

УДК 004.89

Н.А. Дергачева

Национальный аэрокосмический университет им. Н.Е. Жуковского «ХАИ»

## ПОИСК ЗНАНИЙ В БАЗАХ ДАННЫХ ДЛЯ ОЦЕНКИ ФИНАНСОВЫХ РИСКОВ

*В статье сформулированы основные задачи поиска знаний в базах данных для оценки финансовых рисков, связанных с оплатой за обучение в ВУЗе. Приведены возможные методы решения задач, основанные на применении технологий Data Mining.*

**обучение, финансовые риски, интеллектуальный анализ данных, база данных, технология Data Mining**

### Введение

**Постановка проблемы.** В настоящее время в ВУЗах широко распространена контрактная форма подготовки специалистов, которая позволяет увеличить количество студентов и обеспечить качественный уровень работы профессорско-преподавательского и инженерно-технического составов. Применение данной формы обучения связано с рядом финансовых рисков для ВУЗов и банков, кредитующих обучение студентов, связанных с несвоевременной оплатой за обучение и погашением кредитных задолженностей. Таким образом, становится актуальным решение задачи разработки методики оценки финансовых рисков, связанных с оплатой за контрактное обучение.

**Анализ литературы.** Одним из важных этапов управления рисками является сбор и обработка первичных данных [1]. Полученная информация является основой применения различных технологий интеллектуального анализа данных (ИАД) [2, 3]. В последнее время получили распространение такие методы ИАД как: методы поиска ассоциативных правил [4, 5], применение нейронных сетей [6, 7], деревьев решений [8], самоорганизующихся карт [9, 10]. Для повышения эффективности использования методов ИАД к рассматриваемым данным целесообразно применять сразу несколько алгоритмов [2]: деревья решений, нейронные сети и выводы путем сопоставления (memory-based reasoning). Если несколько моделей приводят к одинаковым результатам, значит, найден наилучший сценарий. Это особенно полезно, когда инструмент выдает ответ, но не предлагает интуитивного объяснения (наиболее часто встречающаяся проблема нейронных сетей).

**Целью статьи** является обоснование актуальности постановки задачи оценки финансового риска по характеристикам сведений о студентах, обучающихся в ВУЗах.

### Изложение основного материала исследования

Рассмотрим основные бизнес-процессы [11] в системе оплаты контрактного обучения (рис. 1). Оплата может проводиться прямыми платежами (1); при помощи кредитования (2 – 3), а также при применении смешанной формы оплаты.

Прямые платежи производятся непосредственно студентом, полностью за весь период обучения, либо периодически по частям (раз в месяц, раз в триместр, раз в год). Финансовые риски, в этом случае, связаны с несвоевременным внесением оплаты за обучение самим студентом. Если студент оплачивает всю сумму за обучение сразу, то финансовый риск в процессе обучения студента отсутствует. Если оплаты осуществляются по частям, то финансовый риск распределяется на весь период обучения.

Студент может получить банковский кредит на обучение как физическое лицо (2). В большинстве случаев для получения кредита необходим поручитель, а финансовым рискам подвергаются банки, выдавшие кредит на обучение.

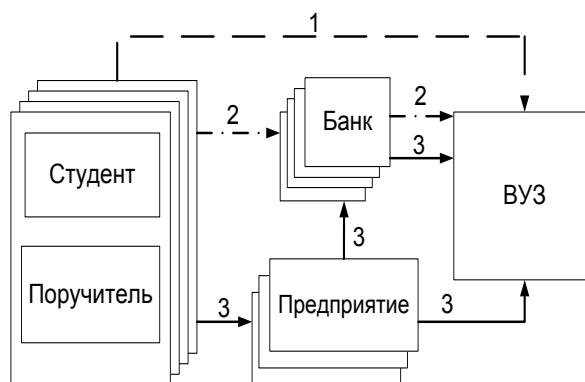


Рис. 1. Бизнес-процессы при оплате за обучение по контракту

При оформлении кредита на обучение студента могут привлекаться предприятия с различной формой собственности (3). При этом заключается дого-

вор между ВУЗом, студентом и предприятием на получение кредита. Предприятие выступает в качестве поручителя и оплачивает обучение студента в соответствии с условиями договора. Финансовым рискам подвергается непосредственно ВУЗ как субъект хозяйственной деятельности.

Анализ бизнес-процессов в системе оплаты за контрактное обучение показал, что финансовые риски при различных формах оплаты имеют различные источники, степени и величины риска, вероятности наступления рисков событий, следовательно, необходимы различные приемы управления риском и способами его снижения. Таким образом, оценку финансовых рисков необходимо проводить для каждого варианта оплаты за обучение в отдельности.

Процесс управления финансовым риском можно разделить на два этапа: этап – анализ риска; этап принятия мер по устранению и минимизации. Анализ финансового риска включает: сбор и обработку данных по аспектам риска, качественный и количественный анализ риска.

Меры по устранению и минимизации финансового риска включают: выбор и обоснование предельно допустимых уровней риска, выбор методов снижения риска, формирование вариантов рискового

вложения капитала, оценку их оптимальности на основе сопоставления ожидаемой прибыли и величины риска.

При оценке финансовых рисков может быть применена технология поиска знаний в базах данных Data Mining (DM). Технология DM заняла свои позиции в последнее десятилетие, получив центральную роль во многих сферах бизнеса. Эта технология позволяет существенно повысить возможности организации в достижении целей. Ее популярность растет все больше, поскольку инструменты совершенствуются, получают все более широкое применение, дешевеют и становятся проще в использовании.

В основу технологии поиска знаний в данных положена концепция шаблонов (образов). Эти шаблоны представляют собой функциональные, логические и другие закономерности, свойственные подвыборкам данных и могут быть компактно выражены в понятной человеку форме.

Важным требованием к разыскиваемым шаблонам является их неожиданность, нетривиальность, неочевидность (unexpected regularities).

Весь процесс выделения знаний из данных принято делить на этапы, которые представлены на рис. 2.

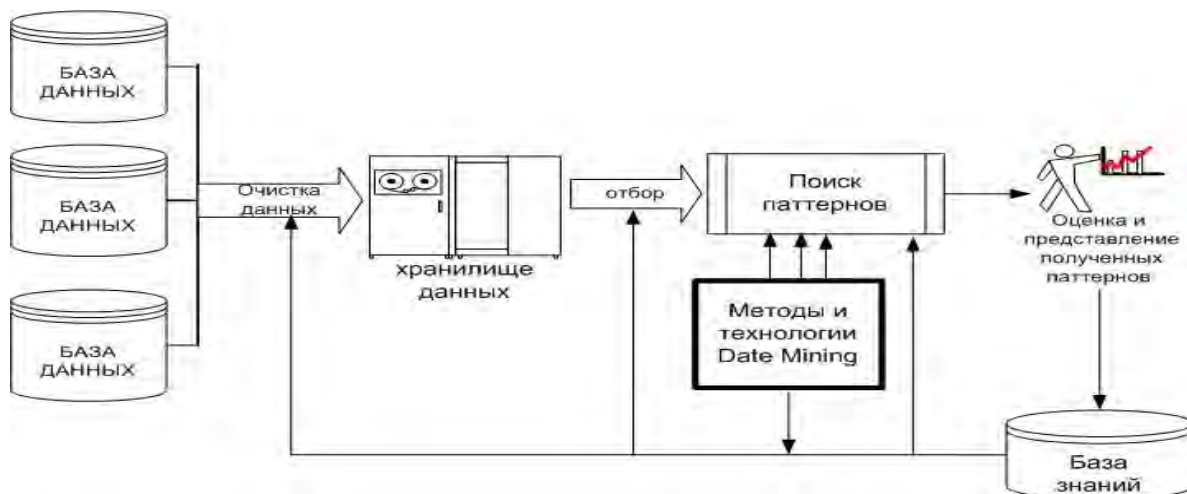


Рис. 2. Процесс обнаружения знаний в базах и хранилищах данных

Большинство авторов приводит классификацию задач Data Mining по типам производимой информации [2, 3, 5] и выделяет четыре наиболее типичные задачи извлечения знаний из данных в виде нахождения регулярностей в больших массивах информации:

- 1) построение классификатора;
- 2) кластеризация;
- 3) вывод правил ассоциации;
- 4) восстановление моделей зависимости.

Для решения реальной практической задачи, как правило, требуется применение нескольких подходов и методов.

Задача построения классификатора заключается в построении решающего правила на основании уже классифицированных объектов (обучающей выборки).

С помощью решающего правила (набора правил) объект неизвестного класса может быть отнесен к какому-либо из имеющихся классов.

Под задачей кластеризации понимается разбиение множества объектов на заданное или неизвестное число классов (задача кластеризации) на основании некоторого математического критерия качества классификации, напрямую связана с классификацией.

Вывод ассоциативных правил заключается в нахождении шаблонов имплицитивного вида, отражающих набор взаимно сопутствующих событий. Каждое правило должно иметь меру определенности не ниже заданного уровня. Ассоциативное правило является, таким образом, логической закономерностью вида "Если...то...", свойственной подвыборке данных и обладающей статистической состоятельностью. Для поиска ассоциативных правил применяются методы ограниченного перебора или строят деревья покрытий

Так в случае оплаты за обучение при непосредственных прямых платежах возможно применение методов построения ассоциативных правил, с целью получения ассоциативного портрета «потенциально неплательщика за обучения», на основе анализа базы данных о сроках оплаты и должниках за обучения. И наоборот, построить ассоциативный портрет «идеального плательщика». Подобные решения могут быть получены на основе применения нейронных сетей обратного распространения (back propagation) или самоорганизующихся картах Кохонена, которые помогут зная информацию о некоторой из части исследуемых объектов (студентах, совершающих оплату за обучение), достаточно достоверно судить по ряду общих признаков об объектах, с которыми мы мало знакомы.

Рассматривая финансовые риски при кредитовании физических лиц для оплаты за обучение, необходимо отметить, что известна классическая задача Data Mining – задача о предоставлении кредита и ее решение методом построения дерева решений. На сегодняшний день известно достаточно много методик кредитного скоринга. Одной из самых известных является модель Дюрана. Однако, известные методы принятия решения в большинстве случаев неприменимы [8] по причине плохой адаптируемости, ее высокой стоимости и большой вероятности ошибки модели при определении кредитоспособности потенциального заемщика, обусловленной субъективным мнением специалиста. Таким образом, необходим новый надежный инструмент для эффективного решения задач о предоставлении кредита, основывающийся на применении технологий Data Mining. Также, по мнению авторов, может явиться интересной задачей об оценке финансовых рисков при предоставлении кредита в условиях недостоверной информации (например, о доходах за последнее время).

При анализе финансового риска предоставления кредита для оплаты за обучение студента с привлечением различных предприятий, наряду с классическим решением задачи о предоставлении кредита, необходим анализ финансово-экономической деятельности предприятия, выступающего гарантом

возвращения кредита за обучение студента с привлечением технологий Data Mining. На основе анализа имеющихся данных о выплатах предприятиях, задолженностях по выплатам, их известных финансово-экономических показателях возможно построение ассоциативных правил, и на их основе, ассоциативного портрета «предприятия не являющегося гарантом выплаты кредита» за обучение студента и наоборот «предприятия – надежного гаранта выплаты кредита».

## Выводы

В статье сформулированы основные задачи поиска знаний в базах данных для оценки финансовых рисков, связанных с оплатой за обучение в ВУЗе. Приведены возможные методы решения задач, основанные на применении технологий Data Mining.

## Список литературы

1. Data Mining -- добыча данных [Электрон. ресурс]. – Метод доступа: <http://www.basegroup.ru/tasks/datamining.htm>.
2. Колесов А.А. Интеллектуальный анализ данных в "1С:Предприятии 8.0" [Электрон. ресурс]. – Метод доступа: [http://www.visual.2000.ru/kolesov/byte/2006/1c-ia/iad\\_1c.htm](http://www.visual.2000.ru/kolesov/byte/2006/1c-ia/iad_1c.htm)
3. Введение в анализ ассоциативных правил [Электрон. ресурс]. – Метод доступа: <http://www.basegroup.ru/rules/intro.htm>.
4. Выявление обобщенных ассоциативных правил - описание алгоритма [Электрон. ресурс]. – Метод доступа: <http://www.basegroup.ru/rules/generalized.htm>.
5. Нейронные сети - математический аппарат [Электрон. ресурс]. – Метод доступа: <http://www.basegroup.ru/neural/math.htm>.
6. Нейронные сети как средство добычи данных [Электрон. ресурс]. – Метод доступа: <http://www.basegroup.ru/neural/ns.htm>
7. Использование деревьев решений для оценки кредитоспособности физических лиц. – [Электрон. ресурс]. – Метод доступа: <http://www.basegroup.ru/practice/solvency.htm>
8. Самоорганизующиеся карты - математический аппарат [Электрон. ресурс]. – Метод доступа: <http://www.basegroup.ru/neural/som.htm>.
9. Пастухов Е.С. Кутыин В.М. Методика оценки изменений в банковской среде на основе технологии самоорганизующихся карт признаков [Электрон. ресурс]. – Метод доступа <http://www.bankclub.ru/seminar-article.htm>.
10. Торнтуэйт У. Как начать внедрение технологии data mining, не теряя ее ценных возможностей [Электрон. ресурс]. – Метод доступа: <http://citycity.ru/12996/>.

Поступила в редколлегию 22.12.2006

**Рецензент:** канд. техн. наук, доцент И.В. Шостак, Национальный аэрокосмический университет им. Н.Е. Жуковского «ХАИ», Харьков.