Потенциал цифровых технологий и искусственного интеллекта в управлении персонализированным обучением

Е. Е. Котова
Санкт-Петербургский государственный электротехнический университет «ЛЭТИ» им. В.И. Ульянова (Ленина)
eekotova@gmail.com

***Аннотация*. Развитие цифровых технологий открывает новые возможности для управления персонализированным обучением с применением интеллектуальных методов анализа данных. Целью всестороннего анализа данных, получаемых в процессе обучения, является создание управляемой среды обучения с учетом индивидуальных потребностей, целей и личностных особенностей обучающихся. В ходе нашего исследования были проанализированы различные подходы к организации образовательного процесса, основанные на применении интеллектуальных технологий. На основе изучения передового опыта применения новых цифровых инструментов для моделирования процессов обучения на основе анализа данных, в докладе представлена разработанная автоматизированная система интеллектуальной поддержки персонализированного обучения.**

Ключевые слова: цифровые технологии, трансформация образования, методы анализа данных, моделирование прогресса обучения, прогнозирование результатов

I. Введение

Персонализированное обучение еще не распространено в российских вузах несмотря на то, что призыв к персонализации процесса обучения, к управлению индивидуальным обучением, представлению индивидуальных траекторий достаточно популярен в академическом сообществе. В реальном образовательном процессе подход персонализированного обучения влечет за собой серьезные проблемы как для учащихся, которым придется научиться обучаться большей частью самостоятельно, так и для преподавателей, которым необходимо будет их поддерживать. Обучение сохраняет массовый характер, и эта форма достаточно стабильна в высшей школе, в то время, как контингент становится все более разнородным по уровню подготовки, потребностям, по целям и задачам, по личностным и когнитивным характеристикам познавательной, мыслительной сферы. Данное противоречие является практически основным, наряду с противоречием между большим потенциалом технологий, беспрецедентным внедрением цифровых средств во все аспекты образовательного процесса и недостаточными возможностями традиционных дидактических средств по их использованию при оказании эффективной индивидуальной поддержки и руководства студентами.

II. Основные подходы к исследованию персонализированного обучения

С тех пор, как персонализированное обучение с использованием технологий было определено как одно из новых областей, прошло порядка 30 лет. Вместе с тем, экспоненциальное увеличение доступности цифровых данных и вычислительных мощностей за последнее десятилетие способствовало притоку новых исследований в направлении интеллектуального анализа образовательных данных и аналитики обучения (learning analytics) [1].

Влияние технологий меняет понимание персонализированного обучения. То, что в ранних исследованиях называлось обучением, ориентированным на учащегося, индивидуальным обучением дифференцированным обучением, обучением, основанным на компетенциях, к настоящему времени уже включает проектирование гибких индивидуальных планов, проектное обучение, управляемое с применением интеллектуальных и адаптивных технологий.

Многочисленные описания персонализированного обучения в публикациях частично совпадают, иногда противоречат, но имеют много общих черт [2]. К ним относятся:

- цели – повышение вовлеченности и успеваемости учащихся;
- дифференциация – акцент на удовлетворении индивидуальных учебных потребностей, интересов, устремлений, культурных традиций;
- гибкость – способность меняться в соответствии с меняющимися потребностями;
- переменный темп – признание того, что люди прогрессируют по-разному.

Все перечисленное говорит о сложности персонализированного обучения, реализация которого практически затруднительна без применения передовых технологий.

В аналитическом докладе Юнеско и Шанхайского Открытого Университета (The UNESCO Institute for Information Technologies in Education (UNESCO IITE) and Shanghai Open University (SOU), 2022 год [3]) об использовании передовых технологий ИКТ/ИИ для цифровой трансформации образования отмечаются изменения, которые происходят в методах преподавания и обучения, начиная от организации обучения с

использованием простейших цифровых инструментов и заканчивая внедрением персонализированных обучающих платформ.

На основе анализа современных тенденций в образовании отмечаются цифровые технологии, обладающие наибольшим потенциалом в цифровой трансформации образования:

- искусственный интеллект (artificial intelligence, AI);
- виртуальная реальность (virtual reality, VR);
- дополненная реальность (augmented reality, AR);
- Интернет вещей (Internet of Thing, IoT);
- большие данные (big data) и
- блокчейн (blockchain).

Использование методов цифровых технологий открывает новые возможности для оценки не только знаний и компетенций учащихся, но и других важнейших характеристик интеллектуальной и познавательной деятельности человека с применением интеллектуальных методов анализа данных и методов инженерии знаний.

Центральными компонентом в системах искусственного интеллекта в образовании (Artificial Intelligence in Education, AIEd) являются модели предметной области обучения, предназначенные для представления знаний и модели учащихся, предназначенные для использования в управлении ростом знаний отдельных учащихся.

Исследователи сосредоточились на двух основных подходах, поддерживающих открытые исследования – подход, основанный на знаниях и подход, основанный на данных, которые также могут быть объединены [4].

В подходах, основанных на знаниях, модели учащихся основаны на подробном анализе целевого исследовательского взаимодействия, предметной области и соответствующих когнитивных и метакогнитивных способностей, тогда как подходы, основанные на данных, основаны на методах интеллектуального анализа данных и машинного обучения для автоматического определения соответствующих качеств учащегося.

Отмечаются различия в подходах к персонализированному обучению в разных странах. В США – нацеленность на индивидуализм, политику и культуру, в Европе – на личностное развитие, развитие социальных и эмоциональных навыков [2]. Также различаются подходы и к трактовке компетенций. И, если определенные компетенции предлагается измерять с помощью компьютерного тестирования (что достаточно развито в США), то навыки, на которые больший акцент делается в Европейских странах, редко поддаются оценке с множественным выбором в системах тестирования [2]. Многие авторы используют термины «индивидуализированное обучение», «дифференциация» и «индивидуализация» как синонимы. Поэтому часто используется более общее понятие «индивидуализированной среды обучения» (“personalised learning environments” – PLE).

Подчеркивается важность эффективной организации образовательного процесса в цифровой среде. В то же время отмечены основные вопросы, которые требуют решений [1]:

- консолидации разнородных образовательных ресурсов;
- управления дистанционным обучением;
- адаптивного управления образовательными траекториями и
- персонифицированного подбора курсов.

III. Задачи ителлектуального анализа образовательных данных

Одни из первых, так называемые интегрированные системы обучения (Integrated Learning Systems, ILS), появились после 80-х годов, и считалось, что «система должна быть интегрирована, или приспособлена, к обычной практике учителя» [2]. В последующие годы благодаря развитию технологий стали разрабатываться интеллектуальные системы обучения (Intelligent Tutoring Systems, ITS), исследовательские среды обучения (Exploratory Learning Environments, ELE), интеллектуальные системы управления обучением (Smart Learning Management Systems, LMS), обучающие сети (Learning Network Orchestrators), системы обучения на основе игр (Digital Games-based Learning). С целью обработки накопленных данных возникла необходимость использовать аналитику данных и методы машинного обучения.

Алгоритмы машинного обучения предназначены для анализа больших наборов данных, чтобы обнаружить — или «изучить» скрытые правила и шаблоны, которые могут помочь принять необходимое решение в будущем [5]. Значительные достижения в области машинного обучения существенно расширили диапазон методов и моделей ИИ для достижения целей обучения [1].

Приведем примеры применения некоторых подходов к построению моделей образовательных процессов.

В задачах интеллектуального анализа образовательных данных используются разные подходы к построению моделей. Например, используется вероятностная модель (Bayesian Knowledge Tracing, BKT). BKT предлагает динамический процесс моделирования вероятности того, что учащийся обладает знаниями каждого навыка на протяжении обучения [6 и др.]. Другой подход к моделированию текущего состояния знаний учащихся основан на теории тестирования (Item Response Theory, IRT). Оба подхода направлены на извлечение информации и понимание закономерностей из данных об учащихся, полученных в ходе учебного процесса. IRT в первую очередь фокусируется на оценке способностей и сложности заданий. BKT моделирует динамику обучения и включает параметры, связанные с начальным уровнем знания, скоростью обучения, вероятностью угадывания и вероятностью ошибки в тестировании [6]. Байесовское отслеживание знаний (BKT) является самым популярным подходом к построению темпоральных моделей обучения студентов [7]. Однако, эти модели не учитывают, как различия в обучении могут влиять на результаты обучения [6].

Другие модели предлагают общую структуру нейронной когнитивной диагностики (NeuralCD) для определения уровня владения учащимися конкретными концепциями знаний, где целью когнитивной

диагностики является обнаружение владения навыками, например, по результатам тестов [8].

Для моделирования сложных данных отслеживания знаний учащихся могут быть применены методы глубокого обучения (Deep Knowledge Tracing, DKT). DKT обучает рекуррентную нейронную сеть взаимодействию учащихся, чтобы моделировать состояния скрытых знаний и прогнозировать их будущую производительность [7]. Цель обучения по отслеживанию знаний заключается в долгосрочном прогнозировании будущей успеваемости учащегося на основе их прошлой деятельности.

Область интеллектуального анализа образовательных данных и аналитики обучения постоянно развивается. Изучением факторов, влияющих на результаты обучения (например, в виде средних баллов) занимаются ученые во многих университетах. Исследования по прогнозированию результатов обучения, которые можно привести в пример и в наши дни, уже проводились в Калифорнийском университете, начиная с 80-х годов [9, 10]. Несмотря на различающиеся данные о контингенте обучающихся (например, этническая принадлежность, социально-экономическое положение, доход семьи и многие другие в [9–11]), методы исследования используются подобные – это статистика, регрессионный анализ, машинное обучение: линейная регрессия, Linear Regression; регрессия дерева решений, Decision Tree; регрессия случайного леса, Random Forest Regression; регрессия XGBoost; нейронные сети, neural network, NN.

Среди задачи прогнозирования также исследуется влияния иных, кроме стандартизированного тестирования успеваемости, дополнительных факторов, повышающих прогностическую достоверность.

IV. Опыт разработки автоматизированной системы интеллектуальной поддержки персонализированного обучения

Система, которую мы называем Open-OntoMASTER, последовательно разрабатывалась, начиная с версии исследовательского прототипа [12–14], в течение нескольких последних лет с полевыми испытаниями в учебном процессе и с учетом отзывов студентов, преподавателей, администраторов и исследователей. В программный комплекс включена среда аналитических исследований, среда тестирования, подсистема экспресс-диагностирования когнитивного потенциала, подсистема структурирования методических материалов, построения учебных онтологий, управления когнитивной нагрузкой и поддержки распределения дидактических ресурсов. Поддерживается совместимость с системой Moodle (moodle.otg), совместимость с редактором онтологий Protégé (protege.stanford.edu), средствами аналитики. К настоящему времени по мере отладки исследовательского прототипа возникла необходимость создания удобных информационных панелей для пользователей системы, с наличием виртуального агента-помощника, в связи с чем вводится новая надстройка в виде встроенной подсистемы-помощника Chat-help для создания упреждающей онлайн-помощи.

Создание открытых образовательных ресурсов и подсказок в рамках адаптивных обучающих систем требует определенных усилий в первую очередь со стороны экспертов-преподавателей и методистов, задачей которых является решение, возможно, с некоторой технической поддержкой, как дифференцировать дидактические ресурсы, адаптировать к потребностям отдельного учащегося [2].

Интерфейс системы Open-OntoMASTER изображен на рис. 1.

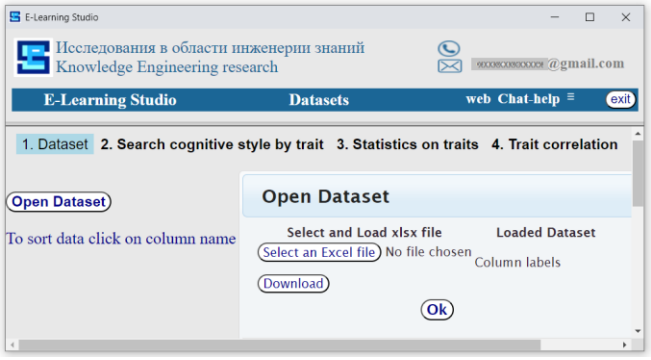

Рис. 1. Пример интерфейса программного комплекса Open-OntoMASTER

Функции следующие:

Представление банка данных. Выбор файла из банка данных (из системы обучения, например, Moodle, из системы экспресс-диагностирования когнитивно-стилевого потенциала, из любого файла в формате Excel-таблицы).

Анализ данных по запросам. Выбор параметров и метода анализа: статистика, корреляционный анализ, регрессионный анализ, метод главных компонент.

Визуализация результатов. Поддержка персонализированной динамической визуализации прогресса обучения.

V. Пример исследования рейтинга обучающихся на начальном этапе обучения

Дизайн процесса обучения смоделирован в виде правил и представлен в виде электронной таблицы (рис. 2 и 3).

	1 НЕДЕЛЯ		04.09.23-10.09.23			
	Представление курса	Дата	Абстракт-Инженерия знаний-2023	Дата	Краткие правила обучения	Дата
Ст 1	3	08.09.23	3	07.09.23	3	10.09.23
Ст 2	0		2,5	18.09.22	3	10.09.23
Ст 3	0		2,5	09.09.22	3	06.09.23
Ст 4	0				3	06.09.23
Ст 5	3	06.09.23	3	06.09.23	3	06.09.23
Ст 6	2	11.09.23	3	07.09.23		
Ст 7	3	06.09.23	3	06.09.23	3	07.09.23
Ст 8	3	06.09.23	3	06.09.23	0	
Ст 9	3	07.09.23	3	10.09.23	1	13.09.23

Рис. 2. Пример (фрагмент) дизайна открытой модели ученика для дисциплины «Инженерия знаний» первой недели обучения

В столбцах таблицы перечислены виды деятельности и типы оценивания, которые для лучшего восприятия отмечены различными цветами с применением цветового кодирования. Участие и индивидуальные достижения обучающихся фиксируются и оцениваются баллами. Компоненты модели студента представлены в каждой ячейке таблицы.

Например, на первой неделе обучения студенту необходимо ознакомиться с представлением курса (просмотреть видео), самостоятельно изучить материал «Абстракт курса», изучить краткие правила обучения с указанием рекомендованных сроков (рис. 2).

На второй неделе обучения необходимо приступить к изучению теоретического материала первого раздела курса, изучить необходимые понятия, выполнить первое практическое задание в установленные сроки (рис. 3). При нарушении сроков вводятся условные «штрафы», но и при опережении сроков или при выполнении заданий на высокие баллы вводятся «бонусы», поощрения. Все условия прописываются в правилах обучения.

2 НЕДЕЛЯ		04.09.23-10.09.23		
Материалы раздела 1	Необходимые понятия для изучения Раздела 1	1.1.1. Место и роль онтологии в Семантическом Веб	Задание 1	1.1.2. Разметка данных Data labeling
3	3	3	10	3
3	3	3	8	
3	3	3	4	
3	3	3	2	
3	3	3	4	
0	0	3	0	
3	2	2	6	
3	2	0	10	
0	2	0	10	

Рис. 3. Пример (фрагмент) дизайна открытой модели ученика для дисциплины «Инженерия знаний» второй недели обучения

После каждого действия студента пересчитывается его рейтинг. Заметно, что Ст1 опережает план обучения (выполнил изучение раздела 1.1.2, оценено максимальным баллом), поэтому на текущий момент его деятельность оценивается выше 100 % и он занимает первую строчку в рейтинге на второй неделе обучения (рис. 4). На каждом шаге студент может оценить свой прогресс, сравнить с общими показателями группы, получить рекомендации от системы по выполнению индивидуальных заданий с целью повышения своего уровня.

	Рейтинг	
	балл	%
Ст 1	31	110,7
Ст 2	22,5	80,4
Ст 5	22	78,6
Ст 7	22	78,6
Ст 8	21	75,0
Ст 9	19	67,9
Ст 3	18,5	66,1
Ст 4	14	50,0
Ст 6	8	28,6
.....	....	.....
Среднее	19,8	70,6
Max	28	100,0

Рис. 4. Пример отражения динамического рейтинга по группе

Формируется текущий рейтинг в виде графиков (рис. 5).

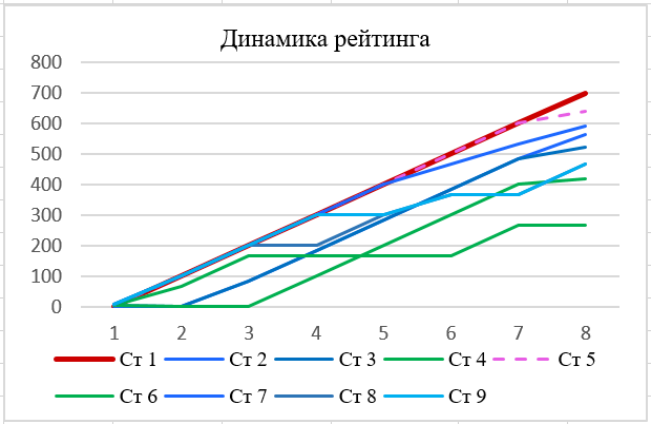

Рис. 5. Пример динамики рейтинга по группе

На рис. 2 и 3 приведены фрагменты заполнения таблиц данных. Так как формат обучения смешанный, т. е. онлайн плюс работа в аудитории, плюс самостоятельная работа, работа в группах, первичные данные для каждого учащегося поступают из разных источников – из системы Moodle (порядка 140 фиксируемых данных), из системы ОнтоМАСТЕР (50 данных), из Zoom (порядка 20–30 значений), из работы в аудитории и др. Дополнительно фиксируются так называемые мягкие навыки, проявленные в процессе учебной деятельности. Ориентировочно для группы из 25 человек по одной дисциплине формируется таблица, включающая 6000 – 8000 первичных данных (данные низкого уровня). Соответственно для трех групп, которые учатся в одном потоке, собирается набор более 24000 данных. Приведены примеры (фрагменты) таблиц самого «низкого» уровня сбора данных. На следующих уровнях представлены данные, полученные статистическими методами и методами анализа (рис. 6 – пример текущего рейтинга).

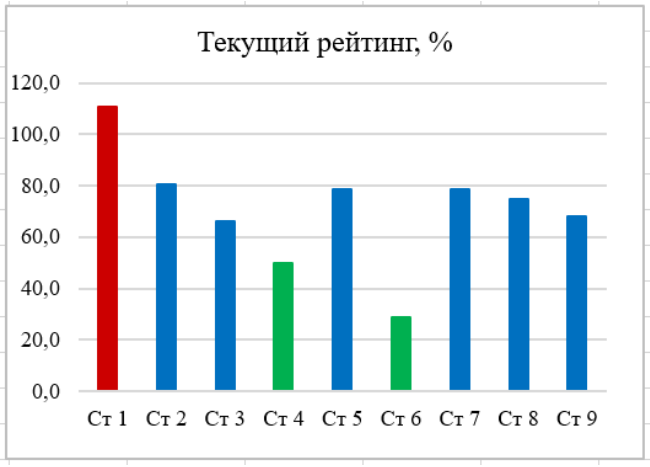

Рис. 6. Пример текущего рейтинга по группе

В примере представлен анализ рейтинга студентов, целью которого является выявление на ранних этапах обучения студентов группы риска. В данном случае мы наблюдаем разный прогресс: студенты Ст4 и Ст6 попадают в группу риска после первой недели обучения.

Визуальное представление позволяет учащемуся увидеть свой прогресс по неоцененным аспектам, которые можно наверстать в любую неделю. Подробная информация о других компонентах в настоящее время скрыта, она может быть получена по специальному

запросу в зависимости от исследовательского или образовательного контекста и целей анализа.

Помимо оценивания академической успеваемости в системе вводится система оценивания когнитивных характеристик деятельности, к которым относятся показатели работы с информацией (полнота и точность восприятия, имплицитная обучаемость, понимание инструкций, визуальное восприятие, способность осваивать сложные навыки, самодисциплина, исполнительные функции, и др.). Включение дополнительных данных о студентах, может повысить точность прогнозов.

С применением методов анализа данных массив данных существенно увеличивается, так как решаются задачи получения новых данных, таких как прогнозирование результатов обучения по каждому студенту.

VI. Заключение

Система Open-OntoMASTER представляет интегрированную технологическую основу, которая включает банк данных, аналитические методы и инструменты Learning Analytics Tools. В системе на основе принципов аналитики данных и практики обучения обеспечивается:

1. инфраструктура для обработки данных;
2. поддержка процесса обучения с помощью журнала планирования и демонстрации результатов в динамике;
3. обмен данными;
4. визуализация результатов.

Предложенный подход дизайна обучения (проектирования учебного плана) по конкретным дисциплинам проиллюстрирован примером из практики. Процесс проектирования ориентирован в первую очередь на учащегося и на его поддержку для получения более высоких результатов обучения, но также и на преподавателей и исследователей.

В связи с постоянным увеличением разнообразных данных обучения, в дальнейшем необходимо продолжить исследования по применению методов анализа. Также необходимо изучить возможность расширения упреждающего чата для активного взаимодействия с пользователями.

Благодарность

Автор выражает благодарность к.т.н., доц. Писареву А.С. за оказанную помощь при проведении данного исследования и практической реализации автоматизированной среды аналитики образовательных данных Open-OntoMASTER.

Список литературы

[1] Eynon R. The Future Trajectory of the AIED Community: Defining the ‘Knowledge Tradition’in Critical Times. International Journal of Artificial Intelligence in Education. 2023. Jul 5. Pp. 1-6. Available at: URL: https://link.springer.com/article/10.1007/s40593-023-00354-1

[2] Holmes W, Anastopoulou S, Schaumburg H, Mavrikis M. Technology-enhanced personalised learning: Untangling the evidence. 113 p. Available at: URL: https://oro.open.ac.uk/56692/1/TEPL_en.pdf

[3] Analytical Report on the Use of Advanced ICT/AI for Digital Transformation of Education Available at: URL: https://iite.unesco.org/publications/analytical-report-on-the-use-of-advanced-ict-ai-for-digital-transformation-of-education/

[4] du Boulay B., Mitrovic A., Yacef K., editors. Handbook of Artificial Intelligence in Education. Edward Elgar Publishing; 2023. 696 p.

[5] Gillani N., Eynon R., Chiabaut C., & Finkel K. Unpacking the “Black Box” of AI in Education. Educational Technology & Society, 2023. Vol. 26. No. 1. Pp. 99-111. Available at: URL: https://doi.org/10.30191/ETS.202301_26(1).0008

[6] Bulut O., Shin J., Yildirim-Erbasli S.N., Gorgun G., Pardos Z.A. An Introduction to Bayesian Knowledge Tracing with pyBKT. Psych. 2023. Vol. 5(3). Pp. 770-786. Available at: URL: https://doi.org/10.3390/psych5030050

[7] Piech C., Bassen J., Huang J., et al. Deep knowledge tracing. Advances in neural information processing systems. 2015. 8 p. Available at: URL: https://proceedings.neurips.cc/paper/2015

[8] Wang F., Liu Q., Chen E., Huang Z., Yin Y., Wang S., Su Y. NeuralCD: a general framework for cognitive diagnosis. IEEE Transactions on Knowledge and Data Engineering. 2022 Aug 23. Available at: https://ieeexplore.ieee.org/document/9865139

[9] Geiser S., Santelices M.V. Validity of High-School Grades in Predicting Student Success beyond the Freshman Year: High-School Record vs. Standardized Tests as Indicators of Four-Year College Outcomes. Research & Occasional Paper Series: CSHE. 6.07. Center for studies in higher education. 2007. Available at: https://files.eric.ed.gov/fulltext/ED502858.pdf

[10] Goldman R.D., Widawski M.H. An analysis of types of errors in the selection of minority college students. Journal of Educational Measurement. 1976. Vol. 13. No. 3. Pp. 185-200.

[11] Kurlaender M., Cohen K. Predicting College Success: How Do Different High School Assessments Measure Up?. Policy Analysis for California Education. PACE. 2019. Available at:URL: https://files.eric.ed.gov/fulltext/ED594712.pdf

[12] Котова Е.Е. База данных когнитивно-стилевого потенциала студентов AnalyticSYSTEMS-hDS (StudCSP-DataSet). Свидетельство о государственной регистрации базы данных № 2019621974 от 30.10.2019 г

[13] Котова Е.Е., Писарев А.С. Программа интеллектуального анализа продуктивности решения когнитивных задач в электронной среде (Эксперт-Аналитик ART). Свидетельство о государственной регистрации программы для ЭВМ № 2020667345 от 22.12.2020 г.

[14] Котова Е.Е., Писарев А.С., Писарев И.А. Программа оценки и анализа групповых результатов решения серий перцептивно-когнитивных задач в электронной среде «E-LEARNING-4E». Свидетельство о государственной регистрации программы для ЭВМ № 2021669274 от 25.11.2021 г.