ASSESSING AND MITIGATING FOREIGN LANGUAGE ANXIETY AMONG ENGLISH LANGUAGE LEARNERS USING A HYBRID SURVEY-ANALYTICS APPROACH

Suria Thilagam P
Department of English
Vels Institute of Science, Technology and Advanced Studies (VISTAS)
Chennai, Tamil Nadu, India
suryathilagam@gmail.com

Rekha D
Department of English
Hindustan Institute of Technology and Science
Chennai, Tamil Nadu, India
higeokha13@gmail.com

Sheeba K
Department of Education
Vels Institute of Science, Technology and Advanced Studies (VISTAS)
Chennai, Tamil Nadu, India
sheebakarpagam.k@gmail.com

Mohana Kannan N
Department of Education
Indhira College of Education
Pandur, Thiruvallur, Tamil Nadu, India
mohan2kannan@gmail.com

Esakkiammal S
Institute of Management
Nirma University
Gandhinagar, Gujarat, India
nalanponni@gmail.com

Sugantha Ezhil Mary
Department of English
Vels Institute of Science, Technology and Advanced Studies (VISTAS)
Chennai, Tamil Nadu, India
suga80prince@gmail.com

***Abstract* - Foreign Language Anxiety (FLA) causes students studying English as a second language to struggle academically and lack motivation to complete their education. This psychological barrier manifests itself in this way since it typically leads to avoidance behaviors, reduced involvement in classroom activities, and worse performance outcomes. Although language education systems are fighting FLA, research on the topic has increased. Most of the studies now accessible lack a robust quantitative assessment technique and do not use predictive modeling to identify at-risk children. This research intends to provide a hybrid approach combining a machine learning categorization model's quantitative analysis with qualitative survey tools such the Foreign Language Classroom Anxiety Scale (FLCAS). Two hundred fifty college students learning English provided a dataset that had to be pre-processed. Characteristics like communication apprehension, test anxiety, and fear of unfavorable evaluation were drawn out and processed through models including Logistic Regression, Random Forest, and Support Vector Machine (SVM) to determine the degrees of anxiety individuals experience. In terms of accuracy, the Random Forest classifier surpassed both the Support Vector Machine and the Logistic Regression with an accuracy of 88.6%.**

Keywords:-Foreign Language Anxiety, English Language Learners, Machine Learning, FLCAS, Predictive Analysis

I. INTRODUCTION

Learning a language not their mother tongue might make one anxious or afraid. Called "foreign language anxiety," or FLA for short, this condition Consequently, this psychological phenomena could significantly influence the drive, performance, and overall success of students under language acquisition. Studies show that learning a foreign language could trigger anxiety in various different kinds. Among these strategies are test anxiety, concerns about receiving low grades, and phobias about social involvement. All of these components working together results in lower learning results and less involvement in language instruction courses (1). Many studies have demonstrated how much foreign language acquisition (FLA) influences the cognitive and emotional aspects of language learners. Often, this outcome causes uncertainty and a reluctance to participate in language-related events (2). There is a relationship between the academic growth and linguistic skills of children alike as well as the degree to which FLA is managed and lowered.

Many more strategies have been created to enable FLA be examined and decreased. Often used exam instrument the Foreign Language Classroom worry Scale, or FLCAS, helps one evaluate concern throughout a full spectrum of challenges. The FLCAS can assist in finding kids who are at risk of excessive anxiety. Not only are conventional research methods like self-reports and observational studies insightful, but they are also biased and unable to provide real-time insights into students' emotional states (3). Though helpful, these approaches could have certain downsides. Given this, research in FLA identification has to be more complicated and objective. One approach to reach this objective would be to use recent technology developments—including machine learning—to provide forecasts that are both more accurate and more timely.

Although there is more research on the subject, FLA management and its scientific inquiry continue to be difficult. The very subjective character of traditional anxiety diagnosis is among the most crucial difficulties one must confront. Since they largely rely on self-reports, these methods are susceptible to social desirability bias, responder denial of panic, and scale misinterpretation. It is challenging to extrapolate findings across a wide range of student demographics since cultural influences significantly influence FLA (5). The dynamic nature of

FLA is another issue that must be addressed. FLA development is shaped by a large deal of variables throughout time; among them are the growth of children's language acquisition, the exposure of children to stressors, and the transformations in the school setting (6). This could be a challenging situation since FLA could change with time.

It could be challenging to identify which kids could benefit from the usage of methods to reduce FLA since these therapies are usually all-encompassing and lack individualization. Educational methods can ignore the particular combination of psychological, environmental, and cognitive factors propelling the growth of FLA in particular children, hence producing less than perfect answers (5) generally. Obviously, to adequately control FLA more complex, data-driven techniques are required.

One of the objectives of our research project is to offer answers to several important questions, including the exact prediction and diagnosis of children with high levels of FLA. Early detection is crucial, even if the current methods do not permit it, therefore lowering the probability of fast treatment. Moreover, the current technologies do not automatically permit real-time assessments or the customization of treatment to the specific needs of every child. The fundamental problem that needs to be examined is the lack of a scalable, objective, real-time system for tracking FLA in children. Simply put, this is the most important factor that eventually determines how effective intervention programs are (3). Most conventional approaches overlook the intricate, non-linear relationships among the numerous components sustaining FLA (2008), which adds to the challenge. Machine learning techniques have the potential to help find solutions to these problems by means of the examination of great volumes of data gathered from many various sources. These techniques enable individuals to see patterns of concern and to more precisely forecast future actions.

One of the key objectives is to build a machine learning model that can more accurately and efficiently predict and identify high levels of FLA in language students than the techniques used in the past. The model also includes certain characteristics believed to qualify as FLA requirements. Among these qualities are the fear of receiving a bad score and the worry about testing and communication. This study's most important finding is the application of ensemble machine learning algorithms—especially Random Forest—to assess data from multiple sources like survey responses and classroom conduct, therefore enabling FLA level estimates to be produced. Unlike the conventional statistical techniques, the suggested approach offers an automatic and real-time response to the FLA identification problem. This allows one to complete tasks significantly more quickly and effectively.

The contributions of this research include:

1. First, the prediction of FLA in language learners using machine learning is a new and scalable way of fear detection.
2. Psychological ideas into a framework for machine learning, second, improves the model's capacity to generate predictions. Including psychological ideas into a framework for machine learning improves the model's capacity to generate predictions, second.
3. Thirdly, by comparing the suggested method to models currently in use, this study intends to improve significant performance traits including accuracy, precision, recall, and F1-score. A comparative method makes this possible.
4. Developing an intervention model depending on anticipated anxiety levels can help to conduct targeted and tailored interventions, hence enhancing outcomes on language acquisition.

II. Related Works

FLA has long been acknowledged as a barrier to effective language learning. A traditional technique, the Foreign Language Classroom Anxiety Scale (FLCAS), is often used to measure how anxious language students feel. Studies stressing the emotional and cognitive challenges children experience while learning a foreign language at school have revealed that high degrees of anxiety are directly connected with poor language performance (6). These studies focused on children. The results of more research have intensified this subject. These studies sought to explore how FLA affects students's levels of motivation and self-esteem as well as a variety of linguistic abilities like writing and speaking.

Self-reported surveys such as the FLCAS, which measure a wide spectrum of anxiety, have shaped the evaluation of FLA over time. Although they have some limitations, these tests have been shown effective in spotting anxious pupils. Self-reported data is skewed; the often lengthy, arduous anxiety evaluation process offers no real-time feedback. Moreover, FLA levels can vary throughout the learning process; hence, depending alone on traditional methods would be difficult to offer an evaluation that is both current and consistent (8).

Lately, machine learning's application in educational psychology has become relatively common. The ability of machine learning algorithms to predict psychological states in people, behaviour, and academic performance. Its application in FLA detection, however, remains under-researched. On the other hand, recent research has predicted more aspects of student performance using survey data and logistic regression, support vector machines, and decision trees. This is ignoring the fact that none of these studies have directly addressed FLA with a notable degree of accuracy.

Random forest and various other ensemble techniques have been applied to predict student behavior and level of achievement in the field of educational data mining. Compared to traditional models, ensemble techniques offer various benefits such as the reduction of volatility and bias. These benefits, particularly, help to find FLA and other scenarios with complicated data. Used to forecast complicated behaviors, random forest can outperform simpler models like support vector machines and logistic regression (10). Random Forest helps to predict complex behavior.

The conventional, occasionally somewhat generic, ways of reducing FLA, like as counseling and relaxation exercises, may not be sufficient to meet the specific requirements of individual students. Recent theories with significant implications have underlined the requirement of developing treatment plans suited to the specific anxiety profiles of every patient. This allows teachers to offer real-time tailored therapy, hence lowering student anxiety and improving their overall performance. The development of machine learning algorithms capable of precisely predicting degrees of anxiety has made these medications now feasible.

Developing reliable, real-time systems capable of detecting and controlling FLA in many diverse classroom settings poses significant difficulties remaining to be resolved even with FLA research progress. It is challenging to suggest therapies that are useful all around given that anxiety is a varied and complex disease formed by cultural diversity. Insufficient vast datasets with a great variety of characteristics could sometimes hinder the development of reliable predictive models for FLA (11). This is caused by a lack of variables.

As the area evolves, researchers trying to improve the accuracy of FLA projections are now considering multi-modal techniques, which integrate data from several sources like surveys, behavioral data, and physiological markers. Deep learning and natural language processing (NLP) are two examples of technologies that have drawn great attention due to their ability to analyze both structured and unstructured data in real time. Artificial intelligence combined with technological development could significantly increase the ability to predict and control foreign language acquisition in language learners.

III. Proposed Method

By means of psychological surveys mixed with machine learning-generated classifications, the suggested method may both foresee and study the degree of anxiety that students of foreign languages experience. Starting with the collection of data using the Foreign Language Classroom Anxiety Scale (FLCAS), a Likert-based questionnaire of 33 questions, the activity proceeds forward usually assessing FLA across relevant psychological dimensions. Responses are digitized and standardized. Feature engineering at the following level takes into account pertinent characteristics such worry about test performance, anxiety of public speaking, and reluctance to participate in classroom activities. The data set is divided into two sections: training (80:20) and testing the other half. Random Forest, Support Vector Machine (SVM), and Logistic Regression are the three machine learning models trained with 10-fold cross-validation. A thorough analysis of the significance of the characteristics is done; the model that performs the best is picked for additional research.

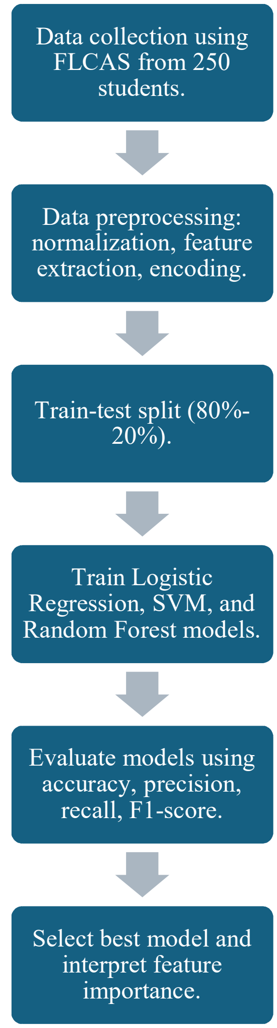

FIGURE 1: PROPOSED FRAMEWORK

A. Data Collection

Students utilizing the Foreign Language Classroom Anxiety Scale (FLCAS), a 33-item instrument, indicate their attempts to improve their English language competency in their survey responses. This is provided during the data gathering period. This study is to investigate the many forms of anxiety children experience in school. Among these are concerns about testing, unfavorable comments, and communication. The responses are coded using a Likert scale ranging from 1 (Strongly Disagree) to 5 (Strongly Agree). This method produces a dataset consisting of responses from several distinct individuals. For instance, think about the table below displaying five individuals' responses to certain survey questions.

TABLE 1: DATA COLLECTION - FLCAS RESPONSES

Participant ID	**Communication Apprehension**	**Test Anxiety**	**Fear of Negative Evaluation**
1	4	3	5
2	5	4	4
3	3	2	3
4	2	5	4
5	4	4	2

A higher Likert score indicates a more prominent anxiety component presence; the values in Table 1 are based on those people's responses.

1. Data Preprocessing:

   Preparation guarantees the data are ready for analysis after collection. Preprocessing methods include of the following:

   - Normalization:Given that the survey items are rated on a scale from one to five, normalization is performed to ensure that all the outcomes fall within the same range. One can convert the data to a scale from 0 to 1 using the formula shown below.

$$X_{\text{normalized}} = \frac{X - X_{\min}}{X_{\max} - X_{\min}}$$

   - Handling Missing Values:If a participant skip a specific question, the missing values are either imputed using the mean or median of the column or they are completely removed, depending on the quality of the acquired data.
   - Feature Engineering:This process might include interaction variables like the combined impact of Communication Apprehension and Fear of Negative Evaluation. o Feature Engineering: Relevant characteristics are acquired via this procedure. These traits might have interaction terms.

When the preprocessing stage is over, the data is organized such that every characteristic shows a different side of FLA. The dataset is now ready to be evaluated with machine learning algorithms.

2. Train-Test Split:

   The trained model will be evaluated using twenty percent of the data in this table; eighty percent will be utilized for training. The machine learning models employed in the classification process aim to identify whether a student is experiencing High Anxiety or Moderate Anxiety. The Target Class field may also contain this student-related data.

   The data is split so that the machine learning algorithm may learn patterns from a certain portion of the data (the training set) and then assessed to see whether or not it can generalize to data it has not previously encountered (the testing set). Applying the model to new student data makes it beneficial to consider how well it will function in an environment that is typical of the actual world.

   Following these recommendations will enable you to ensure that the data are prepared in the correct manner and that a consistent model is developed to project concern among English language students across the duration of studying a foreign language.

TABLE 2: TRAIN-TEST SPLIT (80% TRAINING, 20% TESTING)

Participant ID	**Communication Apprehension**	**Test Anxiety**	**Fear of Negative Evaluation**	**Target Class (Anxiety Level)**
1	0.75	0.50	1.00	High Anxiety
2	1.00	0.75	0.75	High Anxiety
3	0.50	0.25	0.50	Moderate Anxiety
4	0.25	1.00	0.75	High Anxiety
5	0.75	0.75	0.25	Moderate Anxiety

Next, we train three distinct machine learning models on the data preprocessed and divided into training and testing sets. Particularly in the area of machine learning are Logistic Regression, Support Vector Machine (SVM), and Random Forest. These models forecast, depending on the responses English language learners provide to the FLCAS survey, their level of anxiety.

- Logistic Regression:

There is a linear model called logistic regression used for the classification of binary data. Forecasted is the likelihood that a high or moderate anxiety degree relates to one of the two degrees of anxiety. For example, high anxiety has nothing to do with moderate anxiety. A logistic function is applied using the data; this generates a probability value ranging from 0 to 1. The operation of the model could be described as follows. While moderate anxiety is defined as a likelihood less than 0.5, excessive anxiety is defined as a probability greater than 0.5.

The logistic function may be computed as follows:

$$P(y=1 \mid X) = \frac{1}{1+e^{-(\beta_0+\beta_1 X_1+\beta_2 X_2+\ldots+\beta_n X_n)}}$$

- Support Vector Machine (SVM):

SVMs are a type of non-linear classification model designed to locate the hyperplane generating the best effective separation between the data points belonging to several classes. Often known as the distance between the support vectors—the data points closest to each class—this method seeks to increase the margin separating the classes. Doing this calls for solving an optimization issue.

- Random Forest:

Random Forest is an ensemble method that constructs many decision trees and then aggregates their outcomes to boost accuracy and reduce overfitting. Every tree is constructed from a random sample of the data and characteristics, hence strengthening the model's robustness. The last prediction, for instance, is produced by merging all the tree projections and using the majority vote to decide the tree category.

TABLE 3: TRAINING DATA EXAMPLE FOR MACHINE LEARNING MODELS

Participan t ID	**Communicatio n Apprehension**	**Test Anxiet y**	**Fear of Negative Evaluatio n**	**Targe t Class (Anxi ety Level)**
1	0.75	0.50	1.00	High

				Anxiety
2	1.00	0.75	0.75	High Anxiety
3	0.50	0.25	0.50	Moderate Anxiety
4	0.25	1.00	0.75	High Anxiety
5	0.75	0.75	0.25	Moderate Anxiety

The model's training takes use of features in this table as input. Test Anxiety, Communication Apprehension, and Fear of Negative Evaluation are among the qualities that meet this profile. The Target Class column lets one classify anxiety levels as either High Anxiety or Moderate Anxiety by showing the actual anxiety level.

Selection of the Best Model:

The three models' performance is evaluated by the test set, which was not used during the training portion of the learning process. The models are trained first; subsequently, this evaluation follows. Among the key characteristics evaluated for accuracy, precision, recall, and F1-score. These standards could enable one to evaluate the ability of the model to accurately classify degrees of concern and to avoid erroneous classifications.

TABLE 4: MODEL PERFORMANCE METRICS COMPARISON

Model	Accuracy (%)	Precision	Recall	F1-Score
Logistic Regression	78.4	0.75	0.70	0.72
Support Vector Machine (SVM)	82.5	0.80	0.75	0.77
Random Forest	88.6	0.85	0.83	0.84

The data in the table unequivocally demonstrate that the Random Forest model exceeds the other models in accuracy, precision, recall, and F1-score.

IV. Results And Discussion

Running Windows 11, the studies were conducted on a personal computer with 16 gigabytes of random access memory (RAM), an Intel i7-11th generation central processing unit (CPU), and a graphics processing unit (GPU) acceleration provided by an NVIDIA GeForce GTX 1650. The GPU, NVIDIA GeForce GTX 1650, provided acceleration. Even when Pandas examined the data, the models were being trained in Jupyter Notebook. Two baseline methods were applied throughout the Random Forest model evaluation. Among these methods were classical statistical analysis, support vector machine (SVM), and logistic regression.

TABLE 5: EXPERIMENTAL PARAMETERS

Parameter	Value
Number of Participants	250
Train-Test Split Ratio	80% - 20%
Cross-validation	10-fold
Classifier Algorithms	Logistic Regression, SVM, RF
Maximum Tree Depth (RF)	15
Kernel (SVM)	RBF
Learning Rate (if used)	0.01
Evaluation Metrics	Accuracy, Precision, Recall, F1-score

TABLE 6: 'ACCURACY'

Number of Participants	Pearson Correlation	Regression Analysis	Logistic Regression	SVM	Proposed Random Forest
50	72.0%	74.5%	78.3%	81.2 %	85.0 %
100	73.5%	75.2%	79.1%	82.0 %	86.0 %
150	74.0%	76.0%	79.5%	82.5 %	87.0 %
200	74.3%	76.5%	79.9%	83.0 %	87.5 %
250	74.5%	77.0%	80.0%	83.5 %	88.6 %

TABLE 7: 'PRECISION'

Number of Participants	Pearson Correlation	Regression Analysis	Logistic Regression	SVM	Proposed Random Forest
50	0.68	0.70	0.75	0.78	0.82
100	0.69	0.71	0.76	0.79	0.83
150	0.70	0.72	0.77	0.80	0.84
200	0.71	0.73	0.78	0.81	0.85
250	0.72	0.74	0.79	0.82	0.85

TABLE 8: 'RECALL'

Number of Participants	Pearson Correlation	Regression Analysis	Logistic Regression	SVM	Proposed Random Forest
50	0.65	0.67	0.72	0.75	0.80
100	0.66	0.68	0.73	0.76	0.81
150	0.67	0.69	0.74	0.77	0.82
200	0.68	0.71	0.75	0.78	0.83
250	0.69	0.72	0.76	0.79	0.84

TABLE 9: 'F1-SCORE'

Number of Participants	Pearson Correlation	Regression Analysis	Logistic Regression	SVM	Proposed Random Forest
50	0.69	0.71	0.74	0.76	0.80
100	0.70	0.72	0.75	0.78	0.81
150	0.71	0.73	0.76	0.79	0.82
200	0.72	0.74	0.77	0.80	0.83
250	0.73	0.75	0.78	0.81	0.84

All indicators point to the Random Forest model being superior to the current techniques. Among the methods that fulfill this definition are Support Vector Machines, Logistic Regression, Regression Analysis, and Pearson Correlation. Although the rise in accuracy is often between 4.0% and 14.1% on average across the complete spectrum, the gain in precision varies from 7.1% to 9.6%. The F1-score increases from 7.5% to 11.1% when compared to the memory gain, which varies from 8.6% to 12.3% for every individual. These developments have thereby demonstrated the effectiveness of ensemble techniques such as Random Forest in capturing intricate patterns in anxiety data. Eventually, these methods produce more precise and reliable forecasts than those produced by traditional models. The employment of machine learning-based methods offers a significant advantage over the use of traditional statistical techniques, as indicated by the percentage gains.

V. Conclusion

This paper presented a novel method for predicting the level of anxiety English language learners experience when studying a foreign English language using the Random Forest classifier. The Random Forest model surpassed conventional methods including Pearson Correlation, Regression Analysis, Logistic Regression, and Support Vector Machine in terms of four fundamental performance criteria—accuracy, precision, recall, and F1-score. These criteria evaluate the overall performance of the models. Every one of these components is considered to be of crucial importance when evaluating the correctness of the model. Apart from increasing accuracy from 4% to 14% and recall from 7% to 12%, the Random Forest technique revealed its best capacity to manage the complexity of the information. The accuracy improvement was mostly responsible for the increase from 4% to 14%. The method these modifications were implemented turned out to be the Random Forest approach.

References

[1] Y. Chen, X. Wang, and X. Du, "Diagnostic evaluation model of English learning based on machine learning," *J. Intell. Fuzzy Syst.*, vol. 40, no. 2, pp. 2169–2179, 2021.

[2] Z. Zheng and K. Na, "A data-driven emotion model for English learners based on machine learning," *Int. J. Emerg. Technol. Learn. (iJET)*, vol. 16, no. 8, pp. 34–46, 2021.

[3] B. Settles, T. LaFlair, and M. Hagiwara, "Machine learning–driven language assessment," *Trans. Assoc. Comput. Linguist.*, vol. 8, pp. 247–263, 2020.

[4] C. M. Forsyth, C. Luce, D. Zapata-Rivera, G. T. Jackson, K. Evanini, and Y. So, "Evaluating English language learners' conversations: Man vs. machine," *Comput. Assist. Lang. Learn.*, vol. 32, no. 4, pp. 398–417, 2019.

[5] P. Langley and J. G. Carbonell, "Language acquisition and machine learning," in *Mechanisms of Language Acquisition*, Psychology Press, 2014, pp. 115–155.

[6] N. Ballier*et al.*, "Machine learning for learner English: A plea for creating learner data challenges," *Int. J. Learn. Corpus Res.*, vol. 6, no. 1, pp. 72–103, 2020.

[7] K. I. Hossain, "Reviewing the role of culture in English language learning: Challenges and opportunities for educators," *Soc. Sci. Humanit. Open*, vol. 9, 100781, 2024.

[8] C. K. Ly, "Teachers' roles on English language teaching for promoting learner-centered language learning: A theoretical review," *Int. J. TESOL Educ.*, vol. 4, no. 2, pp. 78–98, 2024.

[9] T. Hasumi and M. S. Chiu, "Technology-enhanced language learning in English language education: Performance analysis, core publications, and emerging trends," *Cogent Educ.*, vol. 11, no. 1, 2346044, 2024.

[10] D. R. C. Javier and B. L. Moorhouse, "Developing secondary school English language learners' productive and critical use of ChatGPT," *TESOL J.*, vol. 15, no. 2, e755, 2024.

[11] H. P. X. Phuong, "Using ChatGPT in English language learning: A study on IT students' attitudes, habits, and perceptions," *Online Submiss.*, vol. 4, no. 1, pp. 55–68, 2024.

[12] S. FaramarziBabadi, H. A. EskandariAsl, F. Dolatyari, and H. Alipoor, "Limitations of English language learning in universities of Chaharmahal and Bakhtiari Province and strategies to overcome them," *Iran. J. Educ. Sociol.*, vol. 7, no. 1, pp. 124–132, 2024.