© AUG 2026 | IRE Journals | Volume 10 Issue 2 | ISSN: 2456-8880
DOI: https://doi.org/10.64388/IREV10I2-1722647

Applied Artificial Intelligence for Predictive Student Performance and Early Academic Intervention in Saudi Secondary Schools

ABDUL REHMAN BAROK
ORCiD: 0009-0007-4548-0107

Abstract- Predictive analytics in the field of education has moved from providing retrospective descriptions to the use of applied artificial intelligence systems that are able to identify academic at-risk students while there is still time to take action. This review looks at the ways in which that shift can be implemented in Saudi secondary schools, stressing that early identification will only be worthwhile if prediction is linked to fair, interpretable and privacy-respecting interventions. It draws together research published between 2020 and 2025 on educational data mining, machine learning, deep learning, early warning systems, explainable artificial intelligence, and the adoption of artificial intelligence in K-12 education, as well as Saudi policy documents relating to the development of human skills and personal data. Instead of focusing on which classifier achieves the highest individual accuracy, the analysis addresses five implementation questions: which predictors are actionable, when predictions become reliable enough to trigger intervention, how model effectiveness should be validated, how teachers and counselors should make use of the explanations, and what kind of governance is needed regarding data from minors. It is consistently found that previous achievement and continuous assessment are strong predictors, while attendance, engagement, traces from the learning-management system, and certain contextual variables can provide additional useful information. Although ensemble learning and deep models might improve discrimination in certain datasets, their usefulness depends on temporal validation, calibration, impartiality across subgroups, interpretability, and incorporation into everyday operations. A framework customized to the Saudi context is suggested in which school data go through privacy and quality checks, the models produce calibrated risk estimates, the explanations assist professional judgment, and tiered interventions are recorded and evaluated. The review concludes by stating that the most defensible aim of applied artificial intelligence in secondary education should not be automated student classification, but rather earlier, more consistent and auditable human support.

Keywords: educational data mining; predictive analytics; early warning systems; machine learning; secondary education; Saudi Arabia; explainable AI; academic intervention

I. INTRODUCTION

Secondary school is a critical stage where academic choices increasingly influence graduation, post-secondary opportunities, skill development, and future employment. In Saudi Arabia, this stage corresponds with national strategies focused on future-ready skills, enhanced educational results, as well as alignment with emerging economic needs. The Human Capability Development Program focuses on education, skills, lifelong learning, and global competitiveness within Vision 2030 (Kingdom of Saudi Arabia, 2021). In this context, applied artificial intelligence can help schools identify emerging problems before they become difficult to address.

Educational data mining and learning analytics deliver the foundation for early recognition of student risk. Since 2020, research on predicting student performance, identifying at-risk students, forecasting dropouts, and supporting academic decisions has expanded (Alyahyan & Düştegör, 2020; Namoun & Alshanqiti, 2021; Albreiki et al., 2021; Batool et al., 2023). Typically, historical academic or behavioral data are transformed into features, used to train predictive systems, and then applied to assign risk labels or rankings. However, having a prediction pipeline does not guarantee an effective intervention system. Even technically robust models may be educationally inadequate if they detect risk too late, rely on non-actionable variables, provide poorly calibrated probabilities, perform inconsistently across student groups, or present results in ways educators do not trust.

The case study on Saudi Arabia carried out by Alghamdi and Rahman provides a important starting point for local research. It used the academic and demographic features of secondary-school leavers from Al-Baha, applied the Random Forest, Naive Bayes, and J48 models, dealt with the problem of class imbalance using SMOTE, employed correlation-based feature selection, and assessed the models by means of cross-validation and direct partitioning (Alghamdi & Rahman, 2023). The study found that data from Saudi secondary schools are capable of supporting high-performing predictive systems and identified semester grades as among the most important factors. At the same time, it points out the further research issue of shifting from retrospectively predicting final attainment to providing prospective, time-sensitive and accountable support throughout the school years.

This distinction is fundamental to applied artificial intelligence; early warning systems should not only tackle the question of "who may struggle?" but also answer "how early can the risk be estimated", "why has the system raised the alert?", "what response is appropriate", and "does that response lead toward improved outcomes?" Research into early intervention has demonstrated that risk estimates can be produced at various stages of a student's progress and that the timing of the prediction has an impact on the actual opportunity to intervene (Adnan et al., 2021; Alturki & Alturki, 2021). Likewise, research concerning trustworthy early-warning infrastructure stresses that a model must be incorporated into a repeatable workflow rather than being provided as a standalone score (Bañeres et al., 2021). Such issues are especially important in secondary education since the predicted outcomes concern minors and could affect counseling, expectations, subject support, and access to opportunities.

Recent evaluations of K-12 education confirm that artificial intelligence can be used to enable personalization, assessment, feedback, and administrative decision-making, while at the same time raising worries regarding teachers' readiness, transparency, data quality, equity, and responsible use (Crompton & Burke, 2022; Crompton et al., 2024; Chiu et al., 2023). Research into explainable artificial intelligence brings an additional need: stakeholders should receive explanations that are meaningful in an educational context, not just ones that are technically accurate about the model (Khosravi et al., 2022). Fairness research also warns that overall accuracy might mask unequal benefits or disadvantages for different subgroups (Herrmann & Weigert, 2024). In the case of secondary schools in Saudi Arabia, these issues are linked to the country's requirements about personal data since academic records, attendance, family information, and digital traces all constitute personal data that must be processed on a justified and proportionate basis (Saudi Data & AI Authority, 2023).

This review frames student performance prediction as a socio-technical intervention challenge. It synthesizes evidence on data, algorithms, validation, explainability, and intervention design to propose an implementation architecture for Saudi secondary schools. The focus is practical: prediction is positioned as one segment of a broader, human-led process involving detection, diagnosis, support, monitoring, and reexamination. Unlike reviews that primarily compare classifiers or accuracy, this analysis addresses the gap in literature by stressing the educational actions resulting from projections.

II. AIM, OBJECTIVES AND REVIEW QUESTIONS

This review intends to create a framework for using artificial intelligence to predict student performance and enable early academic intervention in Saudi secondary schools, focusing on responsible implementation rather than proposing new classifiers.

The synthesis is guided by five objectives: (1) identify academic, behavioral, digital, and contextual predictors that are both empirically useful and actionable in secondary schools; (2) compare machine-learning and deep-learning methods based on temporal usefulness, robustness, interpretability, and transferability, not just accuracy; (3) assess how prediction links to tiered intervention and outcome monitoring; (4) examine governance requirements for privacy, data limitation, fairness, explanation, specialized oversight, and appeals; and (5) propose a Saudi-specific operating model for prospective testing in various schools and regions.

This review considers four key questions: Which data streams provide the strongest and most defensible basis for early risk prediction? Which modeling and verification procedures best support deployment before outcomes are known? How can predictive outputs be translated into appropriate academic interventions without stigmatizing students? What governance and evaluation mechanisms guarantee trustworthy use in Saudi secondary education?

III. REVIEW METHODOLOGY

A structured integrative review was selected due to the diversity of empirical prediction studies, methodological reviews, early-warning system research, explainable AI, K-12 implementation evidence, and Saudi policy. A narrow meta-analysis was unsuitable given the variation regarding educational levels, outcome definitions, feature domains, sampling frames, algorithms, class balance, validation strategies, and outcome measures. The goal was to conduct a comparative synthesis to identify convergent findings, recurring methodological weaknesses, and broadly applicable design principles.

The review covers publications from 2020 to 2025 to reflect the rapid digitization of education and current developments in applied AI. Sources were identified using targeted scholarly searches and publisher or indexing metadata employing terms such as "student performance prediction," "educational data mining," "learning analytics," "at-risk students," "early warning system," "early intervention," "secondary school," "K-12," "machine learning," "deep learning," "explainable AI," and "Saudi Arabia." Preference was given to peer-reviewed studies and reviews directly related to prediction, intervention, model trustworthiness, or school-level AI. Two authoritative Saudi sources were included, recognizing that pragmatic deployment must integrate national capability policy and personal data governance with technical design.

To be included, a study had to have been published between 2020 and 2025, include verifiable bibliographic information, and be substantially relevant to at least one of the review questions. Empirical studies that had used machine-learning, deep-learning, or educational-data-mining methods to predict academic achievement, failure, attrition, or an outcome closely related to education were included. Reviews were included if they dealt with predictive analytics, educational machine learning, K-12 artificial intelligence, explainability, or other related implementation issues. Those studies which focused on higher education were kept selectively since a large amount of the early-warning evidence has developed in that area; nevertheless, the findings were treated as providing design evidence that could be transferred to Saudi secondary schools rather than as direct proof for such schools. Studies that were purely generative-AI and did not have a predictive or intervention function, publications prior to 2020, duplicate reports, studies that made predictions in a non-educational context, and sources with unverifiable bibliographic details were excluded.

The final corpus comprised twenty-eight scholarly publications and two Saudi policy or regulatory documents. Each source was analyzed across seven dimensions: educational context, predicted outcome, data and features, modeling approach, validation method, reported performance criteria, and evidence on explainability, intervention, fairness, or deployment. Special attention was given to whether predictors were available before the outcome and to the risk of temporal leakage, as models using information generated near or after the outcome may appear accurate but offer little practical lead time for intervention.

Six criteria were used instead of a single numerical score: sampling adequacy and outcome definition, transparency of preprocessing, handling of imbalance and missing data, appropriateness of validation, use of performance measures beyond accuracy, and attention to actionability or accountable deployment. This approach corresponds with recent reviews, which note that accuracy comparisons are unreliable when studies use different datasets and evaluation methods (Namoun & Alshanqiti, 2021; Albreiki et al., 2021; Fahd et al., 2022; Almalawi et al., 2024). Consequently, performance figures were interpreted individually rather than combined into a misleading ranking.

The synthesis followed three stages: (1) grouping empirical findings by predictor and model families; (2) using studies on interventions and trustworthy AI to identify needs beyond discrimination, such as lead time, calibration, explanation, human review, and feedback cycles; and (3) adapting requirements to the Saudi secondary school context, considering workflow, data availability, regional differences, and legal governance. As a literature-based design proposal, the framework calls for prospective testing before high-stakes adoption.

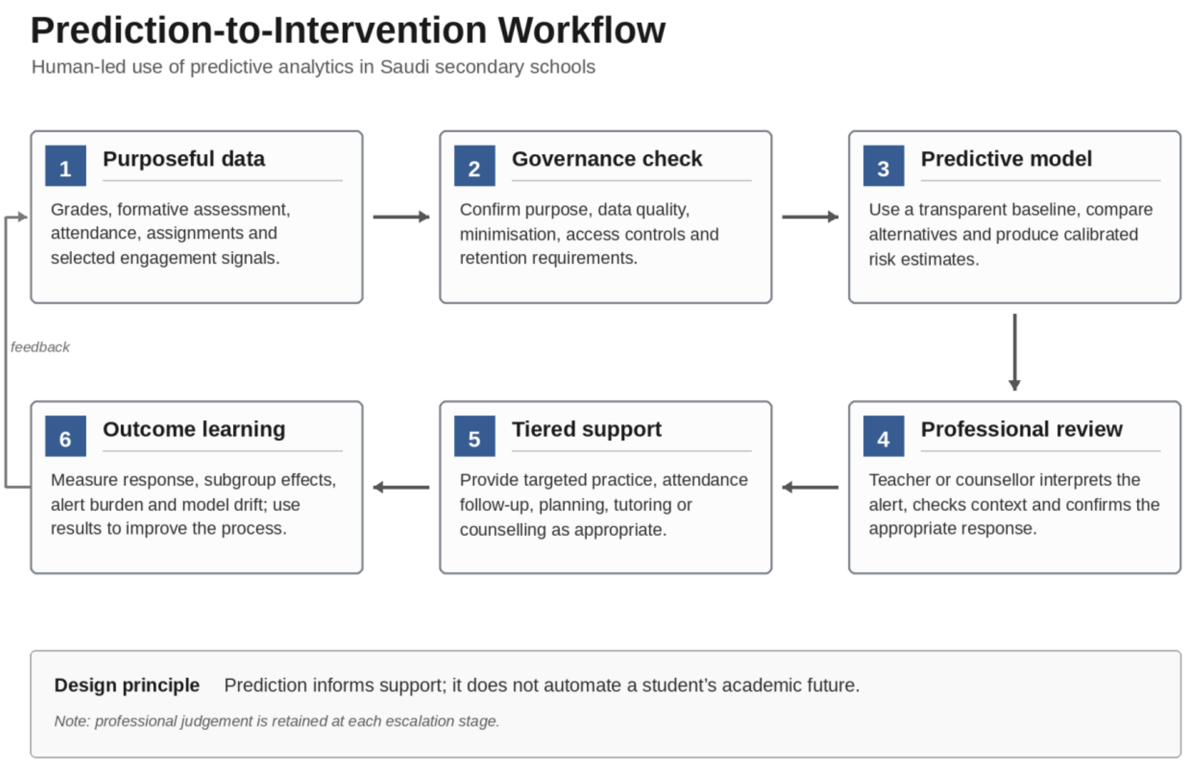

Figure 1. The cycle of evidence leading to intervention in Saudi secondary schools involving applied AI.

IV. EVIDENCE SYNTHESIS

4.1 Predictive factors: from past grades to signals that can be acted upon

Across the literature, prior academic performance is the most accurate predictor. Semester grades, assessment scores, and course marks consistently indicate accumulated mastery and study behavior (Yağcı, 2022; Feng et al., 2022; Nabil et al., 2021). The Saudi Al-Baha study also identified semester grades as key correlates of final achievement (Alghamdi & Rahman, 2023). However, the strongest predictor is not always the most useful for early intervention. Research shows that risk can be predicted at various points during a course. Adnan et al. (2021) demonstrated the benefits of combining demographic, clickstream, and assessment data as a course advances, while Alturki and Alturki (2021) highlighted educational data mining for early intervention. In secondary schools, features should be organized by their availability: baseline variables at term start, attendance and engagement updated weekly, formative assessments after each unit, and cumulative achievement at checkpoints. Risk scores must only use information available at the time action is expected. Behavioral and engagement data improve timeliness, as changes in attendance, punctuality, assignment completion, platform access, and interaction trends often precede grade declines. In online settings, clickstream and activity data support early prediction, though their value depends on course design and platform use (Alhothali et al., 2022; Karalar et al., 2021). Multimodal learning analytics combines digital traces with other signals, but increased information complexity raises risks of overcollection, unstable proxies, and intrusive surveillance (Chango et al., 2022). Saudi secondary schools ought to follow the principle of "minimum sufficient signal": use the smallest, least sensitive feature set that meaningfully improves early detection. Certain contextual attributes, such as family income, housing, parental characteristics, gender, and residence, may correlate with outcomes and reveal structural differences, as shown in various studies. However, these variables may be unsuitable for individual risk decisions. Their primary use should be at the population level, for auditing model fairness, identifying resource gaps, or planning school support. Sensitive or weakly actionable attributes should not be treated as causal; support should be based on current learning evidence, not demographic profiles. Recent research uses logistic regression, decision trees, random forests, support vector machines, kNN, boosting, ensembles, neural networks, and deep learning. Reviews consistently find no single algorithm outperforms all others, as performance depends on data size, feature engineering, outcome definition, class imbalance, and validation method (Khan & Ghosh, 2021; Batool et al., 2023; Almalawi et al., 2024). Random forests and ensembles are often competitive due to their capacity to capture nonlinear associations and handle mixed features, while simpler models are preferred for transparency, stability, and local deployment. Comparative modeling is important. Hussain and Khan (2023) and Zafari et al. (2021) demonstrated machine learning for secondary-level prediction, and Alghamdi and Rahman (2023) found Naive Bayes and Random Forest achieved high accuracy after balancing and feature selection. However, these results should not be generalized; algorithm accuracy may decline with curriculum changes, assessment reforms, or fluctuations in student behavior. Deep neural networks can model complex connections and sequences, as shown by Nabil et al. (2021), but added complexity must be justified. If a calibrated logistic model or gradient-boosted tree offers similar early-warning utility, their simpler maintenance and explanation requirements may make them more desirable. AI in schools should optimize for discrimination, calibration, lead time, interpretability, computational cost, stability, and equity.

4.2 Validation: the distinction between a model and a reliable warning

The assessment of educational predictions is a frequent weakness. Depending exclusively on accuracy can be misleading, especially in cases where the number of at-risk students is less than the number of successful ones. Precision shows how many of the warnings relate to real risk, recall shows how many genuinely at-risk students are identified, and the F1 score takes account of both of these measures. The area under the receiver-operating characteristic curve can be used to summarise the model's ranking performance, while the area under the precision-recall curve is generally more informative in situations where the risk is rare. Additionally, the model has to be calibrated so that when it assigns a 30 per cent risk, about that proportion of similar students should in fact experience the specified outcome.

The problem of class imbalance should be addressed directly. Techniques involving oversampling, such as SMOTE, can help improve learning from the minority groups, as shown in the study from Saudi Arabia, but resampling has to be carried out only on the training data in order to avoid information leakage. Although cross-validation is useful for the development of the model, the use of random folds may lead to an overestimation of the performance when the student cohorts are either temporally or institutionally clustered. A more robust evaluation design in Saudi Arabia should involve training the model on earlier cohorts, testing it on later cohorts, and then carrying out external validation in schools or regions that were not used in the model development process. Research into generic models also cautions that predictive performance can decline when models are applied to different courses or contexts (Ramaswami et al., 2022).

The threshold should be set in line with the capacity of the intervention. A counselling team which is able to establish meaningful contact with fifty students should not be given five hundred alerts that have low specificity. On the other hand, a high threshold that results in few false positives might fail to identify students who could benefit from cheap support. This is a decision problem, not a problem for which a leaderboard is relevant. Different thresholds can be

© AUG 2026 | IRE Journals | Volume 10 Issue 2 | ISSN: 2456-8880
DOI: https://doi.org/10.64388/IREV10I2-1722647

used according to the level of intervention: a low threshold could trigger an automated message reminding students about study resources, a moderate threshold might lead to teacher review, and a high and persistent risk level could result in outreach by a counsellor. Human confirmation should be obtained before any action is taken at each stage of escalation.

4.3 Explainability, fairness and trust

It is particularly important for explainable artificial intelligence that human decisions be influenced by predictions. Khosravi and others (2022) claim that educational explanations need to meet the requirements of a variety of stakeholders and serve different purposes; an explanation that is useful to a data scientist might be of no use to a teacher, a learner, or a parent. In the case of early warning, the type of explanation that is preferred should be one that is actionable and set in time, for example by referring to a decline in formative scores, repeated absences, unsubmitted assignments, or a drop in engagement over a specific recent period. Simply giving feature significance is not enough since it may only show an association at the level of the population and not a causal reason for that individual.

Fairness also has to be considered in the light of the entire system. Herrmann and Weigert (2024) demonstrate that AI-driven predictions can lead to an uneven distribution of benefits even if the average performance appears positive. In secondary schools, subgroup audits need to look at the false-negative rates, incorrect positive rates, calibration, the receipt of interventions, and the later outcomes among different groups and regions. It may be necessary to use protected or sensitive variables when carrying out an audit even if such variables are not included in the functional forecast. The aim is not to achieve mathematical equality at any price, but rather to provide evidence that the system does not consistently fail to offer useful support or impose disproportionate burdens.

Reliable systems also need to have documentation. Bañeres et al. (2021) regard early warning as a form of infrastructure, stressing the importance of repeatability and confidence in the process. It follows that a deployment in Saudi Arabia should include model cards, data dictionaries, version histories, the reasons for the thresholds, activity logs, and records of the decisions made by teachers or counsellors. If a professional chooses to override an alert, then that decision should be both permitted and recorded. Eventually, the overrides may show up model blind spots, data-quality issues, or local knowledge which should lead to retraining.

Table 1. Evidence of a representative nature that informed predictive early intervention (2020–2024).

Study	Setting / data focus	Analytical approach	Review implication
Alghamdi & Rahman (2023)	Saudi secondary graduates; academic + demographic data	NB, RF, J48; SMOTE; feature selection	Local feasibility is strong, but prospective and cross-region validation is still required.
Adnan et al. (2021)	Risk prediction at successive course-completion points	ML/DL using demographic, clickstream and assessment data	Lead time is part of model utility; prediction should occur while support can still change outcomes.
Alturki & Alturki (2021)	Saudi university records; early academic prediction	Six data-mining classifiers + feature correlation	Demonstrates a Saudi early-intervention framing and the value of prior academic indicators.
Bañeres et al. (2021)	Operational early-warning infrastructure	Predictive analytics pipeline and EWS architecture	Trust depends on workflow, data handling, repeatability and user-facing decision support.
Ifenthaler & Yau (2020)	Systematic review of learning analytics for study success	Evidence synthesis	Analytics becomes educationally useful when it delivers timely feedback and scaffolding.

© AUG 2026 | IRE Journals | Volume 10 Issue 2 | ISSN: 2456-8880
DOI: https://doi.org/10.64388/IREV10I2-1722647

Karalar et al. (2021)	Distance learning; synchronous and asynchronous activity	Ensemble learning	Engagement indicators can provide earlier signal than final grades, but their meaning is context dependent.
Hussain & Khan (2023)	Secondary and intermediate-level students	Comparative machine learning	School-age prediction is feasible; local validation remains necessary before deployment.
Zafari et al. (2021)	High-school student performance	Machine learning with feature-importance analysis	Prediction is more useful when influential features can support educational interpretation.
López-García et al. (2023)	Early failure detection	Machine-learning classification	Operational value depends on identifying risk before the target outcome is fixed.
Ramaswami et al. (2022)	Cross-context prediction models	Generic educational-data-mining models	Transportability should be tested rather than assumed across cohorts, courses or schools.
Khosravi et al. (2022)	Explainable AI in education	XAI conceptual and empirical synthesis	Explanations must be tailored to educational stakeholders and decision purposes.
Herrmann & Weigert (2024)	Fairness of AI-based academic prediction	Subgroup-focused predictive evaluation	Aggregate accuracy can conceal unequal support or disadvantage across student groups.

V. MOVING FROM PREDICTION TO EARLY ACADEMIC INTERVENTION IN SAUDI SCHOOLS

The main idea behind the design is to keep the process of making predictions separate from dealing with the consequences. Instead of the AI model producing an irreversible academic decision, it should provide evidence for people to review. A practical workflow for a school can be divided into four stages. The first stage involves continuous monitoring using ordinary academic and attendance data. The second stage consists of a mild reaction to emerging risks, for example by giving specialized practice, arranging study plans, following up on attendance, or offering teacher feedback. In the third stage, if the risk continues to show up across several different indicators, a counselor or someone providing learning support looks at the case. In the fourth stage, individualized support is arranged in consultation with the student, and where appropriate also with their parents or guardians. The need to escalate should be based on the current evidence and on the student's response to the support given, not on a fixed label.

Since this approach gives greater importance to timing than to maximum retrospective accuracy, research into early detection has shown that useful warning systems are able to pick up possible failures before the outcomes take place (López-García et al., 2023). The school should therefore set up prediction points that match actual opportunities to take action—for instance, following the first assessment cycle, after a sustained change in attendance, or after a number of instances of incomplete work. Although alerts issued just before the final exams may be statistically striking, they are educationally inadequate.

The intervention plan should be matched to the predictor as well. If a student has low formative attainment, then they should be given subject-specific tutoring or adaptive practice. Where there is repeated absence, attendance outreach and a barrier assessment should be carried out. When assignments are missed, the student should receive support with executive function, have deadlines planned, or be contacted by the teacher. In the case of a drop in engagement on the platform, it is appropriate to check whether the problem lies with access, workload, or motivation rather than issuing a general academic warning. This approach of linking predictors to responses stops the

© AUG 2026 | IRE Journals | Volume 10 Issue 2 | ISSN: 2456-8880
DOI: https://doi.org/10.64388/IREV10I2-1722647

AI from becoming a broad deficit detector and turns the evidence into a hypothesis that instructors can verify against the student.

When putting into practice the model it should take account of variations between regions and between schools. The Al-Baha example is so valuable because it shows that it is possible to make local predictions, covering both rural and suburban situations, yet a national model must not assume that the distribution of features or instructional methods in one area applies to all schools. External validation therefore needs to involve different regions, different sizes of school, varying degrees of urbanisation, gender-segregated environments in which these are relevant, different curricular tracks, and different patterns of technology use. It is necessary to monitor for model drift since changes to the assessment policy, to the learning platforms, or to attendance practices can affect the relationships between the predictors and the outcomes.

Data governance should not be treated as a simple compliance measure since it has an impact on the design of the system. The Saudi Personal Data Protection Law and its implementing framework mandate that personal data be handled in a lawful, purpose-specific, and accountable manner (Saudi Data & AI Authority, 2023). In the case of educational analytics that involve minors, schools must implement data limitation, access controls, retention schedules, clearly documented purposes, and visibility based on role. Whenever possible, raw data should be kept separate from the analytics outputs, and the dashboards should show only the information necessary for each role. Any sensitive contextual variables should be subject to a particular justification, particularly when a less intrusive academic or behavioural signal could achieve the same intervention objective.

Students and families should be able to see human monitoring. A school must be able to explain that an alert is a probabilistic indication based on the student's most recent educational records, that no automated decision determines what the student's future will be, and that teachers or counselors look at the evidence. Students should have a way of correcting inaccurate data in the sources, for example data showing a wrong absence or missing work. This operates as both a fairness measure and a mechanism to improve data quality. It also helps to reduce the chance that a prediction becomes self-fulfilling since educators may lower their expectations after seeing a risk label.

Table 2. Matrix of the Saudi intervention and assurance in secondary schools.

Trigger	Proportionate response	AI function	Human safeguard	Evaluation indicator
Formative attainment decline	Subject-specific tutoring, practice or feedback	Estimate near-term academic risk from recent attainment	Teacher reviews trend and instructional context	Calibration; grade recovery; false-alert rate
Sustained attendance change	Attendance outreach and barrier assessment	Detect change over a defined time window	Attendance team verifies records and context	Attendance recovery; corrected-record rate
Repeated missing assignments	Planning, deadline support and teacher contact	Flag completion-pattern risk	Teacher/student confirms workload and due dates	Completion recovery; alert persistence
Digital engagement decline	Check access, workload, motivation or platform use	Treat digital trace as a trajectory signal, not surveillance	Student conversation before escalation	Signal stability; intervention acceptance

Persistent multi-domain risk	Counselor-led individualized support plan	Combine calibrated evidence across domains	Counselor confirms need; family involved where appropriate	Outcome improvement; alert burden; student feedback
Model drift or subgroup disparity	Review thresholds, features or retraining need	Surface monitoring metrics and audit results	Governance group approves model changes	Subgroup error gaps; calibration drift; override patterns

VI. A SUGGESTED FRAMEWORK FOR APPLIED-AI AND A RESEARCH AGENDA

The framework that has been proposed consists of six interconnected layers. The first of these is a data layer which includes only those inputs that have been approved and that are relevant to the purpose in question: previous grades, formative assessments, attendance, completion of assignments, and certain engagement indicators that have been carefully chosen. The second layer is a governance checkpoint that ensures the purpose is legal, that the data quality is adequate, that data minimisation has been achieved, that access is properly controlled, that data retention is appropriate, and that the data can only be used in permitted ways. The third layer is a modelling layer in which a transparent baseline is compared with tree-based ensembles and, in appropriate cases, with deep models. The fourth layer is a validation layer and it involves temporal holdout, testing using data from other schools, calibration, an examination of the precision-recall behaviour, assessment of fairness inside subgroups, and the detection of drift. The fifth layer is a decision-support layer that turns risk probabilities into explanations and provides tiered recommendations. The sixth layer is an intervention-learning layer and it keeps a record of the type of support that was given and whether or not the student's future progress improved.

This last stage is the most important aspect of research. While most studies which aim to predict performance assess whether a model accurately forecasts an outcome, only a small number examine whether acting on that forecast leads toward improved outcomes. It follows that a well-developed Saudi research programme should shift from carrying out prediction studies to conducting prospective intervention trials; schools could compare standard support with AI-assisted triage while still permitting teachers their discretion and then assess not just grades but also attendance recovery, assignment completion, pupil engagement, counselor workload, the amount of false alerts, and students' perceptions of fairness.

The following three methodological priorities arise. First, temporal generalization ought to be adopted. Models should be trained on earlier cohorts and then tested prospectively on subsequent cohorts, as opposed to depending primarily on random divisions. Second, alongside measures of discrimination, assessments of calibration and use of decision curves should be carried out so that schools can decide whether the alerts result in a net educational benefit when delivered at realistic levels of capacity. Third, research should make a distinction between predictive features and intervention targets; a feature may be statistically useful without being suitable for action, and an intervention may still be valuable even if the feature that triggers it is only a proxy for a more general learning difficulty.

Research should also look into the transportability of models within Saudi Arabia. It should quantify which approach—a single national model, region-specific models, or models that are locally fine-tuned—best achieves a balance involving consistency and relevance. Although federated or privacy-preserving learning might one day reduce the need to centralise detailed student records, such approaches should only be adopted once simpler methods of governance and modelling have been put in place. Technical innovation should not go beyond the institution's capacity.

In the end, the evaluation must take into account the people who use and experience the system. Teachers require concise explanations and a reasonable number of alerts; counselors need access to the history and

© AUG 2026 | IRE Journals | Volume 10 Issue 2 | ISSN: 2456-8880
DOI: https://doi.org/10.64388/IREV10I2-1722647

information about escalation; school leaders should be given overall trends without being exposed to individual cases; and students should receive reasons that are easy to understand along with adequate support. By using participatory design it is possible to determine whether a dashboard that seems to be useful might cause stigma, give rise to surveillance concerns, or result in additional administrative work. The criterion for success is thus not a higher model score by itself, but rather a more effective support process.

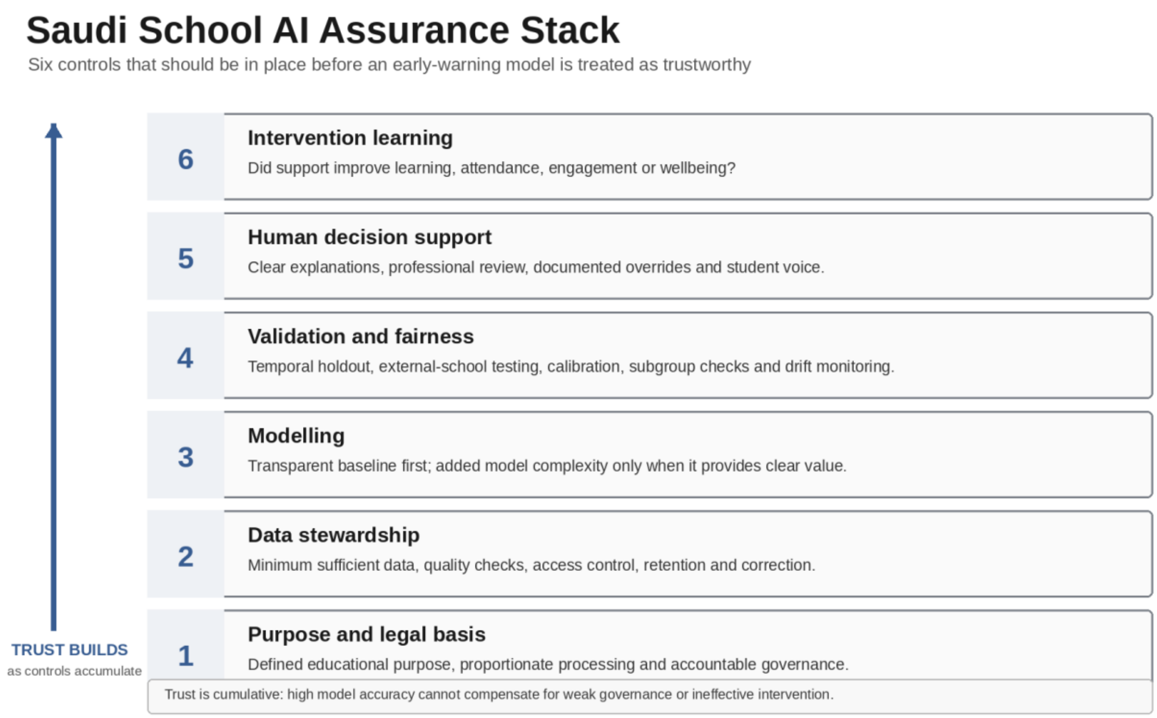

Figure 2. Six-layer assurance stack

The literature indicates that predictive student modelling is technically advanced enough to enable the use of early-warning systems that have been designed, but not advanced enough to warrant autonomous educational decisions. In various studies, previous attainment, continuous assessment, attendance, and level of engagement act as consistent predictive indicators, whereas the algorithmic rankings depend on the specific context. Numerous reviews cautions against considering isolated accuracy as evidence of general superiority. Evidence from Saudi Arabia shows that data obtained from local secondary schools can yield good predictive outcomes, but it also points out the necessity of more extensive external validation before the approach can be adopted nationally The main contribution of this review is to shift the function of artificial intelligence from that of a classifier to that of a coordination mechanism; in this capacity, it enables schools to detect changes earlier, decide what to focus on, gather the relevant evidence, and assess whether the support had any effect. Although this approach is in line with research into explainable AI, with early-warning systems, and with the current body of evidence from the K-12 sector—all of which support a human-centred integration rather than replacement—professional judgment is still responsible for interpretation and intervention (Crompton et al., 2024; Khosravi et al., 2022; Susnjak, 2024).thAnother consequence is that the quality of intervention determines the value of prediction; if schools do not have the capacity to provide tutoring, counseling routes, attendance support, or systems for monitoring responses, improved prediction can result in a more accurate identification of unmet needs. Implementation research should therefore assess an organisation's readiness together with the performance of the model. This is especially pertinent to the human-capability objectives of Vision 2030 since educational AI should be intended to expand students' opportunities and helping them make successful progress rather than just optimising institutional reporting. support successful progression rather than merely optimize institutional reporting.

VII. CONCLUSION

Applied artificial intelligence can strengthen early academic support in Saudi secondary schools when predictive analytics is designed as part of a governed human involvement cycle. Evidence from 2020-2025 shows that academic history remains the strongest predictor family, while attendance, engagement, formative assessment, and selected contextual signals can improve early warning. Machine-learning and deep-learning methods are useful, but no method is universally best; deployment quality depends on temporal validation, calibration, external testing, explainability, fairness, and sustainable school workflows.

The proposed framework therefore places privacy and data quality before modeling, professional review before consequence, and intervention outcomes after every alert. It extends the Saudi secondary-school evidence base from whether performance can be predicted to the more consequential question of whether timely, transparent prediction may help educators deliver better support. Future studies should test this system prospectively throughout diverse Saudi regions and should report educational benefit, subgroup effects, alert burden, and student experience alongside conventional predictive metrics. A successful national approach should therefore reward interventions that clearly improve learning, while continuously retiring features, thresholds, or models that no longer reliably serve students well.

The appropriate endpoint for school AI is not automated classification. It is earlier, more consistent, and accountable assistance that preserves human decision and student dignity.

REFERENCES

[1] Adnan, M., Habib, A., Ashraf, J., Mussadiq, S., Raza, A. A., Abid, M., Bashir, M., & Khan, S. U. (2021). Predicting at-risk students at different percentages of course length for early intervention using machine learning models. *IEEE Access, 9*, 7519–7539. Crossref

[2] Albreiki, B., Zaki, N., & Alashwal, H. (2021). A systematic literature review of student' performance prediction using machine learning techniques. *Education Sciences, 11*(9), 552. Crossref

[3] Alghamdi, A. S., & Rahman, A. (2023). Data mining approach to predict success of secondary school students: A Saudi Arabian case study. *Education Sciences, 13*(3), 293. Crossref

[4] Alhothali, A., Albsisi, M., Assalahi, H., & Aldosemani, T. (2022). Predicting student outcomes in online courses using machine learning techniques: A review. *Sustainability, 14*(10), 6199. Crossref

[5] Almalawi, A., Soh, B., Li, A., & Samra, H. (2024). Predictive models for educational purposes: A systematic review. *Big Data and Cognitive Computing, 8*(12), 187. Crossref

[6] Alturki, S., & Alturki, N. (2021). Using educational data mining to predict students' academic performance for applying early interventions. *Journal of Information Technology Education: Innovations in Practice, 20*, 121–137. Crossref

[7] Alyahyan, E., & Düştegör, D. (2020). Predicting academic success in higher education: Literature review and best practices. *International Journal of Educational Technology in Higher Education, 17*, 3. Crossref

[8] Batool, S., Rashid, J., Nisar, M. W., Kim, J., Kwon, H.-Y., & Hussain, A. (2023). Educational data mining to predict students' academic performance: A survey study. *Education and Information Technologies, 28*(1), 905–971. Crossref

[9] Bañeres, D., Guerrero-Roldán, A. E., Rodríguez-González, M. E., & Karadeniz, A. (2021). A predictive analytics infrastructure to support a

trustworthy early warning system. *Applied Sciences, 11*(13), 5781. Crossref

[10] Chango, W., Lara, J. A., Cerezo, R., & Romero, C. (2022). A review on data fusion in multimodal learning analytics and educational data mining. *WIREs Data Mining and Knowledge Discovery, 12*(4), e1458. Crossref

[11] Chiu, T. K. F., Xia, Q., Zhou, X., Chai, C. S., & Cheng, M. (2023). Systematic literature review on opportunities, challenges, and future research recommendations of artificial intelligence in education. *Computers and Education: Artificial Intelligence, 4*, 100118. Crossref

[12] Crompton, H., & Burke, D. (2022). Artificial intelligence in K-12 education. *SN Social Sciences, 2*(7), 113. Crossref

[13] Crompton, H., Jones, M. V., & Burke, D. (2024). Affordances and challenges of artificial intelligence in K-12 education: A systematic review. *Journal of Research on Technology in Education, 56*(3), 248–268. Crossref

[14] Fahd, K., Venkatraman, S., Miah, S. J., & Ahmed, K. (2022). Application of machine learning in higher education to assess student academic performance, at-risk, and attrition: A meta-analysis of literature. *Education and Information Technologies, 27*(3), 3743–3775. Crossref

[15] Feng, G., Fan, M., & Chen, Y. (2022). Analysis and prediction of students' academic performance based on educational data mining. *IEEE Access, 10*, 19558–19571. Crossref

[16] Herrmann, L., & Weigert, J. (2024). AI-based prediction of academic success: Support for many, disadvantage for some? *Computers and Education: Artificial Intelligence, 7*, 100303. Crossref

[17] Hussain, S., & Khan, M. Q. (2023). Student-Performulator: Predicting students' academic performance at secondary and intermediate level using machine learning. *Annals of Data Science, 10*(3), 637–655. Crossref

[18] Ifenthaler, D., & Yau, J. Y.-K. (2020). Utilising learning analytics to support study success in higher education: A systematic review. *Educational Technology Research and Development, 68*(4), 1961–1990. Crossref

[19] Karalar, H., Kapucu, C., & Gürüler, H. (2021). Predicting students at risk of academic failure using ensemble model during pandemic in a distance learning system. *International Journal of Educational Technology in Higher Education, 18*, 63. Crossref

[20] Khan, A., & Ghosh, S. K. (2021). Student performance analysis and prediction in classroom learning: A review of educational data mining studies. *Education and Information Technologies, 26*(1), 205–240. Crossref

[21] Khosravi, H., Buckingham Shum, S., Chen, G., Conati, C., Tsai, Y.-S., Kay, J., Knight, S., Martinez-Maldonado, R., Sadiq, S., & Gašević, D. (2022). Explainable artificial intelligence in education. *Computers and Education: Artificial Intelligence, 3*, 100074. Crossref

[22] Kingdom of Saudi Arabia. (2021). Human Capability Development Program Delivery Plan 2021–2025: Saudi Vision 2030. *Vision 2030.*

[23] López-García, A., Blasco-Blasco, O., Liern-García, M., & Parada-Rico, S. E. (2023). Early detection of students' failure using machine learning techniques. *Operations Research Perspectives, 11*, 100292. Crossref

[24] Nabil, A., Seyam, M., & Abou-Elfetouh, A. (2021). Prediction of students' academic performance based on courses' grades using deep neural networks. *IEEE Access, 9*, 140731–140746. Crossref

[25] Namoun, A., & Alshanqiti, A. (2021). Predicting student performance using data mining and learning analytics techniques: A systematic literature review. *Applied Sciences, 11*(1), 237. Crossref

[26] Ramaswami, G., Susnjak, T., & Mathrani, A. (2022). On developing generic models for predicting student outcomes in educational data mining. *Big Data and Cognitive Computing, 6*(1), 6. Crossref

[27] Saudi Data & AI Authority. (2023). The Implementing Regulation of the Personal Data Protection Law. Kingdom of Saudi Arabia.

[28] Susnjak, T. (2024). Beyond predictive learning analytics modelling and onto explainable artificial intelligence with prescriptive analytics

and ChatGPT. *International Journal of Artificial Intelligence in Education, 34*, 452–482. Crossref

[29] Yağcı, M. (2022). Educational data mining: Prediction of students' academic performance using machine learning algorithms. *Smart Learning Environments, 9*, 11. Crossref

[30] Zafari, M., Sadeghi-Niaraki, A., Choi, S.-M., & Esmaeily, A. (2021). A practical model for the evaluation of high school student performance based on machine learning. *Applied Sciences, 11*(23), 11534. Crossref