

ПРИМЕНЕНИЕ ТЕХНОЛОГИЙ DATA MINING ПРИ РАЗРАБОТКЕ КОНЦЕПЦИЙ РЕКЛАМНЫХ КАМПАНИЙ

Минзов Анатолий Степанович¹, Букин Виктор Сергеевич²

¹ Доктор технических наук, профессор;
ГОУ ВПО Международный Университет природы, общества и человека «Дубна»,
Институт системного анализа и управления;
141980, Московская обл., г. Дубна, ул. Университетская, 19;
e-mail: a@minzov.ru.

² Студент;
ГОУ ВПО Международный Университет природы, общества и человека «Дубна»,
Институт системного анализа и управления;
141980, Московская обл., г. Дубна, ул. Университетская, 19;
e-mail: bukinvictor@inbox.ru.

В статье на примере рассматривается подход к обработке маркетинговой информации с использованием технологии Data Mining. Даются рекомендации по практическому использованию этих технологий в работе маркетолога.

Ключевые слова: Data Mining, WizWhy, конкурентные стратегии в маркетинге.

USE OF DATA MINING TECHNOLOGY TO CREATE ADVERTISING CONCEPTS

Minzov Anatoliy¹, Bukin Victor²

¹ Doctor of Science in Engineering, Professor;
Dubna International University of Nature, Society, and Man,
Institute of system analysis and management;
141980, Dubna, Moscow reg., Universitetskaya str., 19;
e-mail: a@minzov.ru.

² Graduate student;
Dubna International University of Nature, Society, and Man,
Institute of system analysis and management;
141980, Dubna, Moscow reg., Universitetskaya str., 19;
e-mail: bukinvictor@inbox.ru.

In the article approach to marketing data processing with the use of the Data Mining technology is examined and demonstrated, using an example. Recommendations for practical application of these technologies in work of marketing specialist are given.

Keywords: Data Mining, WizWhy, competition strategies in marketing.

Массовое внедрение информационных технологий в производственную, экономическую и другие сферы деятельности хозяйствующих субъектов привело к появлению баз данных и архивов большой размерности, обработка которых обычными статистическими методами представляет собой сложную технологическую задачу. Основная трудность заключается в невозможности систематизировать огромные объемы данных и находить в них закономерности. В связи с этим во многих компаниях появился интерес к средствам обнаружения закономерностей в базах данных большой размерности. В частности, аналитикам в области маркетинга при разработке конкурентных стратегий необходимо систематизировать большие объемы данных и находить в них скрытые закономерности и

взаимозависимости. Для решения этих задач обычно используют статистические методы обработки информации. Однако при этом выявляются только статистические линейные связи между анализируемыми переменными. Нелинейные преобразования, используемые для статистического анализа переменных, не всегда очевидны. Поэтому сегодня аналитиков и маркетологов привлекают методы обработки статистической информации, позволяющие выявлять широкий набор закономерностей. Среди этих технологий, на наш взгляд, наибольший интерес представляют методы извлечения знаний из статистических данных Data Mining (DM).

Сам термин Data Mining переводится как «извлечение информации» или «добыча данных». Нередко рядом с Data Mining встречаются слова knowledge discovery – «обнаружение знаний» и Data Warehouse – «хранилище данных». Возникновение указанных терминов, которые являются неотъемлемой частью Data Mining, связано с новым витком в развитии средств и методов обработки и хранения данных [2].

Технологии Data Mining представляют собой множество различных математических методов, предназначенных для обнаружения в информации объективных, неочевидных и в то же время практически полезных закономерностей и взаимозависимостей [3].

Неочевидных – это значит, что найденные закономерности не обнаруживаются стандартными статистическими методами обработки информации или даже опытными экспертами. Дело в том, что стандартные статистические методы преимущественно ориентированы лишь на обобщение информации, а не на ее глубокий анализ. Эксперты же будут искать закономерности на основе своего прошлого опыта. Если закономерность не укладывается в его представление, он ее никогда не обнаружит.

Объективных – это значит, что обнаруженные закономерности будут полностью соответствовать действительности в отличие, например, от экспертного мнения, которое всегда основано на субъективном и, следовательно, ограниченном видении ситуации.

Практически полезных – это значит, что полученные выводы имеют свое конкретное бизнес-значение, которое позволит повысить прибыльность бизнеса [1].

Для оценки возможности использования Data Mining при обработке маркетинговой информации нами была использована система WizWhy, выпускаемая фирмой WizSoft.

Система WizWhy была применена к данным опроса 959 студентов одного российского ВУЗа с целью, выяснить наиболее эффективные источники информации, используемые ими при выборе факультета. В качестве исходных данных были использованы материалы социологического опроса приемной комиссии негосударственного высшего образовательного учреждения. Так как эти материалы относятся к конфиденциальным данным, они были обезличены путем исключения информации о времени, месте, участниках и условиях проведения социологического исследования. Кроме того, было исключено наименование ВУЗа и другая информация, позволяющая провести его идентификацию.

В качестве целевой переменной для исследования был взят источник информации. Студенты назвали 22 источника, которыми они пользовались при выборе факультета для поступления в ВУЗ. При проведении исследований наиболее популярными оказались следующие СМИ:

- журнал «Куда пойти учиться?» (КПУ);
- «Московский Комсомолец» (МК);
- студенты данного ВУЗа, рекомендующие ВУЗ (СВ);
- справочник (СПР);
- журнал «Образование и карьера» (ОК);
- рекламный щит (РЩ).

Остальные источники информации менее значимы и объединены в категорию «другие» (см. рис. 1).

Независимыми переменными в ходе исследования являлись: пол студента (1 – мужской, 2 – женский), факультет (1 – психологический, 2 – юридический, 3 – экономический) и форма обучения (1 – очная, 2 – очно-заочная, 3 – заочная, 4 – сокращенная).

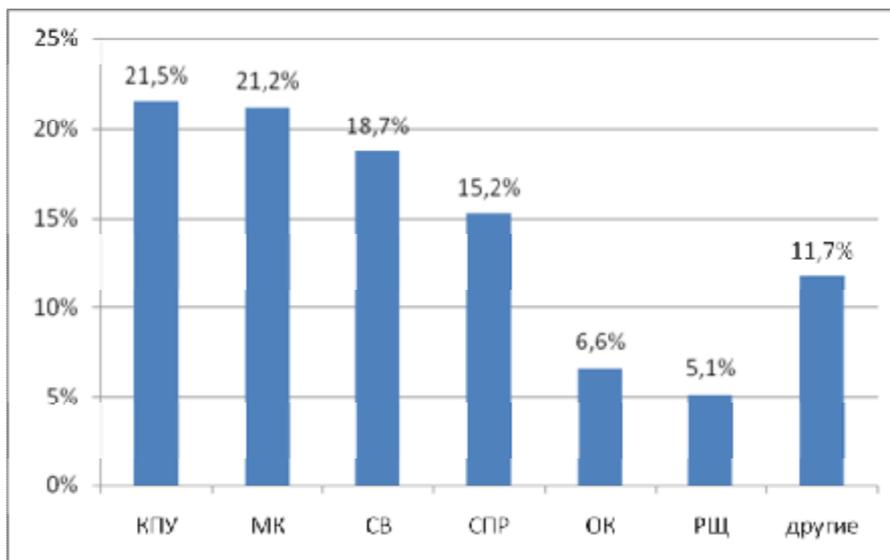


Рис. 1. Распределение СМИ по степени значимости с позиции поступивших студентов; %

Анализ рис.1 показывает, что наибольшее влияние на абитуриентов оказали следующие СМИ:

КПУ – 21,5%;

МК – 21,2%;

СВ – 18,7%.

Более глубокие результаты выявления закономерностей мотивации абитуриентов при поступлении в ВУЗ были получены в программе WizWhy в форме правил.

Система выдала одно прямое правило и 28 обратных правил (т.е. при каких случаях не рекомендуется использовать тот или иной источник информации).

Правила были представлены в следующем виде:

If Pol is 2.000000

and Fakultet is 3.000000

Then

Istochnik inform is not 13.000000 (РЦ)

Rule's probability: **0,985**

The rule exists in **258** records.

Significance Level: Error probability < 0,001

Positive Examples (records' serial numbers):

2, 3, 4, 6, 8, 9, 11, 12, 13, 14

Negative Examples (records' serial numbers):

596, 599, 763, 765.

Обобщенные результаты исследования по факультету № 1 (психологический) представлены в таблице 1.

Таблица 1. Результаты анализа по психологическому факультету

Факультет	Психологический							
Форма обучения	Очная		Заочная	Очно-заочная		Сокращенная		
Пол	м	ж	м	м	ж	м	ж	
Номер СМИ/ Достоверность правила	-	-	-	-	-	-	+	-
	СПР/0,9 РЩ/1	КПУ/0,9 РЩ/1	СПР/0,9 Др/0,9	СПР/0,9 КПУ/0,9 РЩ/1	КПУ/0,9 РЩ/1 СВ/0,9	СПР/0,9	КПУ/ 0,3	РЩ/1

В таблице используются следующие обозначения:

м – мужской пол;

ж – женский пол;

«+» – рекомендуемые СМИ;

«-» – не рекомендуемые СМИ.

По оставшимся факультетам (экономический и юридический) был проведен аналогичный анализ и все полученные результаты сведены в единую таблицу.

Таблица 2. Результаты обработки данных

Факультет	психологический		юридический		экономический	
СМИ	+	-	+	-	+	-
	КПУ, СПР, СВ	МК, РЩ, ОК	МК, КПУ, СВ, РЩ	СПР, ОК	МК, КПУ, СВ	СПР, РЩ, ОК

В таблице используются следующие обозначения:

«+» – рекомендуемые СМИ;

«-» – нерекондуемые СМИ.

Анализ таблицы 1 показал, что восприятие рекламы у абитуриентов различно в зависимости от пола и формы обучения. Последний фактор определяется, скорее всего, наличием определенного жизненного опыта у поступающих в ВУЗ. Аналогичные исследования, проведенные для других факультетов и специальностей, показали, что влияние СМИ на абитуриентов, поступающих на разные специальности, так же различно (табл. 2). Это означает, что можно повысить эффективность рекламных кампаний за счет ориентации на систему предпочтений СМИ у различных абитуриентов. По существу, у каждой специальности и формы обучения должна быть «своя» реклама.

Далее была проведена оценка эффективности СМИ с учетом стоимости рекламы по всем трем факультетам – психологическому, экономическому и юридическому. На рис. 2 показана эффективность рекламы в различных СМИ по психологическому факультету. Под эффективностью рекламы понимается относительное количество студентов (в % от общего количества опрошенных студентов факультета), поступивших на факультет по рекомендации той или иной рекламы. Стоимость рекламы указана в относительных единицах.

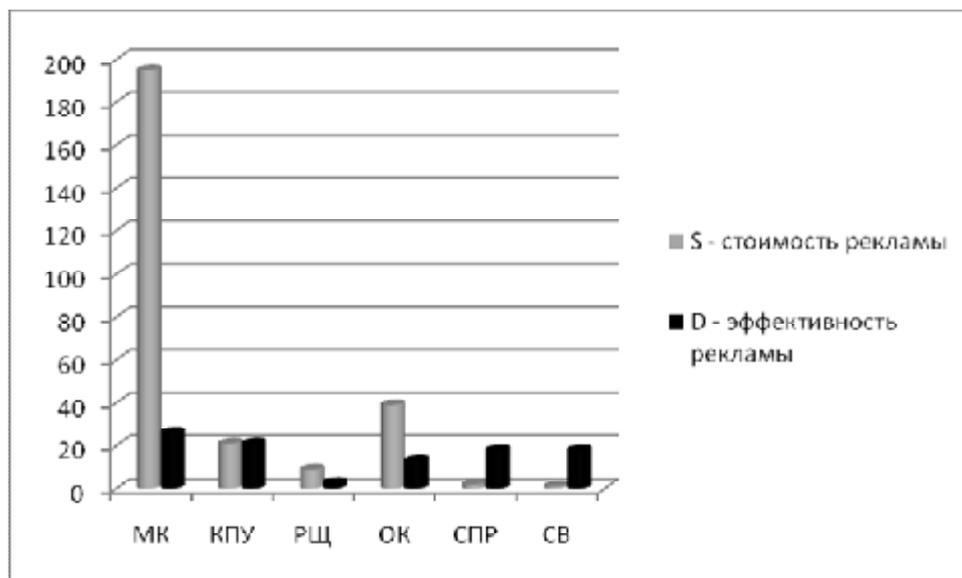


Рис. 2. Анализ эффективности рекламы для психологического факультета

С точки зрения экономической эффективности более эффективными являются малобюджетные формы СМИ.

Можно отметить, что существуют явно неэффективные средства массовой информации, на рекламу в которых потрачены большие средства, а количество студентов, пришедших на факультет по этой рекламе достаточно мало. Однако рекламная емкость этих СМИ значительно превышает емкость других. Очевидно, что при разработке стратегий рекламной кампании этот фактор необходимо обязательно учитывать. Поэтому распределение бюджета рекламных кампаний ВУЗа (S) определяется следующими факторами:

$$S = \langle O, Z, H \rangle,$$

где O – рекламная емкость СМИ;

Z – затраты на привлечение одного студента СМИ;

H – фактор предпочтений выбора СМИ абитуриентом, который, как мы показали ранее, является различным для каждой специальности (факультета) и формы обучения.

Заключение

В результате проведенного исследования выявлено, что технологии Data Mining позволяют выявить скрытые закономерности среди исследуемых переменных, в том числе нелинейные статистические связи. Эти технологии могут эффективно использоваться в маркетинговых исследованиях при разработке конкурентоспособных стратегий проведения рекламных кампаний.

Список литературы

1. Чубукова И. А. Data Mining: учебное пособие. – М.: Интернет-университет информационных технологий: БИНОМ: Лаборатория знаний, 2006. – С. 382.
2. Fayyad U., Piatetsky-Shapiro G., Smyth P., Uthurusamy R. Advances in Knowledge Discovery and Data Mining. – AAAI/MIT Press, 1996.
3. Parsaye K. A Characterization of Data Mining Technologies and Processes // The Journal of Data Warehousing. – 1998. – № 1. – С. 7-11.