

Data Mining и статистические методы при оценивании качества курсов повышения квалификации

В статье рассматриваются кейс анализа данных, с использованием специализированного ПО, связанных с образовательной статистикой, а именно – результатом опроса слушателей курсов повышения квалификации. Показывается необходимость расширенной статистической обработки результатов, схема проведения анализа. Показаны выводы по исследуемому кейсу.

Ключевые слова: курсы повышения квалификации, статистика, Data Mining, R Language, KNIME.

DATA MINING AND STATISTICS METHODS USAGE FOR ADVANCED TRAINING COURSES QUALITY MEASUREMENT: CASE STUDY

In the article we consider a case of the analysis of the data connected with educational statistics, namely – result of professional development courses students survey with specialized software usage. Need for expanded statistical results processing, the scheme of carrying out the analysis is shown. Conclusions on a studied case are presented.

Keywords: professional development courses, statistics, Data Mining, R Language, KNIME.

1. Введение. Постановка задачи

Оценка любого мероприятия, позволяющего повысить компетентность слушателей, состоит из двух компонент: собственно оценка полученных знаний, умений и навыков и оценка пользователями проведенных мероприятий. Если в случае высших образовательных учреждений, школ оценка подачи материала, самооценка полученных знаний, ведения занятий, качества преподавания зачастую связана с огромной долей субъективизма и по этой причине не может быть признана валидной, то в случае курсов повышения квалификации можно ожидать более ответственного подхода к такого рода оценкам.

По этой причине отдел кадров одной из сетевых компаний проводит мониторинг процесса обучения слушателей с целью получения обратной связи на уровне слушателей и корректировки выявляемых проблем. Сбор информации произво-

дится после недельного обучения на очных курсах повышения квалификации по темам, связанным с обслуживанием электроэнергетического оборудования («Организация ремонтного, оперативного и технического обслуживания оборудования подстанций 35–110 кВ», «Организация ремонтного, оперативного

и технического обслуживания оборудования подстанций 35–110 кВ» и т. д.) путем заполнения слушателями опросного листа.

Для заполнения слушателям предлагались опросные листы, содержащие 15 характеристик, названия которых были кодированы для последующей обработки (табл. 1).

Таблица 1.

Оцениваемые параметры и их кодировка для последующей обработки

Оцениваемый параметр	Кодировка
Общая организация процесса обучения	obshorg
Оснащение и оборудование учебной аудитории	oboryd
Соответствие времени проведения занятий заявленному по расписанию	cootvraspis
Соответствие содержания обучения заявленной программе	sootvprogr
Практическая полезность проведенного обучения	polezn
Доступность изложения материала	dostypmat
Использование современных методов подачи материала	sovrmetod
Индивидуальный подход к слушателю	indpodxod
Эффективность взаимодействия преподавателя с аудиторией	vzaimodayd
Получение обратной связи от преподавателей	obrcvyazprepod
Получение обратной связи от работников кафедры	obrcvyazkaf
Взаимодействие слушателей между собой	vzaimodslsh
Наличие, качество, полезность раздаточных материалов	razdatmat
Качество питания в столовой	pitanie
Общие впечатления о проведенном обучении	obsh



Максим Иванович Гальченко,
ст. преподаватель
Тел.: (911) 761-67-41
Эл. почта:
maxim.galchenko@gmail.com
Санкт-Петербургский
государственный аграрный
университет

Maxim I. Galchenko,
Senior lecturer
Tel.: (911) 761-6741
E-mail: maxim.galchenko@gmail.com
Saint-Petersburg State
Agrarian University



Александр Геннадьевич Гущинский,
руководитель учебного центра
ОАО «Ленэнерго»
Тел.: (921) 958-76-70
Эл. почта:
Guschinskiy.AG@nwenargo.com
Ленэнерго, ОАО

Alexander G. Gushchinskij,
Head of the Training Center
of "Lenenergo"
Tel.: (921) 958-76-70
E-mail: Guschinskiy.AG@nwenargo.com
"Lenenergo", JSC

Каждый из показателей оценивался по десятибалльной шкале от 0 (очень плохо) до 10 (отлично). Следует отметить, что последний параметр «Общие впечатления о проведенном обучении» по сути является обобщающим для всех прочих (целевой переменной).

Всего было получено 316 заполненных опросников.

В результате отдел кадров стал обладателем большого массива данных, обработка которого проводилась в Excel и носила весьма общий, описательный характер. Целью нашей работы, которая и описывается в настоящей статье, было выполнение максимально полного анализа данных и фактически построение протокола воспроизводимого исследования.

Задачи, ставившиеся при проведении анализа, были сформулированы отделом кадров следующим образом:

- оценить качество полученных в результате опроса данных;
- выделить основные факторы, влияющие на оценку курса слушателями;
- дать рекомендации касательно улучшения характеристик курсов.

В качестве инструментов нами были выбраны: язык статистического программирования R [1] и платформа для интеллектуального анализа данных KNIME [2].

Ниже мы представим описание самого кейса с целью максимально полно проиллюстрировать причины, по которым были выбраны именно такой набор методов и последовательность шагов.

2. Статистическая обработка данных

Изначально данные подверглись очистке от неполных данных, так как часть применяемых методов не может быть использована в случае наличия пропущенных значений. В результате выборка уменьшилась до 315 записей.

Проверка данных на нормальность с помощью критерия Шапиро – Уилка дала отрицательный результат по всем полям. Пик приходится на значение 10 (отлично) по всем показателям (что уже вызывает вопрос о наличии «мусора» – формально заполненных опросных листов). Так, для показателя «Общие впечатления о проведенном обучении» гистограмма показывает преобладающее в выборке значение 10, при низких частотах ниже 6 (рис. 1), что является характерным для всех остальных полей.

Для оценки связей между переменными была вычислена матрица корреляций Спирмена, впоследствии подвергнутая фильтрации на уровне 0.6, для чего был использован пакет corrplot языка R [3] (рис. 2).

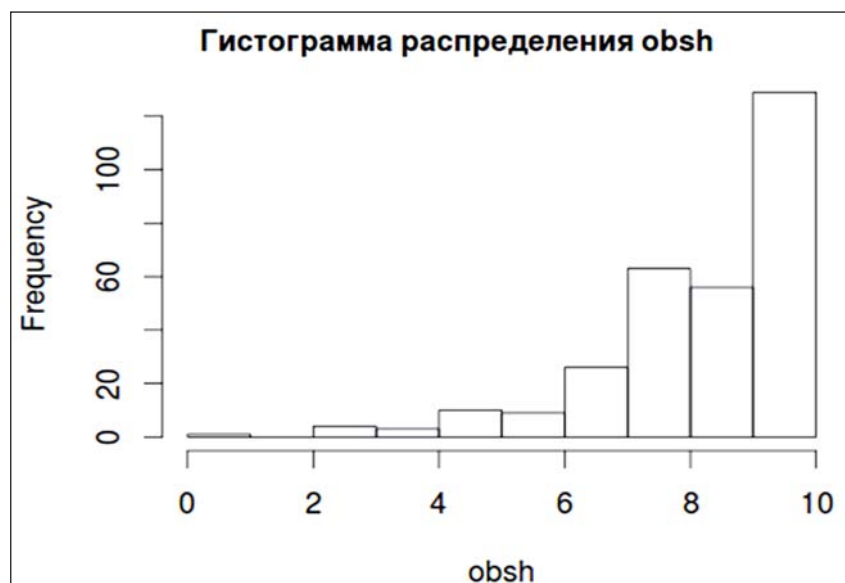


Рис. 1. Гистограмма распределения оценок по полю «Общие впечатления о проведенном обучении» (обозначения по меткам табл. 1)

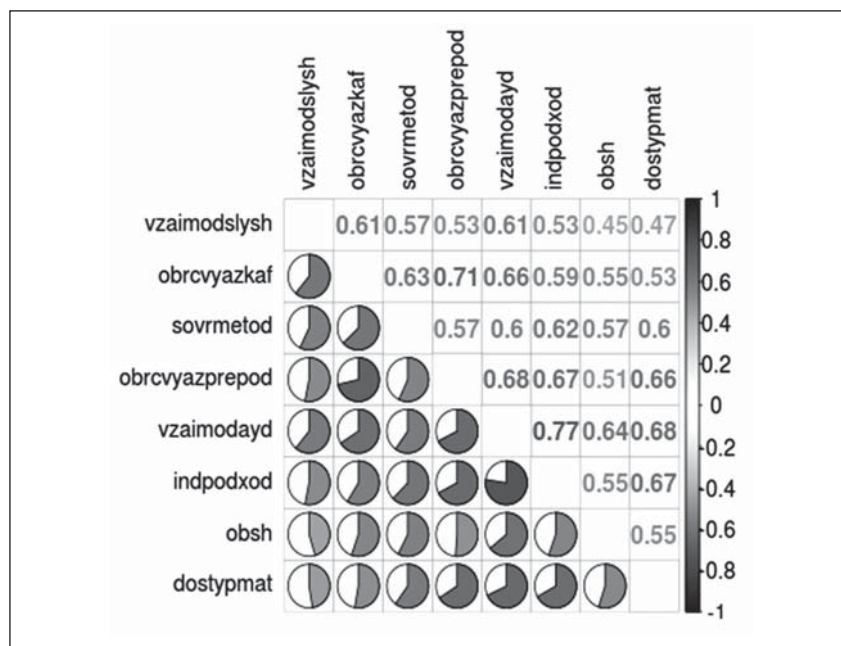


Рис. 2. Матрица корреляций по Спирмену

В указанном случае корреляции значимы на уровне $p = 0.05$.

Исходя из полученных результатов на полном наборе данных, можно утверждать, что присутствует высокий уровень корреляции между показателями «Индивидуальный подход к слушателю» и «Эффективность взаимодействия преподавателя с аудиторией» и «Получение обратной связи от преподавателей» и «Получение обратной связи от работников кафедры». Таким образом, можно с уверенностью говорить о том, что слу-

шатели практически не разделяют индивидуальный подход и эффективность преподавателя, а также видят в преподавателе представителя кафедры.

Стоит отметить большую долю корреляций в рассматриваемом наборе, значения которых могут быть трактованы как средние (значения от 0.5 до 0.7).

При оценивании процессов часто возникают группы лиц, действующих сходным образом, поэтому следующим вопросом, который решался в рамках поставленной за-

дачи, было получение информации о наличии либо отсутствии таких групп. К тому же результаты первых шагов дают повод к размышлениям о наличии «мусора» – формально заполненных опросных листов.

Для этих целей к выборке была применена процедура кластерного анализа с использованием алгоритма fuzzy C-means с предопределенным количеством кластеров, равным 4. В результате было получено четыре группы, одна из которых составила порядка 40%, остальные группы – от 17 до 20% от мощности исходной выборки (табл. 2).

Для оценки различий между полученными группами были рассчитаны медианы и межквартильный размах по всем показателям (табл. 3).

В целом стоит отметить достаточно высокие оценки по всем параметрам процесса слушателями групп 2–4.

Остановимся на некоторых выводах, которые позволяет сделать имеющаяся информация.

Группа 2 слушателей оценила все параметры на 10, а «Оснащение и оборудование учебной аудитории» и «Качество питания в столовой» в большинстве случаев на 9–10 баллов. Скорее всего, эта группа слушателей подошла формально к заполнению бланков и должна быть исключена из дальнейшего рассмотрения.

Таблица 3

Межквартильный размах и медианы показателей по группам

Группа	Межквартильный размах				Медиана			
	1	2	3	4	1	2	3	4
Общая организация процесса обучения	1,0	0,0	2,0	1,0	7,5	10,0	9,0	8,0
Оснащение и оборудование учебной аудитории	3,0	1,0	3,5	2,0	7,0	10,0	8,0	8,0
Соответствие времени проведения занятий заявленному по расписанию	2,0	0,0	0,5	1,0	9,0	10,0	10,0	10,0
Соответствие содержания обучения заявленной программе	1,0	0,0	1,0	2,0	8,0	10,0	10,0	9,0
Практическая полезность проведённого обучения	3,0	0,0	2,5	2,0	7,0	10,0	8,0	8,0
Доступность изложения материала	1,0	0,0	1,0	1,0	8,0	10,0	10,0	8,0
Использование современных методов подачи материала	1,8	0,0	1,0	1,0	7,0	10,0	10,0	8,0
Индивидуальный подход к слушателю	2,0	0,0	1,0	2,0	7,0	10,0	10,0	8,0
Эффективность взаимодействия преподавателя с аудиторией	1,0	0,0	1,0	1,0	8,0	10,0	10,0	9,0
Получение обратной связи от преподавателей	1,0	0,0	0,0	2,0	8,0	10,0	10,0	9,0
Получение обратной связи от работников кафедры	2,0	0,0	0,0	2,0	7,0	10,0	10,0	9,0
Взаимодействие слушателей между собой	1,0	0,0	0,0	2,0	8,0	10,0	10,0	9,0
Наличие, качество, полезность раздаточных материалов	3,0	0,0	3,0	2,0	6,5	10,0	8,0	8,0
Качество питания в столовой	2,8	1,0	2,5	4,0	7,0	10,0	9,0	8,0
Общие впечатления о проведенном обучении	2,0	0,0	2,0	1,5	7,0	10,0	9,0	8,0

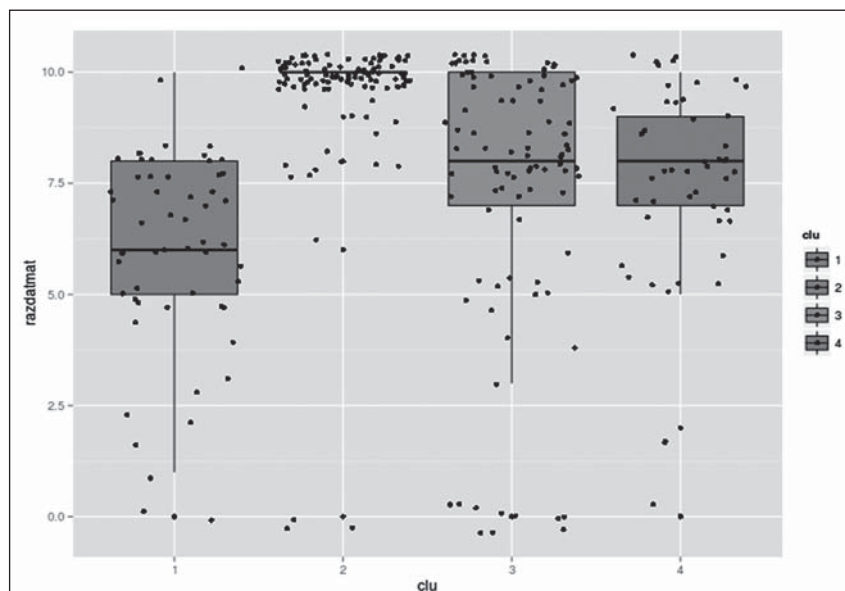


Рис. 3. Оценка раздаточных материалов в группах 1–4

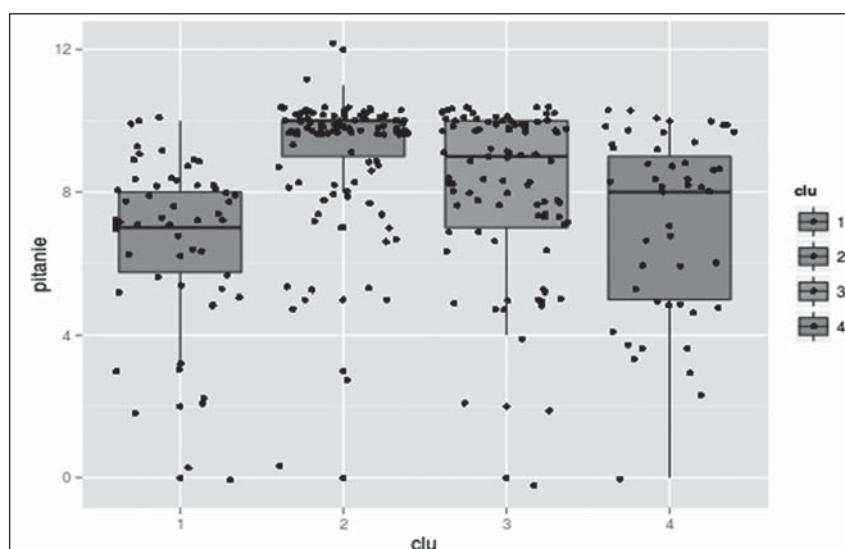


Рис. 4. Оценки фактора «Качество питания в столовой» в различных группах

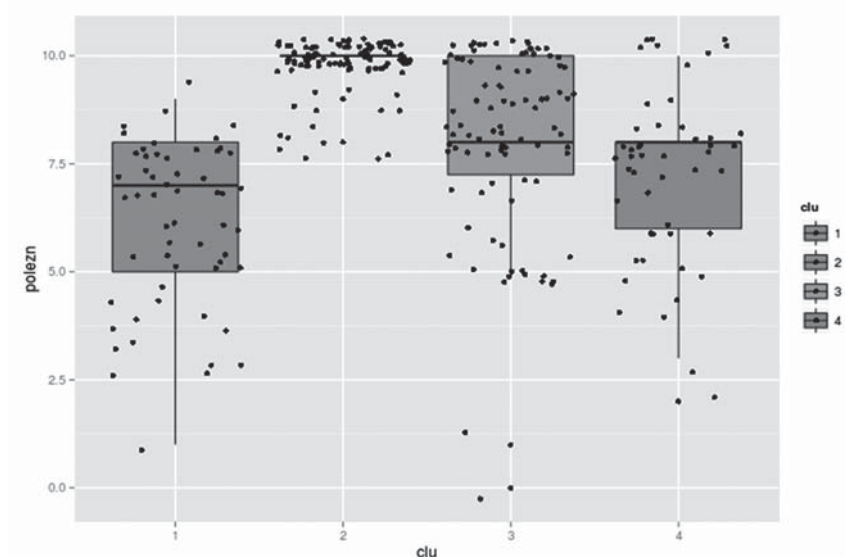


Рис. 5. Оценки фактора «Практическая полезность проведённого обучения» в различных группах.

Группа 1 наиболее критично отнеслась к оценке курсов. Особо стоит отметить низкую оценку по параметру «Наличие, качество, полезность раздаточных материалов» в этой группе (рис. 3). Вообще, если говорить о раздаточных материалах, то различия между группами по данному фактору значимы на уровне 0.05 (Kruskal – Wallis chi-squared = 20.5337, df = 2, p = 3.477e-05). Построение диаграмм проводилось с помощью пакета ggplot2 языка R [4].

Аналогичная картина наблюдается по параметру «Качество питания в столовой». Различия между группами по данному фактору значимы на уровне 0.05 (Kruskal – Wallis chi-squared = 20.0733, df = 2, p = 4.377e-05), при этом в группе 4 размах очень велик – первая квартиль равна 5 баллам (рис. 4).

Различия по показателю «Практическая полезность проведённого обучения» также статистически значимы на уровне 0.05 (Kruskal-Wallis chi-squared = 38.5622, df = 2, p = 4.23e-09), при этом первая группа оценивает в 25% случаев данный фактор на уровне, меньшем пяти баллов (рис. 5).

Аналогичная картина наблюдается и по параметрам «Оснащение и оборудование учебной аудитории», «Доступность изложения материала» и «Индивидуальный подход к слушателю».

Около трети аудитории находится в группах 1 и 4, в которых и наблюдается относительно низкое оценивание по описанным выше параметрам (оценки всё же редко опускаются ниже 5, т.е. в область полной неудовлетворенности сложившимся положением).

Для выявления скрытых факторов и их относительной важности был использован факторный анализ. Предварительно из рассмотрения была исключена 2-я группа слушателей, как, вероятно, подошедшая формально к заполнению листов.

Фильтрация привела к потере значимой части линейных связей (рис. 6).

Анализ статистически значимых показателей приводит к выводу о том, что слушатели тесно увя-

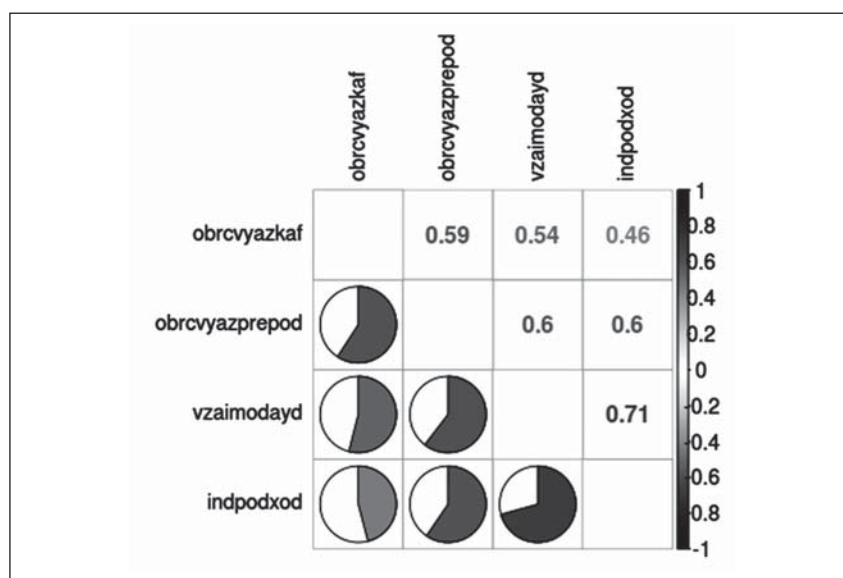


Рис. 6. Корреляция по Спирмену после фильтрации 2-й группы слушателей

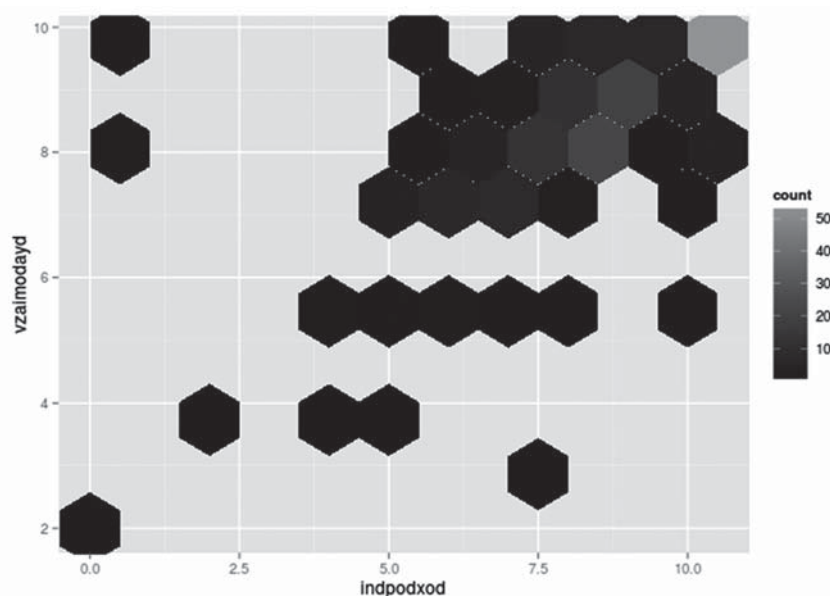


Рис. 7. Диаграмма рассеяния для показателей «Индивидуальный подход к слушателю», «Эффективность взаимодействия преподавателя с аудиторией»

зывают «Индивидуальный подход к слушателю», «Эффективность взаимодействия преподавателя с аудиторией», «Получение обратной связи от преподавателей» (коэффициент корреляции не менее 0.6). Сохраняется средний уровень связи (0.59) между факторами «Получение обратной связи от преподавателей» и «Получение обратной связи от работников кафедры».

Особо выделим сильную связь между параметрами «Индивидуальный подход к слушателю» и «Эффективность взаимодействия преподавателя с аудиторией». Связь не тривиальна по структуре, что показывает диаграмма рассеяния (рис. 7).

Результаты факторного анализа (пакет psych [5]) позволяют говорить о наличии пяти скрытых факторов (MLE Chi Square = 39.804, $p < 0.134$) (табл. 4).

Данные пять факторов объясняют порядка 50% вариации выборки, при этом факторы 1 и 2 объясняют по 16% вариации, факторы 4, 5 и 3 – по 8, 5, 4% соответственно.

Графическая интерпретация результатов (рис. 8) дает наглядное представление о структуре факторов.

Первый фактор можно трактовать как «Индивидуальное мастерство преподавателя» (индивидуальное мастерство, эффективное взаимодействие с аудиторией, доступность изложения материала), второй – «Эффективность работы кафедры» (общая организация, современные методы преподавания,

Таблица 4

Нагрузки параметров

Параметр	MR1 (Фактор 1)	MR2 (Фактор 2)	MR4 (Фактор 4)	MR3 (Фактор 3)	MR5 (Фактор 5)
Общая организация процесса обучения	–0,003	0,717	0,289	0,228	0,026
Оснащение и оборудование учебной аудитории	0,137	0,501	–0,028	0,002	0,021
Соответствие времени проведения занятий заявленному по расписанию	0,001	0,416	0,175	0,035	0,358
Соответствие содержания обучения заявленной программе	0,332	0,131	0,554	–0,011	0,327
Практическая полезность проведённого обучения	0,228	0,135	0,468	0,050	0,066
Доступность изложения материала	0,571	0,135	0,501	–0,009	–0,048
Использование современных методов подачи материала	0,203	0,693	0,274	–0,135	–0,063
Индивидуальный подход к слушателю	0,786	0,187	0,143	–0,066	–0,029
Эффективность взаимодействия преподавателя с аудиторией	0,761	0,169	0,166	0,147	0,241
Получение обратной связи от преподавателей	0,654	0,192	0,361	0,147	0,136
Получение обратной связи от работников кафедры	0,254	0,596	0,063	0,194	0,273
Взаимодействие слушателей между собой	0,186	0,547	0,042	–0,009	0,301
Наличие, качество, полезность раздаточных материалов	0,068	0,079	0,074	–0,221	0,326
Качество питания в столовой	0,101	0,109	0,059	0,716	–0,096

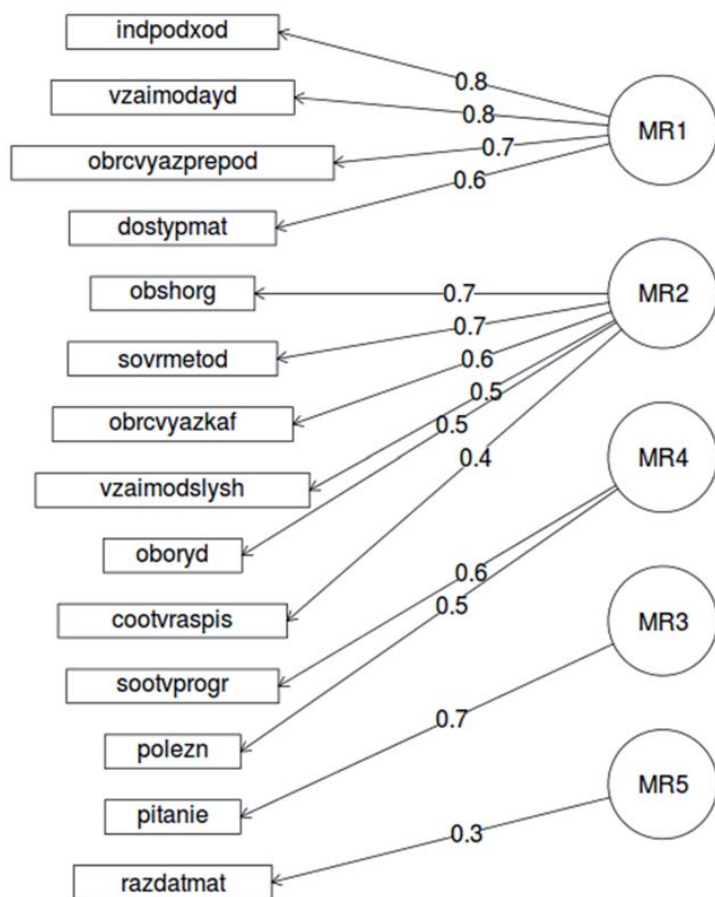


Рис. 8. Графическая интерпретация результатов факторного анализа

обратная связь, взаимодействие слушателей, оборудование, соответствие расписанию), четвертый – «Полезность обучения» (соответствие программе, полезность). Отдельно идут качество питания и раздаточных материалов.

В соответствии с ранжированием факторов наибольший вес имеет индивидуальное мастерство преподавателя и эффективность работы кафедры. Управление, нацеленное на решение именно этих проблем, даст наибольший эффект. Полезность является также достаточно важным фактором, который требует особого внимания, с учетом полученных ранее результатов.

3. Data Mining

Для оценки возможности прогнозирования общего впечатления от курсов было проведено исследование выборки с помощью различных алгоритмов классификации Data Mining в KNIME.

Предварительно было проведено изменение шкалы по полю

«Общие впечатления о проведенном обучении» с целью увеличения частот по возможным вариантам ответа: значения от 0 до 6 кодировались значением 1 (Плохо), от 7 до 8–2 (Удовлетворительно), от 9 до 10–3 (Отлично).

Результат следует признать не очень впечатляющим: оценка аккуратности при кросс-валидации методом leave-one-out дала результат в 63% при применении алгоритма Decision Tree (оценка качества – Gain Ratio, минимально количество записей на узел – 10). Оценка F-measure по 20%-й тестовой выборке дала относительно хороший результат для класса «1» (0.8) и удовлетворительный для классов «2» и «3» (0.6 и 0.632 соответственно, что объясняется низкой специфичностью класса «3» – 0.565 и низкой чувствительностью класса «2» – 0.529).

Полученная модель дает общее представление о структуре оценивания слушателями курсов (рис. 9).

Так, большая часть всех слушателей (14 из 20), оценивших от 0 до 4 общие впечатления от курса, пос-

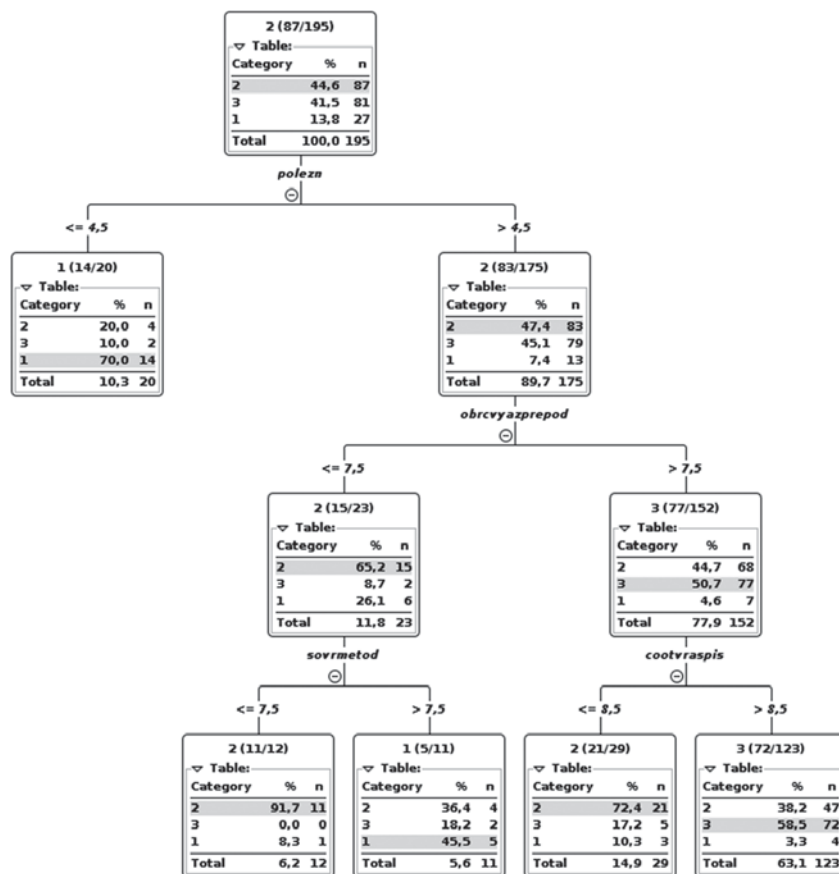


Рис. 9. Дерево решений, первые три деления

читали курс бесполезным для себя (оценки ниже 5 баллов).

С другой стороны, полезность, оцененная в 9–10 баллов, при оценках обратной связи с кафедрой в 9–10 баллов, соответствии занятий расписанию в 9–10 баллов, обратной связи с преподавателем в 8–10 баллов дает общую оценку в 9–10 баллов (84%, 36 из 43 слушателей).

Если же при выполнении всех предыдущих условий полезность оценивается в диапазоне от 6 до 8 баллов, то основную роль играют питание и общая организация процесса: питание, оцененное в 8–10 баллов, и общая организация, оцененная в 9–10 баллов, дают результат «Отлично» (73%, 19 из 26 слушателей). Если же питание оценено в 7 баллов и ниже, результат в большинстве случаев – «Удовлетворительно» (78%, 14 из 18 слушателей).

Остальные диапазоны указанных выше факторов дают неустойчивый результат ввиду малых частот.

4. Выводы

На основании проведенного анализа для нашей, конкретной задачи, поставленной отделом кадров, можно сформулировать следующие рекомендации:

- акцентировать внимание слушателей на вдумчивое заполнение опросного листа, так как количество брака существенно и достигает 40% от общего числа. Стоит рассмотреть вопрос об электронной форме тестирования и заполнении дополнительных полей (возраст, стаж, образование и пр.) с анонимизацией опросника;

- как первоочередные для улучшения показатели можно выделить: «Наличие, качество, полезность раздаточных материалов», «Качество питания в столовой», «Практическая полезность проведенного обучения», «Оснащение и оборудование учебной аудитории», «Доступность изложения материала» и «Индивидуальный подход к слушателю»;

- при найме преподавателей стоит обращать внимание на готовность работать в сетке расписания, максимальную нацеленность на индивидуальную работу, эффек-

тивное взаимодействие со слушателями;

- хорошее питание, общая организация процесса и эффективная обратная связь с кафедрой способны значительно повысить общую удовлетворенность слушателей;

- особенно стоит отметить важность факторов «Качество питания в столовой» и «Общая организация процесса обучения» в тех случаях, когда прямая польза курсов для слушателей не очевидна;

- высокий уровень корреляции между параметрами «Индивидуальный подход к слушателю» и «Эффективность взаимодействия преподавателя с аудиторией» может говорить о том, что студенты будут тяжело воспринимать элементы дистанционного образования в курсах.

Как видно, использовавшийся набор показателей позволяет достаточно полно оценить удовлетворенность слушателей, выявить узкие места и дать рекомендации по улучшению процесса подготовки слушателей на курсах повышения квалификации.

С точки зрения формирования протокола воспроизводимого исследования мы можем предложить следующую последовательность шагов для обработки подобных данных:

- на первом этапе выполнить очистку данных от неполных записей и выбросов. Очистка производится только тогда, когда соответствующий метод не способен учитывать возможные пропуски в данных. В противном случае очистка данных от неполных записей может быть вредным действием;

- на втором этапе оценить нормальность данных по каждому показателю с целью определения методов статистического исследования (допустимости использования методов параметрической статистики);

- исследовать результат оценивания с помощью кластерного анализа. Этот вид анализа позволяет достаточно эффективно выделить группы пользователей, отвечавших сходным образом. Как показал рассмотренный выше кейс, кластерный анализ позволяет выделить

подозрительные на «мусор» записи (в нашем случае – высший балл по всем параметрам). В нашем случае хорошие результаты дал метод fuzzy C-means с использованием Манхэттенского расстояния. При анализе удобно визуализировать результат кластерного анализа с помощью ящиков с усами с условием (conditional boxplot) с наложением диаграммы разброса. Статистические методы, используемые в анализе: тест Манна – Уитни – Уилкоксона, тест Крускала – Уоллиса;

- при обнаружении «мусора» в записях целесообразно исключить его из дальнейшего анализа. Если доля таких записей велика, анализ далее не проводится;

- целесообразно провести корреляционный анализ с учетом типа данных. В нашем случае мы применили коэффициент корреляции Спирмена;

- для исследования структуры в ответах целесообразно провести факторный анализ (результаты такого анализа необходимо воспринимать с большой осторожностью, так как данные измерены в ранговой шкале). Существуют методы, позволяющие проводить полноценный факторный анализ и в условиях ранговой шкалы, но в данном исследовании они не использовались в силу вспомогательной роли анализа, и методы data mining. В нашем случае хороший результат показали деревья решений – этот метод имеет значимое преимущество перед такими, как нейронные сети, SVM и т. п., выражающееся в высокой степени возможности интерпретации. В случае относительно малого количества записей целесообразно огрублять шкалу для целевой переменной.

Следует отметить, что для обработки данных и подготовки отчета и данной статьи проприетарное программное обеспечение не применялось вообще. Применение R и KNIME дает возможность обрабатывать массивы данных с применением широко спектра методов статистики и Data Mining, при этом не жертвуя удобством аналитика. Следует отметить и отличную визуализацию в R: получаемые диаграммы вполне подходят как для внутрен-

него аналитического отчета, так и для презентации.

Интересен тот факт, что такой инструмент, как R, ставший де-факто стандартом в статистических расчетах и data mining за

рубежом, мало используем в РФ. Фактически, можно говорить об одном-двух изданных учебниках [6, 7] и небольшом количестве материалов, свободно обращающихся в Сети.

Указанная обработка легко реализуема и в случае дистанционного проведения курсов, а построение обработки на базе R и KNIME позволяет легко добиться воспроизводимости.

Литература

1. R Core Team (2013) R: A language and environment for statistical computing. R Foundation for Statistical Computing [Electronic resource] // Vienna, Austria. URL: <http://www.R-project.org/>.
2. Michael R. Berthold, Nicolas Cebron, Fabian Dill, Thomas R. Gabriel, Tobias Kotter, Thorsten Meinl, Peter Ohl, Christoph Sieb, Kilian Thiel, Bernd Wiswedel. Studies in classification, data analysis, and knowledge organization. knime: the konstanz information miner // Springer. 2007.
3. Taiyun Wei corrplot: Visualization of a correlation matrix. R package version 0.73 [Electronic resource]. – URL: <http://CRAN.R-project.org/package=corrplot>
4. H. Wickham. ggplot2: elegant graphics for data analysis // Springer New York. 2009.
5. Revelle, W. psych: Procedures for Personality and Psychological Research 73 [Electronic resource] // Northwestern University, Evanston, Illinois, USA. – URL: <http://CRAN.R-project.org/package=psych> Version = 1.3.10.
6. Буховец А.Г. Статистический анализ данных в системе R: учебное пособие / А.Г. Буховец, П.В. Москалев, В.П. Богатова, Т.Я. Бирючинская; под ред. проф. А.Г. Буховца. – Воронеж: ВГАУ, 2010.
7. Наглядная статистика. Используем R! / А.Б. Шипунов, Е.М. Балдин, П.А. Волкова, А.И. Коробейников, С.А. Назарова, С.В. Петров, В.Г. Суфиянов. – М.: ДМК Пресс, 2012.