

## ПРИМЕНЕНИЕ КАТЕГОРИАЛЬНОГО АНАЛИЗА В ПРОФЕССИОНАЛЬНОЙ ОРИЕНТАЦИИ АБИТУРИЕНТОВ

© С.В.Чернова, А.А.Салмин

Поволжский государственный университет телекоммуникаций и информатики

Статья поступила в редакцию 25.01.2010

Рассматриваются вопросы профориентационной работы с абитуриентами на основе их индивидуальных свойств и качеств. Отражены достоинства анализа личностных характеристик абитуриента. Приводится категориальный анализ свойств абитуриента на основе байесовского подхода.

Ключевые слова: профессиональная ориентация, категориальный анализ, абитуриент, личностная характеристика.

*Введение.* В выпускном классе подростков особенно остро волнует проблема профессионального самоопределения. На этом этапе многие начинают осознать, что адекватность выбора профессии в дальнейшем может существенно повлиять на качество жизни, поэтому важное значение в этом случае приобретает профориентация.

Профессиональная ориентация – процесс определения того вида трудовой деятельности, в которой подросток хочет себя проявить, осознание своих склонностей и способностей к этому виду деятельности и осведомленность о каналах и средствах приобретения знаний, умений и навыков для овладения конкретной профессией. Профориентационная работа ставит перед собой такие задачи, как пропаганда профессий, информация о потребности в кадрах. Важным показателем правильной организации работы по профессиональной ориентации является намерение молодых людей работать или продолжить учебу в соответствии с приобретенной ими профессией. Неудовлетворенность полученной профессией или невозможность трудоустройства по имеющейся профессии зачастую ведут к новому витку процесса профессиональной ориентации и пересмотру своего выбора. Особое значение развитие системы профориентации имеет в период экономических преобразований в стране, обостряющих безработицу, когда резко меняется востребованность отдельных профессий на рынке труда и у индивида возникает потребность в выборе и приобретении не одной, а двух или трех профессий

для успешного трудоустройства<sup>1</sup>. На сегодняшний день существует множество методик и тестов по профориентации<sup>2</sup>.

В 2009 году в рамках проекта «Профориентация» был осуществлен сбор и анализ существующих методик, применяемых в различных регионах России при работе с молодежью, обращающейся за профессиональным консультированием<sup>3</sup>. В исследовании принимали участие ведущие эксперты и психологи в области профессиональной ориентации молодежи. В ходе исследования были запрошены методики из всех центров профориентации действующих в крупнейших регионах России, в том числе в исследовании приняли участие: Новосибирск, Самара, Челябинск, Саратов, Чебоксары и многие другие. По результатам исследования присланных списков был сформирован перечень наиболее популярных и используемых в региональных центрах методик. В результате анализа этих методик были сделаны заключения, отражающие факт того, что исследованные методики, традиционно используемые в профориентационной работе, неплохо зарекомендовали себя в профконсультировании школьников 10 – 11 классы. С их помощью можно оценить как склонности и познавательные способности оптантов, так и их индивидуально-личностные качества. Однако, было отмечено, что в ходе исследования выявлены следующие недостатки:

<sup>1</sup>Чернова Светлана Владимировна, ассистент кафедры информационных систем и технологий.

E-mail: [Chernova\\_s\\_v@bk.ru](mailto:Chernova_s_v@bk.ru)

Салмин Алексей Александрович, кандидат технических наук, доцент кафедры информационных систем и технологий. E-mail: [ist@psati.ru](mailto:ist@psati.ru)

1 Социология: Энциклопедия (Мир энциклопедий) / Сост. А.А.Грищанов, В.Л.Абушенко, Г.М.Евелькин, Г.Н.Соколова, О.В.Терещенко. – Мн.: 2003.

2 Профориентация. Тесты профориентации. Комплекс методик для профессиональной ориентации. URL: <http://www.effecton.ru/730.html> (дата обращения 27.09.2009)

3 Федеральное агентство по делам молодежи. Росмолодежь. URL: <http://www.fadm.gov.ru/about/mission/> (27.10.2009).

1) методики зарубежных авторов (Айзенка)<sup>4</sup>, Басса-Дарки<sup>5</sup>, Лири, Кеттела и др.) не проходят процедур регулярной реадаптации в нашей стране, на их использование нет разрешения правообладателей; 2) подбор методик эклектичен, ряд из них «схватывают» одну и ту же реальность в разных контекстах, что неизбежно ведет к противоречиям в полученных результатах, неоднозначности выводов; 3) большая часть перечисленных методик рассчитаны на юношеский и взрослый возраст оптантов и не могут быть использованы с детьми моложе 16 лет.

По итогам анализа эксперты сделали вывод о том, большинство методик используемых для работы с современной молодежью являются устаревшими и на данный момент альтернатив их замены не существует. Так же можно отметить, что используемые в настоящее время методики и тесты имеют обобщенные задачи, не учитывающие особенности абитуриента и его предпочтения в зависимости от личностных характеристик. Для исследования процесса профессионо-лично-личностного самоопределения абитуриента необходимо рассматривать как его личностные характеристики, так и профессионо-мотивационную составляющую развития абитуриента. Игнорирование данных факторов у абитуриента может привести к неблагоприятному прогнозу на дальнейшее развитие обучения по выбранной специальности.

Важной задачей является сориентировать абитуриента, например, в пределах вуза на конкретную специальность. В этом случае необходимо выявить предрасположенность абитуриента к выполнению определенной профессиональной деятельности, при помощи анализа интересов, наклонностей и личностных качеств абитуриента. Такой анализ тесно связан с использованием новых информационных технологий, формирующих оценки гипотез о склонности абитуриента к той или иной специальности. Далее описываются основы категориального анализа, который позволяет нам использовать информацию об абитуриенте на разных уровнях.

*Категориальный анализ.* Под категорией понимается группа или множество объектов, обладающих определенным набором свойств. В качестве таких объектов могут выступать абитуриенты. Процесс категориального анализа заключается в аналитическом разделении объектов на однородные группы. Применительно к профориентационной работе категориальный

анализ подразумевает процесс деления абитуриентов на группы со схожими параметрами, называемыми *сегментами*. В качестве таких сегментов в контексте профориентационной работы вуза с абитуриентами могут выступать специальности, на которые ведется набор.

Для того чтобы произвести процесс сегментации с целью выявления предрасположенности абитуриента к определенной специальности, необходимо использовать его индивидуальные характеристики, определяющиеся *свойствами*, которые описывают определенные черты, присущие абитуриенту. Каждое из свойств состоит из набора *качеств*, описывающих его состояние. Число качеств зависит от самого свойства. Например, свойство «медаль» – абитуриент закончил школу с медалью (золотой или серебряной), при этом качество «медалиста» может характеризоваться типом медали или просто ее наличием. В случае наличия медали *качество* имеет две категории – «да» и «нет».

Свойства абитуриента могут быть тесно связаны с интересами вуза (например, посещал ли абитуриент подготовительные курсы) или представлять второстепенный интерес (например, район проживания и т.д.). Следует также отметить, что второстепенные свойства, не имеющие, на первый взгляд, особого значения, неожиданно могут стать важнее основных. Основу категориального анализа абитуриентов определяет таблица свойств и категорий. Пример такой таблицы представлен в таб. 1. В данной таблице свойство «Общий результат ЕГЭ» выражено в баллах от 89 до 93, в связи с тем, что проходной балл 91, свойство «Способность к углубленному изучению предметов (УКЛОН ШКОЛЫ)» выражено категориями: Экономический, Гуманитарный, Математический или Без уклона. Свойства «Знание предмета» по четырем дисциплинам выражено оценками из аттестата в баллах от 3 до 5. Путем выборки нужных категорий в таб. 1 (данные категории выделены *подчеркиванием*) формируются личностные характеристики абитуриента (ЛХА). В результате данной выборки охарактеризована абитуриентка, имеющая результат ЕГЭ в диапазоне  $\geq 93$  балла, окончившая школу с экономическим уклоном, посещавшая ПОДГОТОВИТЕЛЬНЫЕ КУРСЫ ВУЗа, не имеющая медали, не посещавшая ШКОЛУ ПРОГРАММИСТОВ ВУЗа, знающая следующие предметы на оценку: ИНФОРМАТИКА=4, АЛГЕБРА=5, ФИЗИКА=5, ОБЩЕСТВОЗНАНИЕ=4. В итоге: ЛХА= (Пол=Ж) & (Общий результат ЕГЭ $\geq 93$ ) & (Посещение ПОДГОТОВИТЕЛЬНЫХ КУРСОВ=ДА) & (Посещение ШКОЛЫ ПРОГРАММИСТОВ=НЕТ) (Медаль=НЕТ) & (Способность к углубленному

4 Бурлачук Л.Ф., Морозов С.М. Словарь-справочник по психодиагностике. – Киев: 1989. – С. 8 – 11.

5 Клиническая психология. Словарь / Под ред. Н.Д.Твороговой. – М.: 2007; Психологический лексикон. Энциклопедический словарь в шести томах / Ред.-сост. Л.А.Карпенко. Под общ. ред. А.В.Петровского.

изучению предметов (УКЛОН ШКОЛЫ)=Экономический) & (Знание предмета: ИНФОРМАТИКА=4) & (Знание предмета: АЛГЕБРА=5) & (Знание предмета: ФИЗИКА=5) & (Знание предмета: ОБЩЕСТВОЗНАНИЕ=4). В данном случае ЛХА представлена как набор из 10-ти фактов, связанных конъюнктивной связью. Данное сочетание свойств позволит идентифицировать «место» конкрет-

ного абитуриента в процессе сегментации и сформировать стереотип качества, который максимально приблизит абитуриента к определенному сегменту. В конечном счете, для вуза важны такие свойства и качества абитуриентов, которые обеспечат максимально правильную принадлежность к тому или иному сегменту специальности.

**Таб. 1.** Свойства и категории

Свойства	Категории			
	М	Ж		
Общий результат ЕГЭ	<=89	>=90, <=92	>=93	-
Знание предмета: ИНФОРМАТИКА	3	4	5	-
Знание предмета: АЛГЕБРА	3	4	5	-
Знание предмета: ФИЗИКА	3	4	5	-
Знание предмета: ОБЩЕСТВОЗНАНИЕ	3	4	5	-
Способность к углубленному изучению предметов (УКЛОН ШКОЛЫ)	Гуманитарный	Экономический	Математический	Без уклона
Посещение ШКОЛЫ ПРОГРАММИСТОВ	ДА	НЕТ	-	-
Медаль	ДА	НЕТ	-	-
Посещение ПОДГОТОВИТЕЛЬНЫХ КУРСОВ	ДА	НЕТ	-	-

**Таб. 2.** Пример базы данных учебного заведения

Пол	ЕГЭ общ.	Медаль	Алгебра	Обществознание	Информ.	Физика	Уклон шк.	Подг. курсы	Дополнит. уч. нагр. вуза	Направл.
1	2	3	4	5	6	7	8	9	10	11
1	90		3	4	4	3	Э	Нет	Да	ИТ
1	90		4	5	5	5	Э	Да	Да	ПО
0	92		4	4	5	4		Да	Да	СО
1	89		5	4	5	5		Нет	Нет	ИТ
1	91	Да	5	5	5	5	М	Нет	Да	ПО
1	90		5	5	4	4		Нет	Да	ПО
1	89		5	4	5	4		Нет	Да	ПО
0	90		3	3	4	4		Нет	Нет	ПО
1	92		4	3	5	3		Нет	Нет	СО
1	92		3	4	4	3	М	Нет	Нет	ИТ
0	92		4	4	5	3	Г	Да	Нет	СО
1	92		4	5	4	4		Нет	Нет	ПО
1	92		4	5	3	4	М	Нет	Нет	СО
0	92		3	4	3	3	М	Да	Нет	ИТ
0	92		3	3	4	4	Э	Нет	Нет	ИТ
1	92		4	5	4	4	Э	Нет	Да	ИТ
1	92		4	4	4	4		Нет	Нет	СО
1	92		5	5	4	4		Нет	Нет	СО

Комплексный анализ личностных характеристик абитуриента с целью определения сегмента, в который его можно отнести на основе индивидуальных характеристик, может позволить выявить насколько абитуриент склонен к определенной специальности.

*Распределение абитуриентов по сегментам.* Применительно к профориентационной задаче,

как уже упоминалось выше, под сегментами подразумеваются специальности, на которые ведется прием в учебном заведении. В этом случае для проведения категориального анализа используется база данных приемной комиссии учебного заведения, которая образуется на основе статистики предыдущих лет с учетом вновь поступающих абитуриентов. В этих дан-

ных содержится информация, которую можно рассматривать как шаблоны для оценки свойств абитуриента, на предмет его тяготения к определенному сегменту.

В таб. 2 представлен фрагмент общей базы данных абитуриентов ГОУВПО «Поволжского государственного университета телекоммуникаций и информатики» (ПГУТИ). В представленной таблице отражены свойства и определенные категории качества некоторых абитуриентов предыдущих лет. Столбец, содержащий фамилию, имя и отчество абитуриента в таблице опущен. Здесь представлены как общие характеристики абитуриентов, например, столбец 1, характеризующие свойство «Пол» с категориями 1 – «мужской», 0 – «женский», так и характеристики, отражающие его успеваемость, например, столбцы 2 – 7. Так в столбце 2 отражены суммарные баллы результатов ЕГЭ (Единого государственного экзамена), в столбце 3 отражено наличие или отсутствие медали у абитуриента, в столбцах 3 – 7 оценки из аттестата по соответствующим предметам по пятибалльной шкале. Кроме того, в исходной базе данных может быть представлена информация, представляющая интерес для конкретного учебного заведения. Например, столбец 8 характеризует уклон школы, в которой обучался абитуриент: «Э» – «экономический», «Г» – гуманитарный, «М» – математический или без уклона или дополнительное обучение в пределах учебного заведения (столбцы 9 – 10): посещал ли абитуриент подготовительные курсы или какие-либо другие дополнительные учебные программы с категориями «Да» / «Нет». Столбец 11 отражает направления, на которые были определены абитуриенты в соответствии с их качествами по статистике предыдущих лет. В качестве примера, представлено 3 направления, на которые ведется прием в ПГУТИ: ИТ – «Информационные системы и технологии», ПО – «Программное обеспечение вычислительной техники и автоматизированных систем», СО – «Связи с общественностью».

Следует также отметить, что набор свойств и качеств может варьироваться в зависимости от требований учебного заведения. Для того, чтобы произвести категориальный анализ, необходимо из общей таблицы выделить данные, относящиеся к определенному сегменту. Данный факт позволит сформировать безусловные априорные вероятности принадлежности абитуриента к сегменту. Кроме того, каждый сегмент должен состоять из определенного набора свойств и качеств, отличного от остальных. На основе статистики об абитуриентах внутри каждого из сегментов можно сформировать априорные условные вероятности их принадлежно-

сти к тому или иному сегменту. В совокупности комплекс характеристик для каждого из сегмента образуют общую сегментарную базу, позволяющую производить анализ абитуриента с учетом его личностных свойств для определения сегмента, к которому он наиболее склонен. Следует также отметить, что личностные характеристики как внутри сегмента, так и в случае категориального анализа абитуриента не имеют связи между собой, например, свойства «пол» и «ЕГЭ» не связаны между собой и т.д.

Выдвижение и оценка гипотез о том, попадет ли конкретный абитуриент в тот или иной сегмент, связана с задачей классификации, при этом количественное значение такой оценки должно определяться вероятностной мерой. При этом вероятность следует интерпретировать как степень возможного «попадания» абитуриента в соответствующий сегмент. В описываемом подходе используется байесовская вероятность<sup>6</sup>, под которой понимается оценка влияния свойств и категорий абитуриента на процесс сегментации.

Анализ, основанный на формуле Байеса, является универсальным, так как при внесении новых переменных принцип сегментации не изменяется, а всего лишь усложняет определение условных вероятностей, в связи с увеличением исследуемых факторов. На основании базы данных приемной комиссии учебного заведения (таб. 2) определяем общую априорную статистику сегментов ПО, ИТ, СО. Для проведения категориального анализа на основе байесовского перерасчета вероятностей рассматриваются следующие исходные данные: 1) личностная характеристика абитуриента (ЛХА); 2) общая посегментная статистика абитуриентов (или априорные вероятности попадания абитуриента в сегмент  $P(S_i)$ ); 3) посегментная статистика категорий абитуриентов (условные вероятности принадлежности абитуриента к категориям  $K_{NM}$ , составляющие ЛХА, для всех сегментов  $S_i, i=1,2,...,Z$ ). Пример посегментной статистики для специальности ПО показан в таб. 3. В данной таблице представлена посегментная статистика предыдущих лет для специальности ПО, такая же статистика имеется для специальностей ИТ, СО.

Объем выборки для каждой таблицы представляет из себя количество человек, подавших документы на соответствующую специальность и используется для систематизации свойств и качеств. Отношение количества объектов свя-

6 Байесовские процедуры квалификации: Вводный обзор [Электронный ресурс] – Режим доступа [http://www.spc-consulting.ru/DMS/Machine%20Learning/Machine Learning/Overvie](http://www.spc-consulting.ru/DMS/Machine%20Learning/Machine%20Learning/Overvie) свободный. – Загл. с экрана (27.10.2009).

занных с категорией, к общему количеству объектов определяет эмпирические вероятности или частоты категориальных событий. Число сегментов, в данном случае – 3, сегменты альтернативны. Таб. 3 представляет сегмент специальности ПО, для других сегментов будут использоваться аналогичные таблицы, но стати-

стика, собираемая в них, будет отличаться для различных сегментов. Каждый сегмент может быть образован при анализе свойств конкретных абитуриентов и должен иметь четко определенные границы, связанные с множеством атрибутов объекта.

**Таб. 3.** Посегментная статистика

Сегмент ПО						
Свойства	Категории качества				Объем вы- борки	Вероятность
	1	2	3	4		
Пол	М	Ж	-	-	570	1
	0,596491	0,403509			570	1
Общий результат ЕГЭ	<=89	>=90, <=92	>=93	-	570	1
	0,145614	0,798246	0,05614	-	570	1
Алгебра	5	4	3	-	570	1
	0,235088	0,538596	0,226316	-	570	1
Обществознание	5	4	3	-	570	1
	0,261404	0,575439	0,163158	-	570	1
Информатика	5	4	3	-	570	1
	0,235088	0,564912	0,2	-	570	1
Физика	5	4	3	-	570	1
	0,298246	0,517544	0,184211	-	570	1
Уклон школы	Э	Г	М	Без укло- на	570	1
	0,059649	0,003509	0,077193	0,859649	570	1
Подготовительные курсы	Да	Нет		-	570	1
	0,107018	0,892982		-	570	1
Дополнительная учебная нагрузка вуза	Да	Нет		-	570	1
	0,119298	0,880702		-	570	1
Медаль	Да	Нет		-	570	1
	0,003509	0,996491		-	570	1

Для реализации пересчета апостериорных вероятностей с учетом вышеперечисленных исходных данных формула Байеса примет вид:

$$P(S_i / ЛХА) = \frac{P(ЛХА / S_i)P(S_i)}{\sum_{i=1}^Z P(ЛХА / S_i)P(S_i)}, \quad (1)$$

где  $P(S_i)$  – априорная вероятность попадания абитуриента в сегмент  $S_i$ ;

$P(S_i/ЛХА)$  – условная вероятность тяготения абитуриента к сегменту  $S_i$ , при условии, что он имеет определенный набор ЛХА;

$P(ЛХА/S_i)$  – условная вероятность того, что абитуриент с набором ЛХА находится в сегменте  $S_i$ .

После предъявления комплекса характеристик для каждого из сегмента и определенного набора ЛХА, для которого производится анализ, согласно формуле (1) можно произвести категориальный анализ, основанный на формуле Байеса и пересчитать апостериорные вероятности принадлежности абитуриента к определенному сегменту. Последняя операция будет обновлять априорные вероятности сегментов и заменять их апостериорными, которые были получены на основе ЛХА. Набор сегментов со-

держит в себе статистику (стереотип сегмента), в которую входят данные об общем количестве абитуриентов в сегменте и количестве абитуриентов, дифференцированных по значениям личностных характеристик. Применительно к профориентационной работе вуза, стереотипы сегментов образуют общую сегментарную базу приемной комиссии вуза, на основе пересчета вероятностей с использованием формулы Байеса на основе ЛХА.

*Заключение.* Исследование показало, что проблемы формирования профессиональной компетентности у будущих специалистов лежат в плоскости использования новейших подходов к профессиональной ориентации абитуриентов. Описываемый подход, основанный на категориальном анализе личностных характеристик абитуриента, может быть использован в приемных комиссиях как высших, так и в средних профессиональных учебных заведениях. В результате проведенного категориального анализа было выявлено, что успеваемость студентов повысилась, количество отчислений уменьшилось. Попадание абитуриента в определенный сегмент повысит эффективность обучения, что в даль-

нейшем приведет к более качественному выпуску специалистов. В конечном итоге, можно говорить о том, что профессиональная ориента-

ция абитуриентов, основанная на байесовском подходе существенно повысит желание выпускников работать по выбранной специальности.

## **APPLICATION OF CATEGORY ANALYSIS IN APPLICANTS' VOCATIONAL ORIENTATION**

©2010 S.V.Chernova, A.A.Salmin<sup>°</sup>

Povolzhskiy State University of Telecommunications and Informatics

The questions of career-oriented work with applicants on the basis of their individual qualities are considered. Advantages of the applicants' personal characteristics analysis are reflected. A category analysis of an applicant qualities on the basis of Bayesian approach is given.

Key words: career-oriented, category analysis, applicants, personal characteristics.

---

<sup>°</sup>*Chernova Svetlana Vladimirovna, Assistant of the Information systems and technologies department.*

*E-mail: [Chernova\\_s\\_v@bk.ru](mailto:Chernova_s_v@bk.ru)*

*Salmin Aleksey Aleksandrovich, Cand. Sc. in Engineering, Associate Professor of the Information systems and technologies department.*

*E-mail: [ist@psati.ru](mailto:ist@psati.ru)*