Advancements In English Teaching: Integrating Digital Media And Machine Learning For Interactive Learning

Hongli Niu

College of Foreign Language, Inner Mongolia University of Technology, Hohhot, 010050, China

Corresponding author. E-mail: hongli_niu@outlook.com; ghhl_0822@163.com

Received: January 05, 2026; Accepted: March 22, 2026

Integrating digital media and interactive learning has transformed English instruction, enhancing teaching and learner engagement. Limited resources hinder spoken proficiency, but advances in digital media and ML provide solutions. This study proposes an ML-based pronunciation model using Intelligent Honeybee Mating tuned FSVM (IHM-FSVM) in an online interactive framework. Students used multimedia instruction via apps, videos, and gamified exercises. Speech data were preprocessed with Wiener filtering and features extracted with MFCC. IHM-FSVM outperformed conventional methods, achieving 97.9% accuracy, 95.9% recall, 96.2% precision, and 96.5% F1-score, demonstrating its effectiveness in improving interactive English learning and pronunciation.

Keywords: English Teaching; Digital Media; Interactive Learning; Intelligent Honeybee Mating tuned Flexible Support Vector Machine (IHM-FSVM)

http://dx.doi.org/10.6180/jase.202609_32.049

1. Introduction

English is a global language, valued as a second language. Simulation-based learning supports engagement [1] and effective teaching improves motivation and competence [2]. Many learners struggle with spoken English due to accent interference, limited speaking practice, and delayed or insufficient feedback, reducing pronunciation accuracy, fluency, and confidence. Teaching is now learner-centered [3]. Digital technologies support goals [4], boost engagement, enable personalized learning [5, 6], strengthen interaction, and develop skills [7], showing how interactive learning and DM transform English instruction [8]. Most ML-based pronunciation models need large datasets and high computation, lacking adaptive feedback. ML enables scalable, adaptive, learner-centered instruction with personalized assessment. Digital technologies and ML enhance traditional English classrooms, overcome resource limitations and enable interactive, personalized learning to improve students' spoken English proficiency. Contribution

- Predict English pronunciation using IHM-FSVM with personalized feedback.
- Use gamified exercises and video lessons for engaging learning.
- Analyze speech data with MFCC and Wiener filtering for accuracy.

The organization of this research consists of the following structure: Part 2 covers Materials and Methods, Part 3 presents the Results, Part 4 provides the Discussion and Part 5 concludes the study.

2. Materials and methods

An online training course [9] helped English teachers use digital media (DM) via needs analysis, course development, trials, and evaluation. Seventy-one teachers were selected with G*Power, and data were collected from assessments, exams, and questionnaires. Training improved DM proficiency. AI applications in ELT [10–12] supported accessibility, personalized instruction, and virtual assistance. DM access improved competence and reduced cultural anxiety [13], while teacher media use was studied [14]. Emerging technologies and models [15–21] enhanced learning, supervision, engagement, and outcomes.

2.1. Methodology

English data were collected, denoised with Wiener filtering, features extracted via MFCC, and IHM-FSVM predicted pronunciation. The methodology is shown in Fig. 1.

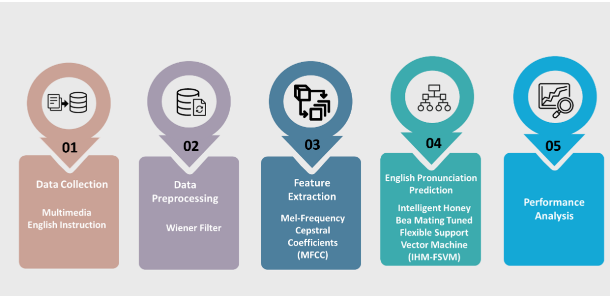

Fig. 1. Methodology Flow Diagram

2.2. Data Collection

Data were collected via multimedia platforms, including videos, apps and gamified exercises, to improve engagement and pronunciation. Students practiced key words and audio recordings captured each attempt.

2.3. Data preprocessing using Weiner Filtering

Wiener filtering reduces noise, improves audio and recognition, supporting listening and engagement in ML-based English learning. WF is widely used in speech enhancement, retaining essential speech features for reliable analysis and effective pronunciation assessment. It filters noise using spectral features of the signal and noise as linear stochastic processes. This linear filter is applied to the expected signal using the coefficients X_L. The input signaly(m), consists of noise $w(m)$ as shown in Eq. (1).

$$y(m) = e(m) + w(m) \tag{1}$$

The adaptive method attempts to modify the weights X_L such that the mean square error is reduced by Eqs. (2) and (3).

$$f = \min\left(F\left(f(m)^2\right)\right) \tag{2}$$

$$f(m) = z(m) - e(m) \tag{3}$$

Eq. (4) is used by a l tap discrete Wiener filter to determine the value of $z(m)$.

$$z(m) = \sum_{l=0}^{M-1} X_L(e(m-l) * w(m-l)) \tag{4}$$

The Wiener filter uses the Wiener-Hopf equation for weights (Eq. (5)).

$$\sum_{k=0}^{q-1} X_{ol} r_{yy}(l-1) = r_{y\varepsilon}(-k) \tag{5}$$

Where r_{yy} is the autocorrelation function of $y(m)$, and r_{ye} is the cross-correlation function between $y(m)$ and $e(m)$.

2.4. Feature Extraction using Mel-frequency cepstral coefficients (MFCC)

MFCCs capture phonetic and articulatory variations, preserve subtle pronunciation cues, resist noise and speaker variability, and provide a compact representation for accurate pronunciation assessment. MFCC, widely used in speech recognition, is extracted via channel banks, framing, pre-emphasis, windowing, logarithmic scaling, filter banks, and discrete cosine transform with FIR pre-emphasis. The basic layout of the first-order FIR channel change work in the area was provided in Eq. (6).

$$F(y) = 1 - \beta y^{-1} \tag{6}$$

The Coefficient β that lies between the range of $0 \leq \delta \leq 1$ is described, as shown in Eqs. (7) and (8).

$$f\left(w_j^{s'}\right) = \rho\left(w_j^{s'}\right) - \beta\rho\left(w_j^{s'} - 1\right) \tag{7}$$

$$f\left(w_j^{u'}\right) = \rho\left(w_j^{u'}\right) - \beta\rho\left(w_j^{u'} - 1\right) \tag{8}$$

Eq. (7) trains and Eq. (8) tests. Short-time analysis preserves features, hamming windows reduce discontinuities, and Mel-scale triangular filter banks compute weighted spectral outputs. The Mel is assessed for the pre-set recurrence f in HZ by utilizing the resultant condition through Eq. (9).

$$F(\text{Mel}) = 2595 \times \log_{10}\left(1 + \frac{f}{700}\right) \tag{9}$$

Logarithmic scaling compresses channel outputs, shown in Eq. (10).

$$Y_{f_b s}(\ln) = \ln\left(Y_{f_b s}\right), 1 \leq Y_{f_b s} \leq Y_{f_b s} \tag{10}$$

2.5. English pronunciation prediction using IHM-FSVM

IHM-FSVM uses ML and optimization to improve English pronunciation prediction, with FSVM handling complex non-linear linguistic patterns. In FSVM, each training sample is assigned a flexible membership value that reflects its reliability, allowing the model to reduce the influence of noisy or ambiguous pronunciation data while preserving meaningful speech patterns. Unlike traditional SVM, HMM, or neural network models, IHM-FSVM uses adaptive honeybee mating to adjust kernel and error-insensitive parameters, modeling subtle pronunciation variations [22]. MFCC features are normalized and input to FSVM, while IHM optimizes parameters over 100 iterations, providing scaffolded, personalized feedback and adjusting tasks based on real-time performance.

2.5.1. Flexible Support Vector Machine (FSVM)

FSVM adjusts error tolerances, enabling accurate assessment, personalized feedback, and interactive English teaching. Eq. (11) uses measured and expected runoff values.

$$R_p = \frac{1}{m}\sum_{j=1}^{m} u_j \tag{11}$$

$$u_j = \begin{cases} 1, \frac{|z_j - z_j'|}{z_j} \times 100\% \leq 20\% \\ 0, \frac{|z_j - z_j'|}{z_j} \times 100\% \leq 20\% \end{cases}, (j = 1, 2, ..m) \tag{12}$$

Eq. (12) updates the dual function.

$$\text{Max}\left[\sum_{j=1}^{k}\left(\propto_j^+ - \propto_j^-\right) z_j - \varepsilon\left(z_j\right)\sum_{j=1}^{k}\left(\propto_j^+ - \propto_j^-\right) \right.$$
$$\left. -\frac{1}{2}\sum_{j,k}\left(\propto_j^+ - \propto_j^-\right)\left(\propto_k^+ - \propto_k^-\right) L\left(Y_j, Y_k\right)\right]$$

Where $L\left(Y_j, Y_k\right)$ is the kernel function, while $\propto_j^+$and $\propto_j^-$ are Lagrange multipliers. To discusses the two common functions, $\varepsilon\left(z_j\right)$, which are displayed in Eqs. (13) and (14).

$$\varepsilon_1\left(z_j\right) = b_1 z_j + b_2 \tag{13}$$

$$\varepsilon_1\left(z_j\right) = c_1 f^{c_2 z_j} + C_3 \tag{14}$$

2.5.2. Intelligent Honeybee Mating (IHM)

In the proposed model, IHM was employed to iteratively tune key FSVM hyperparameters, including the penalty factor, kernel width, and adaptive error-insensitive threshold, using pronunciation classification error as the fitness criterion to select optimal solutions. By refining these parameters, the optimization process directly improves the model's ability to distinguish subtle phonetic deviations, enabling more precise pronunciation error detection and corrective feedback. Mutation reduces stagnation by generating four mutant vectors per step using distinct rules (Eqs. (15) to (19)) for diverse, adaptable learning.

$$Y^j_{\text{mutant 1}} = Y^j_{r1} + F_1 \times \left(Y^j_{\text{queen}} - Y^j_{r2}\right) + F_1 \times \left(Y^j_{r3} - Y^j_{r4}\right) \tag{15}$$

$$Y^j_{\text{mutant 2}} = Y^j_{\text{queen}} + F_2 \times \left(Y^j_{r1} - Y^j_{r2}\right) \tag{16}$$

$$Y^j_{\text{mutant 3}} = Y^j_{r1} + F_3 \times \left(Y^j_{r2} - Y^j_{r3}\right) + F_3 \times \left(Y^j_{r1} - Y^j_{r4}\right) \tag{17}$$

$$Y^j_{\text{mutant 4}} = \frac{\left(Y^j_{r2} + Y^j_{r2} + Y^j_{r3}\right)}{3} + (\beta_2 - \beta_1)\left(Y^j_{r1} - Y^j_{r2}\right) + (\beta_3 - \beta_2)\left(Y^j_{r2} - Y^j_{r3}\right) + (\beta_1 - \beta_3)\left(Y^j_{r3} - Y^j_{r1}\right) \tag{18}$$

$$\beta_1 = \frac{\left|f\left(Y^j_{r2}\right)\right|}{\beta^*}, \beta_2 = \frac{\left|f\left(Y^j_{r2}\right)\right|}{\beta^*}, \beta_3 = \frac{\left|f\left(Y^j_{r3}\right)\right|}{\beta^*} \tag{19}$$

Where $\beta^* = \left|f\left(Y^j_{r1}\right)\right| + \left|f\left(Y^j_{r2}\right)\right| + \left|f\left(Y^j_{r3}\right)\right|$. Mutant vector replaces target if cost is lower (Eq. (20)).

$$Y^{j+1}_{\text{queen}} = \begin{cases} Y^j_{\text{mut,Best}} \text{ if } f\left(Y^l_{\text{new,Best}}\right) \leq f\left(Y^l_{\text{queen}}\right) \\ Y^j_{\text{queen}} \text{ otherwise} \end{cases} \tag{20}$$

Algorithm 1: IHM- FSVM

Input: Dataset for English pronunciation (Z), initial FSVM parameters,

population size (P), max iterations (MaxIter)

Output: Optimized FSVM parameters for pronunciation prediction

Step 1:Initialize:

a. Generate initial population $Y_{\text{population}}$ of size P with random vectors (Y).

b.For each vector Y_i in $Y_{\text{population}}$, evaluate the fitness using FSVM:

fitness (Y_i) = FSVM_prediction_accuracy (Y_i)

c.Set Y_{queen} as the best solution (best fitness).

d. Set tolerance for projected mistake $\varepsilon\left(z_j\right)$ as $0.2 * z_j$.

Step 2: Repeat until MaxIter is reached:

a. Mutation Phase:

For each individual vector Y_i in $Y_{\text{population}}$:

Select four random vectors Y_{r1}, Y_{r2}, Y_{r3}, and Y_{r4} from $Y_{\text{population}}$ (ensuring $Y_{r1} \neq Y_{r2} \neq Y_{r3} \neq Y_{r4}$).

Generate mutant vectors using Eqs. (15) to (18):

Compute $\beta 1, \beta 2$, and $\beta 3$ using Eq. (19)

Clip mutant vectors:

If any element of mutant vectors exceeds bounds, set $Y^j_{\text{mutant } k}$ to $Y_{\max}$

Saved to this PC

or $Y_{\min}$ based on restrictions.

Evaluate fitness for each mutant vector using FSVM:

fitness $(Y_{\text{mutant 1}})$, fitness $(Y_{\text{mutant 1}}2$, fitness $(Y_{\text{mutant 3}})$, fitness $(Y_{\text{mutant 4}})$.

b.Selection Phase:

For each mutant vector Y_mutant:

If fitness (Y_{mutant}) < fitness (Y_i), then update Y_i to Y_{mutant}.

Saved to this PC
If fitness $(Y_{\text{mutant}}) \geq$ fitness (Y_i), retain Y_i.
c. Update Queen Vector Using Eq. (20)
Step 3: Fitness Function Calculation
Calculate the qualifying rate R_p using Eq. (11)
Step 4:Generate FSVM Parameters Using IHM Eq. (12):
a. Initialize the queen vector Y_{queen} and other population members using random initialization.
b. Evaluate the fitness of each member of the population using FSVM.
Step 5: Apply the Crossover Operation
a. For each pair of vectors Y_i and Y_j in the population, perform crossover:
If random_value < crossover_probability:
Use Eq. (13)
b. Evaluate the fitness of the newly generated vectors.
Step 6: Apply the Inversion Operation:
a. For each individual vector Y_i :
If random_value < inversion_probability:
Use Eq. (14)
b. Evaluate the fitness of the newly modified vector.
Step 7: Final Output:
Return the optimized FSVM parameters Y_queen

3. Result and discussion

Learners practiced with ML-driven feedback, receiving personalized guidance to reinforce pronunciation and improve skills.

3.1. Experimental Setup

Python 3.9 was used in a Jupyter notebook environment with GPUs such as NVIDIA Tesla V100 and cloud platforms like Google Colab or AWS EC2.

3.2. Performance Metrics

IHM-FSVM was compared with HMM-ET [21] using the below metrics.

- **Accuracy:** Accuracy measures the percentage of correct predictions, reflecting overall performance. Eq. (21) shows the calculation, with results in Table 1 and Fig. 2.

$$\text{Accuracy} = \frac{\text{True Positives} + \text{True Negatives}}{\text{Total Predictions}} \tag{21}$$

Table 1. Accuracy's Outcomes.

Methods	Accuracy (%)
HMM-ET [21]	96.0%
IHM-FSVM [Proposed]	97.9%

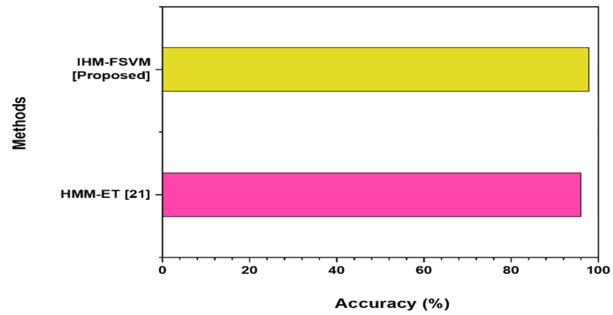

Fig. 2. Accuracy.

- **Precision:** Precision measures correct positive predictions, minimizing false positives. Eq. (22) shows the calculation with results in Table 2 and Fig. 3.

$$\text{Precision} = \frac{\text{True Positives}}{\text{True Positives} + \text{False Positives}} \tag{22}$$

Table 2. Precision's Outcomes.

Methods	Precision (%)
HMM-ET [21]	93.0%
IHM-FSVM [Proposed]	96.2%

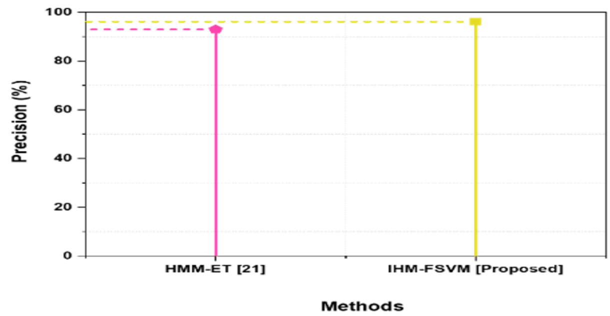

Fig. 3. Precision.

- **Recall:** Recall measures detected errors, minimizing false negatives; Eq. (23) shows calculation, results in Table 3 and Fig. 4.

$$\text{Recall} = \frac{\text{True Positives}}{\text{True Positives} + \text{False Negatives}} \tag{23}$$

Table 3. Recall's Outcomes.

Methods	Recall (%)
HMM-ET [21]	94.0%
IHM-FSVM [Proposed]	95.9%

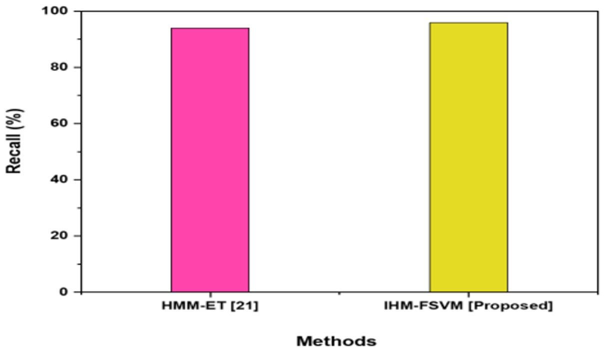

Fig. 4. Recall.

- **F1-Score:** F1-score, the harmonic mean of precision and recall, evaluates error detection reliability; Eq. (24), Table 4, and Fig. 5.

$$\mathrm{F1}-\mathrm{score}=2\times\frac{\mathrm{Precision}\times\mathrm{Recall}}{\mathrm{Precision}+\mathrm{Recall}} \tag{24}$$

Table 4. F1-Score's Outcomes.

Methods	F1-Score (%)
HMM-ET [21]	95.0%
IHM-FSVM [Proposed]	96.5%

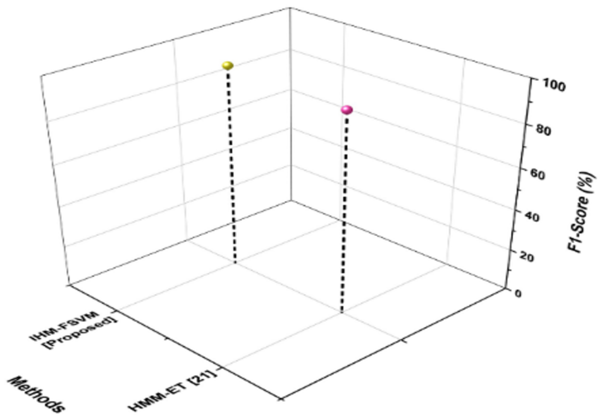

Fig. 5. F1-Score.

4. Discussion

The model's adaptive feedback promotes self-regulated learning and pronunciation improvement. Students reported ease of use, clear error identification, increased confidence, and reduced frustration. The interactive design encouraged participation, and teachers can use it as a supplementary tool with basic training. IHM-FSVM overcomes HMM-ET limits, using adaptive feedback to enhance engagement and pronunciation.

5. Conclusion

The IHM-FSVM with interactive digital media improved English teaching. It outperformed previous models with 97.9% accuracy, 95.9% recall, 96.2% precision, and 96.5% F1-score. This demonstrates how combining digital media and ML enables efficient, responsive, and scalable interactive learning environments. Students benefited from personalized feedback that reinforced correct pronunciation and accelerated learning. Limitations include equitable technology access, large datasets, digital infrastructure, and data privacy. Future work addresses accents and adaptive learning.

References

[1] L. Ma, (2021) *"An immersive context teaching method for college English based on artificial intelligence and machine learning in virtual reality technology"* **Mobile Information Systems *2021***(1): 2637439. DOI: 10.1155/2021/2637439.

[2] W. Pranawengtias, (2022) *"Undergraduate students' motivation on English language learning at Universitas Teknokrat Indonesia"* **Journal of English Language Teaching and Learning** *3*(2): 27–32. DOI: 10.33365/jeltl.v3i2.1956.

[3] B. B. Dash, (2022) *"Digital tools for teaching and learning English language in 21st century"* **International Journal of English and Studies** *4*(2): 8–13. DOI: 10.47311/IJOES.2022.4202.

[4] N. R. Putri and F. M. Sari, (2021) *"Investigating English teaching strategies to reduce online teaching obstacles in the secondary school"* **Journal of English Language Teaching and Learning** *2*(1): 23–31. DOI: 10.33365/jeltl.v2i1.780.

[5] B. Mandasari and D. Aminatun, (2022) *"Investigating teachers' belief and practices toward digital media of English learning during COVID-19 pandemic"* **English Review: Journal of English Education** *10*(2): 475–484. DOI: 10.25134/erjee.v10i2.6248.

[6] C. Rwodzi, L. De Jager, and N. Mpofu, (2020) *"The innovative use of social media for teaching English as a second language"* **Theological Studies** *16*(1): DOI: 10.4102/td.v16i1.702.

[7] T. Eli, (2021) *"Students' perspectives on the use of innovative and interactive teaching methods at the University of Nouakchott Al Aasriya, Mauritania: English department as a case study"* **International Journal of Technology, Innovation and Management** *1*(2): 90–104. DOI: 10.54489/ijtim.v1i2.26.

[8] K. R. N. Malikovna, S. Z. Mirsharapovna, S. M. Shadjalilovna, and A. A. Kakhramonovich, (2022) *"Types of interactive methods in teaching English to students"* **Texas Journal of Multidisciplinary Studies *14***: 1–4.

[9] W. Simpol, P. Paulina, P. Sinthong, and P. Praraksa, (2022) *"Research and development an online training course to enhance teacher competency in creating digital media for teaching English"* **UMGESHIC Conference Proceedings *1***(2): 965–978. DOI: 10.30587/umgeshic.v1i2.3505.

[10] M. Anis, (2023) *"Leveraging artificial intelligence for inclusive English language teaching: Strategies and implications for learner diversity"* **Journal of Multidisciplinary Educational Research *12***(6): 54–70.

[11] Z. Sun, M. Anbarasan, and D. J. C. I. Praveen Kumar, (2021) *"Design of online intelligent English teaching platform based on artificial intelligence techniques"* **Computational Intelligence *37***(3): 1166–1180. DOI: 10.1111/coin.12351.

[12] W. Villegas-Ch, M. Román-Cañizares, and X. Palacios-Pacheco, (2020) *"Improvement of an online education model with the integration of machine learning and data analysis in an LMS"* **Applied Sciences *10***(15): 5371. DOI: 10.3390/app10155371.

[13] A. S. G. Samosir, (2023) *"The meaning of well-being for undergraduate medical students: Perspectives from Indonesia"* **PhD Thesis, University of Glasgow**: DOI: 10.5525/gla.thesis.83604.

[14] F. Wirawan, (2020) *"A study on the teaching media used by the English teacher at SMP Muhammadiyah 2 Malang"* **Jurnal Ilmiah Profesi Pendidikan *5***(2): 89–95. DOI: 10.29303/jipp.v5i2.115.

[15] A. Almufarreh and M. Arshad, (2023) *"Promising emerging technologies for teaching and learning: Recent developments and future challenges"* **Sustainability *15***(8): 6917. DOI: 10.3390/su15086917.

[16] M. Ayu and Z. F. Pratiwi, (2021) *"The implementation of online learning in English language teaching during pandemic: The teachers' voice"* **Journal of Research on Language Education *2***(2): 93–99. DOI: 10.33365/jorle.v2i2.1316.

[17] H. Guo and W. Gao, (2022) *"Metaverse-powered experiential situational English-teaching design: An emotion-based analysis method"* **Frontiers in Psychology *13***: 859159. DOI: 10.3389/fpsyg.2022.859159.

[18] W. Lu, G. N. Vivekananda, and A. Shanthini, (2023) *"Supervision system of English online teaching based on machine learning"* **Progress in Artificial Intelligence *12***(2): 187–198. DOI: 10.1007/s13748-021-00274-y.

[19] P. Zhang, (2023) *"Cloud computing English teaching application platform based on machine learning algorithm"* **Soft Computing**: 1–13. DOI: 10.1007/s00500-023-08589-9.

[20] Q. Yang, (2022) *"Analysis of English cultural teaching model based on machine learning"* **Computational Intelligence and Neuroscience *2022***(1): 7126758. DOI: 10.1155/2022/7126758.

[21] Y. Chen, (2024) *"Application of multimedia technology in teaching English in colleges and universities"* **International Journal of Maritime Engineering *1***(1): 333–346. DOI: 10.5750/ijme.v1i1.1365.

[22] L. Qian, (2022) *"Research on college English teaching and quality evaluation based on data mining technology"* **Journal of Applied Science and Engineering *26***(4): 547–556. DOI: 10.6180/jase.202304_26(4).0011.