

УДК 336.774.3

UDC 336.774.3

**ОЦЕНКА ПЛАТЕЖЕСПОСОБНОСТИ
РОССИЙСКИХ ПРЕДПРИЯТИЙ НА ОСНОВЕ
СОВРЕМЕННЫХ ТЕХНОЛОГИЙ
СТАТИСТИЧЕСКОГО МОДЕЛИРОВАНИЯ**

**ESTIMATION OF RUSSIAN ENTERPRISES
SOLVENCY ACCORDING TO THE MODERN
TECHNIQUES OF STATISTIC MODELLING
RESULTES**

Халафян Алексан Альбертович
к.т.н., доцент

Khalaphyan Aleksan Albertovich
Cand. Tech. Sci., associate professor

Пелипенко Екатерина Юрьевна
аспирант
*Кубанский государственный университет,
Краснодар, Россия*

Pelipenko Ekaterina Yurievna
postgraduate student
Kuban State University, Krasnodar, Russia

В статье предложен метод определения величины риска банкротства российских предприятия на основе статистических методов классификационного анализа, если известны определенные экономические показатели предприятия. Для автоматизации процесса классификации создан программный модуль на языке STATISTICA VISUAL BASIC (SVB)

In the present work Russian enterprises non-payment risk method is suggested. This one is based on methods of classification analysis in the case of financial indices patterns of the enterprise are known. The STATISTICA VISUAL BASIC (SVB) programm-module was created for automatization of classificaton process

Ключевые слова: КЛАССИФИКАЦИОННЫЙ АНАЛИЗ, КЛАСТЕРНЫЙ АНАЛИЗ, ДИСКРИМИНАНТНЫЙ АНАЛИЗ, ДЕРЕВЬЯ КЛАССИФИКАЦИИ, ЛИНГВИСТИЧЕСКАЯ ШКАЛА, ПЛАТЕЖЕСПОСОБНОСТЬ, ОБУЧАЮЩАЯ ВЫБОРКА, АПОСТЕРИОРНЫЕ ВЕРОЯТНОСТИ

Keywords: CLASSIFICATION ANALYSIS, CLUSTER ANALYSIS, DISCRIMINANT ANALYSIS, CLASSIFICATION TREES, LINGUAL SCALE, SOLVENCY, INSTRUCTION SAMPLE, POSTRIOR PROBABILITIES

Вопрос оценки платежеспособности предприятия в настоящее время является актуальным для большинства кредитных организаций. Используемые методы часто не дают точного результата из-за недостатка исходных данных и неустойчивости финансовых показателей (параметров) предприятия. Значения используемых в данных методах финансовых показателей не только изменяемы во времени, но и зависимы от территориального фактора самого предприятия. Объясняется это тем, что условия развития предприятий в различных регионах страны различаются. Например, метод Альтмана определения уровня риска предприятия не является устойчивым к вариациям исходных данных, так как эти данные обладают сильной динамикой изменения как во времени, так и в пространстве.

В настоящей работе предложен метод определения уровня риска неплатежеспособности предприятия, в основе которого лежит

статистический анализ виртуальной клиентской базы. Значения экономических показателей, занесенных в эту базу, устойчивы как во времени, так и в пространстве. Клиентская база виртуальных предприятий, принадлежащих к различным группам риска, методом дерева классификации или дискриминантным анализом позволяет прогнозировать принадлежность предприятия к той или иной группе риска неплатежеспособности. Для удобства использования предлагаемого метода в среде программы Statistica был создан модуль, позволяющий автоматизировать процесс классификации предприятий.

Наиболее удачным способом решения проблемы выбора устойчивых во времени и в пространстве финансовых показателей, по мнению авторов, является метод комплексного финансового анализа предприятий, предложенный в статье [1]. Это связано с тем, что комплексные финансовые оценки были выбраны из поля качественных, а не количественных признаков. Взаимосвязь комплексного показателя и количественных признаков проанализирована в статье [2]. В работе [3] впервые были предложены 16 показателей, наилучшим образом характеризующие состояние предприятий в условиях российской экономики. Именно эти, наиболее весомые экономические показатели, легли в основу создания клиентской базы, образующие четыре группы:

1. Рентабельность: **R1** – общая рентабельность; **R2** – рентабельность активов; **R3** – рентабельность собственного капитала; **R4** – рентабельность продукции; **R5** – рентабельность оборотных активов.

2. Ликвидность и платежеспособность: **L1** – быстрый коэффициент ликвидности; **L2** – коэффициент покрытия запасов; **P1** – текущий коэффициент ликвидности.

3. Деловая активность: **A2** – оборачиваемость активов; **A4** – оборачиваемость кредиторской задолженности; **A5** – оборачиваемость дебиторской задолженности; **A6** – оборачиваемость запасов.

4. Финансовая устойчивость: **F1** – коэффициент финансовой зависимости; **F2** – коэффициент автономии собственных средств; **F3** – обеспеченность запасов собственными оборотными средствами; **F4** – индекс постоянного актива.

Для качественной и более наглядной интерпретации уровней выбранных показателей использовалась лингвистическая шкала [4] – пятиуровневый классификатор (пенташкала), на которой уровни интервальных значений параметров представлены в пяти вариациях: «Очень низкий – Низкий – Средний – Высокий – Очень высокий». Для пенташкалы была построена система функций, которые соответствуют нечетким подмножествам. Простейшим способом задания является система трапециевидных нечетких чисел. На рисунке 1 приведен пример такой системы с нормированным значением показателя (диапазон изменения значения финансового показателя находится в пределах от 0 до 1). Чем ниже уровень показателя (рис.1), тем выше уровень риска банкротства предприятия.

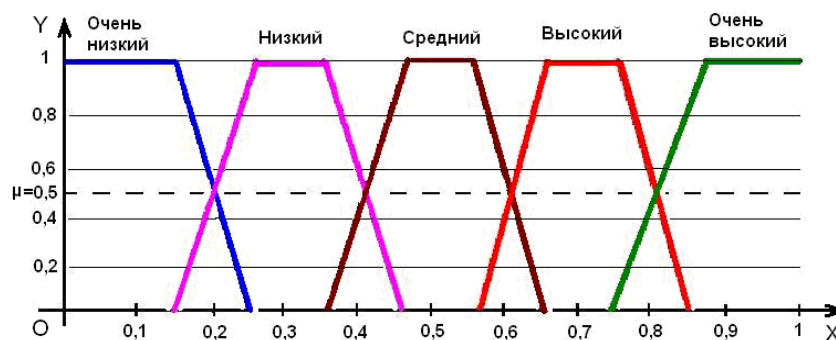


Рисунок 1 – Пример системы пенташкалы с нормированным значением экономического показателя

Ось **OX** – нормированное значение экономического показателя. Ось **OY** - уверенность специалиста при отнесении параметра к одному из пяти уровней риска.

При помощи рассмотренной классификации было выделено пять групп предприятий: *1 группа* – предприятия очень высокого уровня риска; *2 группа* – предприятия, находящиеся в зоне высокого риска; *3 группа* –

предприятия, находящиеся в зоне среднего риска; 4 группа – предприятия, находящиеся в зоне низкого риска; 5 группа – предприятия, находящиеся в зоне очень низкого риска (кредитоспособные предприятия).

Используя пенташкалу для каждого из показателей, была составлена таблица финансовых показателей, четко определяющих принадлежность предприятия к той или иной группе (табл.1).

Значение параметра приписывается к тому уровню, у которого ордината соответствующей трапеции для данного значения больше. Если ординаты совпадали, то значение параметра относили к «более низкому» лингвистическому уровню.

Таблица 1 – Диапазон значений экономических показателей по группам уровня риска

Группы уровня риска	Экономические показатели					
	L1	L3	P1	F1	F2	
Размерность показателя	нет	%	нет	нет	нет	
Очень высокий	0,1 – 0,2	40 – 77	0,1 – 0,4	2,70 – 2,50	0,25 – 0,31	
Высокий	0,2 – 0,6	77 – 144	0,4 – 0,8	2,50 – 1,70	0,31 – 0,54	
Средний	0,6 – 1,2	144 – 267	0,8 – 1,5	1,70 – 1,17	0,54 – 0,83	
Низкий	1,2 – 2,0	267 – 454	1,5 – 2,4	1,17 – 0,90	0,83 – 0,93	
Очень низкий	2,0 – 2,2	454 – 580	2,4 – 2,8	0,90 – 0,50	0,93 – 0,95	
Группы уровня риска	Экономические показатели					
	F3	F4	R1	R2	R3	
Размерность показателя	нет	нет	%	% за кв.	% за кв.	
Очень высокий	-2,00 – -1,66	1,60 – 1,51	0 – 1	-3,00 – -2,50	-4,00 – -3,33	
Высокий	-1,66 – -0,33	1,51 – 1,18	1 – 5	-2,50 – -0,50	-3,33 – -0,66	
Средний	-0,33 – 0,92	1,18 – 0,83	5 – 16	-0,50 – 1,57	-0,66 – 3,17	
Низкий	0,92 – 1,84	0,83 – 0,56	16 – 27	1,57 – 4,34	3,17 – 7,84	
Очень низкий	1,84 – 3,50	0,56 – 0,20	27 – 30	4,34 – 8,00	7,84 – 20,00	
Группы уровня риска	Экономические показатели					
	R4	R5	A2	A4	A5	A6
Размерность показателя	%	% за кв.	раз в кв.	раз в кв.	раз в кв.	раз в кв.
Очень высокий	0 – 1	-5,00 – -4,42	0,06 – 0,08	0,40 – 0,50	0,6 – 0,7	1,0 – 1,3
Высокий	1 – 5	-4,42 – -0,58	0,08 – 0,15	0,50 – 0,90	0,7 – 1,1	1,3 – 2,3
Средний	5 – 15	-0,58 – 2,58	0,15 – 0,24	0,90 – 1,54	1,1 – 1,74	2,3 – 3,5
Низкий	15 – 27	2,58 – 10,27	0,24 – 0,29	1,54 – 2,07	1,74 – 2,3	3,5 – 5,5
Очень низкий	27 – 46	10,27 – 18,0	0,29 – 0,58	2,07 – 5,80	2,3 – 5,8	5,5 – 15

При помощи средств программы Statistica 6 функцией генерации нормально распределенных случайных чисел была составлена база данных для 1000 предприятий пяти уровней риска (по 200 предприятий каждого класса) по всем шестнадцати экономическим показателям.

Чтобы убедиться в том, что полученные данные определяют принадлежность к обозначенным классам, был проведён кластерный анализ.

Задача исследований заключалась в проверке схожести предприятий, принадлежащих одному классу и максимальной различимости предприятий различных классов. Кластерный анализ проводился с различным заданием количества кластеров.

Сравнение процесса генерации виртуальной базы предприятий с полученным разбиением на кластеры позволило убедиться, что была проведена 100 % успешная кластеризация: все предприятия попали в кластеры, соответствующие 5 классам уровней риска. Это значит, что сгенерированные данные качественно определяют принадлежность предприятий к обозначенным классам.

Используя средства программы Statistica был проведен дискриминантный анализ виртуальной клиентской базы предприятий и построено дерево классификации.

Исследование итоговой таблицы анализа данных (табл.2) показало, что коэффициенты рентабельности активов (**R2**), финансовой зависимости (**F1**) и автономии собственных средств (**F2**) являются главными переменными, которые позволяют производить дискриминацию между различными группами предприятий, так как они дают наиболее значительный вклад в общую дискриминацию.

Таблица 2 – Итоговая таблица дискриминантного анализа данных

Число переменных: 16; Группирующие переменные: Классы (5 групп); Лямбда Уилкса: 0,00000 F(64,3838)=1300,3 p<0,0000						
	Лямбда Уилкса	Частная лямбда	Значение F-критерия	Уровень значимости F – критерия	Толерантность	Коэффициент множественной корреляции
L1	0,000006	0,827377	51,1165	0,000000	0,971886	0,028114
L2	0,000006	0,845276	44,8461	0,000000	0,981104	0,018896
P1	0,000006	0,840156	46,6124	0,000000	0,984020	0,015980
F1	0,000008	0,626187	146,2570	0,000000	0,989247	0,010753
F2	0,000007	0,664263	123,8299	0,000000	0,995379	0,004621
F3	0,000006	0,883961	32,1615	0,000000	0,966743	0,033257
F4	0,000006	0,877814	34,1025	0,000000	0,984351	0,015649
R1	0,000006	0,764040	75,6640	0,000000	0,979979	0,020021
R2	0,000013	0,366750	423,0309	0,000000	0,988742	0,011258
R3	0,000005	0,910658	24,0364	0,000000	0,974302	0,025698
R4	0,000006	0,893515	29,1980	0,000000	0,975140	0,024860
R5	0,000006	0,882433	32,6416	0,000000	0,971746	0,028254
A2	0,000005	0,934974	17,0395	0,000000	0,972912	0,027088
A4	0,000005	0,930492	18,3016	0,000000	0,981252	0,018748
A5	0,000006	0,861200	39,4867	0,000000	0,963720	0,036280
A6	0,000006	0,883622	32,2679	0,000000	0,972059	0,027941

Величина единичного вклада соответствующей переменной в разделительную силу модели определяется значением *Partial Lambda* (частной лямбды): чем меньше ее величина, тем больше вклад параметра в общую дискриминацию.

Толерантность (*Toler*) является мерой избыточности переменной в модели (чем меньше ее значение, тем избыточнее переменная в модели, тем меньшую дополнительную информацию несет переменная). Из таблицы 2 видно, что значения толерантности всех показателей достаточно большие, значит можно говорить об их успешном включении в модель.

Для получения дальнейших результатов о природе дискриминации был проведен канонический анализ. Для оценки того, как шестнадцать переменных разделяют различные совокупности (группы предприятий), были вычислены четыре дискриминантные функции (число совокупностей – групп минус один). Результат пошагового критерия для

дискриминантных функций (канонических корней) позволил определить, что все эти корни статистически значимы.

Стандартизованные коэффициенты дискриминантных функций, характеризующие направление и вклад переменных (экономических показателей) в значения дискриминантных функций, показали, что в дискриминантную функцию 1 (Root 1) наибольший вклад вносят переменные L3, P1; в дискриминантную функцию 2 (Root 2) – R2, F2; в дискриминантную функцию 3 (Root 3) – R2, F1; в дискриминантную функцию 4 (Root 4) – R1, R2.

Из результатов анализа следует, что функция 1 ответственна за 96,61% объясненной дисперсии, т.е. 96,61% всей дискриминирующей мощности определяется этой функцией. Поэтому эта функция наиболее «важна». Функция 2 ответственна за 2,65% объясненной дисперсии, функция 3 – за 0,72%, а функция 4 – за 0,02%.

Таблица 3 средних значений для дискриминантных функций позволила определить группы, лучше всего идентифицируемые конкретной дискриминантной функцией. Функция 1 идентифицирует в основном группы 1 и 5, т.е. «ОТКАЗ» и «КРЕДИТ», так как им соответствуют наибольшие значения функции 1. Функция 2 – группы 4 и 5, т.е. «НИЗКИЙ РИСК» и «КРЕДИТ». Функция 3 – группы 3 и 4, т.е. «НИЗКИЙ РИСК» и «СРЕДНИЙ РИСК». Функция 4 – группы 1 и 2, т.е. «ОТКАЗ» и «ВЫСОКИЙ РИСК».

Таблица 3 – Средние значения канонических переменных

Группы риска	Средние значения канонических переменных			
	Корень 1	Корень 2	Корень 3	Корень 4
Очень высокий	36,3172	-3,65463	-2,57269	-0,416201
Высокий	24,3966	-0,88211	1,06452	0,755099
Средний	3,8340	2,84699	4,24179	-0,408506
Низкий	-18,0988	7,82423	-2,67280	0,052658
Очень низкий	-46,4491	-613447	-0,06081	0,016950

Матрица классификации, которая содержит информацию о количестве и проценте корректно классифицированных наблюдений в каждой группе, показала, что исходные (строки матрицы) и предсказанные (столбцы матрицы) классы полностью совпали. Это также означает, что дискриминация была проведена успешно.

В таблице 4 приведен фрагмент таблицы апостериорных вероятностей, в которой отображена вероятность принадлежности каждого предприятия к группам риска банкротства.

Таблица 4 – Значения апостериорных вероятностей

Случаи	Исходная классификация	Очень высокий p=0,2000	Высокий p=0,2000	Средний p=0,2000	Низкий p=0,2000	Очень низкий p=0,2000
1	Очень высокий	1,000000	0,000000	0,000000	0,000000	0,000000
2	Очень высокий	1,000000	0,000000	0,000000	0,000000	0,000000

Клиент приписывается к той группе риска неплатежеспособности, которой соответствует наибольшая апостериорная вероятность классификации.

При помощи языка STATISTICA VISUAL BASIC (SVB), была создана первая часть программного модуля анализа данных, определяющая предприятие по введенным значениям его финансовых показателей в одну из пяти рассматриваемых групп уровня риска неплатежеспособности исходя из анализа таблицы значений апостериорных вероятностей.

Заключительным этапом работы стало построение дерева классификации, как альтернативного метода оценки платежеспособности предприятия. Данный метод является независимым от уже описанного ранее метода дискриминантного анализа, но в совокупности они могут показать более полную картину финансового состояния предприятия.

Процесс построения дерева классификации включил в себя три основных этапа:

1. Выбор критерия точности прогноза;
2. Выбор типа ветвления;
3. Определение «подходящего размера» дерева.

Критерий точности прогноза определяется наименьшей ценой (т.е. долей неверно классифицированных наблюдений). Наиболее «важным» классом является 5 группа риска неплатежеспособности, так как ошибка отнесения клиента к группе кредитоспособных предприятий говорит о высокой вероятности невыплаты денежных средств по долговому обязательству заемщика. Поэтому цену ошибок для предприятий, относящихся к группе очень низкого уровня риска неплатежеспособности, принимаем равную 0. А для остальных групп – равную по 0,25. Все внедиагональные элементы матрицы цен ошибок полагаются равными 1 (табл. 5).

Таблица 5 – Матрица цен ошибок

Уровень риска	Очень высокий	Высокий	Средний	Низкий	Очень низкий
Очень высокий	0,25	1	1	1	1
Высокий	1	0,25	1	1	1
Средний	1	1	0,25	1	1
Низкий	1	1	1	0,25	1
Очень низкий	1	1	1	1	0

В модуле дерева классификации программы Statistica реализованы два алгоритма ветвления: QUEST и CART.

Анализ метода QUEST (метод дискриминантного многомерного ветвления по линейной комбинации) показал более качественную классификацию, чем метод CART. Для него удалось построить дерево классификации (рис.2) с нулевой матрицей глобальной кросс-проверки.

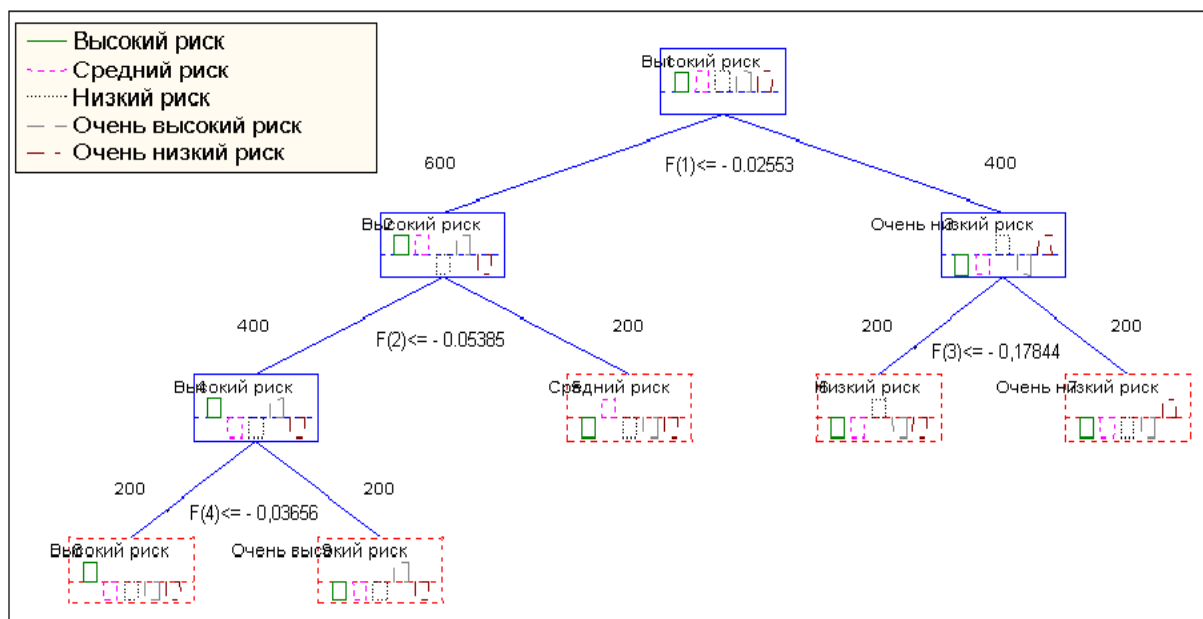


Рисунок 2 – Граф дерева классификации при выборе типа ветвления по методу QUEST

Цену глобальной кросс-проверки и стандартное отклонение цены получили равными 0, что говорит о более точном построении дерева классификации.

Так как дерево бинарное, то из каждой вершины выходят две ветви. Под вершинами указаны условия перехода по ветвям (левая ветвь отвечает за выполнение условия, правая – за нарушение). В верхней части вершин указано название класса, над вершиной – количество предприятий этого класса.

Красным пунктиром обозначены терминальные вершины (листья) дерева, дальнейшее ветвление из которых невозможно. Высота гистограмм, изображенных в вершинах дерева, соответствует количеству предприятий определенного класса в данной вершине.

Для выбора дерева классификации «подходящего размера» все значения параметров были выбраны согласно методу автоматического построения дерева Бримана [5]: стандартная ошибка равна 1, остановка определяется ошибкой классификации. Минимальное число неправильно классифицированных наблюдений в терминальной вершине было принято равным 5.

Результаты классификации, отражающие количество объектов, отнесенных к той или иной группе риска, показали, что исходные и предсказанные группы полностью совпали.

Числовые значения из таблицы решающих правил (условий ветвления) были использованы в качестве коэффициентов для составления линейных функций $F(1)$, $F(2)$, $F(3)$, $F(4)$, представляющих собой условия перехода алгоритма второй части программного модуля анализа данных (см. рис.2), написанного в среде SVB.

Таким образом, был создан программный модуль, состоящий из двух частей, который позволяет производить оценку платежеспособности реально существующего предприятия методами деревьев классификации и дискриминантного анализа. По введенным значениям финансовых показателей предприятий модуль классифицирует их по пяти группам уровня риска неплатежеспособности. Для удобства использования модуля были созданы пользовательские диалоговые окна ввода и вывода данных, содержащие результаты классификации для каждого из методов.

Созданный программный модуль автоматизирует процесс классификации предприятий по уровню платежеспособности, тем самым значительно сокращая затраты времени пользователя, предупреждая ошибки кредитования. Используя данный модуль можно оперативно прогнозировать уровень риска неплатежеспособности реально существующего предприятия. От пользователя не требуется специальных знаний в области современных технологий анализа данных, так как программа имеет простой интерфейс, состоящий из окон ввода и вывода данных.

Список литературы

1. Недосекин А.О., Бессонов Д.Н. Сводный финансовый анализ российских предприятий за 2000-2003 г.г. // Аудит и финансовый анализ. 2005. № 1. С 53 - 60.
2. Недосекин А.О. Применение теории нечетких множеств к задачам управления финансами // Аудит и финансовый анализ. 2000г. №2. С. 32-34.

3. Давыдова Г.В., Беликов А.Ю. Методика количественной оценки риска банкротства предприятий // Управление риском. 1999 г. № 3. С 13-20.

4. Недосекин А.О., Фролов С.Н. Лингвистический анализ гистограмм экономических факторов // Ученые Записки Российского Государственного Социального Университета – Издательство РГСУ, 2008.

5. Breiman L.J. Random Forests–Random Features // UC Berkeley, Statistics Department, Technical Report. 1999. N 567.