Enhancing the Performance of Multiple Intelligence Learning Styles Prediction in e-Learning Systems Using Machine Learning Techniques

Gerzon J. Maulany [1,2], Paulus I. Santosa [1,*], and Indriana Hidayah [1]

[1]Department of Electrical Engineering and Information Technology, Faculty of Engineering, Universitas Gadjah Mada, Yogyakarta, Indonesia

[2]Department of Information System, Faculty of Engineering, Universitas Musamus, Merauke, Indonesia

Email: gerzonjokomenmaulany@mail.ugm.ac.id (G.J.M.); insap@ugm.ac.id (P.I.S.); indriana.h@ugm.ac.id (I.H.)

*Corresponding author

Manuscript received March 19, 2025; revised April 8, 2025; accepted May 13, 2025; published December 12, 2025

***Abstract*—Learning style detection is essential for developing adaptive learning tailored to learner preferences. Extensive research has been conducted to address the limitations of direct learner engagement, behavioral-based approaches, and the lack of learner-teacher interaction caused by the shift to online learning. Many learning style models have been widely used and evaluated for their performance in supporting adaptive learning. However, implementing Multiple Intelligence Learning Styles (MILS) remains limited in online learning due to low detection performance. This study employs machine learning approaches to enhance the detection performance of Multiple Intelligence Learning Styles. This study proposes an integrated machine learning framework to enhance MILS classification by combining data preprocessing, log-modulus transformation, and class imbalance handling via Synthetic Minority Over-Sampling Technique (SMOTE), Adaptive Synthetic Sampling Approach (ADASYN), and SMOTE-Tomek. Recursive Feature Elimination (RFE) with Support Vector Machine (SVM) is used for feature selection while eight machine learning models are used for classification. Interactions between learners and the Learning Management System (LMS) comprise the dataset, which are labeled using the Multiple Intelligence Inventory. The SVM model using ADASYN outperforms previous studies and other models, achieving the highest F1-Score of 89%, according to experimental data using 5-fold cross-validation. Statistical tests (Shapiro-Wilk, Analysis of variance (ANOVA), and paired t-tests) confirm significant differences in performance between models. The proposed approach demonstrates enhanced detection performance and can be adapted to broader e-learning applications, supporting adaptive and personalized learning systems.**

***Keywords*—personalized learning, e-learning, learning styles, machine learning, multiple intelligence**

I. Introduction

The rapid development of information and communication technology fosters the advancement of learning technology; online learning enables the implementation of tailored or flexible learning according to learners' needs and preferences [1]. A key element of personalized or adaptive learning is the learner model, and learning style is one factor that supports learner preference [2]. Learning style is essential in developing a personalized learning system and other aspects such as cognitive ability, prior knowledge, and learner's profile [3]. Learning style has been proven to enhance learning performance, learner satisfaction, learning motivation, while also reducing learning time [4–6]. Learning styles are a way for learners to engage in learning activities and processes. Learners' experiences and contexts influence learners' learning styles. These can change over time and across different learning situations [7]. Identifying the learner's learning style is crucial, as it allows for adapting the educational environment to suit each learner's unique traits and needs, producing more effective adaptive learning experiences for individuals [8].

Psychologists have developed numerous conceptual models of learning styles. These models, such as those by Kolb [9], Honey and Mumford [10], the modalities-based approach [11], Gregorc, Visual, Auditory, and Kinesthetic (VAK)/Visual, Auditory, Read/write, and Kinesthetic (VARK) Neil Fleming's learning style model [12], Felder-Silverman, and Gardner's theory of multiple intelligences, are widely recognized and used in e-learning environments [13].

Multiple Intelligences Theory (MIT) acknowledges that individuals learn in a variety of ways. Cognitive study suggests that learners have various mental abilities and process information in diverse ways. According to the theory, learners can have several strong intelligences. These diverse intelligences can help the learning process by providing a range of learning experiences. Additionally, intelligence is not fixed but can be enhanced through learning.

In contrast, intelligence may decline or fade if not exercised and reinforced through continued learning. MIT identifies nine intelligences: musical, naturalistic, existential, interpersonal, intrapersonal, verbal-linguistic, visual-spatial, bodily-kinesthetic, and mathematical-logical. These intelligences also correspond to distinct learning styles. Mathematical-Logical: learning through quantitative, logical, and mathematical methods while demonstrating excellent problem-solving, reasoning, and pattern-recognition abilities. Visual-Spatial: learning using visual aids like diagrams, photos, and images. Verbal-Linguistic: learning is best accomplished through oral or written means, such as word manipulation and expressiveness. Bodily-kinesthetic: learning is best done through hands-on activities, simulations, and direct experiences with concrete materials. Interpersonal: choosing to learn through group communication, collaborative thinking, and discussion-based approaches. Intrapersonal: exhibiting a preference for self-directed learning, focusing on concepts and theories. Musical-rhythmic: learning through musical elements such as patterns, rhythms, and listening to music. Naturalistic: learning by making connections between concepts and the environment, with a focus on practical, application-oriented

doi: 10.18178/ijiet.2025.15.12.2463

approaches. And Existential: learning by grasping the overall conceptual framework and big picture understanding, rather than focusing solely on details and specifics [14].

As one of the learning styles widely used in traditional learning and has demonstrated exemplary performance, studies have shown that more than 80% of evaluations on the role of multiple intelligences in supporting education found that groups utilizing multiple intelligences outperform those that do not [15]. Few studies have examined the use of Multiple Intelligence Learning Styles in online learning environments, in contrast to the extensive usage of these learning styles in traditional learning contexts, as noted in the literature review in Section II. According to the study's findings, the Felder-Silverman model, VARK, and Kolb's model are the three primary learning style models most frequently used in online education [16].

The primary obstacle to utilizing Multiple Intelligence Learning Styles (MILS) in online education is the subpar performance of learning type recognition models. In order to enhance the accuracy of learning style detection for different intelligences in online learning environments, this study employs machine learning approaches. This study used a rigorous research method to design the research, implementation, validation, and assessment to offer insightful information that will aid in developing personalized learning systems.

This study presents a novel innovation in the detection method for MILS compared to previous studies, through a combination of resampling techniques with feature selection—Recursive Feature Elimination (RFE) using linear kernel Support Vector Machine (SVM) to assess feature importance. In addition to enhancing the prediction performance of Multiple Intelligence Learning Styles, combining multiple resampling techniques with the RFE method contributes to a structured classification pipeline that uses statistical analysis to verify the accuracy of the model results.

The sections that follow describe the structure of this paper. Section II examines existing models and methods for automatically identifying learning styles, their methods, outcomes, validation procedures, and assessment metrics. Section III details the research methodology, including dataset creation, data preprocessing, resampling strategies, machine learning models, validation procedures, and model performance evaluation. Section IV discusses the resampling methodologies, the recursive feature elimination process, the resultant model performance, and the evaluation outcomes. Finally, Section V concludes the paper, with limitations and recommendations for future research.

II. Literature Review

Collecting and analyzing data from learner interactions in online learning environments is becoming a fundamental basis for utilizing Learning Style Model (LSM). The literature review highlights the application of diverse machine learning models for the detection of learning styles in e-learning, including K-Nearest Neighbor (KNN), Naïve Bayes (NB), Decision Tree (DT), Random Forest (RF), Neural Network (NN), Linear Discriminant Analysis (LDA), SVM, Logistic Regression (LogR), Linear Regression (LR), Fuzzy C-Means (FCM), K-Means, Classification and Regression Trees (CART), and Ensemble Learning (EL) techniques.

Kamal and Abo-Rizka [17] detected learning styles and knowledge levels using a literature-based approach that has shown promising detection performance, which can effectively support adaptive learning. Another study utilized a hybrid approach to identify learning styles and assign projects based on learners' preferences. This study used Honey and Mumford's learning style, and this identification process showed greater efficiency than traditional methods [18]. Another study used a combination of techniques to identify the learning styles of elementary school learners. The approach utilized learner interaction data with the LMS, collaborative filtering methods, and a modified matrix factorization algorithm to detect learning styles based on the VAK LSM [19]. The study of VARK learning style detection, which employed a prior knowledge approach and utilized ANN models, successfully enhanced the accuracy of learning style detection [20]. Furthermore, this study evaluated datasets with and without multimodal features, and the results showed that non-multimodal datasets achieved much higher accuracy than multimodal datasets [21]. Altamimi *et al.* [22] found that linear regression models outperformed classification models in predicting learners' learning styles with multiple dominant learning styles, and that the RF model achieved the best prediction performance. A study that used learner clickstream data to predict VAK-LSM in an online learning environment demonstrated how accurately the RF algorithm identified learners' VAK learning styles [23]. In order to identify VARK-LSM, Lokare *et al*. [24] used Electroencephalogram (EEG) data that included assessments of cognitive load, attention, facial expression, meditation, and emotional state. According to the study, the RF model outperformed the DT and EL models [24]. Costa *et al*. [8] utilized artificial intelligence to detect Kolb's-LSM based on learner interactions with the LMS. However, the researchers found low accuracy in classifying learning styles and identified one learning style that falls outside of Kolb's model [8]. Fadili *et al*. [25] used demographics, educational background, learning strategies, and social involvement to identify the learning styles of nursing learners in Morocco. According to the study's findings, the Neural Network model outperformed the Naive Bayes model in terms of prediction [25].

The Gravitational Search-Based Back Propagation Neural Network model used the labels produced by the FCM approach to predict the learning styles of incoming learners in later research. These studies used Web Usage Mining and the Felder-Silverman learning style paradigm to determine learners' styles [26]. A study examined how to identify learner learning styles through e-learning attributes determined from learner activities, such as time spent on learning materials, frequency of access, and navigation patterns. The Felder-Silverman learning styles of the pupils were predicted using machine learning models [27].

According to the findings, the SVM model showed the most reliable and consistent predicting capacity in this task, whereas the Naive Bayes model performed the least well [28].

The last two studies used multiple intelligence and

Felder-Silverman learning style detection approaches based on learner interaction data with the LMS and machine learning models for classification tasks. The first study, conducted by Rasheed and Wahid [29], used a Neural Network model to effectively recognize learner learning styles based on multiple intelligences. However, further enhancement in model convergence is required. The SVM performed the best for several intelligence learning styles (logical-mathematical, visual-spatial, verbal-linguistic, interpersonal-intrapersonal, and bodily-kinaesthetic) in the second study by Rasheed and Wahid [13], which included seven machine learning models. With an average of 86.30%, the SVM model performed best across all four dimensions for the Felder-Silverman learning style detection.

The current research is designed and implemented based on the research gap analysis compared to previous studies, as shown in Table 1. The focus is on enhancing the performance of Multiple Intelligence Learning Style detection using learner and LMS interaction data. Specifically, the five dominant intelligences detected from the Multiple Intelligence Learning Styles are logical-mathematical, visual-spatial, verbal-linguistic, interpersonal, and intrapersonal.

While previous studies have explored various learning style models and machine learning methods, few studies use feature selection and data resampling techniques separately to improve the classification of different intelligence learning styles. This study proposes integrating three types of resampling techniques with RFE and statistically evaluating the performance results of the best models.

Table 1. Comparison of previous learning style detection studies with the present study

Article	Machine learning model	Validation & evaluation method	Feature selection method	Resampling technique	Statistical test & model comparison	Learning style model
[17]	DT, KNN, NB	Ratio Time and Ratio Visit	-	-	-	FSLSM
[18]	NN	-	-	-	-	Honey and Mumford's LSM
[29]	NN	Accuracy	-	-	-	MILSM
[20]	NN	Accuracy	-	-	-	VARK LSM
[26]	FCM	Precision, Recall	-	-	-	FSLSM
[8]	NN	Accuracy	-	-	-	Kolb's LSM
[21]	DT, SVM, LDA, KNN	CV, F1-Score, ROC, AUC, Accuracy, Precision, and Recall	-	-	-	VARK LSM
[13]	KNN, NB, DT, RF, LDA, SVM, LogR	F1-Score, CV, Accuracy, Precision, and Recall	Feature Selection	-	-	MILSM and FSLSM
[19]	MF	CV, RMSE	-	-	t-test	VAK LSM
[28]	LogR, LDA, KNN, CART, NB, SVM	F1-Score, Accuracy, Precision, Recall, and CV	-	-	-	FSLSM
[22]	LR, SVM, NN, DT, RF, KNN	RMSE, MdAE, MAE, Accuracy	-	-	Wilcoxon	VARK LSM
[23]	KNN, RF, SVM, LogR	CV, F1-Score, Accuracy, Precision, Recall, and AUC	-	-	-	VAK LSM
[25]	SVM, DT, NN, NB	CV, F1-Score, ROC, AUC, Accuracy, Precision, and Recall	-	-	-	Kolb's LSM
[27]	SVM	Random Hold-out Method	RFE+SVM	-	-	FSLSM
[24]	DT, RF, EL	Accuracy, Precision, Recall, F1-Score	-	-	-	VARK LSM
Present work	KNN, NB, DT, RF, ANN, LDA, SVM, LogR, XGBoost	Accuracy, Precision, Recall, F1-Score	RFE+SVM	ADASYN, SMOTE, SMOTE Tomek link	Shapiro-Wilk, ANOVA, and t-test	MILSM

III. Materials and Methods

This paper attempts to improve the performance of MILS classification in online learning by building a dataset of online learning interaction activities using the LMS-Moodle platform. The process includes data preprocessing, covering selecting relevant features, labelling the dominant intelligence class based on the results of the Multiple Intelligence Inventory instrument, overcoming multi-class imbalanced, applying feature selection technique, implementing classification prediction using machine learning models, validating classification prediction models, and evaluating the results of prediction models. The suggested approach for identifying various intelligence learning styles in online courses is depicted in Fig. 1.

A. Apparatuses

This study employed equipment to support the achievement of its objectives, including a PC with an Intel Core i7 processor, 8 GB of RAM, and a 256 GB hard disk. This study utilized Python version 3.10.12 to implement data preprocessing, resampling techniques, feature selection, classification tasks, and statistical analyses. Multiple Intelligence Inventory (MII), developed by McKenzie based on MIT, was used to survey the dominant intelligence possessed by learners; MII was used to assign a dominant intelligence label value to the dataset.

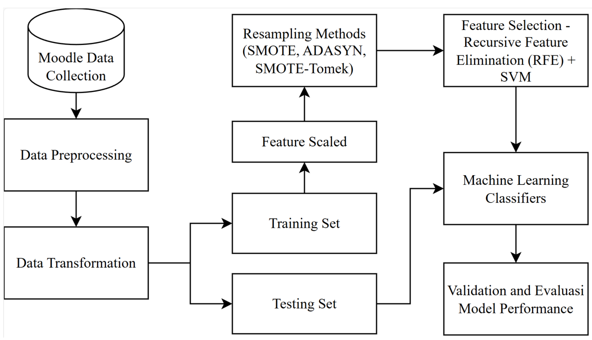

Fig. 1. Proposed methodology.

B. Dataset Information

The study involved 66 learners enrolled in teacher professional education program at X University. The data were collected from the learners' interactions with the Moodle-LMS platform to facilitate distance education. All participants completed a survey using a Multiple Intelligence Inventory instrument to assess their dominant intelligence. The study used learner and LMS interaction data based on a log file containing the following features: user_full_name, time, affected_user, event_name, component, description, origin, event_context, and IP_address. From the features in the log file, two key aspects were calculated to construct the features of the MILS dataset: the duration learners spent accessing learning objects or activities and the frequency or quantity of their access to learning objects or specific activities on the LMS. Table 2 presents the mapping of learning management system components and activities used to form multiple intelligence datasets and its attributes, supporting the identification dominant intelligence classes. The frequency of a user's access to various activities and components within the LMS can be calculated using formulas based on the available log file features, as in Eq. (1). These features include time (t), event_context (ec), user_full_name (u), event_name (en), and description (d).

$$F(u,ec,en)=\sum_{i=1}^{N}\delta(u_i,ec_i,en_i) \tag{1}$$

The duration (d_i) of an individual's engagement with specific activities or resources within the LMS was calculated as the difference in timestamp (t_i) between the current activity and the subsequent activity (t_i+1) in the sequence [30], as shown in Eq. (2). Based on the calculation concept adapted to the context of the dataset in the change study, the formula is applied as shown in Eq. (3).

$$d_i=t_{i+1}-t_i \tag{2}$$

$$D(u,ec,en)=\sum_{i=2}^{N(u,ec,en)}(T_i-T_{i-1}) \tag{3}$$

Table 2. Mapping of LMS activity components to support identification of dominant intelligence

Dominant intelligence	LMS activity component	Detailed LMS activity	Description	Attribute name
Logical-Mathematical	Assignment	Course module viewed, submission created, file upload, a submission has been submitted	Resolving problem-solving concepts and solutions (completing concept tasks)	Att01_T, Att01_N
	Forum	Course module viewed, discussion created, discussion viewed, post created	Interacting to develop innovative learning models (Developing learning products)	Att02_T, Att02_N
Verbal-Linguistic	Book/Material	Course module viewed, chapter viewed	Studying material deepening (Studying material in depth)	Att03_T, Att03_N
	URL	Course module viewed, course activity completion updated	Studying/listening to audio files of learning guidelines (learning materials from audio)	Att04_T, Att04_N
Visual Spatial	Folder/Material	Course module viewed, Zip archive of folder downloaded, course activity completion updated	Studying learning materials in PPT/PDF format	Att05_T, Att05_N
	Forum, Ou Blog, Questionnaire	Course module viewed, discussion created, discussion subscription created, discussion viewed, post created, comment created, participation viewed, post viewed, responses saved, responses submitted	Interacting and engaging in learning simulation activities (Engage in simulation activities)	Att06_T, Att06_N
	URL	Course activity completion updated, course module viewed	Accessing recorded learning videos (Recorded learning video files)	Att07_T, Att07_N
Interpersonal	Forum	Course module viewed, course activity completion updated, discussion created, discussion viewed	Engaging in collaboration and discussion to explore the causes of the problem (Collaboration to examine the causes of the problem)	Att08_T, Att08_N
	URL	Course activity completion updated, course module viewed	Studying, discussing results, and reinforcing problem identification (Studying discussion results and reinforcement)	Att09_T, Att09_N
Intrapersonal	Assignment	Course module viewed, submission created, a submission has been submitted, course activity completion updated, a file has been uploaded	Completing the action plan and reflecting on the assignment	Att10_T, Att10_N
	UO Blog	Course module viewed, course activity completion updated, post created, Post viewed	Conducting self-assessment	Att11_T, Att11_N

The outcomes of the dataset creation are shown in this section. The outcome used the LMS's learning components' activity mapping procedure connected to the dominant intelligence (Table 2). The dataset includes twenty-two main numeric attributes that support each dominant intelligence of the Multiple Intelligence Learning Styles. The attributes with the suffix "T" denote the time spent by the learner on a specific activity or learning object. In contrast, those with the suffix "N" indicate the frequency of the learner's activity on the learning object. The explanation of the numeric attributes supporting the dominant intelligence class/target in the multiple intelligence dataset (Table 2) are Att01_T: Time spent solving problem-solving concepts and solutions, Att01_N: Number of accesses to problem-solving concepts and solutions, Att02_T: Time spent interacting to develop problem-solving models and solutions, Att02_N: Number of accesses to interaction involving problem-solving models and solutions, Att03_T: Time spent studying material for deepening understanding, Att03_N: Number of accesses to deepening material, Att04_T: Time spent studying/reading the study guide and accessing audio files, Att04_N: Number of accesses to study guide audio files, Att05_T: Time spent studying educational resources in PPT and PDF formats, Att05_N: Number of accesses to educational resources in PPT and PDF formats, Att06_T: Time spent interacting and engaging in learning simulation activities, Att06_N: Number of accesses to learning simulation activities, Att07_T: Time spent studying and accessing learning recordings, Att07_N: Number of accesses to study and access learning recordings, Att08_T: Time spent engaging in collaboration and discussion exploring the causes of the problem, Att08_N: Number of accesses to collaborative discussions on problem causes, Att09_T: Time spent studying discussion results and reinforcement of learning, Att09_N: Number of accesses to discussions and reinforcement materials, Att10_T: Time devoted to finishing tasks, making plans for action, and reflecting, Att10_N: Number of accesses to assignment completion, action planning, and reflection activities, Att11_T: Time spent on self-assessment, Att11_N: Number of times doing self-assessment activities.

The identity attribute is not shown in Table 2 but remains part of the Multiple Intelligence Learning Styles dataset. The dominant intelligence class values were obtained through a survey using the MII instrument.

C. Imbalanced Dataset Problem

The issue of imbalanced datasets is frequently encountered in various real-world applications, especially as the volume and complexity of data grow. This situation is characterized by one or more classes with substantially greater instances or samples than the other classes. Due to this imbalance problem, the classification model may begin to favor the majority class and struggle to accurately predict the minority class, which could negatively impact its performance. Addressing the issue of imbalanced datasets is crucial in many real-world applications since it can lead to inaccurate forecasts and suboptimal decision-making [31]. This imbalance in class representation can negatively impact the classifier's overall performance, resulting in skewed and unreliable classification results [32]. There are various strategies to solve the dataset imbalance problem; with resampling techniques being among the most commonly used. There are three methods to implement resampling techniques: over-sampling, under-sampling, and hybrid-sampling (combining over-sampling and under-sampling). Three resampling strategies are used in this study: one hybrid methodology, SMOTE-Tomek [33], two over-sampling techniques, namely SMOTE [34] and ADASYN [35].

ADASYN is a resampling technique that balances the dataset by focusing on the creation of synthetic data in areas that need it the most or are the most difficult, which are those with the majority class [35].

The SMOTE-Tomek link hybrid resampling technique produces better synthetic data than the standard SMOTE technique because it not only oversamples the minority class using SMOTE but also addresses the issue of noisy or borderline samples. The Tomek-link method removes samples close to the decision boundary and does not have nearby samples of the same class. By combining these techniques, the resulting synthetic data is of higher quality and can positively impact the performance of the classification models.

However, when using these hybrid techniques, it is essential to carefully discard the samples considered noisy by Tomek-link to ensure that correct data is recovered. Additionally, since these are combination methods, the K-neighbor values for the SMOTE and Tomek-link components are self-determined, allowing the technique to adapt the optimal parameters for the specific dataset. Fig. 2 in the paper shows the results of combining different parameter settings for the ADASYN, SMOTE and SMOTE-Tomek techniques to identify the optimal resampling approach that supports the best model performance. Importantly, the classes on the testing set must remain uneven, and the resampling technique is only used on the training dataset.

The SMOTE, ADASYN, and SMOTE-Tomek resampling techniques are chosen because they are commonly recommended to reduce multiclass imbalance and are effective even on relatively small datasets, as they improve the representation of the minority class without requiring a large number of original samples.

D. Data Transformation and Feature Scaling

Data transformations improve data quality, increase model accuracy, adjust data formats, and enhance computational efficiency. There are various techniques and approaches to perform data transformations. One technique that can be used is Log Transformation and its variants (Log Modulus Transformation), which can handle negative values, maintain a relatively smaller scale compared to the original data, reduce data skewness, and can be applied to variables with an extensive range [36].

One technique for standardizing the range of independent variables or features is feature scaling, often known as data normalization. Since numerous machine learning models rely on the Euclidean distance between data points, feature scaling is essential to ensure optimal performance [37]. There are several ways to implement scaled features: Min Max Scaling, Z-Score Standardization, Robust Scaling, Unit Vector Scaling, Log Transformation, and Quantile Transformation. This paper uses the Z-Score Standardization technique to normalize the range of features in different multiple intelligence datasets representing the time and number of accesses used in performing activities. Z-Score Standardization, using a statistical approach to normalize data, technically subtracts the data mean from each value and divides the result by the standard deviation. Z-Score Standardization produces a distribution of data with a mean of zero and a variance of one. Z-Score Standardization is a common and highly beneficial normalization method as it preserves the original data distribution and is less susceptible to the influence of outliers.

The formula used to scale the feature's value is as Eq. (4): generating the average (mean) μ and σ is the standard deviation of the mean/average [38]:

$$Z = \frac{X - \mu}{\sigma} \tag{4}$$

E. Recursive Feature Elimination (RFE)

Recursive Feature Elimination (RFE) is a dimensionality reduction technique that involves removing features deemed unimportant by evaluating the performance of the resulting machine learning model. The steps of the RFE process are as follows:

1) Initialization
   a) Dataset: (X, y) with $X \in R^{nxp}$
   b) Model: $F(X, \theta)$
   c) Final feature count: P_{final}
2) Iteration t:
   a) Train model: $F^{(t)}(X;\theta^{(t)})$
   b) Calculate feature importance: $W^{(t)} = [w_1^{(t)}, w_2^{(t)}, \dots, w_p^{(t)}]$

c) Sort features Ranking: $r_j^{(t)} \alpha \left| w_j^{(t)} \right|$

d) Eliminate the k lowest-ranked features:

$$X^{(t+1)} \leftarrow X^{(t)} \setminus \{j_1, j_2, ..., j_k\}$$

3) Stop the iteration if: $\left| X^{(t_1)} \right| = P_{fianl}$

F. Machine Learning Model

A variety of machine learning methods can handle classification challenges. The following machine learning models are used in this study for multi-class classification tasks: KNN [39], NB [40], DT [41], RF [23], Artificial Neural Network (ANN) [42, 43], LDA, SVM [44], and LogR [45]. The selection of the eight machine learning classifiers is based on their ability to perform well on small to medium-sized datasets, where simpler models such as SVM and KNN can perform well without overfitting, the LogR model is able to work with multi-class imbalanced datasets, and more complex models such as RF and ANN are intended to assess the potential for performance improvement. To ensure optimal performance for each machine learning model, we use the Grid Search hyperparameter tuning strategy. Specific parameters and their ranges are systematically defined and explored. For example, for the SVM model, the search space includes kernel type (linear, rbf, poly), C value in the range (0.01, 0.1, 1, 10, 100), and gamma in (scale, auto). For RF, we vary the parameters n_estimators (100, 200, 300), max_depth (10, 20, None), and min_samples_split (2, 5). This parameter grid is evaluated through 5-fold Cross-Validation on each resampled training set to select the configuration with the best performance. The final hyperparameters selected are summarized in Table 3.

G. Model Validation

One method for assessing model performance is cross-validation, which divides the dataset into several subgroups. This technique guarantees that the model performs consistently on unseen data while reducing the chance of overfitting. This study splits the dataset into five sections using a 5-fold shuffle cross-validation technique. The model is then trained using four folds, with the final fold being utilized for testing. Each of the five looped iterations this validation uses serves as a test set. A thorough measure of model performance is the average of these five evaluation outcomes.

Table 3. Machine learning models with their best hyperparameter settings

Model	Hyperparameter settings		
	SMOTE	**ADASYN**	**SMOTE-Tomek**
LogR	C: 10, penalty: l1	C: 100, penalty: l1	C: 10, penalty: l1
SVM	C: 100, gamma: scale, kernel: poly	C: 10, gamma: auto, kernel: rbf	C: 10, gamma: scale, kernel: rbf
RF	min_samples_split: 5, max_depth: 10, n_estimators: 100	max_depth: 20, min_samples_split: 2, n_estimators: 300	min_samples_split: 5, max_depth: 10, n_estimators: 200
NB	No parameter Setting	No parameter Setting	No parameter setting
DT	Criterion: gini, max_depth: None, min_samples_split: 2	Criterion: entropy, max_depth: None, min_samples_split: 10	Criterion: gini, max_depth: None, min_samples_split: 2
LDA	Solver: svd	Solver: svd	Solver: svd
KNN	Metric: euclidean, n_neighbors: 3, weights: distance	Metric: euclidean, n_neighbors: 3, weights: distance	Metric: euclidean, n_neighbors: 3, weights: distance
ANN	Activation: tanh, alpha: 0.0001, hidden_layer_sizes: (50, 50), solver: adam	Activation: relu, alpha: 0.0001, hidden_layer_sizes: (100, 50), solver: adam	hidden_layer_sizes: (50,), activation: relu, alpha: 0.01

H. Evaluation Methods

Machine learning models' performance quality can be examined using the confusion matrix. The confusion matrix graphically depicts the model's accurate and inaccurate predictions for every class in a multi-class classification problem. The matrix values stand for True Positive (TP), False Positive (FP), False Negative (FN), and True Negative (TN). Classes correctly predicted as the class under consideration are represented by TP, classes incorrectly predicted as the class in question by FP, classes incorrectly predicted as an alternative class by FN, and classes correctly not predicted as the class under consideration by TN. The TN value is derived by subtracting the total number of cases for the given class from the overall number of cases. Moreover, these TP, FP, FN, and TN values are instrumental in computing model evaluation metrics, including Accuracy, Precision, Recall, and F1-Score. The computation of Accuracy, Precision, Recall, and F1-Score metrics for classification models is based on the subsequent formulas:

$$Accuracy = \frac{\sum_{i=1}^{n} TP_i + \sum_{i=1}^{n} TN_i}{Total\ number\ of\ instances} \tag{5}$$

where n is the number of classes, TP_i is True Positive for class i, and TN_i is True Negative for all classes other than class i.

Micro-averaging and macro-averaging are the primary methods for calculating Precision, Recall, and F1-Score for a multi-class classification model. Micro-averaging assigns equal weight to each instance by combining the contributions from all classes to determine the average measure. Additionally, it indicates that the metric is computed by considering the sum of TP, FP, and FN for all classes without making a distinction between the performance of each class. In contrast, the macro-averaging gets the average of the metric calculations for each class separately, thus giving equal weight to the performance of each class rather than each instance. The macro-averaging approach involves calculating each class's Precision, Recall, and F1-Score separately and then averaging those values to get the overall metric. Eq. 6–8 are used to calculate macro average Precision, Recall, and F1-Score.

$$Macro\ Precision = \frac{1}{N} \sum_{i=1}^{N} \left(\frac{TP_i}{TP_i + FP_i} \right) \tag{6}$$

$$Macro\ Recall = \frac{1}{N} \sum_{i=1}^{N} \frac{TP_i}{TP_i + FN_i} \tag{7}$$

$$Macro\ F1-Score = \frac{1}{N}\sum_{i=1}^{N}(2\times\frac{Precision_i \times Recall_i}{Precision_i + Recall_i}) \quad (8)$$

In the context of multi-class imbalanced datasets, it is inappropriate to use accuracy alone as it can be misleading since the majority class can dominate accuracy. Therefore, to give a more balanced view of model performance across classes, greater emphasis is placed on Precision, Recall, and particularly the Macro-Averaged F1-Score. Models trained on resampled datasets and arbitrarily altered class distributions can be evaluated well with the Macro F1-Score since it treats each class equally.

Evaluating and comparing the performance of classification models is crucial for obtaining reliable and accurate comparative results, as otherwise, it may lead to misleading analysis. Consequently, determining the best model or approach based on its capabilities is crucial. Addressing this evaluation analysis problem can be facilitated using statistical techniques or methods. However, to apply appropriate statistical techniques, it is necessary to assess whether the data meets the required statistical assumptions, particularly regarding the normality of the data distribution.

The normality of the resulting dataset was evaluated in this article using the Shapiro-Wilk test. This study selected the Shapiro-Wilk test because it is commonly used in many scientific domains to assess the normality of data and works well with small to medium sample sizes [46]. If the sample comes from normally distributed data, it is indicated by a p-value $>$ 0.05; otherwise, the data is not normally distributed.

The classification performance models for the resulting datasets were compared for the significance of difference for each pair of models using the ANOVA test [47, 48] and further post hoc tests using paired t-tests [49].

IV. Result and Discussion

This section presents and discusses the results of the steps and methods described in the previous section. The main objective of this study is to improve the performance of MILS prediction in an online learning environment using machine learning techniques. The steps discussed earlier provide the direction to achieve this research goal. The results of the implemented design are divided into three sections of experimental results: dataset and data preprocessing results, validation and evaluation of the performance in predicting Multiple Intelligence Learning Styles, and statistical comparison of model performance across datasets using resampling techniques on the training set. Then, a subset of the results is analyzed using the feature selection technique, RFE technique, in a combination with an SVM model with the linear kernel to assess feature importance.

A. Implementation Resampling Technique Data-Level Approach Result

We use Log-Modulus Transformation, which aims not to eliminate outliers because the dataset has a small sample and to keep the expected data distribution close to normal. The stratified split is used for data separation to preserve the proportionality of class distribution in the training and testing sets and to normalize the data, particularly in the training set. Three resampling techniques, namely SMOTE, ADASYN, and SMOTE-Tomek, are used to overcome the problem of multi-class imbalanced datasets. The results of the resampling technique settings are presented in Fig. 2.

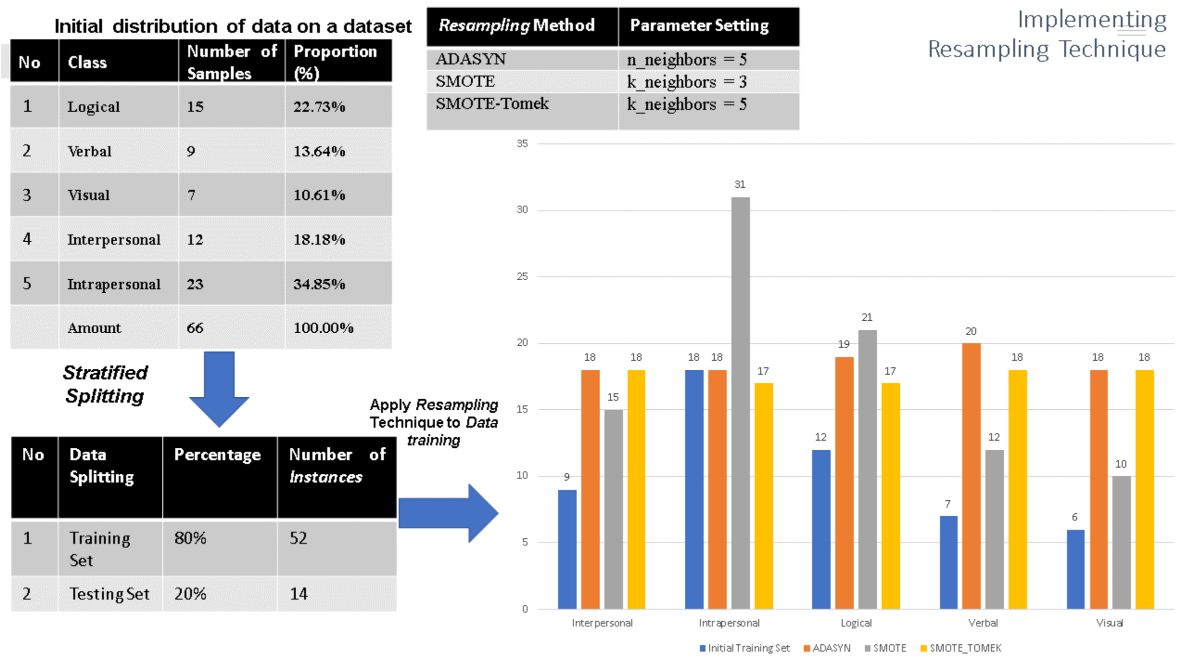

Fig. 2. Implementation of resampling technique on training set.

By assessing the RF model's F1-Score performance and adjusting each resampling technique's hyperparameters, resampling approaches are applied to the training set. The RF Model's F1-Score performance evaluation for the ADASYN resampling approach yields the highest results (61.67%), followed by SMOTE (74.18%) and Tomek's SMOTE (68.33%).

Based on the evaluation of parameter settings for resampling techniques and RF models, SMOTE obtains the best performance results with the parameter k_neighbors = 3, namely 74.18%. The analysis of the training set before and after the application of the resampling technique is that initially, the feature distribution is very diverse, and some show deep skewness; the SMOTE results still maintain the diversity of the feature distribution, but with more synthetic samples. The ADASYN results show a more diverse pattern by adding more instances in the minority class. Finally, SMOTE-Tomek produces the same pattern as SMOTE, with noise near different classes removed, produces more stable and cleaner results, while also reducing some outliers. However, the overall evaluation of the results of applying the three resampling techniques in the next section should be able to evaluate their effect on the performance of the resulting model.

B. Implementation of Recursive Feature Elimination Using SVM and Linear Kernel Result

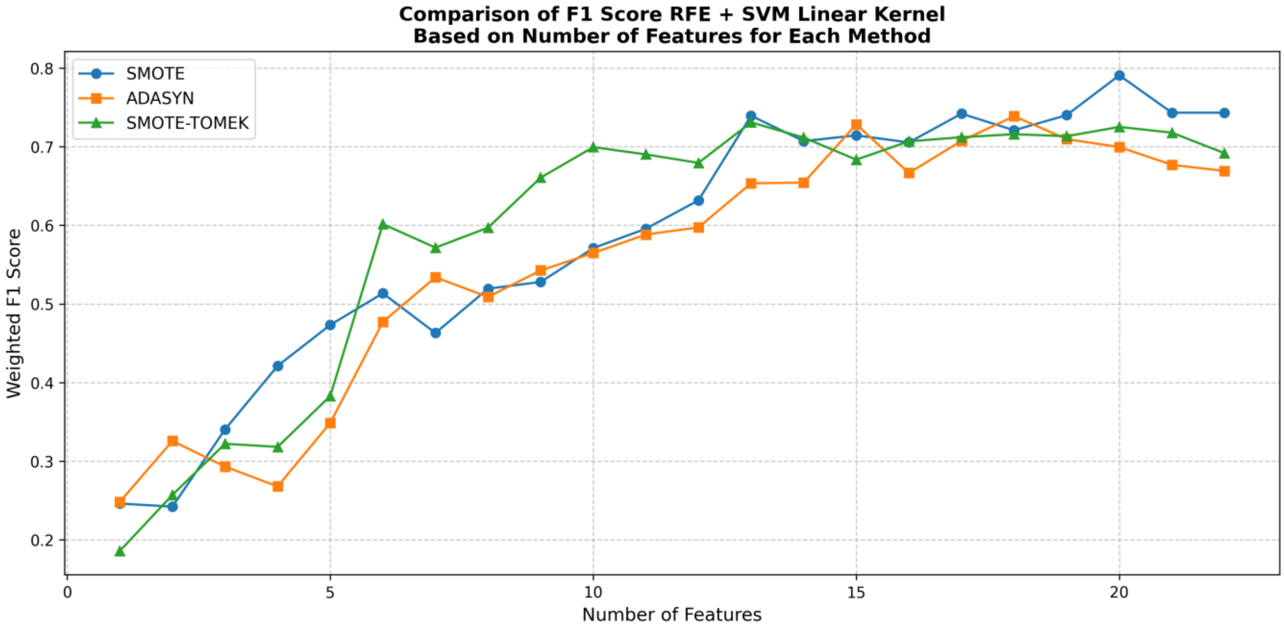

Fig. 3. Comparison of F1-Score metric value with number of features using RFE on three resampling techniques.

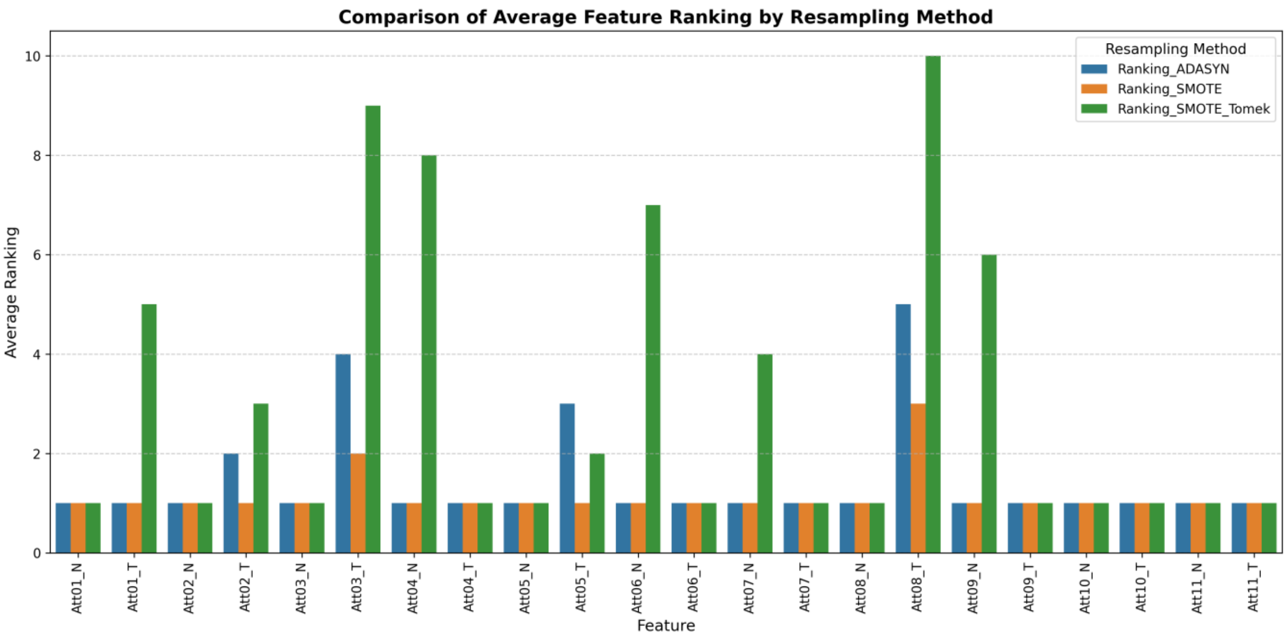

Fig. 4. Comparison of the average ranking of features based on resampling techniques.

Similar to the application of resampling techniques, Feature Selection Techniques are also applied only to the training set. In this case, it is done on the training set that has gone through the resampling process before. Applying the feature selection technique-wrapped method is helpful to produce a subset with features that strongly influence the classification results; the performance is indicated by the value of the F1-Score metric selected to evaluate the classification performance on multi-class imbalanced datasets. The process is performed iteratively by initially including all the features. Evaluating the classification performance value of the lowest contributing features is removed. This process continues until only high contributing features are left. Fig. 3 presents the results of applying RFE using the SVM model with Linear Kernel on three training sets using SMOTE, ADASYN, and SMOTE-Tomek resampling techniques. The results show that for the SMOTE technique, the highest F1-Score performance is 0.790828924 with the number of features 20, ADASYN achieves an F1-Score value of 0.738814815 with the number of features 18, and SMOTE-Tomek achieves the highest F1-Score value of 0.730925926 with the number of features used as many as 13. These results indicate that the number of features selected by the RFE technique varies according to the characteristics of the data and the resampling technique used; the trend of increasing F1-Score performance increases after the number of features increases (above 10). SMOTE shows a gradual increase, ADASYN is more progressive, and SMOTE-Tomek is more stable because it uses the SMOTE technique for oversampling and the Tomek Link technique for under-sampling. Fig. 4 compares the average ranking of features resulting from calculating feature importance RFE using SVM with a linear kernel across the three resampling techniques. A low-ranking value indicates a feature with a high contribution to classification performance, and conversely, a feature with a high ranking indicates that it contributes low to classification performance. SMOTE and ADASYN resampling techniques retain many features. In contrast, Tomek's SMOTE filters feature aggressively, especially in removing noise, but also open the possibility of eliminating helpful information.

Attributes such as Att01_T, Att04_N, Att06_N, Att06_T, Att07_N, Att07_T, Att08_N, Att09_N, Att09_T, Att10_N, Att10_T, and Att11_T rank highest in all three resampling techniques indicating that all resampling techniques consider all attributes to be necessary. Therefore, for the classification process, these features should be regarded as key features. On the other hand, for Att03_T and Att08_T, all three resampling techniques rank them as the features with the lowest contribution, so it can be considered to exclude these two features from the classification process. In contrast, the other features need further investigation to determine the priority contribution to the classification performance.

C. Evaluation of Multiple Intelligence Learning Style Classification Using Resampling and RFE Techniques

The evaluation performance results for detecting Multiple Intelligence Learning Style using eight machine learning models applied to features selected through RFE with an SVM linear kernel, and incorporating three resampling techniques, 5-fold cross-validation, and F1-Score metrics; are presented in Table 4.

Table 4 also compares classification performance results from previous studies on multiple intelligences. ADASYN Resampling technique shows the highest F1-score performance on SVM, ANN, and KNN models. This technique significantly improves the performance compared to the baseline (previous studies) on models sensitive to data distribution, such as ANN and SVM. However, it harms DT and NB models, possibly due to the noise from synthetic data generated by ADASYN. SMOTE improves the performance of DT, SVM, and KNN, showing success in overcoming imbalance with "cleaner" synthetic data compared to ADASYN. However, performance decreases for RF, LogR, and NB, indicating that SMOTE is inappropriate for highly sensitive models to overfitting. SMOTE-Tomek produces the best performance on LogR, LDA, and DT. However, it significantly degrades the performance of RF and ANN, indicating that Tomek link data cleaning is too aggressive, removing important data. It is suitable for linear and tree-pruning-friendly models (LogR, LDA) but is detrimental to complex models such as ANN and RF.

Table 4. Classification results of multiple intelligences using machine learning model using RFE+SVM with linear kernel

Model	F1-Score			
	Previous Study	ADASYN	SMOTE	SMOTE-Tomek
LogR	72%	71%	55%	75%
SVM	73%	89%	77%	87%
RF	73%	65%	59%	53%
NB	65%	42%	48%	42%
DT	43%	23%	71%	68%
LDA	69%	56%	57%	77%
KNN	67%	82%	78%	72%
ANN		83%	77%	46%

In addition, increasing the number of samples with greater size and more proportionate distribution is crucial, as the performance of classifiers is significantly impacted when the number of classes in a dataset rises. The SVM model is the best, as it shows the highest F1-Score among all resampling techniques, surpassing the performance of previous studies, followed by the KNN model, which can outperform previous studies. These results need to be reinforced with statistical test results so that models and resampling techniques can be optimally used to improve classification performance.

These findings highlight the trade-offs among the resampling techniques: SMOTE provides stable improvements but can risk overfitting through excessive synthetic instances; ADASYN adapts to difficult regions but may introduce noise and performance variability; SMOTE-Tomek balances oversampling and cleaning but may discard valid borderline data. RFE improves performance by focusing on features important in determining classification. Moreover, RFE can be sensitive to data characteristics and potentially eliminate highly correlated features. However, when assessed individually, it does not have a high level of importance. These conditions need to be considered when choosing a feature selection technique because they affect the robustness and reproducibility of the model.

While the classification results presented in this study were based on evaluation using the resampled and feature-selected datasets, it should be noted that the models were not assessed on the original 20% hold-out testing set from the initial stratified split. This omission occurred due to the primary focus on constructing a robust modelling pipeline through systematic preprocessing, resampling, feature selection, and model optimization.

However, a follow-up evaluation conducted using the original testing set—which consisted of only 14 samples and exhibited severe class imbalance—revealed a notable drop in performance across all models. This suggests that while the proposed pipeline is effective under balanced and engineered data conditions, its generalization to real-world imbalanced data remains a challenge and should be explored further in future research.

D. Statistical Test Result

Testing and determining the best model approach using resampling techniques to predict students' Multiple

Intelligence Learning Styles must be proven by statistical significance tests on differences in performance. We first assessed whether the F1-Score distribution from the 5-fold cross-validation results followed a normal distribution to statistically validate the performance evaluation across models and resampling techniques. We applied the Shapiro-Wilk test, confirming the normality of the data if the p-value > 0.05 (Table 5, p-value = 0.120602). The test results showed a normally distributed distribution, so we used the Repeated Measures ANOVA test to prove the significance of performance differences between the models. We ensured an inference-based approach based on the influence of the model and the resampling strategy to address the problem of imbalanced multi-class datasets. Table 5 presents the summary of the Shapiro-Wilk test results.

Table 5. Normality test results using the Shapiro-Wilk test for overall testing

Mean	Standard Deviation	Df	Number of Samples	w-statistic	p-value
0.59708	0.156621	119	120	0.982466	0.120602

Table 6. Repeated measures ANOVA test result

	Sum_sq	Df	F	PR(>F)
C(Model)	0.881226	7.0	6.758415	0.000002
C(Resampling_Technique)	0.055134	2.0	1.479951	0.232796
C(Model):C(Resampling_Technique)	0.194530	14.0	0.745956	0.723295
Residual	1.788199	96.0	NaN	NaN

To complement the evaluation metrics, statistical analysis was conducted. A brief discussion of Table 6 is provided in this section to contextualize the statistical test results. First, the C (Model) test, which shows the significance of the model on F1-Score performance, yielded a p-value of 0.000002, proving a significant difference in F1-Score performance between models. Second, the C (Resampling_Technique) test, produced a p-value of 0.232796, suggesting no significant difference in F1-Score performance across resampling techniques. Third, the C (Model): C (Resampling_Technique) test, which shows the interaction between the model and resampling technique, resulted in a p-value of 0.723295, suggesting no significant interaction between the machine learning models and the resampling techniques used. Finally, the residual error showed a Sum_sq of 1.788199 and Df of 96, indicating the presence of variation in the data that cannot be explained by the model and resampling techniques, which may be attributed to other factors such as preprocessing, hyperparameter tuning, or the characteristics of the dataset.

The distribution of classification performance for each model employing three distinct resampling methods is shown in Fig. 5 as the F1-Score metric. The findings demonstrate that the SVM model operates well on resampled data, with low variation and good performance, particularly when using the SMOTE and ADASYN resampling strategies. ANN has fairly high and relatively stable performance with SMOTE and ADASYN resampling techniques, but there are still some outliers with the SMOTE Tomek technique. Both SVM and ANN work well and are resilient to outliers. LogR also has a narrow distribution range, suggesting stable performance with all resampling techniques. Models with high variance are RF, DT, and NB. Overall, the SMOTE resampling technique produces stable performance for most models with fewer outliers. ADASYN tends to cause an increase in variance, especially for RF and DT models, due to the addition of more adaptive synthetic samples while SMOTE-Tomek tends to result in the broader distribution with some models experiencing outliers, indicating that this technique is more prone to noise in the data.

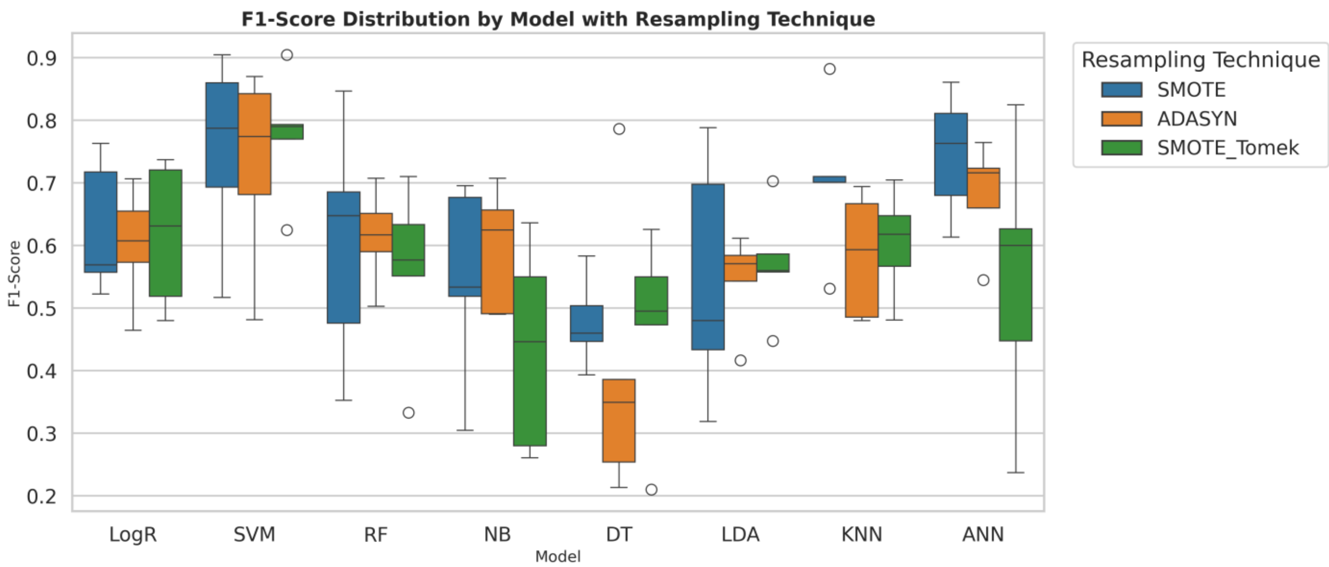

Fig. 5. Model-based F1-Score distribution visualisation with resampling technique.

The pairwise t-test results between the models are displayed in Fig. 6 to illustrate the relevance of the variation in F1-Score performance. A substantial difference in performance between the two compared models is indicated if the p-value is less than 0.05. In the graphic, significant results are represented by red dots, whilst non-significant results are indicated by blue dots. The significance value increases with the size of the dot.

It should be noted that implementing different resampling techniques will affect the evaluation metrics with different approaches. As an illustration, SMOTE can increase Recall but decrease Precision due to the synthetic oversampling process. In contrast, the ADASYN adaptive sampling approach can increase performance variability, as seen in the significant difference in F1-Score values. Meanwhile, the SMOTE-Tomek approach and a noise-cleaning procedure to stabilize performance can reduce Recall by eliminating adjacent instances from different classes. The results of these different approaches emphasize the importance of interpreting evaluation metrics carefully in the context of implementing different resampling scenarios.

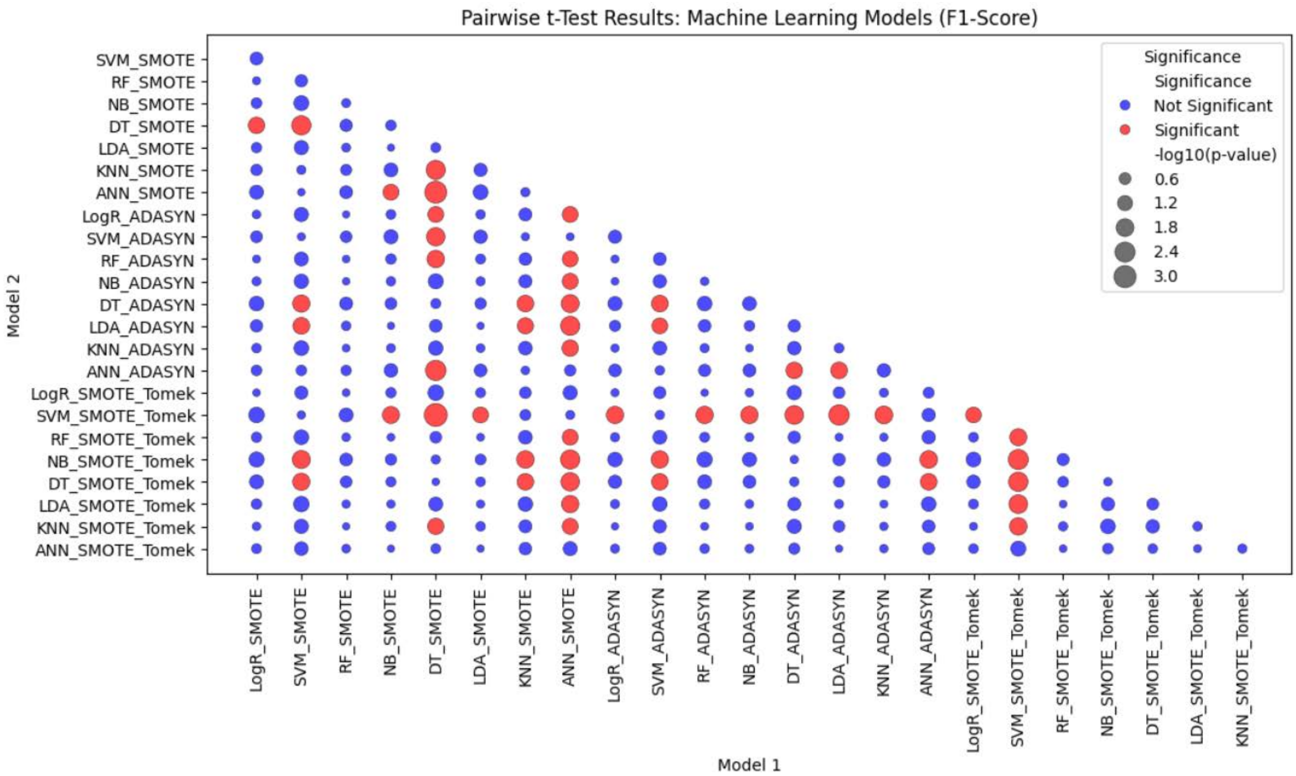

Fig. 6. Pairwise t-test results of performance between machine learning models.

E. Findings and Comparison with Existing Literature

The findings of this study show that high-performing models must address the challenges of multidimensional datasets [13]; overcoming the problem of imbalanced datasets with multiple classes [50] and using efficient feature selection methods that can enhance the output machine learning models' performance [27], especially in the context of detecting Multiple Intelligences Learning Styles. The SVM model consistently correlates with previous studies on detecting Multiple Intelligence Learning Styles [13]. This study has also successfully improved the performance of Multiple Intelligences Learning Styles detection compared to previous studies [2, 13]. A data transformation approach is required to bridge the different data scales in the features extracted from the learner data. Selecting the correct model requires strong characteristics, resilience to outliers, and the ability to work well on resampled data. In other words, models with good capabilities and characteristics will not be influenced by resampling techniques to demonstrate their performance. Even statistically, the resampling technique do not affect classification performance, the SMOTE resampling technique can be recommended to address imbalanced datasets as it shows stable results across all models. Finally, the model proposed in this study has contributed to improving the classification performance of Multiple Intelligences Learning Styles in online learning.

V. Conclusion

This study develops an integrated machine learning framework approach to improve MILS detection in online learning environments by integrating log-modulus data transformation, class imbalance handling using (SMOTE, ADASYN, and SMOTE-Tomek) resampling techniques, feature selection using RFE, and hyperparameter tuning using Grid Search. The proposed method greatly enhances the model's classification performance. Eight machine learning models were tested for performance, and the SVM classifier with ADASYN had the highest F1-Score (89%). Statistical validation using Shapiro-Wilk, Repeated Measures ANOVA, and paired t-test confirmed the significance of differences in model performance. These findings can contribute to adaptive learning systems requiring robust learner profiles based on limited or imbalanced data.

This study's shortcomings include the limited sample size, the multi-class imbalanced dataset, and the survey instrument still used to identify students' dominant intelligence labels despite notable performance improvements. Because of the class imbalance and comparatively limited sample size, the results may not be as representative or generalizable to a broader context. The results of resampling techniques can add unnatural variation to the data or reduce the original data useful for representing learner interactions. Therefore, this technique should be considered carefully. The potential bias in using instruments used to obtain the dominant intelligence labels perceived by students can result in subjective self-perceptions.

In addition, a limitation of this study is that the final model evaluation was conducted on data that had undergone resampling and feature selection, rather than on the original hold-out test set. This result may have led to optimistic performance estimations. Although the modelling pipeline was designed for robustness, its evaluation on imbalanced and limited real-world data needs further investigation. Future work should involve performance testing on a larger and more representative hold-out set that maintains natural class distribution, as well as employing repeated stratified cross-validation or nested validation techniques to strengthen the reliability of performance outcomes.

Beyond the specific application of Multi Intelligence Learning Style detection, the proposed methodology that combines resampling, feature selection, and machine learning classification can be generalized to the broader context of adaptive learning. The methodology can be used in learner behaviour modelling, recommending personalized learning materials, and improving engagement detection across various e-learning platforms, especially in the case of imbalanced or limited interaction data.

Conflict of Interest

The authors declare no conflict of interest.

Author Contributions

Gerzon J. Maulany: conceptualization, methodology, data gathering, and writing: preparing the initial draft. Paulus I. Santosa: writing, reviewing, editing, and supervising. Indriana Hidayah: monitoring, composing, editing, and reviewing. All authors have approved the final version of this paper.

Funding

The Indonesia Endowment Fund for Education (Lembaga Pengelola Dana Pendidikan/LPDP) provided funding for this study.

References

[1] E. Yilmaz İnce and M. İnce, "Automatic e-content sequencing system for personalised learning environments by using fuzzy AHP based on multiple intelligences," *J. Inf. Sci.*, vol. 47, no. 6, pp. 821–837, Dec. 2021. doi: 10.1177/01655515211041333

[2] C. Lwande, L. Muchemi, and R. Oboko, "Identifying learning styles and cognitive traits in a learning management system," *Heliyon*, vol. 7, no. 8, e07701, Aug. 2021. doi: 10.1016/j.heliyon.2021.e07701

[3] G. George and A. M. Lal, "Review of ontology-based recommender systems in e-learning," *Comput. Educ.*, vol. 142, 103642, Dec. 2019. doi: 10.1016/j.compedu.2019.103642

[4] M. Soflano, T. M. Connolly, and T. Hainey, "An application of adaptive games-based learning based on learning style to teach SQL," *Comput. Educ.*, vol. 86, pp. 192–211, Aug. 2015. doi: 10.1016/j.compedu.2015.03.015

[5] C. Troussas, F. Giannakas, C. Sgouropoulou, and I. Voyiatzis, "Collaborative activities recommendation based on students' collaborative learning styles using ANN and WSM," *Interact. Learn. Environ.*, vol. 31, no. 1, pp. 54–67, May 2020. doi: 10.1080/10494820.2020.1761835

[6] A. Gonzalez-Nucamendi, J. Noguez, L. Neri, V. Robledo-Rella, R. M. G. García-Castelán, and D. Escobar-Castillejos, "The prediction of academic performance using engineering student's profiles," *Comput. Electr. Eng.*, vol. 93, 107288, Jul. 2021. doi: 10.1016/j.compeleceng.2021.107288

[7] N. B. Afini Normadhi, L. Shuib, H. N. Nasir, A. Bimba, N. Idris, and V. Balakrishnan, "Identification of personal traits in adaptive learning environment: Systematic literature review," *Comput. Educ.*, vol. 130, pp. 168–190, Mar. 2019. doi: 10.1016/j.compedu.2018.11.005

[8] R. D. Costa, G. F. Souza, T. B. Castro, R. A. Valentim, and A. Dias, "Identification of learning styles in distance education through the interaction of the student with a learning management system," *IEEE Rev. Iberoam. Tecnol. Aprendiz.*, vol. 15, no. 3, pp. 148–160, Aug. 2020. doi: 10.1109/RITA.2020.3008131

[9] A.Y. Kolb and D. A. Kolb, "Learning styles and learning spaces: enhancing experiential learning in higher education," *Acad. Manag. Learn. Educ.*, vol. 4, no. 2, pp. 193–212, Jun. 2005. doi: https://doi.org/10.5465/amle.2005.17268566

[10] A. Mumford and P. Honey, "Questions and answers on learning styles questionnaire," *Ind. Commer. Train.*, vol. 24, no. 7, Jul. 1992. doi: 10.1108/00197859210015426

[11] M. J. Meyers, "The significance of learning modalities, modes of instruction, and verbal feedback for learning to recognize written words," *Learn. Disabil. Q.*, vol. 3, no. 3, pp. 62–69, Aug. 1980. doi: 10.2307/1510633

[12] H. R. Mozaffari, M. Janatolmakan, R. Sharifi *et al.*, "The relationship between the vark learning styles and academic achievement in dental students," *Adv. Med. Educ. Pract.*, vol. 11, pp. 15–19, Jan. 2020. doi: 10.2147/amep.s235002

[13] F. Rasheed and A. Wahid, "Learning style detection in e-learning systems using machine learning techniques," *Expert Syst. Appl.*, vol. 174, 114774, Jul. 2021. doi: 10.1016/j.eswa.2021.114774

[14] A. Cichocki and A. P. Kuleshov, "Future trends for human-ai collaboration: a comprehensive taxonomy of AI/AGI using multiple intelligences and learning styles," *Comput. Intell. Neurosci.*, vol. 2021, pp. 1–21, Feb. 2021. doi: 10.1155/2021/8893795

[15] M. Ferrero, M. A. Vadillo, and S. P. León, "A valid evaluation of the theory of multiple intelligences is not yet possible: Problems of methodological quality for intervention studies," *Intelligence*, vol. 88, 101566, Sep. 2021. doi: 10.1016/j.intell.2021.101566

[16] A. Wijaya, N. A. Setiawan, and M. I. Shapiai, "Mapping research themes and future directions in learning style detection research: A bibliometric and content analysis," *Electron. J. e-Learn.*, vol. 21, no. 4, pp. 274–285, Sep. 2023. doi: 10.34190/ejel.21.4.3097

[17] N. Kamal and M. Abo-Rizka, "Detecting students learning styles and knowledge level in e-learning environment," in *Proc. 2019 29th International Conference on Computer Theory and Applications (ICCTA)*, 2019, pp. 73–78. doi: 10.1109/ICCTA48790.2019.9478835

[18] L. Alfaro, E. Apaza, J. Luna-Urquizo, and C. Rivera, "Identification of learning styles and automatic assignment of projects in an adaptive e-learning environment using project based learning," *Int. J. Adv. Comput. Sci. Appl.*, vol. 10, no. 11, pp. 697–704, Jan. 2019. doi: 10.14569/IJACSA.2019.0101191

[19] B. Pardamean, T. Suparyanto, T. W. Cenggoro, D. Sudigyo, and A. Anugrahana, "AI-based learning style prediction in online learning for primary education," *IEEE Access*, vol. 10, pp. 35725–35735, Mar. 2022. doi: 10.1109/ACCESS.2022.3160177

[20] M. S. Hasibuan, L. E. Nugroho, and P. I. Santosa, "Model detecting learning styles with artificial neural network," *J. Technol. Sci. Educ.*, vol. 9, no. 1, p. 85, Feb. 2019. doi: 10.3926/jotse.540

[21] M. A. M. Sahagun, "Prediction of electronics engineering student's learning style using machine learning," in *Proc. 2021 1st Conference on Online Teaching for Mobile Education (OT4ME)*, Alcalá de Henares, Spain, 2021, pp. 42–49. doi: 10.1109/OT4ME53559.2021.9638868

[22] A. M. Altamimi, M. Azzeh, and M. Albashayreh, "Predicting students' learning styles using regression techniques," *Indones. J. Electr. Eng. Comput. Sci.*, vol. 25, no. 2, pp. 1177–1185, Feb. 2022. doi: 10.11591/ijeecs.v25.i2

[23] A. Rashad Sayed, M. Helmy Khafagy, M. Ali, and M. Hussien Mohamed, "Predict student learning styles and suitable assessment methods using click stream," *Egypt. Inform. J.*, vol. 26, 100469, Jun. 2024. doi: 10.1016/j.eij.2024.100469

[24] V. T. Lokare and P. M. Jadhav, "An AI-based learning style prediction model for personalized and effective learning," *Think. Ski. Creat.*, vol. 51, 101421, Mar. 2024. doi: 10.1016/j.tsc.2023.101421

[25] S. Fadili, M. Ertel, A. Mengad, and S. Amali, "Predicting optimal learning approaches for nursing students in Morocco," *Int. J. Adv. Comput. Sci. Appl.*, vol. 15, no. 4, Apr. 2024. doi: 10.14569/IJACSA.2024.0150410

[26] Y. E. Allioui, "Advanced prediction of learner's profile based on Felder-Silverman learning styles using web usage mining approach and fuzzy c-means algorithm," *Int. J. Comput. Aided Eng. Technol.*, vol. 11, no. 4–5, pp. 495–512, Mar. 2019. doi: 10.1504/IJCAET.2019.100447

[27] A. Lestari, A. Lawi, S. A. Thamrin, and N. Hidayat, "Automated detection of learning styles using online activities and model indicators," *Int. J. Adv. Comput. Sci. Appl.*, vol. 15, no. 6, Jan. 2024. doi: 10.14569/IJACSA.2024.0150661

[28] Z. Mehenaoui, Y. Lafifi, and L. Zemmouri, "Learning behavior analysis to identify learner's learning style based on machine learning techniques," *JUCS-J. Univers. Comput. Sci.*, vol. 28, no. 11, pp.

1193–1220, Nov. 2022. doi: 10.3897/jucs.81518
[29] F. Rasheed and A. Wahid, "Learning style recognition: a neural network approach," in *Proc. First Int. Conf. Artif. Intell. Cogn. Comput*, vol. 815, pp. 301–312, 2019. doi: 10.1007/978-981-13-1580-0_29
[30] D. Rotelli and A. Monreale, "Processing and understanding Moodle log data and their temporal dimension," *J. Learn. Anal.*, vol. 10, no. 2, pp. 126–141, Aug. 2023. doi: 10.18608/jla.2023.7867
[31] S. Goswami and A. K. Singh, "A literature survey on various aspect of class imbalance problem in data mining," *Multimed. Tools Appl.*, vol. 83, no. 27, pp. 70025–70050, Feb. 2024. doi: 10.1007/s11042-024-18244-6
[32] M. Kim and K. B. Hwang, "An empirical evaluation of sampling methods for the classification of imbalanced data," *PLOS ONE*, vol. 17, no. 7, e0271260, Jul. 2022. doi: 10.1371/journal.pone.0271260
[33] M. Khushi, K. Shaukat, T. M. Alam *et al.*, "A comparative performance analysis of data resampling methods on imbalance medical data," *IEEE Access*, vol. 9, pp. 109960–109975, Aug. 2021. doi: 10.1109/ACCESS.2021.3102399
[34] N. V. Chawla, K. W. Bowyer, L. O. Hall, and W. P. Kegelmeyer, "SMOTE: synthetic minority over-sampling technique," *J. Artif. Intell. Res.*, vol. 16, pp. 321–357, Jun. 2002. doi: 10.1613/jair.953
[35] R. O. Öztornaci, H. Syed, A. P. Morris, and B. Taşdelen, "The use of class imbalanced learning methods on ULSAM data to predict the case–control status in genome-wide association studies," *J. Big Data*, vol. 10, no. 1, p. 174, Nov. 2023. doi: 10.1186/s40537-023-00853-x
[36] J. A. John and N. R. Draper, "An alternative family of transformations," *Appl. Stat.*, vol. 29, no. 2, p. 190, 1980. doi: 10.2307/2986305
[37] S. Aksoy and R. M. Haralick, "Feature normalization and likelihood-based similarity measures for image retrieval," *Pattern Recognit. Lett.*, vol. 22, no. 5, pp. 563–582, Apr. 2001. doi: 10.1016/S0167-8655(00)00112-4
[38] K. K. Pal and K. S. Sudeep, "Preprocessing for image classification by convolutional neural networks," in *Proc. 2016 IEEE International Conference on Recent Trends in Electronics, Information & Communication Technology (RTEICT)*, 2016, pp. 1778–1781. doi: 10.1109/RTEICT.2016.7808140
[39] N. Marchang and R. Tripathi, "KNN-ST: Exploiting spatial-temporal correlation for missing data inference in environmental crowd sensing," *IEEE Sens. J.*, vol. 21, no. 3, pp. 3429–3436, Feb. 2021. doi: 10.1109/JSEN.2020.3024976
[40] C. K. Aridas, S. Karlos, V. G. Kanas, N. Fazakis, and S. B. Kotsiantis, "Uncertainty based under-sampling for learning naive bayes classifiers under imbalanced data sets," *IEEE Access*, vol. 8, pp. 2122–2133, Dec. 2020. doi: 10.1109/ACCESS.2019.2961784
[41] S. Shrivastava and R. Anju, "Spam mail detection through data mining techniques," in *Proc. 2017 International Conference on Intelligent Communication and Computational Techniques (ICCT)*, Jaipur, India, 2017, pp. 61–64. doi: 10.1109/INTELCCT.2017.8324021
[42] J. Bernard, E. Popescu, and S. Graf, "Improving online education through automatic learning style identification using a multi-step architecture with ant colony system and artificial neural networks," *Appl. Soft Comput.*, vol. 131, 109779, Dec. 2022. doi: 10.1016/j.asoc.2022.109779
[43] S. J. Lehmler, M. Saif-ur-Rehman, T. Glasmachers, and I. Iossifidis, "Understanding activation patterns in artificial neural networks by exploring stochastic processes: Discriminating generalization from memorization," *Neurocomputing*, vol. 610, 128473, Dec. 2024. doi: 10.1016/j.neucom.2024.128473
[44] P. I. Santosa and Rianto, "Modelling the classifier of online learning effectiveness using machine learning," *ICIC International*, vol. 16, no. 10, pp. 1079–1086, Oct. 2022. doi: 10.24507/icicel.16.10.1079
[45] G. Gayathri, M. F. Begam, and G. Aashika, "Machine learning based learning style detection using felder-silverman framework," in *Proc. 2023 International Conference on Innovative Computing, Intelligent Communication and Smart Electrical Systems (ICSES)*, Chennai, India, 2023, pp. 1–7. doi: 10.1109/ICSES60034.2023.10465313
[46] B. He, R. Tang, and Q. Tagn, "Identifying the best common factor model via exploratory eactor analysis," *Appl. Math.-J. Chin. Univ.*, vol. 39, no. 1, pp. 24–33, Mar. 2024. doi: 10.1007/s11766-024-3544-7
[47] K. K. Zhang, Y. M. Chen, Y. G. Li, S. Yao, Y. Su, and I. L. Wang, "Different drop heights in bilateral asymmetry and interjoint coordination during repetitive drop-jumps," *Symmetry*, vol. 13, no. 9, 1590, Aug. 2021. doi: 10.3390/sym13091590
[48] P. J. Senden, L. W. Sabelis, M. L. Zonderland, E. H. Hulzebos, E. Bol, and W. L. Mosterd, "The effect of physical training on workload, upper leg muscle function and muscle areas in patients with chronic heart failure," *Int. J. Cardiol.*, vol. 100, no. 2, pp. 293–300, Apr. 2005. doi: 10.1016/j.ijcard.2004.10.039
[49] W. T. Huang, C. C. Lu, and J. F. Dang, "Improving the return loading rate problem in northwest China based on the theory of constraints," *Mathematics*, vol. 9, no. 12, 1397, Jun. 2021 doi: 10.3390/math9121397
[50] R. Ghorbani and R. Ghousi, "Comparing different resampling methods in predicting students' performance using machine learning techniques," *IEEE Access*, vol. 8, pp. 67899–67911, Apr. 2020. doi: 10.1109/ACCESS.2020.2986809