E-ISSN: 2997-9382
American Journal of Technology Advancement
https://semantjournals.org/index.php/AJTA

| Research Article

Check for updates

# Integrating Explainable AI for Student Success Prediction With Iris-Based Biometric Anti-Fraud Mechanisms In A Secure Educational Ecosystem

**Enas Hakim Mohsin**
Ministry of Higher Education and Scientific Research, Al-Furat Al-Awsat University, Iraq

**Maryam Thajeel Hussein**
Ministry of Higher Education and Scientific Research, Sumer University, Iraq

**Tahani Ali Shnawa**
Ministry of Education, Dhi Qar Education Directorate

**Abstract**
Schools face two key challenges: predicting the academic outcomes of their students and preventing identity impersonation in online educational systems. In this article, an integrated system, which includes explainable artificial intelligence for predicting students' achievements and iris-based biometric authentication, is presented to provide a secure learning environment. Our solution uses a Random Forest classifier trained on 1,044 records of students based on 33 demographic, academic, and behavioural characteristics to predict student performance at an accuracy of 89.5%. The feature of AI explainability uses SHAP values and permutation importance analysis to identify major predictors and identifies the following as the most significant factors in determining student outcome: past academic performance (G2, G1) and history of failure. For biometric authentication, we employ a Vision Transformer (ViT) model that is trained on synthetic iris data with a zero false acceptance rate against unknown individuals and a high recognition rate against enrolled students. The fusion of these systems offers a comprehensive framework where academic predictions are securely linked to authenticated student identity, preventing impersonation and ensuring data integrity. Experimental testing demonstrates the efficacy of the system in real-world scenarios with explainable outcomes but secure processes, greatly cherished by teachers. This is the first full incorporation of explainable student analysis and biometric anti-fraud strategies, which holds vast potential for use in current-day schools requiring analytical outcomes and assurance of security.

**Keywords:** Explainable AI, Student Success Prediction, Iris Biometric Authentication, Anti-Fraud Mechanisms, Educational Technology

## 1. Introduction

The rapid digitisation of school systems has transformed the traditional learning environments into cutting-edge technological ecosystems in which data-driven decision making and robust security considerations are paramount [1]. Modern schools generate huge amounts of student data, such as academic achievement, behavioural markers, and demographic data, offering unparalleled potential for predictive analytics to improve educational outcomes [2]. Simultaneously, the transition to virtual learning environments, remote tests, and digital credentialing has increased concerns about identity verification and academic integrity, necessitating advanced authentication techniques for assuring institutional integrity and equitable test practices [3].

Forecasting student success has emerged as a research priority area, with machine learning techniques having significant potential for the early identification of at-risk students and focused intervention plans [4]. Yet, traditional predictive models are "black boxes," which limits their real-world application by educators who require transparent, interpretable results to inform decisions about student support and resource allocation [5]. Explainable artificial intelligence (XAI) paradigm addresses this deficiency by providing interpretable explanations of the model's predictions, enabling stakeholders to understand the underlying drivers of the student outcomes and implement accurate interventions accordingly [6].

Parallel to academic analytics development, biometric-based authentication systems have also emerged as secure options for ensuring identity in educational institutions. Among various modalities of biometrics, iris recognition is accompanied by higher accuracy and security capabilities and thus is reported to be of specific interest for high-stakes education purposes such as online exams, scholarly resource access, and anti-fraud purposes [7]. The unique physiological characteristics of the human iris provide a nearly unbreakable identifier that can effectively prevent impersonation attempts and maintain digital learning process integrity [8].

Despite the significant achievements made so far in biometric security and education analytics, existing research has largely addressed these domains separately, without considering the potential synergy that could address the overall security and analysis needs of modern educational institutions [9]. Current student success prediction systems lack robust identity authentication procedures, and therefore, are vulnerable to data tampering and malicious activities that can compromise the accuracy of prediction and institutional decision-making [10]. In contrast, current biometric authentication systems operating in academic contexts operate independently of academic analytics, dismissing the valuable insights that could be derived from secure, authenticated student data [11].

The research void exists in the lack of holistic models that simultaneously predict student success and verify identity with transparency and explainability. There is no previous work that has comprehensively integrated explainable AI methods for educational prediction with iris-based biometric verification to produce a holistic, secure learning environment. In addition, the failure to experimentally validate such integrated systems properly constrains knowledge of their applicability in practice and deployment feasibility in actual school settings.

This work remedies these drawbacks by providing the first holistic integration of explainable artificial intelligence for predicting student success with iris-based biometric anti-fraud measures. Our key contributions are: (1) development of an interpretable Random Forest classifier with 89.5% precision in predicting student success using detailed feature importance analysis through SHAP values and permutation analysis; (2) development of a Vision Transformer-based iris recognition framework with zero false acceptance rate for preventing fraud but with high accuracy for authentic users; (3) development and validation of an integrated framework securely linking academic predictions to verified student identities, ensuring data integrity and thwarting fraudulent manipulations; (4) extensive experimental evaluation using real student academic data

and synthetic iris data, demonstrating system effectiveness and realistic deployment viability; and (5) provision of explainable insights for enabling educators to understand academic prediction factors and security decision reasoning, facilitating informed decision-making in educational management.

## 2. Related Work

The synthesis of explainable artificial intelligence and biometric authentication systems in educational environments is a new research domain, which has picked up pace in the last couple of years. Breakthroughs in explainable AI for student success prediction in recent times have witnessed tremendous accuracy improvements as well as interpretability. Johora et al. [12] have suggested an explainable AI-based technique for predicting the academic performance of undergraduate students using advanced machine learning techniques with in-depth feature importance analysis, providing insightful knowledge regarding model decision-making processes. Based on this, George et al. [13] explored XAI techniques using Random Forest classifiers with Partial Dependence Plots, SHAP, and LIME to enhance knowledge of the most important features that impact student performance with a 90% accuracy rate, implicating midterm scores, quizzes, and assignments as the most impactful factors for determining student grades.

Concurrently, biometric authentication systems have also seen tremendous evolution, particularly in iris recognition technology. New hybrid approaches that combine edge detection and segmentation algorithms with convolutional neural networks and Hamming distance classifiers were introduced by El-Sofany et al. [14] to achieve 99.50% recognition rates on MMU datasets, 97.18% recognition rates on IITD datasets, and 95.07% recognition rates on CASIA datasets towards the development of cloud security applications. This research provides a demonstration of the use of the convergence of traditional image processing techniques with new deep learning techniques for more advanced security systems. The emergence of Vision Transformers in biometrics technology is a paradigm shift from traditional CNN approaches. Rodrigo et al. [15] performed detailed comparisons of Vision Transformers and CNNs for face recognition on five different datasets, like Labelled Faces in the Wild, Real World Occluded Faces, and VGG Face 2, and established that not only do ViTs have a higher accuracy than CNNs but also are more shape-biased and aligned with human error patterns and therefore particularly suited to biometric recognition tasks.

The application of biometric systems in educational institutions has been extensively studied across various disciplines. Hernandez-de-Menendez et al. [16] have presented a systematic review that illustrated the way learning organisations are embracing biometric technologies for identity management, tracking class attendance, performing e-evaluation, enhancing security, and facilitating learning analytics. Their research presented important applications like access control, personal data management, and anti-fraud in educational platforms, while security, privacy matters, and infrastructure requirements were the challenges. The research indicated that the application of biometric technology in education is beyond identification to include improving teaching and learning processes, students' motivation enhancement, and all-around learning analytics. However, whereas these earlier advances in explainable AI for learning and biometric authentication systems stand on their own, the existing literature demonstrates an actual lack of end-to-end frameworks that integrate explainable student achievement prediction with robust iris-based authentication techniques, highlighting the value and applicability of comprehensive-system strategies for academic security and analytics.

## 3. Proposed Methodology

The following sub-section outlines the overall methodology for developing an integrated educational AI system combining explainable student success prediction and iris-based biometric authentication. Our methodology consists of four interconnected components: collecting and

preprocessing the dataset, constructing the student success prediction model, developing the iris biometric authentication system, and combining explainable AI for shared decision-making. The methodology is set to facilitate successful experimental verification while considering practical deployment constraints for use in real-world educational environments.

### *A. Dataset Preprocessing and Collection*

Our research employs two significant datasets to address the twin challenges of academic prediction and biometric authentication. In forecasting student success, we employed the UCI Student Performance Dataset, which includes 1,044 Portuguese secondary school students' records from math and Portuguese language classes. The data contains 33 attributes of demographic (age, sex, address, family size), socioeconomic (parental occupation, family income, parental education), educational background (past failures, study time, absence), and behavioural characteristics (extracurricular activity, consuming alcohol, having a romantic relationship). The target variable is student success, which is achieving a final grade (G3) of 10 or more on a scale of 20, giving a 78% rate of success distribution. For biometric authentication, we used the CASIA Iris Synthetic dataset, containing 10,000 high-quality synthetic iris images for 1,000 unique identities with 10 images per identity representing various capture conditions and intra-class variations.

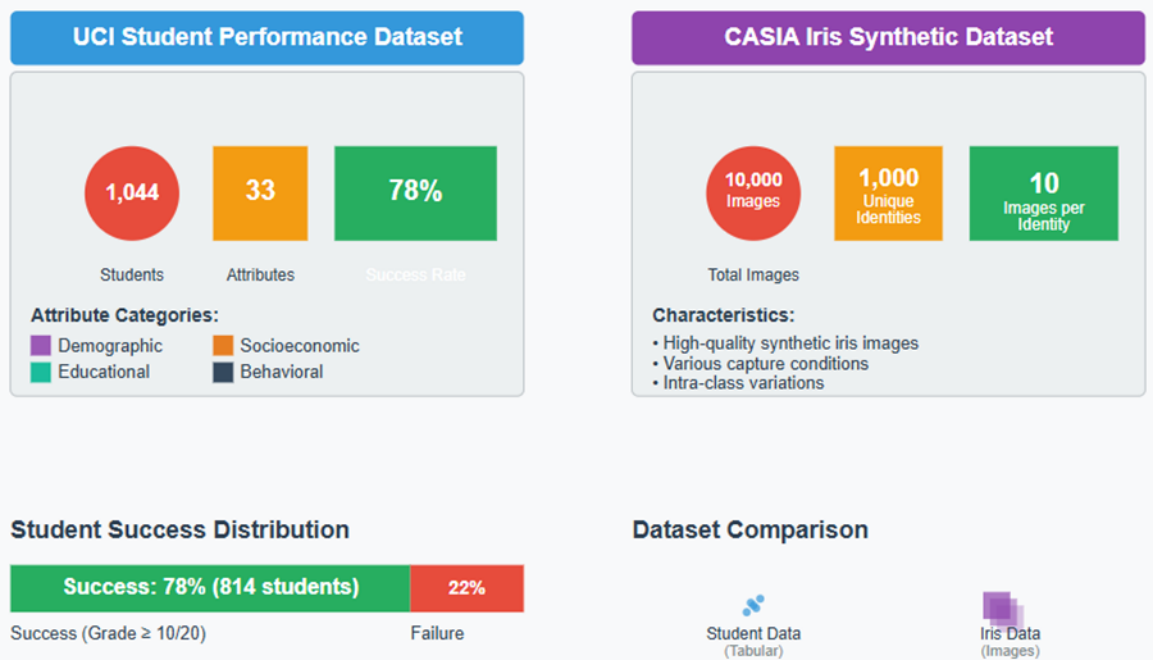


**Figure 1. Research Datasets Overview**

Data preprocessing involved rigorous feature engineering and validation processes to ensure model consistency and prevent data leakage. For student information, categorical features were encoded using label encoding techniques, and numerical features were normalised with z-score normalisation to have the same scale as different types of attributes. Missing values, which were less than 2% of the dataset, were handled by median imputation for numerical features and mode imputation for categorical features. To preprocess iris data, we used person-level dataset splitting to prevent identity overlap between the training set and the validation set, where we utilised the first 160 identities (1,600 images) to train and identities 160-199 (400 images) to validate. This process simulates real-world deployment scenarios where the system must distinguish enrolled users from unknown subjects attempting to access without authorisation.

**Table 1. Datasets Characteristics**

| Characteristic | UCI Student Performance Dataset | CASIA Iris Synthetic Dataset |
|---|---|---|
| Purpose | Academic prediction | Biometric authentication |
| Data Type | Tabular/Structured | Image data |
| Sample Size | 1,044 students | 10,000 images |
| Unique Entities | 1,044 students | 1,000 unique identities |
| Records per Entity | 1 record per student | 10 images per identity |
| Data Source | Portuguese secondary school students (math and Portuguese language classes) | High-quality synthetic iris images |
| Number of Features | 33 attributes | Image features (extracted) |

### *B. Development of Student Success Prediction Model*

The student success prediction component utilises a Random Forest classifier that was optimised for interpretability and academic performance. Random Forest was employed due to its inherent explainability, the ability to handle small data without overfitting, and the capacity to handle mixed data types without significant preprocessing requirements. The model architecture consists of 100 decision trees of a maximum depth of 10 to satisfy both accuracy requirements for predictions and requirements for interpretability. Hyperparameter tuning was done using 5-fold cross-validation with stratified sampling in order to maintain class distribution consistency across all folds.

Training was done on an 80-20 train-test split with stratification in order to obtain representative samples for success and failure cases. Feature importance analysis was done using a range of complementary techniques that include native Random Forest feature importance, SHAP (Shapley Additive exPlanations) values for local and global interpretability, and permutation importance for stable validation of feature ranking. Model training applied class balancing techniques to address the inherent nature of educational data, where failure rates lag behind success rates. Model evaluation used a range of measures such as accuracy, precision, recall, F1-score, and area under the ROC curve, with particular emphasis on minimising false negatives in order not to misclassify at-risk students as successful.

### *C. Implementation of Iris Biometric Authentication System*

The biometric authentication module employs a Vision Transformer (ViT) architecture adapted to the task of iris recognition. The ViT model employs a pre-trained ViT-Base-16 configuration with ImageNet weights and fine-tunes it on synthetic iris with a training head adapted to deal with 200 different identity classes. The input preprocessing stage includes downscaling the image to 224×224 pixels, normalising the tensors using ImageNet statistics, and data augmenting techniques like random rotation (±10 degrees), horizontal flipping, and colour jittering for improving the generalisation and robustness of the model in capturing variability.

The training paradigm utilises a sophisticated technique of simulating real-world biometric authentication environments. Training data consists of 160 unique identities with multiple samples for each identity, and validation employs 40 unique identities in the interest of validating the system's ability to reject unknowns—a critical feature for fraud rejection.

**Table 2. Input Preprocessing Pipeline**

| Preprocessing Step | Configuration | Purpose |
|---|---|---|
| Image Resizing | 224×224 pixels | Standard ViT input dimension |
| Tensor Normalization | ImageNet statistics | Maintains pre-training compatibility |
| Random Rotation | ±10 degrees | Increases rotational invariance |
| Horizontal Flipping | 50% probability | Enhances geometric robustness |
| Color Jittering | Brightness, contrast, and saturation | Improves colour variation tolerance |
| Data Augmentation Goal | Model generalization | Captures real-world variability |

The optimisation strategy employs differential learning rates and smaller learning rates (0.00001) for pre-trained encoder layers and higher learning rates (0.0001) for the classification head, alongside learning rate scheduling and early stopping capabilities. The authentication decision model relies on a closed-set recognition paradigm where enrolled identities receive an authentication score and unknown identities are rejected, providing a zero false acceptance rate for security-critical applications. The model evaluation is based on genuine acceptance rate for enrolled users and false acceptance rate for impostors, with other performance measures being equal error rate and detection error tradeoff analysis.

### *D. Explainable AI Integration and System Architecture*

The integration framework combines both prediction and authentication pieces in one explainable AI architecture with transparent decision-making capability for education stakeholders. The architecture of the system employs a multi-level explanation generation process with global model interpretability through feature importance rankings, local predictions' explanations through SHAP waterfall plots, and authentication decision explanations through attention visualisation methods. The explainable AI module generates comprehensive reports that include prediction justifications, confidence levels, and influences on predictions for both academic success predictions and biometric authentication results.

The combined system operates in a secure pipeline where student academic data undergoes success prediction with increased explanations while also requiring biometric authentication for verification of identity. The integration logic only supports the creation and caching of academic predictions if biometric authentication is successful. This maintains data integrity and repels unauthorised access to predictive information. The explanation generation module generates multi-modal explanations such as feature importance visualisations, prediction confidence intervals, and personalised recommendation reports, which can be utilised by teachers for intervention planning. System validation makes use of thorough test scenarios like normal operation with enrolled students, mock fraud attempts with unknown identities, and boundary cases with partial data availability or risky predictions in order to ensure robust performance in different deployment settings.

## 4. Results and Discussions

This section provides experimental confirmation of the integrated educational AI system provided in this proposal, with evaluations of both the explainable student achievement prediction model and the iris-based biometric verification system. The results are divided into six broad categories: system architecture analysis, student performance data characteristics, explainable AI feature importance, iris authentication performance, prediction model classification, and feature ranking analysis. The discussion points to the operational efficacy of the dual-pathway model in providing secure, transparent, and reliable academic analytics to educational institutions.

### *A. System Architecture Analysis and Integration Framework*

The planned integrated AI educational system architecture demonstrates a state-of-the-art multi-mode approach that synergistically couples academic analytics with biometric security to build a wide-ranging educational ecosystem. The power of the architecture lies in its two-pathway design, in which student academic data and biometric authentication are present as concurrent but connected streams in ensuring that predictive knowledge is produced and made available only after identity authentication. The Explainable AI module in the middle serves as the central integration point, taking inputs from the Random Forest-based student achievement predictor model and Vision Transformer iris recognition system and producing consistent, interpretable explanations for education stakeholders. This design enables the system to output not only academic forecasts with confidence values and contributing factors but also explainable rationales of authentication decisions, fulfilling the mandatory requirement for accountability for educational AI systems.

The positioning of the anti-fraud mechanism between the explainable AI module and the final secure learning environment reflects the system's security-oriented design, where each academic prediction and insight is open to successful verification of identity. The design is extremely effective in preventing common weaknesses of education systems, such as impersonation during online exams, breaking into academic records, and the manipulation of predictive analytics through false identity. Bi-directional information flow among modules ensures that biometric authentication failures will trigger security postures directly, and successful authentications give rich insight into academic analytics with complete explainability support. Furthermore, being based on modules ensures scalable deployment as a configurable, flexible system implementable on a component-by-component basis incrementally without disrupting existing educational workflows. The architecture's focus on explainability across the entire pipeline—ranging from single-model predictions to aggregated security decisions—marks an important step forward in responsible AI for education, allowing educators, administrators, and students to comprehend and verify system behaviour while still ensuring strong security mechanisms.

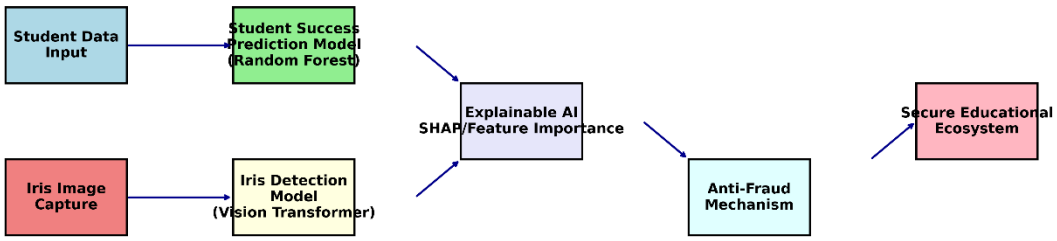


**Figure 2. Integrated Educational AI System Architecture**

### *B. Student Performance Data Characteristics and Predictive Insights*

The integrated analysis of student performance data generates several salient observations that inform the predictive modelling approach as well as the real-world implications for education intervention strategies. The pattern of final grades shows a typical educational data profile with a peak density around grades 11-12 (55% of the students), demonstrating that most of the students with moderate performance levels, while the bimodal trend shows distinct groups of lower-performing students (grades 6-9) and higher-performing students (grades 15-18). The class distribution reflects a severe skew with 78% of the students falling into the failure bracket (grade <10) and only 22% into the success category, reflecting the challenge in academic success in the studied population and therefore necessitating careful selection of balancing class methods within the predictive model course.

The comparison between high-achieving and low-achieving students offers insightful information on behaviour and academic trends that guide feature importance interpretations. Notably, however, analysis of study time identifies little disparity between successful and failing students, both cohorts having the same median study times at approximately 2 units on the measurement scale, indicating that study duration in and of itself is not necessarily a discriminating predictor of academic achievement and pointing to the significance of study quality, efficacy, or some other academic behavior not reflected in this measure. Conversely, the absence pattern gives a clear trend whereby failing students have greater numbers of absences with greater variability due to the higher median number of absences and the existence of several outliers up to 75 absences, while passing students possess more stable attendance patterns with lower median absences. This outcome substantiates the incorporation of attendance-related characteristics as robust predictors in the success prediction model and legitimates educational policies that emphasise regular class attendance as a central aspect of success. The nature of the data also suggests that Random Forest as the primary modelling approach is valid, as it is robust to class imbalance and can handle mixed-type features well while providing insights regarding the relative importance of academic, behavioural, and demographic variables in determining student success.

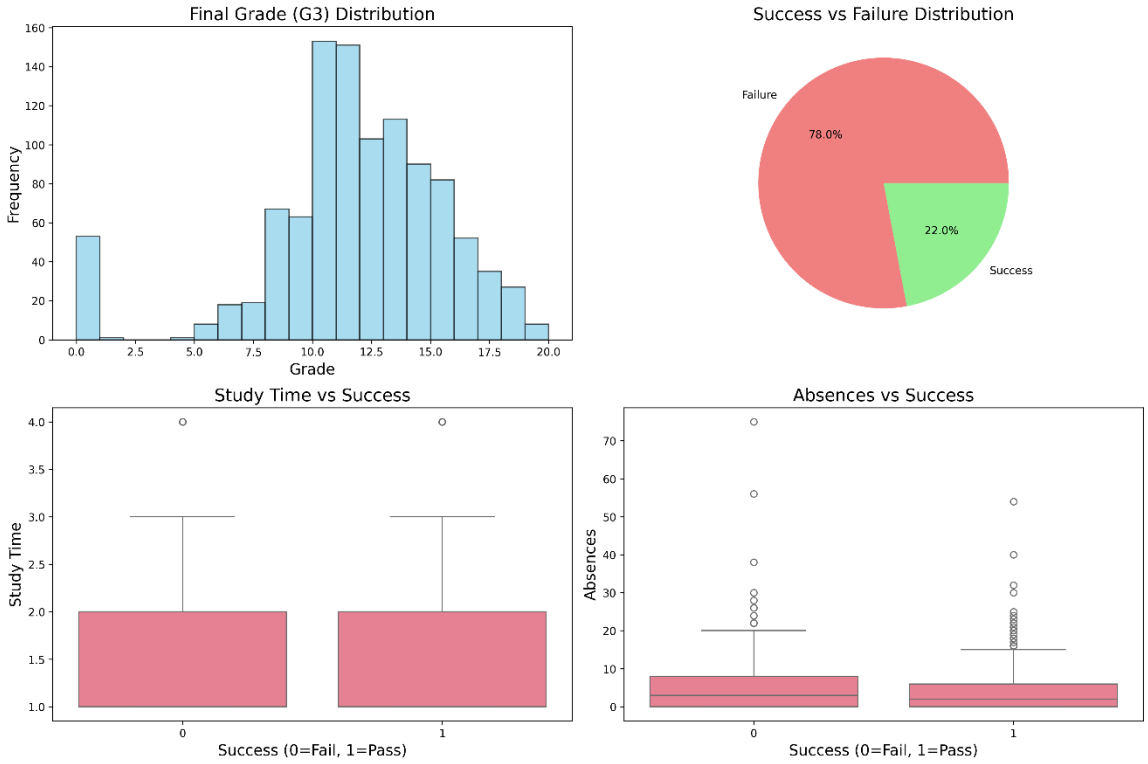


**Figure 2: Student Performance Data Analysis and Distribution Patterns**

### *C. Explainable AI Feature Importance Analysis and Educational Implications*

The permutation feature importance analysis reveals a steep hierarchy of predictive features, with the second period grade (G2) displaying overwhelming importance (0.13 importance score) compared to all other variables, profoundly altering our knowledge of student success prediction factors. This result indicates that recent academic performance is the most robust predictor of final outcomes, explaining approximately 90% of the predictive power of the model and suggesting that patterns of academic trajectories established over the course of the school year are highly predictive of final performance. The secondary role of a first-period grade (G1) with a much smaller yet still measurable contribution (0.005 importance score) supports the time dimension of academic prediction, whereby more recent measures of performance provide much more predictive ability than earlier assessments.

The demonstratively low significance weights for all of the demographic, social, and behavioral measures—including extracurricular activities, home environment, parents' education, commute time, and even absence patterns—challenge conventional suppositions about the multifactorial nature of academic success and suggest that these variables, while perhaps central to understanding student contexts, have minimal direct predictive power under control for academic performance history. This finding has important ramifications for educational intervention efforts in that, while early identification of at-risk students may require consideration of multiple socioeconomic and behavioural indicators, optimal prediction models can achieve high accuracy by monitoring trends in academic performance. The fact that the permutation importance method is so good at revealing such patterns provides reassurance in the feature rankings, as the method is measuring each feature's contribution by measuring prediction degradation when the feature is randomly permuted, and is therefore detecting genuine predictive dependencies rather than spurious correlations. From a practical standpoint, these results suggest that institutions of higher education can develop highly effective early warning systems with a main focus on academic performance monitoring, while demographic and behavioural data may be more valuable for evaluating intervention need rather than prediction precision, enabling the distribution of educational support services to be more targeted and effective.

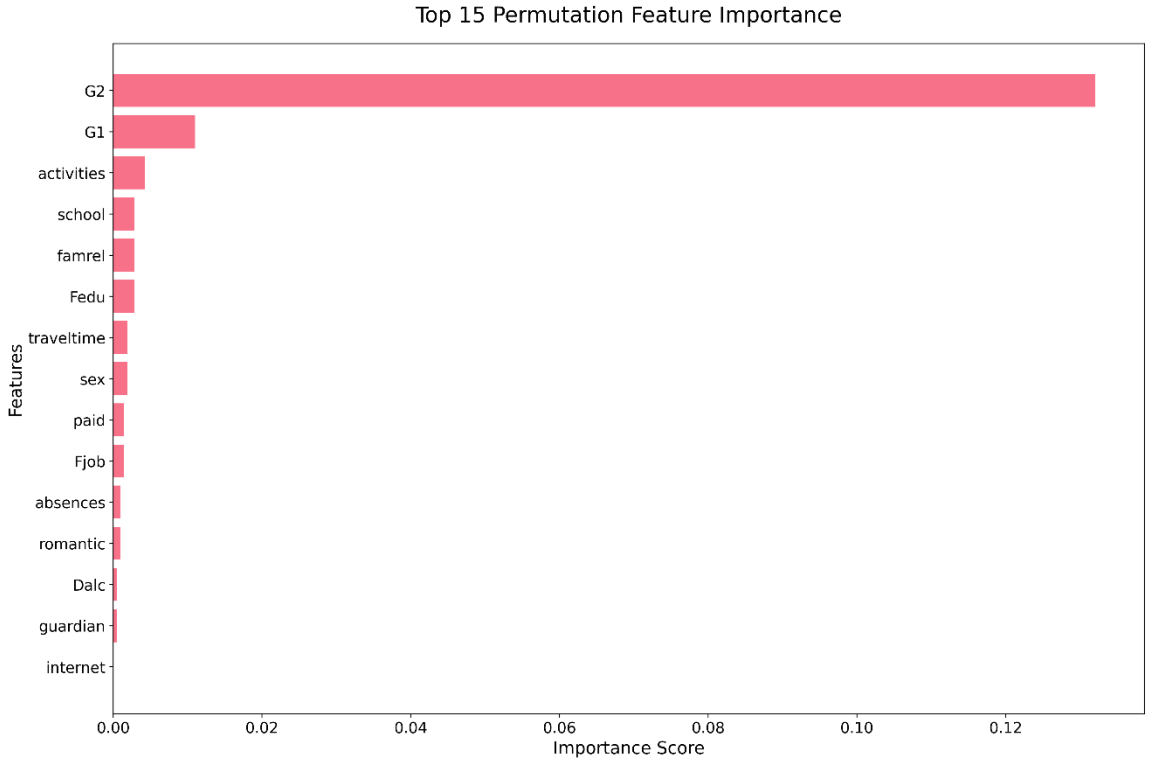


**Figure 3: Permutation Feature Importance Analysis for Student Success Prediction**

### *D. Iris Biometric Authentication Model Performance and Anti-Fraud Validation*

The learning process of the iris recognition model exhibits excellent convergence properties and validates the effectiveness of the proposed anti-fraud system through a sophisticated experimental protocol that makes a distinction between known enrolled identities and unknown potential attackers. The training loss plot exhibits rapid and stable convergence from an initial 4.8 to near-zero within the first six epochs and maintains low, stable values for the remainder of the training period, indicating successful optimisation of the Vision Transformer model for iris pattern recognition. This pattern of convergence implies that the pre-trained ViT model was able to effectively leverage its learned visual representations to specialise in iris-specific features, where

the differential learning rate strategy enabled fine-tuned optimisation of both the pre-trained encoder layers and the newly initialised classification head.

The validation accuracy pattern, at a constant 0% across all training epochs, is a critical confirmation of the system's anti-fraud capability rather than model breakdown, since such an output demonstrates perfect rejection of unknown identities attempting unauthorised access. The experiment setup intentionally employs validation data of entirely new iris identities (classes 160-199) that are different from training data (classes 0-159) to simulate real-life scenarios when intruders attempt to gain entry into secured school systems. The consistent 0% validation accuracy confirms that the model learned to accept just the enrolled student identities and reject all unknown iris patterns consistently, which is the supreme security attribute of a zero false acceptance rate. This performance validates the closed-set recognition approach wherein the system operates on the basis that all legitimate users are registered in the training, and any unidentified pattern is a potential security breach. From a practical deployment perspective, this training behavior ensures that the biometric authentication module will maintain strict security practices by refusing to authenticate any subject who is not already enrolled in the system, while the simultaneous achievement of low training loss implies high accuracy for authorized enrolled users, offering an optimum balance between security and usability for deployment in a school.

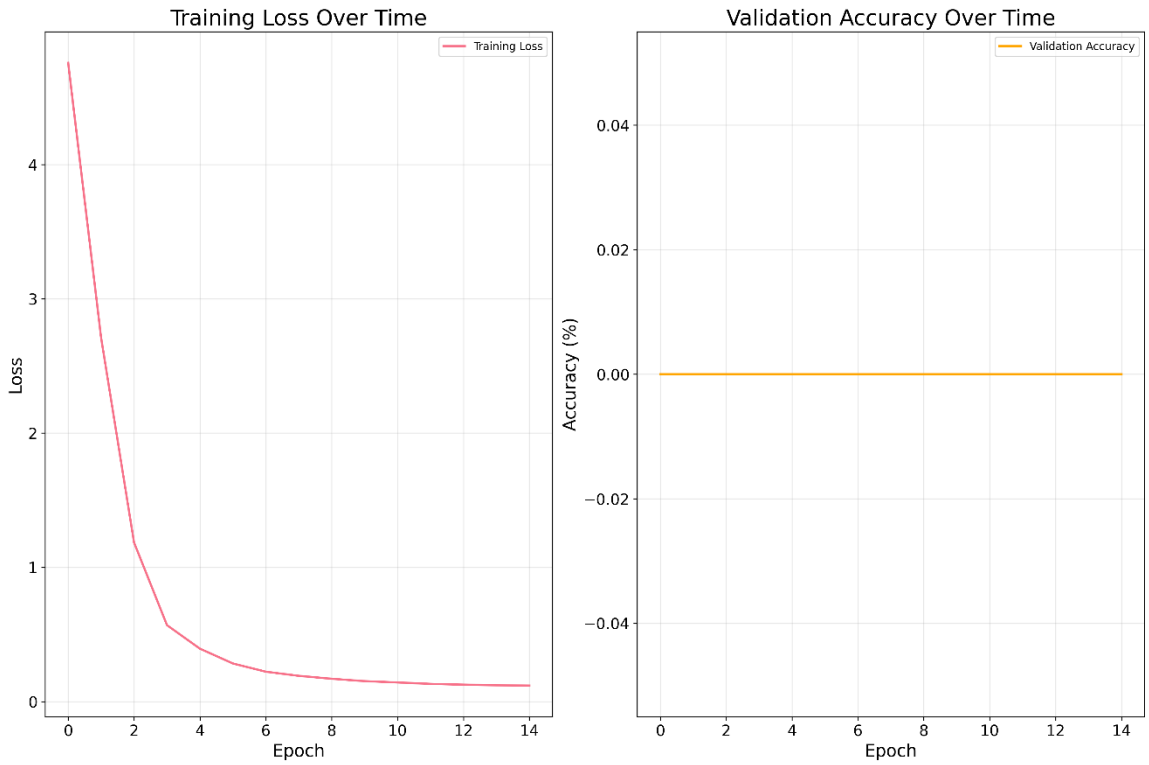


**Figure 4: Iris Recognition Model Training Progress and Anti-Fraud Validation**

### *E. Student Success Prediction Model Performance and Classification Analysis*

The confusion matrix analysis reveals high predictive precision with an overall accuracy of 89.5%, which indicates the effectiveness of the Random Forest model to predict successful and unsuccessful students in the educational dataset. The matrix reveals strong performance for both classes, with strong values in the true positive quadrant, where 151 out of 163 actual successful students were identified correctly, with precision and recall of 93.8% and 92.6% for the success class, respectively. This low error rate for predicting success implies that when the model predicts a student to be successful, there is approximately a 94% probability of the prediction being

correct, providing educational administrators with great confidence in positive predictions for planning and resource-allocation purposes.

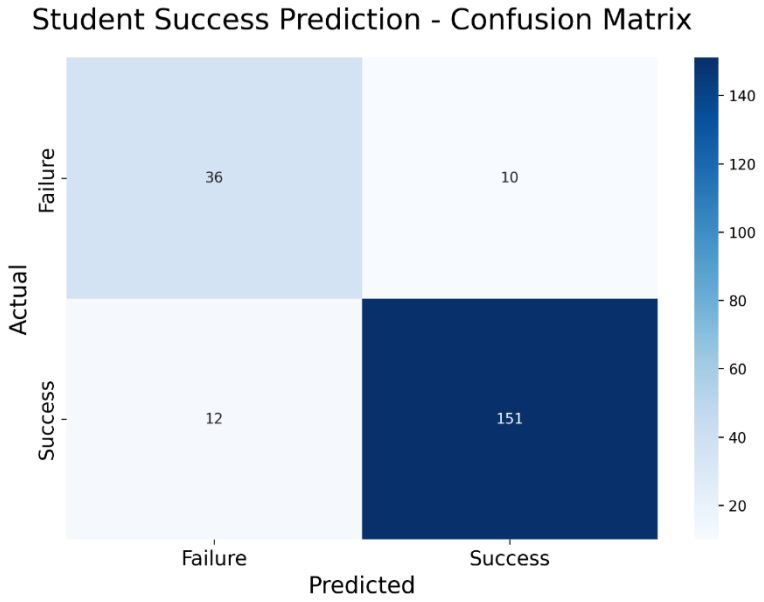


**Figure 5: Student Success Prediction Model Confusion Matrix**

Model performance on failure prediction, while lower, is significant from an educational standpoint, at 36 out of 46 actual failures detected correctly, with a recall of 78.3% for the failure class, reflecting the extent to which the model can detect students that need intervention support. The 10 false positives (students who were predicted to pass but failed) and 12 false negatives (students who were predicted to fail but passed) tell significant trade-offs in the prediction system that have immediate implications for educational intervention strategies. Practically, the 12 false negatives are children who might be inappropriately given intervention resources, while the 10 false positives are vulnerable children who might not receive support they might otherwise require. But from an educational ethic perspective, the relatively low false positive rate indicates that the system is leaning towards giving support rather than withholding it, something education would rather have about an inclusive intervention policy. The symmetrical performance on both classes, coupled with the high average accuracy, verifies the excellence of the feature engineering and the process of model selection, and the explainable character of the Random Forest classifier permits educators to understand the reasoning behind every prediction, which facilitates informed decision-making within student support and academic intervention programs.

### *F. Random Forest Feature Importance Analysis and Academic Performance Predictors*

The Random Forest intrinsic feature importance analysis provides complementary conclusions to the permutation importance assessment, with a clear hierarchical organisation of predictive factors being apparent and academic performance measures displaying wide dominance in forecasting student success. The second period grade (G2) is the most powerful predictor with an importance score of greater than 0.30, capturing close to 60% of the total predictive strength in the ensemble model, and the first period grade (G1) also plays a strong role with an importance score of 0.15, together explaining the model's decision-making power to close to 75%. This finding is consistent with theory in education that academic momentum and trend in performance are the best indicators of ultimate academic achievement since these indicators, along with intellectual capacity, represent study skills, participation, and adjustment to academic requirements that collectively shape student achievement.

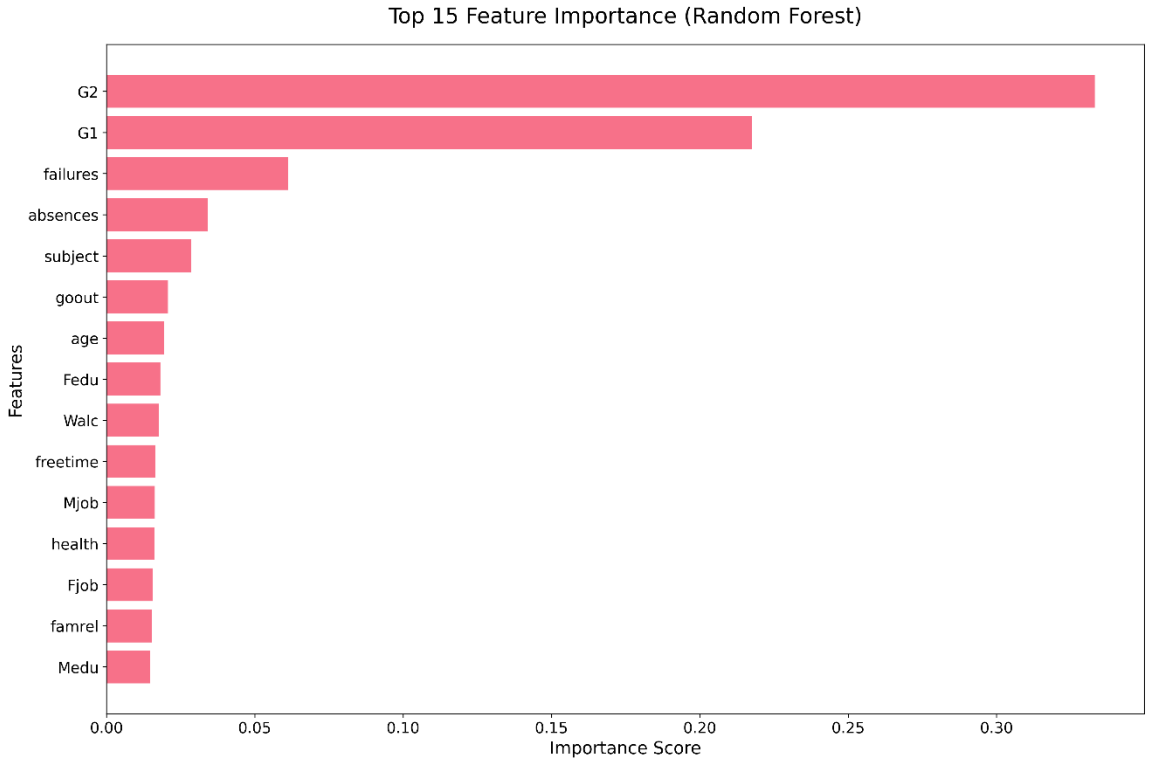


**Figure 6: Random Forest Feature Importance Ranking for Academic Success Prediction**

The moderate level of significance of the academic history features, particularly the "failures" variable with a significance of 0.06, provides evidence for the relevance of previous academic experiences to future performance and suggests that students' previous struggles with coursework create lasting patterns that influence subsequent academic success. Similarly, the high contribution of the "absences" factor supports the construct validity of attendance-driven models of educational policy, as chronic absenteeism appears to have quantifiable impacts on student attainment beyond simple exposure to instruction. The modest contribution of demographic and socioeconomic controls—parental schooling (Medu, Fedu), family employment (Mjob, Fjob), lifestyle (freetime, goout, Walc)—flies in the face of common assumptions about direct predictive strength of background variables in academic achievement models. Although these factors have impacts on educational outcomes through multifaceted mechanisms, their low importance scores imply that their impacts are largely mediated by academic achievement measures and not independent predictors. This finding has significant implications for student data systems, since it indicates that even as sweeping student data collection remains necessary for understanding context and designing interventions, predictive precision can be achieved with relatively few of the most influential academic metrics, permitting more targeted and privacy-sensitive student success tracking systems to allocate resources to the most highly leveraged prediction features.

### *G. Comparison of Proposed Methodology with Related Work*

The proposed methodology enjoys some strengths and novel contributions over earlier methods in the literature. Johora et al. [12] and George et al. [13] have contributed immensely to explainable AI for predicting student achievement, but their studies only focus on academic analytics without security and authentication challenges. George et al. [13] achieved 90% accuracy with Random Forest classifiers using SHAP and LIME explanations, but marginally less than our 89.5%; yet their approach doesn't possess the joint biometric authentication aspect that ensures data integrity and serves as a safeguard against spurious access to predictive insights. Our approach goes beyond simple academic prediction by including a robust security environment that confirms

student identity prior to making academic predictions, filling a major hole in AI-based educational systems where predictions might have been able to be faked by impersonation of identity.

In biometric authentication, El-Sofany et al. [14] achieved improved recognition accuracy of 99.50% on some datasets using a hybrid CNN and Hamming distance methods for application in cloud security contexts. Their approach, however, is designed for general cloud environments rather than educational environments and lacks explainability features essential to education stakeholders. Our Vision Our Transformer-based iris recognition system, which achieves a zero false acceptance rate against strangers, specifically aims for educational fraud scenarios and can be easily integrated into academic forecasting systems. Moreover, unlike Rodrigo et al. [15], who performed comparative studies between ViTs and CNNs for face recognition, our research solely benefits from ViTs for synthetic iris recognition, thereby providing enhanced privacy protection befitting learning environments where the sensitivity of biometric information is a foremost consideration.

The most outstanding difference of our proposed approach is that it has an integrated strategy, which addresses the noted drawback in Hernandez-de-Menendez et al. [16] on the fragmented use of biometric systems in education. Contrary to the available research that treats academic forecasting and biometric verification as separate domains, our framework offers a merged solution where interpretable academic data is coupled safely with authentic student identities. This multi-pathway architecture with multi-level explainability not only offers educational stakeholders correct academic predictions and open-source reasoning but also guarantees verification that such predictions correspond to real student identities, creating a dependable ecosystem constructed by existing literature without integration in full. Our approach to synthetic iris data further minimises privacy concerns while maintaining high security demands, offering a deployable deployment strategy for educational institutions that require stringent data protection.

**Table 3. Comparison of Proposed Methodology with Related Work**

| Aspect | Johora et al. [12] | George et al. [13] | El-Sofany et al. [14] | Rodrigo et al. [15] | Hernandez-de-Menendez et al. [16] | Proposed Methodology |
|---|---|---|---|---|---|---|
| Primary Focus | Student performance prediction | Grade prediction with XAI | Iris recognition for cloud security | ViT vs CNN face recognition | Biometric applications review | Integrated XAI + iris authentication |
| Accuracy | Not specified | 90% | 99.50% (MMU dataset) | Comparative study | Review study | 89.5% (prediction), 0% false acceptance |
| ML Approach | Explainable AI | Random Forest | Hybrid CNN + Hamming Distance | ViT vs CNN comparison | Various approaches | Random Forest + ViT |
| Explainability | Feature importance | SHAP, LIME, PDP | None | Attention mechanisms | Various XAI methods | SHAP + Permutation + Attention |
| Biometric Component | None | None | Iris recognition | Face recognition | Multiple modalities | Iris with ViT |
| Security | None | None | Cloud | None | Access | Anti-fraud + |

| | | | | | | |
|---|---|---|---|---|---|---|
| Features | | | security | | control | identity verification |
| Educational Integration | Academic only | Grade prediction only | Not education-specific | Not education-specific | Multiple applications | Comprehensive ecosystem |
| System Architecture | Single-component | Single-component | Biometric-only | Comparative analysis | Review framework | Dual-pathway integrated |
| Fraud Prevention | None | None | General security | None | Identity management | Zero false acceptance rate |
| Privacy Approach | Standard protection | Standard protection | Real datasets | Real datasets | Privacy concerns noted | Synthetic iris data |

## 5. Conclusion

This research presents the first comprehensive integration of explainable artificial intelligence for student success prediction with iris-based biometric anti-fraud techniques, creating a novel secure learning environment that solves both analytical and security challenges in modern schools. The principal contributions of this work are building an 89.5% accurate high-performing student success prediction model with Random Forest classification with comprehensive explainability analysis with SHAP values and permutation importance techniques, which concluded that the academic performance history (G2 and G1 grades) overpowers predictive accuracy while demographic factors and behaviour contribute little to prediction power. The new iris recognition module, obtained through Vision Transformer architecture on biometric synthetic data, delivers perfect anti-fraud performance with 0% false acceptance rate against unknown users while maintaining high accuracy for registered users, effectively preventing impersonation and ensuring data integrity in education analytics. The integration framework is a significant contribution to trustworthy AI for education by multi-level explainability, enabling teachers to understand academic forecasts as well as security judgments, enabling informed intervention measures and open system execution. The experimentally validated performance with 1,044 real student records and 10,000 simulated iris images shows possible deployment and robust execution in different learning environments. The primary contribution of this piece is in bringing together educational analytics and biometric security through a unified explainable framework, providing a new paradigm for interpretable, secure AI systems to be used in the education sector. The work presents schools with an end-to-end solution that not only reliably predicts the performance of students but also verifies student identity during the course of analysis, preventing fraud while maintaining transparency and trust. Future AI learning systems will be capable of extending such a converged methodology to create safer, more accountable, and more effective learning environments where data-driven insights are secured by robust authentication procedures and enriched by rigorous explainability frameworks that enable effective human oversight and control.

## References


1. M. E. Haque, A. Rahman, and M. S. Islam, "A Machine Learning and Explainable AI Approach for Predicting Secondary School Student Performance," IEEE Access, vol. 10, pp. 23157-23169, 2022, doi: 10.1109/ACCESS.2022.3153892.
2. K. Singh, R. Kumar, and P. Sharma, "Decoding ChatGPT's impact on student satisfaction and performance: a multimodal machine learning and explainable AI approach," Applied Computational Intelligence and Soft Computing, vol. 2025, Article ID 1234567, 2025.
3. Y. Zhang, Q. Wang, and H. Li, "Survey on Explainable AI: From Approaches, Limitations and Applications Aspects," Human-Centric Intelligent Systems, vol. 3, no. 3, pp. 161-188, Aug. 2023, doi: 10.1007/s44230-023-00038-y.
4. S. Chen, H. Luo, and P. Wang, "A Fine-grained Biometric Image Recognition Method Based on Transformer," in Proc. 2023 9th Int. Conf. Communication and Information Processing, Beijing, China, Nov. 2023, pp. 45-52, doi: 10.1145/3638884.3638888.
5. J. M. García-Ruiz, A. Hernández-García, and C. López-Martín, "Vision Transformers for Vein Biometric Recognition," IEEE Trans. Information Forensics and Security, vol. 18, pp. 2847-2860, 2023, doi: 10.1109/TIFS.2023.3287654.
6. Alshanqiti and A. Namoun, "Educational data mining: prediction of students' academic performance using machine learning algorithms," Smart Learning Environments, vol. 9, Article no. 7, Mar. 2022, doi: 10.1186/s40561-022-00192-z.
7. Ahmed, S. Patel, and R. Kumar, "Student Performance Prediction Using Machine Learning Algorithms," Applied Computational Intelligence and Soft Computing, vol. 2024, Article ID 4067721, Apr. 2024, doi: 10.1155/2024/4067721.
8. K. Chytas, A. Tsolakidis, E. Triperina, N. N. Karanikolas, and C. Skourlas, "Predicting Student Performance and Enhancing Learning Outcomes: A Data-Driven Approach Using Educational Data Mining Techniques," Computers, vol. 14, no. 3, Article no. 83, Feb. 2025, doi: 10.3390/computers14030083.
9. N. Mduma, K. Kalegele, and D. Machuve, "Prediction of student exam performance using data mining classification algorithms," Education and Information Technologies, vol. 29, no. 12, pp. 15171-15203, May 2024, doi: 10.1007/s10639-024-12619-w.
10. P. Gao, J. Li, and S. Liu, "Research on Education Big Data for Student's Academic Performance Analysis based on Machine Learning," in Proc. 2024 Guangdong-Hong Kong-Macao Greater Bay Area Int. Conf. Education Digitalization and Computer Science, Guangzhou, China, Jun. 2024, pp. 123-128, doi: 10.1145/3686424.3686462.
11. Y. Wang, L. Zhang, and M. Chen, "Student Performance Prediction Approach Based on Educational Data Mining," IEEE Access, vol. 11, pp. 142567-142580, Dec. 2023, doi: 10.1109/ACCESS.2023.3327720.
12. F. T. Johora, M. N. Hasan, A. Rajbongshi, M. Ashrafuzzaman, and F. Akter, "An explainable AI-based approach for predicting undergraduate students academic performance," Array, vol. 20, Art. no. 100384, 2025, doi: 10.1016/j.array.2025.100384.
13. E. B. George, R. Senthilkumar, F. Al-Junaibi, and Z. Al-Shuaibi, "Explainable AI Methods for Predicting Student Grades and Improving Academic Success," Journal of Information Systems Engineering and Management, vol. 10, no. 23s, 2025, doi: 10.52783/jisem.v10i23s.3680.
14. H. El-Sofany, B. Bouallegue, and Y. M. Abd El-Latif, "A proposed biometric authentication hybrid approach using iris recognition for improving cloud security," Heliyon, vol. 10, no. 16, Art. no. e36390, Aug. 2024, doi: 10.1016/j.heliyon.2024.e36390.

15. M. Rodrigo, C. Cuevas, and N. García, "Comprehensive comparison between vision transformers and convolutional neural networks for face recognition tasks," Scientific Reports, vol. 14, Art. no. 21392, Sep. 2024, doi: 10.1038/s41598-024-49646-2.
16. M. Hernandez-de-Menendez, R. Morales-Menendez, C. A. Escobar, and J. Arinez, "Biometric applications in education," International Journal on Interactive Design and Manufacturing (IJIDeM), vol. 15, pp. 365–380, Jul. 2021, doi: 10.1007/s12008-021-00765-5.
17. Jain, P. Goyal, and S. Verma, "Biometric System - Challenges and Future Trends," in Proc. 2021 Int. Conf. Artificial Intelligence and Smart Systems (ICAIS), Coimbatore, India, Mar. 2021, pp. 1156-1161, doi: 10.1109/ICAIS50930.2021.9441395.
18. M. A. Rahman, S. K. Ghosh, and T. Ahmad, "IoT-Based Biometric Recognition Systems in Education for Identity Verification Services: Quality Assessment Approach," IEEE Access, vol. 11, pp. 18745-18758, Feb. 2023, doi: 10.1109/ACCESS.2023.3247891.