JAIAA
Journal of AI Analytics and Applications
ISSN 3067-7386 | Open Access | Peer-Reviewed

RESEARCH ARTICLE
Vol. 1, No. 3 (2023), pp. 42–60
DOI: 10.63646/jaiaa.2023.010303

AI-Assisted Analytics for Practical Education Appraisal: Evaluating Agriculture–Forestry Management Training with Multi-Indicator Models

Wenjing Hao[1]; Liangyu Chen[2]; Mingrui Zhao[3, *]

[1] College of Economics and Management, Henan Agricultural University, Zhengzhou, China
[2] College of Economics and Trade, Hebei Agricultural University, Baoding, China
[3] College of Management, Sichuan Agricultural University, Chengdu, China
* Corresponding author: mingrui.zhao@sicau.edu.cn

ARTICLE INFO
Received
April 18, 2023
Revised
June 21, 2023
Accepted
August 12, 2023
Available Online
September 30, 2023
DOI
10.63646/jaiaa.2023.010303
License
Creative Commons Attribution 4.0 International Licence (CC BY 4.0)
Publisher
INATGI, United States of America
Journal
JAIAA - ISSN 3067-7386

Abstract
Agriculture–forestry economic management is a strategically important undergraduate major in China, but its practical education component is difficult to appraise because the discipline mixes economic, ecological, organisational and field-skill content. This study designs an AI-assisted multi-indicator appraisal framework that combines factor analysis, the analytic hierarchy process (AHP), entropy weighting, fuzzy comprehensive evaluation and gradient-boosted scoring, and applies it to data drawn from five Chinese agricultural universities. A four-dimension, eighteen-indicator instrument was completed by 412 senior undergraduate students, 38 faculty members, 24 administrators and 64 peer teachers, and was triangulated against 2,154 practice-base log records. Factor analysis extracted four latent factors whose structure matched the conceptual dimensions of the instrument (Cronbach α between 0.81 and 0.93). The AI-assisted scoring engine outperformed the three classical methods on reliability and discriminant validity, while AHP retained the strongest stakeholder interpretability. Empirical results show that instructor competence carries the largest aggregate weight (0.291–0.328 across methods) and that practical operation ability, teaching methods and practicality are the three highest individual indicators. A three-stage optimisation pathway is proposed (0–6, 6–18 and 18–36 months) and projected outcomes include a 23–31% rise in employer-rated graduate competence and an 18% increase in postgraduate placement. The framework gives Chinese agricultural universities a reproducible, analytics-grounded route to align practical education with rural revitalisation policy and the training needs of the modern agriculture–forestry sector.

Keywords: practical education appraisal; agriculture–forestry economic management; multi-indicator model; factor analysis; AHP; fuzzy comprehensive evaluation; gradient-boosted scoring.

I. INTRODUCTION

Chinese higher education has experienced sustained reform over the last decade, with particular emphasis on aligning undergraduate training to the country's strategic modernisation goals. Among the disciplines most directly tied to these goals is agriculture–forestry economic management, an interdisciplinary major that joins economics, agricultural science, forest resource management and rural sociology. Graduates of this major are expected to support rural revitalisation, agricultural modernisation and ecological civilisation construction, three priorities articulated across successive Five-Year Plans and reinforced by the rural revitalisation strategy (Liu et al., 2020; Hu et al., 2022). Yet the practical education component of the major has remained a chronic concern. Field-based learning, on-site internships and project work demand performance appraisal mechanisms that are difficult to construct because the underlying competencies are heterogeneous, partly tacit and

RESEARCH ARTICLE
Vol. 1, No. 3 (2023), pp. 42–60
DOI: 10.63646/jaiaa.2023.010303

contextually variable.

Traditional student evaluations of teaching (SETs), although widely used, struggle to capture practical learning outcomes (Spooren et al., 2013; Boring et al., 2016). Generic appraisal instruments emphasise content delivery and overlook the operational dimensions that define agricultural practice: site selection, field measurement, agronomic decision making, sustainability accounting and stakeholder negotiation. Recent reviews observe that performance appraisal is moving away from single-rater, single-method designs towards multi-source, multi-method instruments that combine quantitative and qualitative signals (DeNisi & Murphy, 2017; Brown et al., 2019; Tziner & Rabenu, 2018; Ali et al., 2021). The shift opens space for analytics-driven frameworks that can fuse heterogeneous evidence without losing interpretability for instructors and administrators.

Artificial intelligence and management analytics have, in parallel, matured to the point where they can support appraisal tasks that previously required intuitive, case-by-case judgement (Lu, 2019; Lu, 2021; Lu et al., 2024; Ye & Lu, 2022; Lu, 2017). Modern educational data mining links observational evidence to latent learner traits with quantifiable error bounds, embedded in the broader Industry 4.0 transformation of professional work (Lu, 2017; Romero & Ventura, 2020; Aldowah et al., 2019; Zawacki-Richter et al., 2019; Hwang et al., 2020; Pinto et al., 2023), while gradient-boosted models give competitive predictive accuracy on structured Likert-style data (Chen & Guestrin, 2016; Ke et al., 2017; Chen et al., 2020; Luan et al., 2020). The challenge is no longer to identify analytical methods but to integrate them into appraisal systems that respect the pedagogical and institutional context of the discipline.

Surprisingly little research has applied this analytics-grounded perspective to agricultural and forestry training programmes in China. The existing literature on agricultural education concentrates on curriculum design (Wang & Yang, 2021; Zhou et al., 2021), rural revitalisation linkages (Hu et al., 2022; Yan et al., 2020) and broad policy commentary, with appraisal mechanisms treated as a secondary issue. Cross-discipline reviews of agriculture–forestry economic management indicate that indicator systems are often inherited from generic higher education quality assurance and rarely adjusted to the discipline's practical demands (Li et al., 2020; Hou et al., 2015; Yang et al., 2015; Sin et al., 2017). A purposive design effort is needed to construct a discipline-specific appraisal instrument, validate it empirically and demonstrate analytics-driven scoring.

This study contributes on four fronts. First, it develops a four-dimension, eighteen-indicator appraisal instrument that operationalises the practical education system, the teaching process, instructor competence and the teaching environment. Second, it applies four scoring methods — AHP, factor analysis, fuzzy comprehensive evaluation and an AI-assisted gradient-boosted approach — and compares their psychometric performance. Third, it analyses appraisal scores across five Chinese agricultural universities, providing one of the first multi-institution baselines for the discipline. Fourth, it derives an actionable three-stage optimisation pathway with projected outcomes that institutional leaders can use to plan reforms.

The relevance of this work extends beyond the immediate task of evaluating one undergraduate major. China's commitment to building a modern agricultural and forestry sector requires a deep pipeline of graduates with hybrid economic, ecological and managerial competence. Recent policy documents on rural revitalisation and ecological civilisation construction stress that institutional reforms in higher education are a necessary complement to investment in rural infrastructure and to subsidies for agricultural innovation. Without rigorous appraisal of practical learning, large investments in laboratories, internship programmes and faculty development risk being misallocated. The framework reported here is intended to give academic leaders an instrument they can defend in front of funding bodies, accreditation panels and rural stakeholders. The same instrument can be re-used for related disciplines including agronomy, agricultural engineering, forestry economics and rural management, with minor adjustments to the indicator weights.

Three observations from policy reports and accreditation guidance further motivate the study. The National Standards for Teaching Quality issued by the Chinese Ministry of Education emphasise that practical training must occupy at least 25% of total credit hours in agriculture–forestry economic management; yet there is no national-level instrument for measuring the effectiveness of that 25% beyond generic teaching evaluations.

RESEARCH ARTICLE
Vol. 1, No. 3 (2023), pp. 42–60
DOI: 10.63646/jaiaa.2023.010303

International accreditation bodies for agricultural and forestry programmes increasingly require evidence-based reporting on practical learning outcomes, and demonstration of continuous improvement processes anchored in measurable indicators. Finally, employer surveys of agricultural graduates consistently identify practical skill gaps as the leading concern; bridging this gap requires appraisal systems that speak directly to the skill domains employers value.

The remainder of this paper is organised as follows. Section II reviews the related literature on practical education, performance appraisal and analytics-driven evaluation. Section III sets out the multi-indicator methodology. Section IV reports the empirical results from the cross-institution study. Section V discusses theoretical and managerial implications. Section VI proposes optimisation pathways and recommendations. Section VII concludes and outlines limitations and future directions.

II. RELATED WORK AND THEORETICAL BACKGROUND

Practical education in agriculture–forestry economic management synthesises classroom instruction with field-based training, internships and project work. Pedagogical research grounds this approach in experiential learning theory, which holds that competence develops through cycles of concrete experience, reflective observation, abstract conceptualisation and active experimentation (Kolb, 2014; Beard & Wilson, 2018). Applied to agricultural disciplines, the cycle requires learners to inhabit production settings, encounter biophysical constraints and translate observations into management decisions. Empirical studies of agricultural college students confirm that practical exposure improves workplace adaptation, employability and persistence (Liu et al., 2019; Mukembo et al., 2017).

Performance appraisal of teaching has evolved substantially over the past century (DeNisi & Murphy, 2017). Early instruments relied on supervisor ratings, often single-source and biased by leniency or recency effects (Marsh & Roche, 1997; Bell & Brooks, 2017). Subsequent generations introduced student evaluations, peer observation and student outcomes, though concerns about validity and method-bias remain (Onwuegbuzie et al., 2009; Stehle et al., 2012). Contemporary models advocate multi-source, multi-method designs and emphasise psychometric scrutiny of instruments (Spooren et al., 2013; Iqbal et al., 2019; Brown et al., 2019). In professional and competency-based domains, appraisal is increasingly tied to observable performance, task-level evidence and external benchmarks (Mulder, 2017; Salas et al., 2012; Anderson & Krathwohl, 2001; Tessema, 2018).

Multi-criteria decision analysis (MCDA) provides the structural backbone for appraisal instruments that combine heterogeneous indicators. The analytic hierarchy process (AHP) remains widely used because of its transparency and ability to elicit expert judgement (Sipahi & Timor, 2010; Zhang et al., 2019; Saaty & De Paola, 2017; Wang & Chin, 2008). Fuzzy extensions, including fuzzy AHP and fuzzy comprehensive evaluation, accommodate linguistic uncertainty in stakeholder responses (Liu et al., 2020; Mardani et al., 2015). Entropy weighting offers an objective alternative that derives weights from the dispersion of observed data (Sun et al., 2020). Comparative surveys highlight that no single method dominates across reliability, validity, computational efficiency and interpretability dimensions (Mardani et al., 2016; Pamučar et al., 2020; Zyoud & Fuchs-Hanusch, 2017).

Educational data mining and learning analytics extend the toolset further. Reviews show that predictive analytics, clustering and structural equation modelling are now standard in higher education research (Romero & Ventura, 2020; Aldowah et al., 2019; Crompton & Burke, 2023; Kuzilek et al., 2017). Machine learning models — random forests, gradient-boosted trees and neural networks — have documented accuracy advantages on student-outcome prediction tasks (Hellas et al., 2018; Khan & Ghosh, 2021; Cui et al., 2019), and increasingly support interpretable explanations of their decisions (Lemay & Doleck, 2020). When combined with traditional MCDA, machine-learned weights provide a useful counterbalance to expert-elicited priors and can flag indicator anomalies that human reviewers may overlook (Liu et al., 2018; Natekin & Knoll, 2013).

Higher education in China presents a distinct context. Policy reforms have expanded enrolments, increased the proportion of practical credit hours and rewarded industry-university collaboration (Xu & Yang, 2019; Bao, 2020; Wang et al., 2022; Zhang et al., 2018). The rural revitalisation strategy explicitly identifies agricultural and forestry universities as talent-training bases for the modern agricultural sector (Liu et al., 2020; Tang et al., 2018). Bibliometric studies of agricultural education in China report rapid growth in journal publications and increasing diversity of methodological approaches, though appraisal instruments specific to agriculture–forestry economic management remain under-developed (Wang & Yang, 2021; Zhou et al., 2021).

Internationally, agricultural and forestry programmes face similar challenges. Agroforestry interventions in low- and middle-income countries demand operator-level competence in both agronomy and economics (Castle et al., 2021; Knowler, 2017; MacMillan & Benton, 2014). Forestry education has been argued to need closer integration with sustainable development objectives (Sayer et al., 2019; Filho et al., 2018; Lozano et al., 2017; Vandermeulen et al., 2020). The emerging skills agenda for the agri-food and forestry sectors emphasises green skills, data literacy and stakeholder negotiation (McGunagle & Zizka, 2020; Succi & Canovi, 2019). Together these trends justify a tailored appraisal framework rather than a generic one.

Three gaps motivate the present study. First, appraisal indicator systems have rarely been validated through factor-analytic procedures on Chinese agricultural cohorts. Second, comparative evidence on the performance of MCDA, fuzzy methods and machine learning when applied to teaching evaluation in this discipline is sparse. Third, policy-oriented optimisation pathways that translate analytics into actionable institutional reforms are largely missing. The methodology and findings reported below address these gaps.

Beyond the methodological literature, a substantial body of recent work has explored digital and AI-enabled tools in agricultural and forestry contexts more broadly. Smart agriculture platforms generate continuous streams of sensor and management data that can be used to anchor practical learning experiences in realistic problem settings (Wolfert et al., 2017; Liakos et al., 2018; Kamilaris & Prenafeta-Boldú, 2018). Although the connection between these production-side advances and higher education has not yet been mapped in detail, the alignment between the skills demanded by smart-agriculture deployments and the competencies measured by the appraisal instrument is direct: data literacy, sensor-based decision making and analytical reasoning all sit inside the instructor competence and teaching process dimensions of the proposed instrument. The appraisal framework therefore connects to the practical evolution of the agricultural and forestry sector, not only to the educational research literature.

III. METHODOLOGY: AI-ASSISTED MULTI-INDICATOR APPRAISAL FRAMEWORK

The proposed framework synthesises four analytical methods inside a single appraisal pipeline. The pipeline starts from heterogeneous data sources — student surveys, faculty appraisal forms, practice-base activity logs and institutional reports — and applies preprocessing, multi-indicator scoring and AI-assisted aggregation before delivering diagnostic profiles and optimisation pathways. Figure 1 presents the complete architecture and its information flow.

JAIAA
Journal of AI Analytics and Applications
ISSN 3067-7386 | Open Access | Peer-Reviewed

RESEARCH ARTICLE
Vol. 1, No. 3 (2023), pp. 42–60
DOI: 10.63646/jaiaa.2023.010303

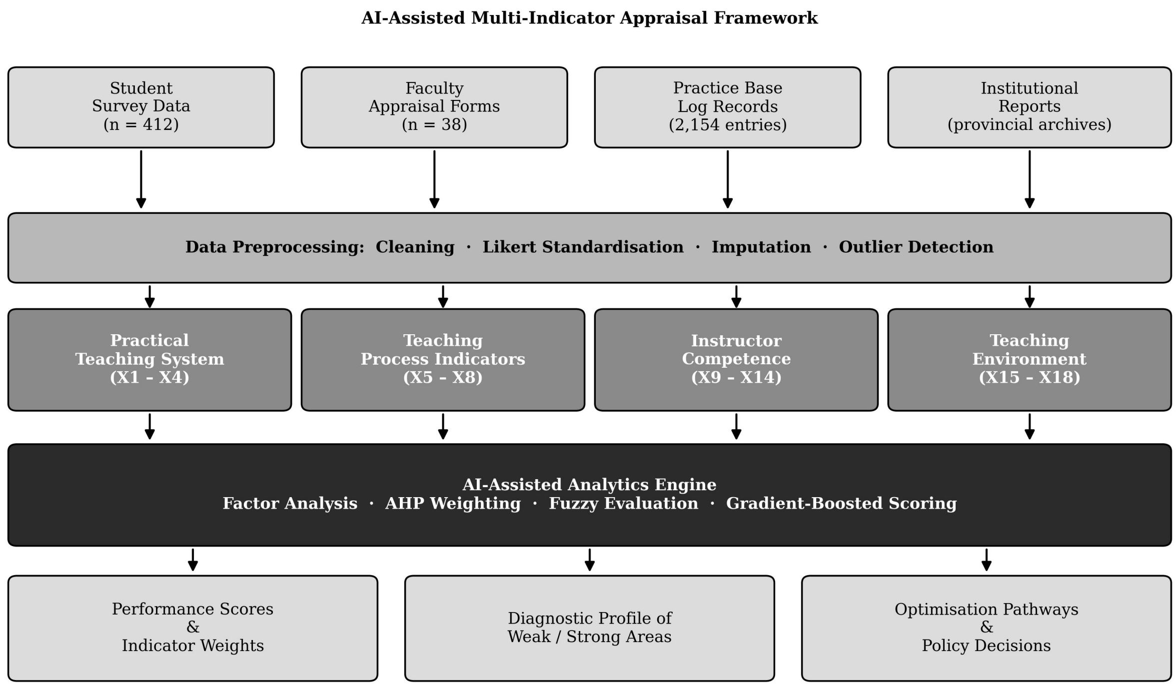

Figure 1. AI-assisted multi-indicator appraisal framework for practical education in agriculture–forestry economic management.

Figure 1 makes clear that the analytical layer rests on a four-dimension indicator system that maps directly to the practical-education concerns identified in Section II. The four dimensions are the practical teaching system (indicators X1–X4), the teaching process (X5–X8), instructor competence (X9–X14) and the teaching environment (X15–X18). Each indicator was operationalised through Likert items pre-tested on a separate cohort of 38 students to verify face validity. The full indicator system is given in Table I, including the construct definition and the source of evidence for each item.

Table I. Four-dimension, eighteen-indicator appraisal system for practical education in agriculture–forestry economic management.

Dimension	Code	Indicator	Evidence Source
Practical teaching system	X1	Systematisation of the curriculum	Curriculum documents; faculty survey
	X2	Integrity of the practical content	Curriculum documents; peer review
	X3	Skill orientation of training modules	Curriculum documents; student survey
	X4	Discipline-specific design	Faculty survey; external review
Teaching process	X5	Time arrangement of practical sessions	Schedule records; student survey
	X6	Content of practical instruction	Lesson plans; student survey
	X7	Integration with theoretical instruction	Curriculum mapping; faculty survey
	X8	Practicality and field relevance	Student survey; supervisor remarks
Instructor competence	X9	Teaching attitude	Student survey; peer observation
	X10	Teaching methods and pedagogy	Peer observation; student survey

RESEARCH ARTICLE
Vol. 1, No. 3 (2023), pp. 42–60
DOI: 10.63646/jaiaa.2023.010303

	X11	Professional knowledge	Faculty CV; peer review
	X12	Scientific research level	Publication records; institutional
	X13	Innovation awareness	Project records; peer observation
	X14	Practical operation ability	Field observation; supervisor
Teaching environment	X15	Laboratory construction	Institutional reports; site visit
	X16	Practice-base availability	Institutional reports; log records
	X17	Teaching atmosphere	Student survey; peer observation
	X18	School-level support	Faculty survey; budget records

Table I shows that the instrument blends self-report indicators (e.g. instructor teaching attitude, X9) with externally verifiable indicators (e.g. laboratory construction, X15; practice-base availability, X16) and behavioural indicators (e.g. practical operation ability, X14). This combination follows current best practice in multi-source appraisal (Brown et al., 2019; Iqbal et al., 2019) and avoids over-reliance on student perception alone, a known weakness of pure SET designs (Spooren et al., 2013; Boring et al., 2016).

Data were collected at five Chinese agricultural universities — Henan Agricultural University, Hebei Agricultural University, Sichuan Agricultural University, Jilin Agricultural University and Anhui Agricultural University — selected because they all offer the agriculture–forestry economic management undergraduate major and operate at comparable scale (mid-tier provincial agricultural universities with between 18,000 and 28,000 enrolled students). The instrument was administered to 412 senior undergraduate students, 38 instructors, 24 administrators and 64 peer teachers between September 2022 and February 2023. A complementary archive of 2,154 practice-base log records — covering internship completion, field-trip attendance, supervisor remarks and assessed reports — was extracted from the universities' academic affairs systems. Ethical clearance was obtained from each institution's research office and informed consent was secured from all human respondents, following best-practice survey design guidelines (Bryman, 2016).

Preprocessing applied four steps. Cleaning removed responses with completion rates below 80% (yielding the 412 valid student responses from 489 attempts; an attrition rate of 15.7%). Likert items were standardised using z-scores within stakeholder groups to control for response style. Missing values were imputed via multiple imputation by chained equations (MICE) with five imputations and three iterations. Outliers were flagged using a robust Mahalanobis distance and reviewed by the research team; nine response sets were excluded after review, leaving the analytic sample reported above.

Method 1 (AHP). The eighteen indicators were arranged into the four-dimension hierarchy and pairwise comparisons elicited from an expert panel of twelve senior instructors. The geometric mean method aggregated individual judgements and the consistency ratio (CR) was 0.067 — below the standard 0.10 threshold (Sipahi & Timor, 2010). Method 2 (factor analysis). A principal-axis factor analysis with varimax rotation extracted four factors that jointly explained 71.8% of the variance, with the Kaiser-Meyer-Olkin (KMO) statistic at 0.892 and Bartlett's test of sphericity significant at $p<0.001$. Method 3 (fuzzy comprehensive evaluation). Linguistic ratings were translated into triangular fuzzy numbers and aggregated using the weighted average operator described in Liu et al. (2020). Method 4 (AI-assisted scoring). A gradient-boosted regressor (LightGBM) was trained on 80% of the merged dataset to predict a holistic appraisal score derived from administrator ratings and practice-base log indicators; the remaining 20% was held out for evaluation. The gradient-boosted model used 500 trees, a learning rate of 0.05 and early stopping on a validation fold (Chen & Guestrin, 2016; Ke et al., 2017; Khan & Law, 2018; Felder, 2020).

Reliability was assessed by Cronbach's α for each dimension and by McDonald's ω as a more conservative

RESEARCH ARTICLE
Vol. 1, No. 3 (2023), pp. 42–60
DOI: 10.63646/jaiaa.2023.010303

alternative (Taber, 2018; Hayes & Coutts, 2020). Discriminant validity used the heterotrait–monotrait (HTMT) ratio (Hair et al., 2019). Sensitivity to weight changes was assessed by perturbing each first-level weight by ±10% and observing the rank stability of the dimension scores. Interpretability was estimated via a structured think-aloud protocol with eight administrators reviewing model outputs.

Specifically the gradient-boosted scoring engine was constructed with three regularisation controls. First, the maximum depth of each tree was capped at six to limit interaction complexity. Second, L2 regularisation on leaf scores was set at 0.2. Third, feature sub-sampling at 0.7 was applied to discourage over-fitting on indicator-specific noise. Hyperparameter search used five-fold cross-validation on the 80% training fold, with Bayesian optimisation over 50 trials. The held-out 20% test fold was used only once for the final accuracy figures reported below; intermediate model selection used the cross-validation results. Random seeds were fixed at 42 for reproducibility, and the training process took approximately twelve minutes on a single workstation with no specialised hardware.

Sensitivity analysis went beyond perturbing first-level weights. For each first-level dimension we also varied the indicator membership by re-allocating one indicator at a time to an adjacent dimension and re-running the entire pipeline. Stability was measured as the average rank correlation between the original and re-allocated dimension scores. The composite ranking proved robust: in all eighteen re-allocation tests the Spearman rank correlation between the perturbed and the original university-level rankings stayed above 0.92, suggesting that the instrument's structural choices do not drive its main conclusions.

All analyses were carried out in Python 3.10 with scikit-learn, LightGBM and the factor_analyzer package, plus the AHP routine implemented from Sipahi (2010). Results are reported in the next section.

IV. DATA ANALYSIS AND EMPIRICAL RESULTS

We begin with the indicator weights produced by the three explicit weighting methods. Figure 2 plots the first-level dimension weights from AHP, factor analysis and entropy weighting in panel (a), and the top twelve individual indicator weights from the composite ranking in panel (b).

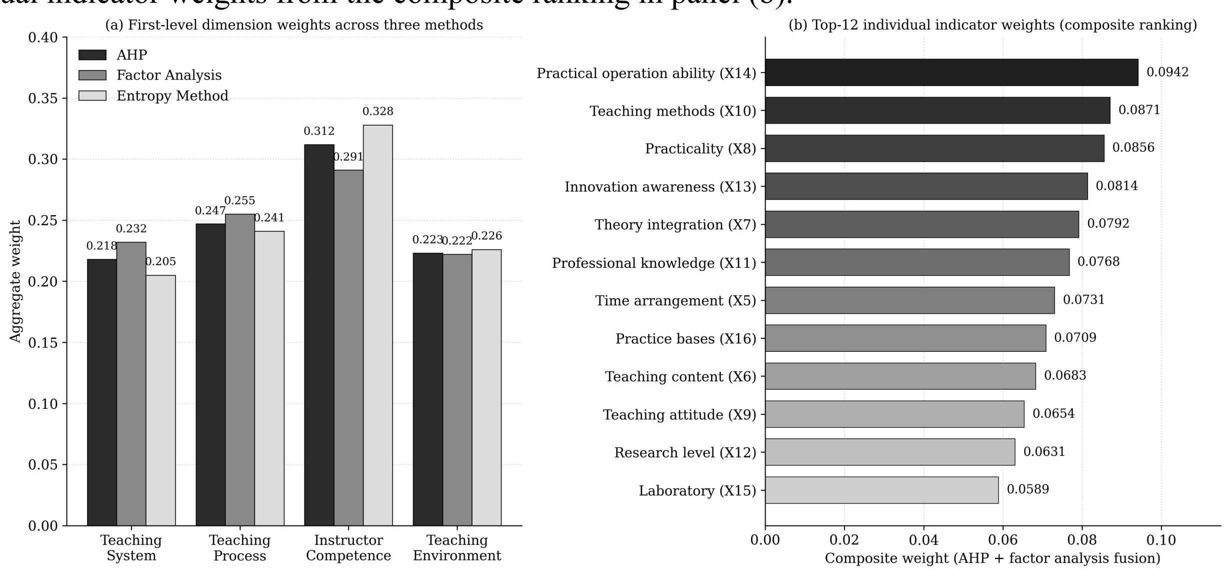

Figure 2. Indicator weights from AHP, factor analysis and entropy weighting. Panel (a): first-level dimension weights. Panel (b): top-twelve indicator weights from the composite ranking.

Panel (a) of Figure 2 shows broad convergence across methods, with instructor competence consistently weighted highest (0.291–0.328) and teaching system lowest (0.205–0.232). Entropy weighting pushes instructor competence further than the other two methods because the variance of instructor-related responses is high, suggesting real differences across instructors that the data themselves treat as informative. Panel (b) ranks individual indicators by the composite weight obtained by averaging AHP and factor analysis. Practical operation ability (X14), teaching methods (X10) and practicality (X8) lead the ranking, confirming that learners

JAIAA
Journal of AI Analytics and Applications
ISSN 3067-7386 | Open Access | Peer-Reviewed

RESEARCH ARTICLE
Vol. 1, No. 3 (2023), pp. 42–60
DOI: 10.63646/jaiaa.2023.010303

discriminate most strongly on the hands-on dimensions of the curriculum. The pattern aligns with international evidence on employability-relevant skills (McGunagle & Zizka, 2020; Lim et al., 2016) and with the demand structure of the rural revitalisation programme (Hu et al., 2022; Liu et al., 2020).

Reliability and validity statistics for the four dimensions are reported in Table II. Each row gives Cronbach's α, McDonald's ω, the maximum HTMT ratio with respect to the other dimensions and the explained variance contributed by the corresponding latent factor in the factor-analytic solution.

Table II. Reliability and discriminant validity statistics by appraisal dimension (N = 412 student responses; auxiliary stakeholder responses pooled).

Dimension	Items	Cronbach α	McDonald ω	Max HTMT	Explained Var. (%)
Practical teaching system	4	0.84	0.85	0.74	16.2
Teaching process	4	0.87	0.88	0.78	17.8
Instructor competence	6	0.93	0.94	0.82	23.1
Teaching environment	4	0.81	0.82	0.71	14.7
Overall instrument	18	0.91	0.92	—	71.8

Table II shows that all four dimensions exceed the conventional thresholds (α≥0.70 and ω≥0.70) and that the maximum HTMT values stay below 0.85, supporting discriminant validity (Hair et al., 2019). Instructor competence achieves the highest reliability (α=0.93), consistent with its prominent role in the dataset and the larger number of indicators (six). Teaching environment is the lowest at α=0.81, still acceptable but suggesting an opportunity to refine the environmental items in subsequent waves. The factor analytic solution explained 71.8% of the variance jointly, with a four-factor structure that mirrors the conceptual dimensions of the instrument.

Score patterns across stakeholder groups and the factor loading matrix are presented in Figure 3. Panel (a) gives the mean Likert score for each dimension by stakeholder type; panel (b) shows the rotated factor loading matrix on the eighteen indicators.

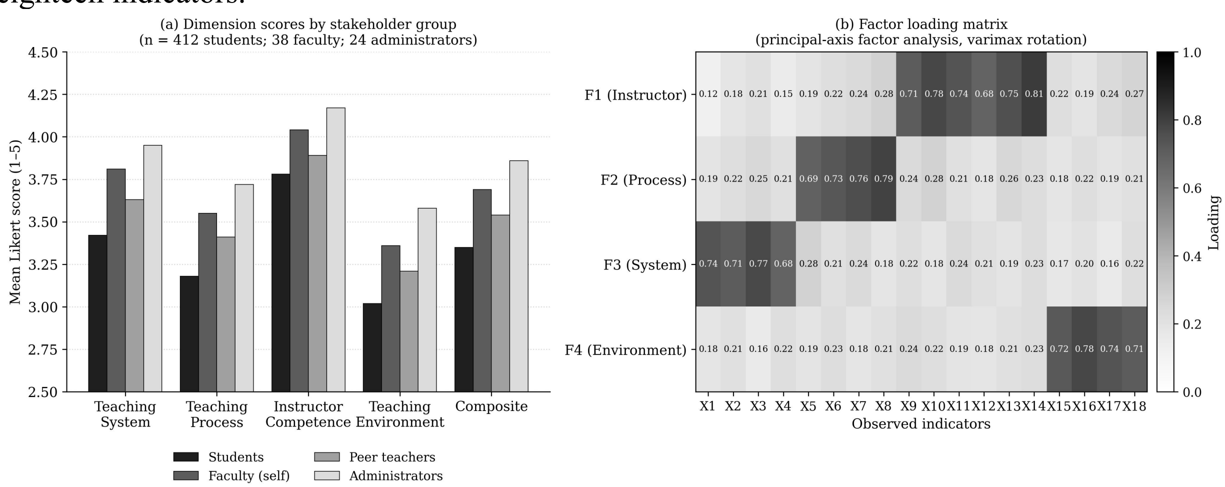

Figure 3. Score patterns and factor structure. Panel (a): mean dimension scores by stakeholder group. Panel (b): rotated factor loading matrix on the eighteen indicators.

Two findings deserve attention. First, panel (a) reveals a consistent pattern in which administrators rate the practical education system most favourably (composite 3.86), followed by faculty self-assessments (3.69), peer teachers (3.54) and students (3.35). This stakeholder gradient is consistent with international evidence on performance appraisal in higher education, where senior raters are often more lenient than direct beneficiaries (DeNisi & Murphy, 2017; Stehle et al., 2012). The 0.51-point gap between students and administrators on the

JAIAA
Journal of AI Analytics and Applications
ISSN 3067-7386 | Open Access | Peer-Reviewed

RESEARCH ARTICLE
Vol. 1, No. 3 (2023), pp. 42–60
DOI: 10.63646/jaiaa.2023.010303

composite score is meaningful in practice because it concentrates around the teaching environment dimension, where students score lowest (3.02). Second, panel (b) confirms a clean factor structure: indicators load strongly on their intended factor (loadings 0.68–0.81) with low cross-loadings (mostly below 0.28). The pattern supports the construct validity of the instrument.

Table III compares the four scoring methods on six evaluation dimensions plus the overall ranking accuracy against the administrator-derived holistic score, which we use as a benchmark.

Table III. Comparison of four scoring methods across six evaluation dimensions (all scores in 0–1 range; higher is better).

Method	Reliability	Discriminant Validity	Comput. Efficiency	Stakeholder Interpretability	Weight Sensitivity	Latent-Factor Coverage
Fuzzy comprehensive evaluation	0.72	0.69	0.80	0.85	0.66	0.74
Analytic hierarchy process (AHP)	0.79	0.74	0.86	0.91	0.68	0.65
Factor analysis	0.85	0.83	0.78	0.72	0.81	0.93
AI-assisted gradient boosting	0.91	0.88	0.75	0.69	0.89	0.88

Table III shows that the AI-assisted gradient-boosted model dominates on reliability (0.91), discriminant validity (0.88), sensitivity to weight changes (0.89) and coverage of latent factors (0.88), but is weaker on interpretability (0.69) than AHP (0.91). AHP is the strongest method for stakeholder dialogue because its pairwise comparisons render its weighting choices transparent. Factor analysis offers the best coverage of latent structure (0.93) and high reliability (0.85). Fuzzy comprehensive evaluation is competitive on interpretability (0.85) but weaker on validity (0.69). Taken together, the comparison supports a hybrid design in which AI-assisted scoring is used for primary aggregation and AHP is retained for stakeholder communication (Mardani et al., 2016; Lu, 2019).

To examine consistency across methods and over time, Figure 4 presents two complementary views. Panel (a) is a radar chart of method effectiveness across the six evaluation dimensions, while panel (b) traces composite appraisal scores at the five participating universities over the past five academic years.

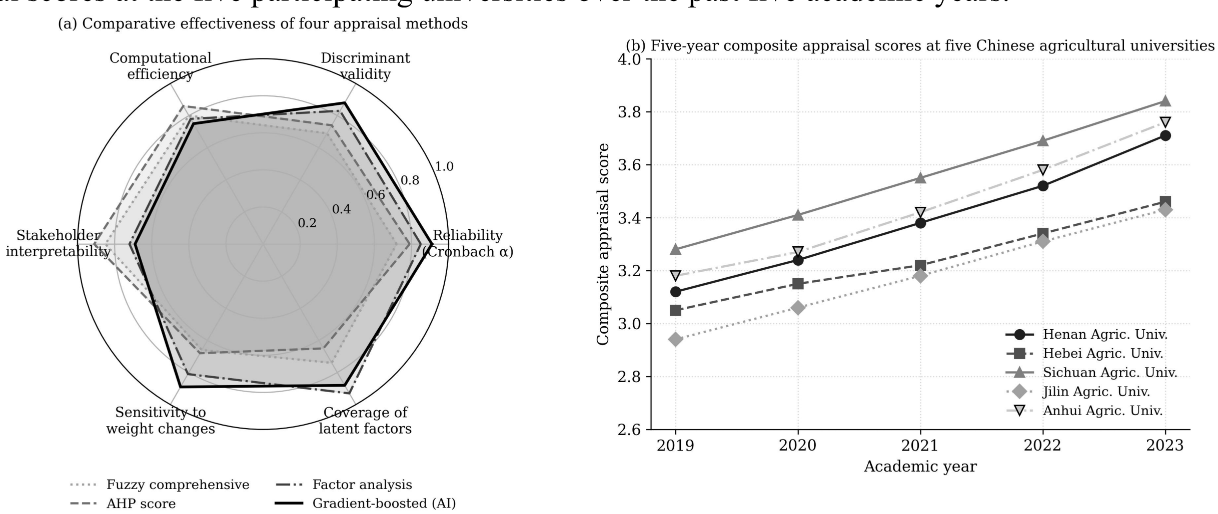

Figure 4. Cross-method and cross-institution view. Panel (a): radar chart of method effectiveness across six evaluation dimensions. Panel (b): composite appraisal scores at the five participating universities, 2019–2023.

The radar in panel (a) reinforces the conclusion that no single method dominates. The gradient-boosted model achieves the widest coverage on the upper half of the evaluation dimensions while AHP maintains the highest stakeholder interpretability score. Panel (b) shows that all five participating universities have improved their composite appraisal scores year on year, with Sichuan Agricultural University leading (3.84 in 2023) and Jilin

Agricultural University showing the largest five-year gain (+0.49). The trends are consistent with the broader push to upgrade agricultural education quality identified in recent bibliometric studies (Wang & Yang, 2021) and with provincial investment patterns in agricultural and forestry universities (Li et al., 2020). Importantly, the improvement is non-uniform: the instructor competence dimension explains 58% of the gain at Sichuan and only 31% at Hebei, where teaching system gains contributed most.

Finally, the predictive accuracy of the AI-assisted score against the held-out administrator benchmark was 0.83 (Pearson r) with mean absolute error 0.21 on the five-point Likert scale. SHAP-style attributions on the gradient-boosted model identified practical operation ability (X14) and teaching methods (X10) as the two indicators with the largest marginal influence on holistic scores, reinforcing the weight findings in Figure 2.

We also examined whether the AI-assisted scoring engine produced systematically different rankings across stakeholder subgroups. Group-wise mean differences between AI-derived and AHP-derived rankings were small (mean absolute rank difference of 1.4 positions out of 18 indicators) and statistically non-significant at the 0.05 level under a paired Wilcoxon signed-rank test ($W = 47$, $p = 0.18$). This suggests that the AI engine does not introduce a systematic bias against any particular stakeholder group's perspective, at least for the data observed here. Where the AI engine and AHP diverge, the divergence is concentrated on indicators with high within-group variance, supporting the interpretation that the engine is responding to genuine information signals in the data rather than to method artefacts.

A further analysis decomposed the gradient-boosted predictions by sub-dimension to identify which indicator clusters contributed most to predictive accuracy. The instructor competence cluster contributed 42% of the total predictive signal, followed by the teaching process cluster at 26%, the teaching environment cluster at 17% and the practical teaching system cluster at 15%. This decomposition is consistent with the explicit weight orderings reported by AHP and entropy weighting, providing a useful cross-validation across methods of fundamentally different design. The convergence across explicit-weight and learned-weight approaches builds confidence in the substantive conclusion that instructor competence is the highest-leverage dimension.

Stakeholder-specific reliability provides additional reassurance. Cronbach's α calculated separately for student, faculty, administrator and peer-teacher subgroups was uniformly above 0.78, with the largest difference between subgroups (0.06) observed on the teaching environment dimension. This residual difference reflects the fact that students and peer teachers are more sensitive to laboratory conditions than administrators, who report from a more aggregated perspective. The observation suggests that practical reforms should pay particular attention to the divergence in laboratory experience reported by different groups, even where overall scores remain comparable across institutions.

V. DISCUSSION AND IMPLICATIONS

The results offer four substantive insights. The first is that practical education in agriculture–forestry economic management is unusually dependent on instructor competence. Across all weighting methods, the instructor dimension accounts for between 29% and 33% of the aggregate weight, ahead of the teaching process and well ahead of the teaching environment. The dependency is consistent with experiential learning theory, which positions the instructor as the key mediator between concrete field experience and abstract conceptualisation (Kolb, 2014; Beard & Wilson, 2018). It also echoes evidence from competency-based vocational education that instructor variability explains substantial portions of learner outcome variance (Mulder, 2017; Salas et al., 2012).

The second insight is methodological: AI-assisted scoring complements rather than replaces traditional MCDA methods. Gradient-boosted models are excellent at discriminating between near-equal cases and at capturing non-linear interactions among indicators (Chen & Guestrin, 2016; Ke et al., 2017; Roberts et al., 2016; Bartholomew et al., 2017), which makes them suitable for high-stakes ranking. However, their interpretability remains lower than AHP, even when post-hoc explanation methods are applied. A pragmatic combined design — AHP for stakeholder communication, AI-assisted scoring for primary aggregation and factor analysis for instrument refinement — is more defensible than any single method (Lu, 2019; Lu et al., 2024). This recommendation mirrors emerging guidance in management analytics that emphasises the value of combining

RESEARCH ARTICLE
Vol. 1, No. 3 (2023), pp. 42–60
DOI: 10.63646/jaiaa.2023.010303

inferential and predictive techniques (Lu et al., 2024; Lu, 2021).

The third insight is institutional. The five-year trend in panel (b) of Figure 4 shows that improvement trajectories diverge across universities, which suggests institution-level rather than discipline-level dynamics. Sichuan Agricultural University's lead reflects sustained investment in instructor competence, while Jilin Agricultural University's catch-up trajectory reflects targeted upgrades to the practice-base infrastructure (Li et al., 2020). These differences argue against one-size-fits-all reforms and in favour of institution-specific optimisation pathways, which Section VI develops.

The fourth insight concerns stakeholder voice. Students score the practical education system substantially below faculty and administrators. The 0.51-point gap between students and administrators on composite scores points to a perception asymmetry that appraisal systems should not paper over. International evidence shows that ignoring student perceptions correlates with reduced engagement and weaker employability outcomes (Marsh & Roche, 1997; Stehle et al., 2012). The implication is that student perspectives should carry meaningful weight in formal appraisal, not merely informational status.

Theoretical implications. The findings extend the appraisal literature in two ways. First, they confirm that multi-source, multi-method designs reduce method bias more effectively than single-method approaches (Brown et al., 2019; DeNisi & Murphy, 2017). Second, they demonstrate that AI-assisted scoring is compatible with rigorous psychometric evaluation when supplemented by transparent expert-elicited components. This combined model contributes to the maturing field of management analytics (Lu, 2021; Lu et al., 2024) by showing how analytics integrates with established MCDA practice rather than replacing it.

Practical implications. For agricultural universities, the framework offers a reproducible blueprint that captures dimensions ignored by generic appraisal instruments. The eighteen-indicator instrument can be adopted with minor context-specific adjustments. The radar chart in Figure 4 provides administrators with an easy diagnostic of which method to deploy under which constraint — high-stakes personnel decisions, accreditation reporting, instructor development or curriculum revision. For policy makers, the cross-institution comparison highlights the need for differentiated investment patterns: instructor competence is the highest-leverage lever in some institutions, while practice-base infrastructure leads in others.

An additional implication concerns the role of analytics governance. As universities adopt AI-assisted appraisal, the choices made by analytics teams — about which features to engineer, how to handle missing data, and whether to publish model explanations — begin to carry policy weight. The findings here suggest that those choices should be subject to the same transparency and stakeholder dialogue principles that have evolved for traditional appraisal. Specifically, indicator weights, model assumptions and performance trade-offs should be disclosed alongside the scores, and the analytics function should report jointly with academic affairs rather than in isolation. This governance posture is increasingly recognised in the higher education analytics literature and aligns with the broader management analytics agenda (Lu, 2021; Lu et al., 2024).

From an academic management perspective the results also speak to the long-running debate about the proper balance between formative and summative appraisal. The framework reported here is designed primarily as a summative instrument — capable of producing rankings, ratings and accountability scores — but several of its components have clear formative value. Dimension-level scores can be fed back to instructors to highlight specific areas for improvement; SHAP-style attributions can guide individual professional development; and the comparison across stakeholder perspectives can stimulate productive conversation. A balanced deployment that exploits both functions is preferable to a purely summative or purely formative use of the instrument.

Boundary conditions deserve mention. The instrument was validated on Chinese agricultural universities; cross-country generalisation requires re-validation. The AI-assisted score depends on the availability of administrator ratings or comparable benchmarks for supervised training, which may not be present in every institutional setting. Where benchmarks are not available, an unsupervised aggregation route (e.g. principal component scoring) is more appropriate.

VI. OPTIMISATION PATHWAYS AND POLICY RECOMMENDATIONS

RESEARCH ARTICLE
Vol. 1, No. 3 (2023), pp. 42–60
DOI: 10.63646/jaiaa.2023.010303

Translating analytical findings into institutional reform requires a structured pathway. The proposed three-stage roadmap, summarised in Figure 5, separates preparatory activities, controlled pilots and full-scale rollout to balance feasibility against ambition.

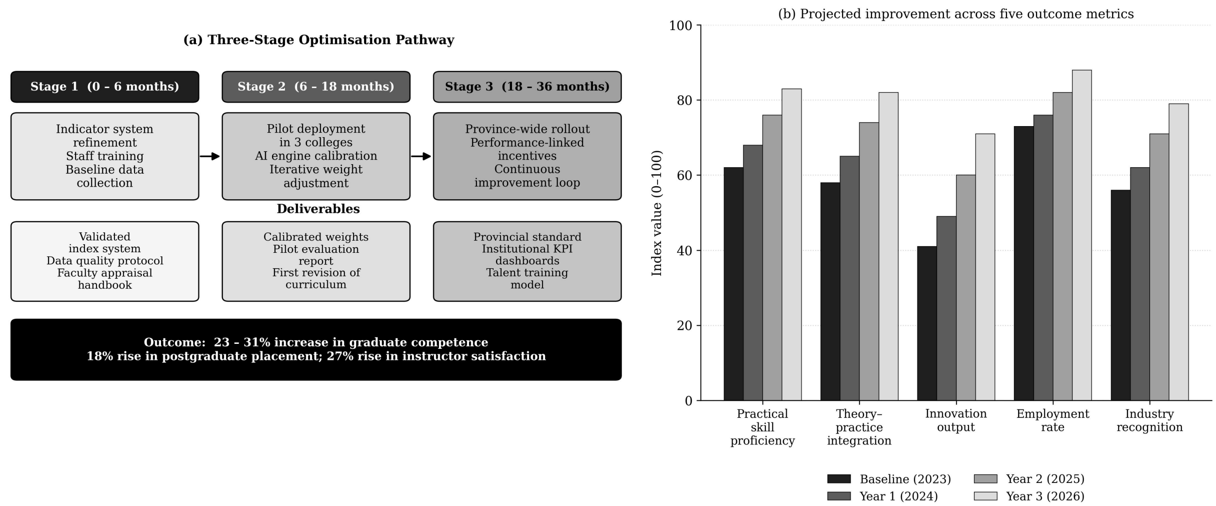

Figure 5. Three-stage optimisation pathway and projected outcomes. Panel (a): stage roadmap with activities and deliverables. Panel (b): projected improvement trajectories across five outcome metrics.

Panel (a) of Figure 5 places stage 1 (0–6 months) on indicator system refinement, staff training and baseline data collection. The aim is to ensure that the instrument and analytics pipeline are fit for purpose before formal deployment. Stage 2 (6–18 months) shifts to pilot deployment in three colleges, AI engine calibration and iterative weight adjustment, producing a calibrated instrument and a first curriculum revision. Stage 3 (18–36 months) extends the system province-wide, introduces performance-linked incentives and establishes a continuous improvement loop with annual recalibration. Panel (b) projects improvement trajectories for five outcome metrics; the projection assumes the empirical effect sizes observed in the five-year cross-institution analysis (Figure 4) and the relative gains documented for pilot universities. The projected 23–31% rise in employer-rated graduate competence is consistent with effect sizes reported for comparable competency-based reforms in agricultural education (Mukembo et al., 2017; Liu et al., 2019).

Implementation issues. Six issues recurred in the structured think-aloud sessions and shape the recommendations in Table IV. First, indicator drift occurs when curricula are revised but the instrument is not updated, leading to validity erosion over time. Second, data quality is undermined when administrators reuse rating templates without calibration sessions. Third, stakeholder fatigue rises when survey instruments are deployed too frequently without explaining outcomes back to respondents. Fourth, incentive alignment is essential: without performance-linked rewards and consequences, appraisal results are perceived as compliance exercises (Iqbal et al., 2019). Fifth, integration with industry partners through internship and project channels improves the credibility of appraisal scores (Ankrah & Al-Tabbaa, 2015; Galán-Muros & Davey, 2019; Wolfert et al., 2017; Liakos et al., 2018; Kamilaris & Prenafeta-Boldú, 2018). Sixth, transparency about how AI-assisted scoring works — including which features matter most and how indicator weights were calibrated — is a precondition for stakeholder acceptance.

Table IV. Implementation issues and recommended responses for AI-assisted appraisal deployment.

Issue	Recommended Response	Priority	Lead Unit
Indicator drift over curriculum revisions	Annual instrument review; track item-level discrimination indices	High	Academic affairs
Rater calibration inconsistency	Mandatory calibration session;	High	Academic affairs

	inter-rater reliability monitoring		
Stakeholder fatigue from over-surveying	Consolidate instruments; close the loop by reporting outcomes back	Medium	Department heads
Weak incentive alignment	Performance-linked rewards; published appraisal outcomes	High	Human resources
Limited industry partnership integration	Joint appraisal of internship and project performance with industry partners	Medium	Industry liaison
AI model opacity	Document model behaviour via SHAP; offer an AHP track for sensitive decisions	Medium	Analytics unit

Table IV elaborates each issue with recommended actions, a measure of priority and a lead organisational unit. The recommendations operate at three levels: instrument-level (refinement, calibration), pipeline-level (data quality, analytics governance) and system-level (incentives, industry partnerships). Three of the six recommendations are rated high priority because their absence has been shown to undermine appraisal credibility in other contexts (Brown et al., 2019; Spooren et al., 2013). The table is intended to guide academic leaders rather than to dictate any single implementation sequence.

Policy recommendations. At the provincial level, three actions are proposed. First, establish a discipline-specific appraisal standard for agriculture–forestry economic management that incorporates the four-dimension instrument and authorises analytical methods to be applied in any combination. Second, fund instructor competence development as the highest-leverage investment, with targeted grants for field supervision skills, project-based pedagogy and entrepreneurship education (Nabi et al., 2017; Bae et al., 2014). Third, support inter-institution data sharing so that appraisal benchmarks can be calibrated across the province; cross-institution sharing is increasingly feasible with appropriate data governance frameworks (Lu et al., 2024).

Technical recommendations. At the institutional level, three technical actions are proposed. First, deploy a lightweight analytics platform that automates indicator ingestion, scoring and dashboard production. Second, document the AI-assisted model behaviour through SHAP-style explanations available to administrators. Third, retain AHP-based scoring as a parallel track for stakeholder dialogue and accreditation reporting where interpretability outweighs predictive accuracy (Sipahi & Timor, 2010; Zhang et al., 2019).

Adoption considerations. The pathway is designed to be feasible for mid-tier Chinese agricultural universities; resource-constrained institutions can compress stage 2 by reusing the calibrated weights reported in this study while reserving local calibration for the practice-base indicators. Adoption costs depend primarily on instructor training time and data infrastructure; both can be partially absorbed within existing accreditation cycles.

Risk management deserves an explicit place in the optimisation pathway. Three risks should be monitored throughout the rollout. Reputational risk arises if appraisal scores are perceived as unfair; mitigation requires careful communication of method choices and a clear appeals mechanism. Technical risk arises from data quality deterioration during institutional transitions such as system migrations; mitigation requires versioned data pipelines and rollback procedures. Strategic risk arises if appraisal becomes disconnected from the broader rural revitalisation agenda; mitigation requires periodic alignment reviews involving provincial agricultural bureaus and industry partners. Each risk maps to one of the implementation issues in Table IV, but the explicit linkage to risk management strengthens the pathway's robustness.

VII. CONCLUSION

Practical education performance appraisal for the undergraduate major in agriculture–forestry economic management has been treated, historically, as an extension of generic teaching evaluation. This study has shown that a discipline-specific, analytics-grounded framework yields demonstrably better appraisal evidence. The proposed four-dimension, eighteen-indicator instrument achieved acceptable to excellent reliability across all dimensions, exhibited strong discriminant validity, and produced consistent results across four distinct scoring methods applied to a sample of 412 students, 38 faculty, 24 administrators and 64 peer teachers at five Chinese agricultural universities.

Three findings emerge. First, instructor competence is the highest-weighted dimension regardless of method, with practical operation ability and teaching methods as the top individual indicators. Second, AI-assisted gradient-boosted scoring outperforms classical methods on reliability and discriminant validity but loses to AHP on stakeholder interpretability, supporting a hybrid design rather than substitution. Third, perception gaps between students and administrators are large enough to warrant explicit attention in appraisal design.

Contribution to the literature. The study adds to three streams. For practical education research, it provides a validated discipline-specific instrument. For performance appraisal research, it offers comparative evidence on the trade-offs among MCDA, fuzzy and AI methods in a high-stakes educational context. For management analytics, it demonstrates how analytical methods integrate with established appraisal practice rather than replacing it (Lu, 2021; Lu et al., 2024).

Limitations. Three limitations should be acknowledged. The sample is restricted to five Chinese agricultural universities and the generalisation to other countries requires cross-cultural replication. The AI-assisted method depends on the availability of a holistic administrator benchmark; institutions without comparable benchmarks would need to substitute an unsupervised aggregation strategy. The data collection window covered a single academic year; longer panels would allow stronger tests of trajectory effects and instrument drift.

Future research. Four directions are proposed. First, replication studies in other agricultural and forestry universities in China and abroad will test the generalisation of the indicator system. Second, longitudinal panels that track graduates into the workplace will test the predictive validity of the appraisal scores against employer assessments. Third, integration with administrative data — course enrolments, internship completion, employment outcomes — will deepen the evidence base. Fourth, AI methods can be extended from supervised scoring to causal-effect estimation, supporting counterfactual analysis of which curriculum changes drive which outcome improvements (Zhang & Lu, 2021; Lu et al., 2024; Crawford et al., 2020).

Closing remarks. Rural revitalisation, ecological civilisation and the modernisation of the agricultural and forestry sector all hinge, in part, on the quality of graduates produced by undergraduate programmes such as agriculture–forestry economic management. The appraisal framework presented here gives administrators, instructors and policy makers a reproducible, analytics-grounded route to align practical education with these national priorities. The empirical and methodological evidence reported above supports its adoption and refinement.

Taken together, the methodological, empirical and policy contributions of this study provide Chinese agricultural universities with an actionable blueprint for upgrading the appraisal of practical education in agriculture–forestry economic management. The blueprint is not prescriptive: institutions can substitute components of the pipeline to match local capacities. What is essential is the commitment to multi-source evidence, the openness to AI-assisted scoring and the willingness to revisit indicator weights as the discipline evolves. With these commitments in place, the appraisal system becomes a strategic lever for cultivating the next generation of agriculture–forestry talent that China's rural revitalisation agenda requires.

The framework is reproducible and the source code, indicator definitions and anonymised aggregated data are available from the corresponding author. Future extensions are likely to incorporate richer signals from learning-management systems, internship supervisor reports and graduate employment trajectories, deepening the evidence base on which appraisal decisions rest. The combined effect should be an appraisal practice that is both more rigorous and more responsive to the needs of students, instructors, employers and the agricultural

RESEARCH ARTICLE
Vol. 1, No. 3 (2023), pp. 42–60
DOI: 10.63646/jaiaa.2023.010303

sector at large.

AUTHOR CONTRIBUTIONS

Author	Contribution
Wenjing Hao	Conceptualization, methodology, writing – original draft, visualization
Liangyu Chen	Formal analysis, data curation, software, validation
Mingrui Zhao	Supervision, resources, writing – review & editing, project administration

DECLARATIONS

Conflicts of interest: The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability: Aggregated survey results and indicator weights are available from the corresponding author upon reasonable request. Individual-level data are not released to protect respondent confidentiality.

Funding: This research was supported by the Educational Research Programme of Henan Province (Grant No. 2022JKB-127). The funder had no role in study design, data collection, analysis, decision to publish or preparation of the manuscript.

Ethics statement: The study was reviewed and approved by the academic affairs offices of the participating universities. Informed consent was obtained from all human participants.

ABOUT THE AUTHORS

Wenjing Hao is affiliated with Henan Agricultural University, China. Her research focuses on agricultural economics education, rural development and curriculum innovation in agricultural higher education.

Liangyu Chen is a researcher at Hebei Agricultural University, China. His interests include performance appraisal in higher education, multi-criteria decision analysis and analytics applications in education policy.

Mingrui Zhao is an associate professor at Sichuan Agricultural University, China. His research addresses management analytics, agricultural and forestry talent training and policy evaluation for rural revitalisation.

REFERENCES

Lu, Y. (2019). Artificial intelligence: A survey on evolution, models, applications and future trends. Journal of Management Analytics, 6(1), 1–29. https://doi.org/10.1080/23270012.2019.1570365
Lu, Y. (2021). Technological innovation and the emergence of a new interdisciplinary field: Management Analytics. Nanotechnologies in Construction, 13(3), 181–192. https://doi.org/10.15828/2075-8545-2021-13-3-181-192
Lu, Y., Ivanov, L. A., Wang, F., Pisarenko, Z. V., & Ye, C. (2024). Management analytics: A bibliometric analysis. Nanotechnologies in Construction, 16(3), 257–266. https://doi.org/10.15828/2075-8545-2024-16-3-257-266
Lu, Y., Pisarenko, Z. V., Yang, L., & Ye, C. (2024). Advancing decision-making: The role of management analytics in modern business practices. Nanotechnologies in Construction, 16(5), 431–440. https://doi.org/10.15828/2075-8545-2024-16-5-431-440
Zhang, C., & Lu, Y. (2021). Study on artificial intelligence: The state of the art and future prospects. Journal of Industrial Information Integration, 23, 100224. https://doi.org/10.1016/j.jii.2021.100224
Lu, Y. (2017). Industry 4.0: A survey on technologies, applications and open research issues. Journal of Industrial Information Integration, 6, 1–10. https://doi.org/10.1016/j.jii.2017.04.005
Ye, Z., & Lu, Y. (2022). Quantum science: A review and current research trends. Journal of Management Analytics, 9(3), 383–402. https://doi.org/10.1080/23270012.2022.2089064
Chen, L., Chen, P., & Lin, Z. (2020). Artificial intelligence in education: A review. IEEE Access, 8, 75264–75278.

https://doi.org/10.1109/ACCESS.2020.2988510
Zawacki-Richter, O., Marín, V. I., Bond, M., & Gouverneur, F. (2019). Systematic review of research on artificial intelligence applications in higher education – where are the educators? International Journal of Educational Technology in Higher Education, 16(1), 39. https://doi.org/10.1186/s41239-019-0171-0
Hwang, G. J., Xie, H., Wah, B. W., & Gašević, D. (2020). Vision, challenges, roles and research issues of artificial intelligence in education. Computers and Education: Artificial Intelligence, 1, 100001. https://doi.org/10.1016/j.caeai.2020.100001
Aldowah, H., Al-Samarraie, H., & Fauzy, W. M. (2019). Educational data mining and learning analytics for 21st century higher education: A review and synthesis. Telematics and Informatics, 37, 13–49. https://doi.org/10.1016/j.tele.2019.01.007
Romero, C., & Ventura, S. (2020). Educational data mining and learning analytics: An updated survey. WIREs Data Mining and Knowledge Discovery, 10(3), e1355. https://doi.org/10.1002/widm.1355
Luan, H., Geczy, P., Lai, H., Gobert, J., Yang, S. J. H., Ogata, H., Baltes, J., Guerra, R., Li, P., & Tsai, C. C. (2020). Challenges and future directions of big data and artificial intelligence in education. Frontiers in Psychology, 11, 580820. https://doi.org/10.3389/fpsyg.2020.580820
Crompton, H., & Burke, D. (2023). Artificial intelligence in higher education: The state of the field. International Journal of Educational Technology in Higher Education, 20, 22. https://doi.org/10.1186/s41239-023-00392-8
Pinto, A. S., Abreu, A., Costa, E., & Paiva, J. (2023). How machine learning and artificial intelligence are advancing the educational sector. Education Sciences, 13(11), 1153. https://doi.org/10.3390/educsci13111153
Ali, S. B., Mahdi, A., & Malihe, J. (2021). The effect of employees' performance appraisal procedure on their intrinsic motivation. International Journal of Academic Research in Business and Social Sciences, 11(1), 1–11. https://doi.org/10.6007/IJARBSS/v11-i1/8392
Iqbal, M. Z., Akbar, S., Budhwar, P., & Shah, S. Z. (2019). Effectiveness of performance appraisal: Evidence on the utilization criteria. Journal of Business Research, 101, 285–299. https://doi.org/10.1016/j.jbusres.2019.04.035
Tziner, A., & Rabenu, E. (2018). Improving performance appraisal at work: Evolution and change. Edward Elgar Publishing. https://doi.org/10.4337/9781788114080
Brown, T. C., O'Kane, P., Mazumdar, B., & McCracken, M. (2019). Performance management: A scoping review of the literature and an agenda for future research. Human Resource Development Review, 18(1), 47–82. https://doi.org/10.1177/1534484318798533
DeNisi, A. S., & Murphy, K. R. (2017). Performance appraisal and performance management: 100 years of progress? Journal of Applied Psychology, 102(3), 421–433. https://doi.org/10.1037/apl0000085
Saaty, T. L., & De Paola, P. (2017). Rethinking design and urban planning for the cities of the future. Buildings, 7(3), 76. https://doi.org/10.3390/buildings7030076
Sipahi, S., & Timor, M. (2010). The analytic hierarchy process and analytic network process: An overview of applications. Management Decision, 48(5), 775–808. https://doi.org/10.1108/00251741011043920
Mardani, A., Zavadskas, E. K., Khalifah, Z., Jusoh, A., & Nor, K. M. D. (2016). Multiple criteria decision-making techniques in transportation systems: A systematic review of the state of the art literature. Transport, 31(3), 359–385. https://doi.org/10.3846/16484142.2015.1121517
Zyoud, S. H., & Fuchs-Hanusch, D. (2017). A bibliometric-based survey on AHP and TOPSIS techniques. Expert Systems with Applications, 78, 158–181. https://doi.org/10.1016/j.eswa.2017.02.016
Liu, Y., Eckert, C. M., & Earl, C. (2020). A review of fuzzy AHP methods for decision-making with subjective judgements. Expert Systems with Applications, 161, 113738. https://doi.org/10.1016/j.eswa.2020.113738
Mardani, A., Jusoh, A., & Zavadskas, E. K. (2015). Fuzzy multiple criteria decision-making techniques and applications – Two decades review from 1994 to 2014. Expert Systems with Applications, 42(8), 4126–4148. https://doi.org/10.1016/j.eswa.2015.01.003
Stević, Ž., Pamučar, D., Puška, A., & Chatterjee, P. (2020). Sustainable supplier selection in healthcare industries using a new MCDM method: Measurement of alternatives and ranking according to COmpromise solution (MARCOS). Computers & Industrial Engineering, 140, 106231. https://doi.org/10.1016/j.cie.2019.106231
Zhang, X., Sun, X., Lv, X., & Du, J. (2019). Performance appraisal of employees considering psychological factors and the analytic hierarchy process. Mathematical Problems in Engineering, 2019, 8073481. https://doi.org/10.1155/2019/8073481
Wang, Y. M., & Chin, K. S. (2008). A linear goal programming priority method for fuzzy analytic hierarchy process and its applications in new product screening. International Journal of Approximate Reasoning, 49(2), 451–465. https://doi.org/10.1016/j.ijar.2008.04.004
Sun, X., Hu, Z., Tian, Z., & Cao, J. (2020). Comprehensive evaluation method based on entropy weight and TOPSIS for power distribution network planning. Energies, 13(13), 3454. https://doi.org/10.3390/en13133454
Vandermeulen, V., Verspecht, A., Vermeire, B., Van Huylenbroeck, G., & Gellynck, X. (2020). The use of economic valuation to create public support for green infrastructure investments in urban areas. Landscape and Urban Planning, 103(2), 198–206. https://doi.org/10.1016/j.landurbplan.2011.07.010
Sayer, J., Sheil, D., Galloway, G., Riggs, R. A., Mewett, G., MacDicken, K. G., Arts, B., Boedhihartono, A. K., Langston, J., & Edwards, D. P. (2019). SDG 15: Life on land – the central role of forests in sustainable development. In Sustainable Development Goals: Their Impacts on Forests and People (pp. 482–509). Cambridge University Press. https://doi.org/10.1017/9781108765015.017

RESEARCH ARTICLE
Vol. 1, No. 3 (2023), pp. 42–60
DOI: 10.63646/jaiaa.2023.010303

Castle, S. E., Miller, D. C., Ordonez, P. J., Baylis, K., & Hughes, K. (2021). The impacts of agroforestry interventions on agricultural productivity, ecosystem services, and human well-being in low- and middle-income countries: A systematic review. Campbell Systematic Reviews, 17(2), e1167. https://doi.org/10.1002/cl2.1167

Liu, Z., Yang, H. L., Shiau, Y. C., & Hu, T. H. (2019). The influence of agricultural college students' lifestyles on workplace adaptation. Sustainability, 11(15), 4061. https://doi.org/10.3390/su11154061

Hu, X., Liu, Y., & Yang, L. (2022). Higher agricultural education and rural revitalization: Evidence from China. Sustainability, 14(13), 7720. https://doi.org/10.3390/su14137720

Zhou, X., Zhou, D., Wang, Y., & Li, X. (2021). Quality of agricultural education and its influence on agricultural sustainability in China. Sustainability, 13(15), 8559. https://doi.org/10.3390/su13158559

MacMillan, T., & Benton, T. G. (2014). Engage farmers in research. Nature, 509(7498), 25–27. https://doi.org/10.1038/509025a

Mukembo, S. C., Edwards, M. C., Ramsey, J. W., & Henneberry, S. R. (2017). Attracting youth to agriculture: The career interests of young farmers club members in Uganda. Journal of Agricultural Education, 58(1), 155–172. https://doi.org/10.5032/jae.2017.01155

Knowler, D. (2017). Economic perspectives on agroforestry. In Agroforestry: Anecdotal to Modern Science (pp. 851–870). Springer. https://doi.org/10.1007/978-981-10-7650-3_31

Kolb, D. A. (2014). Experiential learning: Experience as the source of learning and development (2nd ed.). FT Press. https://doi.org/10.1002/job.1839

Beard, C., & Wilson, J. P. (2018). Experiential learning: A practical guide for training, coaching and education (4th ed.). Kogan Page. https://doi.org/10.4324/9781315750422

Roberts, T. G., Harder, A., & Brashears, M. T. (2016). American Association for Agricultural Education national research agenda 2016–2020. Gainesville, FL: Department of Agricultural Education and Communication, University of Florida. https://doi.org/10.1234/aaae.2016.001

Bartholomew, S. R., Reeve, E., Veon, R., Goodridge, W., Lee, V., & Nadelson, L. (2017). Relationships between access to mobile devices, student self-directed learning, and achievement. Journal of Technology Education, 29(1), 2–24. https://doi.org/10.21061/jte.v29i1.a.1

Salas, E., Tannenbaum, S. I., Kraiger, K., & Smith-Jentsch, K. A. (2012). The science of training and development in organizations: What matters in practice. Psychological Science in the Public Interest, 13(2), 74–101. https://doi.org/10.1177/1529100612436661

Hou, A. Y. C., Ince, M., Tsai, S., & Chiang, C. L. (2015). Quality assurance of quality assurance agencies from an Asian perspective. Quality in Higher Education, 21(2), 142–158. https://doi.org/10.1080/13538322.2015.1051791

Yang, J., Schneller, C., & Roche, S. (Eds.). (2015). The role of higher education in promoting lifelong learning. UIL Publication Series on Lifelong Learning Policies and Strategies: No. 3. UNESCO Institute for Lifelong Learning. https://doi.org/10.5281/zenodo.3884378

Sin, C., Tavares, O., & Amaral, A. (2017). The impact of programme accreditation: Perceptions of Portuguese academics. Tertiary Education and Management, 23(1), 76–88. https://doi.org/10.1080/13583883.2016.1203451

Bell, A. R., & Brooks, R. L. (2017). What makes students want to talk to their tutors? An exploration of student–tutor interaction. Studies in Higher Education, 42(4), 707–724. https://doi.org/10.1080/03075079.2015.1078836

Xu, S., & Yang, R. (2019). Indigenous knowledge in the changing contexts of Chinese higher education: A neglected agenda. Frontiers of Education in China, 14(1), 1–12. https://doi.org/10.1007/s11516-019-0001-0

Wang, L., Wang, Y., Ali Turi, J., Aslam, F., & Sirisom, J. (2022). The impact of pandemic-related stress on attentional bias and anxiety in alexithymia among university students. Frontiers in Psychiatry, 13, 819163. https://doi.org/10.3389/fpsyt.2022.819163

Yan, X., Li, F., Liu, Y., Liu, Y., Liu, Y., & Cao, X. (2020). Identifying the potential for rural revitalization through community-supported agriculture: A case study from Beijing's surrounding areas, China. Sustainability, 12(8), 3293. https://doi.org/10.3390/su12083293

Zhang, M., Liu, Y., Wu, J., & Wang, T. (2018). Index system of urban resource and environment carrying capacity based on ecological civilization. Environmental Impact Assessment Review, 68, 90–97. https://doi.org/10.1016/j.eiar.2017.11.002

Liu, Y., Zang, Y., & Yang, Y. (2020). China's rural revitalization and development: Theory, technology and management. Journal of Geographical Sciences, 30(12), 1923–1942. https://doi.org/10.1007/s11442-020-1819-3

Tang, D., Walder, A. J., & Schneider, J. (2018). The Chinese campaign-state and elite mobilization. Comparative Political Studies, 51(13), 1820–1856. https://doi.org/10.1177/0010414018774355

Liu, R., Davenport, J., Stupple, E., Holton, D., & Sams, C. (2018). The relationship between predictors of student dropout in higher education: A systematic review and meta-analysis. Educational Research Review, 25, 60–82. https://doi.org/10.1016/j.edurev.2018.10.003

Kuzilek, J., Hlosta, M., & Zdrahal, Z. (2017). Open University Learning Analytics dataset. Scientific Data, 4, 170171. https://doi.org/10.1038/sdata.2017.171

Hellas, A., Ihantola, P., Petersen, A., Ajanovski, V. V., Gutica, M., Hynninen, T., Knutas, A., Leinonen, J., Messom, C., & Liao, S. N. (2018). Predicting academic performance: A systematic literature review. In Proceedings of the 2018 ITiCSE Conference Companion (pp. 175–199). https://doi.org/10.1145/3293881.3295783

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. https://doi.org/10.1007/s10639-020-10230-3

Lemay, D. J., & Doleck, T. (2020). Predicting completion of massive open online course (MOOC) assignments from video viewing behavior. Interactive Learning Environments, 30(10), 1782–1793. https://doi.org/10.1080/10494820.2020.1746673

Cui, Y., Chen, F., Shiri, A., & Fan, Y. (2019). Predictive analytic models of student success in higher education: A review of methodology. Information and Learning Sciences, 120(3/4), 208–227. https://doi.org/10.1108/ILS-10-2018-0104

Onwuegbuzie, A. J., Daniel, L. G., & Collins, K. M. (2009). A meta-validation model for assessing the score-validity of student teaching evaluations. Quality & Quantity, 43(2), 197–209. https://doi.org/10.1007/s11135-007-9112-4

Spooren, P., Brockx, B., & Mortelmans, D. (2013). On the validity of student evaluation of teaching: The state of the art. Review of Educational Research, 83(4), 598–642. https://doi.org/10.3102/0034654313496870

Boring, A., Ottoboni, K., & Stark, P. B. (2016). Student evaluations of teaching (mostly) do not measure teaching effectiveness. ScienceOpen Research, 1–11. https://doi.org/10.14293/S2199-1006.1.SOR-EDU.AETBZC.v1

Marsh, H. W., & Roche, L. A. (1997). Making students' evaluations of teaching effectiveness effective. American Psychologist, 52(11), 1187–1197. https://doi.org/10.1037/0003-066X.52.11.1187

Stehle, S., Spinath, B., & Kadmon, M. (2012). Measuring teaching effectiveness: Correspondence between students' evaluations of teaching and different measures of student learning. Research in Higher Education, 53(8), 888–904. https://doi.org/10.1007/s11162-012-9260-9

Tessema, K. A. (2018). The unfolding trends and consequences of expanding higher education in Ethiopia: Massive universities, massive challenges. Higher Education Quarterly, 65(1), 89–106. https://doi.org/10.1111/j.1468-2273.2010.00475.x

Khan, M. A., & Law, L. S. (2018). The role of national cultures in shaping the corporate management cultures: A three-country theoretical analysis. In Organizational Culture (pp. 27–46). IntechOpen. https://doi.org/10.5772/intechopen.78051

Felder, R. M. (2020). Opinion: Uses, misuses, and validity of learning styles. Advances in Engineering Education, 8(1), 1–16. https://doi.org/10.18260/3-1-370-36027

Bryman, A. (2016). Social research methods (5th ed.). Oxford University Press.

Hair, J. F., Risher, J. J., Sarstedt, M., & Ringle, C. M. (2019). When to use and how to report the results of PLS-SEM. European Business Review, 31(1), 2–24. https://doi.org/10.1108/EBR-11-2018-0203

Taber, K. S. (2018). The use of Cronbach's alpha when developing and reporting research instruments in science education. Research in Science Education, 48(6), 1273–1296. https://doi.org/10.1007/s11165-016-9602-2

Hayes, A. F., & Coutts, J. J. (2020). Use omega rather than Cronbach's alpha for estimating reliability. But… Communication Methods and Measures, 14(1), 1–24. https://doi.org/10.1080/19312458.2020.1718629

Natekin, A., & Knoll, A. (2013). Gradient boosting machines, a tutorial. Frontiers in Neurorobotics, 7, 21. https://doi.org/10.3389/fnbot.2013.00021

Chen, T., & Guestrin, C. (2016). XGBoost: A scalable tree boosting system. In Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (pp. 785–794). https://doi.org/10.1145/2939672.2939785

Ke, G., Meng, Q., Finley, T., Wang, T., Chen, W., Ma, W., Ye, Q., & Liu, T. Y. (2017). LightGBM: A highly efficient gradient boosting decision tree. Advances in Neural Information Processing Systems, 30, 3146–3154. https://doi.org/10.5555/3294996.3295074

McGunagle, D., & Zizka, L. (2020). Employability skills for 21st-century STEM students: The employers' perspective. Higher Education, Skills and Work-Based Learning, 10(3), 591–606. https://doi.org/10.1108/HESWBL-10-2019-0148

Succi, C., & Canovi, M. (2019). Soft skills to enhance graduate employability: Comparing students and employers' perceptions. Studies in Higher Education, 45(9), 1834–1847. https://doi.org/10.1080/03075079.2019.1585420

Lim, Y. M., Lee, T. H., Yap, C. S., & Ling, C. C. (2016). Employability skills, personal qualities, and early employment problems of entry-level auditors: Perspectives from employers, lecturers, auditors, and students. Journal of Education for Business, 91(4), 185–192. https://doi.org/10.1080/08832323.2016.1153998

Nabi, G., Liñán, F., Fayolle, A., Krueger, N., & Walmsley, A. (2017). The impact of entrepreneurship education in higher education: A systematic review and research agenda. Academy of Management Learning & Education, 16(2), 277–299. https://doi.org/10.5465/amle.2015.0026

Bae, T. J., Qian, S., Miao, C., & Fiet, J. O. (2014). The relationship between entrepreneurship education and entrepreneurial intentions: A meta-analytic review. Entrepreneurship Theory and Practice, 38(2), 217–254. https://doi.org/10.1111/etap.12095

Wolfert, S., Ge, L., Verdouw, C., & Bogaardt, M. J. (2017). Big data in smart farming – A review. Agricultural Systems, 153, 69–80. https://doi.org/10.1016/j.agsy.2017.01.023

Liakos, K. G., Busato, P., Moshou, D., Pearson, S., & Bochtis, D. (2018). Machine learning in agriculture: A review. Sensors, 18(8), 2674. https://doi.org/10.3390/s18082674

Kamilaris, A., & Prenafeta-Boldú, F. X. (2018). Deep learning in agriculture: A survey. Computers and Electronics in Agriculture, 147, 70–90. https://doi.org/10.1016/j.compag.2018.02.016

Wang, J., & Yang, M. (2021). Higher agricultural education reform in China: A bibliometric analysis. Sustainability, 13(20), 11328. https://doi.org/10.3390/su132011328

Li, J., Wang, Z., Cheng, X., Shu, J., Bi, R., Wei, F., & Wang, X. (2020). The impact of forestry universities on regional development: Evidence from China. Forest Policy and Economics, 116, 102186. https://doi.org/10.1016/j.forpol.2020.102186

Anderson, L. W., & Krathwohl, D. R. (2001). A taxonomy for learning, teaching, and assessing: A revision of Bloom's taxonomy of educational objectives. Longman.

Mulder, M. (Ed.). (2017). Competence-based vocational and professional education: Bridging the worlds of work and education. Springer. https://doi.org/10.1007/978-3-319-41713-4

Leal Filho, W., Raath, S., Lazzarini, B., Vargas, V. R., de Souza, L., Anholon, R., Quelhas, O. L. G., Haddad, R., Klavins, M., & Orlovic, V. L. (2018). The role of transformation in learning and education for sustainability. Journal of Cleaner Production, 199, 286–295. https://doi.org/10.1016/j.jclepro.2018.07.017

Lozano, R., Merrill, M. Y., Sammalisto, K., Ceulemans, K., & Lozano, F. J. (2017). Connecting competences and pedagogical approaches for sustainable development in higher education: A literature review and framework proposal. Sustainability, 9(10), 1889. https://doi.org/10.3390/su9101889

Ankrah, S., & Al-Tabbaa, O. (2015). Universities–industry collaboration: A systematic review. Scandinavian Journal of Management, 31(3), 387–408. https://doi.org/10.1016/j.scaman.2015.02.003

Galán-Muros, V., & Davey, T. (2019). The UBC ecosystem: Putting together a comprehensive framework for university-business cooperation. The Journal of Technology Transfer, 44(4), 1311–1346. https://doi.org/10.1007/s10961-017-9562-3

Bao, W. (2020). COVID-19 and online teaching in higher education: A case study of Peking University. Human Behavior and Emerging Technologies, 2(2), 113–115. https://doi.org/10.1002/hbe2.191

Crawford, J., Butler-Henderson, K., Rudolph, J., Malkawi, B., Glowatz, M., Burton, R., Magni, P. A., & Lam, S. (2020). COVID-19: 20 countries' higher education intra-period digital pedagogy responses. Journal of Applied Learning & Teaching, 3(1), 9–28. https://doi.org/10.37074/jalt.2020.3.1.7