Jurnal Infotel
Vol. 17, No. 4, November 2025, pp. 872–890. doi:10.20895/infotel.v17i4.1409

Research Article

Identifying the Learning Style of Students Using Reinforcement Learning Techniques: An Approach to the Felder-Silverman Learning Style Model

Nesi Syafitri[1], Suzani Mohamad Samuri[2,*], and Yudhi Arta[3]

[1,2,3]Department of Software Engineering and Smart Technology, Faculty of Computing and Meta-Technology, Sultan Idris Education University, 35900, Malaysia
[1,3]Department of Informatics Engineering, Engineering Faculty, Universitas Islam Riau, 28284, Indonesia

*Corresponding email: suzani@meta.upsi.edu.my

Received: July 07, 2025; Revised: October 13, 2025; Accepted: November 24, 2025.

Abstract: To make education better, it is essential to know how students learn. This study presents a novel automated approach to identifying students' learning styles using artificial intelligence. This study applies the Felder–Silverman Learning Style Model (FSLSM) to investigate the' approaches of students to processing information, their input preferences (visual or verbal), and their conceptual understanding. To achieve this, a Q-Learning agent, a reinforcement learning approach, was trained to identify learning patterns directly from questionnaire data, thereby reducing the need for students to complete lengthy forms repeatedly. This study evaluated this methodology with data from 799 students from different colleges in Indonesia. The findings suggested that the model could reliably estimate learning styles in nearly all circumstances, especially with regard to how children receive and comprehend information. In a follow-up test with 50 students, the model obtained 100% accuracy, matching the results of the standard FSLSM exams. This shows that Q-learning can be an effective method of identifying individual learning styles. It opens new avenues for creating personalized and adaptive learning systems that tailor materials and procedures to each student's learning style. The system can be improved in the future to handle situations better where some learning styles are underrepresented.

Keywords: adaptive learning systems, felder–silverman learning style model, learning style identification, reinforcement learning, Q-learning

---

© Jurnal Infotel

1 Introduction

Technological developments have substantially changed the landscape of digital education, giving rise to innovative teaching methodologies and tailored learning platforms. Central to the efficiency of personalized learning is the identification and comprehension of distinct learning styles. These styles reflect the' different preferences and cognitive processes of learners for perceiving, processing and storing information. Learning styles are the cognitive, affective and psychological features of individuals who participate in the learning process [1].

Students differ in their learning styles; some like to engage with concepts, theories, and experiments, while others prefer practical problem-solving or reading-based tasks. As such, a "one-size-fits-all" teaching paradigm generally does not suit students' different needs, especially with respect to performance feedback and incentives [2]. Felder and Silverman (1988) note that students demonstrate varied learning styles [3]. If instructional practices do not accommodate these variances, it may lead to less engagement, less attention, and uneven academic performance. In such circumstances, students may be slow to understand the course material [4], feel disengaged or demotivated [5], and eventually struggle to maintain the academic pace [6].

According to Abu Baharin, the direct impact of learning styles on student' academic achievement shows that teachers should consider these styles as a foundation when planning instructional activities [4]. Teachers who are well-informed about their students' learning preferences and modify their teaching approaches accordingly are more likely to promote meaningful and compelling learning experiences [7]. In this context, personalized learning models offer a viable option that allows students to achieve their learning objectives more efficiently. Crucially, the first step in implementing such models is precisely identifying the learning type of each student [8]. To facilitate this, numerous learning-style models have been developed and widely used. These models are essential to help educators understand how children best receive, process, and retain information, helping to create more effective teaching and learning strategies. Numerous investigations have highlighted the importance of recognizing diverse learning styles as a foundation for creating effective educational interventions [9]. These theoretical frameworks try to classify learners based on various cognitive and behavioral features that differentiate one individual from another [10]. Several prominent learning style models have been introduced in the literature, including Fleming's VARK model—Visual, Auditory, Reading/Writing and Kinesthetic [11]; Kolb's Learning Style Inventory (LSI) [12]; the Felder-Silverman Learning Style Model [13]; and the Honey and Mumford Learning Styles Model [14]. Each model offers a unique perspective on how learners interact with educational content and provides a valuable framework for building more personalized and adaptive instructional strategies.

Learning styles can be assessed using three primary approaches: traditional, automated, and hybrid methods. During the last decade, substantial research has focused on the automated identification of learning styles. Various studies have explored this domain, employing diverse machine learning techniques that use supervised and unsupervised learning across multiple learning systems. Decision Trees (DT), Support Vector Machines (SVM), K-Nearest Neighbors (K-NN), Naïve Bayes, Random Forest, Regression, and Linear Discriminant Analysis (LDA) have been employed to ascertain learning styles according to the Felder-Silverman Learning Style Model (FSLSM) [6], [8], [15–17]. Neural networks have been employed to identify learners' learning styles according to Kolb's Model within an adaptive e-learning system [18]. The VARK Model has been utilized to determine the styles of learners' in an e-learning system through the Naïve Bayes algorithm [19]. The FSLSM Model has been used, in conjunction with the fuzzy C-means clustering technique, to identify learners' styles in e-learning systems [20–22]

Despite promising results, traditional supervised machine learning methods have some limitations. These include dependence on labeled data, vulnerability to irrelevant or noisy characteristics, increased computational costs, and challenges in dealing with dynamic and non-standard learning environments [23]. In addition, learning styles can evolve, making rigid categorization less effective in promoting continuous personal growth and learning.

In light of these problems, subsequent investigations have studied reinforcement learning (RL) as a more adaptive option. Reinforcement learning has demonstrated significant achievements in fields such as robotics, game play, and natural language processing, due to its ability to learn and adapt through ongoing feedback [24]. This flexibility highlights reinforcement learning as a valuable method for identifying learning styles, especially in evolving educational environments where students' behaviors and preferences can shift over time. By engaging in repeated interactions with the learning environment, reinforcement learning continuously develops its strategies to meet the changing requirements of each student [25].

Given the constraints of traditional machine learning methods and the evolving nature of personal learning preferences, reinforcement learning offers a more adaptable, unbiased, and context-sensitive approach. This study focuses primarily on developing a learning agent utilizing Q-Learning to identify students' learning styles in real time. This study focuses on three primary objectives: (1) to create a state–action–reward framework based on the ILS questionnaire that is appropriate for the implementation of Q-Learning, (2) to train the Q-Learning agent to categorize each dimension of the FSLSM through reward-based interactions with learner response patterns, and (3) to evaluate the model's ability to accurately and adaptively forecast learning styles.

2 Research Method

This study presents a reinforcement learning method that employs the Q-Learning algorithm to detect learning styles based on the Felder-Silverman Learning Style Model (FSLSM). The research approach is depicted in the diagram below, which displays the successive stages of the procedure. The study begins with the construction and design of the dataset, in which learner data are collected using the Index of Learning Styles (ILS) questionnaire, aligned with the FSLSM framework. The survey was distributed digitally using Google Forms. Once data collection was complete, the next step was preprocessing, which aimed to clean and prepare the dataset for effective use by the Q-learning model. The third phase involves developing a reinforcement learning model that leverages the Q-Learning algorithm to identify learning styles and evaluate their performance. The evaluation step carefully assesses the model's precision and ability to recognize learning styles from previously unexplored data. Figure 1 offers a comprehensive breakdown of each phase.

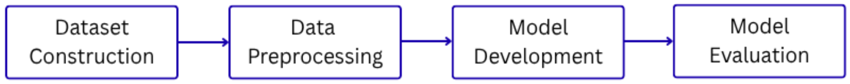

Figure 1: Stages of research for identifying learning styles using Q-learning.

2.1 Data Construction

The dataset for this study was collected from 814 undergraduate students enrolled in various universities in Indonesia. The participants represented a diverse range of academic disciplines, including education, engineering, computer science, and the social sciences, ensuring a wide coverage of learning behavior patterns. All participants, aged 18 to 22, were instructed to respond sincerely and reflect on their typical study habits and learning preferences. Data collection was conducted using the Index of Learning Styles (ILS) questionnaire, which is based on the Felder–Silverman Learning Style Model (FSLSM). The instrument was administered online through Google Forms and distributed through digital communication platforms such as WhatsApp and email to facilitate wide accessibility and participation. Participants were informed that there were no right or wrong answers, as the objective was to accurately identify their learning style tendencies on the four dimensions of the FSLSM. The data collection process lasted six months, from December 2024 to May 2025.

To ensure the validity and reliability of the collected data, the administration process was conducted in a controlled and supervised manner. The questionnaire was distributed to universities located in West Sumatra, Medan, Jakarta, and Riau through a network of lecturers who were professional colleagues of the research team. This trusted network enabled clear communication and effective coordination during data collection. The lecturers were instructed to guide and supervise their students as they completed the questionnaire during class sessions. Specifically, the survey was administered approximately 15 minutes before the end of the lecture to ensure that only active students participated in a monitored academic setting. This controlled procedure was designed to maintain the authenticity of the response and strengthen the validity and reliability of the dataset.

After data collection, the next stage in dataset construction involved a systematic data cleaning process. All responses were screened for completeness and any incomplete entries were removed, resulting in a final valid sample of 799 participants. Subsequently, the cleaned data were labeled and organized by mapping each response to its corresponding FSLSM dimension. The structured dataset was then stored in *.CSV format to facilitate subsequent computational modeling and analysis. Table 1 provides an overview of the metadata of the dataset.

2.2 Data Preprocessing

After data collection, a systematic data cleaning process was conducted to ensure the integrity of the dataset. Incomplete or inconsistent responses were removed, resulting in 799 valid samples. The learning style label of each participant was then derived from the four dimensions of the FSLSM. The 44-item ILS instrument was converted to binary format, with “A” coded as zero and “B” as 1. The responses were grouped into sets of 11 per dimension—active–reflective, sensing–intuitive, visual–verbal, and sequential–global—to

Table 1: Metadata specifications of the dataset

Feature	Description	Value
Q1	Answer to question 1	Optional answer (A or B)
Q2	Answer to question 2	Optional answer (A or B)
...	...	...
Q44	Answer to question 44	Optional answer (A or B)
Label1	Label of Processing Dimension	A = active, B= reflective)
Label2	Label of Perception Dimension	A = sensing, B= intuitive)
Label3	Label of Input Dimension	A = visual, B= verbal)
Label4	Label of Understanding Dimension	A = sequential, B= global)

generate the corresponding binary labels. A detailed mapping of the item numbers to the dimensions of FSLSM is presented in Table 2.

Table 2: Mapping of questionnaire items

Dimension	Option	Number of Items	Items Number
Processing	Active-Reflective	11	1, 5, 9,13, 17, 21, 25, 29, 33, 37, 41
Perception	Sensing-Intuitive	11	2, 6, 10, 14, 18, 22, 26,30, 34, 38, 42
Input	Visual-Verbal	11	3, 7, 11, 15, 19, 23, 27, 31, 35, 39, 43
Understanding	Sequential-Global	11	4, 8, 12,16, 20, 24, 28, 32, 36, 40, 44

The ILS questionnaire is grounded in the theoretical framework of FSLSM, which conceptualizes learning preferences in four distinct dimensions. The Processing dimension (Active–Reflective) captures how learners engage with information, either through direct participation or through introspective reflection. The Perception dimension (Sensing–Intuitive) reflects how individuals acquire information, whether through concrete, detail-oriented observation or through abstract conceptualization. The Input dimension (Visual–Verbal) refers to the preferred mode of receiving information, such as visual imagery or verbal explanation. Lastly, the Understanding dimension (Sequential–Global) represents how learners organize and synthesize knowledge, progressing step-by-step or through holistic integration. These dimensions collectively form the conceptual basis for the subsequent data transformation and model training.

Building upon this theoretical foundation, the dataset was reformulated for use in a reinforcement learning (RL) context, specifically within a Q-learning framework. Unlike conventional classification methods that treat questionnaire responses as independent input attributes, the RL formulation represents learning style identification as a sequence of state–action interactions. In this framework, a state represents the learner's condition within a particular FSLSM dimension. At the same time, an action denotes the model's decision to classify the learner into one of two possible categories (e.g., Active or Reflective). Directly constructing states from all 11 binary responses per dimension would yield 211 or 2048 possible state combinations—an excessively large and sparse state space that could hinder learning efficiency.

To address this issue, the state representation was simplified by aggregating the number of "0" and "1" responses across the 11 items within each dimension. This aggregation reduced the state–action space to 12 possible combinations per dimension while retaining the essential behavioral information captured by the ILS responses. The resulting abstraction not only improved computational tractability and learning stability, but also preserved interpretability in accordance with FSLSM theory. Table 3 summarizes the configurations of the state-action pair used in the Q-learning process.

This compact state formulation is closely aligned with the theoretical premise of the FSLSM, where the cumulative pattern of responses within each dimension better reflects the dominant preference of the learner than the variations at the individual item-level. Consequently, the data processing design ensured both theoretical coherence and computational efficiency, allowing the reinforcement learning model to effectively identify and generalize learning style patterns across diverse student profiles

Table 3: The State-Action pair configurations

State		**Action**
Number of 0s	**Number of 1s**	
0	11	Option 2
1	10	Option 2
2	9	Option 2
3	8	Option 2
4	7	Option 2
5	6	Option 2
6	5	Option 1
7	4	Option 1
8	3	Option 1
9	2	Option 1
10	1	Option 1
11	0	Option 1

2.3 Model Development

Reinforcement learning is a branch of machine learning that differs from supervised and unsupervised learning. Instead of relying on labeled data, it enables a model to learn by interacting with its environment and adapting to different contexts. Reinforcement learning (RL) is an area of machine learning that is concerned with what actions an agent,i.e., an intelligent program, should take in an environment to maximize cumulative reward [24]. The key components of RL include the agent, environment, policy, reward signal, value function, and environment model. Reinforcement Learning (RL) provides various models for identifying the' learning styles of the students. This research employs the Q-learning algorithm to develop FSLSM-based prediction models for the identification of learning styles. The Q-learning algorithm is classified as a model-free approach. Model-free algorithms operate without needing a predefined model of the environment, instead learning the optimal policy or value function directly through interaction with the environment. These approaches can be further categorized into value-based and policy-based methods. Value-based techniques aim to learn value functions that estimate expected rewards associated with specific states or pairs of states of action [24]. The Q-value in Q-learning methods is updated using Eq. (1).

$$Q(s,a) = Q(s,a) + \alpha \left[r + \gamma \max_{a'} Q(s',a') - Q(s,a) \right] \tag{1}$$

Where α is the learning rate, γ is the discount factor, r is the reward, and s' is the next state. $Q(s,a)$ is the Q-value for the current state–action pair, and $Q(s',a')$ is the Q-value for the next state–action pair. The process of how the Q-learning algorithm identifies student' learning styles is illustrated in Figure 2.

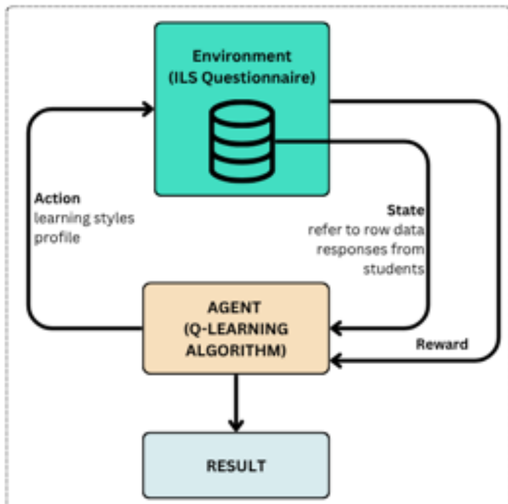

Figure 2: Overview of the Q-learning model development.

This environment may include the ILS questionnaire and the algorithms that support it. The questionnaire is part of the environment by providing input for the agent's learning process. The agent, represented by an algorithm, identifies or predicts the' learning styles of students based on their responses to the ILS questionnaire. By interacting with the environment (the questionnaire), the agent continuously gathers knowledge and enhances its decision-making capabilities over time.

The state represents the students' responses to specific questions in the ILS questionnaire. This study uses an existing dataset to determine the initial state of the agent. Actions refer to the agent's systematic processes for analyzing and interpreting questionnaire responses to identify student' learning styles. The reward serves as feedback on the accuracy of the identification.

2.4 Model Evaluation

The performance of the Q-learning models was rigorously evaluated using multiple quantitative metrics, including accuracy, cumulative reward, precision, recall (sensitivity), specificity, and F1-score. To ensure robustness and reduce overfitting, a cross-validation approach was employed, with the training data partitioned into multiple subsets. The model was trained iteratively on selected folds and validated on the remaining ones, providing a reliable estimate of its generalizability in classifying the' learning styles of students according to the FSLSM framework.

To address the significant class imbalance across several FSLSM dimensions, class weighting was applied during model training. This technique adjusts the loss function to assign higher weights to minority classes, thereby reducing the model's bias toward majority classes and promoting more balanced learning across categories. Each evaluation metric provides a distinct perspective on the performance of the model. Accuracy measures the general accuracy of the predictions, whereas cumulative reward captures the agent's ability to make optimal decisions throughout the learning process. Precision assesses the proportion of correctly identified positive cases among all positive predictions, while recall (sensitivity) measures the model's ability to identify true positives. Specificity quantifies the accuracy of recognizing true negatives, reflecting the model's ability to avoid false positives. The F1-score, computed as the harmonic mean of precision and recall, provides a balanced indicator that accounts for both types of classification errors.

All performance metrics were calculated from the confusion matrix, which consists of True Positives (TP), True Negatives (TN), False Positives (FP) and False Negatives (FN) [26]. The mathematical formulations for accuracy, precision, sensitivity, specificity, F1-Score and cumulative reward are summarized below, serving as the foundation for the quantitative evaluation of the effectiveness of the Q-learning model.

$$\text{Accuracy} = \frac{\text{TP} + \text{TN}}{\text{TP} + \text{FP} + \text{TN} + \text{FN}} \tag{2}$$

$$\text{Precision} = \frac{\text{TP}}{\text{TP} + \text{FP}} \tag{3}$$

$$\text{Sensitivity/Recall} = \frac{\text{TP}}{\text{TP} + \text{TN}} \tag{4}$$

$$\text{Specificity} = \frac{\text{TN}}{\text{TN} + \text{FP}} \tag{5}$$

$$\text{F1-Score} = \frac{2 \times \text{Recall} \times \text{Precision}}{\text{Recall} + \text{Precision}} \tag{6}$$

$$G_t = \sum_{k=0}^{\infty} \gamma^k R_{t+k} \tag{7}$$

Where G_t is the cumulative return starting at time t, R_{t+k} is the reward received at time $t + k$, and γ is the discount factor ($0 \leq \gamma \leq 1$) that determines how critical future rewards are. Stable and high rewards indicate that agents understand the environment strongly [24].

3 Results

3.1 Data Collection and Preprocessing

This study used the 44-questions ILS questionnaire based on the FSLSM model. The questionnaire was distributed to 799 learners via an online form. The results are presented in Table 4.

Table 4: The results of questionnaires

ID	Q1	Q2	Q3	Q4	Q5	Q6	Q7	Q8	Q9	Q10	...	Q44
0001	A	B	B	A	B	A	A	A	B	A	...	A
0002	A	B	A	A	B	B	A	A	A	A	...	B
0003	A	A	A	B	B	A	A	B	B	A	...	B
0004	A	B	B	A	B	A	B	A	B	A	...	A
...	...	...	...	...	...	...	...	...	...	...	...	
0799	A	A	B	B	B	A	B	A	B	A	...	A

After the data collection phase, a preprocessing step was performed to confirm its suitability for further analysis. This process involved transforming the labels 'A' and 'B' into binary values (0 and 1). The students' learning styles were subsequently classified based on their responses, adhering to the FSLSM framework for each dimension. The results of the questionnaire preprocessing are shown in Table 5.

Table 5: Preprocessing results

ID	Question							Dimension			
	Q1	Q2	Q3	Q4	Q5	...	Q44	Processing	Perception	Input	Understanding
0001	0	1	1	0	1	...	0	0	0	0	0
0002	0	1	0	0	1	...	1	0	1	1	0
0003	0	0	0	1	1	...	1	1	0	0	1
0004	0	1	1	0	1	...	0	0	0	1	0
0005	0	1	1	1	1	...	0	0	1	0	0
...	...	...	...	...	...	...	...	...	...	...	...
0799	0	0	1	1	1	...	0	1	1	1	1

In the FSLSM framework, there are 16 possible learning style profiles (24), as it comprises four dimensions: Processing (active or reflective), Perception (sensing or intuitive), Input (visual or verbal), and Understanding (sequential or global). Figure 3 illustrates the distribution of the 799 sample data in these 16 combinations of learning styles.

Meanwhile, Table 6 presents a summary of the dataset, which covers 799 learners in the four FSLSM categories.

3.2 Model Development using Q-Learning Algorithm

The goal of this phase is to develop a prediction model for the identification of learning-styles using the FSLSM and reinforcement learning. The process is divided into two main phases: training and testing.

During the training phase, Q-learning agents learn to recognize the' learning styles of the students from their responses to the ILS questionnaire. The process begins with the agent observing the initial state of the student, which represents the student's current learning characteristics. This study uses a dataset derived from responses to the ILS questionnaire to define the starting state of the agent.

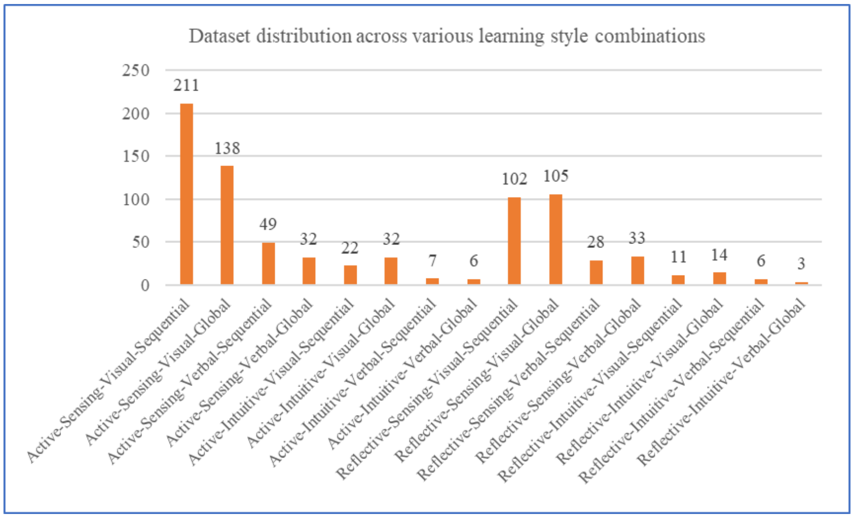

Figure 3: Distribution of the dataset in various combinations of learning styles.

Table 6: Summary of the dataset by FSLSM categories

FSLSM Dimension	Category	Number of learners
Processing	Active	497 (62.20%)
	Reflective	302 (37.80%)
Perception	Sensing	698 (87.35%)
	Intuitive	101 (12.65%)
Input	Visual	635 (79.47%)
	Verbal	164 (20.53%)
Understanding	Sequential	436 (54.56%)
	Global	363 (45.44%)

Each agent was designed to select between two possible actions, representing the binary learning style labels (e.g. Active versus Reflective for the Processing dimension). Action selection was guided by the -greedy strategy, which balances the exploration of new alternatives with the exploitation of the best known choices. After selecting an action, the agent received a class-weighted reward, where higher values were assigned to correctly classified minority classes and lower values to majority classes. The Q-values were then updated following the Q-learning rule. Through this iterative process, the agents progressively refined their policies, leading to more accurate identification of the' learning styles of students over time.

After the training phase, the Q-learning agent determines the optimal Q-values associated with each state–action pair. By analyzing these Q-values relative to the observed state of the student, the agent can identify the learning style that is most suited to that individual. For example, if the Q-value for the action "Visual learning style" is highest in a specific state, the agent infers that the student prefers a visual learning style.

The Q-learning model was evaluated through two experiments: one using a subset of 409 samples and another using the complete dataset of 799 samples. Table 7 summarizes the Q-values from training in both datasets for the processing dimension. The results indicate that the agent trained on the larger dataset achieved non–zero Q-values in all states, suggesting stronger learning and improved adaptability.

Table 7: Q-value for processing dimension

State		Q-value (409 Dataset)		Q-value (799 Dataset)	
Number of Feature 0s	**Number of Feature 1s**	**Active**	**Reflective**	**Active**	**Reflective**
0	11	7.945	9.992	3.923	9.999
1	10	7.935	9.982	8.960	10.000
2	9	7.904	9.928	8.999	10.000
3	8	7.860	9.740	9.000	10.000
4	7	7.943	9.964	9.000	10.000
5	6	7.856	9.967	9.000	10.000
6	5	0	0	10.000	9.000
7	4	0	0	10.000	9.000
8	3	0	0	10.000	9.000
9	2	0	0	10.000	9.000
10	1	0	0	10.000	8.999
11	0	9.827	7.895	10.000	8.966

Table 8 Table 9 and Table 10 present the Q-value results for the perception, Input, and understanding dimensions.

Table 8: Q-value for perception dimension

State		Q-value (409 Dataset)		Q-value (799 Dataset)	
Number of Feature 0s	**Number of Feature 1s**	**Sensing**	**Intuitive**	**Sensing**	**Intuitive**
0	11	7.930	9.967	0.008	0.008
1	10	7.808	9.723	7.035	9.999
2	9	7.879	9.946	8.836	10.000
3	8	7.922	9.925	8.999	10.000
4	7	7.909	9.916	8.999	10.000
5	6	7.877	9.672	9.000	10.000
6	5	0	0	10.000	9.000
7	4	0	0	10.000	9.000
8	3	0	0	10.000	9.000
9	2	0	0	10.000	9.000
10	1	0	0	10.000	9.000
11	0	9.999	7.934	10.000	9.000

Based on this experiment, it can be concluded that as the agent interacts more with the environment, its ability to adapt to diverse conditions improves. This is demonstrated in the 799-dataset experiment, where the Q-values for all states are non-zero. Consequently, the agent becomes more effective in identifying the most suitable learning style for each student based on their observed state. After training the agent, the testing phase is carried

Table 9: Q-value for input dimension

State		Q-value (409 Dataset)		Q-value (799 Dataset)	
Number of Feature 0s	**Number of Feature 1s**	**Visual**	**Verbal**	**Visual**	**Verbal**
0	11	7.890	9.757	0.008	0.007
1	10	7.903	9.751	8.769	10.000
2	9	7.902	9.940	8.999	10.000
3	8	7.913	9.895	9.000	10.000
4	7	7.867	9.972	9.000	10.000
5	6	7.922	9.985	9.000	10.000
6	5	9.652	3.486	10.000	9.000
7	4	9.678	2.911	10.000	9.000
8	3	9.647	4.198	10.000	9.000
9	2	0	0	10.000	9.000
10	1	9.892	3.650	10.000	9.000
11	0	9.760	7.894	10.000	9.000

Table 10: Q-value for dimension understanding

State		Q-value (409 Dataset)		Q-value (799 Dataset)	
Number of Feature 0s	**Number of Feature 1s**	**Sequential**	**Global**	**Sequential**	**Global**
0	11	7.928	9.840	6.271	9.999
1	10	7.934	9.880	8.998	10.000
2	9	7.913	9.962	9.000	10.000
3	8	7.924	9.984	9.000	10.000
4	7	7.958	9.958	9.000	10.000
5	6	7.823	9.891	9.000	10.000
6	5	9.997	4.916	10.000	9.000
7	4	9.592	4.078	10.000	9.000
8	3	9.999	5.186	10.000	9.000
9	2	0	0	10.000	9.000
10	1	9.660	1.357	10.000	9.000
11	0	9.910	7.855	10.000	9.000

out to evaluate how effectively the Q-learning agent can automatically predict student' learning styles. The performance metrics reported in this section reflect the application of class-weighting during training to mitigate class imbalance. This adjustment ensured that minority classes contributed more substantially to the loss, leading to more balanced predictions in the learning style categories. Consequently, the reported results more accurately represent the model's ability to generalize across both dominant and underrepresented learning style dimensions. The testing phase assesses the agent's performance in environments not encountered during training. The results of cross-validation with 3, 5, and 10 folds (k=3, k=5, k=10) are presented in Table 11, which show the performance of the agent during the testing.

Table 11 summarizes the performance of the Q-learning agent's in all FSLSM dimensions under 3, 5 and 10-fold cross-validation using a data set of 799 samples. The model consistently achieved near-perfect results, with accuracy, precision, recall, and specificity

Table 11: Agent's performance result for each dimension in the 799 dataset

K-Fold	Dimension	Cumulative Reward	Reward Max	Avg Accuracy (%)	Avg Precision (%)	Avg Recall / Sensitivity (%)	Avg Specificity (%)	Avg F1-Score (%)
3-Fold	**Processing** (Active/Reflective)	245759.22	500.94	100	100	100	100	100
	Perception (Sensing/Intuitive)	115417.19	235.29	100	100	100	100	100
	Input (Visual/Verbal)	170497.01	347.57	100	100	100	100	100
	Understanding (Sequential/Global)	259186.70	528.21	99.88	100	100	100	100
5-Fold	**Processing** (Active/Reflective)	295013.81	601.13	100	100	100	100	100
	Perception (Sensing/Intuitive)	138490.43	282.34	100	100	100	100	100
	Input (Visual/Verbal)	204644.082	417.08	100	100	100	100	100
	Understanding (Sequential/Global)	311054.28	633.86	99.87	100	99.72	100	99.86
10-Fold	**Processing** (Active/Reflective)	331789.43	676.19	99.88	100	99.67	100	99.83
	Perception (Sensing/Intuitive)	155869.09	317.64	100	100	100	100	100
	Input (Visual/Verbal)	230195.01	469.17	100	100	100	100	100
	Understanding (Sequential/Global)	349881.01	713.09	99.88	100	99.72	100	99.86

values exceeding 99.8% in all configurations. The high and stable F1-scores confirm that the model maintained a strong balance between precision and recall across different validation folds. Among the dimensions of FSLSM, the perception dimensions (Sensing–Intuitive) and the input dimensions (Visual–Verbal) showed the most stable performance, suggesting that the formulation of state action and the integration of class-weight effectively captured the distinct behavioral tendencies in these learning categories. Minor fluctuations observed in the Processing and Understanding dimensions, although negligible, reflect the inherent cognitive overlap within these styles rather than the model instability. The near-perfect accuracy observed across folds does not indicate overfitting or data leakage, but rather reflects the theoretical soundness of the reinforcement learning framework employed. The use of a compact 12-state representation per dimension allowed the Q-learning agent to fully explore the state–action space and achieve optimal policy convergence. Furthermore, integrating class weighting effectively mitigated class imbalance by penalizing the majority classes and improving decision boundaries for underrepresented learning styles. The steadily increasing cumulative and maximum rewards further support the stable learning dynamics of the agent. In general, these findings demonstrate that the Q-learning model

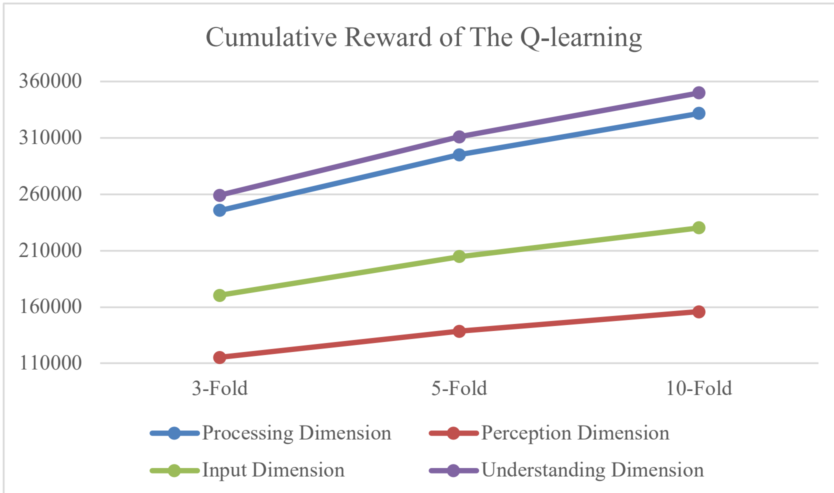

Figure 4: Cumulative reward of the Q-learning agent.

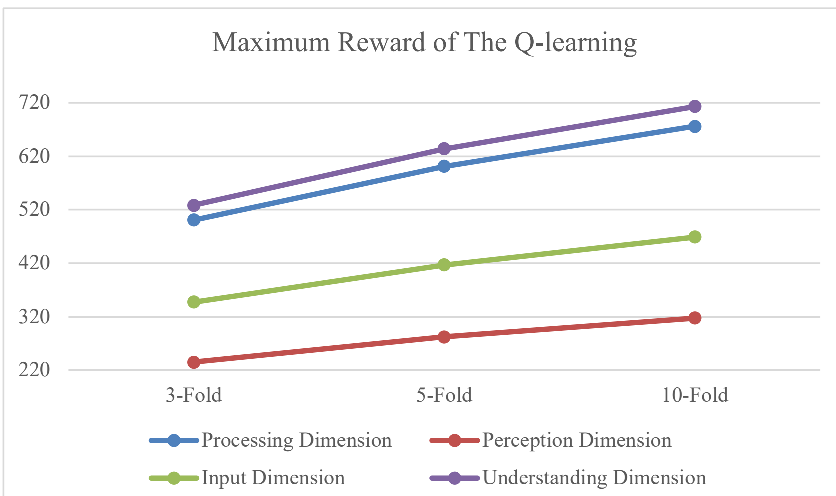

Figure 5: Maximum reward of the Q-learning agent.

generalizes effectively across diverse dimensions of the learning style, combining theoretical coherence with empirical robustness in modeling FSLSM-based learning behaviors. Beyond standard evaluation indicators such as accuracy, precision, recall, and F1-score, the Q-learning model was further evaluated using specific reinforcement learning metrics. These included the cumulative reward, which quantifies the total reward accumulated in all training episodes, and the maximum reward, which reflects the highest reward achieved within a single episode. Together, these measures provide deeper insight into the agent's learning efficiency and decision-making stability. The results of these evaluations are presented in Figure 4 and Figure 5.

As shown in Figure 4 and Figure 5, both the cumulative and maximum rewards increase steadily from 3-fold to 10-fold cross-validation, indicating consistent improvement in the Q-Learning agent's learning performance. The Understanding and Processing dimensions achieve the highest reward values, reflecting more stable and optimal policy convergence. In contrast, the Perception dimension yields the lowest rewards, suggesting greater variability and less distinct feedback signals during learning. Overall, the upward trends across folds confirm that extended training exposure enhances the agent's capacity to optimize Q-values and generalize effectively across learning style dimensions.

These findings collectively highlight the robustness and theoretical consistency of the Q-learning framework in identifying FSLSM-based learning styles. The model's near-perfect performance across all folds and dimensions demonstrates its ability to generalize effectively without overfitting, attributed to the balanced state–action design and class-weighting mechanism. The strong alignment between cumulative rewards, maximum rewards, and predictive accuracy further confirms that the agent's policy optimization reflects genuine learning behavior rather than artifact-driven performance. Consequently, the results validate the suitability of reinforcement learning for modeling cognitive tendencies in personalized learning environments.

3.3 Model Evaluation

At this stage, the Q-learning model was evaluated using data from 50 students. This evaluation compares the observed accuracy with the expected accuracy using the Kappa statistic. The Kappa statistic measures agreement by comparing observed and expected accuracy (random chance). In this study, Cohen's Kappa [?] is used to interpret the Kappa statistic (κ) based on the following scale: 0.81–1.00 (almost perfect), 0.61–0.80 (substantial), 0.41–0.60 (moderate), 0.21–0.40 (fair), 0.01–0.20 (slight), and less than 0.0 (poor or no agreement). Cohen's Kappa coefficients were measured using Equation (8).

$$\kappa = \frac{P_0 - P_e}{1 - P_e} \tag{6}$$

Table 12 presents the results of the testing of the Q-learning model for the identification of learning styles in 50 students.

The Q-learning model demonstrated strong performance in accurately identifying students' learning style combinations, closely aligning with the expected results based on the FSLSM framework. A kappa coefficient of 1 indicates outstanding precision, significant consistency, and reliability in identifying learning forces across the dimensions of the FSLSM, thereby further substantiating the model's efficacy. These results emphasize the robust validity of Q-Learning as an automated method for identifying learning styles.

Although the 50-student validation yielded a perfect accuracy rate, this result should be interpreted in the context of the Q-Learning framework rather than as evidence of overfitting. The state–action design adopted in this study enables the agent to exhaustively explore all 12 possible state combinations, leading to stable Q-value convergence and consistent policy generation. Once convergence is achieved, the agent's performance becomes deterministic, explaining the 100% correspondence with the FSLSM-based ground truth. Nevertheless, the dataset's limited size and homogeneous characteristics constrain the model's generalizability. Future experiments involving larger, more diverse datasets

Table 12: Evaluation result

Combination of Learning Styles	Number of Participants	Number identified by		Accuracy (%)
		the FSLSM model	the Q-Learning model	
Active–Sensing–Visual–Sequential	11	11	11	100%
Active–Sensing–Visual–Global	9	9	9	100%
Active–Sensing–Verbal–Sequential	4	4	4	100%
Active–Sensing–Verbal–Global	3	3	3	100%
Active–Intuitive–Visual–Sequential	2	2	2	100%
Active–Intuitive–Visual–Global	4	4	4	100%
Active–Intuitive–Verbal–Sequential	4	4	4	100%
Active–Intuitive–Verbal–Global	2	2	2	100%
Reflective–Sensing–Visual–Sequential	1	1	1	100%
Reflective–Sensing–Visual–Global	1	1	1	100%
Reflective–Sensing–Verbal–Global	1	1	1	100%
Reflective–Intuitive–Visual–Sequential	2	2	2	100%
Reflective–Intuitive–Visual–Global	3	3	3	100%
Reflective–Intuitive–Verbal–Sequential	3	3	3	100%

and stochastic environments are essential to validate the scalability and adaptability of the proposed approach.

4 Discussion

This study developed and evaluated a Q-learning agent to automatically identify learning styles based on the Felder–Silverman Learning Style Model (FSLSM) using ILS questionnaire data. The design of the state–action–reward structure was a critical component, with each FSLSM dimension represented as a 12-state model corresponding to the binary distribution of responses. This representation enabled the agent to systematically explore all possible state transitions and optimize its decision policy through reward feedback. The framework not only simplified the computational process but also established a meaningful link between the agent's internal learning process and students' behavioral patterns, providing interpretability often lacking in other machine learning approaches.

Unlike conventional classification models that rely on static decision boundaries, Q-learning introduces a dynamic and adaptive learning mechanism. Traditional classifiers, such as Logistic Regression or Support Vector Machines, depend on large, balanced datasets and remain fixed after training, limiting their ability to adapt to new or changing data. In contrast, Q-learning iteratively updates its decision policy based on experience, optimizing the Q-value function using accumulated rewards rather than explicit parameter fitting. This allows the model to adapt even under data scarcity or imbalance and to capture probabilistic patterns inherent in human learning behavior. Therefore, although the classification task in this study was essentially static, the application of Q-learning remains theoretically justified, as it mirrors real educational environments in which decision-making evolves based on learner interaction and feedback.

The results confirm that the Q-learning agent achieved stable and interpretable convergence across all FSLSM dimensions. The generated Q-tables displayed distinct value distributions, with higher Q-values corresponding to dominant response types, especially in the Processing and Understanding dimensions. As the number of samples increased, Q-value convergence became smoother, indicating improved learning stability. Performance metrics across all folds demonstrated near-perfect accuracy, with the Perception and Input dimensions consistently yielding flawless outcomes. Minor fluctuations observed in the Processing and Understanding dimensions were not signs of instability but reflected natural variability in student responses. The agent's success can thus be attributed to complete state–action exploration and optimized reward-driven learning rather than overfitting or data leakage, as evidenced by consistent cumulative reward growth and stable validation outcomes. Although the model achieved perfect accuracy and a Kappa coefficient of 1.00 across all 16 learning style combinations, these results should be interpreted in light of the Q-learning framework's deterministic nature. The fully explored state–action space ensures internal stability but limits generalization to more diverse and stochastic environments. The current dataset, while sufficient for demonstrating feasibility, remains relatively small and homogeneous, potentially constraining the scalability of the approach. Future research should therefore extend this work by employing larger, more diverse datasets, introducing probabilistic state transitions, and testing the framework in dynamic, context-aware learning scenarios. In addition, comparative analyses with traditional classifiers such as Logistic Regression, Random Forest, and Support Vector Machines are essential to validate Q-learning's advantages in terms of robustness and adaptability. Ultimately, this study demonstrates that reinforcement learning provides a powerful alternative for modeling adaptive learning behavior, offering both theoretical and practical foundations for the development of intelligent educational systems that evolve alongside learner diversity and behavior.

5 Conclusion

This study demonstrates the effectiveness of reinforcement learning, particularly the Q-Learning algorithm, in accurately identifying students' learning styles according to the Felder–Silverman Learning Style Model (FSLSM). By designing a compact, fully explored state–action representation, the Q-Learning model successfully achieved Q-value convergence and stable policy formation across all FSLSM dimensions. The model's performance—achieving 100% accuracy and a perfect Kappa coefficient-indicates a strong alignment between the learned policies and the expected FSLSM outcomes. These results highlight the potential of reinforcement learning as a viable computational approach for modeling adaptive and personalized learning behaviors, even with relatively small datasets. However, the near-perfect accuracy obtained in both training and validation suggests that the model's performance may be influenced by the deterministic structure of the state–action space rather than true generalizability. The findings, therefore, underscore the need for further experimentation with larger, more diverse datasets and for integrating stochastic or dynamic state representations to better capture the complexity of real-world learning behaviors. Future work will also explore hybrid frameworks that combine reinforcement learning with deep or transformer-based models to enhance interpretability, scalability, and adaptability in dynamic educational environments.

Acknowledgments

The authors would like to express their sincere appreciation to Sultan Idris Education University (UPSI) and Universitas Islam Riau (UIR) for their generous support, resources, and academic guidance throughout this research. Their contributions were essential to the successful completion of this study.

References

[1] J. Keefe, "Learning style: An overview," *NASSP's Student learning styles: Diagnosing and proscribing programs/National Association of Secondary School Principles*, 1979.

[2] F. Demir, C. Bruce-Kotey, and F. Alenezi, "User experience matters: does one size fit all? evaluation of learning management systems," *Technology, Knowledge and Learning*, vol. 27, no. 1, pp. 49–67, 2022.

[3] R. M. Felder, L. K. Silverman, *et al.*, "Learning and teaching styles in engineering education," *Engineering education*, vol. 78, no. 7, pp. 674–681, 1988.

[4] A. Hidayat, K. Adi, and B. Surarso, "Detection of student learning styles using the index of learning style," in *1st International Conference Of Education, Social And Humanities (INCESH 2021)*, pp. 34–38, Atlantis Press, 2021.

[5] J. Bernard, T.-W. Chang, E. Popescu, and S. Graf, "Learning style identifier: Improving the precision of learning style identification through computational intelligence algorithms," *Expert Systems with Applications*, vol. 75, pp. 94–108, 2017.

[6] M. Hasibuan, R. Abdul Aziz, D. Dewi, T. Kurniawan, and N. Syafira, "Recommendation model for learning material using the felder silverman learning style approach," *HighTech and Innovation Journal*, vol. 4, no. 4, pp. 811–820, 2023.

[7] C. S. Claxton and P. H. Murrell, *Learning Styles: Implications for Improving Educational Practices. ASHE-ERIC Higher Education Report No. 4, 1987.* ERIC, 1987.

[8] W. Waam and P. Hks, "Identifying the learning style of students using machine learning techniques: An approach of felder silverman learning style model (fslsm)," *Asian Journal of Research in Computer Science*, vol. 17, no. 3, pp. 15–37, 2024.

[9] M. S. Hasibuan and O. W. Purbo, "Learning motivation increased due to a relaxed assessment in a competitive e-learning environment," in *Proceeding of Electrical Engineering, Computer Science and Informatics*, vol. 5, Nov. 2018.

[10] A. M. Altamimi, M. Azzeh, and M. Albashayreh, "Predicting students' learning styles using regression techniques," *arXiv preprint arXiv:2209.12691*, 2022.

[11] N. D. Fleming and C. Mills, "Not another inventory, rather a catalyst for reflection," *To improve the academy*, vol. 11, no. 1, pp. 137–155, 1992.

[12] D. A. Kolb, *Experiential learning: Experience as the source of learning and development*. FT press, 2014.

[13] R. M. Felder, L. K. Silverman, *et al.*, "Learning and teaching styles in engineering education," *Engineering education*, vol. 78, no. 7, pp. 674–681, 1988.

[14] P. Honey and A. Mumford, "Styles of learning," *Gower Handb. Manag. Dev*, vol. 101, pp. 101–111, 1994.

[15] M. S. Hasibuan and R. A. Aziz, "Detection of learning styles with prior knowledge data using the svm, k-nn and naã¯ ve bayes algorithms," *Jurnal Infotel*, vol. 14, no. 3, pp. 209–213, 2022.

[16] M. Jebbari, B. Cherradi, W. Moutaouakil, O. El Gannour, S. Hamida, and A. Raihani, "Prediction and classification of learning styles using machine learning approach," in *2023 3rd International Conference on Innovative Research in Applied Science, Engineering and Technology (IRASET)*, pp. 1–7, IEEE, 2023.

[17] R. R. Maaliw III, "Classification of learning styles in virtual learning environment using data mining: A basis for adaptive course design," *International Research Journal of Engineering and Technology (IRJET)*, vol. 3, no. 7, pp. 56–61, 2016.

[18] C. Waladi, M. Khaldi, and M. L. Sefian, "Machine learning approach for an adaptive e-learning system based on kolb learning styles.," *Int. J. Emerg. Technol. Learn.*, vol. 18, no. 12, pp. 4–15, 2023.

[19] S. Sharuni, R. Dhana Lakshmi, and A. Murugappan, "Automatic detection of learner's learning style," in *International Conference on Data, Engineering and Applications*, pp. 77–87, Springer, 2022.

[20] I. Azzi, A. Radouane, L. Laaouina, A. Jeghal, A. Yahyaouy, and H. Tairi, "Fuzzy classification approach to select learning objects based on learning styles in intelligent e-learning systems," in *Informatics*, vol. 11, p. 29, MDPI, 2024.

[21] O. El Aissaoui, Y. El Madani El Alami, L. Oughdir, and Y. El Allioui, "A hybrid machine learning approach to predict learning styles in adaptive e-learning system," in *International conference on advanced intelligent systems for sustainable development*, pp. 772–786, Springer, 2018.

[22] B. A. Muhammad, W. Jianping, G. Gao, C. Zhao, Q. Li, X. Zuo, and Y. Lv, "A fuzzy c-means algorithm to detect learning styles in online learning environment," *Journal of Networking and Network Applications*, vol. 4, no. 1, pp. 39–47, 2024.

[23] R. S. Sutton, A. G. Barto, *et al.*, *Reinforcement learning: An introduction*, vol. 1. MIT press Cambridge, 1998.

[24] M. Boussakssou, B. Hssina, and M. Erittali, "Towards an adaptive e-learning system based on q-learning algorithm," *Procedia Computer Science*, vol. 170, pp. 1198–1203, 2020.

[25] K. Fahd, S. Venkatraman, S. J. Miah, and K. Ahmed, "Application of machine learning in higher education to assess student academic performance, at-risk, and attrition: A meta-analysis of literature," *Education and Information Technologies*, vol. 27, no. 3, pp. 3743–3775, 2022.

[26] J. L. Fleiss and J. Cohen, "The equivalence of weighted kappa and the intraclass correlation coefficient as measures of reliability," *Educational and psychological measurement*, vol. 33, no. 3, pp. 613–619, 1973.