IEEE Access
Multidisciplinary ⋮ Rapid Review ⋮ Open Access Journal

Received 18 February 2024, accepted 16 April 2024, date of publication 19 April 2024, date of current version 30 April 2024.

Digital Object Identifier 10.1109/ACCESS.2024.3391418

RESEARCH ARTICLE

Multi-Class Adaptive Active Learning for Predicting Student Anxiety

AHMAD ALMADHOR[1], SIDRA ABBAS[2], (Graduate Student Member, IEEE), GABRIEL AVELINO SAMPEDRO[3,4], (Member, IEEE), SHTWAI ALSUBAI[5], STEPHEN OJO[6], (Member, IEEE), ABDULLAH AL HEJAILI[7], AND LUBOMIRA STRAZOVSKA[8]

[1]Department of Computer Engineering and Networks, College of Computer and Information Sciences, Jouf University, Sakaka 72388, Saudi Arabia
[2]Department of Computer Science, COMSATS University Islamabad, Islamabad 44000, Pakistan
[3]Faculty of Information and Communication Studies, University of the Philippines Open University, Los Baños 4031, Philippines
[4]Gokongwei College of Engineering, De La Salle University, Malate, Manila 1004, Philippines
[5]College of Computer Engineering and Sciences, Prince Sattam bin Abdulaziz University, Al-Kharj 11942, Saudi Arabia
[6]Department of Electrical and Computer Engineering, College of Engineering, Anderson University, Anderson, SC 29621 USA
[7]Faculty of Computers and Information Technology, Computer Science Department, University of Tabuk, Tabuk 71491, Saudi Arabia
[8]Faculty of Management, Comenius University in Bratislava, 82005 Bratislava, Slovakia

Corresponding authors: Lubomira Strazovska (lubomira.strazovska@fm.uniba.sk) and Sidra Abbas (sidraabbas@ieee.org)

This work was supported by the Deanship of Scientific Research at Jouf University under Grant DSR-2021-02-0383.

ABSTRACT This research delves into applying active and machine learning techniques to predict student anxiety. This research explores how these technologies can be explored to understand and predict student anxiety levels. This study utilizes active learning strategies to increase the effectiveness of machine learning models in predicting anxiety levels among students. Additionally, this adaptation emphasizes the usefulness of active learning methodologies in enhancing the precision of machine learning models for student anxiety prediction. This study uses two datasets containing information on student behavior and leverages machine learning methods to construct predictive models for student anxiety. This study uses various machine learning models: K-Nearest Neighbors (KNN), Logistic Regression (LR), XGBoost (XGB), Naive Bayes (NB), and Random Forest (RF). Experiments revealed that active learning-based LR yielded a score of 0.61, and RF performed well with an average accuracy of 0.60 on the first dataset. Similarly, for the second dataset, RF is the most effective model, achieving an accuracy of 0.83. These results provide valuable insights into the models' performance across key metrics. Further, this research highlights the potential of employing machine learning techniques and active learning methodologies to predict and manage student anxiety.

INDEX TERMS Active learning, student anxiety, student wellbeing, inclusive education, machine learning, deep learning.

I. INTRODUCTION

In modern education, nurturing students' well-being and mental health has gained unprecedented importance. A dynamic interaction of social, technological, and pedagogical shifts characterizes the modern educational landscape [1]. In light of these changes, student anxiety has become a prominent concern. Likewise, the present age has witnessed a discernible surge in reported instances of student anxiety, as documented by comprehensive studies [2], [3]. This surge underscores the compelling need to effectively address the various challenges posed by anxiety within educational contexts [4]. Moreover, the contours of student anxiety exhibit diverse forms, encompassing the vision of test anxiety, the intricacies of social anxiety, and the weight of performance anxiety. In addition, the clear consequences of anxiety are far-reaching, rippling through the educational landscape in the form of diminished academic performance, diminished enthusiasm for learning, and an overall decline in the quality of the educational experience [5].

The associate editor coordinating the review of this manuscript and approving it for publication was Alba Amato.

These challenges appeal to creating a fine understanding of the complex stressors and pressures students face inside and outside the classroom [6]. There has been a significant shift towards a more comprehensive approach to predicting student anxiety. Beyond the usual surveys and online behavior analysis, exploring new prospects includes diving into biometrics, including measuring heart rate variability, skin conductance, and eye movement tracking [7]. Likewise, these biometric measures give us a much deeper understanding of how anxiety affects students physically, making predictive models more accurate and insightful [8]. More precisely, how students interact with each other and their instructors in physical classrooms and online settings has become a major focus. Overall, studying peer dynamics, teacher influence, and support networks to understand student anxiety.

Combining machine learning and active learning methodologies to predict and reduce student anxiety represents a powerful approach to improving educational outcomes. Machine learning algorithms can analyze various data sources, such as students' academic performance, participation levels, and physiological data like heart rate or skin conductance [9]. These algorithms can then identify patterns and early warning signs of anxiety [10]. Integrating active learning and machine learning techniques into the educational process makes it possible to address these indicators [11]. This approach can help students better engage in the learning process, leading to better outcomes. For instance, personalized interventions can be designed to provide additional support and resources to at-risk students, or course materials can be adapted based on real-time feedback to reduce anxiety triggers. In other words, this harmony between machine learning and active learning enhances the understanding of student anxiety. It offers tailored solutions to promote a healthier and more productive learning environment, ultimately improving students' well-being and academic success.

In sum, this research undertakes a journey, recognizing that in today's dynamic educational landscape, students' well-being is intrinsically tied to their academic success. Reducing anxiety levels promotes resilience, engagement, and inclusivity in education while easing student anxiety and making education a supportive arena where students feel empowered and belong. Their anxieties subside, leading to academic success.

A. CONTRIBUTIONS

- **Applying Active Learning in the Context of Predicting Student Anxiety:** This study utilizes an active learning approach to increase the effectiveness of machine learning models in predicting anxiety levels among students. Additionally, this adaptation emphasizes the usefulness of active learning methodologies in enhancing the precision of machine learning models for Student Anxiety Prediction.
- **Enhancing Prediction Rate of Student Anxiety with Active Learning:** The primary goal is to enhance the predictive capabilities of student anxiety prediction, focusing on accuracy, precision, and recall as critical performance metrics. Results revealed that active learning-based LR yielded a score of 0.61, and RF performed well with an average accuracy of 0.60 on the first dataset. Similarly, for the second dataset, RF is the most effective model, achieving an accuracy of 0.83. These results provide valuable insights into the models' performance across key metrics. Further, this research highlights the potential of employing machine learning techniques and active learning methodologies to predict and manage student anxiety.

B. PAPER ORGANISATION

This study starts by explaining the research methodology employed in predicting student anxiety. It outlines the data acquisition process, the active learning techniques, and the model evaluation procedures. Subsequently, the study presents its predictive outcomes, demonstrating its efficacy in identifying student anxiety levels, supported by visualizations and statistical metrics for clarity. Following this, the study delves into the practical implications of these findings. Likewise, it explores how the predictive model can enhance student well-being and the quality of education by providing timely interventions and support. Finally, this study explores potential future machine learning applications in student anxiety. It contemplates how this technology can contribute to a more proactive approach to addressing anxiety among students, potentially revolutionizing how educational institutions support their students' mental health.

II. RELATED WORK

Recognizing student anxiety as a significant concern within educational settings has prompted the development of predictive models [12]. Combining active learning techniques with machine learning algorithms offers a promising approach to identifying and supporting at-risk students. This review delves into the evolution of this interdisciplinary field. Moreover, in the initial stages of research, the focus was on identifying factors contributing to student anxiety [13]. Studies on academic stress and anxiety among college students laid the groundwork by highlighting stressors related to academic workload, peer pressure, and self-esteem. Researchers began integrating machine learning into predicting student anxiety [8]. Likewise, predicting student anxiety in online learning environments employed machine learning algorithms to analyze online learning behaviors and uncover anxiety-associated patterns. As well as active learning techniques have become essential for gathering relevant data [14]. Using crowdsourcing for collecting student anxiety data explored the potential of crowdsourcing to collect real-time student anxiety data, facilitating more dynamic predictions [15].

Researchers are pushing the boundaries of predicting student anxiety by looking at more than surveys and online behaviors. They investigate biometrics like heart rate variability, skin conductance, and eye-tracking data. These measurements offer a deeper understanding of how anxiety affects the body and can make predictive models even more accurate. Moreover, another exciting development is considering the role of social interactions in physical classrooms and online learning environments. Scientists are studying how students interact with their peers, their support networks and how instructors influence their anxiety levels [16]. Most importantly, this approach gives a complete picture of what causes student anxiety. Similarly, these innovations are not just about academic curiosity. They have real-world applications that could improve the lives of students. These advanced techniques can create personalized interventions and support systems for each student's unique needs. Moreover, through the smart use of data and machine learning, teachers can gain valuable insights into student anxiety, helping them offer timely support and create a more supportive learning atmosphere [17], [18]. Likewise, integrating environmental monitoring systems in educational spaces can provide students with a healthier and more comfortable atmosphere and reduce anxiety triggers [19].

Advancements led to hybrid models that combine active learning and machine learning [20]. A Hybrid Approach to predicting student anxiety stands as a unique example. The active feedback mechanisms within machine learning models adapt to individual student needs, enhancing prediction accuracy. Subsequently, as predictive models matured, ethical concerns surfaced. In the ethical implications of predicting student anxiety, researchers critically examined issues surrounding data privacy, informed consent, and the potential stigmatization of students with anxiety [21]. Furthermore, ethical guidelines emerged as a crucial aspect of research and implementation [22]. To ensure that predictive models are applicable across different cultures, cross-cultural studies have become increasingly important [23]. Similarly, a cross-cultural analysis of student anxiety prediction explored how cultural differences impact anxiety prediction and identified strategies for developing culturally sensitive models [24]. Recent efforts have transitioned from theoretical models to real-world applications. Implementing anxiety prediction in educational institutions discussed the challenges and successes of deploying predictive models in educational settings, demonstrating the potential impact on student well-being [25].

This section discusses emerging trends, including integrating additional factors, such as social interactions and academic performance, to enhance prediction accuracy. It also explores the potential for predictive models to adapt to changing learning environments, such as online and hybrid learning. In conclusion, the intersection of active learning and machine learning in predicting student anxiety has witnessed significant growth. This comprehensive review highlights the field's evolution, emphasizing ethical considerations and real-world implementations. Hence, as research advances, it holds promise for improving students' mental health and academic success worldwide.

III. METHODOLOGY

This study adopts a systematic approach to predicting student anxiety by leveraging active learning and machine learning methodologies. The Student Anxiety Dataset[1] contains essential information about student anxiety levels. The goal is to predict and understand factors contributing to anxiety among students, ultimately helping in better support and intervention strategies. Moreover, the initial phase of this research involves meticulous preparation of the anxiety dataset. This critical step includes cleaning data to eliminate inconsistencies, addressing missing information, and transforming features to ensure they are in the appropriate scale or format. Likewise, the study conducts a comprehensive exploratory data analysis to gain insights into the dataset's characteristics. These insights guide subsequent decisions as the research progresses. Furthermore, this study collects and organizes data related to various student-related parameters to better comprehend the factors associated with student anxiety. These parameters encompass academic performance, extracurricular involvement, social interactions, and personal attributes. To provide a more comprehensive analysis, this study analyses the data, recognizing the significant impact of these factors on student anxiety levels. This well-organized dataset is a valuable resource for predicting and addressing student anxiety, enabling educational institutions to tailor their support and intervention strategies effectively. In addition, a secondary dataset from Kaggle was employed to assess the model performance. This dataset involves 787 responses from University of Lahore undergraduate students inspired by the Beck Depression and Beck Anxiety inventories. It is designed to evaluate machine learning models in classifying depression and anxiety severity, with a commitment to ethical considerations such as informed consent and confidentiality during data collection. Similarly, data preprocessing for the secondary dataset was completed, encompassing all necessary steps to ensure its preparation for model training. Figure 1 outlines the steps and approach employed in this research.

Algorithm 1 provides the overview of the whole pipeline of the proposed approach that starts with data collection and the pre-processing of this data. Next, we perform label encoding and apply normalization to scale the dataset. Next, we initialize active learning-based algorithms and run the experiments in three iterations. Each iteration provides results according to evaluation metrics, and later, we extract the confusion matrix and roc curves.

This study initiated the analysis by acquiring a dataset and loading it into a Pandas data frame. The dataset encompassed a variety of columns, each representing distinct attributes or features. An initial examination of the dataset's columns was conducted to understand its structure comprehensively. This examination involved iterating through the columns using a for loop. Moreover, after the initial examination, certain columns were identified as irrelevant or redundant

[1] https://www.kaggle.com/datasets/petalme/student-anxiety-dataset

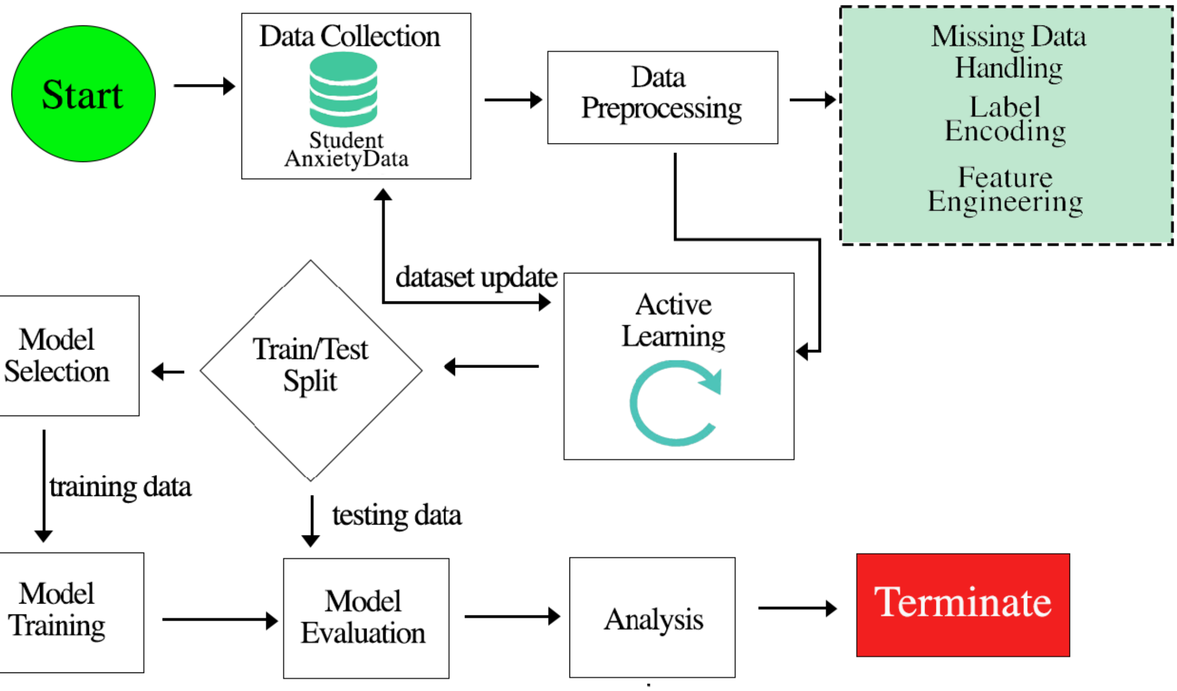


FIGURE 1. Employed methodology.

for the analysis. Specifically, the 'S. No.' and 'Timestamp' columns were deemed non-contributory to the research objectives and subsequently removed from the data frame. Likewise, a feature engineering process enhanced the dataset's richness, creating a novel feature called 'Hours_streams.' This new feature was generated by aggregating the values from the 'Hours' and 'Streams' columns. Subsequently, data cleaning procedures were employed to eliminate rows in which the 'Hours_streams' value exceeded 115 or equal to 0. This rigorous data-cleaning process was instrumental in ensuring that the dataset remained free from outliers and data points that were not pertinent to the analysis. To standardize the numerical features, the study applied Min-Max scaling to specific columns, including 'Hours,' 'streams,' 'Age,' 'GAD_T,' 'SWL_T,' and 'SPIN_T.' This transformation harmonized these attributes to a consistent scale, ranging from 0 to 1. Similarly, for exploratory data analysis, the study computed the Spearman correlation matrix to analyze relationships among variables within the dataset [26]. Furthermore, a heatmap was generated to visually display the extent of association between different variables to enhance the interpretability of these relationships. In terms of feature selection, a new data frame named 'df1' was introduced, encompassing columns that pertained to anxiety scores ('GAD1' through 'GAD7'), the incremental anxiety score ('GAD_T'), social phobia scores ('SPIN1' through 'SPIN17'), and the aggregate social phobia score ('SPIN_T'). This strategic feature selection enabled a focused analysis of specific attributes relevant to the research objectives. In addition, label encoding was applied to the 'GADE' column to prepare the dataset for machine learning models, converting categorical data into numerical form [27]. Most importantly, this encoding process assigned a distinct numerical value to each category within the 'GADE' column, facilitating subsequent analyses. In a final step, the researchers addressed missing data by identifying and removing rows with missing values in 'df1.' This data imputation procedure ensured the dataset's integrity, rendering it suitable for further analyses. In preparation for model training and evaluation, the dataset was partitioned into input features ('X,' excluding 'GADE') and the target variable ('Y,' representing 'GADE').

The research adopted a unique approach, incorporating active learning and following a step-by-step process. Initially, the models were trained using a small dataset from the anxiety dataset. Subsequently, the models actively identified samples from the same dataset that were uncertain or provided valuable information, contributing to an improved understanding of student anxiety over time. To assess the models' performance, the team utilized standard metrics such as accuracy, F1 score, precision, and recall [28]. These metrics provided valuable insights into the effectiveness of recognizing and understanding student anxiety using the anxiety dataset. In the results and analysis section, the research presented the findings from these experiments, highlighting the impact of active learning techniques on the accuracy of student anxiety identification. Furthermore, the research discussed the broader implications of these findings in the context of student well-being and educational support while also acknowledging the limitations of their approach and suggesting potential avenues for future research. Overall, the

Algorithm 1 Anxiety Recognition Algorithm

Data Collection:
Data = $\mathcal{C}$(*StudentAnxietyData*)
Data Preprocessing:
Initialize dataset preprocessing.
Handle missing data.
Remove rows where 'Hours_streams'>115 or equals 0.
Label Encoding:
Apply label encoding to the 'GADE' column to convert categorical data into numerical form.
Standardization (Min-Max Scaling):
Apply Min-Max scaling to specific columns ('Hours,' 'streams,' 'Age,' 'GAD_T,' 'SWL_T,' 'SPIN_T').
Dataset Preparation for Machine Learning:
Partition the dataset into input features 'X' (excluding 'GADE') and the target variable 'Y' ('GADE').
Initialize labeled dataset with seed instances for active learning approach
for Each n-iterations **do**
    Choose instances for manual annotation.
    Process and update in the labeled dataset.
end for
Initialize each ML model.
for Each ML model **do**
    Initiate the training on the labeled dataset.
    Model Training and Evaluation:
    Generate accuracy, precision, recall, confusion matrix and ROC curve.
end for

methodology employed by the research team, in conjunction with the anxiety dataset and active learning, represents a comprehensive approach to improving the well-being of students.

IV. RESULTS AND EXPERIMENTATION

In predicting student anxiety through active learning and machine learning, this study thoroughly analyzes experiments using various machine learning models to discern daily physical behaviors in older adults. The primary objective is to identify the most effective model for this task, all the while assessing the model's consistency and reliability. Table 1 shows the experimental results of the trained models in three different iterations of active learning. Initially, the K-Nearest Neighbors model consistently achieved an accuracy rate of 56.70% across three iterations, showcasing its reliability. Subsequently, the LR model yielded an accuracy of 61.13% across all iterations, surpassing KNN in terms of accuracy. However, a closer examination of precision and recall values reveals potential limitations in capturing nuanced aspects of daily behaviors in older adults. Furthermore, the NB model, with an accuracy rate of 44.12% across iterations, demonstrated slightly reduced performance compared to LR. Nevertheless, it exhibited higher precision, signifying its proficiency in correctly classifying positive instances, even though its slightly lower recall indicated room for improvement in capturing all relevant activities. In contrast, the RF model emerged as a standout performer with an impressive accuracy rate of 60.97%. Its consistency and effectiveness, supported by closely aligned F1 Score, precision, and recall values, make it a strong contender for activity recognition. Most importantly, considering that other models displayed lower accuracy, precision, recall, and F1 Score values, it suggests the possibility of superior choices for this particular task. Experiments with XGBoost demonstrated an accuracy rate of 59.44%, closely mirroring the RF model's performance. This shows that both models offer comparable

TABLE 1. Experimental results on first dataset using active learning (all iterations).

Model	Accuracy	F1-Score
K-Nearest Neighbors		
Iteration 1	0.5670	0.5442
Iteration 2	0.5670	0.5442
Iteration 3	0.5670	0.5442
Logistic Regression		
Iteration 1	0.6113	0.5721
Iteration 2	0.6113	0.5721
Iteration 3	0.6113	0.5721
Naive Bayes		
Iteration 1	0.4412	0.4547
Iteration 2	0.4412	0.4547
Iteration 3	0.4412	0.4547
Random Forest		
Iteration 1	0.6031	0.5793
Iteration 2	0.6097	0.5858
Iteration 3	0.6058	0.5828
XGBoost		
Iteration 1	0.5944	0.5793
Iteration 2	0.5944	0.5793
Iteration 3	0.5944	0.5793

TABLE 2. Experimental results on secondary dataset using active learning (all iterations).

Model	Accuracy	F1-Score	Precision	Recall
K-Nearest Neighbors				
Iteration 1	0.7500	0.7479	0.7762	0.7500
Iteration 2	0.7390	0.7323	0.7695	0.7390
Iteration 3	0.7574	0.7462	0.7989	0.7574
Logistic Regression				
Iteration 1	0.6176	0.6180	0.6238	0.6176
Iteration 2	0.5441	0.5435	0.5451	0.5441
Iteration 3	0.5188	0.5552	0.5655	0.5588
Naive Bayes				
Iteration 1	0.6287	0.6288	0.6289	0.6287
Iteration 2	0.6066	0.6060	0.6067	0.6066
Iteration 3	0.5662	0.5531	0.5841	0.5662
Random Forest				
Iteration 1	0.8199	0.8190	0.8451	0.8272
Iteration 2	0.8235	0.8222	0.8400	0.8199
Iteration 3	0.8125	0.8106	0.8506	0.8235
XGBoost				
Iteration 1	0.7978	0.7976	0.8110	0.7978
Iteration 2	0.8088	0.8080	0.8160	0.8088
Iteration 3	0.7757	0.7730	0.7839	0.7757
DT				
Iteration 1	0.8199	0.8192	0.8402	0.8199
Iteration 2	0.8346	0.8318	0.8618	0.8346
Iteration 3	0.7941	0.7915	0.8039	0.7941

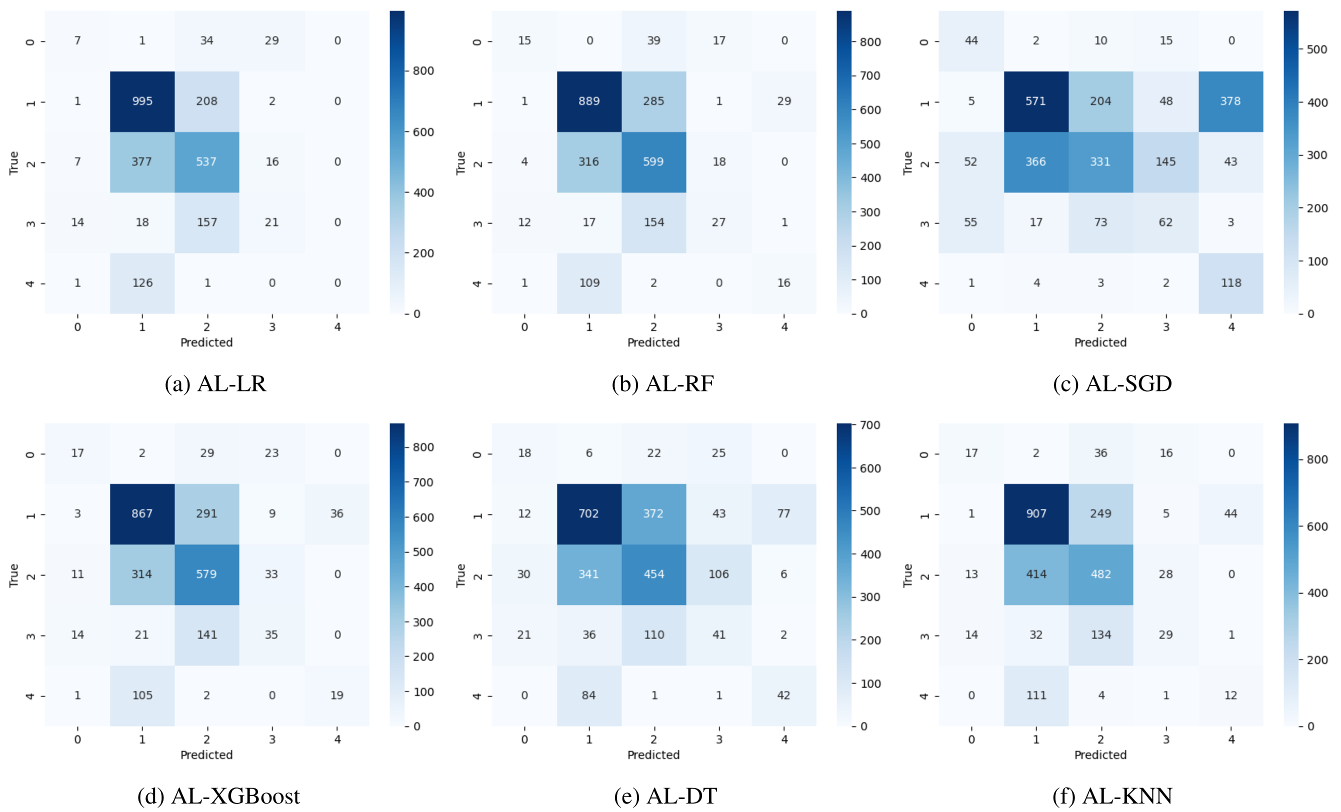
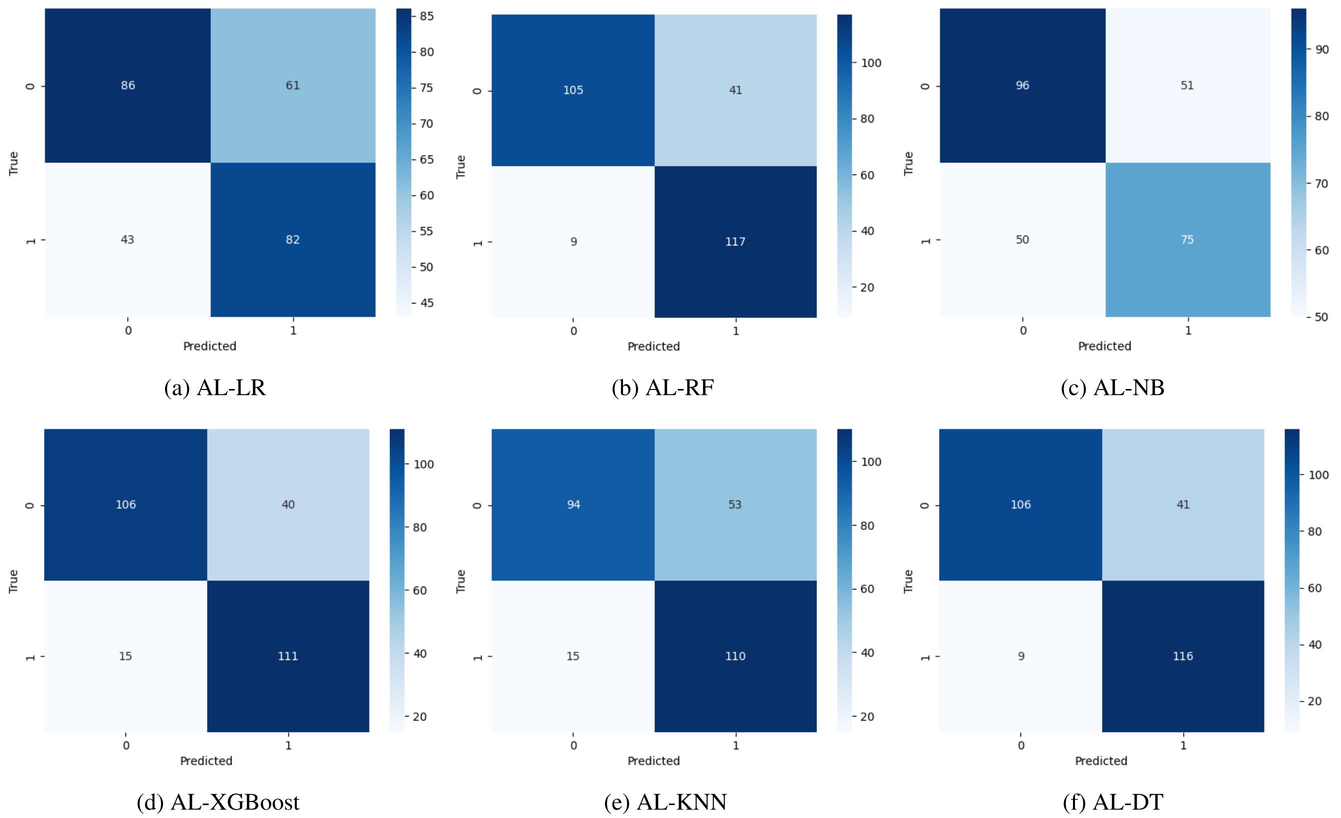


FIGURE 2. **Confusion matrix for active learning models on primary dataset.**

levels of accuracy, affording flexibility in selecting the most appropriate model for specific usage scenarios. In conclusion, this extensive experimentation highlights KNN, RF, and XGBoost as potential models for recognizing daily physical behaviors in older adults. KNN maintains solid accuracy and consistency, while RF and XGBoost provide competitive accuracy levels, making them preferable for more intricate scenarios. The choice of the most suitable model hinges on the specific requirements and priorities of eldercare and healthcare technology applications. Further analysis, including fine-tuning and feature engineering, is recommended to optimize model performance. Receiver Operating Characteristic Curves (ROC) scores for different models help assess the performance of binary classification models. Likewise, these curves display the balance between True Positive and False Positive rates at different model thresholds, providing insight into a model's ability to distinguish between positive and negative classes. A greater area under the curve indicates enhanced discriminatory power [29]. For instance, the LR model exhibits excellent discriminative ability (AUC = 0.94) for class 0, very good performance (AUC = 0.78) for class 1, and reasonable discrimination (AUC = 0.75) for class 2. At the same time, it demonstrates strong discrimination (AUC = 0.85) for class 3 and class 4. On the other hand, XGB achieved high ROC curve areas (AUC) for classes 0 (0.92), 3 (0.83), and 4 (0.94), indicating strong predictive performance, while classes 1 (0.76) and 2 (0.73) had lower AUC values, suggesting relatively weaker predictive accuracy.

Figure 2 depicts the confusion metrics of all active learning-based classifiers. Class 0 is usually confused with classes 2 and 3 for all classifiers. Classes 1 and 2 are easily classifiable, resulting in good classification scores. Class 4 seems to be highly confused with class 1 as well.

Moreover, Table 2 presents the experimental results on a secondary dataset, evaluating the performance of five machine learning models across three iterations. For K-Nearest Neighbors, the metrics–accuracy, F1 score, precision, and recall–remain consistently around 0.75-0.79 in all iterations, indicating a stable but moderate performance. Subsequently, LR, on the other hand, exhibits lower overall performance with an average accuracy of approximately 0.56 across all iterations, suggesting that it may not be as effective on this dataset as KNN. Similarly, NB demonstrates modest performance with the best accuracy of about 0.62 in all three iterations. On the other hand, RF and DT outperform the other models, with higher accuracy values reaching around 0.82 and 0.83, respectively. Likewise, the F1 score, precision, and recall metrics for RF and DT consistently show relatively high and stable values, indicating their effectiveness on this dataset. Overall, RF and DT emerge as the best models for this dataset, while KNN, LR, NB,



(a) AL-LR (b) AL-RF (c) AL-NB

(d) AL-XGBoost (e) AL-KNN (f) AL-DT

FIGURE 3. Confusion matrix for active learning models on secondary dataset.

and XGB exhibit similar and lower performance levels. The stability of metrics across iterations suggests consistent model performance rather than significant variations from one iteration to another. Figure 3 depicts the confusion metrics for the 2nd dataset. It can be seen that the majority of the classifiers are capable of classifying both classifiers.

V. CONCLUSION AND FUTURE WORK

This study has made significant strides in predicting student anxiety through active learning and machine learning, utilizing various models. The research has provided valuable insights into the performance and reliability of various machine learning models in this context. Results demonstrate that active learning-based LR yielded a score of 0.61, and RF performed well with an average accuracy of 0.60 on the first dataset. Similarly, for the second dataset, RF is the most effective model, achieving an accuracy of 0.83. These results provide valuable insights into the models' performance across key metrics. Further, this research highlights the potential of employing machine learning techniques and active learning methodologies to predict and manage student anxiety.

Furthermore, this research enhances the educational experience by integrating active learning techniques to predict student anxiety. Most importantly, one key contribution lies in developing personalized intervention strategies. By predicting student anxiety, the model enables educators to tailor interventions, such as counseling services and stress management workshops, to meet individual student needs effectively. Another crucial contribution is the early identification of students at risk of heightened anxiety, facilitated by the model's advanced predictive capabilities. Likewise, this allows educators to implement timely and proactive support measures, maintaining students' overall well-being and academic performance. Moreover, the research contributes to promoting mental health awareness within educational settings.

While integrating active learning techniques to predict student anxiety holds promise for educational settings, certain limitations and areas for future work should be considered. Firstly, the model's effectiveness may depend on the diversity of educational contexts and student populations in the training data. Addressing this limitation requires a more comprehensive dataset that spans various educational environments to ensure the generalizability of the model's predictions across different settings. Moreover, translating predictive insights into actionable interventions within the dynamic landscape of educational institutions presents a significant challenge. In addition, biasness in predictive models for student anxiety in educational settings is crucial due to its potential impact on students' emotional well-being. A biased model may produce inaccurate predictions, leading to misidentification of at-risk students or mislabeling those not needing support. Specifically, it affects the individual student and has broader implications for the overall trust in the system. Students may feel misunderstood, stigmatized, or unfairly targeted, potentially exacerbating existing mental health concerns. Moreover, misallocating resources based on inaccurate predictions may strain the educational system's capacity to provide effective support where it is genuinely needed.

Future research should explore the practical implementation of the model's findings, considering the diverse resources and support systems available in various educational settings. Additionally, while valuable, the model's interpretability may require further refinement to align with educators' and students' needs and understanding. Enhancements in interpretability mechanisms could facilitate more effective communication of the model's predictions, promoting informed decision-making in educational contexts. Furthermore, ethical considerations and privacy concerns must be thoroughly addressed, considering the sensitivity of mental health issues in educational settings. Future work should focus on developing ethical frameworks that guide the deployment of predictive models, ensuring the confidentiality and well-being of students. Additionally, exploring the broader spectrum of mental health factors beyond anxiety within the educational context is crucial. Future research could extend the scope to include factors like stress, depression, and overall well-being, providing a more holistic understanding of students' mental health needs. This study discusses predicting student anxiety through active learning and machine learning. Future research efforts should focus on mitigating limitations, refining models, and exploring innovative data sources and methodologies to advance the understanding and application of predictive models in this domain.

REFERENCES

[1] J. Southworth, K. Migliaccio, J. Glover, J. Glover, D. Reed, C. McCarty, J. Brendemuhl, and A. Thomas, "Developing a model for AI across the curriculum: Transforming the higher education landscape via innovation in AI literacy," *Comput. Educ., Artif. Intell.*, vol. 4, Aug. 2023, Art. no. 100127.

[2] X.-Q. Liu, Y.-X. Guo, and Y. Xu, "Risk factors and digital interventions for anxiety disorders in college students: Stakeholder perspectives," *World J. Clin. Cases*, vol. 11, no. 7, pp. 1442–1457, Mar. 2023.

[3] M. K. Khaira, R. L. R. Gopal, S. M. Saini, and Z. M. Isa, "Interventional strategies to reduce test anxiety among nursing students: A systematic review," *Int. J. Environ. Res. Public Health*, vol. 20, no. 2, p. 1233, Jan. 2023.

[4] K. Meeks, A. S. Peak, and A. Dreihaus, "Depression, anxiety, and stress among students, faculty, and staff," *J. Amer. College Health*, vol. 71, no. 2, pp. 348–354, Feb. 2023.

[5] D. H. Halat, A. Soltani, R. Dalli, L. Alsarraj, and A. Malki, "Understanding and fostering mental health and well-being among University Faculty: A narrative review," *J. Clin. Med.*, vol. 12, no. 13, p. 4425, Jun. 2023.

[6] R. D. Godsil and L. S. Richardson, "Racial anxiety," *Iowa Law Rev.*, vol. 102, p. 2235, Sep. 2016.

[7] J. Santhosh, D. Dzsotjan, and S. Ishimaru, "Multimodal assessment of interest levels in reading: Integrating eye-tracking and physiological sensing," *IEEE Access*, vol. 11, pp. 93994–94008, 2023.

[8] R. Qasrawi, S. P. Vicuna Polo, D. A. Al-Halawa, S. Hallaq, and Z. Abdeen, "Assessment and prediction of depression and anxiety risk factors in schoolchildren: Machine learning techniques performance analysis," *JMIR Formative Res.*, vol. 6, no. 8, Aug. 2022, Art. no. e32736.

[9] P. J. Bota, C. Wang, A. L. N. Fred, and H. P. Da Silva, "A review, current challenges, and future possibilities on emotion recognition using machine learning and physiological signals," *IEEE Access*, vol. 7, pp. 140990–141020, 2019.

[10] P. Meshram and R. K. Rambola, "Diagnosis of depression level using multimodal approaches using deep learning techniques with multiple selective features," *Expert Syst.*, vol. 40, no. 4, p. e12933, May 2023.

[11] E. Mosqueira-Rey, D. Alonso-Ríos, and A. Baamonde-Lozano, "Integrating iterative machine teaching and active learning into the machine learning loop," *Proc. Comput. Sci.*, vol. 192, pp. 553–562, Jan. 2021.

[12] J. C. Cassady, E. E. Pierson, and J. M. Starling, "Predicting student depression with measures of general and academic anxieties," *Frontiers Educ.*, vol. 4, p. 11, Feb. 2019.

[13] H. S. Yi and W. Na, "How are maths-anxious students identified and what are the key predictors of maths anxiety? Insights gained from Pisa results for Korean adolescents," *Asia Pacific J. Educ.*, vol. 40, no. 2, pp. 247–262, Apr. 2020.

[14] J. A. Strangfeld, "Promoting active learning: Student-led data gathering in undergraduate statistics," *Teaching Sociol.*, vol. 41, no. 2, pp. 199–206, Apr. 2013.

[15] J. Chandler and D. Shapiro, "Conducting clinical research using crowdsourced convenience samples," *Annu. Rev. Clin. Psychol.*, vol. 12, no. 1, pp. 53–81, Mar. 2016.

[16] M. Kachaturoff, M. Caboral-Stevens, M. Gee, and V. M. Lan, "Effects of peer-mentoring on stress and anxiety levels of undergraduate nursing students: An integrative review," *J. Prof. Nursing*, vol. 36, no. 4, pp. 223–228, Jul. 2020.

[17] A. E. Layne, G. A. Bernstein, and J. S. March, "Teacher awareness of anxiety symptoms in children," *Child Psychiatry Human Develop.*, vol. 36, no. 4, pp. 383–392, Jun. 2006.

[18] S. Rodger, H. G. Murray, and A. L. Cummings, "Effects of teacher clarity and Student anxiety on Student outcomes," *Teaching Higher Educ.*, vol. 12, no. 1, pp. 91–104, Feb. 2007.

[19] T. Inada, "The causes of anxiety in highly anxious students in Japanese communicative EFL classes," *English Lang. Teach.*, vol. 15, no. 2, pp. 1–8, 2022.

[20] X. Zhu, J. Lafferty, and Z. Ghahramani, "Combining active learning and semi-supervised learning using Gaussian fields and harmonic functions," in *Proc. Workshop Continuum Labeled Unlabeled Data Mach. Learn. Data Mining*, vol. 3, 2003.

[21] S. C. DeFreitas, T. Crone, M. DeLeon, and A. Ajayi, "Perceived and personal mental health stigma in Latino and African American college students," *Frontiers Public Health*, vol. 6, p. 49, Feb. 2018.

[22] B. H. Ximenes and G. L. Ramalho, "Concrete ethical guidelines and best practices in machine learning development," in *Proc. IEEE Int. Symp. Technol. Soc. (ISTAS)*, Oct. 2021, pp. 1–8.

[23] N. D. Jones, "Socially sustainable development: Capturing black voices for creating an appropriate livability index for the black American community," Ph.D. dissertation, Univ. North Carolina Greensboro, Greensboro, NC, USA, 2021.

[24] S. G. Hofmann, M. A. A. Asnaani, and D. E. Hinton, "Cultural aspects in social anxiety and social anxiety disorder," *Depression Anxiety*, vol. 27, no. 12, pp. 1117–1127, Dec. 2010.

[25] S. Mutalib, N. S. M. Shafiee, and S. Abdul-Rahman, "Mental health prediction models using machine learning in higher education institution," *Turkish J. Comput. Math. Educ.*, vol. 12, no. 5, pp. 1782–1792, Apr. 2021.

[26] J. C. F. de Winter, S. D. Gosling, and J. Potter, "Comparing the Pearson and Spearman correlation coefficients across distributions and sample sizes: A tutorial using simulations and empirical data," *Psychol. Methods*, vol. 21, no. 3, pp. 273–290, Sep. 2016.

[27] X. Yang, L. Hou, Y. Zhou, W. Wang, and J. Yan, "Dense label encoding for boundary discontinuity free rotation detection," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2021, pp. 15814–15824.

[28] R. Yacouby and D. Axman, "Probabilistic extension of precision, recall, and F1 score for more thorough evaluation of classification models," in *Proc. 1st Workshop Eval. Comparison NLP Syst.*, 2020, pp. 79–91.

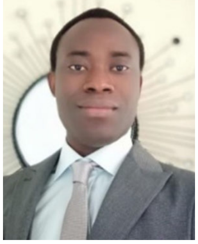
[29] T. Fawcett, "An introduction to ROC analysis," *Pattern Recognit. Lett.*, vol. 27, no. 8, pp. 861–874, Jun. 2006.

AHMAD ALMADHOR received the B.S.E. degree in computer science from Aljouf University (formerly Aljouf College), Aljouf, Saudi Arabia, in 2005, the M.E. degree in computer science and engineering from the University of South Carolina, Columbia, SC, USA, in 2010, and the Ph.D. degree in electrical and computer engineering from the University of Denver, Denver, CO, USA, in 2019. From 2006 to 2008, he was a Teaching Assistant and the College of Sciences Manager with Aljouf University, where he was a Lecturer, from 2011 to 2012. Then, he became a Senior Graduate Assistant and a Tutor Advisor with the University of Denver, from 2013 to 2019. He is currently an Assistant Professor of CEN and VD with the Computer and Information Science College, Jouf University, Saudi Arabia. His research interests include AI, blockchain, networks, smart and microgrid cyber security, integration, image processing, video surveillance systems, PV, EV, machines, and deep learning. His awards and honors include Aljouf University Scholarship (Royal Embassy of Saudi Arabia in D.C.), Aljouf's Governor Award for Excellency, and others.

SIDRA ABBAS (Graduate Student Member, IEEE) received the B.S. degree from the Department of Computer Science, COMSATS University Islamabad, Pakistan. Her research interests include computer forensics, machine learning, criminal profiling, software watermarking, intelligent systems, and data privacy protection.

GABRIEL AVELINO SAMPEDRO (Member, IEEE) is currently with the Faculty of Information and Communication Studies, University of the Philippines Open University, Laguna, Calabarzon, PH, USA. His current research interests include the Internet of Things and artificial intelligence.

SHTWAI ALSUBAI, photograph and biography not available at the time of publication.



STEPHEN OJO (Member, IEEE) received the B.Sc. degree (Hons.) in electrical and electronics engineering from the Federal University of Technology Akure, Nigeria, in 2014, and the M.Sc. and Ph.D. degrees in information systems from Girne American University, Cyprus, in 2017 and May 2021, respectively. He is currently an Assistant Professor with the Department of Electrical and Computer Engineering, College of Engineering, Anderson University, Anderson, SC, USA. Before Joining Anderson University, he was a Lecturer with Girne American University, where he taught courses in distributed computing, advanced programming, and electric circuits. He was a Research Scholar with Vodafone Telecommunication Company, Cyprus, where he developed a multiplicative-based model for signal propagation in wireless networks. He was awarded the Mobil and full Ph.D. scholarships throughout his undergraduate program. He is also a full-time Faculty Member with Anderson University, USA, where he teaches computer programming, electric circuits, machine learning, artificial intelligence in wireless mobile networks, and biomedical applications. He has authored and coauthored several peer-reviewed journals. His research interests include wireless networks, machine learning for wireless mobile networks, machine learning, and AI in biomedical devices.

ABDULLAH AL HEJAILI received the bachelor's degree in computer science from Tabuk Teachers College, Saudi Arabia, in 2007, and the master's degree in computer science from CLU, USA, in 2011. He is currently pursuing the Ph.D. degree with the Informatics School, University of Sussex. He is currently a Lecturer of computer science with the University of Tabuk. His research interests include technology-enhanced learning, image processing, virtual and augmented reality, motion capture, and education applications.

LUBOMIRA STRAZOVSKA is currently a Docent with the Faculty of Management, Department of Economics and Finance, Comenius University, Bratislava. He was with the Faculty of Philosophy, Comenius University, and the Faculty of Business, University of Economics in Bratislava. He has long been devoted to management and artificial intelligence in research and publishing activities. At the same time, he also conducts final theses on these topics. He was the author of the first monograph on family business in Slovakia. He is the author of many valuable scientific outputs on the mentioned topics. In publishing, he is the author and the coauthor of several university textbooks and texts and professional articles focused on the economic activity of small and medium-sized enterprises.