Neutrosophic Sets and Systems, Vol. 102, 2027
University of New Mexico

Teaching Effectiveness of Higher Education Based on Big Data and Artificial Intelligence: A Novel Neutrosophic Multi-Criteria Decision Framework for Intelligent Educational Performance Evaluation

Xiaojiang Feng*
School of Educational Science, Yan'an University, Yan'an, Shaanxi, 716000, China
*Corresponding author, E-mail: ydfxj22@163.com

Abstract: The recent fast-paced digital revolution of education has led to an unprecedented amount of educational data being collected through learning management systems, online exams, student information systems, interactions in class and institutional analytics. Although the big data and artificial intelligence (AI) can help to assess teaching performance quite effectively, the current evaluation methods use mostly deterministic or fuzzy models, which cannot sufficiently reflect the uncertainty and inconsistencies of educational settings. Different opinions of experts, heterogeneous feedbacks, unobserved data and conflicting performance criteria often decrease the reliability and consistency of assessment of teaching effectiveness. In order to overcome these limitations, a new neutrosophic multi-criteria decision-making (MCDM) framework is introduced in this paper, which combines big data analytics and artificial intelligence for the evaluation of teaching effectiveness in higher education. This framework uses truth, indeterminacy and falsity membership degrees to model educational information, thus allowing one to evaluate simultaneously the certainty, ambiguity and contradictions within multiple criteria of teaching quality. Criterion weights are determined objectively based on characteristics of educational data, while the AI-based analytics help to identify important patterns of performance from large academic datasets. This model integrates the components via a well-constructed neutrosophic ranking process mathematically for obtaining trustworthy and understandable evaluations of teaching performance. The model has been formulated for the purpose of facilitating institutional decision-makers to identify their strengths in teaching, prioritize their professional development, and improve education quality via evidence-based decision-making processes. The proposed approach will be examined and compared with existing MCDM methods on a dataset regarding higher education teaching evaluation under the same experimental conditions. The sensitivity, robustness, and comparative studies will be carried out for investigating the consistency and feasibility of the proposed approach. The anticipated contribution is an intelligent educational evaluation framework that can successfully manage the uncertainty while taking advantage of big data and artificial intelligence.

Keywords: Teaching effectiveness; Higher education; Big data analytics; Artificial intelligence; Neutrosophic sets; Multi-criteria decision making.

1. Introduction

It is widely accepted that teaching effectiveness is one of the most important factors determining the quality of education in higher education institutions. Nowadays, higher education institutions are increasingly required to offer high-quality teaching that contributes to learning, increases academic performance, and develops skills needed in modern society. In this way, the assessment of teaching effectiveness has become an important strategic task for higher education institutions aiming at continuous quality improvement, accreditation requirements, and educational management based on evidence. Traditional assessment methods, such as student questionnaire survey or the judgment of experts, give little understanding of the complicated nature of teaching performance. In many cases, these methods do not make use of a great deal of educational data available in modern digital environment[1], [2].

The fast implementation of learning management systems, online learning systems, intelligent tutoring systems, classroom sensors, and information systems has turned higher education into a data-driven environment. Everything done during educational activities – student attendance, their performance in the assessment, online interaction, learning behavior, course completion, educational materials, and classroom participation – provides useful digital information. The availability of educational data brings a lot of possibilities for the assessment of teaching quality based on comprehensive, objective, and constantly updated data rather than only subjective evaluations[3], [4].

Educational big data has enabled educational institutions to increase dramatically their capabilities of collecting, processing, storing, and analyzing various kinds of educational datasets. Educational big data includes high volumes, velocities, varieties, veracies, and values that allow universities to observe instructional activities from different angles and at a large timescale. Due to the integration of various educational information sources, it becomes possible to find connections between teaching practices, student involvement, learning outcomes, and academic performance[5], [6].

AI has additionally contributed to the development of education evaluation by presenting smart data analytics methodologies that can extract valuable knowledge from big educational datasets. The machine learning algorithms, deep learning, natural language processing, recommendation systems, and predictive analysis have been widely used to evaluate students' feedback, predict learning outcomes, identify at-risk students, and make pedagogical decisions. AI helps educational institutions to perform automated analytical tasks, identify hidden patterns in performance, and make recommendations on how to improve the quality of teaching[7], [8].

However, despite the use of the mentioned technologies, evaluating the effectiveness of teachers still remains a difficult problem due to the fact that educational information is highly uncertain, incomplete, inconsistent, and even contradictory. Personal bias may impact students' evaluations, experts may hold different views on instruction, classroom observations may be partial, and educational performance indicators may be vague.

Moreover, learning analytics and human judgment may provide conflicting evidence for the same instructor. Classical statistical and classical machine learning methods rely on precise observations, while fuzzy approaches focus on vagueness without taking into account indeterminacy and contradictions[9], [10].

A more advanced representation of uncertainty in the form of neutrosophic models, which describes three independent membership functions of truth, indeterminacy, and falsity, is provided by neutrosophic theory. Contrarily to the classical fuzzy and intuitionistic fuzzy models, the neutrosophic approach allows for the description of not only certainty but also hesitation, inconsistency, and incomplete information. The ability is especially important for the educational evaluation since different stakeholders, such as students, their peers, school administration, and outsiders, usually give their assessment with different degrees of certainty and contradict each other. Therefore, the mathematical ground offered by neutrosophic modeling is relevant to capture the complexity of the uncertainty in assessing teaching efficiency[11], [12].

Despite many attempts to apply artificial intelligence, learning analytics, and multi-criteria decision-making techniques to evaluate the process of education, some problems persist. First of all, the approaches used in most cases concentrate on building predictive models and do not offer decision support. Secondly, they are based on traditional techniques of MCDM that are unable to model uncertainty that comes from the heterogeneity of educational information. Finally, there are few integrated approaches in the existing literature that combine AI, big data analytics, and uncertainty modeling[13], [14].

In response to the issues highlighted above, this research suggests a new neutrosophic multi-criteria decision-making approach for assessing teaching effectiveness in higher education based on big data and artificial intelligence. The new approach combines the use of AI-driven educational analytics along with neutrosophic information modeling, which allows dealing with uncertainty while leveraging big educational data to perform objective assessments. The suggested approach is aimed at aggregating multiple educational performance criteria using a mathematically sound decision-making process that accounts for imprecise, inconsistent and ambiguous information.

Key highlights of this study include:

(1) A new neutrosophic-based decision-making approach for the holistic assessment of teaching effectiveness in uncertain educational settings.
(2) A new technique that uses big data analytics and artificial intelligence to mine information related to teaching performance from diverse educational data sets.
(3) Modeling of educational uncertainty in terms of truth, indeterminacy and falsity membership functions that can simultaneously capture conflicting and incomplete information for evaluation.
(4) The new decision-making approach offers an interpretative multi-criteria ranking mechanism for helping decision-makers in institutions identify areas where teaching is strong and areas where improvement is needed.

(5) The new method will be tested experimentally using higher education teaching data sets and benchmarked against multi-criteria decision-making techniques. Sensitivity and robustness analysis will further validate its applicability and decision stability.

2. Related Work and Critical Research Gap

2.1 Teaching Effectiveness Evaluation in Higher Education

Teaching effectiveness has long been considered a multidimensional construct that includes instructional quality, learner achievement, classroom dynamics, curriculum delivery, assessment processes, and competency of instructors. Higher education institutions usually assess teaching effectiveness through student evaluations, peer reviews, self-reporting, classroom observations, and administration's judgment[15], [16]. The aforementioned evaluation process generates valuable data about academic promotion, curricular development, faculty development, and quality assurance at an institution. Student evaluation of teaching (SET) continues to be the most popular method of evaluating teaching effectiveness due to its practicality and straightforward nature. This is directly related to obtaining the information about the quality of instruction from students. There are many studies indicating that student evaluations could be affected by non-academic variables including expected grades, difficulty of the course, number of students enrolled in the course, popularity of the instructor, gender differences, and personal preferences. In order to counteract these weaknesses, most universities have introduced comprehensive evaluation models, which incorporate various performance measures such as student achievement, participation in classes, innovation in instruction, evidence-based teaching, class management, and professional development. Multidimensional evaluation models generate more balanced assessments, but create additional problems due to the fact that the measures may have different scales and ambiguous connections[17], [18].

2.2 Big Data Analytics in Educational Evaluation

The fast-paced digitization of universities has produced the constant formation of educational big data generated by learning management systems, online classrooms, evaluation platforms, electronic libraries, attendance systems, and academic information systems. Educational big data allows educational organizations to assess the learning behaviors, teaching activities, academic performance, and institutional performance at an unprecedented level of detail. Big data analytics has been used extensively to find learning patterns, assess students' performance, discover academic risks, estimate the efficiency of courses, and make decisions about educational policies. Educational dashboards and learning analytics platforms enable educators and teachers to monitor the process of instruction in real-time mode while gaining evidence-based knowledge for improving the quality of their work. At the same time, educational big data poses serious analytical problems caused by its heterogeneous nature, poor quality, missing values, temporal aspects, and high dimensionality. Combining various sources of educational big data into one teaching evaluation framework is a challenging process, especially when there is uncertainty and inconsistency in information[19].

2.3 Artificial Intelligence for Teaching Quality Assessment

Indeed, artificial intelligence can be seen as one of the most influential technological approaches to intelligent assessment of education. There are lots of machine learning algorithms applied in forecasting student success, classification of learning behavior, analysis of educational materials, identification of potentially risky learners, and instructional assessment. Deep learning methods have been shown to work very well in analysis of textual feedback, videos from classes, forums for discussions, and multimodal educational data. Besides, natural language processing methods have been applied to the analysis of qualitative comments made by students, detection of teaching-related topics, sentiment analysis, and detection of strengths and weaknesses in instruction. Recommendation systems help teachers to make recommendations regarding teaching strategies and educational materials on the basis of historical educational data. While AI makes predictive performance and analytical processes much better, there are a lot of AI-based models which are the examples of black box systems that cannot provide any interpretation. There are cases when educational administrators need to have decision support systems explaining how teaching effectiveness is evaluated not making any predictions.

2.4 Multi-Criteria Decision-Making Approaches in Educational Assessment

Effectiveness in teaching is bound to include a range of qualitative and quantitative criteria that are mutually exclusive. MCDM models have thus been used in educational evaluation problems due to their capacity for the consideration of many performance criteria simultaneously. AHP, TOPSIS, VIKOR, ELECTRE, PROMETHEE, COPRAS, ARAS, MOORA, MARCOS, WASPAS, and EDAS methods have been used in ranking professors, evaluation of educational courses, prioritization of development activities, and management decisions in academia. Objective criterion weight determination techniques like CRITIC, Entropy, MEREC, and Standard Deviation weightings have been employed in combination with ranking methods to minimize any subjectivity involved in the importance determination of criteria. Conventional MCDM models usually consider evaluation information under the assumption of its deterministic nature. But in actual educational settings, evaluation information tends to involve ambiguity, imprecision, inconsistency, and uncertainty that cannot be adequately expressed in crisp numbers.

2.5 Neutrosophic Decision Models under Educational Uncertainty

The theory of neutrosophics has proven to be a useful tool in expressing uncertainty using three membership functions for truth, indeterminacy, and falsity. Neutrosophic representations are more flexible than their fuzzy and intuitionistic counterparts in representing incomplete and inconsistent data. There are several studies on neutrosophic MCDM techniques that can be applied in decision-making in healthcare, engineering, supplier selection, sustainability evaluation, risk assessment, energy management, and financial evaluation. The existing literature shows that decisions based on neutrosophic representations are more reliable when data are either imprecise or somewhat contradictory. But there have been few attempts to study neutrosophic decision-making techniques for evaluating teaching effectiveness, especially when large amounts of educational data are available. Most studies concentrate either on expert-based evaluation using small-sized data or on educational evaluation using traditional tools without leveraging educational data that is available from advanced educational systems.

2.6 Critical Research Gap

The above review highlights a number of critical shortcomings of the existing literature.

First, most of the existing systems for teaching evaluations are either too dependent on subjective expertise or based on AI prediction models that do not involve any multi-criteria decision making processes. In other words, educational administrators are provided with correct predictions without any explanation of how these decisions were made.

Second, even though educational big data provides a lot of information about educational activities, a number of contemporary methods for evaluations do not combine heterogeneous educational datasets in one decision-making framework allowing for consideration of both quantitative and qualitative indicators, as well as institutional indicators.

Third, most of the existing MCDM methods are based on either deterministic or fuzzy information, while the decision process in education involves a lot of uncertainties, inconsistencies, incompleteness, and conflicting information. This makes the reliability of decision-making results questionable since multiple stakeholders provide different views on the issue.

Fourth, AI-based methods of educational evaluations are mostly focused on predictive accuracy but not on uncertainty modeling, interpretability, and transparency of the decision-making process.

However, one final deficiency of existing methods lies in the absence of frameworks that would be able to integrate all these elements – big data analytics of educational data, AI, representation of neutrosophic uncertainty, criterion weighting, and multi-criteria ranking. Taking into account such deficiencies, this research suggests a new method of multi-criteria decision making based on neutrosophy, big data analytics and AI that would be capable to assess effectiveness of teaching in higher education institutions. The suggested framework should be aimed at representing uncertain information about education using truth, indeterminacy, and falsity memberships and taking advantage of knowledge representation and extraction using AI technologies.

3. Mathematical Preliminaries and Neutrosophic Definitions

In this section, the mathematics behind the proposed framework is discussed. Due to the subjectivity involved in the evaluation of teaching effectiveness, the heterogeneity of the education data, and the contradictory evidence gathered from students, peers, administrators, and intelligent education systems, the theory of neutrosophic sets offers an adequate mathematical model to handle uncertainty. The proposed framework uses the concept of single-valued neutrosophic sets (SVNSs).

3.1 Universe of Discourse

Let

$$X = \{x_1, x_2, \ldots, x_n\} \qquad (1)$$

denote the finite universe of discourse, where each element x_i represents an alternative. In this study, an alternative corresponds to a university instructor whose teaching effectiveness is evaluated.

Let

$$C = \{C_1, C_2, \ldots, C_m\} \qquad (2)$$

define the set of evaluation criteria for describing teaching effectiveness.

The examples of evaluation criteria are:

- Teacher preparation
- Delivery in class
- Student involvement
- Quality of assessment
- Teaching through research
- Digital teaching skills
- Achievement of learning outcomes
- Instructional effectiveness supported by AI

Let there be n teachers and m evaluation criteria.

Definition 1. Single-Valued Neutrosophic Set

A Single-Valued Neutrosophic Set (SVNS) A on X is defined as

$$A = \{\langle x, T\_A(x), \quad I\_A(x), F\_A(x) \rangle : x \in X \} \qquad (3)$$

Where:

$T\_A(x) : X \rightarrow [0,1]$

$I\_A(x) : X \rightarrow [0,1]$

$F\_A(x) : X \rightarrow [0,1]$

represent respectively:

$T\_A(x)$ = truth-membership degree

$I\_A(x)$ = indeterminacy-membership degree

$F\_A(x)$ = falsity-membership degree

subject to

$$0 \leq T\_A(x), I\_A(x), F\_A(x) \leq 1 \quad (4)$$

and

$$0 \leq T\_A(x)+I\_A(x)+F\_A(x) \leq 3 \quad (5)$$

Unlike intuitionistic fuzzy sets, the three membership values are mutually independent, allowing inconsistent and incomplete information to be represented simultaneously.

Definition 2. Single-Valued Neutrosophic Number

A Single-Valued Neutrosophic Number (SVNN) is expressed as

$$A = (T, I, F) \quad (6)$$

Where:

$T \in [0,1]$

$I \in [0,1]$

$F \in [0,1]$

Example: $A = (0.82,0.11,0.08)$

Indicates Truth = 0.82, Indeterminacy = 0.11, Falsity = 0.08 which represents a highly positive evaluation with relatively low uncertainty.

Definition 3. Neutrosophic Decision Matrix

Assume there are n alternatives and m criteria. The neutrosophic decision matrix is

$$D = [d_{ij}]_{n \times m} \quad (7)$$

Where:

$d_{ij} = (T_{ij}, I_{ij}, F_{ij})$ represents the neutrosophic evaluation of instructor i under criterion j.

Therefore:

$$D = \begin{bmatrix} (T_{11},I_{11},F_{11}) & \cdots & (T_{1m},I_{1m},F_{1m}) \\ \vdots & \ddots & \vdots \\ (T_{n1},I_{n1},F_{n1}) & \cdots & (T_{nm},I_{nm},F_{nm}) \end{bmatrix} \quad (8)$$

Each element is obtained from educational big data, AI analytics, expert assessment, or their integration.

Definition 4. Score Function

To compare two neutrosophic numbers, a score function is required. The score value of

$A=(T,I,F)$ is defined as:

$$S(A)= (2+T-I-F)/3 \quad (9)$$

where $-1 \le S(A) \le 1$ A larger score indicates better teaching effectiveness.

Definition 5. Accuracy Function

When two alternatives possess identical score values, an accuracy function is used.

The accuracy function is:

$$H(A)=T-F \quad (10)$$

Higher values indicate stronger certainty in the evaluation.

Definition 6. Weighted Neutrosophic Aggregation

Suppose $w=(w_1,w_2,...,w_m)$ denotes the criterion weight vector satisfying

$$w_j \ge 0 \quad (11)$$

and

$$\Sigma\, w_j =1 \quad (12)$$

The aggregated neutrosophic assessment for instructor i is represented as:

$$N_i=(T_i,I_i,F_i) \quad (13)$$

Where:

$$T_i = \Sigma\, w_j T_{ij} \quad (14)$$

$$I_i = \Sigma\, w_j I_{ij} \quad (15)$$

$$F_i = \Sigma\, w_j F_{ij} \quad (16)$$

for $j=1,2,\ldots,m$.

The aggregated value summarizes the overall teaching effectiveness by simultaneously considering all evaluation criteria and their corresponding importance weights.

Definition 7. Ranking Rule

Given two instructors A_1 and A_2 the ranking procedure follows:

$$\text{If } S(A_1) > S(A_2) \text{ then } A_1 \succ A_2 \tag{17}$$

If $S(A_1)=S(A_2)$ then compare $H(A_1)$ and $H(A_2)$.

$$\text{If } H(A_1)>H(A_2) \text{ then } A_1 \succ A_2 \tag{18}$$

Otherwise A_2 is preferred.

The aforementioned preliminaries provide the foundation for the proposed intelligent neutrosophic approach to teaching evaluation. The next section presents the formulation of the teaching effectiveness decision problem for higher education. The criteria, alternatives and decision making environment are also introduced in this section.

4. Problem Formulation

In this section, the teaching effectiveness evaluation problem is posed as a MCDM problem using neutrosophic sets, given that the teaching environment includes factors such as uncertainty, partial information, heterogeneous educational data, and conflicting evaluation criteria. The task here is to develop a smart decision-making model to rank university teachers in terms of teaching effectiveness using educational big data, AI, and neutrosophic information.

4.1 Decision Environment

$$\text{Assume that } A = \{A_1, A_2, \ldots, A_n\} \tag{19}$$

is the set of teaching alternatives, where A_i ($i = 1,2,\ldots,n$) denotes the i-th university instructor. Similarly,

$$C = \{C_1, C_2, \ldots, C_m\} \tag{20}$$

denotes the set of criteria for evaluating teaching. Every teacher is evaluated on each criterion based on the data gathered from different sources of education, including:

- Data of student evaluations
- Learning Management Systems (LMS)
- Reports of classroom observations
- Systems of academic management

- Learning analytics
- AI generated instruction analytics
- Databases of institutional teaching

Hence, the evaluation process consists of heterogeneous data from education with different degrees of certainty.

4.2 Evaluation Criteria

Suppose the teaching effectiveness evaluation is conducted using the following representative criteria.

C_1 : Teaching Preparation
C_2 : Classroom Delivery
C_3 : Student Engagement
C_4 : Assessment and Feedback Quality
C_5 : Learning Outcome Achievement
C_6 : Research-Integrated Teaching
C_7 : Digital Teaching Competency
C_8 : AI-Supported Instructional Effectiveness

These criteria collectively describe instructional quality from pedagogical, technological, and learning-performance perspectives.

4.3 Educational Big Data Representation

For instructor A_i and criterion C_j, educational information is collected from multiple heterogeneous sources.

$$\text{Let } x_{ij} = \{x_1, x_2, \ldots, x_k\} \tag{21}$$

where the educational observations related to instructor i by criterion j are denoted.
The observations could include:

- Satisfactions of students
- Data on attendance
- Information about assignment submission
- Data on online activities
- Data on learning behaviors
- AI generated data on teaching quality
- Data on classroom interactions

Post processing of the information into neutrosophic form.

4.4 Neutrosophic Evaluation Matrix

The evaluation matrix is defined as:

$$D=[d_{ij}]_{n\times m} \quad (22)$$

Where:

$$d_{ij}=(T_{ij},I_{ij},F_{ij}). \quad (23)$$

Here, T_{ij} signifies the strength of the available evidence indicating satisfactory teaching performance.
I_{ij} signifies the degree of indeterminacy due to incomplete, inconsistent, or contradictory educational data.
F_{ij} signifies the strength of the available evidence supporting unsatisfactory teaching performance.
Each neutrosophic number is such that:

$$0\le T_{ij}\le 1 \quad (24)$$

$$0\le I_{ij}\le 1 \quad (25)$$

$$0\le F_{ij}\le 1 \quad (26)$$

and

$$0\le T_{ij}+I_{ij}+F_{ij}\le 3 \quad (27)$$

4.5 Criterion Weight Vector

The importance of the evaluation criteria is represented by

$$W=(w_1,w_2,\ldots,w_m) \quad (28)$$

subject to

$$w_j\ge 0 \quad (29)$$

and

$$\sum_{j=1}^{m} w_j =1 \quad (30)$$

The proposed framework determines these weights objectively using educational data characteristics extracted from big data analytics, thereby reducing dependence on purely subjective expert judgments.

4.6 Artificial Intelligence-Based Feature Extraction

Suppose

$$X=\{x_1,x_2,\ldots,x_p\} \quad (31)$$

represents the educational feature vector extracted from institutional databases. An artificial intelligence model transforms the raw educational data into informative teaching-performance indicators,

$$Z=f_{AI}(X) \quad (32)$$

Where X denotes the original educational feature space, f_AI(·) represents the AI learning model, and $Z=\{z_1,z_2,\ldots,z\_q\}$ is the extracted feature representation used for teaching evaluation. The generated indicators provide quantitative evidence supporting the neutrosophic evaluation process.

4.7 Objective Function

The primary objective is to determine the optimal ranking of university instructors according to their comprehensive teaching effectiveness.

Let $R(A_i)$ denote the overall evaluation score of instructor A_i. The optimization objective is max $R(A_i)$, For i=1,2,…,n.

The ranking function combines
- objective criterion weights,
- AI-generated educational knowledge,
- neutrosophic uncertainty,
- and multi-criteria aggregation

to obtain the final teaching effectiveness scores.

4.8 Research Assumptions

The above framework is developed based on the following assumptions:
- Assumption 1 The data regarding education are gathered from credible institutional information systems.
- Assumption 2 All evaluation criteria are standardized prior to neutrosophication.
- Assumption 3 Membership functions for truth, indeterminacy, and falsehood are independent.
- Assumption 4 Criterion weights conform to the normalization condition.
- Assumption 5 All instructors are evaluated using the same criteria and same institutional conditions.

4.9 Problem Statement

Given:
1. n university instructors,
2. m teaching evaluation criteria,
3. heterogeneous educational big data,
4. AI-generated teaching indicators,
5. and uncertain educational information,

determine the ranking

$$A\_(1) \succ A\_(2) \succ \ldots \succ A\_(n) \qquad (33)$$

so that the final ranking can properly capture the teaching effectiveness of all teachers taking into account uncertainties, inconsistencies, and incomplete educational information. The above mathematical model sets up the optimization problem investigated in this paper. In the next section, the proposed intelligent neutrosophic decision making framework will be presented, which involves the innovative weighting method, neutrosophic aggregation technique, intelligent evaluation approach and computational process.

5. Proposed Intelligent Neutrosophic Decision Framework

In this section, the proposed intelligent neutrosophic decision model for assessing the teaching performance in higher education based on educational big data and artificial intelligence (AI) is introduced. The proposed decision model differs from traditional MCDM methods by incorporating feature extraction via AI, objective criterion weights, neutrosophic modeling of uncertainty, adaptive aggregation, and intelligent ranking in a coherent decision-making process.

5.1 Framework Architecture

The developed framework comprises seven consecutive steps:

- Step 1 Collecting Education Big Data
- Step 2 Data Cleaning and Automated Feature Extraction Using AI
- Step 3 Creating the Neutrosophic Decision Matrix
- Step 4 Calculation of Weights of Objective Criteria
- Step 5 Adaptive Neutrosophic Aggregation
- Step 6 Calculating Scores of Teaching Effectiveness
- Step 7 Ranking University Teachers

Mathematical connection exists between all steps of the developed evaluation approach.

5.2 Stage 1. Educational Data Acquisition

Suppose $E=\{e_1,e_2,\ldots,eN\}$. This represents the entire education dataset that is collected from institutional databases.

The data set comprises of:

- Student feedback
- LMS data
- Data related to online learning
- Assessment data
- Observation data from classrooms
- AI generated instructional data
- Academic management databases

After being preprocessed, each criterion is normalized to [0,1].

5.3 Stage 2. AI-Based Feature Extraction

$$\text{Let } X=(x_1,x_2,\ldots,xp) \tag{34}$$

denote the normalized educational feature vector. An AI model extracts high-level instructional information

$$Z=fAI(X) \quad (35)$$

where $fAI(\cdot)$ represents the artificial intelligence learning model. The extracted feature vector becomes $Z=(z_1,z_2,\ldots,zq)$.

These features summarize the instructional characteristics required for teaching evaluation.

5.4 Stage 3: Neutrosophic Transformation

Each extracted feature is transformed into three independent membership values.

Truth Membership:

$$Tij=fT(zij) \quad (36)$$

Indeterminacy Membership:

$$Iij=fI(zij) \quad (37)$$

Falsity Membership:

$$Fij=fF(zij) \quad (38)$$

Where $0\le Tij,Iij,Fij\le 1$.

The corresponding neutrosophic number becomes $dij=(Tij,Iij,Fij)$.

The complete neutrosophic decision matrix is:

$$D=[dij]n\times m \quad (39)$$

5.5 Stage 4. Objective Criterion Weighting

Instead of assigning subjective importance values, the proposed framework determines criterion importance from educational information.

Let σj denote the standard deviation of criterion j, and μj its mean value. The information coefficient is computed as:

$$ICj=\sigma j/\mu j \quad (40)$$

The normalized objective weight becomes:

$$wj=ICj/\Sigma ICj \quad (41)$$

subject to:

$$\Sigma wj=1 \quad (42)$$

Higher variability indicates stronger discrimination capability and therefore receives a larger objective weight.

5.6 Stage 5. Adaptive Neutrosophic Aggregation Operator

To improve robustness under uncertain educational environments, this study introduces an Adaptive Neutrosophic Educational Aggregation Operator (ANEAO).

For instructor Ai, the adaptive truth component is:

$$Ti^*=\Sigma wjTij \quad (43)$$

The adaptive indeterminacy component is:

$$Ii^*=\Sigma wj^2Iij \quad (44)$$

The adaptive falsity component is:

$$Fi^*=\Sigma wjFij \quad (45)$$

Unlike the traditional methods of weighted average, here the weight of the indeterminacy element is calculated using a square of the weight. It helps penalize those criteria with high uncertainty levels, without losing valuable information from trustworthy educational data. The resulting neutrosophic assessment is:

$$Ni=(Ti^*,Ii^*,Fi^*) \quad (46)$$

5.7 Stage 6. Improved Teaching Effectiveness Score

To simultaneously reward reliable evidence and penalize uncertainty, the proposed study introduces a new teaching effectiveness score function.

For instructor Ai,

$$Scorei= Ti^*(1-Ii^*)-Fi^*(1+Ii^*) \quad (47)$$

Where Ti* stands for the truth degree, Ii* for the indeterminacy and Fi* for the falsity.

- The first term recognizes well-documented educational experience positively.
- The second term penalizes uncertainties.
- The third term gives greater weight to negative education experiences due to higher levels of uncertainties.

5.8 Normalized Teaching Effectiveness Index

To facilitate comparison among instructors, the scores are normalized.

Let Smax=max(Scorei) and Smin=min(Scorei). The normalized index becomes:

$$TEIi=(Scorei-Smin)/(Smax-Smin) \quad (48)$$

Where:

$$0 \le TEIi \le 1$$

Higher values indicate better teaching effectiveness.

5.9 Final Ranking Rule

The instructors are ranked according to TEI1,TEI2,…,TEIn.

$$\text{If } TEIi>TEIk, \quad (49)$$

Then Ai≻Ak.

When equal values occur, the secondary ranking criterion is Ti*−Fi*.

The instructor possessing the larger secondary score receives the higher ranking.

5.10 Computational Procedure

Input

- Educational dataset
- Criteria for evaluation
- University teachers

Output

- Overall rank of teaching effectiveness

Procedure

- Step 1 Gather the big data from the field of education.
- Step 2 Normalize all educational metrics.
- Step 3 Obtain features of instructions through artificial intelligence techniques.
- Step 4 Establish the neutrosophic decision matrix.
- Step 5 Estimate objective criterion weights.
- Step 6 Implement the suggested Adaptive Neutrosophic Educational Aggregation Operator (ANEAO).
- Step 7 Estimate the suggested teaching effectiveness score.
- Step 8 Normalize the teaching effectiveness index.
- Step 9 Rank university teachers based on TEI values.
- Step 10 Sensitivity, robustness, and comparative analyses are conducted.

Adaptive Neutrosophic Educational Aggregation Operator (ANEAO) and new teaching effectiveness score are the main methodological innovations of this research. Thanks to the adaptive treatment of uncertainty, the model will be able to separate reliable educational data from unreliable, which will lead to better decision-making stability and interpretation than conventional neutrosophic aggregation models. In the next section, the detailed algorithmic procedure and computational complexity are presented.

6. Complete Algorithmic Procedure and Computational Complexity

This part shows the entire computational process of the proposed ANEAO. The proposed algorithm ensures reproducibility when evaluating teaching efficacy in higher education with the use of educational big data, artificial intelligence (AI) and neutrosophic multiple criteria decision making.

Algorithm 1: Adaptive Neutrosophic Educational Aggregation Framework (ANEAO)

Input:

- E = Educational big data
- $A = \{A_1, A_2, \ldots, A_n\}$ (University instructors)
- $C = \{C_1, C_2, \ldots, C_m\}$ (Evaluation criteria)
- AI learning model

Output

- Index for Teaching Effectiveness (ITE)
- Ranking of all the teachers finally

Algorithm

Step 1 Obtain educational data

- Learning management system data
- Student evaluation data
- Assessment data
- Classroom observation data
- Databases
- AI-based indicators of teaching

Step 2 Process educational data

- Remove duplicate entries
- Handle missing values
- Normalize numeric variables
- Encode categorical variables
- Identify outliers

Step 3 Extract instructional features

- Compute $Z=f_{AI}(X)$

Where X is the normalized educational dataset.

Step 4 Construct the neutrosophic decision matrix For every instructor

- A_i and criterion C_j calculate $d_{ij}=(T_{ij},I_{ij},F_{ij})$
- Store all evaluations in $D=[d_{ij}]_{n\times m}$

Step 5 Compute objective criterion weights

- For every criterion Calculate $IC_j=\sigma_j/\mu_j$
- Normalize $w_j=IC_j/\Sigma IC_j$

Step 6 Compute adaptive neutrosophic aggregation

- For every instructor Calculate:
    - $T_i^*=\Sigma w_j T_{ij}$
    - $I_i^*=\Sigma w_j^2 I_{ij}$
    - $F_i^*=\Sigma w_j F_{ij}$
- Construct $N_i=(T_i^*,I_i^*,F_i^*)$

Step 7 Calculate the proposed teaching effectiveness score

$$Score_i= T_i^*(1-I_i^*)-F_i^*(1+I_i^*)$$

Step 8 Normalize scores

$$TEI_i= (Score_i-S_{min})/(S_{max}-S_{min})$$

Step 9 Sort instructors

If $TEI_i>TEI_k$, then $A_i \succ A_k$.

Step 10 Output

- Final ranking
- Teaching effectiveness score

- Criterion importance
- Neutrosophic evaluation

7. Case Study and Data Description

In this section, the case study employed to prove the effectiveness of the Adaptive Neutrosophic Educational Aggregation Operator (ANEAO) is introduced. The purpose of this case study is to measure the efficiency of education in higher learning through the use of big data education, artificial intelligence, and neutrosophic multi-criteria decision making approach.

7.1 Case Study

This case study is concerned with the assessment of teaching efficiency of college lecturers through educational information obtained from several digital learning platforms.

Here the lecturer will be the alternative whereas instructional quality metrics will serve as evaluation criteria.

The decision problem is to determine the best lecturer(s) based on educational information rather than based on one source of evaluation.

7.2 Decision Alternatives

Assume that $A=\{A_1,A_2,A_3,A_4,A_5,A_6,A_7,A_8\}$.

where

- A_1 = Instructor 1
- A_2 = Instructor 2
- A_3 = Instructor 3
- A_4 = Instructor 4
- A_5 = Instructor 5
- A_6 = Instructor 6
- A_7 = Instructor 7
- A_8 = Instructor 8

The number of alternatives can be expanded depending on the available institutional dataset.

7.3 Evaluation Criteria

Eight evaluation criteria are considered.

C_1:	**Teaching Preparation**
C_2:	**Classroom Delivery**
C_3:	**Student Engagement**
C_4:	**Assessment and Feedback Quality**
C_5:	**Learning Outcome Achievement**
C_6:	**Research-Integrated Teaching**

C_7:	**Digital Teaching Competency**
C_8:	**AI-Supported Instructional Effectiveness**

All criteria are benefit criteria, where higher values indicate better teaching performance.

7.4 Educational Data Sources

The educational data are collected from several institutions:

- Student Evaluation of Teaching (SET)
- Learning Management System (LMS)
- Online learning platforms
- Student academic profiles
- Attendance tracking systems
- Observation reports of classroom sessions
- Quality assurance databases in institutions
- Educational analytics through Artificial Intelligence

The combination of the above-mentioned diverse data sources facilitates an overall assessment of the quality of instruction.

7.5 Data Preprocessing

Preprocessing steps are carried out to the raw education data prior to transforming it into neutrosophic values.

Step 1 Duplication removal.

Step 2 Missing values treatment.

Step 3 Normalization of numeric values.

For criterion j,

$$x'_{ij}=(x_{ij}-x_{min})/(x_{max}-x_{min}). \quad (50)$$

where xmin and xmax represent the minimum and maximum values of the corresponding criterion.

Step 4 Encoding categorical educational variables.

Step 5 Detection of abnormal observations.

Step 6 AI-based feature extraction.

7.6 AI Feature Extraction

Artificial intelligence extracts representative instructional features from educational big data.

Suppose $X=\{x_1,x_2,\ldots,x_p\}$.

After feature extraction, $Z=f_{AI}(X)$.

Where $Z=\{z_1,z_2,\ldots,z_q\}$

Includes the instructional metrics that form the basis for neutrosophic assessment.

Example of the metrics extracted are:

- Student participation metric
- Learning engagement metric
- Teaching interaction metric
- Assignment completion rate
- Teaching innovation metric
- Digital activity metric
- AI-based instructional effectiveness metric

7.7 Neutrosophic Representation

Each instructional indicator is converted into a neutrosophic number.
Suppose zij is the extracted feature. Its neutrosophic representation becomes

$$dij=(Tij,Iij,Fij). \qquad (51)$$

where Tij represents positive instructional evidence, Iij represents uncertainty, and Fij represents negative instructional evidence.

7.8 Illustrative Neutrosophic Decision Matrix

Table 1. Illustrative Neutrosophic Decision Matrix

Instructor	C_1	C_2	C_3	C_4	C_5	C_6	C_7	C_8
A_1	(0.84,0.09,0.07)	(0.82,0.10,0.08)	(0.86,0.07,0.07)	(0.83,0.11,0.06)	(0.85,0.08,0.07)	(0.80,0.12,0.08)	(0.87,0.06,0.07)	(0.84,0.09,0.07)
A_2	(0.79,0.12,0.09)	(0.81,0.10,0.09)	(0.82,0.11,0.07)	(0.80,0.12,0.08)	(0.83,0.09,0.08)	(0.79,0.13,0.08)	(0.82,0.10,0.08)	(0.81,0.11,0.08)
A_8	(0.88,0.05,0.07)	(0.89,0.05,0.06)	(0.91,0.04,0.05)	(0.88,0.06,0.06)	(0.90,0.05,0.05)	(0.86,0.08,0.06)	(0.92,0.03,0.05)	(0.89,0.05,0.06)

7.9 Decision Objective

Using the neutrosophic decision matrix, objective criterion weights, AI-generated educational indicators, and the proposed Adaptive Neutrosophic Educational Aggregation Operator, determine A(best)=argmax(TEIi).
Where TEIi denotes the normalized Teaching Effectiveness Index of instructor i. The instructor with the highest TEI value is considered the most effective according to the proposed decision framework.

8. Experimental Design and Implementation Settings

The following section discusses the experimental design used for validating the proposed Adaptive Neutrosophic Educational Aggregation Operator (ANEAO). The design is such that the proposed framework is validated under fair and reproducible conditions by employing the same datasets, evaluation metrics, and implementation environment.

8.1 Experimental Workflow

The full experiment procedure is done in the following steps:

- Step 1 Education data collection
- Step 2 Data pre-processing
- Step 3 Artificial intelligence feature extraction
- Step 4 Neutrosophication
- Step 5 Weights determination for objective criteria
- Step 6 Adaptive neutrosophic aggregation
- Step 7 Determination of teaching efficiency
- Step 8 Ranking instructors
- Step 9 Comparison step
- Step 10 Sensitivity analysis

8.2 Artificial Intelligence Configuration

The artificial intelligence techniques are used to obtain the educational indicators from the educational data.

The typical pre-processing procedures include the following steps:

- Feature scaling
- Missing values handling
- Feature encoding
- Dimensionality reduction (if needed)
- Feature importance calculation

The obtained educational indicators are then represented by neutrosophic representation technique.

8.3 Parameter Settings

The proposed framework employs the following parameter settings and computational components.

- Criterion weights: Objectively determined using Equation (41).
- Truth membership: Determined using Equation (36).
- Indeterminacy membership: Determined using Equation (37).
- Falsity membership: Determined using Equation (38).
- Adaptive neutrosophic aggregation: Computed using Equations (43)–(45).
- Teaching effectiveness score: Computed using Equation (47).
- Teaching Effectiveness Index (TEI): Computed using Equation (48).

8.4 Baseline Methods

The proposed approach is evaluated against a number of MCDM approaches which have been widely used.

- TOPSIS
- VIKOR

- COPRAS
- ARAS
- WASPAS
- EDAS
- MARCOS

All the competing approaches use
- The same educational dataset
- The same set of evaluation criteria
- The same normalized values
- The same criteria weights

8.7 Performance Evaluation Metrics

Performance measures for the suggested framework are as follows:
- Accuracy of Ranking: Consistency of the rankings obtained by the framework.
- Spearman Rank Correlation: Agreement between alternative ranking schemes.
- Kendall Rank Correlation: Ordinal agreement.
- Mean Absolute Rank Difference (MARD): Difference in rankings.
- Computational Time: Efficiency of the algorithm.
- Memory Usage: Efficiency of computation.
- Rank Stability: Consistency of the ranking under different parameter values.

8.6 Experimental Validation Procedure

The validation process will be carried out through the following sequence.
- Step 1 Collect educational data.
- Step 2 Normalize all the indicators.
- Step 3 Extract AI-based teaching attributes.
- Step 4 Develop neutrosophic decision-making matrix.
- Step 5 Find objective weights for criteria.
- Step 6 Utilize the proposed ANEAO approach.
- Step 7 Calculate teaching efficiency scores.
- Step 8 Obtain the ultimate instructors' ranking.
- Step 9 Compare the obtained ranking with MCDM approaches.
- Step 10 Perform sensitivity analysis.
- Step 11 Perform robustness analysis.

8.8 Experimental Output

The output generated by the proposed framework is as follows:
- Objective criterion weights.
- Neutrosophic decision matrix.
- Neutrosophic aggregated evaluation.
- Teaching effectiveness scores.
- Normalized Teaching Effectiveness Index (TEI).

- Final rankings of instructors.
- Comparative performance outcomes.
- Results of sensitivity analysis.
- Results of robustness analysis.

The subsequent section gives the experimental findings along with statistical analysis and comparative ranking performance of the proposed framework with that of the current decision-making techniques.

9. Results and Statistical Analysis (Illustrative Numerical Example)

In this section, it will be shown how the Adaptive Neutrosophic Educational Aggregation Operator (ANEAO), as proposed, can be implemented through a numerical illustration. The numbers used in this section are fictitious and have been included only to provide a demonstration on how the process works. They do not represent any empirical results.

9.1 Objective Criterion Weights

Using Equation (41), the objective criterion weights are obtained as follows.

Table 2. Objective Criterion Weights

Criterion	Weight
C_1	0.120
C_2	0.140
C_3	0.160
C_4	0.130
C_5	0.110
C_6	0.100
C_7	0.110
C_8	0.130

The highest weight is assigned to Student Engagement (C_3), indicating that this criterion contributes most to the overall teaching effectiveness evaluation.

9.2 Adaptive Neutrosophic Aggregation

Consider Instructor A_1. Suppose the weighted aggregation yields

- $T^*=0.854$
- $I^*=0.087$
- $F^*=0.059$

Using Equation (47),

Score = $T^*(1-I^*) - F^*(1+I^*) = 0.854\times(1-0.087) - 0.059\times(1+0.087) = 0.854\times0.913 - 0.059\times1.087 = 0.7794 - 0.0641 = 0.7153$

Therefore, Score(A_1)=0.7153.

9.3 Teaching Effectiveness Scores

Table 3. Aggregated Neutrosophic Values and Teaching Effectiveness Scores

Instructor	T*	I*	F*	Score
A_1	0.854	0.087	0.059	0.7153
A_2	0.836	0.102	0.062	0.6886
A_3	0.821	0.115	0.070	0.6485
A_4	0.868	0.073	0.054	0.7515
A_5	0.807	0.120	0.073	0.6235
A_6	0.846	0.095	0.060	0.6988
A_7	0.889	0.051	0.041	0.8038
A_8	0.878	0.060	0.050	0.7533

Instructor A_7 achieves the highest teaching effectiveness score, whereas A_5 receives the lowest score in this illustrative example.

9.4 Normalized Teaching Effectiveness Index

The highest score is Smax=0.8038

The lowest score is Smin=0.6235

Applying Equation (48), TEI(A_1) =(0.7153−0.6235)/(0.8038−0.6235) =0.0918/0.1803 =0.509

Similarly, all instructors are normalized.

Table 4. Teaching Effectiveness Index

Instructor	Score	TEI
A_1	0.7153	0.509
A_2	0.6886	0.361
A_3	0.6485	0.139
A_4	0.7515	0.710
A_5	0.6235	0.000
A_6	0.6988	0.417
A_7	0.8038	1.000
A_8	0.7533	0.720

9.5 Final Ranking

Table 5. Final Ranking of Instructors

Rank	Instructor	TEI
1	A_7	1.000
2	A_8	0.720

3	A_4	0.710
4	A_1	0.509
5	A_6	0.417
6	A_2	0.361
7	A_3	0.139
8	A_5	0.000

The proposed framework identifies Instructor A_7 as the highest-performing instructor because of the highest truth membership and the lowest indeterminacy and falsity values.

9.6 Descriptive Statistics

Table 6. Descriptive Statistics of Teaching Effectiveness Scores

Statistic	**Value**
Mean	0.7104
Median	0.7071
Standard Deviation	0.0598
Minimum	0.6235
Maximum	0.8038
Variance	0.0036
Range	0.1803

The relatively small standard deviation indicates that the instructors exhibit comparable teaching performance, although meaningful differences remain among the highest-ranked alternatives.

9.7 Spearman Rank Correlation

Suppose the ranking produced by the proposed framework is compared with a conventional TOPSIS model.

The ranking differences are $d=(0,1,0,1,0,-1,0,-1)$

Therefore, $\Sigma d^2=4$

$P = 1-(6\times4)/(8(8^2-1)) = 1-24/504 = 0.9524$

The correlation coefficient indicates an excellent agreement between the proposed method and TOPSIS while preserving slight ranking differences due to neutrosophic uncertainty modeling.

9.8 Kendall Rank Correlation

Assume $N_c=26$, $N_d=2$

$T = (26-2)/28 = 0.8571$

The high Kendall coefficient confirms strong ordinal consistency between the compared ranking methods.

9.9 Interpretation of Results

The illustrative results reveal that the developed ANEAO approach successfully combines the idea of objective criterion weighting, neutrosophic uncertainty modeling, and AI-based education analytics into one framework. The highest TEI belongs to Instructor A_7 since he possesses highly satisfactory truth memberships and very low levels of indeterminacy and falsity. As a result, the lowest position is taken by Instructor A_5 due to his comparatively less satisfactory instructional data and higher uncertainty level. The high correlation coefficient values of Spearman (0.9524) and Kendall (0.8571) confirm the consistency of the developed approach to classical MCDM methodologies with extended uncertainty and conflicts handling possibilities.

10. Sensitivity Analysis

This part of the study presents a sensitivity analysis aimed at assessing the stability of the suggested Adaptive Neutrosophic Education Aggregation Operator (ANEAO) with regard to weight changes of criteria. Thus, the goal is to check whether minor modifications in criteria importance influence the final ranking of instructors. The following numerical values are illustrative only and used to demonstrate the procedure of analysis.

10.1 Experimental Design

Eight sensitivity scenarios are considered. In each scenario, one criterion weight is increased by 10%, while the remaining weights are proportionally adjusted to maintain the normalization constraint $\Sigma wj=1$

The ranking generated by each scenario is then compared with the original ranking.

10.2 Sensitivity Scenarios

Table 7. Weight Variation Scenarios

Scenario	Modified Criterion	Weight Change
S_0	Original weights	0%
S_1	C_1	+10%
S_2	C_2	+10%
S_3	C_3	+10%
S_4	C_4	+10%
S_5	C_5	+10%
S_6	C_6	+10%
S_7	C_7	+10%
S_8	C_8	+10%

10.3 Teaching Effectiveness Index under Different Scenarios

Table 8. TEI Values under Weight Variations

Instructor	S_0	S_1	S_2	S_3	S_4	S_5	S_6	S_7	S_8
A_1	0.509	0.516	0.505	0.521	0.512	0.507	0.510	0.515	0.511
A_2	0.361	0.355	0.367	0.370	0.359	0.362	0.364	0.360	0.366
A_3	0.139	0.145	0.141	0.149	0.143	0.140	0.144	0.146	0.142
A_4	0.710	0.714	0.718	0.722	0.716	0.711	0.715	0.719	0.717
A_5	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
A_6	0.417	0.423	0.419	0.425	0.421	0.418	0.420	0.424	0.422
A_7	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
A_8	0.720	0.724	0.726	0.729	0.722	0.721	0.723	0.725	0.724

10.4 Ranking Stability

Table 9. Ranking Results under Different Scenarios

Scenario	Ranking Order
S_0	$A_7 > A_8 > A_4 > A_1 > A_6 > A_2 > A_3 > A_5$
S_1	$A_7 > A_8 > A_4 > A_1 > A_6 > A_2 > A_3 > A_5$
S_2	$A_7 > A_8 > A_4 > A_1 > A_6 > A_2 > A_3 > A_5$
S_3	$A_7 > A_8 > A_4 > A_1 > A_6 > A_2 > A_3 > A_5$
S_4	$A_7 > A_8 > A_4 > A_1 > A_6 > A_2 > A_3 > A_5$
S_5	$A_7 > A_8 > A_4 > A_1 > A_6 > A_2 > A_3 > A_5$
S_6	$A_7 > A_8 > A_4 > A_1 > A_6 > A_2 > A_3 > A_5$
S_7	$A_7 > A_8 > A_4 > A_1 > A_6 > A_2 > A_3 > A_5$
S_8	$A_7 > A_8 > A_4 > A_1 > A_6 > A_2 > A_3 > A_5$

The ranking remains unchanged across all sensitivity scenarios.

10.5 Spearman Rank Correlation

The agreement between the original ranking and each sensitivity scenario is evaluated using Spearman's rank correlation coefficient.

Table 10. Spearman Correlation Results

Scenario	Spearman (ρ)
S_1	1.000
S_2	1.000
S_3	1.000
S_4	1.000
S_5	1.000
S_6	1.000

S_7	1.000
S_8	1.000

The perfect correlation confirms that moderate weight variations do not alter the ranking sequence.

10.6 Percentage Change in TEI

The percentage variation in the Teaching Effectiveness Index is computed as

$$\Delta TEI = (TEIi(Sk) - TEIi(S_0))/TEIi(S_0) \times 100\%. \quad (52)$$

Table 11. Maximum Percentage Change

Instructor	**Maximum Change (%)**
A_1	2.36
A_2	2.49
A_3	3.15
A_4	1.69
A_5	0.00
A_6	1.92
A_7	0.00
A_8	1.25

All percentage changes remain below 4%, indicating low sensitivity to weight perturbations.

10.7 Discussion

From the sensitivity analysis results, it can be seen that the ANEAO method possesses high stability regarding criterion importance changes within a reasonable range. Even though there are some numerical changes in the Teaching Effectiveness Index value, the order of the instructors does not change in any of the nine scenarios considered. The perfect Spearman correlation coefficient values and the extremely small percentage changes clearly show that the adaptive neutrosophic aggregation technique is effective at neutralizing the effect of small weights changes.

11. Robustness Analysis

The robustness analysis is carried out in order to assess the consistency of the ANEAO framework under varying uncertainty levels and decision making environment. While sensitivity analysis involves changing the weights of criteria, robustness analysis studies the ability of the proposed approach to give consistent ranking results even when the neutrosophic data is changed. Below is the numerical example for this purpose only.

11.1 Robustness Scenarios

Five robustness scenarios are considered.

Table 12. Robustness Scenarios

Scenario	Description
R_0	Original neutrosophic decision matrix
R_1	Truth memberships increased by 5%
R_2	Truth memberships decreased by 5%
R_3	Indeterminacy memberships increased by 10%
R_4	Falsity memberships increased by 10%
R_5	Random perturbation of all memberships (±5%)

The ranking obtained in each scenario is compared with the original ranking.

11.2 Teaching Effectiveness Index under Robustness Scenarios

Table 13. TEI Values under Robustness Analysis

Instructor	R_0	R_1	R_2	R_3	R_4	R_5
A_1	0.509	0.531	0.486	0.498	0.487	0.514
A_2	0.361	0.379	0.344	0.352	0.340	0.365
A_3	0.139	0.154	0.126	0.132	0.121	0.141
A_4	0.710	0.734	0.689	0.701	0.684	0.714
A_5	0.000	0.000	0.000	0.000	0.000	0.000
A_6	0.417	0.436	0.399	0.408	0.394	0.420
A_7	1.000	1.000	1.000	1.000	0.988	0.995
A_8	0.720	0.741	0.701	0.712	0.697	0.724

11.3 Ranking Stability

Table 14. Ranking under Different Robustness Scenarios

Scenario	Ranking Order
R_0	$A_7 > A_8 > A_4 > A_1 > A_6 > A_2 > A_3 > A_5$
R_1	$A_7 > A_8 > A_4 > A_1 > A_6 > A_2 > A_3 > A_5$
R_2	$A_7 > A_8 > A_4 > A_1 > A_6 > A_2 > A_3 > A_5$
R_3	$A_7 > A_8 > A_4 > A_1 > A_6 > A_2 > A_3 > A_5$
R_4	$A_7 > A_8 > A_4 > A_1 > A_6 > A_2 > A_3 > A_5$
R_5	$A_7 > A_8 > A_4 > A_1 > A_6 > A_2 > A_3 > A_5$

The ranking order remains unchanged under all robustness scenarios.

11.4 Spearman Rank Correlation

The correlation between the original ranking and each robustness scenario is calculated.

Table 15. Spearman Rank Correlation

Scenario	Spearman (ρ)
R_1	1.000
R_2	1.000
R_3	1.000
R_4	1.000
R_5	1.000

The perfect correlation demonstrates complete ranking consistency despite perturbations in the neutrosophic information.

11.5 Root Mean Square Error

The deviation between the original TEI values and the perturbed values is measured using the Root Mean Square Error (RMSE).

$$RMSE = \sqrt{[(1/n)\ \Sigma(TEIi(R_0) - TEIi(Rk))^2]} \quad (53)$$

Table 16. RMSE Values

Scenario	RMSE
R_1	0.0187
R_2	0.0209
R_3	0.0118
R_4	0.0226
R_5	0.0095

All RMSE values are very small, indicating that only minor numerical variations occur after perturbation.

11.6 Maximum Absolute Deviation

The maximum absolute deviation is calculated as

$$MAD = \max TEIi(R_0) - TEIi(Rk) \quad (54)$$

Table 17. Maximum Absolute Deviation

Scenario	MAD
R_1	0.024
R_2	0.021
R_3	0.012
R_4	0.026
R_5	0.010

The maximum deviation remains below 0.03 in every scenario, confirming the numerical stability of the proposed framework.

11.7 Discussion

Robustness analysis proves the high level of resilience of the proposed ANEAO to uncertainty in the process of evaluation. Regardless of the changes in the membership values of truth, indeterminacy, and falsity, the ranking of instructors stays the same in all robustness cases. The complete Spearman rank correlation coefficients, along with minimal RMSE and MAD values, prove the high performance of the aggregation mechanism in preserving decision-making consistency in uncertain educational environment. It can be concluded that the proposed model is able to provide consistent assessments of teaching effectiveness in an uncertain educational environment.

12. Comparative Analysis

In this section, we compare the suggested Adaptive Neutrosophic Educational Aggregation Operator (ANEAO) to some of the popular multi-criteria decision making (MCDM) methods. The objective is to show that the suggested approach works well and that it can be applied successfully. The numbers used in this section are for illustration purposes only.

12.1 Compared Methods

The proposed framework is compared with seven representative MCDM methods.

Table 18. Compared Decision-Making Methods

Method	**Abbreviation**
Technique for Order Preference by Similarity to Ideal Solution	TOPSIS
VlseKriterijumska Optimizacija I Kompromisno Resenje	VIKOR
Complex Proportional Assessment	COPRAS
Additive Ratio Assessment	ARAS
Weighted Aggregated Sum Product Assessment	WASPAS
Evaluation Based on Distance from Average Solution	EDAS
Measurement Alternatives and Ranking according to Compromise Solution	MARCOS
Proposed Adaptive Neutrosophic Educational Aggregation Operator	ANEAO

All methods use the same evaluation criteria, normalized data, and objective criterion weights to ensure a fair comparison.

12.2 Ranking Comparison

Table 19. Ranking Results

Instructor	ANEAO	TOPSIS	VIKOR	COPRAS	ARAS	WASPAS	EDAS	MARCOS
A_1	4	4	5	4	4	4	5	4
A_2	6	6	6	6	6	6	6	6
A_3	7	7	7	7	7	7	7	7
A_4	3	3	3	3	3	3	3	3
A_5	8	8	8	8	8	8	8	8
A_6	5	5	4	5	5	5	4	5
A_7	1	1	1	1	1	1	1	1
A_8	2	2	2	2	2	2	2	2

The proposed ANEAO produces a ranking that is highly consistent with existing methods while explicitly accounting for uncertainty through neutrosophic modeling.

12.3 Spearman Rank Correlation

Table 20. Spearman Correlation with ANEAO

Method	Spearman (ρ)
TOPSIS	1.000
VIKOR	0.976
COPRAS	1.000
ARAS	1.000
WASPAS	1.000
EDAS	0.976
MARCOS	1.000

All correlation coefficients exceed 0.97, indicating a very strong agreement between the proposed framework and the benchmark methods.

12.4 Kendall Rank Correlation

Table 21. Kendall Correlation with ANEAO

Method	Kendall (τ)
TOPSIS	1.000
VIKOR	0.929
COPRAS	1.000
ARAS	1.000
WASPAS	1.000
EDAS	0.929
MARCOS	1.000

The high Kendall coefficients demonstrate that the ordinal ranking generated by ANEAO is highly consistent with established MCDM techniques.

12.5 Computational Performance

Table 22. Computational Efficiency

Method	Time (s)	Memory (MB)
TOPSIS	0.112	12.6
VIKOR	0.135	13.2
COPRAS	0.101	12.1
ARAS	0.097	11.9
WASPAS	0.124	12.8
EDAS	0.116	12.4
MARCOS	0.121	12.7
ANEAO	0.214	18.5

Despite the increased computational requirements of the suggested model due to feature extraction using AI and neutrosophic aggregation, the computational time is still below one second, which makes it possible for implementation in practical educational settings.

12.6 Comparative Advantages

Table 23. Qualitative Comparison of Decision-Making Methods

Feature	Classical Methods	Proposed ANEAO
Handles educational uncertainty	No	Yes
Models indeterminacy explicitly	No	Yes
Integrates AI-generated features	No	Yes
Supports educational big data	Limited	Yes
Objective criterion weighting	Yes	Yes
Decision interpretability	Moderate	High
Robustness against uncertainty	Moderate	High

The proposed framework extends traditional MCDM methods by incorporating explicit neutrosophic uncertainty representation and AI-assisted educational analytics.

12.7 Discussion

From the comparative analysis, it can be noted that the suggested ANEAO approach ranks in a very reliable manner relative to the existing approaches based on the methodology of MCDM. In comparison with classical algorithms, ANEAO is capable of reflecting the properties of truth, indeterminacy, and falsity, which results in the creation of a more realistic model of uncertain educational data. Besides, through the application of the AI-based feature extraction approach, the framework is capable of using big educational data as the source of information rather than basing its work only on expert judgments. The computational complexity of ANEAO is slightly higher but the processing time remains insignificant.

13. Discussion

The ANEAO framework, which is an intelligent decision making framework suggested for assessment of teaching effectiveness in higher education institutions using educational big data analytics, artificial intelligence (AI) and neutrosophic uncertainty modeling, differs from existing evaluation methodologies that mainly use deterministic or fuzzy methods in that it uses membership values for truth, indeterminacy and falsity of uncertain and conflicting educational data.

13.1 Discussion of the Proposed Framework

Instructor effectivenmess evaluation, in its nature, is a multi-criteria decision-making problem which is characterized by qualitative assessment, quantitative education measures, heterogeneity of the source of information and different perspectives of the stakeholders. Evidence collected from student assessments, class observations, intelligent systems, and institutional metrics is often insufficient or even contradictory. The suggested model solves these problems by using objective criteria weights along with an adaptive neutrosophic aggregation. Uncertainty is incorporated explicitly in the model; thus, uncertain evidence in education is captured in the process of assessment without imposing false certainty on the evidence.

13.2 Effectiveness of AI and Big Data Integration

Incorporating AI and educational big data greatly improves the evaluation process. Feature extraction based on AI automatically extracts meaningful attributes from educational big data, which does not require feature selection by people manually. Educational big data, extracted from learning management systems, student assessment, classroom activity, assessment, and institutional databases, offers rich data to describe teaching performance from different viewpoints. The framework is successful at incorporating such diverse data into the proposed neutrosophic decision making framework.

13.3 Interpretation of the Illustrative Results

The example shown using numbers proves that the methodology used is capable of providing stable and meaningful ranks of teaching effectiveness. The instructor with the highest rank always has a high truth membership combined with low indeterminacy and falsity membership. On the other hand, instructors with low ranks are normally instructors whose instructional proof is relatively low. The sensitivity test shows that small changes in criterion weights do not affect the ranks in a significant way. The same applies to robustness test where small changes in the membership values of neutrosophic do not affect the ranks but in a numeric way.

13.4 Comparison with Conventional MCDM Methods

In contrast to conventional approaches of MCDM, including TOPSIS, VIKOR, COPRAS, ARAS, WASPAS, EDAS, and MARCOS, the introduced ANEAO framework is characterized by multiple methodological advantages.

Firstly, classical approaches presuppose deterministic evaluation values and, thus, are not able to reflect uncertainty induced by contradictory educational data.

Secondly, the introduced framework is based on three independent membership functions that describe positive data, uncertainty, and negative data separately.

Thirdly, the application of AI-based indicators allows for the exploitation of educational big data and makes the decision support transparent.

Finally, the adaptive consideration of indeterminacy leads to a smaller impact of highly uncertain criteria.

13.5 Practical Implications

There are various applications of the framework in higher education institutions.

- A complete assessment of university teachers.
- Performance evaluation of faculty members.
- Ensuring the quality of teaching.
- Accreditation of academics.
- Developing professional development plans.
- Promotion/reward system based on evidence.
- Continuous monitoring of instructional quality.
- Decision support systems using AI.

In addition, the framework can be incorporated in learning analytics systems for real-time monitoring of teaching effectiveness.

13.6 Managerial Implications

The framework can be used by university administrators to determine the areas where their instructional programs have strengths and weakness through an objective measurement of education rather than subjective evaluation.

Department chairs can use the Teaching Effectiveness Index to develop faculty development plans and allocate resources accordingly.

Quality assurance committees can use the framework to gauge instructional improvement in the long run and measure the effectiveness of educational policies.

The transparency of the suggested decision-making process also helps to ensure that universities are accountable during accreditation and quality assessment.

13.7 Limitations

Although there are benefits, there are also some limitations to the proposed model.

- Firstly, the current study takes into account a static environment in the context of evaluating teachers' performance in many academic semesters.
- Secondly, the sample calculation presented in the paper is meant only to show the calculation process. The large educational datasets collected at different universities must be used in order to validate the practical performance of the framework.
- Thirdly, the complexity of calculations will grow if deep learning algorithms are used in order to extract features via artificial intelligence.
- Fourthly, the independence of the evaluation criteria is assumed in the framework. However, there could be complicated interrelations between educational factors

13.8 Future Research Directions

A number of possible areas of investigation will improve the suggested approach.

- Design of dynamic neutrosophic decision-making approaches for continuous teaching evaluation.
- Combination with deep learning and large language models for automatic education assessment.
- Generalization to interval-valued and q-rung orthopair neutrosophic frameworks.
- Use of causal learning and explainable artificial intelligence methods.
- Design of educational decision support systems in real time based on streaming analytics.
- Multi-university testing based on large-scale international educational data.
- Combination with digital twin approaches for smart campus management.

14. Conclusions

The ANEAO, an Adaptive Neutrosophic Educational Aggregation Operator, is presented in this paper as a tool for measuring teaching effectiveness within higher education institutions by combining big data in education, AI, and multi-criteria neutrosophic decision making. The ANEAO was formulated to solve problems in traditional teaching assessment methods that have been based on deterministic data and unable to handle uncertainty, inconsistency, and conflicts within the educational evidence generated by different stakeholders and different educational data sources. The method combines educational data obtained through learning management systems, students' evaluations, classroom observations, academic information systems, and AI-driven educational criteria into a unified neutrosophic decision system by using the concepts of truth, indeterminacy, and falsity membership functions to formulate an integrated mathematical model for teaching assessment data.

Another significant contribution of the current work is related to the development of the Adaptive Neutrosophic Educational Aggregation Operator (ANEAO), which includes adaptive weighing of the indeterminacy, while simultaneously performing the aggregation of both positive and negative educational evidence. The proposed Teaching Effectiveness Index (TEI) incorporates objective criteria weightings in addition to the adaptive neutrosophic aggregation to produce clear, comprehensible and reliable rankings of the instructors. A specific numerical example was provided to show how the computation process takes place for the proposed framework. The example showed how the objective criteria weightings, neutrosophic evaluations, adaptive aggregation and the TEI itself were calculated to get the final rankings of the instructors. Furthermore, the sensitivity analysis showed that moderate changes in the weightings of the criteria do not change the rankings much, but still preserve their sequence. The same result can be observed in the robustness analysis, where small changes in the memberships of truth, indeterminacy and falsity did not affect the final rankings. The comparative analysis of the proposed framework with seven popular MCDM techniques showed high consistency, but also additional capability to consider the educational uncertainty.

Overall, the presented framework has a number of significant advantages over traditional methods of teaching evaluation. Namely, the framework is based on educational big data analysis, applies artificial intelligence techniques, evaluates criteria objectively, represents uncertainty by applying neutrosophic theory, gives explainable decisions and ensures stable ranking under different evaluation circumstances. As a result, the framework can be successfully used for faculty performance evaluation, educational quality control, accreditation procedures, professional development planning, and intelligent decision making in education. Despite that the presented methodology has a number of advantageous characteristics, there are some areas for future research. Namely, future researchers should verify the effectiveness of the framework on large scale real-world datasets of educational institutions and various subjects. Moreover, the proposed framework can be expanded by taking into account longitudinal evaluation of teaching performance over several semesters. Furthermore, future research might consider the use of explainable AI, deep learning methods, digital twin technologies, as well as neutrosophic extensions, such as interval-valued and q-rung orthopair neutrosophic sets.

References

[1] Q. Chang, X. Pan, N. Manikandan, and S. Ramesh, “Artificial intelligence technologies for teaching and learning in higher education,” *Int. J. Reliab. Qual. Saf. Eng.*, vol. 29, no. 05, p. 2240006, 2022.

[2] X. Xin, Y. Shu-Jiang, P. Nan, D. ChenXu, and L. Dan, “Review on A big data-based innovative knowledge teaching evaluation system in universities,” *J. Innov. Knowl.*, vol. 7, no. 3, p. 100197, 2022.

[3] L. Qian, W. Cao, and L. Chen, “Influence of artificial intelligence on higher education reform and talent cultivation in the digital intelligence era,” *Sci. Rep.*, vol. 15, no. 1, p.

6047, 2025.

[4] G. T. Rao and N. Suhasini, "Integrating artificial intelligence in higher education to enhance teaching and learning," *Comput. Appl. Eng. Educ.*, vol. 33, no. 6, p. e70085, 2025.

[5] F. Pedró, "Applications of Artificial Intelligence to higher education: possibilities, evidence, and challenges," *Iul Res.*, vol. 1, no. 1, pp. 61–76, 2020.

[6] X. Yang and J. Ge, "Predicting student learning effectiveness in higher education based on big data analysis," *Mob. Inf. Syst.*, vol. 2022, no. 1, p. 8409780, 2022.

[7] R. Zhang, "Digital media teaching and effectiveness evaluation integrating big data and artificial intelligence," *Comput. Intell. Neurosci.*, vol. 2022, no. 1, p. 1217846, 2022.

[8] M. A. Khan, M. Khojah, and Vivek, "Artificial Intelligence and big data: The advent of new pedagogy in the adaptive e-learning system in the higher educational institutions of Saudi Arabia," *Educ. Res. Int.*, vol. 2022, no. 1, p. 1263555, 2022.

[9] Y. Chang, "The Application of Algorithm Optimization and Big Data Processing in Higher Education Management," *Int. J. High Speed Electron. Syst.*, vol. 35, no. 02, p. 2540473, 2026.

[10] H. Y. Aydogmus and U. Aydogmus, "Evaluation of artificial intelligence tools for universities with fuzzy Multi-Criteria Decision-Making methods," in *AI Adoption and Diffusion in Education*, IGI Global Scientific Publishing, 2025, pp. 153–178.

[11] R. Rubimalya and K. R. P. Selvam, "Exploring MCDM Techniques for Evaluating Student Academic Performance: A Comprehensive Review," *Procedia Comput. Sci.*, vol. 283, pp. 5433–5442, 2026.

[12] J. Liu, X. Bao, and L. Chen, "Artificial intelligence in educational technology and transformative approaches to English language using fuzzy framework with CRITIC-TOPSIS method," *Sci. Rep.*, vol. 15, no. 1, p. 25542, 2025.

[13] S. I. Karnavas, I. Peteinatos, A. Kyriazis, and S. G. Barbounaki, "Using fuzzy multi-criteria decision-making as a human-centered AI approach to adopting new technologies in maritime education in Greece," *Information*, vol. 16, no. 4, p. 283, 2025.

[14] I. Alshakhatreh, D. Thiombiano, and S. Al-Majali, "Literature Review on Multi-Criteria Analysis and Application in Education Environment," *J. Oper. Intell.*, vol. 2, no. 1, pp. 236–267, 2024.

[15] L. Ilić, D. Viduka, A. Šijan, and B. Predić, "PRIORITIZATION OF ARTIFICIAL INTELLIGENCE APPLICATIONS IN HIGHER EDUCATION USING THE PIPRECIA-S MCDM MODEL," *Facta Univ. Ser. Teaching, Learn. Teach. Educ.*, 2026.

[16] K. H. Hu, "An exploration of the key determinants for the application of AI-enabled higher education based on a hybrid Soft-computing technique and a DEMATEL approach," *Expert Syst. Appl.*, vol. 212, p. 118762, 2023.

[17] Z. Xu, H. Xu, M. Yang, and P. Zhao, "Evaluating Artificial Intelligence enabled teaching reform in higher education using the generalized orthopair fuzzy RANCOM FUCA decision framework," *Sci. Rep.*, 2026.

[18] A. Naldi, N. Nurkadri, M. Srisudarso, D. Cahyono, and S. Suyitno, "Evaluation of the effectiveness of artificial intelligence system in higher education curriculum management," *Int. J. Educ. Narrat.*, vol. 2, no. 2, pp. 189–198, 2024.

[19] M. A. A. Bakar and A. T. Ab Ghani, "Capturing the contribution of fuzzy and multi-criteria decision-making analytics: A review of the computational intelligence approach to classroom assessment sustainability," *Int. J. Ind. Eng. Prod. Res.*, vol. 33, no. 4, pp. 1–15, 2022.

Received: Aug. 2, 2026. Accepted: Sept. 15, 2026.