

УДК: 330.4; 519.237

Оценка кредитного риска на основе методов многомерного анализа

Е. В. Орлова

Уфимский государственный авиационный технический университет,
Институт экономики и управления, кафедра экономики предпринимательства,
Россия, 450000, г. Уфа, ул. К. Маркса, д. 12

E-mail: ekorl@mail.ru

*Получено 19 августа 2013 г.,
после доработки 15 ноября 2013 г.*

В статье предложена авторская методика многомерного анализа для формирования прогнозной оценки кредитного риска организаций, основанная на использовании информации кредитных историй, учитывающая объемы и сроки предоставляемых кредитов. Рассмотрен пример оценки кредитного риска на статистических данных кредитной организации.

Ключевые слова: оценка и прогнозирование риска, многомерный анализ данных, кластерный анализ, факторный анализ, кредитные организации

Credit risk assessment on the basis of multidimensional analysis

E. V. Orlova

Ufa state aviation technical university Organization, 12 K. Marks st., Ufa, 450000, Russia

Abstract. — The article is devoted to description the author's method of multidimensional analysis for generate an predictive assessment of organizations' credit risk, based on the credit history information, which taking into account value and period of credit. An example of credit risk assessment is given.

Keywords: risk assessment and prediction, data multidimensional analysis, cluster analysis, factor analysis, credit organizations

Citation: *Computer Research and Modeling*, 2013, vol. 5, no. 5, pp. 893–901 (Russian).

© 2013 Екатерина Владимировна Орлова

Введение

Кредитные риски являются одними из базовых рисков кредитных организаций и связаны с неплатежами по взятым обязательствам при осуществлении заемной деятельности. Данный вид риска проявляется в полном или частичном невозврате заемных ресурсов или процентов по ним. Этот риск можно определить как вероятность и величину ущерба, понесенного кредитной организацией в связи с невозможностью или нежеланием заемщика обеспечить соответствующие денежные потоки для погашения кредитного долга и процентов по нему [Гузаиров, Орлова, 2012, С. 227–228; Орлова, 2012а, С. 123–139].

Традиционно для проведения мониторинга платежеспособности своих клиентов кредитные организации используют скоринговые модели, с помощью которых на основе анализа кредитных историй предыдущих клиентов составляется рейтинг заемщиков и определяется вероятность возврата кредита в срок потенциальным заемщиком. Основные проблемы и решаемые в научных исследованиях задачи, связанные с построением и использованием скоринговых моделей при принятии решений, можно объединить в две группы.

Первая группа проблем связана с подбором адекватного сложности решаемых задач инструментария, формированием и обоснованием системы факторов, включаемых в модель. Известные модели для оценки кредитного риска используют статистический подход и базируются на обработке эмпирической информации прошлых кредитных историй, но отличаются способами построения аппроксимирующих зависимостей — нейросетевые, нечеткие и гибридные алгоритмы для построения приближения [Лебедев, 2007; Макаренко, 2012; Пятковский, 2008], эконометрические методы [Банкова, 2011; Глинкина, 2011; Crone, 2012; Crook, 2007; Mircea, 2011; Ong, 2005]. Обсуждению подвергаются способы сбора необходимой информации, необходимость и количество включаемых качественных характеристик в модельные конструкции для более точного отражения портрета заемщика, методы идентификации моделей, анализ их качества и прогностических свойств и другие.

Вторая группа проблем связана с разработкой систем автоматизации процессов сбора, обработки и хранения информации о заемщиках, проектированием систем поддержки принятия инвестиционных решений банка [Бесулин, 2013; Пятковский, 2008], формированием баз данных клиентов. В условиях большого потока разнородных клиентов основным требованием при разработке таких систем является скорость принятия решений.

Анализ существующих методических подходов и аналитического инструментария показал, что существующие модели оценки риска невозврата кредитов не позволяют выявить тенденции в поведении определенных категорий клиентов, имеющих схожий экономический портрет [Исмагилова, Орлова, 2012, С. 174–175; Орлова, 2012б, С. 187]. Формирование таких групп клиентов позволит, с одной стороны, выявлять общие закономерности экономического поведения агентов, с другой, сформировать совокупность дифференцированных требований со стороны кредитной организации, предъявляемых к заемщикам отдельных групп с учетом их специфики, в-третьих, учесть склонность к риску лица, принимающего решения по элементам кредита — объемам, срокам, процентам.

Предлагаемая методика

Методика оценки кредитного риска основана на применении методов многомерного анализа данных кредитных историй заемщиков — кластеризации, факторного анализа (дисперсионного и метода главных компонент). Этапы методики:

- i) снижение размерности выборки о кредитных историях клиентов, исследование значимых факторов, влияющих на уровень кредитоспособности заемщиков;
- ii) классификация клиентов, выявление однородных групп с одинаковым уровнем риска (надежные и рискованные группы клиентов);
- iii) прогнозирование уровня кредитного риска для потенциального клиента и обоснование решения об отказе или выдаче кредита.

Классификация клиентов кредитной организации

Для эффективного управления рисками в организации, занимающейся кредитованием, следует разработать такие требования к клиентам-заемщикам, которые были бы достаточно привлекательны для них и в то же время гарантировали возврат кредитов. Однако для каждого конкретного заемщика разрабатывать индивидуальные требования нецелесообразно. Для решения этой задачи следует выделить несколько основных групп клиентов, обладающих похожими свойствами. А затем, исходя из свойств, характеризующих клиентов в однородных группах, выработать адекватные для каждой группы клиентов требования.

Классификация как метод объединения исследуемых элементов системы в однородные группы, имеющих схожие свойства. Для этого требуется выбрать такой метод классификации/кластеризации, который наиболее точно отражает структуру исходных данных и дает более точное разбиение. Традиционно для этого используются методы кластерного анализа, нейросетевые методы. Существуют два принципиально различных способа кластеризации — обучение без учителя (иерархические методы), при котором количество кластеров заранее не задано, обучение с учителем (метод k -средних) — количество кластеров задано. Задача кластеризации хорошо решается также с помощью методов нейросетевого моделирования. При этом обучающая выборка может быть задана (многослойный персептрон), или не известна (самоорганизующиеся карты Кохонена). В [Исмагилова, Орлова, 2004, С. 51–55] было показано, что в условиях больших многомерных выборок данных эти два метода дают примерно одинаковое разбиение объектов на классы. В задаче моделирования рисков будем использовать методы кластерного анализа.

Можно отметить, что методы кластеризации применимы для решения следующих экономических задач:

- i) управление проектами, классификация инвестиционных (инновационных) проектов;
- ii) анализ инвестиционных и кредитных рисков, построение скоринговых моделей, потребительский кредитный мониторинг, классификация задолженностей, классификация кредиторов;
- iii) анализ и управление потребительским поведением, сегментация рынка, дифференциация показателей по степени удовлетворенности услугами;
- iv) выявление потенциальных отраслевых кластеров, определение кластерных эффектов и др.

Факторный анализ и снижение размерности выборки

Для выявления значимых факторов, определяющих изменение кредитного рейтинга предприятий-заемщиков, оценки уровня риска неплатежеспособности предприятий используются метод факторного анализа, позволяющего решить задачи. Во-первых, снизить размерность исходных данных, описывающих кредитную историю предыдущих клиентов, не утратив важной информации. Во-вторых, объединить в несколько факторов исследуемые показатели, описываемые разными свойствами клиентов. При этом в факторы объединяются коррелирующие показатели, то есть такие, коэффициент корреляции между которыми превосходит 0.7. Такое объединение позволяет перераспределить дисперсию между отдельными факторами (компонентами) и получить наглядную структуру данных.

Будем использовать метод главных компонент как процедуру снижения размерности исходных данных без потери информативности данных. Суть метода заключается в преобразовании исходных коррелирующих переменных в другой набор некоррелирующих переменных на основе вращения факторов.

Модель оценки риска

Для оценки риска инвестиционной деятельности необходима статистика нарушений клиентами условий договоров и размер ущерба, причиненного кредитной организации, вследствие

каждого такого нарушения. При этом каждый из выявленных классов клиентов характеризуется рядом показателей — численность класса, удельный объем оборотных средств, приходящихся на каждого клиента, коэффициент вариации денежного потока при нарушении договорных обязательств и другие. Величину риска как размер ущерба (*risk* — невыплата в срок основного долга) можно рассматривать как регрессионную зависимость от факторов среднего размера кредита x_1 , срока, на который выдается кредит x_2 , и ряд других выявленных факторов. Спецификация и идентификация такой регрессии должна проводиться с использованием информации об ущербе и характеристиках кредита для клиентов каждого класса. Прогноз уровня риска потенциального клиента может быть проведен на основе одной из этих моделей.

Численный эксперимент

Модельные эксперименты проводятся с использованием статистической информации о кредитных историях клиентов кредитной организации г. Уфы, занимающейся потребительским кредитованием, с использованием программы статистической обработки данных Statistica 7.0. Особенности изучаемых данных состоят в их разнородности и многомерности. Исследуемая модельная выборка включает данные по 38 клиентам и состоит из следующих показателей, характеризующих заемщиков: срок кредита (*month*), сумма кредита (*value*), пол (0 — женщина, 1 — мужчина), возраст (*age*), количество детей (*children*), средний доход (*income*), а также каждому заемщику присвоим переменную *problems*, характеризующую наличие или отсутствие проблем с возвратом кредита (0 — проблема отсутствует, 1 — проблема существует), и величину экономического ущерба *risk*. Фрагмент таблицы с исходными для анализа данными представлен ниже.

Таблица 1. Исходные данные о клиентах-заемщиках

Номер клиента	Срок кредита, мес.	Сумма кредита, руб.	Возраст	Число детей	Пол (0 — муж., 1 — женц.)	Среднемесячный доход, руб.	Наличие проблем с возвратом кредита (0 — нет, 1 — да)
1	5	10000	25	1	0	15000	1
2	12	10000	36	2	0	10000	1
3	12	13000	56	3	0	8000	0
4	12	15000	45	4	0	12000	0
5	6	15000	39	2	0	11256	0
6	12	17000	42	0	0	18000	1
7	12	25000	44	0	1	23000	1
8	12	25000	48	0	0	25035	1
9	6	25000	20	0	1	30000	0
10	12	25000	21	0	1	15065	0

Факторный анализ заемщиков

Решим задачу факторного анализа по снижению размерности данных и прогнозированию риска невозврата кредита. Для снижения размерности исходных данных и для выявления наиболее значимых факторов, влияющих на кредитный риск, будем использовать модуль факторного анализа, включающего методы главных компонент, дисперсионного и корреляционного анализа. Процедура осуществляется пошагово.

Шаг 1. Задание исходных параметров задачи. Определим количество факторов, равное количеству исходных переменных, т. е. шести, переменные *risk* и *problems* при проведении факторного анализа не учитываются.

Шаг 2. Вычисление собственных значений факторов. Собственные значения отражают дисперсию вновь выделенного фактора, таблице 2. В третьем столбце для каждого фактора приводится процент от общей дисперсии.

Таблица 2. Собственные значения выделенных факторов

Фактор	Собственное значение	Объясненная вариация	Накопленная объясненная вариация
1	1.997931	33.29886	33.29886
2	1.394679	23.24464	56.54350
3	1.134050	18.90083	75.44433

Из таблицы видно, что первый фактор объясняет 33,3 % общей дисперсии, фактор 2 — 23,2 %, а третий — 18,9 %. На основе полученной информации об объясненной каждым фактором дисперсии можно перейти к вопросу о количестве факторов, которые следует оставить. Для этого используются факторные нагрузки, которые можно интерпретировать как корреляционные связи между выделяемыми факторами и исходными переменными.

Шаг 3. Исследование факторных нагрузок. Сначала оценим факторные нагрузки без вращения для всех шести исходных переменных. Результаты анализа факторных нагрузок без вращения показаны в таблице 3.

Таблица 3. Значения факторных нагрузок для операции «без вращения»

Переменная	Фактор 1	Фактор 2	Фактор 3
month	0.540868	0.104905	0.655439
value	0.818261	-0.333261	-0.195093
age	-0.405546	-0.251135	0.639473
children	-0.205777	-0.778772	-0.377427
sex	0.330074	0.703828	-0.324087
income	0.848576	-0.328151	0.111284

Выделение факторов происходит так, чтобы последующие факторы включали в себя все меньшую и меньшую дисперсию. Фактор 1, как видно из таблицы 2, имеет наивысшие значения нагрузок для переменных, относящихся к экономическим характеристикам клиентов. Фактор 2 имеет максимальные нагрузки для переменных, связанных с социальным статусом клиента.

Шаг 4. Уточнение количества факторов. Для сравнения и окончательного принятия решения о количестве выделяемых факторов производится вращение факторов. Используем метод *Varimax row* (Варимакс вращение) — наиболее распространенный метод вращения, при котором факторы остаются независимыми по отношению друг к другу, так что значения переменных по одному фактору не коррелируют со значениями по другим факторам. Результаты вращения факторов показаны в таблице 4.

Таблица 4. Значения факторных нагрузок для операции «варимакс вращение»

Переменная	Фактор 1	Фактор 2	Фактор 3
month	0.461165	0.674559	0.255813
value	0.875882	-0.108511	-0.199325
age	-0.238427	0.136053	0.749069
children	0.140263	-0.840554	0.250341
sex	-0.016102	0.359075	-0.761683
income	0.909886	0.109644	0.015300

Уточнение описательных характеристик выявленных факторов показывает, что первый фактор связан с финансово-экономическими параметрами заемщика (средний доход, объем кредита), третий отражает его личностные (возрастные, принадлежность к определенному по-

лу) особенности, второй связан с социальными параметрами (количество детей). Кроме того, три фактора наиболее полно описывают 76 % вариации исходных данных. Поэтому целесообразно продолжить анализ на основе трех выделенных факторов.

Шаг 5. Оценка адекватности решения. Для проверки правильности выбора числа выделенных факторов необходимо построить воспроизведенную корреляционную матрицу, которая по своим коэффициентам должна быть близка к исходной корреляционной матрице в том случае, если факторы выделены правильно. Для определения степени возможного отклонения элементов этой матрицы от исходной формируется матрица остаточных корреляций, элементы которой равны разности между элементами исходной и воспроизведенной матриц. Исходная и остаточная корреляционные матрицы показаны в таблицах 5, 6.

Таблица 5. Исходная корреляционная матрица

Переменная	month	value	age	children	sex	income
month	1.00	0.20	0.01	-0.25	0.05	0.37
value	0.20	1.00	-0.20	0.13	0.18	0.70
age	0.01	-0.20	1.00	0.06	-0.25	-0.17
children	-0.25	0.13	0.06	1.00	-0.34	-0.04
sex	0.05	0.18	-0.25	-0.34	1.00	-0.02
income	0.37	0.70	-0.17	-0.04	-0.02	1.00

Таблица 6. Матрица остаточных корреляций

Переменная	month	value	age	children	sex	income
month	0.27	-0.08	-0.17	0.18	0.01	-0.13
value	-0.08	0.18	0.17	-0.04	0.08	-0.08
age	-0.17	0.17	0.36	0.02	0.26	0.02
children	0.18	-0.04	0.02	0.21	0.16	0.08
sex	0.01	0.08	0.26	0.16	0.29	-0.03
income	-0.13	-0.08	0.02	-0.08	-0.03	0.16

Входы в матрице остаточных корреляций могут быть интерпретированы как суммарные корреляции, за которые не могут быть ответственны полученные факторы. Диагональные элементы матрицы содержат стандартные отклонения, за которые не могут отвечать эти факторы, и равны квадратному корню из единицы минус соответствующие общности для двух факторов. Общностью переменной является дисперсия, которая объяснена выбранным числом факторов. Тщательный анализ остаточной матрицы показывает, что здесь фактически не имеется остаточных корреляций, больших по модулю 0.26. Следовательно, выделенные факторы адекватно отражают исходную информацию.

Классификация клиентов

Сначала формируют однородные группы клиентов по двум показателям, определенным первым фактором, выделив переменные «объем кредита» и «срок кредита». Кластеризация проводится в два этапа — качественный анализ с помощью иерархических методов и анализ с помощью метода *k*-средних. Разведочный анализ по выяснению возможного количества групп проводится с помощью иерархической классификации с использованием разных мер сходства и различия объектов в группах — евклидово расстояние, расстояние Манхэттена, расстояние Чебышева — для оценки степени близости объектов внутри групп и мер расстояний между кластерами — одиночной, полной связи. Изменяя меры расстояния, можно качественно оценить возможный состав кластеров и количество этих кластеров. Анализ различных разбиений исходной выборки методом иерархической классификации показал, что можно образовать от трех до шести кластеров. Для более обоснованной группировки объектов необходимо использование методов кластеризации, использующих количественные критерии для оценки ка-

чества разбиения. К таким методам можно отнести метод k -средних. Ниже представлены результаты разбиения, в котором четыре группы ($k = 4$) и показано значимое отличие образованных классов между собой. Результаты однофакторного дисперсионного анализа для выявления сходства/различия групп представлены в таблице 7.

Таблица 7. Результаты однофакторного дисперсионного анализа

Переменная	Межгрупповая сумма квадратов отклонений для числа степеней свободы, равного 3	Внутригрупповая сумма квадратов отклонений для числа степеней свободы, равного 34	F-критерий
month	781505.7	6325.706	14.0017
value	489827.5	4565.484	121.5945

Таблица исходных данных дополнилась информацией о кластере, к которому принадлежит конкретный клиент, евклидовом расстоянии и номере наблюдения. Теперь для каждого кластера можно вычислить основные описательные статистики, строить модели регрессии и формировать соответствующие прогнозы. Статистические характеристики — среднее значение, стандартное отклонение для каждого класса по каждому из исследуемых признаков показаны, а также количество объектов в каждом классе показаны в табл. 8.

Таблица 8. Описательные статистики выделенных классов

Класс-тер	Количество клиентов	Номер клиента	Описательные статистики переменной month		Описательные статистики переменной value	
			Среднее значение	Среднеквадратическое отклонение	Среднее значение	Среднеквадратическое отклонение
1	7	1–16, 24, 30	33.4	15.44	327142.9	76313.9
2	6	17–21, 37	16	15.95	125000	41833.0
3	7	25, 31–36	51.4	13.35	52142.9	11495.3
4	18	22, 23, 26–29, 38	14.6	12.26	22222.2	8292.8

Итак, на данном этапе получена информация о количестве классов заемщиков, их размере, доле оборотных средств каждого класса в общей их сумме, информация о размере удельной суммы оборотных средств, приходящейся на каждого элемента класса. Самый многочисленный класс (18 заемщиков) — четвертый, характеризуется самыми низкими значениями кредита, от 10 до 35 тыс. руб., причем колеблемость среднего кредита довольно высока (37 %), что определяет нестабильность этого класса, причем этот класс клиентов определяет всего 10 % от всего объема выданных кредитов, таблица 9.

Таблица 9. Итоговая информация о полученных классах клиентов

Показатель	Значение показателя в классе			
	1	2	3	4
Численность класса, чел.	7	6	7	18
Суммарная величина кредита, руб.	2 290 000	750 000	365 000	400 000
Доля от суммарного кредита, %	60.2	19.7	9.6	10.5
Минимальный кредит, руб.	290 000	100 000	40 000	10 000
Максимальный кредит, руб.	500 000	200 000	70 000	35 000
Средний размер кредита, руб.	327 143	125 000	52 143	22 222
Средний срок кредита, мес.	33.4	16	51.4	14.6
Стандартное отклонение, руб.	76 314	41 833	11 495	8 293
Коэффициент вариации размера кредита	0.23	0.33	0.22	0.37

Третий кластер довольно стабилен — коэффициент вариации составляет 22 %, сумма кредита небольшая — от 40 до 70 тыс. руб. и составляет 9,6 % от всех кредитов, но сроки кредита клиентов данного класса самые большие — в среднем 51,4 месяца. Второй класс немногочислен, имеет довольно большую величину среднего кредита, может характеризоваться как средне стабильный. Самый высокодоходный класс — первый, это около 20 % всех клиентов, которые берут значительные кредиты на не очень длительные сроки — в среднем 33 месяца, класс стабилен.

Для более глубокого анализа каждого класса клиентов для принятия решений о кредитной политике следует исследовать статистику по возврату кредитов. Можно для каждого класса заемщиков построить модели зависимости возврата кредита в сроки, установленные договорными обязательствами, от суммы кредита, срока кредита и других значимых факторов.

Была построена регрессионная зависимость уровня риска *risk* (ущерба от невозврата кредита) от факторов: суммы кредита (*value*), среднего дохода (*income*), срока кредита (*month*) для клиентов четвертого класса:

$$\text{risk} = -2841 + 36.9 \cdot \text{month} - 0.11 \cdot \text{value} + 0.27 \cdot \text{income}, \quad R^2 = 0.98; \quad DW = 2.9.$$

ст.ошибка	641	53	0.02	0.03
-----------	-----	----	------	------

Идентифицированная модель является статистически значимой и достаточно точной, стандартные ошибки параметров модели не превышают 1/4 от значения соответствующего параметра, коэффициент детерминации, равный 0,98, объясняет вариацию значений ущерба изменением включенных в модель факторов на 98 %, остаточная дисперсия составляет всего 2 %. Разработанная эконометрическая модель является основой для оценки и прогнозирования уровня риска, причиненного потенциальными клиентами, относящимися к четвертому классу.

Эконометрическое моделирование рисков можно реализовать с помощью одной регрессионной модели, охватывающей данные об объектах все групп. Для этого в модель вводятся фиктивные переменные $k_1, k_2, \dots, k_i, \dots, k_n$, где n — количество однородных групп объектов и $k_i = 1$ — для объектов i — группы и 0 — для всех остальных [Орлова, 2013, с. 72].

Заключение

Предложена методика многомерного анализа для формирования прогнозной оценки кредитного риска организаций, основанная на использовании информации кредитных историй, учитывающая объемы и сроки предоставляемых кредитов. Методика основана на применении методов кластеризации, метода дисперсионного анализа и анализа главных компонент и позволяет выявлять общие закономерности экономического поведения клиентов кредитной организации, сформировать совокупность дифференцированных требований к заемщикам отдельных групп с учетом их специфики, а также учесть склонность к риску лица, принимающего решения по объемам, срокам, процентам выдаваемого кредита.

Список литературы

- Банкова К. В. Использование скоринговых моделей для оценки кредитоспособности заемщиков в России / К. В. Банкова // Известия Академии управления: теория, стратегии, инновации. — 2011. — № 4. — С. 14–16.
- Бесулин А. М. Анализ программного обеспечения «sas credit scoring» для коммерческого банка / А. М. Бесулин // Инновационные информационные технологии. — 2013. — Т. 4, № 2. — С. 32–37.
- Глинкина Е. В. Кредитный скоринг как инструмент эффективной оценки кредитоспособности / Е. В. Глинкина // Финансы и кредит. — 2011. — № 16. — С. 43–47.

- Гузаиров М. Б., Орлова Е. В. Моделирование инновационных процессов региональных систем в условиях риска / М. Б. Гузаиров, Е. В. Орлова // Вестник Уфимского государственного авиационного технического университета. — 2012. — № 1. — С. 226–232.
- Исмагилова Л. А., Орлова Е. В. Модели и алгоритмы оперативного управления финансовыми ресурсами предприятия / Л. А. Исмагилова, Е. В. Орлова // Научно-технические ведомости Санкт-Петербургского государственного политехнического университета. Экономические науки. — 2012. — № 5. — С. 172–175.
- Исмагилова Л. А., Орлова Е. В. Нейросетевые технологии в экономике: сравнение с классическими методами / Л. А. Исмагилова, Е. В. Орлова // Нейрокомпьютеры: разработка и применение. — 2004. — № 9. — С. 49–57.
- Лебедев Е. А. Синтез скоринговой модели методом системно-когнитивного анализа / Е. А. Лебедев // Политематический сетевой электронный научный журнал Кубанского государственного аграрного университета. — 2007. — № 29. — С. 17–30.
- Макаренко Т. М. Сочетание сценарного прогнозирования с процедурами динамического ранжирования экспертов при оценке кредитного риска заемщика — физического лица в банке / Т. М. Макаренко // Вестник Ленинградского государственного университета им. А. С. Пушкина. — 2012. — Т. 6, № 3. — С. 56–63.
- Орлова Е. В. Экономико-математический инструментарий управления экономической системой в условиях неопределенности / Е. В. Орлова. — Уфа: УГАТУ. — 2012. — 172 с.
- Орлова Е. В. Эконометрическое моделирование и прогнозирование / Е. В. Орлова. — Уфа: УГАТУ. — 2013. — 250 с.
- Орлова Е. В. Имитационная модель управления стохастическими финансовыми потоками предприятия / Е. В. Орлова // Научно-технические ведомости Санкт-Петербургского государственного политехнического университета. Экономические науки. — 2012. — № 5. — С. 185–189.
- Оценка кредитоспособности физических лиц [Электронный ресурс]. — URL: http://www.basegroup.ru/download/presentations/what_if_evaluation.pdf
- Пятковский О. И. Скоринговая система оценки кредитоспособности физических лиц на основе гибридных экспертных систем / О. И. Пятковский, Д. В. Лепчугов, В. В. Бондаренко // Ползуновский альманах. — 2008. — № 8. — С. 127–129.
- Crone S.F., Finlay S. Instance sampling in credit scoring: An empirical study of sample size and balancing // International Journal of Forecasting. — 2012. — no. 28. — P. 224–238.
- Crook J.N., Edelman D.B., Thomas L.C. Recent developments in consumer credit risk assessment // European Journal of Operational Research. — 2007. — no. 183(3). — P. 1447–1465.
- Mircea G., Pirtea M., Neamtu M., Băzăvan S. Discriminant analysis in a credit scoring model // Recent Advances in Applied & Biomedical Informatics and Computational Engineering in Systems Applications. — 2011. — no. 2. — P. 56–69.
- Ong C., Huang J., Tzeng G. Building credit scoring models using genetic programming // Expert Systems with Applications. — 2005. — no. 9. — P. 41–47.