



DOI 10.22363/2313-2329-2017-25-2-242-254

УДК 004.413.4

ОЦЕНКА РИСКА БАНКРОТСТВА СУБЪЕКТОВ МАЛОГО ПРЕДПРИНИМАТЕЛЬСТВА НА ОСНОВЕ МЕТОДОВ МАШИННОГО ОБУЧЕНИЯ

И.В. Ариничев, И.В. Богдашев

Кубанский государственный университет
ул. Ставропольская, 149, Краснодар, Россия, 350040

В статье рассматривается методика построения алгоритма определения риска наступления банкротства предприятия с использованием методов машинного обучения. Преимуществом данной методики является использование не только количественных, но и качественных индикаторов финансовой устойчивости субъектов бизнеса, а также возможность исключения факторов, слабо влияющих на итоговый рейтинг. Предполагается, что разработанная математическая модель будет полезна представителям малого и среднего бизнеса и позволит получить объективную и точную картину о финансовом положении предприятия, текущих угрозах и риске банкротства.

Ключевые слова: риск наступления банкротства, финансовая устойчивость, машинное обучение, системы интеллектуальной аналитики, бинарное дерево, малый бизнес

Введение

Банкротство — это ситуация, когда фирма неспособна разрешить свои денежные обязательства, ведущая к правовой угрозе. Финансовые активы компаний продаются для погашения задолженности, что приводит к огромным потерям как для собственников, так и для инвесторов. В этой связи компаниям необходимо разрабатывать эффективные стратегии прогнозирования банкротства на более раннем этапе, чтобы избежать финансового кризиса. Заинтересованными лицами в определении финансовой устойчивости организации могут выступать не только собственники и инвесторы. Различные физические и юридические лица нуждаются в информации о «реальном положении дел в компании, с которой они связаны»: от поставщиков и партнеров, до сотрудников настоящих и потенциальных.

С институциональной точки зрения, чем большей информацией обладают все экономические агенты, тем меньше будет риск возникновения трансакционных издержек. Другими словами, благодаря полноте информации, полученной в результате применения простой и надежной методики оценивания финансовой устойчивости и отслеживания угрозы наступления банкротства, ключевые игроки рынка смогут проявлять высокий уровень доверия к данной организации: кли-

енты будут уверены в качестве продукции, сотрудники — в оплате труда и премиях за результат, поставщики, подрядчики и партнеры — в благонадежности и кредитоспособности организации в качестве заемщика, что, в свою очередь, положительно скажется на деловой репутации данной компании.

Обзор литературы

В настоящее время опубликовано значительное количество как отечественной, так и зарубежной литературы, посвященной данной теме. Среди моделей и методов оценки риска банкротства компаний, отраженных в данных работах и получивших широкое распространение в нашей стране и за рубежом, можно отметить модель Z-счета Альтмана (Altman E., 1968), относящуюся к классу моделей множественного дискриминантного анализа и позволяющую на основе показателей финансовой и бухгалтерской отчетности разделить предприятия на потенциальных банкротов и не банкротов (Altman E., 1968; 2005) KMV — модель Мертона (Merton R.C., 1974), используемая для расчета вероятности дефолта компаний, в основном для предприятий, которые торгуются на фондовых рынках; модель сокращенных форм (Jarrow R.A., Turnbull S., 1995) является обобщением модели Мертона и оценивает вероятность банкротства; макроэкономическая модель Уилсона на основе экзогенных показателей (Wilson T., 1997), которая легла в основу программного продукта CreditPortfolioView, предназначенного для оценки кредитного риска и разработанного консалтинговой группой McKinsey & Co; модель однофакторного дискриминантного анализа Бивера (Beaver W.H., 1966), где для оценки риска банкротства рассчитывается отношение чистого денежного потока компании к заемным средствам; модель Чессера из класса дискриминантных моделей (Chesser D., 1974), позволяющая предсказать не только дефолт как таковой, но и невыполнение клиентом условий договора и др.

Многие из перечисленных моделей, несмотря на солидный возраст выдержали проверку временем и сейчас используются в мировой практике. Большинство из них были проверены в глобальном масштабе и сосредоточены на крупных многонациональных компаниях, акции которых котируются на фондовых рынках и не адаптированы для малого бизнеса. Кроме этого, значительная часть представленных в обзоре моделей реализуется с помощью дискриминантного анализа, на основе которого принимается решение об оценке степени риска. Одним из существенных ограничений такого анализа служит невозможность учета качественных показателей, играющую не последнюю роль в оценке риска банкротства малых предприятий. К таким факторам можно отнести деловую репутацию компании, уровень менеджмента на предприятии, удовлетворенность потребителей и др. В то же время применение более продвинутых инструментов, таких как байесовские методы, нечеткая логика (Korol T., Korodi A., 2011), нейронные сети (Bredart X., 2014), логистическая регрессия (Ohlson J., 1980) и др. требуют специального, чаще всего дорогостоящего программного обеспечения, а также определенной квалификации менеджеров, что для малых предприятий недоступно.

В настоящей работе предлагается новая архитектура оценки риска компании, использующая методологию машинного обучения, относящуюся к системам ин-

теллектуальной аналитики. Основываясь на данных бухгалтерского учета, корпоративных характеристик, главным образом, структуры долга и капитала, а также ряда качественных показателей прогнозируется риск банкротства организации.

Методы машинного обучения используются для обнаружения и анализа скрытых структур в данных путем разработки и применения специальных алгоритмов (Воронцов К.В., 2014; Донской В.И., 2014; Загоруйко Н.Г., 1999). Общая схема процесса машинного обучения, а также формализация постановки задачи описана далее.

Методология машинного обучения оценки риска банкротства в организациях

Объектом x из множества всех объектов X в данной задаче будем считать юридическое лицо (малое предприятие), риск банкротства которого необходимо оценить. Чтобы задать объект, введем его признаковое описание $f_j: x_i \rightarrow D_j$, где $f_j(x_i)$ — значение j -го частного критерия для i -го предприятия x ; $i = 1, m, j = 1, n$. Например, в качестве группы признаков могут быть выбраны показатели, характеризующие бизнес-риски предприятия (срок функционирования бизнеса, зависимость от поставщиков и потребителей, наличие судебных разбирательств, конкурентная позиция заемщика на рынке и др.); другая группа содержит факторы характеризующие финансовые риски (денежное покрытие, ликвидность, обеспеченность собственным капиталом и др.); третья группа признаков может отвечать за анализ рисков кредитной истории и др. В этом случае упорядоченный набор $[f_1(x_i), f_2(x_i), \dots, f_n(x_i)]$ задает полное признаковое описание объекта. Для удобства восприятия, может быть записана матрица «объекты-признаки»

$$\begin{pmatrix} f_1(x_1) & \dots & f_1(x_m) \\ \dots & \dots & \dots \\ f_n(x_1) & \dots & f_n(x_m) \end{pmatrix},$$

столбцы которой содержат признаковое описание j -го объекта (малого предприятия). В зависимости от того, в какой шкале измеряется тот или иной признак (интервальная, порядковая, номинальная), множество D_j может быть конечным или бесконечным, с отношением порядка или нет. К примеру, если f_j — среднесписочная численность сотрудников в организации, то соответствующее множество будет конечным и упорядоченным $D_j = \{1, 2, 3, \dots, 100\}$. Для показателя, характеризующего долю рынка, логично в качестве D_j рассмотреть отрезок $[0; 100]$.

Для определения зависимой переменной, разделим предприятия на три класса с высоким, средним и низким риском банкротства. Присвоив каждому классу метку, получим множество всевозможных ответов для каждого предприятия соответственно $Y = \{-1; 0; 1\}$. Предложенная номинальная шкала для зависимой переменной не освобождена от критики и выбрана авторами для конкретизации проводимого исследования. В общем случае она без может быть определена произвольно. Например, задавая Y как множество натуральных чисел, фактически получаем финансовую модель оценки риска скорингового типа.

Чтобы успешно оценить риск произвольно выбранного малого предприятия методами машинного обучения, необходимо рассмотреть два этапа (Eksi I., 2011; Falahpour S., 2005):

1) этап обучения — это этап, на котором по обучающей выборке X^t (выборка признаковых описаний объектов, для них заранее известны ответы Y^t) на основе некоторого метода μ строится алгоритм классификации $a = \mu(X^t \times Y^t)$;

2) этап тестирования, когда построенный на первом этапе алгоритм a классифицирует объекты, которые не входили в обучающую выборку.

Если на втором этапе, алгоритм a демонстрирует допустимую ошибку (менее 15%), то он может использоваться в дальнейшем для диагностики риска банкротства произвольного малого предприятия. В противном случае производится переобучение алгоритма, путем подбора для него новых параметров. Чтобы измерить точность ответа алгоритма на одном отдельном объекте рассматривается функция потерь

$$L(a, x) = \begin{cases} 1, & a(x) \neq y(x), \\ 0, & a(x) = y(x), \end{cases}$$

которая равна 0, если алгоритм верно проклассифицировал объект и 1 в противном случае. Тогда величина ошибки в задаче классификации определяется с помощью функционала качества

$$Q(a, X^{test}) = \frac{1}{|X^{test}|} \sum_i L(a, x_i), \quad (1)$$

где X^{test} — тестовая выборка.

Учитывая специфику задачи и качественный характер некоторых признаков, был выбран логический алгоритм классификации малых предприятий, который представляет собой бинарное дерево (ациклический граф), каждой внутренней вершине $v \in V_{внутр}$ которого приписан предикат $\beta_v: X \rightarrow \{0, 1\}$, а каждой листовой (терминальной) вершине $v \in V_{лист}$ — метка класса $c_v \in Y$. В настоящей работе были использованы одномерные предикаты вида $\beta_v(x) = \{x_j \leq \theta_j\}$, где θ_j — некоторое пороговое значение j -го признака объекта.

Бинарный решающий алгоритм стартует из корневой вершины v_0 и вычисляет значение предиката β_{v_0} . Если оно равно нулю, то алгоритм переходит в левую дочернюю вершину, иначе — в правую, вычисляет значение предиката в новой вершине и делает переход или влево, или вправо. Процесс продолжается, пока не будет достигнута листовая вершина; алгоритм возвращает тот класс, который приписан этой вершине («риск высокий», «риск средний» и «риск низкий»).

В настоящей работе при построении решающего дерева была использована рекурсивная процедура Induction of Decision 3 (LearnID3), псевдокод которой представлен на рис. 1.

Наиболее важным шагом рассмотренного алгоритма является поиск предиката с максимальной информативностью. При определении предиката, который

бы максимально хорошо выделял бы какую-то группу классов от всех остальных классов, в зависимости от выбора критерия ветвления, возникает большое разнообразие методов построения решающих деревьев.

1.	ПРОЦЕДУРА $LearnID3(S \subseteq X)$	
2.	ЕСЛИ $\forall s \in S \Rightarrow y_s = c \in Y$	если все объекты подвыборки лежат в одном классе;
3.	ВЕРНУТЬ $v, c_v = c$	вернуть новую листовую вершину и присвоить ей метку класса;
4.	$\beta = \arg \max_{\beta} I(\beta, S)$	найти предикат с максимальной информативностью;
5.	$S_0 = \{x \in S : \beta(x) = 0\}$ $S_1 = \{x \in S : \beta(x) = 1\}$	разбить выборку на две подвыборки $S = S_0 \cup S_1$ по предикату максимальной информативности β ;
6.	ЕСЛИ $S_0 = \emptyset$ ИЛИ $S_1 = \emptyset$	если разбиения не произошло;
7.	ВЕРНУТЬ $v, c_v = c$	вернуть новую листовую вершину, присвоить метку того класса объектов которого больше в подвыборке;
8.	$v : \beta_v = \beta$ $L_v = LearnID3(S_0)$ $R_v = LearnID3(S_1)$	создать новую внутреннюю вершину и поместить в нее предикат; построить левое поддерево; построить правое поддерево;
9.	ВЕРНУТЬ v ;	вернуть решающее дерево;

Рис. 1. Рекурсивная процедура построения решающего бинарного дерева
[Fig. 1. Recursive procedure for constructing a solving binary tree]

В данной работе был использован наиболее распространенный и часто используемый критерий Джини, показывающий, сколько пар объектов, лежащих в одном и том же классе, одновременно попадут либо в левую, либо в правую дочернюю вершину дерева (значения предиката на них совпадают):

$$I(\beta, X) = \#\{(x_i, x_j) : y_i \neq y_j, \beta(x_i) = \beta(x_j)\}.$$

После того, как дерево построено, в целях упрощения его структуры и понижения сложности восприятия результатов моделирования можно провести его «стрижку». Существует ряд исследований, показывающих, что «стрижка» позволяет достичь лучшего качества по сравнению с ранним останом построения дерева на основе различных критериев. Тем не менее, на данный момент методы «стрижки» редко используются и не реализованы в большинстве библиотек для анализа данных (Воронцов К.В.). Другим подходом к оптимизации дерева служит сокращение размерности исходной задачи.

Чтобы понизить размерность задачи, сохранив при этом максимум информации в меньшем количестве переменных может быть использован метод главных компонент, широко использующийся в анализе данных [Айвазян С.А., Бухштабер В.М. и др., 1989; Gorban A.N., Kegl B., Wunsch D., Zinovyev A.Y., 2007; Jolliffe I.T., 2002]. Суть анализа главных компонент заключается в том, что все признаки-колонки преобразуются в компоненты, причем наибольшую информацию о разнообразии объектов несет первая компонента, вторая несет меньше информации, третья — еще меньше и т.д. Данный метод позволяет уменьшить количество признаков выбрав «самые изменчивые».

По своей математической сути данный метод представляет собой линейную замену переменных или переход из n -мерного пространства признаков $[f_1(x), f_2(x), \dots, f_n(x)]$ в k -мерное пространство главных компонент $[g_1(x), g_2(x), \dots, g_k(x)]$, где $k \leq n$:

$$\begin{aligned} g_1(x) &= v_{11}f_1(x) + v_{12}f_2(x) + \dots + v_{1n}f_n(x), \\ g_2(x) &= v_{21}f_1(x) + v_{22}f_2(x) + \dots + v_{2n}f_n(x), \\ &\dots\dots\dots \\ g_k(x) &= v_{k1}f_1(x) + v_{k2}f_2(x) + \dots + v_{kn}f_n(x). \end{aligned}$$

Веса v_{ij} подбираются так, чтобы выборочная дисперсия новой главной компоненты на каждом шаге была максимально возможной и выборочные корреляции $r(g_i(x), g_j(x)) = 0, i \neq j$. Известно, что если значения исходных признаков центрированы, то $g_j(x) = X^j \cdot v_j, |v_j|^2 = \lambda_j$, где v_j — собственные вектора, соответствующие собственным значениям λ_j матрицы $(X^j)^T X^j$, упорядоченным по убыванию $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_n \geq 0$.

Если данные, лежащие в матрице X^j действительно укладываются в пространстве меньшей размерности, то для определения эффективной размерности выборки используют следующее правило: оптимальное значение новой выборки — это наименьшее целое k , при котором

$$\frac{\lambda_{k+1} + \lambda_{k+2} + \dots + \lambda_n}{\lambda_1 + \lambda_2 + \dots + \lambda_n} > 0,5. \tag{2}$$

Указанное соотношение позволяет принять решение о том, какое количество последних главных компонент должно быть исключено из выборки без последствий, редуцировав тем самым размерность анализируемого пространства.

С учетом сказанного общая алгоритма оценки риска банкротства малых предприятий укладывается в определенную схему (рис. 2).

Численный эксперимент оценки риска банкротства на основе обученного алгоритма классификации

Для реализации рассмотренной методики оценки риска (рис. 2) было рассмотрено 48 малых предприятий, по каждому из которых зафиксированы значения по 23 признакам (табл. 1).

Учитывая, что разные признаки имеют разные единицы измерения, все они должны быть предварительно стандартизированы, например, по формуле

$$f_j^*(x_i) = \frac{f_j(x_i) - \overline{f_j(x_i)}}{sd[f_j(x_i)]},$$

где $sd[f_j(x_i)]$ — стандартное отклонение j -го признака; $j = 1, n; i = 1, m$. Применение процедуры стандартизации переменных необходимо перед использованием метода главных компонент, так как в противном случае будут выбраны не наиболее информативные признаки и их комбинации, а признаки с максимальными размахами, которые в свою очередь

определяются, в том числе, единицами измерения. Так, незначительное колебание прибыли, выраженной в рублях перекроет существенный разброс доли рынка, выраженный в процентах.

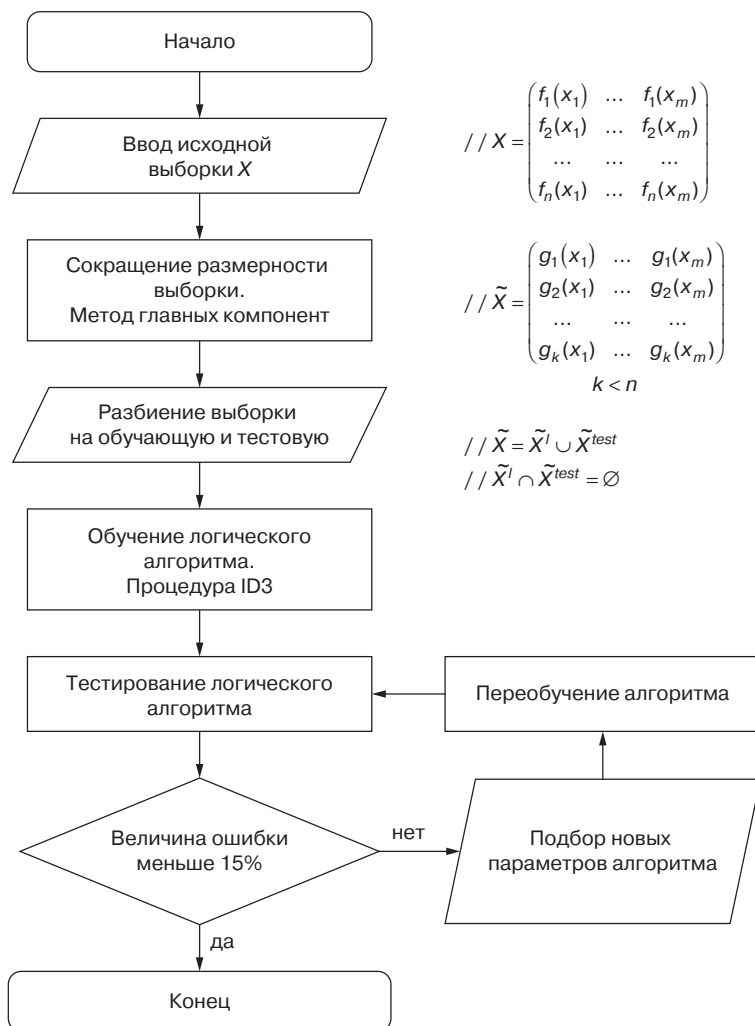


Рис. 2. Методика разработки алгоритма оценки риска банкротства малых предприятий
[Fig. 2. Methodology for the development of the algorithm for assessing the risk of bankruptcy of small enterprises]

Таблица 1

Факторы риска банкротства на малых предприятиях
[Risk Factors for Bankruptcy in Small Enterprises]

Группа факторов	Порядковый номер	Наименование	Комментарий
Деловые факторы	1	Business_period	Срок функционирования бизнеса — период в годах от начала регистрации компании до 2017 г.
	2	Legalization	Позиция на рынке, кредитная история, качество работы, степень легализации бизнеса

Окончание табл. 1

Группа факторов	Порядковый номер	Наименование	Комментарий
Деловые факторы	3	Management quality	Компетентность, уровень образования, сплоченность руководства и его эффективность
	4	Market_share	Доля рынка — доля компании на рынке
	5	Supplier_dependence	Зависимость от поставщиков — наличие крупнейшего поставщика (>30%)
	6	Customer_dependence	Зависимость от покупателей — наличие крупнейшего покупателя (>30%)
Показатели финансового риска	7	Net_assets	ВБ — долгосрочные и краткосрочные заемные средства + доходы будущих периодов
	8	Current_liquidity	Оборотные активы / Краткосрочные пассивы
	9	Financial_autonomy	Собственный капитал / Суммарный пассив (ВБ)
	10	Equity_to_debt_relation	Собственный капитал / Суммарные обязательства
	11	Sales_profitability	Прибыль от реализации / Выручка от реализации
	12	Activity_profitability	Чистая прибыль за последний отчетный квартал / выручка от продажи товаров, продукции, работ, услуг за последний отчетный квартал
	13	Cash_cover	Размер кредитных обязательств на текущую дату / среднемесячная выручка
	14	Interest_cover	Квартальная прибыль от продаж за последний отчетный квартал / сумма процентов по кредитам и займам, уплаченная за последний отчетный квартал
Активы предприятия	15	Fixed_assets	Внеоборотные активы — труднореализуемые активы (А4)
	16	Stock	Запасы — медленно реализуемые активы (А3)
	17	Receivables	Дебиторская задолженность — быстро реализуемые активы (А2)
	18	Shortterm_investments	Краткосрочные финансовые вложения — наиболее ликвидные активы (А1)
	19	Cash	Денежные средства — наиболее ликвидные активы (А1)
Пассивы предприятия	20	Equity	Собственный капитал — постоянные пассивы (П4)
	21	Longterm_borrowings	Долгосрочные кредиты и займы — долгосрочные пассивы (П3)
	22	Shortterm_borrowings	Краткосрочные кредиты и займы — краткосрочные пассивы (П2)
	23	Payables	Кредиторская задолженность — наиболее срочные обязательства (П1)

Реализация метода главных компонент была осуществлена с помощью функции **prcomp()** (*principal component*) в статистической среде вычислений **R**, простейший формат которой имеет вид: `prcomp(x, scale = TRUE)`, где x — матрица «объекты-признаки», $scale$ — логическое значение, указывающее следу-

ет ли нормировать переменные (Шипунов А.Б., Балдин Е.М.). В выводах значений данной функции содержатся, как сами значения главных компонент PC_j , так и веса v_{ij} через которые новые, синтетические переменные, выражаются через старые переменные $f_j(x)$. Из описания результатов применения метода (табл. 2) в среде R (первые 7 компонент), видно, что первая главная компонента $PC1$ улавливает 58% дисперсии совокупного исходного набора данных (Cumulative Proportion), первая и вторая главные компоненты $PC1, PC2$ объясняют 74% дисперсии в вариации рассматриваемой выборки. Первые четыре главные компоненты «ловят» 94% разброса исходных данных. Таким образом, в нескольких первых главных компонентах содержится практически вся информация, что и в 23 исходных признаках.

Таблица 2

Выводы метода главных компонент в среде статистических вычислений R
[Conclusions of the principal component method in the statistical computation environment R]

Показатель	$PC1$	$PC2$	$PC3$	$PC4$	$PC5$	$PC6$	$PC7$
Стандартное отклонение	4,383	3,617	2,094	1,587	0,983	0,814	0,801
Доля дисперсии	0,5836	0,1575	0,1268	0,0745	0,0251	0,0054	0,0017
Накопленная доля дисперсии	0,5836	0,7411	0,8679	0,9424	0,9775	0,9829	0,9846

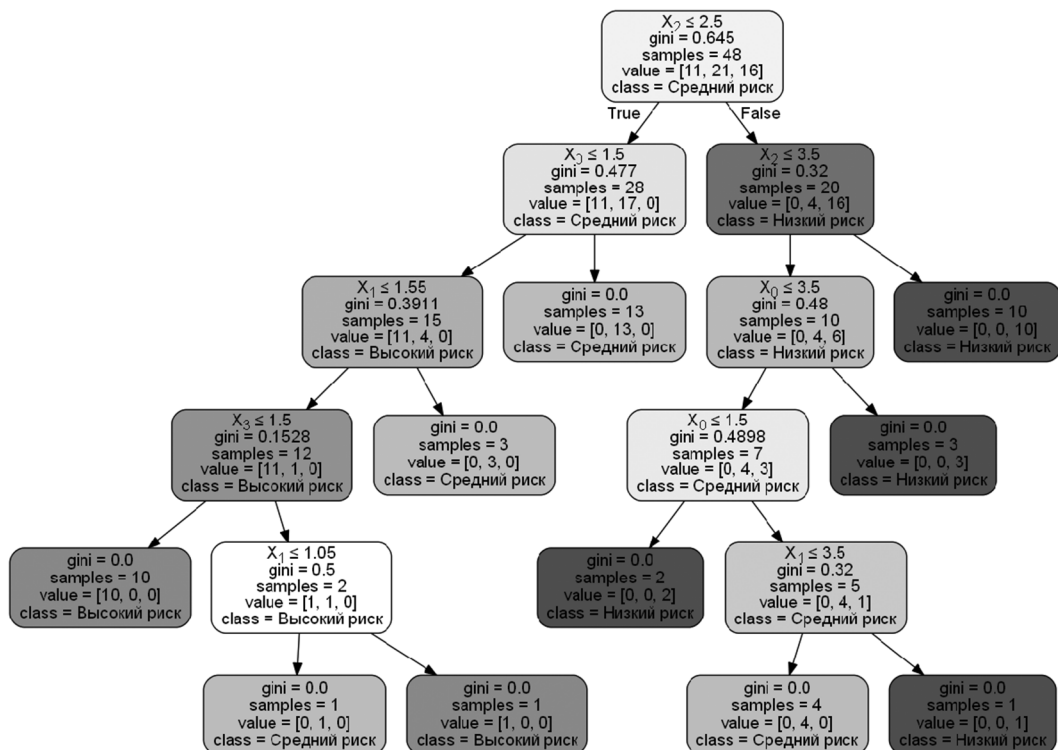


Рис. 3. Логический алгоритм классификации малых предприятий по уровням риска банкротства
[Fig. 3. Logical algorithm for classification of small enterprises by risk levels of bankruptcy]

С учетом результатов применения метода (табл. 2) и правила определения эффективного размера выборки (2), новая размерность признакового пространства составила $k = 4$. В результате применения процедуры *ID3* было построено бинарное решающее дерево (рис. 3), двигаясь из начальной вершины которого, в зависимости от значений главных компонент, любой объект (предприятие) может быть отнесен к одному из трех классов банкротства.

Функционал качества (1), рассчитанный для построенного алгоритма (см. рис. 3) по тестовой выборке, показал допустимые значения ошибки (<15%), что дает возможность использовать его в качестве инструмента поддержки принятия решений при выявлении субъектов малого предпринимательства, ориентированных на риск банкротства в целях предотвращения финансовых потерь.

Заключение

Два предварительных вывода вытекают в случае применения машинного обучения в секторе финансовых услуг. Во-первых, способность методов машинного обучения анализировать очень большие объемы данных, предлагая высокую степень детализации и глубину интеллектуального анализа, может значительно улучшить аналитические возможности в областях управления рисками и риска банкротства, в частности. Во-вторых, применение подходов машинного обучения в секторе финансовых услуг сильно зависит от качества и доступности исходных данных. Что более важно, прогнозная сила некоторых методов сильно зависит от сложности модели и объема исходных данных. Частично проблема решается путем сокращения размерности задачи.

СПИСОК ЛИТЕРАТУРЫ

- Айвазян С.А., Бухштабер В.М., Енюков И.С., Мешалкин Л.Д.* Прикладная статистика. Классификация и снижение размерности. М.: Финансы и статистика, 1989. 607 с.
- Воронцов К.В.* Математические методы обучения машин по прецедентам (теория обучения машин) [Электронный ресурс]. URL: <http://www.machinelearning.ru/wiki/images/6/6d/Voron-ML-1.pdf>
- Донской В.И.* Алгоритмические модели обучения классификации: обоснование, сравнение, выбор. Симферополь: ДИАЙПИ, 2014. 228 с.
- Загоруйко Н.Г.* Прикладные методы анализа данных и знаний. Новосибирск: ИМ СО РАН, 1999.
- Шупунов А.Б., Балдин Е.М.* Анализ данных с R [Электронный ресурс]. URL: <http://www.soc.univ.kiev.ua/sites/default/files/course/materials/r1.pdf>
- Altman E.* Financial ratios. Discriminant analysis, and the prediction of corporate bankruptcy // *Journal of Finance*, September. 1968.
- Altman E.* Corporate Financial Distress and Bankruptcy, 3rd edition. John Wiley and Sons, 2005.
- Beaver W.H.* Financial ratios as predictors of failure // *Journal of Accounting Research*. 1966. Vol. 4. Pp. 71–111.
- Bredart X.* Bankruptcy Prediction Model Using Neural Networks // *Accounting and Finance Research*. 2014. Vol. 3(2). Pp. 124–128.

- Chesser D. Predicting loan noncompliance // *The Journal of Commercial Bank Lending*. 1974. Pp. 28—38.
- Eksi I. Classification of firm failure with classification and regression trees // *International Research Journal of Finance and Economics*. 2011. 76. Pp. 113—120.
- Falahpour S., Raie R. Application of support vector machine to predict financial distress using financial ratios // *Journal of Accounting and Auditing Studies*. 2005. 53. Pp. 7—34.
- Gorban A.N., Kegl B., Wunsch D., Zinovyev A.Y. (Eds.) *Principal Manifolds for Data Visualisation and Dimension Reduction*, Series: Lecture Notes in Computational Science and Engineering 58, Springer, Berlin — Heidelberg — New York, 2007, XXIV, 340 p. 82.
- Jarrow R.A., Turnbull S. Pricing derivatives on financial securities subject to credit risk // *Journal of Finance*. 1995. Vol. 50. Pp. 53—85.
- Jolliffe I.T. *Principal Component Analysis*, Series: Springer Series in Statistics, 2nd ed., Springer, NY, 2002, XXIX, 487 p. 28.
- Korol T., Korodi A. An evaluation of effectiveness of fuzzy logic model in predicting the business bankruptcy // *Romanian Journal of Economic Forecasting*. 2011. № 3. Pp. 92—107.
- Merton R.C. On the pricing of corporate debt: the risk structure of interest rates // *Journal of Finance*. 1974. Vol. 29(2). Pp. 449—470.
- Ohlson J. Financial ratios and the probabilistic prediction of bankruptcy // *Journal of Accounting Research*. 1980. 18(1). Pp. 109—131.
- Wilson T. Portfolio Credit Risk: part I // *Risk Magazine*. 1997. 23. Pp. 111—117.

© Ариничев И.В., Богдашев И.В., 2017

История статьи:

Дата поступления в редакцию: 05 июня 2017

Дата принятия к печати: 30 июня 2017

Для цитирования:

Ариничев И.В., Богдашев И.В. Оценка риска банкротства субъектов малого предпринимательства на основе методов машинного обучения // *Вестник Российского университета дружбы народов. Серия: Экономика*. 2017. Т. 25. № 2. С. 242—254. DOI: 10.22363/2313-2329-2017-25-2-242-254

Сведения об авторах:

Ариничев Игорь Владимирович, кандидат экономических наук, доцент, доцент кафедры теоретической экономики экономического факультета Кубанского государственного университета. *Контактная информация:* e-mail: iarinichev@gmail.com

Богдашев Илья Владимирович, кандидат экономических наук, доцент, доцент кафедры теоретической экономики экономического факультета Кубанского государственного университета. *Контактная информация:* e-mail: ilbogdashev@gmail.com

ESTIMATION OF BANKRUPTCY RISK OF SMALL BUSINESS COMPANIES BASING METHODS OF MACHINE LEARNING

I.V. Arinichev, I.V. Bogdashev

Kuban State University
Stavropolskaya str., 149, Krasnodar, Russia, 350040

Abstract. The article deals with the methodology for constructing an algorithm for determining the probability of bankruptcy of an enterprise using machine learning methods. The advantage of this methodology is the use of not only quantitative, but also qualitative indicators of financial stability of business entities, as well as the possibility of excluding factors that have little effect on the final rating. It is assumed that the created mathematic model will be useful to representatives of small and medium-sized businesses and will provide an objective and precise picture of the financial situation of the enterprise, including current threats and the risk of bankruptcy.

Key words: risk of bankruptcy, financial stability, machine learning, systems of intellectual analytics, binary tree, small business

REFERENCES

- Aivazyan S.A., Bukhshtaber V.M., Enyukov I.S., Meshalkin L.D. (1989). *Prikladnaya statistika. Klassifikatsiya i snizhenie razmernosti*. M.: Finansy i statistika, 607 p. (In Russ)
- Vorontsov K.V. *Matematicheskie metody obucheniya mashin po pretsendentam (teoriya obucheniya mashin) [Elektronnyi resurs]*. Access Mode: <http://www.machinelearning.ru/wiki/images/6/6d/Voron-ML-1.pdf>
- Donskoi V.I. (2014). *Algoritmicheskie modeli obucheniya klassifikatsii: obosnovanie, sravnenie, vybor*. Simferopol': DIAIPI, 228 s. (In Russ)
- Zagoruiko N.G. (1999). *Prikladnye metody analiza dannykh i znanii*. Novosibirsk: IM SO RAN (In russ)
- Shipunov A.B., Baldin E.M. *Analiz dannykh s R [Elektronnyi resurs]*. Rezhim dostupa: <http://www.soc.univ.kiev.ua/sites/default/files/course/materials/r1.pdf>
- Altman E. (1968). Financial ratios. Discriminant analysis, and the prediction of corporate bankruptcy. *Journal of Finance*, September.
- Altman E. (2005). *Corporate Financial Distress and Bankruptcy*, 3rd edition. John Wiley and Sons.
- Beaver W.H. (1996). Financial ratios as predictors of failure. *Journal of Accounting Research*. Vol. 4. Pp. 71–111.
- Bredart X. (2014). Bankruptcy Prediction Model Using Neural Networks. *Accounting and Finance Research*. Vol. 3(2). Pp. 124–128.
- Chesser D. (1974). Predicting loan noncompliance. *The Journal of Commercial Bank Lending*. P. 28–38.
- Eksi I. (2011). Classification of firm failure with classification and regression trees. *International Research Journal of Finance and Economics*. Vol. 76. Pp. 113–120.
- Falahpour S., Raie R. (2005). Application of support vector machine to predict financial distress using financial ratios. *Journal of Accounting and Auditing Studies*. Vol. 53. P. 7–34.
- Gorban A.N., Kegl B., Wunsch D., Zinovyev A.Y. (Eds.) (2007). *Principal Manifolds for Data Visualisation and Dimension Reduction*, Series: Lecture Notes in Computational Science and Engineering 58, Springer, Berlin — Heidelberg — New York, XXIV, 340 P. 82.

- Jarrow R.A., Turnbull S. (1995). Pricing derivatives on financial securities subject to credit risk. *Journal of Finance*. Vol. 50. P. 53—85.
- Jolliffe I.T. (2002). *Principal Component Analysis*, Series: Springer Series in Statistics, 2nd ed., Springer, NY, XXIX, 487 p.
- Korol T., Korodi A. (2011). An evaluation of effectiveness of fuzzy logic model in predicting the business bankruptcy. *Romanian Journal of Economic Forecasting*. № 3. P. 92—107.
- Merton R.C. (1974). On the pricing of corporate debt: the risk structure of interest rates. *Journal of Finance*. Vol. 29(2). P. 449—470.
- Ohlson J. (1980). Financial ratios and the probabilistic prediction of bankruptcy. *Journal of Accounting Research*. Vol. 18(1). P. 109—131.
- Wilson T. (1997). Portfolio Credit Risk: part I. *Risk Magazine*. 23, pp. 111—117.

Article history:

Received: 05 June 2017

Revised: 20 June 2017

Accepted: 30 June 2017

For citation:

Arinichev I.V., Bogdashev I.V. (2017) Estimation of bankruptcy risk of small business companies basing methods of machine learning. *RUDN Journal of Economics*, 25 (2), 242—254. DOI: 10.22363/2313-2329-2017-25-2-242-254

Bio Note:

Arinichev I.V. Cand. Ec. Sci., Associate Professor, Department of Theoretical Economy, Kuban State University. *Contact information:* e-mail: iarinichev@gmail.com

Bogdashev I.V. Cand. Ec. Sci., Associate Professor, Department of Theoretical Economy, Kuban State University. *Contact information:* e-mail: ibogdashev@gmail.com