JURNAL EMACS
[Engineering, MAthematics and Computer Science] Vol.6 No.1 January 2024: 45-53
e-ISSN: 2686-2573
DOI: 10.21512/emacsjournal.v6i1.10938

Design of An Intelligent Tutoring System – Student Model: Predicting Learning Style

Nubli Hawari[1*], Tanty Oktavia[2]

[1,2] Information Systems Management Department,
BINUS Graduate Program – Master of Information Systems Management,
Bina Nusantara University,
Jakarta, Indonesia 11480
nubli.hawari@binus.edu, toktavia@binus.edu

*Correspondence: nubli.hawari@binus.edu

Abstract - *Education is very important for everyone, not only for acquiring knowledge but also for improving quality of life and well-being. An Intelligent Tutoring System (ITS) is a computer system that can provide personalized and adaptive learning assistance and support to students. This system is designed to offer effective guidance to students based on their individual abilities and learning styles. ITS utilizes artificial intelligence (AI) technology to understand students' abilities and provide guidance tailored to their needs. Recently, there have been methods to predict learning styles, such as through questionnaires on the EducationPlanner website, but these determinations are often too general. This study aimed to predict the learning styles used by specific students for specific subjects. Researchers conducted this study at XYZ University to determine the learning styles of certain students or groups. With this information, instructional materials and methods can be uniquely designed to cater to the needs of these groups. Based on the evaluation results, the study found that the Logistic Regression model was the best, with a precision of 0.5653 and a hamming loss value of 0.3468. This research demonstrates that information from six selected subjects (English, Religion, Civics, Arts, Physics, and Geography) can be used to determine students' learning styles.*

Keywords: *Intelligent Tutoring System; Student Module; Mechine Learning, Prediction*

I. INTRODUCTION

The Change is a necessity, the inevitability that is needed and desired is a change accompanied by progress. In the industrial era 4.0, industry players are connected and communicate with each other which will make decisions without human involvement. A number of industrial sectors are ready to enter the industrial era 4.0, including the education sector at universities. Discussing the education sector will be closely related to digitalization. Many learning and teaching activities have also been carried out through video conferencing platforms and learning management systems that can help communicate and discuss material provided by instructors interactively without being limited by distance.

Teaching and learning process in tertiary institutions cannot be separated from the role of the lecturer who will provide lecture material. In Indonesia, a person can teach if he meets the academic qualifications determined by the Indonesian government, one of which is in Ministry of Education decree 44 2015 concerning National Higher Education Standart Article 26, as follows (Kementerian Riset, Teknologi, 2015),shown in Table I:

Table I. Lecturer academic qualifications

Program	Lowest academic qualification (must be relevant to the study program)	Graduates may also be used (relevant to the study program)
Diploma I and II	Master Degree	Diploma III + work experience + equal to 6 KKNI (level equivalent to Bachelor's degree level, covering aspects of work ability, master of knowledge, managerial ability and attitudes and values)
Diploma III and IV	Master Degree	Proffesional certificate + equal to 8 KKNI (proffesional education graduates)
Bachelor Degree	Master Degree	Certification + equal to 8 KKNI (proffesional education graduates)
Proffesion	Master Degree + work experience more than two years	Proffesional certificate + equal to 8 KKNI (proffesional education graduates) + work experience more than two years
Master	Doctoral	Proffesional certificate + equal to 9 KKNI (specialist education graduates)
Specialist and Subspecialist	Doctoral + work experience more than two years	-
Doctoral	Doctoral	Proffesional certificate + equal to 9 KKNI (specialist education graduates)

From the data provided above, it can be seen that university lecturers in Indonesia have academic standards and workloads regulated by the government. However, in reality, many students still encounter difficulties in understanding the course materials, and one of the contributing factors is the less interactive or passive teaching methods (Gunawan Tambunsaribu & Yusniaty Galingging, 2021).

On the other hand, each student is unique and diverse. The objective of this research is to explore alternative methods by creating a predictive model in machine learning that can predict students' learning styles. This allows educators to understand the needs of students in determining learning strategies for each specific type of learner. This component is included in the student model which is one of the components in a larger system known as the Intelligent Tutoring System (ITS).

Several studies suggest the use of resources offered by artificial intelligence (AI) to provide adaptative learning environment based on student characteristics and enable high-level interactions (Sein Minn, 2022). ITS allows for the modeling of a student's knowledge level and utilizes teaching strategies to enhance or correct a student's knowledge. They are based on the development and implementation of artificial intelligence (AI) methods and techniques, and as a result, the content and presentation methods of topics can be tailored to individual student abilities, shown in Figure 1.

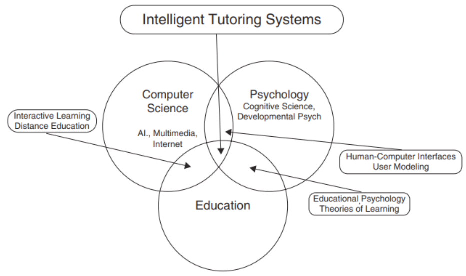

Figure 1. The ITS field is based on three disciplines: computer science, psychology and education (Woolf, 2009)

Figure 1 illustrates how ITS applies the theory of the teaching and learning process while incorporating various disciplines and educational domains to select effective teaching tactics for students. Psychology studies can be used to examine how students behave, understand how they learn, and determine the best ways to motivate them. Computer science is utilized to develop applications and determine the technology needed to assist students (Woolf, 2009). ITS is built based on four interconnected software modules, shown in Figure 2.

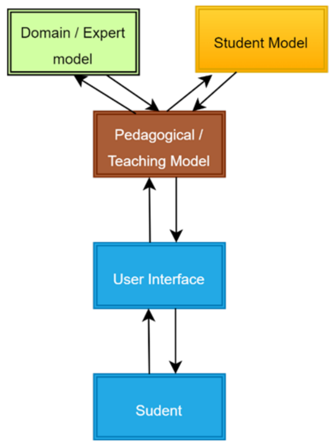

Figure 2. Intelligent Tutoring System Architecture Component Features

The primary goal of the Student Module is to ensure that the system possesses fundamental knowledge about each student so that it can respond appropriately, engage the student's interest, and enhance learning. The Student Module will generate teaching strategies, which can be seen in the Figure 3 below.

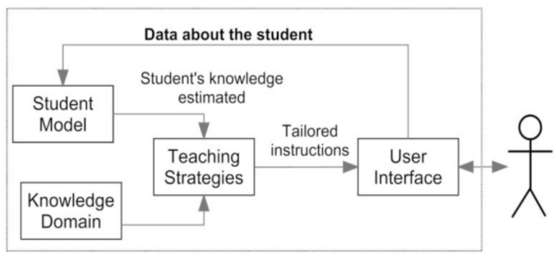

Figure 3. Basic architecture of an ITS (Morales-rodríguez et al., n.d.)

In the real classroom, teachers need to know understand the learning styles of their students to fit their teaching strategies (Cantina, 2021). ITS can adapt to learning styles use a formal questionnaire to collect information about learners (Yotta, 2023).

To achieve the highest level of understanding in learning outcomes, every learner possesses a unique learning style. Some textbooks described learning style as an individual style, the other half defined it as a choice or strategy. Texts for beginners studying education tended to take a more positive view of the use of learning styles, whereas texts for beginners studying educational psychology took a more neutral position. Matching instructional strategies to student learning styles was advised in 25% of the textbooks (Wininger et al., 2019). There are three main types of learning, namely visual, auditory, and tactile/kinesthetic, according to (Yaseen et al., 2021).

Machine Learning (ML) techniques in Intelligent Tutoring Systems (ITS) are used to achieve various goals, including: Building Predictive Models for Student Learning Styles: This assists educational institutions in understanding and integrating learning styles to maximize student learning outcomes (Marouf & Abu-naser, 2019).

Identify Effective Teaching Strategies: ML targets individual skills and assesses student knowledge (through practices or teaching actions - hints or explanations) to provide better teaching interventions (Woolf, 2009).

SMOTE works by performing oversampling on the minority class. The technique involves the following steps: Selecting a sample from the minority class. Choosing several nearest neighbors of the selected sample (k-nearest neighbors). Creating new synthetic examples by combining the selected sample with its nearest neighbors through line segments connecting them.

By following these steps, SMOTE increases the number of samples in the minority class by generating synthetic examples that represent the variations within that class.

Imputation in machine learning refers to the process of replacing or filling in missing values in a dataset with estimated values. Missing values can occur in a dataset for various reasons, such as data collection errors, faulty sensors, or the inherent unavailability of certain data. One type of imputation in the context of filling missing values in a dataset is Multivariate Imputation.

Multivariate Imputation refers to techniques that use information from multiple variables (features) to predict and fill in missing values in the variable being imputed. One of the techniques of multivariate imputation is Iterative Imputation, which involves an iterative approach where missing values are filled based on model predictions that take into account other variables (Fouad et al., 2021).

According to (Berrar, 2018), K-Fold Cross-validation is one of the most widely used methods of data sampling for assessing the generalization ability of predictive models. Its utility includes aiding in reducing bias from relying solely on a single data split and assisting in optimizing model parameters. In this method, the training data is divided into k subsets of the same size or nearly the same size (k-fold), where k is a positive integer, shown in Figure 4.

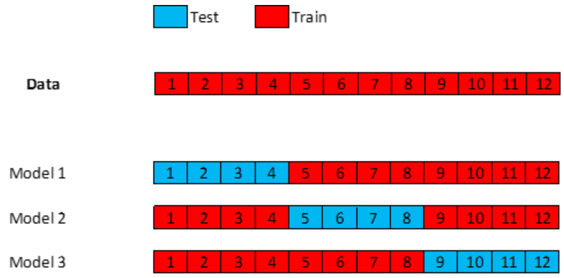

Figure 4. K-Fold Cross Validation ilustration

The data will be partitioned into k sub-sections or folds. For example, in Figure 6 using 3-fold cross-validation with 12 data points, there will be 3 iterations. In the first iteration, the first subset will use the first fold for testing, while the other folds will be used for training. In the second iteration, the second subset will use the second fold for testing, and the other folds for training. Similarly, in the third iteration, the third subset will use the third fold for testing, and the other folds for training. Although, lack of a formal framework for selecting the number of the neighborhood k is problematic (Melek et al., 2017).

A well-established and independent industrial process model for data mining tasks is CRISP-DM. It consists of six iterative phases, spanning from business understanding to the deployment process, shown in Figure 5.

Phase	Short description
Business Understanding	The business situation should be assessed to get an overview of the available and required resources. The determination of the data mining goal is one of the most important aspect in this phase. First the data mining type should be explained (e. g. classification) and the data mining success criteria (like precision). A compulsory project plan should be created.
Data understanding	Collecting data from data sources, exploring and describing it and checking the data quality are essential tasks in this phase. To make it more concrete, the user guide describe the data description task with using statistical analysis and determining attributes and their collations.
Data preparation	Data selection should be conducted by defining inclusion and exclusion criteria. Bad data quality can be handled by cleaning data. Dependent on the used model (defined in the first phase) derived attributes have to be constructed. For all these steps different methods are possible and are model dependent.
Modeling	The data modelling phase consists of selecting the modeling technique, building the test case and the model. All data mining techniques can be used. In general, the choice is depending on the business problem and the data. More important is, how to explain the choice. For building the model, specific parameters have to be set. For assessing the model it is appropriate to evaluate the model against evaluation criteria and select the best ones.
Evaluation	In the evaluation phase the results are checked against the defined business objectives. Therefore, the results have to be interpreted and further actions have to be defined. Another point is, that the process should be reviewed in general.
Deployment	The deployment phase is described generally in the user guide. It could be a final report or a software component. The user guide describes that the deployment phase consists of planning the deployment, monitoring and maintenance.

Figure 5. CRISP-DM Model Process (Schröer et al., 2021)

Logistic Regression is one of the most popular machine learning algorithms in supervised learning techniques. This algorithm is used to analyze the relationship between one or more independent variables (predictor factors) and a dependent variable that is binary (two categories) or nominal that can be treated as binary (for example, yes/no, success/failure, 1/0). Logistic Regression can be employed to classify observations using various types of data and can easily determine the most effective variables for classification. The Figure 6 below depicts the logistic function.

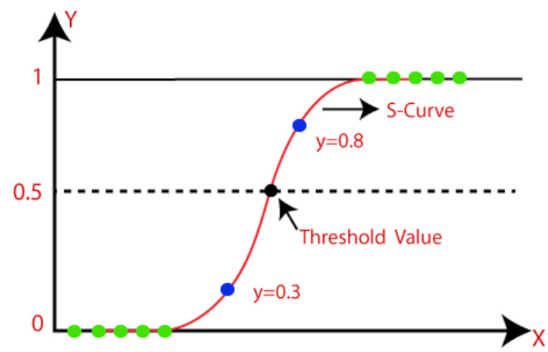

Figure 6. Logistic Regression Function
(Man Singh Basnet et al., 2023)

II. METHODS

The research is conducted by creating a machine learning model using a supervised learning algorithm. The method to design the machine learning model will follow the CRISP-DM approach, with the following Figure 7 stages:

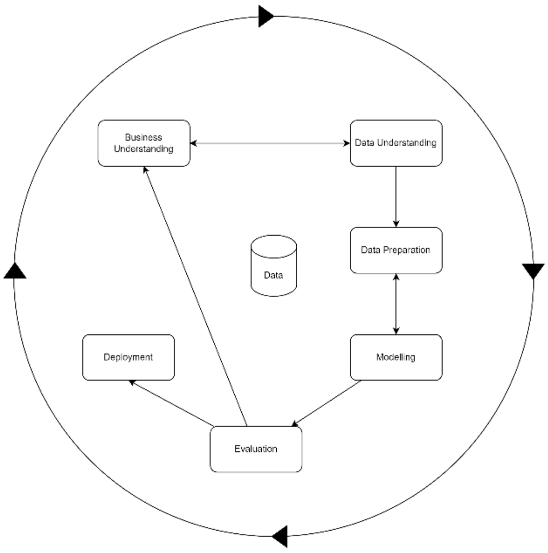

Figure 7. CRISP-DM Diagram

The stages of the research are carried out by analyzing the business needs that have been explained previously, analyzing the data requirements needed for making machine learning models, tidying up the data that has been collected, processing and preparing data, making machine learning models, evaluating the results of making machine models learning, and system design to implement machine learning models as student models in predicting student learning styles at XYZ University.

2.1 Business Understanding

The exploration process is conducted through interviews with several undergraduate lecturers at XYZ University. Currently does not have a mapping of students based on their learning styles. Referring to the Education Planner website, which provides self-assessment for determining learning styles, the complexity of implementation would be higher if the university wants to integrate this self-assessment application with the existing Learning Management System (LMS). Additionally, the input process would take longer as it requires answering 20 questions in the assessment form.

Therefore, predictive modeling using logistic regression in machine learning is needed to predict learning styles more efficiently. This is because the input data is derived from students' actual high school (SMA) grades and is more time-efficient since predictions can be made with only a few student grades inputted in a relatively short time.

2.2 Data Understanding

Data was obtained from the research subjects by administering a questionnaire using an online survey tool. The data sample consists of active student data for the year 2023 from XYZ University. The questionnaire data is divided into two parts:

- The first part consists of 20 questions sourced from the EducationPlanner website questionnaire at: *http://www.educationplanner.org/students/self-assessments/learning-styles-quiz.shtml#* to obtain the student labels, shown in Figure 8.

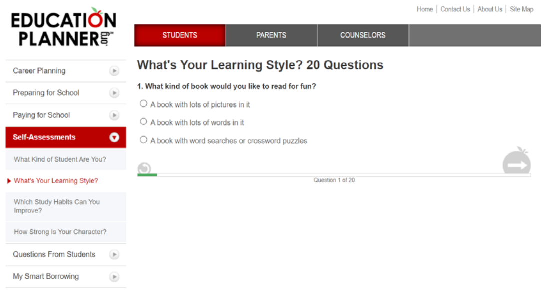

Figure 8. Selft assessment questionnaire

- Second part consists of 14 questions regarding the high school subject scores, including the grades for Indonesian Language, English, Mathematics, Physical Education, History, Civic Education, Arts, Religion, Physics, Chemistry, Biology, Sociology, Geography, and Economics, shown in Figure 9.

Visual	Auditory	Taktil	BahasaIndonesia	BahasaInggris	Matematika	Sejarah	Agama	PPKN	Kesenian	Olahraga	Fisika	Kimia	Biolo
1	0	0	80.0	85.0	88.0	80	80.0	86	0.0	80	83	88.0	
1	0	0	80.0	85.0	93.0	90	88.0	90	87.0	92	86	88.0	
1	0	1	85.0	90.0	85.0	85	80.0	85	80.0	90	80	80.0	
1	0	0	90.0	95.0	96.0	88	94.0	93	91.0	92	98	98.0	
1	0	0	85.0	87.0	81.0	80	92.0	82	79.0	82	78	75.0	

Figure 9. Data with label

If a student receives the label "Visual" from the self-assessment learning style, the researcher assigns 1 point to the Visual column, 0 points to the Auditory and Tactile columns. A student can receive one or two labels with different combinations. If a student receives both "Visual" and "Tactile" labels, the researcher assigns 1 point to both the Visual and Tactile columns and assigns 0 points to the Auditory column. In Figure 11, there are 3 columns for the label variable and 14 columns for the predictor variables.

2.3 Data Description

The results of the data collected have the Table II following attributes:

Table II. Data Attribute

Attributes	Description
Label	Describes label of each student (Visual, Auditory, Tactile)
Bahasa Indonesia	Describes student's score of Indonesia language
Bahasa Inggris	Describes student's score of English language
Matematika	Describes student's score of Math
Sejarah	Describes student's score of History
Agama	Describes student's score of Religion
Kewarganegaraan	Describes student's score of Civic education
Kesenian	Describes student's score of art and culture

Olahraga	Describes student's score of physical
Fisika	Describes student's score of physics
Kimia	Describes student's score of chemistry
Biologi	Describes student's score of biology
Geografi	Describes student's score of Geoghraphy
Ekonomi	Describes score of economics

2.4 Data Preparation

In this phase, data from non-active students will be excluded from the dataset as an exclusion criterion. Non-standard data quality can be addressed through data cleansing techniques such as removing rows of data that have a value of 0 for all subjects.

2.5 Modelling

The creation of the machine learning model is carried out using two methods: Logistic Regression and Random Forest, shown in Figure 10.

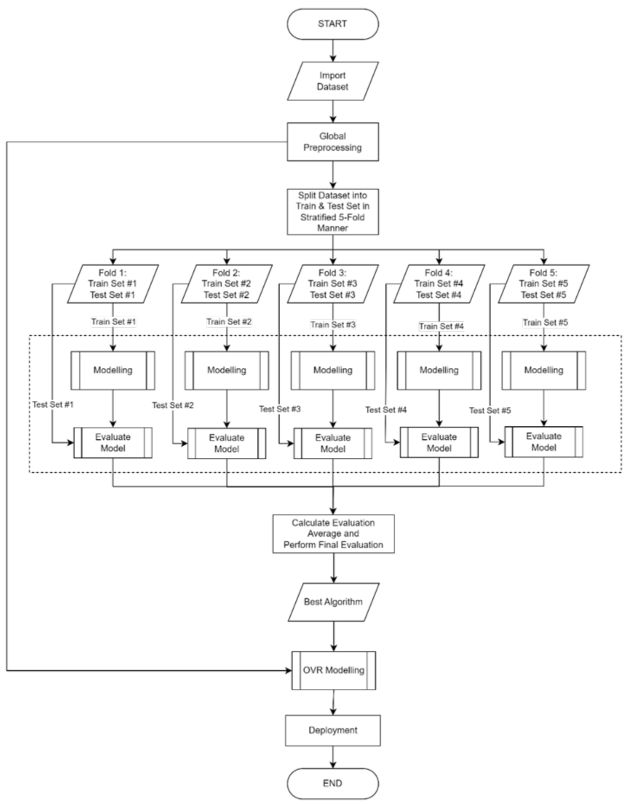

Figure 10. Model Process

Creation of a machine learning model involves various stages, namely:

- Import dataset.
- Perform global preprocessing.
- Split the data into training and testing sets using k-fold cross-validation; the average evaluation scores will be taken from 5-fold.
- Next, conduct a final evaluation to determine the best algorithm, whether it is logistic regression or random forest.
- The best algorithm will be used as the final model for deployment.
- The model selection involves 5-fold cross-validation, with each fold having its training and test sets. The training set is utilized for binary relevance modeling, and the resulting model will then be used to evaluate its corresponding test set.

2.6 Model Evaluation

After obtaining the final model from binary relevance, an evaluation will be conducted using the precision score and Hamming loss metrics.

2.7 Deployment

Once the machine learning model has been established, the model can be implemented to predict students' learning styles. The output of each prediction will be stored in a database and used for retraining the model so that it can continue to update if there are changes in the prediction patterns. This approach ensures that the model remains accurate and relevant as new data becomes available.

III. RESULTS AND DISCUSSION

One student can have more than one learning style, so the classification problem addressed in this research is a multi-label classification with three response variables, each having a cardinality of 2 (0 and 1). Two machine learning algorithms will be employed: Logistic Regression representing a linear model, and Random Forest representing an ensemble model. Hyperparameter tuning using Bayesian Search and 10-Fold cross-validation will also be utilized to select the best hyperparameters for each model. Furthermore, model selection will be carried out using 5-Fold cross-validation.

3.1 Logistic Regression

One of the folds from the outer cross-validation (CV) will be used for further exploration of the model. Here is a list of coefficients from the L2 Logistic Regression model built on Fold-5, shown in Table III.

Table III. Logistic Regression Coefficient

Variable	Coefficient		
	Visual	Auditory	Tactile
Bahasa Indoensia	-0.00359074	0.03200845	-0.01835429
Bahasa Inggris	0.04699108	-0.10693914	0.01180992
Matematika	-0.01308189	0.01624803	-0.02262123
Sejarah	-0.02060919	0.04205086	0.01392154
Agama	-0.03730958	0.05331494	-0.01146261
PPKN	0.01903321	-0.06193345	0.01411877
Kesenian	-0.00118276	0.03953388	-0.10429466
Olahraga	-0.0173282	0.00400155	0.00600421
Fisika	-0.03179912	0.01926379	0.0356225
Kimia	0.02496756	-0.00664707	-0.01192179
Biologi	0.02243378	-0.00772617	0.00796527
Geografi	-0.01697232	0.00105862	0.03170774
Sosiologi	0.00192061	-0.02199386	0.02205411
Ekonomi	0.03063554	-0.00661398	0.02404205

Each variable is already on the same scale, so the absolute values of the coefficients above can measure the impact of each variable on the prediction outcome. Even small changes in the predictor variable values with large absolute coefficients can have relatively significant effects on the response variable's outcome. Therefore, in this case, variables with large absolute coefficient values will be used as conditions for feature selection in the logistic regression model. Based on the logistic regression model built on Fold-5, the best features obtained are as follows:

- $BF_{1/Visual} = \{BahasaInggris, Agama, Fisika\}$
- $BF_{2/Auditory} = \{BahasaInggris, PPKN, Agama\}$
- $BF_{3/Taktil} = \{Kesenian, Fisika, Geografi\}$
- $BF_{ML} = BF_1 \cup BF_2 \cup BF_3 = \{BahasaInggris, Agama, PPKN, Kesenian, Fisika, Geografi\}$

Therefore, English, Religion, Civics, Arts, Physics, and Geography scores are the values that have the most significant influence on determining the learning style with the Logistic Regression model. Here are the best hyperparameters for the Logistic Regression model, shown in Table IV.

Table IV. LR - Best Hyper Parameter

Penalty	Features	Best Hyperparameters					
		Visual		Auditory		Tactile	
		C	solver	C	solver	C	solver
L1	All	36.917	saga	0.1486	liblinear	0.9437	saga
	Selected	12.4812	liblinear	4.8508	liblinear	15.8359	Liblinear
L2	All	22.1992	saga	8.3071	sag	23.7208	saga
	Selected	42.6639	liblinear	13.8436	sag	27.1906	liblinear

The best hyperparameters for the optimization algorithm 'solver' in the logistic regression model are all among 'liblinear,' 'sag,' or 'saga' with varying regularization strength 'C.' The L2 logistic regression model tends to prefer larger values of 'C,' while the L1 logistic regression model tends to prefer smaller values of 'C.'.

3.2 Random Forest

One of the folds from the outer cross-validation (CV) will be used for further exploration of the model. Here is the list of feature importance from the Random Forest model built on Fold-5, shown in Figure V.

Table V. Random Forest Coefficient

Variabel	Feature Importance		
	Visual	Auditory	Tactile
Bahasa Indonesia	0.01859165	0.07890281	0.05034051
Bahasa Inggris	**0.24879417**	**0.31391531**	0.05374463
Matematika	0.06860703	0.0185848	0.01429079
Sejarah	0.0612633	0.03313976	0.06222984
Agama	0.02431341	0.09349668	0.04438501
PPKN	0.09474635	0.09418422	0.03520156
Kesenian	0.02666827	0.01335426	**0.22743013**
Olahraga	0.08523781	0.07735166	0.07383624
Fisika	0.06479397	0.05881023	0.0052808
Kimia	0.0384599	0.05701213	0.03529789
Biologi	0.00450708	0.02477298	0.15571119
Geografi	0.10256678	0.04721726	0.11062312
Sosiologi	0.05977588	0.01645328	0.05710513
Ekonomi	0.1016744	0.07890281	0.07452317

Based on the feature importance of the trained random forest model, it can be observed that English and Arts scores are the most crucial features in distinguishing learning styles. When each feature is sorted based on its feature importance from high to low, a significant difference in values can be seen between the most important variables and others. This indicates the importance of English and Arts scores for this random forest model, shown Figure 11.

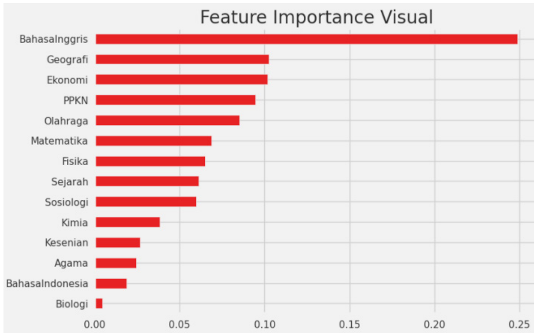

Figure 11. Feature Importance Visual

In the binary Visual model using random forest, the feature importance score for English is more than twice the feature importance score for Geography, which is the second most important variable, shown Figure 12.

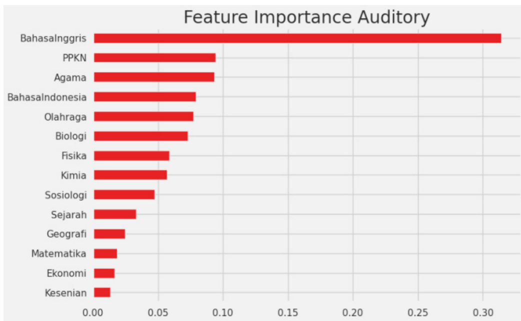

Figure 12. Feature Importance Auditory

In the binary Auditory model using random forest, the feature importance score for English is more than three times the feature importance score for Civic Education (PPKN), which is the second most important variable, shown Figure 13.

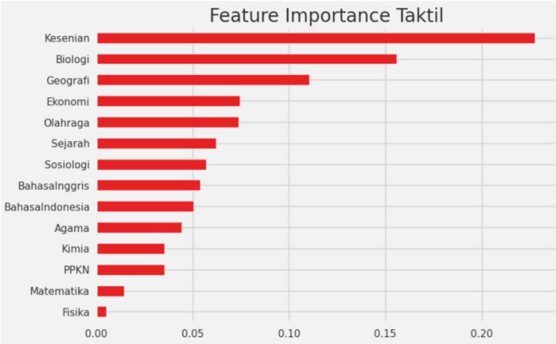

Figure 13 : Feature Importance Tactile

In the binary Tactile model using random forest, the feature importance score for English is only around 1.4 times the feature importance score for Biology, which is the second most important variable. Overall, it is indicated that English scores play the most crucial role in distinguishing Visual from Non-Visual learning styles and Auditory from Non-Auditory learning styles. Additionally, Arts scores play the most crucial role in distinguishing Tactile from Non-Tactile learning styles. The results obtained from these feature importance scores align well with the findings from the data exploration.

Therefore, in this case, variables with high feature importance values will be used as conditions for feature selection in the random forest model. Based on the random forest model constructed in Fold-5, the following features are identified as the best:

- $BF_{1/Visual}$ = {*BahasaInggris, Geografi, Ekonomi*}
- $BF_{2/Auditory}$ = {*BahasaInggris, PPKN, Agama*}
- $BF_{3/Taktil}$ = {*Kesenian, Boilogi, Geografi*}
- $BF_{ML} = BF_1 \cup BF_2 \cup BF_3$ = {*BahasaInggris, Agama, PPKN, Kesenian, Fisika, Biologi, Geografi, Ekonomi*}

Thus, the scores for English, Religious Studies (Agama), Civic Education (PPKN), Arts, Physics, Biology, Geography, and Economics are the values that have the most significant impact in determining learning styles using the Random Forest model.

Here are the optimal hyperparameters for the Random Forest model, shown Table VI.

Table VI. RF - Best Hyper Parameter

Features	Visual		Auditory		Tactile	
	All	Selected	All	Selected	All	Selected
Best Hyperparameters: Criterion	gini	gini	gini	entropy	entropy	gini
Max_depth	16	5	25	50	9	27
Max_features	13	1	14	14	12	14
Min_samples_leaf	9	1	11	11	5	2
Min_samples_split	5	11	8	11	2	11
N_estimators	16	11	21	10	10	11

The selected optimal hyperparameter values for each binary Random Forest model vary significantly. The Auditory model tends to build a relatively complex model, as seen from the relatively large 'max_depth' values both with and without feature selection.

3.3 Evaluation

The evaluation was performed using 5-fold cross-validation. In each fold, the multilabel model was evaluated using commonly used metrics for multilabel cases, namely hamming loss. In addition to hamming loss, precision score was also used because it is quite suitable for the case being studied. Therefore, precision score is more recommended than specificity for this case. Here are the evaluation results of the multilabel model with 5-fold CV using hamming loss metric.

Table VII. Hamming Loss Metric

Algorithm	Fold-1	Fold-2	Fold-3	Fold-4	Fold-5	Mean	Std
Logistic Regression (Penalty: L1)	0.345679	**0.333333**	**0.390805**	0.386667	**0.346154**	0.360527	0.023523
Logistic Regression (Penalty: L2)	**0.333333**	0.356322	0.402299	**0.360000**	**0.346154**	**0.359622**	0.023259
Random Forest	0.382716	**0.333333**	0.436782	0.400000	0.423077	0.395182	0.036090

From the perspective of hamming loss, Logistic Regression with L2 penalty outperforms Logistic Regression with L1 penalty and Random Forest, although the difference in loss between Logistic Regression L1 and L2 is not very significant.

However, when each model is trained with only the best features, the hamming loss values become lower, indicating that feature selection plays a role in improving performance. When all models are compared, both models built with feature selection and those without it, it is shown that the Logistic Regression L2 model with feature selection outperforms the others in terms of hamming loss, shown Table VIII.

Table VIII. Hamming Loss Metric (feature selection)

Algorithm	Fold-1	Fold-2	Fold-3	Fold-4	Fold-5	Mean	Std
Logistic Regression (Penalty: L1)	**0.345679**	0.379310	0.367816	0.360000	**0.307692**	0.352100	0.024751
Logistic Regression (Penalty: L2)	**0.345679**	0.390805	**0.356322**	**0.333333**	**0.307692**	**0.346766**	0.027346
Random Forest	0.370370	**0.367816**	0.413793	0.346667	**0.307692**	0.361268	0.034549

Here are the evaluation results of the multilabel model with 5-fold cross-validation using the precision score metric, shown Table IX.

Table IX. Precission Score Metric

Algorithm	Fold-1	Fold-2	Fold-3	Fold-4	Fold-5	Mean	Std
Logistic Regression (Penalty: L1)	0.555556	0.586207	0.448276	0.480000	**0.576923**	0.529392	0.055120
Logistic Regression (Penalty: L2)	0.555556	0.551724	**0.482759**	**0.540000**	**0.576923**	**0.541392**	0.031653
Random Forest	**0.586420**	**0.637931**	0.402299	0.500000	0.461538	0.517638	0.084809

From the perspective of precision score, Logistic Regression with L2 penalty outperforms Logistic Regression with L1 penalty and Random Forest. In addition to having the highest average precision score, Logistic Regression L2 also has the lowest standard deviation of precision score.

However, when each model is trained with only the best features, the precision score values become higher, indicating that feature selection plays a role in improving performance. When all models are compared, both models built with feature selection and those without it, it is shown that Logistic Regression L1 model with feature selection outperforms the others in terms of precision score. However, the difference in precision score between the models built with feature selection is relatively small, shown in Table X.

Table X. Precission Metric (feature selection)

Algorithm	Fold-1	Fold-2	Fold-3	Fold-4	Fold-5	Mean	Std
Logistic Regression (Penalty: L1)	0.555556	0.500000	**0.517241**	0.600000	0.653846	**0.565329**	0.056059
Logistic Regression (Penalty: L2)	**0.574074**	0.448276	0.500000	0.600000	0.653846	0.555239	0.072903
Random Forest	0.549383	**0.551724**	0.408046	**0.620000**	**0.692308**	0.564292	0.094111

Based on the cross-validation results with precision scores, Random Forest manages to outperform 3 out of 5 experiments, with the highest score reaching nearly 70% achieved on fold-5. Although Random Forest has a fairly impressive performance, its score variance is also the highest among the models. On fold-3, its precision score is the lowest compared to the other models. On the other hand, Logistic Regression has lower variance in performance with a higher average precision score compared to the other models.

IV. CONCLUSION

Machine learning modeling with feature selection can predict students' learning styles by inputting 6 subject grades (English, Religion, Civics, Arts, Physics, and Geography). This can facilitate educational institutions, allowing educators to design teaching strategies tailored to each student. Therefore, it can be concluded that the binary relevance model based on logistic regression with feature selection is the best model according to the results of K-Fold Cross Validation, with an average precision score of 0.5653 and an average hamming loss of 0.3468, outperforms the binary relevance model based on random forest, which only has an average precision score of 0.5643 and an average hamming loss of 0.3613 Researcher hopes that this study can be extended by obtaining primary data directly from educational institutions (high schools) and by employing other prediction models such as XGBoost and Multi-layer Neural Network. By utilizing such primary data and different prediction models, the author anticipates that the performance of the constructed models could be further optimized.

REFERENCE

Berrar, D. (2018). Cross-validation. *Encyclopedia of Bioinformatics and Computational Biology: ABC of Bioinformatics*, *1–3*(January 2018), 542–545. https://doi.org/10.1016/B978-0-12-809633-8.20349-X

Cantina, J. (2021). *Students ' Learning Styles : Basis for Module Development in Literature STUDENTS ' LEARNING STYLES : BASIS FOR MODULE DEVELOPMENT IN LITERATURE. December 2016*.

Fouad, K. M., Ismail, M. M., Azar, A. T., & Arafa, M. M. (2021). Advanced methods for missing values imputation based on similarity learning. *PeerJ Computer Science*, *7*, 1–38. https://doi.org/10.7717/PEERJ-CS.619

Gunawan Tambunsaribu, & Yusniaty Galingging. (2021). Masalah Yang Dihadapi Pelajar Bahasa Inggris Dalam Memahami Pelajaran Bahasa Inggris. *Dialektika: Jurnal Bahasa, Sastra Dan Budaya*, *8*(1), 30–41. https://doi.org/10.33541/dia.v8i1.3110

Kementerian Riset, Teknologi, dan P. (2015). Standar Nasional Pendidikan Tinggi (SN Dikti). *Produk Hukum*, *49*, 21–23.

Man Singh Basnet, P., Mahtab, S., & Jin, A. (2023). A comprehensive review of intelligent machine learning based predicting methods in long-term and short-term rock burst prediction. *Tunnelling and Underground Space Technology*, *142*(September), 105434. https://doi.org/10.1016/j.tust.2023.105434

Marouf, A. M., & Abu-naser, S. S. (2019). Intelligent Tutoring System for Teaching Computer Science I in Al-Azhar University , Gaza. *International Journal of Academic and Applied Research (IJAAR)*, *3*(3), 31–53. https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3368968

Melek, M., Karadeniz, T. K., & Melek, N. (2017). A novel simple method to select optimal k in k-nearest neighbor classifier. *International Journal of Computer Science and Information Security (IJCSIS)*, *15*(2), 464–469.

Morales-rodríguez, M. L., Ramírez-saldivar, J. A., Hernández-ramírez, A., Sánchez-solís, J. P., & Martínez-flores, J. A. (n.d.). *Architecture for an Intelligent Tutoring System that Considers Learning Styles*. *47*(2012), 37–47.

Schröer, C., Kruse, F., & Gómez, J. M. (2021). A systematic literature review on applying CRISP-DM process model. *Procedia Computer Science*, *181*(2019), 526–534. https://doi.org/10.1016/j.procs.2021.01.199

Sein Minn. (2022). AI-assisted knowledge assessment techniques for adaptive learning environments. *Computers and Education: Artificial Intelligence*, *3*(January), 100050. https://doi.org/10.1016/j.caeai.2022.100050

Wininger, S. R., Redifer, J. L., Norman, A. D., & Ryle, M. K. (2019). Prevalence of Learning Styles in Educational Psychology and Introduction to Education Textbooks: A Content Analysis. *Psychology Learning and Teaching*, *18*(3), 221–243. https://doi.org/10.1177/1475725719830301

Woolf, B. P. (2009). Building Intelligent Interactive Tutors. In *Building Intelligent Interactive Tutors*. https://doi.org/10.1016/B978-0-12-373594-2.X0001-9

Yaseen, S., Chohan, B. I., Komal, W., Hussain, M., & Jamal, R. (2021). A Study of Learning Style Preferences and Challenges Faced by the Students in Virtual Education. *International Journal of Innovation, Creativity and Change*, *15*(7), 46–68.

Yotta, E. G. (2023). Accommodating students' learning styles differences in English language classroom. *Heliyon*, *9*(6), e17497. https://doi.org/10.1016/j.heliyon.2023.e17497