

Интеллектуальный анализ персональных предпочтений студентов в ходе подготовки кадров в высшей школе

Е.Г. Антоненков¹, Ю.А. Крюков²

(Государственное бюджетное образовательное учреждение Московской области «Университет «Дубна», Дубна)

Данная статья посвящена интеллектуальному анализу персональных предпочтений обучающихся. Контроль успеваемости является основным способом оценивания хода образовательного процесса, как отдельного студента, так и всего контингента обучающихся. Своевременный контроль успеваемости обучающегося является одним из важных элементов для повышения качества образования. Для успешного освоения дисциплины обучающийся должен не только выполнять домашние задания, успешно сдавать зачеты и экзамены, он также должен посещать и занятия, тем более, если учится на очной форме обучения. Система образования создает внушительный объем данных, которые никаким образом не анализируются, кроме основных показателей. Однако, как раз незначительная, на первый взгляд, информация может влиять на конечный результат, то есть уровень подготовки высококвалифицированного специалиста. Современные интеллектуальные системы способны прогнозировать успеваемость студента, указать на его предрасположенности. Предпочтения обучающегося можно выявить также и из иных источников, например, из социальных сетей, где студенты делятся своими впечатлениями, следят за последними новостями науки и т.п. С помощью интеллектуального анализа больших данных станет возможным создание индивидуальных траекторий обучения, адаптацию процесса обучения с целью получения высококвалифицированного специалиста.

Ключевые слова: информационная система, большие данные, интеллектуальный анализ, сбор данных.

Президентом страны неоднократно поднимался вопрос необходимости укрепления человеческого потенциала как одного из важнейших приоритетов стратегии развития России. В послании Президента Федеральному собранию поставлены задачи правительству по увеличению объема инвестиций уже к 2020

¹ Антоненков Евгений Геннадьевич, аспирант (eugene.antonenkov@gmail.ru).

² Крюков Юрий Алексеевич, к.т.н., доцент (kua@uni-dubna.ru).

году на уровень 6-7 процентов, что предусматривает развитие новых производств и увеличение количества высококвалифицированных рабочих мест. Однако страна входит в сложный демографический период, связанный с огромными людскими потерями во время Великой Отечественной войны и драматический период конца двадцатого века. Повышение качества подготовки специалистов с учетом внимательного отношения к студенческому сообществу является основной задачей организаций высшего образования.

Одним из важных элементов качества образования является своевременный контроль успеваемости обучающихся.

Образовательные организации всё чаще задаются вопросом: «Как организовать постоянный мониторинг процесса обучения с учетом личностного фактора каждого из студентов?» Современные интеллектуальные системы способны обрабатывать огромные потоки малозначимой, но разнообразной информации о текущей успеваемости студента и, при необходимости, информировать кафедры о проблемах, которые или уже назрели у конкретного обучающегося или могут у него появиться в ближайшей перспективе.

Контроль успеваемости предназначен для регулярного и систематического оценивания хода освоения студентами дисциплин и выполнения других видов учебной работы во время контактных занятий преподавателя со студентами [4]. В настоящее время контроль успеваемости студента состоит из анализа данных в рамках промежуточной аттестации. Однако, в ряде случаев, получаемые данные по окончании семестра не позволяют своевременно начать взаимодействие с отстающими студентами для ликвидации текущей задолженности по учебному графику. Отчисление по результатам экзаменационной сессии может стать проблемой как для самого обучающегося, так и для образовательного учреждения. С проблемой несвоевременного получения картины успеваемости студентом сталкиваются как образовательные учреждения, так и государственные органы управления. Государственное задание по выпуску специалистов с требуемым уровнем качества подготовки является главной задачей образовательной организации. Основной же задачей

органов государственного управления является поддержка баланса на рынке труда и обеспечение развития предприятий региона с формированием когорты молодых специалистов в соответствии с номенклатурой востребованных специальностей. Современные технологии аналитики больших данных, собираемых из множества открытых источников, могут стать основой для организации системы координации усилий педагогического состава по целевой корректировке текущей успеваемости студентов.

Основной формой контроля обучающегося является промежуточная аттестация. Промежуточная аттестация регламентируется локальными нормативными актами, созданными в соответствии с Федеральным законом от 29 декабря 2012 г. № 273-ФЗ «Об образовании в Российской Федерации» [6]. Чаще всего промежуточная аттестация в высших учебных заведениях проходит раз в семестр. Таким образом, образовательная организация, факультет или кафедра, получают информацию об успеваемости студента с периодичностью раз в полгода. Некоторые организации вводят в течение семестра некие дополнительные формы контроля, например «контрольные точки» на 8 и 12 неделе обучения в семестре. Такой подход позволяет организации реагировать на проблемные ситуации быстрее, чем при промежуточной аттестации, однако не может с точностью определить уровень знаний обучающегося в определенный момент обучения.

Во время перехода к цифровой экономике каждое учебное заведение вынуждено использовать те или иные информационные системы, для своевременного получения разнородной информации, связанную как с участниками образовательного процесса, так и с самим образовательным процессом.

Образовательный процесс, если рассматривать его структуру, построен на строгом учебном плане и на рабочих программах дисциплин. В первом указывается весь перечень дисциплин в определенном семестре, и вид промежуточной аттестации по дисциплине. Во-втором указывается содержание дисциплины, темы семинарских занятий и объемы самостоятельной работы студента, которую в последующем оценивает преподаватель.

Если собирать эти оценки по результатам каждого проведенного задания, например, используя «электронные дневники», то функционал существующей информационной системы, в итоге, необходимо будет расширить. Таким образом, ИС, существующая в образовательной организации, будет хранить в себе не только оценки за промежуточную аттестацию (сессии), но и оценки за задания, по которым ведется контроль в семестре (Рис.1).

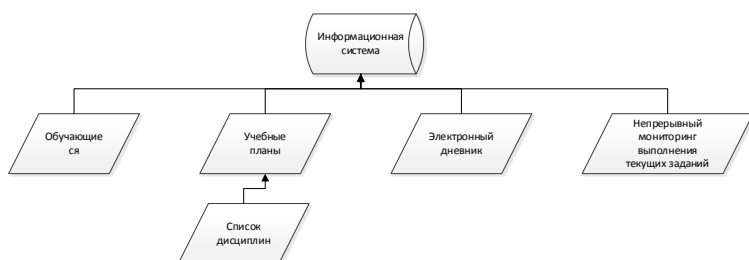


Рис. 1. Данные, хранящиеся в информационной системе

Ещё одним критерием успешного и своевременного усвоения дисциплины является посещаемость занятий. Если обучающийся не посещает занятия, то вероятность того, что он успешно освоит дисциплину, уменьшается. Однако такое суждение не может быть однозначным. Как существуют дисциплины, которые построены на самостоятельном обучении студентов или частично перенесены в статус электронного курса, так и существуют студенты, которые при самостоятельном изучении постигают ту или иную технологию более успешно, чем на занятиях, ведь за несколько часов преподаватель не сможет дать всю глубину

знаний. Основными факторами, влияющими на учебную посещаемость, являются:

- личностные факторы, исходящие от самого учащегося (мотивация учения, несостоятельность в учебе, здоровье учащегося);
- внешние факторы (востребованность профессий на рынке труда, работа во время учебы);
- институционные факторы (организация учебно-воспитательного процесса в учебном заведении, наличие мониторинга посещаемости, уровень профессионализма преподавателей, имидж образовательного учреждения, режим работы и дисциплина в учебном заведении, наличие стимулирующих программ и адекватные условия обучения) [3].

Для получения информации о посещаемости необходим источник данных, например, система контроля и управления доступом (СКУД) посредством видеофиксации человека. СКУД – это совокупность программных и технических средств, а также организационно-методических мероприятий, с помощью которых решается задача контроля и управления посещением охраняемого объекта [5]. Такая система позволяет избежать действий от объекта (обучающегося) и от субъекта (преподавателя и т.п.): обучающемуся не нужно отмечаться у преподавателя, либо прислонять

ключ-карту, которую он может благополучно забыть дома. Система автоматически собирает информацию о том, в какой момент времени студент появился на занятии. Если спарить данные из СКУД с ИС, в которой имеется информация о всех физических лицах, расписание занятий (в котором также указано время начала занятия), то это позволит получить все необходимые данные по посещаемости обучающегося (Рис.2). Аналогично можно поступить и с преподавателями, чтобы контролировать выполнение ими работы.

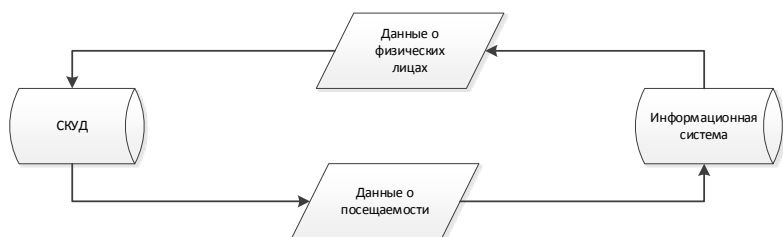


Рис. 2. Связь СКУД и ИС

Таким образом мы можем получить некое хранилище исходных данных, в котором сосредоточена и накапливается информация о результате освоения дисциплины и о посещаемости по каждому студенту. Но хотелось бы использовать эту информацию, а не просто собирать и хранить. Например, проанализировать: насколько эффективно проходят очные занятия со студентами, какие курсы дисциплин студенты с минимальным посещением осваивают лучше, получить картину

способностей студента (выявить одаренных и т.п.). Для этого уже придется прибегнуть к интеллектуальному алгоритму обработки исходных данных.

Не для кого не секрет, что мы живем в эпохе больших данных. Любое наше действие, которое можно считать или зафиксировать, начиная с информации о том, когда мы выходим из дома выкидывать мусор, и заканчивая поисковыми запросами в интернете, может использоваться для определенной задачи по анализу больших данных. Уже сейчас поисковые системы фиксируют вспышку заболеваний быстрее, чем к этому выводу приходят департаменты здравоохранения. Магазины научились определять беременность женщин по информации о покупках или интересах, и своевременно использовать эту информацию, направляя купоны или специальные предложения клиентам. Сайтами-агрегаторами бронирования гостиниц или заказов авиабилетов уже никого не удивить.

Система образования также создает внушительный объем данных. В связи с чем возникают вопросы как их анализировать и какие принимать решения на их основе. Несомненно, в мире существует много примеров успешного применения технологий больших данных. В частности, одна из сфер применения больших данных - прогнозное моделирование. Американские колледжи и уни-

верситеты сами направляют письма-обращения будущим студентам, приглашая поступить в то или иное учебное заведение. Каждый вуз стремится пригласить наиболее перспективных студентов, которые наверняка поступят. Чтобы облегчить работу приемной комиссии, аналитики из компании *ForecastPlus* собрали и проанализировали несколько типов данных студентов: этническая принадлежность, успеваемость по ряду предметов, выпускные работы, оценки. Прогнозное моделирование *ForecastPlus* доказало свою эффективность более чем в ста кампусах США. Так, Университет Крейтон в штате Небраска смог исключить 35000 не самых перспективных студентов и не направлять им письма, что позволило сэкономить более 30 тысяч долларов [8].

Источником большого количества разнообразной информации может служить социальная сеть. Всероссийский центр изучения общественного мнения (ВЦИОМ) подсчитал, что 45% граждан Российской Федерации пользуются хотя бы одной из социальных сетей почти каждый день. Лидером опроса стала социальная сеть «ВКонтакте» [1]. У данной социальной сети существует API (*application programming interface*), у которого имеется функционал получения информации о пользователе. С помощью API можно получить информацию о группах, в которых состоит пользователь, записи на стене пользователя, заметки

пользователя, обсуждения, информацию об интересах и т.д. Подобная информация из социальных сетей позволяет установить связи обучаемого со своими сверстниками, перемещения по сети, его интересы, частоту вхождения в социальную сеть и время нахождения в сети и др. В основе анализа данных социальных сетей лежит математическая теория графов, которая представлена в работе Эрдоса [7]. Математически, сеть – то набор узлов (в нашем случае обучающиеся), соединенных линиями, характеризующими отношения между узлами.

Первым шагом в сетевом анализе является визуализация данных. Визуализированные графы позволяют выявить узлы, являющиеся ближайшими, найти плотные кластеры активности. Одним из основных параметров графа является «центральность по степени», выражающий отношение количества связей определённого узла к общему количеству других узлов. Если для некоторого узла этот параметр равен 1, это означает, что этот узел связан со всеми остальными узлами сети, если он равен 0, то этот узел является изолированным. Этот показатель показывает степень «знаменитости» узла, показывает, что студент оказывает большое влияние на остальных. На рисунке 4 показан фрагмент визуализированного графа для анализа взаимодействия учащихся в социальной сети. Из рисунка видно, что пользователь сети R1

обладает более высокой степенью центральности, чем пользователь R3 (из R3 отходит меньше линий связи, чем из узла R1) [2].

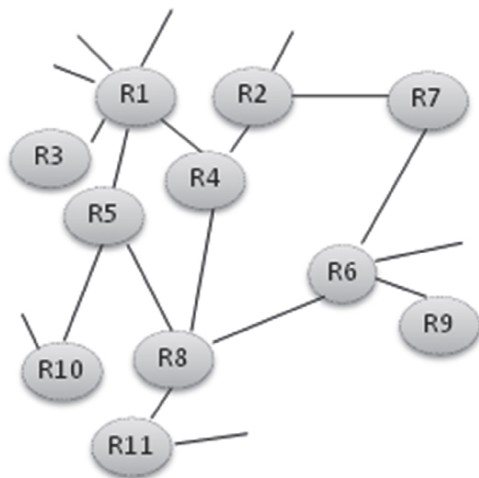


Рис. 3. Фрагмент визуализации взаимодействия студентов в социальной сети

На сегодняшний день разработано множество программных средств для интеллектуального анализа данных. Для анализа социальных сетей (Social Network Analysis, SNA) зачастую используют программное приложение MySpace [9]. Анализ социальных сетей направлен на изучение отношений между людьми, рассматривает социальное взаимодействие в терминах теории сетей. MySpace может быть использован для интерпретации и анализа структуры и отношений студентов при решении совместных заданий или при

взаимодействиях различными средствами коммуникации. Таким образом, при анализе лишь одной социальной сети возможно получить некую картину интересов пользователя, которую впоследствии можно использовать для создания комфортных условий обучения, повысив уровень заинтересованности у обучающегося, например, дать задание, связанное с предметной областью его интересов и т.п.

Технологии больших данных и методы их анализа используются повсеместно, зачастую незаметно для нас, т.е. тех, кто является их источниками. Данные технологии могут быть применены и к образовательной сфере, включая мониторинг успеваемости, вопросы принятия решений и финансовое планирование. Большие данные позволят сохранить опыт обучения, дадут картину обучения каждого обучающегося. Специалистам будет проще создавать индивидуальные траектории обучения и адаптировать процесс обучения с целью удовлетворения потребностей каждого обучаемого, или отрасли.

Литература

1. ВЦИОМ: Какими социальными сетями пользуются россияне [Электронный ресурс]. – Режим доступа: <https://wciom.ru/index.php?id=236&uid=116691>.

2. МАМЕДОВА Г.А., Л.А. ЗЕЙНАЛОВА, Р.Т. МЕЛИКОВА *Технологии больших данных в электронном образовании // Открытое образование.* – 2017. – Т.21, №6. – С. 41–48.
3. МИСЕЛИМЯН Т.Л., МЕТЕЛИЦА Н.Т. *Влияние посещаемости занятий учащимися на качество образовательного процесса // Успехи современного естествознания.* – 2005. - №5. – С.76–79.
4. ПРОКОФЬЕВА Е.В., ПРОКОФЬЕВА О. Ю., ШАРКЕВИЧ Н.В. *Контроль успеваемости в образовательных учреждениях // Приоритетные направления развития науки и образования.* – 2014. - №3(3). – С.111–112.
5. *Системы контроля и управления доступом* [Электронный ресурс]. – Режим доступа: <http://ksb-ul.ru/uslugi/sistema-kontrolya-i-upravleniya-dostupom-skud>.
6. *Федеральный закон РФ №273-ФЗ «Об образовании в Российской Федерации»* [Электронный ресурс]. – Режим доступа: <http://xn--273--84d1f.xn--p1ai/zakonodatelstvo/federalnyy-zakon-ot-29-dekabrya-2012-g-no-273-fz-ob-obrazovanii-v-rf>.
7. ERDOS P., RENYI A. *On the evolution of random graphs* // Publication of Mathematics Institute Hungary Academy of the Science. – 1960. –V.5. – P.17–61.
8. *Increase enrollment management precision and efficiency* [Электронный ресурс]. – Режим доступа: <https://www.ruffalonl.com/complete-enrollment-management/cultivating-applicants/forecast-plus-student-recruitment-predictive-modeling>.
9. <https://myspace.com> (Дата обращения: 20.07.2019).

Intellectual analysis of personal preferences of students in the course of training in higher education

Evgeniy Antonenkov, Dubna State University, Dubna, postgraduate (eugene.antonenkov@gmail.com)

Yuri Kryukov, Dubna State University, Dubna, Candidate of Engineering Sciences, Associate Professor (kua@uni-dubna.ru)

Abstract: This article is devoted to the intellectual analysis of personal preferences of students. Monitoring of progress is the main way to assess the progress of the educational process, as an individual student, and the entire contingent of students. Timely monitoring of student progress is an important element for improving the quality of education. For the successful development of the discipline, the student must not only perform homework, successfully pass tests and exams, he must also attend classes, especially if he is studying full-time. The education system generates an impressive amount of data that is not analysed in any way other than the basic indicators. However, just a small, at first glance, information can affect the final result, that is, the level of training of highly qualified specialists. Modern intellectual systems are able to predict the student's progress, to point to his predisposition. Preferences of the student can also be identified from other sources, for example, from social networks, where students share their impressions, follow the latest news of science, etc. With the help of intellectual analysis of big data it will be possible to create individual learning trajectories, to adapt the learning process in order to obtain a highly qualified specialist.

Keywords: information system, big data, data mining, data collection.

УДК 004.8

ББК 32.813.5

Статья представлена к публикации
членом редакционной коллегии ...заполняется редактором...

Поступила в редакцию ...заполняется редактором...
Опубликована ...заполняется редактором...