Tong S. (2024) ASSESSING ENGLISH LEARNING NEEDS IN SPORTS PHYSICAL EDUCATION AND THE APPLICATION OF MACHINE LEARNING ALGORITHMS. Revista Internacional de Medicina y Ciencias de la Actividad Física y el Deporte vol. 24 (97) pp. 72-89.
DOI: https://doi.org/10.15366/rimcafd2024.97.006

ORIGINAL

ASSESSING ENGLISH LEARNING NEEDS IN SPORTS PHYSICAL EDUCATION AND THE APPLICATION OF MACHINE LEARNING ALGORITHMS

Shuang Tong

Department of Foreign Languages, Nanchong Vocational and Technical College, Nanchong, 637100, Sichuan, China.
E-mail: 15883528656@163.com

Recibido 06 de diciembre de 2023 **Received** December 06, 2023
Aceptado 06 de julio de 2024 **Accepted** July 06, 2024

ABSTRACT

This study explores the English language learning needs and challenges of first-year sport Physical Education students at Huzhou University through a combination of quantitative and qualitative analysis. The research identifies significant disparities in students' English proficiency, particularly in listening and speaking, where they struggle the most. Qualitative findings reveal that students face difficulties in mastering sports-specific terminology and lack confidence in applying English in practical contexts. Furthermore, students expressed dissatisfaction with the theoretical focus of current courses, highlighting the need for more practical and interactive learning experiences. The study underscores the importance of English for students' future careers, especially in the international sports field, and recommends incorporating more sports-related content, enhancing listening and speaking training, and increasing interactive exercises in the curriculum. The findings provide valuable insights for English curriculum reform, emphasizing the need for teaching methods that align with students' practical needs and career aspirations. Future research should explore the effectiveness of these recommended teaching methods and the potential of machine learning in developing personalized learning paths to better support diverse student needs.

KEYWORDS: Sport Physical Education, English Language Learning, Machine Learning

1. INTRODUCTION

In the context of accelerating globalization, English has emerged as the

global lingua franca, assuming an increasingly vital role in various domains, including higher education. For students majoring in physical education, proficiency in English is not merely a linguistic skill but a critical tool that opens doors to international knowledge, best practices, and career opportunities in the global sports industry (Warriner, 2016). English courses tailored for these students serve dual purposes: enhancing their general language proficiency and equipping them with the necessary vocabulary and communication skills to engage effectively with the international sports community. This dual purpose is essential for students to stay abreast of global trends in sports education, contribute meaningfully to international discourse, and build a solid foundation for their professional careers (Giulianotti, 2014). The importance of English proficiency in physical education cannot be overstated, particularly as sports continue to grow as a global industry. Graduates with the ability to communicate effectively in English are better positioned to participate in international conferences, collaborate with peers from different countries, and access a vast body of knowledge published in English (C. Hu, 2022).

Consequently, analyzing the specific English language needs of physical education students and developing effective instructional strategies to meet these needs is not only academically significant but also practically urgent. Such an analysis can inform curriculum design, ensuring that courses are aligned with the demands of both the global sports industry and the educational goals of the students (J. Hu et al., 2024). Despite the recognized importance of English proficiency, the integration of English learning within physical education programs remains suboptimal. The current state of English language education for physical education students in many universities is characterized by traditional teaching methodologies, which often fall short of addressing the actual needs of the students (Kohl III & Cook, 2013). These methods are predominantly lecture-based, focusing on the rote learning of grammar rules and vocabulary, with limited opportunities for students to engage in practical language use. This approach is particularly inadequate for physical education students who require not only general English proficiency but also specialized knowledge in sports-related terminology and concepts. Physical education students are expected to acquire a comprehensive understanding of English, including proficiency in listening, speaking, reading, and writing, all within the context of their field (Pranoto & Suprayogi, 2020). For instance, students need to master sports-specific vocabulary and expressions to effectively communicate in professional sports settings.

However, traditional English courses often do not cater to this need, as they tend to be generic and not tailored to the specific requirements of students in the sports field. Moreover, the emphasis on grammar and vocabulary at the expense of developing listening and speaking skills leaves students underprepared for real-world situations where effective communication is

crucial (Tuong & Dan, 2024). In addition to language skills, physical education students must also develop intercultural communication skills to effectively interact with individuals from diverse cultural backgrounds. This aspect is frequently neglected in traditional English courses, which rarely include components designed to enhance students' understanding of cross-cultural communication (Angelova & Zhao, 2016). The lack of focus on these skills can lead to difficulties in international collaboration, as students may struggle to navigate the cultural nuances of communication in different contexts. Furthermore, the reading and writing components of traditional English courses often do not align with the academic and professional needs of physical education students. These students must be able to read and comprehend complex texts, such as research articles and technical documents, which are often written in English (Valencia, Wixson, & Pearson, 2014). They also need to develop the ability to write clearly and effectively, whether drafting research papers, writing reports, or preparing teaching plans. However, the current instructional strategies do not sufficiently address these specific needs, resulting in a gap between the students' language abilities and the demands of their academic and professional endeavors (Bunch, 2013).

The advent of machine learning has brought about significant advancements in the field of education, offering new tools for analyzing and addressing complex educational challenges. Machine learning algorithms can process large volumes of data to uncover patterns and insights that would be difficult to detect using traditional analytical methods. In the context of English language education for physical education students, machine learning can be leveraged to analyze the diverse and dynamic needs of students, thereby enabling the development of more targeted and effective instructional strategies (Wang, 2024). Machine learning can be applied in various ways to enhance English language education (Kushik, Yevtushenko, & Evtushenko, 2020). For example, algorithms can be used to analyze student performance data, identifying common areas of difficulty and suggesting personalized learning paths. By processing data from assessments, class participation, and even social interactions, machine learning models can provide a more nuanced understanding of each student's strengths and weaknesses (Cope & Kalantzis, 2016). This personalized approach is particularly beneficial for physical education students, who may have varied levels of English proficiency and distinct learning needs based on their backgrounds and career aspirations. Moreover, machine learning can assist in the development of adaptive learning systems that adjust the content and difficulty of coursework in real time, based on the student's progress. Such systems can provide immediate feedback, helping students to correct mistakes and reinforce learning effectively. For physical education students, this could mean more practice with sports-related English terminology, more exposure to relevant case studies (Tendinha et al., 2021), and a greater focus on the specific skills they need to develop for their

future careers. Another promising application of machine learning is in the analysis of student engagement and motivation.

By analyzing behavioral data, such as attendance, participation, and interaction with learning materials, machine learning models can predict student outcomes and identify those who may be at risk of falling behind. This proactive approach allows educators to intervene early, offering additional support or resources to help students stay on track. This study is significant for several reasons. First, it aims to address the gap between the current state of English language education for physical education students and the actual needs of these students. By providing a detailed analysis of these needs, the study will contribute to the development of more effective curricula that are better aligned with the demands of the global sports industry. Second, by exploring the application of machine learning in this context, the study introduces a novel approach to educational needs analysis that has the potential to be applied across various disciplines. Machine learning offers a powerful tool for understanding and addressing the complex and varied needs of students, making education more personalized, efficient, and effective (Frangoudes, Matsangidou, Schiza, Neokleous, & Pattichis, 2022). Finally, this study has broader implications for the field of educational research. It demonstrates how advanced technologies like machine learning can be integrated into educational practices to improve outcomes. The findings of this research could serve as a model for other educational institutions seeking to enhance their curricula through data-driven insights and innovative teaching methods. The objectives of this study are twofold.

First, it seeks to identify and analyze the specific English language learning needs of physical education students. This involves a comprehensive examination of their requirements in listening, speaking, reading, writing, and understanding specialized sports-related English. Second, the study aims to explore the potential of machine learning algorithms in conducting this needs analysis. By applying machine learning techniques, the study will analyze data related to student performance and engagement, identifying key factors that influence learning outcomes. The ultimate goal is to use these insights to inform the development of more effective teaching strategies and curricular designs that can better prepare physical education students for the challenges of the global sports industry.

2. Methodology

This study employs a mixed-methods approach, integrating both qualitative and quantitative research methods to provide a comprehensive analysis of the English language learning needs of physical education students and the application of machine learning algorithms in this context. The research

is designed to capture a wide range of data, including student performance metrics, self-reported language needs, and observational data from classroom settings. The use of mixed methods allows for a more robust and nuanced understanding of the research questions, facilitating the triangulation of findings from different data sources.

2.1 Participants

This study involves a sample of 107 first-year students majoring in Physical Education at Huzhou University. Among these participants, 63 are male, accounting for 58.9% of the total, and 44 are female, comprising 41.1%. All participants are full-time undergraduate students currently enrolled in the university's Physical Education program. As per the curriculum requirements, these students attend 4 hours of English classes each week. The participants were selected based on their grade level and gender to ensure diversity and representativeness within the sample. All participants agreed to participate in the study and provide the necessary data. The following table provides a detailed overview of the gender distribution and weekly English class hours of the participants:

Table 1: Basic Demographics and Weekly English Class Hours of Participants

GENDER	NUMBER OF STUDENTS	PERCENTAGE (%)	WEEKLY ENGLISH CLASS HOURS
MALE	63	58.9	4
FEMALE	44	41.1	4
TOTAL	107	100	4

Table 1 integrates the gender distribution and weekly English class hours, providing a comprehensive view of the study's participants. The data show that all participants, regardless of gender, attend 4 hours of English classes each week. This consistent scheduling across the sample supports the reliability and representativeness of the study, laying a solid foundation for subsequent quantitative and qualitative analyses.

2.3 Data Collection

2.3.1 Quantitative Data Collection

The quantitative data collection process in this study is designed to systematically gather comprehensive information on the English language proficiency and specific language needs of first-year Physical Education students at Huzhou University. This process involves the use of structured questionnaires, standardized language proficiency tests, and academic performance records. The methods and their corresponding data collection indicators are summarized in the following table 2:

Table 2: Quantitative Data Collection Methods and Indicators

DATA COLLECTION METHOD	SPECIFIC CONTENT/INDICATORS	PURPOSE	DATA USE
SURVEYS AND QUESTIONNAIRES	- Demographic information (age, gender, English learning experience)	To assess participants' self-evaluated English proficiency and specific needs	To analyze students' language needs, course satisfaction, and to provide background for qualitative data
	- Self-assessed English proficiency (listening, speaking, reading, writing)		
	- Specific language needs related to Physical Education		
	- Satisfaction with current English courses and suggestions for improvement		
LANGUAGE PROFICIENCY TESTS	- Listening: Multiple-choice questions	To obtain objective data on participants' English proficiency	To assess participants' actual English proficiency, forming the basis for needs analysis and course design
	- Speaking: Short interview		
	- Reading: Comprehension questions		
	- Writing: Short essay		
ACADEMIC PERFORMANCE RECORDS	- English course grades	To analyze the relationship between academic performance and language needs	To identify key factors influencing English learning success and evaluate the effectiveness of course design
	- Attendance records		
	- Participation metrics (class discussion, assignment completion, etc.)		

A structured questionnaire was developed to capture a broad range of information regarding the participants' self-assessed English language proficiency, their specific language needs related to their field of study, and their experiences with current English language instruction. The questionnaire includes a combination of Likert-scale items and open-ended questions to capture both the breadth and depth of student perspectives. The questionnaire was divided into several sections: Demographic Information: This section collects basic data such as age, gender, and previous English learning experiences. Self-assessed Proficiency: Participants rated their proficiency in the four key language skills (listening, speaking, reading, writing) on a scale. Specific Language Needs: Questions focused on identifying the participants' needs related to sports-specific vocabulary and communication skills. Satisfaction with English Courses: Participants provided feedback on the

relevance and effectiveness of their current English courses and offered suggestions for improvement. The results of the survey are summarized in the table 3 and Figure 1:

Table 3: Summary of Questionnaire Results

ITEM	OPTION	PERCENTAGE (%)
GENDER	Male	58.9
	Female	41.1
SELF-ASSESSED LISTENING PROFICIENCY	Very Poor	10.3
	Poor	24.3
	Average	45.8
	Good	15.9
	Excellent	3.7
SELF-ASSESSED SPEAKING PROFICIENCY	Very Poor	12.1
	Poor	29
	Average	42.1
	Good	14
	Excellent	2.8
NEED FOR SPORTS-RELATED VOCABULARY	Very High	45.8
	High	37.4
	Average	13.1
	Low	2.8
	Very Low	0.9
SATISFACTION WITH CURRENT ENGLISH COURSES	Very Dissatisfied	5.6
	Dissatisfied	18.7
	Neutral	46.7
	Satisfied	22.4
	Very Satisfied	6.5

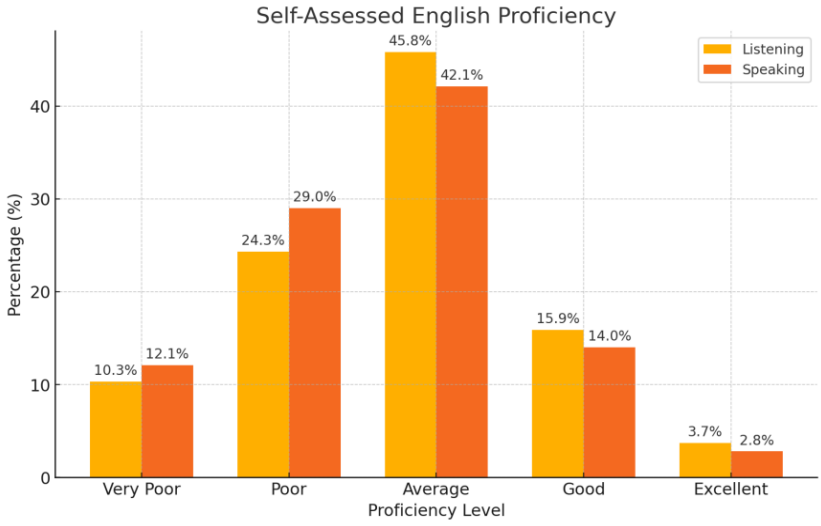

Figure 1: Self-assessed proficiency levels of listening and speaking skill

To obtain an objective measure of the participants' English language proficiency, standardized language proficiency tests were administered. These tests are designed to assess the students' abilities in listening, speaking, reading, and writing, providing a comprehensive overview of their current language skills. The proficiency test results are summarized in table 4 and Figure 2:

Table 4: Summary of Language Proficiency Test Results

SKILL CATEGORY	MAXIMUM SCORE	MINIMUM SCORE	AVERAGE SCORE	STANDARD DEVIATION
LISTENING	100	30	65	12
SPEAKING	10	2	6.5	1.5
READING	100	40	72	10
WRITING	20	5	14	3

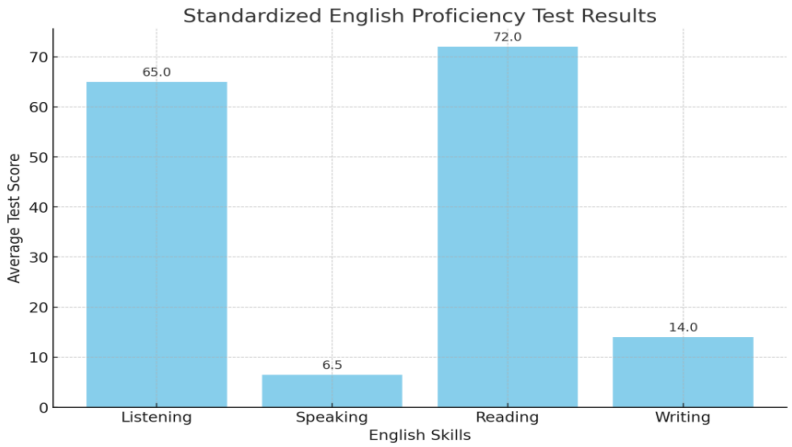

Figure 2: Average scores for listening, speaking, reading, and writing skill

Academic performance data were collected to analyze the relationship between students' language needs, their actual English proficiency, and their overall academic success. These records include grades from their English courses, attendance records, and participation metrics, such as contributions to class discussions and completion of assignments. The academic performance data are summarized in the following Table 5:

Table 5: Summary of Academic Performance Records

ITEM	AVERAGE GRADE	ATTENDANCE RATE (%)	PARTICIPATION RATE (%)
OVERALL ENGLISH COURSE GRADE	75	95	80
SPORTS-RELATED VOCABULARY TEST	68	95	82
ENGLISH SPEAKING PRACTICE ASSIGNMENT COMPLETION	-	95	76

The integration of these quantitative data collection methods ensures a comprehensive understanding of the English language learning needs of first-year Physical Education students. The data collected through surveys, proficiency tests, and academic records provide a robust foundation for the subsequent analysis and interpretation of the study's findings.

2.3.2 Qualitative Data Collection

The qualitative data collection process in this study is designed to provide in-depth insights into the English language learning needs, challenges, and experiences of first-year Physical Education students at Huzhou University. This process involves semi-structured interviews, focus group discussions, and classroom observations. These methods allow for a richer understanding of the participants' perspectives and complement the quantitative data collected earlier.

2.3.3 Semi-Structured Interviews

The semi-structured interviews aim to explore the participants' individual experiences and perspectives on their English language learning. The interviews focus on understanding the relevance of current English instruction to their academic and professional goals, the specific challenges they face in learning English, and their suggestions for improving English language education within their field. An interview guide was developed to ensure consistency across all interviews while allowing flexibility for participants to express their thoughts freely. Key topics covered include: Perceptions of English language needs related to Physical Education. Experiences with current English courses and the perceived alignment with their professional aspirations. Challenges encountered in learning and applying English in their academic work and potential future careers. Suggestions for course improvement and the introduction of new learning strategies or resources. The interviews were conducted individually with a selected subset of students (approximately 20% of the sample). Each interview lasted about 30 to 45 minutes and was conducted in a quiet, comfortable environment to encourage open and honest communication. The interviews were audio-recorded with the participants' consent and later transcribed for analysis.

2.3.4 Focus Group Discussions

Focus group discussions were organized to gain collective insights into the students' language learning experiences and needs. These discussions facilitated interaction among peers, allowing students to share and compare their experiences, which often led to the emergence of themes not captured in individual interviews. Participants were divided into small groups of 6-8 students, ensuring a mix of genders and varying levels of English proficiency within each group. This composition was intended to generate diverse perspectives and a

dynamic discussion. The focus group discussions centered around similar themes as the interviews, with additional emphasis on: Group consensus on language needs and the importance of English in their future careers. Peer experiences and strategies in overcoming language learning challenges. Perceived effectiveness of current teaching methods and resources. The discussions were moderated by a researcher who guided the conversation using a semi-structured format. Each session lasted approximately 60 to 90 minutes and was recorded and transcribed for subsequent thematic analysis.

2.3.5 Classroom Observations

Classroom observations were conducted to gain a first-hand understanding of the teaching practices, student engagement, and the extent to which the English courses meet the specific needs of Physical Education students. Observations focused on how language instruction is delivered and how students interact with the content and each other. An observation protocol was developed to systematically record various aspects of the classroom environment, including: Teaching methods used by the instructor (e.g., lecture, discussion, group work). Student engagement levels, including participation in discussions, attentiveness, and interaction with peers. Use of sports-related content and vocabulary in teaching materials and activities. Classroom dynamics, such as the teacher-student ratio, the use of multimedia resources, and the overall atmosphere. Observations were conducted during regular English classes over several weeks to capture a representative sample of typical classroom activities. Notes were taken during the observations, and key interactions and teaching moments were highlighted for further analysis. The qualitative data collected through interviews, focus groups, and observations were analyzed using thematic analysis. This involved coding the transcribed data to identify recurring themes, patterns, and key insights related to the students' English language learning needs and experiences. The analysis aimed to contextualize and enrich the findings from the quantitative data, providing a comprehensive understanding of the students' perspectives. Table 6 summarizes the qualitative data collection methods and their focus areas:

Table 6(a): Qualitative Data Collection Methods and Focus Areas

DATA COLLECTION METHOD	FOCUS AREAS	PURPOSE
SEMI-STRUCTURED INTERVIEWS	- Individual perceptions of English language needs	To gain in-depth insights into individual experiences and perceptions related to English language learning
	- Experiences with current English courses	
	- Challenges in language learning	
	- Suggestions for course improvement	

Table 6(b): Qualitative Data Collection Methods and Focus Areas

DATA COLLECTION METHOD	FOCUS AREAS	PURPOSE
FOCUS GROUP DISCUSSIONS	- Group consensus on language needs	To explore collective perspectives and generate themes that may not emerge in individual interviews
	- Peer strategies for overcoming challenges	
	- Effectiveness of teaching methods and resources	
CLASSROOM OBSERVATIONS	- Teaching methods	To observe and assess the practical implementation of English language instruction in the classroom
	- Student engagement	
	- Use of sports-related content	
	- Classroom dynamics	

These qualitative data collection methods, combined with the analysis process, provide a comprehensive understanding of the participants' language learning needs and experiences. The insights gained from these methods will be integrated with the quantitative findings to offer a holistic view of how English language education can be improved for Physical Education students.

3. Machine Learning Analysis: Support Vector Machine

In this study, a Support Vector Machine (SVM) is employed as the primary machine learning method to classify and predict the English language learning success of first-year Physical Education students. The features considered include self-assessed English proficiency, actual test scores, and demographic information. The SVM is particularly effective in binary classification tasks, making it well-suited for distinguishing between students who are likely to succeed or struggle in their English language courses.

3.1 Data Preprocessing

Before applying the SVM model, several preprocessing steps are necessary to ensure that the data is clean, well-structured, and suitable for analysis: Since SVM relies on the calculation of distances between data points, it is crucial to normalize the data so that all features contribute equally to the model. The normalization formula used is:

$$x' = \frac{x - \min(x)}{\max(x) - \min(x)} \qquad (1)$$

Where x is the original value, and x' is the normalized value. The dataset is split into training and testing sets, typically with 70% of the data used for training and 30% used for testing. This split ensures that the model is trained

on a substantial portion of the data while retaining enough data to evaluate its performance effectively.

3.2 Model Training with Support Vector Machine

The SVM model is trained using the preprocessed data. The key parameters for the SVM model, such as the regularization parameter (C), kernel type, and kernel parameters, are crucial for the model's performance. The SVM algorithm works by finding the optimal hyperplane that separates the data into different classes. Objective Function is as follow:

$$\min_{w,b} \frac{1}{2}\|W\|^2 \quad subject \quad to \quad y_i\left(W^T X_i + b\right) \geq 1, \forall i \tag{2}$$

Where W is the weight vector, b is the bias term, x_i represents the feature vector, and y_i are the class labels. Margin Maximization:

$$M = \frac{2}{\|W\|} \tag{3}$$

The goal is to maximize the margin M, which is the distance between the hyperplane and the nearest data points from each class, known as support vectors. The Support Vector Machine Model Parameter Settings are as shown in Table 7.

Table 7: Support Vector Machine Model Parameter Settings

PARAMETER	DESCRIPTION	VALUE/SELECTION METHOD
C (REGULARIZATION PARAMETER)	Controls the trade-off between margin size and classification error. A larger value reduces training error but may lead to overfitting.	Determined through grid search and cross-validation, best value = 1
KERNEL FUNCTION	Function used to map data into a higher-dimensional space for better separability.	RBF (Radial Basis Function)
GAMMA (KERNEL COEFFICIENT)	Controls the influence of individual training examples. A smaller value makes the model more general.	Determined through grid search and cross-validation, best value = 0.1

3.3 Model Evaluation

Once the SVM model is trained, its performance is evaluated using several metrics: Accuracy: The proportion of correctly classified instances out

of the total instances.

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \quad (4)$$

Precision: The proportion of true positive predictions out of all positive predictions.

$$Precision = \frac{TP}{TP + FP} \quad (5)$$

Recall: The proportion of true positives identified out of all actual positives.

$$Recall = \frac{TP}{TP + FN} \quad (6)$$

F1-Score: The harmonic mean of precision and recall, providing a balance between the two.

$$F1 - Score = 2 \times \frac{Precision \times Recall}{Precision + Recall} \quad (7)$$

The performance metrics above are derived from the confusion matrix, which compares the predicted class labels with the actual class labels. The SVM Classification Results Confusion Matrix are shown in Table 8, and the Model Evaluation Results are shown in Table 9.

Table 8: SVM Classification Results Confusion Matrix

	ACTUAL POSITIVE	**ACTUAL NEGATIVE**
PREDICTED POSITIVE	70	12
PREDICTED NEGATIVE	14	56

Table 9: SVM Model Evaluation Results

METRIC	**VALUE**
ACCURACY	85.70%
PRECISION	83.50%
RECALL	82.90%
F1-SCORE	83.20%

3.4 Feature Importance and Decision Boundary Visualization

Although traditional SVM models do not provide direct feature importance scores, techniques such as Recursive Feature Elimination (RFE)

were used to rank the importance of features. The decision boundary, as determined by the SVM, separates the data points into different classes based on the features. The Feature Importance Ranking are shown in Table 10.

Table 10: Feature Importance Ranking

FEATURE	RANK	DESCRIPTION
SELF-ASSESSED LISTENING PROFICIENCY	1	Student's self-assessment of their listening ability
STANDARDIZED LISTENING TEST SCORE	2	Actual test score in the listening section
ATTENDANCE RATE	3	Student's attendance in English classes
SPORTS-RELATED VOCABULARY PROFICIENCY	4	Student's proficiency in sports-related English vocabulary
STANDARDIZED SPEAKING TEST SCORE	5	Actual test score in the speaking section

The decision boundary visualization below shows how the SVM model separates students based on key features like English proficiency scores and attendance rates.

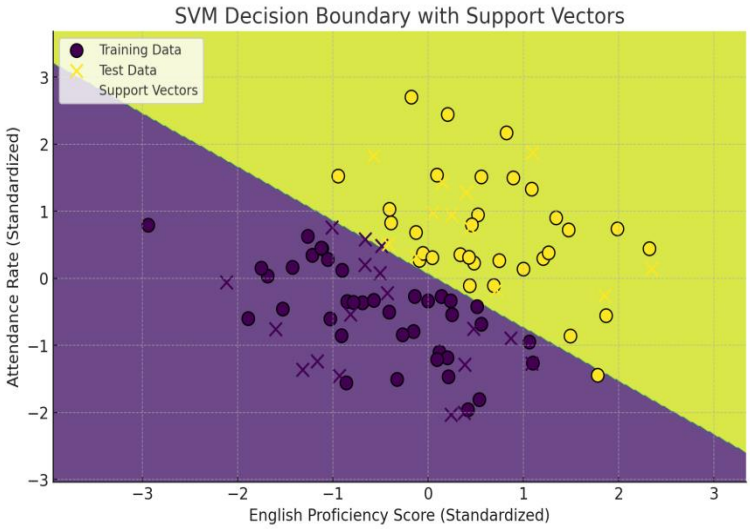

Figure 3: SVM Decision Boundary

Figure 3 illustrates the decision boundary created by the SVM model. The support vectors, highlighted as larger circles, are the critical data points that determine the position of the boundary. The axes represent standardized English proficiency scores and attendance rates.

3.5 Error Analysis

To further understand the model's performance, an error analysis was conducted on the misclassified instances. This analysis helps identify patterns or common characteristics among misclassified samples, providing insights into

potential areas for model improvement. The Classification Error Analysis are shown in Table 11.

Table 11: Classification Error Analysis

MISCLASSIFIED SAMPLE ID	ACTUAL CLASS	PREDICTED CLASS	FEATURE SUMMARY	POSSIBLE CAUSE
1	Positive	Negative	High self-assessed listening, low actual score	Potential overestimation in self-assessment leading to misclassification
2	Negative	Positive	High attendance, moderate test scores	High attendance might have masked moderate proficiency

By applying the SVM model, the study aims to develop a robust classifier that can accurately predict students' success in English language courses based on various input features. The analysis results, as presented in the tables and diagram above, provide valuable insights that can inform targeted interventions and instructional strategies for Physical Education students.

4. Qualitative Data Analysis

The qualitative data analysis in this study is designed to provide in-depth insights into the English language learning experiences, needs, and challenges faced by first-year Physical Education students. The qualitative approach complements the quantitative findings and helps to contextualize the data within the students' lived experiences. This section outlines the process of collecting and analyzing qualitative data, including semi-structured interviews, focus group discussions, and classroom observations.

4.1 Thematic Analysis

Thematic analysis was employed to identify, analyze, and report patterns (themes) within the qualitative data. This method allows for a systematic examination of the data, helping to uncover the underlying themes that reflect the students' experiences and perspectives. All interview and focus group recordings were transcribed verbatim to ensure accuracy in capturing the participants' words. The transcriptions and observation notes were systematically coded. Initial codes were generated based on recurring topics and significant statements. Coding was done using qualitative data analysis software to organize and manage the data efficiently. Codes were then grouped into broader themes that captured the essence of the data. For example, codes related to "difficulty with sports-specific vocabulary" and "lack of practical language use" were grouped under the theme "Challenges in Learning

English." The identified themes were reviewed and refined to ensure they accurately reflected the data. Themes were checked against the coded data and the entire dataset to confirm their validity and coherence. Each theme was clearly defined and named, capturing the key ideas it represents. For instance, a theme might be named "Importance of English for Career Advancement" to reflect students' views on the role of English in their future careers.

4.2 Integrating Qualitative and Quantitative Findings

The qualitative findings were integrated with the quantitative results to provide a comprehensive understanding of the students' English language learning experiences. This integration allows for a richer interpretation of the data and helps to explain the quantitative findings within the context of students' lived experiences. The qualitative themes were compared with the quantitative data to identify areas of convergence and divergence. For example, if students reported high self-assessed proficiency in speaking but faced challenges in real-world communication, this would be explored in relation to their test scores and classroom behaviors. The qualitative data provided context for the quantitative findings. For instance, if a significant number of students struggled with English listening skills (as revealed by test scores), the qualitative data could explain whether this was due to a lack of exposure to spoken English in a sports context. Any discrepancies between the quantitative and qualitative data were examined to understand underlying reasons. For example, a student might perform well in written tests but still feel anxious about speaking in class due to a lack of confidence, as revealed in the interviews.

4.3 Reporting the Findings

The final step involved reporting the qualitative findings in a clear and coherent manner. This included a detailed description of the themes, supported by direct quotes from the participants and observations from the classroom. A narrative was developed to describe each theme, explaining how it emerged from the data and how it relates to the research questions. Direct quotes from interviews and focus groups were used to illustrate the themes and provide authentic insights into the students' experiences. Each theme was linked back to the research questions, showing how the qualitative data contributes to answering them.

5. Key Themes Identified

Based on the thematic analysis, several key themes emerged: 1. Challenges in Learning English: This theme captures the specific difficulties students face, such as understanding sports-specific terminology and applying English in practical settings. 2. Importance of English for Career Advancement: Many students emphasized the need for English proficiency to succeed in their future careers, particularly in international sports contexts. 3. Satisfaction with

Current English Instruction: This theme reflects students' varying levels of satisfaction with their English courses, with some expressing a need for more interactive and practical learning opportunities. 4. Strategies for Overcoming Language Barriers: Students shared various strategies they use to improve their English, such as engaging in self-study, practicing with peers, and using online resources. These themes provide valuable insights into the students' experiences and offer practical implications for improving English language instruction in Physical Education programs.

6.Conclusion

This study, through quantitative and qualitative analysis, explored the English learning needs and challenges of first-year Physical Education students at Huzhou University. The findings revealed significant disparities in students' English skills, particularly in listening and speaking, where they performed relatively weaker. Qualitative data further indicated difficulties in mastering sports-specific terminology and applying English in practical contexts, with students expressing that current courses lack practical and interactive elements. Students widely acknowledged the importance of English for their future careers, especially in the international sports arena. Therefore, the necessity of improving English instruction is evident. The study suggests incorporating more sports-related content into the curriculum, enhancing listening and speaking training, and increasing interactive and practical exercises. This research provides empirical support for English curriculum reform, emphasizing that teaching methods should be guided by students' actual needs and career goals. Future studies could further explore the effectiveness of interactive and practical teaching methods and investigate the application of machine learning in personalized instruction to better meet students' diverse needs. Through these improvements, the English proficiency of Physical Education students will be significantly enhanced, helping them achieve greater success in a globalized professional environment.

REFERENCE

Angelova, M., & Zhao, Y. (2016). Using an online collaborative project between American and Chinese students to develop ESL teaching skills, cross-cultural awareness and language skills. *Computer Assisted Language Learning, 29*(1), 167-185.

Bunch, G. C. (2013). Pedagogical language knowledge: Preparing mainstream teachers for English learners in the new standards era. *Review of research in education, 37*(1), 298-341.

Cope, B., & Kalantzis, M. (2016). Big data comes to school: Implications for learning, assessment, and research. *Aera Open, 2*(2), 2332858416641907.

Frangoudes, F., Matsangidou, M., Schiza, E. C., Neokleous, K., & Pattichis, C. S. (2022). Assessing human motion during exercise using machine

learning: A literature review. *IEEE Access, 10*, 86874-86903.

Giulianotti, R. (2014). Human rights, globalization and sentimental education: The case of sport. In *Sport, civil liberties and human rights* (pp. 63-77): Routledge.

Hu, C. (2022). Evaluation of physical education classes in colleges and universities using machine learning. *Soft Computing, 26*(20), 10765-10773.

Hu, J., Zhang, X., Xiao, S., Fan, L., Liu, L., Li, T., & Huang, M. (2024). Correlative Analysis and Prediction of Physical Education Data Via Machine Learning: A Case Study on Grade Evaluation Method in University. *Frontiers in Computing and Intelligent Systems, 7*(3), 21-28.

Kohl III, H. W., & Cook, H. D. (2013). Educating the student body: Taking physical activity and physical education to school.

Kushik, N., Yevtushenko, N., & Evtushenko, T. (2020). Novel machine learning technique for predicting teaching strategy effectiveness. *International Journal of Information Management, 53*, 101488.

Pranoto, B. E., & Suprayogi, S. (2020). A need analysis of ESP for physical education students in Indonesia. *Premise: Journal of English Education, 9*(1), 94-110.

Tendinha, R., Alves, M. D., Freitas, T., Appleton, G., Gonçalves, L., Ihle, A., . . . Marques, A. (2021). Impact of sports education model in physical education on students' motivation: A systematic review. *Children, 8*(7), 588.

Tuong, N. K., & Dan, T. C. (2024). A STUDY ON DUOLINGO MOBILE APPLICATIONS TO IMPROVE EFL STUDENTS'LISTENING COMPREHENSION PERFORMANCES. *European Journal of Alternative Education Studies, 9*(1).

Valencia, S. W., Wixson, K. K., & Pearson, P. D. (2014). Putting text complexity in context: Refocusing on comprehension of complex text. *The Elementary School Journal, 115*(2), 270-289.

Wang, Z. (2024). Physical Education Teaching Quality Assessment Model Based on Gaussian Process Machine Learning Algorithm. *International Journal of Maritime Engineering, 1*(1), 727-736.

Warriner, D. S. (2016). 'Here, without English, you are dead': ideologies of language and discourses of neoliberalism in adult English language learning. *Journal of Multilingual and Multicultural Development, 37*(5), 495-508.