



INTERNATIONAL JOURNAL OF CREATIVE RESEARCH THOUGHTS (IJCRT)

An International Open Access, Peer-reviewed, Refereed Journal

Student Anxiety Prediction Using Adaptive Algorithms And Multi-Class Active Learning

Dr. MUVVA VENKATESWARA RAO 1, ANKANI USHA SRI 2,

GANJINABOINA GOPARAJU 3, DUNGALA PRANEETH 4,

#1Professor in Department of Information Technology, NRI Institute Of Technology, Agiripalli.

#2#3#4 B.Tech with Specialization of Information Technology in NRI Institute Of Technology, Agiripalli.

Abstract: This study explores the use of active learning and machine learning techniques to predict student anxiety, aiming to improve mental health management in educational settings. By leveraging active learning, the system identifies the most informative data points, enhancing the efficiency and accuracy of predictive models. Various machine learning algorithms, including K-Nearest Neighbours (KNN), Logistic Regression (LR), XGBoost (XGB), Naive Bayes (NB), and Random Forest (RF), are evaluated for their ability to detect anxiety patterns. The results indicate that Random Forest consistently performs well due to its robustness in handling complex data relationships. Active learning further refines predictions by iteratively selecting valuable data points, reducing the need for extensive labelled datasets. This research highlights the potential of data-driven approaches in educational psychology, contributing to the development of effective student mental health monitoring tools.

Index terms - Active Learning, Machine Learning, Predictive Modelling, Student Anxiety, Multi-Class Classification, Mental Health Prediction, Data-Driven Approaches, Student Behaviour Analysis, Anxiety Detection,

Adaptive Learning, Educational Psychology, Behavioural Data, K-Nearest Neighbours (KNN), Random Forest (RF), Logistic Regression (LR), Imbalanced Data Handling, Psychological Well-being, Automated Anxiety Prediction.

1. INTRODUCTION

Student anxiety is a significant concern in educational institutions, impacting academic performance, emotional well-being, and overall learning experiences. Traditional methods for detecting anxiety rely on self-reported surveys and manual assessments, which can be subjective and inconsistent. These approaches often fail to provide timely interventions, making it necessary to explore automated solutions that can predict anxiety more efficiently. Machine learning techniques offer a data-driven alternative, leveraging behavioural and contextual data to identify patterns associated with student anxiety.

However, accurately predicting anxiety presents challenges due to the complexity of mental health factors and the variability in student responses. To improve prediction efficiency, this study integrates active learning, a technique

that selectively chooses the most valuable data points for training models. By focusing on uncertain or informative samples, active learning enhances model performance while reducing the need for large labelled datasets.

This research evaluates multiple machine learning algorithms, including K-Nearest Neighbours (KNN), Logistic Regression (LR), XGBoost (XGB), Naive Bayes (NB), and Random Forest (RF), to determine their effectiveness in anxiety prediction. Among these, Random Forest demonstrates superior performance due to its ability to handle complex relationships in data. Additionally, active learning helps mitigate issues related to noisy and imbalanced datasets, ensuring more reliable predictions.

By integrating machine learning and active learning strategies, this study aims to develop an intelligent system capable of assisting educators and mental health professionals in monitoring student well-being. The findings highlight the potential of automated anxiety prediction in creating a more supportive and responsive educational environment.

2. LITERATURE SURVEY

2.1 Promoting Active Learning: Student-Led Data Gathering in Undergraduate Statistics

<https://www.researchgate.net/publication/27498596>

[6 Promoting Active Learning Student-Led Data Gathering in Undergraduate Statistics](#)

ABSTRACT: Scholarship on teaching undergraduates increasingly emphasizes the benefits of providing students with an active role in their education whereby instructors are more aptly described as facilitators of knowledge rather than merely providers of it. Additionally, recommendations from the American Sociological Association aimed specifically at the undergraduate sociology curriculum argue that students must engage as practitioners of sociology at each level of their program development. In short, undergraduates should do sociology-not just read or write sociology. Applying this recommendation to teaching statistics, I suggest organizing a course around a student-led research project in which

students generate the topics, questions, and data that are then used to complete a substantive research paper. Students are given the opportunity to be actively engaged in all stages of the research and data-gathering process. Students also present their work in student roundtables, further legitimating their important contributions as researchers.

2.2 Conducting Clinical Research Using Crowdsourced Convenience Samples:

<https://pubmed.ncbi.nlm.nih.gov/26772208/>

ABSTRACT: The advent of crowdsourcing has greatly increased the efficiency and scope of scientific inquiry. Amazon Mechanical Turk (MTurk) is a well-liked marketplace for crowdsourcing workers, and clinical scientists have reaped the benefits of its easily accessible research study participants and its simplified recruitment and payment processes. This function has been performed by MTurk for over five years. Researchers have conducted a great deal of methodological study on the platform because to its novelty and popularity, making it the most researched nonprobability sample currently accessible. With a focus on results applicable to clinical psychology research, this article provides a concise overview of what is known on the makeup of the MTurk sample and the quality of the data collected. The article continues by discussing the methodological concerns with MTurk, which are prevalent with other nonprobability samples but are new to researchers in clinical science. It then provides practical solutions to these problems or at least ways to lessen their influence.

2.3 Effects of peer-mentoring on stress and anxiety levels of undergraduate nursing students: An integrative review:

<https://pubmed.ncbi.nlm.nih.gov/32819548/>

ABSTRACT: This integrative review examines the effect of peer mentoring on stress and anxiety levels among undergraduate nursing students, guided by Bandura's Self-Efficacy Theory. Nursing students face high levels of stress and anxiety due to the competitiveness of their education.

While peer mentoring has been linked to increased retention, engagement, and competence, its impact on stress and anxiety remains underexplored. A literature search of PubMed, CINAHL, PsychINFO, and ERIC identified eight peer-reviewed studies published between 2000 and 2018. Findings suggest that peer mentoring reduces stress and short-term anxiety among nursing students, though the limited number of studies necessitates cautious interpretation.

2.4 Teacher awareness of anxiety symptoms in children:

<https://pubmed.ncbi.nlm.nih.gov/16736381/>

ABSTRACT: The present study aimed to determine which anxiety symptoms in children are associated with teacher awareness and whether teacher awareness differs according to student age and gender. The Multidimensional Anxiety Scale for Children (MASC) was completed by 453 second through fifth grade students and teachers nominated the three most anxious students in their classrooms. A multivariate analysis of variance was conducted with MASC scale scores as the dependent variables. Children identified by teachers as anxious had significantly higher levels of overall anxiety, physiological anxiety, social anxiety, and separation anxiety. Overall, teacher awareness did not differ based on student age or gender.

2.5 Effects of teacher clarity and student anxiety on student outcomes

<https://www.researchgate.net/publication/23346406>

[2 Effects of teacher clarity and student anxiety on student outcomes](#)

ABSTRACT: A laboratory experiment was carried out with 120 undergraduate students to examine a possible aptitude–treatment interaction between teacher clarity and student test anxiety in relation to two outcome measures, namely student achievement and student motivation, with student intelligence statistically controlled. Students completed measures of intelligence and test anxiety and were randomly assigned to high teacher clarity or low teacher clarity conditions, defined by the presence or absence of specific

teaching behaviours in a videotaped lecture with content held constant across conditions. Measures of motivation and self-efficacy for learning the material were completed immediately post-treatment, then one week later participants completed an achievement test based on the material contained in the lecture and assigned homework. Results revealed significant beneficial main effects for high vs. low teacher clarity for both achievement and motivation measures, but no aptitude–treatment interaction between teacher clarity and student test anxiety

3. METHODOLOGY

i) Proposed Work:

The proposed system leverages active learning and machine learning techniques to enhance the prediction of student anxiety in educational settings. By integrating active learning, the system prioritizes the most uncertain and informative data points for labelling, improving model efficiency with minimal labelled data. Various machine learning models, including K-Nearest Neighbours (KNN), Logistic Regression (LR), XGBoost (XGB), Naive Bayes (NB), and Random Forest (RF), are employed to analyze student behaviours and identify anxiety patterns. Among these, Random Forest consistently demonstrates superior performance due to its robustness in handling complex data relationships. Active learning further refines predictions by addressing challenges such as data imbalance and noise, ensuring improved accuracy. This adaptive approach enables educators and mental health professionals to monitor student well-being effectively, providing early interventions to create a supportive learning environment.

ii) System Architecture:

The proposed system architecture consists of interconnected modules that facilitate the prediction of student anxiety using active learning and machine learning techniques. The data collection module gathers student behavioural, academic, and psychological data from various sources, which is then processed in the preprocessing module to clean, normalize, and handle missing values while addressing data imbalances. The feature selection and

extraction module identifies the most relevant factors influencing anxiety levels. The active learning module optimizes model training by selecting the most uncertain and informative data points for labelling, improving efficiency. Multiple machine learning models, including KNN, LR, XGB, NB, and RF, are employed to analyze patterns and predict anxiety levels. The prediction and classification module categorizes students based on identified anxiety patterns, with results visualized through the reporting module for educators and mental health professionals. Additionally, a feedback and model refinement mechanism continuously updates the system with new data, ensuring improved accuracy and adaptability. This structured approach enhances predictive capabilities, enabling timely interventions for student well-being.

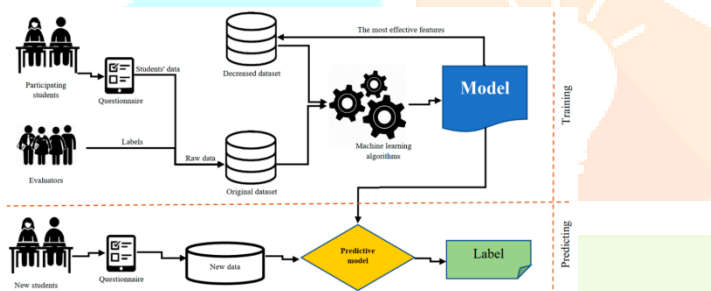


Fig 1 Proposed architecture

iii) Modules:

a) Data Collection Module

This module gathers data from various sources, including student surveys, academic records, digital interactions, and psychological assessments. The collected data includes behavioural patterns, academic performance, social interactions, and self-reported anxiety levels. The accuracy and quality of this data play a crucial role in effective prediction.

b) Data Preprocessing Module

Raw data often contains missing values, inconsistencies, and noise, which can affect model performance. This module handles missing data,

removes duplicates, normalizes numerical values, and balances the dataset to ensure equal representation of different anxiety levels. Techniques such as data scaling, outlier detection, and handling categorical variables are applied in this stage.

c) Feature Selection and Extraction Module

Not all collected data points contribute equally to anxiety prediction. This module identifies the most relevant features that have a significant impact on student anxiety. It applies statistical methods, correlation analysis, and feature importance techniques to retain only the most informative attributes, reducing computational complexity and improving model performance.

d) Active Learning Module

Active learning optimizes the efficiency of machine learning models by prioritizing the most uncertain or informative data points for labelling. Instead of labelling an entire dataset, this module focuses on critical instances that improve prediction accuracy. This is particularly beneficial when dealing with imbalanced datasets or when labelled data is limited.

e) Machine Learning Model Training Module

This module implements and trains various machine learning models, including K-Nearest Neighbours (KNN), Logistic Regression (LR), XGBoost (XGB), Naive Bayes (NB), and Random Forest (RF). Each model is evaluated based on performance metrics such as accuracy, precision, recall, and F1-score. The best-performing model is selected for anxiety prediction.

f) Prediction and Classification Module

Once trained, the models classify students into different anxiety levels based on detected patterns. The classification results help in identifying students who may require intervention. The system ensures that predictions are interpretable and can be effectively used by educators and mental health professionals.

g) Visualization and Reporting Module

This module presents prediction results in a user-friendly format, including graphical visualizations, dashboards, and detailed reports. The goal is to provide insights that educators and psychologists can use to track student well-being and take necessary actions.

h) Feedback and Model Refinement Module

The system continuously improves through feedback loops by incorporating new data and retraining the models. This module ensures adaptability, allowing the system to adjust to changing student behaviours and evolving mental health trends. Regular updates help maintain high accuracy and reliability in anxiety prediction.

iv) Algorithms:

a) K-Nearest Neighbours (KNN)

- A distance-based algorithm that classifies students' anxiety levels by comparing them with the closest data points (neighbors).
- Suitable for small datasets but may struggle with high-dimensional data.

b) Logistic Regression (LR)

- A statistical method used for binary and multi-class classification, predicting anxiety levels based on student behaviour patterns.

- Works well with linearly separable data but may not capture complex relationships.

c) XGBoost (XGB)

- An optimized gradient boosting algorithm that enhances prediction accuracy by iteratively refining decision trees.
- Handles noisy and imbalanced data effectively but requires careful tuning.

d) Naive Bayes (NB)

- A probabilistic classifier based on Bayes' theorem, assuming feature independence to predict anxiety levels.
- Fast and efficient but may oversimplify relationships in complex datasets.

e) Random Forest (RF)

- A robust ensemble learning method that constructs multiple decision trees and combines their predictions.
- Offers high accuracy, effectively handles complex data, and minimizes overfitting.

4. EXPERIMENTAL RESULTS

Naive Bayes Classifier:
Training Accuracy: 0.795134443821767
Testing Accuracy: 0.7895716945996276

Classification Report (Naive Bayes):

	precision	recall	f1-score	support
0	0.69	0.75	0.72	383
1	0.78	0.59	0.67	344
2	0.88	0.86	0.87	618
3	0.79	0.90	0.84	244
4	0.77	0.81	0.79	559
accuracy			0.79	2148
macro avg	0.78	0.78	0.78	2148
weighted avg	0.79	0.79	0.79	2148

Fig 2. Accuracy table

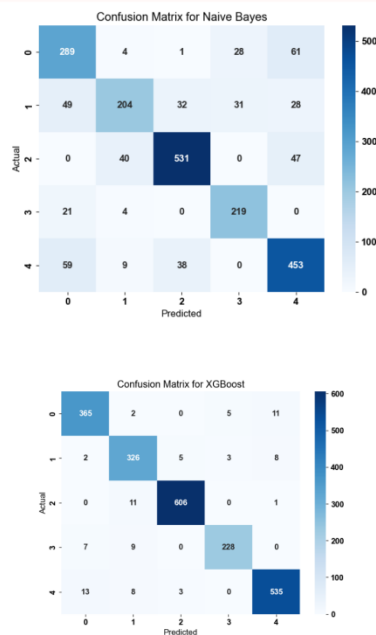


Fig 3. Confusion matrix

XGBoost Classifier:
Training Accuracy: 1.0
Testing Accuracy: 0.9598316573556797

Classification Report (XGBoost):

	precision	recall	f1-score	support
0	0.94	0.95	0.95	383
1	0.92	0.95	0.93	344
2	0.99	0.98	0.98	618
3	0.97	0.93	0.95	244
4	0.96	0.96	0.96	559
accuracy			0.96	2148
macro avg	0.96	0.95	0.95	2148
weighted avg	0.96	0.96	0.96	2148

Fig 4 accuracy graph

5. CONCLUSION

This research demonstrates the effectiveness of active learning and machine learning techniques in predicting student anxiety, enabling early intervention in educational settings. By leveraging models like Random Forest, KNN, XGBoost, Naive Bayes, and Logistic Regression, the system efficiently analyzes behavioural data to identify anxiety levels. The integration of active learning enhances predictive accuracy while reducing the need for extensive labelled data. This approach provides a data-driven solution to support educators and mental health professionals in monitoring student well-being, fostering a healthier learning environment.

6. FUTURE SCOPE

Future work can focus on integrating deep learning models for improved accuracy and handling complex behavioural patterns. Real-time anxiety monitoring using wearable devices and IoT-based sensors can enhance data collection. Additionally, incorporating personalized interventions and chatbot-based counselling systems can provide immediate support to students. Expanding the system to diverse educational settings and cross-cultural datasets will further improve its adaptability and effectiveness.

REFERENCES

- [1] J. A. Strangfeld, "Promoting active learning: Student-led data gathering in undergraduate statistics," *Teaching Sociol.*, vol. 41, no. 2, pp. 199–206, Apr. 2013.
- [2] J. Chandler and D. Shapiro, "Conducting clinical research using crowd sourced convenience samples," *Annu. Rev. Clin. Psychol.*, vol. 12, no. 1, pp. 53–81, Mar. 2016.
- [3] 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.
- [4] 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.
- [5] 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.
- [6] 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.
- [7] X. Zhu, J. Lafferty, and Z. Ghahramani, "Combining active learning and semi-supervised learning using Gaussian fields and harmonic functions," in *Proc.*

WorkshopContinuumLabeledUnlabeledDataMach.Learn.Data Mining, vol. 3, 2003.

[8] 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.

[9] 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.

[10] 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.

[11] 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.

Author's Profiles



Dr. MUVVA VENKATESWARA RAO is presently working as a Professor in the Department of Information Technology at NRI

Institute of Technology, Vijayawada, specializing in Artificial Intelligence and Machine Learning. With over 20 years of teaching experience, he has been instrumental in shaping the careers of numerous students in AI and ML. He holds a strong academic background and has contributed significantly to the field through his expertise and mentorship. His dedication to research and teaching has made a lasting impact in academia, inspiring students to excel in the ever-evolving field of technology.



ANKANI USHA SRI is a B. Tech student specializing in Information Technology at NRI Institute of Technology. She has gained valuable experience through multiple internships, including a Data Science role at BIST Technologies, a Python Programming internship at CodSoft, and a DevOps internship at Advaita Global. During her time at Advaita Global, she worked hands-on with cloud infrastructure and automation, creating custom Virtual Private Clouds (VPCs) and configuring subnets for efficient network management. She has also earned certifications in Java Full Stack, Python (HackerRank), and Cloud Computing (NPTEL), showcasing her

commitment to continuous learning and technical expertise.



GANJINABOINA GOPARAJU is a B. Tech student specializing in Information Technology at NRI Institute of Technology. He has earned certifications in Java Full Stack from Wipro and Cloud Computing from NPTEL. Through internships, including DevOps at Advaita Global - IT Labs Pvt Ltd and Data Science at BIST Technology Pvt Ltd, he has gained hands-on experience and practical knowledge. He is passionate about programming and has proficiency in Python and Java. With a strong interest in AI tools and their applications across various domains, he enjoys learning new technologies, building scalable applications, and solving real-world problems through technology.



DUNGALA PRANEETH is a B. Tech student specializing in Information Technology at NRI Institute of Technology. He has earned a certification in Java Full Stack Programming from Wipro and has hands-on experience with DevOps tools, particularly Jenkins and AWS. During his internship at Advaita - IT Labs Pvt. Ltd, he worked on automating CI/CD processes, gaining valuable practical knowledge. He also explored Data Science during his internship at BIST Technology Pvt Ltd. With a strong interest in programming languages like Java and Python, he is passionate about building scalable applications and solving real-world problems through technology.