Automated Detection of Learning Styles using Online Activities and Model Indicators

Alia Lestari[1], Armin Lawi[2], Sri Astuti Thamrin[3], Nurul Hidayat[4]
Department of Mathematics, Hasanuddin University, Makassar, Indonesia[1]
Department of Mathematical Education, Institut Agama Islam Negeri Palopo, Palopo, Indonesia[1]
Department of Information Systems, Hasanuddin University, Makassar, Indonesia[2]
Department of Statistics, Hasanuddin University, Makassar, Indonesia[3]
Department of Informatics, Jenderal Soedirman University, Purwokerto, Indonesia[4]

***Abstract*—Understanding learning styles is essential for learners and instructors to identify strengths and weaknesses in the education system. Although the Felder-Silverman Learning Style Model (FSLSM) is commonly used for this purpose, its reliance on in-person surveys can be time-consuming and prone to inaccuracies. This paper proposes an automated approach using Machine Learning (ML) to detect learning styles. This method extracts features from online activity data in Learning Management System (LMS) databases, aligning them with FSLSM indicators to label different learning styles. The dataset is divided into training and testing groups, respectively, to build and evaluate Support Vector Machine (SVM) classifiers. Feature selection is performed using the Recursive Feature Elimination (RFE) algorithm to improve the performance of the classifier, which results in the SVM-RFE algorithm. The experimental results showed promising accuracy for all model dimensions, i.e., 95.76% for processing, 85.88% for perception, 93.16% for input, and 96.42% for understanding dimensions. This approach offers a robust framework for automated learning style detection, which significantly reduces reliance on manual surveys and improves efficiency in educational settings.**

***Keywords*—Learning style; Felder-Silverman Learning Style Model; machine learning; support vector machine; recursive feature elimination; accuracy**

I. Introduction

The profound impact of the global pandemic caused by COVID-19, which lasted for nearly four years from 2020 to 2023, has substantially changed digital behavior in the education system. The study in [1] noted that nearly six billion students from 200 countries in the world were affected by this pandemic. The government's policy of closing schools and universities to prevent the spread of the virus has forced students to familiarize themselves with online learning, where e-learning is the most feasible solution to help schools and colleges facilitate student learning during the pandemic. Currently, most traditional learning has been juxtaposed with online learning, facilitated through learning management systems (LMS). Initially implemented as a way to limit the spread of the virus, government mandates requiring students to utilize online learning have been and are still being carried out in many places around the world. This underscores the effectiveness of the LMS in supporting student education both during and after the pandemic [2]. Unfortunately, the utilization of LMS features in the learning process remains limited, particularly in adapting to students' learning styles.

The LMS integration into modern educational frameworks can greatly enrich the learning experience for participants involved in e-learning. The LMS, as a web-based application, is specifically engineered to administer learning materials, facilitate learner interactions, deploy assessment tools, and generate reports on learner progress and activities [3]. Accessible online learning materials via the LMS empower learners to engage with educational resources seamlessly through web browsers across various operating systems, computers, or mobile devices. Moreover, apart from its core functionalities, LMSs encompass learning systems, classroom management systems, materials management systems, portals, and instructional management systems [4]. Furthermore, LMSs serve to facilitate learners' access to educational content through course guides, assignment submission, and retrieval mechanisms, interactive communication between learners and instructors, peer-to-peer collaboration, interaction with learning tools, knowledge sharing, as well as the administration of online assessments and quizzes [5].

It is imperative to acknowledge the unique preferences and requisites of each student within the learning milieu. Thus, conscientious consideration of individual learning styles becomes pivotal at every phase of the educational journey. Learning style denotes a consistent and habitual methodology in assimilating information, notably pertaining to cognitive processes such as cognition, retention, and problem-solving [6], [7], [8]. Students predominantly adhere to their distinct learning and information-processing modalities, thereby manifesting a myriad of learning style paradigms [9]. Theoretically, the absence of learner-centric support mechanisms within the educational framework may precipitate learner attrition throughout the learning trajectory. Consequently, learners are encouraged to discern and accommodate their learning styles to optimize learning efficacy [10].

Ongoing research endeavors in adaptive e-learning, which integrate considerations of learning styles, persist in grouping students according to specific learning style typologies. Nonetheless, critiques have been levied against learning style models by various scholars, such as [11] those who contend that many iterations of these models lack empirical validation [12], casting doubts on the validity and reliability of associated assessment tools. Notwithstanding, [13] assert the Index of Learning Style (ILS) as a valid and dependable instrument for gauging learning styles. ILS has been instrumental in automating the detection of learning styles [14], [15], [16], [17], [18]. Additionally, [7] ascertain a correlation between learning styles, learning strategies, and academic performance. Building upon these insights, [19] advocate for the development of a learning style framework to optimize the learning process's efficacy and outcome.

Over 70 learning style models have been proposed, exhibiting varying degrees of overlap and integration, with some models amalgamating or refining existing frameworks [20]. Predominantly employed within online learning systems are models such as Gregorc's Mind Styles Model (GMSM), Riding Cognitive Style (RCS), Myers-Briggs Type Indicator (MBTI), Kolb's Experiential Learning Theory (KELT), Honey and Mumford model, and the Felder-Silverman Learning Style Model (FSLSM) [21]. Notably, the FSLSM has emerged as a favored model for automated learning style detection [22], [23]. Its popularity stems from its comprehensive depiction of learning styles, coupled with established validity and reliability. Moreover, the FSLSM-based learning style assessment tool is straightforward, presenting respondents with only two opposing options [24]. In the context of the COVID-19 pandemic, literature suggests that the Felder-Silverman learning style model is particularly suitable for online learning environments [25]. The model's accessible and effective variables purportedly enhance students' learning abilities [26], [27].

Historically, identifying students' personal learning styles has relied on labor-intensive questionnaire analyses, especially burdensome in large-enrollment courses. Consequently, automated learning style modeling has garnered attention in both computational and educational spheres. Numerous studies have explored automatic learning style detection utilizing data from Learning Management Systems (LMS) and the FSLSM, employing various Machine Learning techniques. A comprehensive review by [24] covering machine learning approaches for automatic learning style detection from 1999 to 2011 concluded that the FSLSM model is most conducive to educational contexts. Among Machine Learning techniques, Neural Networks have exhibited the highest accuracy, according to [24]. However, recent research comparing automatic learning style detection techniques highlighted Naïve Bayes as the most accurate [18], [22]. Moreover, [18] achieved an 87% accuracy rate by modifying the Decision Tree algorithm to detect learning styles in 300 online course participants, while [28] employed Twin Support Vector Machine to classify MBTI learning style models. Furthermore, [29] developed models capable of simultaneously detecting learning styles and cognitive traits.

The Support Vector Machine (SVM) remains underutilized in the automatic detection of student learning styles, despite its capability as a linear model for both classification and regression problems, adept at addressing linear and non-linear complexities, and demonstrating efficacy in practical scenarios. SVM operates by identifying a hyperplane that effectively separates two sets of data belonging to distinct classes, with its efficiency further bolstered by the utilization of support vectors to expedite computation [30]. Moreover, SVM exhibits versatility in modeling non-linear data structures [31]. Comparative analyses of SVM against alternative machine learning methods for automated learning style detection consistently underscore its advantages [32], [33]. Nonetheless, the alignment of data availability within Learning Management Systems (LMS) with indicators specified by learning style models often presents a challenge. Additionally, past research predominantly focused on individual dimensions of the Felder-Silverman Learning Style Model (FSLSM) in isolation, neglecting potential interrelations among features affecting multiple dimensions concurrently.

This paper advocates for the automatic detection of student learning styles through a machine learning framework, leveraging features extracted from the mapping of online activities within LMS databases onto FSLSM indicators. This approach introduces three novel contributions. Firstly, a feature identification methodology for classifier models is introduced. This encompasses the direct extraction of original features from database attributes, alongside the derivation of synthetic features through the aggregation or accumulation of multiple attributes corresponding to FSLSM learning style indicators within the classifier model. Secondly, the mapping of identified features onto classes (learning styles) for each FSLSM dimension is proposed, followed by the labeling of each learning style. Subsequently, the identified and mapped feature dataset is partitioned into training data, utilized for model construction, and test data, employed to evaluate the performance of the resultant classifier. Initially, the SVM classifier model is adopted, with feature selection facilitated by the Recursive Feature Elimination (RFE) algorithm to enhance classifier efficacy. The ensuing SVM-RFE algorithm operationalizes these two stages, constituting the third contribution, culminating in the generation and validation of a high-performance classifier model.

II. Material and Methods

A. Data Source

The primary data source originates from the Learning Management System (LMS) Online Learning System (SPADA LMS), a program under the purview of the Directorate General of Higher Education (DIKTI) within the Ministry of Education and Culture. The overarching goal of SPADA LMS is to enhance equitable access to quality higher education. SPADA LMS is accessible via the following link: (https://lmsspada.kemdikbud.go.id/). It is noteworthy that SPADA is administered by DIKTI and encompasses diverse subjects across social sciences, natural sciences, engineering, and health disciplines from various Indonesian universities.

B. Support Vector Machine Recursive Feature Elimination

Feature selection or dimensionality reduction techniques aim to alleviate the challenge posed by an abundance of features in training data that exhibit limited statistical correlation with class labels, thereby augmenting efficiency and accuracy [34]. SVM-RFE, an SVM-based feature selection algorithm introduced in [35], functions by identifying critical feature subsets. As a result, SVM-RFE optimizes the computational time necessary for classification tasks while concurrently enhancing classification accuracy [35].

SVM-RFE Algorithm [35] :

1. Input
   a. Training data feature, $X_0 = [x_{i1}, x_{i2}, \ldots, x_{im}]$.
   b. Training data label, $y = [y_1, y_2, \ldots, y_n]^T$.
   c. Current feature set, $s = [1, 2, \ldots, m]$.
   d. Features with sorted weight, $r = \emptyset$.
2. Feature Sorting
   a. Perform steps 2.a. to 2.h. to $s = \emptyset$.
   b. A new training data matrix is obtained from the remaining features, $X = X_0(:, s)$.
   c. Training data classifier, $\alpha = SVM - train(X, y)$
   d. Calculate weights, $w = \sum_k \alpha_k y_k x_k$.
   e. Calculate sorting base value, $c = (w)^2$.
   f. Determine the feature with smallest weight, $f = \min(c)$.
   g. Updating the sorted feature list, $r = [s(f), r]$.
   h. Remove the feature with the smallest weight, $s = s\big(1{:}-1, f + 1{:}\,length(s)\big)$.
3. Output: Sorted features r.

In each iteration, the feature with the minimum weight value is systematically excluded. Subsequently, SVM retrains on the remaining features to generate a new sorted feature list. This iterative process is repeated until a finalized sorted list of features is attained. During each iteration, SVM constructs a model utilizing training data derived from the subset of sorted features, evaluating their respective performances. Ultimately, this iterative approach facilitates the acquisition of an optimal subset of features [36].

C. Proposed Detection Method

The stages outlined in the proposed method during the investigation are delineated as follows:

1) Feature identification: This stage entails identifying the features utilized to construct the SVM classifier model, leveraging attributes from the SPADA LMS database that correspond to indicators across each dimension within the FSLSM model. Feature identification involves two methods: firstly, selecting SPADA LMS database attributes directly linked to FSLSM model indicators; secondly, synthesizing features through the aggregation or accumulation of multiple SPADA LMS database attributes to fulfill the indicators for each FSLSM model dimension.

2) Feature mapping: Identified features are mapped to learners' online behaviors, aligning them with learning styles (or classes) corresponding to each of the four dimensions (labels) within the FSLSM model.

3) Learning style labeling: In this stage, each sample or observation data is assigned a learning style label based on the results of online behavioral mapping across each dimension. This process annotates each sample or learner's data into binary classes representing their learning styles across all four dimensions.

4) Data splitting: The formed dataset, after mapping features and labeling for the four dimensions, is split into two groups. The dataset is divided into 80% and 20% proportions, with this ratio uniformly applied to each of the eight learning styles (2 learning styles with four dimensions) for all data. The subset comprising 80% of the data constitutes the training data, while the remaining 20% forms the test data. The training data is utilized to construct the classifier model, while the test data is employed to evaluate the performances of the final classifier.

5) Initial classifier construction: Utilizing the training data, the initial classifier model is constructed by employing the SVM algorithm. All features are incorporated into the initial classifier, with feature selection conducted in the subsequent stage using the Recursive Feature Elimination (RFE) algorithm.

6) Feature selection: Feature selection is performed through the SVM-RFE procedure, recursively eliminating features in each iteration based on the sequentially stored lowest weight value for each FSLSM dimension.

7) Final classifier evaluation: Upon obtaining the final classifier model using SVM-RFE with selected features for each FSLSM dimension, its performance is evaluated using the test data. Performance metrics such as accuracy, precision, sensitivity/recall, and F1-score are computed by assessing correctly and misclassified learning styles. These metrics are presented in a confusion matrix table to facilitate performance evaluation.

The workflow of the proposed method is described in Fig. 1.

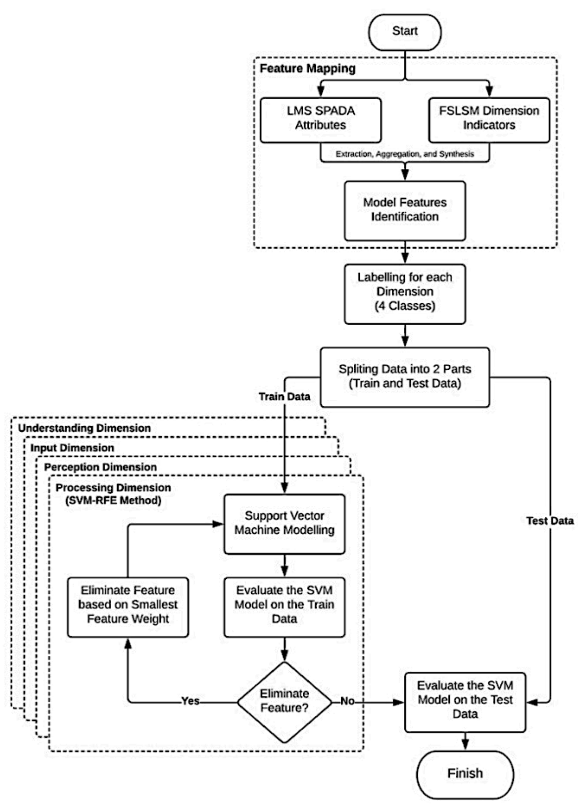

Fig. 1. Workflow diagram of the proposed detection method.

III. Results

A. Features Identification

1) Directly extracted features from the LMS attributes: Data is extracted from SPADA LMS through direct data retrieval. This process is predicated on the assumption that the selected features directly influence the determination of learners' learning styles. The outcomes of feature extraction are detailed in Table I.

TABLE I. Classifier Features Extracted from the SPADA LMS Database Attributes based on the FSLSM Model

Feature	LMS Attribute Database	FSLSM Indicators
Quiz Revisions (X1)	*Total Quiz Attempts*	The number of quiz revisions. Features are retrieved from columns in the SPADA LMS dataset.
Exercise Visit (X2)	*Number of Assignment Submission*	The number of duty visits. Features are derived from the "Number of Assignment Submissions" column in the SPADA LMS dataset.
Content Visit (X3)	*Content Completed*	The number of contents visited. Features are extracted from the "Content Completed" column in the SPADA LMS dataset.
Content Stay (X4)	*Time Spent in Content*	The length of time on content. Features are derived from the "Time spent in a content" column in the SPADA LMS dataset.
Forum Visit (X5)	*Discussion Post Read*	The visited discussion forums. Features are extracted from the "Discussion Post Read" column in the SPADA LMS dataset.
Forum Posts (X6)	*Discussion Post Created*	The discussion forum created. The feature is extracted from the "Discussion Post Created" column in the SPADA LMS dataset.

2) Synthesized features derived from multiple attributes within the LMS: Features can also be generated through the aggregation or accumulation of several attributes within the SPADA LMS database, culminating in synthesized features utilized as inputs in the SVM classifier model. The outcomes of several synthetic attributes are detailed in Table II.

Thirteen features have been generated, presumed to influence learning styles. These features are derived from the extraction and synthesis of various attributes within the SPADA LMS database, as demonstrated in Tables I and II. Moreover, all features will be aligned with the four dimensions of the FSLSM model.

TABLE II. Classifier Features Synthesized from the Attributes of the SPADA LMS Database, Guided by the FSLSM Model

Features	LMS Database Attributes	FSLSM Indicators
Forums Stay (X7)	*Discussion Post Read, Discussion Post Replies*	Long duration of time spent in the forum. The features are derived from the "Discussion Post Read" and "Discussion Post Replies" columns in the SPADA LMS dataset. The durations are categorized by comparing "Discussion Post Read" and "Discussion Post Replies" with their respective averages.
Question Graphics Points (X8)	*Quiz Completed Content Stay*	Question points in graphical form. The features are extracted from the "Quiz Completed" and "Content Stay" columns in the SPADA LMS dataset. Points are allocated by comparing "Quiz Completed" with the average, and the average "Content Stay" for each class.
Question Text Points (X9)	*Quiz Completed Content Stay*	Question points in textual format. Points are allocated based on the comparison of "Quiz Completed" with the average, and the average "Content Stay" for each class. The features are extracted from the "Quiz Completed" and "Content Stay" columns in the SPADA LMS dataset.
Question Facts Points (X10)	*Quiz Completed Number of Assignment*	Question points based on factual data. The features are derived from the "Quiz Completed" and "Number of Assignment Submissions" columns in the SPADA LMS dataset. Points are assigned by comparing "Quiz Completed" and "Number of Assignment Submissions" with their respective averages.
Question Concepts Points (X11)	*Quiz Completed Content Completed*	Question points based on a conceptual framework. The features are extracted from the "Quiz Completed" and "Content Completed" columns in the SPADA LMS dataset. Points are allocated by comparing "Quiz Completed" and "Content Completed" with their respective averages.
Question Details Points (X12)	*Quiz Completed Discussion Post Replies*	Question points presented as granular details. The features are extracted from the "Quiz Completed" and "Discussion Post Replies" columns in the SPADA LMS dataset. Points are assigned by comparing "Quiz Completed" and "Discussion Post Replies" with their respective averages.
Question Overview Points (X13)	*Quiz Completed Discussion Post Read*	Question points presented in an overarching manner. The features are derived from the "Quiz Completed" and "Discussion Post Read" columns in the SPADA LMS dataset. Points are allocated by comparing "Quiz Completed" and "Discussion Post Read" with their respective averages.

B. Features Mapping

The outcomes of feature mapping are presented in Table III, categorized by LMS activities corresponding to learning styles within the FSLSM dimensions. Table III illustrates the roles of the 13 features obtained in the preceding process, each associated with specific dimensions in the FSLSM. Consequently, each feature in Table III will be utilized for labeling units.

TABLE III. FSLSM ONLINE BEHAVIOR PATTERN MAPPING

Features	FSLSM Dimension							
	Processing		Perception		Input		Understanding	
	Active	Reflective	Sensing	Intuitive	Visual	Verbal	Sequential	Global
Quiz Revisions (X1)			+	-				
Exercise Visit (X2)	+	-	+	-				
Content Visit (X3)	-	+	-	+	-	+	-	+
Content Stay (X4)	-	+	-	+	+	-		
Forums Visit (X5)	-	+			-	+		
Forum Posts (X6)	+	-			-	+		
Forum Stay (X7)					-	+		
Question Graphics. Points (X8)					+	-		
Question Text Points (X9)					-	+		
Question Facts Points (X10)			-	+				
Question Concepts Points (X11)			+	-				
Question Details Points (X12)			+	-			+	-
Question Overview Points (X13)							-	+

C. SVM-RFE Modelling

The labeling process involves establishing the threshold for the learning behavior pattern pertinent to each dimension in the FSLSM, as outlined in Table IV. Subsequently, a value of (-1.1) is assigned to each feature based on the threshold, and the total value for each learning style within each dimension in the FSLSM is computed. The resulting total values are then categorized based on their sign.

Within the Processing dimension, a total score with a positive sign (+) signifies an active learning style, whereas a negative sign (-) indicates a reflective learning style. In the Perception dimension, a positive sign (+) denotes a sensing learning style, while a negative sign (-) characterizes an intuitive learning style. In the Input dimension, a positive sign (+) signifies a visual learning style, whereas a negative sign (-) denotes a verbal learning style. Finally, within the Understanding dimension, a positive sign (+) describes a sequential learning style, whereas a negative sign (-) represents a global learning style. The outcomes of the labeling process are depicted in Fig. 2. According to Fig. 2, there are 52,027 learners with active learning styles, 57,780 with reflective styles, 78,713 with sensing styles, 31,094 with intuitive styles, 62,216 with visual styles, 47,591 with verbal styles, 36,397 with sequential styles, and 73,410 with global styles.

1) Building the initial SVM classifier: Following the labeling process, the data is modeled using the SVM method, incorporating all the features obtained in the preceding approach (Feature Identification for SVM Classifiers in SPADA LMS). However, prior to SVM modeling, it is essential to partition the data into training and test sets. The training data is utilized to construct a model, whereas the test data is employed to evaluate the performance of the resultant model. Previous studies have demonstrated favorable accuracy with a training data percentage of 70% and test data percentage of 30%, and 80% and 20%, respectively, for substantial datasets (comprising thousands or millions of entries). In this study, a proportion of 80% of training data and 20% of test data will be employed. The outcomes of the confusion matrix for the SVM model utilizing all the features are detailed in Table IV.

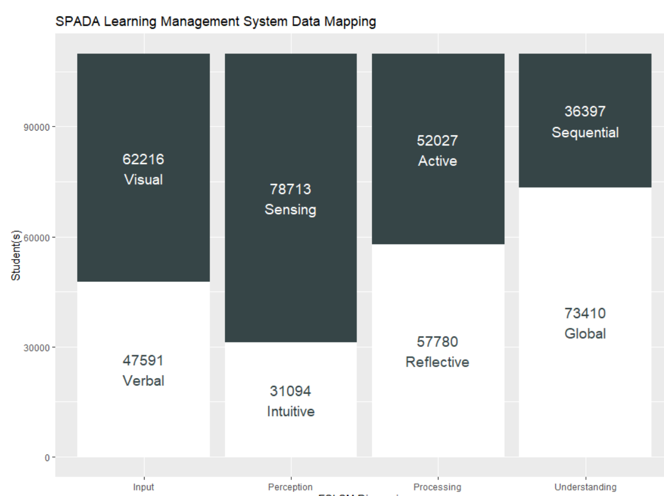

Fig. 2. Composition of dimension/learning style labeling.

2) *Initial SVM classifier model*: The researchers utilize the confusion matrix presented in Table IV to assess the performance of the SVM model constructed. The performance evaluation results of the model are outlined in Table V. The performance metrics obtained include accuracy, sensitivity, specificity, and F1-Score values.

TABLE IV. CONFUSION MATRIX SVM

Confusion Matrix		Reference				Dimension
		Train		Test		
		+	-	+	-	
Prediction	+	39906	2362	9974	540	Processing
	-	1765	44008	382	10870	
	+	58611	8213	14568	2022	Perception
	-	4482	16735	1052	4124	
	+	46481	2798	11479	672	Input
	-	3439	35323	817	8798	
	+	26466	301	6480	83	Understanding
	-	2754	58520	697	14506	

According to Table V, the metrics' values within each dimension indicate satisfactory performance. However, further analysis is warranted to discern the most influential features in each dimension. One appropriate approach to depict this is the SVM-RFE method, aligning with the classifier method employed in this study.

TABLE V. SVM MODEL PERFORMANCE

Dimension Model	Data Split	Accuracy	Sensitivity	Specificity	F1-Score
Processing	Train	95.31%	95.76%	94.91%	95.08%
	Test	95.76%	96.31%	95.27%	95.58%
Perception	Train	85.58%	92.90%	67.08%	90.23%
	Test	85.88%	93.27%	67.10%	90.46%
Input	Train	92.92%	93.11%	92.66%	93.71%
	Test	93.16%	93.36%	92.90%	93.91%
Understanding	Train	96.53%	90.58%	99.49%	94.54%
	Test	96.42%	90.29%	99.43%	94.32%

3) *Feature selection*: Feature selection is conducted using the Recursive Feature Elimination (RFE) method. As the base classifier model employs the Support Vector Machine (SVM), the technique is termed SVM-RFE. This method entails iteratively modeling the SVM method with a linear kernel. During each iteration, weights obtained in the SVM model are calculated, and the feature with the lowest weight is eliminated. Subsequently, a new SVM model is constructed by excluding the removed features. This process is repeated recursively until the specified number of features is attained. Feature importance is determined based on the sequence in which features are eliminated during the recursive process. Features removed earlier indicate lower significance in the resultant SVM model. The results of SVM-RFE for feature selection from each dimension are presented in the Appendix at the end of this paper. Table VI delineates the outcomes of the feature selection process utilizing the SVM-RFE method.

TABLE VI. SVM-RFE FEATURE RANKING RESULTS

Rank	Processing	Perception	Input	Understanding
1	Forums Visit (X5)	Content Stay (X4)	Question Graphics Points (X8)	Content Visit (X3)
2	Content Visit (X3)	Question Concepts Points (X11)	Question Overview Points (X13)	Exercise Visit (X2)
3	Content Stay (X4)	Content Visit (X3)	Question Details Points (X12)	Forums Stay (X7)
4	Question Text Points (X9)	Exercise Visit (X2)	Content Visit (X3)	Question Overview Points (X13)
5	Question Overview Points (X13)	Quiz Revisions (X1)	Content Stay (X4)	Question Details Points (X12)

The results of feature ranking in Table VI are derived from the weight values obtained through the SVM-RFE method, accessible in the Appendix of this paper. Based on the experimental findings, it can be inferred that the "Content Visit" (X3) feature significantly influences all four dimensions of FSLSM. This is evidenced by the fact that the "Content Visit" (X3) feature is ranked within the top 5 in each FSLSM dimension. Furthermore, apart from the "Content Visit" (X3), the "Content Stay" (X4) feature also exhibits a considerable influence, albeit not in the Understanding dimension. This observation is supported by Table VI, where the "Content Stay" (X4) feature is not among the top 5 rankings in the Understanding dimension.

4) *SVM-RFE classifier performance evaluation*: In addition to the confusion matrix, metric values such as accuracy, sensitivity, specificity, and F1-Score are presented in Table VIII. The SVM model utilizing the top 5 variables in the training data for the Processing, Perception, Input, and Understanding dimensions yields respective F1-Score values of 94.49%, 71.00%, 89.18%, and 97.46%. Conversely, the SVM model employing the top 5 variables in the validation data for the Processing, Perception, Input, and Understanding dimensions exhibits F1-Score values of 94.94%, 71.48%, 89.61%, and 97.38%, respectively. The negligible discrepancy between F1-Score values in the training and test data indicates that SVM yields a reasonably robust model on the SPADA LMS dataset. Table VII shows the confusion matrix SVM-RFE.

TABLE VII. CONFUSION MATRIX SVM-RFE

Confusion Matrix		Reference				Dimension
		Train		Test		
		+	-	+	-	
Prediction	+	39906	2362	9974	540	Processing
	-	1765	44008	382	10870	
	+	58611	8213	14568	2022	Perception
	-	4482	16735	1052	4124	
	+	46481	2798	11479	672	Input

	-	3439	35323	817	8798	
	+	26466	301	6480	83	Understanding
	-	2754	58520	697	14506	

TABLE VIII. SVM-RFE MODEL PERFORMANCE

Dimension Model	**Data Split**	**Accuracy**	**Sensitivity**	**Specificity**	**F1-Score**
Processing	Train	94.78%	94.53%	95.00%	94.49%
	Test	95.18%	95.12%	95.22%	94.94%
Perception	Train	84.48%	67.06%	91.37%	71.00%
	Test	84.75%	67.69%	91.47%	71.48%
Input	Train	90.68%	88.72%	92.17%	89.18%
	Test	91.01%	89.16%	92.44%	89.61%
Understanding	Train	96.53%	99.49%	90.57%	97.46%
	Test	96.41%	99.43%	90.27%	97.38%

The results presented in Table IX reveal a relatively low level of correlation between features within the Processing dimension. The highest correlation coefficient, at 0.516, is observed between features X3 and X4. Consequently, it can be inferred that the features comprising the SVM-RFE model in the Processing dimension contain independent information, except for the moderate level of association between X3 and X4. Similarly, in the Input and Understanding dimensions, the highest correlation coefficient values between features range from 0.516 to 0.553. Notably, the highest correlation coefficient within these dimensions occurs between pairs of features, namely X3 and X4 in the Input dimension, and X12 and X13 in the Understanding dimension. In contrast, the Perception dimension exhibits a slightly different pattern, with the highest correlation coefficient ranging from 0.516 to 0.748 observed among two pairs of features: X4 and X3, and X1. Consequently, it can be deduced that the features constituting the SVM-RFE model in the Perception dimension contain information with a relatively high level of association. This phenomenon is presumed to contribute to the slightly lower performance of the model in the Perception dimension compared to the other three dimensions.

TABLE IX. FEATURE CORRELATION

Dimension	**Correlation**					
Processing		X5	X3	X4	X9	X13
	X5	1.000	0.352	0.371	0.008	0.323
	X3	0.352	1.000	0.516	0.173	0.157
	X4	0.371	0.516	1.000	0.056	0.220
	X9	0.008	0.173	0.056	1.000	0.258
	X13	0.323	0.157	0.220	0.258	1.000
Perception		X4	X11	X3	X2	X1
	X4	1.000	0.307	0.516	-0.016	0.488
	X11	0.307	1.000	0.145	0.056	0.167
	X3	0.516	0.145	1.000	-0.063	
	X2	-0.016	0.056	-0.063	1.000	-0.051
	X1	0.488	0.167		-0.051	1.000
Input		X8	X13	X12	X3	X4
	X8		0.223	0.202	-0.201	0.273
	X13	0.223		0.553	0.157	0.220
	X12	0.202	0.553		0.165	0.194
	X3	-0.201	0.157	0.165		0.516
	X4	0.273	0.220	0.194	0.516	
Understanding		X3	X2	X7	X13	X12
	X3	1.000	-0.063	0.011	0.157	0.165
	X2	-0.063	1.000	-0.007	-0.081	-0.068
	X7	0.011	-0.007	1.000	0.007	0.002
	X13	0.157	-0.081	0.007	1.000	0.553
	X12	0.165	-0.068	0.002	0.553	1.000

IV. DISCUSSION

Based on the available data from LMS-SPADA, six features correspond to FSLSM indicators. Subsequently, we generated features by extracting and synthesizing several attributes assumed to determine learning styles, following the methodology outlined by [37]. Consequently, we obtained 13 features that will be mapped onto FSLSM learning styles based on [38]. Notably, several features correspond to more than one dimension in the FSLSM learning style model during this feature identification process. For instance, the content visit feature (X3) can detect learning styles across all four dimensions of FSLSM, while the content stay feature is applicable to the processing, perception, and input dimensions. Unlike previous research, which independently mapped each feature onto each dimension, as observed in studies by [31] and [39], our approach considers the holistic mapping of features onto multiple dimensions. Furthermore, most previous researchers utilized the ILS instrument for labelling the four dimensions of FSLSM for modelling using machine learning techniques.

The Appendix presents the feature selection results sorted by feature weight value using the SVM-RFE algorithm for each dimension in the FSLSM model. Table X illustrates the sorting features using the SVM-RFE method for the Processing dimension. It is apparent that removing features X10, X11, X6, X2, X1, X12, X7, X8, X13, and X9 does not significantly impact Accuracy, Sensitivity, Specificity, and F1-Score. This indicates that eliminating these features can expedite the computational process without significantly affecting the SVM model's performance in classification. However, for features X4, X3, and X5, the model's performance also decreased, albeit more noticeably compared to the earlier features. This suggests that the contribution or influence of features X4, X3, and X5 is more significant than that of the previous features on the Processing dimension.

The data reveals that X3 and X11 play a crucial role in classifying learning styles within the Perception dimension. Table XI presents the sorted features using the SVM-RFE method for the Perception dimension. Notably, the performance of the SVM model remained relatively stable until the removal of the X3 feature. This suggests that features released before the removal of X3 have a minimal contribution to determining the learning style within the Perception dimension. The significant drop in the Specificity value subsequent to the removal of X3 and X11 provides further evidence that these features substantially influence the differentiation between Sensing and Intuitive learning styles within the Perception dimension.

Similar to the Perception dimension, the Input dimension, as depicted in Table XII, exhibited a notable reduction in Specificity upon the removal of the X12 and X13 features. This indicates the substantial contribution of the X12 and X13 features to identifying the learning style within the Input dimension. Therefore, at minimum, the inclusion of the X8, X12, and X13 features is necessary to achieve a reasonably effective SVM model, as evidenced by the high F1-Score value in the 11th iteration.

In contrast to the preceding three dimensions, the Understanding dimension, as presented in Table XIII, demonstrates a minor shift in the SVM model's performance during the SVM-RFE process. As discussed earlier, the majority of features significantly influence each dimension within the model, indicating that the final feature, X3, holds the most prominent sway over the learning styles within the Understanding dimension. Appendix 4 further underscores the independence of X3 from other features, given the negligible alteration in SVM model performance upon its removal. These findings underscore the necessity for additional analysis during feature selection in each dimension using SVM-RFE to mitigate the risk of SVM models exhibiting high bias or variance.

The accuracy of the SVM-RFE model exhibits the lowest value in the perception dimension, whereas it remains relatively consistent across the other dimensions. This contrasts with the findings of [39], where the SVM model achieved its highest score in the perception dimension and nearly identical values across the other dimensions.

V. Conclusion

This paper has presented a framework for extracting features from LMS-SPADA, the largest higher-education LMS in Indonesia, to align with the learning style indicators of the FSLSM model. These features were identified based on the indicators of the Felder-Silverman Learning Style Model (FSLSM). We utilized the mapping results as independent variables to automatically detect students' learning styles using the SVM-RFE method. The SVM-RFE classifier integrates features from LMS-SPADA database attributes with FSLSM dimension indicators, enhancing the accuracy of learning style detection. Our experiments yielded accuracy results of 95.76% for the Processing dimension, 85.88% for the Perception dimension, 93.16% for the Input dimension, and 96.42% for the Understanding dimension. Additionally, the SVM-RFE method identified the top five features contributing to learner learning styles in each dimension: for the Processing dimension, these features are Forums Visit (X5), Content Visit (X3), Content Stay (X4), Question Text Points (X9), and Question Overview Points (X13); for the Perception dimension, they are Content Stay (X4), Question Concepts Points (X11), Content Visit (X3), Exercise Visit (X2), and Quiz Revisions (X1); and for the Input dimension, they are Question Graphics Points (X8), Question Overview Points (X13), and Question Detail Points (X12).

We have identified several limitations in our study, including the lack of comparison with other classification techniques. Further research is necessary to validate our findings in different contexts using standard machine learning methods. In future work, it would be beneficial to compare various classification techniques across different machine learning models to determine the most suitable model for detecting learning styles. One strategy to improve model performance is through the use of ensemble techniques, which combine the outputs of multiple weak learner algorithms, whether similar or disparate. These ensemble techniques include averaging, voting, stacking, boosting, and other similar approaches.

References

[1] S. Pokhrel and R. Chhetri, “A Literature Review on Impact of COVID-19 Pandemic on Teaching and Learning,” Higher Education for the Future, vol. 8, no. 1, pp. 133–141, Jan. 2021, doi: 10.1177/2347631120983481.

[2] S. Subedi, S. Nayaju, and S. Subedi, “Impact of E-learning during COVID-19 Pandemic among Nursing Students and Teachers of Nepal,” International Journal of Science and Healthcare Research, vol. 5, no. 3, pp. 68–76, 2020.

[3] Ramzi Nasser, M. Cherif, and M. Romanowski, “Factors that Impact Student Usage of the Learning… – International Review of Research in Open and Distributed Learning – Érudit.” Accessed: Jan. 29, 2022. [Online]. Available: https://www.erudit.org/en/journals/irrodl/1900-v1-n1-irrodl05121/1067477ar/abstract/

[4] H. coates, R. james, and G. baldwin, “A Critical Examination Of The Effects Of Learning Management Systems On University Teaching And Learning,” Tert Educ Manag, vol. 11, no. 1, pp. 19–36, Mar. 2005, doi: 10.1007/s11233-004-3567-9.

[5] T. Jurubescu, “Learning Content Management Systems,” Revista Informatica Economică, p. 4, 2008.

[6] R. D. Costa, G. F. Souza, R. A. M. Valentim, and T. B. Castro, “The theory of learning styles applied to distance learning,” Cognitive Systems Research, vol. 64, pp. 134–145, Desember 2020, doi: 10.1016/j.cogsys.2020.08.004.

[7] Y. Feng, F. Iriarte, and J. Valencia, “Relationship Between Learning Styles, Learning Strategies and Academic Performance of Chinese Students Who Learn Spanish as a Foreign Language,” Asia-Pacific Edu Res, vol. 29, no. 5, pp. 431–440, Oct. 2020, doi: 10.1007/s40299-019-00496-8.

[8] N. Shamsuddin and J. Kaur, “Students’ Learning Style and Its Effect on Blended Learning, Does It Matter?,” International Journal of Evaluation and Research in Education, vol. 9, no. 1, pp. 195–202, Mar. 2020.

[9] R. S. Vaishnav, “Learning Style And Academic Achievement Of Secondary School Students,” Learning Style And Academic Achievement, vol. 1, no. 4, Mar. 2013.

[10] S. Graf, “Adaptivity in Learning Management Systems Focussing On Learning Styles,” Vienna University of Technology, 9801086 Neulinggasse 22/12A 1030 Vienna, 2007.

[11] H. Pashler, M. McDaniel, D. Rohrer, and R. Bjork, “Learning Styles: Concepts and Evidence,” Psychol Sci Public Interest, vol. 9, no. 3, pp. 105–119, Desember 2008, doi: 10.1111/j.1539-6053.2009.01038.x.

[12] P. A. Kirschner, “Stop propagating the learning styles myth,” Computers & Education, vol. 106, pp. 166–171, Mar. 2017, doi: 10.1016/j.compedu.2016.12.006.

[13] R. M. Felder and J. E. Spurlin, “Applications, Reliability and Validity of the Index of Learning Styles,” International Journal of Engineering Education, vol. 21, no. 1, pp. 103–112, 2005.

[14] O. E. Aissaoui, Y. E. A. El madani, L. Oughdir, and Y. E. Allioui, “Combining supervised and unsupervised machine learning algorithms to predict the learners’ learning styles,” Procedia Computer Science, vol. 148, pp. 87–96, Jan. 2019, doi: 10.1016/j.procs.2019.01.012.

[15] L. D. Ferreira, G. Spadon, A. C. Carvalho, and J. F. Rodrigues, “A comparative analysis of the automatic modeling of Learning Styles through Machine Learning techniques,” in 2018 IEEE Frontiers in Education Conference (FIE), Oct. 2018, pp. 1–8. doi: 10.1109/FIE.2018.8659191.

[16] N. Hidayat, R. Wardoyo, A. Sn, and H. D. Surjono, “Enhanced Performance of the Automatic Learning Style Detection Model using a Combination of Modified K-Means Algorithm and Naive Bayesian,” International Journal of Advanced Computer Science and Applications

(IJACSA), vol. 11, no. 3, Art. no. 3, 40/30 2020, doi: 10.14569/IJACSA.2020.0110380.

[17] I. Karagiannis and M. Satratzemi, "An adaptive mechanism for Moodle based on automatic detection of learning styles," Educ Inf Technol, vol. 23, no. 3, pp. 1331–1357, May 2018, doi: 10.1007/s10639-017-9663-5.

[18] A. Kika, L. Leka, S. Maxhelaku, and A. Ktona, "Using data mining techniques on Moodle data for classification of student?s learning styles," International Institute of Social and Economic Sciences, 9211567, Jul. 2019.

[19] L. A. Dantas and A. Cunha, "An integrative debate on learning styles and the learning process," Social Sciences & Humanities Open, vol. 2, no. 1, p. 100017, Jan. 2020, doi: 10.1016/j.ssaho.2020.100017.

[20] E. Gomede, R. Miranda de Barros, and L. de Souza Mendes, "Use of Deep Multi-Target Prediction to Identify Learning Styles," Applied Sciences, vol. 10, no. 5, Art. no. 5, Jan. 2020, doi: 10.3390/app10051756.

[21] O. Zine, A. Derouich, and A. Talbi, "A Comparative Study of the Most Influential Learning Styles used in Adaptive Educational Environments," International Journal of Advanced Computer Science and Applications (IJACSA), vol. 10, no. 11, Art. no. 11, 51/30 2019, doi: 10.14569/IJACSA.2019.0101171.

[22] L. X. Li and S. S. Abdul Rahman, "Students' learning style detection using tree augmented naive Bayes," Royal Society Open Science, vol. 5, no. 7, p. 172108, 2018, doi: 10.1098/rsos.172108.

[23] S. M. Nafea, F. Siewe, and Y. He, "On Recommendation of Learning Objects Using Felder-Silverman Learning Style Model," IEEE Access, vol. 7, pp. 163034–163048, 2019, doi: 10.1109/ACCESS.2019.2935417.

[24] J. Feldman, A. Monteserin, and A. Amandi, "Automatic detection of learning styles: state of the art," Artif Intell Rev, vol. 44, no. 2, pp. 157–186, Aug. 2015, doi: 10.1007/s10462-014-9422-6.

[25] S. Shrestha and M. Pokharel, "Determining Learning Style Preferences of Learners," Journal of Computer Science Research, vol. 3, no. 1, Art. no. 1, Feb. 2021, doi: 10.30564/jcsr.v3i1.2761.

[26] R. E. Caceffo, E. Valle, R. Mesquita, and R. Azevedo, "Assessment of Predictive Power of The Felder & Silverman Learning Styles Model on Students' Performance in an Introductory Physics Course," European Journal of Physics Education, vol. 10, no. 2, pp. 1–22, Apr. 2019, doi: 10.20308/ejpe.v10i2.227.

[27] M. M. El-Bishouty et al., "Use of Felder and Silverman learning style model for online course design," Education Tech Research Dev, vol. 67, no. 1, pp. 161–177, Feb. 2019, doi: 10.1007/s11423-018-9634-6.

[28] T. Sheeba and R. Krishnan, "Automatic Detection of Students Learning Style in Learning Management System," in Smart Technologies and Innovation for a Sustainable Future, A. Al-Masri and K. Curran, Eds., in Advances in Science, Technology & Innovation. Cham: Springer International Publishing, 2019, pp. 45–53. doi: 10.1007/978-3-030-01659-3_7.

[29] J. Nasiri, A. M. Mir, and S. Fatahi, "Classification of learning styles using behavioral features and twin support vector machine," Technology of Education Journal (TEJ), vol. 13, no. 2, pp. 316–326, Mar. 2019, doi: 10.22061/jte.2018.3358.1859.

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

[31] E. S. Amir, M. Sumadyo, D. I. Sensuse, Y. G. Sucahyo, and H. B. Santoso, "Automatic detection of learning styles in learning management system by using literature-based method and support vector machine," in 2016 International Conference on Advanced Computer Science and Information Systems (ICACSIS), Oct. 2016, pp. 141–144. doi: 10.1109/ICACSIS.2016.7872770.

[32] V. Vapnik, S. E. Golowich, and A. J. Smola, "Support Vector Method for Function Approximation, Regression Estimation and Signal Processing," in Advances in Neural Information Processing Systems 9, M. C. Mozer, M. I. Jordan, and T. Petsche, Eds., MIT Press, 1997, pp. 281–287.

[33] D. Ifenthaler, "Toward automated computer-based visualization and assessment of team-based performance," Journal of Educational Psychology, vol. 106, no. 3, pp. 651–655, 2014, doi: https://doi.org/10.1037/a0035505.

[34] A. Pal and J. Maiti, "Development of a hybrid methodology for dimensionality reduction in Mahalanobis–Taguchi system using Mahalanobis distance and binary particle swarm optimization," Expert Systems with Applications, vol. 37, no. 2, pp. 1286–1293, Mar. 2010, doi: 10.1016/j.eswa.2009.06.011.

[35] I. Guyon, J. Weston, S. Barnhill, and V. Vapnik, "Gene Selection for Cancer Classification using Support Vector Machines," Machine Learning, vol. 46, no. 1, pp. 389–422, Jan. 2002, doi: 10.1023/A:1012487302797.

[36] M.-L. Huang, Y.-H. Hung, W. M. Lee, R. K. Li, and B.-R. Jiang, "SVM-RFE Based Feature Selection and Taguchi Parameters Optimization for Multiclass SVM Classifier," The Scientific World Journal, vol. 2014, p. e795624, Sep. 2014, doi: 10.1155/2014/795624.

[37] S. Graf, Kinshuk, and T.-C. Liu, "Identifying Learning Styles in Learning Management Systems by Using Indications from Students' Behaviour," in 2008 Eighth IEEE International Conference on Advanced Learning Technologies, Jul. 2008, pp. 482–486. doi: 10.1109/ICALT.2008.84.

[38] P. Pitigala Liyanage, L. Gunawardena, and M. Hirakawa, "Detecting Learning Styles in Learning Management Systems Using Data Mining," Journal of Information Processing, vol. 24, pp. 740–749, Jul. 2016, doi: 10.2197/ipsjjip.24.740.

[39] F. Rasheed and A. Wahid, "Learning style detection in E-learning systems using machine learning techniques," Expert Systems with Applications, vol. 174, p. 114774, Jul. 2021, doi: 10.1016/j.eswa.2021.114774.

APPENDIX

TABLE X. SVM-RFE RESULTS FOR PROCESSING DIMENSION

Iteration	Sorted Feature													Deleted Feature	Metric			
															Acc.	Sensv.	Specf.	F1
1	X10	X11	X6	X2	X1	X12	X9	X13	X8	X7	X4	X3	X5	-	95.40%	95.88%	94.96%	95.18%
	0.0017	0.0067	0.0166	0.0277	0.1925	0.5041	0.6614	0.6702	0.8872	1.0383	2.95	13,956	21,675					
2	X11	X6	X2	X1	X12	X9	X13	X8	X7	X4	X3	X5	-	X10	95.38%	95.86%	94.94%	95.16%
	0.0061	0.0171	0.0292	0.1931	0.5262	0.6543	0.6931	0.8839	1.0799	2,931	13,951	21,553	-					
3	X6	X2	X1	X12	X13	X9	X8	X7	X4	X3	X5	-	-	X11	95.35%	95.82%	94.92%	95.12%
	0.016	0.0291	0.1915	0.4595	0.619	0.6491	0.8204	0.9649	3.1501	13.91	21.3	-	-					
4	X2	X1	X12	X13	X9	X8	X7	X4	X3	X5	-	-	-	X6	95.33%	95.76%	94.94%	95.10%
	0.0287	0.1867	0.3724	0.5123	0.5445	0.6866	0.7724	3.2194	13,667	21.13	-	-	-					
5	X1	X12	X9	X13	X8	X7	X4	X3	X5	-	-	-	-	X2	95.00%	95.28%	94.75%	94.75%
	0.1739	0.3554	0.5986	0.6056	0.7106	0.8265	2.5906	10.262	24,357	-	-	-	-					
6	X12	X13	X9	X8	X7	X4	X3	X5	-	-	-	-	-	X1	94.89%	95.15%	94.66%	94.64%
	0.3118	0.5536	0.5782	0.5823	0.7457	2.5022	11.893	23,994	-	-	-	-	-					
7	X7	X8	X13	X9	X4	X3	X5	-	-	-	-	-	-	X12	94.87%	94.80%	94.94%	94.60%
	0.0218	0.0533	0.0981	0.1064	3.4126	10,902	24,991	-	-	-	-	-	-					
8	X8	X13	X9	X4	X3	X5	-	-	-	-	-	-	-	X7	94.83%	94.72%	94.93%	94.55%
	0.0276	0.0382	0.8001	3.5356	10.77	25,272	-	-	-	-	-	-	-					

9	X13	X9	X4	X3	X5	-	-	-	-	-	-	-	-	X8	94.85%	94.65%	95.04%	94.57%
	0.0243	0.0429	3.9216	10.337	24,295	-	-	-	-	-	-	-	-					
10	X9	X4	X3	X5	-	-	-	-	-	-	-	-	-	X13	94.80%	94.51%	95.07%	94.52%
	0.0272	3.8354	10,563	21.482	-	-	-	-	-	-	-	-	-					
11	X4	X3	X5	-	-	-	-	-	-	-	-	-	-	X9	94.80%	94.48%	95.10%	94.51%
	3.7407	10,876	21.105	-	-	-	-	-	-	-	-	-	-					
12	X3	X5	-	-	-	-	-	-	-	-	-	-	-	X4	90.66%	89.63%	91.58%	90.09%
	7.1102	14,296	-	-	-	-	-	-	-	-	-	-	-					
13	X5	-	-	-	-	-	-	-	-	-	-	-	-	X3	82.29%	81.42%	83.08%	81.33%
	25,037	-	-	-	-	-	-	-	-	-	-	-	-					

TABLE XI. SVM-RFE RESULTS FOR PERCEPTION DIMENSION

Iteration	Sorted Feature													Deleted Feature	Metric			
															Acc.	Sensv.	Specf.	F1
1	X6	X10	X5	X1	X12	X13	X9	X3	X11	X2	X7	X8	X4	-	85.64%	92.97%	67.10%	90.28%
	0.0004	0.002	0.0054	0.2246	0.2418	0.2913	0.4059	0.5351	0.5762	0.5858	0.5886	0.789	2.0794					
2	X10	X5	X1	X12	X13	X9	X3	X11	X7	X2	X8	X4	-	X6	85.65%	92.97%	67.12%	90.28%
	0.002	0.0054	0.224	0.2372	0.2874	0.4009	0.5364	0.5759	0.5788	0.5834	0.78	2.0725	-					
3	X5	X12	X1	X13	X9	X3	X7	X11	X2	X8	X4	-	-	X10	85.65%	92.97%	67.09%	90.28%
	0.005	0.2209	0.2272	0.2705	0.4039	0.5236	0.5463	0.5823	0.6244	0.7882	2.0797	-	-					
4	X12	X1	X13	X9	X3	X7	X11	X2	X8	X4	-	-	-	X5	85.63%	92.97%	67.07%	90.27%
	0.2341	0.2349	0.2634	0.4029	0.5149	0.5803	0.5808	0.6114	0.8051	2.0465	-	-	-					
5	X13	X7	X9	X8	X1	X2	X11	X3	X4	-	-	-	-	X12	85.42%	92.95%	66.37%	90.14%
	0.028	0.0362	0.0553	0.1745	0.2251	0.3798	0.4627	0.7035	1.2827	-	-	-	-					
6	X7	X9	X8	X1	X2	X11	X3	X4	-	-	-	-	-	X13	85.43%	93.05%	66.16%	90.16%
	0.0043	0.0247	0.1043	0.2273	0.3371	0.4353	0.7518	1.1402	-	-	-	-	-					
7	X9	X8	X1	X2	X11	X3	X4	-	-	-	-	-	-	X7	85.44%	93.01%	66.29%	90.16%
	0.0269	0.1056	0.2358	0.321	0.44	0.7482	1.1091	-	-	-	-	-	-					
8	X8	X1	X11	X3	X2	X4	-	-	-	-	-	-	-	X9	85.07%	92.25%	66.90%	89.86%
	0.0743	0.2743	0.64	0.7276	0.86	1.3062	-	-	-	-	-	-	-					
9	X1	X11	X4	X3	X2	-	-	-	-	-	-	-	-	X8	84.51%	91.37%	67.14%	89.43%
	0.3384	0.6908	0.9448	0.9488	1.8042	-	-	-	-	-	-	-	-					
10	X2	X3	X11	X4	-	-	-	-	-	-	-	-	-	X1	83.83%	90.25%	67.58%	88.89%
	0.4342	0.4966	0.7918	0.8622	-	-	-	-	-	-	-	-	-					
11	X3	X11	X4	-	-	-	-	-	-	-	-	-	-	X2	83.39%	91.90%	61.84%	88.81%
	0.5734	0.7046	0.7698	-	-	-	-	-	-	-	-	-	-					
12	X11	X4	-	-	-	-	-	-	-	-	-	-	-	X3	78.89%	91.98%	45.76%	86.20%
	0.4001	1.2444	-	-	-	-	-	-	-	-	-	-	-					
13	X4	-	-	-	-	-	-	-	-	-	-	-	-	X11	74.72%	94.42%	24.83%	84.26%
	0.4577	-	-	-	-	-	-	-	-	-	-	-	-					

TABLE XII. SVM-RFE RESULTS FOR INPUT DIMENSION

Iteration	**Sorted Feature**													**Deleted Feature**	**Metric**			
															Acc.	**Sensv.**	**Specf.**	**F1**
1	X10	X11	X9	X2	X1	X6	X4	X3	X5	X7	X12	X13	X8	-	92.96%	93.17%	92.67%	93.75%
	1e-04	0.0016	0.0295	0.0562	0.1252	0.2396	1.3902	1.8188	2.1678	6.029	8.7485	9,607	44,569					
2	X11	X9	X2	X1	X6	X4	X3	X5	X7	X12	X13	X8	-	X10	92.96%	93.16%	92.69%	93.75%
	0.0016	0.0301	0.0553	0.1261	0.2389	1.3896	1.8242	2.1635	5.9656	8,696	9.5479	44,453	-					
3	X9	X2	X1	X6	X4	X3	X5	X7	X12	X13	X8	-	-	X11	92.94%	93.15%	92.65%	93.73%
	0.0276	0.0551	0.1263	0.2393	1.4249	1.8254	2.1934	5.9969	8.6975	9.5287	44,829	-	-					
4	X2	X1	X6	X4	X3	X5	X7	X12	X13	X8	-	-	-	X9	92.	92%93.22 %	92.54%	93.72%
	0.0595	0.1305	0.2563	1.2888	1.7414	2.3125	7.4043	9.8369	10,826	.165	-	-	-					
92.20%	X1	X6	X4	X5	X3	X7	X12	X13	X8	-	-	-	-	X2		93.26%	90.82%	93.13%
	0.1395	0.1801	0.9783	1.4124	1.5644	4.9187	7.4545	8.2854	35,699	-	-	-	-					
6	X6	X3	X4	X5	X7	X12	X13	X8	-	-	-	-	-	X1	92.44%	93.62%	90.91%	93.35%
	0.1786	1.2235	1.2727	1.2777	4.3821	7.1945	8.739	35,625	-	-	-	-	-					
7	X5	X3	X4	X7	X12	X13	X8	-	-	-	-	-	-	X6	92.33%	93.56%	90.72%	93.25%
	1.1361	1.2183	1.232	3.3407	6.2367	7.6352	30.6	-	-	-	-	-	-					
8	X7	X4	X3	X12	X13	X8	-	-	-	-	-	-	-	X5	91.16%	92.84%	88.97%	92.25%
	0.529	0.6628	1.2901	2.8709	3.9324	12.44	-	-	-	-	-	-	-					
9	X4	X3	X12	X13	X8	-	-	-	-	-	-	-	-	X7	90.74%	92.25%	88.76%	91.86%
	0.6686	1.2561	1.4523	2.1813	9.5629	-	-	-	-	-	-	-	-					
10	X3	X12	X13	X8	-	-	-	-	-	-	-	-	-	X4	88.90%	94.72%	81.29%	90.63%
	2E-09	1.1737	1.2721	8.7165	-	-	-	-	-	-	-	-	-					
11	X12	X13	X8	-	-	-	-	-	-	-	-	-	-	X3	88.90%	94.72%	81.29%	90.63%

	1.173	1.2714	8.7127	-	-	-	-	-	-	-	-	-	-					
12	X13	X8	-	-	-	-	-	-	-	-	-	-	-	X12	82.20%	96.59%	63.38%	86.01%
	1,271	2.1775	-	-	-	-	-	-	-	-	-	-	-					
13	X8	-	-	-	-	-	-	-	-	-	-	-	-	X13	63.97%	99.88%	36.51%	53.45%
	2.1798	-	-	-	-	-	-	-	-	-	-	-	-					

TABLE XIII. SVM-RFE RESULTS FOR UNDERSTANDING DIMENSION

Iteration	Sorted Feature													Deleted Feature	Metric			
															Acc.	Sensv.	Specf.	F1
1	X11	X4	X5	X9	X1	X8	X10	X6	X12	X13	X7	X2	X3	-	96.48%	90.44%		94.46%
	1e-11	2e-11	1e-10	3e-09	4e-09	8e-09	4e-05	4e-05	0.0076	0.0083	0.016	0.0167	28,752					
2	X5	X1	X4	X8	X9	X10	X6	X12	X13	X7	X2	X3	-	11	96.48%	90.44%	99.48%	94.46%
	1e-09	5e-09	1e-08	2e-08	3e-08	3e-05	4e-05	0.0076	0.0082	0.0159	0.0167	28,753	-					
3	X4	X1	X9	X8	X10	X6	X12	X13	X7	X2	X3	-	-	X5	96.48%	90.44%	99.48%	94.46%
	8e-11	2e-09	6e-09	2e-08	4e-05	4e-05	0.0076	0.0082	0.0159	0.0167	28,754	-	-					
4	X1	X9	X8	X10	X6	X12	X13	X7	X2	X3	-	-	-	X4	96.48%	90.44%	99.48%	94.46%
	4e-10	9e-10	6e-09	4e-05	4e-05	0.0076	0.0083	0.016	0.0167	28,752	-	-	-					
5	X9	X8	X10	X6	X12	X13	X7	X2	X3	-	-	-	-	X1	96.48%	90.44%	99.48%	94.46%
	8e-12	3e-10	4e-05	4e-05	0.0076	0.0083	0.016	0.0167	28,751	-	-	-	-					
6	X8	X10	X6	X12	X13	X7	X2	X3	-	-	-	-	-	X9	96.48%	90.44%	99.48%	94.46%
	2e-09	4e-05	4e-05	0.0076	0.0082	0.0159	0.0167	28,754	-	-	-	-	-					
7	X10	X6	X12	X13	X7	X2	X3	-	-	-	-	-	-	X8	96.48%	90.44%	99.48 %	94.46%
	4e-05	4e-05	0.0076	0.0083	0.016	0.0167	28,754	-	-	-	-	-	-					
8	X6	X12	X13	X2	X7	X3	-	-	-	-	-	-	-	X10	96.48%	90.43 %	99.48%	94.46%
	4e-05	0.0068	0.0069	0.0136	0.0142	28,754	-	-	-	-	-	-	-					
9	X12	X13	X7	X2	X3	-	-	-	-	-	-	-	-	X6	96.50%	90.49%	99.48%	94.48%
	0.0043	0.0044	0.009	0.0134	28,745	-	-	-	-	-	-	-	-					
10	X13	X7	X2	X3	-	-	-	-	-	-	-	-	-	X12	96.51%	90.52 %	99.48 %	94.50%
	6e-08	1e-06	0.009	28,716	-	-	-	-	-	-	-	-	-					
11	X7	X2	X3	-	-	-	-	-	-	-	-	-	-	X13	96.51%	90.52%	99.48%	94.50%
	1e-06	0.0089	28,709	-	-	-	-	-	-	-	-	-	-					
12	X2	X3	-	-	-	-	-	-	-	-	-	-	-	X7	96.51%	90.52%	99.48%	94.50%
	0.0086	28,753	-	-	-	-	-	-	-	-	-	-	-					
13	X3	-	-	-	-	-	-	-	-	-	-	-	-	X2	96.49%	90.48%	99.48%	94.48%
	28,754	-	-	-	-	-	-	-	-	-	-	-	-					