JOURNAL OF TECHNICAL EDUCATION AND TRAINING
ISSN: 2229-8932 e-ISSN: 2600-7932

JTET

Vol. 16 No. 1 (2024) 71-83
https://publisher.uthm.edu.my/ojs/index.php/jtet

Improving English Foreign Language (EFL) Performance using Artificial Intelligence in Vocational Education and Training (VET)

Cristina de-la-Peña[1]*, Jacobo Roda-Segarra[2], Beatriz Chaves-Yuste[3]

[1] *Universidad Internacional de la Rioja, SPAIN*
[2] *Universitat de Valencia, SPAIN*
[3] *Universidad Complutense de Madrid, SPAIN*

*Corresponding Author: cristina.delapena@unir.net
DOI: https://doi.org/10.30880/jtet.2024.16.01.006

Article Info

Received: 17th January 2024
Accepted: 31st March 2024
Available online: 30th June 2024

Keywords

Artificial intelligence, performance, EFL, affective variables, VET

Abstract

Internationalisation is one of the strategies for improving the technical qualifications and employability of trainers in initial and continuing vocational education and training. It is based on the full development of linguistic competence in a foreign language such as English, which is influenced by various factors, including affective factors. Currently, one resource for detecting poor performance in English is artificial intelligence to the extent that it can predict academic performance. This research aims to predict performance in English as a foreign language based on affective variables such as willingness to communicate orally in English, self-efficacy and English language anxiety. The experimental result shows that the prediction model trained with a decision tree algorithm (J48) provides the best data for predicting performance in English in terms of accuracy = 0.74, precision = 0.70, recall = 0.678 and F-score = 0.68. Analysing the influence of the variables and eliminating the data for the affective variable willingness to communicate orally in English yields the best accuracy = 0.76. This finding has relevant practical implications for the early identification of underachievement in English and for personalising educational interventions to improve learning and performance in English as a foreign language among vocational education and training students.

1. Introduction

Vocational Education and Training (hereafter, VET) has proven to be relevant for economic growth, labour market improvement and social welfare. As stated by the United Nations General Assembly in the 2030 Agenda (UNESCO, 2019), VET should provide high-quality training to ensure the professionalisation of the European workforce. The Ministry of Education and Vocational Training (2020) is committed to improving the quality of this level of education to ensure optimal integration of its students into the international labour market. The European Commission's guidelines aim to increase the production of skills that promote and improve the progress of VET students in the European Union. It is committed to greater flexibility in VET by promoting quality and professional excellence. In this way, VET or vocational students could create specific value chains and industrial ecosystems that are part of knowledge triangles in collaboration with the educational, scientific, creative and business communities.

To achieve this, it is essential to improve language training in foreign languages, generally in English, which is considered the lingua franca (Jenkins, 2009). Vocational students must master this foreign language in order to be successful in their future profession. However, the academic level of Spanish vocational students is far from the desired parameters. As the Plataforma Observatorio de la Formación Profesional (2021) states, there is a clear learning plateau among Spanish vocational students and their English as a foreign language (EFL). To successfully improve EFL skills, it is highly recommended for educational institutions to use a tool that predicts the performance of different students in EFL. In this way, teachers could propose the most appropriate pedagogical measures at the beginning of the school year to improve the EFL learning and internationalisation of vocational students by monitoring the success of their integration into the international labour market.

The Horizon Report 2019 (Educause, 2019) states that the use of artificial intelligence (AI) in education will increase by more than 43%. AI is steadily increasing in the education sector due to its importance for personalised recognition and recommendation processes to improve the quality of education (Xiao et al., 2018).

AI integrates a heterogeneous variety of technologies and methods to simulate tasks associated with a human mind to generate learning (Baker & Smith, 2019). One of these methods is machine learning, which enables the generation of predictions from data, e.g. about academic performance (Popenici & Kerr, 2017). In this context, several studies (Alamri & Alharbi, 2021; El Shazly, 2021) have focused on predicting student performance to identify at-risk situations. Lu et al. (2018) found that the decision tree model was the most accurate in predicting academic performance. This prediction was significant for students, teachers and school leaders as it enabled them to anticipate recommendations to improve the teaching-learning process (Alamri & Alharbi, 2021).

In the educational context of EFL learning and performance, AI is currently used in foreign language learning support systems (Zhang & Zou, 2020). Research has focused on facilitating foreign language acquisition and development resources, such as language identifiers, translators, intelligent tutors, etc. (Zhang & Zou, 2020). Dizon (2020) argued that using chatbots significantly positively affects foreign language development by promoting performance, interaction and affectivity (motivation). These chatbots could allow learners to follow at their own pace, self-regulate, speak without being criticised, practice repeatedly, etc. Nevertheless, there are no studies to date that focus on predicting EFL performance using affective variables. Several studies, such as the one conducted by de-la-Peña and Chaves-Yuste (2022), concluded that vocational students have an affective barrier that hinders their EFL learning, as they feel inadequate to perform their work in this foreign language. Furthermore, Arnold and Douglas (2000) and Kutuk et al. (2022) found that affective variables modulate EFL learning and performance. Therefore, this study predicts EFL performance based on well-studied affective variables such as willingness to communicate orally in English, self-efficacy, and English language anxiety.

Within the framework of Bandura's (2001) social-cognitive theory, several papers (Bouih et al. 2021; Duckworth et al., 2007; Raoofi et al., 2012) have shown that self-efficacy is a reliable predictor of performance and learning. For Köseoğlu (2015), self-efficacy directly explained 30% of academic performance. Saka and Merç (2021) demonstrated that self-efficacy has a strong relationship with foreign language anxiety (FLA), as students with high FLA tend to have low self-efficacy concepts. The most commonly used model to explain FLA (MacIntyre, et al., 1998) assumes that anxiety can significantly affect foreign language performance. For Atkinson et al. (1994), students might feel anxious because they believe they are not proficient in the foreign language. They expect that they will perform poorly and are afraid of the negative consequences. In fact, FLA could be considered as an affective factor that most hinders the learning process (Arnold & Douglas, 2000) and academic performance (Onwuegbuzie et al., 2000). In addition, willingness to communicate orally in English (WTC) is also associated with self-efficacy and anxiety. According to MacIntyre et al.'s model (1998), WTC improves performance and language acquisition. In this context, self-efficacy can be defined as the "beliefs in one's capabilities to organise and execute the courses of action required to produce given attainments" (Bandura, 1997, p.3); FLA as "a distinct complex of self-perceptions, beliefs, feelings, and behaviours related to classroom language learning arising from the uniqueness of the language learning process" (Horwitz & Young, 1991, p.128); and WTC as the "readiness to enter into discourse at a particular time with a specific person, or persons, using a L2" (MacIntyre et al., 1998, p. 547).

This study aims to predict EFL outcomes based on a set of affective variables such as willingness to communicate orally in English, self-efficacy, and English language anxiety. To this end, the field of AI, within computer science, has been widely used for predictive purposes in education (Adnan et al., 2021; El Shazly, 2021; He et al., 2020). The questions raised in the study are the following:

i) Are the predictions of the AI models accurate in predicting EFL performance using affective variables?

ii) Which AI model is most accurate in predicting EFL performance?

This is the first research to analyse the effectiveness of predicting different AI models for vocational students' EFL performance and to identify the most effective models. The results provide data that can help teachers and school leaders make decisions to improve vocational students' EFL performance. In particular, it provides information on the most accurate AI model that can be applied in an educational institution in any country to optimise students' internationalisation and employability.

2. Methods

A supervised learning process must be performed to predict an outcome based on input values, in which the AI is given a set of registers whose outcome is already known. Using an algorithm that represents a finite sequence of mathematical rules and operations (Knuth, 1997), the AI changes its internal structure in a process that could be simplified and compared to human learning. This learning is precisely the necessary condition for a computer program to be considered AI (Mira et al., 2003).

2.1 Participants

The students who participated in the experiment (N=100) were all enrolled in vocational training for early childhood education, aged 18-40 years (X: 20; SD: 1.51). 75% were women and 25% were men from a middle socio-economic class, enrolled in two different schools where they had to take the compulsory module of EFL in Vocational Education. 44% of the students were enrolled in Center A, while the remaining 56% were enrolled in Center B. Participants were selected through a non-probabilistic, intentional sample. All participants are native Spanish speakers and their English level in L2 is B1+ according to the Common European Framework of Reference for Languages (CEFRL) (Council of Europe, 2001) and its Companion Volume (Council of Europe, 2020). Only 20% have an official EFL certificate and the remaining 80% have no certificate. As for their English proficiency at the time of the experiment, 5% had a zero level in English, 36% a low level, 48% an acceptable level and 11% an excellent level. The selection criteria for the sample were: being enrolled in early childhood education vocational training, taking all the tests, not having previously taken a test and not suffering from a learning disability or a physical or psychopathological sensory problem. Students volunteered to participate in the study and all students gave their ethical approval. The ethics committee of the University XXXXX approved the study with code XXX/2022.

2.2 Instruments

Our data were obtained through a series of tests administered during the 2022 school year:

i) Foreign Language Classroom Anxiety Scale (FLCAS) (Horwitz et al., 1986). This scale is the most widely used scale for assessing anxiety about oral communication in a foreign language and includes 33 indicators that assess anxiety about negative evaluations in EFL classes and fear of speaking in English. Each indicator is rated on a Likert scale from 1 to 5, where 1 stands for" totally agree", 2 for "agree", 3 for "neither agree nor disagree", 4 for "disagree" and 5 for "totally disagree". The total score is calculated by adding up the 33 indicators. The higher the score on the scale, the greater the student's anxiety. The score can be between 33 and 165 points. 165 points means that the student is very anxious in EFL lessons, while 33 points means that the student shows no anxiety in EFL lessons. Pérez-Paredes and Martínez-Sánchez (2001) identified four factors in the Spanish version of the FLCAS: negative attitude towards learning (items 6, 17), anxiety about situations and processes (items 4, 7, 15, 16, 23, 25, 29, 30), communication anxiety (items 1, 3, 9, 12, 13, 18, 20, 24, 31, 33) and discomfort with English inside and outside the classroom (items 8, 14, 32). Pérez-Paredes and Martínez-Sánchez (2001) have demonstrated reliability and validity with an internal consistency of 0.889. The reliability with the ordinal omega coefficient for this sample is 0.963.

ii) Willingness to Communicate Questionnaire (WTC) in English (McCroskey & Baer, 1985). This questionnaire is the most widely used to assess willingness to communicate in English. The Spanish version of Santos (2014) is used for the first nine questions and the version of Díaz-Pinto (2010) for the tenth question. These versions are best adapted to the Spanish educational system and the Spanish language. It comprises 10 questions on 10 different situations that provide information about the students' willingness to communicate orally in English. The answers are rated on a Likert scale from 0 to 5 points, where 0 stands for "never", 1 for "almost never", 2 for "sometimes", 3 for "about 50% of the time", 4 for "usually" and 5 for "almost always". The final score is the sum of the 10 indicators and can range from 10 points (low willingness) to 50 points (highest willingness). The reliability for this sample with the ordinal omega coefficient is 0.934.

iii) *Ad hoc* questionnaire on students' perceptions of EFL: There is no measure to assess the language skills of vocational students. Therefore, questions were asked about gender, age, previous study, EFL certificates and level, and reasons for this EFL level. In addition, information about the students' self-perceived level of EFL self-efficacy is collected with some items that attempt to assess their EFL language skills (listening comprehension, speaking, reading, writing and reading comprehension) (created according to the CEFRL (Council of Europe, 2001, 2020): to indicate the level of difficulty of the ability to listen to English, speak English, read English, write English, and understand English language texts using a Likert scale from 1, the easiest, to 5, the most difficult: 1 easiest, 2 fairly easy, 3 neither easy nor difficult, 4 fairly difficult and 5 most difficult. Answers with a value of 1 therefore mean that the corresponding language skill is easy, and a value of 5 means that the corresponding language skill is maximally difficult. The reliability for this sample with the ordinal omega coefficient is 0.862.

iv) Current EFL grade: This grade is given by the EFL teacher in the final EFL test and ranges from 0 fail (F), 5-6 pass (C), 7-8 fairly good (B) and 9-9.9 excellent (A) and 10 excellent with distinction (A+). The exam consists of five parts with a maximum value of two points. The parts are reading comprehension (text and five open-ended questions and five multiple choice questions), written expression (writing an essay on a free topic), listening comprehension (audio and five open-ended questions and five multiple choice questions), speaking (interview in pairs choosing one of five possible topics) and use of English (five gap-fill exercises, verb conjugation, word formation and paraphrasing).

First, ethics committee approval is obtained and permission is sought from the centers to conduct the research. Second, the students' teacher is contacted to determine the date and time to administer the instruments. Third, two researchers go to the centres and classes and explain the research to the students and their need to consent or not to participate voluntarily. All students give their consent and the two researchers are present when all instruments are administered. Finally, the tests are administered to all students in the morning hours in the same order: FLCAS, WTC and *ad hoc* questionnaire on students' perception of EFL for approximately 30 minutes in a classroom with appropriate volume and lighting conditions. On the day of application, the teachers share the students' results with us.

2.3 Data Processing

The tool chosen to implement the AI was Weka, a free software developed by the University of Waikato in New Zealand (https://www.cs.waikato.ac.nz/ml/weka/). It can be used to apply various algorithms to a sample of data sets to make predictions or classifications (Frank, Hall & Witten, 2016). The data obtained from the different tests applied to the sample of students have been processed to adapt them to the requirements of Weka.

The data obtained from each participant consisted of 56 variables that included biological aspects, previous training and self-perceptions. The variable selected for performing the prediction was the current English grade, and the remaining 55 variables were those to be used as input data for the AI to predict the grade. In order to reduce the input data as much as possible, certain groups of variables were combined into a single data type: Questions on willingness to talk (10 questions) and anxiety (33 questions). In both cases, these were Likert scale responses, so it was decided to standardise them by adding each of the individual values so that the 10 willingness to talk variables and the 33 anxiety variables were reduced to a single numerical figure. Furthermore, the data were grouped to subsequently analyse their impact on the predictive power of the AI (3.3), so that groups of personal and academic data, self-perceptions, willingness to communicate and anxiety were formed. The input variables, the groupings and the number of variables per grouping are shown in Table 1.

Table 1 *Data obtained from the sample*

Grouping	Data	Included Variables
Personal	Age	1
Personal	Gender	1
Academic	Course	1
Academic	EFL certificate	1
Academic	Studies	1
Self-efficacy	EFL level	1
Self-efficacy	Reason for the EFL level	1
Self-efficacy	Listening comprehension	1
Self-efficacy	Speaking	1
Self-efficacy	Writing	1
Self-efficacy	Reading	1
Self-efficacy	Reading comprehension	1
Willingness to communicate	Willingness to communicate	10
Anxiety	Anxiety	33

The input data (textual categories with the exception of listening comprehension-speaking-writing-reading-reading comprehension, willingness to communicate, and anxiety), were expressed in numerical values using the following procedure, where the number in parentheses corresponds to the value into which the data was transformed:

i) Age: 18 years old (1), 19 years old (2), 20 years old (3), 21-25 years old (4), 26-30 years old (5), 31-35 years old (6) and 36-40 years old (7). The reason for the division into these age groups corresponds to the age range of the questionnaire design. In the questionnaire design, the students did not indicate their current age, but selected the group to which they belonged from a predefined list. For the AI training, these age groups were converted into a discrete variable (with a range of 1-7). It should be noted that these discrete values are not relevant for the training of the AI from a numerical point of view, as the algorithm looks for correlations between the inputs and the expected output, but without taking into account their cardinality or ordinality. In other words, the values that are input to the AI for training should be understood as the numerical encoding of discrete variables.

i) Gender: female (1) and male (2).
ii) Course: no (0), medium (1) and higher (2).
iii) English certificate: no (0) and yes (1).
iv) vi) English training: school (1) and language school (2).
v) English level: none (0), low (1), acceptable (2) and high (3).
vi) Reason for the level: (0) no interest, (1) poor teaching performance, (2) attending language schools and/or bilingual schools, and (3) own interest (movies, music, etc.).

There were five categories of response options for the variable chosen for prediction: failing, passing, good, remarkable and excellent. The balance of responses, i.e. the percentage in which each of them occurred, was analysed, since a high degree of imbalance is one of the main problems of AI-based classification and prediction systems (Su et al., 2006), as the results provided by the algorithm have the highest frequency of occurrence. Thus, assuming that the variable could have 5 categories, the ideal swing would be 20% for each of them.

The predicted variable is unbalanced with respect to the desirable 20%, with its lower end in the "fail" or "insufficient" category (8%) and the upper end in the "pass" or "sufficient" category (32%). This would lead to a bias of detections in the "pass" category, reducing the number of detections in the "fail" category. In order to reduce this problem as much as possible, the scores were grouped in order to obtain a sample that was as balanced as possible. Thus, the "fail" and "pass" categories were combined to form the "low qualifications" group, while "good", "notable" and "outstanding with distinction" formed the "high qualifications" group. This also made the variable to be predicted a dichotomous variable, which increased the number of algorithms that could be used for prediction. The result of this grouping, and its approximation to the ideal balance (which is now 50%).

After these transformations and data processing, a file in ARFF format was created, which was required by the Weka software to perform the supervised learning processes in the AI. This file format required header information specifying the variables to be included and the types of data they may contain. After the header, the individual data sets were inserted into the fine line by line using the processing described in this section.

To globally identify the algorithm that best predicted high or low skills in EFL, a single ARFF file was first created with all groups of input variables (personal, academic, self-perception, willingness to communicate and anxiety). A supervised learning procedure was performed on this file after the comparison algorithms were selected.

2.4 Supervised Learning

Four prediction and classification algorithms widely used in education (Bujang et al., 2021; Nabil et al., 2021; Uliyan et al., 2020) with different performances were selected to determine the most effective one. The first one, J48, belonged to the decision trees that classify according to certain rules identified in the input data (Gil et al., 2018). The second algorithm chosen was Naïve Bayes (NB), which is based on Bayes' theorem, by which it can be assumed that the effect of a particular input value is independent of the values of the other inputs. The third algorithm, logistic regression (LR), used a logistic function that determines the output as a function of the input. The last and fourth algorithm chosen, multilayer perceptron (MLP), is based on the artificial neural network model, which imitates the biological functioning of neuron layers in a simplified form.

To measure the effectiveness of the individual algorithms in predicting the EFL score, the variables true-positive (TP), true-negative (TN), false-positive (FP) and false-negative (FN) were determined and presented in confusion matrices. The confusion matrices obtained after applying the four algorithms to the previously described ARFF file with all groups of input variables (personal, academic, self-perception, willingness to communicate, and anxiety) for the prediction of EFL grade category (low skill, high skill). The data were validated by cross-validation with 10 subsets.

With the values obtained and presented in the confusion matrices, the effectiveness indicators of each of the algorithms, such as accuracy, precision, recall and f1-score, were calculated according to the following formulas:

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

$$Precision = \frac{TP}{TP + FP}$$

$$Recall = \frac{TP}{TP + FN}$$

$$F1 = \frac{2 * Precision * Recall}{Precision + Recall}$$

The results of these metrics were analyzed in Table 2 and show that the J48 algorithm achieved the best results in terms of overall AI accuracy.

Table 2 *Comparison of the different algorithms*

Algorithm	Accuracy	Precision	Recall	F1
J48	0.74	0.70	0.67	0.68
NB	0.72	0.65	0.74	0.69
LR	0.72	0.67	0.67	0.67
MLP	0.70	0.64	0.64	0.64

2.5 Impact of the Prediction Variables

When looking at the accuracy metric using all input variables in supervised learning, the J48 algorithm predicted high or low skills in EFL more efficiently. However, as described earlier, the input variables were grouped into 5 sets (personal, academic, self-perception, willingness to communicate and anxiety) to determine the influence of variable on AI prediction.

For this purpose, 5 supervised learning sections were performed using the J48 algorithm, deleting one group of variables at a time, according to the following scheme:

• First variant: academic + self-perception + willingness to communicate + anxiety (personal data was removed).

• Second variant: Personal + self-perception + willingness to communicate + anxiety (academic data was removed).

• Third variant: Personal + academic + willingness to communicate + anxiety (self-perception has been removed).

• Fourth variant: Personal + academic + self-perception + anxiety (willingness to communicate data was removed).

• Fifth variant: Personal + academic + self-perception + willingness to communicate (anxiety data was removed).

Finally, the accuracy, precision, recall and f1 score metrics were calculated using the data from the confusion matrices of the 5 variations performed on the data supplied to the AI. These data are shown in Table 3. The best results were obtained when the willingness to communicate data were eliminated, while the worst results were obtained when the self-efficacy data were eliminated.

Table 3 *Comparison of the different variations*

Variation	Accuracy	Precision	Recall	F1
Without personal data	0.71	0.67	0.62	0.64
Without academic data	0.73	0.69	0.64	0.67
Without self-perception data	0.61	0.54	0.48	0.51
Without willingness to communicate data	0.76	0.73	0.69	0.71
Without anxiety data	0.65	0.58	0.62	0.60

3. Discussion

In recent years, AI has been used to predict the academic performance of students in different educational levels and disciplines. To date, there is no empirical research that has been conducted in the context of EFL learning of vocational students where prediction can be based on affective variables. Moreover, whether AI models can be accurate when using affective variables and identify the best predictive model for EFL performance is an open question for the context of vocational students in a single country, in this case from the education sector. Moreover, there could be further empirical evidence for vocational students from other sectors in the country and even for all educational levels and international settings.

This study aims to predict outcomes in EFL performance based primarily on affective variables such as willingness to communicate orally in English, self-efficacy, and EFL anxiety. In particular, this research has shown that the AI models are reasonably accurate for machine learning. The J48 model has been shown to be the most effective and accurate model for predicting vocational students' EFL performance, followed by NB, LR and MLP.

Are the predictions of the AI models accurate in predicting EFL performance using affective variables? The research results show that the machine learning AI models used work well in predicting EFL performance of vocational learners. This finding is consistent with Bilal et al.'s (2022) and Tomasevic et al.'s (2020), who indicated that the models used provide predictions with high reliability and effectively predict academic performance. Alamri and Alharbi (2021) stated, in their systematic review, that the most commonly used machine learning models for predicting student performance were decision trees and rule-based learning. However, there is a lack

J. of Technical Education and Training Vol. 16 No. 1 (2024) p. 71-83

of studies on predicting EFL performance, so our results are not comparable. This work points the way for future work.

This study utilizes data on personal, academic, and affective variables, with the greatest emphasis on affective variables, which is the novelty of the research. Academic and personal (socio-demographic) factors are most commonly used to analyze the prediction of student achievement and profiles (Hamim et al., 2021). Other studies have used socio-demographic and social data (Singh & Kaur, 2016). Mishra et al. (2014) used academic and emotional data to predict academic performance in the third semester. Mitrofan and Ion (2013) used personality factors such as self-confidence, achievement orientation, self-discipline, self-organization, and persistence to predict performance. Anam and Stracke (2020) showed that it is necessary to deepen the affective variables related to students' response in L2 learning, as these affective factors have been shown to significantly influence academic performance (Kutuk et al. 2022).

Knowledge and control of affective variables facilitates mastery of foreign language skills in all its domains (writing, comprehension and productive skills), and leads to the guidelines proposed in the 2030 Agenda (UNESCO, 2019) to effectively promote language proficiency in order to optimize internationalization. These findings are related to several studies that have analyzed the use of affective approaches in EFL teaching, such as Deb (2018), who finds that affective factors play an essential role in EFL learning, as there is a correlation between self-efficacy and anxiety in tasks such as writing. In the same vein, Zahibi (2018) states that a confident student who does not feel anxious is able to take more risks when performing writing activities. Authors such as Guglielmino (1986) also go further by drawing on music when it comes to enhancing these affective aspects to improve the teaching-learning processes in EFL classes.

More generally, Sison's (2022) study shows that the use of affective strategies, such as reducing anxiety, encouraging oneself, and taking an emotional temperature are widely used in ESL classrooms, although they are used more by teachers who are young women. Thus, the findings in the present study are consistent with several scholarly studies that demonstrate the use of affective strategies to improve EFL instruction and provide an empirical basis for these initiatives that may be used intuitively.

With regard to the question “Which AI model is more accurate in predicting EFL performance in English?”, in this paper, a comparison was made between 4 AI models. The J48 model showed the best accuracy in predicting academic performance in EFL among vocational students. This result is consistent with studies in other disciplines (Hamim et al., 2021; Hussain et al., 2018) that found that the decision tree was the most efficient in the context of predicting academic performance. Khan et al. (2021) found that the most efficient model for predicting performance was the J48 model. Imran et al. (2019) used language and math data and compared different AI models. They found that J48 showed the highest accuracy. Pandey and Taruna (2014) argued that the J48 model has the best predictive power for academic achievement. The J48 algorithm was said to have the best overall accuracy. It was chosen to use this metric as a comparison because it is suitable when the sample is balanced (Ruiz & Srinivasan, 2002). In the current study, the high and low skill were very close to a balanced situation of 50%.

After choosing the J48 algorithm as the one that provided the best prediction data, 5 variations were made with the data fed to the algorithm to analyze how its performance varied. It was found that the worst results were obtained when the self-efficacy data were eliminated (accuracy = 0.61), while the best results were obtained when the willingness to communicate data were eliminated (accuracy = 0.76). It can be inferred that the information provided by the students in the self-efficacy questionnaires is significant in predicting future outcomes in EFL. In this paper, self-efficacy refers to the different levels of English proficiency, the level of education and the reasons that led the students to that level. Scientific literature (Bouih et al. 2021; Haerazi & Irawan, 2020; Köseoğlu, 2015) points to the importance of self-efficacy for EFL learning and achievement has been indicated. For Shi (2016), students with high self-efficacy were able to achieve better academic results, use more learning strategies and keep language anxiety low. Woodrow (2006) argued that students with high self-efficacy are more likely to use metacognitive strategies both inside and outside the classroom than students with low self-efficacy. Different data emerged for the variable willingness to communicate orally in English, as its presence caused a significant decrease in the predictive power of AI. This result reflected the fact that students with better or worse willingness to communicate performed equally well.

Moreover, on the one hand, there are studies indicating a positive relationship between WTC and academic achievement (Muamaroh & Prihartanti, 2013) and students with high WTC showing a higher EFL performance (Cong-Lem & Thu-Hang, 2018). On the other hand, there is evidence that speaking is the language skill that students dislike the most, perform worse and feel more anxious (Al Hosni, 2014). Shen and Chiu (2019) found that the main cause of difficulties in speaking skills in EFL learners was anxiety, nervousness, etc. and Nasiruddin and Hum (2018) found that students in EFL classes were not prepared to communicate. Oral proficiency in a foreign language has been an important issue for decades, and researchers have attempted to develop various strategies to promote its development. Yang (2014) has asked her students to watch English dramas, imitate famous people speaking, listen to the news in English, and talk to friends. These types of strategies led them to be less afraid of making mistakes in English and improved their WTC in EFL.

4. Conclusion

This research provides scientific evidence for the benefits of using AI to predict EFL performance using affective variables. 4 AI models for prediction were identified as accurate, with J48 showing the best accuracy compared to NB, LR, and MLP. The practical implications can be directed to educational researchers using J48 as a predictor of EFL using affective variables in vocational students. As for educational managers, they can benefit indirectly by making decisions about interventions before the start of the academic year so that students can improve their learning and performance in EFL.

On the one hand, early warning systems can be developed to detect "low skills" (they know that their decision is based on the most effective algrothym (J48) with affective variables). Du et al. (2020) have demonstrated a more effective deep learning-based system than machine learning to detect warning signs in at-risk K-12 students.

A practical implication for instructional design is to incorporate affective strategies into the classroom by integrating them into all EFL activities to create a positive and supportive environment in which students can feel relaxed. For example, strategies to control anxiety levels, to focus attention on content rather than physiological manifestations, to develop motivation and a positive attitude towards English, to boost confidence without fear of making mistakes, or to perform collaborative or paired tasks. Furthermore, these affective strategies can be included in EFL course materials. These strategies and techniques can easily be transferred to other subjects and educational levels. Pedagogical training should be about combining the cognitive and affective aspects in the teaching process. As Immordino-Yang (2018) stated, emotions can drastically influence the learning process and teachers should provide students with emotionally engaging learning materials, such as stories or videos that evoke empathy or emotional connection to help students activate memory and learning and facilitate an optimized learning environment.

A practical implication for teacher training is to train teachers in emotional intelligence and its application to classroom subjects, as caring but fair teachers, with positive expectations for their students and with an invisible force that drives students to progress. In addition, teachers can be trained in digital resources that allow them to create activities or materials for students outside the classroom to practice and improve emotional self-regulation in EFL lessons.

EFL teachers need to didactically start from the communicative theoretical framework to move towards current active learning methodologies (i.e. service learning, cooperative learning, gamification, project learning) where active learning strategies, such as group discussions or problem-solving activities enhance the students' understanding. Teachers should move away from traditional English language passive learning methods based on grammar and translation which are not student-centered and do not respect students' needs, learning styles and learning pace. Communicative competence includes different areas (linguistic, sociolinguistic, discourse and strategic) so that grammar understanding, cultural knowledge, conversational skills, and the ability to maintain control over language gaps should be adapted to the specific needs of each student and to the teaching activity. It is crucial to train teachers to understand students (which helps them to learn EFL better) as this is the basis for effective EFL teaching in the classroom.

On the other hand, personalized pedagogical interventions can be developed for each student to optimize their EFL learning process in learning centers. This is relevant for these vocational students in the education sector as they will be teaching English to 0–3-year-olds in children's centers. Personalized learning can be developed with guided digital tutoring programs that work affective strategies and techniques with digital interactions with other people, benefiting all students and, especially, those students who have greater affective difficulties (more anxiety and less willingness) when confronted with English. Another possibility would be to integrate personalized learning in the classroom, based on collaborative projects that students must carry out in the classroom.

Moreover, the possibility of agreements and internships in other countries presupposes an adequate level of English. The urgent need for countries to improve the professional qualification of VET students by promoting internationalization is reflected in the 2030 Agenda goals (Organization for Economic Cooperation and Development [OECD] 2021). VET must be consolidated as a sustainable system that provides skills to optimize the professional qualification and successful employability of individuals (UNESCO, 2019).

In general, AI in education will increasingly be used to benefit the teaching-learning process. This means that school leaders, teachers and policymakers should begin to consider it in educational curricula as a screening and improvement tool that can promote the individualization of education. This finding encourages further research with the aim of improving the educational quality and professional qualification of vocational students and, moreover, of students at other educational levels.

4.1 Limitations and Prospective

For the interpretation of the results of this research, the limitation of the sample size (100 students of a specific educational level such as VET) may affect the generalization of the results, which must be understood in this educational context among the future technicians of early childhood education from 0-3 years. The heterogeneity of the sample in terms of age is typical of this type of study on VET in all countries, as is the gender distribution in the predominantly female early childhood education specialties.

Nevertheless, the sample of 100 students was chosen so that the AI results can be considered from a scientific point of view. EFL learning is necessary for the training and development of students in all levels of education, but especially at the VET level because of the rapid access into the labour market and for teaching in English to 0-3-year-old children.

Another limitation related to the sample and data collection is the potential risk of sampling bias, as the sample was selected based on accessibility rather than randomly across the country. The tests used to assess the affective variables are self-report questionnaires, which makes them susceptible to response bias due to social desirability, but they are the most commonly used questionnaires for the selected sample. This possible social desirability bias could influence the results by not reflecting students' actual perceptions of anxiety, willingness to communicate in English, and perceived English proficiency and being slightly more positive than is actually the case. One solution could be to use at least two tests to assess each dimension -anxiety, willingness to communicate and perception- as there are currently no objective affective tests specifically for EFL. Another possible solution could be to change the format of the questions in the questionnaires to make them as neutral as possible and to validate the questionnaire and perform the scoring procedure.

Limitations in terms of data analysis could be the type of algorithms used (J48, NB, LR, MLP), although they are the most common for prediction in the educational context. This may influence because there could be an algorithm not analyzed in this study that would provide a higher level of efficiency in predicting EFL performance. A possible solution could be to analyze more algorithms but using the same variables since the selection of variables to generate data for the AI model affects the performance of the prediction models (Tsao et al., 2017). However, the selected algorithms cover different approaches to artificial learning. Although others such as Decision Trees or Random Forest (RF) could have been used, these two together with J48 fall into the group of decision tree algorithms, with common principles that probably would not have made too much difference. Another limitation in analysis could be the high degree of imbalance. This is one of the main problems of AI-based classification and prediction systems, as the results provided by the algorithm are biased towards the answers with the highest frequency of occurrence (Su et al., 2006). To mitigate this, the categories of performance variables were reduced to two: high and low skills. This choice was considered the most appropriate as it had already been used by other researchers (Kulkarni et al., 2020). However, a slight bias towards the low-skill cases was already observed in the confusion matrices. This was possible because despite the attempt to make the sample as balanced as possible by standardizing the categories of the grade variables, the low-skill group was slightly higher than the high-skill group (58% vs. 42%).

As for the interpretation of the data, a limitation is that there is no research and theoretical approach for predicting EFL performance with AI using affective variables. The studies that exist to predict academic performance in EFL use demographic, academic and/or social variables. This implies the lack of a theoretical model on which to interpret the data obtained and points to the need for studies in this area that use predictive models based on affective variables.

Future research should focus on the application of J48 as a more accurate model for predicting students' academic performance in foreign language classes. This would allow educational institutions to personalize the EFL teaching-learning process and improve students' learning before the beginning of the academic year. At the same time, further research is needed in other VET sectors and at lower and higher levels of education to optimize EFL performance.

Another research direction would be to consider other types of affective variables such as motivation as well as cognitive and language variables in relation to EFL performance. An example of this is the studies looking at language awareness and how oral production, metacognition and other factors play a role in EFL learning (Kieseier et al., 2022). The use of these variables in an AI prediction model would be recommended in early schooling in monolingual, bilingual, and multilingual contexts.

Acknowledgements

This study is linked to the research agreement entitled Perception of English as a Foreign Language in Vocational Education and Training, in which a study of vocational students' perceptions of EFL was conducted.

Conflict of Interest

Authors declare that there is no conflict of interests regarding the publication of the paper.

Author Contribution

The authors confirm contribution to the paper as follows: ***study conception and design:*** *Cristina de-la-Peña;* ***data collection:*** *Jacobo Roda-Segarra;* ***analysis and interpretation of results:*** *Beatriz Chaves-Yuste;* ***draft manuscript***

preparation: *Cristina de-la-Peña: Jacobo Roda-Segarra. All authors reviewed the results and approved the final version of the manuscript.*

References

Adnan, M., Habib, A., Ashraf, J., Mussadiq, S., Raza, A. A., Abid, M., et al. (2021). Predicting at-risk students at different percentages of course length for early intervention using machine learning models. *Ieee Access, 9*, 7519-7539. https://doi.org/10.1109/ACCESS.2021.3049446

Alamri, R. & Alharbi, B. (2021). Explainable student performance prediction models: a systematic review. *IEEE, 9,* 33132-33143. https://doi.org/10.1109/ACCESS.2021.3061368

Al Hosni, S. (2014). Speaking difficulties encountered by young EFL learners. *International Journal on Studies in English Language and Literature, 2*(6), 22-30. https://www.researchgate.net/publication/270340628_Speaking_Difficulties_Encountered_by_Young_EFL_Learners

Arnold, J. & Douglas, H. (2000). Mapa del terreno [Terrain Map]. In J. Arnold (Ed.), *La dimensión afectiva en el aprendizaje de lenguas* [Affective Dimension in Language Learning] (pp.19-41). Cambridge University Press.

Atkinson, R. L., Atkinson, R. C., Smith, E. E., Bem, D. J., & Nolen-Hoeksema, S. (1994). *Pszichológia* [Psychology]. Budapest: siris – Századvég.

Baker, T., & Smith, L. (2019). *Educ-AI-tion rebooted? Exploring the future of artificial intelligence in schools and colleges.* Nesta. https://media.nesta.org.uk/documents/Future_of_AI_and_education_v5_WEB.pdf

Bilal, M., Omar, M. Anwar, W., Bokhari, R.H. & Choi, GG.S. (2022). The role of demographic and academic features in a student performance prediction. *Scientific reports, 12*, 12508. https://doi.org/10.1038/s41598-022-15880-6

Bouih, A., Nadif, B., & Benattabou, D. (2021). Assessing the effect of general self-efficacy on academic achievement using path analysis: a preliminary study. *J. Eng. Lang. Teach. Appl. Linguist. 3,* 18–24. https://doi.org/10.32996/jeltal.2021.3.4.3

Bujang, S., Selamat, A., Ibrahim, R., Krejcar, O., Herrera-Viedma, E., Fujita, H., & Ghani, N. (2021). Multiclass prediction model for student grade prediction using machine learning. *IEEE Access, 9.* https://doi.org/10.1109/access.2021.3093563

Cong-Lem, N., & Thu-Hang, N. T. (2018). The interplay among L2 willingness to communicate, speaking test anxiety and speaking proficiency. *Conference: 9th International Conference on TESOL: Breakthroughs in English language teaching in the 21st century*. SEAMEO RETRAC & Curtin University. http://www.vnseameo.org/TESOLConference2018/wp-

Council of Europe. (2001). *Common European Framework of Reference for Languages: Learning, teaching, assessment.* Council of Europe Publishing.

Council of Europe. (2020). *Common European Framework of Reference for Languages: Learning, teaching, assessment – Companion Volume.* Council of Europe Pubhishing.

Deb, J. (2018). Affective Factors in Second Language Writing: Is It a Matter of Concern? *The Morning Watch: Educational and Social Analysis, 46*(1-2 Fall).

De-la-Peña, C. & Chaves-Yuste, B. (2022). Claves de éxito para mejorar la competencia lingüística en inglés (L2) en FP: percepción de los estudiantes. *Didáctica Innovación y Multimedia, 40,* 1-10. https://www.raco.cat/index.php/DIM/article/view/402721

Díaz-Pinto, E. R. (2009). *Estudio sobre las inteligencias inter- e intrapersonales como instrumentos de desarrollo de la disposición a comunicarse en el aula. [Study on inter- and intrapersonal intelligences as instruments for the development of the willingness to communicate in the classroom*]. [Doctoral dissertation, Universidad de Huelva]. http://rabida.uhu.es/dspace/handle/10272/157

Dizon, G. (2020). Evaluating intelligent personal assistants for L2 listening and speaking development. *Language, Learning and Technology, 24*, 16-26. https://doi.org/10.1002/tesj.353

Du, X., Yang, J. & Hung, J.L. (2020). An integrated framework based on latent variational autoencoder for providing early warning of at-risk students. *IEEE Access, 8*, 10110-10122. https://doi.org/10.1109/ACCESS.2020.2964845

Duckworth, A., Peterson, C., Matthews, M., & Kelly, D. (2007). Grit: perseverance and passion for long term goals. *Journal of personality and social psychology, 92*, 1087-1101. https://doi.org/10.1037/0022-3514.92.6.1087

Educase (2019). *Horizon report: 2019 higher education edition.* https://library.educause.edu/-/media/files/library/2019/4/2019horizonreport.pdf

El Shazly, R. (2021). Effects of artificial intelligence on English speaking anxiety and speaking performance: a case study. *Expert Systems*, *38*: e12667. https://doi.org/10.1111/exsy.12667

Frank, E., Hall, M.A., & Witten, I.H. (2016). *The WEKA Workbench*. *Online Appendix for Data Mining: Practical Machine Learning Tools and Techniques*. Morgan Kaufmann, Fourth Edition.

Gil, D., Fernández-Alemán, J. L., Trujillo, J., García-Mateos, G., Luján-Mora, S., & Toval, A. (2018). The effect of green software: a study of impact factors on the correctness of software. *Sustainability, 10*(10), 3471. https://doi.org/10.3390/su10103471

Guglielmino, L. M. (1986). The Affective Edge: Using Songs and Music in ESL Instruction. *Adult literacy and basic education, 10*(1), 19-26.

Haerazi, H., & Irawan, L. (2020). The effectiveness of ECOLA technique to improve reading comprehension in relation to motivation and self-efficacy. *Int. J. Emerg. Technol. Learn. 15,* 61–76. https://doi.org/10.3991/ijet.v15i01.11495

He, Y., Chen, R., Li, X., Hao, C., Liu, S., Zhang, G., & Jiang, B. (2020). Online at-risk student identification using RNN-GRU joint neural networks. *Information, 11*(10), 474. https://doi.org/10.3390/info11100474

Horwitz, E. K., Horwitz, M. & Cope, J. (1986). Foreign Language Classroom Anxiety. *The Modern Language Journal, 70*(2), 125-132. https://doi.org/10.2307/327317

Hussain, S., Dahan, N.A., Ba-Alwib, F.M., & Ribata, N. (2018). Educational data mining and analysis of students' academic performance using Weka. *Indonesian Journal Electrical Engineering and Computer Science, 9*, 447-459. https://doi.org/10.11591/ijeecs.v9.i2.pp447-459

Immordino-Yang, M. H. (2018). *Emotions, Learning, and the Brain: Exploring the Educational Implications of Affective Neuroscience*W. W. Norton & Company.

Imran, M., Latif, S., Mehmood, D. & Sha, M. S. (2019). Student academic performance prediction using supervised learning techniques. *International Journal of Emerging technologies in learning, 14*(14), 92-104. https://doi.org/10.3991/ijet.v14i14.10310

Jenkins, J. (2009). English as a lingua Franca: Interpretations and Attitudes. *World Englishes, 28*(2), 200-207. https://doi.org/10.1111/j.1467-971X.2009.01582.x

Khan, I., Ahmad, A.R., Jabeur, N. & Mahdi, M.N. (2021). Machine learning prediction and recommendation framework to support introductory programming course. *International Journal of Emerging Technologies, 16*(17), 42-59. https://doi.org/10.3991/ijet.v16i17.189

Kieseier, T., Thoma, D., Vogelbacher, M. & Holger, H. (2022). Differential effects of metalinguistic awareness components in early foreign language acquisition of English vocabulary M., and grammar. *Language Awareness, 31*(4), 495-514. https://doi.org/10.1080/09658416.2022.2093888

Knuth, D.E. (1997). *The art of computer programming, volume 1 (3rd Ed.): fundamental algorithms*. Addison Wesley Longman Publishing Co. Inc.

Köseoğlu, Y. (2015). Self-efficacy and academic achievement-a case from Turkey. *Journal of Education and Practice, 6*(29), 131-141. https://www.iiste.org/Journals/index.php/JEP/article/view/26683/27333

Kulkarni, A., Chong, D. & Batarseh, F. (2020). Foundations of data imbalanced and solutions for a data democracy. *Data Democracy*, 83-106. https://doi.org/10.1016/B978-0-12-818366-3.00005-8

Kutuk, G., David W., Linda K., & Bethan, G. (2022). Relations between Gender Stereotyping and Foreign Language Attainment: The Mediating Role of Language Learners' Anxiety and Self-efficacy. *British Journal of Educational Psychology, 92*(1), 212–235. https://doi.org/10.1111/bjep.12446

Lu, O., Huang, A., Huang, J. Lin, A., Ogata, H. & Yan, S. (2018). Applying learning analytics for the early prediction of students' academic performance in blended learning. *Educational Technology and Society, 21*(2), 220-232. https://doi.org/10.2307/26388400

MacIntyre, P. D., Clément, R., Dörnyei, Z. & Noels, K. A. (1998). Conceptualizing willingness to communicate in a L2: A situational model of L2 confidence and affiliation. *The Modern Language Journal, 82*(4), 545-562. https://doi.org/10.2307/330224

McCroskey, J. & Baer, J. (1985). Willingness to communicate: the construct and its measurement. Paper presented at the *Annual Meeting of the speech communication association*. Denver.

Ministerio de Educación y Formación Profesional. (2020). *Plan de Modernización de la Formación Profesional* [Modernization Plan of Vocational Education and Training]. Ministerio de Educación y Formación Profesional.

Mira, J., Delgado, A.E., Boticario, J.G., & Díez, F.J. (2003). *Aspectos básicos de la inteligencia artificial*. Sanz y Torres.

Mishra, T., Kumar, D., & Gupta, S. (2014). Mining students' data for prediction performance. In *Fourth International Conference on Advanced Computing & Communication Technologies,* (pp. 255-262). IEEE. https://doi.org/10.1109/ACCT.2014.105

Mitrofan, N., & Ion, A. (2013). Predictors of academic performance. The relation between the big five factors and academic performance. *Procedia-Social and Behavioral sciences, 78*, 125-129. https://doi.org/10.1016/j.sbspro.2013.04.264

Muamaroh, & Prihartanti, N. (2013). Willingness to communicate in English: A case study of Indonesian university students. *Kajian Linguistik dan Sastra Universitas Muhammadiyah Surakarta, 25*(1), 71-81.

https://publikasiilmiah.ums.ac.id/xmlui/bitstream/handle/11617/4564/6.pdf?sequence=1&isAllowed=y

Nabil, A., Seyam, M., & Abou-Elfetouh, A. (2021). Prediction of students' academic performance based on courses' grades using deep neural networks. *IEEE Access, 9*, 140731-140746. https://doi.org/10.1109/ACCESS.2021.3119596

Nasiruddin, U. A., & Hum, S.M. (2018). Strategies to support foreign language learning in large classes: a systematic literature review. *XLinguae* [XLanguages]*, 11,* 1, 3 – 15. https://doi.org/10.18355/XL.2018.11.01.01

Onwuegbuzie, A. J., Bailey, P., & Daley, C. E. (2000). Cognitive, affective, personality, and demographic predictors of foreign-language achievement. *The Journal of Educational Research, 94,* 3–15. https://www.jstor.org/stable/27542298.

Organization for Economic Cooperation and Development. (2021). A Balancing Act. Retrieved August 8, 2023 from https://www.oecd.org/economic-outlook/december-2021/

Pandey, M., & Taruna, S. (2014). A multi-level classification model pertaining to the student's academic performance prediction. *International Journal of Advances in Engineering & Technology, 7*(4), 1329-1341.

Pérez-Paredes, P., & Martínez-Sánchez, F. (2001). A Spanish version of the foreign language classroom anxiety scale: Revisiting Aida's factor analysis. *Revista Española de Lingüística Aplicada RESLA, 14*, 337-352.

Plataforma Observatorio de la Formación Profesional. (2021). Informe 2021. *La FP como clave de desarrollo y sostenibilidad* [Report 2021. VET as a Development and Sustainability Key]. https://www.observatoriofp.com/downloads/2021/informe-completo-2021.pdf

Popenici, S., & Kerr, S. (2017). Exploring the impact of artificial intelligence on teaching and learning in higher education. *Research and Practice in Technology Enhanced Learning, 12,* 22. https://doi.org/10.1186/s41039-017-0062-8

Raoofi, S., Bee Hoon T., & Swee, C. (2012). Self-Efficacy in Second/Foreign Language Learning Contexts. *English Language Teaching, 5*(11), 60-73. https://doi.org/10.5539/elt.v5n11p60

Ruiz, M. E., & Srinivasan, P. (2002). Hierarchical text categorization using neural networks. *Information Retrieval, 5*(1), 87-118. https://doi.org/10.1023/A:1012782908347

Saka, I. & Merç, A. (2021). The relationship among Turkish EFL learners' willingness to communicate in English, self-efficacy perceptions and linguistic self-confidence. *The Literacy Trek, 7*, 2, 1-36. https://doi.org/10.47216/literacytrek.770371

Santos, E. (2014). *Ansiedad y disposición a comunicarse en el aprendizaje del inglés como segunda lengua: estudio de las influencias del modelo formative (AICLE y Enseñanza formal). [Anxiety and willingness to communicate in learning English as a second language: A study of the influences of the formative model (CLIL and formal teaching*). [Doctoral dissertation, Universitat de les Illes Balears] https://www.tesisenred.net/handle/10803/291816#page=1

Shen, M., & Chiu, T. (2019). EFL learners' English speaking difficulties and strategy use. *Education and Linguistic Research, 5*(2), 88. https://doi.org/10.5296/elr.v5i2.15333

Shi, L. (2017). Empirical study on learners' self-efficacy in ESL/ EFL context. *College Student Journal, 50*(3), 454-465. https://www.thefreelibrary.com/Empirical+study+on+learners%27+self-efficacy+in+ESL%2fEFL+context.-a0466783194

Singh, W., & Waur, P. (2016). Comparative analysis of classification techniques for predicting computer engineering students' academic performance. *International Journal of Advanced Research in Computer Science, 7*(6), 31-36. https://doi.org/10.26483/ijarcs.v7i6.2793

Sison, N. R. M. (2022). Affective strategies in teaching and learning English as a Second Language (ESL). *International Research Journal of Science, Technology, Education, and Management, 2*(2), 79-90. 0. https://doi.org/10.5281/zenodo.6951410

Su, J.S., Zhang, B. F., & Xu, X. (2006). Advances in machine learning based text categorization. *Ruan Jian Xue Bao (Journal of Software), 17*(9), 1848-1859. https://www.jos.org.cn/josen/article/abstract/20060902

Tsao, N. L., Kuo, C. H., Guo, T. L., & Sun, T. J. (2017, July). Data Consideration for At-Risk Students Early Alert. In *Advanced Applied Informatics (IIAI-AAI), 2017 6th IIAI International Congress on* (pp. 208-211). IEEE.

Tomasevic, N., Gvozdenovic, N., & Vranes, S. (2020). An overview and comparison of supervised data mining techniques for student exam performance prediction. *Computers & Education, 143*, 103676. https://doi.org/10.1016/j.compedu.2019.103676

Uliyan, D., Aljaloud, A. S., Alkhalil, A., Al Amer, H. S., Mohamed, M. A. E. A., & Alogali, A. F. M. (2021). Deep learning model to predict students' retention using BLSTM and CRF. *IEEE Access, 9*, 135550-135558. https://doi.org/10.1109/ACCESS.2021.3117117

UNESCO (2019). *Evaluation of the Future of UNESCO's Education Sector: The Normative vs Operational Role in the Context of Agenda 2030.* https://unesdoc.unesco.org/ark:/48223/pf0000373235?posInSet=11&queryId=476fae8f-1ecc-4232-8121-c4f02af014f3.

Woodrow, L. (2006). A model of adaptive language learning. *Modern Language Journal, 90.* 297-319. https://doi.org/10.1111/j.1540- 4781.2006.00424.x

Xiao, J., Wang, M. Jiang, B., & Li, J. (2018). A personalized recommendation system with combinational algorithm for online learning. *Journal of Ambient Intelligence and Humanized Computing, 9*(3), 667-677. https://doi.org/10.1007/s12652-017-0466-8

Yang, Z. (2014). The role of input, interaction and output in the development of oral fluency. *English Language Teaching, 2*(4), 91-100. https://doi.org/10.5539/elt.v2n4p91

Zabihi, R. (2018). The role of cognitive and affective factors in measures of L2 writing. *Written Communication, 35*(1), 32-57. https://doi.org/10.1177/0741088317735836