Design Of Decision Support System Final Project Topic In Computer And Network Engineering Study Program State Polytechnic Of Ujung Pandang

Fadli Tamrin[1)], **Meylanie Olivya** [2)], **Maryam Azzahra** [3)]
[1] Teknik Informatika dan Komputer, Politeknik Negeri Ujung Pandang, Jl. Perintis Kemerdekaan KM.10, 90245, Makassar
[2] Teknik Informatika dan Komputer, Politeknik Negeri Ujung Pandang, Jl. Perintis Kemerdekaan KM.10, 90245, Makassar
[3] Teknik Informatika dan Komputer, Politeknik Negeri Ujung Pandang, Jl. Perintis Kemerdekaan KM.10, 90245, Makassar
[a] fadlitamrin@poliupg.ac.id, [b] meylanie@poliupg.ac.id, [c] mayyamazz04@gmail.com

***Abstract*—This research is motivated by the difficulties often experienced by students in determining the topic of their final assignment that suits their interests and abilities, which results in delayed graduation. The purpose of this study is to create an accurate final assignment topic decision support system model using the Random Forest Algorithm, design a website-based decision support system to assist students in determining the topic of their final assignment, and evaluate the efficiency and quality of the system. The data used are academic grades and interests of students in the Computer and Network Engineering study program at the Ujung Pandang State Polytechnic. The results of the study show that the Random Forest Model is able to produce final assignment topic recommendations with a Cross-validation accuracy of 61.36%. The website-based decision support system designed can assist students in determining the topic of their final assignment according to their interests and academic abilities, as well as improve the efficiency and quality of the final assignment topic selection process. This research contributes to making it easier for students to find the right final assignment topic, as well as providing guidance for the development of similar recommendation systems in the future.**

Keywords—Final Project Topic, Random Forest, Decision Support Sistem

I. Introduction

Final Project (TA) is a graduation requirement that must be prepared by students as part of their academic education and reflects their ability to conduct research in accordance with their field of study [1]. However, many students often have difficulty in determining a TA topic that matches their interests and expertise. In the Computer and Network Engineering Study Program of the Ujung Pandang State Polytechnic, many students do not consider supporting and inhibiting factors comprehensively when choosing a thesis topic that they will research, which can ultimately hinder the process of completing their TA [2].

Based on a survey of 14 alumni and 30 final year students in the study program, 47.7% revealed that difficulty in determining a TA topic that matches their abilities is the main inhibiting factor in completing the final assignment. This causes some students to experience a delay in graduation of up to 15% from what it should be. These difficulties include mastery of theory and practice, data collection, and communication with supervisors [3]. Therefore, choosing the right TA topic, not only in accordance with the interests but also the academic abilities of students, is very important to support the smooth process of completing the study[4]. To overcome these problems, a classification-based decision support system is needed that can help students choose TA topics that match their academic abilities and interests, based on their academic track records[5].

Several studies have shown the superiority of the Random Forest algorithm in predicting graduation. One study showed that Random Forest has a higher accuracy (99.49%) compared to SVM (98.98%) in predicting MAS student graduation [6]. Another study showed that Random Forest achieved the highest accuracy of 77.35% compared to other algorithms such as Naïve Bayes, Decision Tree, K-NN, and SVM in predicting student graduation [7]. Another study proved that Random Forest is superior to Gradient Boosting in analyzing the timeliness of student graduation, with an accuracy of 82.64% [8]. Based on these results, this study uses the Random Forest algorithm for graduate profile classification because of its ability to manage complex data, tolerance to overfitting, and produce good accuracy estimates [9].

Similar research has been conducted by applying the Random Forest algorithm to map MSIB fields based on student course grades, and produced 80% accuracy, 80% precision, and 82% recall [10]. This research focuses on the application of Random Forest in a final project topic recommendation system, in order to test its effectiveness

in providing relevant recommendations based on academic interests and abilities, while also filling the research gap in higher education related to determining final project topics.

The Random Forest algorithm was chosen in this study because of its superior ability to manage complex data and provide accurate classification results. The implementation of this system is expected to increase efficiency and accuracy in selecting TA topics, as well as help reduce the number of students who experience delays in graduation due to difficulty finding suitable topics.

II. Research Methodology

Research methods require a research approach to ensure the right structure and direction, so that the research results are in line with the objectives that have been set. System design refers to the conceptual arrangement that clearly describes the desired system design in the context of this research. The flow of this research system can be seen in the figure below.

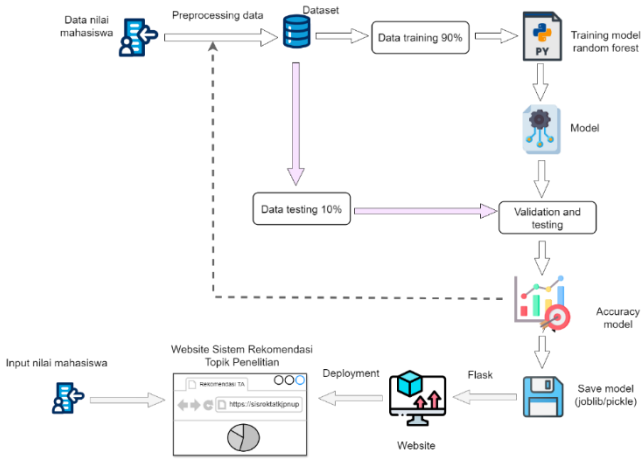

Figure 1. System Architecture

Figure 1 is the system architecture that will be used in this study. Student academic and interest data are processed and preprocessed to form a dataset, which is then separated into training data (90%) and testing data (10%). Training data is used to train the Random Forest Model. The model obtained after training using Random Forest is then evaluated using testing data. After evaluation, the Random Forest Model is saved for integration with the recommendation system application. This system receives student data and uses the model to provide recommendations for final project topics. The recommendation results are presented through a user interface built using Flask. This allows users to easily use the system to select final project topics.

III. Results and Discussion

Data of 213 students were collected from SIMPONI and questionnaires, consisting of grades of 30 vocational courses grouped by field of study and graduate profile, while general courses were not included. The data overview can be seen in Figure 2

Nim	Nama Mahasiswa	Arsitektur Komputer	Komunikasi Data	Bengkel Komputer & Jaringan	Dasar Pemrograman	Struktur Data & Analisis Algoritma	Sistem Digital	Jaringan Komputer	Statistik dan Probabilitas	Metode Komputasi	Desain Algoritma	Teknologi Jaringan Komputer	Basis Data	Sistem Operasi	Rekayasa Perangkat Lunak	Rekayasa Jaringan Komputer Dasar	Administrasi Basis Data	Pemrograman Berorientasi Objek
42517002	MUHAMMAD RAHMAT	B+	C	A	C	A	B+	B	B+	B	B	B	B+	B	A	C+	A	C
42517003	RIKA HARIANTI	B+	C+	A	B+	B	B+	B+	A	A	B	A	A	A	A	B+	A	B+
42517004	MUHAMMAD YUSRIL ZAENAL	B	C+	A	B+	A	A	B+	A	B+	B+	A	A	A	A	B+	A	B
42517005	ANDI FACHRUL REZA	B+	C+	B+	C	C+	B+	C+	A	A	B+	B+	B+	A	A	B+	B	B
42517006	SEPTY AINUN ARIATI	B+	B	A	C+	A	A	A	A	A	B	A	A	A	A	B+	A	B
42517007	A. M. YUSRIL IKA RAMADHAN	A	B	A	B+	A	C+	B	A	B	B	A	B+	A	A	B	A	B+
42517008	HASRI AINUN BR A PERTIWI	B	A	A	C	B	B+	B+	B+	C	B+	A	A	A	A	B+	A	B

Figure 2 Academic Grade Data

The main purpose of this design is to ensure that the system runs according to specifications, is efficient, and easy to use. After data collection, preprocessing is carried out by converting grade values to numbers using a dictionary, filling missing values with averages, and encoding categorical columns such as interests using LabelEncoder. Furthermore, the determination of graduate profile labels is carried out using rule-based logic based on the analysis of average academic grades for classification using Random Forest. An overview of the course groups can be seen in Table 1

Table 1 Course Group

No	Subject	Network Admin	Software Engineer	Data Engineer	IoT Engineer
1	arsitektur_komputer	✔			
2	komunikasi_data	✔			
3	bengkel_komputer_jaringan	✔			
4	jaringan_komputer	✔			
5	teknologi_jaringan_komputer	✔			
6	sistem_operasi	✔			

7	rekayasa_jaringan_komputer_dasar	✔			✔
8	administrasi_jaringan	✔			
9	rekayasa_jaringan_komputer_lanjut	✔			✔
10	manajemen_dan_desain_jaringan	✔			
11	virtualisasi_dan_komputasi_awan	✔			
12	cyber_security	✔			✔
13	jaringan_wireless_mobile	✔			✔
14	dasar_pemrograman		✔		
15	sistem_digital		✔		
16	rekayasa_perangkat_lunak		✔		
17	pemrograman_berorientasi_objek		✔		
18	desain_dan_pemrograman_web		✔		
19	sistem_terdistribusi		✔		✔
20	aplikasi_mobile		✔		
21	rekayasa_web		✔		
22	sistem_cerdas	✔	✔	✔	
23	struktur_data_analisis_algoritma			✔	
24	statistik_dan_probabilitas			✔	
25	metode_komputasi			✔	✔
26	desain_algoritma			✔	
27	basis_data			✔	
28	administrasi_basis_data			✔	
29	big_data			✔	
30	sistem_embedded_dan_iot				✔

DATASET GABUNG CLUSTER.c...

clustered_data_berd_kondisi.csv

	Minat Bidang Ilmu	Minat Profile Lulusan	Highest_Avg_Field	\
0	3	0	Avg_Data_Engineer	
1	7	3	Avg_Data_Engineer	
2	4	1	Avg_Network_Admin	
3	7	3	Avg_Data_Engineer	
4	3	1	Avg_IoT_Engineer	
..	...	...	...	
208	7	3	Avg_Network_Admin	
209	4	0	Avg_Network_Admin	
210	7	3	Avg_Network_Admin	
211	2	0	Avg_Network_Admin	
212	4	0	Avg_Network_Admin	

	Predicted Profile
0	Associate Data Engineer
1	Associate Data Engineer
2	Network Administrator
3	Associate Data Engineer
4	IoT Engineer
..	...
208	Network Administrator
209	Network Administrator
210	Network Administrator
211	Network Administrator
212	Network Administrator

Figure 3 Initial Data Files and Grouping Results

Before programming Random Forest, additional preprocessing is performed to ensure the dataset is ready for use. Identity features such as "NIM" are removed because they can interfere with the training process, while "nama_mahasiswa" is also removed to avoid errors during programming. The "Predicted Profile" label comes from rule-based logic based on the highest average value. The dataset used is stored with the program and its image can be seen in the following image.

Table 2 Convert Grade to Numeric Value

Grade	Numeric Value
A	4
A-	3.7
B+	3.3
B	3
B-	2.7
C+	2.3
C	2
C-	1.7
D	1.3
D-	1
E	0
T	0

After the data is processed, the categorical columns are encoded using Label Encoding into unique numeric values, such as in 'Interest in Field of Study' and 'Interest

in Graduate Profile'. The data is then separated into features (academic grades, interests) and targets (Predicted Profile). The numeric data is normalized so that the features have the same scale, improving model performance. To handle class imbalance, RandomOverSampler is used to add minority class samples to balance. Next, the best hyperparameter search is carried out with Randomized Search CV, testing combinations such as the number of trees, depth, and other parameters for model optimization. The model is evaluated with cross-validation to prevent overfitting and ensure stability. The model and encoder are saved using joblib so that they can be reused. The training model is saved in pickle format in Google Colaboratory for deployment with Flask on the website, as can be seen in the following image.

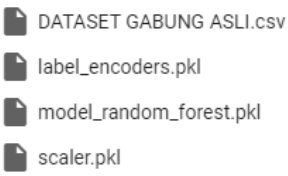

Figure 4 Random Forest Dataset and Model Results

This website application is built with Python and uses the Random Forest model as an asset called in the application. Python code is created to predict the topic of the final project based on academic grades and student interests. The system processes the data according to the model, then displays the prediction results as a classification output in the final project decision support application.

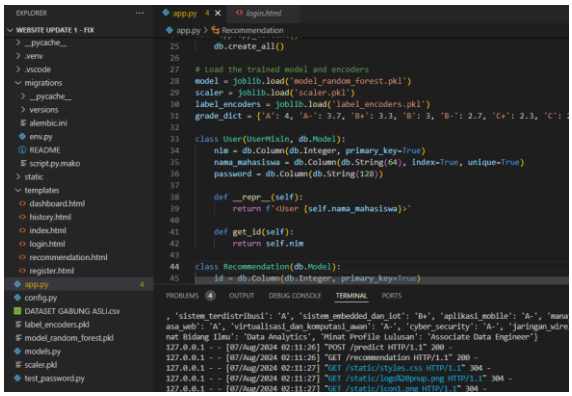

Figure 5 Deployment Model

This system is a Final Project Topic Decision Support System that allows students to know the recommendations for final project topics based on student academic grade data, and data on interests in the field of science and interests of their graduate profiles. Then the results of the final project topic recommendations will be displayed based on the fields that match the recommended graduate profile based on the predictions of the trained Random Forest algorithm.

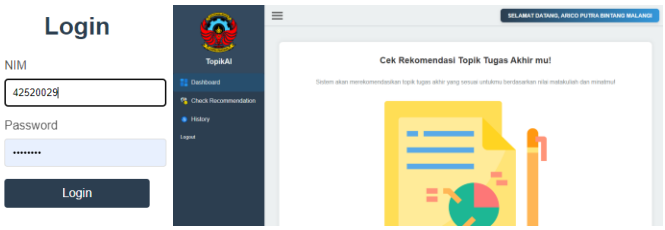

Figure 6. System Login and Dashboard

This login page is a page for users to enter the Final Project Topic Decision Support System application. Where users will enter the NIM and password that have been registered in the database. And the main page or dashboard that is displayed after the user has successfully logged in to the website. Where there is a "Check Recommendation" button to go to the page where users can check the recommendations for their final project topics. Can be seen in Figure 6.

Figure 7 Value View Page

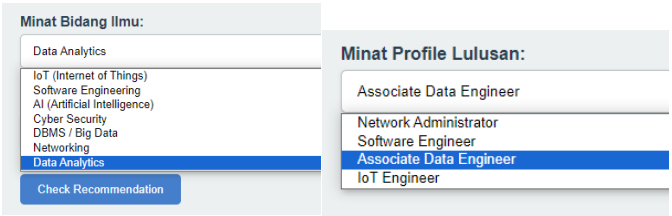

Figure 8 Interest Input Form

This page displays student grades for 30 courses (see Figure 7), as well as an input form for academic interests and graduate profiles for prediction purposes (see Figure

8). There is also a “Check Recommendation” button to see the results of the final project topic recommendations.

Figure 9. System Recommendation Results View

In Figure 9 This page is a page where users will see the results of the model predictions that are recommendations for graduate profiles and final project topics that are in accordance with the field of science, where by knowing the recommendations of both, it is hoped that students can more easily decide on a suitable final project topic. The results of the final project topic recommendations are based on course grades, for interests in the field of science and graduate profiles to be points that increase the percentage when recommendations based on grades are in accordance with their interests.

System testing aims to ensure the accuracy, efficiency, and ease of use of the final project topic recommendation system. The results of RandomizedSearchCV show that the best combination of hyperparameters is: `bootstrap=False`, `max_depth=8`, `min_samples_leaf=5`, `min_samples_split=7`, and `n_estimators=348`. Cross-validation with 5 folds produces an average score of 0.61. The fitting process was carried out 250 times to find the best configuration.

3.1 Cross-Validation Testing

The cross-validation results on 5 folds show the following accuracies for each fold:

1) Fold 1: 0.53846154
2) Fold 2: 0.47058824
3) Fold 3: 0.58823529
4) Fold 4: 0.68627451
5) Fold 5: 0.78431373

The average score from the cross-validation is 0.6136, which gives a general idea of the model's performance on data unseen during training.

3.2 Black Box Testing

Black Box testing on the website runs well and according to the expected function. All tested scenarios were successful with the expected results. Based on these results, further testing of users will be carried out, and the results will be explained in the questionnaire results testing section.

3.3 Questionnaire Testing

Based on the questionnaire testing using the TAM (Technology Acceptance Model) model, which consists of 9 questions and 47 respondents of final year Computer and Network Engineering students, the aspects of benefits, convenience, and user satisfaction towards the Decision Support System for the Student Final Project Topic that has been created are "very good".

The results of the alignment analysis showed that 78.17% of alumni who graduated on time chose thesis topics from the top two system recommendations, while only 21.81% chose outside of that. This shows that the system has a high level of accuracy in providing thesis topic recommendations.

To assess the effectiveness of the thesis topic recommendation system, a validation test was conducted on alumni from the study program where the research took place. The purpose of this validation was to evaluate how well the system’s recommendations aligned with the actual research topic choices of students who graduated on time. Data were collected from 55 alumni whose graduation status and chosen research topics were available through academic records and interview/questionnaire responses.

The analysis compared each alumnus' selected topic with the topics recommended by the system. The level of alignment was then calculated as a percentage by dividing the number of alumni whose chosen topics matched the system’s top or second recommendation by the total number of alumni, multiplied by 100%. The results showed that 34.54% of the alumni selected the top recommended topic, 43.63% chose the second recommendation, and 21.81% selected a topic outside the top two suggestions. This means that 78.17% of on-time graduates selected topics that aligned with the system's first or second recommendation, indicating that the

system provides accurate and beneficial research topic suggestions.

Table 3 Validation of Effectiveness

On-Time Graduation Status	Number of Alumni	Topic Prediction According to Recommendation (%)	Topic Prediction in Second Place Recommendation (%)	Predicted Topics Outside Top Two (%)
Ya	55	34.54%	43.63%	21.81%
Total		**78.17%**		**21.81%**

IV. Conclusion

1. The Random Forest model has an accuracy of 61.36% based on the average Cross-validation.
2. The website-based system effectively helps students choose topics according to their abilities and interests.
3. The features of academic grades, interests in fields of study, and interests in graduate profiles increase model accuracy.
4. The data "interests in fields of study" and "interests in graduate profiles" increase recommendation accuracy by 2-3%.
5. The application speeds up the topic selection process and increases efficiency.
6. Validation with a larger dataset is needed to avoid overfitting and ensure model generalization.

Acknowledgement

The author would like to express his deepest gratitude to God Almighty, his parents, friends and companions, his two supervisors, and all the lecturers and staff of the Informatics and Computer Engineering Department, Computer and Network Engineering Study Program, Ujung Pandang State Polytechnic and all parties who have helped in compiling this research until completion.

References

[1] R. Maryadi, "Analisis Dan Perancangan Sistem Rekomendasi Topik Tugas Akhir Menggunakan Metode Hybrid And Ensemble-Based Recommender Systems," 2019.

[2] N. R. I. Fahmi, A. C. Prihandoko, and W. E. Y. Retnani, "Implementasi Metode Fuzzy AHP pada Sistem Penunjang Keputusan Penentuan Topik Skripsi (Studi Kasus : Program Studi Sistem Informasi Universitas Jember) (Implementation of Fuzzy AHP Method in Decission Support System Determination of Topic Thesis) (Case S," *Berk. Sainstek*, vol. 2, pp. 76–81, 2021.

[3] H. Haviluddin, S. J. Patandianan, G. M. Putra, N. Puspitasari, and H. S. Pakpahan, "Implementasi Metode K-Means Untuk Pengelompokkan Rekomendasi Tugas Akhir," *Inform. Mulawarman J. Ilm. Ilmu Komput.*, vol. 16, no. 1, p. 13, 2021, doi: 10.30872/jim.v16i1.5182.

[4] M. R. Muttaqin and M. Defriani, "Algoritma K-Means untuk Pengelompokan Topik Skripsi Mahasiswa," *Ilk. J. Ilm.*, vol. 12, no. 2, pp. 121–129, 2020, doi: 10.33096/ilkom.v12i2.542.121-129.

[5] A. Salam, F. P. Albahri, and Fathurrahmad, "Sistem Rekomendasi Tugas Akhir Mahasiswa pada AMIK Indonesia untuk Mendukung Merdeka Belajar-Kampus Merdeka Menggunakan Metode Collaborative Filtering (CF)," *J. JTIK (Jurnal Teknol. Inf. dan Komunikasi)*, vol. 6, no. 2, pp. 281–288, 2022, doi: 10.35870/jtik.v6i2.420.

[6] A. Darmawan, I. Yudhisari, A. Anwari, and M. Makruf, "Pola Prediksi Kelulusan Siswa Madrasah Aliyah Swasta dengan Support Vector Machine dan Random Forest," *J. Minfo Polgan*, vol. 12, no. 1, pp. 387–400, 2023, doi: 10.33395/jmp.v12i1.12388.

[7] J. Zeniarja, A. Salam, and F. A. Ma'ruf, "Seleksi Fitur dan Perbandingan Algoritma Klasifikasi untuk Prediksi Kelulusan Mahasiswa," *J. Rekayasa Elektr.*, vol. 18, no. 2, pp. 102–108, 2022, doi: 10.17529/jre.v18i2.24047.

[8] M. L. Mu'tashim and A. Zaidiah, "Klasifikasi Ketepatan Lama Studi Mahasiswa Dengan Algoritma Random Forest Dan Gradient Boosting (Studi Kasus Fakultas Ilmu Komputer Universitas Pembangunan Nasional Veteran Jakarta)," *Semin. Nas. Mhs. Bid. Ilmu Komput. dan Apl.*, vol. 4, no. 1, pp. 155–166, 2023.

[9] T. Agnes *et al.*, "Sistem Rekomendasi Pencarian Indekos di Surabaya Menggunakan Random Forest," vol. 2023, no. Senada, pp. 172–177, 2023.

[10] N. Aini, M. Arif, I. T. Agustin, and Z. B. Toyibah, "Implementasi Algoritma Random Forest untuk Klasifikasi Bidang MSIB di Prodi Pendidikan Informatika," *J. Inform.*, vol. 11, no. 1, pp. 11–16, 2024, doi: 10.31294/inf.v11i1.20637.