

3. В исследовании были проанализированы взятые из базы данных Orbis на 2007 г. 37 млн. инвестиционных групп и бизнес-компаний из которых были выделены 43 тыс. транснациональных корпораций. С помощью методов системного анализа была построена сеть, а углубление анализа позволило сделать опубликованные выводы. В том числе, были выделены 1318 наиболее влиятельных компаний, связанных настолько тесно, что делят между собой и фонды, и прибыль. Дальнейший анализ и позволил выделить те 147 компаний, которые в сущности и делают погоду на рынке. Забавно, что в итоговый список на 34 месте значится печально известный американский инвестиционный банк Lehman Brothers, который еще существовала в 2007 г. (обанкротился в результате кризиса в 2008 г.). Изложено на основании данных: Coghlan A., MacKenzie D. Revealed – the capitalist network that runs the world // New Scientist, 2011, 24 October. // <http://www.newscientist.com/article/mg21228354.500-revealed—the-capitalist-network-that-runs-the-world.html> (Режим доступа 3 ноября 2011 г.).

4. В действительности их больше, чем в исследовании швейцарских ученых. Они анализировали базу Orbis 2007 г., содержащую 37 млн. единиц, однако на июль 2011 года эта база содержала уже свыше 85 млн. единиц (<http://www.bvdinfo.com/Products/Economic-and-M-A/Economic-data/EIU-CityData>) (Режим доступа 27 августа 2011 г.).

5. Коуз Р. Фирма, рынок и право: сборник статей. – М.: Новое издательство, 2007. – 224 с. Впервые работа была опубликована в 1937 г.

6. Уильямсон О.И. Экономические институты капитализма: Фирмы, рынки, «отношенческая» контракция. – СПб.: Лениздат; CEV Press, 1996. – 702 с.

7. Колодия Г.В. Трансакционные издержки как фактор эволюции фирмы: автореф. дис... д.э.н.: 08.00.01. – М.: Финансовая акад. при Правительстве РФ, 2008. – 43 с.

8. Литвинцева Г. Словарь терминов по институциональной экономике: Национальная экономическая энциклопедия // <http://vocabulary.ru/dictionary/96> (Режим доступа 15 февраля 2011 г.).

*М.А. Зильберглейт, д-р. хим. наук,
М.В. Самойлов, канд. техн. наук, БГЭУ (Минск)*

ИСПОЛЬЗОВАНИЕ МЕТОДОВ DATE MINING ДЛЯ ОЦЕНКИ ДЕЙСТВИЙ ЭКСПЕРТОВ ПО ЦЕННЫМ БУМАГАМ

Настоящая статья посвящена использованию категориальных переменных в виде преобразованной шкалы интервалов для оценки действий экспертов по ценным бумагам при использовании методов ассоциативных правил и деревьев решений. Обычное использование методов Date Mining предполагает обработку значительного числа числовых данных с целью «выявления знаний». Однако на практике встречаются случаи, когда количество экспериментальных данных не велико и тогда, как правило, использование данных из шкалы отношений к успеху не приводит. Переход к категориальным переменным позволяет сгруппировать данные и тем самым улучшить их структуру. Так как понятие категориальных данных в литературе до конца не определено, то в нашем случае к ним мы относим данные, полученные при разбиении числовой шкалы на интервалы и приписыванию значение переменных, попавших в один и тот же интервал собственного имени. В этом случае можно говорить о номинальных переменных.

Данные заимствованы нами из базы данных по применению пакета Statgraphics, в котором приведены действия экспертов ряда инвестиционных фондов. Эксперт при-
104

нимает следующие действия: купить (Buy), продать (Sell), держать (Hold). В качестве факторов, которыми он пользуется, являются – доходность за пятилетний период (Five_Yr), риск (Risk), средний процент дохода за пятилетний период (Perf 90-94), расходная часть (Expense), налоговый рейтинг (Tax). Фрагмент данных по действию экспертов представлен в табл. 1. Соответственно были образованы следующие категориальные переменные Var 01: Доходность за пятилетний период – больше 14000 (R3), 12000-14000 (R2), менее 12000 (R1); Var 02: Риск 1-2 (Z1), 3-5 (Z2); Var 03: Средний процент дохода за пятилетний период – до 15% (D1), больше 15% (D2); Var 04: Расходная часть – менее 1,1 (A1), более 1,1 (A2), Var 05: налоговый рейтинг – менее 70 (B1), 70-85 (B2), более 85 (B3).

Как известно алгоритм ассоциативных правил входит в набор методов Data Mining и позволяет сформулировать набор правил при анализе транзакции в виде «Если А, то В». Вывод правил сопровождается поддержкой правила (%), достоверностью правила (%) и т.н. называемой корреляцией (%), что в данном методе можно характеризовать как нормированную величину достоверности.

В результате применения метода ассоциативных правил при минимальной поддержке и достоверности в 15% были получены следующие ассоциативные правила.

Таблица 1. Фрагмент данных об инвестиционных фондах

Fund	Five_Yr	Risk	Perf94	Perf93	Perf92	Perf91	Perf90	Expense	Tax	Recommend
F. Chip	16476	2	10	25	6	55	4	1,22	89	Buy
F. Contra	15476	2	-1	21	16	55	4	1,03	90	Buy
F. Destiny	14757	3	4	26	15	39	-3	0,7	69	Buy
Vista A	15145	4	-1	20	13	71	-6	1,49	96	Hold
Berger 100	15596	5	-7	21	9	89	-6	1,7	95	Hold
Gab. Assett	13640	1	0	22	15	18	-6	1,33	85	Buy
Neub Focus	14081	3	1	16	21	25	-6	0,85	75	Buy
F. Magellan	13827	3	-2	25	7	41	-5	0,96	73	Buy
Janus	13187	2	-1	11	7	43	-1	0,91	85	Sell
L. Mason Value	13029	4	1	12	11	35	-17	1,82	92	Hold
Gabelli Growth	12301	3	-3	11	4	34	-2	1,41	80	Buy
Franklin Growth	11793	2	3	7	3	27	2	0,77	90	Sell
Janus 20	12441	4	-7	3	2	69	1	1,02	95	Sell
AARP Capital	11728	4	-10	16	5	41	-16	0,97	68	Sell
Kemper Growth A	11386	4	-6	2	-2	67	4	1,09	86	Sell
20th Cent. Growth	11258	4	-8	15	-4	32	0	1	60	Buy

1. Если средний процент дохода за пятилетний период равен D2 и расходная часть A2 и налоговый рейтинг B2, то покупать (поддержка 15%, достоверность 100%, корреляция 54,8%).

2. Если риск равен Z2 и расходная часть A2 и налоговый рейтинг B3, то держать (поддержка 15%, вероятность 100%, корреляция 100%).

3. Если доходность за пятилетний период R2 и средний процент дохода за пятилетний период равен D1 и расходная часть A1, то продавать (поддержка 15%, достоверность 75, корреляция 56,7%).

Естественно, что в данном случае мы приводим только самые «сильные» правила, т.е. правила имеющие наибольшую достоверность.

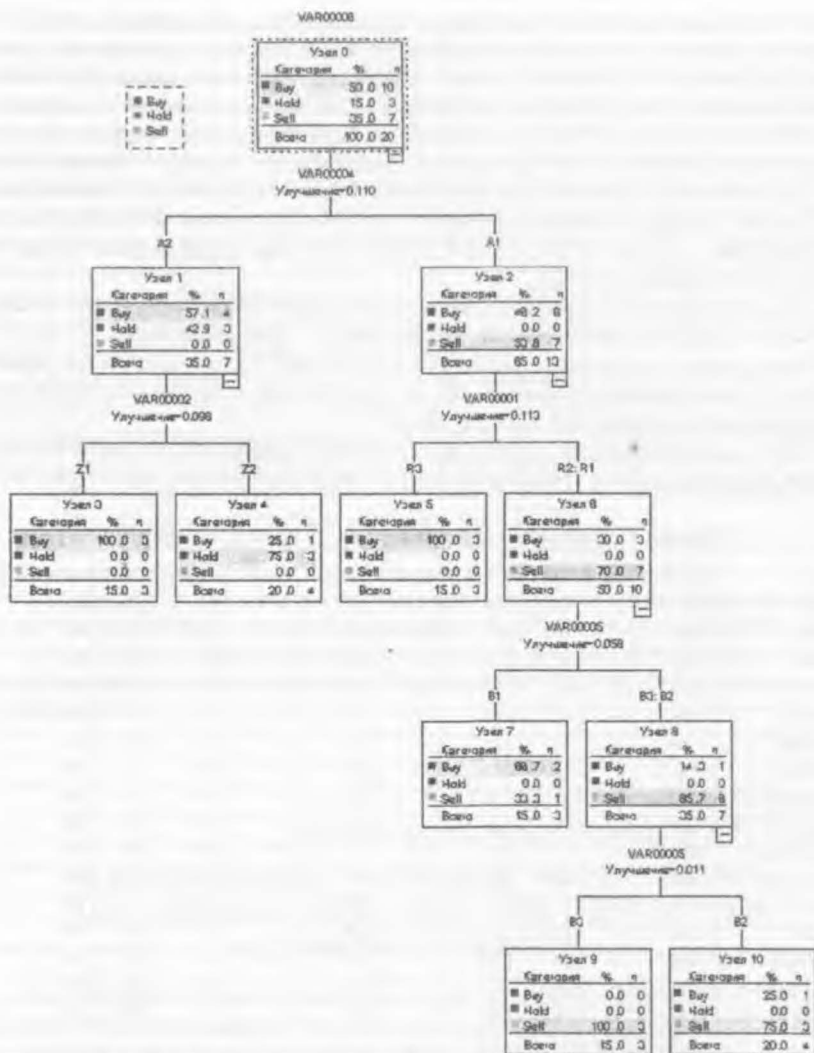


Таблица 2. Классификация методом дерева решений

Наблюдаемое	Предсказанное			Процент корректных
	Buy	Hold	Sell	
Buy	8	1	1	80,0%
Hold	0	3	0	100,0%
Sell	1	0	6	85,7%
Общий процент	45,0%	20,0%	35,0%	85,0%

Метод построения: CRT

Зависимая переменная: Recommend

Таблица 3. Важность независимых переменных

Независимая переменная	Важность	Нормализованная важность
VAR05	0,165	100,0%
VAR01	0,119	72,1%
VAR04	0,110	67,1%
VAR02	0,096	58,5%
VAR03	0,010	6,1%

Метод построения: CRT

Зависимая переменная: Recommend

Из полученных данных следует, что пользуясь деревом решений доля корректно распознанных объектов составляет 85%. При использовании числового решения данной задачи с применением линейной дискриминантной функции эффект конечно выше – 94%, однако и этот результат можно считать вполне приемлемым, так как наряду с деревом решения мы получаем ранжированное значение переменных из которых следует, что переменная – налоговый рейтинг вносит определяющий вклад в мнение эксперта, а переменная – средний процент дохода за пятилетний период, наименьший.

Некоторые программные комплексы реализующие данные алгоритмы имеют дополнительную опцию «если..., то..», которая в отличие от изложенных выше, уже по результатам предыдущего анализа позволяют пользователю ввести свои данные и получить решение частной задачи.

Зубик В.Б., кандидат экономических наук, доцент БГЭУ (Минск)

Дроздова С.Н., УО БФ «БГЭУ»

АКТИВИЗАЦИЯ МАЛЫХ ИННОВАЦИЙ

Важным стратегическим направлением развития республики в целом, и промышленных предприятий в частности, является необходимость активизации инновационной деятельности. Инновационная деятельность позволяет предприятиям, независимо от сферы их деятельности и формы собственности, переходить к выпуску новых, более качественных товаров, снижать себестоимость выпуска единицы продукции, увеличивать эффективность и рентабельность производства, расширять производственные мощности и т.д.