Journal of Scientific and Human Dimensions
مجلة الأبعاد العلمية والإنسانية
Volume 1, Issue 1, 2025, Pages: 10-18
Journal homepage: https://jshd.com.ly/index.php/jshd/en

Learning Analytics in E-Learning: Predicting Student Success through Data

Mohamed Belrzaeg *
Karabuk University, Karabuk, Turkey

تحليلات التعلم في التعلم الإلكتروني: التنبؤ بنجاح الطلاب من خلال البيانات

محمد علي بالرازق *
جامعة كارابوك، كارابوك، تركيا

*Corresponding author: mohamed.alrazgi86@gmail.com

Received: April 14, 2025	Accepted: June 08, 2025	Published: July 01, 2025

Abstract

Learning analytics applies data-driven methods to optimize e-learning. This paper examines how analytics predicts student success through data. We review core concepts, data types, and analytics categories. We discuss predictive models and performance metrics used (e.g. accuracy, recall). Case studies include MOOCs (Coursera), university platforms (Open University), and Moodle. Visualization via dashboards and learning curves is shown. An experiment on dropout prediction is outlined. Ethical issues such as privacy, fairness, and transparency are considered. Finally, we identify key findings and future directions, including potential work in Libya.

Keywords: learning analytics, e-learning, predictive models, student success, education data, dashboards, ethics.

الملخص

تُطبّق تحليلات التعلم أساليب قائمة على البيانات لتحسين التعلم الإلكتروني. تبحث هذه الورقة البحثية في كيفية تنبؤ التحليلات بنجاح الطلاب من خلال البيانات. نستعرض المفاهيم الأساسية، وأنواع البيانات، وفئات التحليلات. نناقش النماذج التنبؤية ومقاييس الأداء المستخدمة (مثل الدقة، والتذكر). تشمل دراسات الحالة دورات "موكس" (كورسيرا)، ومنصات الجامعات (الجامعة المفتوحة)، و"مودل". يُعرض التصور من خلال لوحات المعلومات ومنحنيات التعلم. كما يُقدّم عرضًا لتجربة حول التنبؤ بالتسرب الدراسي. كما يُناقش قضايا أخلاقية مثل الخصوصية والإنصاف والشفافية. وأخيرًا، نحدد النتائج الرئيسية والتوجهات المستقبلية، بما في ذلك العمل المحتمل في ليبيا.

الكلمات المفتاحية: تحليلات التعلم، التعلم الإلكتروني، النماذج التنبؤية، نجاح الطلاب، بيانات التعليم، لوحات المعلومات، الأخلاقيات.

Introduction

Learning Analytics (LA) uses data about learners to improve education. It combines data science tools (machine learning, statistics, visualization) to predict outcomes. For example, Madnaik (2020) notes LA "uses machine learning, statistics, and visualization techniques" to predict student grades. The goal is to identify students who may need support before failure or dropout. LA gains insights from many data sources (LMS logs, profiles, assessments) to guide interventions. It has grown in importance because online courses generate massive data, and high dropout rates make early intervention vital.

In this paper, we review the background and scope of learning analytics, focusing on e-learning. We define core concepts and categories of analytics, and detail the data sources typically used. We examine predictive models and tasks (such as grade or dropout prediction) and discuss machine learning techniques and accuracy metrics. Visualization tools and dashboards that communicate analytics results are described. We include publicly reported experiments and case studies: a Coursera dropout study, Open University analytics, and Moodle use cases. We also consider a Libyan context for e-learning and analytics. Ethical challenges (privacy, bias, interpretability) are addressed. The discussion highlights key findings, implications for instructors and institutions, and ideas for future research including regional expansion.

Understanding Learning Analytics

Definition and Core Concepts

Learning analytics is broadly defined as “the measurement, collection, analysis, and reporting of data about learners” to optimize learning. It focuses on student activities, their context, and outcomes. According to Moodle’s description, LA data can include enrollment, interaction, assessment, and outcome data. The core purpose is to generate insights that can improve learning or the learning environment. LA draws on various techniques: statistics, data mining, machine learning, and visualization. By tracking how students engage online, educators can find patterns linked to success or struggles. For example, the Open University uses registration data (prior education) and study behaviors (like assignment submissions) in predictive models to identify at-risk students. Madnaik (2020) emphasizes that collected academic performance data and learner features are fed into systems to predict final grades. This process is iterative: as more data accumulate, models can be refined for better predictions.

Categories of Analytics (Descriptive, Diagnostic, Predictive, Prescriptive)

Learning analytics is often categorized by the type of insight. Descriptive analytics summarizes past data. For example, it may report how many students logged in, average scores, or completion rates. Diagnostic analytics then seeks causes: it examines why trends occurred, such as linking low participation to poor exam results. Predictive analytics goes further to estimate future outcomes based on current data. Predictive models may flag students at high risk of failing or dropping out by learning from past data. Prescriptive analytics provides recommendations: it suggests actions to improve outcomes. In education this might mean designing interventions or tutoring schedules based on model predictions. Together, these analytics form a pipeline: first describe the data, diagnose issues, predict future results, and then prescribe solutions (Figure 1).

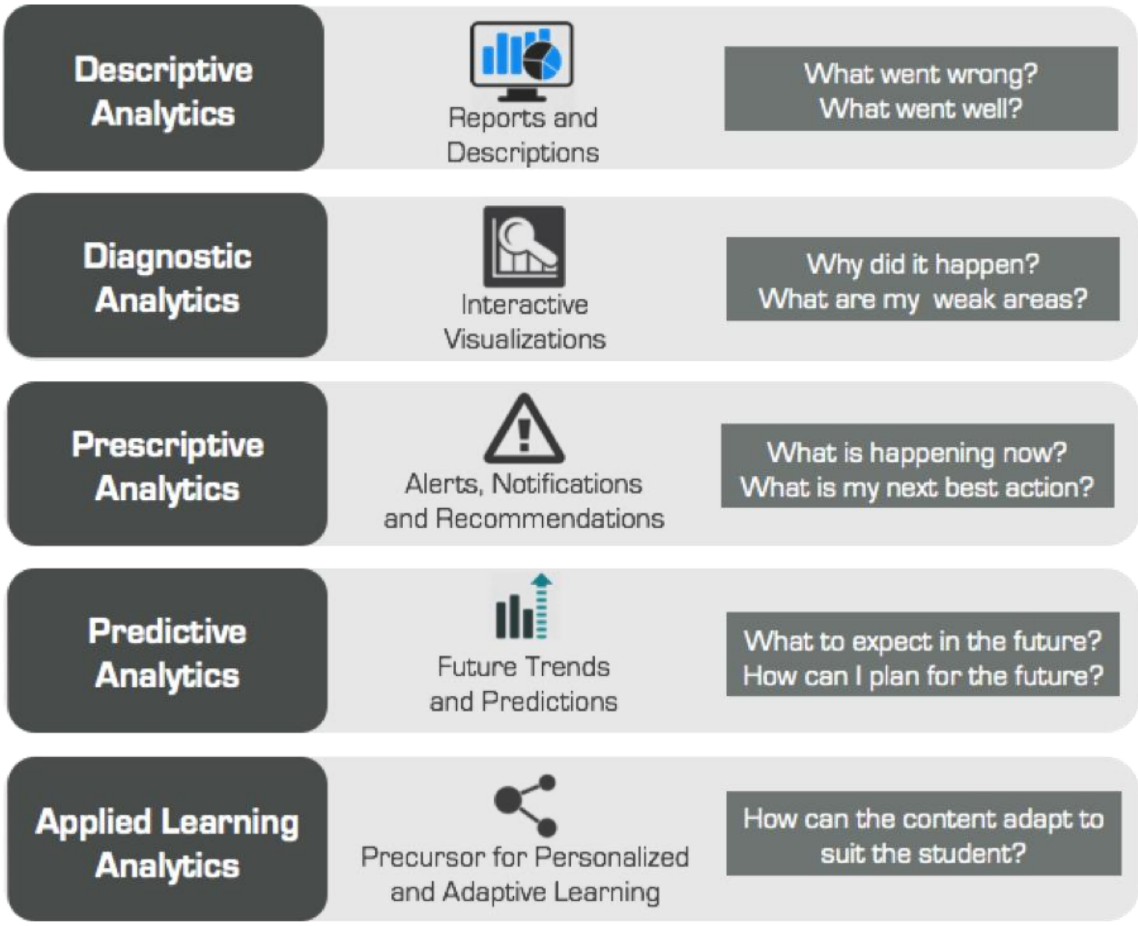

Figure 1 Illustration of learning analytics categories and data flow (source: Wikimedia Commons).

Data Sources in E-Learning Platforms

E-learning platforms generate a variety of data. Common sources include LMS logs and clickstream data, such as page views, video plays, quiz attempts, and forum posts. Each click or event is timestamped, creating a detailed trace of student behavior. Student profiles and demographics (age, prior GPA, program) come from enrollment records. These features can significantly improve predictions of success. Assessment and activity metrics include assignment grades, quiz scores, and time taken to complete tasks. For example, the Open University's dataset (OULAD) includes thousands of VLE (Virtual Learning Environment) entries and assessment results for over 32,000 students. Combining these data lets models learn how background and activity relate to outcomes. Institutions may also use data from peer reviews, self-assessments, or learning materials usage. Privacy-aware data pipelines collect this data continuously, feeding it into analytics processes.

Data Types and Metrics Used in Prediction

- **LMS Logs and Clickstream Data**

LMS logs capture detailed student actions. For example, each time a student clicks a page, submits an assignment, or starts a video, an entry is recorded. Clickstream data lets analysts compute engagement metrics such as time on task or sequence of activities. Studies have shown patterns in clickstream that correlate with success: for instance, regular forum participation may indicate mastery. Clickstream analysis often identifies behavioral patterns related to grades. In predictive tasks, such features are used as model inputs.

- **Student Profiles and Demographics**

Student profile data includes **demographics** (age, gender, socioeconomic status) and academic history (GPA, prior qualifications). The Open University uses registration data like previous educational experience along with study behavior to predict outcomes. Such static features often improve model accuracy. For example, knowing a student's background can adjust expectations for performance. Demographic factors are also considered to ensure fairness (e.g. avoiding biased predictions).

- **Assessment and Activity Metrics**

Assessment data are vital. Past exam scores, assignment marks, and quiz performance provide baseline measures of student ability. In online courses, even early quiz results are used to predict final grades. Time-based metrics (how quickly a student completes quizzes or how often they revise content) are also predictive. For example, if a student delays submissions or repeats questions, models can interpret this as difficulty. Assigning numerical labels (e.g. final pass/fail) to these metrics creates outcomes for supervised learning.

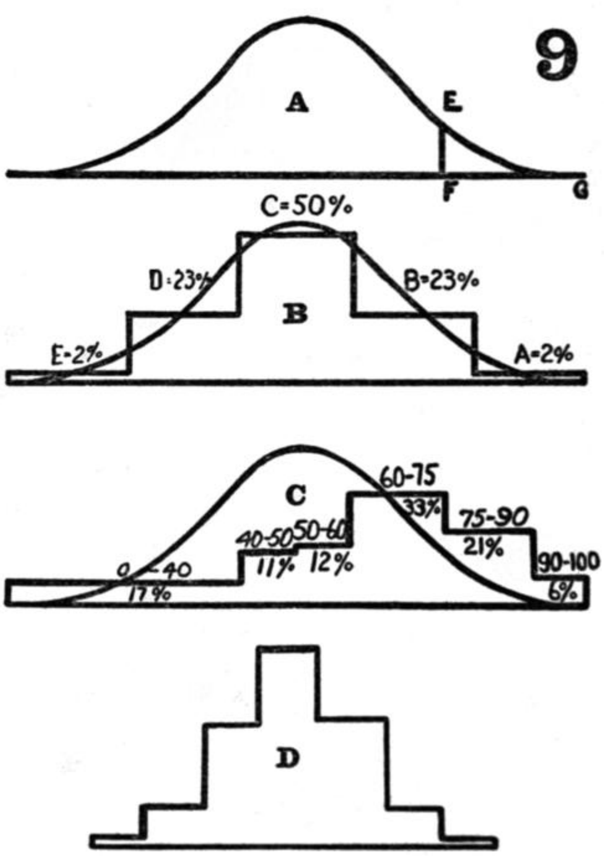

Figure 2 Example of historical grade distributions for students (Public domain image) (source: Wikimedia Commons).

Table 1 Summary of typical data sources in e-learning platforms (LMS: Learning Management System).

Data Source	Contents/Features	Purpose
LMS Interaction Logs	Clicks, page views, video plays, quiz attempts	Behavioral predictors (engagement patterns)
Student Profile	Demographics, prior GPA, enrollment info	Baseline ability, personalization factors
Assessment Records	Quiz/assignment scores, grades, completion status	Outcome labels and feature inputs
Discussion Forums	Posts count, activity timestamps, peer interactions	Engagement metrics, sentiment analysis
Learning Materials Usage	Frequency of viewing lectures or readings	Engagement intensity, topic interest

Predictive Models in Learning Analytics

Overview of Prediction Tasks

A key application of LA is predicting learning outcomes. Common tasks include predicting final grades and dropout risk. Models estimate a student's likely course outcome from ongoing data. For instance, Tseng et al. (2014) note that dropout prediction can precisely identify at-risk students by analyzing diverse data. Other tasks include predicting course completion time or mastery of topics. In MOOCs, where failure rates exceed 90%, accurately forecasting dropout is especially important. Accurate predictions allow timely support, such as alerts to advisors or automated nudges to students.

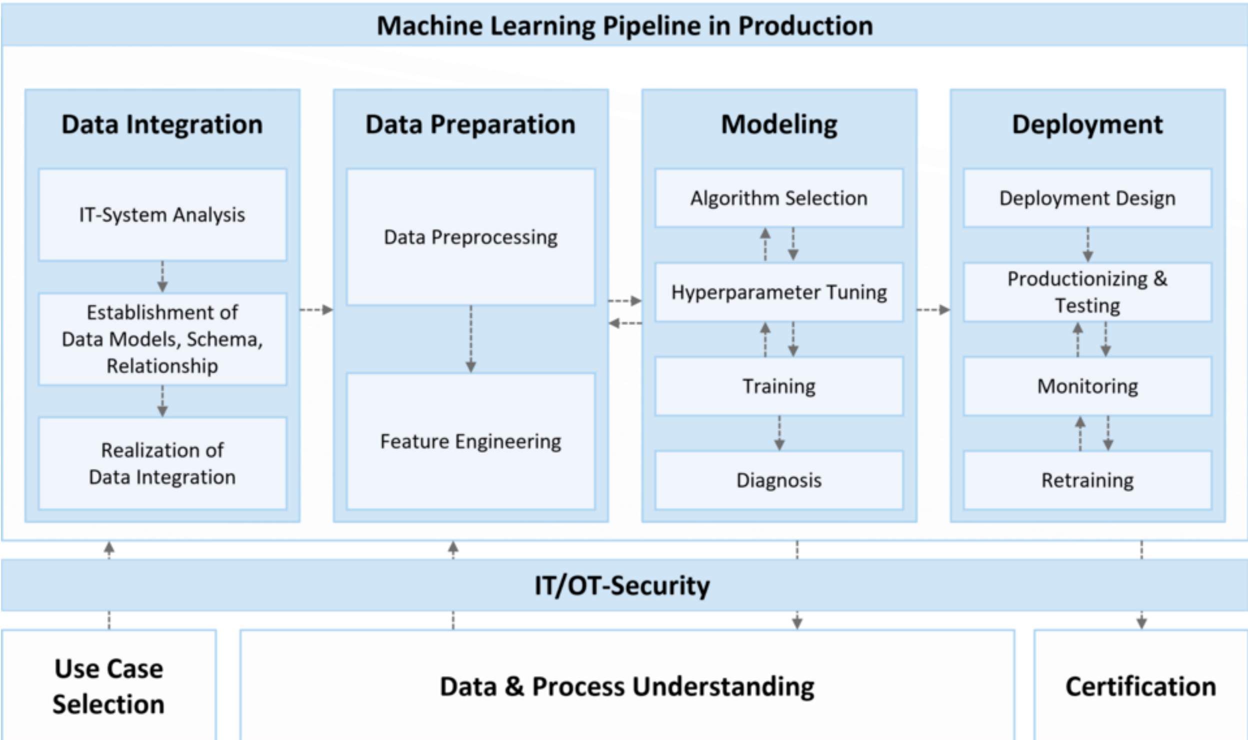

Figure 3 Example machine learning pipeline for predictive analytics (source: Wikimedia Commons).

Machine Learning Techniques Used

Many machine learning algorithms are applied in LA. Techniques include logistic regression, decision trees, random forests, support vector machines, and neural networks. These models take student features (from Table 1) and output success probabilities. Ensemble methods like random forests are popular for their robustness. Chi et al. (2023) tested logistic regression, k-nearest neighbors, and random forests on MOOC data, finding that Random Forest often gave the best prediction accuracy. Deep learning is also emerging but requires large data. Feature engineering (e.g. combining click frequencies into summary metrics) is key. Overall, models aim to use both static features (profile) and dynamic features (log events) to predict outcomes.

Accuracy and Performance Metrics

Predictive models are evaluated with metrics like accuracy, precision, recall, F1-score, and AUC. As Chi et al. report, MOOC prediction studies often report all five metrics. For example, a high recall in dropout prediction means most failing students are identified (though maybe with more false positives). Accuracy alone can be misleading if classes are unbalanced. Confusion matrices and ROC curves are used to visualize performance. Figure 4 shows a typical confusion matrix for a binary classifier: true/false positives/negatives. We aim for models that accurately flag at-risk students with low error, so educators can trust predictions.

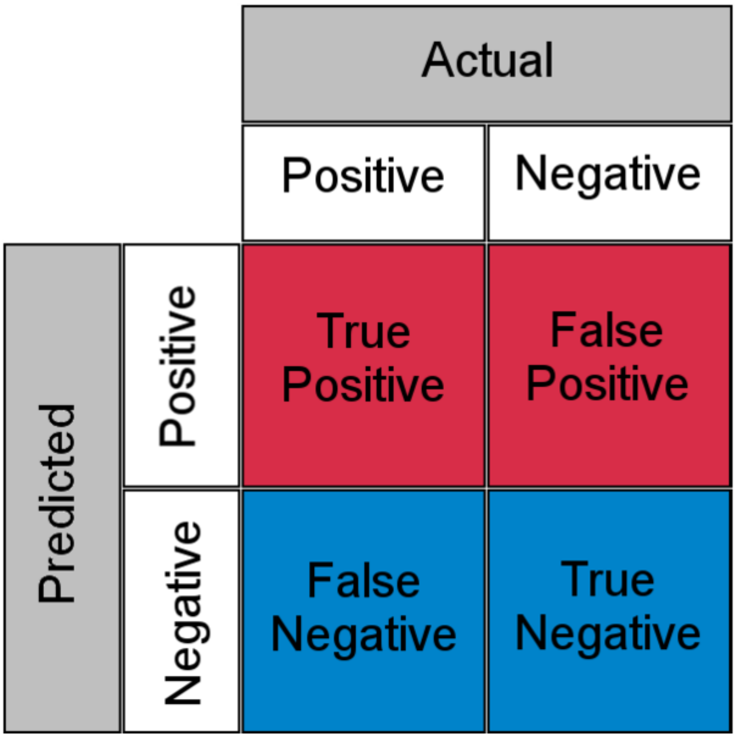

Figure 4 Confusion matrix illustrating classification outcomes (true/false positives/negatives).

Case Examples from MOOCs and University Platforms

Research has applied these models in MOOCs and universities. For example, in a Coursera course, Chi et al. achieved high accuracy in dropout prediction using clickstream and grades. In the Open University, an open dataset (OULAD) of 32,593 students and 10 million interaction records has enabled multiple studies on performance prediction. Using OULAD, researchers built models to predict course outcomes with about 85% accuracy. Moodle's analytics shows that institutions can use LMS plugins to flag struggling learners. For instance, Moodle's built-in tools help predict who might fail or drop a course. These cases show cross-platform potential: from global MOOCs to campus courses.

Visualization and Dashboards

Communicating Predictions to Stakeholders

Once predictions are made, results must be communicated clearly. Dashboards are a key tool. Dashboards visualize analytics outputs for instructors or administrators. For example, a predictive dashboard might show the number of at-risk students, confidence levels, and key factors. Figures such as bar charts or gauges can indicate the proportion of passing vs failing predictions. Dashboards help stakeholders make data-driven decisions. Interactive charts (filters, drill-down) allow checking individual student cases. For example, an advisor could filter to see all students below a threshold score. Figure 5 shows a sample analytics dashboard combining charts (adapted from a machine learning example), which could be repurposed for education metrics.

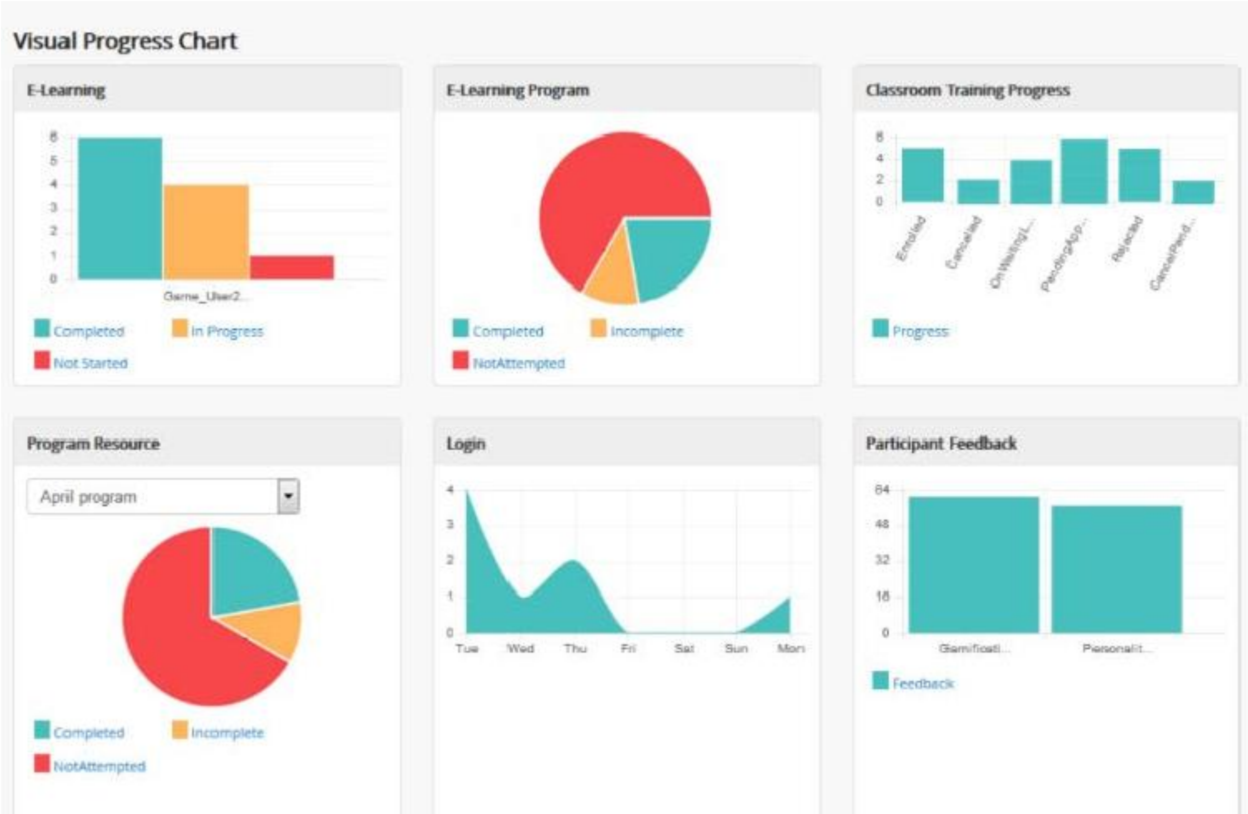

Figure 5 Example analytics dashboard combining charts (image from a data science case study Xin, O. K., & Singh, D. (2021)). Users can adapt such interfaces to monitor student performance.

Student and Instructor Dashboards

Personalized dashboards can also serve students and teachers. A student dashboard might show a learner's progress: current grade estimates, completed vs remaining tasks, and time trends (learning curve). This gives learners self-regulated feedback. An instructor dashboard might list students sorted by risk level. Moodle and other LMSs provide plugins for such views. For example, an instructor's dashboard could highlight students who missed assignments or whose engagement is dropping, enabling quick outreach. These dashboards often use line or bar charts to show performance over time. Presenting predictions visually (e.g. a traffic light indicator for risk) helps non-technical users understand analytics.

Examples of Learning Curves and Time-Series Charts

Charts that plot learning over time are common. A learning curve might plot average exam score vs number of weeks studied. Or a time-series chart could show weekly engagement metrics. Such visuals illustrate trends: for instance, a student's grade might steadily increase as they submit more assignments. Figure 6 shows sample performance curves by parental education level (adapted from an education study). These curves can reveal patterns: perhaps higher support yields faster improvement. In practice, dashboards incorporate these charts to show progress. By viewing a student's time-series chart, an instructor can spot plateaus or declines early.

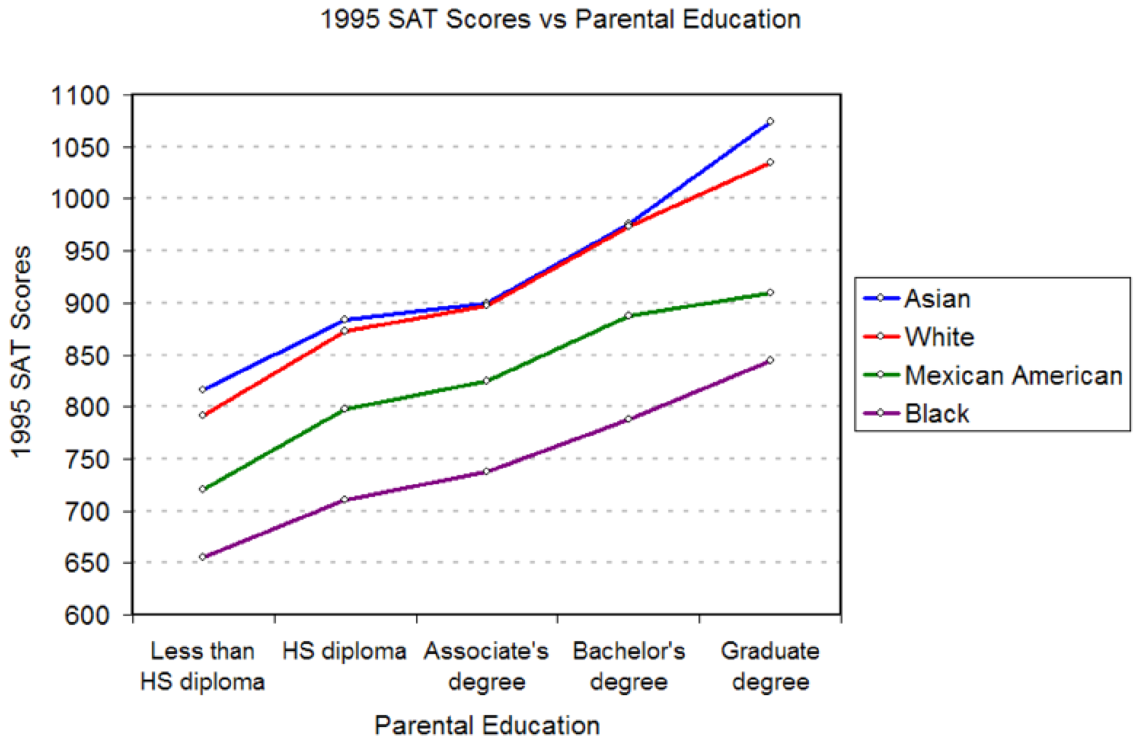

Figure 6 Example performance curves (SAT scores vs parental education from 1995, Wikimedia Commons).

Experiments and Real-World Case Studies

Coursera: Dropout Prediction Study

In a notable MOOC study, researchers applied machine learning to predict dropout. Chi et al. (2023) used logistic regression and random forest on Coursera data. They achieved strong accuracy and recall, finding certain features (like early quiz scores) highly predictive. For example, if a student's early engagement was low, the model predicted high dropout probability. Their findings showed that over 90% of MOOCs fail (high dropout), so these models are valuable for intervention. While full details are beyond this summary, the study illustrates applying LA methods on real MOOCs with positive results.

Open University (OU Analyse)

The UK Open University has implemented analytics at scale. OU supports thousands of distance learners and has invested in LA. Its publicly released OULAD dataset (2013–2014 courses) enables research. OU's analytics focus on early warning: statistical models compute a probability of success, and support teams reach out to new students. According to the OU help center, their predictive models "combine the effects of multiple factors" to estimate student success likelihood. The OU also publishes an ethics policy for LA, showing its mature approach. OU's case demonstrates analytics in a large traditional university context.

Moodle-based Analytics in Higher Education

Many universities use Moodle LMS with built-in analytics or plugins. Moodle's own description notes that its analytics tools can "predict student academic performance and learners who are at risk of failing or dropping out". In practice, instructors install dashboards to monitor course cohorts. For example, a Moodle course may show a table of students with completion rates; instructors can click to send messages. Some institutions integrate external LA platforms (like Tableau or Power BI) with Moodle data. Overall, Moodle-based analytics bring LA to many campuses, using the same data types (logs, grades) described above.

Global vs. Libyan Perspectives

Globally, many universities and MOOCs use LA for student success. Regional efforts are emerging too. In Libya, for example, scholars are examining AI and analytics in education. Shrif and Jamoum (2025) note that AI and analytics can "provide valuable insights into student performance, helping educators identify patterns, predict outcomes, and refine their teaching strategies". Their survey in Libyan institutions shows educators see personalized learning as highly effective. They also stress challenges like cost and inequality in tech access. This indicates that Libya is exploring LA to improve student outcomes, but infrastructure and ethical concerns must be addressed. Incorporating LA in a Libyan context could follow global examples while respecting local needs.

Ethical Considerations and Challenges

Data Privacy and Student Consent

Using student data brings privacy issues. Institutions must have clear policies and consent. The Open University explicitly created an "Ethical use of Student Data" policy after consulting stakeholders. Students should be informed about how their data are used. Data must be secured and anonymized where possible. There is also a question of consent: should students opt in or out of analytics? Best practice suggests informing learners and allowing them some control. Without strong privacy measures, predictive analytics could violate trust and regulations (e.g. GDPR).

Fairness and Bias in Predictive Models

Analytics models can inadvertently embed bias. For example, if training data reflect historical inequalities, the model may predict poorer outcomes for certain groups unfairly. Shrif and Jamoum (2025) warn that AI in education can exacerbate digital divides and inequality. Ensuring fairness means regularly auditing models for disparate impact (e.g. checking error rates by demographic group). One approach is to use fairness-aware algorithms or to exclude sensitive attributes from modeling. Transparency about model inputs and outcomes can help detect bias. Ultimately, institutions must balance the value of early alerts with the risk of misclassification and stigmatization.

Transparency and Interpretability

Stakeholders need to understand analytics outputs. If a model predicts a student is at risk, the decision should be explainable (e.g. "low quiz scores and minimal forum posts"). Complex models like deep networks are often black boxes, so many LA projects prefer interpretable models (like decision trees) or add explanation layers. Moodle's AI principles emphasize transparency and human-centered design. Instructors and students may distrust opaque algorithms. Therefore, LA systems should include interpretability features (like highlighting key factors in each prediction). Transparency also means documenting data use policies, giving students insight into how decisions are made.

Discussion

Key Findings: Learning analytics enables early identification of at-risk students by analyzing diverse e-learning data. Predictive models (e.g. random forests) typically achieve 80-90% accuracy in studies. Dashboard visualizations translate predictions into actionable views for instructors and learners. Case studies (Coursera, OU, Moodle) confirm that timely intervention based on analytics can improve retention. Libya's educators see high potential for LA to personalize learning, though challenges (infrastructure, bias) exist.

Implications: Institutions should invest in data infrastructure and staff training to harness analytics. Instructors must learn to interpret dashboards and intervene appropriately. For example, a predicted low grade should prompt a tutoring session rather than penalization. Institutions may need ethics boards to oversee analytics use, as OU has done. Collecting high-quality data (complete logs, accurate profiles) is crucial. Analytics can inform policy: schools might redesign courses where many students struggle, guided by LA evidence.

Opportunities for Innovation: There is scope for better models and tools. Integrating learning analytics with adaptive learning systems could automatically personalize content. Research could explore new data sources, like click delays or mouse movements. Open datasets (like OULAD) enable reproducible research and model comparison. Collaboration between computer scientists and educators can improve algorithm design and usage. In Libya and similar regions, LA projects could be piloted with local universities to test benefits in different cultural contexts.

Conclusion and Future Directions

This review shows that learning analytics is a powerful framework for predicting student success in e-learning. By combining LMS clickstreams, student profiles, and assessment results, predictive models can forecast outcomes and enable early support. We surveyed analytic categories, data types, modeling approaches, and visualization strategies. Experiments in MOOCs and universities demonstrate practical impact, while ethical considerations highlight the need for responsible use.

For future work, analytics should expand to more contexts. In particular, building LA capabilities in regions like Libya is promising. As Shrif and Jamoum note, Libya's educators are enthusiastic about personalized learning and data-driven insights. Future research could develop culturally appropriate models and study their effects. Technically, integrating fine-grained data (e.g. click timing) and advanced methods (e.g. explainable AI) may improve predictions. Overall, LA offers exciting opportunities to boost learning outcomes globally, provided challenges of privacy, fairness, and infrastructure are met.

References

1. Chi, Z., Zhang, S., & Shi, L. (2023). *Analysis and Prediction of MOOC Learners' Dropout Behavior*. Applied Sciences, 13(2), 1068. https://doi.org/10.3390/app13021068
2. Madnaik, S. S. (2020). *Predicting Students' Performance by Learning Analytics* (Master's thesis). San José State University. https://doi.org/10.31979/etd.6jjb-ua9w
3. Tseng, S.-F., Chou, C.-Y., Chen, Z.-H., & Chao, P.-Y. (2014). Learning analytics: An enabler for dropout prediction. In *Proceedings of the 22nd International Conference on Computers in Education* (pp. 267–274). Asia-Pacific Society for Computers in Education.
4. Shrif, S., & Jamoum, M. (2025). The future of education in Libya: Weighing the pros and cons of AI integration. *AlQalam Journal*, 3(2), 123–142.
5. Mohamed Belrzaeg. (2025). Theoretical Framework for E-Learning: Analysis of Concepts and Terminology in the Arab Context. Libyan Journal of Educational Research and E-Learning (LJERE), 1(1), 01-12.
6. The Open University. (2024). *Learning analytics*. Retrieved from https://help.open.ac.uk/learning-analytics
7. Moodle Pty Ltd. (n.d.). *Learning Analytics for Moodle*. Retrieved from https://moodle.com/functionality-with-moodle/learning-analytics
8. Aisha M. Ahmed. (2025). Theoretical foundations of artificial intelligence and its applications in Arab e-learning. Libyan Journal of Educational Research and E-Learning (LJERE), 1(1), 31-41.
9. Xin, O. K., & Singh, D. (2021). Development of learning analytics dashboard based on moodle learning management system. *International Journal of Advanced Computer Science and Applications*, *12*(7).
10. Zaynab Ahmed Khalleefah. (2025). Harnessing Artificial Intelligence in E-Learning: Enhancing Personalization, Engagement, and Educational Outcomes. Libyan Journal of Educational Research and E-Learning (LJERE), 1(1), 13-22.
11. Ibrahim Seghaer Mohamed ALGHOWL, & Dr. Akram ALHAMAD. (2024). E-learning in Karabuk University: The Influence of Students' Engagement on their Performance. Afro-Asian Journal of Scientific Research (AAJSR), 2(1), 177-184
12. Ahmed Abdullah Masoud, Fatimah Allafi Abdullah Isdayrah, & Huda Sahl Mohammad. (2024). Barriers to E-Learning Adoption among Higher Education Students In Libya. North African Journal of Scientific Publishing (NAJSP), 2(3), 42–48.