

АЛГОРИТМ ФОРМИРОВАНИЯ АДАПТИВНОЙ ТРАЕКТОРИИ ОБУЧЕНИЯ ШКОЛЬНИКОВ НА ОСНОВЕ ДАННЫХ ОБ ИХ УСПЕВАЕМОСТИ В ЦЕНТРЕ ПОДГОТОВКИ К ЕГЭ И ОГЭ

А. А. Драган^а, А. В. Гавриленко^б

Сургутский государственный университет, г. Сургут, Российская Федерация

^а ORCID: <https://orcid.org/0009-0009-9589-9808>, ✉ dragan.alexandr.902@gmail.com

^б ORCID: <https://orcid.org/0000-0003-1837-5698>, gavrilenko.anna.v@gmail.com

Аннотация: персонализированное обучение в центрах подготовки к ЕГЭ и ОГЭ позволяет с помощью современных технологий адаптировать учебный процесс под каждого ученика. Разработан алгоритм формирования адаптивной траектории обучения на основе данных об успеваемости учеников. Для сбора данных при анализе успеваемости ученика использовалась IRT-модель. Для оценки образовательных возможностей ученика применялся расчет вероятности правильного ответа с помощью трехпараметрической модели Бирнбаума. Полученное значение использовалось при расчете уровня знаний учащегося с применением метода максимального правдоподобия. Для распределения учеников по группам успеваемости использовался метод кластеризации K-means.

Ключевые слова: IRT-модель для сбора данных, метод кластеризации K-means, метод максимального правдоподобия, трехпараметрическая модель Бирнбаума, адаптивная образовательная траектория.

Для цитирования: Драган А. А., Гавриленко А. В. Алгоритм формирования адаптивной траектории обучения школьников на основе данных об их успеваемости в центре подготовки к ЕГЭ и ОГЭ. *Успехи кибернетики*. 2026;7(2):151–156.

Поступила в редакцию: 03.02.2026.

В окончательном варианте: 27.04.2026.

A METHOD FOR BUILDING ADAPTIVE LEARNING PATHS FOR SCHOOL STUDENTS BASED ON ACADEMIC PERFORMANCE IN STANDARDIZED SCHOOL EXAMS

A. A. Dragan^a, A. V. Gavrilenko^b

Surgut State University, Surgut, Russian Federation

^a ORCID: <https://orcid.org/0009-0009-9589-9808>, ✉ dragan.alexandr.902@gmail.com

^b ORCID: <https://orcid.org/0000-0003-1837-5698>, gavrilenko.anna.v@gmail.com

Abstract: we studied personalized learning in school exam preparation settings and developed a method for constructing adaptive learning paths based on student performance in standardized school exams. The approach targets individual learning needs and uses modern data-driven techniques to support instructional planning.

We collected student performance data and applied Item Response Theory (IRT) to model test responses. We estimated the probability of a correct answer using a three-parameter logistic model. We then inferred student ability levels using maximum likelihood estimation.

We grouped students according to their estimated ability using the K-means clustering method. The resulting clusters provided a basis for adapting learning trajectories to different performance levels.

Keywords: Item Response Theory (IRT), K-means, maximum likelihood estimation, three-parameter logistic model, adaptive learning path.

Cite this article: Dragan A. A., Gavrilenko A. V. A Method for Building Adaptive Learning Paths for School Students Based on Academic Performance in Standardized School Exams. *Russian Journal of Cybernetics*. 2026;7(2):151–156.

Original article submitted: 03.02.2026.

Revision submitted: 27.04.2026.

Введение

В образовательном процессе индивидуализация обучения является ключевым фактором для успешного повышения уровня знаний учеников. Актуальной задачей представляется внедрение современных информационных технологий с использованием алгоритмов формирования адаптивных траекторий обучения на основе анализа данных об успеваемости в центрах подготовки к ЕГЭ и ОГЭ. Применение индивидуальной образовательной траектории позволяет интенсифицировать и индивидуализировать образовательную деятельность обучающихся за счет повышения уровня их субъектности, стимулируя, таким образом, их становление активными участниками системы непрерывного образования [1].

В настоящей работе описывается метод сбора и анализа данных с использованием ИРТ-модели, расчет уровня знаний учащихся на основе трехпараметрической модели Бирнбаума и метода максимального правдоподобия, а также последующая обработка данных посредством кластеризации методом K-means.

Сбор данных и оценка уровня знаний учащихся

В целях сбора данных ученикам предлагаются задания фиксированной сложности, чтобы провести стартовое тестирование с использованием ИРТ-модели для определения вероятности правильного ответа на задания. Получившееся значение используется для расчета уровня знаний ученика (θ) на основе метода максимального правдоподобия с оценкой по шкале от -1 до 1 со следующей градацией:

1. Уровень «Низкий», диапазон [-1; -0.5]. На данном уровне знаний ученик справляется с решением легких заданий, с заданиями средней сложности периодически возникают трудности.

2. Уровень «Средний», диапазон [-0.5; 0.5]. На данном уровне знаний ученик может решать легкие задания и задания средней сложности с периодическими ошибками, а также изредка решает сложные задания.

3. Уровень «Высокий», диапазон [0.5; 1]. На данном уровне ученик выполняет задания вне зависимости от уровня сложности с периодическими ошибками.

ИРТ (Item Response Theory) — это модель, которая показывает, что вероятность правильного ответа на задание j у хорошо подготовленного испытуемого должна быть больше вероятности правильного ответа у слабо подготовленного испытуемого: чем выше подготовка испытуемого, тем выше может быть вероятность правильного ответа на задание данного уровня сложности [2]. При использовании данной модели необходимо выделить основные параметры для дальнейшего определения уровня знаний учеников:

1) A (дискриминативность). Параметр дискриминативности в диапазоне от -1 до 1 помогает определить, насколько задание различает учеников с разным уровнем знаний.

2) B (сложность). Данный параметр отражает сложность заданий в диапазоне от 0 до 2.

3) C (угадывание). Данный параметр отражает вероятность угадывания учеником правильного ответа на задания в диапазоне от 0 до 0.99.

Для расчета вероятности правильного ответа необходимо найти параметр дискриминативности:

1) нужно определить группы учеников: выделить группу «сильных» испытуемых и группу «слабых» испытуемых;

2) посчитать для каждого задания количество правильных ответов у «сильной» и «слабой» группы;

3) рассчитать разницу между количеством правильных ответов в «слабой» группе и количеством правильных ответов в «сильной» группе;

4) разделить полученную разницу на количество учеников в обеих группах.

Также необходимо определить уровень сложности заданий по следующей шкале:

1) уровень «0»: легкое задание;

2) уровень «1»: задание средней сложности;

3) уровень «2»: задание повышенной сложности.

Кроме того, необходимо найти вероятность угадывания правильного ответа, для этого необходимо выполнить следующие действия:

1) определить параметр n — количество возможных вариантов ответа;

- 2) определить параметр m — количество правильных ответов;
- 3) рассчитать вероятность по формуле (1):

$$P(C) = \frac{m}{n}, \quad (1)$$

где C — это параметр угадывания.

Также в заданиях могут присутствовать текстовые ответы, для которых необходимо зафиксировать значения с учетом сложности заданий со следующей градацией:

- 1) для заданий низкой сложности определить вероятность угадывания как 0.4;
- 2) для заданий средней сложности определить вероятность угадывания как 0.2;
- 3) для заданий высокой сложности определить вероятность угадывания как 0.05.

Следующим этапом является расчет вероятности правильного ответа с помощью трехпараметрической модели Бирнбаума [3]:

$$P(x) = c + (1 - c) * \frac{1}{1 + e^{-a(\theta - b)}}, \quad (2)$$

где $P(x)$ — вероятность правильного ответа, θ — уровень знаний ученика, a — параметр дискриминативности, b — параметр сложности, c — параметр угадывания.

Метод максимального правдоподобия — это метод оценивания неизвестного параметра путем максимизации функции правдоподобия. Основан на предположении о том, что вся информация о статистической выборке содержится в функции правдоподобия [4]. Для определения уровня знаний учеников используем метод максимального правдоподобия для одного ученика:

$$L(\theta) = \prod_{j=1}^M P_j(\theta)^{x_j} (1 - P_j(\theta))^{1-x_j}, \quad (3)$$

где $P_j(\theta)$ — вероятность правильного ответа на задание j , рассчитанная по трехпараметрической модели Бирнбаума, x_j — ответ ученика на задание j (1 — правильно, 0 — неправильно).

Для составления адаптивной траектории обучения, кроме уровня знаний ученика, необходимо учитывать предмет и скорость обучения. Предмет определяется на основе содержания тестовых заданий с учетом измеренных параметров (дискриминативность, сложность, угадывание). Для определения скорости обучения используется лонгитюдный метод сбора данных — длительное и систематическое изучение одного и того же объекта для выявления динамики его существования и прогнозирования его дальнейшего развития [5]. Для получения параметра скорости обучения необходимо следовать следующему плану:

1. Проведение тестирования в одинаковых интервалах времени, например раз в неделю.
2. Сбор результатов.
3. Расчет уровня знаний.
4. Построение графика:
 - необходимо построить график с осями X (время) и Y (уровень знаний ученика);
 - следует рассчитать значения переменных m и b для построения линии регрессии, используя метод наименьших квадратов [6]:

$$m = \frac{N * \sum xy - \sum x * \sum y}{N * \sum x^2 - (\sum x)^2}, \quad (4)$$

$$b = \frac{\sum y - m * \sum x}{N}, \quad (5)$$

где m — наклон линии регрессии, что будет являться скоростью обучения;

b — пересечение с осью Y ;

N — количество точек данных;

x — значение времени;

y — значение параметра знаний ученика (θ).

После расчета значений переменных m и b необходимо построить линию регрессии [5]:

$$y = mx + b. \quad (6)$$

Линия регрессии будет показывать динамику скорости обучения ученика.

Результатом сбора данных и их оценки будет являться определение уровня знаний, скорости и предмета обучения каждого ученика.

Анализ данных и кластеризация учеников

На этапе анализа необходимо учитывать параметры уровня знаний (θ), предмета и скорости обучения, полученные на предыдущем этапе, для кластеризации учеников [7]:

1. Для начального определения группы учеников необходимо задать центры кластеров.
2. Определить для каждого ученика ближайший кластер с использованием вычисления ближайшего расстояния евклидовой метрикой:

$$d = \sqrt{(x_1 - x_2)^2 + (y_1 - y_2)^2}, \quad (7)$$

где x — уровень знаний ученика (θ), а y — скорость обучения.

3. Пересчитать центры кластеров на основании параметров учеников, распределенных по группам:

$$K_i = \left(\frac{\sum_{j=1}^n x_j}{n}, \frac{\sum_{j=1}^n y_j}{n} \right), \quad (8)$$

где x — уровень знаний ученика (θ), а y — скорость обучения;

i — порядковый номер кластера (K_1 — низкий уровень, K_2 — средний уровень, K_3 — высокий уровень);

n — количество учеников в кластере;

K — значение центра кластера.

В результате получения значения кластеризации формируются три типовые группы учеников с низким, средним и высоким уровнем знаний.

Формирование адаптивной траектории обучения

При формировании адаптивной траектории обучения используются данные о текущем уровне знаний учеников, предмете и скорости обучения, а также кластер, в котором ученик находится, для выполнения заданий, находящихся в диапазоне сложности $\theta \pm 0.5$. В случае если вероятность успешного выполнения заданий превышает 80% для текущего уровня, сложность повышается; если она ниже 50%, то сложность понижается. После выполнения заданий на основе их результатов необходимо пересчитать значения уровня знаний ученика, скорости обучения и определить новый кластер для каждого ученика.

Пример использования алгоритма формирования адаптивной траектории обучения

В центре подготовки к ЕГЭ и ОГЭ существует группа из 6 человек по обществознанию. Для оценки уровня знаний ученикам предлагается тест, состоящий из 10 заданий различной сложности.

Для подсчета вероятности правильного ответа на задание необходимо определить значение дискриминативности, сложность задания и вероятность угадывания. Для определения значения дискриминативности разделим группу на сильных и слабых учеников следующим образом:

- «сильные ученики»: ученик 1, ученик 3, ученик 4;
- «слабые ученики»: ученик 2, ученик 5, ученик 6.

Для каждого задания следует определить уровень дискриминативности исходя из ответов всех учеников на конкретное задание. Например, расчет уровня дискриминативности для задания Q1:

1. Подсчет правильных ответов по группам:

- «сильные ученики»: [1, 1, 0];
- «слабые ученики»: [1, 0, 0].

2. Подсчет разницы между количеством правильных ответов в группах, равной 1.

3. Нормализация:

- общее количество учеников в обеих группах: 6;
- A (параметр дискриминативности) для $Q_1 = 1/6 \approx 0,17$.

Таким образом необходимо просчитать значение дискриминативности для каждого задания теста.

Таблица 1

Параметры заданий теста

Задание	Дискриминативность (A)	Сложность (B)	Угадывание (C)
Q_1	0.17	1	0.05
Q_2	0.66	2	0.40
Q_3	-0.58	0	0.05
Q_4	-0.64	1	0.05
Q_5	-0.63	1	0.40
Q_6	-0.39	2	0.40
Q_7	0.05	1	0.05
Q_8	-0.14	1	0.20
Q_9	-0.42	2	0.05
Q_{10}	0.22	2	0.05

Для каждого ученика необходимо высчитать вероятность правильного ответа с использованием трехпараметрической модели Бирнбаума исходя из ответов на задания и их параметров (таблица 1). Расчет вероятности для ученика 1 из «сильной» группы:

1. Результат теста: [1, 0, 1, 1, 1, 0, 0, 1, 1, 0];

2. Расчет вероятности правильного ответа с использованием трехпараметрической модели Бирнбаума.

Для Q_1 : $A = 0.17$, $B = 1$, $C = 0.05$, $\theta = 0$ (при начальном тестировании используется значение по умолчанию).

Подставляем в формулу для расчета вероятности правильного ответа:

$$P(x = 1) = 0.05 + (1 - 0.05) * \frac{1}{1 + e^{-0.17(0-1)}} \approx 0.4847.$$

Необходимо рассчитать вероятность правильного ответа для каждого задания ученика 1.

Подставляем данные из таблицы 2 для расчета уровня знаний ученика:

$$L(\theta) = P_1(\theta)^{x_1} * P_2(\theta)^{x_2} * \dots * P_{10}(\theta)^{x_{10}} = 0.485 * 0.473 * 0.525 * 0.378 * 0.608 * 0.414 * 0.486 * 0.572 * 0.337 * 0.578 = 0.00062.$$

Рассчитываем скорость обучения с помощью лонгитюдного метода сбора данных с использованием метода наименьших квадратов:

$x = 8$ часов, $y = 0.00062$,

$$m = \frac{2 * (8 * 0.00062) - 8 * 0.00062}{2 * 8^2 - 0.00062^2} \approx 0.000039.$$

После нахождения скорости обучения и уровня знаний ученика необходимо определить кластер ученика:

- кластер 1 находится в диапазоне [-1; -0.5];
- кластер 2 находится в диапазоне [-0.5; 0.5];
- кластер 3 находится в диапазоне [0.5; 1].

Таблица 2

Вероятности правильного ответа на тест ученика 1

Задание	Вероятность правильного ответа
Q_1	0.485
Q_2	0.527
Q_3	0.525
Q_4	0.378
Q_5	0.608
Q_6	0.586
Q_7	0.514
Q_8	0.572
Q_9	0.337
Q_{10}	0.422

Рассчитываем расстояние до кластеров 1, 2 и 3:

$$d = \sqrt{(0.00062 - 0)^2 + (0.000039 - 0)^2} = 0.000615.$$

Делаем вывод, что ученик 1 относится к кластеру 2.

В дальнейшем подбор заданий необходимо осуществлять с учетом значения уровня знаний ученика.

Заключение

В ходе выполнения данной работы реализован алгоритм формирования адаптивных траекторий обучения с использованием ИРТ-модели для сбора данных и метода правдоподобия для оценки уровня знаний ученика. Исходя из полученных значений происходит подбор заданий для учеников.

ЛИТЕРАТУРА

1. Шеманаева М. А. Индивидуальная образовательная траектория как форма синхронно-асинхронной образовательной деятельности. *Научно-методический электронный журнал «Концепт»*. 2017;9:43-47.
2. Аванесов В. С. Item Response Theory. Основные понятия и положения. *Педагогические измерения*. 2007;3:23-28.
3. Муратова Л. А. Модели Бирнбаума для оценки качества теста «Линейная алгебра, аналитическая геометрия». *Вестник Самарского государственного технического университета. Серия: Психолого-педагогические науки*. 2017;3:83-91.
4. Клепиков Н. П., Соколов С. Н. *Анализ и планирование экспериментов методом максимума правдоподобия: практическое пособие*. Москва: Наука; 1964. 190 с.
5. Корнилов С. А. Лонгитюдные исследования: теория и методы. *Экспериментальная психология*. 2011;4(4):101-116.
6. Линник Ю. В. *Метод наименьших квадратов и основы математико-статистической теории обработки наблюдений*. Москва: Физматгиз; 1962. 90 с.
7. Кугаевских А. В., Муромцев Д. И., Кирсанова О. В. *Классические методы машинного обучения: учебное пособие*. СПб.: НИУ ИТМО; 2022. 14 с.
8. Чернова Н. И. *Математическая статистика: учебное пособие*. Новосибирск: СибГУТИ; 2009. 90 с.