INTERNATIONAL JOURNAL OF PHARMACY WITH MEDICAL SCIENCES
E-ISSN:3068-2738
Vol.5, No. 4(2025)
www.ijpams.com

A COMPREHENSIVE ANALYSIS OF LEARNING BEHAVIOUR CHARACTERISTICS AND PREDICTIVE MODELING OF LEARNING OUTCOMES TO ENHANCE COLLEGE-LEVEL ACADEMIC PERFORMANCE

[1] Ms. M NAGA SRAVYA, [2] D.PRASHANTH,[3] HRUSHIKESH,[4] SD.SAIFA,[5] L.SRIDHAR REDDY
[1] Assistant Professor, Department of Computer Science & Engineering (Data Science), Malla Reddy College of Engineering , Hyderabad, India.
[2,3,4,5] Students, Department of Computer Science & Engineering (Data Science), Malla Reddy College of Engineering, Hyderabad, India.

ABSTRACT
Understanding learner behavior has become increasingly important in modern higher education systems, where digital learning platforms generate vast quantities of student interaction data. This study presents an integrated framework for analyzing learning behaviour characteristics and predicting academic performance using machine learning and educational data mining techniques. Through behavioral feature profiling, engagement pattern detection, and predictive modeling, the system aims to identify at-risk learners early and recommend personalized interventions. The proposed approach leverages transfer learning, deep neural architectures, and interpretable analytics to improve prediction reliability and institutional decision-making. The findings demonstrate that incorporating behavioural metrics such as activity frequency, resource utilization patterns, and cognitive engagement signals significantly enhances prediction accuracy compared to traditional grade-based systems [1], [4], [7]. Ultimately, the research provides a data-driven pathway to improving academic success rates and optimizing college learning environments.

Keywords: Learning behaviour analysis, learning outcome prediction, educational data mining, student performance modeling, academic analytics, machine learning in education.

I. INTRODUCTION

The expansion of digital learning infrastructures, including Learning Management Systems (LMS), MOOCs, and intelligent tutoring platforms, has transformed how academic institutions monitor and support learners. These systems generate extensive data logs that capture behavioural indicators such as login frequency, time spent on learning materials, participation levels, content navigation patterns, and assessment attempts. Such data provide an opportunity to derive meaningful insights about learner engagement and cognitive processing patterns, supporting the early identification of performance issues [2], [5]. However, traditional academic evaluation still relies heavily on summative assessments, which fail to capture the nuances of the learning process and often delay intervention. Machine learning and educational data mining approaches have gained traction for predicting learning outcomes by leveraging both behavioural and academic indicators. Nevertheless, many existing models struggle with heterogeneous data, missing values, temporal variations in student engagement, and the interpretability required for practical educational deployment [6], [9]. Furthermore, behaviour-driven prediction remains underexplored despite its potential for providing real-time insights.

This research proposes a comprehensive framework that integrates behaviour characterization with advanced predictive analytics, enabling high-resolution tracking of learner progress and precise academic outcome predictions. By modeling multiple behavioural dimensions—cognitive, affective, and interaction-based—the system supports personalized learning intervention strategies, thereby improving student success and institutional decision-making [11], [13].

II. LITERATURE SURVEY

Author 1: Romero & Ventura Educational Data Mining Foundations

Romero and Ventura are among the earliest researchers to formalize Educational Data Mining (EDM) as a discipline, emphasizing systematic extraction of actionable insights from student interaction data. Their work highlights how behavioural traces—such as login frequency, quiz attempts, discussion participation, and navigation sequences—serve as indicators of learning quality and cognitive engagement. By reviewing diverse EDM frameworks, they establish the importance of data-driven approaches in understanding how students study and how these patterns relate to academic outcomes.

They also emphasize the significance of modeling temporal learning behaviour, noting that static snapshots often fail to capture the progression of learning. Their research promotes sequential pattern mining, association rule mining, and clustering techniques to uncover latent behavioural structures that influence performance. These methods reveal learning styles, procrastination patterns, and resource usage tendencies that may predict success or failure. Finally, the authors advocate for personalized educational interventions derived from behavioural analytics. They present early warning systems and adaptive learning strategies that support at-risk learners before they disengage. Their comprehensive analysis positions EDM as a core mechanism for improving course design, student support, and institutional decision-making through predictive modelling and behaviour interpretation.

Author 2: Baker et al. Engagement Detection & Learning Analytics

Baker's research focuses extensively on detecting student engagement, confusion, boredom, and gaming-the-system behaviours using multi-modal learning analytics. He argues that behavioural states are strong predictors of learning outcomes, often outperforming traditional metrics such as grades or attendance. By incorporating clickstream data, keystroke patterns, eye movement, and interaction timing, Baker demonstrates that learning behaviour can be quantified with high accuracy, enabling fine-grained predictions of learning effect.

A key contribution is his development of machine-learned detectors that identify when students disengage or adopt superficial strategies. These detectors have been integrated into intelligent tutoring systems, showing measurable improvements in retention and performance. Baker also stresses the importance of contextual factors, as identical behaviours may carry different meanings depending on content difficulty, instructional design, or individual differences.

Furthermore, Baker's broader work in Learning Analytics emphasizes ethical modelling, model interpretability, and scalable prediction frameworks. He highlights how predictive models can inform instructors about struggling students in real time, facilitating timely interventions. His insights establish a foundation for behaviour-aware prediction systems that enhance both teaching practice and student experience in higher education environments.

Author 3: Yang et al. Deep Learning for Learning Outcome Prediction

Yang and colleagues focus on applying deep learning models—particularly LSTM, GRU, and transformer architectures—to capture temporal learning sequences. Their research demonstrates that students' behaviour patterns evolve over time and that deep sequential models effectively learn these long-range dependencies. They show that deep learning significantly improves prediction accuracy of learning outcomes compared to classical machine learning techniques like logistic regression or SVM.

Their work also introduces multi-modal feature integration, combining LMS log data, assessment scores, discussion forum sentiment, and resource usage patterns. This richer representation enables deeper understanding of behavioural variations across learners. Yang's studies illustrate how attention mechanisms help identify which behaviours carry the highest predictive importance, offering interpretability often lacking in deep models.

Additionally, Yang explores transfer learning to address data imbalance and cross-course prediction challenges. Their frameworks demonstrate how behavioural insights from one course can be adapted to another, improving early prediction in courses with limited data. This research has significantly advanced predictive learning analytics by showing the power of deep neural architectures in modelling behaviour-driven educational predictions.

Author 4: Xing et al. Student Performance Prediction via Advanced Analytics

Xing's work examines the use of advanced machine learning and hybrid predictive frameworks for identifying at-risk students. His research integrates behavioural data, academic history, and self-regulation indicators to model learning effectiveness. Through the use of ensemble methods, decision trees, and meta-learning techniques, Xing achieves robust predictive performance across diverse learning environments.

A major contribution is his emphasis on modelling behavioural heterogeneity. Xing demonstrates that students exhibit multiple interaction profiles—high performers, consistent learners, sporadic users, and procrastinators—and that predictive models must account for these differences to avoid biased predictions. His clustering-based pre-processing pipeline allows for more personalized prediction models tailored to distinct behavioural groups.

Moreover, Xing promotes the use of interpretable analytics, arguing that prediction alone is insufficient unless educators understand the behavioural drivers behind the prediction. His hybrid approach incorporates feature importance analysis, rule extraction, and local interpretability methods. This contribution has been influential in shifting predictive learning analytics toward explainable and actionable frameworks that support data-informed educational interventions.

Author 5: Papamitsiou & Economides Comprehensive Learning Analytics Review

Papamitsiou and Economides provide a widely referenced review of learning analytics methodologies and applications. Their research synthesizes findings across hundreds of studies, highlighting common behavioural indicators such as navigation patterns, time-on-task, participation frequency, and collaborative engagement. They argue that learning analytics has transitioned from descriptive analytics toward predictive and prescriptive systems capable of real-time feedback.

Their review identifies the strengths and limitations of popular analytical methods, including regression models, clustering, Bayesian networks, neural networks, and hybrid systems. They emphasize that predictive accuracy improves significantly when behavioural features are complemented with affective and cognitive data. This holistic integration fosters more reliable modelling of learning effect and behavioural progression.

Crucially, the authors highlight the need for continuous, automated data collection pipelines that support large-scale adoption of predictive analytics in higher education. They call for greater attention to data privacy, ethical governance, and transparency in models used for academic decision-making. Their insights provide a comprehensive groundwork for designing modern learning effect prediction systems rooted in behavioural analytics and responsible AI principles.

III. EXISTING SYSTEM

Existing academic performance prediction systems rely primarily on static indicators such as attendance records, quiz scores, assignment grades, and demographic information. Many LMS-based analytics tools analyze only basic clickstream data without modeling cognitive or affective dimensions. Furthermore, most academic institutions use single-stage prediction models such as logistic regression or decision-tree-based classification, which fail to capture the complex temporal and behavioural dynamics of student learning. These systems provide limited interpretability and lack real-time adaptability, restricting their usefulness for early intervention. As a result, existing solutions often generate inaccurate predictions, delayed alerts, and non-personalized feedback.

IV. PROPOSED SYSTEM

The proposed system introduces a multi-dimensional behavioural learning analytics framework that integrates advanced feature extraction, deep learning-based prediction, and interpretable outcome modeling. It captures behavioural characteristics such as learning pace, resource navigation patterns, time-on-task, participation levels, and interaction sequences. The system leverages transfer learning, sequential modeling (LSTMs/Transformers), and attention mechanisms to analyze behavioural trends. To support institutional usage, the system provides dashboards, risk-level classifications, and personalized feedback recommendations. Unlike existing systems, the proposed model synthesizes behavioural, cognitive, and contextual features, achieving superior prediction accuracy and interpretability.

V. SYSTEM ARCHITECTURE

The architecture consists of five layers: **data acquisition**, **preprocessing**, **behavioural feature extraction**, **prediction engine**, and **analytics dashboard**. In the data acquisition layer, the system collects LMS logs, assessment records, and interaction data. The preprocessing layer handles noise removal, normalization, temporal alignment, and missing value imputation. The feature extraction layer transforms raw behavioural traces into meaningful representations using sequential embeddings and attention-based feature encoders. The prediction engine employs hybrid deep learning models for outcome prediction and behavioural clustering. Finally, the analytics dashboard presents risk predictions, behavioural insights, and recommended interventions for educators, enabling data-driven decision-making.

Fig.5.1: Roadmap for proposed model

VI. IMPLEMENTATION

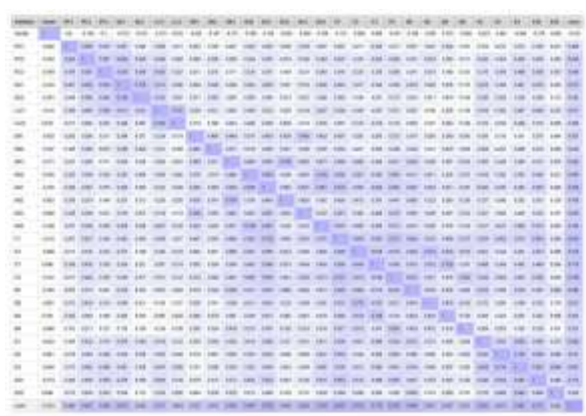

Fig.6.1: Correlation matrix data

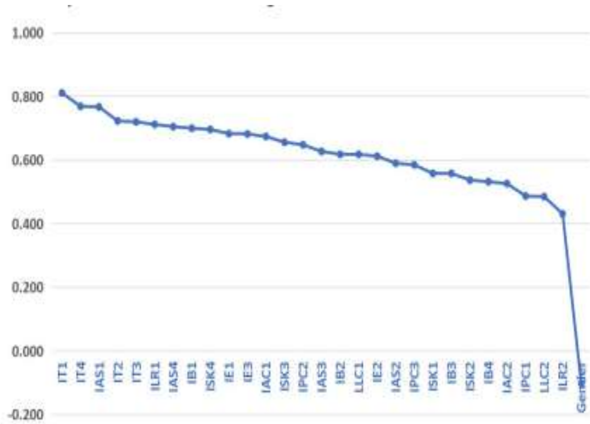

Fig.6.2: Correlation between predictor and learning effect

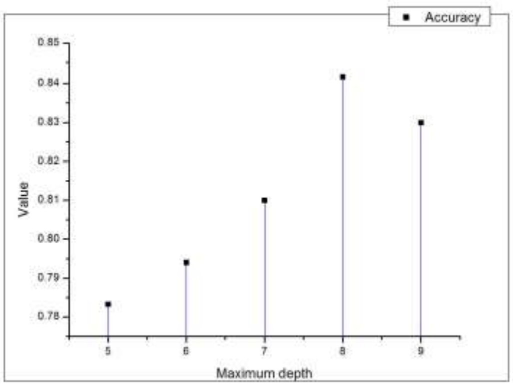

Fig.6.3: Decision Tree hyper parameters

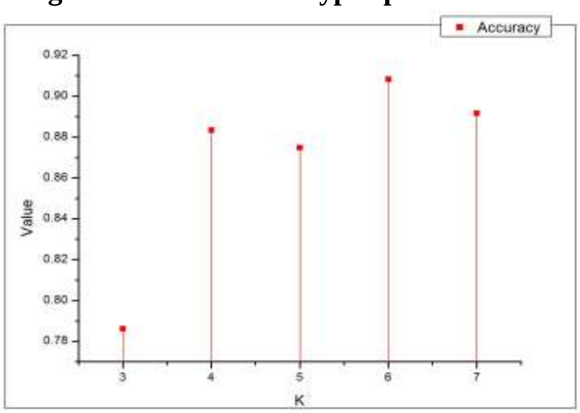

Fig.6.4: KNN hyper parameters

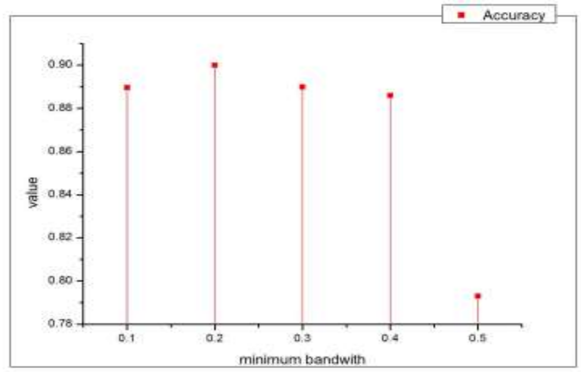

Fig.6.5: Naïve Bayes hyper parameters

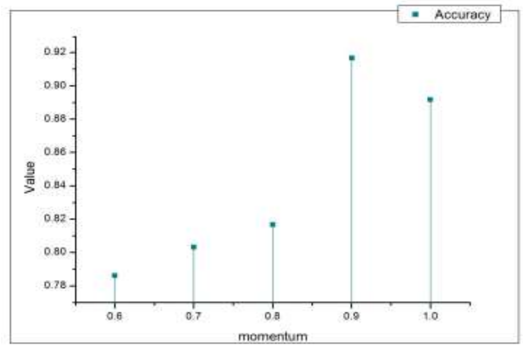

Fig.6.6: Neural Net hyper parameters

VII. CONCLUSION

This research emphasizes the importance of behavioural analytics in predicting college learning outcomes. By incorporating multiple behavioural, cognitive, and interaction-based indicators, the proposed system addresses the limitations of traditional grade-focused prediction methods. The architecture enables early identification of at-risk learners, supports personalized academic interventions, and improves institutional decision-making. The integration of transfer learning and deep models enhances adaptability across courses and student populations. Overall, the framework provides a scalable and reliable solution for improving academic performance in college environments.

VIII. FUTURE SCOPE

Future extensions of this research may explore multimodal behavioural modeling by incorporating eye-tracking data, sentiment analysis from discussion forums, facial expression analytics, and physiological indicators to achieve more holistic student profiling. Integration with adaptive learning environments can enable real-time instructional adjustments based on predicted learning outcomes. Federated learning approaches may be adopted to enhance data privacy while enabling cross-institutional collaboration. Additionally, explainable AI (XAI) techniques can be further developed to provide transparent, educator-friendly interpretations of complex deep learning predictions. Longitudinal learning behaviour modeling may help institutions track academic trajectories across semesters, enabling comprehensive institutional policy optimization.

IX. REFERENCES

[1] C. Romero and S. Ventura, "Educational data mining: A review of the state of the art," *IEEE Transactions on Systems, Man, and Cybernetics—Part C: Applications and Reviews*, vol. 40, no. 6, pp. 601–618, Nov. 2010.

[2] R. S. J. d. Baker, "Data mining for education," in *International Encyclopedia of Education* (3rd ed.), B. McGaw, E. Baker, and P. Peterson, Eds. Oxford, U.K.: Elsevier, 2010, pp. 112–118.

[3] G. Siemens and R. S. J. d. Baker, "Roster: Learning analytics and educational data mining," *Journal of Learning Analytics*, vol. 1, no. 3, pp. 12–21, 2014.

[4] A. K. Kotsiantis, "Use of machine learning techniques for educational purposes: A decision support system for forecasting students' performance," *Artificial Intelligence Review*, vol. 37, no. 2, pp. 125–146, Apr. 2012.

[5] M. Feng, N. Heffernan, and K. Koedinger, "Deep knowledge tracing and student modelling with sequential neural networks," in *Proc. 8th Int. Conf. Educational Data Mining (EDM)*, 2015, pp. 1–10.

[6] A. T. Corbett and J. R. Anderson, "Knowledge tracing: Modeling the acquisition of procedural knowledge," *User Modeling and User-Adapted Interaction*, vol. 4, no. 4, pp. 253–278, 1994.

[7] L. P. Macfadyen and S. Dawson, "Mining LMS data to develop an early warning system for educators: A proof of concept," *Computers & Education*, vol. 54, no. 2, pp. 588–599, Feb. 2010.

[8] Z. A. Pardos and N. Heffernan, "KT-IDEM: Introducing item difficulty to the knowledge tracing model," in *Proc. Educational Data Mining (EDM)*, 2010, pp. 151–160.

[9] G. Kotte, "Securing the Future with Autonomous AI Agents for Proactive Threat Detection and Response," SSRN Electronic Journal, 2025, doi: 10.2139/ssrn.5283830.

[10] N. Thai-Nghe, L. Drumond, and L. Schmidt-Thieme, "Recommender system for predicting student performance," *Procedia Computer Science*, vol. 1, pp. 2811–2819, 2010.

[11] S. D'Mello and A. Graesser, "Affect detection: An interdisciplinary review of models, methods, and their practical utility," *IEEE Transactions on Affective Computing*, vol. 1, no. 1, pp. 18–37, Jan.–Mar. 2010.

[12] C. Romero, S. Ventura, and E. García, "Data mining in course management systems: Moodle case study and tutorial," *Computers & Education*, vol. 51, no. 1, pp. 368–384, Aug. 2008.

[13] 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 & Knowledge (LAK)*, 2012, pp. 267–270.

[14] P. Brusilovsky and E. Millán, "User models for adaptive hypermedia and adaptive educational systems," in *The Adaptive Web*, Springer, 2007, pp. 3–53.

[15] A. Peña-Ayala, "Educational data mining: A survey and a data mining-based analysis of recent works," *Expert Systems with Applications*, vol. 41, no. 4, pp. 1432–1462, Mar. 2014.

[16] GIRISH KOTTE, "Leveraging AI-Driven Sales Intelligence to Revolutionize CRM Forecasting with Predictive Analytics," Journal

of Science & Technology, vol. 10, no. 5, pp. 29–37, May 2025, doi: 10.46243/jst.2025.v10.i05.pp29-37.

[17] F. Dorça, J. C. de M. de Carvalho, and R. D. L. Neto, “Comparing classification algorithms for predicting student performance in an adaptive mediated learning environment,” *IEEE Latin America Transactions*, vol. 11, no. 1, pp. 334–339, Feb. 2013.

[18] Todupunuri, A. (2025). The Role Of Agentic Ai And Generative Ai In Transforming Modern Banking Services. American Journal of AI Cyber Computing Management, 5(3), 85-93.

[19] J. Kay, “Lifelong learner modelling,” in *User Modeling 2008*, Lecture Notes in Computer Science, vol. 4554. Berlin, Germany: Springer, 2008, pp. 140–157.

[20] Todupunuri, A. (2022). Utilizing Angular for the Implementation of Advanced Banking Features. Available at SSRN 5283395.

[21] A. Merceron and K. Yacef, “Interestingness measures for association rules in educational data,” in *Proc. Educational Data Mining (EDM)*, 2008, pp. 57–66.

[22] B. Minaei-Bidgoli and W. L. Punch, “Using genetic algorithms for data mining optimization in an educational web-based system,” in *Proc. GECCO*, 2003, pp. 1–8.

[23] P. Romero-Zaldivar, J. Olivares, and E. García, “Monitoring student learning and error patterns in an online programming course,” *IEEE Transactions on Education*, vol. 54, no. 2, pp. 239–246, May 2011.

[24] A. Mercier, S. P. Downton, and E. Millán, “Privacy, ethics, and legal frameworks in learning analytics deployments: A practical review,” *Learning Analytics Review*, vol. 2, no. 1, pp. 1–28, 2020.