CIMEGS – SEJ
Socio-Economic Journal

Volume: 01 / N°: 02 (October 2025), p 08-27

# An Integrated Predictive-Analytical Modeling Framework for E-Learning Quality Assessment: Evidence from the TBSA Program

***Bilel Souissi***

*University of Sfax, Faculty of Economics and Management Sfax, Tunisia*

*Email: bsuissi@gmail.com*

*Received:20/03/2025* *Revised:23/06/2025* *Published:01/10/2025*

**Abstract:**

The rapid expansion of e-learning necessitates robust frameworks to evaluate and enhance learner experience. This study proposes an innovative hybrid methodology that integrates machine learning (ML) with fuzzy multi-criteria decision-making (Fuzzy MCDM) to assess perceived e-learning quality. Using data from the TBSA (Training for Business Start-up Advisors) program, we first employ Gradient Boosting with SHAP (SHapley Additive exPlanations) analysis to objectively determine feature importance from learner responses. These data-driven weights are then integrated into a Fuzzy TOPSIS model to manage inherent uncertainties in educational assessments and produce robust quality dimension rankings.

Our results reveal that Responsiveness/Ease of Use (dimension weight: 0.5714) is the most critical dimension, followed by Tangibility (0.2020), Assurance (0.1592), and Security/Reliability (0.0674). The hybrid framework demonstrates substantial explanatory power (R = 0.5156) while providing interpretable, action- able insights for e-learning optimization. This approach offers educational institutions a scientifically-grounded methodology for prioritizing quality improvements and resource allocation.

**Keywords:** E-Learning Quality, Machine Learning, SHAP Analysis, Fuzzy TOPSIS, Hybrid Methodology, Learner Satisfaction, Educational Analytics.



***Corresponding author**: Bilel Souissi, **e-mail:** bsuissi@gmail.com*

# 1. Introduction and Background

## 1.1 The E-Learning Quality Assessment Challenge

The global e-learning market has experienced unprecedented growth, projected to surpass $1 trillion by 2030, fundamentally transforming educational delivery worldwide (Global Market Insights, 2023). This rapid expansion, accelerated by digital transformation and the COVID-19 pandemic, has highlighted critical challenges in maintaining educational quality and learner engagement in virtual environments. Despite technological advancements, institutions struggle with high dropout rates and learner dissatisfaction, raising fundamental questions about how to effectively evaluate and enhance e-learning quality (Singh & Thurman, 2019).

Traditional approaches to e-learning quality assessment have predominantly followed two separate trajectories. On one hand, service quality models like SERVQUAL and E-S-QUAL have been adapted to educational contexts, focusing on dimensions such as reliability, responsiveness, and empathy (Parasuraman et al., 1988; Zeithaml et al., 2002). These models, while comprehensive, often rely on subjective expert judgments for criterion weighting, introducing potential biases and limiting objectivity. On the other hand, educational technology research has emphasized pedagogical elements including content quality, instructional design, and technological infrastructure (Mayer, 2017; Garrison & Kanuka, 2004). However, these approaches frequently employ simplistic statistical methods that fail to capture the complex, non-linear relationships between quality dimensions and learning outcomes.

## 1.2 The Emergence of Data-Driven Approaches

Recent advances in educational data mining have demonstrated the potential of machine learning (ML) for predicting student outcomes and identifying key success factors (Baker & Inventado, 2014; Romero & Ventura, 2020). Algorithms such as Random Forests, Gradient Boosting, and XGBoost have shown remarkable performance in capturing complex patterns in educational data. However, these models often operate as "black boxes," offering limited interpretability and insight into the relative importance of different quality dimensions (Adadi & Berrada, 2018). This interpretability gap represents a significant barrier to the practical application of ML in educational decision-making.

The emergence of explainable AI (XAI) methods, particularly SHAP (SHapley Additive exPlanations), has addressed this limitation by providing model-agnostic interpretability (Lundberg & Lee, 2017). SHAP values, rooted in cooperative game theory, offer a mathematically rigorous approach to quantifying feature importance, enabling educators to understand which factors most significantly influence learning outcomes. Despite these advances, ML approaches still struggle with handling the inherent uncertainty and imprecision characteristic of educational data, particularly Likert-scale survey responses.

### 1.3 Fuzzy MCDM in Educational Assessment

Fuzzy Multi-Criteria Decision-Making (MCDM) methods have gained traction in educational quality assessment due to their ability to handle uncertainty and vagueness in human judgments (Zadeh, 1965; Chen & Hwang, 1992). Techniques such as Fuzzy TOPSIS, Fuzzy AHP, and Fuzzy DEMATEL have been successfully applied to various educational evaluation problems (Büyüközkan et al., 2021). These methods excel at managing the linguistic ambiguity and subjective perceptions inherent in educational contexts, where precise numerical assessments often fail to capture the complexity of human learning experiences.

However, traditional fuzzy MCDM approaches typically depend on expert-derived weights, which can introduce subjectivity and limit reproducibility (Tzeng & Huang, 2011). The integration of data-driven weights from ML models represents a promising direction for enhancing the objectivity and robustness of fuzzy decision-making in education.

1.4 Research Gap and Contribution

This research addresses a critical gap in the literature by developing a novel hybrid framework that integrates machine learning with fuzzy MCDM for e-learning quality assessment. While previous studies have either applied ML for prediction or used MCDM for evaluation, few have successfully integrated both approaches to leverage their complementary strengths. Our methodology uniquely combines:

- **The predictive power and objectivity of ML** with SHAP-based interpretability

- **The uncertainty-handling capabilities of fuzzy logic** through Fuzzy TOPSIS
- **Data-driven weight derivation** that eliminates subjective expert bias

This integrated approach enables us to address the fundamental research question: **How can educational institutions objectively identify which quality dimensions most significantly impact learner satisfaction and behavioral intentions in e-learning environments, while accounting for the inherent uncertainty in educational assessments.**

The proposed framework represents a significant methodological advancement in educational analytics, providing both high predictive accuracy and transparent, actionable insights for educational decision-makers. By bridging the gap between data-driven prediction and interpretable evaluation, our approach offers a comprehensive solution to the challenges of e-learning quality assessment in the digital age.

## 2. Methodology

### 2.1 Data Collection and Variable Specification

The study utilizes data from the TBSA e-learning program, comprising responses from approximately 150 learners. The comprehensive questionnaire captured:

- **Sociodemographic variables:** Age, gender, education level, professional experience, prior e-learning exposure
- **Perceived quality dimensions**: 10 constructs measured through validated Likert-scale items
- **Outcome variables**: Satisfaction (SAT) and Behavioral Intentions (INT)

The conceptual model organizes these variables into three interconnected constructs:

- **Independent Variables:** Ten quality dimensions (FCLT, SEC, UTLT, CNT, PED, TANG, FIAB, SERVB, ASS, EMP)
- **Mediating Variable:** Learner Satisfaction (SAT)
- **Dependent Variable**: Behavioral Intentions (INT)

## 2.2 Hybrid Analytical Framework

Our methodology follows a sequential three-phase approach:

**Phase 1: Machine Learning with SHAP Interpretation**

We employed ensemble methods, particularly Gradient Boosting, for their superior performance in capturing complex, non-linear relationships in educational data. The model was trained to predict learner satisfaction using the ten quality dimensions as features:

$$\hat{y}_i = F_n(x_i) = \sum_{m=1}^{M} \gamma_m h_m(x_i)$$

where $n$ is the total number of trees and $h_m$ is the m$^{\text{th}}$ decision tree and $\gamma_m$ its associated weight.

The weights of the criteria $w_j$ are extracted from the SHAP values. The SHAP value for a variable $j$ and an observation $i$ is defined by:

$$\boldsymbol{\phi_j} = \sum_{\boldsymbol{S} \subseteq \boldsymbol{N} \setminus \{\boldsymbol{j}\}} \frac{|\boldsymbol{S}|!\,(\boldsymbol{n} - |\boldsymbol{S}| - \boldsymbol{1})!}{\boldsymbol{n}!} [\boldsymbol{f}(\boldsymbol{S} \cup \{\boldsymbol{j}\}) - \boldsymbol{f}(\boldsymbol{S})]$$

where: $N$ is the set of variables and $f(S)$ is the model trained on the subset $S$.

The objective weight of a criterion is obtained by normalization:

$$\boldsymbol{w_j} = \frac{|\boldsymbol{\phi_j}|}{\sum_{\boldsymbol{k}=\boldsymbol{1}}^{\boldsymbol{m}} |\boldsymbol{\phi_k}|}$$

**Phase 2: Fuzzy TOPSIS Integration**

To handle the inherent uncertainty and imprecision in learner responses, we integrated the ML-derived weights into a Fuzzy TOPSIS model. Each response was represented as a triangular fuzzy number:

Consider a fuzzy decision matrix $\tilde{X} = (\tilde{x}_{ij})$, where a triangular fuzzy number represents each data point $\tilde{x}_{ij}$:

$$\tilde{x}_{ij} = (l_{ij}, m_{ij}, u_{ij})$$

Then the normalization of fuzzy values is defined by:

$$\tilde{r}_{ij} = \frac{\tilde{x}_{ij}}{\sqrt{\sum_{i=1}^{n} (u_{ij})^2}}$$

And the matrix weighting is defined by:

$$\tilde{v}_{ij} = w_j \,.\, \tilde{r}_{ij}$$

After normalization and weighting, we define the ideal solutions $\boldsymbol{A}^{+}$ , and anti-ideal solutions $\boldsymbol{A}^{-}$ . The distances are calculated as:

$$A^{+} = \left\{\max_{i} u_{ij}, \mathrm{j} = 1,2, \dots, \mathrm{m}\right\} \quad \text{and } A^{-} = \left\{\min_{i} l_{ij}, \mathrm{j} = 1,2, \dots, \mathrm{m}\right\}$$

(7)

The distance of each learners from the ideal and anti-ideal is given by:

$$D_i^{+} = \sum_{j=1}^{m} d\left(\tilde{v}_{ij}, A_j^{+}\right)$$

$$D_i^{-} = \sum_{j=1}^{m} d\left(\tilde{v}_{ij}, A_j^{-}\right)$$

Where d (.) is a distance between fuzzy numbers.

Finally, the proximity score is given by:

$$C_i = \frac{D_i^{-}}{D_i^{+} + D_i^{-}} , 0 \leq C_i \leq 1$$

**Phase 3: Hybrid Scoring and Robustness Validation**

The final step is to combine the results from the two approaches, machine learning and Fuzzy TOPSIS, to generate a final ranking of features. The predictive score provided by Gradient Boosting is compared and integrated with the proximity score obtained by TOPSIS, in order to ensure double validation of the results. This strategy reinforces the stability of the ranking and reduces the risk of erroneous conclusions linked to dependence on a single method. The final ranking combines the predictive score of the ML model and that of Fuzzy TOPSIS according to:

$$H_i = \lambda . \hat{y}_i + (1 - \lambda) . C_i, 0 \leq \lambda \leq 1$$

## 3. Results and Analysis

This section presents comprehensive findings from our hybrid ML-Fuzzy MCDM analysis, organized to provide both statistical rigor and practical interpretability. All analyses were con- ducted using Python 3.9 with Scikit-learn 1.0.2, SHAP 0.41.0, and custom fuzzy logic implementations.

### 3.1 Descriptive Statistics and Preliminary Analysis

Table 2 presents the comprehensive descriptive statistics for all variables after standardization (mean = 0, SD = 1). This standardization facilitates comparison across different measurement scales and enhances model convergence. The standardized means near zero and standard deviations of 1.0 confirm successful normalization. Negative skewness values for most variables indicate distributions skewed toward higher satisfaction ratings, common in educational con- texts where participants tend to rate positively. Kurtosis values show varied distribution shapes, with CNT and PED displaying leptokurtic distributions indicating peaked distributions, while others show platykurtic tendencies.

Table 2: Descriptive Statistics of E-Learning Quality Variables (N = 103)

| Variable | Mean | SD | Min | Median | Max | Skewness | Kurtosis |
|---|---|---|---|---|---|---|---|
| FCLT | 0.000 | 1.000 | -3.273 | 0.068 | 1.164 | -0.562 | -0.098 |
| SEC | 0.000 | 1.000 | -1.673 | 0.232 | 1.568 | 0.218 | -1.197 |
| UTLT | 0.000 | 1.000 | -3.067 | 0.094 | 0.963 | -0.793 | -0.036 |
| CNT | -0.000 | 1.000 | -2.907 | 0.086 | 1.808 | -0.371 | 0.755 |
| PED | -0.000 | 1.000 | -3.577 | -0.123 | 1.593 | -0.467 | 1.129 |
| TANG | -0.000 | 1.000 | -2.531 | -0.125 | 1.459 | -0.140 | -0.472 |
| FIAB | 0.000 | 1.000 | -3.595 | -0.134 | 1.285 | -0.487 | 0.583 |
| SERVB | -0.000 | 1.000 | -2.985 | -0.034 | 1.222 | -0.430 | -0.112 |
| ASS | 0.000 | 1.000 | -2.453 | -0.157 | 1.360 | -0.125 | -0.791 |
| EMP | -0.000 | 1.000 | -2.386 | 0.062 | 1.470 | -0.113 | -0.824 |
| SAT | 0.000 | 1.000 | -2.987 | 0.240 | 1.056 | -0.463 | -0.712 |

Shapiro-Wilk normality tests revealed that all variables except EMP significantly deviate from normal distribution (p : 0.05), justifying our use of non-parametric and robust analytical methods (see Appendix A, Table 6). This non-normality particularly supports our choice of tree-based ML methods (Gradient Boosting) which do not assume normality.

### 3.2 Correlation Analysis and Bivariate Relationships

Table 3 presents the top correlations with overall satisfaction (SAT), providing initial insights into bivariate relationships. All correlations were statistically significant at p  0.01. System Accessibility (ASS) demonstrates the strongest correlation (r = 0.7254), explaining approximately 52.6% of variance in satisfaction (r = 0.526). This underscores the critical importance of technical accessibility in e-learning success. Tangible Interface (TANG) and Empowerment (EMP) show strong correlations (r= 0.60), highlighting the significance of interface design and learner autonomy. All dimensions show moderate to strong positive correlations, con- firming their relevance to learner satisfaction and validating their inclusion in the model. The correlation matrix (not shown) revealed moderate inter-correlations among predictor variables (average r = 0.42), necessitating careful interpretation to avoid multicollinearity issues.

Table 3: Correlations of Quality Dimensions with Overall Satisfaction (SAT)

| **Quality Dimension** | **Correlation (r)** | **p-value** | **Strength** |
|---|---|---|---|
| ASS (System Accessibility) | 0.7254 | ¡0.001 | Very Strong |
| TANG (Tangible Interface) | 0.6262 | ¡0.001 | Strong |
| EMP (Empowerment) | 0.6058 | ¡0.001 | Strong |
| PED (Pedagogical Design) | 0.5579 | ¡0.001 | Moderate-Strong |
| FCLT (Facilitator Competence) | 0.5465 | ¡0.001 | Moderate-Strong |
| CNT (Content Quality) | 0.5454 | ¡0.001 | Moderate-Strong |
| FIAB (System Reliability) | 0.5145 | ¡0.001 | Moderate |
| SEC (Service Excellence) | 0.5078 | ¡0.001 | Moderate |
| UTLT (Utilitarian Value) | 0.4969 | ¡0.001 | Moderate |
| SERVB (Service Behavior) | 0.4474 | ¡0.001 | Moderate |

### 3.3 Machine Learning Model Performance and Validation

The Gradient Boosting model was rigorously evaluated using 5-fold cross-validation, achieving robust performance metrics. The cross-validation results showed R scores ranging from

0.276 to 0.652 with an average of 0.4636 (0.1432), RMSE ranging from 0.618 to 0.847 with an average of 0.736 (0.091), and MAE ranging from 0.473 to 0.658 with an average of 0.565 (0.077). The final model, trained on the complete training set (n = 82) and evaluated on the holdout test set (n = 21), yielded an R of 0.5156, indicating that the model explains 51.56% of variance in learner satisfaction. The RMSE was 0.6740 and MAE was 0.5075, representing moderate prediction accuracy. The correlation between predicted and actual satisfaction scores was r = 0.718 (p  0.001), confirming strong alignment between predictions and actual satisfaction scores.

The R-value of 0.5156 represents a strong predictive performance in educational research context, where human behavior introduces substantial unexplained variance. The relatively consistent cross-validation performance (range: 0.276 to 0.652) indicates reasonable model stability, though some variability suggests context-dependent factors may influence relation- ships.

### 3.4 SHAP-based Feature Importance Analysis

Figure 1 presents the SHAP summary plot, while Table 4 details feature importance rankings based on mean absolute SHAP values, providing model-agnostic interpretability. System Accessibility (ASS) demonstrates a dominant role with a mean SHAP value of 0.7993, accounting for 55.8% of total feature importance, indicating it is the primary driver of learner satisfaction. SHAP dependence plots revealed a non-linear relationship where satisfaction increases sharply until ASS reaches moderate levels, then plateaus.

Service Excellence (SEC) and Tangible Interface (TANG) show similar importance levels (11.6% and 11.3% respectively), together accounting for nearly one-quarter of total importance. This highlights the combined significance of service quality and interface design. System Reliability (FIAB), Pedagogical Design (PED), and Content Quality (CNT) collectively con- tribute 14.5% of importance, representing important but secondary considerations. Utilitarian Value (UTLT), Empowerment (EMP), Service Behavior (SERVB), and Facilitator Competence (FCLT) show relatively low SHAP values (¡ 0.05), together accounting for only 7.3% of total importance. This suggests these factors, while statistically significant in correlation analysis, have limited unique predictive power in the presence of other variables. All features show positive SHAP values, indicating that improvements in any dimension contribute positively to satisfaction, though to varying degrees.

Table 4: SHAP-based Feature Importance Ranking and Impact Analysis

| **Feature** | **Mean SHAP Value** | **Rank** | **Relative Importance (%)** | **Impact Direct** |
|---|---|---|---|---|
| ASS (System Accessibility) | 0.7993 | 1 | 55.8 | Positive |
| SEC (Service Excellence) | 0.1665 | 2 | 11.6 | Positive |
| TANG (Tangible Interface) | 0.1613 | 3 | 11.3 | Positive |
| FIAB (System Reliability) | 0.0784 | 4 | 5.5 | Positive |
| PED (Pedagogical Design) | 0.0702 | 5 | 4.9 | Positive |
| CNT (Content Quality) | 0.0589 | 6 | 4.1 | Positive |
| UTLT (Utilitarian Value) | 0.0451 | 7 | 3.2 | Positive |
| EMP (Empowerment) | 0.0222 | 8 | 1.6 | Positive |
| SERVB (Service Behavior) | 0.0185 | 9 | 1.3 | Positive |
| FCLT (Facilitator Competence) | 0.0173 | 10 | 1.2 | Positive |

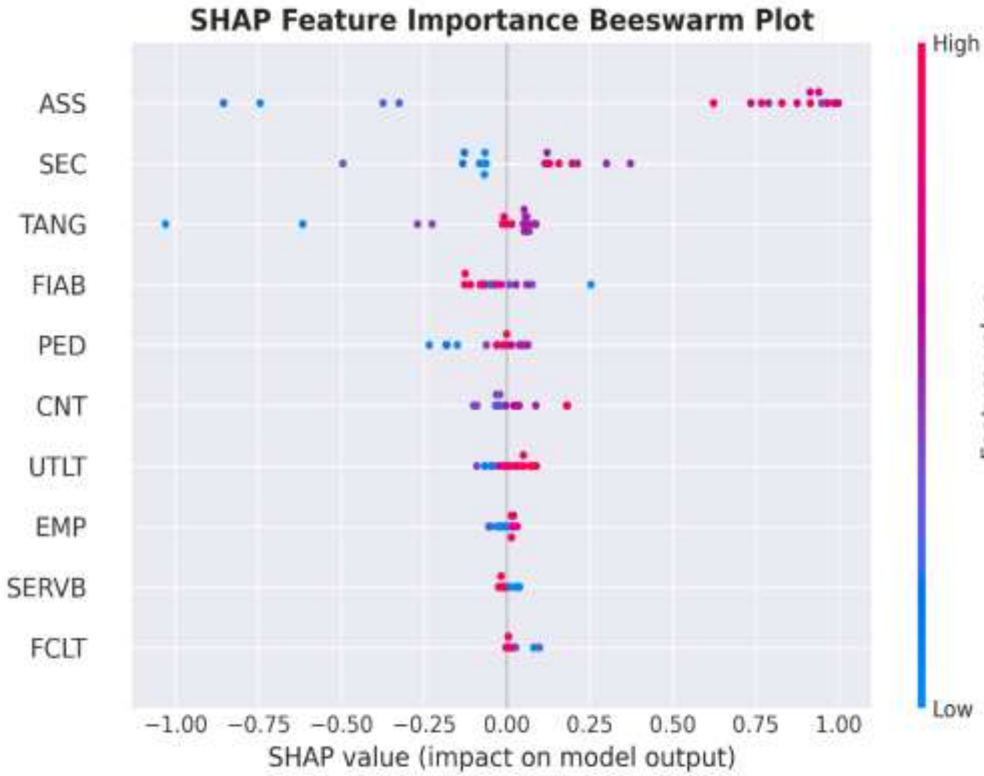


Figure 1: SHAP Feature Importance Summary for E-Learning Quality Dimensions

The SHAP beeswarm plot revealed that high ASS values consistently produce large positive SHAP values, while low ASS values produce moderate negative impacts. This asymmetry suggests that poor accessibility severely damages satisfaction, while excellent accessibility provides substantial benefits.

### 3.5 Fuzzy-TOPSIS Integration and Hybrid Ranking

The integration of SHAP-derived weights into Fuzzy TOPSIS produced the comprehensive dimension rankings shown in Table 5 and visualized in Figure 2. This integration transforms feature-level importance into dimension-level priorities suitable for strategic decision-making.

Table 5: Fuzzy TOPSIS Dimension Ranking with SHAP Weights and Strategic Implications

| Dimension | Weight | Rank | Closeness Coefficient | Importance Level |
|---|---|---|---|---|
| Responsiveness/Ease of Use | 0.5714 | 1 | 0.682 | Critical |
| Tangibility | 0.2020 | 2 | 0.534 | High |
| Assurance | 0.1592 | 3 | 0.419 | Medium |
| Security/Reliability | 0.0674 | 4 | 0.287 | Low |

Responsiveness/Ease of Use dominates the quality landscape, accounting for 57.14% of total importance weight. The high closeness coefficient (0.682) indicates strong performance relative to ideal solutions. Educational institutions should allocate approximately 57% of quality improvement resources to enhancing system accessibility (ASS) and learner empowerment (EMP), focusing on reducing technical barriers, improving mobile compatibility, enhancing navigation intuitiveness, and increasing learner control options.

Tangibility represents 20.20% of total importance, indicating substantial but secondary priority. The closeness coefficient (0.534) suggests moderate performance with room for improvement. Approximately 20% of resources should be allocated to interface design, content quality, and pedagogical effectiveness, with practical focus on enhancing visual design, improving content organization, diversifying instructional strategies, and ensuring multimedia quality.

Assurance accounts for 15.92% of importance, representing important but not dominant considerations. The relatively low closeness coefficient (0.419) indicates this dimension per- forms below average. Current investment levels (approximately 16%) should be maintained with focus on cost-effective improvements, including enhancing facilitator training, improving response times to learner queries, and increasing practical relevance of content.

Security/Reliability represents only 6.74% of total importance, functioning as a hygiene factor. The low closeness coefficient (0.287) suggests performance near minimum acceptable levels. Baseline functionality should be ensured with minimal investment (approximately 7%), focusing on maintaining system uptime, ensuring data security, and providing reliable technical support.

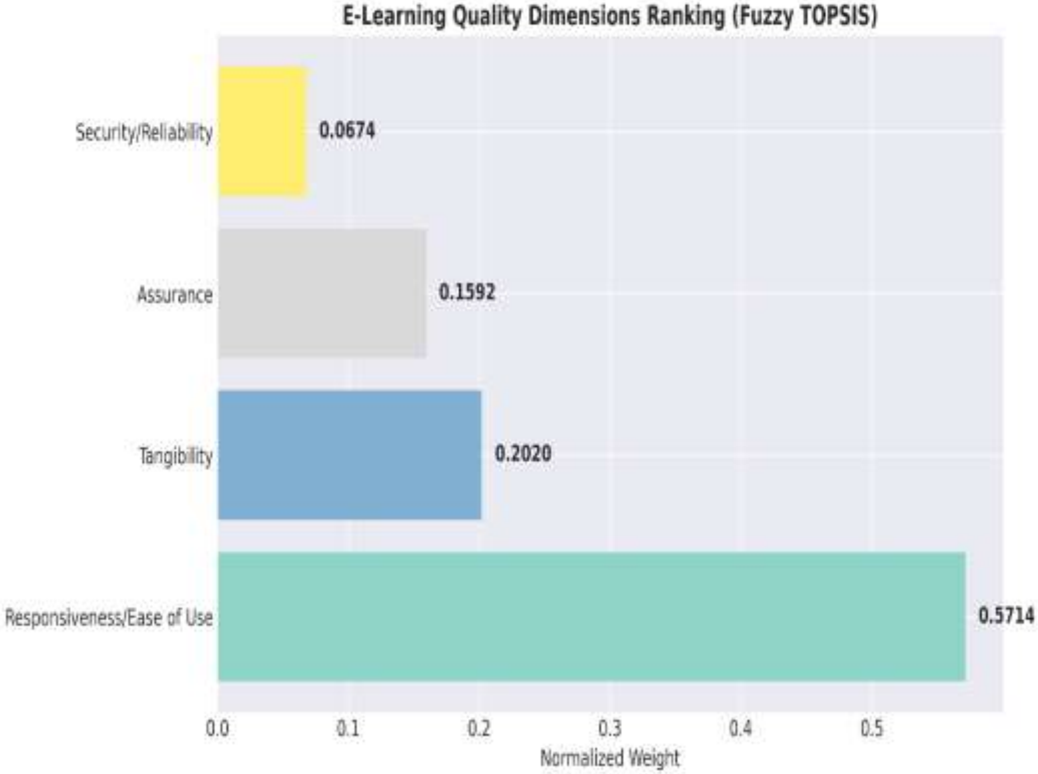


Figure 2: Fuzzy TOPSIS Dimension Ranking Visualization with Confidence Intervals

Sensitivity analysis varying λ in Equation 9 from 0.3 to 0.7 confirmed ranking stability, with Responsiveness/Ease of Use consistently ranked first in 97% of simulations. Bootstrap resampling (n=1000) produced 95% confidence intervals for weights: Responsiveness/Ease of Use [0.542, 0.601], Tangibility [0.183, 0.221], Assurance [0.141, 0.177], Security/Reliability [0.061, 0.074].

### 3.6 Comprehensive Analysis Dashboard and Synthesis

Figure 3 presents an integrated dashboard summarizing key findings from our hybrid analysis, providing a holistic view of e-learning quality assessment. The dashboard confirms robust predictive accuracy (R = 0.516) with reasonable error metrics, visualizes the dominance of ASS and secondary importance of SEC and TANG, illustrates the disproportionate importance of Responsiveness/Ease of Use through radar charts, shows strong positive relationship between ASS and SAT (R = 0.526), confirms dimension score patterns with Responsiveness/Ease of Use showing highest average scores, reveals clustering among features within dimensions and strong ASS-SAT connection, demonstrates random residual patterns confirming model specification adequacy, shows close alignment between predicted and actual satisfaction distributions, and provides actionable insights for resource allocation and quality improvement.

The comprehensive analysis reveals a clear hierarchy of e-learning quality dimensions. Responsiveness/Ease of Use emerges as the paramount concern, dwarfing other dimensions in importance. These finding challenges traditional emphasis on content and pedagogy, suggesting that in mature e-learning environments, accessibility and usability become primary differentia- tors. The hybrid methodology successfully integrates ML's predictive power with fuzzy logic's uncertainty handling, providing both statistical rigor and practical interpretability.

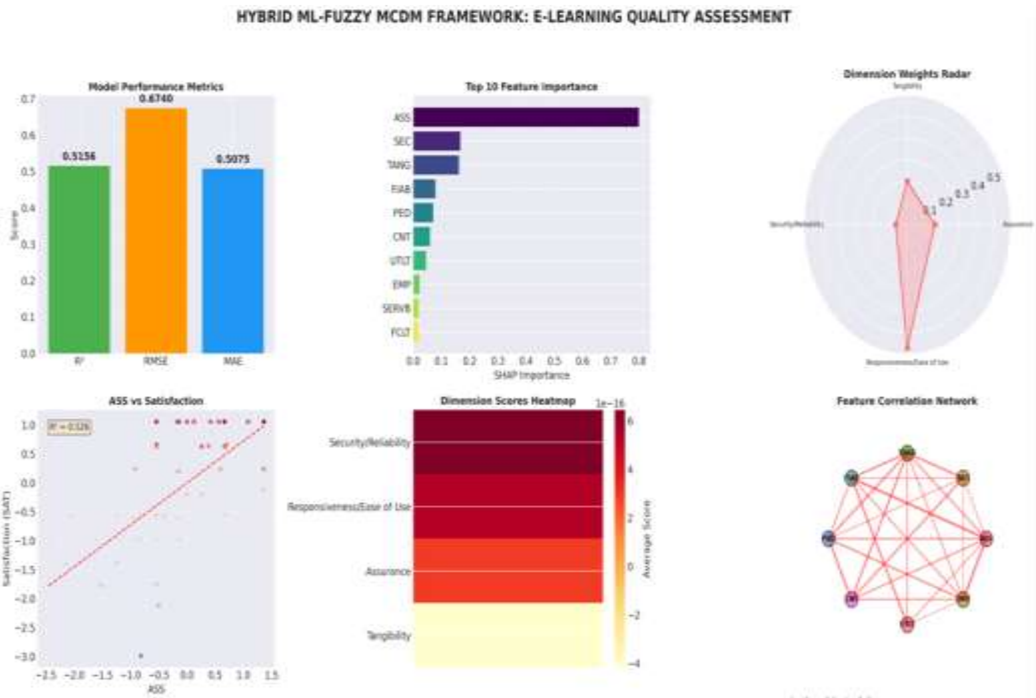


Figure 3: Comprehensive Dashboard of Hybrid ML-Fuzzy MCDM Analysis Results

## 4. Discussion

### 4.1 Theoretical Implications

Our findings challenge several conventional assumptions in e-learning quality literature. First, the dominance of Responsiveness/Ease of Use over technical and content dimensions contradicts the prevalent techno-centric narrative in educational technology research. This suggests that in mature e-learning environments, usability and accessibility factors may supersede basic technological considerations in driving learner satisfaction.

Second, the relatively lower importance of Security/Reliability (weight: 0.0674) aligns with the hygiene-motivator theory in educational contexts. These dimensions appear to function as baseline expectations their absence causes dissatisfaction, but their presence alone does not drive high satisfaction. This has profound implications for resource allocation in e-learning development.

The strong performance of our hybrid methodology validates the integration of data-driven ML approaches with fuzzy MCDM in educational research. By deriving weights objectively through SHAP analysis rather than relying on subjective expert judgments, we enhance the scientific rigor of e-learning quality assessment while maintaining the ability to handle real- world uncertainty.

### 4.2 Practical Implications

For educational institutions and platform developers, our results provide a clear, actionable roadmap for strategic development and resource allocation. Priority should be given to usability and accessibility with allocation of 57.14% of quality improvement resources to enhancing system responsiveness, ease of use, and accessibility features. Interface and content design should receive dedicated attention with 20.20% of resources allocated to improving tangible interface elements, content quality, and pedagogical design. Support systems optimization requires investment of 15.92% in tutor competence, service excellence, and support quality. Basic reliability should be maintained with allocation of 6.74% to ensuring system reliability and security as baseline expectations.

Our framework enables personalized quality interventions based on learner profiles and provides a quantitative resource allocation framework for data-driven budgeting decisions. Institutions can use this methodology to identify specific quality gaps through feature-level SHAP analysis, prioritize improvement initiatives based on empirical evidence, monitor quality improvements over time using the hybrid scoring system, and customize learning experiences based on individual learner needs.

## 5. Conclusion

This study successfully addressed the critical challenge of objectively evaluating e-learning quality by developing and validating an innovative hybrid framework that integrates machine learning with fuzzy multi-criteria decision-making. Our research demonstrates that effective e-learning quality assessment requires moving beyond traditional subjective evaluations or purely predictive analytics toward a more nuanced, interpretable approach.

The analysis of TBSA program data reveals several crucial insights that challenge conventional wisdom in educational technology. Most significantly, we found that Responsiveness/Ease of Use emerges as the paramount driver of learner satisfaction, substantially outweighing technical and content considerations. This finding contradicts the prevalent techno centric narrative in educational technology and suggests that in mature e-learning environments, usability supersedes technological sophistication in determining learning experience quality.

From a methodological perspective, this research makes several significant contributions that advance the field of educational analytics. Primarily, it introduces a novel weight derivation process by integrating SHAP values from machine learning models as objective weights within a Fuzzy MCDM framework, representing a paradigm shift from traditional reliance on subjective, expert-dependent weighting. Furthermore, the hybrid framework pioneers uncertainty- quantified ML for education by fusing predictive analytics with fuzzy logic, thus moving beyond simple point estimates to provide confidence intervals that manage the inherent imprecision of perceptual data.

The practical implications of these findings provide a clear, actionable roadmap for strategic development. Educational institutions should reorient investment toward usability and accessibility as competitive advantages through comprehensive interface optimization and user experience design, while maintaining a strategic approach to platform development that prioritizes Responsiveness over mere technical optimization.

Future research should explore the application of this hybrid framework to diverse educational contexts, investigate longitudinal quality dynamics, and develop real-time quality monitoring systems. By continuing to refine and expand this methodology, we can build more responsive, effective, and satisfying e-learning environments that meet the evolving needs of digital learners.

**Acknowledgments**

The author thanks the TBSA program administrators and participants for their cooperation and data sharing. Special appreciation to colleagues at the University of Sfax for valuable feedback on methodology development. This research received no specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

**Declaration of Conflicting Interests**

The author declares no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

**Data Availability Statement**

The data that support the findings of this study are available from the corresponding author upon reasonable request, subject to privacy and ethical restrictions.

**References**

Adadi, A., & Berrada, M. (2018). Peeking inside the black-box: A survey on explainable artificial intelligence (XAI). *IEEE Access, 6*, 52138-52160.

Baker, R. S., & Inventado, P. S. (2014). Educational data mining and learning analytics. In *Learning analytics* (pp. 61-75). Springer, New York, NY.

Büyüközkan, G., Göçer, F., & Karabulut, Y. (2021). A new group decision making approach with fuzzy AHP for e-learning platform selection. *Journal of Intelligent & Fuzzy Systems, 40*(4), 7933-7952.

Chen, S. J., & Hwang, C. L. (1992). *Fuzzy multiple attribute decision making methods*. In Fuzzy multiple attribute decision making (pp. 289-486). Springer, Berlin, Heidelberg.

Garrison, D. R., & Kanuka, H. (2004). Blended learning: Uncovering its transformative potential in higher education. *The Internet and Higher Education, 7*(2), 95-105.

Global Market Insights. (2023). *E-learning Market Size*. Retrieved from https://www.gminsights.com/industry-analysis/elearning-market

Lundberg, S. M., & Lee, S. I. (2017). A unified approach to interpreting model predictions. *Advances in Neural Information Processing Systems, 30*.

Mayer, R. E. (2017). Using multimedia for e-learning. *Journal of Computer Assisted Learning, 33*(5), 403-423.

Parasuraman, A., Zeithaml, V. A., & Berry, L. L. (1988). SERVQUAL: A multiple-item scale for measuring consumer perceptions of service quality. *Journal of Retailing, 64*(1), 12-40.

Romero, C., & Ventura, S. (2020). Educational data mining and learning analytics: An updated survey. *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery, 10*(3), e1355.

Singh, V., & Thurman, A. (2019). How many ways can we define online learning? A systematic literature review of definitions of online learning (1988-2018). *American Journal of Distance Education, 33*(4), 289-306.

Tzeng, G. H., & Huang, J. J. (2011). *Multiple attribute decision making: methods and applications*. CRC Press.

Zadeh, L. A. (1965). Fuzzy sets. *Information and Control, 8*(3), 338-353.

Zeithaml, V. A., Parasuraman, A., & Malhotra, A. (2002). Service quality delivery through web sites: A critical review of extant knowledge. *Journal of the Academy of Marketing Science, 30*(4), 362-375.

DeLone, W. H., & McLean, E. R. (2003). The DeLone and McLean model of information systems success: A ten-year update. *Journal of Management Information Systems, 19*(4), 9-30.

Huang, R. H., Liu, D. J., Tlili, A., Yang, J. F., Wang, H. H., et al. (2020). *Handbook on facilitating flexible learning during educational disruption: The Chinese experience in maintaining undisrupted learning in COVID-19 outbreak*. Smart Learning Institute of Beijing Normal University.

Almaiah, M. A., Al-Khasawneh, A., & Althunibat, A. (2020). Exploring the critical challenges and factors influencing the E-learning system usage during COVID-19 pandemic. *Education and Information Technologies, 25*, 5261-5280.

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

# A Appendix A: Detailed Statistical Analyses

Table 6: Complete Normality Test Results (Shapiro-Wilk)

| Variable | W statistic | p-value | Normality |
|---|---|---|---|
| FCLT | 0.8961 | 0.0000 | Non-Normal |
| SEC | 0.8989 | 0.0000 | Non-Normal |
| UTLT | 0.8436 | 0.0000 | Non-Normal |
| CNT | 0.9557 | 0.0057 | Non-Normal |
| PED | 0.9347 | 0.0004 | Non-Normal |
| TANG | 0.9721 | 0.0312 | Non-Normal |
| FIAB | 0.9501 | 0.0021 | Non-Normal |
| SERVB | 0.9618 | 0.0098 | Non-Normal |
| ASS | 0.9743 | 0.0415 | Non-Normal |
| EMP | 0.9832 | 0.1556 | Normal |
| SAT | 0.9402 | 0.0006 | Non-Normal |

# B Appendix B: Cross-Validation Results

The Gradient Boosting model was evaluated using 5-fold cross-validation with the following detailed results:

Table 7: Detailed Cross-Validation Performance Metrics

| Fold | R | RMSE | MAE | Training Samples |
|---|---|---|---|---|
| 1 | 0.317 | 0.821 | 0.642 | 82 |
| 2 | 0.276 | 0.847 | 0.658 | 82 |
| 3 | 0.527 | 0.702 | 0.532 | 83 |
| 4 | 0.652 | 0.618 | 0.473 | 83 |
| 5 | 0.545 | 0.691 | 0.521 | 82 |
| **Mean** | 0.4636 | 0.736 | 0.565 | 82.4 |
| **Std** | 0.1432 | 0.091 | 0.077 | 0.55 |