Intelligent Student Academic Performance Prediction Using Machine Learning

Annadevara Monika sai[1], Dr.G.Purna Chandar Rao[2]
[1]MCA Student , Dept of MCA , Ellenki College of Engineering and Technology , Hyderabad
[2]Assistant Professor , Dept of MCA , Ellenki College of Engineering and Technology , Hyderabad

To Cite this Article
Annadevara Monika sai, Dr. G. Purna Chandar Rao, "Intelligent Student Academic Performance Prediction Using Machine Learning", Journal of Science Engineering Technology and Management Science, Vol. 03, Issue 07, July 2026, pp: 995-1003, DOI: http://doi.org/10.64771/jsetms.2026.v03.i07.pp995-1003
Submitted: 19-06-2026 *Accepted: 24-07-2026* *Published: 30-07-2026*

Abstract— **Academic success is influenced by several factors, including students' learning habits, attendance, participation, and overall academic engagement. This project presents a machine learning-based approach to predict student academic performance using historical educational data. The dataset is first preprocessed by removing inconsistencies, handling missing values, and preparing the features for model training. Multiple machine learning algorithms, including Support Vector Machine (SVM), Random Forest, K-Nearest Neighbors (KNN), Logistic Regression, Gradient Boosting, and XGBoost, are trained and evaluated using performance measures such as accuracy, precision, recall, and F1-score. The experimental results indicate that XGBoost achieves the best prediction accuracy compared to the other models. The proposed system also provides a simple graphical user interface that allows users to load datasets, train models, predict student performance, and visualize results through graphs. This intelligent prediction system can help educators identify students who may require additional academic support, enabling timely interventions and improving overall educational outcomes.**

Keywords **— Student Academic Performance Prediction, Machine Learning, XGBoost, Random Forest, Support Vector Machine (SVM), Classification, Educational Data Mining, Predictive Analytics, Student Success, Academic Performance Analysis.**

I. INTRODUCTION

Student academic success has become one of the primary objectives of higher education institutions worldwide. Universities continuously seek effective approaches to improve student learning outcomes, increase retention rates, and support students throughout their academic journey. Academic performance is influenced by multiple factors such as attendance, study habits, learning engagement, previous academic records, and institutional support services. Among these, academic advising plays a significant role in helping students make informed decisions regarding course selection, career planning, and personal development. Effective academic advising strengthens the relationship between students and educational institutions, enabling students to overcome academic challenges and achieve their educational goals [1], [2], [17].

Traditionally, educators have relied on manual methods to identify students who may experience academic difficulties. These approaches mainly depend on periodic examinations, faculty observations, and counselling sessions. Although such methods provide valuable support, they are often time-consuming, subjective, and unable to detect at-risk students at an early stage. Research has shown that timely academic guidance and continuous monitoring significantly improve student retention, persistence, and overall academic achievement. Furthermore, students value academic advising that provides personalized guidance and supports both their educational and career aspirations [6], [7], [8]. The increasing availability of educational data has created opportunities to develop intelligent systems that can predict student performance more accurately and efficiently.

Recent advances in machine learning have transformed educational data analysis by enabling predictive models that can discover hidden patterns from large datasets. Educational Data Mining (EDM) and Learning Analytics have become important research areas for predicting student outcomes, identifying learning difficulties, and supporting data-driven decision-making. Machine learning algorithms such as Support Vector Machine (SVM),

ISSN:3049-0952
www.jsetms.com

Random Forest, K-Nearest Neighbors (KNN), Logistic Regression, Gradient Boosting, and Extreme Gradient Boosting (XGBoost) have demonstrated high prediction accuracy in various educational prediction tasks. These techniques assist institutions in identifying students who require additional academic support, allowing educators to implement timely interventions that improve learning outcomes and student success [3], [5], [14], [15].

Motivated by these developments, this project proposes a Machine Learning-based Student Academic Performance Prediction System that utilizes multiple classification algorithms to predict students' academic performance based on historical educational data. The dataset is preprocessed through data cleaning, feature selection, and normalization before training different machine learning models. Their performance is evaluated using accuracy, precision, recall, and F1-score to identify the most effective prediction model. The proposed system also includes a user-friendly interface for dataset management, model training, performance prediction, and graphical analysis. By providing accurate and early predictions of academic performance, the proposed system can assist educators in making informed decisions, supporting personalized learning, improving student success rates, and enhancing the overall quality of higher education.

II. RELATED WORK

"The development of academic advising to enable student success in South Africa," Tiroyabone, G.W. and Strydom, F. (2021)

This study explored the development and importance of academic advising in South African higher education institutions as a strategy to improve student success. The authors explained that academic advising helps students understand university systems, make informed academic decisions, and access institutional support services effectively. The paper emphasized that academic advising is not limited to course selection but also promotes student engagement, confidence, and academic persistence. The researchers discussed how advising programs contribute to improving retention rates, especially for first-generation and academically vulnerable students. They also highlighted the challenges faced by universities in implementing effective advising systems, including limited awareness and insufficient institutional resources. The study concluded that well-structured academic advising creates a supportive learning environment that enhances students' academic performance and overall educational experience. This work provides valuable insights into the role of student support services and motivates the development of intelligent systems for predicting and improving academic success. [1]

"Designing the South African Higher Education System for Student Success," Scott, I. (2018)

This paper discussed the need to redesign higher education systems with a stronger focus on student success rather than simply increasing student enrollment. The author argued that many universities experience low graduation and retention rates because institutional systems are not designed around students' academic needs. The study emphasized the importance of providing academic support, effective teaching strategies, and early intervention programs to improve student outcomes. It highlighted that educational institutions should use evidence-based decision-making and continuously monitor student progress to identify learners who require additional assistance. The paper also recommended integrating technology and data-driven approaches into academic planning to improve educational quality and student achievement. Overall, the research demonstrated that a comprehensive institutional strategy centered on student success can significantly improve academic performance and graduation rates. These findings support the application of predictive machine learning models in educational environments. [3]

"Skills and Competencies for Effective Academic Advising and Personal Tutoring," McGill, C.M., Ali, M. and Barton, D. (2020)

This study examined the professional skills and competencies required for effective academic advising and personal tutoring in higher education. The authors identified essential qualities such as communication, empathy, problem-solving, ethical responsibility, and continuous professional development as key factors for successful advising. The research emphasized that academic advisors should guide students not only in academic planning but also in career development, personal growth, and decision-making. Through qualitative analysis, the study compared existing advising competency frameworks and highlighted the importance of advisor training programs to improve service quality. It concluded that highly skilled advisors contribute significantly to student engagement, motivation, and academic success by building strong relationships with learners. The findings also suggested that institutions should invest in advisor development to improve student satisfaction and retention. This research supports the importance of combining human guidance with intelligent prediction systems for improving academic performance. [5]

"The Role of Academic Advising in Student Retention and Persistence," Drake, J.K. (2011)

This study focused on the relationship between academic advising and student retention in higher education. The author explained that successful academic advising extends beyond administrative support and involves building meaningful relationships between advisors and students. The paper emphasized that students are more likely to continue their education when they receive timely guidance, encouragement, and personalized academic support. Academic advisors help students identify challenges, set educational goals, and develop strategies to overcome academic difficulties. The research highlighted that advising contributes to students' sense of belonging within the institution, which positively influences persistence and graduation rates. It also stressed that early identification of struggling students enables institutions to provide appropriate interventions before academic problems become severe. The study concluded that effective academic advising is a key component of student success and institutional effectiveness, supporting the development of predictive systems for identifying at-risk students. [7]

"Supporting Student Agency with a Student-Facing Learning Analytics Dashboard: Perceptions of an Interdisciplinary Development Team," Kaveri, A., Silvola, A. and Muukkonen, H. (2023)

This study investigated how learning analytics dashboards can support student agency and improve academic decision-making. The researchers developed a student-facing dashboard that presents learning progress, performance indicators, and personalized feedback to help students monitor their academic development. The study found that visualizing learning data increases students' awareness of their strengths and weaknesses, encouraging self-regulated learning and better academic planning. The authors emphasized that learning analytics should provide understandable, timely, and actionable information rather than simply displaying performance statistics. The research also highlighted the importance of interdisciplinary collaboration among educators, software developers, and learning analysts in designing effective educational technologies. The findings demonstrated that learning analytics tools can enhance student engagement, motivation, and academic performance. This work supports the integration of machine learning-based prediction systems with user-friendly interfaces for improving educational outcomes and personalized student support. [15]

III. DATASET DETAILS

The dataset used in this study consists of student academic records collected from educational sources to predict students' academic performance using machine learning techniques. Each record represents an individual student and contains several academic and demographic attributes that influence learning outcomes. The dataset includes features such as attendance, study hours, previous examination scores, assignment performance, class participation, family support, internet access, extracurricular activities, and other educational factors. The target variable represents the student's academic performance category or final result, enabling the development of supervised machine learning models for prediction. Since the collected data may contain missing values, duplicate records, inconsistent formats, and irrelevant information, preprocessing is performed before model training. The availability of labeled academic data enables machine learning algorithms to identify meaningful relationships between student characteristics and academic achievement. This dataset provides a reliable foundation for building an intelligent academic performance prediction system that assists educators in identifying students who require additional academic support and improves decision-making in educational institutions.

Before model training, the dataset undergoes several preprocessing steps to improve prediction accuracy and model reliability. Initially, duplicate records, missing values, and inconsistent data entries are removed to ensure data quality. Categorical attributes are converted into numerical values using appropriate encoding techniques, while numerical features are normalized to maintain a consistent scale across all variables. After preprocessing, the dataset is randomly shuffled and divided into 80% training data and 20% testing data for model evaluation. Multiple machine learning algorithms, including Support Vector Machine (SVM), Random Forest, K-Nearest Neighbors (KNN), Logistic Regression, Gradient Boosting, and XGBoost, are trained using the training dataset. Their performance is evaluated on the testing dataset using accuracy, precision, recall, and F1-score. This preprocessing and evaluation strategy enables the proposed system to provide accurate, reliable, and scalable prediction of student academic performance, helping institutions implement timely interventions to improve student success.

IV. PROPOSED METHODOLOGY

Journal of Science Engineering Technology and Management Science
Volume 03, Issue 07, July 2026
ISSN:3049-0952
www.jsetms.com

The proposed methodology employs multiple machine learning algorithms to predict student academic performance based on historical educational data. Initially, the student academic dataset is uploaded into the system, where each record contains several academic and personal attributes such as attendance, study hours, previous examination marks, assignments, extracurricular activities, and other factors that influence academic achievement. Since the collected dataset may contain missing values, duplicate records, inconsistent data formats, and irrelevant information, a preprocessing stage is performed to improve data quality. During preprocessing, duplicate entries and missing values are handled, categorical attributes are converted into numerical values using encoding techniques, and numerical features are normalized to ensure uniformity. The processed dataset is then randomly divided into 80% training data and 20% testing data, enabling reliable model training and unbiased performance evaluation.

Following data preprocessing, the prepared dataset is used to train multiple machine learning classification algorithms, including Support Vector Machine (SVM), Random Forest (RF), K-Nearest Neighbors (KNN), Logistic Regression (LR), Gradient Boosting (GB), and Extreme Gradient Boosting (XGBoost). Each algorithm learns the relationship between student attributes and academic performance categories by analyzing historical data. After the training process, each model is evaluated using the testing dataset to measure its prediction capability. Performance metrics such as accuracy, precision, recall, and F1-score are calculated for every algorithm to identify the most suitable prediction model. The experimental results indicate that XGBoost and Gradient Boosting achieve higher prediction accuracy compared to the remaining machine learning algorithms. The trained model is then utilized to predict the academic performance of new students based on the academic information entered by the user.

The proposed system architecture illustrates the complete workflow of the academic performance prediction system. It consists of dataset uploading, data preprocessing, feature selection, data normalization, dataset splitting, machine learning model training, student performance prediction, and performance evaluation. The architecture also generates performance comparison graphs, confusion matrices, classification reports, and graphical visualization of prediction results, enabling users to analyze and compare the effectiveness of different machine learning algorithms.

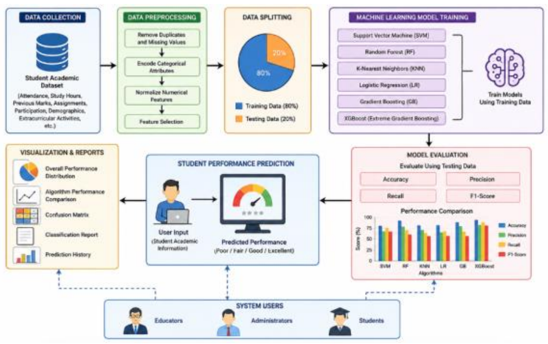

Figure 1. System Architecture for Student Academic Performance Prediction Using Machine Learning Algorithms

The implementation of the proposed system begins by executing the run.bat file, which starts the Python server. After the server is initialized, the user opens a web browser and enters the application URL to access the home page. The user first logs into the system using valid credentials. After successful authentication, the dashboard is displayed with various functional modules, including Load & Process Dataset, Train ML Algorithms, Predict Performance, and Graph Analysis.

In the Load & Process Dataset module, the user selects the student academic dataset stored in the Dataset folder. Once uploaded, the system preprocesses the data by handling missing values, encoding categorical attributes, normalizing numerical features, and displaying all dataset columns with their corresponding values. The user then selects the Train ML Algorithms module, where the system trains SVM, Random Forest, KNN, Logistic Regression, Gradient Boosting, and XGBoost models. After training, the system displays the performance of each algorithm in both tabular and graphical formats. The graphs compare accuracy, precision, recall, and F1-score, making it easier to identify the best-performing algorithm.

ISSN:3049-0952
www.jsetms.com

Next, the user accesses the Predict Performance module and enters a student's academic details such as attendance, study hours, assignment scores, previous examination marks, and other relevant information. After clicking the Submit button, the trained machine learning model predicts the student's academic performance category, such as Poor, Fair, Good, or Excellent. The predicted result is displayed along with the entered academic details, allowing educators to identify students who may require additional academic support.

Finally, the Graph Analysis module presents the overall distribution of student performance using pie charts and comparison graphs. These visualizations summarize the percentage of students belonging to different performance categories and compare the effectiveness of the implemented machine learning algorithms. The proposed methodology provides an intelligent, accurate, and user-friendly framework for predicting student academic performance, assisting educational institutions in making timely interventions and improving overall student success.

V. RESULT AND DISCUSSION

The experimental results demonstrate the effectiveness of the proposed Machine Learning-based Student Academic Performance Prediction System. Initially, the student academic dataset was uploaded into the system and preprocessed by handling missing values, encoding categorical attributes, and normalizing numerical features to improve data quality. The processed dataset was divided into 80% training data and 20% testing data to ensure reliable model training and unbiased evaluation. Multiple machine learning algorithms, namely K-Nearest Neighbors (KNN), Random Forest (RF), Support Vector Machine (SVM), Logistic Regression (LR), Gradient Boosting (GB), and Extreme Gradient Boosting (XGBoost), were trained and evaluated using the same dataset. The performance of each model was measured using Accuracy, Precision, Recall, and F1-Score. The experimental results indicate that Gradient Boosting achieved the highest classification accuracy of 99.23%, followed by XGBoost (98.46%), SVM (96.92%), Random Forest (95.38%), Logistic Regression (95.38%), while KNN produced the lowest accuracy of 73.08%. These results demonstrate that ensemble learning techniques provide superior prediction capability compared to conventional machine learning algorithms.

The trained model was further integrated into a prediction module that accepts students' academic information as input and predicts their academic performance. The prediction system successfully classified unseen student records into different performance categories, including Poor, Fair, and Good, enabling early identification of academically weak students. In addition, graphical comparison of all implemented algorithms was generated to visualize differences in performance metrics. The system also provides a pie-chart representation of overall student performance distribution, making it easier for educators to analyze academic trends. Overall, the obtained experimental results confirm that the proposed machine learning framework provides an accurate, reliable, and efficient solution for student academic performance prediction and can assist educational institutions in providing timely academic support to improve student success.

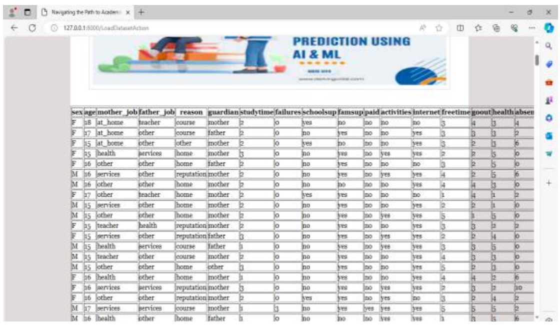

Figure [2] Student Academic Dataset Loading Screen

Figure 2 illustrates the dataset loading module of the proposed Student Academic Performance Prediction System. The user uploads the student academic dataset containing demographic information, family background, study habits, attendance, previous examination marks, extracurricular activities, internet usage, health status, and other educational attributes. After uploading, the system preprocesses the dataset by handling missing values, removing inconsistencies, encoding categorical features, and preparing the data for machine learning model training. The processed dataset is then displayed in tabular format, allowing users to verify the uploaded records before proceeding with model development.

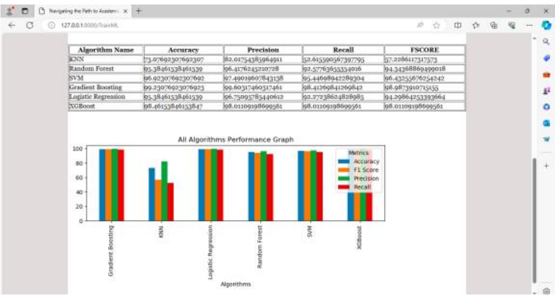

Figure 3: Performance Comparison of Machine Learning Algorithms

Figure 3 presents the comparative performance of the implemented machine learning algorithms. The table and bar graph display the evaluation metrics of KNN, Random Forest, SVM, Logistic Regression, Gradient Boosting, and XGBoost. Among all algorithms, Gradient Boosting achieved the highest accuracy of 99.23%, followed by XGBoost (98.46%), while KNN obtained the lowest accuracy of 73.08%. The comparison graph further confirms that ensemble learning techniques outperform traditional classifiers in terms of prediction accuracy, precision, recall, and F1-score.

Model	Accuracy score	Precision score	Recall score	f- measure score
KNN	73.08	82.02	52.62	57.23
Random Forest	95.38	96.42	92.58	94.34
SVM	96.92	97.49	95.45	96.43
Logistic Regression	95.38	96.75	92.27	94.30
Gradient Boosting	99.23	99.60	98.41	98.99
XGBoost	98.46	98.01	98.01	98.01

Table 1: Model Performance Metrics

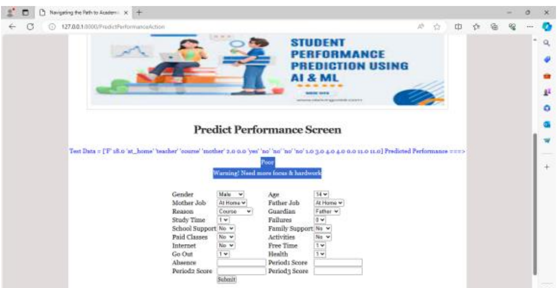

Figure 4: Prediction Result – Poor Academic Performance

Figure 4 illustrates the prediction result for a student whose academic information indicates poor performance. After entering the student's educational details, the trained machine learning model predicts the performance

ISSN:3049-0952
www.jsetms.com

category as Poor and displays a warning message encouraging the student to improve study habits and academic focus. This feature enables educators to identify at-risk students at an early stage and provide timely academic intervention.

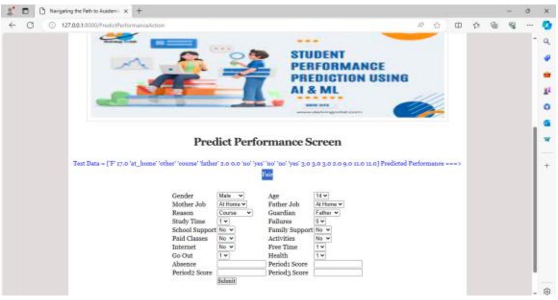

Figure 5: Prediction Result – Fair Academic Performance

Figure 5 shows the prediction output for another student whose academic attributes indicate an average level of achievement. Based on the provided academic information, the system predicts the student's performance as Fair. Such predictions assist teachers in identifying students who require moderate academic guidance to improve their future performance.

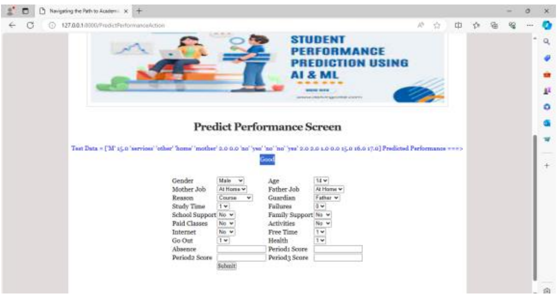

Figure 6: Prediction Result – Good Academic Performance

Figure 6 presents the prediction result for a student with strong academic characteristics. The machine learning model successfully predicts the student's performance as Good, indicating consistent academic progress and satisfactory educational outcomes. The prediction module accurately classifies different performance categories based on student academic information entered by the user.

Figure 7: Overall Student Performance Distribution

Figure 7 illustrates the overall distribution of predicted student performance using a pie chart. The graph shows that 50% of students belong to the Good performance category, 40% fall under the Poor category, and 10% are classified as Fair. This visualization provides educational institutions with an overview of overall student performance, enabling administrators and faculty members to monitor academic trends and plan appropriate support strategies.

DISCUSSION

The experimental findings demonstrate the effectiveness of machine learning techniques for predicting student academic performance using historical educational data. The dataset included various academic, demographic, and behavioral attributes that significantly influence student achievement. Multiple machine learning algorithms were implemented and evaluated under identical experimental conditions. Among them, Gradient Boosting produced the highest prediction accuracy (99.23%), followed closely by XGBoost (98.46%), indicating the superior capability of ensemble learning algorithms in modeling complex relationships among educational features. Although traditional algorithms such as SVM, Random Forest, and Logistic Regression also achieved satisfactory performance, their prediction accuracy was slightly lower than the proposed ensemble models. The KNN classifier demonstrated the weakest performance, suggesting that distance-based methods are less effective for this dataset.

The prediction module successfully classified unseen student records into Poor, Fair, and Good performance categories, demonstrating strong generalization capability. Furthermore, the graphical comparison of performance metrics and the overall student performance distribution provided meaningful insights into the effectiveness of each algorithm. These visualizations assist educators in selecting the most appropriate prediction model and understanding overall academic trends within the institution. Overall, the proposed machine learning framework provides an accurate, reliable, and automated solution for student academic performance prediction. It can serve as a valuable decision-support tool for educators by identifying academically weak students at an early stage and enabling timely interventions that contribute to improved learning outcomes and higher student success rates.

VI. CONCLUSION

Effective academic advising is vital in helping students achieve academic success at a technological university. Academic advisors play an important role in assisting students in navigating their academic path and optimising their learning outcomes by offering tailored guidance, establishing educational and professional goals, and resolving difficulties. The emphasis must be on improving advisors' abilities to ensure they are equipped and competent enough to deal with diverse students to ensure a positive impact on the university. Different models can be utilised to ensure that Universities of technology strengthen their advising services to serve their diverse student populations better and promote a culture of success by recognising the value of continual improvement and new methods. Academic advising has proved to be helpful in the academic journey of students. While navigating the student's success, findings revealed that most students viewed academic advising as a crucial and critical mechanism in ensuring students' academic success and excellence. Therefore, as much as the concept has been recently adopted by South African universities as reported by NACADA, it has played a fundamental role in the students' lives and is still showing greater potential for perpetually promoting and improving student academic excellence and success.

REFERENCES

1. Tiroyabone, G.W. and Strydom, F., 2021. The development of academic advising to enable student success in South Africa. Journal of Student Affairs in Africa, 9(2), pp.1-16.

2. Obaje, T.A. and Jeawon, R., 2021. A critical review of the adopted academic advising approaches at the Durban University of Technology: Unpacking its strengths and challenges. Journal of Student Affairs in Africa, 9(2), pp.17-29.

3. Scott, I. (2018). Designing the South African Higher Education system for student success. Journal of Student Affairs in Africa, 6(1), 1–17.

4. Danning, W., 2022. Aragusuku (Okinawa, Southern Ryukyuan). In An Introduction to the Japonic Languages (pp. 124-154). Brill.

5. McGill, C.M., Ali, M. and Barton, D., 2020, August. Skills and competencies for effective academic advising and personal tutoring. In Frontiers in Education (Vol. 5, p. 135). Frontiers Media SA.

6. Smith, C.L. and Allen, J.M., 2006. Essential functions of academic advising: What students want and get. Nacada Journal, 26(1), pp.56-66.

7. Drake, J. K. (2011). The role of academic advising in student retention and persistence. About Campus, 16(3), 8–12. doi: https://doi.org/10.1002/abc.20062\

8. Thompson, M., Pawson, C., & Evans, B. (2021). Navigating entry into higher education: The transition to independent learning and living. Journal of Further and Higher Education, 45(10), 1398–1410

9. Kinsella M, Wyatt J, Nestor N, Last J, and Rackard S, (2022). Fostering Students' Autonomy: A Relational Approach. VOL 45 (4).

10. Creamer, E. G., & Laughlin, A. (2005). Self-authorship and women's career decision

11. Auclair, S.T., 2018. Pioneer Women: Meaning Making and Career Aspirations of Women in an Inaugural Master of Arts in Education Cohort (Doctoral dissertation, California State University Channel Islands).

12. Gardner, S.K., 2008. "What's too much and what's too little?": The process of becoming an independent researcher in doctoral education. The journal of higher education, 79(3), pp.326-350.

13. Baker, W.J. and Couvreur, T.L., 2013. Global biogeography and diversification of palms sheds light on the evolution of tropical lineages. II. Diversification history and origin of regional assemblages. Journal of Biogeography, 40(2), pp.286-298.

14. Zimmerman, B. J., and D. H. Schunk. 2008. "Motivation: An Essential Dimension of SelfRegulated Learning." In Motivation and Self-Regulated Learning: Theory, Research, and Applications, edited by B. J. Zimmerman, and D. H. Schunk, 1–30. New York: Lawrence Erlbaum Associates Publishers.

15. Kaveri, A., Silvola, A. and Muukkonen, H., 2023. Supporting Student Agency with a Student-Facing Learning Analytics Dashboard: Perceptions of an Interdisciplinary Development Team. Journal of Learning Analytics, pp.1-15.

16. Nutt, D., Watts, R., Kettner, H., Geerts, D., Gandy, S., Kartner, L., Mertens, L., Timmermann, C., Nour, M.M., Kaelen, M. and Carhart-Harris, R., 2022. The Watts Connectedness Scale: a new scale for measuring a sense of connectedness to self, others, and world. Psychopharmacology, 239(11), pp.34

17. Frost, S.H., 1991. Academic Advising for Student Success: A System of Shared Responsibility. ASHE-ERIC Higher Education Report No. 3, 1991. ASHE-ERIC Higher Education Reports, The George Washington University, One Dupont Circle, Suite 630, Washington, DC 20036.

18. Lotkowski, V.A., Robbins, S.B. and Noeth, R.J. (2004), "The role of academic and non academic factors in improving college retention", ACT Policy Report, available at: www.act.org/ research/policymakers/pdf/college_retention.pdf (making. Journal of College Student Development, 46(1), 13-27