

УДК 004.046

ПОСТРОЕНИЕ АЛГОРИТМА ИНТЕЛЛЕКТУАЛЬНОГО АНАЛИЗА ПОВЕДЕНЧЕСКИХ КЛАССИФИКАТОРОВ ПРИ ПЕРСОНАЛИЗАЦИИ ОБУЧЕНИЯ

Тлегенова Т.Е., Шардаков В.М.

*ФГБОУВО «Оренбургский государственный университет», Оренбург,
e-mail: shardakov_vm@mail.osu.ru*

В условиях активного развития электронного обучения одним из ключевых направлений современных исследований является персонализация образования. Персонализированное обучение позволяет адаптировать образовательный процесс под конкретного студента, что максимально индивидуализирует его образовательные потребности, когнитивные стили и карьерные устремления. С целью оптимизации персонализированной системы обучения авторы предлагают применить интеллектуальный анализ данных, обеспечивающий обнаружение скрытых шаблонов, отношений и правил для анализа, классификации и прогнозирования результатов обучения. В частности, в работе предлагается провести анализ поведенческих характеристик обучающихся, полученных на основе результатов предварительного анкетирования в целях обнаружения в полученных результатах, посредством интеллектуального анализа данных, скрытой информации, для извлечения которой разработан алгоритм интеллектуального анализа поведенческих классификаторов при персонализации обучения. Таким образом, спектр функциональных задач, решаемых технологиями интеллектуального анализа данных, чрезвычайно широк, начиная от планирования и проведения самостоятельной работы обучающихся до развития компетенций. Ключевыми задачами при разработке алгоритма интеллектуального анализа данных при персонализированном обучении являются: разработка входного контента, построение нормативной модели для обучающегося, а также проектирование математической модели. Для решения каждой из перечисленных задач требуется формализация входного набора параметров, разработка подсистемы определения педагогических методов и принципов, позволяющей осуществлять построение персонализированного обучения. Применяя разработанный алгоритм анализа данных, авторы прогнозируют итоговые оценки успешных студентов и студентов, подверженных риску неудач.

Ключевые слова: анализ данных, персонализированное обучение, обучающиеся, интеллект, эффективность

BUILDING AN ALGORITHM FOR INTELLIGENT ANALYSIS OF BEHAVIORAL CLASSIFIERS FOR TRAINING PERSONALIZATION

Tlegenova T.E., Shardakov V.M.

Orenburg State University, Orenburg, e-mail: shardakov_vm@mail.osu.ru

In the context of active development of e-learning, one of the key areas of modern research is the personalization of education. Personalized learning allows you to adapt the educational process to a specific student, which maximally individualizes their educational needs, cognitive styles and career aspirations. In order to optimize a personalized learning system, the authors suggest using data mining to detect hidden patterns, relationships, and rules for analyzing, classifying, and predicting learning outcomes. In particular, the paper proposes to analyze the behavioral characteristics of students obtained from the results of a preliminary survey in order to detect hidden information in the results obtained through data mining, for which an algorithm for the intellectual analysis of behavioral classifiers has been developed for personalization of training. Thus, the range of functional tasks solved by data mining technologies is extremely wide, ranging from planning and conducting independent work of students to developing competencies. The key tasks in developing a data mining algorithm for personalized learning are: developing input content, building a normative model for the learner, and designing a mathematical model. To solve each of these tasks, it is necessary to formalize the input set of parameters, develop a subsystem for defining pedagogical methods and principles that allows you to build personalized learning. Using the developed data analysis algorithm, the authors predict the final grades of students and students at risk of failure.

Keywords: data analysis, personalized learning, learners, intelligence, efficiency

Развитие электронного обучения обозначено как обязательная составляющая формирования информационного пространства знаний в соответствии с Указом Президента РФ «О Стратегии развития информационного общества в Российской Федерации на 2017–2030 годы» [1].

В условиях активного развития электронного обучения одним из ключевых направлений современных исследований является персонализация образования. Персонализированное обучение позволяет адаптировать образовательный процесс под конкретного студента, что максимально

индивидуализирует его образовательные потребности, когнитивные стили и карьерные устремления.

Все большее количество образовательных учреждений во всём мире принимает тенденцию персонализации обучения, используя, как один из методов, интеллектуальный анализ данных.

Посредством применения интеллектуального анализа данных становится возможным обнаружение скрытых шаблонов, отношений и правил для анализа, классификации и прогнозирования в различных реальных условиях при персонализации обучения.

Так, исследователи из Шанхая предложили архитектуру персонализированной системы электронного обучения с применением кластерного подхода. Система электронного обучения обнаруживает и реагирует на содержание обучения в соответствии с когнитивными возможностями студентов и разделяет учащихся на различные группы или кластеры на основе их учебного поведения. На основе проведенного анализа методов кластеризации (K-средние, K-Medoids, DBSCAN, агломеративное иерархическое дерево кластеров, кластеризация путем быстрого поиска и нахождения пиков плотности (CFSFDP) было отмечено, что более надежные результаты могут быть достигнуты методом CFSFDP [2].

W. Villegas-Ch и S. Luján-Moga провели анализ методов интеллектуального анализа данных для персонализированного обучения, для повышения эффективности и результативности обучения, путем выявления закономерностей в успеваемости учащихся, данные предлагается извлекать из систем управления обучением (LMS), в частности платформы электронного обучения Moodle. Данные из этих систем могут быть оценены и преобразованы в полезную информацию с целью индивидуализировать образование и приспособить к потребностям каждого учащегося [3].

Группа ученых под руководством A. Segal представили новый подход к персонализации образовательного процесса на основе алгоритма EduRank, сочетающего в себе коллаборативную фильтрацию на основе метода ближайших соседей с методом ранжирования социальных предпочтений. EduRank строит рейтинг сложности для каждого учащегося, объединяя рейтинги похожих учеников, используя различные аспекты их успеваемости. Алгоритм был протестирован на двух наборах данных, содержащих тысячи студентов и миллион записей, и смог превзойти современные методы ранжирования, а также эксперта по предметной области [4].

В работе [5] авторы предлагают персонализированную модель электронного обучения на основе сервисно-ориентированной архитектуры, позволяющей учитывать особенности обучаемого, его профиль, и используют метод коллаборативной фильтрации для системы прогнозирования. В предлагаемой адаптивной системе применяется компетентностный подход, цель которого адаптировать индивидуальную траекторию обучения к характеристикам обучаемого, что в конечном итоге позволит приобрести желаемые компетенции.

Группа ученых [6] разработала новую модель рекомендации персонализированного пути обучения, основанную на методах кластеризации и машинного обучения. Основываясь на метрике сходства признаков учащихся, формируют группу и обучают на основе модели сети долгой краткосрочной памяти (LSTM), чтобы предсказать их пути обучения и производительность. Затем из результатов прогнозирования пути выбираются персонализированные пути обучения. Экспериментальные результаты показывают, что предлагаемые авторами методы дают надежные рекомендации по соответствующим путям обучения со значительно улучшенными результатами обучения с точки зрения точности и эффективности.

Цель исследования: разработка алгоритма интеллектуального анализа поведенческих классификаторов посредством оптимизации производительности персонализированной системы обучения.

Материалы и методы исследования

Нами была предложена модель процесса обучения, в которой все вычисления основываются на процентном соотношении компетенций, формируемых в высшем учебном заведении: универсальные (UK), общепрофессиональные (OPK) и профессиональные (PK). Данные соотношения можно представить в следующем виде:

$$P = 0,25 \cdot UK + 0,35 \cdot OPK + 0,4 \cdot PK. \quad (1)$$

Весовые коэффициенты выявлены посредством опроса предприятий и организаций. В данном опросе каждому представителю предприятия направлялась анкета, в которой необходимо было отметить ключевые факторы, которые необходимы предприятию при найме специалиста на работу.

Соотношение весовых коэффициентов UK, OPK, PK определяется при входном тестировании обучающихся. Каждый коэффициент определяется в диапазоне [0, 100]. Данный подход позволяет определить уровень подготовки студента.

К примеру, обучающийся может быть отличником по универсальным дисциплинам, но прогнозируемая оценка по дисциплинам, относящимся к профессиональной компетенции, у него может быть удовлетворительной, поэтому следующим шагом определим возможный класс обучающегося K:

$$K = \begin{cases} D, & P < 55 \\ C, & 55 \leq P < 70 \\ B, & 70 \leq P < 90 \\ A, & P \geq 90 \end{cases} \quad (2)$$

В соответствии с формулой (2) можно сделать вывод, что если обучающийся набрал менее 55 баллов, то он относится к классу D, т.е. к потенциальным студентам, которые не сдадут экзамен ни по одной дисциплине, относящейся к определенным компетенциям. Рекомендациями разработанной авторами системы будет являться необходимость дополнительных занятий для такого обучающегося и предоставление дополнительных методических материалов. При получении баллов обучающимся в диапазоне [55, 70] его можно отнести к классу C. Средний прогнозируемый уровень освоения компетенций является удовлетворительным. Рекомендациями такому классу обучающихся будет являться предоставление дополнительных материалов для изучения. К классу B отнесены значения, набранные в диапазоне [70, 90]. Уровень усвоения компетенций приравнивается к отметке «Хорошо». Для обучающихся данного класса возможны дополнительные разъяснения по определенной компетенции.

Обучающиеся, набравшие баллы в диапазоне [90, 100], приравниваются к классу A, и прогнозируется, что уровень освоения компетенций у обучающихся является отличным.

Средняя погрешность, которая возможна при определении класса обучающегося,

представлена среднеквадратическим отклонением каждого из возможных результатов тестирования:

$$S = \sum_{i=1}^n \frac{1}{P} S_i, \quad (3)$$

где S_i – среднеквадратическое отклонение на интервале i , оно находится для каждого из возможных результатов по формуле:

$$S_i = \sum_{j=0}^n \frac{1}{UK_j} \cdot \sum_{j=0}^m \frac{1}{OPK_j} \cdot \sum_{j=0}^l \frac{1}{PK_j} \cdot \sum_{k=1}^r S_k, \quad (4)$$

S_k – среднеквадратическое отклонение результата тестирования, которое вычисляется по формуле:

$$S_k = \sqrt{\sum_{i=1}^n \frac{(x_i - \bar{x}_i)^2}{n-1}}, \quad (5)$$

x_i – результат тестирования, \bar{x}_i – среднее значение между минимальным и максимальным оцениванием, n – количество результатов тестирования с оценкой k .

На основе описанной модели авторами разработан алгоритм интеллектуального анализа поведенческих классификаторов при персонализации обучения, который показан на рис. 1.

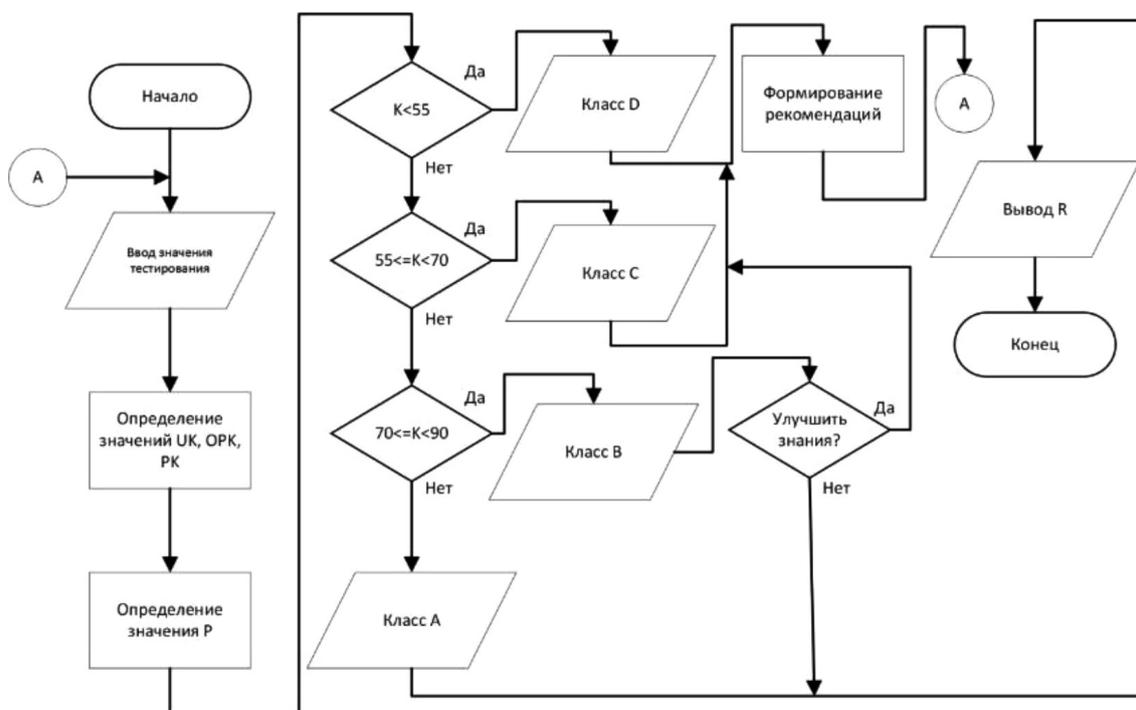


Рис. 1. Алгоритм интеллектуального анализа поведенческих классификаторов при персонализации обучения

Начальным значением, которое получает система для обработки информации, является проведенное преподавателем входное тестирование. Посредством обработанной информации определяются компетенции: универсальные, общепрофессиональные и профессиональные. Следующим этапом является определение его класса (диапазон значения является принятым для пороговых оценок в Оренбургском государственном университете). Считаем, что студентам класса С и D для успешной сдачи дисциплин в обязательном порядке необходимо предоставлять дополнительный материал и проводить дополнительные занятия. После прохождения дополнительных занятий обучающиеся проходят новое тестирование. Классу В предоставлен выбор, же-

лают ли они подтянуть свои знания, чтобы перейти в класс А, который считается, что обучающийся в нём знает более 90 процентов материала.

Результаты исследования и их обсуждение

Для того чтобы получить значения P , авторами работы было проведено входное тестирование обучающихся посредством Гугл-формы (рис. 2), с помощью которого определены значения UK , OPK , PK .

Итоговые результаты экспортируются в формат Microsoft Excel, с указанием номера зачетной книжки и баллов после прохождения тестирования по всем компетенциям и соотношения компетенций. Результаты входного тестирования показаны на рис. 3.

Рис. 2. Пример разработанных тестов

Номер зачетной книжки (ID)	КОМПЕТЕНЦИИ			Количество баллов (P)	Класс (K)
	универсальные (UK)	общепрофессиональные (OPK)	профессиональные (PK)		
	Весовой коэффициент				
	0,25	0,35	0,4		
638200	16	37	41	33,35	D
637559	37	30	35	33,75	D
635690	30	47	36	38,35	D
638081	36	40	37	37,8	D
635240	40	48	48	46	D
616295	59	68	56	60,95	C
616281	82	81	75	78,85	B
638183	80	89	90	87,15	B
615031	94	91	93	92,55	A
638185	92	96	93	93,8	A

Рис. 3. Результаты входного тестирования

Номер зачетной книжки (ID)	КОМПЕТЕНЦИИ			Количество баллов (P)	Класс (K)
	универсальные (UK)	общепрофессиональные (OPK)	профессиональные (PK)		
	Весовой коэффициент				
	0,25	0,35	0,4		
638200	60	59	61	60,05	C
637559	49	67	54	57,3	C
635690	58	77	75	71,45	B
638081	63	80	68	70,95	B
635240	68	77	72	72,75	B
616295	75	75	72	73,8	B
616281	92	88	91	90,2	A

Рис. 4. Результаты повторного тестирования

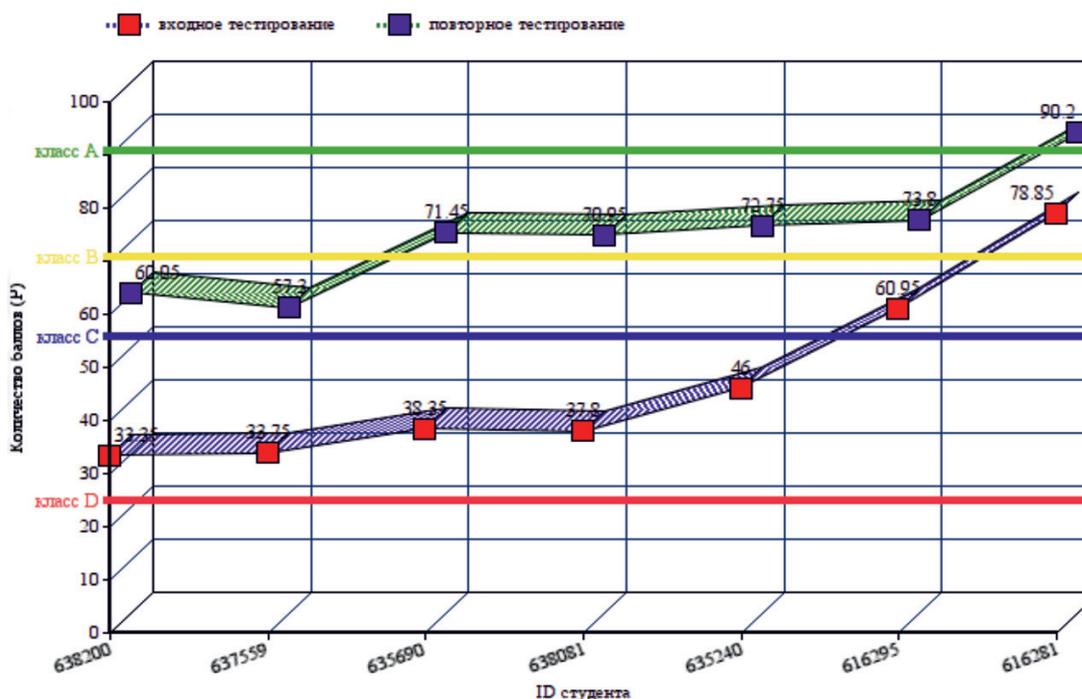


Рис. 5. График изменения академической успеваемости группы

Согласно каждому классу в зависимости от тех компетенций, которые вызывают проблему, авторами разработаны методические рекомендации, соответствующие определенному классу обучающихся. После чего проводится новое тестирование для обучающихся класса C, D и для изымавшихся обучающихся класса B. На рис. 4 представлены результаты изменения уровня освоения универсальных, общепрофессиональных и профессиональных компетенций.

В ходе эксперимента наблюдалась положительная динамика изменения академической

успеваемости студентов, представленная на рис. 5.

Заключение

В представленной работе проведен интеллектуальный анализ данных посредством применения персонализированного обучения, акцентируя внимание на образовательные компетенции. Входными данными являются характеристики студентов, полученные в результате предварительного тестирования посредством оценки уровня их компетентности. На основе полученных значений и разработанного алгоритма ана-

лиза данных можно получить классификаторы студентов, с помощью которых каждого студента предварительно система относит к определенной группе: отличников (A) или к группе студентов, которым необходимо серьезно улучшить свои знания (D). Разработаны методические рекомендации для студентов и показаны результаты эффективности предложенного метода. Используя разработанный алгоритм анализа данных, авторы прогнозируют итоговые оценки студентов и тех, кто рискует провалить экзамен. Полученные результаты свидетельствуют о повышении эффективности процесса обучения, ориентированного на формирование образовательных компетенций посредством применения интеллектуального анализа данных, и рекомендованы для внедрения при построении и планировании образовательной программы.

Список литературы

1. Указ Президента РФ от 9 мая 2017 г. № 203 «О Стратегии развития информационного общества в Российской Федерации на 2017–2030 годы» [Электронный ресурс]. URL: <http://www.garant.ru/products/ipo/prime/doc/71570570/#ixzz4jhpjoKJn> (дата обращения: 24.05.2020).
2. Kausar S., Huahu X., Hussain I., Wenhao Z., Zahid M. Integration of Data Mining Clustering Approach in the Personalized E-Learning System. In proceedings IEEE Access. 2018. Vol. 6. P. 72724–72734.
3. Villegas-Ch W., Luján-Mora S. Analysis of data mining techniques applied to LMS for personalized education. In proceedings IEEE World Engineering Education Conference (EDUNINE). Santos. 2017. Vol. 64. P. 85–89.
4. Segal A., Gal K., Shani G., Shapira B. A difficulty ranking approach to personalization in E-learning. In proceedings International Journal of Human-Computer Studies. 2019. Vol. 130. P. 261–272.
5. Bendahmane M., Falaki B.E., Benattou M. Toward a Personalized learning Path through a Services-Oriented Approach. In proceedings International Journal of Emerging Technologies in Learning (IJET). 2019. Vol. 14. No. 15. P. 52–66.
6. Zhou Y., Huang C., Hu Q., Zhu J., Tang Y. Personalized learning full-path recommendation model based on LSTM neural networks. In proceedings Information Sciences. 2018. Vol. 444. P. 135–152.