

## Research Article

# Evaluation and Promotion of a Multidimensional Information Intelligent Speech System in Dialect Teaching

Ya Pang 

School of Foreign Languages, Hainan Normal University, Haikou 571158, China

Correspondence should be addressed to Ya Pang; 030102@hainnu.edu.cn

Received 5 January 2022; Revised 23 January 2022; Accepted 17 February 2022; Published 8 March 2022

Academic Editor: Yanqiong Li

Copyright © 2022 Ya Pang. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Due to the wide scale of learners, large individual differences and scattered distribution, dialect teaching is difficult to carry out effectively by traditional school education. In order to improve the teaching level of dialect, taking Cantonese as an example, this paper constructs a teaching evaluation system based on multidimensional information. Through the questionnaire and investigation of large Cantonese training institutions in Guangdong Province, the data set is formed, the CMA-ES algorithm with efficient optimization ability is selected to optimize the SVM, and the model is compared with ACO and SVM without optimization algorithm. Experimental results show that the average accuracy of CMA-ES algorithm is 95.85% and the average running time is 21.0 ms on 8 data sets, which has obvious advantages relatively. Based on the evaluation model, the basis for teaching optimization is found through sensitivity analysis, and *student's language expression* is the most important index. And with the help of the intelligent voice system, the improvement measures for Cantonese teaching are proposed from the aspects of scene, oral, and scoring.

## 1. Introduction

UNESCO pointed out that among the more than 6000 Chinese languages in the world, about 96% of the language users' account for less than 3% of the total human population. On average, two languages disappear every month. As one of the countries with the richest language resources in the world today, our country's language resources are facing two basic facts: one is that there are many kinds of minority languages and Chinese dialects, and the language and culture is very rich. The other is that due to the rapid progress of urbanization and modernization with the continuous construction of China and the migration and exchanges between regions, minority languages and Chinese dialects are on the verge of disappearing. The protection of language resources, especially dialect resources, should not be limited to academic activities. The public should be involved in the protection of language resources. Evaluation is an important means to promote education reform, and teaching evaluation has become a hot issue in the field of education at home and abroad. Teaching evaluation is divided into formative evaluation and process evaluation. The former pays more

attention to students' academic performance, while the latter pays more attention to the status of students in the learning process. With the continuous advancement of education reform, process evaluation is more and more widely used in teaching evaluation. Process evaluation is carried out through the conditions of teachers and students in the classroom, which can reflect the scientific nature of the teaching process, and at the same time, judge the relationship between the teaching process and the teaching results based on the students' academic performance. The most important thing in dialect teaching is language cognition, reading, and comprehension. Bloom [1] thought that summative evaluation is designed to judge the achievements that students have achieved at the end of the course, and formative evaluation is intended to provide feedback and corrections at each stage of teaching and learning. Grant and Jay (2003) [2] put forward the concept of "reverse instructional design", and from the perspective of curriculum design dedicated to promoting students' understanding, six dimensions of understanding are proposed: explanation, paraphrase, application, insight, empathy, and self-awareness. The famous contemporary curriculum theory and education research expert, Anderson

(2008) [3]. The research team composed of nearly 10 experts further revised the idea of taxonomy of educational objectives initiated by Bloom et al. The original single-dimensional cognitive domain was changed to a two-dimensional division, namely, “Knowledge dimension” and “cognitive dimension,” the knowledge dimension includes four specific classification levels, namely, factual knowledge, conceptual knowledge, procedural knowledge, and reflection cognitive knowledge. The cognitive process follows cognitive complexity is ranked from low to high, including six categories: memory, comprehension, application, analysis, evaluation, and innovation, a total of 19 specific cognitive processes.

In terms of teaching optimization, with the support of intelligent technology, the classroom management system has developed rapidly. Considering that the classroom management system that provides classroom teaching evaluation must have the basic functions of real-time analysis, evaluation of the learning process, providing teachers with feedback information, etc. Lynnette is an intelligent guidance system that teaches students to solve linear equations [4]; Lumilo is a smart glass that can visualize information and perform certain virtual interactions [5]. Both can help teachers perceive students’ information in real time through a certain mechanism. FACT system is an intelligent classroom management system that can support students to solve mathematical problems collaboratively, and at the same time enhance teachers’ perception and control ability [6]. Spinoza provides guidance and assistance to students and teachers by analyzing the real-time codes and operating behaviors of students in the process of programming (Nisenbaum et al., 2019) [7]. MT-Classroom [8] is also a smart classroom, which enables students to learn on a touch-screen desktop. The built-in software analyzes the students’ learning in real time and sends them to the teacher. The teacher uses the tablet computer’s dashboard program to view student status and control the activity process in real time. In the behavioral evaluation model of teaching and learning, it mainly includes individual abnormal state, individual cognitive state, individual noncognitive state, group problem solving progress, group writing learning state, and teaching intervention. Common abnormal states of individual students include offline, silence, and abusive prompts, among which, abusive prompts are students’ “crafty” behaviors [9]. In the evaluation process of individual cognitive state, simple numerical evaluation is usually used, such as the number of correct answers, the number of attempts, and the number of help requests, which is a coarse-grained evaluation [10]. Individual noncognitive states mainly include concentration, emotion, and stress. Studies have shown that immersed in learning, physiologically manifested as happy and pleasant emotions, and have high concentration [11]. Group problem solving progress requires teachers to understand the learning situation and progress of the study group to help students solve problems [7]. Chi et al. [12] believe that monitoring the status of group collaborative learning is equally important. The higher the student’s participation in learning, the better the learning effect. In group collaborative learning, a member does not express opinions, but passively expresses opinions is passive learning [10]. Finally, it is teacher inter-

vention, which can be divided into cognitive intervention and noncognitive intervention according to the different status of students. Cognitive interventions include direct answers, hints, and prompts [10]. Noncognitive interventions are noncognitive feedback to students, including praise, reward, and criticism [13–15]. In the choice of evaluation method, Mullen et al. [16] first analyze the syntax between texts, and then further analyze the sentiment of the text based on the SVM algorithm. Moraes et al. [17] integrated machine learning algorithm and deep learning into teaching evaluation, compared the classification effect of support vector machine algorithm and artificial neural network algorithm, and proved the classification effect of ANN and support vector machine. In addition, fuzzy mathematics theory, analytic hierarchy process, and BP neural network are also often used in language teaching evaluation. Bamakan [18] and Gao [19] respectively, proposed the use of particle swarms, artificial fish swarms, and genetic algorithms for optimization. Although these methods have improved the blindness of support vector machine parameter selection to a certain extent, they all have varying degrees of premature problems, so that the predictive model is not the optimal model. In terms of the application of intelligent speech systems, the Bayesian network-based intelligent teaching system student model established by LAN [20] can not only objectively evaluate students’ cognitive abilities but also infer the student’s next learning behavior. Myers [21] uses an intelligent teaching system to automatically detect the emotional state of students and guide them into an active learning state. Based on previous research, this article found that mathematical algorithms are widely used in process teaching evaluation, but most of them use analytic hierarchy process, expert scoring, and other evaluation methods that are greatly affected by subjective factors, and algorithm optimization needs to be further improved. In this paper, SVM algorithm is used as the evaluation method of Cantonese teaching, and CMA-ES is used to optimize the SVM algorithm. This article constructs an evaluation index system from three perspectives of teaching language, teaching behavior, and teaching emotion, and fully considers the characteristics of procedural evaluation. Compare to SVM, ant colony algorithm optimized SVM and CMA-ES optimized SVM algorithm, select the algorithm with the highest accuracy rate as the evaluation method in this article, and apply the selected evaluation method to practice to obtain the most important influencing factors. The introduction of the intelligent voice system proposes measures for Cantonese teaching, fully considering the influence of the language environment and living environment on the teaching effect, and taking into account the practicality and innovation. In the evaluation process of this article, the effect of Cantonese teaching is divided into four levels: excellent, good, passing, and failing, and the optimized algorithm is used for empirical research. This paper takes a large Cantonese teaching institution as an example, selects the best plan through the comparison of three evaluation methods, evaluates the teaching effect of the institution, and introduces an intelligent voice system to propose an optimization strategy for Cantonese teaching.

## 2. Evaluation Index System of Cantonese Teaching

The Cantonese teaching evaluation system is an intuitive manifestation of multidimensional information processing, and requires a very deep understanding and research on the teaching system [22, 23]. Teaching evaluation is based on teaching activities and learning activities. The evaluation process focuses on improving ability of teachers and classroom teaching quality, and then evaluating classroom teaching design, process, and results. The current traditional teaching evaluation is manifested as a combination of internal and external multiple evaluations, process and performance evaluation judgments, on-site observation, and appraisal and screening by experts and peers. With the continuous development of information technology, teaching evaluation has undergone changes in terms of subject, content, methods, and results.

In terms of evaluation subjects, it is mainly divided into two aspects: internal subjects and external subjects. The internal subjects refer to the subjects that participate in the teaching activities, namely, students and teachers. External subjects refer to evaluators outside of teaching activities, including experts and colleagues. In terms of evaluation content, with the continuous advancement of education reform, education pays more attention to the change of emotional information, and the organic combination of emotional satisfaction and knowledge acquisition can more comprehensively evaluate the teaching effect. With the support of artificial intelligence technology, teachers and students' voices, facial expressions, and body postures can be collected through professional equipment to carry out teaching emotion recognition, and dynamic emotional changes can be obtained. In terms of evaluation methods, evaluation is the "baton" of teaching and provides decision-making materials for education optimization. Peer evaluation methods are widely used in teaching evaluation. The clarity of teacher teaching, the sufficient degree of content setting, and the sufficient degree of teacher-student interaction are often used. Used as an evaluation indicator, with the empowerment of artificial intelligence, cameras can be installed in the classroom to collect teacher and student voice, facial and posture information, carry out topic language analysis, topic behavior analysis, and topic emotion analysis to obtain students' attention, knowledge mastery, and interaction. Analyze the teaching effect based on the situation, emotional state, and other learning situations.

In terms of evaluation results, teaching evaluation can provide decision-making materials for improving education by judging and discovering value, and ultimately achieve value enhancement. Through the feedback of the evaluation results, it points to the multifaceted development of teachers and students, and exerts the function of their development value. Process evaluation aims to identify the teaching style by grasping the characteristics of the teacher's language structure, identify the classroom teaching structure through the teacher-student language interaction, and identify the classroom state through the voice intonation. From the perspective of students, it is possible to build a multidimen-

sional and multilevel teaching evaluation system by recording the number of times students have raised their hands to speak, their head-up rate, and their participation in discussions.

Based on the above analysis, it can be obtained that the classification of Cantonese teaching evaluation indicators is based on the two evaluation objects of teachers and students, and the evaluation indicators are divided into language, behavior, and emotion. In terms of language analysis, the indicators include the teacher's pronunciation standards, the teacher's intonation, the teacher's classroom organization structure, the teaching clarity, and the students' language expression. In terms of behavior analysis, the indicators include the teacher's body posture, student's head-up rate, classroom interaction, and participation in discussions. The indicators of sentiment analysis include teacher's emotional state, student's emotional state, student's attention, and satisfaction of students. The evaluation index system is shown in Figure 1 and the interpretation and definition of indicators are shown in Table 1.

The developed evaluation index system in Figure 1 can more comprehensively summarize the actual situation of Cantonese teaching. Using Cantonese teaching effect evaluation index data as the input sample of the evaluation model can realize the evaluation of Cantonese teaching effect.

## 3. The Method of Cantonese Teaching Evaluation

As one of the most difficult dialects in China, Cantonese has a complex composition and a variety of factors affecting Cantonese teaching, which belongs to the category of multidimensional information. As a common model for nonlinear relationship processing, SVM is consistent with the evaluation object of this paper. In order to make up for its inherent shortcomings, CMA-ES is also selected as the optimization algorithm.

*3.1. Evaluation Principle of SVM.* Support vector machines (SVM), as the latest content of statistical learning theory, was first proposed for pattern recognition problems, and it has shown its outstanding advantages over traditional methods such as neural networks in solving problems with limited samples, nonlinearity, and high-dimensional pattern recognition. With only a kernel function satisfying the Mercer condition, SVM can realize the problem solving by linear methods in high-dimensional space, which does not increase the computational complexity compared with the general linear model in low-dimensional space. It can be seen that the proposed kernel, like Radial basis kernel function, B-spline kernel function, and multilayer perceptron function, allows SVM to solve the up-dimensional disaster effectively. There are some problems in the running speed and parameter selection of SVM algorithm yet. The existing optimization methods such as artificial fish swarm and genetic algorithm improve the blindness of parameter selection of SVM to a certain extent, but they have different degrees of precocity simultaneously. In order to solve the problem of prediction accuracy, this

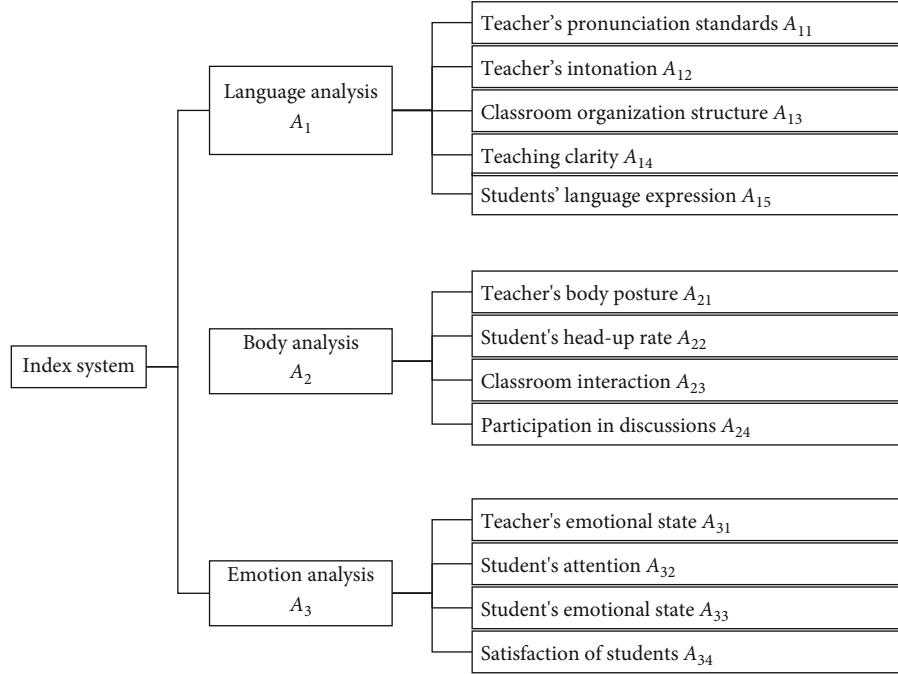


FIGURE 1: Evaluation index system of Cantonese teaching.

TABLE 1: Explanation of Cantonese teaching evaluation index.

| Index    |                                   | Explanation  |
|----------|-----------------------------------|--|
| $A_{11}$ | Teacher's pronunciation standards | The teacher's pronunciation conforms to the pure Cantonese six tones                       |
| $A_{12}$ | Teacher's intonation              | How well the teacher's tone is accepted by the students                                    |
| $A_{13}$ | Classroom organization structure  | The organization and hierarchy of lectures   |
| $A_{14}$ | Teaching clarity                  | Clarity of teaching content  |
| $A_{15}$ | Students' language expression     | Student's language ability   |
| $A_{21}$ | Teacher's body posture            | The teacher's physical movements in the teaching process                                   |
| $A_{22}$ | Student's head-up rate            | The number of times the student raised his head to attend the class                        |
| $A_{23}$ | Classroom interaction             | The frequency of teacher-student interaction in the classroom                              |
| $A_{24}$ | Participation in discussions      | Enthusiasm of student for participating in discussions                                     |
| $A_{31}$ | Teacher's emotional state         | Teachers' emotional state before, during, and after class                                  |
| $A_{32}$ | Student's attention               | Student's attention during class   |
| $A_{33}$ | Student's emotional state         | Students' emotional state before, during, and after class                                  |
| $A_{34}$ | Satisfaction of students          | The degree of satisfaction of the students with the teacher's teaching methods and content |

paper proposes a dialect teaching evaluation model based on covariance matrix adaptive evolutionary strategy (CMA-ES) to optimize SVM, which ensures the accuracy and running speed of the model. The basic idea of SVM is to map the data of sample space to a higher dimension or even infinite dimension feature space through a nonlinear mapping  $\varphi$  based on Mercer kernel expansion theorem, so that the highly nonlinear problems can be solved in feature space, as shown in Equation (1).

$$f(x) = \omega \bullet \varphi(x) + b, \quad (1)$$

where  $\omega$  is the weight vector,  $\omega \in R^n$ , and  $b$  is the offset,  $b \in R$ . It is assumed that all training data  $(x_i, y_i)$  can be fitted with  $f(x)$  within the accuracy  $\varepsilon$ .  $\varepsilon$  is the design parameter of the model, and  $\varepsilon > 0$ . Then the insensitive loss function is defined as Equation (2).

$$|y - f(x)|_\varepsilon = \begin{cases} 0; & |y - f(x)| \leq \varepsilon \\ |y - f(x)| - \varepsilon; & |y - f(x)| > \varepsilon \end{cases} \quad (2)$$

In consideration of the allowable fitting error, when the constraint conditions cannot be fully met, the loose other

variable  $\xi_i$  and  $\xi_i^*$  are introduced, and the optimization problem is like Equation (3).

$$\min = \frac{1}{2} \|\omega\|^2 + \gamma \sum_{i=1}^N (\xi_i + \xi_i^*), \quad (3)$$

where  $\gamma$  is the penalty coefficient, which is another important coefficient of the model, indicating the penalty degree for the training sample data exceeding the loss function.

**3.2. Performance Optimization of SVM.** Because the generalization ability of SVM is limited by the selection of penalty parameter  $c$ , RBF kernel function width  $\delta$  and insensitive loss function parameters  $\varepsilon$ , the optimization of SVM prediction performance is actually to solve the optimal parameter combination problem ( $\gamma$ ,  $\delta$ ,  $\varepsilon$ ). We use the mean square error ( $\Delta MSE$ ) of the trained model for prediction as an individual evaluation metric (i.e., fitness function), and the smaller the value of the metric, the higher the accuracy of the prediction.

CMA-ES shows good performance in solving global optimization problems, so it is used for adaptive parameter optimization selection of SVM parameters. The algorithm mainly consists of sampling and updating.

**3.2.1. Sampling.** CMA-ES uses Gaussian distribution  $N(m, \sigma^2 C)$ , and samples are collected in the solution space of the optimization problem to generate a population distribution composed of  $\alpha$  individuals  $z = (\gamma, \delta, \varepsilon)$ , which corresponds to the population in the optimization algorithm, as shown in Equation (4).

$$z_j^{(h+1)} = m^{(h)} + \sigma^{(h)} N_j \left( 0, C^{(h)} \right) j = (1, 2, \dots, \alpha), \quad (4)$$

where  $z_j^{(h+1)}$  is the  $j$ -th individual generation population  $h+1$  and  $m^{(h)}$  is the mean value of population distribution of generation  $h$ , and  $\sigma^{(h)}$  is the global step size of generation  $h$ , and  $C^{(h)}$  is the covariance matrix of the population distribution of generation  $h$ , and the relationship of each parameter is shown in Equation (5).

$$C^{(h)} = B^{(h)} \left( D^{(h)} \right)^2 \left( B^{(h)} \right)^T, \quad (5)$$

where  $B^{(h)}$  is the orthogonal matrix and its column vector is the orthogonal basis of the eigenvector of  $C^{(h)}$ , which is used for the rotation of the hyper-ellipsoid, and  $D^{(h)}$  is the diagonal matrix, and the diagonal element is the square root of the characteristic of  $C^{(h)}$ , corresponding to each column vector of  $B^{(h)}$ , which is used for scaling the hyper-ellipsoid of population distribution.

**3.2.2. Updating.** The update operation is mainly for parameters  $m$ ,  $C$  and  $\sigma$ , as shown in Equations (6)–(8).

$$m^{(h+1)} = \sum_{i=1}^{\beta} \omega_i z_{1:\alpha}^{h+1}; \quad \sum_{i=1}^{\beta} \omega_i = 1 \quad (\omega_1 \geq \omega_2 \geq \dots \geq \omega_{\beta} \geq 0), \quad (6)$$

where  $\omega_i$  is the set weight and  $z_{1:\alpha}^{h+1}$  is the  $i$ -th optimal individual in the  $h+1$  generation.

$$C^{(h+1)} = (1 - c_1 - c_{\beta}) C^{(h)} + c_1 \left[ p_{\gamma}^{(h+1)} \left( p_{\gamma}^{(h+1)} \right)^T + \delta \left( k_{\sigma}^{h+1} \right) C^{(h)} \right] + c_{\beta} \sum_{i=1}^{\beta} \omega_i y_{1:\alpha}^{h+1} \left( y_{1:\alpha}^{h+1} \right)^T, \quad (7)$$

where  $p_{\gamma}^{(h+1)} = (1 - c_{\gamma}) p_{\gamma}^{(h)} + k_{\sigma}^{(h+1)} \sqrt{c_{\gamma} \beta (2 - c_{\gamma})} (m^{(h+1)} - m^{(h)} / \sigma^{(h)})$  and  $c_{\gamma}$  is the update learning rate of  $p_{\gamma}$ .

$$\sigma^{(h+1)} = \sigma^{(h)} \exp \left( \frac{c_{\sigma}}{d_{\sigma}} \left( \frac{\|p_{\sigma}^{(h+1)}\|}{E(\|N(0, I)\|)} - 1 \right) \right). \quad (8)$$

To sum up, the SVM parameter optimization algorithm based on CMA-ES is as follows.

- (1) Input training set data samples  $S = \{(x_1, y_1), \dots, (x_l, y_l)\} \in R^m \times R$
- (2) Select appropriate parameters  $(\gamma, \delta, \varepsilon)$ .
- (3) Set parameters and initialization. The number of parent and child individuals in the population are  $\mu$  and  $\lambda$ , respectively, and  $\mu < \lambda$ . The maximum number of iterations is  $H$
- (4) Construct the training and test data set required for the experiment
- (5) Sample the population
- (6) Calculate the fitness of individual population
- (7) Update parameters
- (8) If the stop condition is reached, the optimization process stops, and the optimal individual  $(\gamma^*, \delta^*, \varepsilon^*)$  within its optimal fitness value are output. Otherwise, returns (5).
- (9) The working process of the evaluation system is shown in Figure 2

## 4. Experimental Results and Analysis

As a dialect, Cantonese is difficult to learn and start for people whose mother tongue is not it. In order to ensure the scientific and reliable data, we investigated large Cantonese teaching institutions in Guangdong, China. Sample data were collected according to the evaluation index

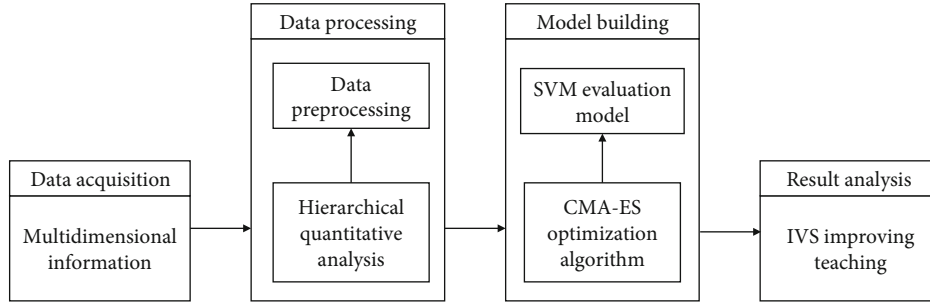


FIGURE 2: System working process block diagram.

TABLE 2: Accuracy comparison.

| Data set | CMA-ES (%) | ACO (%) | SVM (%) |
|----------|------------|---------|---------|
| 1        | 95.49      | 87.68   | 84.72   |
| 2        | 94.32      | 89.67   | 85.24   |
| 3        | 96.38      | 88.52   | 83.95   |
| 4        | 97.15      | 89.94   | 81.89   |
| 5        | 95.82      | 85.51   | 79.87   |
| 6        | 96.84      | 86.08   | 84.65   |
| 7        | 95.08      | 85.79   | 79.68   |
| 8        | 95.7       | 88.24   | 82.43   |

system of Cantonese teaching, and the quality level of large Cantonese teaching institutions and 160 data samples could be obtained for testing by evaluating the actual situation of Cantonese teaching and experts' evaluation of the effectiveness of Cantonese teaching. The 160 data samples are divided into 8 data sets, and each data set contained 20 samples. The training data of each group are input into SVM regressor for learning, and CMA-ES is used to optimize the punished parameter  $\gamma$  and the width of RBF kernel function  $\delta$ . The initial mean (population center)  $m^{(0)} = (\gamma, \delta, \epsilon)^{(0)}$ ,  $\gamma$  and  $\delta$  are set to  $[0.01, 100]$ , the maximum number of evolutionary generations is 100 and the termination threshold is set to 1. Meanwhile, the SVM hyperparameters  $(\gamma, \delta, \epsilon)$  are optimized using the CMA-ES algorithm, and the evaluation model is built with the final output optimal solution  $(\gamma^*, \delta^*, \epsilon^*)$ . After repeated experiments, the optimal solution on all data sets is  $\gamma^* = 129.92$ ,  $\delta^* = 11.63$ ,  $\epsilon^* = 0.003$ . In order to verify the optimization effectiveness of CMA-ES method in Cantonese teaching evaluation model, we also use ant colony optimization (ACO) and SVM model without optimization algorithm for comparative test. The accuracy results of the three schemes are presented in Table 2 and Figure 3. It can be seen that both ACO and CMA-ES have an optimization effect on SVM, but the optimization effect of ant colony algorithm is not obvious, and the accuracy of evaluation model only using SVM is not enough. The average accuracy of the model using CMA-ES method is 95.85%, which has high reliability. Relatively speaking, the accuracy of the three methods in dataset 5 is poor, which may be due to the low discrimination of its data sources.

In addition to accuracy, another important evaluation basis of SVM optimization algorithm is running time. Only

the optimization algorithm that meets both high accuracy and running time is the most cost-effective and most worthy method. Figure 4 shows that although the advantage of CMA-ES for ACO is not obvious in running time, it has a very significant optimization for SVM. It also indicates that CMA-ES has stronger performance when the amount of data is larger.

## 5. Teaching Improvement Measures

The main purpose of constructing a high-precision and efficient teaching evaluation model in this paper is to find the evaluation index that has the greatest impact on the evaluation results. Therefore, we can effectively improve Cantonese learning effect of students who are interested in Cantonese with the help of IVS. According to our data analysis results, we conduct further research in order to find more sensitive evaluation indicators. Our findings show that *teacher's pronunciation standard* ( $A_{11}$ ), *students' language expression* ( $A_{15}$ ), *classroom interaction* ( $A_{23}$ ), and *students' attention* ( $A_{32}$ ) have the higher impact on Cantonese teaching. It is obvious that these four indicators have a positive impact on the final teaching effect, so only their positive changes are considered in this paper when doing sensitivity analysis. Then we increase each index by 5%, 10%, and 15%, respectively, on the original basis, and the impact on the final evaluation results is shown in Figure 5.

Figure 4 clearly shows that  $A_{15}$  has the greatest impact on Cantonese teaching. For a language teaching dominated by interest, the quality of students' expression is the most intuitive embodiment of teaching effect. Teacher's pronunciation also has a great impact on teaching results, and with the increase of indicators, the improvement effect on teaching is more and more significant, indicating that the more standard teacher's pronunciation is, the better students' learning effect is. As for classroom interaction, a small increase has an obvious impact on the results, illustrating that teachers need to control the proportion of classroom interaction.

Cantonese has rich tones and syllables. Modern Putonghua has only four tones, while Cantonese dialect has nine tones and two inflections, and the difference between these nine tones is not obvious, so traditional teaching methods are not very effective in Cantonese teaching. Thus, IVS is more powerful in improving these two important indicators

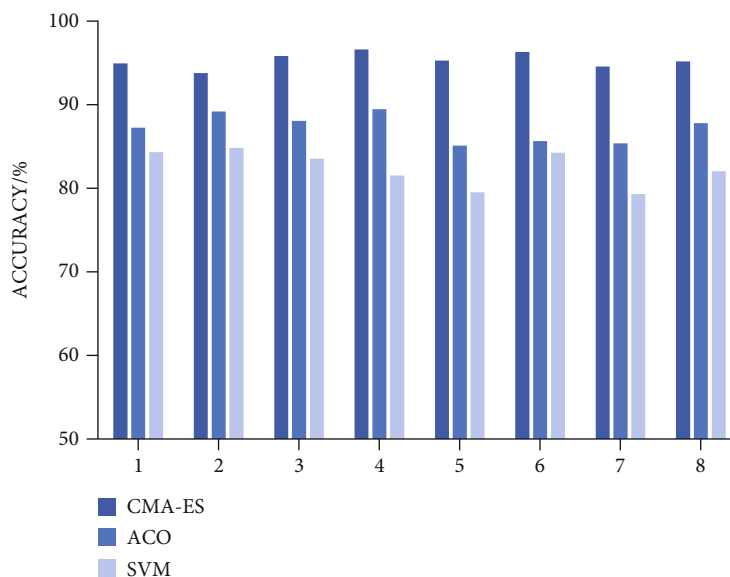


FIGURE 3: Accuracy comparison.

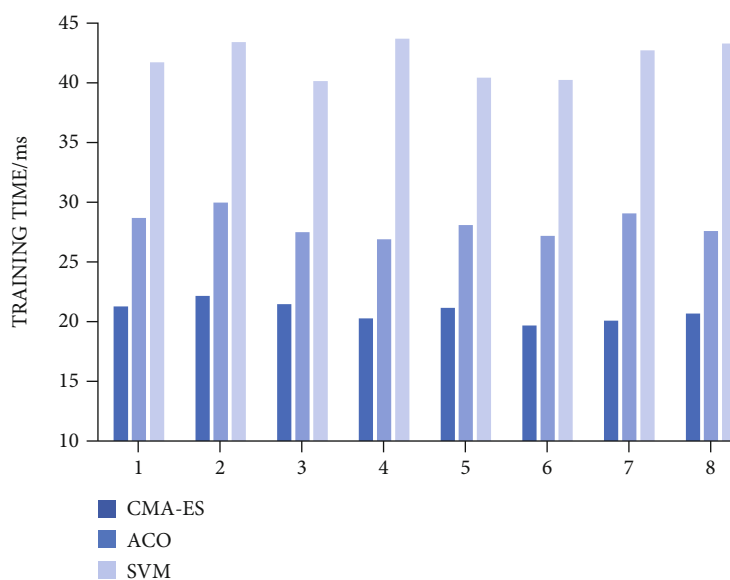


FIGURE 4: Training time comparison.

( $A_{11}, A_{15}$ ). With the support of speech recognition technology and speech synthesis technology, IVS realizes human-computer interaction (HCI) has two technical aspects: input and output. The input of HCI is based on the language information received by the computer, which is recognized and understood by speech recognition technology and then converted into text information. The output of HCI is based on the input text information, which is converted into easily understandable and applicable language information by speech synthesis technology. From the technical aspect, intelligent speech system has the functions of standard reading, speech synthesis, speech evaluation, and audio teaching courseware production. The hardware facilities of IVS support various audio teaching resources. Students can hear

the reading of pronunciation standards with a simple click, so as to lead students to perceive and remember standard and authentic Cantonese. In the environment where modern information technology is widely used in social life, Cantonese teaching must keep pace with the times, combine the content of English teaching, students' age and psychological characteristics, and take advantage of the rich learning resources and vivid knowledge presentation of the intelligent speech system to make students gain aural and visual freshness and endogenize the initiative of Cantonese learning.

*5.1. Optimize Situational Teaching.* Language learning is for communicative purposes, and dialects are more life-like, so only by learning language and using it in a communicative

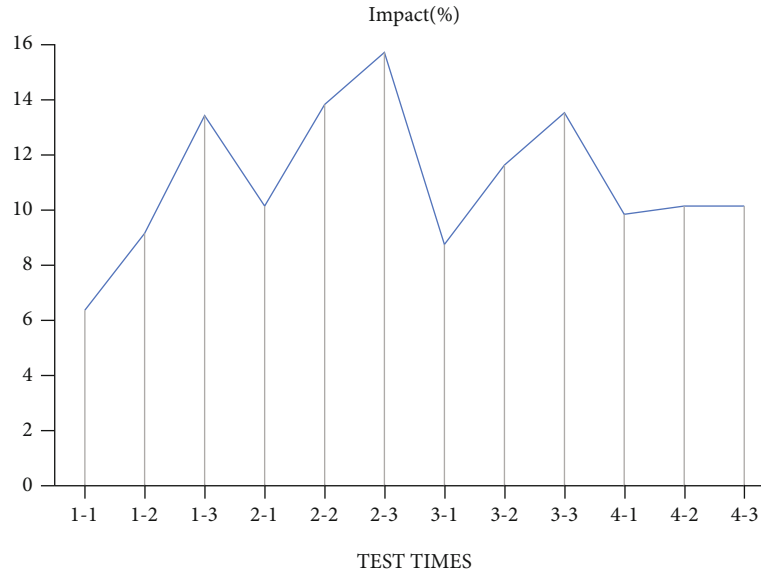


FIGURE 5: Sensitivity analysis results.

context can we have a true understanding of the words, sentences, and texts we learn and clarify the linguistic function in the complete internalization of knowledge. Through IVS, teachers can generate life scenes such as “buying vegetables,” “renting a house,” “traveling,” and “drinking morning tea” according to teaching needs, so that students can deepen their language learning in the context of Cantonese life. Therefore, the teaching can not only provide students with fresh audio-visual perception and sensory experience, mobilize students’ learning initiative, but also enable students to acquire knowledge and improve their ability in communication and interaction. For example, in the characteristic morning tea culture in Guangdong, the teacher first outlines the top view of the tea restaurant on the electronic blackboard. Under the guidance of the waiter, the students divide themselves into groups and complete the learning tasks in the process of “ordering” and “dining.”

**5.2. Optimize Oral Teaching.** Oral communication in Cantonese is far greater than written communication. Most people who are interested in learning Cantonese and can pay the corresponding tuition fees are adults with leisure time. Compared with step-by-step, they prefer to be rapidly improved, which has stronger requirements for teacher’s language level and teaching quality. Cantonese teachers should use IVS to construct and render real situations of oral communication, create more opportunities for students to use spoken language, lead them to be in oral communication situations, and let them start oral communication by role-playing and intelligent imitation with the advantages of technology such as dot reading, following, listening, and discriminating included in IVS. For example, teacher can create multiangle oral Q&A around simple sentence patterns, let students make corresponding answers according to the questions played by IVS, gradually shorten the reserved time for answers, and improve students’ oral response speed.

**5.3. Optimize Teaching Scoring.** IVS not only has a “mouth” that can speak standard Cantonese but also has “ears” that can carefully distinguish whether Cantonese pronunciation is standard. Speech recognition technology endows IVS with the function of automatic evaluation of English pronunciation through auditory processing, information conversion, and output, so as to timely and accurately help teachers and students compare standard pronunciation and correct pronunciation deviation in the form of quantitative score. Students can understand the problems existing in their pronunciation, conduct pronunciation correction exercises through the pronunciation comparison control button, and improve the pronunciation standard in repeated follow-up and evaluation with standard pronunciation. Through the all-round improvement of students’ “listening, speaking, reading, and writing” abilities, they can effectively integrate Cantonese learning into students’ daily lives and feel the charm of depalletizations. Give full play to the functions of the intelligent voice system, and then feedback to the classroom teaching, and conduct a second evaluation of Cantonese teaching, to achieve the PDCA cycle of evaluation.

## 6. Conclusion

The index factors affecting Cantonese teaching come from a wide range of sources, belong to multidimensional information, and there is a complex nonlinear functional relationship between the evaluation results and the evaluation indexes. For the objective science of teaching evaluation, this paper uses the commonly used nonlinear problem processing model SVM to solve it. However, SVM itself has some limitations. In order to improve the applicability of the model, this paper uses CMA-ES algorithm to optimize it. Experimental results show that, in terms of model accuracy, this method is 8% and 13% higher than ACO and SVM, respectively. In terms of running time, it is 7 ms and 21 ms,



respectively. It can be seen that the evaluation method in this article is highly scientific and reasonable. Therefore, relying on the evaluation model constructed in this paper and with the help of IVS, we can put forward targeted optimization measures for Cantonese teaching, which is of great significance to improve the Cantonese level of students with poor foundation. Due to conditions and time constraints, this paper has some deficiencies in data sources and evaluation methods. The evaluation data of this paper adopts the traditional questionnaire survey method. With the strengthening of the function of IVS, it can be combined with computational vision technology to judge the emotional attitude and interactive atmosphere in the classroom through expression and voice intonation recognition, and master the teachers' classroom control ability and emergency handling ability through expression recognition. We hope to provide personalized and targeted methods for the teaching evaluation and improvement of other dialects in future research.

### Data Availability

The labeled dataset used to support the findings of this study are available from the corresponding author upon request.

### Conflicts of Interest

The author declares no competing interests.

### Acknowledgments

This work was supported in part by the "An Acoustic Study of Syllables of Ha Dialect in Li language from the Perspective of Ecological Linguistics", a philosophy and social science planning project in Hainan province in China, 2018. (Grant No. HNSK.QN 18-29).

### References

- [1] B. S. Bloom, "Some theoretical issues relating to educational evaluation," in *Educational Evaluation: New Roles, New Means. The 63rd Yearbook of the National Society for the Study of Education, Part 2*, pp. 26–50, University of Chicago Press, Chicago, IL, 1969.
- [2] W. Grant and M. Jay, *Understanding by design*, vol. 13, Association for Supervisors, 2006.
- [3] L. W. Anderson, "Taxonomy for learning, teaching, and assessing," *European Legacy*, vol. 114, no. 458, pp. 1013–1014, 2001.
- [4] Y. Long and V. Alevan, "Supporting students' self-regulated learning with an open learner model in a linear equation tutor," in *Artificial Intelligence in Education*, pp. 219–228, Springer, Berlin, Heidelberg, 2013.
- [5] K. Holstein, B. McLaren, V. M. Alevan, and V. Alevan, "Co-designing a real-time classroom orchestration tool to support teacher-AI complementarity," *Journal of Learning Analytics*, vol. 6, no. 2, pp. 27–52, 2019.
- [6] K. VanLehn, H. Burkhardt, S. Cheema et al., "Can an orchestration system increase collaborative, productive struggle in teaching-by-eliciting classrooms," *Interactive Learning Environments*, vol. 2, pp. 1–19, 2019.
- [7] M. Tissenbaum, M. Lui, and J. D. Slotta, "Co-designing collaborative smart classroom curriculum for secondary school science," *Journal of Universal Computer Science*, vol. 18, no. 3, pp. 327–352, 2012.
- [8] R. Martinez-Maldonado, A. Clayphan, K. Yacef, and J. Kay, "MTFeedback: providing notifications to enhance teacher awareness of small group work in the classroom," *IEEE Transactions on Learning Technologies*, vol. 8, no. 2, pp. 187–200, 2014.
- [9] V. Alevan, I. Roll, B. M. McLaren, and K. R. Koedinger, "Help helps, but only so much: research on help seeking with intelligent tutoring Systems," *International Journal of Artificial Intelligence in Education*, vol. 26, no. 1, pp. 205–223, 2016.
- [10] K. Vanlehn, "The behavior of tutoring systems," *International Journal of Artificial Intelligence in Education*, vol. 16, no. 3, pp. 227–265, 2006.
- [11] M. Csikszentmihalyi, *Flow: The Psychology of Optimal Experience[M]*, Harper & Row, New York, 1990.
- [12] M. T. Chi and R. Wylie, "The ICAP framework: linking cognitive engagement to active learning Outcomes," *Educational Psychologist*, vol. 49, no. 4, pp. 219–243, 2014.
- [13] A. D. Tillery, K. Varjas, J. Meyers, and A. S. Collins, "General education teachers' perceptions of behavior management and intervention Strategies," *Journal of Positive Behavior Interventions*, vol. 12, no. 2, pp. 86–102, 2010.
- [14] S. Shen and M. Chi, "Reinforcement learning: the sooner the better, or the later the better?," in *Proceedings of the 2016 Conference on user modeling adaptation and personalization (UMAP'16)*, pp. 37–44, Association for Computing Machinery, New York, NY, USA, 2016.
- [15] M. S. Ausin, H. Azizoltani, T. Barnes, and M. Chi, "Leveraging deep reinforcement learning for pedagogical policy induction in an intelligent tutoring system," in *Proceedings of the 12th International Conference on Educational Data Mining (EDM 2019)*, pp. 168–177, Montreal, Canada, 2019.
- [16] T. Mullen and N. Collier, "Sentiment analysis using support vector machines with Di-verse information sources," in *Proceedings of the 2004 Conference on Empirical Methods in Natural Language Processing*, pp. 23–35, Barcelona, Spain, 2004.
- [17] R. Moraes, J. F. Valiati, and W. P. G. Neto, "Document-level sentiment classification: An empirical comparison between SVM and ANN," *Expert Systems with Applications*, vol. 40, no. 2, pp. 621–633, 2013.
- [18] S. M. H. Bamakan, H. Wang, T. Yingjie, and Y. Shi, "An effective intrusion detection framework based on MCLP/SVM optimized by time-varying chaos particle swarm optimization," *Neurocomputing*, vol. 199, pp. 90–102, 2016.
- [19] Y. Gao, Y. Shen, G. Zhang, and S. Zheng, "Information security risk assessment model based on optimized support vector machine with artificial fish swarm algorithm," in *Proc of Software Engineering and Service Science*, pp. 599–602, New York, 2015.
- [20] L. Wu, "Student model construction of intelligent teaching system based on Bayesian network," *Personal and Ubiquitous Computing*, vol. 24, no. 3, pp. 419–428, 2020.
- [21] M. H. Myers, "Automatic detection of a student's affective states for intelligent teaching systems," *Brain Sciences*, vol. 11, no. 3, p. 331, 2021.

- [22] B. Q. Cheng, K. Lu, J. Li, H. Chen, X. Luo, and M. Shafique, "Comprehensive assessment of embodied environmental impacts of buildings using normalized environmental impact factors," *Journal of Cleaner Production*, vol. 334, article 130083, 2021.
- [23] L. Lu, Z. Gu, D. Huang, Z. Liu, and J. Chen, "An evaluation framework for the public information guidance system," *KSCE Journal of Civil Engineering*, vol. 21, no. 5, pp. 1919–1928, 2017.