Journal of Theoretical and Applied Information Technology
31st December 2025. Vol.103. No.24

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

DEVELOPMENT OF INTELLIGENT ADVISORY SYSTEM WITH COGNITIVE TECHNOLOGY

PHANINTORN SUAPRAE[1], PRACHYANUN NILSOOK[2], PANITA WANNAPIROON[3], VITSANU NITTAYATHAMMAKUL[4]

[1] Ph.D. candidate, Division of Information and Communication Technology for Education, Faculty of Technical Education, King Mongkut's University of Technology North Bangkok, Bangkok, Thailand

[2] Professor, Division of Information and Communication Technology for Education, Faculty of Technical Education, King Mongkut's University of Technology North Bangkok, Bangkok, Thailand

[3] Professor, Division of Information and Communication Technology for Education, Faculty of Technical Education, King Mongkut's University of Technology North Bangkok, Bangkok, Thailand

[4] Lecturer, Ratchasuda Institute, Faculty of Medicine Ramathibodi Hospital, Mahidol University, Nakhon Pathom, Thailand

E-mail: [1]Phanintorn.s@psru.ac.th, [1]s6102052910059@email.kmutnb.ac.th, [2]prachyanun.n@fte.kmutnb.ac.th, [3]panita.w@fte.kmutnb.ac.th, [4]vitsanu.nit@mahidol.ac.th

ABSTRACT

Existing research on student retention mainly focuses on risk prediction, with few studies implementing advisory processes that translate predictions into timely, personalized interventions. This study develops and evaluates the intelligent advisory system with cognitive technology (referred to as the IAS-CT system) to improve student retention in higher education. The persisting gap in the literature is that most retention studies stop at risk prediction and rarely operationalize a closed-loop advisory workflow that converts predictions into timely, personalized interventions. Using de-identified institutional records of 2,973 undergraduates from academic years 2019–2022 with 25 academic and socio-demographic features, we trained and compared Decision Trees, Logistic Regression, Random Forest, K-Nearest Neighbors, and Naive Bayes. Preprocessing comprised imputation, normalization, and categorical encoding/selection; evaluation used a stratified split and standard metrics (accuracy, precision, recall, and F1) with confusion matrices. Correlation analysis indicated that GPA ($r = 0.55$), absenteeism ($r = 0.48$), father's income ($r = 0.45$), year of study ($r = 0.38$), and field of study ($r = 0.20$) were the most associated factors with retention. Decision Trees achieved the best predictive performance (accuracy = 98.90%), exceeding Logistic Regression (97.40%), Random Forest (86.10%), K-Nearest Neighbors (85.90%), and Naive Bayes (85.80%). The selected model was integrated into an advisory architecture that issues early-warning alerts, generates personalized study recommendations, and supports advisor–student communication. An expert panel rated the system's suitability at an overall high level. Consequently, the system operationalizes prediction into intervention, providing actionable retention support with practical implications for data governance and institutional scaling.

Keywords: *Cognitive Computing, Educational Data Mining, LINE Official Account, Machine Learning,*

1. INTRODUCTION

The number of educational institutions is increasing due to public and private initiatives, intensifying competition for university admissions. As applicant volumes grow, universities face pressure to differentiate and attract top candidates, making academic quality and standards central to institutional positioning. Leveraging educational data mining (EDM) allows institutions to gain evidence-based insights into academic performance and student success trajectories to inform curriculum enhancement and retention management [1]–[3]. Retaining students not only advances educational goals but also sustains institutional reputation and financial stability. Nevertheless, dropout rates remain a persistent challenge across many contexts. Understanding drivers of retention and designing strategies that both predict risk and enable timely support is therefore critical.

EDM is defined as the application of data mining and learning analytics techniques to educational data to uncover patterns that can inform decision-making and improve outcomes [4]–[7]. Prior work has modeled academic performance and developed retention prediction models; however, the practical translation of predictions into operational advising workflows remains limited**.** A student's early withdrawal reduces tuition revenue and may signal unmet academic or psychosocial needs; predictive insights should thus be coupled with targeted advisory actions that support at-risk students, minimize losses, and improve completion.

In northern Thailand—where students choose among multiple universities, including large regional institutions—competition for new entrants is high and retention is strategically consequential. For Phibunsongkhram Rajabhat University, understanding the factors that contribute to student retention can guide policy, resource allocation, and student services. High dropout rates can erode institutional reputation and perceived supportiveness, affecting attractiveness to prospective students, faculty, and funders. Cognitively enabled advising—that is, advisory processes enhanced by cognitive technology (e.g., pattern recognition, natural-language interaction, and decision support)—offers a promising path to *proactive* retention services when integrated with predictive analytics and sound data governance [8].

Previous studies emphasize data-driven decision-making in educational management [9], [10] and apply machine learning to dropout prediction in higher education and K-12 settings [11]. However, most studies predominantly report predictive accuracy while providing limited mechanisms to operationalize closed-loop advisory interventions. This leaves an actionable gap between risk detection and personalized, timely support (e.g., early warnings, tailored study recommendations, advisor–student follow-up) [12]. Accordingly, we position this study to contribute a retention approach that combines prediction with an implementable advisory workflow.

Our conceptualization integrates academic integration, social integration, student–advisor bonding, and institutional engagement as antecedents of student retention, aligning with classical retention perspectives while incorporating cognitive technology as an enabling layer for insight generation, case triage, and decision support. In this framing, prediction is necessary but not sufficient; institutions require mechanisms that transform model outputs into interventions within routine advising practice.

Existing retention research in our context provides accurate predictions but lacks a closed-loop advisory process that converts predictions into individualized, timely interventions within routine institutional workflows. This problem constrains the practical impact of EDM on day-to-day advising and campus-wide retention outcomes.

This study aims to develop and evaluate the intelligent advisory system with cognitive technology (referred to as the IAS-CT system) for undergraduate retention by linking predictive analytics to implementable advisory workflows, with four research objectives as follows: 1) To analyze the factors that predict the student retention of undergraduates; 2) To benchmark and develop the predictive models for student retention—integrated into the intelligent advisory system with cognitive technology; 3) To design and develop the IAS-CT system that operationalizes early-warning alerts, personalized study recommendations, and advisor-student communication; and 4) To evaluate the developed system's suitability through expert review.

The contribution of this study is twofold. Theoretically, it frames student retention as a closed-loop prediction-to-intervention model that layers cognitive technology onto established constructs (academic/social integration, student-advisor bonding, and institutional engagement). Practically, it presents an implementable IAS-CT system that operationalizes early-warning alerts, personalized study recommendations, and advisor-student communication within routine institutional workflows.

2. RESEARCH HYPOTHESIS

2.1 Under a pre-specified evaluation protocol (stratified train–test split on the training set), the IAS-CT system attains test-set accuracy $\geq$ 85%, with corroborating precision, recall, and F1-score.
2.2 An expert panel rates the system's overall suitability at a high level, defined as a mean Likert score $\geq$ 3.50/5.00.

3. THEORETICAL BACKGROUND

3.1 Intelligent Advisory System

The advisory system refers to a group of teachers officially appointed by a higher education institution to provide advice so that students can succeed in their studies [13]–[15]. In this study, the IAS-CT system is defined as an advising mechanism that not only identifies students at risk but also links risk detection to actionable support through early warnings, personalized study recommendations, and advisor–student communication. The consultation process with students aims to reduce the exit rate and is commonly described in stages—separation, transition, and incorporation—consistent with the student transition perspective [16]. Changes occur naturally with age and due to environmental factors [17]. Environmental impacts on student development include (1) clarity of institutional purpose, policies, practices, and activities; (2) institutional size; and (3) curriculum and teaching evaluations (e.g., course flexibility, evaluable programs, and modes of instruction). Other influential conditions include (4) residential arrangements that foster close relationships between students and institutional personnel; (5) faculty attention to student care and management; and (6) peer groups that effectively support academic life [18], [19]. Accordingly, prediction is necessary but not sufficient; the system must translate predictive outputs into timely, routine advising interventions intended to lower dropout and support student retention.

3.2 Data Mining

Data mining is the process of discovering patterns, correlations, and trends from large datasets using statistical, mathematical, and machine-learning techniques. It involves extracting valuable information to support decision-making and prediction**.** Key aspects include data cleaning (removing noise and inconsistencies), data integration (combining sources into a unified dataset), data selection (choosing relevant variables/records), data transformation (e.g., normalization or aggregation), pattern evaluation (identifying genuinely informative patterns), and knowledge representation (visualization and presentation for end-users). Common techniques include classification (assigning items to predefined classes), regression (predicting continuous values), clustering (grouping similar instances), association-rule learning (discovering relationships among variables), and anomaly detection (identifying outliers) [20]. In this study, educational data mining is applied to institutional undergraduate records to identify factors associated with retention and to train classification models whose outputs are integrated into the intelligent advisory system with cognitive technology.

3.3 Machine Learning (ML)

Machine learning is a subset of artificial intelligence (AI) that develops algorithms and statistical models that learn patterns from data and make predictions without explicit rule-based instructions [21]–[23]. Core paradigms include supervised learning (classification, regression), unsupervised learning (clustering, association), semi-supervised learning, reinforcement learning [24], and deep learning. Common techniques and algorithms include Logistic Regression (binary classification), Decision Trees, Random Forest, K-Nearest Neighbors (KNN), Support Vector Machines (SVM), and Neural Networks [25]–[27]. In supervised settings, models are trained on labelled examples to learn mappings from inputs to outputs; performance is then assessed on unseen data using standard metrics (e.g., accuracy, precision, recall, F1) and confusion matrices.

In this study, we focus on supervised classification for retention prediction and benchmark five classifiers consistent with our objectives—Decision Trees, Logistic Regression, Random Forest, K-Nearest Neighbors, and Naive Bayes—on institutional undergraduate records (Academic Years 2019–2022; n = 2,973; 25 features). Preprocessing (e.g., imputation, encoding, normalization) is applied within the modelling pipeline, and evaluation follows a stratified train–test split with cross-validation on the training set to ensure fair comparison and guard against data leakage. Outputs from the best-performing model are then integrated into the IAS-CT system to support early warnings, personalised study recommendations, and advisor–student communication.

3.4 Student Retention

Student retention refers to an institution's ability to keep enrolled students engaged and progressing to degree completion**,** preventing premature withdrawal from courses or programmes. It encompasses academic success, student satisfaction, institutional support, financial resources, and personal circumstances [20], [28]. In simpler terms, retention indicates how effectively a university sustains students' enrolment and active participation until educational objectives are achieved [29], [30]. In this study, retention is operationalized as a binary outcome (retained vs. not retained) to enable supervised prediction on institutional records and to inform targeted advisory actions within the intelligent advisory system with cognitive technology.

3.5 Cognitive Computing

Cognitive computing is an advanced technology that aims to simulate human thought processes in computerized systems, leveraging artificial intelligence (AI), machine learning, natural language processing, and data mining to create systems that can interact, learn, and support decisions [31]–[34]. Such systems are designed to handle complex, unstructured data and can understand context, reason over evidence, and respond in a human-like manner**.** They exhibit adaptability (learning from new data and experience over time), interactivity (natural-language engagement with users), context-awareness (considering time, location, preferences, and history to personalize outputs), and statefulness (remembering prior interactions to inform subsequent decisions). Cognitive computing has been applied across sectors such as healthcare, finance, and customer service—for example, to analyze large medical datasets or to support risk management [35], [36]. Overall, it enhances efficiency, decision-making, and user experience in tasks that typically require human intelligence [11], [37]. In this study, cognitive technology serves as the enabling layer of the intelligent advisory system: it provides pattern recognition from institutional records, natural-language interaction for advisor–student workflows, and decision support that translates predictive outputs into timely, personalized advising actions within institutional processes.

4. METHODOLOGY

This study adopts a design and development research (DDR) approach and proceeds through four interlinked phases: (1) Factor Analysis, (2) Model Benchmarking, (3) System Development, and (4) System Evaluation. DDR emphasizes iterative refinement of an implementable artifact—here, an intelligent advisory system with cognitive technology—grounded in empirical evidence and embedded within routine institutional workflows.

4.1 Phase 1: Factor Analysis

De-identified undergraduate records from Pibulsongkram Rajabhat University (PSRU), Phitsanulok, Thailand (n = 2,973; Academic Years 2019–2022) comprising 25 academic and socio-demographic features were assembled with a binary target (retained vs. not retained). A reproducible preprocessing pipeline—imputation, categorical encoding, and normalization—was employed; all transformations were fit on training data only and applied to validation/test partitions to prevent data leakage. Pearson's correlation (r) was computed for all 25 features against the student retention outcome; for brevity, we report the five strongest associations—GPA, absence from school, father's income, year of study, and field of study—which were used to prioritize variables for subsequent modeling.

4.2 Phase 2: Model Benchmarking

Findings from Phase 1 informed feature handling in Phase 2. Supervised classifiers—Decision Trees, Logistic Regression, Random Forest, K-Nearest Neighbors, and Naive Bayes—were developed and compared under a stratified 80/20 train–test split. The same preprocessing pipeline was applied across models to ensure parity and guard against leakage; hyperparameters were tuned on the training partition using internal validation. Performance on the held-out test set was reported using accuracy, precision, recall, F1-score, and confusion matrices. The best-performing model identified by this protocol was selected for system integration.

4.3 Phase 3: System Development

The IAS-CT system was developed as a digital intervention and embedded into routine advising practice through the following sequence:

1) co-design and requirements gathering with advisors and students to map advising journeys (risk awareness, study planning, follow-up), define roles/permissions, privacy needs, and communication tone;

2) data mapping and model handoff, aligning institutional sources (student information, grades/attendance, activity and advising records) to application services and specifying the model service contract (inputs/outputs and rationale fields) for the best model from Phase 2;

3) architecture and interaction design, specifying a modular application with institutional data connections and a RESTful model endpoint, and outlining the primary LINE Official Account (LINE OA) interaction alongside a secondary web-portal flow;

4) interface implementation, developing the LINE OA Rich Menu with five entries—(1) Evaluate Retention, (2) Academic Development,

(3) Social Development, (4) Contact Advisor, (5) Activities to Enhance—each bound to its service and relevant databases to enable immediate retrieval, logging, and context-aware recommendations;

5) personalization and advising content, composing rule-based advising notifications and micro-recommendations (e.g., study tips, attendance nudges) that translate model outputs into supportive actions and advisor prompts;

6) identity and access configuration via Single Sign-On (SSO) using institutional credentials and LINE Login, with Facebook and Google enabled only as additional SSO identity providers (not as communication channels); and

7) privacy-by-design controls, including de-identification at rest, consent/opt-out management, role-based access, interaction logging, and message-rate limiting to protect privacy and avoid notification fatigue. Standards-based, vendor-neutral technologies (HTTPS/TLS, OAuth 2.0/OpenID Connect, and REST APIs) were used to ensure portability and reproducibility while keeping advisor workload lightweight and student access seamless on everyday channels.

4.4 Phase 4: System Evaluation

Seven experts (n = 7) evaluate system suitability on a 5-point Likert scale across unified dimensions (implementation quality, data management, user–system interaction, responsiveness, retention-risk forecasting performance, reporting and explainability, data security), with a priori thresholds (e.g., mean ≥ 3.50/5.00 = High). Open-ended comments were collected to capture qualitative feedback for refinement.

5. RESULTS

5.1 Factor Analysis for Student Retention Prediction

To analyze the factors influencing student retention in higher education, this section presents the experimental setup and results obtained from various classification algorithms. The dataset, derived from the Pibulsongkram Rajabhat University Education Services Division, consisted of 2,973 undergraduate records collected between Academic Years 2019 and 2022, containing 25 academic and socio-demographic attributes. The dataset was split into training and testing sets using an 80:20 stratified partition, ensuring balanced representation of both retained and non-retained students. All features were preprocessed through normalization and categorical encoding to ensure data consistency and prevent bias.

Figure 1 illustrates the workflow of the student retention prediction process, which is embedded in the intelligent advisory system with cognitive technology. This framework is designed to assist higher education institutions in monitoring and supporting student retention. It applies artificial intelligence–based analytics to identify students at risk of attrition and provides timely, data-driven advisory interventions. The system's predictive insights are subsequently used to design and deliver personalized recommendations and early-warning alerts for at-risk students.

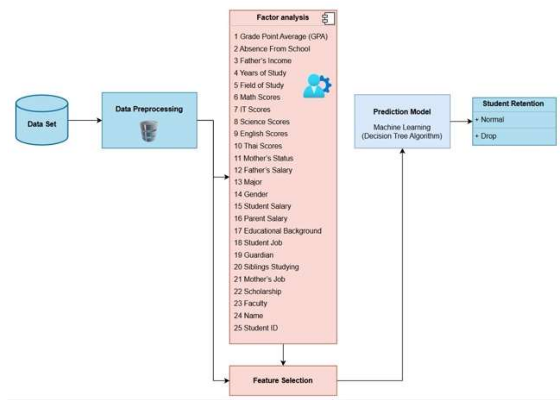

Figure 1: The Workflow of the Student Retention Prediction Process

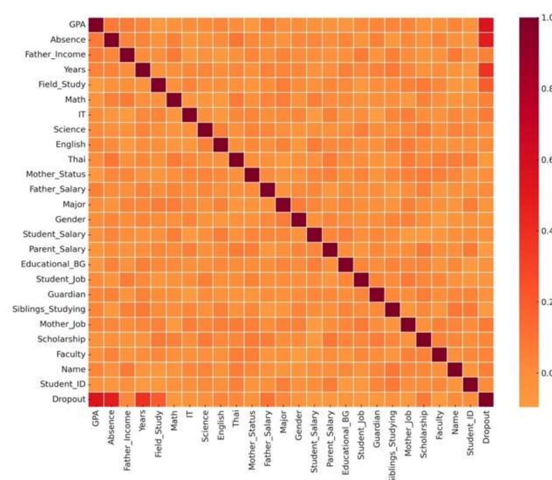

Figure 2: Correlation Analysis of Key Factors Related to Student Retention

Figure 2 presents the correlation analysis among the dataset's key factors, indicating the degree to which each variable is linearly associated with student retention. The correlation coefficient (r) ranges from –1 to +1, where a positive value (+1)

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

represents a perfect direct relationship and a negative value (–1) represents a perfect inverse relationship. A value near 0 indicates a weak or negligible association. This analysis helps identify features that exhibit the strongest relationships with student retention and are therefore prioritized in model training.

Table 1: Correlation Coefficients of the Top Five Predictive Factors

Order	Algorithm	Weights Correlation
1	Grade Point Average (GPA)	0.550
2	Absence from school	0.480
3	Father's income	0.450
4	Year of study	0.380
5	Field of study	0.200

Table 1 shows that the most influential predictors of student retention are, in descending order, GPA, absenteeism, father's income, year of study, and field of study. These findings confirm that academic performance and attendance behavior are the most significant indicators of student retention, while socioeconomic and program-related factors exert moderate influence.

5.2 Development and Evaluation of Predictive Models

To develop the predictive model for the intelligent advisory system with cognitive technology, five supervised machine learning algorithms—Decision Trees, Logistic Regression, Random Forest, K-Nearest Neighbors (KNN), and Naive Bayes—were trained and evaluated. The evaluation metrics included Accuracy, Precision, Recall, and F1-Score, all of which were computed from the confusion matrix, as illustrated in Table 2.

Table 2: Confusion Matrix Structure and Evaluation Metrics

		Predicted condition	
	Total population = P + N	Positive (PP)	Negative (PN)
Actual condition	Positive (P)	True positive (TP)	False negative (FN)
	Negative (N)	False positive (FP)	True negative (TN)

In predictive analytics, the confusion matrix provides a detailed breakdown of classification outcomes, including True Positives (TP), False Positives (FP), True Negatives (TN), and False Negatives (FN). This enables deeper analysis beyond overall accuracy by revealing both correct and incorrect classifications across categories. The following equations were used to compute performance metrics:

$$\text{Accuracy} = \frac{TP + TN}{TP + FP + FN + TN} \quad (1)$$

$$\text{Precision} = \frac{TP}{TP + FP} \quad (2)$$

$$\text{Recall} = \frac{TP}{TP + FN} \quad (3)$$

$$\text{F1 Score} = \frac{2 * Precision * Recall}{Precision + Recall} \quad (4)$$

Accuracy represents the overall proportion of correctly predicted cases; Precision indicates how many of the predicted positive instances were truly positive; Recall (also known as Sensitivity or True Positive Rate) measures how many actual positives were correctly identified; and F1 Score provides a harmonic mean between Precision and Recall, balancing both measures.

Table 3: Comparative Performance of Machine Learning Algorithms

Order	Algorithm	Accuracy (%)	*Precision* (%)	*Recall* (%)	*F1 Score* (%)
1	Decision Trees	98.90	99.50	94.00	96.90
2	Logistic Regression	97.40	91.50	93.50	92.50
3	Random Forest	86.10	89.10	77.50	85.70
4	K-Nearest Neighbor	85.90	89.10	77.10	85.60
5	Naive Bayes	85.80	89.10	77.70	85.60

Table 3 shows that the Decision Tree algorithm achieved the highest predictive performance, with an accuracy of 98.90%, followed by Logistic Regression (97.40%), Random Forest (86.10%), K-Nearest Neighbors (85.90%), and Naive Bayes (85.80%). This indicates that Decision Trees outperformed all other models in correctly classifying student retention outcomes.

The Decision Tree model was therefore selected as the core predictive component of the intelligent advisory system with cognitive technology. Its interpretability, ability to handle mixed data types, and robustness to missing values made it ideal for educational data mining contexts.

Model performance was further validated through cross-validation on the training set to ensure generalizability and to prevent overfitting.

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

The consistent results across folds confirmed that the model's predictive power was stable and reliable.

Consequently, the Decision Tree classifier was integrated into the advisory workflow to enable real-time early-warning alerts, personalized study recommendations, and automated advisor–student interactions. These outputs were used to support institutional decision-making and to guide targeted interventions for at-risk students.

5.3 The IAS-CT System

This section describes the architectural design of an intelligent advisory system that uses cognitive technology to support student retention in higher education. The architecture integrates machine learning–based prediction, cognitive interaction, and personalized advisory services within a unified workflow that facilitates early intervention for at-risk students.

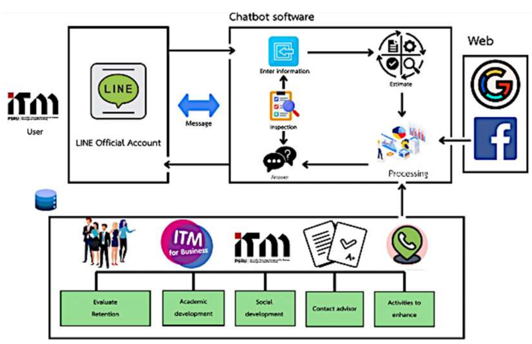

Figure 3: The System Architecture of the Intelligent Advisory System with Cognitive Technology

Figure 3 illustrates the architecture of the IAS-CT system for student retention can be explained as follows.

This section provides a detailed description of the architectural design and operational workflow of the developed IAS-CT system, as illustrated in Figure 3. The architecture is strategically centered around a Chatbot Software engine, which serves as the core processor that intelligently mediates all interactions between the user and the system's diverse functionalities.

The user's journey begins within the LINE Official Account (LINE OA), which functions as the primary, highly accessible user interface. Rather than a simple command-line chat, the user experience is guided by a graphical Rich Menu permanently displayed at the bottom of the chat screen. This menu presents the five core functionalities as distinct, tappable buttons: (1) Evaluate Retention, (2) Academic Development, (3) Social Development, (4) Contact Advisor, and (5) Activities to Enhance. When a user taps one of these buttons, a pre-defined command is transmitted to the Chatbot Software, initiating a systematic four-step internal process:

1) Command Reception and Inspection: The process starts as the system receives the user's command (Enter Information). It immediately proceeds to the Inspection phase, a critical step for validation and security. This involves two key actions: first, it validates the incoming command to understand the user's intent. Second, it authenticates the user's identity by leveraging a Single Sign-On (SSO) protocol that interfaces with external Web services like Google and Facebook, ensuring that only authorized users access the system;

2) Predictive Estimation: Once the user and their request are validated, the system moves to the Estimate phase. This is the analytical core of the system, where the high-performance Decision Tree model is invoked. Based on the specific function requested by the user, the model analyzes relevant data to calculate a predictive outcome. For instance, if "Evaluate Retention" is selected, the model computes the student's real-time retention risk score;

3) Data Processing and Response Formulation: The raw output from the estimation phase (e.g., a risk score) is then passed to the Processing module. This component does more than just relay the prediction; it enriches it by fetching relevant contextual information from the institutional Database. For example, it might combine a high-risk score with the student's recent course performance data to formulate a more comprehensive insight;

4) Answer Generation and Delivery: Finally, the system generates a user-friendly Answer. This is not merely data but a constructed response, which could be a clear summary of their risk status, a set of personalized recommendations, resource links, or a direct prompt to connect with an advisor. This complete, actionable response is then delivered back to the user through the LINE OA interface, effectively closing the interaction loop. This entire workflow demonstrates how the IAS-CT system seamlessly translates a simple tap on a Rich Menu into a data-driven, personalized, and actionable intervention, directly bridging the gap between predictive analytics and practical student support.

5.4 The Developed System's Suitability

The IAS-CT system was evaluated by a panel of seven experts in educational technology and higher education management. The evaluation covered implementation quality, data management, user–system interaction, response speed, forecasting performance, reporting capability, and data security.

The overall suitability score was high (M = 4.48, SD = 0.67). Experts rated data management and reporting capability at a very high level (M = 4.80), indicating that the system efficiently handled institutional data and presented clear, actionable information. Implementation quality (M = 4.70) was also rated very high, confirming the system's readiness for integration into institutional workflows.

Experts further observed that the LINE Official Account interface enhanced communication and accessibility, allowing real-time alerts and personalized recommendations between advisors and students. Qualitative feedback suggested minor improvements, such as refining dashboard visualization and adding longitudinal tracking for ongoing monitoring.

Table 4: The Suitability of the Intelligent Advisory System with Cognitive Technology

Assessment Items	M	SD	Suitability Level
1. Implementation of the developed system.	4.70	0.62	Very High
2. The developed system's data management capability.	4.80	0.67	Very High
3. User–system interaction of the developed system.	4.43	0.62	High
4. The response speed of the developed system.	4.40	0.61	High
5. The retention-risk forecasting performance of the developed system.	4.10	0.73	High
6. The reporting capability of the developed system.	4.80	0.67	Very High
7. The developed system's data security.	3.80	0.72	High
8. Accuracy and consistency of results reporting by the developed system.	4.80	0.71	Very High
Overall	4.48	0.67	High

6. DISCUSSION

This study aimed to develop and evaluate the IAS-CT system that improves student retention in higher education through predictive analytics and real-time advisory support. The system combines data-driven modeling with cognitive interaction, creating a closed feedback loop between prediction and intervention. This integration enables early identification of at-risk students and the delivery of timely, personalized guidance within institutional workflows.

The model benchmarking phase demonstrated that among five algorithms—Decision Trees, Logistic Regression, Random Forest, K-Nearest Neighbors, and Naive Bayes—the Decision Tree achieved the highest accuracy (98.90%). Its interpretability and efficiency make it particularly suitable for educational contexts, where transparency and explainability are critical. This result aligns with Cardona et al. [38], who emphasized that interpretable models are essential for administrative decision-making and advisor trust in data-driven systems.

The correlation analysis identified GPA, absenteeism, father's income, year of study, and field of study as the most influential predictors of student retention. These findings support Tinto's Student Integration Theory [29], which highlights academic and social engagement as the primary factors influencing student retention. The dominance of academic performance variables suggests that learning motivation, attendance, and academic discipline remain fundamental indicators of student retention, even in technologically mediated environments.

The system architecture operationalized these insights by incorporating five functional components—retention assessment, academic and cognitive development, social development, advisor communication, and extracurricular engagement. The integration of the LINE Official Account interface provided an accessible, interactive environment for students and advisors, enabling seamless real-time communication. This feature reflects a paradigm shift from reactive advising toward proactive, continuous, and data-informed educational support. Expert evaluation yielded an overall high suitability rating (M = 4.48, SD = 0.67), particularly in the dimensions of data management and reporting capabilities. These results validate the system's technical stability and pedagogical alignment. The evaluation also affirmed that the cognitive design enhanced usability, allowing advisors to engage efficiently with predictive insights and manage student risk dynamically.

A comparative analysis between prior research and the proposed intelligent advisory system highlights significant advancements in both predictive accuracy and advisory functionalities. Traditional retention studies have largely focused

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

on statistical or standalone machine learning models that predict dropout risk without embedding real-time, adaptive advisory mechanisms [11]. In contrast, the proposed system integrates cognitive technology to personalize recommendations, automate academic alerts, and foster adaptive intervention. This marks a shift from static analytics to interactive educational intelligence, providing immediate, context-aware responses that strengthen advisor–student engagement.

The improvement in advisory functionality demonstrates how cognitive computing bridges human and machine decision-making. Unlike earlier EDM-based systems that treated predictions as final outputs, the proposed system positions them as triggers for intelligent intervention, thereby enhancing institutional responsiveness. The system's adaptive behavior—learning from student interactions and advisor feedback—illustrates the evolution of predictive analytics into a self-improving advisory ecosystem.

When compared with the current state-of-the-art, the IAS-CT system offers a distinct contribution. Contemporary retention solutions often follow one of two paths: (1) predictive dashboards that present risk scores to administrators but lack automated intervention mechanisms, or (2) complex, "black-box" AI models like deep neural networks that may achieve high accuracy but lack the interpretability needed for stakeholder trust and adoption. In contrast, the IAS-CT system uniquely synthesizes three critical elements: first, *interpretable accuracy* through the use of a Decision Tree model; second, *actionable intervention* by translating predictions into concrete recommendations delivered through a familiar channel; and third, an *accessible interface* (LINE OA) that reduces adoption barriers and seamlessly integrates into the daily lives of students and advisors. This integrated approach distinguishes our work from systems that focus solely on prediction or employ less transparent methodologies.

The improvement in advisory functionality demonstrates how cognitive computing bridges human and machine decision-making. Unlike earlier EDM-based systems that treated predictions as final outputs, the proposed system positions them as triggers for intelligent intervention, thereby enhancing institutional responsiveness. The system's adaptive behavior—learning from student interactions and advisor feedback—illustrates the evolution of predictive analytics into a self-improving advisory ecosystem. From a pedagogical standpoint, the system advances the practice of data-informed academic counseling. From a technological perspective, this study underscores the transformative potential of combining AI, machine learning, and cognitive computing in higher education management. Its human-centered design ensures that predictive analytics augment rather than replace human expertise.

Thus, the developed system moves beyond traditional prediction models by closing the loop between analysis, advice, and action. It represents a new generation of intelligent systems that embody both cognitive and adaptive capabilities. The research contributes theoretically by extending retention models with cognitive layers and practically by providing a scalable, explainable, and human-interactive framework that strengthens student success and institutional resilience.

7. LIMITATIONS OF THE STUDY

While this study successfully achieved its objectives, several limitations should be acknowledged to contextualize the findings and guide future research.

First, the study's scope is confined to a single institution (Pibulsongkram Rajabhat University). The predictive factors and model performance may be specific to this demographic and academic context. Consequently, the generalizability of the model to other institutions with different student populations or educational structures may be limited without recalibration and further validation.

Second, although the Decision Tree model achieved a high accuracy of 98.90%, there remains a potential risk of overfitting, despite the use of a stratified train-test split methodology. The model's robustness and true predictive power on entirely new cohorts of students from subsequent academic years have yet to be longitudinally confirmed.

Third, the system's evaluation was conducted by a panel of experts. While their feedback provides a strong validation of the system's technical design and suitability, this study did not include a user-based evaluation involving students and academic advisors. Future research should incorporate usability testing and impact analysis from the perspective of end-users to assess real-world effectiveness and satisfaction.

Finally, the model identifies strong correlations between variables and student retention but does not establish causality. For instance, while GPA is a powerful predictor, it is an indicator, not necessarily the root cause, of a student's decision to withdraw. The system is therefore best utilized as a decision-support tool that empowers human advisors, who must still apply their professional judgment to understand the underlying reasons for a student's risk profile.

8. CONCLUSION

This study developed and evaluated the IAS-CT system to enhance student retention in higher education. The system integrates predictive analytics and cognitive interaction into a single process that transforms student data into meaningful, personalized interventions. Through its adaptive design, the system enables educational institutions to identify students at risk and provide timely academic and social support within existing advisory practices.

The developed system demonstrated high predictive accuracy and usability. The Decision Tree model provided effective classification performance, while the overall system achieved strong expert evaluations in implementation quality, data management, and reporting capability. The inclusion of a conversational interface through the LINE Official Account created a familiar and accessible environment for students, promoting interaction and continuous engagement with their advisors.

Beyond its technical success, the system contributes to a more holistic understanding of student retention. It shows how data-driven prediction, when combined with cognitive technologies, can extend traditional advising toward more proactive, personalized, and responsive academic support. This outcome reflects the evolving role of artificial intelligence in education as a bridge between institutional analytics and individual learning needs.

In practical application, universities can employ the system to monitor student progress, provide adaptive recommendations, and reduce dropout rates through early intervention. The design is flexible and can be adapted for different academic programs and institutional contexts. Its capacity for real-time feedback also supports evidence-based policy decisions and continuous improvement in student services.

For future development, further research could extend the system architecture to include longitudinal data tracking, counselor dashboards, and integration with existing learning management systems. Expanding the dataset and testing across multiple institutions would strengthen scalability and ensure broader applicability. Additional exploration of user experience and ethical considerations would also support the responsible deployment of AI-driven educational systems.

In conclusion, the IAS-CT system demonstrates that predictive analytics and cognitive computing can be effectively merged to create human-centered, intelligent educational support. It represents a practical step toward modern, data-informed student retention strategies, helping institutions foster learning continuity and long-term academic success.

REFERENCES:

[1] H. Dasi, and S. Kanakala, "Student dropout prediction using machine learning techniques", International Journal of Intelligent Systems and Applications in Engineering, Vol. 10, No. 4, 2022, pp. 408-414.

[2] J. S. Gil, A. J. P. Delima, and R. N. Vilchez, "Predicting students' dropout indicators in public school using data mining approaches", International Journal of Advanced Trends in Computer Science and Engineering, Vol. 9, No. 1, 2020, pp. 774-778.

[3] N. Shynarbek, A. Orynbassar, Y. Sapazhanov, and S. Kadyrov, "Prediction of student's dropout from a university program", Proceedings of 2021 16th International Conference on Electronics Computer and Computation (ICECCO), November, 2021, pp. 1-4.

[4] C. Romero, and S. Ventura, "Educational data mining and learning analytics: An updated survey", Wiley interdisciplinary reviews: Data mining and knowledge discovery, Vol. 10, No. 3, 2020, p. e1355.

[5] P. Blikstein, and M. Worsley, "Multimodal learning analytics and education data mining: Using computational technologies to measure complex learning tasks", Journal of learning analytics, Vol. 3, No. 2, 2016, pp. 220-238.

[6] J. E. Beck, and Y. Gong, "Wheel-spinning: Students who fail to master a skill", Proceedings of International conference on artificial intelligence in education, Springer (Berlin, Heidelberg), July, 2013, pp. 431-440.

[7] X. Du, J. Yang, J. L. Hung, and B. Shelton, "Educational data mining: a systematic review of research and emerging trends", Information Discovery and Delivery, Vol. 48, No. 4, 2020, pp. 225-236.

[8] P. Suaprae, P. Nilsook, and P. Wannapiroon, "System framework of intelligent consulting systems with intellectual technology", Proceedings of the 9th International Conference on Computer and Communications Management, July, 2021, pp. 31-36.

[9] A. A. Mubarak, H. Cao, and I. M. Hezam, "Deep analytic model for student dropout prediction in massive open online courses", Computers & Electrical Engineering, Vol. 93, 2021, p. 107271.

[10] S. Rawat, D. Kumar, P. Kumar, and C. Khattri, "A systematic analysis using classification machine learning algorithms to understand why learners drop out of MOOCs", Neural Computing and Applications, Vol. 33, No. 21, 2021, pp. 14823-14835.

[11] S. Issaro, P. Nilsook, and P. Wannapiroon, "Artificial Intelligence Engineering for predictive analytics", Proceedings of 2023 Research, Invention, and Innovation Congress: Innovative Electricals and Electronics (RI2C), August, 2023, pp. 51-58.

[12] V. Moisa, "Adaptive learning management system", Journal of Mobile, Embedded and Distributed Systems, Vol. 5, No. 2, 2013, pp. 70-77.

[13] S. Clarke, "Visible Learning: Feedback", Routledge, 2018.

[14] D. H. Schunk, and M. K. DiBenedetto, "Motivation and social cognitive theory", Contemporary educational psychology, Vol. 60, 2020, p. 101832.

[15] S. Chaiyarak, P. Nilsook, and P. Wannapiroon, "IVUL model: an intelligent learning development process", International Journal of Emerging Technologies in Learning (iJET), Vol. 17, No. 14, 2022, pp. 39-51.

[16] C. I. Bryce, A. M. Fraser, R. A. Fabes, and B. L. Alexander, "The role of hope in college retention", Learning and Individual Differences, Vol. 89, 2021, p. 102033.

[17] S. N. Young, W. R. VanWye, M. A. Schafer, T. A. Robertson, and A. V. Poore, "Factors affecting PhD student success", International journal of exercise science, Vol. 12, No. 1, 2019, p. 34.

[18] M. Rahman, and M. W. Kang, "Safety evaluation of drowsy driving advisory system: Alabama case study", Journal of Safety Research, Vol. 74, 2020, pp. 45-53.

[19] E. D. Churchill, M. R. Rogers, and K. A. Pristawa, "High Schoolers' and Middle Schoolers' Connections in Their Schools: Relation to Tardiness, Absences, Disciplinary Referrals, and Failed Courses", National Youth-At-Risk Journal, Vol. 4, No. 2, 2021, pp. 22-43.

[20] A. W. Astin, "What matters in college?: Four critical years revisited", John Wiley & Sons, 1997.

[21] I. Goodfellow, Y. Bengio, A. Courville, and Y. Bengio, "Deep learning", MIT press, 2016.

[22] T. M. Mitchell, "Machine learning and data mining", Communications of the ACM, Vol. 42, No. 11, 1999, pp. 30-36.

[23] E. Cam, and M. E. Ozdag, "Discovery of Course Success Using Unsupervised Machine Learning Algorithms", Malaysian Online Journal of Educational Technology, Vol. 9, No. 1, 2021, pp. 26-47.

[24] M. I. Jordan, and T. M. Mitchell, "Machine learning: Trends, perspectives, and prospects", Science, Vol. 349, No. 6245, 2015, pp. 255-260.

[25] M. Wang, and W. Deng, "Deep visual domain adaptation: A survey", Neurocomputing, Vol. 312, 2018, pp. 135-153.

[26] D. N. Lu, H. Q. Le, and T. H. Vu, "The factors affecting acceptance of e-learning: A machine learning algorithm approach", Education Sciences, Vol. 10, No. 10, 2020, p. 270.

[27] M. D. C. Nicoletti, and O. L. de Oliveira, "A Machine Learning-Based Computational System Proposal Aiming at Higher Education Dropout Prediction", Higher Education Studies, Vol. 10, No. 4, 2020, pp. 12-24.

[28] J. M. Braxton, A. S. Hirschy, and S. A. McClendon, "Understanding and reducing college student departure: ASHE-ERIC higher education report, volume 30, number 3", John Wiley & Sons, 2011.

[29] V. Tinto, "Leaving college: Rethinking the causes and cures of student attrition", University of Chicago press, 2012.

[30] H. He, Q. Zheng, D. Di, and B. Dong, "How learner support services affect student engagement in online learning environments", IEEE Access, Vol. 7, 2019, pp. 49961-49973.

[31] G. Muhammad, and M. S. Hossain, "Emotion recognition for cognitive edge computing using deep learning", IEEE Internet of Things Journal, Vol. 8, No. 23, 2021, pp. 16894-16901.

[32] A. E. Williams, and N. Petrovic, "The Role of Cognitive Computing and Collective Cognitive Computing in the Future Internet", preprint, 2021, pp. 1-14.

[33] V. L. Patel, J. F. Arocha, and J. S. Ancker, "Cognitive informatics in health and biomedicine: Understanding and modeling health behaviors", Springer, 2017.

[34] W. B. Gevarter, "Artificial intelligence, expert systems, computer vision, and natural language processing", Noyes Data Corporation/Noyes Publications, 1984.

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

[35] M. Minelli, M. Chambers, and A. Dhiraj, "Big Data Big Analytics: Emerging Business Intelligence and Analytic Trends for Todays Businesses", John Wiley, 2013.

[36] M. Haenlein, and A. Kaplan, "A brief history of artificial intelligence: On artificial intelligence's past, present, and future", California Management Review, Vol. 61, No. 4, 2019, pp. 5-14.

[37] G. Marcus, and E. Davis, "Rebooting AI: Building artificial intelligence we can trust", Vintage, 2019.

[38] T. Cardona, E. A. Cudney, R. Hoerl, and J. Snyder, "Data Mining and Machine Learning Retention Models in Higher Education", Journal of College Student Retention: Research, Theory & Practice, 2020