TIFDA JOURNAL
Technology Information and Data Analytic
Journal Website : http://jeptira.id/index.php/tifda
e-ISSN : 3064-0660

Behavioral Biometric-Driven Educational Data Mining: CNN-Based Prediction of Students' On-Time Graduation from Handwritten Signatures

Herianto[1*], Khoirul Mustaan[1], Yahya[2], Nur Syamsiyah[2]
[1] Information Technology, Department Faculty of Engineering, Darma Persada University Jakarta
[2]Information Systems, Department Faculty of Engineering, Darma Persada University Jakarta
[1,2]South Malaka Park Road No. 8, East Jakarta 13450, Indonesia
*herianto@ft.unsada.ac.id

Abstract — Timely graduation is a fundamental metric in higher-education accreditation and a key indicator of institutional efficiency. Conventional prediction models largely rely on longitudinal academic records, which are lagging indicators and often fail to detect risks during the early stages of study. This research proposes a paradigm shift by leveraging behavioral biometrics—specifically, the analysis of handwritten signatures using Deep Learning—to predict students' graduation timelines and academic motivation profiles. Using a dataset from the Undergraduate Information Technology Program at Universitas Darma Persada, the study adopts the Cross-Industry Standard Process for Data Mining (CRISP-DM) framework. A Convolutional Neural Network (CNN) model based on the ResNet-50 architecture was developed, employing transfer learning to extract complex graphological features from signature images. Through rigorous data augmentation and statistical normalization, the model addresses the limitations of a small dataset. Empirical evaluation reports a graduation-prediction accuracy of 65% (Recall: 65%, F1-Score: 64%) and an academic-personality prediction accuracy of 70% (Precision: 74%, F1-Score: 69%). Although its absolute performance remains below transcript-based models, the findings validate the potential of signatures as early leading biometric indicators capable of capturing latent discipline and intrinsic motivation. This approach offers a non-invasive decision-support tool for academic advisors within intelligent education ecosystems.

Keywords – Educational Data Mining; Convolutional Neural Networks; ResNet-50; Behavioral Biometrics; Student Graduation Prediction; Computational Graphology

INTRODUCTION

The landscape of modern higher education faces a dual pressure between meeting increasingly stringent quality assurance standards and ensuring individual student success [1, 2]. In Indonesia, the National Accreditation Board for Higher Education (BAN-PT) explicitly positions on-time graduation rates as one of the most critical performance indicators, and failure to maintain an optimal graduation ratio not only lowers accreditation scores but also influences public perception, institutional financial stability, and, most importantly, students' professional futures [3, 4].

Consequently, the development of predictive mechanisms capable of identifying the risk of delayed graduation at an early stage has become a strategic necessity rather than an optional enhancement [4]. A practical challenge in academic administration is that students often suddenly "disappear" or experience unexpected performance decline without observable early warning in conventional academic records until the academic collapse is already too severe to repair [5]. Existing predictive models are predominantly reactive, relying heavily on lagging indicators such as Grade Point Average (GPA), attendance rate, and demographic information. Although such methods have achieved accuracies exceeding 85%–90% in various studies, their fundamental limitation lies in the fact that poor GPA is a consequence rather than an early predictor of the underlying problem; when students are identified as high-risk based on GPA in the fourth or fifth semester, interventions may be too late to change their graduation trajectory [6, 7].

The research examined in this report proposes a fundamentally different and innovative paradigm: the utilization of student signatures as behavioral biometric predictor variables. This approach is rooted in graphological and behavioral psychology hypotheses positing that signatures are not merely tools for identity authentication but condensed representations of unique personal attributes including personality traits, discipline, emotional stability, and professional ethos. Within the academic context, traits such as perseverance, orderliness, and intrinsic motivation are latent factors that correlate strongly with a student's capacity to complete a final thesis and graduate on time. If micro-patterns in handwriting strokes can be quantified and correlated with historical graduation outcomes, educational institutions would theoretically be able to detect graduation risks even before the first semester begins, using only the signatures collected during enrollment. Advances in Deep Learning, particularly Convolutional Neural Networks (CNN), have enabled automatic extraction of complex image features, replacing manual and subjective graphological assessments [8]. CNNs learn hierarchical representations ranging from simple strokes to abstract structural patterns that may not be perceptible to the human eye but are statistically predictive of target labels. Implementing advanced architectures such as ResNet-50 in this problem domain bridges the gap between behavioral psychology and computational data science, forming a decision-support system that provides new insights for academic advisors at Universitas Darma Persada and other higher education institutions.

Understanding the scientific contribution of this study requires positioning it within the intersecting domains of Educational Data Mining (EDM) and Computational Handwriting Analysis. EDM has evolved rapidly over the last decade driven by the increasing availability of data from Academic Information Systems [9] and Virtual Learning Environments [10-12]. Earlier studies predominantly employed traditional machine learning algorithms such as Support Vector Machine (SVM), particularly the One-Class SVM variant to handle data imbalance in detecting student performance anomalies, and algorithms such as C4.5 and Naïve Bayes [13], which, while effective, require extensive manual feature engineering and depend strongly on structured academic variables [14, 15]. A noticeable shift toward deep learning has emerged in recent years, with models such as Deep Belief Networks (DBN) and Long Short-Term Memory (LSTM) networks enabling more precise modeling of sequential learning behaviors. Comparative studies have shown that LSTM-based and Stacked BiLSTM models that process VLE clickstream data and continuous assessment records can achieve Area Under Curve (AUC) scores ranging from 0.86 to 0.90, outperforming traditional classifiers; however, such approaches rely on temporal learning interactions that are only available after the course has already started. The present study fills this gap by offering a temporally independent predictive mechanism capable of generating risk assessments without waiting for longitudinal academic or behavioral data, but instead using static artifacts available from the beginning of the academic journey [6].

Within the field of Computational Handwriting Analysis, literature demonstrates that handwritten signatures contain rich biometric features associated with personality traits. Prior studies, such as Gornale et al. (2021), reported that combined signature features can predict personality with accuracies of 96.64% for male subjects and 97.64% for female subjects, while other works showed that neural networks can infer Myers-Briggs indicators with accuracies of 83%–91%. The hypothesis linking signatures to on-time graduation is grounded in the Big Five personality construct of conscientiousness[16, 17], which implies self-discipline and emotional stability[18]. Students with consistent, firm, and well-structured strokes are presumed to exhibit higher levels of self-regulation supportive of timely thesis completion, whereas chaotic, inconsistent, or weak-pressure signatures may indicate hesitation or low persistence. The present study operationalizes this theoretical foundation by training a Deep Learning model to learn statistical correlations between pixel-level signature patterns and binary graduation labels (On-Time vs Not On-Time) as well as personality labels (Strong vs Regular Intrinsic Motivation), representing the key novelty of this research by integrating projective psychology with modern neural computation.

METHODOLOGY

This study adopts the Cross-Industry Standard Process for Data Mining (CRISP-DM) as its methodological backbone. This framework ensures that the technical development of the model remains aligned with institutional objectives, particularly accreditation enhancement and educational efficiency [19].

A. Data Understanding and Dataset Acquisition

The data acquisition phase was conducted within the Faculty of Engineering, Information Technology Department, Universitas Darma Persada. The collected dataset is multimodal, comprising handwritten signature images (unstructured data) and academic records (structured data).

a) Signature Data: Obtained from alumni and current students, serving as the primary input variables (studentid).

b) Graduation Labels: The number of semesters completed by each student was used as the ground truth label. Students graduating in ≤8 semesters were labeled "On Time", while those exceeding 8 semesters were labeled "Not n Time."
c) Personality Labels: In addition to graduation labels, each student was assigned a personality label based on academic advisors' subjective assessments of academic motivation, categorized as "Strong Intrinsic Motivation" or "Regular Intrinsic Motivation."

A major challenge in deep learning applications for this specific case is data scarcity. Unlike large-scale public datasets such as ImageNet, student signature data is highly limited. To mitigate overfitting where the model memorizes training samples instead of learning generalizable features this study implements an aggressive data augmentation strategy.

B. Data Preprocessing and Augmentation

The Data Preparation phase is critical for transforming raw images into a format suitable for neural network processing. This study utilizes the torchvision.transforms library to apply a series of geometric and statistical transformations.

Table 1. Data Transformation and Augmentation Strategies

Transformation Technique	Technical Implementation and Impact on the Model
RandomResizedCrop	transforms.RandomResizedCrop(224) Extracts a random region of the image and resizes it to 224×224 pixels. This forces the model to recognize local signature features (curves, strokes, dots) without relying on absolute position or global shape, improving scale invariance.
RandomHorizontalFlip	transforms.RandomHorizontalFlip() Flips the image horizontally with a probability of 0.5. Although handwriting has directional properties, this increases training variability and reduces overfitting to specific orientations.
ToTensor	transforms.ToTensor() Converts PIL or NumPy images (H × W × C) into PyTorch Tensors (C × H × W) with values normalized to [0.0, 1.0].
Statistical Normalization	transforms.Normalize(mean=[0.485, 0.456, 0.406], std=[0.229, 0.224, 0.225]) Standardizes pixel values using ImageNet statistics. This step is crucial for transfer learning, aligning the input distribution with pretrained model weights and accelerating gradient convergence.
CenterCrop (Testing)	transforms.CenterCrop(224) During validation/testing, only resizing and center cropping are applied. This ensures deterministic, distortion-free evaluation.

C. Model Architecture: ResNet-50 and Transfer Learning

Selecting an appropriate deep learning architecture is a key determinant of model performance. This study employs ResNet-50 (a 50-layer Residual Network) as the feature extraction backbone.

ResNet-50 is chosen for its ability to mitigate the vanishing gradient problem commonly found in very deep neural networks. The core mechanism of ResNet is the use of skip connections (shortcut connections). Mathematically, instead of learning a direct mapping H(x), a residual block learns:

$$F(x)=H(x)-x \quad (1)$$

and the block output becomes:

$$F(x)+x \quad (2)$$

This formulation allows gradients to propagate more efficiently during backpropagation, enabling the network to learn highly abstract and complex features from signature patterns without performance degradation. This research does not train ResNet-50 from scratch. Instead, it uses Transfer Learning with pretrained ImageNet weights. This approach is appropriate given the limited dataset size. The pretrained model already contains convolutional filters capable of detecting edges, lines, and basic textures features relevant to signature analysis.

The model's final Fully Connected (FC) layer was modified by replacing the original 1000-class classifier with a new linear layer corresponding to the study's target classes (2 classes: *On Time* vs *Not on Time*).

D. Training Dynamics and Optimization

Model training was conducted using hyperparameters calibrated for stability. The loss function is CrossEntropyLoss, a standard choice for binary or multiclass classification tasks.

The selected optimization algorithm is Stochastic Gradient Descent (SGD) with:

- momentum = 0.9,
- initial learning rate = 0.001.

Momentum accelerates gradient vectors in the correct direction while reducing oscillations.

A Learning Rate Scheduler (StepLR) is applied to reduce the learning rate by a factor of 0.1 (gamma) every 7 epochs. This annealing strategy enables rapid exploration of the loss landscape in early training stages, followed by fine-grained optimization in later epochs. The model was trained for 25 epochs, with a checkpointing mechanism that saves the bestperforming model weights based on validation accuracy.

RESULTS AND DISCUSSION

An essential aspect of applied informatics research is the transition from experimental modeling to deployable systems. In line with the CRISP-DM lifecycle, this study does not conclude at the notebook-based model evaluation stage; instead, it proceeds to full system deployment to ensure practical usability. A graduation prediction application was successfully developed using the Flask micro-framework, enabling the trained deep learning model to operate within a web-based decision-support environment.

A. The deployment architecture

The deployment architecture integrates multiple components that interact seamlessly to support real-time inference. The backend, implemented in Flask, loads the trained PyTorch model (.pth format) into memory at server initialization and exposes HTTP endpoints for model inference. The backend receives uploaded images and processes them following the exact transformation pipeline used during validation—resize, center crop, and normalization—thereby ensuring input consistency during deployment. The resulting tensor is passed to a ResNet-50 classifier to generate class probabilities.

The frontend interface, designed using Bootstrap, provides a responsive and user-oriented interaction layer. Available pages include a dashboard, a prediction input form for image uploads, and a prediction history report. These interfaces are intended to support academic advisors as end-users, allowing quick access to model predictions with minimal technical interaction.

To support system traceability and longitudinal analysis, a MySQL database is integrated into the application. The database stores user information, prediction logs, and personality prediction results across dedicated tables. The relational schema maps student identifiers to prediction outputs along with timestamps, enabling systematic tracking of decision-making records. This design ensures that the system can be extended for institutional analytics and model auditing in future implementations.

The inference workflow demonstrates operational viability in real usage scenarios. Once an advisor uploads a signature image, the system processes the file and returns the label with the highest probability (e.g., "Will Graduate On Time"), accompanied by a confidence score. Both the raw prediction and metadata are stored automatically in the database, ensuring accountability and decision traceability. The deployment results highlight that the proposed model can be effectively embedded into a real-time prediction system with minimal latency and reliable user interaction, marking a successful transition from model experimentation to production-ready application.

To provide a clear representation of the system's functional workflow and the interaction between users and the deployed model, the operational flow of the graduation prediction platform is illustrated in the following use-case diagram on Figure 1.

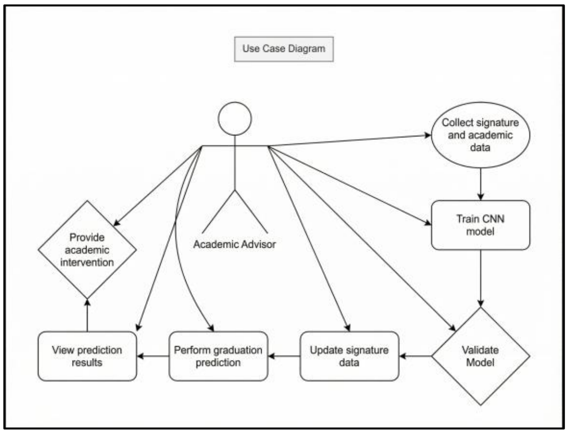

Fig.1. Use-case diagram of the behavioral biometrics–based graduation prediction system from the perspective of the academic advisor.

The diagram shows that the academic advisor acts as the principal system user who uploads student signature images, executes graduation prediction, and reviews the model outcomes to determine appropriate academic interventions. The system also supports continuous data lifecycle management through signature data updates and periodic model validation, thereby forming a feedback loop that contributes to improved model accuracy and supports evidence-based student mentoring practices.

B. Analysis of Empirical Results

The performance of the model was evaluated using standard classification metrics, namely Accuracy, Precision, Recall, and F1-Score. A detailed analysis of these metrics provides insights into the feasibility of using handwritten signatures as biometric predictors of academic performance.

The graduation prediction model achieved an overall accuracy of 65% on the test dataset. The confusion matrix reveals the distribution of prediction errors, as summarized in Table 2.

Table 2. Classification report for graduation prediction

Class	Precision	Recall	F1-Score	Support
Not On Time	0.71	0.5	0.59	10
On Time	0.62	0.8	0.7	10
Accuracy	—	—	0.65	20 (Total)
Macro Avg	0.66	0.65	0.64	—
Weighted Avg	0.66	0.65	0.64	—

The results indicate a performance imbalance across classes. The model achieves a high Recall (0.80) for the "On Time" class, suggesting that it is relatively reliable in identifying students who are likely to graduate on schedule. However, the Recall for the "Not

On Time" class is only 0.50, indicating that half of at-risk students are not detected. In the context of an academic early-warning system, such False Negatives represent critical misclassifications because they eliminate opportunities for timely intervention.

The Precision of 0.71 for the "Not On Time" class demonstrates that when the model predicts a delayed graduation, the prediction is often correct, yet its sensitivity remains insufficient. Although the accuracy of 65% exceeds the random baseline of 50%, it remains below the performance of transcript-based models, which typically achieve 80%–90%. These findings reinforce that signatures should serve as an auxiliary predictive signal, rather than a substitute for academic performance data.

C. Personality Prediction Performance

Interestingly, the model demonstrates superior performance in predicting academic personality traits, achieving an accuracy of 70%. The corresponding classification report is presented in Table 3.

Table 3. Classification report for academic personality prediction.

Class	Precision	Recall	F1-Score	Support
Ordinary Motivation	0.64	0.9	0.75	10
Strong Motivation	0.83	0.5	0.62	10
Accuracy	—	—	**0.7**	**20 (Total)**

To provide a clearer depiction of the distribution of correct and incorrect classifications, a confusion matrix was generated based on the 20 test samples. This visualization facilitates a more intuitive understanding of the model's performance across both classes.

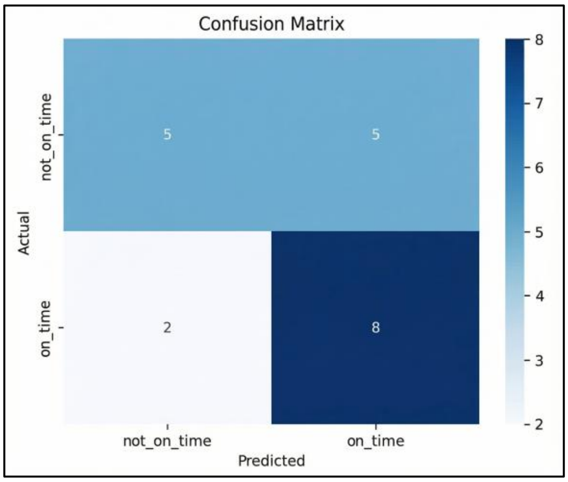

Fig.2. Confusion matrix for the graduation prediction task, illustrating the classification outcomes for the "On Time" and "Not On Time" labels.

As shown in Figure 2, the model correctly identified 8 out of 10 students who graduated on time, but only detected 5 out of 10 students who failed to graduate on time. Moreover, 5 samples in the "Not On Time" category were misclassified as "On Time", reinforcing the earlier observation that the model tends to miss at-risk students, leading to a relatively high false-negative rate. This pattern directly aligns with the precision–recall imbalance reported in the classification metrics.

The higher personality-prediction accuracy (70%) compared to graduation prediction (65%) empirically supports the foundational hypothesis of handwriting analysis. A signature is a direct product of an individual's motor and psychological system, making its correlation with personality traits more first-order in nature. In contrast, on-time graduation is a multifactorial outcome influenced not only by personality but also by external conditions such as financial capacity, health, family dynamics, or work commitments. Consequently, the inference from signature to graduation outcome represents a second-order correlation, naturally exhibiting higher noise. The high Precision (0.83) for the "Strong Motivation" class suggests that the model is particularly specific in recognizing signature features associated with strong intrinsic motivation an observation consistent with previous behavioral-biometrics research.

D. Comparative Study and Positioning of the Research

When compared with studies in the Educational Data Mining (EDM) domain, the proposed approach demonstrates clear novelty despite yielding moderate accuracy. Permana et al. (2024) achieved 89% accuracy using a Naïve Bayes model, while other deep learning research using academic transcripts reported 86.61% accuracy following hyperparameter tuning with Optuna. Although the signature-based model (65%) underperforms in raw accuracy, its strategic value lies in data availability. Transcript-based models require data from semesters four to six to achieve high predictive performance, whereas the signature-based model can be deployed on the first day of university enrolment. Thus, the model is positioned not as a replacement but as an early-stage screening tool capable of identifying latent personality-driven risk factors that may not appear in early GPA records.

E. Limitations

Analysis of the confusion matrix indicates that the size of the test set (20 samples) is relatively small to draw strong statistical conclusions. As a consequence, the variance of the results may increase when the model is applied to a broader population. In addition, the model shows a tendency to bias toward the majority class or toward dominant visual features present in the training samples. Although this limitation was partially mitigated through extensive data augmentation, increasing the availability of real signature data remains the primary recommendation for future development.

Furthermore, the model currently operates as a black box with limited interpretability. The integration of Grad-CAM or other feature-visualization techniques in future studies would be valuable for identifying which specific components of a signature (e.g., underlining strokes, size of capital initials, or writing slant) contribute most strongly to the model's decisions[20]. Such improvements would not only enhance trustworthiness but also increase the explainability of the system within academic advising scenarios.

CONCLUSION

This study has demonstrated both the technical feasibility and operational applicability of using Convolutional Neural Networks (CNN), particularly the ResNet-50 architecture, to predict student graduation timeliness and academic personality traits based on signature biometrics. By systematically applying the CRISP-DM methodology—from business understanding and data preparation to model development and full deployment via a Flask-based web application—the research successfully produced a functional system that can be adopted by Universitas Darma Persada for academic advisory purposes. The empirical results confirm that signatures contain predictive behavioural signals associated with academic performance, reflected in the model's accuracy of 70% for personality prediction and 65% for graduation prediction, thereby indicating that handwriting patterns exhibit meaningful correlations with psychological attributes that influence academic success. The findings further underline the effectiveness of deep learning techniques: ResNet-50, coupled with data augmentation and transfer learning, was capable of extracting discriminative features from a relatively small dataset, overcoming a key challenge in document image analysis. Although the model has not yet reached the accuracy level of transcript-based predictive systems, its value lies in its role as an early-stage decision support tool that operates prior to the availability of academic records. Future improvements should focus on multimodal fusion, integrating signature-derived feature vectors with academic and demographic data to achieve higher predictive performance. Exploration of lighter architectures such as MobileNet or EfficientNet may also enhance deployment efficiency, particularly for mobile-based academic advising applications. Additionally, enlarging the dataset, reducing class imbalance, and employing automated hyperparameter optimisation techniques such as Optuna are projected to improve model reliability and generalisation. Overall, this research offers both theoretical and practical contributions to the fields of Informatics and Educational Management by opening avenues for deeper integration of behavioural biometrics and learning analytics to support higher education quality enhancement.

REFERENCES

[1] Y. Yang, "Rethinking the Challenges Faced by the Quality Assurance System in Chinese Higher Education from the Perspective of Quality Culture," *The Educational Review, USA,* vol. 8, no. 8, 2024.

[2] D. S. Jarvis, "Regulating higher education: Quality assurance and neo-liberal managerialism in higher education—A critical introduction," *Policy and Society,* vol. 33, no. 3, pp. 155-166, 2014.

[3] R. F. Sari and M. R. Luddin, "Performance evaluation of academic services in the university using the balanced scorecard (Study at Indonesia Open University)," *President Pathumthani University,* p. 278, 2020.

[4] V. Sukmayadi and A. Yahya, "Indonesian education landscape and the 21st century challenges," *Journal of Social Studies Education Research,* vol. 11, no. 4, pp. 219-234, 2020.

[5] D. T. Altindag, E. S. Filiz, and E. Tekin, "Is online education working?," *Educational Evaluation and Policy Analysis,* p. 01623737241274802, 2021.

[6] A. Almalawi, B. Soh, A. Li, and H. Samra, "Predictive models for educational purposes: A systematic review," *Big Data and Cognitive Computing,* vol. 8, no. 12, p. 187, 2024.

[7] F. I. Olinmah, B. O. Otokiti, O. Abiola-Adams, and D. E. Abutu, "Integrating predictive modeling and machine learning for class success forecasting in creative education sectors," *Interventions,* vol. 29, p. 31, 2023.

[8] M. Moetesum, M. Diaz, U. Masroor, I. Siddiqi, and G. Vessio, "A survey of visual and procedural handwriting analysis for neuropsychological assessment," *Neural Computing and Applications,* vol. 34, no. 12, pp. 9561-9578, 2022.

[9] N. Sghir, A. Adadi, and M. Lahmer, "Recent advances in Predictive Learning Analytics: A decade systematic review (2012–2022)," *Education and information technologies,* vol. 28, no. 7, pp. 8299-8333, 2023.

[10] A. Rabelo, M. W. Rodrigues, C. Nobre, S. Isotani, and L. Zárate, "Educational data mining and learning analytics: A review of educational management in e-learning," *Information Discovery and Delivery,* vol. 52, no. 2, pp. 149-163, 2024.

[11] S. Sweta, "Educational data mining in e-learning system," in *Modern Approach to Educational Data Mining and Its Applications*: Springer, 2021, pp. 1-12.

[12] S. M. Dol and P. M. Jawandhiya, "Systematic review and analysis of EDM for predicting the academic performance of students," *Journal of The Institution of Engineers (India): Series B,* vol. 105, no. 4, pp. 1021-1071, 2024.
[13] H. Herianto, N. Syamsiyah, A. Arif B, and Y. Yahya, "Evaluasi Kinerja Datamining Pada Dataset Pendaftaran Mahasiswa Baru Dengan Class Yang Tidak Seimbang," *IKRA-ITH Informatika : Jurnal Komputer dan Informatika,* vol. 5, no. 3, pp. 162 - 168, 10/27 2021. [Online].
[14] S. D. A. Bujang *et al.*, "Imbalanced classification methods for student grade prediction: A systematic literature review," *IEEE Access,* vol. 11, pp. 1970-1989, 2022.
[15] P. Umamaheswari, M. Vanitha, P. V. Devi, J. G. Theporal, and B. R. Basha, "Student success prediction using a novel machine learning approach based on modified SVM," *Multidisciplinary Science Journal,* vol. 6, 2024.
[16] J. Erdheim, M. Wang, and M. J. Zickar, "Linking the Big Five personality constructs to organizational commitment," *Personality and individual differences,* vol. 41, no. 5, pp. 959-970, 2006.
[17] V. EGAN, "THE'BIG FIVE': NEUROTICISM, EXTRAVERSION, OPENNESS, AGREEABLENESS AND CONSCIENTIOUSNESS AS AN ORGANISATIONAL SCHEME," *Personality, personality disorder and violence: An evidence based approach,* p. 63, 2009.
[18] M. Yanu, H. Dwi, and H. Jati, "Emotion recognition for improving online learning environments: A systematic review of the literature," *Journal of Electrical Systems,* vol. 20, 2024.
[19] C. Schröer, F. Kruse, and J. M. Gómez, "A Systematic Literature Review on Applying CRISP-DM Process Model," *Procedia Computer Science,* vol. 181, pp. 526-534, 2021/01/01/ 2021, doi: https://doi.org/10.1016/j.procs.2021.01.199.
[20] R. R. Selvaraju, M. Cogswell, A. Das, R. Vedantam, D. Parikh, and D. Batra, "Grad-CAM: Visual Explanations from Deep Networks via Gradient-Based Localization," in *2017 IEEE International Conference on Computer Vision (ICCV)*, 22-29 Oct. 2017 2017, pp. 618-626, doi: 10.1109/ICCV.2017.74.