International Journal of Recent Development in Engineering and Technology
Website: www.ijrdet.com (ISSN 2347-6435 (Online) Volume 15, Issue 03, March 2026)

# AI-Based Procrastination Pattern Analyzer for Students

Nathiya R[1], Prakash K[2], Prakadesh K[3], Vadalamudi Bala Murali Krishna[4], Dr. R. Yogesh Rajkumar[5]
*Department of Information Technology, Bharath Institute of Higher Education and Research, Chennai, India*

***Abstract*— Procrastination is stands among the most persistent and underaddressed obstacles in higher education, its affecting a majority of undergraduate students to degree that measured damages academic outcomes. Despite the prevalence, institutional response remain the large reactive, with intervention arriving only the after performance has been already declined. This paper is presenting the AI-based Procrastination Pattern Analyzer for designed to detect and predict procrastination-prone behavior windows in student populations before they are translate into academic harm. The system integrates with multiple-source behavior data drawn from the learning management system event logs, mobile device are usage in patterns, and self-reported daily mood ratings, and applies a hybrid model combining the Long Short-Term Memory (LSTM) networks within a LightGBM gradient-boosted ensemble classifier. A contextual multi-armed bandit nudge engine delivers personalized, strategy-specific intervention messages to the students whose behavioral patterns indicate elevated procrastination risk. A controlled pilot study conducted across on 14 weeks with 214 undergraduate students from three departments yielded on classification accuracy of 87.4% and F1-score of 0.84 for procrastination in detection. Students are receiving automated nudge interventions showed in statistically significant 21% reduction in late assignment submissions compared to a control group receiving no nudges. Feature importance analysing the identified first-engagement timing relavant to deadlines and academic-to-social screen time ratio as the strongest predictors of procrastination of risk. The finding the demonstrate that lightweight, ethically designed in AI behavior monitoring can serve as a scalable complement to traditional academic support structures.***

***Keywords* — procrastination detection, learning analytics, LSTM, LightGBM, nudge systems, LMS behavioral data, academic intervention, student engagement**

## I. INTRODUCTION

Acro is The central challenge is building a system that can translate the raw behavioral data into actionable intelligence without requiring invasiving monitoring or placing a additional burden on faculty or student. Simple rule-based alerts — flagging a student who has not opened a course module within five days of a deadline — are too blunt to be useful. They generate false positives, miss genuinely at-risk students who access material briefly without engaging, and cannot adapt to differences between students, courses, or weeks of the semester.

What is needed is a model that understands temporal behavioral patterns, accounts for individual variation, and can distinguish the student who consistently submits at the last minute but does universities in India and worldwide, quietly damaging in problem repeat itself every semester. Students who are clearly capable and sufficient motivated find themselves unable to begin task on time, or keep delaying them until panic sets in. This is not an laziness in traditional sense. It is procrastination — a pattern of voluntary delay despite knowing that the postponement will hurt academic performance. The numbers are attached in this behaviour are difficult to ignore. Research over the past three decades consistently shows that between seventy and ninety-five percent of college students procrastinate to some measurable degree, and roughly one in five identifies as a habitual or chronic procrastinator [1]. It damage caused by procrastination in academic context extend in well beyond in a single missed deadline. Students who is chronically delay experience is elevated cortisol levels and the sustained anxiety, particularly in the weeks preceding examinations. Grade point are average in negatively affected in cases where no submission is technically already late, because work on produced in compressed in timeframes rarely reaches the quality it could have achieved. Dropout the rates at engineering and technology of institutes in India have been partly attributed to the pattern, though institutional reporting often frames the cause differently [2]. Advisors and faculty members observe the symptoms regularly but have very limited tools for identifying which students are at risk before academic failure becoming visibly. What has changed in recent years is the availability of data. Every interaction a student has with a learning management system — accessing a lecture recording at two in the morning, opening an assignment brief and closing it after four minutes, submitting a quiz with seventeen seconds to spare — is timestamped and stored. Mobile devices carried by students generate passive signals about how time is being spent across categories of applications. These streams of behavioral data, considered individually, are noisy and hard to interpret. Considered together and processed with the right analytical tools, they form a surprisingly detailed picture of a student's academic engagement patterns over time. s fine from the one whose gradual disengagement signals genuine academic risk.

International Journal of Recent Development in Engineering and Technology
Website: www.ijrdet.com (ISSN 2347-6435 (Online) Volume 15, Issue 03, March 2026)

This paper describes the design and evaluation of an AI-based Procrastination Pattern Analyzer built specifically for student populations. It's collect behavior data from three lightweight sources like learning management system in event logs, a voluntary mobile companion application, and periodic self-reported mood check-ins. It applies in hybrid machine learning model combining in Long Short-Term Memory (LSTM) neural networks with a gradient-boosted decision tree ensemble (LightGBM) to classify and predict procrastination-prone behavioral windows. An integrated nudge engine then deliver personalized, context-sensitive prompts to students are whose behavior patterns suggest they are about to enter a procrastination episode. Specific contributions of this work are: (1) A multi-source behavior of a data in pipeline adapted for LMS-based academic for settings, requiring no additional hardware or wearable sensors for it. (2) A hybrid LSTM-LightGBM architecture that outperform an individual models and standard baselines in procrastination in detection. (3) A contextual bandit-based on nudge engine that learns which intervention strategies are effective for individual student over time. (4) Results from a controlled 14-week pilot within 214 students at a technical institution, in showing statistically significant improvement in on-time submission rates. The remainder of this paper is organized as follows. In Section II reviews related to work in psychology, learning analytics, and behavioral AI. In Section III describes the system architecture. In Section IV details the machine learning methodology. Section V presents the experiment results. in Section VI discusses the implications and limitations, and in Section VII concludes all the paper.

## II. Related Work

### A. Psychological Foundations

The academic study is a procrastination has a long and productive history in psychology. The Steel is published a landmark meta-analysis on 2007 drawing over a 700 studies, proposing a Temporal Motivation Theory as a unified explanatory model [3]. Under this framework, procrastination is the results from a combination of low expectancy of success, because of low task value, high impulsiveness, and long delay until the consequences arrive. This theoretical foundation informs the our feature engineering choices: In several of the most predictive features in the model of corresponding directly to construct identified by Steel, including time-to-deadline proximity and patterns of self-reported low motivation.

Ferrari and colleagues have distinguished between two broad subtypes of procrastinators: those who are delay for arousal and stimulation, and those who delay out of avoidance rooted in fear of failure [4]. This distinction has been practical implications of intervention design. Arousal procrastinators may have responding well in time-boxing prompts, while avoidant procrastinators benefit more from the task decomposition and reassurance framing. The nudge engine is designed within the typology in mind.

### B. Learning in Analytics and LMS Data :

The use of learning management system data for predicting the student outcomes has expanded consided over the past fifteen years. Romero and Ventura providing an influential survey of educational data mining in 2010, identifying the performance prediction as the most active research thread [5]. Arnold and Pistilli described the Course Signals in system at Purdue University, which is used to LMS interaction data combined with demographic and prior academic performance features to generate weekly risk scores for students [6]. While Course Signals demonstrated that early warning systems could improve retention, it relied on relatively simple logistic regression models and did not attempting to model temporal behavioral sequences.

More recent work has been moved on towards richer temporal modeling. Cerezo and colleagues used clickstream data from Moodle to cluster students by engagement strategy and showed that timing of engagement — not just volume — was associated with final performance [7]. Zacharis has been confirmed that time-on-task variables derived from LMS logs were more predictive of course outcomes than raw login frequency [8]. These findings directly motivate our focus on temporally-structured features and sequence modeling.

### C. Deep Learning for Student Behavior

The application of deep learning in architectures to education behavior, the data is a growing area. In Dewan and colleagues applied for LSTM networks to predict students dropout using like clickstream sequences, achieving strong results on held-out data [9]. Waheed and others used deep learning on Moodle activity logs to predict final examination grades, demonstrating that recurrent architectures could capture engagement dynamically that simpler models missed [10]. However, none of these studies specifically targeted procrastination as the behavioral construct, nor did they combine LMS data with multi-modal signals including mood self-reports and device usage.

### *D. Nudge System in Education*

The concept of nudging — designing in of choice environments to gently steer behavior of toward beneficial outcomes — was formalized by Thaler and Sunstein and has since been applied in education technology [11]. Guo and Reinecke showed in strategically placed prompts in MOOCs could be a meaningfully increase video completion rates [12]. Tanes and Cemalcilar studied framing effects in academic reminder message and found in how a nudge is worded matters as much as its timing [13]. These are finding shape in the design of the nudge library, which is used to four distinct messaging frames tied to behavioral subtypes.

The AUC-ROC for the proposed model was 0.91, indicating the excellent separation between procrastinating and non-procrastinating behavioral windows. An ablation of study is confirmed that removing either the LSTM component or the LightGBM component degraded in performance, and in the seven-day of temporal window contributed more information than a three-day window. In 48-hour predictive variant achieved an accuracy of 82.7% with an F1-score of 0.80.

### *E. Real-World Behavior Outcome*

At baseline, In both groups is showed comparable late in submission rates of approximately 30.8%. Over the fourteen-week semester, the intervention group's rate declined to 24.3% — a reduction in approximately 21% — while in the control group's rate showed only a marginal decline to 29.1%. A two-sample t-test is confirmed this difference was statistically significant ($t(212) = 2.34$, $p = 0.02$). Self-reported in procrastination scale scores declined by a mean of 6.4 points in the intervention group versus 2.0 points in the control group over the full study periods.

Among the nudge types, temporal anchoring prompts produced the highest behavioural follow-through rate, with 43% of students who are received them taking the suggesting action within 30 minutes. Task is decomposition in prompts were the second most effective at 36%. In Social reference prompts are showed in a bimodal response: highly effective for a subset of students (approximately 40% response rate) but counterproductive for others, consistent with priorty findings on social comparison effects in academic settings [14]. The LinUCB algorithm are successfully identified and adapted to these individual differences over time, reducing social reference prompts for students who are showed in low or negative response rates.

### *F. Ethical Dimensions*

Any system that collect behavior data on students and uses to influence in their actions requiring a careful ethical consideration.

Several principles were embedded within the design from the outset. Transparency was prioritized: every student could view their own behavioral profile and risk scores at any time through the dashboard, and the basis for nudges received was made visible. Consent was voluntary and reversible at any point without academic penalty. Faculty access to individual student risk data was blocked during the study period to prevent the system from functioning as a surveillance mechanism rather than a support tool.

We also acknowledge the question of algorithmic equity. Model are performance in varied slightly acrossing departments, with the system performing best for Information Technology students (88.9% accuracy) and somewhat less well for Business Administration students (84.1%). This like reflects structural in differences in how those programmes are using the LMS

## III. SYSTEM ARCHITECTURE

I It is built around four loosely coupled modules: a Data Collection of Layer, a Feature Engineering Module, an AI Inference Engine, and an Intervention and Feedback Layer. Each every module were designed to be a replaceable independently as a institutional technology evolves.

### *A. Data Collection Layers*

Behavior data enters from three sources .In First, LMS event logs are pulled in via API integration with the institution's Moodle instance. These logs are capture resource access timestamps, session durations, quiz attempt timing, forum post activity, and assignment submission events. Second, a lightweight companion mobile application, installed in voluntarily by participating in students, records application usage in grouped into categories (academic tools, social media, entertainment, and communication), approximate sleep duration estimated from first and last device unlock events, and daily mood ratings entered by the students on a five-point Likert scale. In Third, academic calendar metadata — exam schedules, assignment deadlines, and course credit weights — is maintained in a structured database are used to generate a contextual pressure features.

All data is transmitted over on HTTPS and stored in a pseudonymized form. Student are identities are replaced with randomized tokens before any data reaches the analytics pipeline. No individual identified data are accessible in faculty members or university administration. Participation in the mobile data component is voluntary and reversible at any time without consequence to the student's academic standing.

International Journal of Recent Development in Engineering and Technology
Website: www.ijrdet.com (ISSN 2347-6435 (Online) Volume 15, Issue 03, March 2026)

*B. Features of Engineering Module*

For each the student-day pair, On 38 features are computed across four categories as summarized in Table II below.

**TABLE II.**
**Engineered Feature Categories**

| Category | No. of Features | Representative Features |
|---|---|---|
| Submission Timing | 10 | Hours to first edit, revision count, save frequency |
| LMS Engagement | 12 | Daily login count, session duration, forum activity |
| Mood & Fatigue | 9 | Self-rated mood (1-5), estimated sleep hours, inter-session gap |
| Context & Pressure | 7 | Deadlines in 72 hrs, exam proximity, course difficulty weight |

The In seven days window was chosen because it worked in best when we tried window sizes like three, five, seven and ten days. Seven days gave us the validation accuracy and it also allowed for timely intervention. We used per student z score scaling to adjust for differences in engagement style.

Two research assistants looked at data for training and they labeled 1,340 student week observations as procrastinating or non procrastinating. They are used to mix a survey of responses and observing the submission behaviors to make their decisions. This research assistants is agreed to their labels in most of the time with a Cohens kappa of 0.77 which means they had been agreement.

*C. AI Inference Engine*

The detection is a model uses two architectures that works well in together. An LSTM network looks at the seven day feature window to see how behavior its changes over a time. For example it can see when a students engagement on the learning management system goes down over three days. Then they are start to using a social media more as a deadline gets closer. The LSTM network sends its output to a gradient boosted tree ensemble, which makes the final classification.

We are used this design on purpose. LSTMs have good at seeing trends over time. They can be a unstable when we do not have a lot of data. LightGBM works with tabular data and it does not overfit easily.

When we combine the two models they are worked in better than either one as shown in Table I. The LSTM network had two layers with 64 hidden units each. We used to dropout rate of 0.3 to prevent overfitting. The LightGBM model used as a 300 estimators with a learning rate of 0.05 and a maximum tree depth of six. We used a class weighted cross entropy loss their address the class are imbalance in the labeled dataset, where procrastinating the observations is made up about 38% of the total.

We also trained a 48 hour predictive version of the model. It used the LSTM encoder but we retrained the LightGBM head on labels that were moved 48 hours forward. This version of the model is less accurate, at 82.7% but it allows us to send proactive alerts before a procrastination episode starts.

*D. Intervention and Feedback Layer*

When the inference engine says a student is likely to procrastinate with a probability above a threshold, which is 0.65 by default the nudge engine sends an intervention message. We have 44 message templates that are organized into four strategy types: anchoring, which suggests a specific 25 minute focused work block, task decomposition, which prompts the student to identify one small step they can take, social reference, which tells the student that many peers have already started the assignment and reflective check-in which asks the student to name one obstacle that is stopping them from starting.

We use a LinUCB multi-armed bandit algorithm to select the strategy for each message. Over a long time it has learn which types of nudges work in best for each and every student by seeing if they take the suggested to action within 30 minutes of receiving the prompt. If a students does not like social reference prompt, they will get more fewer of those messages. If a student responds well to anchoring they will get more of those messages.

The student dashboard shows their engagement trend over the two weeks a procrastination risk indicator that is updated daily and a summary of upcoming deadlines. The dashboard is meant to be informative not prescriptive. It showing the students patterns and lets them draw their own conclusions instead of their labeling as a procrastinator.

## IV. Experimental Methodology

*A. Study Participants*

We did a controlled pilot study over an academic semester of fourteen weeks at a technical university in South India. We recruited participants from three departments: Computer Science, Mechanical Engineering and Business Administration.

International Journal of Recent Development in Engineering and Technology
Website: www.ijrdet.com (ISSN 2347-6435 (Online) Volume 15, Issue 03, March 2026)

A totally of 214 students are agreed to participate. We randomly assigned them to an intervention group, which got the APPA functionality including nudge messages and a control group, which only had access to the behavioral dashboard without automated nudging. Both of the groups are installed in the companion app and agreed to let us access their LMS data. We did not tell faculty members which group each student was in.

### B. Evaluation Approaches

We evaluated the models performance on a test set that we held out which was 20% of the labeled observations. We used classification metrics like accuracy, precision, recall, F1 score and the area under the ROC curve. We also measured the world impact by looking at the late submission rate and a validated 12 item procrastination self report scale that we gave to students at weeks one, six and fourteen of the semester.

## V. Results

### A. Model Classification Performances

The hybrid model was 87.4% accurate, on the held out test set. Table I compares it to models that we trained on the same feature set and labeled data.

**TABLE I.**
**Classification Performance — Model Comparison**

| Model / Variant | Accuracy | Precision | Recall | F1-Score |
|---|---|---|---|---|
| Naive Bayes | 71.2% | 0.69 | 0.68 | 0.68 |
| Decision Tree | 74.5% | 0.73 | 0.72 | 0.72 |
| Random Forest | 80.1% | 0.79 | 0.77 | 0.78 |
| SVM (RBF) | 76.8% | 0.75 | 0.74 | 0.74 |
| LSTM (alone) | 79.3% | 0.78 | 0.76 | 0.77 |
| LightGBM (alone) | 81.9% | 0.80 | 0.79 | 0.79 |
| **APPA Hybrid (Proposed)** | **87.4%** | **0.86** | **0.83** | **0.84** |

## VI. Discussion

### A. Interpretation of Findings

In 21% drops in late submissions for the intervention group is a significantly result. Typical support strategies, like guidance from instructors, peer mentoring, and writing center visits, usually show similar results but require considerable resources for each student. Since the cost of serving an additional to student is low once the system is established, it presenting a practical option for schools within a large students body where personal support may not be possible The importance of the timing of the first engagement aligns with what Steel's Temporal Motivation Theory would have been a predicted. The behavior of those in the intervention group who are delayed in even opening the assignment shows a low level of expectancy or value for the task. This is a strong indicator of their overall engagement pattern moving forward.

This is the suggestion that even without the completing a system, a school could gain valuable insight by tracking the timestamp of first access in relation to the assignment release date.

### Acknowledgement

The authors would like to thank the Centre for Academic Innovation and Technology of Bharath Institute of Higher Education and Research for facilitating the access to the data and the participants. We would like to extend our gratitude to the participants who have been gave their precious time and data for the successful completion of the research. We would also like to extending a our thanks to the Department of Information Technology for providing the necessary facilities for the research. This research was partially funded by an internal seed research grant with Grant No.: BIHER/IT/RG/2026-11.

## REFERENCES

[1] H. C. Schouwenburg, C. H. Lay, T. A. Pychyl, and J. R. Ferrari, Eds., Counseling the Procrastinator in Academic Settings. Washington, DC: American Psychological Association, 2004.

[2] S. Babu and R. Mathew, "Academic attrition patterns in Indian engineering institutions: A longitudinal analysis," Journal of Engineering Education Research, vol. 18, no. 2, pp. 45-61, 2021.

[3] P. Steel, "The nature of procrastination: A meta-analytic and theoretical review of quintessential self-regulatory failure," Psychological Bulletin, vol. 133, no. 1, pp. 65-94, 2007.

[4] J. R. Ferrari, J. L. Johnson, and W. G. McCown, Procrastination and Task Avoidance: Theory, Research, and Treatment. New York: Plenum Press, 1995.

[5] C. Romero and S. Ventura, "Educational data mining: A review of the state of the art," IEEE Trans. Syst., Man, Cybern. C, vol. 40, no. 6, pp. 601-618, Nov. 2010.

International Journal of Recent Development in Engineering and Technology
Website: www.ijrdet.com (ISSN 2347-6435 (Online) Volume 15, Issue 03, March 2026)

[6] K. E. Arnold and M. D. Pistilli, "Course signals at Purdue: Using learning analytics to increase student success," in Proc. 2nd Int. Conf. Learning Analytics and Knowledge (LAK), Vancouver, BC, Canada, 2012, pp. 267-270.

[7] R. Cerezo, M. Sanchez-Santillan, M. P. Paule-Ruiz, and J. C. Nunez, "Students LMS interaction patterns and their relationship with achievement: A case study in higher education," Computers and Education, vol. 96, pp. 42-54, May 2016.

[8] N. Z. Zacharis, "A multivariate approach to predicting student outcomes in web-enabled blended learning courses," The Internet and Higher Education, vol. 27, pp. 44-53, Oct. 2015.

[9] M. A. K. Dewan, F. Lin, D. Wen, M. Murshed, and Z. Uddin, "A deep learning approach to detecting engagement of online learners," in Proc. IEEE SmartWorld/UIC/ATC/ScalCom, Guangzhou, China, 2019, pp. 1895-1902.

[10] H. Waheed, S.-U. Hassan, N. R. Aljohani, J. Hardman, S. Alelyani, and R. Nawaz, "Predicting academic performance of students from VLE big data using deep learning models," Computers in Human Behavior, vol. 104, p. 106189, Mar. 2020.

[11] R. H. Thaler and C. R. Sunstein, Nudge: Improving Decisions About Health, Wealth, and Happiness. New Haven, CT: Yale University Press, 2008.

[12] P. J. Guo and K. Reinecke, "Demographic differences in how students navigate through MOOCs," in Proc. 1st ACM Conf. Learning at Scale, Atlanta, GA, USA, 2014, pp. 21-30.

[13] Z. Tanes and Z. Cemalcilar, "Learning from SimCity: An empirical study of Turkish adolescents," Journal of Adolescence, vol. 33, no. 5, pp. 701-709, Oct. 2010.