

Таблица 1

9.7. МОДЕЛИРОВАНИЕ КРЕДИТНОГО РИСКА ПРЕДПРИЯТИЙ НА ОСНОВЕ КЛАССИФИКАЦИОННОГО ПОДХОДА

Ледовский Д.О., аспирант экономического факультета

Московский государственный университет им. М.В. Ломоносова

Данная статья посвящена проблеме оценки уровня кредитного риска предприятий. В качестве объектов исследования выбраны 40 предприятий-эмитентов отрасли связи. Расчеты проводились по данным публичных финансовых отчетностей за 2000 и 2001 гг. В работе использовались современные экономико-математические методы: корреляционный, кластерный и дискриминантный анализ, нечеткое моделирование. Выполнены сравнения результатов статистической и лингвистической моделей кредитного риска. Предложена лингвистическая модель оценки кредитного риска с обучением.

ВВЕДЕНИЕ

Использование моделей кредитного риска – наиболее удобный способ оценить риск возникновения трудностей с обслуживанием обязательств заемщиком. Большинство моделей не дает универсальных, стопроцентно верных решений о качестве заемщика и поэтому не может заменить профессионального аналитика. Однако они могут и должны служить удобным инструментом для анализа кредитного риска, выявляя некоторые важные стороны финансово-экономического состояния заемщика.

Среди лиц, заинтересованных в оценке платежеспособности предприятий, кредитные организации, рейтинговые аналитические агентства, органы государственной власти, потенциальные и реальные контрагенты, руководство и собственники.

Предметом настоящей работы являются модели экспресс-оценки кредитного риска, основанные на использовании финансовых коэффициентов и классификационных принцип построения.

В данной работе будет проведен анализ кредитного риска по группе предприятий связи. В выборку вошло 40 предприятий-эмитентов, все относятся к отрасли связи. По каждому предприятию было рассчитано 10 показателей:

- X1_1 – коэффициент текущей ликвидности;
- X1_2 – коэффициент критической ликвидности;
- X1_3 – коэффициент абсолютной ликвидности;
- X1_4 – коэффициент обеспеченности оборотных активов собственными средствами (NWC/CA¹);
- X1_5 – коэффициент обеспеченности собственными оборотными средствами (СОК/CA);
- X2_1 – коэффициент автономии;
- X3_1 – оборачиваемость активов;
- X4_1 – рентабельность продаж (по балансовой прибыли),
- X4_2 – рентабельность активов;
- X4_3 – рентабельность собственного капитала.

В качестве источника данных использовались годовые финансовые отчетности компаний за период 2000-2001 гг., опубликованные на www.scrip.ru.

Отсутствие обучающих выборок осложняет задачу моделирования кредитного риска. Большинство западных исследований² располагало обучающими выборками для формирования решающих правил модели. Автору данной работы не удалось сформировать обучающие выборки, что связано с недостатками системы раскрытия информации. Поэтому выбор методов ограничен автоматическими процедурами, не требующими обучения (см. табл. 1).

¹ NWC (net working capital) = текущие активы – краткосрочные обязательства, CA – текущие активы, СОК (собственный оборотный капитал) = собственный капитал – стоимость основных фондов; NWC и СОК связаны следующими соотношениями: $СОК > 0 \Rightarrow NWC > 0$, $NWC < 0 \Rightarrow СОК < 0$.

² Например, E. Altman (1968), D.L. Chesser (1974), E. Altman, G. Marco, F. Varetto (1994), Atiya A. (2001) и др.

СПИСОК ОБЪЕКТОВ

№	Код	Название
1	BISV	ОАО фирма «Башинформсвязь»
2	ESPK	ОАО «Дальневосточная компания электросвязи»
3	KUBN	ОАО «Южная телекоммуникационная компания»
4	NNSI	ОАО «ВолгаТелеком»
5	SPTL	ОАО «Северо-Западный Телеком»
6	SVZK	ОАО «Связь» Республики Коми
7	URSI	ОАО «Уралсвязьинформ»
8	ELKA	ОАО «Электросвязь» Республики Карелия
9	ELRO	ОАО «Электросвязь» Ростовской области
10	ENCO	ОАО «Сибирьтелеком»
11	ESAL	ОАО «Электросвязь» Республики Алтай
12	ESHS	ОАО «Электрическая связь Республики Хакасия»
13	ESIR	ОАО «Электросвязь» Иркутской области
14	ESKI	ОАО «Кировэлектросвязь»
15	ESKK	ОАО «Электросвязь» Красноярского края
16	ESKT	ОАО «Электросвязь» Костромской области
17	ESMO	ОАО «Центральная телекоммуникационная компания»
18	ESOB	ОАО «Электросвязь Оренбургской области»
19	ESOM	ОАО «Электрическая связь» Омской области
20	ESTU	ОАО «Тулателеком»
21	ESVD	ОАО «Электросвязь» Владимирской области
22	ESVL	ОАО «Электросвязь» Вологодской области
23	HTCM	ОАО по электросвязи и телекоммуникациям в Ханты-Мансийском автономном округе
24	IVTL	ОАО «Ивановские телекоммуникационные сети»
25	KCHS	ОАО «Камчатсвязьинформ»
26	MGTS	ОАО «Московская городская телефонная сеть»
27	MUEL	ОАО «Мурманэлектросвязь»
28	NVGT	ОАО «Новгородтелеком»
29	RTKM	ОАО междугородной и международной электрической связи «Ростелеком»
30	SMSI	ОАО связи и информатики Смоленской области «Смоленсксвязьинформ»
31	SVIM	ОАО «Магадансвязьинформ»
32	SVLN	ОАО «Связьинформ» Самарской области
33	TMTK	ОАО «Томсктелеком»
34	TUTE	ОАО «Тюменьтелеком»
35	UDTL	ОАО «Телекоммуникационные сети связи Удмуртской Республики»
36	АНТС	ОАО «Артелеком» Архангельской области
37	ALCO	ОАО «Алтайская телефонно-телеграфная компания»
38	ELCH	ОАО «Электросвязь» Читинской области
39	INSW	ОАО «Инвестиционная компания связи»
40	KGTS	ОАО «Казанская городская телефонная сеть»

Цель проведенного исследования – сопоставление результатов статистической и лингвистической модели, разработка модели с обучением на базе лингвистической модели.

Исследование включает в себя несколько этапов:

- анализ финансовых коэффициентов изучаемой выборки предприятий;
- анализ статистических характеристик выборки и изучение взаимосвязи показателей;
- классификация предприятий по уровню кредитного риска.
- разработка статистической модели сравнительной оценки кредитного риска предприятий;
- разработка лингвистической модели кредитного риска с обучением;
- сопоставление и интерпретация результатов моделей кредитного риска.

В работе изучена структура данных, проведена финансово-экономическая оценка объектов, выявлены взаимосвязи показателей и отобраны наиболее информативные из них. Для этого использованы методы финансового анализа и корреляционного анализа.

На следующем этапе предложено несколько моделей оценки кредитного риска, в основе которых лежат методы прикладного статистического анализа и нечеткого моделирования, предложен подход к разработке лингвистических моделей с обучением. На завершающей стадии выполнено сравнение результатов моделей.

1. АНАЛИЗ КОЭФФИЦИЕНТОВ ГРУППЫ ПРЕДПРИЯТИЙ

На 1-м этапе исследования изучались показатели предприятий выборки, прежде всего, внимание уделено структуре выборки.

Были выявлены проблемы, общие для многих российских предприятий в 1990-х – начале 2000-х гг.: относительно низкие показатели ликвидности. Абсолютное большинство предприятий не соответствовало принятым в практике финансового анализа нормативам: ($K_{тек.ликв-ти} = 2$, $K_{крит.ликв-ти} = 1$, $K_{абс.ликв-ти} = 0.2$)³.

Собственного капитала большинству предприятий было недостаточно для финансирования внеоборотных активов, 32 предприятия имели отрицательный собственный оборотный капитал ($X1_5$). Из них 18 покрывали внеоборотные активы за счет долгосрочных кредитов и займов. Это говорит о значимости для предприятий долгосрочных заемных источников финансирования (см. табл. 2).

Таблица 2

ГРУППИРОВКА ПРЕДПРИЯТИЙ ПО КОЭФФИЦИЕНТАМ ЛИКВИДНОСТИ, 2000-2001 гг.

X1_1	2000 г.		2001 г.	
	число	%	число	%
>2	3	7.5	2	5
[1, 2)	24	60	24	60
<1	13	32.5	14	35
Всего	40	100	40	100
X1_2	2000 г.		2001 г.	
	число	%	число	%
>1	20	52.63	11	27.5
[0.7, 1)	12	31.58	17	42.5
[0.5, 0.7)	4	10.53	7	17.5
<0.5	2	5.26	5	12.5
Всего	38 ⁴	100	40	100
X1_3	2000 г.		2001 г.	
	число	%	число	%
>0.5	2	5.26	2	5
[0.2, 0.5)	8	21.05	6	15
[0.1, 0.2)	10	26.32	12	30
<0.1	18	47.37	20	50
Всего	38	100	40	100

Таблица 3

ГРУППИРОВКА ПО ПОКАЗАТЕЛЯМ ОБЕСПЕЧЕННОСТИ ОБОРОТНЫМ КАПИТАЛОМ.

2001г.	X1_4		X1_5		2000г.	X1_4		X1_5	
	число	%	число	%		число	%	число	%
> 0	26	65	8	20	>0	27	67,5	10	25
< 0	14	35	32	80	<0	13	32,5	30	75
Всего	40	100	40	100	Всего	40	100	40	100

³ См., например, Шеремет А.Д., Сайфулин Р.С. Методика финансового анализа. – М.: ИНФРА-М, 1996, Шеремет А.Д. Комплексный анализ хозяйственной деятельности. – М.: ИНФРА-М, 2006.

⁴ По 2-м предприятиям связи (NVGT, ESOB) нет данных для расчета $X1_2$ и $X1_3$ за 2000 г.

Сопоставление показателей за 2 года говорит о некотором ухудшении ликвидности изучаемой группы предприятий (см. табл. 3). Одновременно выросла доля предприятий с нехваткой чистого ($X1_4$) и собственного оборотного капитала ($X1_5$). Чтобы понять, насколько обновился состав предприятий испытывавших финансовые затруднения, необходим анализ структуры выборки с помощью таблиц сопряженности.

Числа выше главной диагонали матриц (см. табл. 4) показывают переход из более ликвидных и платежеспособных групп. Соответствующие предприятия ухудшили свое финансовое состояние по показателям текущей ликвидности и достаточности собственного и чистого оборотного капитала. Ниже главной диагонали стоят количества предприятий перешедших в «лучшие» группы.

Таблица 4

ДИНАМИКА РАСПРЕДЕЛЕНИЯ ВЫБОРКИ ПО ПОКАЗАТЕЛЯМ ЛИКВИДНОСТИ И ПЛАТЕЖЕСПОСОБНОСТИ.

2000 г	2001 г.			
	X1_1	≥2	(1, 2)	<1
≥2	2	1	0	3
[1, 2)	0	18	6	24
<1	0	5	8	13
Всего	2	24	14	40
-	X1_5>0	X1_5<0, X1_4>0	X1_4<0	Всего
X1_5>0	7	3	0	10
X1_5<0, X1_4>0	1	10	6	17
X1_4<0	0	5	8	13
Всего	8	18	14	40

На протяжении 2-х периодов 8 компаний имеют отрицательный чистый оборотный капитал: KUBN, URSI, ESIR, ESKI, ESMO, ESOB, ESOM, SVLN. 6 компаний ухудшили кредитоспособность (NWC стал отрицательным): NNSI, SVZK, ENCO, ESTU, MGTS, SMSI.

Произошло улучшение кредитоспособности на 5 предприятиях: ESKK, ESVD, IVTL, KCHS, АНТС. Таким образом, на протяжении 2-х лет проблемы с кредитоспособностью возникали на 19 предприятиях из 40.

ELRO, ESAL, NVGT, SVIM, TMTK, ALCO, INSW на протяжении 2-х периодов имели положительный собственный оборотный капитал. У 3-х компаний (ESHS, UDTL, ELCH) собственный оборотный капитал стал отрицательным, что привело к снижению их платежеспособности. Улучшение по показателю $X1_5$ произошло на ESVL.

Как видно из вышеприведенных таблиц, ухудшение произошло на большем числе предприятий, об этом же говорят средние значения показателей ликвидности.

Можно определить несколько возможных причин ухудшения:

- рост краткосрочной или долгосрочной задолженности;
- рост стоимости или доли внеоборотных активов;
- сокращение стоимости или доли оборотных активов.

В данном случае произошел незначительный рост доли оборотных активов, собственного капитала при заметных изменениях структуры долга. Сократилась доля долгосрочных кредитов и займов и выросла доля краткосрочной задолженности, в результате чего уменьшилось отношение долгосрочного к краткосрочному долгу, и снизились показатели ликвидности.

С точки зрения любой кредитной организации, значение $K_{тек.ликв-ти} < 1$ (или $NWC < 0$) означает неудовлетво-

рительную кредитоспособность. Предприятия с такими показателями будут иметь сравнительно высокий уровень кредитного риска. Это должно учитываться моделью кредитного риска.

С другой стороны, положительный собственный оборотный капитал (**СОК**) – благоприятный признак, характеризует платежеспособность предприятия, уменьшает уровень кредитного риска.

Таким образом, выборку условно можно разделить на части:

- финансово устойчивые предприятия (8 компаний) $X1_5 > 0$;
- некредитоспособные компании (14 компаний) $X1_4 < 0$ (или $X1_1 < 1$);
- переходная группа (18 компаний) $X1_5 < 0$ и $X1_4 > 0$.

Долгосрочное долговое финансирование играет важную роль, однако в данной выборке среди источников финансирования преобладает собственный капитал (в среднем по отрасли коэффициент автономии в 2001 году составлял 51,3%, а по выборке – 71,6%).

Показатели оборачиваемости и рентабельности дополняют коэффициенты платежеспособности, характеризуют эффективность и перспективы компании. Так, компания из группы некредитоспособных, но с высокой рентабельностью и оборачиваемостью активов должна быть более привлекательной, чем ее соседи по группе. Обратное относится к низкорентабельным представителям группы финансово устойчивых.

Сравнение со среднеотраслевыми показателями рентабельности может быть не вполне корректным. В-первых, методологические различия расчетов. В-вторых, в состав отрасли входят, кроме традиционных проводных телекомов, компании мобильной, спутниковой, интернет-связи и т.п., рентабельность услуг которых выше.

Компании выборки в целом характеризуются низким и средним уровнем кредитного риска. Большинство из них – крупные предприятия, лидеры региональных рынков услуг связи. Банкротство кого-либо из них практически

исключено, но задержки расчетов по кредитам и займам, кредиторской задолженности вполне возможны.

Анализ финансовых коэффициентов показывает оправданность использования классификационных процедур, дает основания ожидать появления от 2 до 5 групп.

2. АНАЛИЗ ВЗАИМОСВЯЗИ ПОКАЗАТЕЛЕЙ

Анализ и отбор ключевых показателей является важным этапом моделирования кредитного риска. Можно предположить, что некоторые показатели тесно между собой взаимосвязаны. Например, показатели текущей ликвидности ($X1_1$) и обеспеченности оборотных активов чистым оборотным капиталом ($X1_4$) связаны функциональной (но нелинейной) связью:

$$X1_1 = \frac{1}{1 - X1_4}$$

Наличие тесных связей между некоторыми показателями позволяет часть из них исключить из модели без заметной потери информации.

Для проверки предположения о наличии корреляционных связей между отдельными показателями в данном разделе будет проведен корреляционный анализ и группировка показателей.

По результатам корреляционного анализа будет принято решение о включении/исключении переменных в модель.

По всей выборке (за исключением INSW и KGTS) была рассчитана корреляционная матрица (см. табл. 5). Для оценки значимости коэффициентов парной линейной корреляции использовались *t*-статистики, уровень значимости которых или *P*-значения представлены в таблице. Если $P < 0.05$, коэффициент корреляции значим при 5%-ном уровне значимости. Это означает, что между данной парой показателей существует значимая статистическая линейная связь.

Таблица 5

МАТРИЦА ПАРНЫХ КОРРЕЛЯЦИЙ⁵

N=38	X1_1	X1_2	X1_3	X1_4	X1_5	X2_1	X3_1	X4_1	X4_2	X4_3
X1_1	1	0.9605	0.3434	0.9361	0.6401	0.1923	0.3129	0.007	0.1295	0.1438
	p= ---	p=.000	p=.035	p=.000	p=.000	p=.248	p=.056	p=.967	p=.438	p=.389
X1_2	0.9605	1	0.3543	0.9257	0.6589	0.127	0.2573	-0.0356	0.0532	0.0763
	p=.000	p= ---	p=.029	p=.000	p=.000	p=.448	p=.119	p=.832	p=.751	p=.649
X1_3	0.3434	0.3543	1	0.2905	0.1791	-0.3116	0.1931	0.1829	0.2047	0.3791
	p=.035	p=.029	p= ---	p=.077	p=.282	p=.057	p=.245	p=.272	p=.218	p=.019
X1_4	0.9361	0.9257	0.2905	1	0.7363	0.234	0.2677	0.1012	0.1792	0.1621
	p=.000	p=.000	p=.077	p= ---	p=.000	p=.157	p=.104	p=.546	p=.282	p=.331
X1_5	0.6401	0.6589	0.1791	0.7363	1	0.471	0.2512	0.2064	0.2507	0.0777
	p=.000	p=.000	p=.282	p=.000	p= ---	p=.003	p=.128	p=.214	p=.129	p=.643
X2_1	0.1923	0.127	-0.3116	0.234	0.471	1	-0.1704	0.0151	-0.1082	-0.4632
	p=.248	p=.448	p=.057	p=.157	p=.003	p= ---	p=.306	p=.928	p=.518	p=.003
X3_1	0.3129	0.2573	0.1931	0.2677	0.2512	-0.1704	1	-0.0072	0.5001	0.4521
	p=.056	p=.119	p=.245	p=.104	p=.128	p=.306	p= ---	p=.966	p=.001	p=.004
X4_1	0.007	-0.0356	0.1829	0.1012	0.2064	0.0151	-0.0072	1	0.8438	0.7123
	p=.967	p=.832	p=.272	p=.546	p=.214	p=.928	p=.966	p= ---	p=.000	p=.000
X4_2	0.1295	0.0532	0.2047	0.1792	0.2507	-0.1082	0.5001	0.8438	1	0.8767
	p=.438	p=.751	p=.218	p=.282	p=.129	p=.518	p=.001	p=.000	p= ---	p=.000
X4_3	0.1438	0.0763	0.3791	0.1621	0.0777	-0.4632	0.4521	0.7123	0.8767	1
	p=.389	p=.649	p=.019	p=.331	p=.643	p=.003	p=.004	p=.000	p=.000	p= ---

⁵ Оценка корреляционной матрицы выполнена с помощью ПК STATISTICA.

В целом предположения подтверждаются. Показатели $X1_1$, $X1_4$ и $X1_5$ тесно связаны между собой, особенно $X1_1$ и $X1_4$, что объясняется функциональной связью между ними. Достаточно тесно связаны $X1_4$ и $X1_5$, что также не противоречит теоретическим выводам.

В группе 3-х показателей ликвидности сложилась примечательная ситуация. $X1_1$ тесно коррелирует с $X1_2$, но $X1_3$ имеет с остальными 2-мя показателями значимую, но достаточно слабую связь, несмотря на то, что денежные средства и краткосрочные финансовые вложения являются элементами текущих активов, а значит, учитываются в $X1_1$ и $X1_2$. По-видимому, это объясняется отрицательной связью денежных средств и дебиторской задолженности.

Указанные показатели ($X1_1$ – $X1_5$) образуют группу, в которой имеет место частичное дублирование информации. Наличие такой группы подтверждается также теоретическими выводами и результатами классификации переменных. Поэтому число показателей может быть уменьшено без особого вреда для конечных выводов. Наиболее подходящий кандидат на исключение – $X1_1$, поскольку он функционально связан с $X1_4$, коррелирует с $X1_2$ и $X1_5$, кроме того, учитывает в своем составе запасы, роль которых в отрасли связи невелика. Следующий претендент – $X1_5$, так как коррелирует с $X2_1$ ($X1_5$ и $X2_1$ отрицательно связаны с долгосрочным заемным капиталом), $X1_2$ и $X1_4$.

Другая группа взаимосвязанных показателей – коэффициенты рентабельности: $X4_1$, $X4_2$ и $X4_3$.

Корреляционный анализ подтверждает наличие положительной связи оборачиваемости активов ($X3_1$) с рентабельностью активов ($X4_2$) и собственного капитала ($X4_3$), а также отрицательной связи коэффициента автономии ($X2_1$) и рентабельности собственного капитала ($X4_3$), что не удивительно, достаточно вспомнить формулу Дюпона.

Из 3-х показателей рентабельности для дальнейшего анализа можно оставить 1-2. Наилучший выбор – рентабельность активов ($X4_2$) и рентабельность продаж ($X4_1$).

Была подтверждена гипотеза о положительной связи оборачиваемости активов ($X3_1$) с коэффициентом текущей ликвидности ($X1_1$).

Предположение об отрицательной связи оборачиваемости активов ($X3_1$) с коэффициентом автономии ($X2_1$) при 5%-ном уровне значимости не подтверждается (см. табл. 5).

Итак, корреляционный анализ позволил эмпирически подтвердить теоретически выявленные взаимосвязи показателей и уменьшить дублирование информации за счет исключения тесно коррелированных показателей.

Наглядно наличие групп финансовых коэффициентов можно представить с помощью диаграммы дерева классификации (см. рис. 1). Графическое представление очень удобно для изучения взаимосвязей и выделения групп показателей.

Выводы корреляционного анализа были также подкреплены классификацией переменных. Для этого использовалась иерархическая агломеративная процедура кластерного анализа. В качестве меры удаленности каждой пары показателей x_i и x_j использовалась $1-\text{cor}(x_i, x_j)$ ⁶.

⁶ $\text{cor}(x_i, x_j)$ – коэффициент парной линейной корреляции, он же коэффициент корреляции Пирсона.

На диаграмме наглядно представлены группы взаимосвязанных показателей. Как и следовало ожидать, единичные группы образовали показатели рентабельности, показатели ликвидности. Особняком стоит коэффициент автономии ($X2_1$), коэффициент оборачиваемости активов ($X3_1$) ближе всего к показателям рентабельности, однако, связь слишком слаба. Неожиданно «далеко» от группы показателей ликвидности и платежеспособности отстоит коэффициент абсолютной ликвидности ($X1_3$). Очень тесная связь коэффициентов текущей и промежуточной ликвидности ($X1_1$ и $X1_2$), по-видимому, объясняется спецификой изучаемых предприятий, незначительностью запасов предприятий связи.

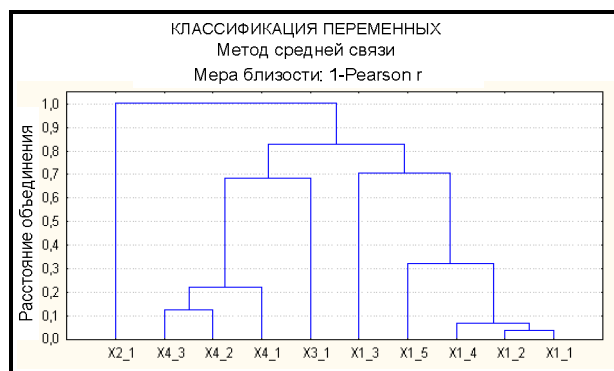


Рис. 1. Группы взаимосвязанных показателей

Для выборки из 38 объектов при 5% уровне значимости критическое значение коэффициента корреляции, согласно статистическим таблицам (распределение Стьюдента), составляет 0.32. Таким образом, критическое расстояние объединения (см. рис. 1) можно установить на уровне примерно 0.7, в этом случае можно выделить 3 группы показателей: рентабельности и оборачиваемости активов ($X4_3$, $X4_2$, $X4_1$, $X3_1$), ликвидности и платежеспособности ($X1_1$, $X1_2$, $X1_3$, $X1_4$, $X1_5$), финансовой устойчивости ($X2_1$).

3. СТАТИСТИЧЕСКАЯ МОДЕЛЬ КРЕДИТНОГО РИСКА

3.1. Классификация объектов

Исходные данные говорят в пользу неоднородности изучаемой выборки. «Связьинвест» и КГТС по своим показателям значительно выделяются из общего ряда, поэтому в дальнейших статистических расчетах они не участвуют.

В ходе исследования было опробовано несколько вариантов классификации по методу k -средних с различным числом кластеров (2-5). Вариант с 2-мя кластерами, несмотря на существенность межклассовых различий и удобство анализа (всего 2 класса), недостаточно чувствителен к особенностям предприятий. В один класс попадают достаточно разнородные объекты с разным уровнем кредитного риска.

Разбиение на 3 класса, по мнению автора, оправдано, содержательными соображениями: во-первых наличием промежуточной группы по уровню кредитного риска, выявленной на стадии анализа показателей выборки, во-вторых, статистическими характеристиками.

Во всех 4-х случаях в качестве входных параметров использовались 6 показателей: $X1_2$, $X1_3$, $X1_4$, $X2_1$, $X3_1$, $X4_2$. Далеко не все из них существенно

влияли на результаты классификации. Дисперсионный анализ показал значимые межклассовые различия для $X1_2$, $X1_4$ и $X3_1$ при 5%-ном уровне значимости (проверка значимости по F-тесту) (см. табл. 6).

Таблица 6

АНАЛИЗ ДИСПЕРСИИ (КЛАССИФИКАЦИЯ ПО ИСХОДНЫМ ДАННЫМ, 3 КЛАССА)

Показатель	Межклассовая дисперсия (Between SS)	Число степеней свободы числителя df	Внутриклассовая дисперсия (Within SS)	Число степеней свободы знаменателя df	F-статистика	Уровень значимости (p-значение)
$X1_2$	1.5841	2	0.7351	35	37.7123	0.0000
$X1_3$	0.0329	2	0.2644	35	2.1746	0.1288
$X1_4$	2.9862	2	1.1042	35	47.3282	0.0000
$X2_1$	0.0159	2	0.4120	35	0.6761	0.5151
$X3_1$	0.7535	2	1.2138	35	10.8636	0.0002
$X4_2$	0.0060	2	0.0929	35	1.1392	0.3316

Выборка разбивается на 3 группы с различным уровнем кредитного риска:

- 1-й кластер – «высокий уровень риска»,
- 2-й – «средний уровень риска»,
- 3-й – «низкий уровень риска».

Средние значения показателей по каждой группе и состав кластеров приведены в табл. 7 и табл. 24 соответственно.

Таблица 7

СРЕДНИЕ ЗНАЧЕНИЯ ПОКАЗАТЕЛЕЙ (3 КЛАССА, МЕТОД К-СРЕДНИХ)

Показатель	Cluster No 1	Cluster No 2	Cluster No 3
$X1_2$	0.4448	0.8555	1.0893
$X1_3$	0.0787	0.1119	0.1757
$X1_4$	-0.5412	0.0924	0.2722
$X2_1$	0.6781	0.7266	0.6954
$X3_1$	0.6769	0.7385	1.1034
$X4_2$	0.0795	0.1029	0.1204
Число объектов	8	24	6

1-я группа – «высокий уровень кредитного риска» – некредитоспособные предприятия ($X1_4 < 0$) с оборачиваемостью активов ниже среднего и низкой ликвидностью (всего 8 компаний). Несмотря на это, некоторые из них имеют высокую рентабельность продаж (ENCO) или абсолютную ликвидность (ENCO, MGTS). Например, коэффициент абсолютной ликвидности MGTS (0.23) соответствует нормативам (0.2-0.5).

2-я группа – кластер 2 занимает промежуточное положение – «средний уровень кредитного риска». С одной стороны, показатели $X4_2$, $X2_1$ сближают его с 3-й группой (см. табл. 7), с другой – по $X1_3$, $X3_1$ он близок к группе 1. Итак, 2-я группа характеризуется достаточно высокой рентабельностью и оборачиваемостью, высокой долей собственного капитала (выше, чем в других группах), соответствующим нормативам коэффициентом промежуточной ликвидности (норматив часто устанавливается в интервале 0.7-1), однако низкими показателями

абсолютной ликвидности и обеспеченности чистым оборотным капиталом (близок к 0). Это наиболее многочисленная группа – 24 компании. В нее вошли как финансово-устойчивые ($X1_5 > 0$), но недостаточно рентабельные, предприятия такие как ESVL, NVGT, ALCO, так и некредитоспособные, но перспективные в плане улучшения состояния: NNSI, SVZK, ESOB, SMSI, SVLN.

3-я группа – «низкий уровень кредитного риска» – наиболее кредитоспособные компании (6 фирм): ELRO, ESAL, ESIR, HTCM, SVIM, ELCH.

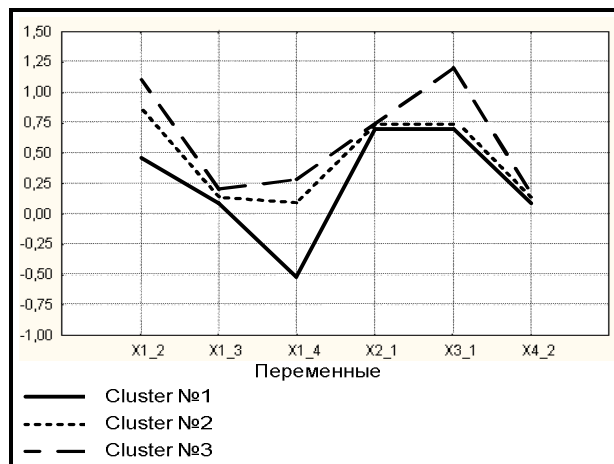


Рис. 2. Средние значение по классам

3.2. Формирование решающего правила модели

Для выработки решающего правила предлагается использовать дискриминантный анализ. Исходные данные и классовая принадлежность каждого объекта (группирующая переменная) рассматриваются как обучающая выборка, к которой применяется стандартная процедура дискриминантного анализа. Классификационные функции, получившиеся в результате, могут использоваться в качестве решающих правил.

В случае 3-х кластеров процедура предлагает 3 функции. Классовая принадлежность изучаемого объекта определяется максимальным значением функции:

$$S_i = \max_k S_{ki}.$$

Кроме того, дискриминантный анализ позволяет проверить корректность результатов кластерного анализа (выявляет некорректно классифицированные объекты).

Данная процедура была применена к результатам классификации по 3-м классам для не стандартизованных данных. В результате были выявлены 3 некорректно отнесенные компании: ENCO, ELRO и ESOM. Они были исключены из выборки при расчете решающих правил.

Были получены классификационные функции:

$$S_1 = -140.988 + 162.166 * X1_2 + 48.838 * X1_3 - 149.455 * X1_4 + 124.691 * X2_1 + 40.788 * X3_1 + 51.106 * X4_2;$$

$$S_2 = -128.345 + 163.138 * X1_2 + 35.780 * X1_3 - 120.592 * X1_4 + 115.755 * X2_1 + 42.225 * X3_1 + 79.829 * X4_2;$$

$$S_3 = -168.494 + 186.301 * X1_2 + 42.568 * X1_3 - 129.346 * X1_4 + 117.166 * X2_1 + 60.834 * X3_1 + 61.411 * X4_2.$$

Скорректированная при помощи дискриминантного анализа классификация приведена в табл. 24 (модель 1 после корректировки).

После корректировки ENCO, ESOM перемещаются из 1-го, ELRO из 3-го во 2-й класс, то есть обладают средним риском. ENCO и ESOM более привлекательны, чем остальные предприятия 1-го класса, несмотря на отрицательный чистый оборотный капитал, имеют относительно высокую абсолютную ликвидность и рентабельность (см. табл. 8).

Таблица 8

СРЕДНИЕ ЗНАЧЕНИЯ ПЕРЕМЕННЫХ ПО КЛАССАМ ПОСЛЕ КОРРЕКТИРОВКИ

Показатель	Классы		
	№1	№2	№3
X1_2	0.4032	0.8423	1.0929
X1_3	0.0852	0.1085	0.1859
X1_4	-0.6275	0.0740	0.2578
X2_1	0.6824	0.7256	0.6700
X3_1	0.6526	0.7452	1.1444
X4_2	0.0668	0.1065	0.1107
Число объектов	6	27	5

Качество классификации несколько выше среднего, о чем говорит показатель «лямбда Уилкса» (Wilks' Lambda) = 0.092 (чем она ближе к 0, тем лучше классификация). С одной стороны, вклад в объяснение классификации некоторых коэффициентов незначим, о чем свидетельствует F-статистика (см. табл. 6). С другой стороны, межклассовые различия значимы при 5%-ном уровне (см. p -значение табл. 9).

Таблица 9

АНАЛИЗ КАЧЕСТВА КЛАССИФИКАЦИИ

Квадрат межклассового расстояния			
	G_1:1	G_2:2	G_3:3
G_1:1	0	23.664	48.368
G_2:2	23.664	0	12.998
G_3:3	48.368	12.998	0
F-статистика; df = 6,27			
	G_1:1	G_2:2	G_3:3
G_1:1	-	13.668	15.115
G_2:2	13.668	-	6.228
G_3:3	15.115	6.228	-
p-значение			
	G_1:1	G_2:2	G_3:3
G_1:1	-	4.4E-07	1.6E-07
G_2:2	4.4E-07	-	0.0003
G_3:3	1.6E-07	0.0003	-

К недостаткам модели можно отнести не высокую устойчивость коэффициентов классификационных функций во времени.

4. ЛИНГВИСТИЧЕСКАЯ МОДЕЛЬ КРЕДИТНОГО РИСКА

В данном исследовании использовались только количественные показатели, поэтому получение итоговых оценок не представляет принципиальных трудностей. Однако такие оценки сами по себе без сравнения с другими объектами, нормативами, номинальными шкалами мало информативны. Как правило, ЛПР (лицо, принимающее решение) оперирует не только количественной, но и лингвистической шкалой. Для принятия решения ЛПР недостаточно знать количественное выражение показателей. Ему необходимы знания о том, велики или малы эти значения. Однако четко определить границы интервалов категорий для каждого показателя не всегда возможно. К тому же возникают вопросы: почему границы жесткие? Что будет, если показатель окажется на малую величину ниже? Насколько малой должна быть эта величина, чтобы не оказать влияния на выводы ЛПР? Можно привести еще много разных доводов против использования жестких границ для номинальных градаций показателя.

Использование лингвистических переменных позволяет снять остроту проблемы. Данный аппарат дает возможность работать в условиях неопределенности, когда нельзя жестко определить границы номинальных категорий, нет уверенности в точности исходных данных.

Данная методика не требует обучающей выборки, результаты анализа и аналитические выражения не зависят от числа наблюдений, не меняются при появлении новых наблюдений, что выгодно отличает ее от процедур кластерного анализа, позволяет проводить оценку риска на единой методической основе, обеспечивает сопоставимость данных.

Однако субъективный выбор весовых коэффициентов, определение лингвистической переменной делают ее чувствительной к качеству экспертной информации.

Лингвистическая модель кредитного риска была предложена А.О. Недосекиным, с некоторыми доработками она представлена в настоящем разделе исследования. Доработки заключались:

- во-первых, в корректировке интегрального показателя, в результате которой, область значений интегрального показателя стала соответствовать носителю нечеткого классификатора (подробнее представлено ниже);
- во-вторых, при моделировании были опробованы 7-уровневый классификатор наряду с 5-уровневым и сплайн функции вместо трапециевидных функций. Результаты моделирования представлены в данном разделе.

Далее описывается принципиально новый подход к моделированию с использованием лингвистических моделей – лингвистическая модель с обучением.

Таблица 10

ЛИНГВИСТИЧЕСКИЕ ПЕРЕМЕННЫЕ «УРОВЕНЬ ФАКТОРА К»

	Очень низкий	Низкий	Средний	Высокий	Очень высокий
X1_2	(0, 0, 0.5, 0.6)	(0.5, 0.6, 0.7, 0.8)	(0.7, 0.8, 0.9, 1)	(0.9, 1, 1.3, 1.5)	(1.3, 1.5, ∞, ∞)
X1_3	(0, 0, 0.02, 0.03)	(0.02, 0.03, 0.08, 0.1)	(0.08, 0.1, 0.3, 0.35)	(0.3, 0.35, 0.5, 0.6)	(0.5, 0.6, ∞, ∞)
X1_4	(-∞, -∞, -0.005, 0)	(-0.005, 0, 0.09, 0.11)	(0.09, 0.11, 0.3, 0.35)	(0.3, 0.35, 0.45, 0.5)	(0.45, 0.5, 1, 1)
X2_1	(0, 0, 0.1, 0.2)	(0.1, 0.2, 0.25, 0.3)	(0.25, 0.3, 0.45, 0.5)	(0.45, 0.5, 0.6, 0.7)	(0.6, 0.7, 1, 1)
X3_1	(0, 0, 0.02, 0.03)	(0.02, 0.03, 0.08, 0.1)	(0.08, 0.1, 0.3, 0.35)	(0.3, 0.35, 0.5, 0.6)	(0.5, 0.6, ∞, ∞)
X4_2	(-∞, -∞, 0, 0)	(0, 0, 0.006, 0.01)	(0.006, 0.01, 0.06, 0.1)	(0.06, 0.1, 0.225, 0.4)	(0.225, 0.4, ∞, ∞)

Источник: Недосекин А.О. Финансовый менеджмент на нечетких множествах. // Аудит и финансовый анализ – М., 2003, №3.

4.1. Описание метода

Для задания лингвистической переменной $\langle \beta, T, X, G, M \rangle$ требуется определить ее основные параметры:

β – наименование;

T – терм-множество (множество номинальных градаций переменной);

X – область определения или универсум;

G – синтаксическая процедура образования новых значений терм-множества;

M – семантическая процедура задания для каждого значения терм-множества нечеткого множества⁷.

Каждое нечеткое множество задается парой (μ_x/X) ,

где

x – любое значение переменной из области определения X ;

μ_x – значение функции принадлежности $(\mu_x \in [0, 1])$.

Задание терм-множеств и соответствующих им нечетких множеств требует привлечения экспертов или анализа достаточно представительной статистики. В обоих случаях число термов может задавать ЛПР. Ключевой момент в этой процедуре – определение носителя и ядра каждого нечеткого множества⁸.

Между интервалами определенности, как правило, задаются интервалы неопределенности, на которых ЛПР не уверено в лингвистической оценке изучаемого показателя. От соотношения размера «ядер» и интервалов неопределенности будет зависеть степень «размытости» нечетких множеств.

В качестве экспертной информации использовались лингвистические переменные «Уровень показателя k » приведенные в работе Недосекина (2003) (см. табл. 10).

В таблице функции принадлежности заданы четвёркой чисел: 1-е и 4-е – границы носителя, 2-ое и 3-е – ядро.

В упомянутой работе предлагаются трапецевидные функции принадлежности вида:

$$\mu_j^{(k)}(X) = \begin{cases} \frac{x-a}{b-a}, & a \leq x \leq b \\ 1, & b \leq x \leq c \\ \frac{d-x}{d-c}, & c \leq x \leq d \\ 0, & \text{иначе,} \end{cases}$$

где

$[a, d]$ – носитель;

$[b, c]$ – ядро.

В рамках данного исследования для анализа кредитного риска были также использованы П-образные сплайн-функции:

$$\mu_j^{(k)}(X) = \begin{cases} 0.5 - 0.5 \cos(\pi \frac{x-a}{b-a}), & a \leq x \leq b \\ 1, & b \leq x \leq c \\ 0.5 - 0.5 \cos(\pi \frac{d-x}{d-c}), & c \leq x \leq d \\ 0, & \text{иначе.} \end{cases}$$

⁷ Обработка нечеткой информации в системах принятия решений. / А.Н. Борисов, А.В. Алексеев, Г.В. Меркурьева и др. – М.: Радио и связь, 1989. – С. 15.

⁸ Носитель нечеткого множества – это такие $x \in X \mid \mu_x(x) > 0$. Ядро нечеткого множества – это такие $x \in X \mid \mu_x(x) = 1$. То есть содержательно ядро задает множество значений изучаемого показателя, в отношении которого эксперт или ЛПР уверены в их точной принадлежности к некоторой данной градации (терму).

Их основное отличие от трапецевидных функций заключается в характере «уверенности» ЛПР. Степень уверенности меняется не линейно, а зависит от положения на интервале неопределенности. Чем ближе к ядру, тем быстрее растет уверенность, чем дальше, тем быстрее она падает. В целом данный тип функций характеризует «более уверенного ЛПР».

Алгоритм оценки включает в себя несколько этапов.

На 1-ом этапе метода проводится лингвистическое распознавание уровней показателей: определяются значения функций принадлежности на каждом терм-множестве по каждому показателю.

Затем с помощью формулы получают интегральный показатель:

$$Y_i = \sum_{k=1}^p q_k \sum_{j=1}^l g_j \mu_{ji}^{(k)},$$

где

для каждого i -го объекта $i=1, \dots, N$;

$k=1, \dots, p$;

p – число показателей;

$j=1, \dots, l$, l – число термов лингвистической переменной

«Уровень фактора»;

$g_j = 0.2j - 0.1$.

Промежуточные результаты $V_{ki} = \sum_{j=1}^l g_j \mu_{ji}^{(k)}$ могут интерпретироваться как индексы-оценки уровня k -ого показателя. Позволяют проводить факторный анализ, выявлять причины, повлиявшие на конкретное значение интегрального показателя.

Для распознавания уровня интегрального показателя и его интерпретации задается лингвистическая переменная «Уровень кредитного риска» (далее нечеткий или лингвистический классификатор). Она ставит в соответствие высоким значениям Y_i низкий уровень риска, низким Y_i – высокий уровень риска. Оценка уровня интегрального показателя L_i производится по принципу:

$$L_i = \max_j \mu_j^{(0)}.$$

Как и переменные «Уровень показателя k », «Уровень кредитного риска» может иметь любые функции принадлежности и любое число термов (в данном случае 5 или 7). Основные требования к классификатору: симметричность размещения термов на 0,1-шкале, непротиворечивость ($\sum_{j=1}^l \mu_j = 1$ – общее для всех переменных требование) (см. табл. 11).

Таблица 11

ПЯТИУРОВНЕВЫЙ КЛАССИФИКАТОР

№ класса	Функция принадлежности	Терм	Z_i
1	$\mu_1^{(0)}$	Очень высокий риск	(0, 0, 0.15, 0.25)
2	$\mu_2^{(0)}$	Высокий	(0.15, 0.25, 0.35, 0.45)
3	$\mu_3^{(0)}$	Средний	(0.35, 0.45, 0.55, 0.65)
4	$\mu_4^{(0)}$	Низкий	(0.55, 0.65, 0.75, 0.85)
5	$\mu_5^{(0)}$	Очень низкий	(0.75, 0.85, 1, 1)

Источник: Недосекин А.О. Финансовый менеджмент на нечетких множествах. // Аудит и финансовый анализ – М., 2003, №3.

Методика Недосекина предлагает использовать интегральный показатель Y_i для распознавания классификатором определенным на 0,1-носителе. Показатель Y_i может принимать значения в интервале [0.1, 0.9], а значит, степени принадлежности значений $Y_i < 0.1$ и > 0.9 равны 0. Поэтому 0,1-носитель для классификатора использовать нельзя. Чтобы исправить этот недостаток,

необходима корректировка: $Z_i = (Y_i - 0.1)/0.8^9$. Скорректированный интегральный показатель определен на 0,1-носителе. Данное преобразование, во-первых, уточняет процедуру, во-вторых, несколько улучшает чувствительность классификатора за счет более равномерного распределения термов.

Веса факторов q_k интегрального критерия определяются предпочтениями ЛПР в отношении важности факторов. В идеале, для определения вида критерия необходимо выявить предпочтения ЛПР (кривые безразличия или нормы замещения в каждой точке факторного пространства). На практике это сделать трудно. При отсутствии информации о предпочтениях или эквивалентности факторов можно использовать равные веса частных критериев или следовать подходу Фишберна, когда факторы ранжированы по значимости:

$$q_j = \frac{2(p-s+1)}{(p+1)p},$$

где

s – рейтинг j -го фактора;

p – число показателей.

Если несколько факторов эквивалентны, их рейтинги усредняются, как это принято в статистике.

Для определения весов можно учитывать вклад k -го фактора в классификацию объектов (результаты кластер-анализа), то есть факторы, по которым объекты разных групп наиболее заметно различаются, получают больший вес.

Возможны и другие подходы. Веса могут быть заданы как нечеткие числа, что, впрочем, на взгляд автора, излишне усложнит модель.

4.2. Результаты анализа

Первоначально расчеты были проведены для случая трапециевидных функций, 5-уровневого классификатора и равных весов факторов.

Совокупность наблюдений представлена 4 группами (из 5 возможных) по уровню кредитного риска:

- Высокий уровень (1 компания);
- Средний (18);
- Низкий (20);
- Очень низкий (1).

В категорию «Очень высокий» не попал ни один объект. В этой процедуре участвовали все 40 объектов, поскольку неоднородность наблюдений не влияет на результаты. Результаты приведены в табл. 24 (Модели 2, 2(A)).

Таблица 12

СРЕДНИЕ ЗНАЧЕНИЯ ПОКАЗАТЕЛЕЙ ПО ГРУППАМ РИСКА

Показатель	Уровень кредитного риска		
	высокий	средний	низкий
$X1_2$	0.2982	0.6311	0.9984
$X1_3$	0.0412	0.0900	0.1426
$X1_4$	-1.0981	-0.1918	0.2143
$X2_1$	0.5842	0.7175	0.7231
$X3_1$	0.4953	0.7332	0.8456
$X4_2$	0.0163	0.1039	0.1022
Число предприятий	1	18	19

Рассмотрим табл. 12, где рассчитаны средние значения показателей по группам. Из 3-ей группы удалена INSW для достижения сопоставимости с результатами

⁹ Y_i и Z_i – эквивалентные критерии, т.е. $Y(X_i) \geq Y(X_j) \Leftrightarrow Z(X_i) \geq Z(X_j)$.

статистических процедур, а также не приведена группа, включающая KGTS.

В состав нижней и верхней групп риска попали наиболее «экстремальные» компании. В группе высокого риска оказалось 1 предприятие, ранее отнесенное к категории некредитоспособных ($X1_4 < 0$): URSI. С помощью факторных индексов V_{ji} возможен факторный лингвистический анализ объектов. Так, объекты группы среднего риска характеризуются низкими и очень низкими значениями коэффициента срочной ликвидности ($X1_2$), очень низкой обеспеченностью чистым оборотным капиталом ($X1_4$), низким и средним уровнем абсолютной ликвидности ($X1_3$), высокой и средней рентабельностью активов ($X4_2$) и при этом высокой и очень высокой финансовой автономией ($X2_1$) и оборачиваемостью активов ($X3_1$).

В группе низкого риска оказались и финансово-устойчивые компании ($X1_5 > 0$), и «промежуточные» компании. Они имеют среднюю и высокую рентабельность, высокую и очень высокую оборачиваемость активов, высокую срочную и среднюю абсолютную ликвидность, высокую и очень высокую долю собственных средств. Группа очень неоднородна по уровню обеспеченности собственным оборотным капиталом: от низкого до очень высокого. В этой группе особо выделяется INSW, для него характерна очень высокая ликвидность, обеспеченность чистым оборотным капиталом и финансовая автономия, но при этом очень низкая оборачиваемость и низкая и средняя рентабельность. KGTS выделена в отдельную группу.

Использование сплайн-функций не меняет ситуацию, дает классификацию, тождественную описанной выше. Единственное различие – значения функций принадлежности и интегральных показателей (различия в сотых и тысячных).

Чтобы иметь более детальную шкалу уровней кредитного риска компаний, предлагается использовать 7-уровневую шкалу классификатора (см. табл. 13).

Таблица 13

СЕМИУРОВНЕВЫЙ КЛАССИФИКАТОР

№ класса		Терм	Z_i
1	$\mu_1^{(0)}$	Очень высокий риск	(0, 0, 0.075, 0.175)
2	$\mu_2^{(0)}$	Высокий	(0.075, 0.175, 0.225, 0.325)
3	$\mu_3^{(0)}$	Выше среднего	(0.225, 0.325, 0.375, 0.475)
4	$\mu_4^{(0)}$	Средний	(0.375, 0.475, 0.525, 0.625)
5	$\mu_5^{(0)}$	Ниже среднего	(0.525, 0.625, 0.675, 0.775)
6	$\mu_6^{(0)}$	Низкий	(0.675, 0.775, 0.825, 0.925)
7	$\mu_7^{(0)}$	Очень низкий	(0.825, 0.925, 1, 1)

В банковской практике для оценки кредитного риска часто используется большое число градаций (9-10)¹⁰, поэтому использование 7 категорий в классификаторе вполне обосновано с практической точки зрения.

Классификатор позволяет задавать любые функции принадлежности.

Указанный классификатор был применен к скорректированным интегральным показателям Z_i , рассчитанным ранее с использованием трапециевидных и сплайн-функций.

На этот раз вид функций принадлежности повлиял на результаты классификации.

¹⁰ См., например, А.М. Santomero, Commercial Bank Risk Management: An Analysis of the Process. // Journal of Financial Services Research, 12:2/3, 1997

В 1-й и 2-й классы не попало ни одного наблюдения. Все объекты разместились в 5 классах с уровнем риска выше среднего (1), средним (14 или 12 для сплайн-функций), ниже среднего (18 или 20), низким (6) и очень низким (1). Состав классов, промежуточных индексов и значения скорректированного интегрального показателя приведены в табл. 24 (Модели 4, 4(А)).

Выбор весов может быть обусловлен результатами статистического анализа. Фактор, наиболее значимо повлиявший на классификацию объектов, по-видимому, обладает большей предсказательной способностью и значимостью для конечных выводов. Поэтому должен иметь больший вес в интегральном показателе.

К сожалению, строго доказать этот факт нельзя, он носит чисто эвристический характер. С таким же успехом (и, главное, также вполне обосновано) можно задавать больший вес факторам с меньшей дисперсией.

Итак, воспользуемся результатами кластерного анализа для определения системы предпочтений: $X1_2 \sim X1_4 > X1_3 > X3_1 \sim X4_2 > X2_1$ (знак « \sim » – эквивалентность, означает, что показатели имеют равную значимость для определения кредитного риска, знак « $>$ » – предпочтение – фактор слева от знака важнее для определения уровня риска, чем стоящий справа). Веса будут рассчитаны по формуле Фишберна (см. табл. 14).

Таблица 14

ФАКТОРНЫЕ ВЕСА ФИШБЕРНА В СООТВЕТСТВИИ С ЗАДАННОЙ СИСТЕМОЙ ПРЕДПОЧТЕНИЙ

Фактор	Рейтинг (s)	Вес (q)
X1_2	1.5	11/42
X1_4	1.5	11/42
X1_3	3	8/42
X3_1	4.5	5/42
X4_2	4.5	5/42
X2_1	6	2/42

При такой системе предпочтений совокупность объектов образует 4 группы по уровню риска:

- высокий (11 компаний);
- средний (17 или 16 для сплайн функций);
- низкий (11 или 12);
- очень низкий (1),

в первую градацию (чень высокий уровень риска не попало ни одного наблюдения).

Существенное изменение результатов по сравнению с предыдущей моделью связано с признанием большей значимости показателей ликвидности, по которым изучаемая выборка имеет достаточно низкие показатели. Одновременно снизилась значимость «сильной стороны» данных предприятий – показателей оборачиваемости и рентабельности активов.

Модель достаточно гибка, позволяет специалисту-аналитику влиять на процесс принятия решения, во-первых, за счет задания лингвистических переменных, во-вторых, за счет выбора весов факторов. Указанные свойства дают возможность приспособлять модель к условиям конкретной отрасли и предпочтениям ЛПР.

Более того, по-видимому, модель способна к обучению. Имея обучающую выборку, в которой представлены номинальные оценки кредитного риска предприятий, можно выбрать веса факторов таким образом, чтобы модель могла объяснять данные выборки, то есть минимизировать некоторый функционал качества, например, сумму квадратов ошибок.

5. ЛИНГВИСТИЧЕСКАЯ МОДЕЛЬ С ОБУЧЕНИЕМ

Пусть каждое предприятие характеризуется не только набором показателей, но и номером класса кредитного риска, к которому оно ранее было отнесено S_i . В качестве оценки класса кредитного риска использовались результаты классификации по модели 1.

Задача заключается в разработке модели, позволяющей учесть априорную информацию о распределении объектов по классам риска, т.е. построении модели с обучением.

5.1. Постановка задачи

В исходной модели лингвистической классификации без обучения решалась следующая задача: по значению интегрального показателя при помощи нечеткого классификатора определялся уровень кредитного риска исследуемого предприятия. В модели с обучением, напротив, требуется по классу риска определить значение интегрального показателя. Однозначно (точно) этого сделать нельзя, но можно дать интервальную оценку интегральному показателю для каждого уровня кредитного риска.

Постановка задачи предусматривает поиск таких весов факторов, которые бы обеспечивали минимум ошибки отнесения к классу кредитного риска. Поскольку интервал значений интегрального показателя Z_i соответствует определенному классу S_i кредитного риска, любое значение из этого интервала будет характеризовать кредитный риск предприятия на уровне S_i . Поэтому задачу можно считать решенной, если удается найти такие веса факторов, что они обеспечивают значения интегрального показателя, попадающие в «нужные» интервалы и при этом показывают наименьшее число некорректно классифицированных предприятий по выборке. Оптимальность решения при этом не гарантирована.

Соответствие интегрального показателя уровню кредитного риска (номеру класса) ставится нечетким классификатором. В данной модели использован нечеткий трехуровневый классификатор (поскольку число классов S_i равно 3). Каждому классу соответствует терм нечеткого классификатора (см. табл. 15, рис. 3).

С помощью нечеткого классификатора осуществляется переход от класса кредитного риска к оценке интегрального показателя (задается интервалом, вообще говоря, нечетким).

Математически задача отыскания факторных весов может быть сформулирована следующим образом: если уровень (класс) риска предприятия $S_i = S_i^*$, то выполняются условия:

$$Z_i' \leq Z_i \leq Z_i'', i = 1, \dots, n; \tag{1}$$

$$\sum_{k=1}^p q_k = 1, q_k \in [0, 1], \tag{2}$$

где Z_i' и Z_i'' нижняя и верхняя границы интервала значений интегрального показателя, соответствующих классу S_i^* , расчетное значение интегрального показателя:

$$Z_i = (Y_i - 0.1) / 0.8, Z_i \in [0, 1]; \tag{3}$$

$$Y_i = \sum_{k=1}^p q_k V_{ki}, Y_i \in [0.1, 0.9]; \tag{4}$$

$$V_{ki} = \sum_{j=1}^l g_j \mu_{ji}^{(k)}, \tag{5}$$

где V_{ki} – частный критерий для k -го фактора (исходные данные).

То есть необходимо решить систему, состоящую из $2 \cdot n$ неравенств вида (1), p условий неотрицательности факторных весов и уравнения (2).

Таблица 15

ТРЕХУРОВНЕВЫЙ КЛАССИФИКАТОР

класс	Фун. прин.	Терм	Z_i				Условная середина интервала
1	$\mu_1^{(0)}$	Высокий риск	0	0	0.25	0.4	0.15
2	$\mu_2^{(0)}$	Средний	0.25	0.4	0.6	0.75	0.5
3	$\mu_3^{(0)}$	Низкий	0.6	0.75	1	1	0.85

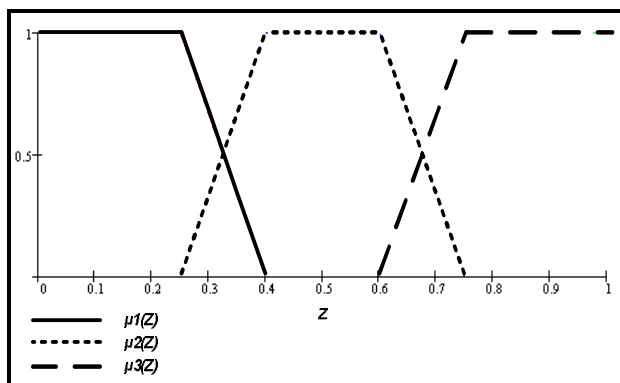


Рис. 3. Трехуровневый классификатор

При решении сформулированной задачи возможно три случая:

- существует единственное решение;
- существует множество решений;
- решений нет (система неравенств несовместна).

Первый случай дает решение изначально поставленной задачи, но он маловероятен.

Во втором случае определяется область решений, но нет никаких критериев для выбора какого-либо набора весов из множества. Все наборы множества обеспечивают корректную классификацию всех объектов выборки.

В третьем случае не существует такого набора весов, который бы удовлетворял всем неравенствам одновременно, значит, неизбежно в выборке появляются некорректно классифицированные объекты, и тогда, выбирая факторные веса, требуется свести к минимуму число некорректно классифицированных объектов.

В последних двух случаях целесообразно использовать некоторый критерий для выбора набора факторных весов.

Таким критерием в точечном варианте обычно служит сумма квадратов ошибок, минимум которой дает искомое решение.

В качестве подхода к отысканию приемлемого набора факторных весов можно предложить перейти от интервальных оценок к точечным и решать для точечных оценок задачу минимизации суммы квадратов ошибок по выборке.

В качестве точечной оценки Z_i были приняты три варианта: условная середина интервала-ядра, нижняя граница интервала-ядра (вторая колонка Z_i , см. 3-уровневый классификатор), верхняя граница интервала-ядра (третья колонка Z_i).

Для каждого варианта решалась задача минимизации суммы квадратов ошибок по обучающей выборке.

Математически это означает, что требуется найти вектор весов (q_1, \dots, q_p) такой, что оптимизирует показа-

тель качества классификации Q . Решается задача математического программирования (6), (7).

$$Q = \sum_{i=1}^n (\hat{Z}(q_1, \dots, q_p)_i - \tilde{Z}_i)^2 \rightarrow \min; \quad (6)$$

$$\sum_{k=1}^p q_k = 1, \quad q_k \in [0, 1]. \quad (7)$$

Методика расчета интегрального показателя соответствует модели лингвистической классификации.

Таким образом, входами модели являются следующие характеристики каждого i -го предприятия O_i : вектор частных критериев V_{ki} , рассчитанный в соответствии с ранее описанной моделью лингвистической классификации, номер класса риска S_i .

Номер класса S_i , к которому относится объект O_i , измеряется в порядковой шкале, интегральная оценка Z_i в количественной шкале, определена на отрезке от 0 до 1, переход от уровня риска к оценке интегрального показателя осуществляется с помощью нечеткого трехуровневого классификатора.

5.2. Методика оценки

Исходная выборка предприятий была разбита на две выборки: обучающую и тестовую. Обучающая выборка служит для оценивания факторных весов, тестовая – для проверки точности классификационной Модели.

Методика формирования выборок представлена ниже.

Всего для разработки Модели использовались данные 38 предприятий, из них 19 предприятий – обучающая выборка, оставшиеся 19 – тестовая выборка. Предприятия отбирались случайным образом с условием, что структура обучающей выборки будет мало отличаться от структуры всей выборки. Данные о распределении выборок по количеству предприятий приведены в табл. 16.

Таблица 16

МЕТОДИКА ФОРМИРОВАНИЯ ОБУЧАЮЩЕЙ И ТЕСТОВОЙ ВЫБОРОК ПРЕДПРИЯТИЙ

Показатели	Класс			Всего
	№1	№2	№3	
Число предприятий (вся выборка)	6	27	5	38
Доля	100.0%	100.0%	100.0%	100.0%
Обучающая выборка	3	13	3	19
Доля	50.0%	48.1%	60.0%	50.0%
Тестовая выборка	3	14	2	19
Доля	50.0%	51.9%	40.0%	50.0%

5.3. Результаты расчетов

Таблица 17

РЕЗУЛЬТАТЫ ОЦЕНКИ ВЕСОВ¹¹

Условие	№	X1 2	X1 3	X1 4	X2 1	X3 1	X4 2	Итог
По нижней границе интервала (ядра) ¹²	p2	0.576	0.292	0.000	0.000	0.000	0.133	1
По условной середине	p1	0.536	0.034	0.000	0.000	0.000	0.430	1
По верхней границе интервала (ядра)	p3	0.500	0.000	0.000	0.000	0.314	0.186	1
Имитационное моделирование	p1 (корр.)	0.548	0.148	0.000	0.000	0.062	0.241	1

¹¹ Оценки весов были получены с помощью ПК MS Excel.

¹² Ядро – интервал носителя (области определения) X , для которого $\mu(x) = 1$.

Задача с целевой функцией (6) и условиями (7) была решена для трех вариантов выбора значения интегрального показателя: условная середина интервала, нижняя граница и верхняя граница интервала.

Значения весов факторов приведены в табл. 17.

Сопоставление исходной классификации с результатами различных вариантов оценивания Модели приведены в таблицах ниже.

Как видно из приведенных таблиц, качество классификации существенно зависит от выбора точечной оценки интегрального показателя по изначально заданному классу риска.

Для уточнения результатов моделирования предлагается использовать метод статистических испытаний (МСИ, метод Монте-Карло).

Таблица 18

РЕЗУЛЬТАТЫ АНАЛИЗА ТОЧНОСТИ КЛАССИФИКАЦИИ ПО ОБУЧАЮЩЕЙ ВЫБОРКЕ

Условие	№	Некорректно классифицированные предприятия		Индекс, %
		Число	Доля, %	
По нижней границе интервала (ядра)	<i>p2</i>	6	31.58	68.42
По условной середине интервала	<i>p1</i>	5	26.32	73.68
По верхней границе интервала (ядра)	<i>p3</i>	9	47.37	52.63
Имитационное моделирование	<i>p1 (корр.)</i>	3	15.79	84.21

Пусть веса факторов изменяются в пределах заданных нижней и верхней границами интервалов (расчеты № *p2*, № *p3*). В каждом имитационном эксперименте с помощью датчика случайных чисел генерируется случайный вектор (q_1, \dots, q_p) факторных весов, используя который рассчитываются интегральные показатели каждого предприятия обучающей выборки и оценивается доля некорректно классифицированных объектов. Выполняется *N* имитаций (в данном исследовании выполнено 100 имитационных расчетов). Из насчитанного массива данных выбирается вектор весов минимизирующий число некорректно классифицированных объектов. Если он не единственный, предлагается оценить усредненный по отобранной группе вектор.

Полученный вектор факторных весов использовался для расчета интегрального показателя каждого предприятия обучающей выборки и оценки доли некорректно классифицированных объектов.

Как показали расчеты, данная процедура позволила существенно улучшить результаты классификации объектов. Число некорректно классифицированных предприятий существенно сократилось (см. табл. 18).

Основные результаты сравнения полученной классификации с исходной, а также некоторые статистические характеристики качества классификации представлены в виде таблицы сопряженности (см. табл. 19).

Под таблицей рассчитаны стат. характеристики:

- коэффициент хи-квадрат – расчетное значение критической статистики, используемой для проверки гипотезы об отсутствии связи между изучаемыми категориями (распределениями по классам риска). В случае если расчетное значение выше табличного (теоретического) значения статистического χ^2 -распределения (хи-квадрат распределения), гипотеза об отсутствии связи не принимается и

вместо нее есть основания принять альтернативную гипотезу о наличии значимой статистической связи. Для проверки этого условия можно пользоваться как теоретическим значением, так и *P*-значением;

- *P*-value – вероятность, так называемой, ошибки первого рода, то есть вероятность того, что принята неверная гипотеза о наличии связи, в то время как верна гипотеза об отсутствии связи;
- ХиПирсон – коэффициент сопряженности Пирсона, измеряет степень взаимосвязи категоризованных данных (распределение по классам риска), изменяется от 0 до 1, чем выше, тем сильнее связь, тем сильнее связаны изучаемые распределения по классам;
- ХиЧупров – коэффициент сопряженности Чупрова, также как коэффициент Пирсона измеряет тесноту статистической связи категоризованных данных;
- Индекс – индекс соответствия, показывает долю корректно классифицированных объектов (предприятий).

Таблица 19

ПО ОБУЧАЮЩЕЙ ВЫБОРКЕ

		Class Adj (Модель 1)			Всего
		№1	№2	№3	
Оценка (Модель 8)	№1	3	0	0	3
	№2	0	11	1	12
	№3	0	2	2	4
	Всего	3	13	3	19
	Хи ²	Индекс	ХиПирсон	ХиЧупрова	P-value
	23.060	84.211	0.740	0.779	0.000

Для анализа уровня соответствия результатов моделей и оценки точности классификации предприятий введен «индекс соответствия» (индекс), который равен отношению суммы численностей объектов в предположительно соответствующих друг другу по уровню риска классах к общему числу объектов. Содержательно, он показывает долю объектов, по которым результаты классификации совпали.

Полученные результаты, а именно: улучшение качества классификации в результате уточнения факторных весов методом статистических испытаний, с одной стороны, свидетельствует о преимуществах численных методов при решении задач, для которых не работают (или работают неэффективно) аналитические методы, с другой – свидетельствует о неоптимальности полученных изначально решений. Строго говоря, найденное в результате применения МСИ решение также не оптимально, хотя и обладает хорошими, по сравнению с прочими решениями, свойствами.

5.4. Тестирование модели

Тестовая выборка была использована для проверки точности модели. Веса факторов, полученные с использованием МСИ, использовались для расчета интегральных показателей и классификации предприятий тестовой выборки.

Таблица 20

ПО ТЕСТОВОЙ ВЫБОРКЕ

		Class Adj (Модель 1)			Всего
		№1	№2	№3	
Оценка (Модель 8)	№1	3	0	0	3
	№2	0	12	1	13
	№3	0	2	1	3
	Всего	3	14	2	19
	Хи ²	Индекс	ХиПирсон	ХиЧупрова	P-value
	20.740	84.211	0.722	0.739	0.000

Результаты анализа точности классификации предприятий тестовой выборки приведены в табл. 20.

Как видно из таблицы, более 84% предприятий классифицированы корректно. Это, по крайней мере, не меньше, чем по обучающей выборке. Полученные результаты свидетельствуют о приемлемой точности оценки факторных весов модели 8.

Итак, разработанная модель показала, что существуют возможности обучения лингвистической модели с целью учета априорной информации об уровне риска предприятия обучающей выборки и определения факторных весов модели. Предложена процедура поиска приемлемого набора весов факторов. К сожалению, предложенное решение не является оптимальным, но вычислимо с помощью доступного инструментария и позволяет сократить число некорректно классифицированных предприятий по сравнению с простым подбором весов.

Таблица 21

ПО ВСЕЙ ВЫБОРКЕ

		Class Adj (Модель 1)			
		№1	№2	№3	Всего
Оценка (Модель 8)	№1	6	0	0	6
	№2	0	23	2	25
	№3	0	4	3	7
	Всего	6	27	5	38
	Chi ²	Индекс	ХиПирсона	ХиЧупрова	P-value
	43.985	84.211	0.732	0.761	0.000

6. СРАВНЕНИЕ И ИНТЕРПРЕТАЦИЯ РЕЗУЛЬТАТОВ

Список моделей (нечетные – статистические модели, четные – лингвистические) приведен на рис. 4.

Модель 1	Статистическая модель, 3 класса, методы <i>k</i> -средних и ДА ¹³ , решающее правило
Модель 2	Лингвистическая модель, 5-уровневый классификатор, трапециевидные функции принадлежности, равные веса факторов
Модель 2А	Лингвистическая модель, 5-уровневый классификатор, сплайн функции принадлежности, равные веса факторов
Модель 4	Лингвистическая модель, 7-уровневый классификатор, трапециевидные функции принадлежности, равные веса факторов
Модель 4А	Лингвистическая модель, 7-уровневый классификатор, сплайн функции принадлежности, равные веса факторов
Модель 6	Лингвистическая модель, 5-уровневый классификатор, трапециевидные функции, веса соответствуют системе предпочтений ЛПП
Модель 6А	Лингвистическая модель, 5-уровневый классификатор, сплайн функции, веса соответствуют системе предпочтений ЛПП
Модель 8	Лингвистическая модель с обучением, 3-уровневый классификатор, трапециевидные функции, веса определены в ходе имитационного моделирования

Рис. 4. Статистические и лингвистические модели

Несмотря на различия методик, некоторые результаты могут сравниваться. Например, статистическая модель с 3-мя классами (модель 1) и лингвистическая модель с 5-уровневым классификатором (модель 2).

¹³ ДА – дискриминантный анализ.

Согласованность результатов моделей можно изучить с помощью таблиц сопряженности. Оценка уровня связи между грациями производится с помощью коэффициентов сопряженности Пирсона и Чупрова.

В данном случае «индекс соответствия» равен 47.5% (=100*(1+12+5+1)/40). Это означает относительно низкую степень соответствия результатов моделей. Главная причина – различие принципов моделей. Статистическая модель основана на сопоставлении объектов выборки, лингвистическая сопоставляет показатели объектов с заложенными в нее лингвистическими переменными. Не менее важны различия значимости показателей для анализа. Модель 2 приписывает равные веса всем факторам. В Модели 1 вклад факторов в классификацию неравный. Анализ дисперсий показывает, что наибольший вклад вносят X1_4, X3_1 и X1_2.

Таблица 22

АНАЛИЗ РАСПРЕДЕЛЕНИЯ ПРЕДПРИЯТИЙ ПО КЛАССАМ ДЛЯ МОДЕЛЕЙ 1 И 2

		Модель 2				
		Высокий риск	Средний риск	Низкий риск	Очень низкий риск	Всего
Модель 1	№1	1	5	0	0	6
	№2	0	12	15	0	27
	№3	0	1	5	1	7
	Всего	1	18	20	1	40
	Chi ²	Индекс	ХиПирсона	ХиЧупрова	P-значение для Chi ²	
	17.6190	47.5	0.5530	0.4241	0.0073	

Примечание: цветом выделены ячейки, содержащие число предприятий в предположительно соответствующих друг другу по уровню риска классах.

Модель 2 дает более оптимистичную оценку уровня риска изучаемых объектов, что связано с их высокими показателями рентабельности.

Подтверждением предположения о влиянии различной весомости факторов на результаты классификации может служить сопоставление моделей 1 и 6 (6А), в скобках приведены результаты модели 6А.

Таблица 23

АНАЛИЗ РАСПРЕДЕЛЕНИЯ ПРЕДПРИЯТИЙ ПО КЛАССАМ ДЛЯ МОДЕЛЕЙ 1 И 6(А)

		Модель 6 (6А)				
		Высокий риск	Средний риск	Низкий риск	Очень низкий риск	Всего
Модель 1	№1	6	0	0	0	6
	№2	5	16 (15)	6 (7)	0	27
	№3	0	1	5	1	7
	Всего	11	17 (16)	11 (12)	1	40
	Chi ²	Индекс	ХиПирсона	ХиЧупрова	P-значение для Chi ²	
	31.38 (30.04)	70 (67.5)	0.663 (0.655)	0.566 (0.554)	0.00002 (0.00004)	

По сравнению с моделью 2, модели 6 и 6А по результатам классификации ближе к модели 1, «индексы соответствия» 70% и 67.5% соответственно.

Оценка согласованности результатов моделей обобщена в форме табл. 25.

Пример с моделью 6 показывает, что при некотором выборе весов факторов совпадение результатов моделей может быть увеличено, это наводит на мысль о возможности «обучения» лингвистической модели.

Реализацией идеи обучения применительно к лингвистическим моделям кредитного риска стала разработка описанной выше модели 8. Как показал анализ, проведенный в предыдущем разделе, обучение модели позволяет так подобрать веса факторов, чтобы учесть априорную информацию о распределении предприятий по классам кредитного риска. Этот подход может использоваться в практических целях для разработки моделей экспресс-оценки кредитного риска предприятий в случае наличия некоторой априорной информации (различных вариантов обучающей выборки) (см. табл 24 и 25).

Таблица 24

РЕЗУЛЬТАТЫ КЛАССИФИКАЦИИ ПРЕДПРИЯТИЙ

№	Код предприятия	Модель №								
		1		2	2А	4	4А	6	6А	8
		До корректировки	После корректировки							
1	BISV	2	2	3	3	5	5	3	3	2
2	ESPK	2	2	4	4	5	5	4	4	2
3	KUBN	1	1	3	3	4	4	2	2	1
4	NNSI	2	2	3	3	4	4	2	2	2
5	SPTL	2	2	4	4	5	5	4	4	3
6	SVZK	2	2	3	3	4	4	3	3	2
7	URSI	1	1	2	2	3	3	2	2	1
8	ELKA	2	2	4	4	5	5	4	4	3
9	ELRO	3	2	4	4	6	6	4	4	3
10	ENCO	1	2	3	3	4	4	2	2	2
11	ESAL	3	3	4	4	6	6	4	4	3
12	ESHS	2	2	3	3	4	4	3	3	2
13	ESIR	3	3	3	3	4	5	3	3	2
14	ESKI	1	1	3	3	4	4	2	2	1

№	Код предприятия	Модель №								
		1		2	2А	4	4А	6	6А	8
		До корректировки	После корректировки							
15	ESKK	2	2	4	4	5	5	3	3	2
16	ESKT	2	2	4	4	5	5	3	4	2
17	ESMO	1	1	3	3	4	4	2	2	2
18	ESOB	2	2	3	3	4	5	3	3	2
19	ESOM	1	2	3	3	4	4	2	2	2
20	ESTU	1	1	3	3	4	4	2	2	1
21	ESVD	2	2	4	4	5	5	3	3	2
22	ESVL	2	2	4	4	5	5	3	3	2
23	HTCM	3	3	4	4	6	6	4	4	3
24	IVTL	2	2	4	4	5	5	3	3	2
25	KCHS	2	2	4	4	5	5	3	3	2
26	MGTS	1	1	3	3	4	4	2	2	1
27	MUEL	2	2	4	4	5	5	3	3	2
28	NVGT	2	2	3	3	4	4	2	2	2
29	RTKM	2	2	4	4	6	6	4	4	3
30	SMSI	2	2	3	3	5	5	3	3	2
31	SVIM	3	3	4	4	6	6	4	4	3
32	SVLN	2	2	3	3	5	5	3	3	2
33	TMTK	2	2	4	4	6	6	4	4	3
34	TUTE	2	2	3	3	4	4	2	2	2
35	UDTL	2	2	4	4	5	5	3	3	2
36	AHTC	2	2	3	3	5	5	3	3	2
37	ALCO	2	2	4	4	5	5	3	3	2
38	ELCH	3	3	4	4	5	5	4	4	3
39.	INSW	NA	3	4	4	5	5	4	4	3
40.	KGTS	NA	3	5	5	7	7	5	5	3

Примечание: последние 2 предприятия не участвовали в разработке решающих правил моделей в силу их заметного отличия по своим показателям от прочих предприятий выборки.

Таблица 25

АНАЛИЗ СОГЛАСОВАННОСТИ РЕЗУЛЬТАТОВ МОДЕЛЕЙ (СРАВНЕНИЕ С МОДЕЛЬЮ 1)

Наименование	Краткое описание модели	Хи ²	Индекс соответствия, %	Коэффициент согласия		Уровень значимости
				Пирсона	Чупрова	
Модель 1	Статистическая модель, 3 класса, методы <i>k</i> -средних и <i>ДА</i> ¹⁴ , решающее правило, по исходным данным	80	100	0.817	1	0.000000
Модель 2	Лингвистическая модель, 5-уровневый классификатор, трапециевидные функции принадлежности, равные веса факторов	17.619	47.5	0.553	0.424	0.007258
Модель 2А	Лингвистическая модель, 5-уровневый классификатор, сплайн функции принадлежности, равные веса факторов	17.619	47.5	0.553	0.424	0.007258
Модель 4	Лингвистическая модель, 7-уровневый классификатор, трапециевидные функции принадлежности, равные веса факторов	20.371	70	0.581	0.424	0.009021
Модель 4А	Лингвистическая модель, 7-уровневый классификатор, сплайн функции принадлежности, равные веса факторов.	24.520	72.5	0.617	0.466	0.001874
Модель 6	Лингвистическая модель, 5-уровневый классификатор, трапециевидные функции, веса соответствуют системе предпочтений ЛПП	31.380	70	0.663	0.566	0.000021
Модель 6А	Лингвистическая модель, 5-уровневый классификатор, сплайн функции, веса соответствуют системе предпочтений ЛПП	30.044	67.5	0.655	0.554	0.000039
Модель 8	Лингвистическая модель с обучением, 3-уровневый классификатор, трапециевидные функции, веса определены в ходе имитационного моделирования	40.051	81.579	0.716	0.726	0.000000

Примечание: нечетные № – статистические модели, четные – лингвистические.

¹⁴ ДА – дискриминантный анализ

ЗАКЛЮЧЕНИЕ

Применение дискриминантного анализа – мощный инструмент моделирования кредитного риска. Но для получения надежных результатов требуется репрезентативная обучающая выборка. В предложенной модели роль обучающей выборки исполняла классификация методом k -средних. Это ослабляет (ухудшает) качество модели, поскольку достоверной информации о реализации кредитного риска нет. В основу модели положено предположение о том, что предприятие с худшими финансовыми показателями по сравнению с другими предприятиями отрасли будет иметь больший уровень кредитного риска. Предположение вполне естественное, используется в финансовом анализе при межфирменных сопоставлениях. Однако на нем нельзя создать точную предсказательную модель кредитного риска. Результаты модели могут носить лишь рекомендательный, но не предписательный характер.

Классификация и решающие правила разрабатывались на основе выборки. Это, с одной стороны, позволило учесть отраслевые особенности, получить инструмент для сравнительной оценки предприятий связи, с другой – кроме отраслевой специфики в данных заключена информация о периоде анализа, то есть нет уверенности в устойчивости модели во времени.

Лингвистическая модель проще с практической точки зрения. Может использоваться для:

- 1) оценки интегрального показателя и составлении на его основе рейтинга;
- 2) распознавания уровня риска с помощью классификаторов.

Кроме того, частные критерии V_{ji} легко интерпретируются, с их помощью можно проводить факторный анализ. Однако на стадии разработки требуются экспертные данные для выбора ключевых показателей, создания лингвистических переменных, выбора весов показателей. Модель явно не учитывает отраслевой специфики, но в то же время не подвержена влиянию фактора времени, так как уровни показателей задаются не точками, а интервалами (изменения в рамках интервала, происходящие во времени не влияют на конечный результат).

Модель очень гибка и может учесть отраслевые особенности, например, на стадии выбора ключевых показателей, или разработки лингвистических переменных, или выбора весов. Гибкость – это и преимущество, и недостаток модели. Преимущество, потому что позволяет модифицировать модель в соответствии с требованиями ЛПР. Недостаток, так как при ошибочных, непропорциональных изменениях может вводить в заблуждение.

Разработанная лингвистическая модель с обучением до некоторой степени является переходной между первой и второй моделями. С одной стороны, позволяет учитывать отраслевые особенности, предпочтения ЛПР, сохраняя, таким образом, гибкость. С другой – реализует возможности обучения по выборочным данным.

Литература

1. Обработка нечеткой информации в системах принятия решений. / А.Н. Борисов, А.В. Алексеев, Г.В. Меркурьева и др. – М.: Радио и связь, 1989.
2. Недосекин А.О., Максимов О.Б. Применение теории нечетких множеств к финансовому анализу предприятий. / www.vtmgroup.ru.
3. Недосекин А.О., Максимов О.Б., Павлов Г.С. Анализ риска банкротства. Методические указания. / www.sedok.narod.ru.
4. Недосекин А.О. Финансовый менеджмент на нечетких множествах. // Аудит и финансовый анализ – М., 2003, №3.
5. Фишберн П. Теория полезности для принятия решений. – М.: Наука, 1978.
6. Шеремет А.Д. Комплексный анализ хозяйственной деятельности. – М.: ИНФРА-М, 2006.
7. Шеремет А.Д., Сайфулин Р.С. Методика финансового анализа. – М.: ИНФРА-М, 1996.
8. E.I. Altman. Financial Ratio, Discriminant Analysis and the Prediction of Corporate Bankruptcy. // The Journal of Finance, Sept. 1968 – pp.589-609.

9. A.F. Atiya. Bankruptcy Prediction for Credit Risk Using Neural Networks: A Survey and New Results. // IEEE TRANSACTIONS ON NEURAL NETWORKS, VOL. 12, NO. 4, JULY 2001 – pp. 929-935.
10. D.L. Chesser. Prediction Loan Noncompliance. // The Journal of Commercial Bank Lending, 56 (12), 1974 – pp.28-38.
11. A.M. Santomero. Commercial Bank Risk Management: An Analysis of the Process. // Journal of Financial Services Research, 12:2/3, 1997.

Ледовский Дмитрий Олегович

РЕЦЕНЗИЯ

Статья Ледовского Д.О. посвящена проблематике анализа кредитного риска предприятий. В статье рассматриваются экономико-математические модели экспресс-оценки кредитного риска на основе классификационного подхода.

Цель проведенного исследования – сопоставление результатов статистической и лингвистической моделей, разработка модели с обучением на базе лингвистической модели.

Автором были рассмотрены две модели оценки кредитного риска. Первая опиралась на методы статистики (кластерный, факторный и дискриминантный анализ). Вторая использовала аппарат нечеткого моделирования. Была предложена лингвистическая модель с обучением.

Модели были апробированы на реальных данных – бухгалтерской отчетности 40 российских предприятий связи. Автор изучил согласованность полученных результатов, провел сопоставление моделей с использованием статистических процедур (анализ сопряженности). На основании проведенного анализа выработаны рекомендации к практическому использованию и повышению предсказательной способности моделей.

Работа выполнена с использованием современных экономико-математических методов, в том числе корреляционного, кластерного и дискриминантного анализа, методов нечеткой логики, имитационного моделирования.

В статье предложен новый вариант лингвистической модели кредитного риска, в основе которого лежит идея оценивания весов факторов по реальным данным – обучение модели.

Предложенный алгоритм поиска весов дает приемлемые, но не оптимальные решения по критерию минимума некорректно классифицированных объектов и нуждается в доработке. Тем не менее, описанный в статье подход представляет научно-практический интерес, как задел для продолжения исследований и основа для проведения экспресс-оценки кредитного риска предприятий.

Рекомендую данную статью к публикации в журнале «Аудит и финансовый анализ».

Грачева М.В., д.э.н., профессор, заведующая кафедрой математических методов анализа экономики Экономического факультета МГУ им. М.В. Ломоносова

9.7. FIRM CREDIT RISK MODELING: CLASSIFICATION APPROACH

D.O. Ledovskij, the Post-graduate Student of Economic Faculty

Moscow State University it. M.B. Lomonosov

The article is devoted to an issue of firm credit risk estimation. The objects of research were 40 communication companies. The public financial reports of companies at 2001 and 2000 were used to credit risk estimation. The research was based on modern economic and mathematic methods: correlation analysis, cluster-analysis, discriminant analysis, fuzzy logic. In the article was implemented comparisons of statistic and linguistic models results, was proposed credit risk linguistic model with training.