have a significant but smaller influence. Organizational activity factors and employment status have a lower contribution to tardiness, but are still relevant in the context of student time management. Based on these results, research suggests the importance of academic guidance support and motivation management to help students overcome obstacles in completing their final assignments on time. This research, which uses the case of ITKES Ika Bina students, is expected to provide recommendations for universities in improving the academic mentoring process to support student graduation.

Keywords : Random Forest, Final Assignment Delay, Frequency Of Guidance, Personal Motivation, Model Accuracy

INTRODUCTION

The final assignment or thesis is the final academic requirement that students must complete to obtain a bachelor's degree. Completion not only measures the student's ability to apply the knowledge they have acquired during their studies, but also reflects the student's responsibility, independence and discipline in completing the overall academic process. However, the phenomenon of late completion of final assignments has become a common problem in various universities, both public and private. In fact, quite a few students fail to complete the maximum study time limit and end up dropping out.

Based on academic data at several universities in Indonesia, on average students complete their final assignments not on time, with an estimated delay of between 1 and 3 semesters. This condition not only harms students personally in the form of additional costs and psychological pressure, but also has an impact on institutions in terms of accreditation, academic efficiency and campus reputation.

Various factors influence a student's success or failure in completing their final assignment on time. These factors, among others, come from internal aspects of students such as the Grade Point Average (GPA), total study load (SKS), level of motivation, discipline and time management. Apart from that, external factors such as part-time work, family responsibilities, organizational activities, and the quality of relationships with supervisors also have an influence. The complexity of these various variables requires an analytical approach that is able to capture interactions between factors simultaneously.

In the current era of digital transformation, data science and machine learning approaches have become very effective tools for identifying patterns, analyzing trends, and predicting academic results more objectively. The Random Forest algorithm was chosen in this research because of its superior ability to classify data, handle large numbers of variables, and provide important information regarding the feature importance of each variable used.

This research uses the case of ITKES Ika Bina students, especially students of the Faculty of Health Sciences, who experienced problems in completing their final assignments. By using the Random Forest algorithm, this research aims to map the most dominant factors influencing the failure of students at ITKES Ika Bina in completing their final assignments on time. This model not only aims to produce predictions, but also as a tool for campuses in designing data-based intervention programs such as more intensive academic guidance, time management training, and early detection of high-risk students.

RESEARCH METHODS

This research is included in the descriptive quantitative category with a data mining approach using the Random Forest algorithm. This approach is used to analyze the factors that contribute to students' delays in completing their final assignments on time at the Faculty of Health Sciences, Ika Bina Institute of Technology and Health. The following is the research methodology carried out which can be seen in Figure 1. The research methodology is as follows.

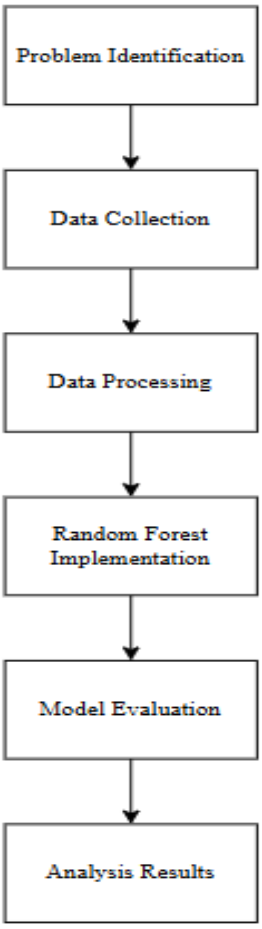

Figure 1. Research Methodology

Data Collection

Data is collected by:

1. Questionnaire: Contains questions related to motivation, organizational activities and work.

2. Academic data: Obtained from the university's academic information system (GPA, credits, TA completion status).
3. Limited interviews (optional): As reinforcement for qualitative data.

Data Processing and Analysis Techniques

Data Preprocessing

1. Normalization of numeric variables
2. Categorical variable encoding (One-Hot Encoding or Label Encoding)
3. Handling empty/null data

Splitting Data

1. Data is divided into training data (70%) and testing data (30%)

Application of the Random Forest Algorithm

1. Using RandomForestClassifier from Scikit-Learn library
2. Parameters such as n_estimators, max_depth, and criterion are adjusted for model optimization

Model Evaluation

1. Using metrics: Accuracy, Precision, Recall, F1-Score

Feature Importance Analysis to determine the dominant variable

RESULTS AND DISCUSSION

Preliminary Data

The following initial data used in this research are as follows:

Table 1. Initial Data

GPA	Number of credits	Employment Status	Tutoring Frequency	Organizational Activities	Personal Motivation	Final Assignment Status
2.75	131	Doesn't work	3	Yes	4	1
3.9	138	Work	9	No	5	0
3.46	148	Doesn't work	3	Yes	4	0
3.2	151	Work	9	No	2	0
2.31	160	Work	2	No	2	1
2.31	131	Doesn't work	2	No	5	0
…	…	…	…	…	…	…
2.05	131	Doesn't work	9	No	1	0
2.22	105	Doesn't work	10	No	4	1

Model Evaluation

Evaluate a model using accuracy, the steps generally involve dividing the dataset into training data and testing data, then training the model using the chosen algorithm (in this case Random Forest) and measuring how good the model is using accuracy metrics.

The following is a model evaluation with accuracy using the Random Forest algorithm which can be seen in Figure 2. Model Evaluation with Accuracy.

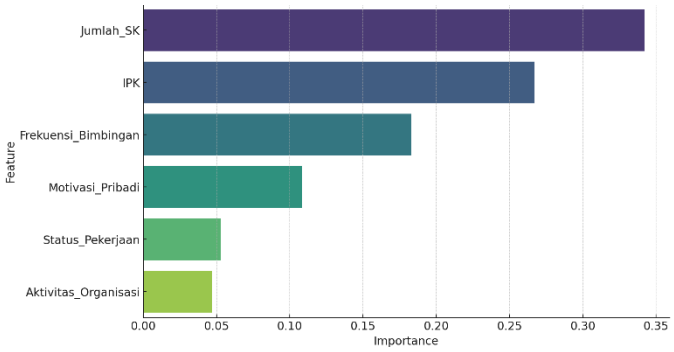

Figure 2. Model Evaluation with Accuracy

The accuracy of the Random Forest model obtained was 63.33% based on test data. The graph above shows the importance of features in predicting delays in completing student final assignments, sorted by their contribution to the model.

Analysis of Results

Based on the evaluation results of the Random Forest model, we can analyze the factors that influence students' delays in completing their final assignments by utilizing the accuracy of the model and the importance of the resulting features.

Model Accuracy

The model accuracy is 63.33%, which shows that the Random Forest model can predict the final task completion status (on time or not on time) with an accuracy level of 63.33%. This accuracy shows that the model is quite good at predicting delays, but there is room for improvement. Usually, to achieve more optimal results, we can perform hyperparameter tuning, enrich the dataset with more features, or use additional data processing methods.

Feature Importance

After observing the feature importance results, it can be seen that some features contribute more to the model predictions. Here is an analysis of some of the most important features:

1. Tutoring Frequency: This feature has the highest importance, which shows that the number of meetings between students and supervisors is a very influential factor in the success of completing final assignments on time. The more often students do tutoring, the more likely they are to complete their final assignments on time.
2. Personal Motivation: Students who have high personal motivation are more likely to complete their final assignments on time. This also shows the importance of emotional support and internal motivation in completing the final assignment.
3. Number of credits and GPA: Although the influence is slightly smaller than the frequency of guidance and personal motivation, the variables Number of credits and GPA still play a role in completing the final assignment on time. Students with a higher GPA and a lower number of credits are more likely to complete their final assignments on time.
4. Employment Status and Organizational Activity: These two features have less influence in the model's predictions, although they remain relevant. Students who work or are involved in campus organizations may have difficulty allocating time to complete their final assignments on time.

Table 2. Analysis Results

No	Feature	Importance (%)
1	Tutoring Frequency	24.51
2	Personal Motivation	19.84
3	Number of credits	18.32
4	GPA	14.79
5	Organizational Activities	12.60
6	Employment Status	10.94

The following table shows the importance of features in the Random Forest model which is used to predict student delays in completing final assignments. Each feature is given an importance value which shows its contribution to the prediction model. The higher the importance value, the greater the influence of the feature in predicting the final task completion status.

Tutoring Frequency (24.51%)

This feature has the highest level of importance which shows that the frequency of meetings between students and supervisors plays a very significant role in determining students' success in completing their final assignments on time. Intense interaction between students and supervisors can provide opportunities for students to receive more in-depth guidance regarding the academic problems they face. With frequent meetings, students have more opportunities to get clarification on topics they don't understand, as well as clearer directions regarding completing their final assignments. Apart from that, good relationships and smooth communication with supervisors can motivate students to stay focused and avoid procrastination. This further increases their chances of completing their final assignment according to the specified schedule, and ultimately increases the graduation rate on time. So, the frequency of meetings is one of the main keys to students' success in completing their studies on time.

Personal Motivation (19.84%)

Students' personal motivation is the second most influential factor in this model, indicating that students' internal drive plays a crucial role in their success in completing their final assignments on time. Students who have a high level of motivation tend to have a strong commitment to completing their final assignment even though they face various challenges, both academic and personal. This motivation not only encourages them to overcome obstacles that arise during the work process, but also helps them stay focused and organized in achieving their goals. This factor is closely related to how much they appreciate the importance of the final assignment as one of the graduation requirements, as well as their understanding of the long-term impact of completing their studies well and on time. Students with high personal motivation will be more persistent in facing difficulties, take the initiative to find solutions, and utilize existing resources to advance the completion of their final assignment. Therefore, personal motivation can be considered as the main driver that supports students' smooth academic process until they achieve successful completion of their final assignment.

Number of credits (18.32%)

The number of credits taken by students also has a significant influence on predicting final assignment completion on time. Students who take fewer credits usually have more free time that they can allocate to focus fully on their final assignments. In this way, they tend to have a greater chance of completing the final assignment on time, because they are not burdened with too many courses that must be done simultaneously. On the other hand, students with a larger number of credits often face difficulties in managing their time efficiently. A larger academic load requires them to divide their attention and time between completing other coursework and work on their final project. In conditions like these, students may find it difficult to give full focus to the final assignment, which can result in delays in completing it. Therefore, the number of credits taken by students influences their ability to manage time and academic priorities, which in turn influences their ability to complete their final assignments according to the set schedule.

GPA (14.79%)

Cumulative Achievement Index (GPA) is an indicator that reflects the quality of a student's overall academic performance, and has a significant influence in predicting final assignment completion on time. Students with higher GPAs generally indicate that they have better time management skills and organizational abilities. This skill really supports them in planning and completing various academic assignments, including final assignments, in a more structured and efficient manner. By having better study habits and more optimal time management, students with a

high GPA tend to be able to allocate their time better between other coursework and final assignments.

Even though its contribution may not be as big as other factors such as the frequency of meetings with supervisors or the level of personal motivation, GPA still plays a role in creating a strong academic foundation for students to complete their final assignments well. Students with high GPAs usually have better maintained academic habits, as well as the ability to deal with pressure better, which allows them to stay focused and productive. Therefore, even though it is not a major factor, GPA still plays an important role in supporting student academic success, including in completing final assignments on time.

Organizational Activities (12.60%)

Student organization activities have less influence in this model, although they can still have an impact on the completion of students' final assignments. Students who are active in campus organizations usually have busier schedules, which can reduce the time available to focus on their final assignments. Involvement in various organizational activities often requires a significant time commitment, so students may find it difficult to divide their attention between academic tasks and organizational obligations. This could risk disrupting the optimal time allocation to complete the final assignment according to schedule.

However, on the other hand, involvement in organizations also provides benefits in developing time management skills, leadership and the ability to work in teams. These skills, although not directly related to completing the final assignment, can help students manage their time more efficiently and complete various assignments in a more structured manner. Apart from that, experience in organizations can also enrich students' abilities in dealing with pressure and challenges, which in turn can support them to remain productive and focused in completing their final assignments. Therefore, even though the influence is smaller, organizational activities still have the potential to make a positive contribution to student academic success in the long term.

Employment Status (10.94%)

Students' employment status, especially whether they work part-time, can affect their ability to complete their final assignments on time. Students who work part-time often face challenges in managing their time, because work obligations reduce the time available for studying or working on final assignments. This puts them at risk of experiencing delays in completing final assignments, especially if part-time work takes a lot of time and energy. Thus, limited time is the main obstacle that must be managed well.

Although its influence is smaller compared to other factors such as frequency of guidance or personal motivation, employment status remains an important consideration in managing time for final assignment completion. Students who work part-time need to have good time management skills so they can divide their time efficiently between work, studies and final assignments. Therefore, although part-time work can provide valuable experience and financial benefits, students must be able to ensure that this work does not hinder them from completing their studies on time.

CONCLUSION

The Random Forest model succeeded in predicting delays in completing student final assignments with an accuracy of 63.33%, with the main factors influencing predictions being the frequency of guidance and personal motivation. Students who provide guidance with lecturers more often and have high motivation tend to complete their final assignments on time. Apart from that, the number of credits and GPA also play a role, although with a smaller contribution. Organizational activities and employment status have a lower influence, but are still relevant in the context of student time management. To increase model accuracy, hyperparameter tuning and adding additional features such as the number of revisions to the final assignment or the duration of the final

assignment are required. These findings provide useful insights for educational institutions, especially at ITKES Ika Bina, to focus more on guidance support, personal motivation, and study load management to help students complete their final assignments on time, and prevent delays in completion that could affect their graduation.

REFERENCES

Alamsyah, H., & Suryanto, A. (2021). The impact of academic stress and time management on university students' success in thesis completion. International Journal of Academic Research in Business and Social Sciences, 11(6), 75-85.

Damayanti, I. A., & Firdaus, M. (2020). Factors affecting the timely completion of final projects: A student perspective. International Journal of Educational Management, 34(2), 150-160.

Ghozali, I. (2018). Aplikasi analisis multivariate dengan program IBM SPSS 25 (9th ed.). Badan Penerbit Universitas Diponegoro.

Hanif, M. S., & Khalid, M. (2020). Analysis of factors influencing academic procrastination among university students. Journal of Educational Psychology, 112(3), 315-324.

Hidayati, I. A., & Firdaus, M. (2020). Psychological factors influencing procrastination in university students. Psychology in Education and Society, 18(4), 54-64.

Kurniawati, D., & Prasetyo, Y. (2019). The impact of motivational factors on student success in completing final projects. Journal of Higher Education Research, 15(2), 121-130.

Marimin, M., & Widodo, S. (2017). Metode pengambilan keputusan dengan aplikasi komputer. Salemba Empat.

Noviana, N., & Suryani, D. (2021). Factors that influence students' academic procrastination in completing their final projects at universities in Jakarta. Journal of Educational Studies, 9(1), 65-73.

Rahman, M., & Hasibuan, A. (2020). Procrastination and academic achievement: An investigation into the academic behavior of students in Indonesia. Journal of Educational Psychology and Counseling, 8(1), 25-39.

Santosa, S. (2016). Statistika untuk penelitian (3rd ed.). Pustaka Pelajar.

Setiawan, H., & Astuti, A. (2021). Academic procrastination and its effect on students' academic achievement: A case study of Indonesian university students. Journal of Educational Psychology, 9(3), 240-250.

Sugiyono, S. (2020). Metode penelitian kuantitatif, kualitatif, dan R&D. Alfabeta.

Widodo, S., & Suhartono, Y. (2021). The role of peer support and self-regulated learning in reducing academic procrastination among final-year students. Journal of College Student Development, 62(4), 387-394.

Yusuf, M., & Mulyadi, A. (2019). Understanding the correlation between social support and academic performance among undergraduate students. Asian Journal of Social Psychology, 22(3), 291-299.

Sarwono, J. (2019). Psikologi pendidikan (3rd ed.). Rajawali Pers