

INFORMATION-EXTREME EVALUATION OF FEATURE INFORMATIVENESS FOR EDUCATIONAL CONTENT ADAPTATION TO LABOR MARKET REQUIREMENTS

Shelehov Igor

*Doctor of Philosophy (PhD) in Technical Sciences, Associate Professor, Sumy National Agricultural University, Ukraine
ORCID ID: 0000-0003-4304-7768*

The rapid development of digital technologies has fundamentally transformed the structure and functioning of modern educational systems. In recent years, particularly under conditions of global instability and accelerated digitalization, higher education institutions have faced the urgent need to continuously adapt their educational programs to dynamically changing labor market requirements. This challenge is especially relevant in the context of emerging economies and post-crisis recovery, where the mismatch between acquired competencies and market demands can significantly hinder economic development and workforce competitiveness.

One of the key problems in this context is the so-called skills gap [1], which reflects the discrepancy between the competencies provided by educational institutions and those required by employers. Traditional approaches to curriculum design are often based on expert judgment, static standards, and infrequent updates, which makes them insufficiently responsive to rapid technological changes and evolving professional requirements. As a result, educational content may become partially outdated, redundant, or misaligned with actual labor market needs.

The transition to digital learning environments, including learning management systems (LMS), online courses, and adaptive educational platforms, creates new opportunities for addressing this problem [2]. In particular, the availability of large volumes of educational and labor market data enables the application of formal methods for analyzing, evaluating, and optimizing educational content. However, despite these opportunities, there remains a lack of rigorous, mathematically grounded approaches for assessing the quality and relevance of educational content in terms of its contribution to learning outcomes and labor market alignment. A promising direction for solving this problem is the application of feature-based analysis, where educational content is described by a set of measurable characteristics (features), and its quality or relevance is evaluated through classification and optimization procedures. Within this framework, a critical task is the evaluation of feature informativeness, i.e., determining the extent to which individual characteristics contribute to the correct classification of educational content according to predefined quality criteria [3].

In this regard, information-extreme approaches [4], originally developed for pattern recognition and decision support systems, offer a powerful theoretical and methodological foundation. These approaches are based on the optimization of information criteria that reflect the functional efficiency of classification systems and allow for the identification of the most informative features while eliminating redundant or misleading ones.

The application of information-extreme evaluation methods to educational content adaptation represents a novel and promising line of research [5]. It enables not only the classification of educational materials by quality level but also the optimization of their structure by selecting the most relevant and informative characteristics. This, in turn, supports the development of adaptive educational programs that are better aligned with labor market requirements.

The aim of this chapter is to develop and investigate an information-extreme approach to evaluating the informativeness of educational content features and to apply this approach to the problem of adapting educational programs to labor market requirements.

To achieve this aim, the following objectives are addressed:

- to formalize educational content as an object of feature-based analysis;
- to define quality classes reflecting the degree of alignment with labor market needs;
- to develop a mathematical model for evaluating feature informativeness based on information-extreme criteria;
- to propose algorithms for feature selection and content optimization;
- to demonstrate the applicability of the proposed approach through a case study.

In order to formalize the problem of adapting educational content to labor market requirements, it is necessary to represent educational programs in a structured and analyzable form. Within the proposed approach, educational content is considered as a set of objects (courses, modules, or learning units), each of which is described by a vector of features reflecting its qualitative and quantitative characteristics.

Let us define a set of educational content objects:

$$\mathcal{Y} = \{Y_1, Y_2, \dots, Y_K\}, \quad (1)$$

where each object Y_k represents a specific course or module within an educational program. Each object is characterized by a set of features:

$$\Omega = \{\omega_1, \omega_2, \dots, \omega_N\}, \quad (2)$$

where each feature ω_i corresponds to a measurable characteristic of the content. These features may include, for example:

- relevance to current technologies and industry standards;
- presence of practical tasks and real-world case studies;
- alignment with job market skill requirements;
- level of complexity and depth;
- degree of interactivity and digital integration.

Thus, each content object Y_k is represented by a feature vector:

$$y_k = \{y_{k,1}, y_{k,2}, \dots, y_{k,N}\} \subseteq \mathbb{R}^N, \quad (3)$$

where $y_{k,i}$ denotes the value of feature ω_i for object Y_k .

To evaluate the quality of educational content, a set of classification classes is introduced:

$$\mathcal{C} = \{C_1, C_2, \dots, C_M\} \quad (4)$$

which represent different levels of alignment with labor market requirements. For example, these classes may correspond to:

- high-quality (fully aligned with labor market needs),
- medium-quality (partially aligned),
- low-quality (not aligned).

The assignment of content objects to these classes may be based on expert evaluation, empirical data (e.g., graduate employment rates), or a combination of both.

Within this framework, the problem can be formulated as a classification task: given a feature vector y_k , determine the class C_m to which the content object belongs. However, a central challenge in this task is that not all features contribute equally to classification accuracy. Some features may be highly informative, while others may be redundant or even negatively affect the classification process.

Therefore, a key problem is the evaluation of feature informativeness [6], which can be defined as the contribution of a given feature to the overall effectiveness of the classification system. Formally, this involves assessing how the inclusion or exclusion of a feature affects a chosen information criterion that reflects the functional efficiency of the system.

Based on this, the main research problems addressed in this chapter are:

1. Feature informativeness evaluation: to determine the contribution of each feature ω_i to the accuracy and reliability of content classification.
2. Feature space optimization: to reduce the dimensionality of the feature space by eliminating non-informative or misleading features while preserving or improving classification performance.
3. Educational content classification: to develop a model that assigns content objects to quality classes based on the most informative features.
4. Content adaptation and optimization: to use the results of feature evaluation and classification to improve the structure and content of educational programs in accordance with labor market requirements.

To solve these problems, this chapter proposes the use of an information-extreme approach, in which the effectiveness of the classification system is evaluated using an information criterion, and feature informativeness is determined based on its impact on this criterion. This approach allows for the systematic

identification of the most significant characteristics of educational content and provides a formal basis for its optimization.

The evaluation of feature informativeness is a fundamental problem in the design of intelligent systems for classification and decision support. In the context of educational content adaptation to labor market requirements, this problem becomes particularly significant, as it determines which characteristics of educational content have the greatest impact on its quality assessment and alignment with external demands.

Feature informativeness can be generally defined as the degree to which a given feature contributes to the discrimination between classes of objects – in this case, levels of educational content quality or relevance to labor market needs. The effectiveness of classification, and consequently the validity of decisions regarding curriculum adaptation, depends directly on the selection of informative features.

Existing approaches to feature informativeness evaluation can be broadly divided into two main categories [6]: filter methods and wrapper methods.

Filter methods evaluate feature informativeness independently of a specific classification algorithm. They rely on general statistical or information-theoretic properties of the data and are typically applied as a preprocessing step.

The main advantage of filter methods is their computational efficiency and universality. Since they do not depend on a particular classifier, they can be applied to a wide range of problems, including those where the final classification model is not yet defined. However, their primary limitation is the inability to directly account for the interaction between features and the classification model, which may reduce their effectiveness in complex tasks.

In the context of educational content analysis, filter methods can be used to preliminarily assess which characteristics – such as relevance to industry standards, presence of practical components, or alignment with job market skills – contain useful information for distinguishing between different quality levels.

Wrapper methods, in contrast, evaluate feature informativeness based on the performance of a specific classification model. In this case, the quality of a feature (or a set of features) is determined by how well the classifier performs when using that feature subset.

The key advantage of wrapper methods is their ability to account for the interaction between features and the classification algorithm. As a result, they often provide more accurate and task-specific evaluations of informativeness. However, this comes at the cost of significantly higher computational complexity, as the classification model must be trained and evaluated multiple times for different feature subsets.

For the problem of educational content adaptation, wrapper methods are particularly valuable because they allow for the direct assessment of how specific content characteristics influence the correctness of classification into quality or relevance classes. This makes them suitable for identifying not only informative features but also those that may negatively affect classification performance.

Feature informativeness can be quantified using different types of measures, which can be grouped into several categories depending on their underlying principles.

1. Distance-Based Measures

Distance-based measures evaluate informativeness by analyzing the separability of classes in the feature space. The basic idea is that informative features increase the distance between objects belonging to different classes while reducing the distance between objects of the same class.

Typical examples include metrics based on Euclidean distance or other similarity measures. These methods are characterized by low computational complexity and general applicability. However, they may not fully capture complex statistical dependencies between features.

In educational content analysis, such measures can be used to assess how well certain characteristics differentiate between high-quality and low-quality content.

2. Information-Theoretic Measures

Information-theoretic measures evaluate feature informativeness based on the concept of uncertainty reduction. A feature is considered informative if its inclusion reduces the uncertainty about the class membership of an object.

Common measures include entropy-based criteria, such as Shannon entropy, and divergence measures, such as Kullback–Leibler divergence. These measures provide a natural and theoretically grounded way to quantify informativeness and are widely used in feature selection tasks.

In the considered problem, information-theoretic measures allow for the evaluation of how much information about the quality or labor market alignment of educational content is gained by considering a particular feature.

3. Correlation-Based Measures

Correlation-based measures assess informativeness by analyzing dependencies between features and between features and class labels. The goal is to identify features that are strongly correlated with the target variable (class) but weakly correlated with each other.

These methods help to eliminate redundant features that do not provide additional information. While they are relatively simple and computationally efficient, they may not capture nonlinear relationships.

In educational applications, correlation-based approaches can be used to identify overlapping or redundant content characteristics, such as multiple indicators of similar competencies.

4. Consistency-Based Measures

Consistency-based measures evaluate the informativeness of features based on the consistency of the data representation. A feature subset is considered informative if it allows for consistent classification of objects with minimal contradictions.

These methods often rely on the concept of the minimal representative subset of features that preserves classification accuracy. They provide a balance between computational complexity and evaluation quality.

In the context of educational content, such measures can be used to determine the minimal set of characteristics required to reliably assess content quality.

5. Classifier-Based Measures

Classifier-based measures directly evaluate informativeness through the accuracy or error rate of a classification model. These measures are typically used within wrapper methods and provide the most direct assessment of feature usefulness.

They are characterized by high accuracy and relevance to the specific task but require significant computational resources. Importantly, they allow for the estimation of classification performance already at the stage of feature selection.

For educational content adaptation, classifier-based measures are particularly important, as they enable the identification of features that directly contribute to correct classification into quality or relevance categories.

The analysis of existing methods shows that there is a trade-off between computational efficiency and evaluation accuracy. Filter methods provide fast and general estimates of feature informativeness but may lack precision in specific tasks. Wrapper methods, on the other hand, offer higher accuracy but require more computational resources.

Given the complexity of educational content as an object of analysis and the importance of accurate classification for decision-making, it is reasonable to prioritize methods that take into account classification performance. In this context, approaches based on classifier performance and information criteria are particularly relevant.

The information-extreme approach combines the advantages of information-theoretic and classifier-based methods. It evaluates feature informativeness through its impact on an information criterion that reflects the functional efficiency of the classification system. This allows not only for the identification of informative features but also for the detection of redundant or misleading ones.

Therefore, the information-extreme framework provides a suitable methodological basis for evaluating feature informativeness in the problem of educational content adaptation to labor market requirements.

The evaluation of feature informativeness is closely related to the problem of feature space optimization, which aims to construct a subset of features that ensures the highest possible classification performance while minimizing redundancy and dimensionality. In the context of educational content adaptation, this problem corresponds to identifying the most relevant characteristics of learning materials that determine their quality and alignment with labor market requirements.

The initial feature space, formed during the problem formulation stage, typically includes a wide range of characteristics describing educational content. These may encompass pedagogical, technological, and labor market-related attributes. However, not all features contribute equally to classification accuracy. Some may be weakly informative, redundant, or even negatively affect the performance of the classification system.

Therefore, the task of feature space optimization can be defined as the selection of an optimal subset of features $\Omega^* \subseteq \Omega$, such that the effectiveness of the classification system, evaluated by a chosen criterion, is maximized.

The optimization of the feature space pursues several interrelated objectives [7]:

1. Reduction of dimensionality. Eliminating non-informative features reduces computational complexity and improves the efficiency of data processing.

2. Elimination of redundancy. Features that duplicate information provided by others do not contribute to improved classification and may introduce noise.

3. Improvement of classification accuracy. Removing irrelevant or misleading features can increase the separability of classes and enhance decision reliability.

4. Interpretability of results. A reduced set of meaningful features facilitates the interpretation of results and supports decision-making in educational management.

In the considered problem, these objectives are directly linked to improving the quality of decisions regarding curriculum design and adaptation.

Feature space optimization methods are commonly classified into three main groups: filter, wrapper, and embedded methods.

In filter-based approaches[7], features are ranked according to a selected informativeness criterion, and only the top-ranked features are retained. This method is computationally efficient and suitable for preliminary dimensionality reduction.

However, since filter methods do not account for interactions between features, they may retain redundant features or discard combinations that are informative only when considered jointly.

Wrapper methods [7] evaluate subsets of features by training and testing a classification model. The optimal subset is selected based on the classification performance.

These methods provide higher accuracy but require significantly more computational resources, especially when the number of features is large. The search for the optimal subset is often performed using heuristic strategies, such as forward selection, backward elimination, stepwise selection.

Embedded methods [7] integrate feature selection directly into the model training process. Examples include regularization techniques and tree-based models that inherently evaluate feature importance.

While these methods are efficient and effective in many applications, they are often tied to specific model architectures and may lack generality.

Given the combinatorial complexity of the feature selection problem, exhaustive search is typically infeasible. Therefore, sequential selection strategies are widely used [8].

Forward Selection: the process begins with an empty feature set and iteratively adds features that provide the greatest improvement in the evaluation criterion.

Advantages:

- simplicity,
- efficiency for large feature spaces.

Limitations:

- may miss optimal combinations due to greedy nature.

Backward Elimination: the process starts with the full feature set and iteratively removes features that contribute the least to the evaluation criterion.

Advantages:

- considers full feature interactions initially.

Limitations:

- computationally expensive,
- sensitive to initial feature set quality.

Hybrid approaches combine forward and backward procedures to balance exploration and computational cost.

Within the information-extreme approach, feature space optimization is performed based on the maximization of an information criterion that reflects the functional efficiency of the classification system.

In this framework, the informativeness of a feature is evaluated by analyzing how its inclusion or exclusion affects the value of the information criterion. This allows for a principled selection of features that contribute positively to class separability.

A key advantage of this approach is the ability to identify not only non-informative features, but also negatively informative features, i.e., those that degrade classification performance.

This property is particularly important in the analysis of educational content, where certain characteristics may introduce bias or distort the assessment of quality.

In the context of adapting educational programs to labor market requirements, feature space optimization has a clear practical interpretation.

The selected optimal feature subset Ω^* represents the most significant characteristics of educational content, which determine its relevance to current industry needs, influence learning outcomes, or contribute to successful employment of graduates.

Conversely, features excluded from the optimal subset may correspond to outdated content elements, redundant topics, ineffective pedagogical approaches.

Thus, feature space optimization provides a formal mechanism for revising curriculum structure, prioritizing key competencies, eliminating unnecessary or low-impact content.

The analysis of feature space optimization methods indicates that no single approach is universally optimal. Filter methods are efficient but may lack precision, while wrapper methods are accurate but computationally expensive.

The information-extreme approach offers a balanced solution by integrating feature evaluation with classification performance, using a unified information criterion, and enabling the detection of both redundant and misleading features.

Therefore, it is well-suited for solving the problem of educational content adaptation, where both accuracy and interpretability are critical.

The information-extreme approach represents a class of methods for modeling intelligent decision-making systems based on the optimization of information criteria that reflect the functional efficiency of classification processes. Originally developed for pattern recognition and diagnostic systems, this approach is characterized by its ability to operate under conditions of uncertainty, limited data, and overlapping classes.

In the context of educational content adaptation to labor market requirements, the information-extreme approach provides a formal framework for evaluating the informativeness of content characteristics, constructing optimal feature spaces, ensuring reliable classification of educational content according to quality and relevance criteria.

The core idea of the information-extreme approach is to maximize the informational separability of classes in the feature space. This is achieved by optimizing a criterion that reflects the balance between:

- inter-class distance (distance between different classes),
- intra-class compactness (similarity within the same class).

In contrast to traditional statistical methods, the information-extreme approach does not rely on strict assumptions about data distribution. Instead, it focuses on the structural organization of data in the feature space and the ability of the system to correctly classify objects under varying conditions.

In the considered problem, educational content objects [9] are represented as feature vectors, and the goal is to construct a classification system that maximizes the distinguishability between different quality levels or degrees of alignment with labor market requirements.

Let the set of educational content objects be represented as (1), where each object is described by a feature vector (3). Each object belongs to one of the predefined classes (4), which correspond to different levels of educational content quality or alignment with labor market requirements.

Within the information-extreme framework, each class is associated with a reference (prototype) vector, which serves as a representative of the class in the feature space. The classification process is based on comparing an object with these reference vectors using a selected similarity or distance measure.

A key element of the approach is the definition of decision rules, which determine class membership based on the position of an object in the feature space relative to class boundaries.

A central component of the information-extreme approach is the information criterion, which serves as a measure of the functional efficiency of the classification system.

This criterion reflects the system's ability to correctly classify objects belonging to different classes, maintain stability under variations in feature values, minimize classification errors.

In general, the information criterion is defined as a function of classification accuracy and reliability. It takes into account both correct and incorrect decisions, as well as the distribution of objects in the feature space.

The optimization process aims to find such parameters of the classification system (including feature subsets and decision thresholds) that maximize the value of the information criterion.

The learning process in the information-extreme approach involves the iterative adjustment of system parameters in order to maximize the information criterion.

This process includes:

1. Formation of the training sample. A set of labeled educational content objects is used to construct the model.

2. Initialization of parameters. Initial values of feature thresholds, class boundaries, and reference vectors are defined.
3. Iterative optimization. System parameters are adjusted step by step to improve the value of the information criterion.
4. Evaluation of feature informativeness. The contribution of each feature is assessed based on its impact on the information criterion.
5. Selection of optimal parameters. The final model is chosen as the one that achieves the maximum criterion value.

An important property of this learning process is its ability to identify informative features, redundant features, negatively informative features.

To apply the information-extreme approach to the problem of educational content adaptation, it is necessary to reinterpret its key components (Table 1)

Table 1. Adaptation of classical interpretation of information-extreme key components

Classical interpretation	Adapted interpretation
Object	Educational content unit
Feature	Content characteristic
Class	Quality / labor market alignment level
Decision rule	Content classification rule

Within this framework, the learning process corresponds to the analysis of educational content and the identification of those characteristics that most strongly influence its classification.

The optimization of the feature space leads to the identification of a subset of characteristics that define the effective structure of educational content in terms of labor market relevance.

The application of the information-extreme approach to educational content analysis provides several important advantages:

- High sensitivity to feature informativeness. The method allows for precise identification of features that influence classification performance.
- Detection of negatively informative features. Unlike many traditional methods, it identifies features that degrade system performance.
- Interpretability. The results can be directly interpreted in terms of educational content characteristics.
- Adaptability. The approach can be applied to different domains and updated as new data becomes available.
- Suitability for decision support. The method provides a formal basis for optimizing educational programs and aligning them with labor market needs.

Despite its advantages, the information-extreme approach has certain limitations:

- it requires a representative training dataset;
- the quality of results depends on the correctness of feature definition;
- computational complexity may increase with the number of features;
- periodic model updating is necessary due to changes in labor market requirements.

However, these limitations are typical for most data-driven approaches and can be mitigated through proper system design and data management.

The information-extreme approach provides a robust theoretical and methodological foundation for evaluating feature informativeness and optimizing educational content. Its ability to integrate feature selection with classification performance makes it particularly suitable for solving the problem of adapting educational programs to labor market requirements.

Within the proposed approach, educational content is represented as a structured set of objects described in a multidimensional feature space. Each object corresponds to a unit of educational content (course, module, or topic), characterized by a vector of features reflecting its qualitative and quantitative properties.

Let the set of educational content objects be defined as (1). Each object Y_k is represented by a feature vector (3), where $y_{k,i}$ denotes the value of the i -th feature $\omega_i \in \Omega$, and (2) is the feature space. The feature values may be binary, discrete, or continuous, depending on the nature of the educational content

characteristics (e.g., presence of practical tasks, level of alignment with labor market skills, degree of interactivity).

To evaluate the relevance of educational content to labor market requirements, a set of classes is introduced (4), where each class represents a level of quality or alignment, for example:

- C_1 : high alignment,
- C_2 : partial alignment,
- C_3 : low alignment.

Each object Y_k in the training set is associated with a class label $\lambda_k \in \{1, 2, \dots, m\}$. The labeled dataset forms the basis for training the classification model and evaluating feature informativeness.

In contrast to conventional approaches where class prototypes are defined as mean vectors in a continuous feature space, the information-extreme approach employs a binary representation of features obtained through a system of control tolerances. This transformation plays a key role in ensuring robustness and improving class separability.

Let the original feature vector of an educational content object be defined as (3). Each feature $y_{k,i}$ is transformed into a binary value using a system of control tolerances:

$$\delta_i = [y_i^{low}, y_i^{up}],$$

which defines the admissible range of values for feature ω_i .

The binarized feature $x_{k,i}$ is defined as:

$$x_{k,i} = \begin{cases} 1, & \text{if } y_{k,i} \in \delta_i, \\ 0, & \text{if } y_{k,i} \notin \delta_i. \end{cases} \quad (5)$$

Thus, each educational content object is represented by a binary vector:

$$x_k = \{x_{k,1}, x_{k,2}, \dots, x_{k,N}\} \subseteq \{0, 1\}^N. \quad (6)$$

This transformation allows the system to operate in a discrete feature space, which simplifies the construction of decision rules and enhances robustness to noise and variability in the data. Within the binary feature space, each class C_m is represented by a reference (prototype) vector:

$$x_m^* = \{x_{m,1}^*, x_{m,2}^*, \dots, x_{m,N}^*\} \subseteq \{0, 1\}^N \quad (7)$$

The reference vector is constructed based on the binarized training samples belonging to class C_m . In the simplest case, it can be defined as the result of aggregation (e.g., majority rule or mean with thresholding):

$$x_{m,i}^* = \begin{cases} 1, & \text{if } \frac{1}{n} \sum_{\lambda_k=m} x_{k,i} \geq \rho, \\ 0, & \text{otherwise,} \end{cases}$$

where ρ is a threshold parameter, n is a number of objects in training set with label $\lambda_k = m$.

Thus, the reference vector represents the typical binary pattern of features associated with a given class of educational content.

In the binary feature space, each class can be interpreted as a region defined around its reference vector. The distance between an object and a class is typically evaluated using a discrete metric, such as the Hamming distance.

This representation enables clear separation of classes, construction of decision regions, and efficient computation of classification rules.

The system of control tolerances δ_i is not fixed but is subject to optimization within the learning process. By adjusting the tolerance ranges, the model effectively reshapes the binary feature space and influences class separability.

This leads to a two-level optimization problem:

1. optimization of tolerance parameters δ_i ;
2. optimization of the feature subset Ω^* .

The joint optimization of these parameters allows the system to achieve maximum classification efficiency according to the information criterion.

In the context of educational content adaptation, the binarization process has a clear interpretation:

- $x_{k,i} = 1$ indicates that the content characteristic satisfies the required standard or falls within an acceptable range;

- $x_{k,i} = 0$ indicates a deviation from desired criteria.

Thus, the reference vector defines a target profile of educational content, corresponding to a specific level of alignment with labor market requirements.

The optimization of tolerances allows the system to adapt to varying requirements, account for acceptable deviations, and ensure flexibility in content evaluation.

Within the information-extreme framework, the classification of educational content objects is performed in a binary feature space using distance-based decision rules. After binarization of features through the system of control tolerances, each object is represented as a binary vector (6). Each class C_m is associated with a binary reference vector (7).

The similarity between an object and a class is evaluated using the Hamming distance, defined as:

$$d_H(x_k, x_j) = \sum_{i=1}^N |x_{k,i} - x_{j,i}|.$$

This distance measures the number of mismatched components between the object and the class prototype.

In the information-extreme approach, each class C_m is represented not only by a reference vector but also by a container (a hypersphere in binary space) K_m , defined by a center x_m^* and a radius d_m^* . Thus, the container K_m of class C_m is defined as:

$$K_m = \{x_j \subseteq \{0,1\}^N \mid d_H(x_m^*, x_j) \leq d_m^*\}.$$

An object belongs to class C_m if its binary representation lies within the corresponding container. The classification decision for an object X_k is made according to the following rule:

$$X_k \in C_m \text{ if } \mu_m = 1 - \frac{d_H(x_m^*, x_j)}{d_m^*} \geq 0.$$

If multiple containers include the object, additional selection criteria may be applied, such as choosing the class with the maximum μ_m :

$$\lambda_j = \arg \max_m \mu_m.$$

If the object does not belong to any container $\forall m \mu_m < 0$, it may be classified as unknown, or assigned to the nearest class based on minimum distance $d_H(x_m^*, x_j)$.

The radii d_m^* are not fixed and are determined during the learning process. Their values are selected in such a way as to maximize the information criterion.

The choice of d_m^* affects classification accuracy, class separability and robustness to noise.

Thus, the optimal radii are obtained as a result of a trade-off, formalized through the maximization of the information criterion.

A key component of the model is the information criterion, which evaluates the effectiveness of the classification system.

Let:

- P_t be the probability of correct classification,
- P_f be the probability of error.

The information criterion E is defined as a function:

$$E = f(P_t, P_f) = 0.5 \log_2 \left(\frac{P_t}{P_f} \right) (P_t - P_f),$$

which increases with higher classification accuracy and decreases with higher error rates.

In the information-extreme framework, the goal is to $E \rightarrow \max$ by optimizing system parameters, including radii d_m^* , system of control tolerances δ_i^* , feature subsets Ω^* :

$$d_m^* = \arg \max_{d_m \in \{0,1,\dots,N\}} E(d_m), \quad (8)$$

$$\delta_i^* = \arg \max_{\delta_i \in [0,\delta_{max}]} \frac{1}{M} \sum_{m=1}^M \max_{d_m \in \{0,1,\dots,N\}} E(d_m), \quad (9)$$

$$\Omega^* = \arg \max_{\Omega_p \subseteq \Omega} \left[\max_{\delta_i \in [0,\delta_{max}]} \frac{1}{M} \sum_{m=1}^M \max_{d_m \in \{0,1,\dots,N\}} E(d_m) \right]. \quad (10)$$

The criterion reflects the ability of the system to distinguish between classes of educational content based on their characteristics.

In the context of educational content adaptation:

- the reference vector represents an ideal profile of content corresponding to a specific quality level;
- the container radius defines the permissible level of deviation from this profile;
- the Hamming distance measures the degree of mismatch between actual content characteristics and the desired profile.

Thus objects inside the container correspond to content that meets the required quality level, and objects outside indicate insufficient alignment with labor market requirements. This interpretation allows decision-makers to identify content that meets standards, detect deviations, and prioritize improvements. The use of Hamming distance and class containers provides computational efficiency, robustness to noise, clear geometric interpretation, compatibility with binary feature representation. These properties make the approach particularly suitable for applications involving structured evaluation criteria, such as educational content quality assessment.

The informativeness of a feature ω_i is defined as its contribution to the value of the information criterion. Formally, the informativeness can be evaluated as:

$$\Delta E_i = E_\Omega - E_{\Omega/\omega_i}, \quad (11)$$

where E_Ω is the criterion value using the full feature set, E_{Ω/ω_i} is the criterion value after removing feature ω_i .

This definition allows for the systematic identification of features that improve or degrade classification performance.

The optimal feature subset is defined as (10), subject to constraints on dimensionality or computational cost. Due to the combinatorial nature of the problem, sequential optimization strategies are typically used, such as forward selection and backward elimination. At each step, the inclusion or exclusion of features is evaluated based on its effect on the information criterion.

In the context of educational content, the proposed model enables:

- identification of key content characteristics that determine alignment with labor market requirements;
- elimination of redundant or outdated features;
- detection of misleading characteristics that negatively affect quality assessment;
- construction of an optimized representation of educational programs.

Thus, the model provides a formal basis for adapting educational content by focusing on the most informative attributes that influence its practical relevance.

The proposed model has the following properties scalability: applicable to large sets of educational content, interpretability: results are expressed in terms of meaningful features, adaptability: can be updated as labor market requirements evolve, robustness: does not rely on strict assumptions about data distribution

The proposed algorithm implements the information-extreme approach to evaluating feature informativeness and optimizing educational content representation. It integrates three interconnected processes:

- binarization of features using control tolerances,
- construction of class containers in binary space,
- maximization of the information criterion.

The algorithm operates on a labeled dataset of educational content objects and produces an optimal feature subset, tolerance system, and classification model.

Input Data and Parameters

Input:

- set of educational content objects $\{Y_1, Y_2, \dots, Y_K\}$,
- feature vectors $y_k = \{y_{k,1}, y_{k,2}, \dots, y_{k,N}\}$
- class labels $\{\lambda_1, \lambda_2, \dots, \lambda_K\}$
- initial tolerance ranges δ_i

Output:

- optimal container $K_m(x_m^*, d_m^*)$ of class C_m ,
- optimal system of control tolerances δ_i^* ,
- optimal feature subsets Ω^* ,
- evaluation of feature informativeness ΔE_i .

Algorithm Structure

The algorithm consists of iterative optimization cycles, each aimed at improving the value of the information criterion.

Step 1. Initialization

- 1.1. Define the initial feature space Ω .
- 1.2. Initialize tolerance system $\delta = \{\delta_i\}$.

Step 2. Binarization of Feature Space

For each object Y_k and feature ω_i construct binary dataset X_k based on (5).

Step 3. Formation of Class Prototypes

For each class C_m

- 3.1. Aggregate binary vectors of training samples.
- 3.2. Construct reference vector x_m^* using thresholding.

Step 4. Optimization of Container Radii

For each class C_m :

- 4.1. Vary radius d_m within range $\{0, 1, \dots, N\}$.
- 4.2. For each value of d_m , compute number of correctly classified objects, number of misclassified objects and estimate P_t be the probability of correct classification, P_f be the probability of error.

4.3. Evaluate information criterion $E = f(P_t, P_f)$.

4.4. Select optimal radius d_m^* based on (8).

Step 5. Optimization of Control Tolerances

For each feature ω_i :

- 5.1. Adjust tolerance parameters δ_i
- 5.2. Recompute binary representation.
- 5.3. Re-evaluate information criterion E_{δ_i} .
- 5.4. Select optimal tolerance δ_i^* based on (9).

Step 6. Evaluation of Feature Informativeness

For each feature ω_i :

- 6.1. Compute criterion using full feature set E_Ω .
- 6.2. Remove feature ω_i and recompute E_{Ω/ω_i} .
- 6.3. Compute informativeness ΔE_i based on (11).
- 6.4. Classify feature:
 - $\Delta E_i > 0$: the feature is informative,
 - $\Delta E_i = 0$: the feature is non-informative,
 - $\Delta E_i < 0$: the feature is negatively informative.

Step 7. Feature Space Optimization

- 7.1. Remove negatively informative features.
- 7.2. Optionally remove weakly informative features.

7.3. Update feature space $\Omega \rightarrow \Omega^*$ based on (10).

Step 8. Iterative Optimization Loop

Repeat Steps 2–7 until:

- convergence of information criterion,
- or reaching stopping condition.

Step 9. Final Model Construction

9.1. Fix optimal parameters $(\Omega^*, \delta_i^*, x_m^*, d_m^*)$.

9.2. Construct final classification model.

To evaluate the effectiveness of the proposed information-extreme approach for educational content adaptation, an empirical dataset was constructed based on the assessment of educational materials developed by the Department of Computer Science at Sumy State University.

The dataset was formed using survey-based evaluations provided by three groups of stakeholders:

- employers,
- graduates,
- current students.

These groups were selected to ensure a comprehensive and balanced assessment of educational content from both academic and labor market perspectives.

Each respondent evaluated individual educational modules using a quantitative scale ranging from 0 to 100 points. The collected data reflect subjective yet structured judgments regarding the quality and relevance of educational content.

All responses were aggregated and stored in a centralized database, forming the basis for the training and evaluation of the classification system.

To provide a more detailed and realistic assessment of educational content quality, the classification scheme was extended from three to four classes. This allows for finer differentiation between levels of alignment with labor market requirements and educational standards.

The set of classes is defined as $\mathcal{C} = \{C_1, C_4\}$, where C_1 : satisfactory content (acceptable level) – partially aligned, requiring improvement; C_4 : unsatisfactory content (low level) – poorly aligned or outdated, requiring substantial revision.

The assignment of objects to classes was performed based on thresholding of aggregated scores. This labeling serves as ground truth for training the information-extreme classification model.

To represent educational content in a structured and analyzable form, a feature dictionary consisting of 28 characteristics was defined. Each feature reflects a specific aspect of content quality, relevance, and effectiveness in the context of alignment with labor market requirements.

The feature space is defined as $\Omega = \{\omega_1, \omega_2, \dots, \omega_{28}\}$. Each feature ω_i corresponds to a measurable attribute evaluated on a quantitative scale and subsequently transformed within the information-extreme framework. For clarity and interpretability, the feature dictionary is organized into five thematic groups: content quality, methodological quality, technical quality, assessment and feedback, and pedagogical value.

Table 2. Description of feature dictionary

No.	Indicator	Group
1	Compliance of content with the educational program	Content Quality
2	Alignment of content with learning outcomes	Content Quality
3	Relevance of learning information	Content Quality
4	Appropriateness to learners' educational level	Content Quality
5	Consideration of labor market needs	Content Quality
6	Compliance with general educational requirements	Technical Quality
7	Completeness and coverage of topics	Content Quality
8	Interdisciplinary connections of the material	Content Quality
9	Clear structuring of learning content	Methodological Quality
10	Correctness of references and bibliography	Content Quality
11	Compliance with educational standards	Content Quality
12	Clarity of instructional wording	Methodological Quality
13	Diversity of content delivery methods	Methodological Quality
14	Use of real-life examples and case studies	Methodological Quality

15	Quality and relevance of learning tasks	Methodological Quality
16	Accessibility of presentation for different learner groups	Methodological Quality
17	Use of active learning methods	Methodological Quality
18	Integrated evaluation of methodological design of materials	Methodological Quality
19	Stimulation of critical thinking	Methodological Quality
20	Development of learners' creativity	Methodological Quality
21	Quality of multimedia support	Technical Quality
22	Compatibility of materials across platforms	Technical Quality
23	Motivational value of content	Pedagogical Value
24	Inclusiveness of learning materials	Pedagogical Value
25	Systematic nature of feedback	Assessment & Feedback
26	Possibility of individualization of learning	Pedagogical Value
27	Flexibility of content updating	Pedagogical Value
28	Consistency of assessment checkpoints	Assessment & Feedback

At the initial (zero) iteration, the optimization of the feature dictionary was performed using the sequential backward selection algorithm. At the first step, the informativeness of each feature from the initial dictionary was evaluated by generating reduced variants of the feature set, each having one feature less than the current dictionary.

For each such variant, the training parameters were optimized, and the maximum averaged value of the information criterion was calculated. The obtained values for the feature set variants at the first step of training are presented in Fig. 1

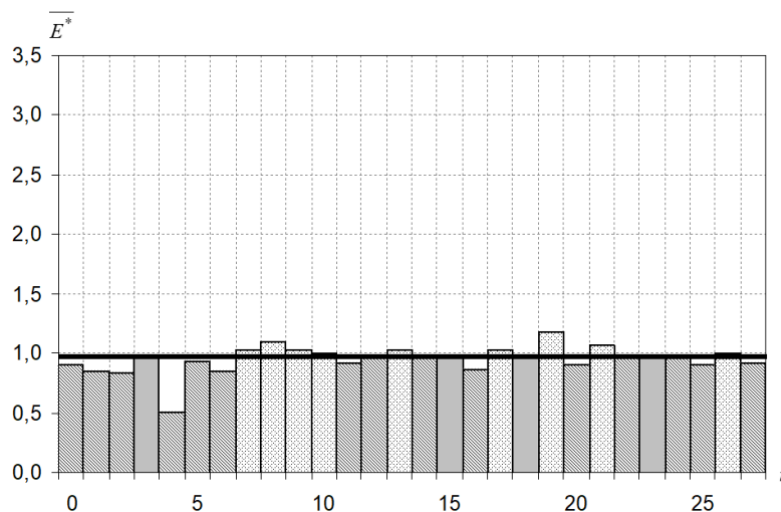


Figure 1. Maximum averaged value of the information criterion for feature set variants at the first step of optimization

In Fig. 2, the gray color indicates feature sets whose criterion values coincide with the averaged maximum value for the current feature set. Single-direction shading denotes variants for which the criterion value decreased relative to the current set, while double-direction shading indicates variants for which the criterion value increased.

As a result, the feature set selected for further iterations was obtained by removing 19 features from the current dictionary. The dynamics of the change in the information criterion are shown in Fig. 2.

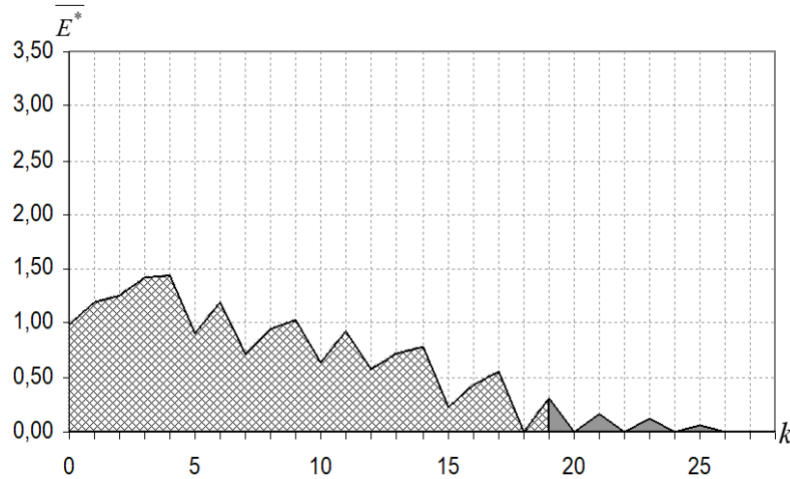


Figure 2. Dynamics of the maximum averaged information criterion value

The analysis of Fig. 2 indicates that the optimal feature set was obtained at the fourth iteration step. After this point, the value of the information criterion decreased significantly, although the process was non-monotonic. This behavior suggests that feature space optimization should continue even when intermediate feature sets demonstrate lower criterion values, as subsequent iterations may lead to new maxima.

The optimal feature set was formed by removing features 18, 19, 23, and 24 from the initial dictionary. The optimal value of the tolerance parameter was $\delta^* = 34$. The optimal geometric parameters of the classification model were determined as follows: for the base class $d_1^* = 8$, for the neighboring class $d_2^* = 9$, and the inter-center distance $d_c^* = 10$.

The maximum values of the information criterion and accuracy characteristics were obtained as: for the base class C_1 : $E_1^* = 2.13037$ ($P_t^* = 0.90$, $P_f^* = 0.10$).

The evaluation of feature informativeness using the modified algorithm was carried out similarly, by generating reduced feature sets and optimizing training parameters, including the feature subset itself. For each variant, the maximum averaged value of the information criterion was computed. The dynamics of the information criterion during this process are presented in Fig. 3.

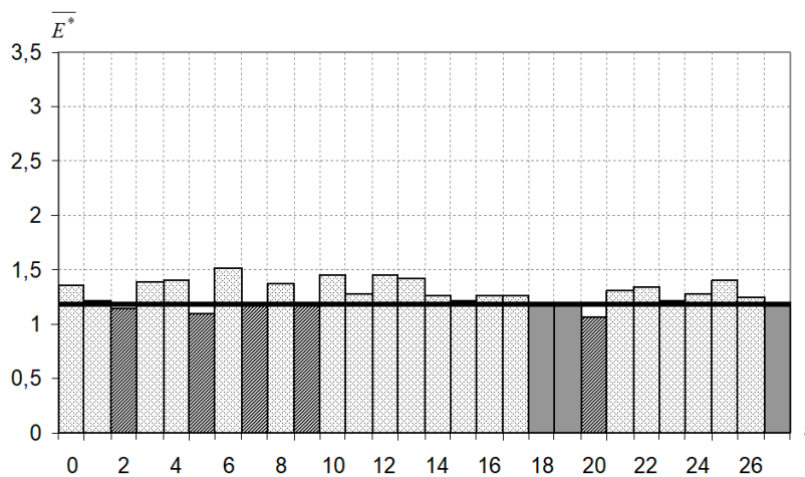


Figure 3. Dynamics of the maximum averaged information criterion during feature informativeness evaluation using the modified algorithm

The analysis of Fig. 3 shows that only five features are informative for the classification system as a whole. Three features are non-informative, as their removal does not affect the value of the information criterion. All remaining features are considered interfering, indicating that the corresponding test items require revision.

The most informative feature is feature 5, with an informativeness value of 0.09767. Its removal leads to changes in the optimal geometric parameters: $d_1^*=7$, $d_2^*=7$, $d_c^*=8$. The performance indicators remain unchanged for the base class but degrade for the neighboring class.

The most negatively informative (interfering) feature is feature 6, with an informativeness value of -0.32173 . Its removal results in improved performance indicators for the base class: $E_1^*=2.98974$ ($P_t^* = 0.95$, $P_f^* = 0.05$).

The evaluation of feature informativeness using the modified algorithm was carried out similarly, by generating reduced feature sets and optimizing training parameters, including the feature subset. For each variant, the maximum averaged value of the information criterion was computed. The dynamics of the information criterion during this process are presented in Fig. 4.

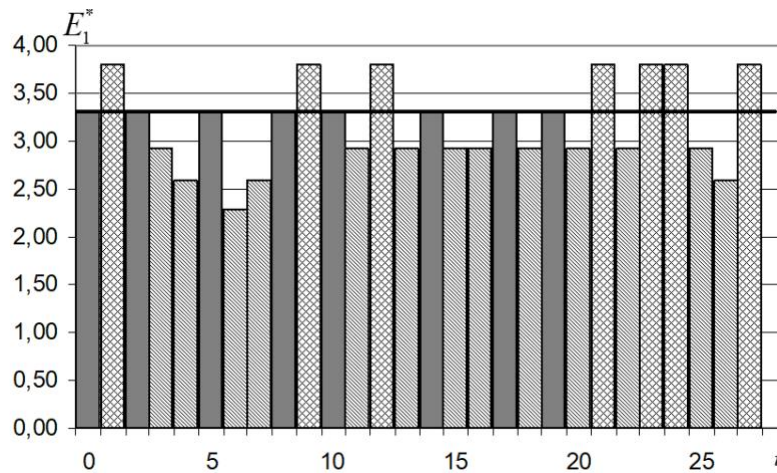


Figure 4. Dynamics of the maximum averaged information criterion during feature informativeness evaluation using the modified algorithm

The analysis shows that the number of non-informative features increased to 7, while the number of informative features decreased to 13, and the number of interfering features increased to 6. Given that error-free classification was achieved for the first and second classes, but not for the base and third classes, the initial feature set for these classes can be considered redundant. Further analysis confirmed that all features for these classes are non-informative. The results of the extended evaluation for the corresponding class are presented in Fig. 5.

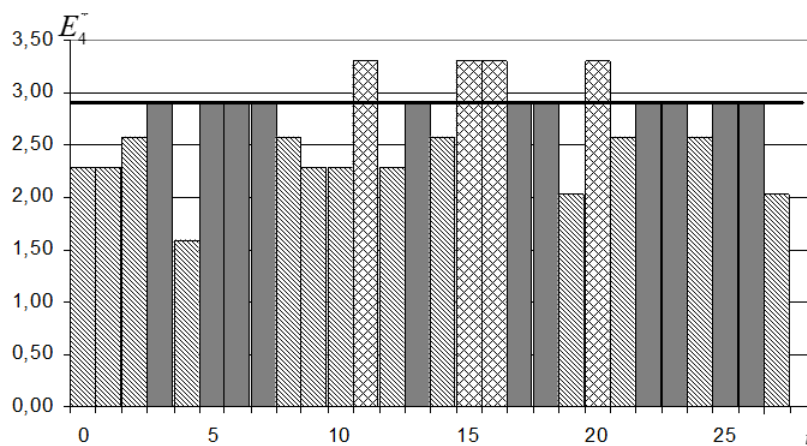


Figure 5. Results of extended feature informativeness evaluation for the corresponding class

The analysis indicates that the number of non-informative features is 11, informative features – 13, and interfering features – 4.

A summary of the feature informativeness evaluation results is presented in Table 3.

Table 3. Summary results of feature informativeness evaluation

Feature				Feature	Informativeness		
	C_1	C_4	Avg		C_1	C_4	Avg
1	0,00	0,63	0,31	15	0,00	0,34	0,17
2	-0,48	0,63	0,08	16	0,39	-0,39	0,00
3	0,00	0,34	0,17	17	0,39	-0,39	0,00
4	0,39	0,00	0,20	18	0,00	0,00	0,00
5	0,73	1,32	1,03	19	0,39	0,00	0,20
6	0,00	0,00	0,00	20	0,00	0,89	0,44
7	1,02	0,00	0,51	21	0,39	-0,39	0,00
8	0,73	0,00	0,36	22	-0,48	0,34	-0,07
9	0,00	0,34	0,17	23	0,39	0,00	0,20
10	-0,48	0,63	0,08	24	-0,48	0,00	-0,24
11	0,00	0,63	0,31	25	-0,48	0,34	-0,07
12	0,39	-0,39	0,00	26	0,39	0,00	0,20
13	-0,48	0,63	0,08	27	0,73	0,00	0,36
14	0,39	0,00	0,20	28	-0,48	0,89	0,20

The results of feature informativeness evaluation reveal a complex and non-trivial structure of the feature space, in which not all characteristics contribute equally to the classification of educational content quality. The elimination of certain features, in particular Feature 6 (compliance with general educational requirements) and Feature 18 (integrated evaluation of methodological design), can be explained by their low discriminative ability. These indicators tend to exhibit minimal variation across quality classes, since most educational materials formally satisfy general institutional and methodological requirements. As a result, they do not provide meaningful differentiation between high- and low-quality content. In addition, such features are typically standardized at the regulatory or accreditation level, which further reduces their sensitivity to actual differences in educational effectiveness. Consequently, their contribution to the information criterion is negligible, leading to their exclusion from the optimal feature subset.

A number of features, including Features 2, 10, 12, 13, 16, 17, 21, 22, 25, and 28, demonstrated mixed informativeness, being useful for certain classes while negatively affecting classification performance for others. This behavior can be attributed to several factors. First, many of these indicators are inherently subjective, as they depend on the perceptions of different stakeholder groups such as students, graduates, and employers. For example, clarity of instructional wording, diversity of content delivery methods, and accessibility of presentation may be interpreted differently depending on individual experience and expectations. Second, some features are context-dependent, meaning that their importance varies depending on the format or conditions of learning. Indicators such as multimedia quality or platform compatibility may be critical in online environments but less relevant in traditional settings, which reduces their overall informativeness. Third, certain features suffer from weak formalization, particularly those related to creativity, feedback, or learner engagement, which are difficult to quantify objectively and therefore introduce noise into the feature space. Finally, redundancy and correlation between features also play a significant role. For instance, alignment with learning outcomes overlaps conceptually with compliance with the educational program, while consistency of assessment checkpoints is partially reflected in feedback-related characteristics. Such overlaps reduce the marginal contribution of individual features to the overall classification performance.

At the same time, the remaining subset of features forms a stable and informative core that ensures effective separation of quality classes. These features are characterized by a direct connection to learning outcomes, a higher degree of measurability, and reduced subjectivity. They also exhibit significant variability across classes, thereby enhancing their contribution to the information criterion. Importantly, many of these characteristics are closely related to the practical applicability of knowledge and skills, which is a key factor in aligning educational materials with labor market demands.

From the perspective of content adaptation, the obtained results indicate that the most informative characteristics are those associated with real-world applicability, completeness of knowledge representation, and alignment with professional competencies. This suggests that instructional design should prioritize

outcome-oriented principles rather than formal compliance with standards. At the same time, the relatively low informativeness of regulated indicators implies that strict adherence to general requirements does not necessarily improve employability-oriented quality. The findings also highlight the need for improved formalization of complex educational characteristics such as creativity, critical thinking, and feedback mechanisms, as their current representation limits their contribution to the classification process [10]. Furthermore, the identification of interfering features provides diagnostic insights into weaknesses in course design. Such characteristics reflect inconsistencies in structure, ambiguity in evaluation criteria, or instability in the learning experience and can serve as indicators for targeted improvement. From a strategic perspective, the results demonstrate that not all traditionally emphasized characteristics are equally important for labor market alignment. Priority should be given to those that directly support the development of practical competencies. In this context, feature selection becomes a powerful tool for curriculum optimization, enabling a data-driven approach to refining educational programs[11].

The proposed information-extreme approach can be effectively implemented within the operational processes of higher education institutions, particularly at the level of graduating departments responsible for curriculum design and continuous improvement. Its practical application is based on transforming stakeholder evaluations into a formalized feature space, followed by automated classification into predefined quality categories. Each educational unit, such as a course or module, is assessed using the defined feature set, forming a quantitative representation processed within the model. The system performs binarization using optimized tolerance parameters, constructs feature vectors, and assigns the content to a corresponding class based on decision rules in the Hamming space. One of the key application scenarios is the periodic assessment of educational programs. In this case, evaluation data are collected from students, graduates, and employers, aggregated, and used as input for the model. The system produces a classification result for each course, providing an objective and interpretable assessment of its alignment with academic standards and labor market expectations. A major advantage of the system is its ability not only to classify content but also to identify specific weaknesses. By analyzing feature informativeness, it determines which characteristics negatively or insignificantly affect the overall evaluation, enabling targeted improvements. For example, if a course is classified as satisfactory or unsatisfactory, the system can indicate deficiencies in practical relevance, alignment with labor market needs, or assessment design. As a result, evaluation evolves from a descriptive process into a diagnostic and decision-support tool. Another important application scenario is adaptive course redesign. Based on classification results and feature analysis, instructors can refine materials by focusing on the most impactful characteristics. Enhancing practical tasks, incorporating real-world cases, or strengthening competency alignment can directly improve quality classification. The iterative use of the model enables tracking the effectiveness of such changes, forming a continuous improvement loop.

The approach can also be integrated into learning management systems, such as Moodle. In this case, evaluation becomes partially automated through the use of data from assessments, student feedback, and learning analytics. The classification module can function as an analytical layer, generating reports for instructors and administrators and enabling real-time monitoring. From an administrative perspective, the system supports strategic curriculum management by identifying systemic weaknesses and highlighting best practices across courses. The approach is particularly relevant in labor market-driven education, where content must be continuously validated against external requirements. Incorporating employer feedback ensures that evaluation reflects real-world expectations rather than purely academic criteria, thereby enhancing graduate employability and strengthening collaboration with industry. The ontological representation of the domain further supports the development of intelligent and explainable systems. Each classification decision can be traced back to specific features, ensuring transparency and interpretability. This is especially important in educational environments, where accountability and trust are essential. The model can be adapted to different domains, extended with new features, and applied to large datasets, enabling the creation of centralized systems for monitoring and improving educational quality at institutional or national levels. An important aspect of implementation is the formalization of the operational workflow. In practice, the process can be organized as a cyclic procedure including data collection, preprocessing, model execution, and decision support. Evaluation data are collected using standardized instruments, normalized, and transformed into binary representations using control tolerances. The classification results are visualized through dashboards and reports, ensuring reproducibility and continuous monitoring over time. A key benefit of the approach is the possibility of quantitative interpretation of improvements. Any modification of the content can be linked to measurable changes in the information criterion, allowing validation of improvements and shifting content development toward evidence-based optimization. The system also enables scenario-based analysis, allowing simulation of potential changes prior to implementation and supporting proactive decision-making in curriculum design. In

addition, explainability is inherently supported, as the contribution of each feature to classification decisions can be traced.

Despite these advantages, several challenges must be considered. The effectiveness of the model depends on data quality and representativeness, while parameter selection requires careful calibration. Organizational factors, such as resistance to change, may also affect implementation and require gradual integration. The approach can be further enhanced through integration with advanced analytics and artificial intelligence tools. Extending the feature space with automatically extracted learning data enables the construction of more comprehensive evaluation models, while the ontology-based representation supports interoperability with semantic systems and knowledge graphs. Long-term application allows accumulation of historical data for trend analysis and forecasting, transforming the system into a strategic tool for continuous development rather than a one-time evaluation mechanism.

Overall, the proposed framework provides a comprehensive solution for evaluating, diagnosing, and improving educational content. It supports continuous adaptation to labor market requirements, enhances decision-making, and contributes to the development of competency-oriented educational programs. The results confirm that the information-extreme approach offers a theoretically grounded and practically applicable solution for content evaluation. By combining feature selection, parameter optimization, and interpretable decision rules, it enables effective classification while reducing feature redundancy. The findings also highlight limitations of traditional evaluation approaches and emphasize the importance of selecting informative features. At the same time, the need for improved formalization of complex educational characteristics remains an important research challenge. Future research should focus on expanding the feature space, automating optimization procedures, and integrating the model into digital learning environments. The development of hybrid approaches combining interpretability and predictive capabilities represents a promising research direction. Additionally, longitudinal studies are required to assess long-term impact on educational quality and employability. In summary, the proposed approach provides a robust and flexible framework for educational content evaluation, supporting explainable decision-making and continuous improvement in dynamic environments.

REFERENCES

1. Japaridze, D., Maghlakelidze, S., Shaverdashvili, E., & Tutberidze, D. (2024). The Role of Higher Education in Reducing the Mismatch between Graduates' Skills and Employer Expectations. *International Scientific Conference «Education, Research, Practice» Proceedings*, 5, 300–318. <https://doi.org/10.52340/erp.2024.05.34>
2. Hernández-García, A., Cuenca-Enrique, C., Del-Río-Carazo, L., Iglesias-Pradas S. (2024) Exploring the relationship between LMS interactions and academic performance: A Learning Cycle approach, *Computers in Human Behavior*, 155, 108183. <https://doi.org/10.1016/j.chb.2024.108183>.
3. Alias M.A.H., Hambali N., Aziz M.A.A., Taib M.N., Jailani R. Feature selection techniques and classification algorithms for student performance classification: a review. *International Journal of Electrical and Computer Engineering (IJECE)*, [S.l.], v. 14, n. 3, p. 3230-3243, doi: <http://doi.org/10.11591/ijece.v14i3.pp3230-3243>.
4. Dovbysh, A.S., Budnyk, M.M., Piatachenko, V.Y. et al. (2020) Information-Extreme Machine Learning of On-Board Vehicle Recognition System. *Cybern. Syst. Anal.* 56, 534–543. <https://doi.org/10.1007/s10559-020-00269-y>
5. Shelehov, I. V., Prylepa, D. V., Khibovska, Y. O., & Otroshchenko M. S. (2023). Machine Learning Decision Support Systems for Adaptation of Educational Content to the Labor Market Requirements. *Radio Electronics, Computer Science, Control*, (1), 62. <https://doi.org/10.15588/1607-3274-2023-1-6>
6. Wang, Z., Zhang, X., Li, Y., Yang, H., Xue, H., Wei, Y., & Qian, Y. (2022). Simulation of Cardiac Flow under the Septal Defect Based on Lattice Boltzmann Method. *Entropy*, 24(2), 187. <https://doi.org/10.3390/e24020187>
7. Barbieri, M. C., Grisci, B. I., & Dorn, M. (2024). Analysis and comparison of feature selection methods towards performance and stability. *Expert Systems with Applications*, 249 (Part B), 123667. <https://doi.org/10.1016/j.eswa.2024.123667>
8. Barrera-García, J., Cisternas-Caneo, F., Crawford, B., Gómez Sánchez, M., & Soto, R. (2024). Feature Selection Problem and Metaheuristics: A Systematic Literature Review about Its Formulation, Evaluation and Applications. *Biomimetics*, 9(1), 9. <https://doi.org/10.3390/biomimetics9010009>
9. Romero C, Ventura S. (2020) Educational data mining and learning analytics: An updated survey. *WIREs Data Mining Knowl Discov*.10:e1355. <https://doi.org/10.1002/widm.1355>
10. Vinoth Kumar, E.S., Augustian Isaac, R., Sundaravadivel, P. et al. (2025) Machine learning-driven development of a behaviour-based student classification system (SCS-B) for enhanced educational analytics. *Sci Rep* 15, 37428. <https://doi.org/10.1038/s41598-025-21332-8>
11. Song, X., Zhang, Y., Zhang, W., He, C., Hu, Y., Wang, J., & Gong, D. (2024). Evolutionary computation for feature selection in classification: A comprehensive survey of solutions, applications and challenges. *Swarm and Evolutionary Computation*, 90, 101661. <https://doi.org/10.1016/j.swevo.2024.101661>