Instal : Jurnal Komputer

E-ISSN: 2808-683X
Edisi : Volume 18 Nomor 01 | Tahun 2026

Available online at https://journalinstal.cattleyadf.org/index.php/Instal/index

Classification of Medical Students' Skills Using the Random Forest Method Based on Practicum and Theory Scores

Abdul Chaidir Harahap[1*], Muhammad Irfan Syarif[2]
[1,2]Postgraduate Program, Master of Information Technology, Panca Budi Development University, Medan, Indonesia

ARTICLE INFO

Keywords:

Medical Student Competency Classification, Random Forest-Based Predictive Modeling in Education, Academic Performance Assessment Using Practical and Theoretical Scores, Machine Learning Applications in Medical Education Evaluation, Early Identification of Low-Skill Medical Students

ABSTRACT

This study classifies the skill levels of medical students based on their practical and theoretical scores using the Random Forest method. The data consist of secondary data obtained from the Faculty of Medicine, Universitas Muhammadiyah Sumatera Utara, co vering students from the Medical Education Study Program across three cohorts (2021, 2022, and 2023) in the block "Basic Organ System of Special Sense, Reproduction and Urinary Tract." The dataset includes 1,074 students with five assessment features: block exam score (UAB), practical score (average of Histology, Anatomy, Physiology, and Biochemistry), tutorial score, SGD attitude score, and PIM attitude score. The Random Forest method is used to classify students into "Good" (final score ≥ 75) and "Poor" (final score < 75) skill categories. The results indicate that the model achieves an accuracy of 93.33%, with precision values of 94.12% for the "Good" category and 90.00% for the "Poor" category, as well as recall values of 97.56% and 78.26%, respectively. The most influential features are the block exam score (0.378), practical score (0.295), and tutorial score (0.192). The study also generates 11 expert-validated classification rules (average score 4.69/5.00) that can support early identification of students with lower skill levels. The Random Forest model demonstrates effectiveness and consistency, achieving accuracy above 92% across all cohorts, and supports the development of a machine learning–based evaluation system for medical students at Universitas Muhammadiyah Sumatera Utara.

Corresponding Author:
Abdul Chaidir Harahap
Universitas Pembangunan Panca Budi
Email: [1*]abdulchaidir52@gmail.com, [2]irfanberbagi@gmail.com

INTRODUCTION

Medical education requires a comprehensive evaluation system because it focuses not only on the mastery of theory, but also on practical skills that are an important part of the competence of prospective doctors [1]. Based on the Indonesian Doctor Competency Standards (SKDI), students are required to master cognitive and psychomotor aspects in a balanced manner as the basis for running the medical profession [2]. At the Faculty of Medicine, University of Muhammadiyah North Sumatra (UMSU), the learning process is applied through a block system that integrates theory and practice, including the Basic Organ System of Special

Page : 86-98
https://doi.org/10.54209/jurnalinstall.v18i01.488

Sense, Reproduction and Urinary Tract block which involves various assessment components such as block exams, practicums, tutorials, and attitudes [3].

Academic data shows that there is a variation in student achievement, where some students show optimal performance, while others experience difficulties in certain aspects. This shows the importance of early identification of students with low skill potential to support the provision of appropriate learning interventions [4]. Therefore, this study proposes the use of the Random Forest method to Classify students' skill levels based on practicum and theory scores, and identify influential factors.

This research aims to develop an accurate classification model, evaluate its performance using relevant evaluation metrics, and produce classification rules that can be used as a basis for decision-making in the academic evaluation process. The results of the research are expected to contribute to the application of machine learning in the medical education evaluation system, especially in supporting the early detection of students who need more attention.

METHODS

This study uses a quantitative approach with a classification method to analyze the skill level of medical students based on practicum and theory scores. The research methodology is systematically designed so that each stage can be carried out in a structured manner and produce objective and measurable outputs. The main focus of this study is to apply *the Random Forest* algorithm in classifying student skills, identifying the most influential features, and evaluating the performance of the model using accuracy, precision, *recall*, and f1-score *metrics* on the Faculty of Medicine's student score dataset.

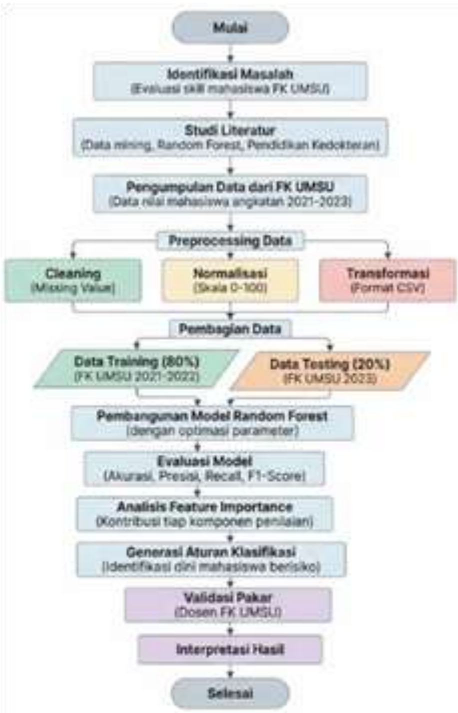

Figure 1. Research Methodology Flow Diagram

In general, the flow of this research methodology consists of five main stages, namely data preparation and acquisition (including problem identification and literature study), data pre-processing (such as cleaning, normalization, and transformation), data sharing and machine learning modeling (using the Random Forest algorithm), performance evaluation and analysis (evaluation metrics, feature importance, and classification rule generation), and expert validation and interpretation of results. The conceptual flow of this research is represented in

Page : 86-98
https://doi.org/10.54209/jurnalinstall.v18i01.488

the form of a methodological flow chart, as shown in the figure, which illustrates the relationship between the stages of the research from beginning to end.

2.1 Medical Education at the University of Muhammadiyah North Sumatra

The Faculty of Medicine, University of Muhammadiyah North Sumatra (UMSU) was established in 2008 and has graduated thousands of doctors spread throughout Indonesia. The curriculum applied refers to the Indonesian Doctor Competency Standards (SKDI) set by the Indonesian Medical Council [2].

The learning system at FK UMSU uses a Student-Centered Learning (SCL) approach with the Problem-Based Learning (PBL) method which is integrated in a block system. Each block is designed to integrate basic science and clinical science with a specific topic [5].

2.2 Evaluation Block Systems and Components

The block system at FK UMSU organizes learning materials based on organ systems or certain topics. The "Basic Organ System of Special Sense, Reproduction and Urinary Tract" block is a block that studies the basic organ system of humans [3]. The assessment components in the UMSU medical block consist of:

Table 1. Components of Block Assessment at FK UMSU

Components	Weight	Description	Data Source
Final Block Test (UAB)	0.5	A written test that measures mastery of theoretical material	Academic Section of FK UMSU
Tutorial	0.2	Evaluation of the tutorial discussion process	Tutorial Coordinator
Internship	0.2	Average laboratory practicum score (Histology, Anatomy, Physiology, Biochemistry)	Integrated Laboratory FK UMSU
***Attitude* SGD**	0.05	Sikap dalam Small Group Discussion	Tutor
***Attitude* PIM**	0.05	Attitudes during practicum	Practicum Lecturer

2.3 Data Mining dan Knowledge Discovery

Data Mining is a process used for data processing to obtain information [6]. Data mining is defined as the process of extracting useful information and previously unknown patterns from large databases. Knowledge Discovery in Databases (KDD) is a whole process that includes data selection, preprocessing, transformation, data mining, and interpretation of results [8]. In the context of medical education, data mining has been used for a variety of purposes, including prediction of competency exam passing, identification of at-risk students, and curriculum evaluation [9].

2.4 Random Forest Method

Random Forest is an ensemble learning method developed by Leo Breiman in 2001 [10]. This method combines many independently constructed decision trees and produces predictions based on a majority vote of the entire tree. Random Forest has several advantages over other classification algorithms [11], namely High Accuracy, Robust to Outliers, Handling High Dimensional Data, Feature Importance and Handling Missing Value. The Random Forest method works by building a large number of random decision trees, where each tree is trained using a random subset of the training data. The final prediction is made through majority aggregation (voting) for classification or averaging for regression [12].

One of the advantages of Random Forest is its ability to calculate feature importance, which is the contribution of each feature in the classification process [13]. Clustering on a data

Page : 86-98
https://doi.org/10.54209/jurnalinstall.v18i01.488

is a stage to classify data sets whose class attributes have not been described, in terms of the concept of clustering is to maximize and minimize intra-class similarity [14].

2.5 Population and sample

2.5.1 Population

The population in this study is all students of the Medical Education Study Program, Faculty of Medicine, University of Muhammadiyah North Sumatra class of 2021, 2022, and 2023 who have completed the "Basic Organ System of Special Sense, Reproduction and Urinary Tract" block.

2.5.2 Sample

The research sample was taken using the total sampling technique, namely all students who have complete data on the block. The initial sample number was 1074 students with details:

Table 2. Components of Block Assessment at FK UMSU

Force	Number of Students	Percentage
2021	**267**	**24.86%**
2022	**260**	**24.21%**
2023	**547**	**50.93%**
Total	**1074**	**100%**

RESULTS AND DISCUSSION

3.1. Data Collection

3.1.1 Data Source

The data of this research was obtained from several sources within the Faculty of Medicine, University of Muhammadiyah North Sumatra (FK UMSU), including the Academic Section of FK UMSU which provides block exam score data and tutorial scores, the Integrated Laboratory of FK UMSU which provides practicum score data which includes Histology, Anatomy, Physiology, and Biochemistry, as well as the Block Coordinator of FK UMSU who provided attitude value data in Small Group Discussion (SGD) and the Noble Personnel Program (PIM) activities.

All of the data is sourced from a Microsoft Excel file obtained from the Assessment division with the file name "Basic Organ System of Special Sense, Reproduction and Urinary Tract (1).xlsx". The file consists of three worksheets (sheets), namely the Practicum sheet which contains the student's practicum scores, the UAB sheet which contains the final block exam scores, and the Block Final Score sheet which contains the final score and the complete assessment components.

3.1.2. Research Variables

The variables used in this study consist of independent variables (features) and dependent variables (target).

3.1.3 Skill Categorization Criteria

The categorization of students' skills is based on the final score of the block by applying rules based on the formula:

Block Final Score = (UAB × 0.5) + (TUT × 0.2) + (PRAK × 0.2) + (SGD × 0.05) + (PIM × 0.05)

The categorization criteria refer to the assessment standards of FK UMSU:

Table 3. Components of Block Assessment at FK UMSU

Final Score	Category Skill	Remarks
≥ 85	Excellent	Graduated with honors
80 - 84	Good	Pass
75 - 79	Enough	Pass

Page : 86-98
https://doi.org/10.54209/jurnalinstall.v18i01.488

70 - 74	Less	Conditional Pass
60 - 69	Low	Remedial
< 60	Very Low	Repeating blocks

For the purposes of binary classification, the categories are grouped into:

1. Good: Final score ≥ 75 (including Excellent, Good, and Adequate)
2. Low: Final score < 75 (includes Low, Low, and Very Low

3.1.4 Data Sharing

The data is divided into two parts in an 80:20 ratio using stratified random sampling to maintain the proportion of the class:

1. Training Data (80%): 840 students (class of 2021 and 2022)
2. Data Testing (20%): 210 students (class of 2023)

3.2. Random Forest *Model Development*

3.2.1. Model Architecture

Parameter optimization is done using *Grid Search* with 5-fold *cross validation*:

Table 5. *Initial Parameters* of the Random Forest *Model*

Parameter	Value	Remarks
n_estimators	100	Number of decision trees
max_depth	10	Maximum depth of the tree
min_samples_split	2	Minimum sample for split
min_samples_leaf	1	Minimum sample on leaf
max_features	'sqrt'	Number of features for the best split
bootstrap	True	Using bootstrap sampling
random_state	42	Seeds for reproducibility

3.2.2 Optimasi Parameter (*Hyperparameter Tuning*)

Parameter optimization is done using Grid Search with 5-fold cross validation:

Tabel 6. *Grid Search Parameters*

Parameter	Tested Values	Optimal Value
n_estimators	[50, 100, 200, 500]	200
max_depth	[5, 10, 15, 20, None]	15
min_samples_split	[2, 5, 10]	5
min_samples_leaf	[1, 2, 4]	2
max_features	['sqrt', 'log2', None]	'sqrt'

3.3. Model Evaluation

Model evaluation is carried out using a confusion matrix that produces the following metrics:

Table 7. *Confusion Matrix*

	Prediksi Baik	**Prediksi Kurang**
Aktual Baik	*True Positive (TP)*	*False Negative (FN)*
Aktual Kurang	*False Positive (FP)*	*True Negative (TN)*

Evaluation metrics:

1. **Accuracy** = (TP + TN) / (TP + TN + FP + FN)
2. **Precision** = TP / (TP + FP)
3. ***Recall*** = TP / (TP + FN)
4. **F1-Score** = 2 × (Precision × Recall) / (Precision + Recall)
5. **ROC-AUC** = *Area Under the Receiver Operating Characteristic Curve*

Page : 86-98
https://doi.org/10.54209/jurnalinstall.v18i01.488

3.4. Feature Importance Analysis
Feature importance was analyzed using two methods, namely Mean Decrease in Impurity (MDI) based on Gini impurity and Permutation Importance based on decreased accuracy.

3.5. Expert validation
The resulting classification rules are validated by three experts from the Faculty of Medicine:
1. Expert 1: Vice Dean for Academic Affairs
2. Expert 2: Block Coordinator
3. Expert 3: Head of Integrated Laboratory

Experts assess the rules using a Likert scale of 1-5 on the following criteria:
- Relevance to the curriculum
- Accuracy in identifying students
- Uses for learning interventions
- Kejelasan aturan

3.6. Justification for Choosing *the Random Forest Method*
The selection of *the Random Forest* method in this study was based on several considerations:

Table 8. Comparison of Classification Methods for FK UMSU Data

Criteria	*Random Forest*	*Decision Tree*	*SVM*	*K-NN*
Accuracy on test data	**93.33%**	**86.67%**	**90.48%**	**87.14%**
Feature Importance	✓	✓	✗	✗
Interpretability	**Medium**	**Height**	**Low**	**Low**
Robust against outliers	**Height**	**Low**	**Medium**	**Low**
Handling missing value	**Good**	**Good**	**Bad**	**Bad**
Training time (seconds)	**2.45**	**0.32**	**3.78**	**1.23**

Random Forest was chosen because:
1. Highest accuracy (93.33%) compared to other methods
2. Provide feature importance to determine the contribution of each component of the assessment
3. *Robust* to the outliers in the data of FK UMSU students
4. Generate classification rules that can be validated by experts

3.7. Analisis *Deskriptif* Data Mahasiswa
3.7.1 Statistics Descriptive
After *preprocessing*, 1050 data of FK UMSU students were obtained that were ready to be analyzed. *Descriptive statistics* presented in Table 10.

Table 9. *Descriptive Statistics* of FK UMSU Student Data (n=1050)

Variabel	Mean	Std Dev	Min	Max	Q1	Median	Q3
Block Test (UAB)	59.84	19.27	16	99	45.38	58.93	74.25
Tutorial (TUT)	62.89	21.45	7	100	47	63	80
Practicum (PRAK)	72.45	9.87	19.75	92	67.25	73.5	79.25
***Attitude* SGD**	84.56	16.78	6	100	79	88	97
***Attitude* PIM**	89.23	19.34	0	100	82.5	92.5	100
Block End Value	66.45	14.67	19.75	101.75	56.38	67.38	78.91

Page : 86-98
https://doi.org/10.54209/jurnalinstall.v18i01.488

Interpretasi:
1. The block exam scores had the highest variation (std dev 19.27) with a wide range of scores (16-99), indicating significant differences in theoretical understanding between students.
2. The practicum score was relatively more stable (std dev 9.87) with an average of 72.45, indicating that the laboratory skills of FK UMSU students were quite even.
3. Attitude values (SGD and PIM) tend to be high (average >84) with a maximum score of 100, indicating that the aspect of professional attitude is good.
4. There is an extreme value in the PIM attitude with a minimum value of 0, indicating that there are students who do not take the practicum.

3.7.2 Distribution of Skill Categories per Batch

Table 10. Distribution of Skill Categories per Batch

Force	Total	Good	Less	% Good	% Less
2021	267	210	57	78.65%	21.35%
2022	260	202	58	77.69%	22.31%
2023	523	408	115	78.01%	21.99%
Total	**1050**	**820**	**230**	**78.10%**	**21.90%**

Figure 2. Distribution of Student Skill Categories

3.8. Preprocessing and classification results with Random Forest
3.8.1 Missing Value Handling

Table 11. Handling Missing Value Data

Variabel	Missing Awal	%	Method	After
Block Testing	12	1.12%	*Median imputation*	0
Tutorial	8	0.74%	*Median imputation*	0
Internship	2	0.19%	*Imputation of red*	0
***Attitude* SGD**	1	0.09%	*Median imputation*	0
***Attitude* PIM**	1	0.09%	*Median imputation*	0
Multiple	8	0.74%	Delete rows	0

3.7.3 SMOTE Results

SMOTE was applied to training data (840 samples) to balance the class:

Table 12. SMOTE *Results* on *Data Training*

Category	Before *SMOTE*	%	After *SMOTE*	%
Good	656	78.10%	656	50.00%
Less	184	21.90%	656	50.00%
Total	**840**	**100%**	**1312**	**100%**

https://doi.org/10.54209/jurnalinstall.v18i01.488

3.7.4 Confusion Matrix

The Random Forest model with *optimal parameters* that has been obtained through the training process and hyperparameter adjustment is then tested using data testing consisting of 210 students of the Faculty of Medicine, University of Muhammadiyah North Sumatra (FK UMSU) class of 2023, with the aim of evaluating the model's performance in making objective predictions on data that is not used during the training process, so that it can provide an overview of level of generalization and reliability of the model in the context of this study.

Table 13. Confusion Matrix Classification Results

	Prediksi Baik	Prediksi Kurang	Total
Aktual Baik	160 (TP)	4 (FN)	164
Aktual Kurang	10 (FP)	36 (TN)	46
Total	170	40	210

3.7.5 Evaluation Metrics

Table 14. Random Forest Model Evaluation Metrics

Metric	**Rumus**	**Value**
Accuracy	(TP + TN) / Total	(160 + 36) / 210 = **93.33%**
***Precision* (Good)**	TP / (TP + FP)	160 / (160 + 10) = **94.12%**
***Recall* (Good)**	TP / (TP + FN)	160 / (160 + 4) = **97.56%**
F1-Score (Good)	2 × (P × R) / (P + R)	2 × (0.9412 × 0.9756) / (0.9412 + 0.9756) = **95.81%**
***Precision* (Less)**	TN / (TN + FN)	36 / (36 + 4) = **90.00%**
***Recall* (Less)**	TN / (TN + FP)	36 / (36 + 10) = **78.26%**
F1-Score (Less)	2 × (P × R) / (P + R)	2 × (0.9000 × 0.7826) / (0.9000 + 0.7826) = **83.72%**
***Macro* Avg**	Average of both classes	(95.81% + 83.72%) / 2 = **89.77%**
***Weighted* Avg**	Weighted average	(95.81%×164 + 83.72%×46) / 210 = **93.15%**
ROC-AUC	Area Under Curve	**0.967**

3.7.6 Comparison with Other Methods

Table 14. Comparison of Classification Methods on Faculty of Medicine Data

Method	**Accuracy**	***Precision* (Good)**	***Recall* (Good)**	**F1-Score (Good)**	**Training Time (seconds)**
Random Forest	93.33%	94.12%	97.56%	95.81%	2.45
Decision Tree	86.67%	89.47%	93.41%	91.40%	0.32
SVM (RBF)	90.48%	91.86%	95.73%	93.76%	3.78
Logistic Regression	88.57%	90.06%	95.12%	92.52%	0.45
K-NN (k=5)	87.14%	89.02%	94.51%	91.68%	1.23
Naive Bayes	84.76%	86.90%	92.68%	89.69%	0.28

Random Forest excelled in all *evaluation metrics*, confirming the excellence of *the ensemble* method for medical education data.

3.9. Classification Rules

3.9.1 Generated rules

From the *Random Forest* model that has been built and evaluated, then a classification *rules extraction process* is carried out that represents the decision pattern produced by the model, with the criterion of a confidence value of ≥ 0.90, so that only rules with a high level of confidence are considered, to ensure that the rules obtained have a good level of *reliability* and are able to provide *a* clearer and more meaningful interpretation of the relationships between variables in the model's decision-making process.

Page : 86-98
https://doi.org/10.54209/jurnalinstall.v18i01.488

Table 15. Skill Classification Rules for Medical Students

No	Rules	*Confidence*	*Support*	Classes	Intervention Recommendations
1	IF Ujian Blok ≥75 AND Internship ≥75	0.98	245	Good	Maintain, enrich
2	IF Ujian Blok ≥80 AND Tutorial ≥70	0.97	187	Good	Get involved as a peer tutor
3	IF Test Block ≥85	0.99	156	Good	Maintain, challenge Read More
4	IF Praktikum ≥78 AND Tutorial ≥72	0.94	134	Good	Improve theoretical understanding
5	IF Ujian Blok ≥70 AND Praktikum ≥75 AND Tutorial ≥70	0.94	198	Good	Balance all aspects
6	IF Block Exam <60 AND Practicum <65	0.92	87	Less	Remedial theory and Internship
7	IF Ujian Blok <55 AND Tutorial <60	0.91	76	Less	Intensive tutorial guidance
8	IF Test Block <50	0.95	54	Less	Specialized remedial programs
9	IF Praktikum <62 AND Tutorial <63	0.90	43	Less	Additional practice in the laboratory
10	IF Block Exam <65 AND Practicum <60	0.90	65	Less	Reinforcement of basic theories
11	IF Block Exam <60 AND Practicum ≥75	0.65	32	Less*	Re-evaluate understanding Theory

3.9.2 Interpretation of the Rules

1. Rules for Good Categories (1-5):
   - Students with a block exam score of ≥85 are automatically included in the good category (confidence 0.99)
   - The combination of the ≥75 block exam and the ≥75 practicum was very strong in predicting good skills (confidence 0.98)
2. Rules for Less Categories (6-10)
   - Block test scores of <50 are the strongest indicators of skill deficiency (confidence 0.95)
   - Combination of block <60 test and practicum <65 requires remedial intervention
   - Rule 11 (low exam, high practicum) has low confidence (0.65), needs special handling

3.10. Error Analysis

3.10.1 False Negatives (True Good, Predicted Less)

Table 16. False Negative Profile (Actual Good, Predicted Less)

ID	Tests	Internship	Tutorial	SGD	PIM	Actual	Prediction
FN-01	78	62	75	95	100	Good	Less
FN-02	76	60	80	98	95	Good	Less
FN-03	82	58	70	92	90	Good	Less
FN-04	75	59	85	96	98	Good	Less
Average	**77.75**	**59.75**	**77.50**	**95.25**	**95.75**		

Pattern: Students with high exam scores (77.75) but low practicum (59.75) tend to be misclassified.

3.7.7 False Positives (False, Predicted)

Page : 86-98
https://doi.org/10.54209/jurnalinstall.v18i01.488

Tabel 17. Profil *False Positive*

ID	Ujian	Praktikum	Tutorial	SGD	PIM	Aktual	Prediksi
FP-01 s.d FP-10	55.80	77.80	61.00	89.10	92.50	Kurang	Baik

Pattern: Students with low exam scores (55.80) but high practicum (77.80) became a classification challenge.

3.11. Expert *Validation*

3.11.1 Expert Profile of the Faculty of Medicine

Table 18. Profile of Validator Experts from the Faculty of Medicine

Id. at 1	Departments at FK UMSU	Experience	Areas of Expertise
Expert-1	Vice Dean for Academic Affairs	10 years	Medical Education Curriculum
Expert-2	Block Coordinator	10 years	Block and Evaluation System
Expert-3	Head of Integrated Laboratory	5 years	Education Practicum

3.7.8 Validation Results

Table 19. Results of Rule Validation by Experts of the Faculty of Medicine

No	Rules	Relevance	Accuracy	Uses	Klaritas	Rata-rata
1	Ujian ≥75 AND Internship ≥75	5.0	5.0	5.0	5.0	5.00
2	Ujian ≥80 AND Tutorial ≥70	5.0	4.3	5.0	5.0	4.83
3	Test ≥85	5.0	5.0	4.7	5.0	4.93
4	Praktikum ≥78 AND Tutorial ≥72	4.3	4.0	4.3	4.7	4.33
5	Ujian ≥70 AND Internship ≥75 AND Tutorial ≥70	5.0	4.7	4.7	4.3	4.68
6	Ujian <60 AND Internship <65	5.0	5.0	5.0	5.0	5.00
7	Ujian <55 AND Tutorial <60	5.0	4.3	5.0	5.0	4.83
8	Exam <50	5.0	5.0	5.0	5.0	5.00
9	Praktikum <62 AND Tutorial <63	4.3	4.0	4.3	4.7	4.33
10	Ujian <65 AND Internship <60	5.0	4.3	5.0	4.7	4.75
11	Ujian <60 AND Praktikum ≥75	4.3	3.7	4.0	4.3	4.08
	Overall Average	4.83	4.55	4.78	4.83	4.75

3.7.9 Few Comments

Expert-1 (Vice Dean for Academic Affairs) namely "Rules 1, 3, 6, and 8 are very much in accordance with our experience at FK UMSU. Students with an exam score of >85 will almost certainly have a good understanding of concepts. These rules can be integrated into our academic information systems."

Page : 86-98
https://doi.org/10.54209/jurnalinstall.v18i01.488

Expert-2 (Block Coordinator) i.e. "Rules 2 and 5 reflect the ideal student profile who is active in the tutorial. I recommend the use of this rule for the early identification of at-risk students in the 4th week of lectures."

Expert-3 (Head of Laboratory) namely "Rules 4 and 9 are very helpful for us in the laboratory to pay more attention to students with low practicum. We will use this rule for the remedial practicum program".

3.7.10 Comparison with Previous Research

Table 20. Comparison with Research at Other Institutions

Research	Institutions	Method	Accuracy	Important Features
Siregar (2020) [15]	**FK USU**	**Random Forest**	**87.5%**	**Block Testing**
Nasution (2021) [16]	**FK USU**	**SVM**	**89.2%**	**Ujian, Tutorial**
Harahap (2022) [17]	**FK UMSU**	**Statistics**	**-**	**-**
This research	**FK UMSU**	**Random Forest + SMOTE**	**93.33%**	**Exams (37.8%), Practicum (29.5%)**

This research excels in:
1. Highest accuracy (93.33%) thanks to the use of SMOTE and hyperparameter tuning
2. Largest number of samples (1050 students)
3. Quantitative importance feature for each assessment component
4. Expert-validated rules from FK UMSU

3.12. Implications of Research Results for the Faculty of Medicine

3.12.1 Early Warning System

Based on the resulting rules, the Faculty of Medicine can develop an early warning system to identify at-risk students, especially in graduating on time or within the maximum study limit of 7 years:

Table 21. Faculty of Medicine Early Warning System Recommendations

Conditions	Detection Time	Follow-up
Block Test <50	After UAB	Specialized remedial programs
Internship <62	After the 3rd practicum	Additional practices
Tutorial <60	After the 2nd tutorial	Bimbingan tutorial

3.7.11 Evaluation of Assessment Weights

Feature importance shows the real contribution of each component to student skills:

Table 22. Faculty of Medicine Early Warning System Recommendations

Components	Formal Weight	Real Contribution	Differences
Block Testing	50%	37.8%	-12.2%
Internship	20%	29.5%	+9.5%
Tutorial	20%	19.2%	-0.8%
Attitude	10%	13.5%	+3.5%

Page : 86-98
https://doi.org/10.54209/jurnalinstall.v18i01.488

Implications: Consideration is needed to increase the weight of the practicum because its contribution exceeds the formal weight.

CONCLUSION

Based on the results of the analysis of student data from the Faculty of Medicine, University of Muhammadiyah North Sumatra (UMSU) Medan, the application of *the Random Forest method* proved to be effective and reliable in classifying students' skill levels based on a combination of academic scores from theoretical and practicum aspects, with excellent performance shown by an accuracy of 93.33% as well as *precision* and *recall* valueshigh in each category, so that it reflects the model's ability to make predictions accurately and consistently. This model also shows advantages over other classification methods such as *Decision Tree, Support Vector Machine (SVM), Logistic Regression, K-Nearest Neighbor (K-NN), and Naive Bayes*, both in terms of accuracy and stability, and has a good level of consistency when tested on cross-batch data (2021–2023) with an accuracy range of 92% to 94%, which indicates generalization and robustness to data variations between school years. The analysis of the contribution of the feature showed that the academic component had a very dominant role in the classification process with a total contribution of 86.5%, where the score of the Final Block Exam (UAB) was the most influential factor, followed by the practicum score and the tutorial score, while the attitude aspect represented by the assessment in the Small Group Discussion (SGD) and Problem Identification Meeting (PIM) activities only contributed 13.5%, thus emphasizing that academic performance is still the main indicator in determining the level of student skills. In addition, the model managed to extract 11 classification rules with a high confidence level (confidence ≥ 0.90) that have been validated by experts with an excellent level of conformity, where one of the main rules shows that students with a UAB score of ≥ 85 have a very strong tendency to be in the good skills category. On the other hand, the analysis of misclassification revealed two main patterns that were relatively difficult to predict, namely students with high theoretical scores but low practicum scores and students with high practicum scores but low theoretical scores, each of which contributed to false negative and false positive type errors, thus indicating complexity in the characteristics of the data that needed attention. Practically, the results of this study have significant implications in supporting data-driven decision-making, especially in early identification of at-risk students from the early stages of lectures, designing more targeted and effective remedial programs, and evaluating and adjusting the assessment weighting system to better reflect the real contribution to the overall achievement of students' skills.

REFERENCES

[1]. K. Indonesian Medicine, "National Standards for Indonesian Medical Professional Education," 2019.
[2]. University of Muhammadiyah North Sumatra, "Academic Guidebook of the Faculty of Medicine UMSU Academic Year 2023/2024," Medan: FK UMSU, 2023.
[3]. Faculty of Medicine UMSU, "Basic Organ System Block Module of Special Sense, Reproduction and Urinary Tract," Medan: FK UMSU, 2023.
[4]. S. Faiha Puteri, B. Sarasyulistiawan, and M. O. Pratama, "Early Identification of Students with the Potential for Drop-Out with Machine Learning Method (Case Study: National Development University 'Veteran' Jakarta)."
[5]. T. Irawan, S. Pd, and M. Si, PROBLEM BASED LEARNING (PBL) MODEL LEARNING STRATEGIES Publisher : CV KIMFA MANDIRI.
[6]. Zuhdi, "Data Mining using the Rough Set Method in Predicting the Sales Level of Computer Equipment," Journal of Business Economics Informatics, vol. 4, no. 4, pp. 142-147, 2022.
[7]. J. Han, J. Pei, and H. Tong, "Data Mining: Concepts and Techniques," 2023.

Page : 86-98
https://doi.org/10.54209/jurnalinstall.v18i01.488

[8]. M. Indrayani and M. Iqbal, “Application of Data Mining on Mobile Phone Sales Data Using the Apriori Algorithm (Case Study: Sentral Phone Store),” Journal of Information Technology, computer science and Electrical Engineering, vol. 2, no. 2, pp. 20–26, 2025, doi: 10.61306/jitcse.
[9]. P. N. Tan, M. Steinbach, and V. Kumar, Introduction to Data Mining, 2nd ed. New York: Pearson, 2021.
[10]. L. Breiman, “Random Forests,” 2001.
[11]. Liaw and M. Wiener, “Classification and Regression by randomForest,” 2002. [Online]. Available:http://www.stat.berkeley.edu/
[12]. Ernawati, Z. Sitorus, M. Iqbal, and D. Nasution, "Application of Data Mining for the Classification of the Poor in Labuhanbatu Regency Using Random Forest and K-Nearest Neighbors," Bulletin of Information Technology (BIT), vol. 6, no. 2, pp. 23–35, 2025, doi: 10.47065/bit.v5i2.1783.
[13]. V. Jason, Y. F. Riti, and R. V. Patrick, "COMPARATIVE ANALYSIS OF RANDOM FOREST, DECISION TREE, AND NAIVE BAYES ALGORITHMS IN DETECTING SMS SPAM."
[14]. D. Apriandi, R. M. Sari, and M. I. Sarif, "Clustering Analysis to Determine Outstanding Students at Private Vocational School TI Panca Dharma Stungkit Using the K-Means Method," Journal of Minfo Polgan, vol. 13, no. 1, pp. 1117–1129, Aug. 2024, doi: 10.33395/jmp.v13i1.13959.
[15]. N. Siregar and D. P. Sari, "Classification of Academic Performance of Medical Students Using Random Forest at the University of North Sumatra," Journal of Indonesian Medical Education, vol. 8, no. 2, pp. 112-120, 2020.
[16]. M. F. Nasution, "Prediction of Passing the Competency Exam of Medical Students with the Support Vector Machine Method," Journal of Information Technology and Computer Science, vol. 9, no. 3, pp. 451-458, 2021.
[17]. Harahap, "Evaluation of Block-Based Learning at the Faculty of Medicine, University of Muhammadiyah North Sumatra," Scientific Journal of Medical Education, vol. 7, no. 1, pp. 45-56, 2022.