

МАТЕМАТИЧЕСКАЯ МОДЕЛЬ УПРАВЛЕНИЯ ОБРАЗОВАТЕЛЬНЫМ ПРОЦЕССОМ НА ОСНОВЕ ИНФОРМАЦИОННОГО АНАЛИЗА ДАННЫХ

В статье предложена математическая модель управления образовательным процессом, основанная на информационном анализе данных и динамическом дереве решений. Модель позволяет персонализировать обучение за счет кластеризации студентов по уровням учебных достижений и применения индекса Джини для оценки неоднородности распределения студентов по учебным достижениям. Применение модели направлено на динамическую корректировку образовательных стратегий и построение индивидуальных образовательных траекторий студентов на основе текущих результатов контроля учебных достижений. Представлены алгоритмы кластеризации, оценивания качества распределения студентов по кластерам и динамического обновления структуры и содержания дерева решений.

Ключевые слова: математическая модель, динамическое дерево решений, кластеризация, энтропия, индекс Джини, персонализация.

Введение. Современный этап модернизации профессионального образования предполагает разработку и внедрение гибких технологий управления учебным процессом на основе учета индивидуальных особенностей каждого студента [11]. Подходы, основанные на традиционных диагностических технологиях, не всегда позволяют эффективно работать с многочисленными группами обучающихся, участники которых существенно различаются по базовой подготовке и текущим учебным достижениям. В этой связи особую актуальность приобретают технологии, которые позволяют анализировать как индивидуальные, так и обобщенные показатели фактического уровня учебных достижений студентов

для динамического построения персонализированных образовательных траекторий [7; 9; 11].

Цель исследования – разработка и апробация математической модели управления образовательным процессом по конкретной дисциплине, позволяющей персонализировать образовательные траектории студентов на основе информационного анализа фактического уровня их учебных достижений.

Задача исследования – предложить и обосновать математическую модель управления образовательным процессом по дисциплине на основе методов интеллектуального анализа данных, включая разработку алгоритмов кластеризации студентов по уровням

учебных достижений, выявления динамики этих достижений и построения персональных образовательных траекторий на основе динамического дерева решений.

Теоретические основы

Персонализация образовательного процесса. Персонализация обучения позволяет формировать индивидуальные образовательные траектории, которые соответствуют фактическому уровню учебных достижений и потребностям конкретных студентов [1]. Использование методов информационного анализа данных в образовательном процессе способствует повышению качества подготовки специалистов и сокращению доли студентов с низкими академическими показателями [3]. Исследования показывают, что адаптивные подходы к обучению, основанные на анализе как персонализированных, так и обобщенных данных, позволяют более эффективно управлять учебным процессом и повышать учебные достижения студентов [4; 6; 8].

Исследование проводилось в БФУ им. И. Канта с привлечением 76 студентов бакалавриата направления подготовки «Землеустройство и кадастры», изучающих дисциплину «Информатика» [11].

Алгоритм информационного анализа образовательных данных. Предлагаемая математическая модель включает несколько ключевых этапов обработки данных, показанных на рисунке.

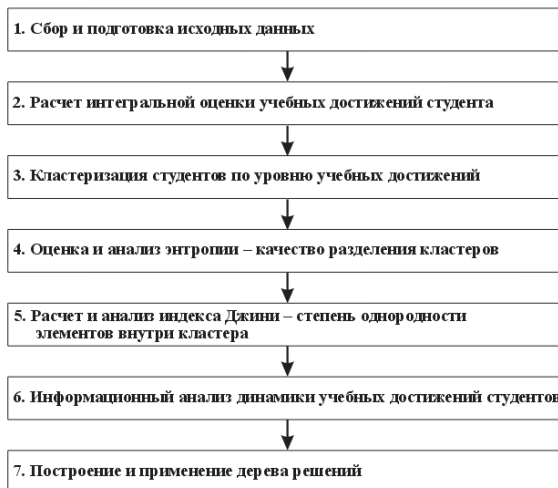


Схема алгоритма информационного анализа образовательных данных

1. Сбор и подготовка исходных данных

- Входные данные: оценки студентов по темам дисциплины, уровень сложности заданий по темам, этапы контроля учебных достижений.
- Формирование массива исходных данных, где каждому студенту соответствует вектор оценок.

2. Расчет интегральной оценки учебных достижений студента

- Рассчитывается взвешенная сумма оценок с учетом сложности тем.
- Позволяет получить обобщенный показатель уровня учебных достижений каждого студента.

3. Кластеризация студентов по уровню учебных достижений

- Экспертный подход (учет педагогического опыта и особенностей учебной группы) и ручная инициализация центроидов каждого кластера.
- Алгоритмический подход (метод K-Means+ и др.) [2].
- Распределение студентов по кластерам.

- Пересчет центроидов по фактическим значениям интегральных оценок учебных достижений студентов, приписанных к соответствующим кластерам.

- Предварительная оценка качества кластеризации.

4. Расчет и анализ энтропии – оценка качества разделения кластеров

- Энтропия рассчитывается для оценки однородности распределения студентов по кластерам.

5. Расчет и анализ индекса Джини – оценка степени однородности элементов внутри кластера

- Индекс Джини отражает разнородность учебных достижений студентов внутри каждого кластера.

- Выявление проблемных кластеров, характеризующихся существенной разнородностью учебных достижений.

6. Информационный анализ динамики учебных достижений студентов

- Сравнение интегральных оценок студента на двух последовательных этапах для выявления позитивной или негативной динамики учебных достижений.

7. Построение и применение дерева решений

- Строится дерево решений с учетом всех рассчитанных оценок качества кластеризации и перестраивается на каждом этапе контроля.

- Дерево обновляется на каждом этапе контроля для адаптации образовательных стратегий и корректировки образовательных траекторий студентов.

- Выходные данные: рекомендации преподавателям по обновлению образовательных стратегий и персонализации траекторий обучения студентов, включая рекомендации по корректировке сложности учебного материала в зависимости от динамики учебных достижений: усложнение программы для студентов с прогрессом и дополнительная поддержка студентов, у которых диагностирован регресс учебных достижений.

Математическая модель информационного анализа образовательных данных, являющаяся основой предложенного алгоритма, базируется на методах информационного анализа данных, включая кластеризацию, энтропийный анализ, индекс Джини и др. Рассмотрим более детальное описание основных компонентов предлагаемой модели.

1. Сбор и подготовка исходных данных

Пусть:

$T = \{T_1, T_2, \dots, T_n\}$ – множество тем дисциплины,

$S = \{S_1, S_2, \dots, S_j\}$ – множество студентов,

$k = \{1, \dots, K\}$ – номер этапа контроля учебных достижений,

$X_{ij,k}$ – оценка студента S_j по теме T_i на k -м этапе,

W_i – коэффициент сложности темы T_i , где $W_i \in (0, 1]$; чем выше сложность, тем больше вклад этой темы в итоговую оценку;

n – количество тем в дисциплине.

2. Расчет интегральной оценки учебных достижений студента

Интегральная оценка учебных достижений студента $Z_{j,k}$ студента S_j

на k -м этапе рассчитывается как взвешенная сумма оценок по всем пройденным темам дисциплины, изученным на текущем и предыдущих этапах обучения, с учетом коэффициентов сложности каждой темы:

$$Z_{j,k} = \sum_{i=1}^n \frac{W_i}{\sum_{i=1}^n W_i} X_{ij,k}, \quad (1)$$

где $Z_{j,k}$ – интегральная оценка учебных достижений студента j на этапе k .

Таким образом, $Z_{j,k}$ формируется нарастающим итогом, позволяя отразить не только текущий уровень подготовки, но и динамику учебных достижений во времени.

3. Кластеризация студентов по уровню учебных достижений

Выбор количества кластеров

Существуют два основных подхода к определению количества кластеров:

Экспертный подход: преподаватель или другое лицо в директивном порядке устанавливает количество кластеров m на основе своих знаний и опыта. Этот подход основан на педагогической практике и эмпирическом учете индивидуальных особенностей членов учебной группы.

Алгоритмический подход: количество кластеров определяется с помощью математических методов, таких как метод K-means, метод локтя или коэффициент силуэта [2]. Описание алгоритмов применения этих методов выходит за рамки настоящей статьи и будет рассмотрено в отдельной публикации.

Ручная инициализация центров:

На начальном этапе кластеризации необходимо задать исходные позиции центроидов каждого кластера μ_r . Например, для пяти кластеров (неявка,

неудовлетворительно, удовлетворительно, хорошо, отлично) могут быть выбраны следующие начальные координаты центроидов:

$$\mu_1 (\text{неявка}) = 0$$

$$\mu_2 (\text{неудовлетворительно}) = 1,5 + (0,5 \times n), \text{ где } n - \text{количество пройденных тем.}$$

$$\mu_3 (\text{удовлетворительно}) = 3,0 + (0,1 \times n)$$

$$\mu_4 (\text{хорошо}) = 4,0 + (0,05 \times n)$$

$$\mu_5 (\text{отлично}) = 4,5 + (0,03 \times n)$$

Координаты центроидов μ_r должны подбираться таким образом, чтобы обеспечить выполнение трех ключевых условий:

1) поддерживать устойчивую дифференциацию между соседними кластерами, предотвращая их смешение или пересечение;

2) учитывать общий объем пройденного учебного материала, что особенно важно при различной продолжительности обучения;

3) отражать традиционную систему оценивания, где каждый последующий уровень (например, от «удовлетворительно» к «хорошо» и «отлично») требует от студента значительно больше усилий для его достижения.

Распределение студентов по кластерам

Для каждого студента S_j рассчитывается расстояние между его интегральной оценкой $Z_{j,k}$ и каждым центроидом μ_r по формуле

$$d(S_j, \mu_r) = \sqrt{(Z_{j,k} - \mu_r)^2}, \quad (2)$$

где $Z_{j,k}$ – интегральная оценка студента S_j на этапе k ; μ_r – центроид кластера r .

Студент S_j включается в кластер C_{rmin} , центроид которого находится

ближе всего, т. е. на минимальном расстоянии от интегральной оценки $Z_{j,k}$:

$$C_{r \min} = \frac{\arg \min d(Z_{j,k}, \mu_r)}{r}.$$

Пересчет центроидов

После первоначального распределения всех студентов по кластерам их центроиды пересчитываются с учетом фактических значений интегральных оценок $Z_{j,k}$ в каждом кластере

$$\mu_r^{(new)} = \frac{1}{|C_r|} \sum_{S_j \in C_r} Z_{j,k}, \quad (3)$$

где $\mu_r^{(new)}$ – обновленный центроид кластера C_r ; $|C_r|$ – количество студентов в кластере с центроидом μ_r ; $Z_{j,k}$ – интегральная оценка студента S_j на этапе k ; $\sum_{S_j \in C_r} Z_{j,k}$ – сумма интегральных оценок всех студентов, включенных в кластер C_r на этапе k .

Процесс повторяется на каждом этапе пересчета учебных достижений, когда появляются новые данные о результатах обучения студентов. Центроиды изменяются, если разность между предыдущим и уточненным значением превышает априори заданный порог чувствительности α :

$$\max_r |\mu_r^{(new)} - \mu_r^{(t)}| > \alpha,$$

где $\mu_r^{(t)}$ – центроид на предыдущей итерации; $\mu_r^{(new)}$ – центроид на текущей итерации; α – заданный порог чувствительности.

Качество кластеризации предварительно оценивается путем расчета внутрикластерной и межкластерной дисперсии [2; 13].

Внутрикластерная дисперсия (Within-Cluster Variance, WCV) характеризует степень однородности оценок студентов внутри кластера C_r . Чем

меньше значение WCV, тем ближе значения оценок к его центроиду. Внутрикластерная дисперсия для кластера C_r рассчитывается по формуле

$$WCV_r = \sum_{S_j \in C_r} (Z_{j,k} - \mu_i)^2. \quad (4)$$

Суммарная внутрикластерная дисперсия по всем кластерам рассчитывается по формуле

$$WCV = \sum_{r=1}^k WCV_r.$$

Чем меньше значение WCV, тем более однородны кластеры.

Межкластерная дисперсия (Between-Cluster Variance, BCV) отражает степень различия между кластерами и вычисляется как

$$BCV = \sum_{r=1}^k |C_r| (\mu_r - \mu)^2, \quad (5)$$

где $|C_r|$ – количество элементов в i -м кластере; μ_r – центроид кластера C_r ; μ – среднее значение всех интегральных оценок:

$$\mu = \frac{1}{m} \sum_{j=1}^m Z_{j,k},$$

где m – общее количество студентов в выборке.

Оценка качества распределения студентов по кластерам определяется по коэффициенту Фишера (F-score). Коэффициент J рассчитывается как отношение межкластерной дисперсии (BCV) к внутрикластерной дисперсии (WCV):

$$J = \frac{BCV}{WCV}. \quad (6)$$

Высокий коэффициент Фишера указывает на то, что группы студентов четко разграничены и между ними существует значительная разница в уровне подготовки. Согласно рекомендациям [16] при $J > 3$ качество разделения кластеров признается высоким,

значение $1,5 \leq J \leq 3$ свидетельствует об **умеренной разграниченности кластеров**, т. е. о приемлемом качестве кластеризации. Малое ($J < 1,5$) значение коэффициента Фишера указывает на размытые границы между кластерами и слабую дифференциацию групп студентов по учебным достижениям, что может говорить о неэффективности выбранных критериев кластеризации или о значительных различиях внутри каждого кластера [14].

При слишком низком значении J качество кластеризации признается неудовлетворительным, преподавателю рекомендуется изменить исходные значения центроидов и повторно провести процедуру распределения студентов по кластерам.

4. Оценка и анализ энтропии – качество разделения кластеров

Энтропия $H(Y)$ характеризует степень неоднородности распределения студентов по кластерам (группам учебных достижений) и используется для оценки качества этого распределения [15].

Если энтропия высокая, это означает, что студенты равномерно распределены по кластерам, границы между которыми представляются недостаточно четкими. Такое распределение препятствует выработке дифференцированных образовательных стратегий для каждой группы студентов, что снижает результативность учебного процесса. В данной ситуации рекомендуется использовать диагностический инструментарий с более высокой дифференцирующей способностью для более точного оценивания учебных достижений каждого студента (в первую оче-

редь, находящихся вблизи границ кластеров). Также для этих студентов можно рекомендовать приоритетное построение индивидуальных образовательных траекторий ввиду высокой вероятности их перехода в смежный кластер.

Низкая энтропия свидетельствует о четко выраженной неоднородности распределения студентов по кластерам ввиду того, что студенты, относящиеся к разным кластерам, характеризуются существенно разными уровнями учебных достижений.

В этой ситуации у преподавателя появляется возможность применять для каждого кластера соответствующую образовательную стратегию, ориентированную на средний уровень учебных достижений всех студентов конкретного кластера. Конечно, такой подход представляется менее трудоемким и более результативным по сравнению с индивидуализацией обучения только части студентов рассматривавшейся выше ситуации с высокой энтропией распределения студентов по кластерам.

Энтропия $H(Y)$ распределения студентов по кластерам рассчитывается по формуле

$$H(Y) = - \sum_{i=1}^K p_i \log_2 p_i, \quad (7)$$

где Y – случайная величина распределения студентов по кластерам; K – количество кластеров; p_i – доля студентов, попавших в кластер C_i :

$$p_i = \frac{|C_i|}{\sum_{j=1}^K |C_j|},$$

где $|C_i|$ – количество студентов в кластере C_i .

Максимальная энтропия достигается, когда все кластеры содержат одинаковое количество студентов [11]:

$$H_{max} = \log_2 K$$

Рост энтропии также может указывать на необходимость уточнения критериев кластеризации, чтобы повысить точность разделения студентов. Если же энтропия снижается, это свидетельствует о формировании устойчивых кластеров, что позволяет адаптировать учебный процесс под потребности каждой группы. В случаях, когда после внедрения новых образовательных методик энтропия не изменяется, целесообразно пересмотреть учебные материалы или подходы к преподаванию.

5. Расчет и анализ индекса Джини
Индекс Джини (*Gini*) – это мера однородности учебных достижений студентов внутри кластера, которая рассчитывается по формуле [12]:

$$Gini = 1 - \sum_{r=1}^K p_r^2, \quad (8)$$

где p_r – доля студентов, принадлежащих кластеру C_r , K – количество кластеров.

Согласно классической работе К. Джини [13] и современным исследованиям в области анализа данных [10], значение индекса Джини позволяет оценить степень однородности учебных групп и сформулировать дифференцированные рекомендации по обучению (см. таблицу).

Интерпретация индекса Джини для кластеров студентов

Interpretation of the Gini index for student clusters

Значение <i>Gini</i>	Интерпретация	Рекомендации
$Gini < 0,3$	Группы студентов однородны	Можно применять единые образовательные стратегии
$0,3 \leq Gini \leq 0,6$	Умеренная разнородность	Рекомендуется частичная дифференциация обучения
$Gini > 0,6$	Высокая разнородность	Требуется индивидуальный подход, адаптация учебных материалов

Если индекс $Gini \rightarrow 0$, это означает, что студенты внутри одного кластера обладают схожими знаниями (кластеры однородны). В этом случае можно использовать единые образовательные стратегии для всех студентов группы.

Если индекс $Gini \rightarrow 1$, это указывает на сильные различия внутри кластера (кластеры неоднородны), что требует гибких образовательных стратегий, например, разделения студентов на подгруппы по уровню подготовки или адаптации учебных материалов

под индивидуальные потребности студентов.

6. Информационный анализ динамики учебных достижений студентов

Чтобы оценить прогресс или регресс учебных достижений студента, можно ввести индекс изменения интегральной оценки на двух последовательных этапах учебного процесса [10]:

$$\Delta Z_j = Z_{j,k} - Z_{j,k-1}, \quad (9)$$

где $Z_{j,k}$ – интегральная оценка студента j на текущем этапе k ; $Z_{j,k-1}$ – интегральная оценка студента j на предыдущем этапе.

Также введем показатель ε – допустимая погрешность оценивания (например, связанная с изменением формата задания, условий контроля и пр.).

Анализ значений $Z_{j,k}$ позволяет сделать следующие выводы:

если $\Delta Z_{j,k} > \varepsilon$, это свидетельствует о **положительной динамике**, то есть прогрессе студента. В этом случае рекомендуется усложнение заданий, расширение объема изучаемого материала, предложение участия в конкурсах, проектной деятельности или перевод в группу с более высоким уровнем подготовки.

если изменение оценки находится в пределах допустимой погрешности (то есть $|\Delta Z_{j,k}| \leq \varepsilon$), наблюдается **стабильность учебных достижений**. Это означает, что студент сохраняет достигнутый уровень подготовки. В этом случае обучение продолжается по текущей траектории, однако при длительно наблюдающейся стагнации учебных результатов целесообразно провести анализ причин отсутствия роста (например, отсутствие мотивации, недостаточная сложность заданий и др.).

если $\Delta Z_{j,k} < -\varepsilon$, – это **регресс**, что может быть связано с потерей интереса, внешними обстоятельствами, ухудшением понимания учебного материала. В этом случае рекомендуются проведение индивидуальных консультаций, корректировка образовательной траектории, назначение дополнительных заданий по проблемным темам, а также временное упрощение содержания и форм контроля.

7. Построение и применение дерева решений

Дерево решений строится с учетом всех рассчитанных оценок качества кластеризации и перестраивается на каждом этапе контроля. Основные элементы [5]:

Корневой узел дерева решений реализует первичное разделение на две основные категории:

- 1) преподаватели;
- 2) студенты.

Ветвь «Преподаватели» содержит рекомендации по совершенствованию образовательной деятельности:

• Стратегии групповой работы

Формируются на основе анализа индикаторов качества кластеризации (коэффициент Фишера J , энтропия, индекс Джини).

Например:

• Если коэффициент Фишера $J > 3$ и индекс Джини $< 0,3$, то использовать дифференцированное обучение по кластерам;

• Если коэффициент Фишера $J \leq 2,5$, то изменить структуру заданий и параметры кластеризации и т. п.

• Персональные рекомендации преподавателям

Выдаются на основе комбинации следующих параметров:

1. Индивидуальные показатели студента:

- динамика успеваемости
- позиция в кластере (нормализованное отклонение от центроида) и т. п.

2. Характеристики кластера:

- индекс Джини
- внутрикластерная дисперсия.

Анализ ключевых показателей результативности образовательного процесса:

- изменение успеваемости каждого студента (улучшение/ухудшение);
- позиция студента в своей группе (слабая/средняя/сильная);
- степень однородности кластера по уровню учебных достижений.

Рекомендации для преподавателей:

- В случае высокой однородности показателей кластеризации – формулирование рекомендаций по выработке и применению новых групповых образовательных стратегий.
- В случае низкой однородности показателей кластеризации – формулирование рекомендаций по построению индивидуальных образовательных траекторий студентов, характеризующихся как наибольшими, так и наименьшими учебными достижениями.
- В случае умеренной однородности показателей кластеризации – формулирование рекомендаций по выборочной корректировке применяемых образовательных стратегий с целью повышения их результативности.

Ветвь «Студенты»

Предварительные исследования направленности и характера рекомендаций студентам свидетельствуют о необходимости комплексного учета не только показателей качества кластеризации, но и показателей динамики учебных достижений каждого студента, а также его позиции в конкретном кластере. Однако даже при полной автоматизации обработки информации по всем студентам попытки индивидуально работать с каждым из них приводят к кратному увеличению нагрузки на

преподавателя и превышению его допустимой нагрузки. Поэтому для построения либо корректировки «студенческой» ветви дерева решений представляется целесообразной реализация двухуровневой аналитической процедуры. На ее первом этапе вырабатываются персональные рекомендации для каждого студента, которые могут быть использованы, в частности в процессе самостоятельной подготовки. На втором этапе эти рекомендации обобщаются для групп студентов, показатели которых различаются на величину, не превышающую априори заданную погрешность, – например, для всех студентов, позиционированных на расстоянии не более x от верхней границы соответствующего кластера, либо для всех студентов, регресс учебных достижений которых превысил априори заданную величину y . Для реализации такого обобщения представляется перспективным применение технологий машинного обучения на основе нейронных сетей, выполняющих функции группирования или агрегирования данных [5].

Заключение

В статье предложена математическая модель управления образовательным процессом, основанная на методах информационного анализа данных об учебных достижениях студентов. Модель позволяет:

- формализовать процесс кластеризации студентов по уровням учебных достижений с использованием интегральных оценок;
- оценивать качество распределения студентов по кластерам с помощью дисперсионного и энтропийного анализа, а также индекса Джини;

- выявлять прогресс или регресс учебных достижений каждого студента на основе информационного анализа изменений интегральной оценки;

- формировать рекомендации по корректировке образовательных стратегий и/или построению индивидуальных либо групповых персонализированных образовательных траекторий путем динамического построения дерева решений.

Применение модели способствует повышению эффективности управления учебным процессом за счет:

1) агрегирования и обобщения данных об учебных достижениях студентов;

2) своевременной диагностики проблем – оперативного выявления студентов с регрессом или стагнацией учебных достижений для выработки корректирующих воздействий;

3) персонализации обучения – адаптации образовательных стратегий под индивидуальные потребности студентов.

4) оптимизации педагогических решений – дифференциации учебных материалов и методов преподавания в зависимости от уровня и динамики учебных достижений в фокусных группах.

Предварительные результаты практической апробации модели в Балтийском федеральном университете имени И. Канта при проведении занятий по дисциплине «Информатика» подтверждают ее эффективность. Применение технологии кластеризации и информационного анализа, а также динамическое построение, обучение и использование дерева решений позволяют сократить долю студентов с низкой успеваемостью и повысить общий уровень академических достижений.

Перспективы дальнейших исследований включают:

- интеграцию модели с используемыми в вузе цифровыми образовательными платформами;

- разработку компьютерного приложения для преподавателей, реализующего технологию кластеризации и информационного анализа;

- исследования и детализацию процедуры построения дерева решений;

- расширение сферы применения предложенной модели за счет одновременного и комплексного анализа учебных достижений по нескольким дисциплинам.

Литература

1. Вдовина С. А., Кунгурова И. М. Сущность и направления реализации индивидуальной образовательной траектории // Вестник евразийской науки. 2013. № 6 (19). С. 175.
2. Воронина В. В. Теория и практика машинного обучения : учеб. пособие / В. В. Воронина, А. В. Михеев, Н. Г. Ярушкина, К. В. Святков. Ульяновск : УлГТУ, 2017. 290 с.
3. Грушецкий С. В., Рудинский И. Д. Модель статистического оценивания знаний // Информационные технологии. 2004. № 12. С. 48 – 54.
4. Зайков А. С., Никитин П. В. Комплект учебно-методических материалов (компьютерные мотивационные игры) // Хроники объединенного фонда электронных ресурсов «Наука и образование». 2014. № 12 (1). С. 97 – 97.

5. Современные образовательные технологии в учебном процессе вуза : метод. пособие / Е. А. Касаткина [и др.]. Кемерово, 2011. 237 с.
6. Минеева О. А., Даричева М. В. Организация тестирования в системе Moodle при обучении иностранному языку // Научно-педагогическое обозрение. Pedagogical Review. 2016. № 3 (13). С. 81 – 86.
7. Рудинский И. Д., Давыдов А. В. Гибридные образовательные технологии: анализ возможностей и перспективы применения // Вестник науки и образования Северо-запада России. 2021. № 1 (7). С. 44 – 52.
8. Рудинский И. Д., Клеандрова И. А. Концепция количественного оценивания объективности педагогического тестирования знаний // Информатика и образование. 2003. № 12. С. 100 – 104.
9. Рудинский И. Д., Ли О. Ю. Гибридные образовательные технологии формирования компетенции поиска и критического анализа у будущих бакалавров по землеустройству и кадастрам // Вестник Российского университета дружбы народов. 2024. № 3 (21). С. 373 – 393.
10. Рудинский И. Д., Ли О. Ю. Модель информационно-технологического обеспечения подготовки специалистов по землеустройству // Информатизация образования и науки. 2025. № 65 (1). С. 35.
11. Рудинский И. Д., Ли О. Ю. Применение методов и моделей системного анализа для оптимизации образовательного процесса подготовки специалистов по землеустройству. 2025. № 1 (71). С. С. 33 – 44.
12. Baker R. Data mining for education // International encyclopedia of education. 2010. № 3 (7). С. 112 – 118.
13. Gini C. Variabilità e mutabilità (1912) // Memorie di metodologica statistica. Rome : Libreria Eredi Virgilio Veschi. 1955. (10).
14. MacQueen J. Some methods for classification and analysis of multivariate observations. California : Un-ty of California press, 1967. С. 281 – 298.
15. Shannon C. E. A mathematical theory of communication // The Bell system technical journal. 1948. № 3 (27). С. 379 – 423.
16. Tan P.-N., Steinbach M., Kumar V. Introduction to data mining // Pearson Education India. 2016.

References

1. Vdovina S. A., Kungurova I. M. Sushhnost` i napravleniya realizacii individual`noj obrazovatel`noj traektorii // Vestnik evrazijskoj nauki. 2013. № 6 (19). С. 175.
2. Voronina V. V. Teoriya i praktika mashinnogo obucheniya : ucheb. posobie / V. V. Voronina, A. V. Mixeev, N. G. Yarushkina, K. V. Svyatov. Ul`yanovsk : UIGTU, 2017. 290 s.
3. Grusheczkij S. V., Rudinskij I. D. Model` statisticheskogo ocenivaniya znaniy // Informacionny`e texnologii. 2004. № 12. С. 48 – 54.
4. Zajkov A. S., Nikitin P. V. Komplekt uchebno-metodicheskix materialov (komp`yuterny`e motivacionny`e igry`) // Xroniki ob``edinennogo fonda e`lektronny`x resursov «Nauka i obrazovanie». 2014. № 12 (1). С. 97 – 97.
5. Sovremenny`e obrazovatel`ny`e texnologii v uchebnom processe vuza : metod. posobie / E. A. Kasatkina [i dr.]. Kemerovo, 2011. 237 s.

6. Mineeva O. A., Daricheva M. V. Organizaciya testirovaniya v sisteme Moodle pri obuchenii inostrannomu yazy`ku // Nauchno-pedagogicheskoe obo-zrenie. Pedagogical Review. 2016. № 3 (13). C. 81 – 86.
7. Rudinskij I. D., Davy`dov A. V. Gibridny`e obrazovatel`ny`e texnologii: analiz vozmozhnostej i perspektivy` primeneniya // Vestnik nauki i obrazovaniya Severo-zapada Rossii. 2021. № 1 (7). C. 44 – 52.
8. Rudinskij I. D., Kleandrova I. A. Konceptiya kolichestvennogo ocenivaniya ob`ektivnosti pedagogicheskogo testirovaniya znaniy // Informatika i obrazovanie. 2003. № 12. C. 100 – 104.
9. Rudinskij I. D., Li O. Yu. Gibridny`e obrazovatel`ny`e texnologii formirovaniya kompetencii poiska i kriticheskogo analiza u budushhix bakalavrov po zemleustroystvu i kadastram // Vestnik Rossijskogo universiteta družby` narodov. 2024. № 3 (21). C. 373 – 393.
10. Rudinskij I. D., Li O. Yu. Model` informacionno-texnologicheskogo obespecheniya podgotovki specialistov po zemleustroystvu // Informatizaciya obrazovaniya i nauki. 2025. № 65 (1). C. 35.
11. Rudinskij I. D., Li O. Yu. Primenenie metodov i modelej sistemnogo analiza dlya optimizacii obrazovatel`nogo processa podgotovki specialistov po zemleustroystvu 2025. № 1 (71). C. S. 33 – 44.
12. Baker R. Data mining for education // International encyclopedia of education. 2010. № 3 (7). C. 112 – 118.
13. Gini C. Variabilità e mutabilità (1912) // Memorie di metodologica statistica. Rome : Libreria Eredi Virgilio Veschi. 1955. (10).
14. MacQueen J. Some methods for classification and analysis of multivariate observations. California : Un-ty of California press, 1967. C. 281 – 298.
15. Shannon C. E. A mathematical theory of communication // The Bell system technical journal. 1948. № 3 (27). C. 379 – 423.
16. Tan P.-N., Steinbach M., Kumar V. Introduction to data mining // Pearson Education India. 2016.

I. D. Rudinsky, O. Yu. Li

MATHEMATICAL MODEL OF EDUCATIONAL PROCESS MANAGEMENT BASED ON INFORMATION DATA ANALYSIS

The article proposes a mathematical model for managing the educational process based on information analysis of data and a dynamic decision tree. The model allows personalizing learning by clustering students by levels of academic achievement and using the Gini index to assess the heterogeneity of student distribution by academic achievement. The application of the model is aimed at dynamically adjusting educational strategies and building individual educational trajectories for students based on current results of monitoring academic achievement. Algorithms for clustering, assessing the quality of student distribution by clusters, and dynamically updating the structure and content of the decision tree are presented.

Key words: *mathematical model, dynamic decision tree, clustering, entropy, Gini index, personalization.*