

Многофакторный анализ влияния особенностей студентов на академическую успеваемость для персонализации образовательной среды

Т. В. Авдеенко, Н. В. Пустовалова¹

Современные университеты, разрабатывая цифровую среду, опираются на студенто-центрический подход, актуальные педагогические и дидактические практики, что позволяет учитывать индивидуальные особенности студентов и адаптировать к ним образовательный процесс. Персонализация образовательной среды, таким образом, призвана улучшить качество образовательного процесса, подняв его на новый уровень. В данной статье приведены результаты применения методов многофакторного моделирования для выявления общих связей и закономерностей между успеваемостью студентов и их психометрическими характеристиками. Исследуемая выборка включает результаты тестирования студентов Новосибирского государственного технического университета (НГТУ), определяющие когнитивные, мотивационные и личностные особенности обучающихся, а также информацию о всех прошедших аттестационных испытаниях, предусмотренных учебным планом. Отличительной чертой данной работы является разнообразие тестов, которым подвергались студенты, с целью найти сочетания составляющих их подтестов (факторов), наиболее сильно влияющих на успешность обучения, для последующего использования выявленных закономерностей при проектировании индивидуальных траекторий, например, выдачи образовательного контента с учетом особенности личности студента.

Ключевые слова: персональная образовательная среда, регрессионный анализ, факторный анализ, структурное моделирование, модель обучаемого, психометрические характеристики.

1. Введение

Сегодня в свете таких глобальных тенденций, как студенто-центрический подход, гуманизация, инклюзивное образование [1–3] перед вузами стоит задача не просто создать цифровую среду, а сделать ее комфортной и дружелюбной для пользователей. Современное видение процесса подготовки кадров предполагает, что университеты не просто дают определенный набор сведений, а позволяют каждому обучаемому сформировать систему знаний, умений и навыков (hard, soft и digital skills), обеспечивающих его дальнейшее личностное и профессиональное развитие [4]. Именно персонализация образовательной среды должна улучшить качество образовательного процесса, поднять его на новый уровень.

На сегодняшний день все университеты используют специализированные программы для автоматизации своей деятельности, в том числе для оценки учебных достижений студентов, контроля посещаемости, самодиагностики, сбора и анализа отзывов. На рынке существует много различных программных систем, которые можно приобрести «под ключ» (готовые решения от 1С), или же конструкторы (Битрикс), или пакеты прикладных программ с базовым

¹ Работа выполнена при финансовой поддержке Министерства науки и высшего образования в рамках госзадания (проект № FSUN-2020-0009).

функционалом и открытым кодом, которые можно доработать под свои нужды (MOODLE). Вместе с тем простой автоматизации уже недостаточно.

Современный образовательный процесс уделяет все больше внимания личности обучаемого, его индивидуальным особенностям и потребностям. Персонализированная образовательная среда (Personal learning environment, сокращенно PLE) – «совокупность объектов образовательного процесса, полученная из информационно-коммуникационной образовательной среды путем адаптации в соответствии с целями, содержанием и планируемыми результатами обучения, потребностями и способностями обучаемого и выступающая средством персонализации личности» [5]. Персонализация цифровой образовательной среды происходит параллельно с цифровизацией, внедрением дистанционного обучения и виртуальных технологий. PLE важна не только для обучаемых, но также для аналитиков и исследователей. Данные о персональных особенностях студентов прогрессивные вузы (например, Университет 2035, МИФИ и др.) используют в педагогическом дизайне образовательных программ и результатов. Так, при регистрации на платформе «Университет 2035» обучающимся предлагается пройти входную диагностику. Эта диагностика включает тесты на скорость реакции, умение адаптироваться, мотивацию, стиль взаимодействия с окружающими и т.д. Такие платформы, как «Coursera» используют данные о том, как слушатели проходят разные курсы, для того чтобы управлять контентом и определенным образом мотивировать обучающихся для завершения изучения курса.

В связи с этим перед разработчиками информационных систем для образовательных организаций встает вопрос о том, какие данные о студентах нужно собирать. Ведь неверно сформулированные требования к входной информации могут привести к тому, что система будет оперировать большими объемами данных, что ведет к увеличению затрат, и при этом не будет предоставлять адекватной информации для принятия управленческих решений. Поэтому очень важно тщательно спроектировать входные данные, т.е. отобрать те индивидуальные характеристики обучаемых, которые будут учитываться университетом для улучшения своей деятельности, а также методы их определения и оценки.

Статья организована следующим образом. Раздел 2 содержит постановку задачи, объясняя важность выбора показателей для формирования PLE. В разделе 3 описаны результаты регрессионного анализа данных, полученных в результате тестирования студентов. Раздел 4 включает результаты факторного анализа и структурного моделирования уравнений. Раздел 5 – заключение.

2. Проектирование модели обучаемого для PLE с использованием данных психометрических измерений студентов

Анализ источников [6–8] показал, что для оценки индивидуальных особенностей обучаемых широко используются психометрические характеристики студентов (скорость реакции, эмоциональный интеллект, IQ и т.д.). Этому способствует тот факт, что психология предоставляет целый набор стандартизированных тестов, проводимых по устоявшимся методикам, которые позволяют оценивать некоторые параметры для выборки обучаемых, анализировать и сравнивать их между собой (Амтхауэр, Равен, Кеттл, Айзенк и др). Разные авторы пытаются анализировать самые разные зависимости, для чего широко используется аппарат прикладной статистики.

Проанализировав работы современных исследователей, приходим к выводу, что для совершенствования учебного процесса необходимо правильно выявить и оценить множество психологических факторов, положительно или отрицательно влияющих на успешность образовательной деятельности студентов. Эти сведения необходимы для построения модели обучаемого, которая лежит в основе архитектуры интеллектуальных обучающих систем. Вместе с ней также используются модель тьютора, содержащая сведения о том, как обучать, и модель эксперта, включающая сведения о том, чему обучать [9]. Для создания модели обучаемого

необходимо определить те измеримые характеристики, которые можно собрать в рамках образовательного процесса и которые будут содержать значимую информацию для принятия управленческих решений в ходе его реализации.

При выполнении исследования мы сталкивались с большим количеством различных показателей, которые описывают некие ненаблюдаемые конструкции, включающие в себя наблюдаемые переменные. В качестве примера можно привести понятие «интеллект», его нельзя измерить напрямую, вместо этого исследователи разрабатывают гипотезу об интеллекте и создают инструменты измерения в виде вопросов или заданий, предназначенных для его оценки в соответствии с их гипотезой. В нашем случае имеется множество характеристик, измеренных тестовыми заданиями, объединение которых в более общие факторы может способствовать более качественной интерпретации их влияния на успеваемость студента. Для выявления скрытых факторов, отвечающих за объединение сильно коррелированных наблюдаемых переменных, было принято решение использовать методы факторного анализа.

В ходе исследования психологических особенностей личности определенную сложность представляет собой интерпретация результатов, зависящая от обоснованного структурирования признакового пространства. Последнее осложняется тем, что в области психологии до сих пор нет единого устоявшегося определения того, что такое интеллект и какова его структура. Однако на основании анализа источников и сравнения различных тестовых материалов было принято решение выделить в структуре интеллекта конструкты, отвечающие за когнитивные (память, скорость реакции и т.д.), личностные (когнитивный стиль, нейротизм и т.д.) и мотивационные (стремление к избеганию неудач и т.д.) характеристики. Представляется, что такое разделение позволит выбрать из показателей, полученных в результате выполнения различных тестов, некоторый набор, отражающий значимые составляющие интеллекта, отвечающие не только за «вычислительные» способности человека, но и за его умение работать в команде, ставить цели и добиваться их, что соответствует современной точке зрения на образовательный процесс и отражает необходимость развивать не только *hard*, но и *soft skills*. Эта парадигма заложена в основу модели обучаемого, разрабатываемой для проектирования и последующей реализации PLE (рис. 1).

В процессе проектирования целесообразно использовать возможности анализа и моделирование для повышения качества артефактов данного этапа. Поэтому было принято решение выполнить анализ исходных данных с применением статистических методов, в частности факторного анализа, а также построить структурную модель. Всё это должно позволить выявить те предикторы и факторы, которые оказывают заметное влияние на зависимую переменную — академическую успеваемость, интерпретировать их и на основе полученных результатов определить входные данные для PLE, в частности, реализовать модель обучаемого в виде онтологии, в которой результаты отдельных тестов будут экземплярами классов и будут связаны с экземплярами класса *Learner* (обучаемый) отношениями. В этом случае для проектирования отношений необходима информация о том, какие предикторы оказывают влияние на академическую успеваемость как по отдельности, так и вместе.

Модель образовательного контента, предложенная в [10], также реализована в виде онтологии *Protege* и используется для представления данных о видах, способах реализации, а также стратегиях выдачи образовательного контента. Она используется совместно с онтологией «Обучаемый» для установления соответствия между совокупностью персональных характеристик, присущих студенту, и наиболее подходящими ему формой и стратегией предоставления контента.

Для построения модели обучаемого (рис.1), интегрирующей набор характеристик субъектов обучения, влияющих на возможности восприятия ими (обучаемыми) информации и ее качественного усваивания, а также связей между этими характеристиками и целями обучения, мы предлагаем использовать методы многомерного анализа данных. Так, в качестве входных данных для оценки когнитивных характеристик предполагается использовать результаты тестов Амтхауэра, Равена. Для оценки личностных характеристик — результаты тестов «Большой

шая пятерка», Бачард. Мотивационные характеристики предполагается оценивать посредством краткого опросника Грэя. Данные об академической успеваемости уже собираются в LMS университета на протяжении нескольких лет, методика их сбора и расчета показателей успеваемости устоялась. В качестве зависимой переменной в исследовании выступает академическая успеваемость, которая рассчитывается как балл по шкале ECTS без учета пересдач, полученный за сдачу экзаменов, зачетов и курсовых.



Рис. 1. Использование модели обучаемого при проектировании PLE

Структура классов модели обучаемого «Learner model», в которой отражены составляющие личности студента, оказывающие влияние на академическую успеваемость, приведена на рис. 2. Однако в большинстве работ, посвященных исследованию влияния персональных свойств на успеваемость, используется или какой-то один тест, или несколько разных текстов, оценивающих один и тот же параметр. В настоящей работе мы исследуем, как влияют на успеваемость различные комбинации свойств личности студентов, на основе анализа данных множества тестов, оценивающих когнитивные, мотивационные и личностные свойства, с целью последующего выявления возможностей влияния на эти показатели посредством создания более персонифицированных форм представления образовательного контента.

Исследование организовано следующим образом. Сначала мы проводим анализ данных, в основе которого лежит построение многомерных линейных моделей. Проводится исследование влияния объясняющих переменных на общую успеваемость, а также успеваемость по группам предметов (гуманитарные, математические, информационные), в результате чего определяются значимые связи, которые будут включены в модель обучаемого. Далее, на основе предварительного проведения факторного анализа предикторов, мы предлагаем и исследуем приемлемые с точки зрения критериев качества структурные модели, интегрирующие исследуемые показатели в их сложной взаимосвязи, с целью последующего включения в модель обучаемого как самих показателей, так и их структурной организации.

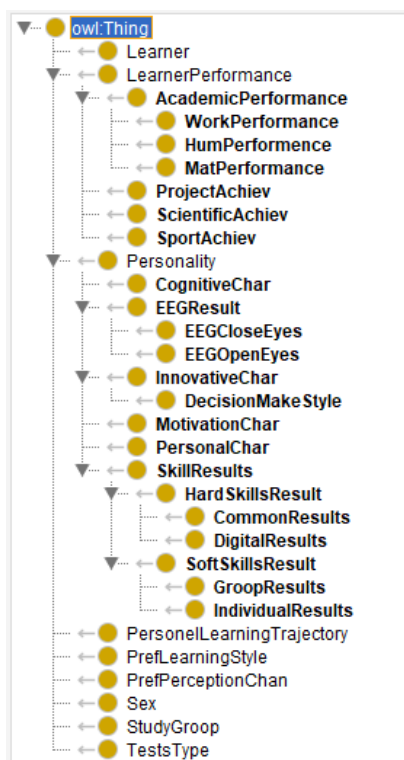


Рис. 2. Онтология «Learner model», реализующая модель обучаемого на основе результатов психометрического тестирования студентов

3. Выделение значимых характеристик обучаемых на основе регрессионного анализа

3.1. Описание и разведочный анализ входных данных

Для исследования на первом этапе были использованы результаты, полученные по следующим методикам оценки (тестовые материалы): Амтхауэра, Равена, Бачард, краткого опросника Грэя и «Большой пятерки». Результат каждого теста представляет собой один или несколько численных показателей (табл. 1). Также в исследуемом наборе данных присутствуют две дискретные переменные: «gender» – пол испытуемого и «group» – группа, в которой обучается студент. При обработке данных каждому студенту предварительно присвоен идентификационный номер (ID), а ФИО удалены.

Выборка включает результаты тестирования 72 человек – студентов третьего курса факультета прикладной математики и информатики НГТУ. Данные содержат 26 переменных, включая три группы характеристик, данные об академической успеваемости. Объекты с пропусками в данных были удалены из анализа.

Таблица 1. Переменные показателей, полученных в результате тестирования обучаемых

Наименование теста	Имя переменной (сокращение)	Описание переменной
Тест Амтхауэра	iq1	логическое мышление и общая осведомленность
	iq2	способность к абстрагированию
	iq3	комбинаторные способности (слова)
	iq4	лексический запас
	iq5	математическое мышление
	iq6	способность оперировать числами и индуктивное мышление
	iq7	комбинаторные способности (фигуры)

	iq8	пространственное воображение
	iq9	способность к кратковременному хранению информации
Тест Равена	RIQ	флюидный интеллект
Тест Бачард	eq1	выражение положительных эмоций
	eq2	выражение отрицательных эмоций
	eq3	внимание к эмоциям
	eq4	роль эмоций в принятии решений
	eq5	отзывчивость на радость
	eq6	отзывчивость к несчастьям
	eq7	эмоциональное участие
Краткий опросник Грэя	BAS	система активации поведения (стимулы награды)
	BIS	система торможения поведения (стимулы наказания)
«Большая пятерка»	Neuroticism (nrt)	нейротизм
	Extraversion (ext)	экстраверсия
	openess_experience (op_)	открытость опыту
	Agreeableness (agr)	доброжелательность
	Conscientiousness (cons)	добросовестность

Для того чтобы более детально оценить влияние нижеперечисленных объясняющих переменных на успеваемость (AVG – средний балл по всем пройденным предметам), учебные предметы дополнительно были разделены на три группы: средний балл по математическим предметам МАТ (напр., «Математический анализ»), предметам, связанным с профессиональной деятельностью PRG (напр., «Основы программирования»), гуманитарным предметам GUM (напр., «Иностранный язык»).

На первом этапе разведочного анализа проведено определение и удаление выбросов, содержащихся в исходных данных, на основе графика boxplot (точки-отскоки на рис. 3).

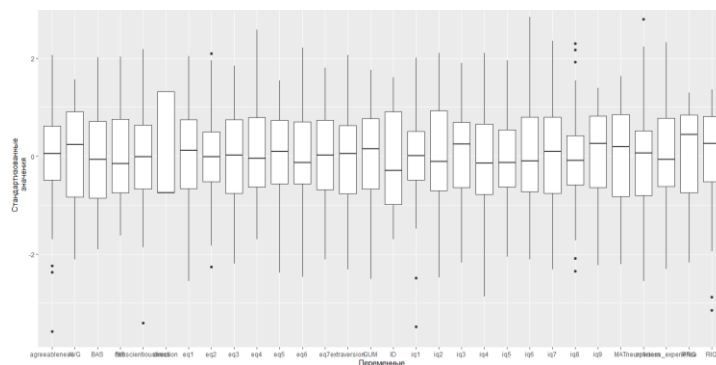


Рис. 3. График boxplot для стандартизованных значений объясняющих переменных

Анализ мультиколлинеарности проведен на основе фактора инфляции дисперсии VIF (табл. 2), в результате которого были исключены избыточные предикторы eq1 и eq6.

Таблица 2. Фактор инфляции дисперсии для исходных данных

Переменная	iq1	iq2	iq3	iq4	iq5	iq6	iq7	iq8	iq9	eq1	eq2	eq3	eq4
VIF	1.39	1.95	2.02	1.67	1.72	1.86	1.88	1.78	1.44	3.04	2.59	1.64	2.03

Переменная	eq5	eq6	eq7	RIQ	BAS	BIS	nrt	ext	op ₋	agr	cons
VIF	2.35	4.42	2.48	1.82	1.60	3.20	2.84	3.02	1.86	2.78	2.69

3.2. Многомерный регрессионный анализ

Регрессионный анализ позволяет построить линейную модель вида

$$y = \beta_0 + \beta_1 x_1 + \dots + \beta_k x_k + \varepsilon, \quad (1)$$

где y – зависимая переменная модели, определяющая качество обучения, x_i – объясняющие переменные (предикторы), представляющие собой характеристики обучаемых, β_i – неизвестные параметры модели, подлежащие оцениванию на основе собранных данных, ε – случайная неизвестная ошибка, удовлетворяющая необходимым требованиям.

Построение регрессионной модели (1) осуществляется при помощи функции `lm` пакета для анализа данных R. Результаты применения функции `lm` ко множеству всех объясняющих переменных с показателем среднего балла *AVG* в качестве зависимой переменной представлено в табл. 3. Как видно, ни одна из переменных не выявляет значимости, что является показателем избыточности предикторов.

Таблица 3. Результат применения функции `lm` для зависимой переменной *AVG*

Предикторы	Коэффициенты	Стандартные ошибки	Значение t-критерия	p-значение
(Intercept)	-78.2106	57.0132	-1.37	0.176
iq1	0.3897	0.2580	1.51	0.137
iq2	-0.1035	0.2730	-0.38	0.706
iq3	0.2720	0.2360	1.15	0.255
iq4	0.2372	0.2720	0.87	0.387
iq5	-0.0592	0.1443	-0.41	0.684
iq6	0.1509	0.2017	0.75	0.458
iq7	0.0141	0.2535	0.06	0.956
iq8	0.1270	0.2183	0.58	0.563
iq9	0.0866	0.1727	0.50	0.618
eq2	0.1189	0.2752	0.43	0.668
eq3	0.1356	0.2167	0.63	0.534
eq4	0.0866	0.2524	0.34	0.733
eq5	-0.2319	0.2124	-1.09	0.280
eq7	-0.0216	0.3510	-0.06	0.951
BAS	0.0518	0.4350	0.12	0.906
BIS	0.3460	0.3718	0.93	0.357
RIQ	0.1074	0.1630	0.66	0.513
neuroticism	-0.0811	0.2421	-0.34	0.739
extraversion	0.0754	0.1897	0.40	0.693
openess experience	-0.1194	0.2358	-0.51	0.615
agreeableness	0.1254	0.2896	0.43	0.667
conscientiousness	0.5504	0.2429	2.27	0.028

Для получения модели незначимые предикторы были последовательно исключены из модели (табл. 4). В результате получаем статистически значимую модель ($F_{4,67} = 7.17$, $p = 0.0000719$) зависимости среднего балла *AVG* от iq1, iq3 теста Амтхауэра, системы торможения поведения и добросовестности. Флюидный и эмоциональный интеллекты не оказывают существенного влияния на средний балл. Функциональный вид модели представлен формулой (2), коэффициент детерминации $R^2 = 0.26$.

Таблица 4. Результат применения функции lm для значимых предикторов

Предикторы	Коэффициенты	Стандартные ошибки	Значение t-критерия	p-значение
(Intercept)	-31.494	26.113	-1.21	0.23204
iq1	0.380	0.208	1.83	0.07174
iq3	0.401	0.157	2.56	0.01260
BIS	0.542	0.222	2.44	0.01726
conscientiousness	0.562	0.146	3.86	0.00026

$$AVG = -31.494 + 0.38 * iq1 + 0.401 * iq3 + 0.542 * BIS + 0.562 * conscientiousness . \quad (2)$$

Аналогичным образом была построена регрессионная модель для среднего балла по математическим предметам. В результате получаем модель (3), отражающую статистически значимую зависимость МАТ от iq3 теста Амтхауэра и добросовестности ($F_{2,69} = 8.55$, $p = 0.000482$). Модель описывает 18 % общей изменчивости успеваемости по математическим предметам. В отличие от общего среднего балла AVG, свою значимость потерял показатель BIS (стимулы наказания) краткого опросника Грэя.

$$MAT = -4.67 + 0.604 * iq3 + 0.443 * conscientiousness . \quad (3)$$

Регрессионная модель для среднего балла PRG по предметам, связанным с профессиональной деятельностью, описана формулой (4). Значимость предикторов iq1 и iq9 невысока, но при их исключении процент общей изменчивости сильно падает, поэтому при условии, что по F-критерию модель является значимой ($F_{3,68} = 4.85$, $p = 0.0409$), она была принята в качестве приемлемой. Процент описываемой изменчивости – 14 %.

$$PRG = -34.864 + 0.536 * iq1 + 0.341 * iq9 + 0.585 * conscientiousness . \quad (4)$$

И, наконец, регрессионная модель (5) отражает зависимость среднего балла по гуманитарным предметам от исследуемых предикторов ($F_{3,68} = 9.33$, $p = 0.0000301$). В отличие от предыдущих результатов, помимо iq3 из теста Амтхауэра и добросовестности, заметно обозначено влияние флюидного интеллекта. Процент объясненной изменчивости – 26 %.

$$GUM = -23.14 + 0.437 * iq3 + 0.368 * RIQ + 0.395 * conscientiousness . \quad (5)$$

Помимо множества количественных объясняющих переменных, рассмотренных ранее, в исследуемой выборке изначально присутствовали две категориальные переменные: пол и направление подготовки, однако они были исключены в силу незначимого влияния на показатели академической успеваемости. При добавлении категориальных предикторов к модели значимого взаимодействия ни с одной из зависимых переменных, определяющих качество обучения, не выявлено.

При интерпретации полученных результатов в первую очередь обращает на себя внимание тот факт, что «добросовестность» (*conscientiousness*) и «комбинаторные способности (слова)» (iq3) оказались значимыми предикторами в трех случаях из четырех. На практике можно наблюдать, что выдающиеся когнитивные способности не приведут к значимым академическим результатам, если обучаемый несвоевременно или не в полном объеме выполняет предусмотренные задания, так как получает меньшие баллы или штрафы, не может вовремя сдать сессию и т.д. При этом не секрет, что обучаемые, не очень хорошо знающие предмет, но обладающие хорошими речевыми навыками, могут получить более высокий балл, создав иллюзию понимания, расположив к себе экзаменатора или имея более выгодный вид на фоне знающего, но «неговорящего» одноклассника. Так что полученный результат не противоречит наблюдениям за обучаемыми. Далее, предиктор «логическое мышление и общая осведомленность» (iq1) оказался значимым как для модели с общей успеваемостью, так и для модели с успеваемостью по профессионально-ориентированным дисциплинам. Стоит отметить, что в современной науке и образовании междисциплинарный подход имеет большое значение. Поэтому успешность (высокая успеваемость) может быть результатом того, что обучаемый обладает кругозором и умеет сопоставлять факты, относящиеся к разным предметным областям,

для решения задач. Примечательно, что для гуманитарных дисциплин существенным оказалось влияние флюидного интеллекта RIQ, который свидетельствует об умении решать новые задачи без опоры на старые знания [14] и считается очень важным для успешного обучения. Также для общей успеваемости значимым оказался предиктор «система торможения поведения (стимулы наказания)» (BIS). При отсутствии значимых предикторов положительной мотивации это может говорить о более существенном вкладе страха наказания или провала, а значит, о меньшей действенности стимулов, связанных с поощрением. Для многих обучаемых очень трудно подобрать действенные положительные мотивационные стимулы (стипендия, похвала, чувство удовлетворенности проделанной работой и т.д.), в то время как отрицательные стимулы действенны и очевидны (звонок родителям, отчисление и т.д.).

4. Многофакторное моделирование

4.1. Факторный анализ

В первую очередь необходимо оценить возможность использования факторного анализа, конкретно, метода главных компонент (principal component analysis, PCA). Для этого определим критерий сферичности Бартлетта и критерий Кайзера–Мейера–Олкина. В результате выполнения теста Бартлетта получено значение 905, $p = 5.64 \times 10^{-68}$, следовательно, PCA имеет перспективы хорошего сжатия. Тест Кайзера–Мейера–Олкина показал значение 0.63, что оценивает возможность выполнения факторного анализа как посредственную. Однако, учитывая результаты теста Бартлетта, проведем анализ.

На основе графика «Каменистая осыпь» (рис. 4) было принято решение о числе главных компонент в модели. Точка на графике каменистой осыпи (рис. 3), где убывание собственных значений максимально замедляется, соответствует числу факторов, равному 4, однако суммарная доля объясненной дисперсии в этом случае равна 47.9 %.

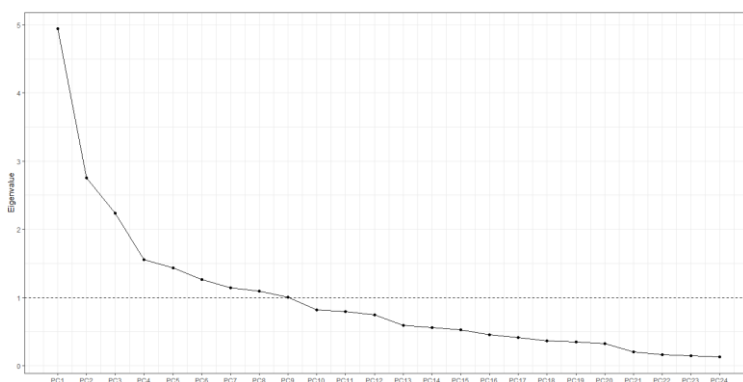


Рис. 4. График «Каменистая осыпь»

После вращения факторов методом ргмах получаем таблицу нагрузок (рис. 5a). Доля общей дисперсии, объясняемой каждым фактором, и накопление этих дисперсий представлены на рис. 6a. Для гипотезы о достаточности заданного числа факторов получаем $p\text{-value} = 0.0693$, что означает, что гипотеза не отвергается, но, возможно, число факторов стоит увеличить для достижения большего уровня значимости и приближения доли общей объясненной дисперсии к 50 %. Поэтому был рассмотрен альтернативный вариант увеличения числа факторов до 6. Результат после вращения методом ргмах отражен на рис. 5b.

Для проверки гипотезы о достаточности заданного числа факторов $p\text{-value} = 0.553$, что с большей уверенностью позволяет сказать, что шести факторов достаточно. Доля общей дисперсии, объясняемой каждым фактором, и накопление этих дисперсий представлены на рис. 6b.

Loadings:					Loadings:						
	Factor1	Factor2	Factor3	Factor4		Factor1	Factor2	Factor3	Factor4	Factor5	Factor6
eq6	0.905	0.658	-0.007	-0.082	eq6	0.799	-0.052	-0.213	-0.030	0.386	0.088
BIS	0.780	0.152	-0.003	0.015	eq7	0.820	0.171	-0.169	-0.021	-0.089	0.104
neuroticism	0.582	-0.040	0.079	-0.213	iq2	-0.237	0.564	-0.036	0.015	-0.137	0.373
eq7	0.421	0.668	0.175	-0.027	iq3	0.235	0.780	0.024	-0.045	0.050	0.093
BAS	0.079	-0.529	0.030	0.098	eq1	0.097	0.075	0.618	-0.131	0.000	0.272
openess_experience	0.016	0.544	0.217	0.106	extraversion	-0.210	-0.105	0.998	0.015	0.072	-0.137
agreeableness	0.233	0.526	0.048	0.313	conscientiousness	0.019	0.038	0.032	-1.024	0.097	0.006
iq2	-0.089	-0.433	0.621	0.308	BIS	0.090	0.007	0.116	-0.091	1.002	0.010
iq3	0.130	0.213	0.816	0.143	eq2	0.099	0.100	-0.018	0.000	0.023	0.998
iq7	-0.113	0.076	0.502	0.118	eq3	0.080	0.110	0.322	-0.093	0.142	0.128
eq1	-0.054	0.079	0.168	0.951	eq4	0.274	-0.332	0.149	0.299	0.021	0.068
eq2	0.413	-0.200	0.116	0.570	eq5	0.422	-0.269	0.250	0.008	-0.090	-0.045
extraversion	-0.228	0.103	-0.177	0.593	BAS	-0.355	0.020	0.120	0.417	0.032	0.121
eq3	0.062	0.125	0.081	0.350	iq1	0.083	0.260	-0.176	0.083	-0.033	-0.115
eq4	0.327	0.134	-0.309	0.151	iq4	-0.007	0.473	-0.035	-0.117	0.212	-0.082
eq5	0.123	0.422	-0.221	0.293	iq5	-0.236	0.466	0.051	0.105	-0.038	-0.096
iq1	0.009	0.123	0.239	-0.302	iq6	-0.255	0.400	0.100	-0.008	0.091	-0.226
iq4	0.069	0.048	0.478	-0.053	iq7	0.185	0.459	0.010	0.129	-0.265	-0.036
iq5	-0.188	-0.175	0.434	-0.077	iq8	0.159	0.315	-0.071	0.047	-0.052	-0.009
iq6	-0.180	-0.158	0.371	-0.063	iq9	-0.005	0.085	0.208	0.154	0.194	0.038
iq8	0.051	0.105	0.304	-0.059	RIQ	0.095	0.224	-0.073	-0.010	0.274	-0.071
iq9	0.270	0.024	0.033	0.040	neuroticism	0.199	0.055	-0.194	0.417	0.302	0.007
RIQ	0.234	0.141	0.257	-0.082	openess_experience	0.495	0.203	0.195	0.029	-0.137	-0.250
conscientiousness	-0.346	0.255	0.026	0.289	agreeableness	0.456	0.085	0.451	0.060	0.044	-0.008

а) Таблица нагрузок для 4 факторов

б) Таблица нагрузок для 6 факторов

Рис. 5. Таблица нагрузок для 4 и 6 факторов

Для визуального сравнения наполнение факторов, полученных в результате построения обеих моделей, приведено в табл. 5. Стоит отметить, что при отнесении переменных к факторам учитывались значения нагрузок более 0.4 (без учета знака).

	Factor1	Factor2	Factor3	Factor4		Factor1	Factor2	Factor3	Factor4	Factor5	Factor6
Proportion Var	0.113	0.101	0.099	0.096	Proportion Var	0.107	0.095	0.085	0.067	0.066	0.06
Cumulative Var	0.113	0.214	0.314	0.410	Cumulative Var	0.107	0.202	0.287	0.354	0.419	0.48

а) для 4 факторов

б) для 6 факторов

Рис. 6. Доля общей дисперсии и ее накопление для 4 и 6 факторов

Таблица 5. Сравнение наполнения факторов двух моделей

Фактор	Переменная	Расшифровка	Фактор	Переменная	Расшифровка
F1	eq6	отзывчивость к несчастьям	F1	eq6	отзывчивость к несчастьям
	BIS	стимулы наказания		eq7	эмоциональное участие
	nr1	нейротизм		eq5	отзывчивость на радость
	eq7	эмоциональное участие		op	открытость опыту
F2	eq7	эмоциональное участие	F2	agr	доброжелательность
	BAS	стимулы награды		iq3	комбинаторные способности (слова)
	op	открытость опыту		iq4	лексический запас
	agr	доброжелательность		iq5	математическое мышление
	eq5	отзывчивость на радость		iq6	способность оперировать числами и индуктивное мышление
				iq7	комбинаторные способности (фигуры)
F3	iq2	способность к абстрагированию	F3	iq2	способность к абстрагированию
	iq3	комбинаторные способности (слова)		eq1	логическое мышление и общая осведомленность
	iq7	комбинаторные способности (фигуры)		ext	экстраверсия
				agr	доброжелательность

	iq4	лексический запас			
	iq5	математическое мышление			
F4	eq1	выражение положительных эмоций	F4	BAS	стимулы награды
	ext	экстраверсия		nrt	нейротизм
	eq2	выражение отрицательных эмоций			
-			F5	BIS	стимулы наказания
-			F6	eq2	выражение отрицательных эмоций

Полученное распределение исходных переменных по группам, определяемым факторами, достаточно разнообразно. Факторы объединяют переменные, полученные для разных групп тестов. Однако попытаемся дать интерпретацию полученных разбиений.

В четырехфакторной модели возможна следующая интерпретация факторов.

Фактор 1. Включает переменные из разных тестов: eq6, BIS, eq2, eq7 и nrt. Все вместе можно интерпретировать как проявления тревожной натуры, чутко реагирующей на неудачи, возможность провала, наказания и т.д.

Фактор 2. Включает переменные, связанные с выражением положительных эмоций от взаимодействия с окружающим миром (eq5, eq6, eq7, op_, agr) и от получаемых стимулов награды (BAS).

Фактор 3. Объединяет наиболее важные переменные из группы когнитивных способностей по Амтхауэру, ответственные за аналитические способности.

Фактор 4. Объединяет переменные, направленные на выражение эмоций окружающим – как положительных, так и отрицательных.

Для шестифакторной модели можно интерпретировать факторы следующим образом.

Фактор 1. Переменные, объединенные этим фактором, связаны с открытостью, доброжелательностью и отзывчивостью. Общим мотивом для этих переменных может быть сотрудничество и эмоциональный отклик от такого рода взаимодействия.

Фактор 2. Объединяет переменные, связанные с различными компонентами теста Амтхауэра. В модели с четырьмя факторами очень схожим по составу получился фактор 3.

Фактор 3. Самая большая нагрузка в составе этого фактора приходится на такие переменные, как «экстраверсия», «логическое мышление и общая осведомленность», а также «доброжелательность». В отличие от фактора 1, такой состав фактора выглядит более ориентированным на взаимодействие с профессиональной точки зрения, так как эмоциональная составляющая в нем сочетается с логической.

Фактор 4. Включает реакцию на стимулы награды – BAS и нейротизм – nrt. Нейротизм как параметр указывает на высокий уровень восприимчивости к стимулам вообще, что говорит об усиленном восприятии стимулов награды.

Фактор 5. В данном факторе самая существенная нагрузка приходится на «реакцию на стимулы наказания» – BIS. Нагрузка на остальные переменные по сравнению с этой очень мала. Таким образом, очевидно сильное влияние стимулов наказания. Выше уже были упомянуты результаты, указывающие на корреляцию между BIS и успеваемостью.

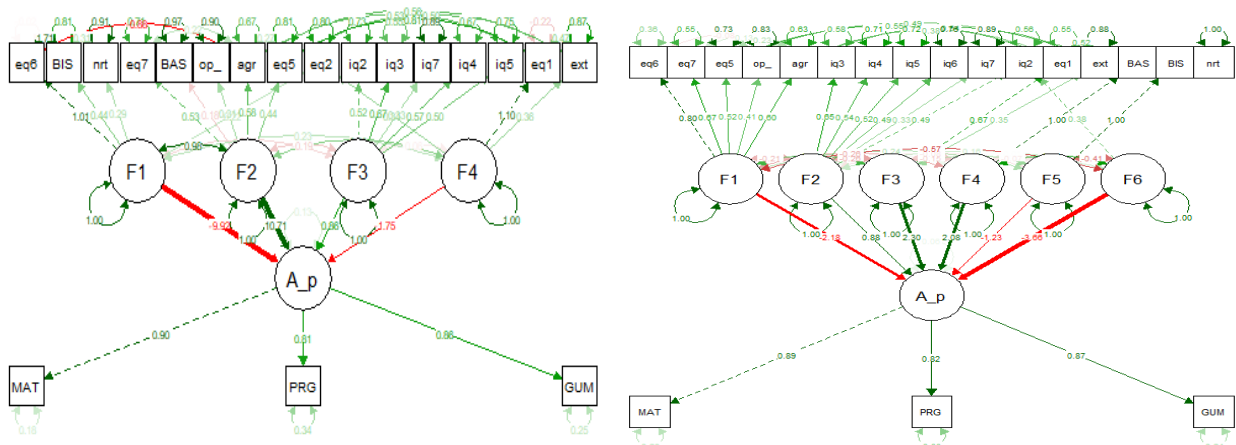
Фактор 6. Основная нагрузка падает на одну переменную – eq2 (выражение отрицательных эмоций), хотя в этом факторе нагрузку, близкую к 0.4, имеют переменные iq2 (способность к абстрагированию) и iq6 (способность оперировать числами и индуктивное мышление).

Говоря о сравнении четырехфакторной и шестифакторной моделей, стоит отметить, что BIS и BAS в обеих моделях относятся к разным факторам. Это говорит о том, что стимулы награды и наказания представляют собой отдельные переменные, не зависящие друг от друга, а также о важности мотивационной составляющей в принципе. Кроме того, в обеих моделях был выделен фактор, в составе которого объединены переменные, отвечающие за аналитические способности в структуре интеллекта по Амтхауэру. Это говорит о важности и независимости группы показателей, отвечающих за когнитивные способности.

4.2. Построение структурной модели

При построении структурной модели зависимости переменных, отвечающих за академическую успеваемость, от когнитивных, личностных и мотивационных характеристик обучаемых воспользуемся результатами факторного анализа. Факторный анализ позволяет построить модель с учетом знания о структуре разбиения предикторов по группам. В структурном моделировании используется несколько статистических тестов, позволяющих определить, соответствует ли модель данным и насколько хорошо. Как правило, мы не можем сделать вывод о том, «правильная» модель или нет, оценив качество соответствия между моделью и данными. Также мы не можем оценить, какую часть изменчивости она объясняет. Правильно подобранная модель означает только то, что модель правдоподобна.

Полученные модели для 4- и 6-факторного представления структуры объясняющих переменных представлены на рис. 7.



а) 4-факторная структурная модель

б) 6-факторная структурная модель

Рис. 7. Структурные модели, отражающие зависимость успеваемости от 4 (а) и 6 (б) факторов

Для оценки состоятельности структурных моделей обычно используются следующие показатели. CFI (сравнительный индекс соответствия) – при изучении исходных сравнений CFI в значительной степени зависит от среднего размера корреляций в данных. Если средняя корреляция между переменными невысока, CFI не будет очень высоким. Желательно, чтобы значение CFI было 0.95 и выше. RMSEA (среднеквадратичная ошибка аппроксимации) – индекс соответствия, где нулевое значение указывает на лучшее соответствие. Принято считать, что RMSEA 0.1 и выше указывает на плохое соответствие модели. SRMR (стандартизованный среднеквадратичный остаток) – популярный индикатор абсолютного соответствия. Значение менее 0.1 говорит о хорошем соответствии модели.

Значения указанных показателей для построенных моделей представлены в табл. 6.

Таблица 6. Сравнение характеристик двух моделей

Индекс	Значение для четырехфакторной модели	Значение для шестифакторной модели
CFI	0.858	0.910
RMSEA	0.082	0.066
SRMR	0.130	0.107

По данным табл. 6 видно, что шестифакторная модель обладает несколько лучшими характеристиками, нежели четырехфакторная. Тем не менее, обе модели можно считать приемлемыми с учетом размера выборки.

В обеих моделях с общей успеваемостью (рис. 9 и 10) присутствует фактор, где наибольшая нагрузка приходится на группу переменных, характеризующих интеллект по методологии Амтхауэра (логическое и математическое мышление, способность оперировать в рамках категорий и т.д.). Отметим тот факт, что и для четырехфакторной, и для шестифакторной модели

эти факторы (F3 и F2 соответственно) напрямую влияют на успеваемость, но при этом есть факторы, вклад которых больше.

Если обратить внимание на F2 и F1 в четырехфакторной и шестифакторной моделях соответственно, можно заметить, что в обоих случаях факторы, включающие переменные стимулов наказания (BIS), имеют знак минус. То есть чем меньше влияние этого фактора, тем лучше показатель успеваемости. Важным кажется тот факт, что в обеих моделях факторы, связанные с эмоциональной составляющей и мотивацией (F1 и F6 соответственно), оказывают существенное влияние на общую успеваемость и при этом входят в модель со знаком минус. Это можно интерпретировать с двух точек зрения: чем меньше факторов, тревожащих обучаемых, тем лучше они учатся; успешные обучаемые умеют интегрироваться в образовательный процесс, понимают, как он устроен, поэтому испытывают меньше тревог и учатся лучше.

Для уточненного анализа построим структурные модели на основе подтверждающего 6-факторного анализа, в которых зависимыми переменными являются средние баллы по различным группам дисциплин (математические, профессионально-ориентированные, гуманитарные), на которые влияют, предположительно, несовпадающие группы объясняющих переменных.

Структурные модели влияния на различные виды успеваемости представлены на рис. 8. Параметры для оценки моделей, приведенных на рис. 8, представлены в табл. 7. Качество моделей приемлемое.

Таблица 7. Значения критериев качества для моделей влияния на успеваемость по группам дисциплин

	Значение, математические предметы	Значение, профессиональные предметы	Значение, гуманитарные предметы
Кол-во параметров	62	62	62
Степени свободы	91	91	91
X ²	126.180	129.476	126.878
CFI	0.880	0.872	0.881
RMSEA	0.073	0.077	0.074
SRMR	0.114	0.111	0.113

При интерпретации шестифакторных моделей в разрезе математических, профессионально-ориентированных и гуманитарных предметов (рис. 8) можно отметить следующее.

На успеваемость по математическим предметам (рис. 8а) наибольшее влияние оказывает фактор, включающий компоненты теста Амтхауэра, соответственно, большее влияние имеют когнитивные способности обучаемого. Аналогичная ситуация наблюдается для модели с успеваемостью по профессионально-ориентированным дисциплинам (рис. 8б), но в этом случае можно видеть, что F1 и F6 не имеют влияния на успеваемость. В этом случае можно предположить, что успешность в дисциплинах, связанных с математикой и профессиональной деятельностью, зависит исключительно от когнитивных способностей обучаемого и требует особого склада ума.

В разрезе гуманитарных предметов (рис. 8с) успеваемость больше всего определяется фактором F6, который также связан с переменной интеллекта. При этом достаточно большое влияние оказывают факторы F1 и F5, где присутствуют предикторы, определяющие различные показатели эмоционального интеллекта (то есть мотивационные и личностные характеристики). F4 присутствует в модели с обратным знаком, что говорит об обратном влиянии этого фактора, где основной вклад дает предиктор BAS – стимулы награды. Можно предположить, что обучение по данным дисциплинам (куда в учебном плане входят родной и иностранный языки, история и т.д.), не вызывает у студентов затруднений. Она воспринимается как рутинная, каждодневная действительность, и награда, похвала за ее успешное выполнение не выглядит особенно ценной в глазах обучаемых.

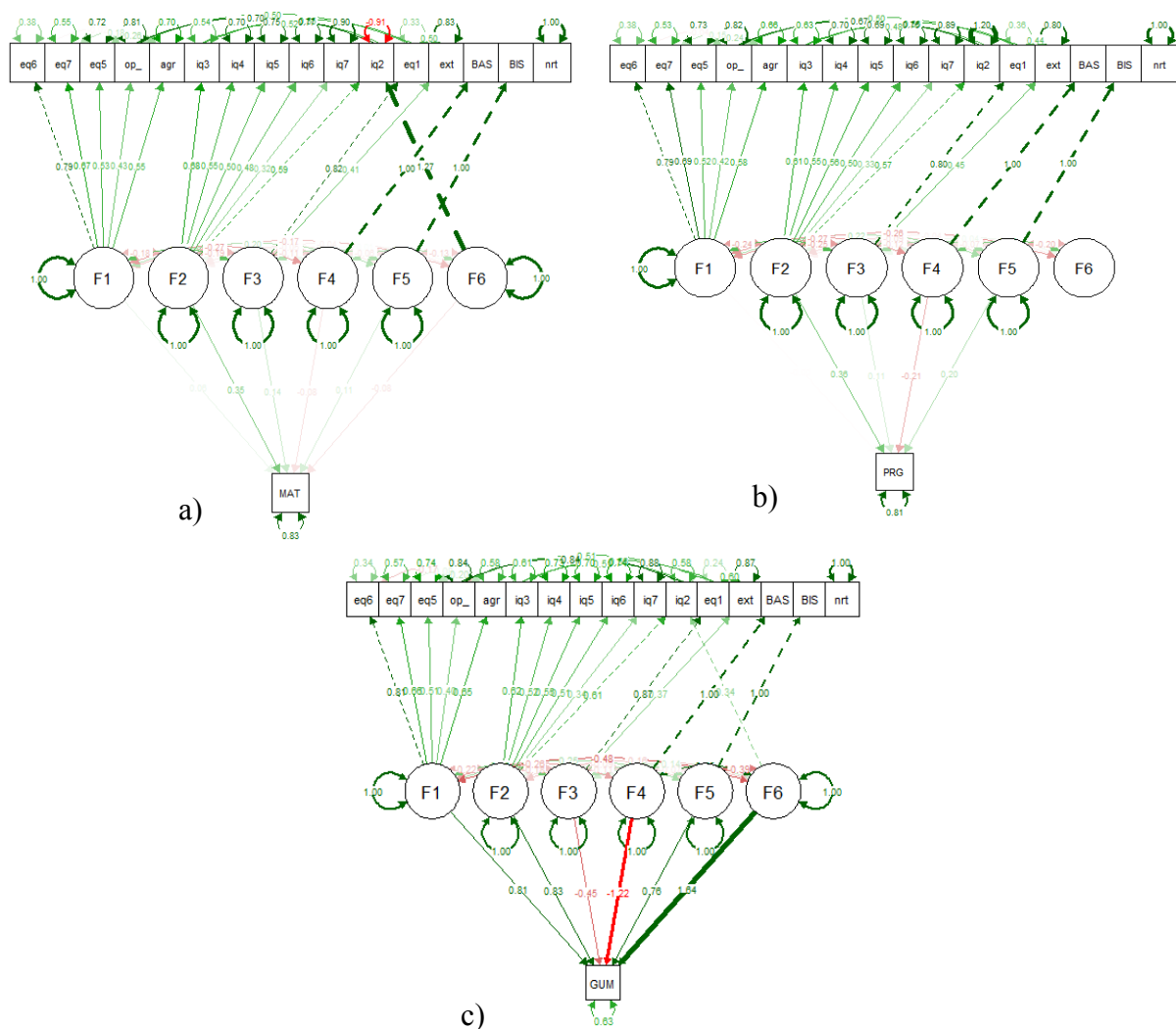


Рис. 8. Структурные модели влияния на успеваемость математических (а), профессиональных (б), и гуманитарных (с) дисциплин

Во всех трех моделях в разрезе групп предметов факторы стимулов награды и наказания представлены в двух разных факторах. Последнее может говорить о том, что использование мотивационных стимулов, как отрицательных, так и положительных, требует определенной стратегии с учётом особенностей обучаемых.

5. Заключение

Выполненное исследование позволило определить степень влияния основных когнитивных, личностных и мотивационных характеристик студентов на различные виды успеваемости. Выявленный набор показателей, оказывающих значимое влияние на успеваемость и при этом не дублирующих друг друга, предполагается использовать для входного диагностического тестирования студентов при адаптации модели обучаемого в реализуемой PLE.

Кроме того, при реализации PLE полученная модель будет включена в архитектуру информационных технологий университета в виде онтологии «Модель обучаемого» (организация интерфейса посредством SWI-Prolog, JSON, API). Совместно с онтологией «Модель образовательного контента» её предполагается использовать для рекомендации студентам образовательного контента, наилучшим образом учитывающего их индивидуальные особенности. Мы предполагаем персонализировать выдачу контента, используя методы интеллектуальной поддержки на основе онтологических моделей. «Модель обучаемого» объединяет характери-

стики, отвечающие за представление различных аспектов личности, что позволяет использовать ее для описания особенностей поведения студентов и последующей их классификации, например, в зависимости от стратегии использования контента. Онтология «Образовательный контент» содержит описание различных признаков современного контента, используемого в образовательном процессе. На основании правил и аксиом, а также знаний, представленных двумя онтологиями в среде Protégé, разрабатываемая система будет рекомендовать форму и стратегию выдачи образовательного контента для конкретного студента или группы студентов, обладающих схожими характеристиками.

Онтология, представляющая модель обучаемого (Learner model), учитывает не только экспертные знания, изначально заложенные в систему, но и результаты анализа статистических данных, которые можно оперативно изменять с учетом изменения контингента студентов. Таким образом, используя стратегию самообучения интеллектуальной системы, можно обновлять знания системы путем добавления экземпляров, а также корректировки связей между классами и экземплярами, что позволит оперативно отслеживать изменения и своевременно отражать их в PLE.

Литература

1. Олейникова О. Н., Редина Ю. Н. Макротренды развития системы высшего образования в мире. [Электронный ресурс]. URL: <http://www.vestnik.vsu.ru/pdf/educ/2018/03/2018-03-02.pdf> (дата обращения: 11.12.2021).
2. Россия 2025: от кадров к талантам. [Электронный ресурс]. URL: http://d-russia.ru/wp-content/uploads/2017/11/Skills_Outline_web_tcm26-175469.pdf (дата обращения: 11.12.2021).
3. Глобальная конкурентоспособность российского образования. [Электронный ресурс]. URL: <https://publications.hse.ru/books/223161350> (дата обращения: 11.12.2021).
4. Драгунова Е. В., Пустовалова Н. В., Вальдман И. А. Инновационные технологии в дизайне образовательных экосистем // Материалы МНТК «Актуальные проблемы электронного приборостроения (АПЭП-2018)». Т. 8. С. 62–69.
5. Васильченко С. Х. Формирование персональной образовательной среды на основе информационных технологий для реализации индивидуальных траекторий обучения (на примере корпоративного обучения): автореф. дисс. ... канд. пед. наук. М., 2012.
6. Кочергина Е. В., Най Дж. В. К., Орёл Е. А. Факторы «Большой пятерки» как психологические предикторы академической успеваемости студентов вузов // Психологические исследования. 2013. Т. 6. № 27. [Электронный ресурс]. URL: <http://psystudy.ru/index.php/num/2013v6n27/776-kochergina27.html> (дата обращения: 11.12.2021).
7. Збанацкая А. Б. Личностные характеристики студентов вуза как основа индивидуального стиля учебной деятельности: дисс. ...канд. псих. наук: М., 2007.
8. Малахова С. И. Связь психометрического интеллекта с личностной саморегуляцией студентов: автореф. дисс. ... канд. пед. наук. М., 2013.
9. Gilmore P., Self J. The application of machine learning to intelligent tutoring systems. Artificial Intelligence and Human Learning. Intelligent Computer-aided Instruction. Chapman and Hall, London, 1988. pp. 179–196.
10. Пустовалова Н. В. Онтологическая модель процесса взаимодействия образовательного контента и обучаемого в архитектуре интеллектуальных обучающих систем // Вестник Астраханского государственного технического университета. Серия: Управление, вычислительная техника и информатика. 2021. № 4. С. 113–125.
11. Ржанова И. Е., Бритова В. С., Алексеева О. С., Бурдукова Ю. А. Флюидный интеллект: обзор зарубежных исследований // Клиническая и специальная психология. 2018. Т. 7,

- № 4. С. 19–43. [Электронный ресурс]. URL: doi:10.17759/cpse.2018070402 (дата обращения: 11.12.2021).
12. Математические методы в психологии. Основы применения (Mathematical Methods in Psychology: Basics of Applying) / Санкт-Петербургский государственный университет. [Электронный ресурс]. URL: <https://www.coursera.org/learn/matematicheskiye-metody-v-psikhologii> (дата обращения: 11.12.2021).
13. RDocumentation [Электронный ресурс]. URL: <https://www.rdocumentation.org/> (дата обращения: 11.12.2021).

*Статья поступила в редакцию 19.12.2021;
переработанный вариант – 18.02.2022.*

Авдеенко Татьяна Владимировна

д.т.н., профессор, профессор кафедры теоретической и прикладной информатики НГТУ (630073, Новосибирск, пр. К. Маркса, 20), тел. (383) 3-460-600, e-mail: tavdeenko@mail.ru.

Пустовалова Наталья Викторовна

аспирант, старший преподаватель кафедры автоматизированных систем управления НГТУ, тел. (383) 3-461-100, e-mail: NVPustovalova@gmail.com.

References

1. Olejnikova O. N., Redina Ju. N. *Makrotrendy razvitija sistemy vysshego obrazovanija v mire. [Jelektronnyj resurs]*. [Macrotrends in the development of the higher education system in the world]. URL: <http://www.vestnik.vsu.ru/pdf/educ/2018/03/2018-03-02.pdf> (access date: 11.12.2021).
2. *Rossija 2025: ot kadrov k talantom. [Jelektronnyj resurs]* [Russia 2025: from personnel to talents]. URL: http://d-russia.ru/wp-content/uploads/2017/11/Skills_Out-line_web_tcm26-175469.pdf (access date: 11.12.2021).
3. *Global'naja konkurentosposobnost' rossijskogo obrazovanija. [Jelektronnyj resurs]* [Global competitiveness of Russian education]. URL: <https://publications.hse.ru/books/223161350> (access date: 11.12.2021).
4. Dragunova E. V., Pustovalova N. V., Val'dman I. A. *Innovacionnye tehnologii v dizajne obrazovatel'nyh jekosistem* [Innovative technologies in the design of educational ecosystems]. *Materialy MNTK «Aktual'nye problemy jelektronnogo priborostroenija (APJeP-2018)»* [Materials of the MNTK "Actual problems of electronic instrumentation (APEP-2018)"]. vol. 8, pp. 62-69.
5. Vasil'chenko S. H. *Formirovanie personal'noj obrazovatel'noj sredy na osnove informacionnyh tehnologij dlja realizacii individual'nyh traektorij obuchenija (na primere korporativnogo obuchenija)* [Formation of a personal educational environment based on information technologies for the implementation of individual learning paths (on the example of corporate training)]: *avtoref. diss. ... kand. ped. nauk*. Moscow, 2012.
6. Kochergina E. V., Naj Dzh. V. K., Orjol E. A. *Faktory «Bol'shoj pjaterki» kak psihologicheskie prediktory akademicheskoy uspevaemosti studentov vuzov* [Factors of the "Big Five" as psychological predictors of academic performance of university students]. *Psihologicheskie issledovanija*. 2013, vol.6, no. 27. [Electronic resource]. URL: <http://psystudy.ru/index.php/num/2013v6n27/776-kochergina27.html> (access date: 11.12.2021).
7. Zbanackaja A. B. *Lichnostnye harakteristiki studentov vuza kak osnova individual'nogo stilja uchebnoj dejatel'nosti* [Personal characteristics of university students as the basis of individual style of learning activity]: *diss. ...kand. psih. nauk*. Moscow, 2007.

8. Malahova S. I. *Svjaz' psihometricheskogo intellekta s lichnostnoj samoreguljaciej studentov* [Relationship of psychometric intelligence with personal self-regulation of students]: *avtoref. diss. ... kand. ped. nauk.* Moscow, 2013.
9. Gilmore P., Self J. The application of machine learning to intelligent tutoring systems. *Artificial Intelligence and Human Learning. Intelligent Computer-aided Instruction*. Chapman and Hall, London, 1988, pp. 179-196.
10. Pustovalova N. V. *Ontologicheskaja model' processa vzaimodejstvija obrazovatel'nogo kontenta i obuchaemogo v arhitekture intellektual'nyh obuchajushhih sistem* [Ontological model of the process of interaction between educational content and the student in the architecture of intelligent learning systems]. *Vestnik Astrahanskogo gosudarstvennogo tehničeskogo universiteta. Serija: Upravlenie, vychislitel'naja tehnika i informatika*. 2021, no. 4, pp. 113-125.
11. Rzhanova I. E., Britova V. S., Alekseeva O. S., Burdukova Ju. A. *Fljuidnyj intellekt: obzor zarubezhnyh issledovanij* [Fluid intelligence: a review of foreign studies] *Klinicheskaja i special'naja psihologija*. 2018, vol. 7, no. 4, pp. 19-43. [Electronic resource]. URL: doi:10.17759/cpse.2018070402 (access date: 11.12.2021).
12. *Matematicheskie metody v psihologii. Osnovy primeneniya* [Mathematical Methods in Psychology: Basics of Applying]. *Sankt-Peterburgskij gosudarstvennyj universitet*. [Electronic resource]. URL: <https://www.coursera.org/learn/matematicheskiye-metody-v-psikhologii> (access date: 11.12.2021).

Multivariate analysis of the students' characteristics influences on academic progress for the personalization of educational environment

Tatiana N. Avdeenko

Doctor of technical sciences, Professor, Novosibirsk State Technical University (Novosibirsk, Russia), tavdeenko@mail.ru.

Natalia V. Pustovalova

Postgraduate student, Novosibirsk State Technical University (Novosibirsk, Russia), NVPustovalova@gmail.com.

Modern universities developing a digital environment rely on a student-centered approach, relevant pedagogical and didactic practices taking into account learners' characteristics and adapting it to the educational process. Personalization of the educational environment should improve the quality of the educational process raising it to a new level. This article presents the results of the multivariate modeling methods to identify standard relations and patterns between academic performance and students' psychometric characteristics. The results of students testing at the Novosibirsk State Technical University (NSTU) determining their cognitive, motivational, and personal characteristics are considered in this paper and information about all past certification tests is provided for the curriculum.

Keywords: regression analysis, principal component analysis, structural equation modeling, learner's model, psychometric characteristics.