Research on the Application of Big Data Analytics in University Education Management

Yuanji Wu
Chongqing Polytechnic University of Engineering, Chongqing, China

Abstract: Universities now generate large volumes of administrative and behavioural data, and education management is expected to make use of it rather than leave it in separate systems. This study examines how big data analytics can be applied to university education management and what conditions make such applications work. Drawing on recent empirical studies and classic analytics literature, the paper first describes the data foundations of education management, covering student behaviour records, teaching process data, and administrative transaction data. It then analyses three application scenarios in detail: student profiling and academic early warning, teaching quality assessment, and resource allocation for administrative services. Evidence from published case studies shows measurable effects, such as improved freshman education scores after a big data based management information system was introduced, and association rule mining results that support enrolment and teaching evaluation decisions. The paper further identifies four implementation conditions, including a unified data platform, explicit data governance rules, methods matched to management questions, and staff who can translate analysis results into action. It also discusses difficulties related to data silos, privacy protection, and analytical capability, and proposes corresponding countermeasures. The study concludes that big data analytics changes education management from experience-based judgement into evidence-based decision making, and that the change depends on technology, institutions, and people advancing together rather than on any single element.

Keywords: Big Data Analytics; Higher Education Management; Learning Analytics; Data-Driven Decision Making; Education Administration

1. Introduction

Universities have accumulated far more data than their management routines can digest. Course selection records, library entries, network logs, examination results, online learning traces, and administrative transactions are stored in dozens of systems, most of which were built separately and used separately. Management decisions in many institutions still rely on periodic reports and the personal judgement of experienced administrators, while the underlying data sits idle. At the same time, higher education is under pressure to improve teaching quality, allocate limited resources fairly, and respond to student needs promptly. The gap between available data and management capability has therefore become a practical problem rather than a technical curiosity.

Research on analytics in higher education has developed along two related lines. Siemens and Long distinguished learning analytics, which uses data about learners and their contexts to understand and improve learning, from academic analytics, which applies data at the institutional level to support management decisions [1]. Picciano traced how these practices evolved in American higher education and noted that operational and academic analytics had already entered planning processes by the early 2010s [2]. Daniel summarised both the opportunities and the challenges, pointing out that institutions must deal with data integration, staff skills, and organisational change before analytics can influence management [3]. Ethics has been a parallel concern: Slade and Prinsloo argued that students should be treated as participants in the data produced about them, which imposes obligations on how institutions collect and use such data [4]. Gašević, Dawson, and Siemens added a methodological caution, namely that the value of analytics lies in its grounding in learning theory and institutional context, not in the volume of data collected [5].

Recent empirical work has moved from

Copyright @ STEMM Institute Press

advocacy to measurable application. Hassna applied a modified value chain model to show how early adopter institutions used big data and analytics across primary activities of higher education, from recruitment to alumni relations [6]. Zhai constructed a big data based education management information system, evaluated it with a fuzzy comprehensive evaluation model, and applied it to teaching effectiveness assessment, teacher workload assessment, scholarship evaluation, and dormitory management, reporting freshman education scores between 4.31 and 4.53 on a five point scale [7]. Jia combined an improved Apriori algorithm with a PSO optimized prediction model for education management data and found that digital innovation in management correlates positively with the quality of student development [8]. Yang and colleagues used association rule mining to support enrolment, learning behaviour analysis, and teaching evaluation in three colleges, where measured teaching levels clustered around a medium score of 3.06 and identified one institution scoring 4.13 as a benchmark for peer learning [9]. This paper builds on that body of work. It organises the application of big data analytics in university education management into data foundations, application scenarios, and implementation conditions, and then discusses the main difficulties and countermeasures. The rest of the paper is arranged as follows: Section 2 analyses data foundations, Section 3 presents application scenarios, Section 4 discusses the technical path, Section 5 examines challenges, and Section 6 concludes.

2. Data Foundations of University Education Management

2.1 Data Sources and Their Characteristics

Education management data comes from three broad sources. The first is behavioural data generated by students in daily activities, including course platform logs, library seat usage, canteen consumption records, campus network access, and attendance signals. These records are fine grained, time stamped, and continuous, which makes them suitable for detecting changes in study habits long before grades drop. The second is teaching and learning process data, such as assignment submissions, examination item responses, classroom interaction records, and course evaluation questionnaires. These data connect directly to the core mission of teaching. The third is administrative transaction data, covering registration, financial aid, housing, and employment files maintained by functional departments. Compared with the first two types, administrative data is more standardised but often locked inside department specific systems with inconsistent coding.

Three characteristics complicate the use of these data. Heterogeneity: data structures differ across systems, and the same student may appear under different identifiers. Longitudinality: the value of most indicators emerges from change over time, so single semester snapshots reveal little. Sensitivity: behavioural data describes minors and young adults in ordinary life, so collection boundaries and access rules matter as much as analytical methods [4]. Any realistic application plan must start from these characteristics instead of assuming a clean, unified data lake.

2.2 From Experience-Based Judgement to Evidence-Based Management

Traditional education management relies on aggregate statistics and layered reporting. A department reports failure rates after examinations; a student affairs office learns about a dropout risk when a counsellor notices repeated absence. The information chain is long, delayed, and lossy. Data driven management shortens the chain by letting indicators surface directly from raw records. The distinction made by Siemens and Long is useful here: learning analytics helps teachers and advisors act on individual student data, while academic analytics helps administrators judge programmes, courses, and resource use at institutional level [1]. Both levels matter for education management, and they share the same underlying data infrastructure. What changes is not merely the speed of information but the position of management in the process: instead of reacting to reported outcomes, administrators can monitor leading indicators and intervene while a problem is still small.

3. Application Scenarios of Big Data Analytics in Education Management

3.1 Student Profiling and Academic Early Warning

Student profiling integrates academic, behavioural, and consumption data into a

 Copyright @ STEMM Institute Press

readable description of each student's campus life, from which management can identify risk and opportunity. The typical technical route combines descriptive indicators with predictive models: grade point trends, library and online learning frequency, and attendance patterns enter a model that estimates the probability of academic difficulty in the coming term. Jia's study illustrates the predictive side, where a particle swarm optimized support vector model achieved relative prediction errors below 0.03 on education management data, and an improved association rule algorithm produced about twenty interpretable rules under a support threshold of 0.22 and a confidence threshold of 0.50 [8]. Interpretable rules matter in practice: an advisor can act on a statement such as "low library frequency combined with failed prerequisite courses predicts probation risk", but not on an opaque score. Early warning then becomes a closed loop, in which identified students receive targeted tutoring within the term rather than remedial notices after it.

3.2 Teaching Quality Assessment and Improvement

Teaching quality assessment benefits from multi source evidence rather than a single end of term questionnaire. Course platform data shows engagement; item level examination data reveals whether specific knowledge points were mastered; evaluation questionnaires supply student perception; peer observation adds professional judgement. Zhai's e-MIS case demonstrates how a big data platform can process such evidence at institutional scale and support evaluation with a fuzzy comprehensive evaluation model, applied to teaching effectiveness, teacher workload, scholarship decisions, and dormitory management [7]. The reported freshman education scores, ranging from 4.31 for life guidance education to 4.53 for learning culture education, show how numeric feedback locates both strengths and weak links in a programme. Yang and colleagues reached a similar conclusion through association rule mining on teaching evaluation data: the overall teaching level of the surveyed institutions was 3.06, and the institution scoring 4.13 became a reference case whose classroom practices could be transferred to peer institutions [9]. Assessment thus shifts from ranking to diagnosis and improvement.

3.3 Enrolment, Resource Allocation, and Administrative Services

The third scenario applies analytics to management decisions outside the classroom. Enrolment planning can use historical score distributions and major level application data; Yang et al. derived suggestions for admitting candidates in the 451 to 580 point band and for expanding capacity in specific majors based on mined admission data [9]. Resource allocation uses consumption and usage records to match supply with demand, for example adjusting library opening hours to occupancy patterns or balancing dormitory assignments against measured living patterns. Administrative services improve through the same platform: Zhai's system moved scholarship evaluation and dormitory management onto data driven procedures, which reduced discretionary space and shortened processing time [7]. Hassna's value chain view frames these cases systematically: analytics improves each primary activity of the institution, and the accumulated improvements change the overall value creation of higher education rather than isolated tasks [6]. Table 1 summarises the scenarios, their data sources, and typical outputs.

Table 1. Application Scenarios of Big Data Analytics in University Education Management

Application Scenario	Main Data Sources	Typical Analytical Output
Student profiling and early warning	Learning platform logs, attendance, consumption records	Risk indicator rules and per student profiles for advisors
Teaching quality assessment	Course platform data, item level results, questionnaires	Diagnostic reports locating weak knowledge points and courses
Enrolment and major planning	Application records, score distributions, employment data	Admission suggestions and capacity adjustment plans
Resource and service allocation	Consumption, space usage, and service request records	Matching of opening hours, housing, and support services to demand

4. Implementation Conditions and Technical Path

4.1 Platform Construction and Data Governance

Application scenarios cannot run on scattered systems. The first condition is a unified data platform that collects, cleans, and stores data from academic, behavioural, and administrative sources under common coding rules. Zhai's e-MIS design shows the practical shape of such a platform: a central system processes education administration and school management data together, and evaluation models operate on the integrated result [7]. The second condition is data governance. Universities need explicit rules on data ownership, access levels, retention periods, and quality responsibilities, otherwise the platform reproduces existing silos under a new name. Governance also covers the ethical boundary drawn by Slade and Prinsloo: students are participants in the data process, so consent, transparency, and the right to contest interpretations belong in the governance document rather than in technical manuals [4]. Experience from early adopters suggests that governance written before platform construction costs little, while governance retrofitted after a data incident costs credibility.

4.2 Matching Analytical Methods to Management Questions

Method selection should start from the management question, not from the algorithm inventory. Descriptive statistics and visual dashboards answer questions about current state; association rule mining produces interpretable rules for counselling and policy; regression and classification models support prediction when sufficient labelled history exists; fuzzy evaluation handles assessment dimensions that are inherently graded rather than binary, as in teaching quality scoring [7, 8, 9]. Two practical rules follow. First, interpretability outranks marginal accuracy in management settings, because decisions must be explained to departments, students, and auditors. Second, models must be validated on the institution's own data before deployment; parameters reported in published cases provide reference points, not guaranteed results. Gašević and colleagues' caution applies directly: without connection to educational theory and institutional context, even technically sound models produce numbers that managers cannot act on [5].

5. Difficulties and Countermeasures

5.1 Data Silos, Quality, and Capability

Three difficulties recur across cases. Data silos persist because departments treat their records as departmental assets; the countermeasure is institutional, requiring top level coordination that assigns integration tasks and rewards data sharing rather than hoarding. Data quality problems, such as missing identifiers and inconsistent coding, undermine models quietly; the countermeasure is to embed quality checks at entry points and to run periodic quality audits with published results, so that cleaning becomes routine maintenance rather than a project. Analytical capability is uneven, and most universities cannot hire large data teams; the realistic countermeasure is a small central team that maintains the platform and models, paired with trained management staff in each department who commission analyses and interpret results. Daniel's summary of challenges remains accurate in this respect: organisational change, not tooling, is the binding constraint [3].

5.2 Privacy Protection and Ethical Boundaries

Behavioural data makes students visible in ways they did not anticipate, and management applications can drift into surveillance if left ungoverned. Countermeasures are concrete. Collection should be minimised to data with a stated management purpose; access should follow role based permissions with logging; profiling outputs should inform human decisions rather than replace them, keeping an advisor or committee in the loop for consequential actions such as academic warnings or aid adjustments. Slade and Prinsloo's principle of students as participants can be operationalised through notices that explain what is collected and how it is used, and through channels for students to query and correct their profiles [4]. Institutions that treat these measures as design requirements avoid the reputational damage that follows poorly governed data use, and they preserve the trust on which data collection depends.

5.3 Sustaining Use After the Pilot Phase

Many analytics initiatives perform well as pilots and fade afterwards. The reasons are organisational: the champion moves on, the dashboard loses its audience, and departments return to old routines. Sustaining use requires attaching analytics outputs to existing management cycles, such as weekly student affairs meetings, semester teaching review sessions, and annual resource planning, so that

Copyright @ STEMM Institute Press

reports have scheduled readers and assigned follow ups. It also requires measuring the analytics system itself, for example by tracking how many early warnings led to interventions and what happened afterwards, which turns the platform into an object of continuous improvement rather than a fixed installation.

6. Conclusion

This study examined the application of big data analytics in university education management across data foundations, application scenarios, and implementation conditions. Three conclusions follow. First, the data foundations exist in most universities, but their value is released only through a unified platform and explicit governance, because heterogeneity, longitudinality, and sensitivity shape what analytics can legitimately do. Second, the application scenarios with clearest evidence are student profiling and academic early warning, multi source teaching quality assessment, and data supported enrolment and resource decisions, and published cases report concrete effects ranging from prediction errors below 0.03 to measurable improvements in freshman education scores. Third, the decisive conditions are organisational rather than algorithmic: interpretability of outputs, privacy protection designed into the system, departmental capability, and routines that keep analysis connected to management cycles. For universities beginning this work, the sensible path is to select one or two scenarios, build the minimum platform those scenarios require, and let governance and capability grow with the applications. Big data analytics will not replace educational judgement, but it gives that judgement a timely and testable evidence base, which is what modern education management demands.

References

[1] Siemens G, Long P. Penetrating the fog: Analytics in learning and education. EDUCAUSE Review, 2011, 46(5): 30-40.

[2] Picciano A G. The evolution of big data and learning analytics in American higher education. Journal of Asynchronous Learning Networks, 2012, 16(3): 9-20.

[3] Daniel B. Big Data and analytics in higher education: Opportunities and challenges. British Journal of Educational Technology, 2015, 46(5): 904-920.

[4] Slade S, Prinsloo P. Learning analytics: Ethical issues and dilemmas. American Behavioral Scientist, 2013, 57(10): 1510-1529.

[5] Gašević D, Dawson S, Siemens G. Let's not forget: Learning analytics are about learning. TechTrends, 2015, 59(1): 61-68.

[6] Hassna G. Big Data and Analytics to Transform Higher Education: A Value Chain Perspective. Perspectives: Policy and Practice in Higher Education, 2023, 27(4): 160-168.

[7] Zhai H. Innovation and Practice of Educational Management Path Building in Colleges and Universities in the Context of Big Data. Applied Mathematics and Nonlinear Sciences, 2024, 9(1).

[8] Jia Y. Digital innovations in higher education management and mechanisms for high-quality training of students. Applied Mathematics and Nonlinear Sciences, 2024, 9(1).

[9] Yang S, et al. Big Data Enables Intelligent Management of College Education. Applied Mathematics and Nonlinear Sciences, 2024, 9(1).