

УДК 519.2:37
DOI

МОДЕЛЬ ЭКЗАМЕНАЦИОННЫХ ОЦЕНОК С ЛАТЕНТНЫМИ ПАРАМЕТРАМИ НА БАЗЕ ФУНКЦИИ SOFTMAX

Братищенко В.В.

Байкальский государственный университет, Иркутск, e-mail: vbrat56@mail.ru

Цель данной работы заключается в изучении влияния на оценки характеристик обучающихся и особенностей экзаменов. Наличие такого влияния констатируется в значительном количестве научных работ. Предложенная модель экзаменационных оценок учитывает академические способности обучающихся и уровни сложности экзаменов в виде латентных параметров. Для моделирования оценок предлагается использовать распределение вероятностей, задаваемое функцией softmax с латентными параметрами. Исходными данными для определения параметров является набор оценок обучающихся. Методом максимального правдоподобия разработан численный алгоритм для оценивания параметров по набору оценок. Получены неравенства для дисперсий оценок латентных параметров, которые можно применять для интервального оценивания. По предложенной методике обработан набор оценок. Адекватность модели проверена методами дисперсионного анализа. Определен коэффициент детерминации, значение которого оказалось близким к теоретически предсказанному. Предложенная модель отличается от аналогов применением функции softmax для описания вероятностей оценок и использованием латентных параметров для каждой градации оценок. Модель может использоваться для определения законов распределения вероятностей оценок по оценкам латентных параметров обучающихся и экзаменов. Распределения вероятностей оценок можно применять для прогнозирования успеваемости и совершенствования экзаменационных процедур.

Ключевые слова: модель экзаменационных оценок, латентные параметры, функция softmax, прогнозирование успеваемости, критерий Фишера, коэффициент детерминации

A LATENT PARAMETER MODEL OF EXAMINATION GRADES BASED ON THE SOFTMAX FUNCTION

Bratishchenko V.V.

Baikal State University, Irkutsk, e-mail: vbrat56@mail.ru

The purpose of this paper is to study the influence of student characteristics and exam features on grades. The presence of such influence is stated in a significant number of scientific papers. The proposed model of examination grades takes into account the academic abilities of students and the difficulty levels of exams in the form of latent parameters. To model grades, it is proposed to use the probability distribution specified by the softmax function with latent parameters. The initial data for determining the parameters is a set of student grades. Using the maximum likelihood method, a numerical algorithm was developed for estimating parameters based on a set of grades. Inequalities were obtained for the variances of latent parameter estimates, which can be used for interval estimation. A set of grades was processed using the proposed methodology. The adequacy of the model was tested using variance analysis methods. The determination coefficient was determined, the value of which turned out to be close to the theoretically predicted one. The proposed model differs from its analogues in that it uses the softmax function to describe the probabilities of grades and uses latent parameters for each grade gradation. The model can be used to determine the probability distribution laws of grades based on the assessments of latent parameters of students and exams. The probability distributions of grades can be used to predict academic performance and improve examination procedures.

Keywords: examination grades model, latent parameters, softmax function, academic performance prediction, Fisher criterion, determination coefficient

Введение

Экзаменационные оценки являются измерителями результатов учебной деятельности. Исследование оценок позволяет выявлять различные недостатки учебного процесса и определять направления его улучшения. Еще одной важной задачей является прогнозирование успеваемости. Прогноз не только позволяет оценить будущее количество обучающихся, но и предложить ряд мероприятий для уменьшения количества неуспевающих.

На экзаменационную оценку влияет большое количество факторов, таких как уровень подготовки обучающегося, личностные характеристики, условия проживания, орга-

низация учебного процесса, взаимоотношения с экзаменатором. Многие факторы являются случайными. Все это приводит к тому, что оценку рассматривают как случайную величину, для которой нужно установить закономерности ее формирования.

Анализ данных успеваемости является значительной частью целого направления исследований – интеллектуального анализа образовательных данных (Educational Data Mining) [1; 2]. Цель такого анализа – выявление различных закономерностей, которые помогли бы прогнозировать показатели успеваемости и повышать качество обучения. В основном применяются различные методы математической статистики. В ра-

боте Бодрякова В.Ю., Торопова А.П. и Фоминой Н.Г. [3] проведен статистический анализ динамики успеваемости для выявления «проблемных зон качества» образовательного процесса и формирования рекомендаций по его улучшению.

Применение двухфакторного дисперсионного анализа [4; 5] позволяет установить, что оценки разных обучающихся (экзаменов) не принадлежат одной генеральной совокупности. Использование соответствующих средних значений не дает полной характеристики процесса оценивания. Обучающийся характеризуется собственным «смещением» оценки в положительную или отрицательную сторону. Это справедливо и в отношении экзаменов, каждый из которых использует принципы измерения, отличные от применяемых на других экзаменах.

Ряд исследований посвящен регрессионным моделям, оценивающим влияние различных факторов на оценки. Применяются линейные модели [6] на основе ретроспективных данных о показателях обучения в школе и вузе. В работах Галимовой Н.С. и Загитовой Л.Р. [7], а также Лебедевой Т.В., Цыпина А.П. и Сидоренко В.С. [8] используют такие факторы, как баллы ЕГЭ, средний балл аттестата, наличие специального профессионального образования, домашний регион обучающегося, пол, источник финансирования учебной деятельности. В исследовании Канапухина П.А., Коротких В.В. и Щекунских С.С. [9] отмечается отсутствие статистически значимых линейных связей между вступительными баллами и успеваемостью обучающихся. В работе Шафоростовой Е.Н. и Лазаревой Т.И. [10] для прогноза успеваемости предлагается использовать оценки по тестам интеллекта и мотивации. Отмечено влияние перечисленных факторов, но не выявлены линейные связи такого влияния. Кроме классических регрессионных моделей используются нейросетевые модели [11; 12], предсказывающие успеваемость по наблюдаемым оценкам и набору социальных факторов. Все перечисленные исследования отмечают влияние личностных характеристик обучающихся на оценки успеваемости.

В исследовании Егоровой Е.С. и Поповой Н.А. [12] рассматривается неоднородность различных подгрупп студентов и строятся модели классификации этих подгрупп для прогнозирования академических результатов. Исходные данные включают социально-демографические (возраст, пол, социально-экономический статус) и академические данные, собранные во время набора студентов в вуз, а также данные об активности, полученные с помощью электронной

информационно-образовательной среды. В работе Сосницкой В.Н. и Потанина Н.И. [13] проверяется наличие функциональной зависимости между средней успеваемостью и вероятностными характеристиками получения каждой из возможных оценок и выдвигается гипотеза о возможном существовании двух психологических характеристик, присущих каждому студенту и определяющих распределение вероятностей его оценок. В данных работах влияние личностных характеристик студента на успеваемость учитывается в виде принадлежности его определенному классу или кластеру.

Среди факторов, влияющих на успеваемость, все чаще рассматриваются характеристики личности, названные в [14] академическими способностями, в которые объединяют когнитивные, психомоторные, коммуникативные, организаторские, волевые способности, для их измерения используются различные психологические тесты.

В анализе данных успеваемости используются в основном характеристики обучающихся. Очевидно, что значительное влияние на оценки оказывают характеристики экзаменатора и технология оценивания знаний. Учет факторов, связанных с экзаменом, должен существенно повысить качество анализа успеваемости. Для более точного описания процесса оценивания знаний модель должна использовать и характеристики обучающихся, и особенности экзаменов.

Цель данной работы заключается в изучении влияния на оценки характеристик обучающихся и особенностей экзаменов. Для этого предлагается использовать модель вероятностного распределения экзаменационной оценки в виде функции Softmax, в определение которой входят латентные параметры, характеризующие уровни подготовленности обучающихся и трудности для каждой градации экзаменационной оценки. Для оценки параметров необходимо построить алгоритм их вычисления по наблюдениям оценок. Адекватность модели должна проверяться статистическими критериями.

Материалы и методы исследования

На экзаменационную оценку влияют, с одной стороны, факторы, связанные с обучающимся, с другой – факторы, связанные с экзаменом, такие как требовательность экзаменатора и сложность вопросов и заданий. Наиболее значимые факторы, такие как академические способности обучающихся, затруднительно измерить непосредственно. Поэтому и факторы обучающихся, и факторы экзаменов предлагается характеризовать латентными переменными. Та-

кой подход давно и успешно применяется в Item Response Theory (IRT) [15, с. 10-18] для обработки данных тестирования. В IRT вероятность

$$P(\theta, \delta) = \frac{e^{\theta}}{e^{\theta} + e^{\delta}}$$

правильного ответа на тестовое задание зависит от «трудности» задания, характеризуемого параметром δ , и уровня «подготовленности» тестируемого, задаваемого параметром θ . Так как эти переменные не наблюдаются непосредственно, то их называют латентными. Вероятность правильного ответа увеличивается при возрастании подготовленности и уменьшается при увеличении трудности.

Оценки обучающихся измеряются в политомической шкале с определенным количеством градаций. В России эта шкала включает следующие градации: «неудовлетворительно», «удовлетворительно», «хорошо» и «отлично». В работе Родионова А.В. [16] для оценки факторов, влияющих на оценки, использовалась Partial Credit Model [17, с. 48-50]. Модель использует латентные параметры трудности для каждой градации оценки, а подготовленность обучающегося описывается одним параметром. Анализ оценок указывает на неравномерность распределения частот для разных градаций, поэтому следует ожидать повышения точности модели при использовании параметра подготовленности для каждой градации оценки.

Для описания вероятностного распределения оценки, измеряемой в политомической шкале, предлагается использовать функцию Softmax [18]. Эта функция широко используется для описания распределений классов, определяемых нейронными сетями. Исходными данными являются наблюдения x_{ij} – оценки обучающихся, i – номер обучающегося, j – номер экзамена. Будем считать наблюдения реализациями случайных величин X_{ij} , независимых в совокупности. Для описания соответствующих распределений вероятностей используем функцию Softmax

$$P(X_{ij} = k) = \frac{\exp(\theta_{ik} - \delta_{jk})}{\sum_{h=1}^K \exp(\theta_{ih} - \delta_{jh})}, \quad (1)$$

где $k = 1, \dots, K$ – номер градации оценки в политомической шкале,

K – количество градаций в шкале,

θ_{ik} – подготовленность i -го обучающегося добиться k -й оценки,

δ_{jk} – трудность добиться k -й оценки на j -м экзамене.

Введем индикатор I_{ijk} , который равен 1, если $x_{ij} = k$, и 0 – в других случаях. В этих обозначениях

$$p_{ijk} = P(X_{ij} = k) = \frac{(\exp(\theta_{ik} - \delta_{jk}))^{I_{ijk}}}{\sum_{h=1}^K \exp(\theta_{ih} - \delta_{jh})}.$$

Логарифм функции максимального правдоподобия

$$L = \sum_{i,j} \left(I_{ijk} (\theta_{ik} - \delta_{jk}) - \ln \left(\sum_{h=1}^K \exp(\theta_{ih} - \delta_{jh}) \right) \right)$$

позволяет получить уравнения для вычисления параметров, приравнивая частные производные к нулю

$$\frac{\partial L}{\partial \theta_{ik}} = \sum_j \left(I_{ijk} - \frac{\exp(\theta_{ik} - \delta_{jk})}{\sum_{h=1}^K \exp(\theta_{ih} - \delta_{jh})} \right) = \sum_j (I_{ijk} - p_{ijk}) = 0, \quad (2)$$

$$\frac{\partial L}{\partial \delta_{jk}} = \sum_i \left(-I_{ijk} + \frac{\exp(\theta_{ik} - \delta_{jk})}{\sum_{h=1}^K \exp(\theta_{ih} - \delta_{jh})} \right) = \sum_i (-I_{ijk} + p_{ijk}) = 0. \quad (3)$$

Пределы суммирования не указаны, так как обучающиеся могут иметь оценки не по всем дисциплинам. Суммирование по i, j подразумевает использование всех оценок, суммирование по i – использование оценок i -го обучающегося, суммирование по j – использование оценок j -го экзамена.

Для итерационного поиска решения таких уравнений в IRT [15, с. 47-50] используется метод касательных

$$x^{(l+1)} = x^{(l)} - \frac{f(x^{(l)})}{f'(x^{(l)})}. \quad (4)$$

Для вычислительных процедур в качестве $f(x)$ используются левые части уравнений (2) и (3), а в качестве $f'(x)$ следующие выражения

$$\frac{\partial^2 l}{\partial \theta_{ik}^2} = - \sum_j \left(\frac{\exp(\theta_{ik} - \delta_{jk})}{\sum_{h=1}^K \exp(\theta_{ih} - \delta_{jh})} \left(1 - \frac{\exp(\theta_{ik} - \delta_{jk})}{\sum_{h=1}^K \exp(\theta_{ih} - \delta_{jh})} \right) \right) = - \sum_j (p_{ijk} (1 - p_{ijk})), \quad (5)$$

$$\frac{\partial^2 l}{\partial \delta_{jk}^2} = - \sum_i \left(\frac{\exp(\theta_{ik} - \delta_{jk})}{\sum_{h=1}^K \exp(\theta_{ih} - \delta_{jh})} \left(1 - \frac{\exp(\theta_{ik} - \delta_{jk})}{\sum_{h=1}^K \exp(\theta_{ih} - \delta_{jh})} \right) \right) = - \sum_i (p_{ijk} (1 - p_{ijk})). \quad (6)$$

В предлагаемом подходе смещение всех латентных параметров на одну и ту же величину не изменяет распределение вероятностей. Таким образом, вычислительная процедура сходится к одному набору значений из возможных.

Для определения начальных значений используем выводы работы [19], в которой показано, что при применении байесовского подхода оценка $x_{ij} = k$ приводит к увеличению $\exp(\theta_{ik})$ на 1 и к увеличению $\exp(\delta_{jk})$ на 1 в противоположном случае. Прямое применение байесовского подхода затруднено, так как начальные значения

латентных параметров неизвестны и определяются с точностью до некоторого смещения. Возможна ситуация, когда большое количество оценок одного обучающегося приведет к тому, что уровень его подготовленности будет существенно больше уровня подготовленности другого обучающегося с небольшим количеством оценок. Аналогичная ситуация характерна для трудностей оценок экзаменов. Для начальных значений латентных параметров такие неточности байесовского оценивания не играют существенной роли, поэтому положим

$$\frac{1}{m_i} \sum_j I_{ijk} = \exp(\theta_{ik}^{(0)}) \quad \text{или} \quad \theta_{ik}^{(0)} = \ln \left(\frac{1}{m_i} \sum_j I_{ijk} \right), \quad (7)$$

где m_i – количество оценок i -го обучающегося,

$$\frac{1}{n_j} \sum_i (1 - I_{ijk}) = \exp(\delta_{jk}^{(0)}) \quad \text{или} \quad \delta_{jk}^{(0)} = \ln \left(\frac{1}{n_j} \sum_i (1 - I_{ijk}) \right), \quad (8)$$

где n_j – количество оценок j -го экзамена.

Усреднение введено, чтобы компенсировать разницу в количествах оценок у разных студентов (экзаменов).

Особенностью вычислительной процедуры является возможность получения нулевого значения под знаком логарифма или бесконечного роста латентного параметра, если соответствующее значение усредненной суммы равняется единице. Такая ситуация возникает, если все оценки одинаковы или отсутствуют оценки по некоторой градации. В таких случаях латентные параметры ограничивают некоторыми минимальными и максимальными значениями

(обычно, предлагается -5 и +5).

Неравенство Рао – Крамера позволяет найти нижние границы дисперсий полученных оценок латентных параметров. Для этого используется информационная матрица Фишера. Для предложенной модели характерно большое количество параметров. Для простоты предположим, что все параметры, кроме θ_{ik} , известны и совпадают с решением уравнений (2) и (3). В этом случае количество информации Фишера для параметра подготовленности обучающегося

$$\begin{aligned} J(\theta_{ik}; \{I_{ijk}\}) &= M \left[\left(\frac{\partial L}{\partial \theta_{ik}} \right)^2 \right] = M \left[\left(\sum_j (I_{ijk} - p_{ijk}) \right)^2 \right] = \\ &= \sum_j M \left[(I_{ijk} - p_{ijk})^2 \right] + \sum_{j \neq h} M \left[(I_{ijk} - p_{ijk})(I_{ihk} - p_{ihk}) \right] = \sum_j D[I_{ijk}] \end{aligned}$$

совпадает с суммой дисперсий индикаторов соответствующей оценки, так как вторая сумма математических ожиданий будет равна 0, поскольку индикаторы $\{I_{ijk}\}$ независимы в совокупности и математическое ожидание произведения будет равно произведению математических ожиданий, каждое из которых равно нулю. Для параметра трудности δ_{jk} количество информации будет равно сумме $\sum_i D[I_{ijk}]$ дисперсий индикаторов оценки k на j -м экзамене. Применяя неравенство Рао – Крамера, получаем неравенства

$$D[\hat{\theta}_{ik}] \geq \frac{1}{\sum_j D[I_{ijk}]},$$

$$D[\hat{\delta}_{jk}] \geq \frac{1}{\sum_i D[I_{ijk}]},$$

которые определяют нижнюю границу дисперсий соответствующих оценок параметров. Для вычисления дисперсий индикаторов $D[I_{ijk}] = p_{ijk}(1 - p_{ijk})$ используем оценки вероятностей $\hat{p}_{ijk}$, вычисленные по оценкам латентных параметров. В результате получаем неравенства

$$D[\hat{\theta}_{ik}] \geq \frac{1}{\sum_j \hat{p}_{ijk}(1 - \hat{p}_{ijk})},$$

$$D[\hat{\delta}_{jk}] \geq \frac{1}{\sum_i \hat{p}_{ijk}(1 - \hat{p}_{ijk})}.$$

Неравенства показывают, что с увеличением объема данных дисперсии полученных оценок латентных параметров уменьшаются.

Результаты исследования и их обсуждение

Для проверки адекватности модели обработан набор, включающий 2363 оценки. Изучались оценки для 55 экзаменов 163 студентов одного направления обучения. Обработка данных выполнялась по следующей схеме.

1. Из набора оценок исключались оценки экзаменов и обучающихся, для которых все оценки одинаковы. В таких случаях соответствующие латентные параметры принимают предельные значения (-20 или +20).

2. По формулам (7) и (8) были определены начальные значения латентных параметров.

3. В цикле до достижения требуемой погрешности решения уравнений (2) и (3) вычислялись суммы левых частей этих уравне-

ний, а также производные по формулам (5) и (6). По этим данным методом касательных (4) пересчитывались латентные параметры обучающихся и экзаменов.

4. Определялись статистические характеристики для проверки адекватности модели.

Применялся F-критерий для сравнения внутригрупповых дисперсий

$$M_2 = \frac{1}{n} \sum_{i=1}^n \frac{1}{m_i - 1} \sum_j (x_{ij} - \bar{x}_{i*})^2,$$

n – количество обучающихся, $\bar{x}_{i*} = \frac{1}{m_i} \sum_j x_{ij}$,

с дисперсией всего массива оценок

$$M_1 = \frac{1}{N - 1} \sum_{i=1}^n m_i (x_{i*} - \bar{x}_{**})^2,$$

N – общее количество оценок, $\bar{x}_{**} = \frac{1}{N} \sum_{i,j} x_{ij}$.

Если влияние факторов на оценки отсутствует, то M_1 и M_2 являются разными оценками дисперсии одной и той же случайной величины. Отношение $F = M_1 / M_2$ в этих условиях будет иметь распределение Фишера со степенями свободы $n - 1$ и $N - n$. Значение F-критерия с уровнем значимости близким к нулю показало несостоятельность этой гипотезы. Такой же результат был получен для групп оценок экзаменов. Это означает, что влияние и обучающихся, и экзаменов на оценки является статистически значимым.

Для проверки того, что предлагаемая модель корректно описывает зависимости законов распределения вероятностей оценок от обучающихся и экзаменов, изучались остатки $x'_{ij} = x_{ij} - M[X_{ij}]$. Математическое ожидание определялось по распределению вероятностей оценки (1), в котором использовались оценки латентных параметров. Гипотезы о принадлежности одной генеральной совокупности остатков оценок и для студентов, и для экзаменов подтвердились с вероятностью близкой к единице. Распределение остатков представлено на рисунке. Гистограмма демонстрирует унимодальность эмпирического распределения. Среднее остатков равно -0,0026. Эмпирическое распределение похоже на распределение Гаусса, хотя эта гипотеза не прошла проверку по критерию хи-квадрат.

Исследование зависимостей в явлениях, подверженных случайным воздействиям, требует оценки доли закономерностей, объясняемых моделью, и доли случайных отклонений. В случае с успеваемостью предлагаемая модель определяет закон распределения вероятностей оценки обучающегося на экзамене, в частности математическое ожидание оценки.



Гистограмма остатков по результатам обработки массива оценок

Насколько точно модель описывает результаты экзаменов, оценивает коэффициент детерминации

$$R^2 = 1 - \frac{D[X|\theta, \delta]}{D[X]} = 1 - \frac{\sum (x_{ij} - M[X_{ij}])^2}{\sum (x_{ij} - \bar{x}_{**})^2},$$

который использует сумму квадратов отклонений оценок от соответствующих математических ожиданий. Если бы оценки отклонялись от математического ожидания незначительно, то коэффициент детерминации был бы близок к единице. Однако это не происходит в силу большого количества случайных факторов, влияющих на процедуру оценива-

ния. Для анализируемого списка оценок коэффициент детерминации составил $R^2 = 0,5916$. Это говорит о том, что модель определяет оценку на 59%, а случайное влияние составляет 41%. Вполне возможно, к случайному влиянию будут отнесены закономерности, которые модель не учитывает. Математическое ожидание суммы квадратов отклонений

$$M\left[\sum (X_{ij} - M[X_{ij}])^2\right] = M\left[\sum (M[X_{ij}^2] - M[X_{ij}])^2\right] = \sum D[X_{ij}] = 2333,036$$

для обработанного набора данных довольно близко к наблюдаемому значению

$$\sum (x_{ij} - M[X_{ij}])^2 = 2241,251.$$

Это означает, что предложенная модель достаточно точно описывает детерминированную составляющую формирования оценки на экзамене.

Латентные параметры не могут заменить традиционные оценки. Они определяются в процессе итерационных вычислений с точностью до некоторого смещения, одинакового для всех параметров. Латентные параметры подготовленности можно применять для сравнения обучающихся, параметры трудности – для сравнения экзаменов. Сравнение можно применять только

в рамках общего массива оценок. Основное назначение латентных параметров заключается в определении законов распределения вероятностей оценок и в прогнозировании успеваемости на этой основе.

Распределения вероятностей экзаменов можно использовать для сопоставления шкал оценивания разных экзаменов [20]. На этой основе возможно введение некоторой стандартной шкалы с усредненными значениями латентных параметров. Применение такой шкалы позволило бы пересчитывать оценки, полученные на разных экзаменах, к оценкам в стандартной шкале и получать, таким образом, более точную картину успеваемости, в частности корректировать результаты экзаменов с завышенными или заниженными оценками.

Заключение

В работе предложена новая модель, которая описывает законы распределения вероятностей экзаменационных оценок в виде функции softmax с латентными параметрами подготовленности обучающихся и трудностей экзаменов. Модель отличается от аналогов [16; 17] применением функции softmax и использованием параметров для каждой градации оценок. Разработаны численные методы определения латентных параметров по набору оценок и методика проверки адекватности модели. Обработка репрезентативного набора оценок авторской программой продемонстрировала возможность применения предложенной модели для решения практических задач.

Преимуществом предложенной модели является более точное описание вероятностного распределения оценок. Практическое применение модели заключается в определении прогнозов оценок и показателей успеваемости по оценкам латентных параметров. Законы распределения оценок можно использовать для выявления экзаменов со значительными отклонениями вероятностей от средних значений для улучшения технологии оценивания.

Список литературы

1. Ширинкина Е.В. Методы интеллектуального анализа данных и образовательной аналитики // Современное образование. 2022. № 1. С. 51-67. URL: <https://cyberleninka.ru/article/n/metody-intellektualnogo-analiza-dannyh-i-obrazovatelnoy-analitiki> (дата обращения: 27.06.2025). DOI: 10.25136/2409-8736.2022.1.37582.
2. Попова Н.А., Егорова Е.С. Интеллектуальный анализ образовательных данных для прогноза успеваемости студентов вуза // Известия Кабардино-Балкарского научного центра РАН. 2023. № 2 (112). С. 18-29. URL: <https://cyberleninka.ru/article/n/intellektualnyy-analiz-obrazovatelnyh-dannyh-dlya-prognoza-uspevaemosti-studentov-vuza> (дата обращения: 27.06.2025). DOI: 10.35330/1991-6639-2023-2-112-18-29.
3. Бодряков В.Ю., Торопов А.П., Фомина Н.Г. Углубленный статистический анализ динамики успеваемости студентов – математиков при обучении в педагогическом вузе // Качество инновации образование. 2009. № 1. С. 6-11.
4. Братищенко В.В. Статистический анализ экзаменационных оценок // Известия Иркутской государственной экономической академии. 2011. № 3. URL: <https://cyberleninka.ru/article/n/statisticheskiy-analiz-ekzamenatsionnyh-otsenok> (дата обращения: 27.06.2025).
5. Гранков М.В., Аль-Габри В.М., Горлова М.Ю. Анализ и кластеризация основных факторов, влияющих на успеваемость учебных групп вуза // Инженерный вестник Дона. 2016. № 4 (43). URL: <https://cyberleninka.ru/article/n/analiz-i-klasterizatsiya-osnovnyh-faktorov-vliyayuschih-na-uspevaemost-uchebnyh-grupp-vuza> (дата обращения: 27.06.2025).
6. Харитонов И.М., Крушель Е.Г., Привалов О.О., Степанченко И.В., Степанченко О.В. Прогнозирование качества обучения в вузе с помощью методов регрессионного анализа // Известия СПбГТИ(ТУ). 2021. № 56 (82). С. 72-80. URL: <https://cyberleninka.ru/article/n/prognostirovanie-kachestva-obucheniya-v-vuze-s-pomoschyu-metodov-regressiionnogo-analiza> (дата обращения: 27.06.2025). DOI: 10.36807/1998-9849-2021-56-82-72-80.
7. Галимова Н.С., Загитова Л.Р. Построение многофакторной модели успеваемости студента // Международный научно-исследовательский журнал. 2020. № 6-3 (96). С. 31-36. URL: <https://cyberleninka.ru/article/n/postroenie-mnogofaktornoy-modeli-uspevaemosti-studenta> (дата обращения: 27.06.2025). DOI: 10.23670/IRJ.2020.96.6.080.
8. Лебедева Т.В., Цыпин А.П., Сидоренко В.С. Статистический анализ факторов, влияющих на успеваемость студентов российских вузов // Интеллект. Инновации. Инвестиции. 2016. № 9. С. 55-58. URL: <https://cyberleninka.ru/article/n/statisticheskiy-analiz-faktorov-vliyayuschih-na-uspevaemost-studentov-rossiyskih-vuzov> (дата обращения: 27.06.2025).
9. Канапухин П.А., Коротких В.В., Щекунских С.С. Статистический анализ факторов академической успеваемости (на примере обучающихся экономического факультета ФГБОУ ВО «ВГУ») // Вестник Воронежского государственного университета. Серия: Экономика и управление. 2020. № 2. С. 27-44. URL: <https://journals.vsu.ru/econ/article/download/2899/2954> (дата обращения: 27.06.2025). DOI: 10.17308/econ.2020.2/2899.
10. Шафоростова Е.Н., Лазарева Т.И. Проблемы внедрения информационной системы контроля качества обучения студентов // Вестник БГТУ имени В. Г. Шухова. 2010. № 3. С. 174-177. URL: <https://cyberleninka.ru/article/n/problemy-vnedreniya-informatsionnoy-sistemy-kontrolya-kachestva-obucheniya-studentov> (дата обращения: 27.06.2025).
11. Русаков С.В., Русакова О.Л., Посохина К.А. Нейросетевая модель прогнозирования группы риска по успеваемости студентов первого курса // Современные информационные технологии и ИТ-образование. 2018. № 4. С. 815-822. URL: <http://sitito.cs.msu.ru/index.php/SITITO/article/view/479> (дата обращения: 27.06.2025). DOI: 10.25559/SITITO.14.201804.815-822.
12. Егорова Е.С., Попова Н.А. Data Mining в образовании: прогнозирование успеваемости учащихся // Моделирование, оптимизация и информационные технологии. 2023. № 11 (2). URL: <https://www.moitvvt.ru/journal/article?id=1325> (дата обращения: 27.06.2025). DOI: 10.26102/2310-6018/2023.41.2.003.
13. Сосницкий В.Н., Потанин Н.И. Вероятностный подход к анализу успеваемости студентов // Фундаментальные исследования. 2014. № 8. С. 734-738. URL: <https://fundamental-research.ru/ru/article/view?id=34625> (дата обращения: 27.06.2025).
14. Аكوпова М.А. Исследование структуры академических способностей студентов педагогического вуза // Психология. Историко-критические обзоры и современные исследования. 2019. Т. 8. № 4А. С. 114-123. URL: <http://publishing-vak.ru/file/archive-psychology-2019-4/12-akopova.pdf> (дата обращения: 27.06.2025). DOI: 10.34670/AR.2019.44.4.042.
15. Нейман Ю.М., Хлебников В.А. Введение в теорию моделирования и параметризации педагогических тестов. М.: Прометей, 2000. 168 с. ISBN: 5-7042-1068-6.
16. Родионов А.В. Модификация рейтинговой параметрической модели оценки латентных факторов для измерения уровня сформированности компетенций // Известия Иркутской государственной экономической академии. 2014. № 6 (98). С. 168-174. URL: <https://cyberleninka.ru/article/n/modifikatsiya-reytingovoy-parametricheskoy-modeli-otsenki-latentnyh-faktorov-dlya-izmereniya-urovnya-sformirovannosti-kompetentsiy> (дата обращения: 27.06.2025). DOI: 10.17150/1993-3541.2014.24(6).168-174.
17. Wright B.D., Masters G.N. Rating scale analysis: Rasch measurement. Chicago, Illinois: MESA Press, 1982. 223 p. ISBN: 0-941938-01-8.
18. Ian Goodfellow, Yoshua Bengio, and Aaron Courville: Deep learning: The MIT Press, 2016. P. 180-184. ISBN: 0262035618. URL: <https://www.deeplearningbook.org/contents/mlp.html#pf15> (дата обращения: 27.06.2025).
19. Братищенко В.В. Пошаговая модель с латентными параметрами для оценок студентов // System Analysis & Mathematical Modeling. 2021. Т. 8. № 3. С. 188-198. URL: <https://samm-bgu.ru/reader/article.aspx?id=24766> (дата обращения: 27.06.2025). DOI: 10.17150/2713-1734.2021.8(3).188-198.
20. Братищенко В.В. Сравнение оценок учащихся на основе вероятностных моделей // Наука. Информатизация. Технологии. Образование: материалы XVI международной научно-практической конференции (Екатеринбург, 27 февраля – 3 марта 2023 г.). Екатеринбург: Издательство РГППУ, 2023. С. 28-34. URL: https://elar.rsvpu.ru/bitstream/123456789/42761/1/978-5-8295-0699-5_2023_006.pdf (дата обращения: 27.08.2025).