Teaching Effect Evaluation Mechanism of English MOOC Combined with LS-SEM Intelligent System

Jing Liu

Basic Teaching Department, Shangqiu Institute of Technology, Shangqiu 476000, China

Correspondence should be addressed to Jing Liu; 161849037@masu.edu.cn

Received 13 April 2022; Revised 8 June 2022; Accepted 17 June 2022; Published 7 July 2022

Academic Editor: Wen Zhou

Copyright © 2022 Jing Liu. 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.

In order to improve the effect of English teaching, this article analyzes the MOOC teaching system, constructs an evaluation mechanism combined with the LS-SEN intelligent system, and obtains an intelligent English MOOC teaching effect evaluation system. According to the actual needs of English MOOC teaching effect evaluation, this article proposes an integrated method LS-Ensem algorithm for continuous modeling from the perspective of gradient descent. At the same time, this article uses the commonly used variance function as the loss function, combined with repeated iterations to improve the running accuracy of the algorithm. In addition, this article constructs an intelligent system structure based on the actual needs of English MOOC teaching. According to the experimental research, the teaching effect evaluation mechanism of English MOOC combined with the LS-SEM intelligent system can play an important role in English MOOC teaching.

1. Introduction

College English teaching urgently needs to enrich the types of courses and improve the quality of teaching, so as to meet the diverse English learning needs of learners, and this demand is particularly prominent due to the impact of MOOCs on traditional classrooms. The themes of MOOCs involve various disciplines, and the language of instruction is mainly English, which can fully meet the diverse needs of learners in terms of professional English, academic English, and general education English. In addition, the MOOC is based on a systematic fragmented learning method and modular short videos, which not only facilitates learners to learn by topic, but also facilitates them to use fragmented time to split and watch. This will attract many students to choose to “skip classes” and go to “Amoy (mo) classes.” At this time, college English teachers will lose their position of knowledge authority to a large extent, and it will be difficult for students to accept traditional teaching methods, which will lead to the danger of student loss in the classroom.

MOOCs can test the effect of college English teaching to a certain extent, and language barriers (i.e., English proficiency) are one of the main reasons that hinder learners from completing MOOCs, which will inevitably exacerbate students’ understanding of the content and teaching of traditional college English courses. Dissatisfaction with quality, to a large extent, “divides” the traditional relationship of domination and dependence between teachers and students, awakening students’ subjectivity. Faced with the impact of MOOCs on traditional college English teaching, colleges and universities should take advantage of MOOCs and explore a hybrid teaching model that integrates MOOCs and classroom teaching through the collaborative innovation of online MOOC education and traditional face-to-face teaching, so as to realize the structure of college English courses upgrade and improve the quality of teaching.

Teacher training, training and maturity, and other related teacher development issues are one of the key factors to realize the transformation of college English follow-up curriculum and the improvement of teaching quality. MOOCs have brought about many changes in college English education concepts, teaching concepts, teaching methods, and teaching content. While challenging the original teaching concepts and knowledge structure of college English teachers, it also breeds opportunities. Professional international MOOC courses improve their own
5. Improving English Microteaching Skills

The development of basic education curriculum reform, training are studied. With the continuous deepening and development of English MOOC teaching effect evaluation system, which provides a theoretical reference for the subsequent improvement of the teaching effect of English MOOCs.

2. Related Work

In recent years, there have been related research studies on microteaching in the context of MOOCs, such as "how to realize the mutual penetration of English microteaching and MOOCs in colleges and universities." The research shows that MOOC teaching provides resources and materials for microteaching; MOOC teaching provides technical support for microteaching [1]. Microteaching in the context of MOOCs can help teachers to further understand and improve their classroom teaching skills, and promote their professional development. With the development of online learning, Xuetong has gradually become a common learning channel in colleges and universities. Teachers can share excellent resources to Xuetong for students to watch and learn [2]. With the help of the Xuetong platform, the effect and problems of English micro-introduction skill training are studied. With the continuous deepening and development of basic education curriculum reform, teachers pay more attention to the effectiveness of classroom teaching [3]. At the beginning of classroom teaching, classroom introduction is related to the success or failure of a class. At the same time, introduction skills are one of the top ten skills in microteaching [4]. Correct and ingenious introduction can prepare for the next teaching and help students to obtain good learning effect. However, the research on the development of teachers' teaching ability is the core part of the research on teachers' professional development. The practical knowledge of English teachers is generally relatively lacking, and the teaching ability of teachers needs to be improved urgently [4]. As one of the compulsory courses for teachers' vocational skills training in higher normal colleges and universities, microteaching just provides a broad practical platform for English teachers, making it possible for normal students to focus on solving a specific teaching behavior or conduct it under controlled conditions [5]. Through English micro-introduction skills training, teachers can further understand microteaching, which will also be a process for teachers to continuously learn, improve teaching skills, and constantly improve themselves [6].

The system provides a wealth of English learning resources. Through videos, students are guided to learn English language knowledge and have a good understanding of their own English language skills for training [7]. The self-learning system can be divided into course selection and self-study. Students can choose their own learning content, including writing, reading, and listening, according to their own interests and real learning needs. Students can choose flexibly to achieve individuality chemical learning [8]. In the self-learning system, students can design their own English learning progress and do not need to study according to the progress of traditional English teaching, so that students with different learning abilities and different learning foundations can be satisfied. Students can choose their weak points of knowledge to strengthen training, further consolidate classroom knowledge, and check and fill gaps [9].

The MOOC education platform is an online learning and education platform for teachers and students [10]. The teaching management system can play a navigational role, guide students to complete the staged learning and then conduct testing, and guide students in the selection of learning content. The supervision function of the teaching management system can better guide and restrict students, mobilize students' enthusiasm for autonomous learning, and encourage students to actively participate in autonomous learning of the MOOC platform [11]. Teachers can use the teaching management system to grasp the online time of students, whether they complete the knowledge test on time, whether they have read the educational resources in the educational platform, etc. [12].

The evaluation feedback system in the MOOC education platform is to better supervise, regulate, and encourage students' self-learning awareness and behavior [13]. The evaluation feedback system will also provide feedback on the students' learning process, such as students' online learning time reports and students' learning process reports on the MOOC platform, which can help teachers better grasp the students' learning process [14]. The system also has a pass detection module. After students complete the learning of a certain skill, the system will pop up a pass test to assess students' English skills online [15]. There are many overall analyses, but the design of empirical research on MOOC teaching and learning, especially the overall analysis of the quality of MOOC teaching, is still extremely lacking, and there is no recognized quality evaluation index system, and even the evaluation standards are in the theoretical level [16].

3. LS-Ensem Algorithm

When traditional algorithms process teaching information, the information processing process is complex, which leads to poor system performance. In the iterative process, the LS-Ensem algorithm no longer selects complex approximation functions, but selects submodels with relaxed conditions and sets a suitable This method can effectively reduce the amount of calculation and the possibility of noise, so this article chooses the LS-Ensem algorithm as the core algorithm of the system.

The partial linear square (PLS) method is a common modeling method, which has been widely used due to its simplicity and speed. However, the PLS method is a linear
algorithm, and most of the actual objects have a certain nonlinearity more or less, which causes the models obtained by the PLS algorithm to often have inaccurate predictions. The following will select the PLS algorithm as the modeling method and explain how to use the AdaBoost.Rm method to improve its prediction accuracy. Because the traditional PLS algorithm does not consider the distribution information in the modeling process. If we want AdaBoost.Rm to be able to call it, it must be improved to model both the sample and the sample distribution.

The traditional PLS algorithm decomposes the sample set $S$ into an input matrix $X = \{x_1, \ldots, x_m\}'$ and an output vector $Y = \{y_1, \ldots, y_m\}'$. Among them, $(x_i, y_i), i = 1, \ldots, m$ is a sample point in $S, W \in \mathbb{R}^{Rp}$ in the resulting model is a weight matrix consisting of weight vectors. However, $P \in \mathbb{R}^{pr}$ and $C \in \mathbb{R}^{sr}$ are the loading matrices of $X$ and $y$, respectively, $b \in \mathbb{R}^{p+1}$ is the parameter vector of the model.

The PLS algorithm itself does not consider the distribution information on the sample set during the execution process. In order to use the AdaBoost.Rm method to improve its accuracy, it must be improved. Therefore, the following work is to improve the PLS so that the distribution information on the sample set is also considered in the calculation process without loss of generality. We assume that the distributions $D(i), i = 1, \ldots, m$ on the sample set $S$ are all in the form of reduced fractions, that is, $D(i) = (D_{ni}/D_{di})$. Among them, $S$ and $D_{di}$ are relatively prime positive integers. The least common multiple $D_{d}$ of all denominators $\{D_{d1}, \ldots, D_{dm}\}$ is calculated. The sample points in $S$ are replicated so that the number of sample points $(x_i, y_i)$ is $(D_{d} \times (D_{ni}/D_{di})), \ldots, (x_m, y_m)D_{d} = (D_{m}/D_{di}) \ldots, (x_m, y_m)D_{d} = (D_{m}/D_{di})$ is obtained. A simple calculation shows that there are $D_{d} \sum_{i=1}^{m} (D_{ni}/D_{di}) = D_{d}$ sample points in the extended sample set $S$. And, for each sample point $(x_i, y_i), i = 1, \ldots, m$, the ratio of its number in the extended sample set $S$ to the total number of samples in $S$ is exactly equal to $D(i)$. That is to say, the extended sample set $S$ not only contains the same sample points as the sample set $S$, but also implies the distribution information on the sample set $S$.

Now, the extended sample set $S$ is decomposed into an input matrix $\tilde{X}$ and an output vector $\tilde{Y}$, and then, the traditional PLS algorithm is invoked. First, the weight vector is calculated, and the input matrix $\tilde{X}$ and output vector $\tilde{Y}$ are brought into the PLS algorithm:

$$\tilde{w} = \frac{\tilde{X}' \tilde{y}}{(\tilde{y}' \tilde{y})}$$

$$\begin{align*}
\tilde{w} &= \frac{X'(D \cdot y)}{y'(D \cdot y)} \\
\tilde{w} &= \frac{X'(D \cdot y)}{y'(D \cdot y)}.
\end{align*}
$$

The symbol in the formula represents the dot product of the vector. The weight vector $\tilde{w}$ obtained in formula (1) is normalized and recorded as $\tilde{w} = (\tilde{w}_1, \ldots, \tilde{w}_p)'$. We set $t = X\tilde{w} = \{t_1, \ldots, t_m\}'$ and then calculate the fractional factor $t$:

$$\begin{align*}
\tilde{t} = \tilde{X}\tilde{w} &= \begin{pmatrix}
Dd & \frac{D_{n1}}{D_{d1}} & \vdots & \frac{D_{n1}}{D_{d1}} \\
\vdots & \vdots & \ddots & \vdots \\
x_1 & \frac{D_{n1}}{D_{d1}} & \vdots & \frac{D_{n1}}{D_{d1}} \\
x_m & \frac{D_{n1}}{D_{d1}} & \vdots & \frac{D_{n1}}{D_{d1}}
\end{pmatrix}
\begin{pmatrix}
t_1 \\
\vdots \\
t_m
\end{pmatrix} = \begin{pmatrix}
t_1 \\
\vdots \\
t_m
\end{pmatrix}.
\end{align*}
$$

Next, the load factor $\tilde{c}$ of $\tilde{y}$ is calculated as

$$\tilde{c} = \frac{\tilde{y}' \tilde{t}}{(\tilde{t}' \tilde{t})} = \frac{Dd \times (D_{n1}/D_{di})y_1t_1 + \cdots + Dd \times (D_{nm}/D_{di})y_mt_m}{Dd \times (D_{d1}/D_{di})t_1^2 + \cdots + Dd \times (D_{d1}/D_{di})t_m^2}.$$  

In the same way, the load factor of $\tilde{x}$ can be obtained as follows:

$$\tilde{p} = \frac{Xt'(D \cdot t)}{(tr(D \cdot t))}.$$

For the final output model of the PL algorithm, the parameter matrix $b$ is only related to the weight vector $W$ and the load factors $P$ and $C$.

The following continuous steady-state system modeling problem is considered. If it is assumed that data $x \in X < \mathbb{R}^p$ is randomly selected according to an unknown but fixed probability distribution $D$, a corresponding real number $g \in \mathbb{R}$ is assigned to each data $x$, so that $(x, g)$ satisfies an unknown functional relationship $y = F(x) \in [-B, B]$. A sample set $S = \{(x_1, y_1), \ldots, (x_m, y_m)\}$ is given, and the goal is to select a model $\tilde{F} \in \tilde{F}$ in a given model set $\tilde{F}$, so that the specified loss function $L(g, F(x))$ is the smallest relative to the probability distribution $D$ on the entire set $X$ value, that is,

$$\tilde{F} = \arg \min_{F \in \tilde{F}} E_{x \in X} L(y, F(x)).$$

The variance function is one of the most commonly used loss functions. In the following discussion, the variance function will be adopted as the loss function, that is $L(y, F(x)) = (1/2)(y - F(x))^2$. The goal then becomes to find a model $F: X \rightarrow \mathbb{R}$ that minimizes the following (generalization) error:
\[ e_{RD}(F) = E_{\gamma,x}\left[ \frac{1}{2}(y - F(x))^2 \right]. \] (6)

In order to solve the problem of minimizing the loss function in formula (5), the model function \( F \) can be regarded as the parameter of the loss function \( L(g, F(x)) \). The gradient descent method in numerical optimization is borrowed to perform a gradient descent search in the model set to find the model function \( \hat{F}(x) = F_{t}(x) = \sum_{i=0}^{T} f_{i}(x) \) that minimizes the loss function \( L(y, F(x)) \). Here, \( f_{0}(x) \) is the initialization function, \( T \) is the number of search iterations, and \( \{ f_{i}(x) \}_{i}^{T} \) is the sequence of incremental functions obtained using the gradient descent algorithm. Among them, \( f_{i}(x) = \rho_{t} \eta_{t}(x) \) is the incremental function obtained by the \( t \)-th search, and \( \eta_{t}(x) = \left[ \frac{\partial F(x)}{\partial F(x)} \right]_{F(x)=F_{t-1}(x)} = \left[ \frac{\partial E_{j}[L(y, F(x)) \mid x]}{\partial F(x)} \right]_{F(x)=F_{t-1}(x)} \) represents the gradient direction of the loss function. We set \( \phi(F) = E_{\gamma,x}[L(y, F(x))] = E_{\gamma}[E_{\gamma}[L(y, F(x)) \mid x]] \) and \( \phi(F(x)) = E_{\gamma}[L(y, F(x)) \mid x] \), and then, we have

\[
\eta_{t}(x) = \left[ \frac{\partial F(x)}{\partial F(x)} \right]_{F(x)=F_{t-1}(x)}
\]

(7)

If sufficient regularity is assumed, the differential and integral operators in formula (7) can be interchanged, and we can get

\[
\eta_{t}(x) = E_{\gamma}\left[ \frac{\partial L(y, F(x))}{\partial F(x)} \right]_{F(x)=F_{t-1}(x)}. \]

(8)

Moreover, \( \rho_{t} \) represents the best step size \( \rho_{t} = \arg\min_{\rho} E_{\gamma,x}[L(y, F_{t-1}(x) - \rho \eta_{t}(x))] \) for searching along the negative gradient direction, and among them, there is \( F_{t-1}(x) = \sum_{i=0}^{T-1} f_{i}(x) \). Since the true functional relationship \( * \) and probability distribution \( D \) are unknown, it is actually impossible to directly use equation (8) to solve \( \hat{g}_{t}(x) \). However, the gradient descent direction at each sample point can be estimated from the sample set:

\[
-\hat{g}_{t}(x_{i}) = \left[ \frac{\partial L(y_{i}, F(x_{i}))}{\partial F(x_{i})} \right]_{F(x)} = y_{i} - F_{t-1}(x_{i}) \]

(9)

They must be generalized over the entire domain. By solving, we get

\[
\alpha_{t} = \underset{a_{t}, \beta_{t}}{\arg\min} \sum_{i=0}^{m} \left[ -g_{t}(x_{i}) - \beta h(x_{i}; a) \right]^{2}. \]

(10)

The (fitting) error of the real-valued function \( F \) on the sample set \( S = \{ (x_{1}, y_{1}), \ldots, (x_{m}, y_{m}) \} \) is defined as

\[
e_{RD}(F) = \frac{1}{2} \sum_{i=1}^{m} \left( y_{i} - F(x_{i}) \right)^{2}. \]

(11)

Then, the gradient descent direction on the sample point represented by formula (9) becomes

\[
-\hat{g}_{t}(x_{i}) = \left[ \frac{\partial L(y_{i}, F(x_{i}))}{\partial F(x_{i})} \right]_{F(x)=F_{t-1}(x)} = y_{i} - F_{t-1}(x_{i}) \]

(12)

A new variable is introduced to represent the degree of association between any submodel \( h_{i}(x) \) in the model set \( H \), and the negative gradient \( \{ y_{i} \}_{i}^{m} \). \( \lambda_{i} \) is defined as follows:

\[
\lambda_{i} = \frac{\left| \sum_{i=1}^{m} h_{i}(x_{i}) \hat{y}_{i} \right|}{\sqrt{\left( \sum_{i=1}^{m} h_{i}^{2}(x_{i}) \right)}}. \]

(13)

Therefore, according to Cauchy’s inequality, we can know \( \lambda_{i} \in [0, 1] \). If \( \hat{y}_{i}, i = 1, 2, \ldots, m \) is seen as a vector with \( m \) components, \( h_{i}(x_{i}), i = 1, 2, \ldots, m \) is seen as another vector with \( m \) components. Then, when there is \( \lambda_{i} = 1 \), it means that the directions of these two vectors are completely consistent or completely opposite, that is, the submodel \( h(x) \) is completely correlated with \( y \). However, when there is \( \lambda_{i} = 0 \), it means that the directions of these two vectors are completely perpendicular; that is to say, there is absolutely no relationship between the submodel \( h(x) \) and \( y \).

A parameter \( \theta_{0} \in [0, 1] \) is set in the LS-Ensem algorithm and \( \lambda_{0} \) is used to ensure that each selected submodel has a certain relationship with the gradient descent direction. Obviously, the minimum degree of correlation between the submodels and the direction of gradient descent is determined by the chosen parameter \( \theta_{0} \).

At the same time, the optimal search step \( \rho_{t} \) is referred to as the iteration coefficient \( \rho_{t} \) in the future, which becomes

\[
\rho_{t} = \arg\min_{\rho} \sum_{i=1}^{m} \left( \hat{y}_{i} - \rho h_{i}(x_{i}) \right)^{2}. \]

(14)

The solution of formula (14) is actually a linear optimization problem, and the analytical solution of this problem can be obtained according to the optimization method:

\[
\rho_{t} = \frac{\sum_{i=1}^{m} h_{i}(x_{i}) \hat{y}_{i}}{\sum_{i=1}^{m} h_{i}^{2}(x_{i})}. \]

(15)

Then, there is

\[
F_{1}(x) = F_{t-1}(x) + \rho_{t} h_{i}(x). \]

(16)

There are still certain requirements for the selection of submodels in the LS-Ensem algorithm; that is, \( \lambda_{i} \geq \lambda_{0} \) needs to be satisfied.

The convergence of the LS-Ensem algorithm is analyzed. It can be proved that the LS-Ensem algorithm has the following properties in the iterative process: the fitting error of the sample decreases monotonically with the increase in the number of iterations.

**Theorem 1.** In the iterative process of the LS-Ensem algorithm, if there is \( \lambda_{0} > 0 \), the following relation holds:
\[ \lambda_i \geq \lambda_0, \forall t. \]  

Then, the fitting error of the sample order \( \varepsilon_{\delta}(F_t) \) decreases at the rate of \( \lambda_i^2 \).

**Proof.** According to the definition of (11) and formula (16), there is

\[
\varepsilon_{\delta}(F) = \frac{1}{2} \sum_{i=1}^{m} (y_i - F(x_i))^2
\]

\[
= \frac{1}{2} \sum_{i=1}^{m} (y_i - F_{t-1}(x)_i - \rho_i h_i(x_i))^2
\]

\[
= \varepsilon_{\delta}(F_{t-1}) - \frac{2\rho_i}{2} \sum_{i=1}^{m} h_i(x_i) \hat{y}_i + 2 \sum_{i=1}^{m} h_i^2(x_i).
\]

According to formula (15), the expression of \( \rho_i \) is substituted into the above formula and referring to the definition of (11), we can get

\[
\varepsilon_{\delta}(F_t) = \varepsilon_{\delta}(F_{t-1}) - \frac{1}{2} \left( \sum_{i=1}^{m} h_i(x_i) \hat{y}_i \right)^2
\]

\[
= \left[ 1 - \frac{(\sum_{i=1}^{m} h_i(x_i) \hat{y}_i)^2}{(\sum_{i=1}^{m} h_i^2(x_i))} \right] \varepsilon_{\delta}(F_{t-1})
\]

\[
= (1 - \lambda_i^2) \varepsilon_{\delta}(F_{t-1}).
\]

According to Cauchy's inequality, we can know that

\[ \lambda_i^2 = (\sum_{i=1}^{m} h_i(x_i) \hat{y}_i) / (\sum_{i=1}^{m} \hat{y}_i^2) (\sum_{i=1}^{m} h_i^2(x_i)) \leq 1, \]

and the algorithm requires that the child obtained in each round of iteration satisfies \( \lambda_i \geq \lambda_0 > 0 \), so there is

\[ \varepsilon_{\delta}(F_t) = (1 - \lambda_i^2) \varepsilon_{\delta}(F_{t-1}) < \varepsilon_{\delta}(F_{t-1}). \]

(20)

It can be seen from Theorem 1 that, under the given conditions, the fitting error on the samples in the LS-Ensem algorithm decreases with the increase in the number of iterations. According to this property, it can be further proved that the fitting error of the LS-Ensem algorithm to the sample can be arbitrarily small after a finite number of iterations.

**Theorem 2.** We set the sample \( (x, y) \) to satisfy some unknown bounded function relation \( y = F^*(x) \in [-B, B] \). If there is \( \lambda_0 > 0 \) so that formula (17) holds, for any given positive real number \( \epsilon \), the fitting error of the corresponding LS-Ensem algorithm on \( FT \) does not exceed \( \epsilon \) after several iterations. Among them, \( \lambda_{\min} \geq \lambda_0 > 0 \) is the infimum of \( \lambda_i, t = 1, \ldots, T \) obtained in the iterative process:

\[
T = \left( \frac{\ln(2/\epsilon/mB^2)}{\ln(1 - \lambda_{\min}^2)} \right).
\]

**Proof.** Since there is \( F_0(x) = \bar{y} = (1/m) \sum_{i=1}^{m} y_i \), there is

\[
e_{\delta}(F_0) = \frac{1}{2} \left[ (y_1 - \bar{y})^2 + (y_2 - \bar{y})^2 + \cdots + (y_m - \bar{y})^2 \right]
\]

\[= \frac{1}{2} \left[ y_1^2 + \cdots + y_m^2 + m\bar{y}^2 - 2(y_1 + \cdots + y_m)\bar{y} \right]
\]

\[< \frac{1}{2} \left[ y_1^2 + \cdots + y_m^2 - m\bar{y}^2 \right]
\]

\[\leq \frac{1}{2} \left( y_1^2 + \cdots + y_m^2 \right) \leq \frac{m}{2}B^2.
\]

(22)

Since there is \( \lambda_i \geq \lambda_0 \) in each iteration, there must be an infimum \( \lambda_{\min} \) that satisfies \( \lambda_{\min} \geq \lambda_0 > 0 \). According to Theorem 1, after \( t \) iterations, there is \( [17] \)

\[e_{\delta}(F_t) < e_{\delta}(F_0) \cdot (1 - \lambda_{\min}^2)^t \]

\[\leq \frac{m}{2}B^2 \cdot (1 - \lambda_{\min}^2)^t. \]

If \( T \) is chosen such that there is \( (m/2)B^2 \cdot (1 - \lambda_{\min}^2)^T \leq \epsilon \), there must be \( e_{\delta}(F_T) \leq \epsilon \). By taking the logarithm of both sides of \( (m/2)B^2 \cdot (1 - \lambda_{\min}^2)^T \leq \epsilon \), we can get

\[T \ln(1 - \lambda_{\min}) \leq \ln\left( \frac{2 \epsilon}{mB^2} \right). \]

Since there is \( T \geq \frac{\ln(2/\epsilon/mB^2)}{\ln(1 - \lambda_{\min})}. \]

(25)

Theorem 2 shows that under the given conditions, the LS-Ensem algorithm is convergent. It can be further proved that although the LS-Ensem algorithm minimizes the fitting error \( e_{\delta}(F) \) for a given sample as the objective function, it still has good generalization performance. That is to say, when the number of samples and the number of iterations meet certain requirements, the model \( F \) obtained by the algorithm will make the generalization error \( \sigma_{\delta}(F_T) \) defined in (6) arbitrarily small over the entire input space.

**Theorem 3.** If it is assumed that all data \( x \in X \subset R^n \) in the input space are randomly selected according to an unknown but fixed probability distribution \( D \), each data \( x \) corresponds to a real number \( y \in R \), which satisfies an unknown functional relationship \( y = F^*(x) \in [-B, B] \). The optimal search size obtained in each iteration satisfies \( \delta \leq C \). Then, for \( V_0 < 0 \), when the number of samples satisfies

\[m = m_L (\epsilon, \delta) \geq \frac{C_1K^4}{\epsilon^4} \ln \left( \frac{4}{\delta} \right) \]

\[+ \frac{C_1K^6q}{\epsilon^4} \left( \frac{C_1K^6}{\epsilon^4} \right) \ln \left( \frac{C_1K^6[C_2K^6/\epsilon^2]}{\epsilon^4} \right). \]

(26)

The model \( F_T \) obtained after
\[ T = \left\lfloor \frac{\ln(\epsilon / \text{mB}^2)}{\ln(1 - \lambda_{\text{min}}^2)} \right\rfloor. \]  

(27)

Iterations satisfy \( e_{\text{erD}}(F_T^r) \leq \epsilon \) with a probability not less than \( 1 - \delta \) in the entire input space.

Proof. \( \epsilon \) and \( \delta \) are given so that the sample set \( S \) contains \( m = m_L(\epsilon, \delta) \) samples. First, according to Theorem 1, after \( T = \left\lfloor \frac{\ln(\epsilon / \text{mB}^2)}{\ln(1 - \lambda_{\text{min}}^2)} \right\rfloor \) iterations, the fitting error on the sample set is called \( e_{\text{erS}}(F_T^r) < \epsilon/2 \).

The sample and model \( F_T^r \) will be normalized below. Since the value range of \( h \in H = [-1, 1] \) and there is \( \rho_i \leq C \), we can get [18]

\[ |F_T(x)| \leq \sum_{i=1}^{T} \rho_i \leq |C|T + |\mathcal{F}|. \]  

(28)

\( y \in [-B, B] \) is in the sample, so we set

\[ K = \max((|C|T + |\mathcal{F}|) + B). \]  

(29)

\( K \) normalizes the sample and model \( F_T \) to get \( y_i' = (1/2K)y + (1/2), F_T = (1/2K)F_T + (1/2) \). Then, there is

\[ e_{\text{erD}}(F_T^r) = E \left( y_T - F_T^r(x) \right)^2 \]

\[ = \frac{1}{4K^2} e_{\text{erD}}(F_T) \]

\[ e_{\text{erS}}(F_T^r) = \frac{1}{4K} e_{\text{erS}}(F_T). \]  

(30)

The model set to which the model \( F_T \) belongs is represented by \( F \), and the function set to which the processed function \( F_T^r \) belongs is represented by \( F' \) [19].

Then, there is

\[ P^n \left\{ \exists F_T \in F: |e_{\text{erD}}(F_T) - e_{\text{erS}}(F_T)| \geq \frac{\epsilon}{2} \right\} \]

\[ = P^n \left\{ \exists F_T \in F: |e_{\text{erD}}(F_T) - e_{\text{erS}}(F_T)| \geq \frac{\epsilon}{8K^2} \right\} \]

\[ \leq 4N_1 \left( \frac{\epsilon}{2K^2}, 2K^2, 2m \right) \exp \left( -\frac{m\epsilon^2}{211K^4} \right) \]

\[ \leq 4N_2 \left( \frac{\epsilon}{2K^2}, 2K^2, 2m \right) \exp \left( -\frac{m\epsilon^2}{211K^4} \right) \]

\[ \leq 4N_3 \left( \frac{\epsilon}{2K^2}, 2K^2, 2m \right) \exp \left( -\frac{m\epsilon^2}{211K^4} \right) \]

\[ \leq 4 \left( \frac{em4k^3}{\epsilon q} \right)^{q[1/2K^6\epsilon^2]} \exp \left( -\frac{m\epsilon^2}{k_3K^4} \right). \]  

(31)

Among them, \( P^n \) represents the joint distribution of \( m \) independent random samples. If we want to make the above formula less than \( \delta \), that is,

\[ 4 \left( \frac{em4k^3}{\epsilon q} \right)^{q[1/2K^6\epsilon^2]} \exp \left( -\frac{m\epsilon^2}{k_3K^4} \right) < \delta. \]  

(32)

Then, there should be

\[ q \left[ \frac{C^2K^6}{\epsilon^2} \right] \ln \left( \frac{em4k^3}{\epsilon q} \right) - \left( -\frac{m\epsilon^2}{k_3K^4} \right) \leq \ln \left( \frac{\delta}{4} \right). \]  

(33)

The value of \( mL \) that makes inequality (33) true is found below. We set \( b = q[C^2K^6/\epsilon^2] \), \( A = mL \), and \( B = (\epsilon^2/2k_3K^4b) \). The above A and B are substituted into \( x \) and \( y \) to get, respectively:

\[ \frac{mL}{2k_3K^4} \geq \ln \left( \frac{4^q}{\delta} \right) + b \ln \left( \frac{\epsilon^2mL}{2k_3K^4b} \right). \]  

(34)

Furthermore, there are

\[ \frac{mL}{2k_3K^4} \geq \ln \left( \frac{4^q}{\delta} \right) + b \ln \left( \frac{4k_42K^6\epsilon^2b}{\epsilon^2} \right). \]  

(35)

If the following inequality holds, there is

\[ \frac{mL}{2k_3K^4} \geq \ln \left( \frac{4^q}{\delta} \right) + b \ln \left( \frac{4k_42K^6\epsilon^2b}{\epsilon^2} \right) \]

\[ = \ln \left( \frac{4^q}{\delta} \right) + q \left[ \frac{C^2K^6}{\epsilon^2} \right] \ln \left( \frac{em4k^3}{\epsilon q} \right). \]  

(36)

Inequalities (35) and (36) are added together to get

\[ \frac{mL}{2k_3K^4} \geq \ln \left( \frac{4^q}{\delta} \right) + b \ln \left( \frac{\epsilon^2mL}{2k_3K^4b} \times \frac{4k_42K^6\epsilon^2b}{\epsilon^2} \right) \]

\[ \geq -\ln \left( \frac{4^q}{\delta} \right) + q \left[ \frac{C^2K^6}{\epsilon^2} \right] \ln \left( \frac{em4k^3}{\epsilon q} \right). \]  

(37)

This is inequality (33). \( C_1 = 2k_3K^4 \) is substituted into inequality (36) to get

\[ mL \geq C_1K^4 \left[ \frac{4^q}{\delta} \right] \ln \left( \frac{em4k^3}{\epsilon q} \right) \]

\[ + C_2K^4 \left[ \frac{C^2K^6}{\epsilon^2} \right] \ln \left( \frac{em4k^3}{\epsilon q} \right). \]  

(38)

This is inequality (26). Therefore, when inequality (26) holds, there is

\[ P^n \left\{ \exists F_T \in F: |e_{\text{erD}}(F_T) - e_{\text{erS}}(F_T)| < \frac{\epsilon}{2} \right\} \geq 1 - \delta. \]  

(39)

Among them, \( T = \left\lfloor \ln(\epsilon / \text{mB}^2)/\ln(1 - \lambda_{\text{min}}^2) \right\rfloor \)

At the same time, because of \( e_{\text{erS}}(F_T) \leq \epsilon/2 \), there is
4. Teaching Effect Evaluation Mechanism of English MOOC Combined with LS-SEM Intelligent System

Based on the research results of the blended teaching model at home and abroad, the model in this study is divided into three stages (as shown in Figure 1): preparation, activity, and evaluation.

This article proposes a new teaching method, also known as the inverted classroom. It refers to disrupting the conventional teaching order, putting the course content after class, allowing students to learn or complete in advance according to the teaching task, and the teacher answering questions in the classroom. The flipped classroom breaks through the traditional teaching mode and allows students to learn independently through the online teaching platform. For example, before class, teachers can publish a video of class introduction on the online teaching platform, and teachers and students should discuss and analyze the questions raised by the video content, so as to promote real-time communication between teachers and students. The teaching process of the combination of flipped classroom and MOOC is shown in Figure 2.

The general introduction of the English MOOC design is shown in Figure 3. (1) Teaching method: online self-learning (microvideo, PPT, audio, and other electronic network resources) + classroom teaching (problem-based classroom discussions, reports, summaries, etc.). (2) Teaching objectives: to learn language skills through a combination of autonomy and practice, and to improve students’ comprehensive English application ability. (3) Teaching content: based on the “Comprehensive Course of College English in the New Century,” a college English MOOC teaching course with eight thematic units has been independently developed and designed. (4) Class schedule: a total of 40 class hours, including 2 class hours for course introduction (detailed introduction to the course, teaching calendar, teaching plan, classroom group discussion methods, and video viewing content, offline Q&A, etc.) and 32 class hours for unit theme teaching (4 class hours per unit), and 6 class hours for final review (the problems in this semester are sorted out, summarized, and the key points and difficulties are summarized). (5) Assessment method: online test + usual performance + final test.

Blended learning curriculum design can be divided into three stages: front-end analysis, activity and resource design, and teaching evaluation design. The flipped classroom teaching mode emphasizes the replanning of preclass, in-class, and after-class links and the subversion of traditional knowledge transfer and knowledge internalization. Based on this, this study designed the basic framework of the MOOC-based flipped classroom teaching model, as shown in Figure 4.

This article proposes a relatively simplified and fast online course production and presentation process. As shown in Figure 5, first, the user uploads the courseware

\[ P^n \exists F_T \in F: |er_D(F_T)| < \epsilon \geq 1 - \delta. \]
made to the server through the web-side platform, and the system automatically extracts the courseware into a picture set by calling a third-party service. After that, users can use mobile devices.

This article divides the “Qing MOOC” system into five functional modules: course management module, user management module, course playback module, course comment and evaluation module, and course production module. As shown in Figure 6, the course management module and the course production module are mainly for the user to complete the whole supply process of the course, including the uploading, saving and conversion of
courseware, course making, publishing, editing, deleting and classifying, and other steps.

The overall function realization process of the system is roughly shown in Figure 7.

The platform is deployed based on the Tomcat6 web server, which effectively reduces the overall deployment cost and supports load balancing technology. The platform needs to publish a large number of courses and at the same time...
needs to support a large number of students to watch, and it is decided that the platform uses Flex4 and Red5 video server technology to play video files. The architecture of the platform is shown in Figure 8.

On the basis of the above research, this article combines the LS-SEM intelligent system to evaluate the effect of the English MOOC teaching constructed in this article and counts the teaching effect of the English MOOC teaching system in this paper, and obtains the evaluation results shown in Table 1 and Figure 9.
From Figure 9, we can draw the evaluation results of the LS-SEM intelligent system on the English MOOC teaching constructed in this article. A total of 54 sets of simulation experiments were carried out. It can be seen from the data in the figure that the minimum evaluation score is not less than 78 points. From the perspective of traditional spinning methods, the LS-SEM intelligent system has a high improvement in the teaching effect of English MOOCs and has certain practicality.

The above research shows that the English MOOC teaching effect evaluation mechanism combined with the LS-SEM intelligent system can play an important role in English MOOC teaching.

5. Conclusion

College English follow-up courses refer to content-based college English courses offered to students by colleges and universities after completing the teaching of the basic stage of college English, including ESP courses, subject English courses, and general education courses in college English. In recent years, the academic circles have carried out a lot of thinking and discussion on the curriculum setting and course orientation of college English follow-up courses, and the analysis of relevant learners’ needs has also confirmed the necessity of carrying out follow-up course teaching in the college English stage. In order to analyze the teaching effect of English MOOC, this article constructs an evaluation mechanism based on the LS-SEN intelligent system and obtains an intelligent MOOC system. The experimental research results show that the English MOOC teaching effect evaluation mechanism combined with the LS-SEM intelligent system can play an important role in English MOOC teaching.

Data Availability

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

Conflicts of Interest

The author declares no conflicts of interest.
Acknowledgments

This work was supported by the Shangqiu Institute of Technology.

References


