

КЛАСТЕРНЫЙ АНАЛИЗ ДАННЫХ БАЛЛЬНО-РЕЙТИНГОВОЙ СИСТЕМЫ (НА ПРИМЕРЕ ДИСЦИПЛИНЫ «МАТЕМАТИКА»)

М.Е. Вайндорф-Сысоева¹, Н.Ю. Фаткуллин², В.Ф. Шамшович²

¹Московский педагогический государственный университет, г. Москва,

²Уфимский государственный нефтяной технический университет, г. Уфа

Балльно-рейтинговая система – широко распространенный и достаточно эффективный инструмент оценки знаний обучающихся. В то же время анализ массива данных рейтинга весьма трудоемок в силу их объема. Обработка данных стандартными статистическими методами зачастую приводит к выведению усредненных показателей, что не отвечает концепции индивидуального подхода к обучающимся. Особую актуальность приобретает реализация индивидуального подхода в виде дифференциации по степени мотивации множества (кластера) обучающихся с низкой успеваемостью, особенно на младших курсах, при сложной демографической ситуации, обуславливающей проблемность набора и необходимость сохранения контингента обучающихся.

В данной работе ставится задача внутренней дифференциации обучающихся в выявленных множествах (подгруппах) исходной учебной группы по степени успешности (величине рейтингового балла) методами кластерного анализа.

В исследовании выдвигается гипотеза: данные балльно-рейтинговой системы в динамике своего формирования несут скрытую информацию о наличии (или отсутствии) тенденций по трансформации кластеров обучающихся. Выявление таких тенденций позволит в дальнейшем определить степень стабильности подобных кластеров, как характеристики мотивированности обучающихся и степени успешности проводимого процесса обучения.

По результатам проведенного кластерного анализа установлено: при сохранении четкой кластеризации на «успешных» и «малоуспешных» обучающихся скрыто присутствует достаточно стабильный кластер «перехода», трудно выявляемый стандартными статистическими методами анализа; кластер «малоуспешных» обучающихся в свою очередь обладает собственной структурой кластеров, аналогичной структуре кластеров полной исходной учебной группы.

С точки зрения теории педагогики данный факт служит подтверждением выдвинутой гипотезы о наличии динамики в формально сформировавшихся кластерах и выявлении потенциала для увеличения контингента «успешных» обучающихся. С практической точки зрения, именно выявленный состав кластера «перехода» – первоочередной контингент обучающихся для работы преподавателей по переводу их на следующую ступень успешности обучения.

Ключевые слова: балльно-рейтинговая система, рейтинговый балл, кластерный анализ, кластеры, тенденции, дифференциация обучающихся.

Состояние проблемы. Вопросы адекватной оценки учебных достижений обучающихся и эффективное построение учебного процесса всегда находились в числе приоритетных. Одним из наиболее удачных и относительно простых в реализации вариантов решения данных проблем, на наш взгляд, является балльно-рейтинговая система (БРС). Данная система может реализовываться в разных форматах и иметь разные алгоритмы вычисления итоговой и промежуточных оценок обучающихся. В качестве примера в кратком изложении приведем действующую модель

БРС по кафедре математики УГНТУ [10, с. 63; 13, с. 255–256].

Балльно-рейтинговая система оценки успеваемости студентов предусматривает возможность для студентов, выполняющих в срок и на требуемом уровне все виды учебной работы по дисциплине «Математика», получить оценку по ней автоматически.

Семестровый рейтинг студентов по дисциплине «Математика» оценивается суммой в 100 максимально возможных баллов и включает:

а) текущий рейтинг;

б) рейтинг по итогам промежуточной аттестации.

Студентам, обучающимся по индивидуальным графикам, а также пропустившим аудиторные занятия по уважительной причине, подтвержденной документально, предоставляется возможность выполнить отчетные работы в дни и часы, установленные кафедрой.

Студентам, набравшим 56 и более баллов за семестр, предоставляется возможность получить оценку по промежуточной аттестации автоматически.

Студентам, набравшим менее 56 баллов за семестр, предоставляется возможность добрать недостающие баллы в процессе промежуточной аттестации в сроки, установленные кафедрой.

Текущий рейтинг студента определяется в рубежных (контрольных) точках согласно приложениям (табл. 1–2).

За выполнение работы в объеме 86–100 % студент получает 5 баллов, выполнение работы в объеме 71–85 % – 4 балла, выполнение работы в объеме 56–70 % – 3 балла.

В УГНГУ курс математики для бакалавриата по унифицированной программе изучается в 1-2 семестрах. Оценка за каждый семестр устанавливается на основании «Положения о рейтинговой оценке знаний студентов на кафедре математики». Согласно этому Положению необходимо выполнить отчетные работы (расчетные задания (РЗ), лабораторные работы (ЛР), аттестационные тестирования (АТ)) в установленный срок и в необходимом объеме по основным разделам курса.

Таблица 1

Таблица БРС за 1 семестр

№	Содержание модуля	Отчетная работа	Примечания	Количество баллов
				максимальное
1	Линейная и векторная алгебра. Аналитическая геометрия	Расчетные задания № 1 (РЗ_1)	Раздел № 1 «Линейная и векторная алгебра». УМК / Материалы для самостоятельной работы (РЗ, с. 95–106); Раздел № 2 «Аналитическая геометрия». УМК / Материалы для самостоятельной работы (РЗ, с. 98–112)	10
		Лабораторная работа № 1 (ЛР_1) «Решение систем линейных уравнений методом Гаусса»	Раздел № 1 «Линейная и векторная алгебра». УМК / Материалы для самостоятельной работы (ЛР, с. 107–117)/	5
		Аттестационное тестирование № 1 (АТ_1)	Разделы № 1 «Линейная и векторная алгебра», № 2 «Аналитическая геометрия». УМК (КИМ)	20
2	Введение в математический анализ. Дифференциальное исчисление функции одной переменной	Расчетные задания № 2 (РЗ_2)	Раздел № 3 «Введение в математический анализ». УМК / Материалы для самостоятельной работы (РЗ, с. 116–139); Раздел № 4 «Дифференциальное исчисление функции одной переменной». УМК	10
		Аттестационное тестирование № 2 (АТ_2)	Разделы № 3 «Введение в математический анализ», № 4 «Дифференциальное исчисление функции одной переменной»	20
3	Дифференциальное исчисление функции нескольких переменных	Расчетные задания № 3 (РЗ_3)	Раздел № 1 «Линейная и векторная алгебра». УМК / Материалы для самостоятельной работы (РЗ, с. 100–116)	10
		Лабораторная работа № 2 (ЛР_2) «Метод наименьших квадратов»	Раздел № 5 «Дифференциальное исчисление функции нескольких переменных». УМК / Материалы для самостоятельной работы (ЛР, с. 117–129)/	5
		Аттестационное тестирование № 2 (АТ_3)	Раздел № 5 «Дифференциальное исчисление функции нескольких переменных». УМК (КИМ)	20

Таблица БРС за 2 семестр

№ модуля	Содержание модуля	Отчетная работа	Примечания	Количество баллов
				максимальное
4	Элементы теории функций комплексного переменного. Интегральное исчисление функции одной переменной	Расчетные задания № 4 (РЗ_4)	Раздел № 11 «Теория функции комплексного переменного». УМК / Материалы для самостоятельной работы (РЗ, с. 153–156); Раздел № 6 «Интегральное исчисление функции одной переменной». УМК	10
		Лабораторная работа № 3 (ЛР_3) «Приближенное вычисление определенных интегралов по формулам прямоугольника, трапеций и Симпсона»	Раздел № 6 «Интегральное исчисление функции одной переменной». УМК / Материалы для самостоятельной работы (ЛР, с. 202–205)	5
		Аттестационное тестирование № 4 (АТ_4)	Разделы № 11 «Теория функции комплексного переменного», № 6 «Интегральное исчисление функции одной переменной». УМК (КИМ)	20
5	Дифференциальные уравнения	Расчетные задания № 5 (РЗ_5)	Раздел № 9 «Дифференциальные уравнения». УМК / Материалы для самостоятельной работы (РЗ, с. 251–277)	10
		Аттестационное тестирование № 5 (АТ_5)	Разделы № 9 «Дифференциальные уравнения». УМК (КИМ)	20
6	Теория вероятностей. Элементы дискретной математики. Математическая статистика	Расчетные задания № 6 (РЗ_6)	Раздел № 13 «Теория вероятностей». УМК / Материалы для самостоятельной работы (РЗ, с. 191–218)	10
		Лабораторная работа № 4 (ЛР_4) «Расчет параметров корреляционной зависимости»	Раздел № 14 «Математическая статистика». УМК / Материалы для самостоятельной работы (ЛР, с. 79–96)	5
		Аттестационное тестирование № 6 (АТ_6)	Разделы № 13 «Теория вероятностей», № 14 «Математическая статистика». УМК (КИМ)	20

В качестве методического обеспечения кафедры математики УГНТУ разработала учебно-методические комплексы (УМК) и контрольно-измерительные материалы (КИМ) по каждому разделу. УМК содержат всю необходимую информацию для полного освоения курса и состоят из следующих элементов: теоретические основы, методические основы, материалы для самостоятельной работы, контрольно-измерительные материалы. КИМ представляют собой банк заданий различной сложности в тестовой форме. По данным заданиям проводятся аттестационные тестирования знаний обучающихся преподавателем. В результате по итогам двух семестров преподаватель имеет таблицу рейтинговых баллов подобную табл. 3.

Постановка задачи. Вопросы расчета рейтинга на основе данных БРС не относятся к классу сложных математических вычислений. В то же время вопрос выбора подходящей модели расчета может стать дискуссионным. На практике тривиальная задача расчета итогового рейтинга решается без затруднений, особенно при использовании табличных процессоров типа Excel, Lotus и т. п. Данная задача не вызовет затруднений даже в случае расчета рейтинга по 2 семестрам в целом или по отдельности. В то же время задача выявления тенденций в мотивации обучающихся есть та сложная задача, решение которой на несколько порядков сложнее арифметического расчета рейтинга [1, с. 74].

Таблица 3

Пример данных БРС 1-2 семестра для учебной группы бакалавров 1 курса

№ студента	Р31	АТ1	Р32	АТ2	Р33	АТ3	Р34	АТ4	Р35(1)	Р35(2)	АТ5	Р36	АТ6(1)	АТ6(2)
1	10,0	18,0	8,0	4,5	9,0	14,0	7,0	14,0	1,0	3,2	12,0	10,0	6,0	9,0
2	10,0	16,0	8,0	13,0	10,0	6,0	7,0	14,0	3,0	4,8	16,0	8,0	10,0	7,0
3	0,0	6,0	4,0	1,5	5,0	12,0	5,0	12,0	1,5	1,6	8,0	8,0	5,0	9,0
4	7,0	18,0	8,0	16,0	0,0	14,0	7,0	14,0	4,0	0,0	14,0	8,0	6,0	7,0
5	6,0	18,0	9,0	17,0	8,0	18,0	5,0	14,0	1,0	2,4	10,0	10,0	7,0	8,0
6	7,0	18,0	7,0	10,0	9,0	16,0	6,0	8,0	0,0	0,8	10,0	5,0	7,0	8,0
7	5,0	10,0	9,0	13,0	2,0	8,0	7,0	20,0	3,0	7,2	16,0	10,0	10,0	8,0
8	9,0	14,0	8,0	18,5	8,0	20,0	7,0	18,0	2,0	4,8	17,0	10,0	10,0	10,0
9	7,0	18,0	10,0	18,5	8,0	20,0	7,0	14,0	3,0	0,8	14,0	3,0	8,0	5,0
10	2,0	8,0	1,0	13,5	2,0	10,0	2,0	6,0	1,0	4,8	12,0	3,0	0,0	5,0
11	6,0	14,0	8,0	10,0	0,0	11,0	7,0	10,0	3,0	3,2	10,0	10,0	9,5	9,0
12	9,0	16,0	7,0	17,0	7,0	14,0	7,0	12,0	2,5	2,4	16,0	7,0	8,5	10,0
13	9,0	17,0	10,0	20,0	10,0	16,0	7,0	16,0	2,0	4,0	18,0	10,0	9,5	6,0
14	10,0	12,0	8,0	18,5	7,0	20,0	6,0	14,0	3,0	2,4	14,0	7,0	9,5	9,0
15	2,0	12,0	5,0	10,0	3,0	12,0	5,0	14,0	1,0	0,8	10,7	2,0	5,0	9,0
16	2,0	8,0	4,0	7,5	6,0	13,0	5,0	8,0	1,0	3,2	10,0	2,0	5,0	9,0
17	6,0	14,0	2,0	1,5	7,0	4,0	4,0	6,0	0,0	0,0	6,0	2,0	1,0	3,0
18	6,0	4,0	5,0	0,0	2,0	3,0	2,0	4,0	0,0	0,8	10,0	3,0	7,0	4,0
19	8,0	18,0	9,0	13,0	10,0	16,0	7,0	14,0	3,0	5,6	16,0	9,0	8,0	10,0
20	7,0	14,0	10,0	20,0	10,0	20,0	6,0	18,0	2,0	5,6	16,0	10,0	10,0	10,0
21	9,0	18,0	8,0	14,5	9,0	12,0	7,0	16,0	4,0	4,0	6,0	7,0	9,0	7,0
22	5,0	10,0	4,0	11,5	7,0	0,0	5,0	4,0	2,0	0,8	10,0	5,0	9,0	6,0
23	7,0	10,0	4,0	7,5	5,0	8,0	5,0	6,0	2,0	2,4	4,0	4,0	4,0	7,0
24	5,0	14,0	8,0	12,5	7,0	16,0	7,0	18,0	2,0	0,8	19,0	10,0	9,0	8,0
25	6,0	6,0	5,0	1,5	4,0	6,0	3,0	1,0	1,0	3,2	8,0	7,0	5,0	4,0
26	7,0	4,0	8,0	13,0	9,0	12,0	6,0	9,0	3,0	0,8	4,0	9,0	7,0	6,0
27	10,0	18,0	10,0	18,5	10,0	17,0	7,0	18,0	2,0	4,0	18,0	10,0	7,5	8,0
28	3,0	6,0	3,0	10,0	7,0	4,0	5,0	4,0	0,0	0,8	4,0	7,0	7,0	6,0
29	4,0	4,0	7,0	6,0	0,0	0,0	0,0	4,0	0,0	1,6	6,0	3,0	4,0	6,0

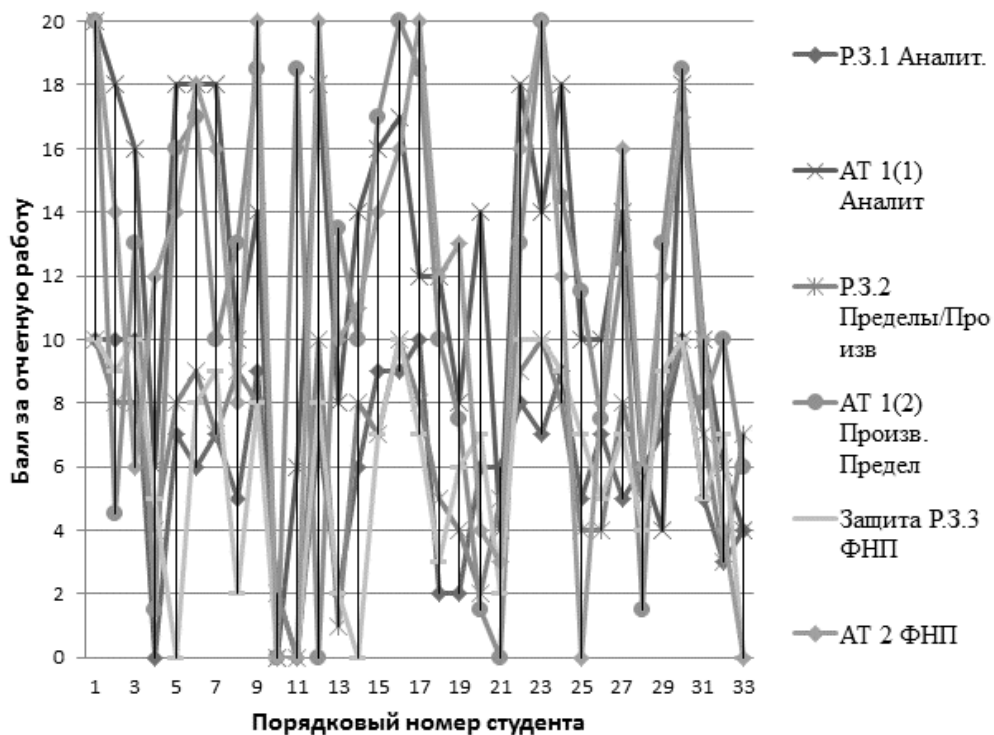


Рис. 1. Визуализация индивидуальных данных БРС по 6 отчетным работам за 1 семестр

Стандартный подход к анализу данных состоит в визуализации данных, построении линейных и нелинейных трендов и применения стандартного пакета описательной статистики. На рис. 1 приведены графики по данным БРС для 33 студентов за 1 семестр всего лишь по 6 отчетным работам.

Построение трендов (линейных и нелинейных) в данном случае лишь усложнит понимание и трактование исходных данных. Следующий этап – вычисление стандартных статистических характеристик для каждого из обучающихся – математическое ожидание ($M(x)$) и дисперсия ($D(x)$) [3, с. 76, 87]. Их применение, на наш взгляд, более оправдано именно в индивидуальном порядке для обучающихся, так как групповые значения лишь скрывают индивидуальные характеристики, так необходимые при определении степени мотивации каждого учащегося и выработки адресных корректирующих воздействий.

Не приводя таблиц расчетов, визуализируем изменения $M(x)$ и $D(x)$ по данным табл. 3 в соответствии с ранжированными значениями итогового рейтингового балла (рис. 2). Очевидно, что данная визуализация мало ин-

формативна и данный шаг вновь не привел к прояснению ситуации.

Таким образом, в подобной ситуации педагог-исследователь приходит к необходимости получения инструментария по обработке значительных массивов информации с целью выявления возможных скрытых тенденций.

Экспериментальное исследование.

В данной работе нами выдвигалась следующая гипотеза: Данные БРС в динамике своего формирования несут скрытую информацию о наличии (или отсутствии) тенденций по трансформации образований (кластеров) обучающихся. Выявление таких тенденций позволяет в дальнейшем определить степень устойчивости подобных кластеров как характеристики мотивированности обучающихся и степени успешности проводимого процесса обучения.

В качестве следующего метода обработки данных нами был выбран кластерный анализ, который, как известно, решает задачу разделения исходного множества элементов на подмножества (кластеры) исходя из меры близости элементов друг другу. Кластерный анализ как инструмент исследования в педа-

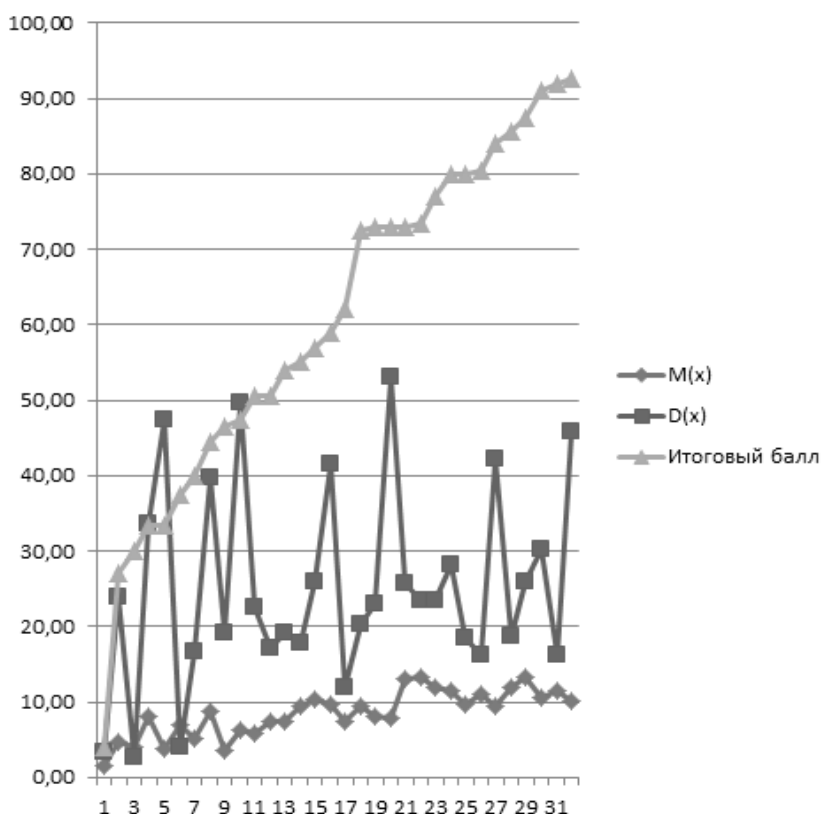


Рис. 2. Сопоставление отсортированных индивидуальных данных БРС, их математических ожиданий и дисперсий

гогике и психологии [5, 6] применяется достаточно эффективно, но несколько ограниченно, в силу требования наличия у исследователя определенных компетенций и навыков [7, с. 274].

Существует множество процедур кластеризации [7, с. 276; 14; 15] и множество программных реализаций их алгоритмов. В данном исследовании был использован лицензионный пакет STATISTICA 8 StatSoft.Inc. [13].

В качестве процедуры кластеризации выбор был остановлен на классическом методе k-средних в силу простоты использования, быстроты расчетов (что критично при больших массивах данных) и отсутствии выбросов в исходной выборке (исключение искажений в расчетах).

Отметим, что тривиальная задача разбиения группы учащихся на подгруппы по успешности обучения не ставилась, так как таблицы рейтингов (табл. 1) при ранжировании результатов решают эту задачу [1, с. 82–83]. Задача исследования заключалась в выявлении скрытых кластеров «переходов» от стадии «малоуспешности» к «успешности» в обучении. На первом этапе исследования был произведено разбиение на 2 кластера (класте-

ры «малоуспешности» и «успешности») по «методу k-средних» (k-means). Результатом явилась дифференциация исходного множества обучающихся (группы бакалавров) на 2 кластера (кластер 1 – «малоуспешные», кластер 2 – «успешные») (рис. 3).

Причем отметим важнейшую характеристику достоверности результатов – значительное различие кластеров и отсутствие пересечений. Данная характеристика указывает на присутствующие значимые различия в подгруппах (кластерах). Кроме того, согласно теории кластерного анализа, рекомендуемого производить процедуру разбиения различными способами, данная процедура была повторена, но с применением квадрата евклидовой метрики (рис. 4).

Получение картины четкой кластеризации по двум процедурам указывает на достоверность полученных результатов. В то же время получение первого тривиального результата не может служить основой для проверки выдвигаемой гипотезы. Поэтому, в дальнейшем нами была предпринята процедура разбиения той же группы на 3 кластера (рис. 5, 6) с целью выявления факта существования кластера «перехода».

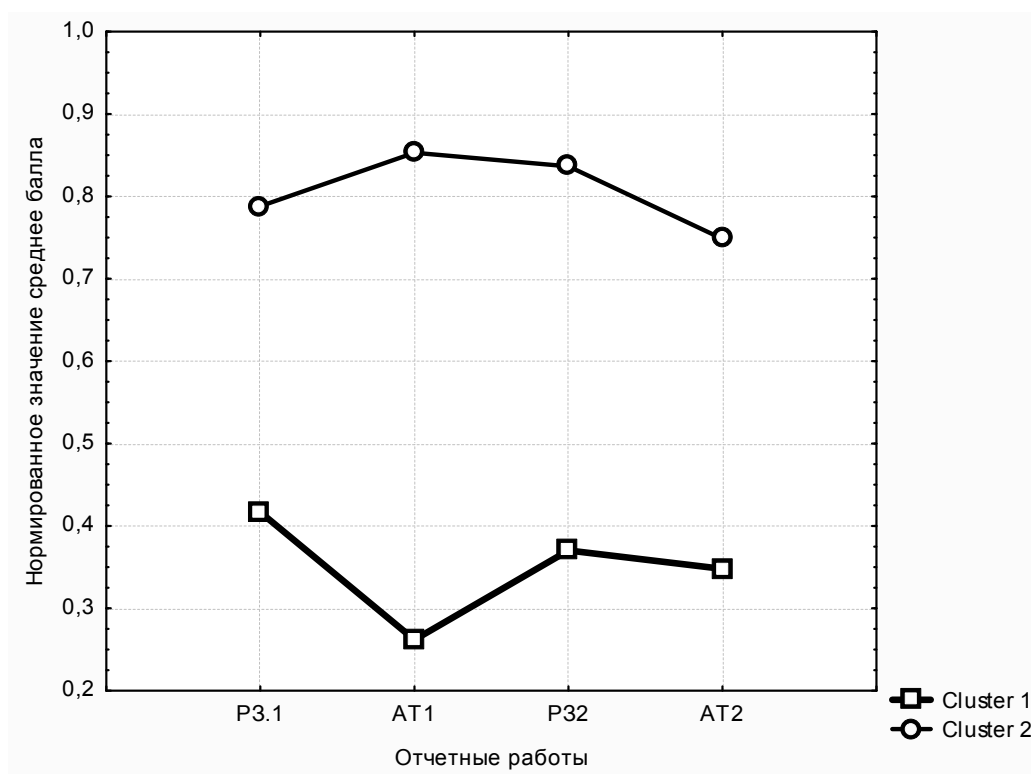


Рис. 3. Разбиение группы на 2 кластера по 4 отчетным работам 1 семестра

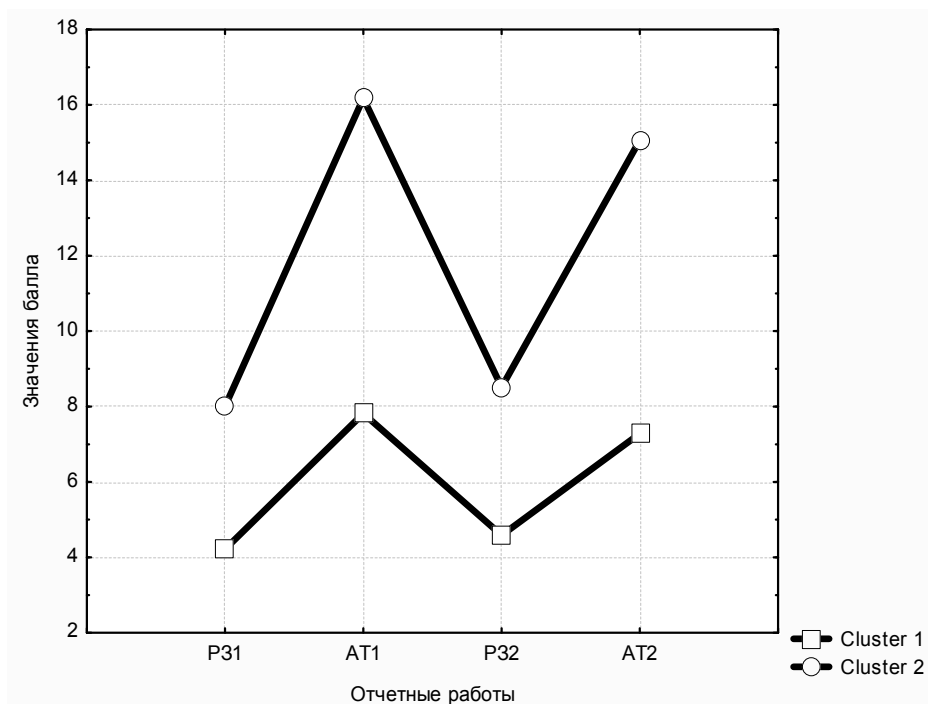


Рис. 4. Разбиение группы на 2 кластера по 4 отчетным работам 1 семестра (квадрат евклидовой метрики)

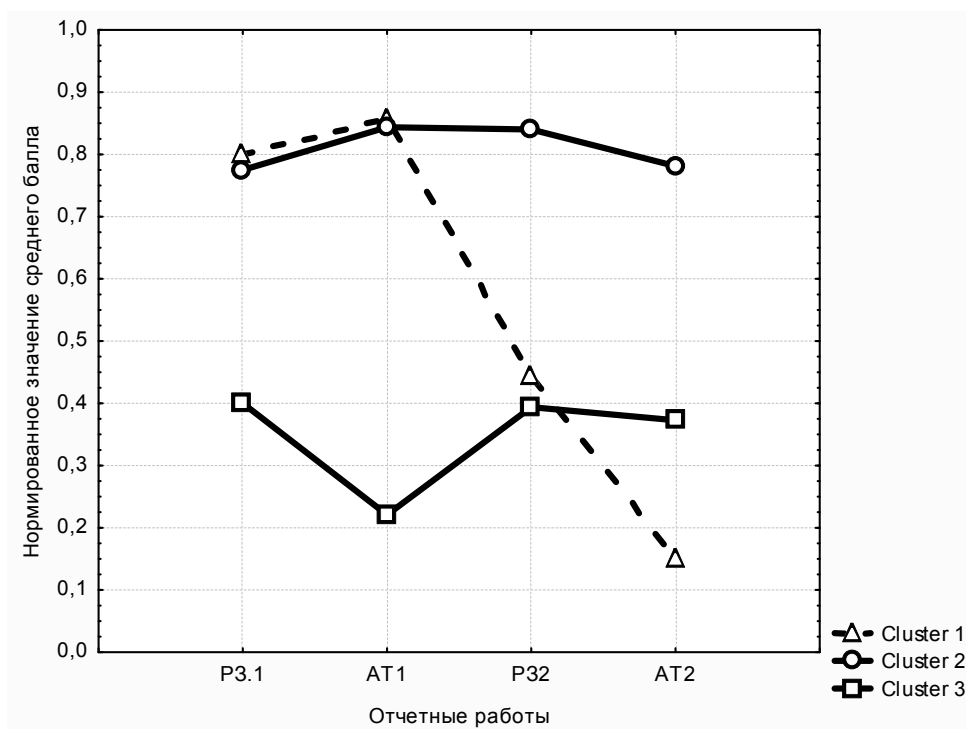


Рис. 5. Разбиение группы на 3 кластера по 4 отчетным работам 1 семестра

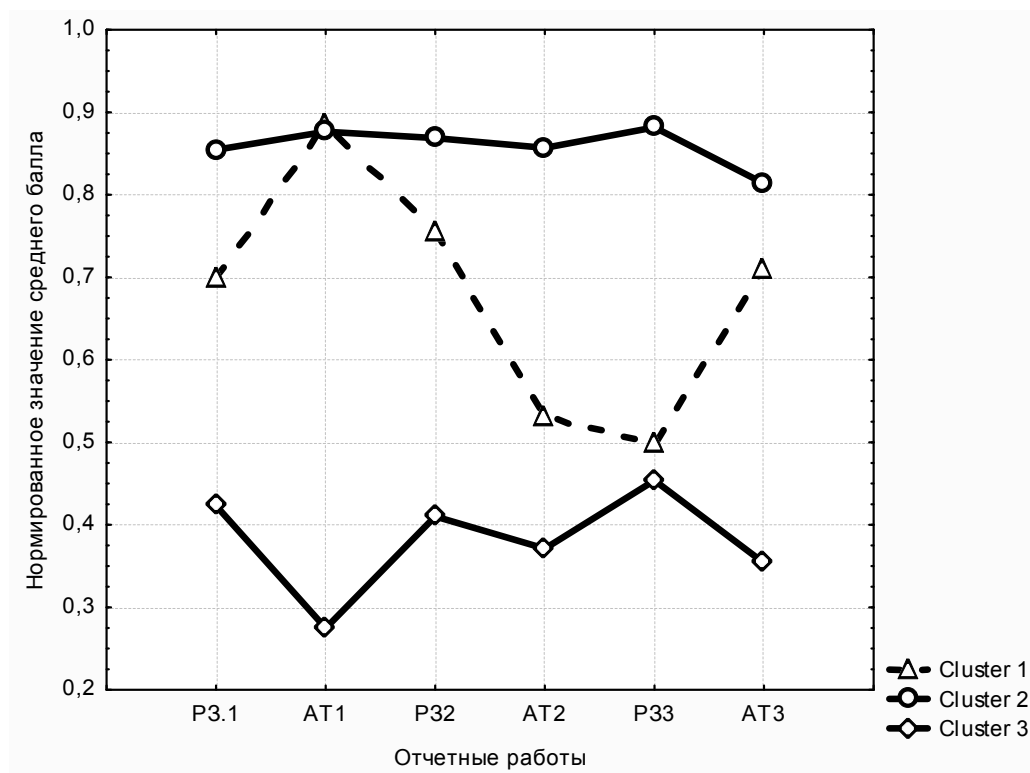


Рис. 6. Разбиение группы на 3 кластера по 6 отчетным работам 1 семестра

Результат новой кластеризации одновременно указал на 2 факта:

1. Сохранение четкой кластеризации на «успешных» (оценки 5 и 4) (45 %) и «малоуспешных» (оценки от 3 и ниже) (48 %);

2. Появление кластера «перехода» (7 %).

Третий кластер является существенно меньшим по объему, но значимость его в том, что он служит индикатором наличия скрытых тенденций в формировании первых двух кластеров.

На следующем этапе произведем увеличение числа отчетных работ до 6 и повторим процедуру кластеризации на 2 и 3 подгруппы. Вторая из процедур снова выявила третий кластер «перехода», который увеличился более чем в 2 раза (17 %) и до этого почти поровну поглощался двумя исходными. Таким образом, удалось выявить контингент учащихся, который не выявлялся другими методами и являлся более нестабильным по сравнению с другими. Очевидно, что именно он представляет потенциальный интерес для педагога как контингент, который обладает достаточными знаниями для непопадания в ряд «малоуспешных» обучающихся.

Таким образом, нами рассмотрены результаты учебного процесса группы бакалавров в 1 семестре. Продолжая исследование по данным БРС для той же группы по отчетным работам 2 семестра (P34-AT6), мы установили ту же тенденцию – наличие устойчивого по объему кластера перехода (рис. 7–10).

Далее проведем кластеризацию по сквозным данным БРС за годовой курс обучения по дисциплине «Математика». Для краткости приведем лишь визуальные итоги разбиения на 3 кластера (рис. 11–13). Во всех случаях объем кластера перехода остался стабильным (15–17 % состава), что свидетельствует об устойчивости данной группы.

Следующим этапом исследования явился кластерный анализ подгруппы «малоуспешных» обучающихся. Гипотезой исследования послужило следующее предположение: Множество «малоуспешных» обучающихся в свою очередь обладает структурой кластеров аналогичной структуре кластеров полной исходной учебной группы. Предполагалось, что кластерный анализ способен дифференцировать данное подмножество обучающихся на 2 или 3 подмножества.

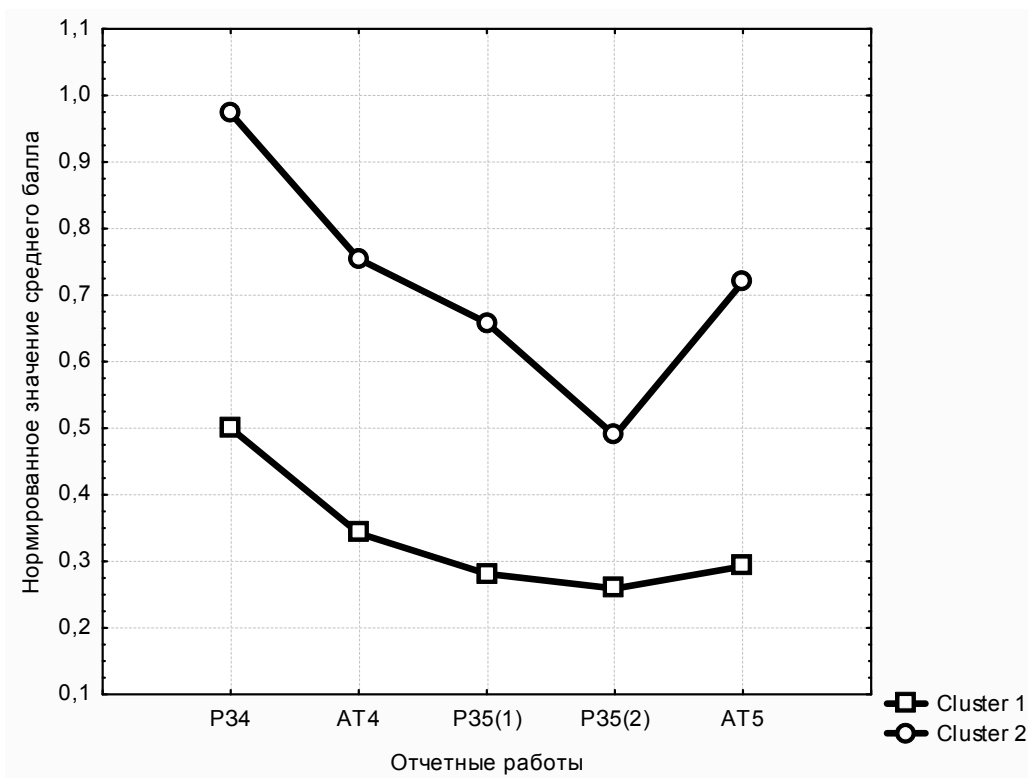


Рис. 7. Разбиение группы на 2 кластера по 4 отчетным работам 2 семестра

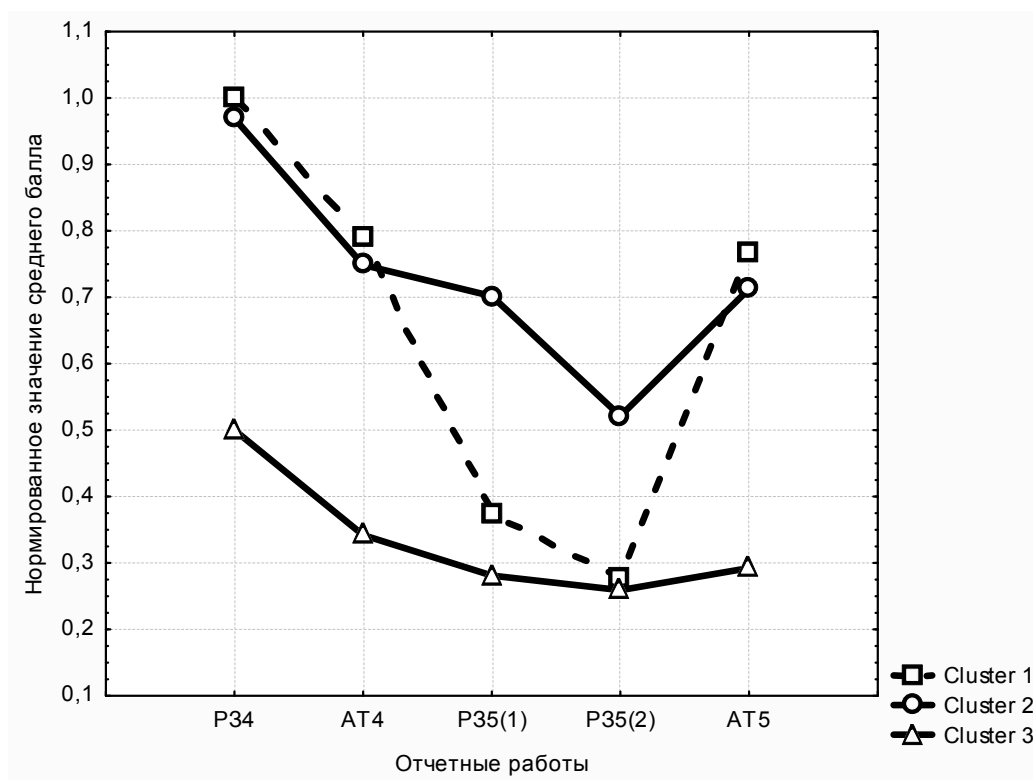


Рис. 8. Разбиение группы на 3 кластера по 4 отчетным работам 2 семестра

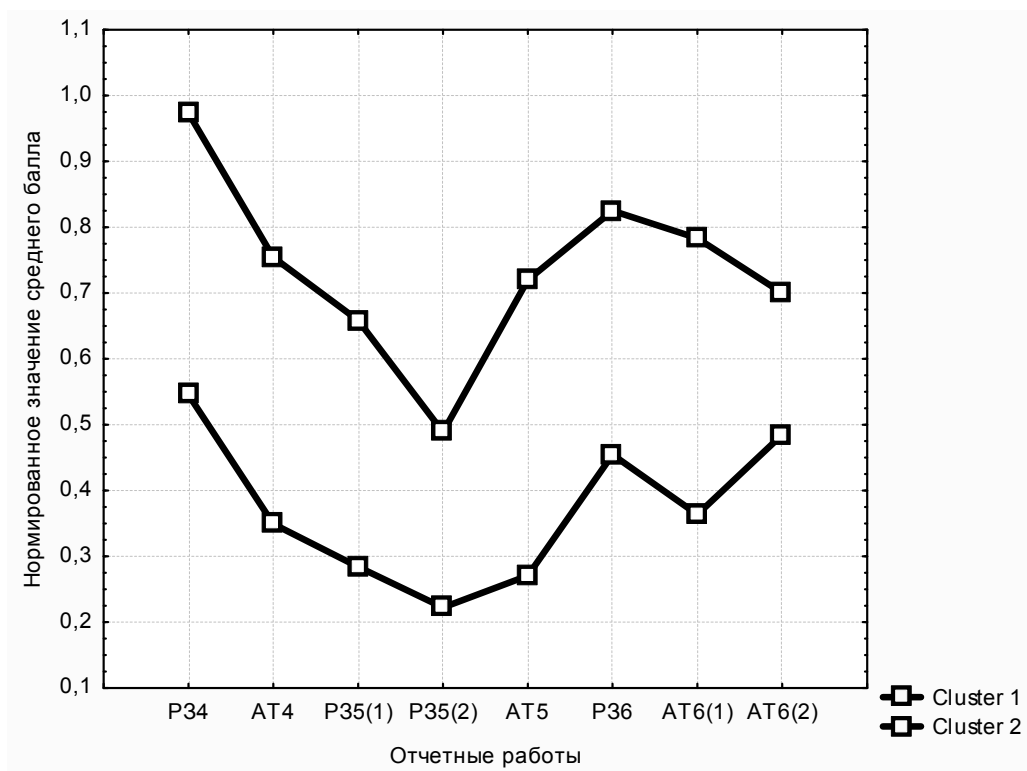


Рис. 9. Разбиение группы на 2 кластера по 6 отчетным работам 2 семестра

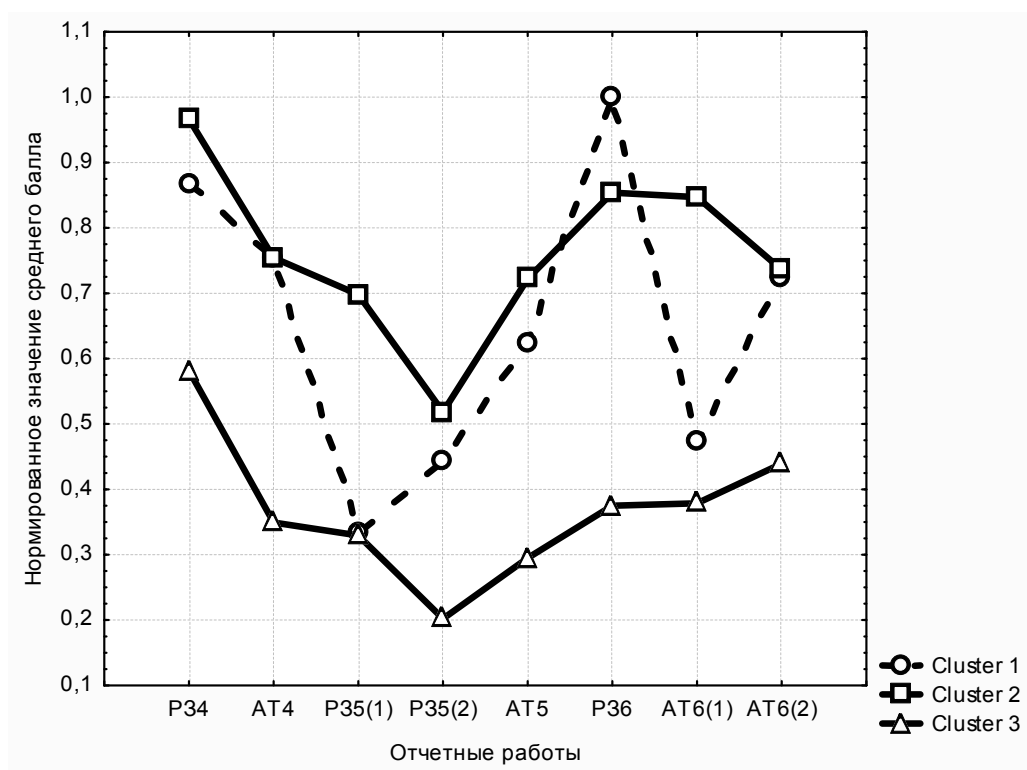


Рис. 10. Разбиение группы на 3 кластера по 6 отчетным работам 2 семестра

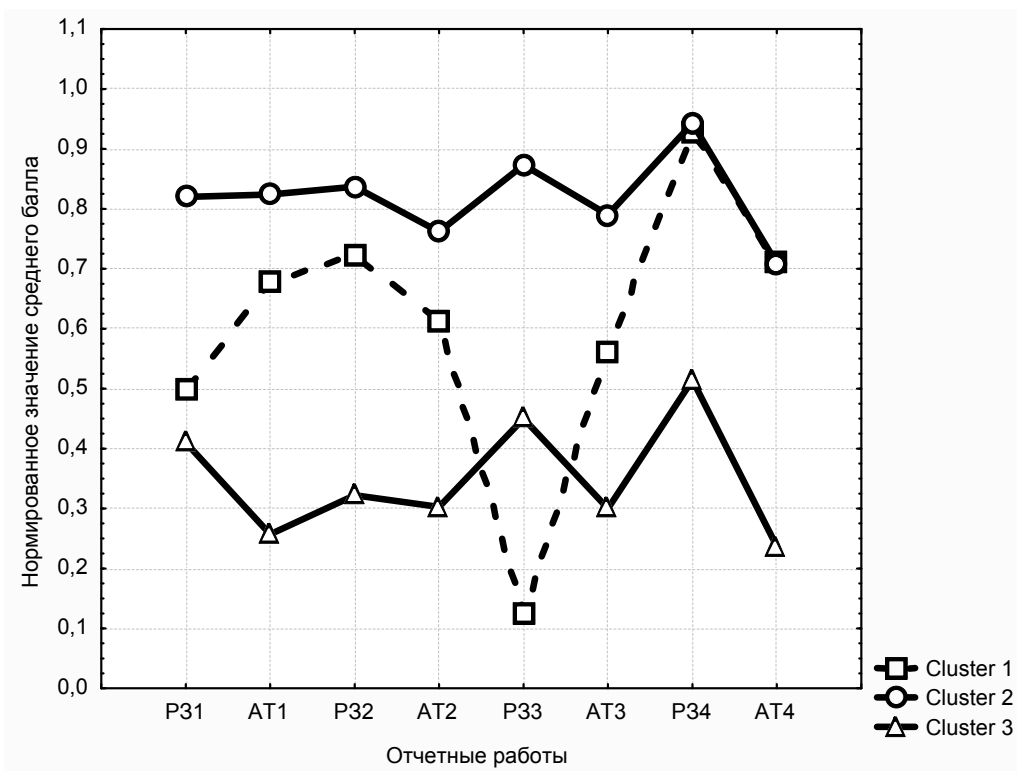


Рис. 11. Разбиение группы на 3 кластера по 8 отчетным работам за год

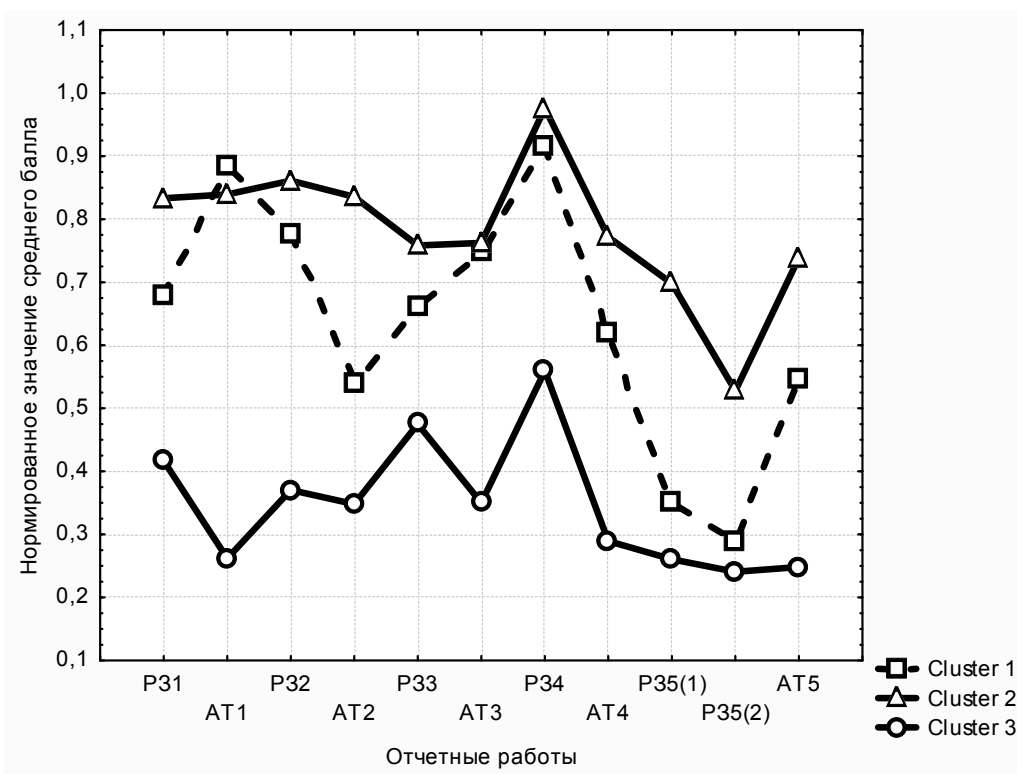


Рис. 12. Разбиение группы на 3 кластера по 10 отчетным работам за год

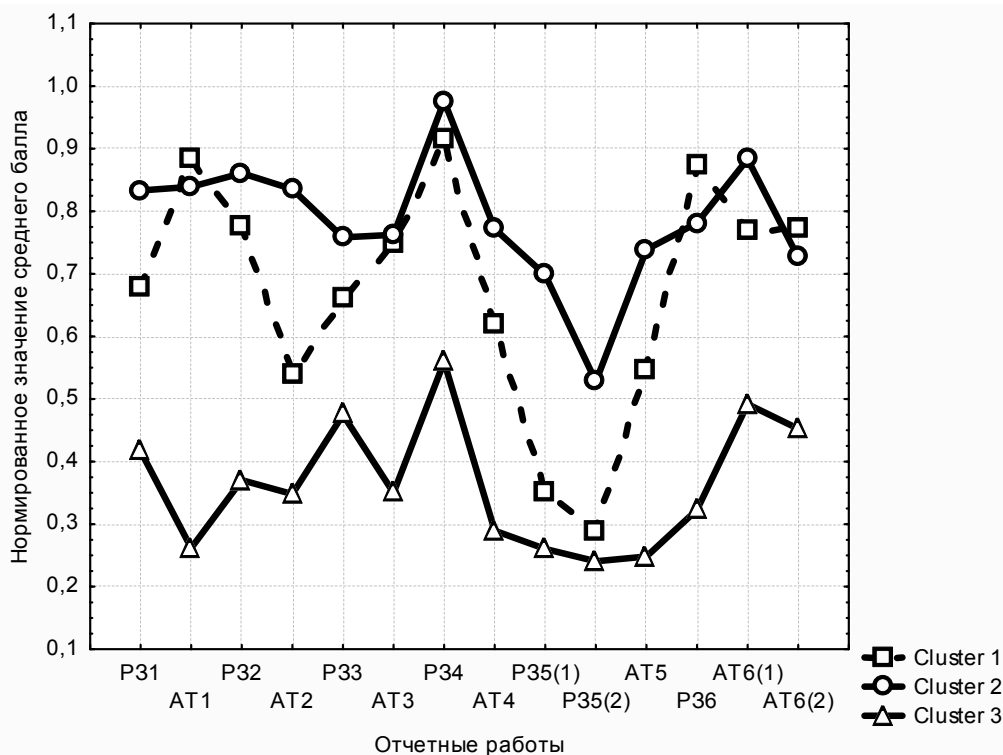


Рис. 13. Разбиение группы на 3 кластера по 12 отчетным работам за год

Результаты исследования показали следующее (рис. 14–15):

1. Разбиение на 2 кластера создало 2 пересекающихся подмножества, что сильно усложняет трактовку результатов;

2. Разбиение на 3 кластера дало четкую дифференциацию на обучающихся:

а) способных к выходу из кластера «малоуспешных» в кластер «успешных» (34 %);

б) практически неспособных к такому переходу (32 %);

в) составляющих кластер «перехода», т. е. обучающихся, способных оторваться от самого нижнего кластера (34 %).

Очевидно, что с точки зрения перспектив успешности обучения первый и третий кластеры есть первоочередной контингент для работы преподавателя [9, с. 137–138].

Работа со вторым кластером является предметом работы психологов, кураторов и других специалистов [4, с. 75–76; 8, с. 44–45; 11]. Для поддержки подобных студентов в УГНТУ проводятся дополнительные занятия по дисциплине, как по вузовской программе, так и по элементам довузовской подготовки, а также консультации, в том числе в дистанционном формате [2, с. 82–83].

Отметим, что аналогичные результаты (с незначительными вариациями) были полу-

чены и по другим учебным группам бакалавров 1 курса обучения.

Заключение. По результатам проведенных исследований установлено:

1. При сохранении четкой кластеризации на «успешных» и «малоуспешных» обучающихся скрыто присутствует достаточно стабильный кластер «перехода», трудно выявляемый стандартными статистическими методами анализа.

2. Кластер «малоуспешных» обучающихся в свою очередь обладает собственной структурой кластеров аналогичной структуре кластеров полной исходной учебной группы.

3. Состав кластеров «перехода» – первоочередной контингент обучающихся для работы преподавателей для перевода их на следующую ступень успешности обучения.

Таким образом, применение методов кластерного анализа, в отличие от стандартных средств математической статистики, оказалось достаточно эффективным при проверке выдвинутой педагогической гипотезы о наличии динамического кластера «перехода» в составе множества «малоуспешных» обучающихся. Результаты проверки (подтверждение) гипотезы могут быть использованы при разработке методики по работе с данным кластером обучающихся.

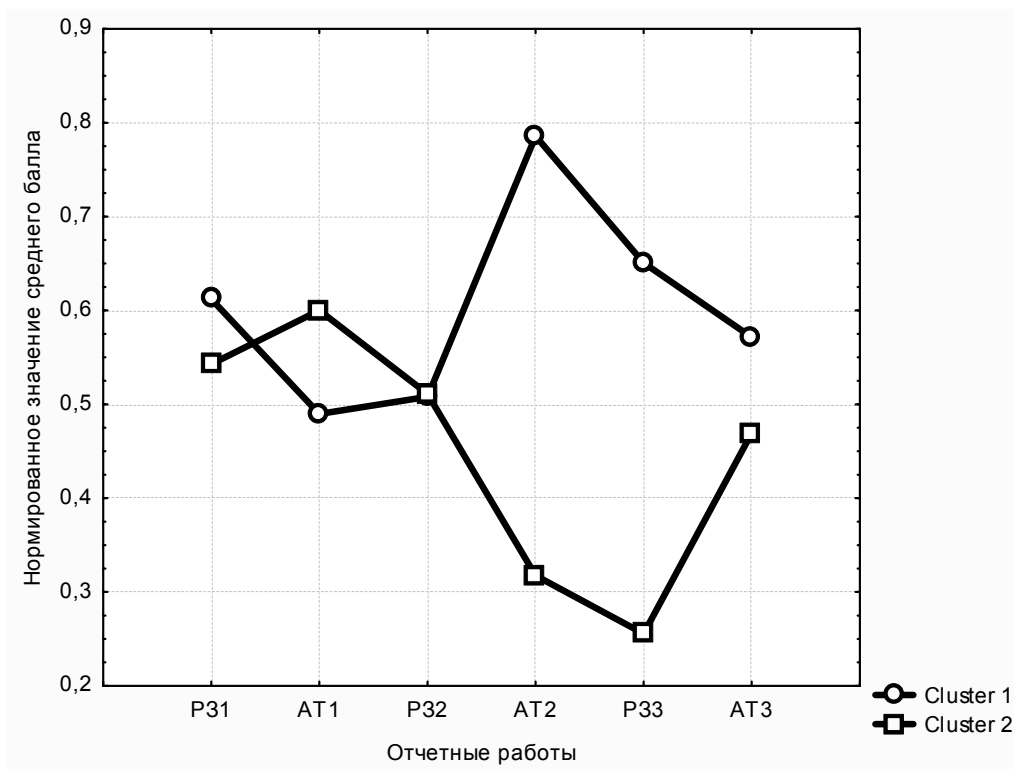


Рис. 14. Разбиение кластера «малоуспешных» обучающихся на 2 кластера по отчетным работам 1 семестра

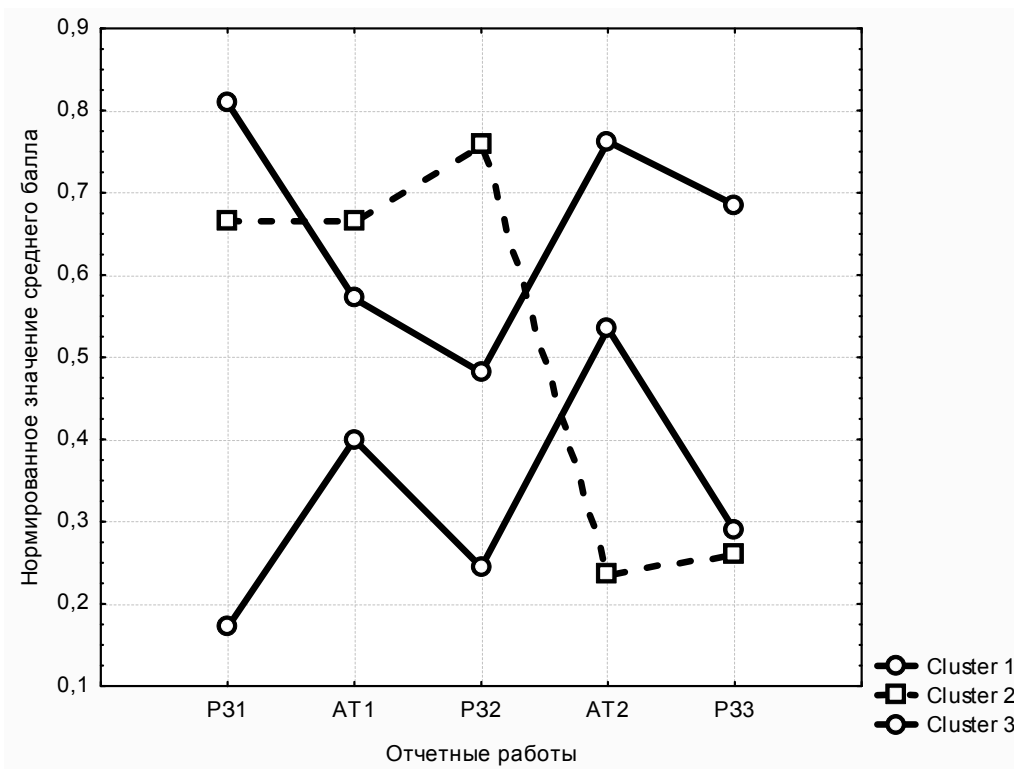


Рис. 15. Разбиение кластера «малоуспешных» обучающихся на 3 кластера по отчетным работам 1 семестра

Литература

1. Вайндорф-Сысоева, М.Е. Дифференциация студентов по признаку изменения рейтинговых достижений на основе анализа трендов прогнозных данных / М.Е. Вайндорф-Сысоева, Н.Ю. Фаткуллин, В.Ф. Шамшович // Вестник ЮУрГУ. Серия «Образование. Педагогические науки». – 2014. – Т. 6, № 2. – С. 74–83.
2. Вайндорф-Сысоева, М.Е. Проверка педагогической гипотезы о повышении рейтинговых характеристик студентов при введении в учебный процесс консультационных занятий в дистанционном формате / М.Е. Вайндорф-Сысоева, Н.Ю. Фаткуллин, В.Ф. Шамшович // Вестник ЮУрГУ. Серия «Образование. Педагогические науки». – 2014. – Т. 6, № 3. – С. 82–88.
3. Гмурман, В.Е. Теория вероятностей и математическая статистика / В.Е. Гмурман. – М.: Юрайт, 2012. – 480 с.
4. Григорьева, Т.В. Психолого-педагогический анализ состояния готовности студентов-первокурсников к учебно-познавательной деятельности в техническом вузе / Т.В. Григорьева, Г.М. Мифтахова // Вестник МГОУ. Серия «Педагогика». – 2009. – № 4. – С. 74–79.
5. Доронин, А.М. Кластерный анализ в структуре информатизации дидактического процесса // Успехи современного естествознания / А.М. Доронин, Д.А. Романов, М.Л. Романова. – 2010. – № 9. – С. 157–159. – <http://www.natural-sciences.ru/ru/article/view?id=8855> (дата обращения: 01.02.2016).
6. Дроздова, С.О. Применение кластерного анализа для оценки качества образования университета Дубна / С.О. Дроздова, Е.А. Пахомова // Устойчивое развитие: наука и практика: междунар. электрон. журн. – 2014. – Вып. 1 (12). – С. 15. – <http://www.yrazvitie.ru/wp-content/uploads/2014/06/15-Rakhomova.pdf> (дата обращения: 01.02.2016).
7. Дьячук, А.А. Математические методы в психологических и педагогических исследованиях: учеб. пособие / А.А. Дьячук. – Красноярск: Изд-во Краснояр. гос. пед. ун-т им. В.П. Астафьева, 2013. – 347 с.
8. Сакович, С.М. Роль информационных и коммуникационных технологий в обеспечении качества и доступности высшего образования / С.М. Сакович // Вестник МГОУ. Серия «Педагогика». – 2010. – № 3. – С. 40–45.
9. Солдатенков, Р.М. Математическая подготовка студентов высших учебных заведений и направления её преобразования / Р.М. Солдатенков // Вестник МГОУ. Серия «Педагогика». – 2010. – № 1. – С. 136–141.
10. Фаткуллин, Н.Ю. Балльно-рейтинговая система оценки знаний студентов по математике с использованием информационно-коммуникационных технологий в ГОУ ВПО «Уфимский государственный нефтяной технический университет» / Н.Ю. Фаткуллин, В.Ф. Шамшович, Р.Н. Бахтизин // Сборник тезисов докладов конференции: материалы четвертой международной конференции по вопросам обучения с применением технологий e-learning «Moscoweducationonline 2010», Москва, гостиничный комплекс «Альфа», 29 сентября – 1 октября 2010 г. – М.: ООО «Globalconferances», 2010. – С. 62–64.
11. Фаткуллин, Н.Ю. Практическая реализация методов дистанционного обучения на основе информационно-коммуникационных технологий при балльно-рейтинговой системе оценки знаний студентов / Н.Ю. Фаткуллин, В.Ф. Шамшович, Р.Н. Бахтизин // Тез. докл. науч.-метод. семинара «Дистанционные технологии в учебном процессе», Иркутск, 28–29 апреля 2010 г. – Иркутск: Изд-во ИГТУ, 2010. – С. 46–47.
12. Халафян, А.А. STATISTICA 6. Статистический анализ данных: учеб. – Изд. 2-е. – М.: Бином, 2009. – 528 с.
13. Шамшович, В.Ф. Внедрение балльно-рейтинговой системы оценки знаний студентов с использованием информационно-коммуникационных технологий и поведение процедур мониторинга и прогнозирования оценки успеваемости студентов по математике методами нейросетевых технологий / Н.Ю. Фаткуллин, В.Ф. Шамшович, Р.Н. Бахтизин // Материалы второй всероссийской научно-практической конференции (Казань, 16–22 апреля 2010 года) «Электронная Казань 2010» / редкол.: К.Н. Пономарев (пред.) и др. – Казань: ЮНИВЕРСУМ, 2010. – С. 255–258.
14. Xu R., Wunsch Ii D.C. Clustering. – New York: Wiley and Sons, 2009.
15. Zhao Y., Karypis G., Fayyad U. Hierarchical clustering algorithms for document datasets // Data mining and knowledge discovery. – 2005. – № 10 (2). – С. 141–168.

Вайндорф-Сысоева Марина Ефимовна, кандидат педагогических наук, профессор кафедры технологии и профессионального обучения, Московский педагогический государственный университет, mageva@yandex.ru.

Фаткуллин Николай Юрьевич, кандидат экономических наук, доцент, доцент кафедры математики, Уфимский государственный нефтяной технический университет, nick_idpo@mail.ru.

Шамшович Валентина Федоровна, кандидат экономических наук, доцент кафедры математики, Уфимский государственный нефтяной технический университет, shamshovich@mail.ru.

Поступила в редакцию 3 февраля 2016 г.

DOI: 10.14529/ped160209

CLUSTER ANALYSIS OF DATA OF SCORE RATING SYSTEM (BASED ON THE SUBJECT OF “MATHEMATICS”)

M.E. Vayndorf-Sysoeva¹, mageva@yandex.ru,

N.Yu. Fatkullin², nick_idpo@mail.ru,

V.Ph. Shamshovich², shamshovich@mail.ru

¹*Moscow Pedagogical State University, Moscow, Russian Federation,*

²*Ufa State Petroleum Technological University, Ufa, Russian Federation*

Score rating system is considered a rather effective assessment tool. At the same time, the analysis of a data file of a rating is very labor-consuming. Data processing by standard statistical methods often leads to the removal of average indicators that do not meet the requirement of an individual approach to students. Individual approach is of particular importance in differentiating clusters of students with low progress on a motivation scale at junior university courses in conditions of difficult demographic situation when there is a need to preserve the number of students.

The aim of the paper is to differentiate the revealed clusters of initial group of students on a success degree (credit score-rating) by cluster analysis.

The hypothesis is formulated: the score-rating data in dynamics of its formation has hidden information on the existence (or absence) of tendencies on transformation of the students' clusters. The identification of such tendencies will allow us further to define the degree of stability of the similar clusters as characteristics of students' motivation and degree of success of the training process.

The results show that when the division into the clusters of “successful” and “unsuccessful” students is preserved, we can observe a rather stable cluster of “transition” students, which is difficult to reveal by standard statistical methods of the analysis. The cluster of “unsuccessful” students in its turn possesses its own structure of clusters similar to the structure of clusters of initial group.

In terms of Pedagogy, this fact can confirm the hypothesis about the existence of dynamics in formally created clusters and identify the potential increase in a number of “successful” students. The revealed structure of a “transition” cluster is the target audience for the teachers to help these students to improve their results and transfer from the cluster of “unsuccessful” to the cluster of “successful” students.

Keywords: score-rating system, rating point, cluster analysis, clusters, tendency, differentiation of students.

References

1. Vayndorf-Sysoeva M.E., Fatkullin N.Yu., Shamshovich V.F. [Differentiation of Students on the Basis of Change of Rating Achievements on the Basis of the Analysis of Trends of Expected Data]. *Bulletin of the South Ural State University. Ser. Education. Pedagogical Sciences*, 2014, vol. 6, no. 2, pp. 74–83. (in Russ.)

2. Vayndorf-Sysoeva M.E., Fatkullin N.Yu., Shamshovich V.F. [Check of a Pedagogical Hypothesis of Increase of Rating Characteristics of Students at Introduction to Educational Process of Con-

sulting Occupations in a Remote Format]. *Bulletin of the South Ural State University. Ser. Education. Pedagogical Sciences*, 2014, vol. 6, no. 3, pp. 82–88. (in Russ.)

3. Gmurman V.E. *Teoriya veroyatnostey i matematicheskaya statistika* [Probability Theory and Mathematical Statistics]. Moscow, Yurayt Publ., 2012. 480 p.

4. Grigor'eva T.V., Miftakhova G.M. [The Psychology and Pedagogical Analysis of a Ready State of First-year Students to Educational Cognitive Activity in Technical College]. *Bulletin of the Moscow State Regional University. Ser. Pedagogics*, 2009, no. 4, pp. 74–79. (in Russ.)

5. Doronin A.M., Romanov D.A., Romanova M.L. [The Cluster Analysis in Structure of Informatization of Didactic Process]. *Achievements of Modern Natural Sciences*, 2010, vol. 9, pp. 157–159. (in Russ.)

6. Drozdova S.O., Pakhomova E.A. [Application of the Cluster Analysis for an Assessment of Quality of Education of University Dubna]. *Sustainable Development: Science and Practice*, 2014, iss. 1(12), pp. 15. (in Russ.)

7. D'yachuk, A.A. *Matematicheskie metody v psikhologicheskikh i pedagogicheskikh issledovaniyakh* [Mathematical Methods in Psychological and Pedagogical Researches]. Krasnoyarsk, KGPU im. V.P. Astaf'eva Publ., 2013. 347 p.

8. Sakovich S.M. [Role of Information and Communication Technologies in Ensuring Quality and Availability of the Higher Education]. *Bulletin of the Moscow State Regional University. Ser. Pedagogics*, 2010, no. 3, pp. 40–45. (in Russ.)

9. Soldatenkov R.M. [Mathematical Training of Students of Higher Educational Institutions and Direction of its Transformation]. *Bulletin of the Moscow State Regional University. Ser. Pedagogics*, 2010, no. 1, pp. 136–141. (in Russ.)

10. Fatkullin N.Yu., Shamshovich V.F., Bakhtizin R.N. [Mark and Rating System of an Assessment of Knowledge of Students of Mathematics with Use of Information and Communication Technologies in Public Educational Institution of Higher Professional Training Ufa State Petroleum Technological University]. *Materialy chetvertoy mezhdunarodnoy konferentsii po voprosam obucheniya s primeneniem tekhnologiy e-learning "Moscow education online 2010"* (29 sentyabrya – 1 oktyabrya 2010 g.) [The Fourth International Conference on Questions of Training with Application of the E-learning "Moscow Education Online 2010" (on September 29 – on October 1 2010)]. 2010, pp. 62–64. (in Russ.)

11. Fatkullin N.Yu., Shamshovich V.F., Bakhtizin R.N. [Practical Realization of Methods of Distance Learning on the Basis of Information and Communication Technologies at Mark and Rating System to an Assessment of Knowledge of Students]. *Distantсионные технологии в учебном процессе: doklady nauchno-metodicheskogo seminara* [Remote Technologies in Educational Process: The Scientific and Methodical Seminar]. 2010, pp. 46–47. (in Russ.)

12. Khalafyan A.A. *Statistica 6. Statisticheskii analiz dannykh* [Statistica 6. Statistical Analysis of Data]. Moscow, Binom Publ., 2009. 528 p.

13. Shamshovich V.F., Fatkullin N.Yu., Bakhtizin R.N. [Introduction of Mark and Rating System of an Assessment of Knowledge of Students with Use of Information and Communication Technologies and Behavior of Procedures of Monitoring and Forecasting of a Gain Score of Students on Mathematics by Methods of Neural Network Technologies]. *Materialy vtoroy vserossiyskoy nauchno-prakticheskoy konferentsii (Kazan', 16–22 aprelya 2010 goda) "Elektronnaya Kazan' 2010"* [Proceedings of the Second All-Russian Scientific-practical Conference (Kazan, 16–22 April 2010) "Electronic Kazan 2010"]. 2010, pp. 255–258. (in Russ.)

14. Xu R., Wunsch II D.C. *Clustering*. New York, Wiley and Sons Publ., 2009.

15. Zhao Y., Karypis G., Fayyad U. Hierarchical Clustering Algorithms for Document Datasets. *Data Mining and Knowledge Discovery*, 2005, no. 10(2), 141–168. DOI: 10.1007/S10618-005-0361-3

Received 3 February 2016

ОБРАЗЕЦ ЦИТИРОВАНИЯ

Вайндорф-Сысоева, М.Е. Кластерный анализ данных балльно-рейтинговой системы (на примере дисциплины «Математика») / М.Е. Вайндорф-Сысоева, Н.Ю. Фаткуллин, В.Ф. Шамшович // Вестник ЮУрГУ. Серия «Образование. Педагогические науки». – 2016. – Т. 8, № 2. – С. 62–77. DOI: 10.14529/ped160209

FOR CITATION

Vayndorf-Sysoeva M.E., Fatkullin N.Yu., Shamshovich V.Ph. Cluster Analysis of Data of Score Rating System (Based on the Subject of "Mathematics"). *Bulletin of the South Ural State University. Ser. Education. Educational Sciences*. 2016, vol. 8, no. 2, pp. 62–77. (in Russ.) DOI: 10.14529/ped160209