

Д.А. Мозгова, О.М. Замятина, Н.А. Семенова, Л.В. Куровская

ДИАГНОСТИКА ПРОФЕССИОНАЛЬНЫХ ДЕФИЦИТОВ И КОМПЕТЕНЦИЙ ПЕДАГОГОВ ОБЩЕГО ОБРАЗОВАНИЯ: КЛАСТЕРНЫЙ АНАЛИЗ

Исследование выполнено при финансовой поддержке Министерства просвещения РФ в рамках государственного задания на 2021 год по теме № 25 «Технология оценки профессионализма и уровня компетенций педагогов общего и профессионального образования» (дополнительное соглашение № 073-03-2021-014/2 от 21.07.2021 г. к соглашению № 073-03-2021-014 от 18.01.2021 г.).

Описан процесс реализации технологии оценки профессионализма и уровня компетенций педагогов общего образования на основе применения методов машинного обучения. Образовательные организации Томской области были поделены на кластеры на основе применения двух алгоритмов кластеризации данных: метода иерархической кластеризации и метода кластеризации k-средних на основании результатов национального единого государственного экзамена. Выявлены «проблемные» задания внутри кластеров, что позволило определить дефицитные компетенции у педагогов для работы над ними.

Ключевые слова: кластеризация данных; компетенции педагога; k-средние (k-means); профессиональные дефициты; оценка уровня компетенции; результаты учеников

Введение

Ежедневно в системе образования формируются огромные потоки информационных данных, касающихся образовательного процесса: успеваемость/неуспеваемость, посещаемость и вовлеченность обучающихся. Часть информации связана с работой педагогов: квалификация и аттестация, результаты внутришкольных, региональных, федеральных мониторингов. Большие данные (Big Data) – большие массивы данных. Они позволяют учесть объем, разнообразие, скорость, вариативность поступающей информации. Такой подход требует использования технологии масштабирования для эффективного хранения, обработки, управления и анализа [1].

Технологический прогресс предоставляет новые возможности для анализа больших данных и применения результатов этого анализа для решения актуального круга задач: от управления системой образования до автоматического составления расписания. Ресурсы использования больших данных (Big Data) в сфере образования перспективны, но данная тенденция только начинает формироваться. В научно-педагогическом сообществе поднимаются вопросы о применении результатов анализа больших данных в сфере образования. Ученые Е.В. Бабенина и О.М. Елкин рассматривают возможность применения технологии нейросети в части управления образовательным процессом [2], В.В. Утёмов и П.М. Горев описывают технологию больших данных как средство развития образовательных систем [3], Б.А. Кондратенко, А.Б. Кондратенко в статье «Перспективы использования больших данных в современном образовании» конкретизируют перспективы использования анализа больших данных среди которых: определение способностей, формирование индивидуальных маршрутов для обучающихся [4].

Для анализа, сегментации и моделирования больших данных применяются алгоритмы машинного обучения. Именно они позволяют интерпретировать массивы данных для составления прогнозов (задачи регрессии), получения категориального ответа (задачи

классификации), распределения данных на группы по заданному количеству признаков (задачи кластеризации), уменьшения размерности данных и определения аномалий. Авторы статьи используют алгоритмы машинного обучения на больших данных для диагностики профессиональных дефицитов педагогов.

Наше исследование рассматривает новую перспективу применения технологий обработки больших данных, которая связана с возможностью оценки профессиональных компетенций педагога через результаты его учеников и созданием предпосылок для устранения профессиональных дефицитов педагогов путем внедрения образовательных программ.

Основная задача Национального проекта «Образование» – обеспечение глобальной конкурентоспособности российского образования и вхождение Российской Федерации в число 10 ведущих стран мира по качеству общего образования [5]. По результатам исследований PISA, PIRLS и TIMSS Российская Федерация занимает 14-е место в мировом рейтинге [6].

В отчете Федерального института оценки качества образования (ФИОКО), посвященном интерпретации данных исследования «Оценка по модели PISA-2020», акцентируется внимание на учителе как на определяющем факторе достижения качественных результатов обучающимися: «...качество работы учителя обуславливает качество образовательных результатов школьников. <...> Создание условий для развития компетенций педагогов и качественного преподавания – важнейшая управленческая задача системы образования» [7].

Результаты исследований компетенций и контекста работы российских учителей указывают на слабое знание, владение и применение учителями современных педагогических технологий, приемов работы с учащимися с рисками образовательной неуспешности и преобладание традиционных урочных форм, а также на слабо развитые предметные компетенции [8]. В работе [9] повышение предметной и методической компетентности педагогов выделяется как основное направление для дальнейшей деятельности по повышению качества общего образования. Основываясь на

анализе возможностей различных мониторингов качества образования и взаимосвязи их результатов с уровнем профессионализма педагогов, исследователи сделали вывод о том, что существует потребность в создании технологии оценки уровня сформированности профессиональных компетенций педагогов для выявления и дальнейшего устранения их профессиональных дефицитов.

На данный момент существуют различные подходы к диагностике профессионализма педагога. Виды диагностики можно разделить по способу взаимодействия:

- аттестация педагогических работников образовательных организаций (ОО) [10], национальная система профессионального роста педагогических работников РФ, включая национальную систему учительского роста [11];

- внешняя экспертная диагностика. Инструменты: карта комплексной диагностики, алгоритм проведения процедуры диагностики, ключевой этап анализ результатов, обеспечение объективности интерпретации результатов диагностики;

- внутришкольная оценка. Оценка производится на диагностической основе, цель: конструктивный поиск причин недостатков и путей их исправления;

- самоанализ деятельности и компетенций педагога.

Вышеперечисленные методы имеют ряд достоинств. Они позволяют произвести экспертную оценку и выявление предметных и методических дефицитов, признание результатов оценки профессиональным сообществом и работодателем. Тем не менее существует противоречие между потребностью оценки профессионализма педагога во взаимосвязи с результатами его учеников и отсутствием технологии, позволяющей это реализовать. Исследование направлено на решение проблемы определения результативности деятельности учителя по итогам оценочных процедур его учеников.

Целью исследования является разработка и апробация технологии, позволяющей выполнить анализ результатов оценочных процедур школьников и установить связи с уровнем профессионализма педагогов для планирования дальнейшей работы по устранению профессиональных дефицитов и развитию компетенций.

В основу технологии оценки профессионализма и уровня компетенций педагога был положен кластерный подход. Анализ имеющихся научных трудов показал, что кластерный подход в образовании реализуется с целью разделения образовательных организаций по характерным и/или основным направлениям деятельности с целью повышения эффективности каждого одиночного элемента и административного управления [12–14].

Методы кластеризации, благодаря их возможностям классификации и стандартизации данных, широко применяются зарубежными специалистами в оценке качества образования. В частности, через кластерный подход можно проанализировать взаимосвязь между полом обучающихся и их мотивацией работы в групповом проекте [15] или оценить уровень сформированности у обучающихся той или иной компетенции [16].

Эффективность процесса обучения во многом зависит от уровня компетенций педагогов. Кластерный

подход позволяет автоматизировать и сделать более объективной экспертную систему оценки компетентности педагогов [17, 18].

В процессе разработки и реализации технологии был применен кластерный анализ данных для создания групп Томских общеобразовательных организаций на основе анализа результатов учеников с применением методов машинного обучения (machine learning): алгоритм кластеризации *k*-средних (*k*-means) и иерархическая кластеризация. В основу исследования было положено предположение о том, что применение кластерного анализа позволит выявить дефициты и наметить содержательные линии по улучшению образовательных результатов школьников через работу с педагогическими коллективами образовательных организаций Томской области.

Алгоритм реализации исследования

Для работы с большими данными используются алгоритмы машинного обучения, если же говорить об обобщенных типах задач машинного обучения, то можно выделить следующие, самые распространенные: регрессия; классификация; кластеризация [19]. Рассмотрим постановку этих задач:

Задача регрессии – приближение неизвестной целевой зависимости на некотором множестве данных [19]. Предполагается, что существует зависимость объясняемой переменной от другой или других переменных, алгоритмы регрессии могут прогнозировать вероятность наступления определенного события или предсказывать значения зависимой переменной по значениям независимой или независимых переменных.

Задача классификации – распределение объектов по заданному множеству групп (классов) [19]. Исходя из классов объектов подмножества (конечное множество объектов), алгоритм присваивает классовую принадлежность любому объекту из этого множества.

Задача кластеризации – разбиение некоторого множества исследуемых объектов и признаков на однородные и непересекающиеся группы, или кластеры [20].

Учитывая цель данного исследования для реализации алгоритма технологии оценки профессионализма и уровня компетенций педагогов на основе анализа результатов учеников были выбраны методы кластерного анализа. Методы кластеризации можно разделить на две большие группы: иерархические и неиерархические. Неиерархические алгоритмы позволяют получить фиксированное разбиение данных, а иерархические – систему вложенных разбиений, соответствующих различным уровням иерархии [21].

Методы иерархической кластеризации основываются на двух идеях: агломерации (AGNES, Agglomerative Nesting), т.е. последовательного объединения индивидуальных объектов или их групп во все более крупные подмножества, или обратном по смыслу процессе разбиения (DIANA, Divise Analysis), который начинается с корня и на каждом шаге делит образующие группы по степени их гетерогенности. В обоих случаях результат работы алгоритма представляет собой древовидную структуру, или дендрограмму [22].

Метод кластеризации *k*-средних (*k*-means) относится к *неиерархическим алгоритмам*. Основной иде-

ей неиерархических алгоритмов кластеризации данных является минимизация расстояний между объектами в кластерах. Это происходит до тех пор, пока минимизирование расстояния между объектами становится невозможным [23].

Алгоритм реализации технологии оценки профессионализма и уровня компетенций педагогов на основе анализа результатов учеников представлен на рис. 1.

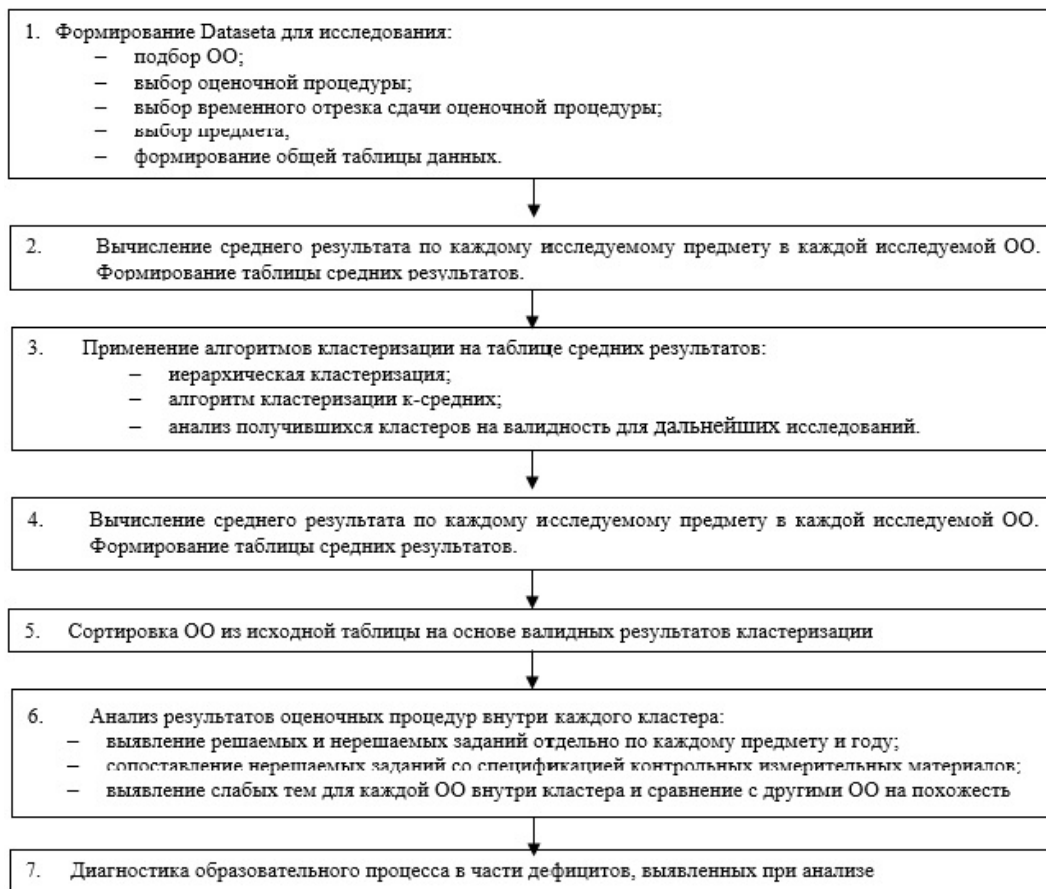


Рис. 1. Алгоритм реализации технологии оценки профессионализма и уровня компетенций педагогов на основе анализа результатов учеников

Данные для исследования

В исследовании для первичной кластеризации были взяты образовательные результаты национального единого государственного экзамена (ЕГЭ) учеников Томской области по предметам: русский язык, профильная математика, физика и обществознание (с 2015 по 2019 г.) (табл. 1). Результаты ЕГЭ рассматриваются с 2015 г. по причине разделения экзамена по математике на базовый и профильный уровни. Базовый уровень сдается выпускниками, которые не планируют продолжать обучение в высших учебных заведениях или для поступления им не требуется предмет «математика». Сдача экзамена по профильной математике необходим выпускникам, которые планируют продолжать обучение в высшем учебном заведении по специальностям, для поступления на которые требуется сдача экзамена по математике.

Обезличенные результаты одиннадцатиклассников с разбиением по школам были занесены в таблицу, выведено среднее значение по каждому из предметов для каждой школы, далее к этим данным были применены два алгоритма кластеризации: k-средних и алгоритм иерархической кластеризации (табл. 2).

Таблица 1
Количество образовательных организаций, участвующих в ЕГЭ с 2015–2019 гг. по предметам

Год	Русский язык	Профильная математика	Физика	Обществознание
2015	223	215	163	197
2016	220	212	170	190
2017	220	204	161	197
2018	222	214	173	201
2019	226	198	150	197

Таблица 2
Вид таблицы для реализации алгоритмов кластеризации

Закодированный номер ОО	Год, предмет				
	2015 ru	2015 math	2015 fi		2019 ob
n ₁	mean_score _{ru}	mean_score _{math}	mean_score _{fi}	-----	mean_score _{ob}
-----	-----	-----	-----	-----	-----
n _m	mean_score _{ru}	mean_score _{math}	mean_score _{fi}	-----	mean_score _{ob}

Результаты исследования

Алгоритмы кластеризации были реализованы с помощью языка программирования Python. Количество кластеров было выбрано методом подбора, n = 15.

Задача применения алгоритмов кластеризации в нашем исследовании направлена не столько на деление образовательных организаций на сильные и слабые, сколько на выявление связей между образовательными организациями на основе результатов учеников за пять лет по нескольким предметам. Результаты работы по кластеризации представлены в табл. 3, 4.

Таблица 3

Результаты работы алгоритма «Иерархическая кластеризация»

№ кластера	Количество ОО
1	4
2	2
3	1
4	11
5	3
6	168
7	8
8	1
9	12
10	3
11	13
12	1
13	2
14	6
15	1

Таблица 4

Результаты работы алгоритма «Метод к-средних (k-means)»

№ кластера	Количество ОО
1	83
2	5
3	6
4	15
5	12
6	10
7	10
8	7
9	10
10	14
11	4
12	3
13	23
14	5
15	29

По результатам применения кластеризации были сделаны промежуточные выводы:

Результаты реализации алгоритма иерархической кластеризации в дальнейшей работе не рассматриваются, так как кластеры, сформировавшиеся после первой итерации, неравномерны (количество образовательных организаций варьируется от 1 до 168) и не валидны для дальнейшего исследования.

Результаты, полученные после применения метода к-средних, являются валидными и применимы для дальнейшего исследования.

Количество образовательных организаций в кластерах варьируется от 3 до 29, а также сформировался большой кластер 83 образовательные организации (табл. 4).

На следующем этапе исследования рассматривались результаты учеников внутри каждого кластера. Эта работа позволила выявить группы «сложных» заданий по предметам «русский язык» и «профильная математика».

В данной работе под «сложными» заданиями понимаются задания, которые не смогли решить 30% и более обучающихся. «Сложные» задания были сопоставлены со спецификацией контрольных измерительных материалов ЕГЭ и представлены в табл. 5. Спецификации контрольных измерительных материалов для проведения в 2015–2019 гг. единого государственного экзамена по русскому языку и математике (профильный уровень) подготовлены и утверждены Федеральным государственным бюджетным учреждением «Федеральный институт педагогических измерений».

Выводы

Работа по сопоставлению внутри кластеров позволила выделить темы, вызывающие наибольшее затруднение у обучающихся по предметам «Русский язык», «Профильная математика» и, следовательно, определить дефициты в знаниях и умениях учеников. Произвести оценку «сложных» заданий по предметам «Физика» и «Обществознание» в рамках данного исследования не представляется возможным из-за малого количества учеников, выбравших эти предметы для сдачи.

Таблица 5

Темы проверяемых элементов содержания, вызвавших наибольшие затруднения у обучающихся по русскому языку и профильной математике*

№ кластера	Русский язык	Профильная математика
1	Текст как речевое произведение. Смысловая и композиционная целостность текста	Уметь использовать приобретенные знания и умения в практической деятельности и повседневной жизни (базовый уровень). Уметь строить и исследовать простейшие математические модели (базовый уровень). Уметь выполнять действия с функциями. Уметь решать уравнения и неравенства (повышенный уровень)
2	Правописание корней. Правописание НЕ и НИ. Правописание -Н- и -НН- в различных частях речи. Знаки препинания в предложениях с обособленными членами. Знаки препинания в сложноподчиненном предложении. Знаки препинания в сложном предложении с разными видами связи	Низкое качество результатов по всем темам, кроме: уметь использовать приобретенные знания и умения в практической деятельности и повседневной жизни (базовый уровень)
3	Текст как речевое произведение. Смысловая и композиционная целостность текста	Низкое качество результатов по всем темам, кроме: уметь использовать приобретенные знания и умения в практической деятельности и повседневной жизни (базовый уровень), уметь выполнять действия с геометрическими фигурами, координатами и векторами (базовый уровень), уметь строить и исследовать простейшие математические модели (базовый уровень)

№ кла-стера	Русский язык	Профильная математика
4	Морфологические нормы. Синтаксические нормы. Нормы согласования. Нормы управления. Знаки препинания в простом осложненном предложении (с однородными членами). Пунктуация в сложносочиненном предложении и простом предложении с однородными членами. Знаки препинания в сложном предложении с разными видами связи. Текст как речевое произведение. Смысловая и композиционная целостность текста	Низкое качество результатов по всем темам, кроме: уметь использовать приобретенные знания и умения в практической деятельности и повседневной жизни (базовый уровень), уметь строить и исследовать простейшие математические модели (базовый уровень)
5	Средства связи предложений в тексте. Орфографические нормы (постановка ударения). Синтаксические нормы. Нормы согласования. Нормы управления. Слитное, дефисное, раздельное написание слов. Знаки препинания в простом осложненном предложении (с однородными членами). Знаки препинания в сложном предложении с разными видами связи. Текст как речевое произведение. Смысловая и композиционная целостность текста. Функционально-смысловые типы речи. Лексическое значение слова. Синонимы. Антонимы. Омонимы. Фразеологические обороты. Группы слов по происхождению и употреблению	Низкое качество результатов по всем темам, кроме: уметь использовать приобретенные знания и умения в практической деятельности и повседневной жизни (базовый уровень), уметь строить и исследовать простейшие математические модели (базовый уровень)
6	Слитное, дефисное, раздельное написание слов. Знаки препинания в предложениях с обособленными членами (определениями, обстоятельствами, приложениями, дополнениями). Текст как речевое произведение. Смысловая и композиционная целостность текста. Текст как речевое произведение. Смысловая и композиционная целостность текста. Лексическое значение слова. Синонимы. Антонимы. Омонимы. Фразеологические обороты. Группы слов по происхождению и употреблению	Низкое качество результатов по всем темам, кроме: уметь использовать приобретенные знания и умения в практической деятельности и повседневной жизни (базовый уровень), уметь строить и исследовать простейшие математические модели (базовый уровень), уметь выполнять действия с геометрическими фигурами, координатами и векторами (продвинутый уровень)
7	Правописание суффиксов различных частей речи (кроме Н и НН). Знаки препинания в предложениях с обособленными членами (определениями, обстоятельствами, приложениями, дополнениями). Знаки препинания в сложноподчиненном предложении. Текст как речевое произведение. Смысловая и композиционная целостность текста. Лексическое значение слова. Синонимы. Антонимы. Омонимы. Фразеологические обороты. Речь. Языковые средства выразительности	Уметь выполнять действия с геометрическими фигурами, координатами и векторами (базовый уровень). Уметь выполнять действия с функциями (базовый уровень). Уметь выполнять вычисления и преобразования (повышенный уровень). Уметь использовать приобретенные знания и умения в практической деятельности и повседневной жизни (повышенный уровень)
8	Знаки препинания в сложноподчиненном предложении. Знаки препинания в предложениях с обособленными членами (определениями, обстоятельствами, приложениями, дополнениями). Текст как речевое произведение. Смысловая и композиционная целостность текста. Средства связи предложений в тексте	Низкое качество результатов по всем темам, кроме: уметь использовать приобретенные знания и умения в практической деятельности и повседневной жизни (базовый уровень), уметь строить и исследовать простейшие математические модели (базовый уровень). Уметь выполнять действия с функциями (базовый уровень)
9	Правописание личных окончаний глаголов и суффиксов причастий	Низкое качество результатов по всем темам, кроме: уметь использовать приобретенные знания и умения в практической деятельности и повседневной жизни (базовый уровень), уметь строить и исследовать простейшие математические модели (базовый уровень), уметь выполнять действия с геометрическими фигурами, координатами и векторами (продвинутый уровень), уметь выполнять действия с функциями (базовый уровень), уметь строить и исследовать простейшие математические модели (базовый уровень)
10	Орфографические нормы (постановка ударения). Синтаксические нормы. Нормы согласования. Нормы управления. Слитное, дефисное, раздельное написание слов. Знаки препинания в сложноподчиненном предложении. Знаки препинания в предложениях с обособленными членами (определениями, обстоятельствами, приложениями, дополнениями). Речь. Языковые средства выразительности	Уметь выполнять вычисления и преобразования (повышенный уровень), уметь выполнять действия с функциями (повышенный уровень), уметь решать уравнения и неравенства (повышенный уровень)
11	Слитное, дефисное, раздельное написание слов. Знаки препинания в предложениях со словами и конструкциями, грамматически не связанными членами предложения. Текст как речевое произведение. Смысловая и композиционная целостность текста. Лексическое значение слова. Синонимы. Антонимы. Омонимы. Фразеологические обороты. Группы слов по происхождению и употреблению	Низкое качество результатов по всем темам, кроме: уметь использовать приобретенные знания и умения в практической деятельности и повседневной жизни (базовый уровень), уметь строить и исследовать простейшие математические модели (базовый уровень), уметь строить и исследовать простейшие математические модели (базовый уровень)
12	Правописание личных окончаний глаголов и суффиксов причастий. Знаки препинания в простом осложненном предложении (с однородными членами). Пунктуация в сложносочиненном предложении и простом предложении с однородными членами	Уметь выполнять действия с геометрическими фигурами, координатами и векторами (базовый уровень), уметь решать уравнения и неравенства (базовый уровень), уметь выполнять действия с геометрическими фигурами, координатами и векторами (повышенный уровень)
13	Средства связи предложений в тексте. Морфологические нормы (образование форм слова). Средства связи предложений в тексте. Знаки препинания в предложениях с обособленными членами (определениями, обстоятельствами, приложениями, дополнениями)	Уметь использовать приобретенные знания и умения в практической деятельности и повседневной жизни (повышенный уровень). Уметь решать уравнения и неравенства (высокий уровень). Уметь строить и исследовать простейшие математические модели (высокий уровень)
14	Знаки препинания в сложном предложении с разными видами связи. Текст как речевое произведение. Смысловая и композиционная целостность текста	Низкое качество результатов по всем темам, кроме: уметь использовать приобретенные знания и умения в практической деятельности и повседневной жизни (базовый уровень), уметь строить и исследовать простейшие математические модели (базовый уровень)

№ кла-стера	Русский язык	Профильная математика
15	Текст как речевое произведение. Смысловая и композиционная целостность текста.	Уметь строить и исследовать простейшие математические модели (профильный уровень), уметь решать уравнения и неравенства (высокий уровень).

* Спецификации контрольных измерительных материалов для проведения в 2015–2019 гг. единого государственного экзамена по русскому языку и математике (профильный уровень) размещены в открытом доступе на сайте <https://fipi.ru/egge/demoversii-specifikacii-kodifikatory>

Анализ показал, что внутри одного кластера учебники имеют схожие затруднения при решении одинаковых заданий, что позволяет акцентировать внимание на соответствующих компетенциях педагога и на его предметных и методических дефицитах и вести дальнейшую работу по оценке его профессиональной деятельности уже узконаправленно. Кроме того, проведенная работа позволяет наметить и пути решения по преодолению дефицитов педагогов. При формиро-

вании содержания курсов повышения квалификации по темам и траекториям можно за основу брать группы школ, входящих в один кластер.

Перспективами исследования является изучение возможностей учета дополнительных сведений при оценке уровня профессионализма: уровня квалификации педагогов, результатов анализа дополнительных оценочных процедур для составления более объективной картины об уровне образовательных результатов.

ЛИТЕРАТУРА

- ГОСТ Р ИСО/МЭК 20546-2021. Информационные технологии – Большие данные – Обзор и словарь: национальный стандарт Российской Федерации: издание официальное: утвержден и введен в действие Приказом Федерального агентства по техническому регулированию и метрологии от 13 июля 2021 г. № 632-ст: введен впервые: дата введения 13-07-2021 / подготовлен Федеральным государственным бюджетным образовательным учреждением высшего образования «Московский государственный университет имени М.В. Ломоносова» (МГУ имени М.В. Ломоносова) в лице Научно-образовательного центра компетенций в области цифровой экономики МГУ и Автономной некоммерческой организацией «Институт развития информационного общества» (ИРИО) на основе собственного перевода на русский язык англоязычной версии стандарта, указанного в пункте 4. М. : Стандартинформ, 2017. 16 с.
- Бегенина Е.В., Елкин О.М. Повышение качества управления образованием с использованием технологии обработки больших данных // Отечественная и зарубежная педагогика. 2020. Т. 1, № 6 (72). С. 22–29.
- Утёмов В.В., Горев П.М. Развитие образовательных систем на основе технологии Big Data // Научно-методический электронный журнал «Концепт». 2018. № 6 (июнь). С. 449–461. URL: <http://e-koncept.ru/2018/181039.htm> (дата обращения: 30.09.2021).
- Кондратенко Б.А., Кондратенко А.Б. Перспективы использования больших данных в современном образовании // Вестник Балтийского федерального университета им. И. Канта. Сер.: Филология, педагогика, психология. 2018. № 1. С. 117–126.
- Российская Федерация. Паспорт национального проекта «Образование»: протокол от 24.12.2018 г. № 16: [утвержден президиумом Совета при Президенте Российской Федерации по стратегическому развитию и национальным проектам].
- Россия опережает намеченный график вхождения в десятку лучших стран мира по качеству общего образования // Федеральная служба по надзору в сфере образования : официальный сайт Рособразования. URL: <http://obrnadzor.gov.ru/news/rossiya-operezhaet-namechennyj-grafik-vhozheniya-v-desyatku-luchshih-stran-mira-po-kachestvu-obshhego-obrazovaniya/> (дата обращения: 30.08.2021).
- Резильентность. Оценка по модели PISA-2020 // Федеральный институт оценки качества образования : официальный сайт. URL: <https://fioco.ru/Media/Default/Documents/Резильентность%202020.pdf> (дата обращения: 30.08.2021).
- Федеральная служба по надзору в сфере образования (Рособразование). Приказ № 847. Об утверждении методики выявления общеобразовательных организаций, имеющих низкие образовательные результаты обучающихся, на основе комплексного анализа данных об образовательных организациях, в том числе данных о качестве образования: [принят 19.08.2020 г.].
- Mozgova D., Zamyatina O. Identification of Learning Outcome Deterioration Factors Among Children, Review of Potential Mitigators. The Case of Tomsk Region // Advances in Social Science, Education and Humanities Research. 29.11.2020. doi: 10.2991/assehr.k.201128.116
- Приказ Министерства образования и науки РФ от 7 апреля 2014 г. № 276 «Об утверждении Порядка проведения аттестации педагогических работников организаций, осуществляющих образовательную деятельность» (с изменениями и дополнениями) // Информационно-правовой портал система Гарант. URL: <http://base.garant.ru/70662982/#ixzz77iFtaAZx> (дата обращения: 02.09.2021).
- Распоряжение Правительства РФ от 31 декабря 2019 г. № 3273-р «Об утверждении основных принципов национальной системы профессионального роста педагогических работников РФ, включая национальную систему учительского роста» // Информационно-правовой портал система Гарант. URL: <https://www.garant.ru/products/ipo/prime/doc/73284005/> (дата обращения: 02.09.2021).
- Семькина Е.Н. Кластерный подход как управленческий в образовании и воспитании // Вестник Тамбовского университета. Серия: Гуманитарные науки. 2010. Вып. 2 (82). С. 141–144.
- Воронин Ф.И. Кластерный подход в организации развития муниципального образования // Экономика. Информатика. 2009. № 9 (64). С. 24–29.
- Боровская М.А., Масыч М.А., Паничкина М.В. Совершенствование системы непрерывного образования: кластерный и экосистемный подход // Гуманитарные Юга России. 2020. Т. 9, № 45. doi: 10.18522/2227-8656.2020.5.1
- Miller E., Bailey R. Basic Psychological Need Fulfillment by Gender in Team Environments // Proceedings - Frontiers in Education Conference, FIE. 2018. Art. № 8659027. doi: 10.1109/FIE.2018.8659027
- Bossolasco M.L., Chiecher A.C., Dos Santos D.A. Profiles of access and appropriation of ICT in freshmen students. Comparative study in two Argentine public universities // Pixel-Bit, Revista de Medios y Educacion. 2020. Vol. 57. P. 151–172. doi: 10.12795/pixelbit.2020.i57.06
- Regueras L.M., Verdu M.J., De Castro J.-P. Clustering Analysis for Automatic Certification of LMS Strategies in a University Virtual Campus // IEEE Access. 2019. Vol. 7. P. 137680–137690. Art. № 8846693. doi: 10.1109/ACCESS.2019.2943212
- Longobardi S., Falzetti P., Pagliuca M.M. Quis custodiet ipsos custodes? How to detect and correct teacher cheating in Italian student data // Statistical Methods and Applications. 2018. Vol. 27, Is. 3. P. 515–543. doi: 10.1007/s10260-018-0426-2
- Воронина В.В., Михеев А.В., Ярушклина Н.Г., Святков К.В. Теория и практика машинного обучения : учеб. пособие. Ульяновск : УлГТУ, 2017. 290 с.
- Ким Дж.-О. Факторный, дискриминантный и кластерный анализ / пер. с англ. А.М. Хотинского и С.Б. Королева ; науч. ред. И.С. Енюкова. М. : Финансы и статистика, 1989. 215 с.
- Пестунов И.А., Рылов С.А., Бериков В.Б. Иерархические алгоритмы кластеризации для сегментации мультиспектральных изображений // Автоматика. 2015. Т. 51, № 4. С. 12–22.
- Шитиков В.К., Мастицкий С.Э. Классификация, регрессия, алгоритмы Data Mining с использованием R. Электронная книга. 2017. URL: <https://github.com/ranalytics/data-mining>
- Кокорева Я.В., Макаров А.А. Поэтапный процесс кластерного анализа данных на основе алгоритма кластеризации k-means // Молодой ученый. 2015. № 13 (93). С. 126–128. URL: <https://moluch.ru/archive/93/20759/> (дата обращения: 06.09.2021).

Статья представлена научной редакцией «Педагогика» 1 октября 2021 г.

Diagnostics of Professional Deficits and Competencies of General Education Teachers: Cluster Analysis

Vestnik Tomskogo gosudarstvennogo universiteta – Tomsk State University Journal, 2021, 472, 189–196.

DOI: 10.17223/15617793/472/22

Darya A. Mozgova, Tomsk State Pedagogical University (Tomsk, Russian Federation); Tomsk Regional Teachers Professional Re-training Institute (Tomsk, Russian Federation). E-mail: mozgovadarya@bk.ru

Oksana M. Zamyatina, Tomsk State Pedagogical University (Tomsk, Russian Federation), Tomsk Regional Teachers Professional Re-training Institute (Tomsk, Russian Federation). E-mail: zamyatina@tpu.ru

Nataliya A. Semenova, Tomsk State Pedagogical University (Tomsk, Russian Federation). E-mail: natalsem@rambler.ru

Larisa V. Kurovskaya, Tomsk State Pedagogical University (Tomsk, Russian Federation). E-mail: l_kurovskaya@mail.ru

The work was supported by the Ministry of Education of Russian Federation, Theme 25 “Technology for Assessing the Professionalism and Level of Competence of Teachers of General and Vocational Education” of state assignment (Additional Agreement No. 073-03-2021-014/2 dated July 21, 2021, to Agreement No. 073-03-2021-014 dated January 18, 2021).

The study examines a new perspective for the use of big data processing technologies that is associated with the possibility of assessing professional competencies of a teacher by means of assessing the learning results of students and contributes to the creation of prerequisites for eliminating professional deficiencies of teachers with the help of new specially designed educational programs. The aim of the study is to develop and test such a technology that makes it possible to analyze the results of assessing procedures of the quality of schoolchildren's education and establish the connection between these results and the level of teachers' professionalism in order to plan further activity to eliminate professional deficiencies and develop teachers' competencies. To implement such a technology, an algorithm has been developed for assessing professionalism and level of teachers' competencies, making possible to assess educational results of students. The technology is based on a cluster approach to data analysis, the methods of clustering k-means and hierarchical clustering have been chosen. In the study, for primary clustering, the learning results of students (Tomsk Oblast) in the national Unified State Exam within the period from 2015 to 2019 were analyzed (subjects: Russian, Mathematics (profile level), Physics, Social Studies). The number of clusters was selected by the selection method, $n = 15$. The next stage of the study analyzed the results of students within each cluster. The stage made possible to identify groups of “difficult” assignments that 30% or more of schoolchildren could not cope with (subjects: Russian and Mathematics (profile level)). As a result, the assignments that caused the greatest difficulty for schoolchildren were fixed, and thus deficiencies in their knowledge and skills were identified. It was not possible to assess “difficult” assignments in Physics and Social Studies in the framework of the study, since a small number of students had chosen these subjects to pass in the exam. The comparative analysis demonstrated that, within the same cluster, schoolchildren had similar difficulties in doing the same assignments, which that allows us to focus on appropriate teachers' competencies and on subject and methodology deficiencies, and to conduct further activity on assessing the professional competence of teachers focused narrowly. In addition, the study allows us to outline ways of solving the problem of overcoming professional deficits of teachers. Thus, it becomes possible to specifically develop additional professional training programs (advanced/refresher training) for teachers of the same cluster, the curricula of the programs containing topics and learning paths that will improve the educational results of their students, and thereby improve the quality of secondary education in Tomsk Oblast.

REFERENCES

1. Russian Federation. (2017) *GOST R ISO/MEK 20546-2021. Informatsionnye tekhnologii – Bol'shie dannye – Obzor i slovar': natsional'nyy standart Rossiyskoy Federatsii: izdanie ofitsial'noe* [GOST R ISO/IEC 20546-2021. Information Technologies – Big Data – Overview and Dictionary: National Standard of the Russian Federation: official edition]. Moscow: Standartinform.
2. Bebenina, E.V. & Elkin, O.M. (2020) Improving the quality of education management with big data processing technology. *Otechestvennaya i zarubezhnaya pedagogika*. 6-1 (72). pp. 22–29. (In Russian).
3. Utemov, V.V. & Gorev, P.M. (2018) Razvitiye obrazovatel'nykh sistem na osnove tekhnologii Big Data [Development of educational systems based on Big Data technology]. *Nauchno-metodicheskiy elektronnyy zhurnal Kontsept*. 6. pp. 449–461. [Online] Available from: <http://e-koncept.ru/2018/181039.htm>. (Accessed: 30.09.2021). DOI: 10.24422/MCITO.2018.6.14501
4. Kondratenko, B.A. & Kondratenko, A.B. (2018) Big data in modern education. *Vestnik Baltiyskogo federal'nogo universiteta im. I. Kanta. Ser.: Filologiya, pedagogika, psikhologiya – Vestnik Ikbfu. Philology, Pedagogy, And Psychology*. 1. pp. 117–126. (In Russian).
5. Rossiyskaya Federatsiya. (2018) *Pasport natsional'nogo proekta “Obrazovanie”*: protokol ot 24.12.2018 g. № 16: utverzhden prezidiumom Soveta pri Prezidente Rossiyskoy Federatsii po strategicheskoy razvitiyu i natsional'nyy proektam [Passport of the national project “Education”: Protocol No. 16 of December 24, 2018 approved by the Presidium of the Council under the President of the Russian Federation for Strategic Development and National Projects]. [Online] Available from: <https://edu.gov.ru/national-project/>.
6. Federal Service for Supervision in Education and Science. (2020) *Rossiya operezhaet namechennyy grafik vkhodzheniya v desyatku luchshikh stran mira po kachestvu obshchego obrazovaniya* [Russia is ahead of the planned schedule of entering the top ten countries in the world in terms of the quality of general education]. [Online] Available from: <http://obrnadzor.gov.ru/news/rossiya-operezhaet-namechennyy-grafik-vkhodzheniya-v-desyatku-luchshih-stran-mira-po-kachestvu-obshchego-obrazovaniya/>. (Accessed: 30.08.2021).
7. Federal'nyy institut otsenki kachestva obrazovaniya: ofitsial'nyy sayt [Federal Institute for Education Quality Assessment: official website]. (2020) *Rezil'entnost'. Otsenka po modeli PISA-2020* [Resistance. Assessment according to the PISA-2020 model]. [Online] Available from: <https://fioco.ru/Media/Default/Documents/Rezil'entnost'%202020.pdf>. (Accessed: 30.08.2021).
8. Federal Service for Supervision in Education and Science. (2020) *Prikaz № 847. Ob utverzhdenii metodiki vyyavleniya obshcheobrazovatel'nykh organizatsiy, imeyushchikh nizkie obrazovatel'nye rezul'taty obuchayushchikhsya, na osnove kompleksnogo analiza dannykh ob obrazovatel'nykh organizatsiyakh, v tom chisle dannykh o kachestve obrazovaniya: prinyat 19.08.2020 g.* [Order No. 847 of August 19, 2020. On the approval of the methodology for identifying general education organizations with low educational results of students, based on a comprehensive analysis of data on educational organizations, including data on the quality of education].
9. Mozgova, D. & Zamyatina, O. (2020) Identification of learning outcome deterioration factors among children, review of potential mitigators. the case of Tomsk region. *Advances in Social Science, Education and Humanities Research*. DOI: 10.2991/assehr.k.201128.116
10. Garant. (2014) *Order of the Ministry of Education and Science of the Russian Federation No. 276 of April 7, 2014 “On Approval of the Procedure for Certification of Teaching Staff of Organizations Engaged in Educational Activities” (with amendments and additions)*. [Online] Available from: <http://base.garant.ru/70662982/#ixzz77iFtaAZx>. (Accessed: 02.09.2021). (In Russian).
11. Garant. (2019) *Decree of the Government of the Russian Federation No. 3273-r of December 31, 2019 “On Approval of the Basic Principles of the National System of Professional Growth of Teachers of the Russian Federation, Including the National System of Teacher Growth”*. [Online] Available from: <https://www.garant.ru/products/ipo/prime/doc/73284005/>. (Accessed: 02.09.2021). (In Russian).

12. Semykina, E.N. (2010) Cluster approach as the administrative resource in education and up-bringing. *Vestnik Tambovskogo universiteta. Seriya: Gumanitarnye nauki – Tambov University Review: Series Humanities*. 2 (82). pp. 141–144. (In Russian).
13. Voronin, F.I. (2009) Klasternyy podkhod v organizatsii razvitiya munitsipal'nogo obrazovaniya [Cluster approach in the organization of municipal education development]. *Ekonomika. Informatika – Economics. Information Technologies*. 9 (64). pp. 24–29. (In Russian).
14. Borovskaya, M.A., Masych, M.A. & Panichkina, M.V. (2020) Improving lifelong learning: cluster and ecosystem approaches. *Gumanitarii Yuga Rossii – Humanities of the South of Russia*. 5 (9). (In Russian). DOI: 10.18522/2227-8656.2020.5.1
15. Miller, E. & Bailey, R. (2018) [Basic Psychological Need Fulfillment by Gender in Team Environments]. *Proceedings of 2018 IEEE Frontiers in Education Conference (FIE)*. San Jose. 03–06 October 2018. New Jersey: IEEE. DOI: 10.1109/FIE.2018.8659027
16. Bossolasco, M.L., Chiecher, A.C. & Dos Santos, D.A. (2020) Profiles of access and appropriation of ICT in freshmen students. Comparative study in two Argentine public universities. *Pixel-Bit, Revista de Medios y Educacion*. 57. pp. 151–172. DOI: 10.12795/pixelbit.2020.i57.06
17. Regueras, L.M., Verdu, M.J. & De Castro, J.-P. (2019) Clustering Analysis for Automatic Certification of LMS Strategies in a University Virtual Campus. *IEEE Access*. 7. pp. 137680–137690. DOI: 10.1109/ACCESS.2019.2943212
18. Longobardi, S., Falzetti, P. & Pagliuca, M.M. (2018) Quis custodiet ipsos custodes? How to detect and correct teacher cheating in Italian student data. *Statistical Methods and Applications*. 3 (27). pp. 515–543. DOI: 10.1007/s10260-018-0426-2
19. Voronina, V.V. et al. (2017) *Teoriya i praktika mashinnogo obucheniya* [Theory and Practice of Machine Learning]. Ulyanovsk: Ulyanovsk State Technical University.
20. Kim, J.-O. et al. (1989) *Faktornyy, diskriminantnyy i klasternyy analiz* [Factorial, discriminant and cluster analysis]. Translated from English by A.M. Khotinskiy & S.B. Korolev. Moscow: Finansy i statistika.
21. Pestunov, I.A., Rylov, S.A. & Berikov, V.B. (2015) Hierarchical clustering algorithms for segmentation of multispectral images. *Avtometriya – Optoelectronics, Instrumentation and Data Processing*. 4 (51). pp. 12–22. (In Russian). DOI: 10.3103/S8756699015040020
22. Shitikov, V.K. & Mastitskiy, S.E. (2017) *Klassifikatsiya, regressiya, algoritmy Data Mining s ispol'zovaniem R* [Classification, regression, Data Mining Algorithms Using R]. [Online] Available from: <https://github.com/ranalytics/data-mining>.
23. Kokoreva, Ya.V. & Makarov, A.A. (2015) Poetapnyy protsess klasternogo analiza dannykh na osnove algoritma klasterizatsii k-means [A step-by-step process of clustering data analysis based on the k-means clustering algorithm]. *Molodoy uchenyy – Young Scientist*. 13 (93). pp. 126–128. [Online] Available from: <https://moluch.ru/archive/93/20759/> (Accessed: 06.09.2021).

Received: 01 October 2021