

УДК 004.896

UDC 004.896

05.00.00 Технические науки

Technical science

**ИНФОРМАЦИОННАЯ СИСТЕМА
ПОДДЕРЖКИ ПРИНЯТИЯ РЕШЕНИЯ В
СФЕРЕ ОЦЕНКИ ФИНАНСОВОГО
СОСТОЯНИЯ ПРЕДПРИЯТИЙ МАЛОГО И
СРЕДНЕГО БИЗНЕСА****DECISION MAKING SUPPORT
INFORMATION SYSTEM IN SPHERE OF
SMALL AND MEDIUM BUSINESS
COMPANIES SOLVENCY**

Пелипенко Екатерина Юрьевна
соискатель, преподаватель кафедры прикладной
математики факультета Компьютерных
технологий и прикладной математики.
тел. 8(961)5804321, e-mail: katrins_notes@mail.ru

Pelipenko Ekaterina Yurievna
postgraduate student, lecturer of the Applied
mathematics department of computer technology and
applied mathematics faculty
Tel. 8(961)5804321, e-mail: katrins_notes@mail.ru

Халафян Алексан Альбертович
д-р. техн. наук, профессор кафедры прикладной
математики факультета Компьютерных
технологий и прикладной математики
*ФГБОУ ВПО Кубанский государственный
университет, г. Краснодар, Россия.*

Khalafyan Aleksan Albertovich
Dr.Sci.Tech., professor of the Applied mathematics
department of computer technology and applied
mathematics faculty
Kuban State University, Krasnodar city, Russia.

Дестабилизация экономической ситуации в России в конце 2014 – начале 2015 годах в первую очередь отразилась на кредитовании малого и среднего бизнеса (МСБ). Одной из существенных причин высоких рисков в кредитовании МСБ, а значит, и высоких ставок кредитования, является отсутствие надежных информационных систем для оценки риска дефолта предприятий МСБ посредством всестороннего анализа результатов его финансово-экономической деятельности. Поэтому в настоящее время достоверная оценка платежеспособности предприятий МСБ – это фундаментальная научная проблема, которая в высшей степени актуальна для всех кредитных учреждений, так как напрямую связана с возможностью банкротства самой кредитной организации. В то же время, высокая конкуренция на рынке кредитования приводит к необходимости появления индивидуальных кредитных продуктов, учитывающих обоюдную выгоду заемщика и кредитора. Данная статья описывает методологию создания эффективных информационно-аналитических систем для оценки платежеспособности предприятий МСБ посредством мониторинга и анализа их финансово-экономического состояния. В основу разработанной системы легла интеграция вероятностно-статистических методов классификационного анализа (дискриминантного анализа, логистической регрессии, деревьев классификации), эвристических процедур (нейронных сетей) в интерактивную оболочку системы с использованием облачных технологий. Авторами была решена проблема малого объема данных, наличия выбросов и несоответствия нормальному закону распределения выборочной

Destabilization of the economic situation in Russia at the end of 2014 and in early 2015 has influenced small and medium businesses (SMB) landing at first. One of the most important reason of high lending risks and, as a result, high lending rates is absence of reliable information systems for assessment of SMB enterprise default according to total analysis of their financial activities. Thus nowadays the reliable assessment of SMB enterprises solvency is the fundamental scientific problem, which one is highly actual for each credit organization because the bankruptcy of a credit institution is depended on it. At the same time high competition at the landing market leads to necessity of individual credit conditions existing, which takes into account borrower's and lender's benefits. In the present work the creating of reliable information and analytical systems for assessment of SMB company default method is suggested. This one is based on integration of probabilistic and statistical classification analysis methods (discriminant analysis, logistic regression, and classification trees), heuristic procedures (neural network) and interactive shell of the system using cloud technology. By the authors, there was solved the problem of small data amount, exception anomalous values and discrepancy normal distribution of sample by the generation of enterprises financial activity model database

совокупности путем генерации модельной базы данных финансовой деятельности предприятий

Ключевые слова: МОДЕЛЬНАЯ БАЗА ДАННЫХ, МНОГОМЕРНЫЙ КЛАССИФИКАЦИОННЫЙ АНАЛИЗ, ФИНАНСОВОЕ СОСТОЯНИЕ ПРЕДПРИЯТИЯ, ДИСКРИМИНАНТНЫЙ АНАЛИЗ, ЛОГИСТИЧЕСКАЯ РЕГРЕССИЯ, ДЕРЕВЬЯ КЛАССИФИКАЦИИ, НЕЙРОННЫЕ СЕТИ, СТАВКА КРЕДИТОВАНИЯ, ВЕЛИЧИНА КРЕДИТА, ОБЛАЧНЫЕ ТЕХНОЛОГИИ

Keywords: MODEL DATABASE, MULTIVARIATE CLASSIFICATION ANALYSIS, FINANCIAL CONDITION OF COMPANY, DISCRIMINANT ANALYSIS, LOGISTIC REGRESSION, CLASSIFICATION TREES, NEURAL NETWORK, LONE RATE, AMOUNT OF LENDING, CLOUD TECHNOLOGY

Сложная экономическая ситуация в стране заставила кредиторов искать новые методы анализа кредитоспособности потенциальных заемщиков. Вопрос достоверной оценки коснулся, в первую очередь, предприятий малого и среднего бизнеса (МСБ) – наиболее рискованного сектора экономики с точки зрения экономической устойчивости [1].

Статистические пакеты прикладных программ существенно расширили возможности практического применения стохастических моделей для мониторинга финансового состояния предприятий. Они позволили успешно использовать достаточно трудоемкие в реализации многомерные вероятностно-статистические методы для построения сложных математических моделей классификации объектов различной природы по эмпирическим данным большого объема. Это кластерный анализ, дискриминантный анализ, деревья классификации, логит-регрессия и другие [2].

Интеграция вероятностно-статистических методов и эвристических процедур – нейронных сетей с методикой анализа финансового состояния предприятия, разработанной на базе теории нечеткой логики и теории матриц, легла в основу предлагаемой в статье методологии построения информационно-аналитических систем оценки риска дефолта предприятий МСБ.

На первом этапе исследований были выделены четыре группы 28 финансовых коэффициентов, характеризующих предприятия по рентабельности, ликвидности, платежеспособности и деловой активности.

К группе коэффициентов рентабельности относятся: R_1 – общая рентабельность, R_2 – рентабельность активов, R_3 – рентабельность собственного капитала, R_4 – рентабельность продукции, R_5 – рентабельность оборотных активов, R_6 – период окупаемости собственного капитала.

К коэффициентам ликвидности и платежеспособности относятся: L_1 – быстрый коэффициент ликвидности, L_2 – коэффициент абсолютной ликвидности, L_3 – коэффициент покрытия запасов, P_1 – текущий коэффициент ликвидности, P_2 – коэффициент обеспеченности собственными средствами. Расчет коэффициентов рентабельности производится на основе формул, указанных в таблице 2.

К коэффициентам деловой активности относятся: A_1 – коэффициент фондоотдачи, A_2 – коэффициент оборачиваемости активов, A_3 – коэффициент оборачиваемости собственного капитала, A_4 – коэффициент оборачиваемости кредиторской задолженности, A_5 – коэффициент оборачиваемости дебиторской задолженности, A_6 – коэффициент оборачиваемости запасов

К коэффициентам финансовой устойчивости относят: F_1 – коэффициент финансовой зависимости, F_2 – коэффициент автономии собственных средств, F_3 – коэффициент обеспеченности запасов собственными оборотными средствами, F_4 – индекс постоянного актива, F_5 – коэффициент концентрации собственного капитала, F_6 – коэффициент концентрации заемного капитала, F_7 – коэффициент маневренности собственного капитала, F_8 – коэффициент соотношения заемных и собственных средств, F_9 – коэффициент долгосрочного привлечения

заемных средств, F_{10} – коэффициент долгового покрытия активов, F_{11} – коэффициент покрытия активов.

Нами был оплачен доступ к данным финансовой деятельности предприятий МСБ Краснодарского края [3]. На основе формул [4] был произведен расчет значений вышеуказанных экономических коэффициентов для 410 предприятий МСБ на основе данных годовой бухгалтерской отчетности формы №1 «Бухгалтерский баланс» [5] и формы №2 «Отчет о прибылях и убытках» [6] за 2011-2013 года.

Из указанной выше совокупности финансовых показателей на основе анализа данных реальных предприятий были исключены из рассмотрения малоинформативные показатели: F_7 , F_9 и F_{10} . Проблема анализа данных в случае неограниченности сверху диапазонов значений коэффициентов была решена путем перехода к рассмотрению обратных показателей: $1/A_3$, $1/A_1$, $1/A_5$, $1/F_8$. Множественный корреляционный анализ позволил исключить из системы коэффициентов сильно коррелированные показатели. Дальнейшие исследования были осуществлены с использованием 21 независимого финансового коэффициента: R_3 , R_4 , R_5 , R_6 , L_1 , L_2 , L_3 , P_1 , P_2 , $1/A_1$, A_2 , $1/A_3$, A_4 , $1/A_5$, A_6 , F_1 , F_2 , F_3 , F_4 , $1/F_8$, F_{11} .

В виду определенной специфики математических моделей, лежащих в основе методов классификационного анализа, на различных исходных данных методы работают по-разному и с различной эффективностью [7]. Поэтому для увеличения достоверности оценки платежеспособности предприятия МСБ, а значит и верности принимаемого решения о кредитовании заемщика, целесообразно реализовать как можно большее количество методов классификационного анализа, которые требуют наличия обучающей выборки – совокупности данных по предприятиям, уровень риска неплатежеспособности которых изначально известен. На практике сбор такой информации является весьма трудоемким и доступ к

ней либо ограничен, либо вообще отсутствует. Для решения проблемы информационной неопределенности и создания обучающей выборки, лишенной недостатков эмпирических данных (малый объем, наличие выбросов и аномалий, несоответствие нормальному закону распределения признаков и т.д.) был предложен метод создания виртуальной модельной клиентской базы предприятий малого и среднего бизнеса [8].

В качестве критерия оценки финансового состояния предприятий использовали риск неплатежеспособности, т.е. риск невозврата денежных средств финансовому учреждению. Было выделено 5 уровней риска неплатежеспособности: очень высокий риск, высокий риск, средний риск, низкий риск, очень низкий риск. Были определены диапазоны изменения для каждого показателя по уровням риска неплатежеспособности согласно их нормативным значениям. На основе выбранной системы показателей сгенерирована модельная база 6000 виртуальных предприятий пяти уровней риска: очень высокого (1517 предприятий); высокого (572 предприятия); среднего (1687 предприятий); низкого (1537 предприятий); очень низкого (687 предприятий). Задача исследования качества построенной модельной базы заключалась в проверке однородности предприятий, принадлежащих одному классу и максимальной неоднородности предприятий различных классов посредством кластерного анализа. Было показано [8], что при числе кластеров равном 5, проведено наиболее успешное разбиение предприятий на кластеры – группы однородности предприятий: 98,4 % предприятий попали в кластеры, соответствующие пяти уровням риска.

Построенную модельную базу использовали в качестве обучающей выборки для решения задачи оценки финансового состояния реально

существующего предприятия МСБ при помощи методов дискриминантного анализа, логит-регрессии, деревьев классификации и нейронных сетей, реализованных в пакете прикладных программ Statistica. Ниже представлены результаты построенных моделей.

Проведенный дискриминантный анализ позволил ранжировать финансовые коэффициенты по степени их значимости в процессе классификации предприятий. Исследование итоговой таблицы анализа данных показало, что коэффициент финансовой зависимости (F_1), быстрый коэффициент ликвидности (L_1) и рентабельность продукции (R_4) являются главными переменными, так как они дают наиболее значительный вклад в общую дискриминацию. Толерантность, как показателя меры избыточности переменной в модели, для всех коэффициентов приняла достаточно большие значения, намного больше нижней границы (0,01), значит можно говорить об успешном включении всех коэффициентов в модель.

Четыре дискриминантные функции (канонических корня), построенные в процессе канонического анализа, оказались статистически значимы. Из результатов анализа был сделан вывод, что функция 1 ответственна за 79,31 % объясненной дисперсии, т.е. 79,31 % всей дискриминирующей мощности определяется этой функцией. Поэтому эта функция наиболее «важна». Функция 2 ответственна за 16,6 % объясненной дисперсии, функция 3 – за 2,77 %, а функция 4 – за 1,32 %. Средние значений для дискриминантных функций определили группы, лучше всего идентифицируемые конкретной дискриминантной функцией.

Были построены математические модели 5 функций классификации для рассматриваемых 5 групп уровней риска неплатежеспособности. Их

можно использовать для классификации произвольного предприятия МСБ. Предприятие приписывают той группе, для которой классификационная функция имеет наибольшее значение. Функции классификации для построенной модели дискриминантного анализа имеют вид:

$$Z_1(\text{«Очень высокий риск»}) = -4,4028 \cdot R_3 - 26,7150 \cdot R_4 - 4,8246 \cdot R_5 + 8,2722 \cdot L_1 + 2,8827 \cdot L_2 + 4,1881 \cdot P_1 + 2,3309 \cdot A_2 + 0,1384 \cdot A_4 + 3,9262 \cdot 1/A_5 + 0,3152 \cdot A_6 + 0,4394 \cdot F_1 + 4,6717 \cdot F_2 + 0,4782 \cdot F_3 - 13,8678 \cdot F_4 - 0,1873 \cdot R_6 + 10,7887 \cdot L_3 + 438,0445 \cdot 1/A_1 - 14,4161 \cdot 1/A_3 - 3,4334 \cdot 1/F_8 - 3,8719 \cdot F_{11} - 2,1385 \cdot P_2 - 19,5877;$$

$$Z_2(\text{«Высокий риск»}) = 2,7524 \cdot R_3 - 18,3192 \cdot R_4 - 4,6661 \cdot R_5 + 7,8225 \cdot L_1 + 0,5735 \cdot L_2 + 3,6651 \cdot P_1 + 1,9874 \cdot A_2 + 0,7286 \cdot A_4 + 4,5010 \cdot 1/A_5 + 1,7189 \cdot A_6 + 0,4544 \cdot F_1 + 16,1608 \cdot F_2 + 0,7868 \cdot F_3 - 0,8070 \cdot F_4 + 1,1117 \cdot R_6 + 17,0085 \cdot L_3 + 533,7231 \cdot 1/A_1 + 48,0572 \cdot 1/A_3 - 1,0619 \cdot 1/F_8 - 3,1474 \cdot F_{11} - 2,2732 \cdot P_2 - 38,0327;$$

$$Z_3(\text{«Средний риск»}) = 2,7375 \cdot R_3 - 12,9802 \cdot R_4 - 5,4501 \cdot R_5 + 8,9234 \cdot L_1 + 9,6643 \cdot L_2 + 4,9120 \cdot P_1 + 3,6014 \cdot A_2 + 2,9001 \cdot A_4 + 4,1383 \cdot 1/A_5 + 2,5409 \cdot A_6 + 0,4861 \cdot F_1 + 12,4259 \cdot F_2 + 2,3585 \cdot F_3 + 1,8791 \cdot F_4 + 1,0925 \cdot R_6 + 17,8306 \cdot L_3 + 492,8463 \cdot 1/A_1 + 36,5480 \cdot 1/A_3 + 0,6445 \cdot 1/F_8 - 1,7093 \cdot F_{11} + 0,7937 \cdot P_2 - 45,6868.$$

$$Z_4(\text{«Низкий риск»}) = 2,5578 \cdot R_3 - 2,8690 \cdot R_4 - 2,0106 \cdot R_5 + 8,9584 \cdot L_1 + 16,3080 \cdot L_2 + 8,1874 \cdot P_1 + 3,3027 \cdot A_2 + 1,9095 \cdot A_4 + 5,6893 \cdot 1/A_5 + 4,8473 \cdot A_6 + 0,2939 \cdot F_1 + 12,0667 \cdot F_2 + 2,9055 \cdot F_3 + 8,6932 \cdot F_4 + 2,0621 \cdot R_6 + 27,6148 \cdot L_3 + 459,6205 \cdot 1/A_1 + 24,9638 \cdot 1/A_3 + 2,3429 \cdot 1/F_8 - 3,4307 \cdot F_{11} + 1,6547 \cdot P_2 - 71,1227;$$

$$Z_5(\text{«Очень низкий риск»}) = 2,783 \cdot R_3 - 0,360 \cdot R_4 - 0,102 \cdot R_5 + 11,585 \cdot L_1 + 20,506 \cdot L_2 + 12,191 \cdot P_1 + 6,616 \cdot A_2 + 2,078 \cdot A_4 + 7,644 \cdot 1/A_5 + 7,375 \cdot A_6 +$$

$$7,375 \cdot F_1 + 10,460 \cdot F_2 + 2,697 \cdot F_3 + 13,640 \cdot F_4 + 2,230 \cdot R_6 + 38,323 \cdot L_3 + 584,406 \cdot 1/A_1 + 24,487 \cdot 1/A_3 + 2,327 \cdot 1/F_8 - 1,283 \cdot F_{11} + 2,429 \cdot P_2 - 143,173;$$

Матрица классификации, которая содержит информацию о количестве и проценте корректно классифицированных наблюдений в каждой группе, показала, что 96,17% от общего числа данных были классифицированы верно. Из оставшихся – 3,8% данных виртуальных предприятий классифицировано смежными к исходным классами и 0,03% – классифицировано ошибочно. Это также означает, что дискриминация была проведена успешно.

Результирующая таблица значений апостериорных вероятностей данных предприятий отображает вероятность принадлежности к каждой из 5 групп риска банкротства. На основе этих значений так же возможно производить классификацию произвольного предприятия МСБ по группам уровней риска неплатежеспособности. Клиент приписывается к той группе, которой соответствует наибольшая апостериорная вероятность.

Метод дерева классификации является менее строгим в отличие от традиционных методов анализа. Его использование не налагает ограничений на исходные данные, но результативность этого метода ниже. Основным преимуществом данного метода является его наглядность, простота и возможность использования без ограничений на исходные данные. Кроме того, метод дерева классификации является более гибким, что отражается в способности последовательно изучать эффект влияния отдельных переменных на дискриминацию и выполнять одномерное ветвление для анализа вклада отдельных переменных [9], что дает возможность работать с предикторными переменными различных типов. В нашем случае, метод дерева классификации используется, как дополнительный, хотя и менее точный. Если в модуле «Дискриминантный

анализ» пакета *Statistica* реализована возможность классификации нового наблюдения, то в модуле «Деревья классификации» такой возможности нет. Пользователь вынужден по полученным программой на обучающей выборке решающим правилам самостоятельно отнести новый объект к тому или иному классу, произведя соответствующие вычисления. Также программа не определяет вероятностных характеристик принадлежности каждого объекта к классам.

Процесс построения дерева классификации включает в себя три основных этапа: выбор критерия точности прогноза, выбор типа ветвления, определение «подходящего размера» дерева.

Критерий точности прогноза определяется наименьшей ценой – долей неверно классифицированных наблюдений, которые определяют элементы матрицы цен ошибок. 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. Для остальных групп – 0,25. Все внедиагональные элементы матрицы цен ошибок полагаются равными 1.

В модуле дерева классификации программы *Statistica* реализованы два алгоритма ветвления: *QUEST* и *CART*. Анализ метода *QUEST* (дискриминантного одномерного ветвления для категориальных и порядковых предикторов) показал более качественную классификацию, чем метод *CART*. Для него удалось построить дерево классификации (рисунок 1) с нулевой матрицей глобальной кросс-проверки, соответствующей наименьшему количеству ошибок классификации.

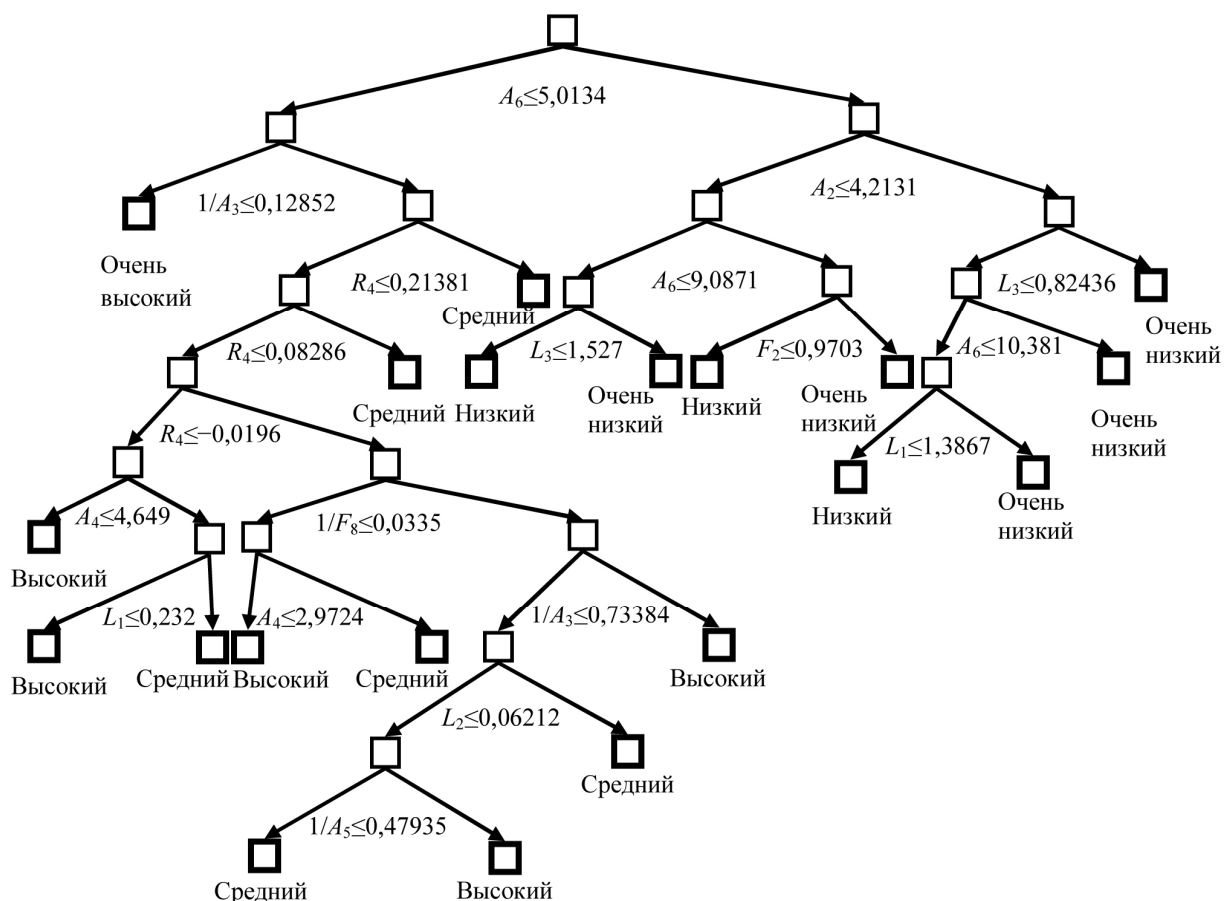


Рисунок 1. – Граф дерева классификации при выборе типа ветвления по методу *QUEST*

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). Более жирной границей обозначены терминальные вершины (листья) дерева, дальнейшее ветвление из которых невозможно. Построенное дерево имеет 19 ветвлений и 20 терминальных вершин.

Для выбора дерева классификации «подходящего размера» все значения параметров были выбраны согласно методу автоматического построения дерева Бримана [10]: 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 имеют высокий процент совпадений.

Числовые значения из таблицы решающих правил (условий ветвления) были использованы в качестве коэффициентов для составления линейных функций, представляющих собой условия перехода алгоритма анализа данных построенной математической модели. По значениям финансовых показателей предприятий алгоритм позволяет произвести классификацию предприятия по пяти группам уровня риска неплатежеспособности.

Классификация предприятий по пяти группам соответствует требованию [11], но отнесение предприятия к одному из двух групп уровней риска банкротства: высокий риск и низкий риск соответствует решению финансового учреждения дать кредит или отказать в кредите. Поэтому была так же построена вероятностно-статистическая модель логит-регрессии [12]:

$$y = \frac{e^z}{1 + e^z}, \quad (1)$$

где $z = b_0 + b_1 \cdot R_3 + b_2 \cdot R_4 + b_3 \cdot R_5 + b_4 \cdot L_1 + b_5 \cdot L_2 + b_6 \cdot P_1 + b_7 \cdot A_2 + b_8 \cdot A_4 + b_9 \cdot 1/A_5 + b_{10} \cdot A_6 + b_{11} \cdot F_1 + b_{12} \cdot F_2 + b_{13} \cdot F_3 + b_{14} \cdot F_4 + b_{15} \cdot R_6 + b_{16} \cdot L_3 + b_{17} \cdot 1/A_1 + b_{18} \cdot 1/A_3 + b_{19} \cdot 1/F_8 + b_{20} \cdot F_{11} + b_{21} \cdot P_2$.

Бинарную переменную «уровень риска», принимающую значение «низкий риск» или «высокий риск» получаем по правилу: если $y \in [0; 0,5]$, то предприятие относится к группе высокий риск неплатежеспособности, а если $y \in (0,5; 1]$, то – к группе низкий риск.

Для построения модели необходимо вычислить коэффициенты b_0, \dots, b_{21} . Для нахождения параметров уравнения регрессии использовали процедуру «Логит-регрессия» модуля «Нелинейное оценивание» пакета прикладных программ Statistica 10. Адекватность построенной модели

оценивали по значениям статистик «Итоговые потери», уровень значимости p критерия χ^2 . В модель включены показатели для которых уровень значимости p критерия Стьюдента не больше 0,5.

Адекватность модели также определяется количеством ошибок классификации наблюдений, осуществленной в соответствии с построенной моделью. Проценты безошибочных классификаций составили 98,6% (3604 предприятий) для предприятий со значением уровня риска банкротства «кредит» и 99,8% (2270 предприятий) – со значениями уровня риска банкротства «отказ». Общий процент верно классифицированных моделью предприятий составил 99,36%. Если эти проценты считать статистическими вероятностями, то вероятность, что предприятие вернет заемные денежные средства при его классификации в группу риска банкротства «кредит» равна 0,986. А вероятность возврата заемных денежных средств при классификации предприятия в группу «отказ» равна $(1 - 0,998)$, т.е. 0,002. Общая вероятность правильного прогноза равна – 0,9936.

Полученная логит-модель задается регрессионным уравнением (1), где $z = -5,658267 + 1,290519 \cdot R_3 + 9,766556 \cdot R_4 + 0,550896 \cdot R_5 - 2,824749 \cdot L_1 + 11,21751 \cdot L_2 + 1,186363 \cdot P_1 - 1,775693 \cdot A_2 + 1,641499 \cdot A_4 + 0,6794485 \cdot 1/A_5 + 0,9686623 \cdot A_6 - 0,05492874 \cdot F_1 - 2,250089 \cdot F_2 + 2,855861 \cdot F_3 + 3,364808 \cdot F_4 + 0,3246109 \cdot R_6 - 1,264199 \cdot L_3 - 117,1512 \cdot 1/A_1 - 7,610508 \cdot 1/A_3 + 2,096704 \cdot 1/F_8 + 0,1700767 \cdot F_{11} + 2,179195 \cdot P_2$.

На основе построенной логит-модели возможно производить анализ финансового состояния предприятия МСБ по значениям 21 финансового коэффициента.

В случае, когда не «работают» традиционные методы классификации, нейронные сети, по мнению многих исследователей [13], дают вполне приемлемый результат, так как работают с любыми данными и воспроизводят сложные нелинейные взаимосвязи. Поэтому нами была

построена математическая модель классификации предприятия МСБ по пяти уровням риска платежеспособности на основе метода нейронных сетей. Для построения нейронной сети использовали модуль *Neural Networks* программы *Statistica 6.1*, а в качестве обучающей выборки мы так же использовали модельную базу клиентов. Строилась сеть при критерии обучения – минимизация квадратичной ошибки с использованием градиента оценки, который вычислялся методом двойственности, а именно, методом обратного распространения ошибки [14], [15], [16]. В качестве функций активации и ошибки приняты логистическая и среднеквадратичная функции. Для определения минимального числа нейронов и структуры сети, обеспечивающих устойчивую работу системы, использовали процедуру контрастирования [17]. В результате, выбран тип сети – многослойный персептрон. Архитектура построенной сети представлена на рисунке 2.

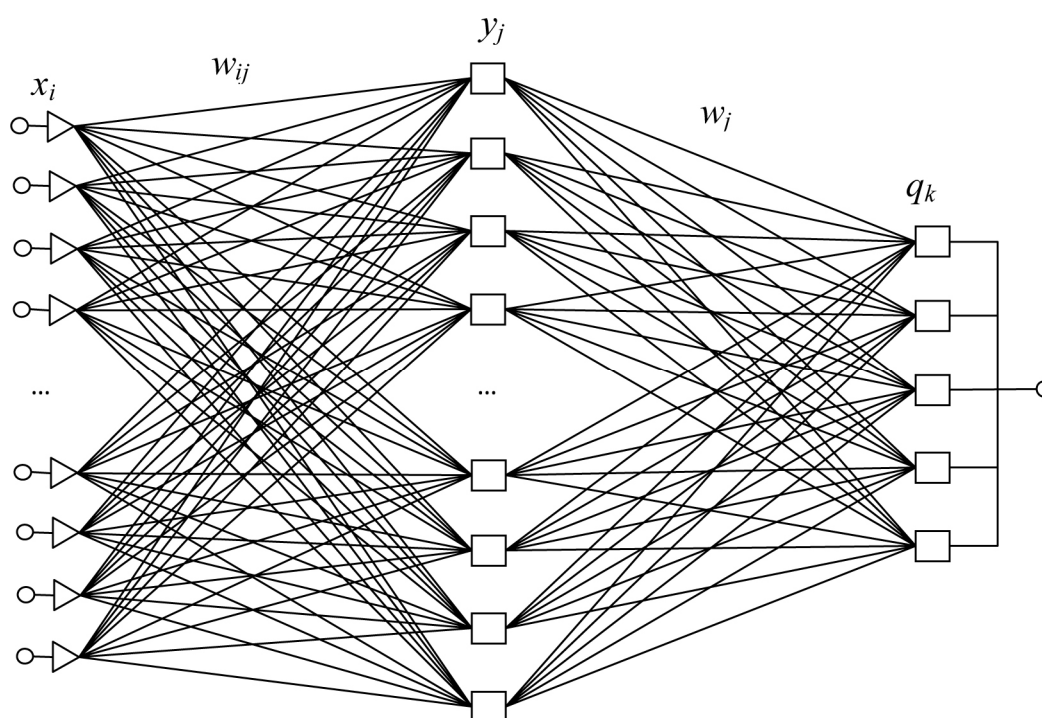


Рисунок 2. – Архитектура нейронной сети

Число слоев в сети 3, промежуточный слой y_j содержит 251 элемент ($j = 251$). На первом слое x_i 21 нейрон, соответствующих финансовым

показателям предприятия ($i = 21$): $R_3, R_4, R_5, R_6, L_1, L_2, L_3, P_1, P_2, 1/A_1, A_2, 1/A_3, A_4, 1/A_5, A_6, F_1, F_2, F_3, F_4, 1/F_8, F_{11}$; на последнем – 5, выходной элемент q_k , ($k = 5$). Связи между нейронами 1 и 2, 2 и 3 слоев на рисунке изображены линиями и обозначены w_{ij} и w_{jk} соответственно. Построение нейронной сети сопровождалось тестированием с помощью тестовой выборки составленной по данным 409 реальных предприятий МСБ.

Свойства модуля *Neural Networks* позволяют классифицировать наблюдение тестовой выборки с соответствующими уровнями доверия для различных классов, которые представляются уровнями активации выходных нейронов, относящимся к этим классам. Данные значения являются доверительными оценками для выходных классов. Их нельзя интерпретировать как оценки вероятностей. Если они выходят за пределы пороговых значений (класс не предсказан), то уровни активации отражают не вероятность принадлежности к классу, а меру ошибки классификации. Для всех предприятий тестовой выборки доверительные значения не больше нижнего предела или не меньше верхнего. Это значит, что уровни доверия отражают принадлежность предприятия к уровням риска неплатёжеспособности.

Для оценки меры производительности сети проанализированы значения статистики классификации. процент правильно классифицированных предприятий для каждого из пяти классов равен 100, что говорит о качестве построенной модели.

В результате выполнения работы нами были созданы в среде программирования *Statistica Visual Basic (SVB)* автономные программные модули классификации предприятий МСБ на основе построенных математических моделей дискриминантного анализа, логит-регрессии, деревьев классификации и нейронных сетей. Классификация производится по группам уровня риска неплатежеспособности на основе данных бухгалтерской отчетности формы №1 «Бухгалтерский баланс» и формы

№2 «Отчет о прибылях и убытках». С целью создания единой системы анализа финансового состояния предприятий разработана оболочка информационной системы поддержки принятия решения (ИСППР), в которую интегрированы созданные модули.

При создании ИСППР придерживались следующих положений:

- доступ к ИСППР должен быть простым и удобным, осуществляться независимо от характеристик ПК пользователя и места его расположения;

- решения должно приниматься ИСППР в максимально короткий промежуток времени, т.е. должен быть автоматизированы все процессы работы с данными;

- ИСППР должна быть простой и иметь интуитивно понятный интерфейс для неопытного пользователя ПК;

- ИСППР должна функционировать на базе недорогого программного обеспечения.

Для решения поставленных задач были применены облачные технологии, которые позволили наделять разработанную ИСППР указанными выше преимуществами. По модели развертывания (или охвата) ИСППР, как инфраструктура, представляет собой общественное облако, предназначенное для конкретной группы пользователей (в большей мере для кредитных и финансовых аналитиков), и имеющее ограниченный доступ. Моделью обслуживания созданной облачной технологии является модель SaaS (англ. Software as a service) – программное обеспечение как услуга.

Использовали структурную модель состава СППР, состоящую из четырех частей (рис. 3):

- подсистемы работы с данными (I);
- подсистемы формирования финансовых коэффициентов (II);
- подсистемы многомерного классификационного анализа (III).

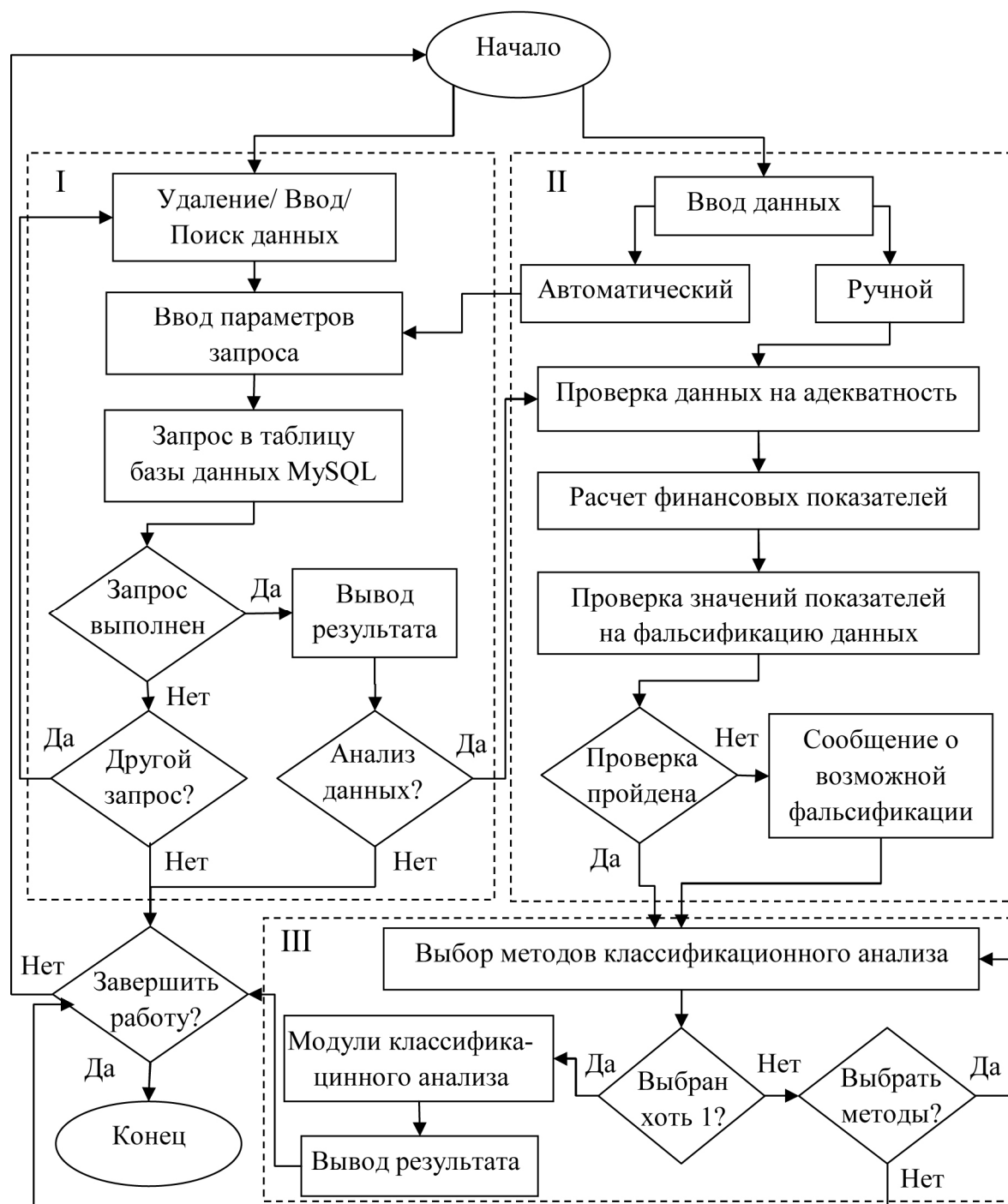


Рисунок 3. – Схема алгоритма ИСППР

Подсистема работы с данными предназначена для хранения информации о предприятии МСБ в СУБД MySQL. База данных позволяет структурировать, систематизировать, организовывать данные для их компьютерного хранения, обработки и представления. Подсистема с

помощью SQL запросов обеспечивает возможность ИСППР внесения, удаления и поиска данных о предприятии МСБ и его финансовой отчетности.

Подсистема формирования финансовых коэффициентов позволяет произвести расчет системы 21 коэффициента, приведенной выше. Анализ корректности введенных данных финансовой отчетности предприятия основан на взаимосвязях между статьями финансовых отчетностей форм 1 и 2. Возможная фальсификация данных идентифицируется ИСППР в случае выхода значений финансовых показателей в разряд аномальных.

Подсистема многомерного классификационного анализа дает возможность решения задачи классификации предприятия МСБ по пяти уровням риска неплатежеспособности при помощи методов дискриминантного анализа, деревьев классификации, логит-регрессии и нейронных сетей.

Таким образом нами была разработана ИСППР, которая поддерживает возможности ввода, удаления и поиска данных предприятия по заданным параметрам, анализа финансового состояния предприятия различными методами из базы или без сохранения в базе данных, а так же другие возможности. Работа системы основана на облачных технологиях, которые позволяют осуществлять доступ к ней в режиме реального времени через Интернет. Это расширяет возможности пользователей, а так же избавляет их от покупки лицензионного программного обеспечения. Система размещена на сайте лаборатории статистических исследований Кубанского государственного университета [18] и доступна для читателей журнала по логину *KubGAU* и паролю 5946721135.

Литература

1. Ф.Ф. Абдуллаев Малый и средний бизнес в современной России// Дискуссия, 2013, №3(33). Доступна по ссылке: <http://journal-discussion.ru/publication.php?id=144>.
2. Халафян А.А. Статистический анализ данных. STATISTICA 6.0: Учебник. 3-е изд., испр. и доп. – М.: Бинوم. 2007. 512 С.
3. База данных Контрагент. Доступна по ссылке: <http://www.k-agent.ru/> (дата обращения: 26.12.2014).

4. Официальный сайт системы комплексного раскрытия информации и новостей. Методика расчета финансовых показателей.
http://www.skrin.ru/info_doc/docs/mehod.html
5. КонсультантПлюс Бухгалтерский баланс (форма 1) (ОКУД 0710001),
http://www.consultant.ru/law/ref/forms/buh_formy/buhgalterskii_balans
6. КонсультантПлюс Отчет о прибылях и убытках (форма 2) (ОКУД 0710002), http://www.consultant.ru/law/ref/forms/buh_formy/otchet_o_pribylyah_i_ubytkah/
7. Халафян А.А. Обзор методов статистического анализа данных// Сайт лаборатории статистических исследований Кубанского государственного университета
<http://www.statlab.kubsu.ru/node/4>
8. Шевченко И.В. Создание Виртуальной клиентской базы для анализа кредитоспособности российских предприятий / И. В. Шевченко, А.А. Халафян, Е.Ю. Пелипенко // Финансы и кредит. 2010. №1. С. 13 – 18.
9. StatSoft. Электронный учебник по статистике. Деревья классификации
<http://www.statsoft.ru/home/textbook/modules/stclatre.html>
10. Breiman L.J. Random Forests–Random Features // UC Berkeley, Statistics Department, Technical Report. 1999. N 567.]:
11. О внесении изменений в положение Банка России от 26.03.2004 №254-П «О порядке формирования кредитными организациями резервов на возможные потери по ссудам, по ссудной и приравненной к ней задолженности»: указание Банка России от 25.10.2013 №3098-У: принят согласно протоколу заседания Совета директоров Банка России от 14.10.2013 № 22. Доступ из справ.-правовой системы «КонсультантПлюс» (дата обращения: 13.09.2014).
12. StatSoft. Электронный учебник по статистике. Нелинейное оценивание
<http://www.statsoft.ru/home/textbook/modules/stnonlin.html>
13. А.Н.Горбань Обобщенная аппроксимационная теорема и вычислительные возможности нейронных сетей //Сибирский журнал вычислительной математики. – 1998. – Т.1, № 1. – С. 12-24.
14. Haykin, S., 1994, “Neural Networks: A Comprehensive Foundation”, Prentice-Hall, Inc., Upper Saddle River, N.J.
15. Fausett, L. V., 1994, “Fundamentals of Neural Networks: Architectures, Algorithms, and Applications”, Prentice-Hall, Inc., Englewood Cliffs, N.J.
16. Hertz, J., et.al., 1991, “Introduction to The Theory of Neural Computations”, Addison-Wesley, Redwood City, California.
17. Миркес Е.М. Нейроинформатика: Учеб. пособие для студентов / Е.М. Миркес. Красноярск: ИПЦ КГТУ, 2002, 347 с. Рис. 58, табл. 59, библиогр. 379 наименований.
18. Система поддержки принятия решения в сфере оценки финансового состояния предприятий МСБ // Сайт лаборатории статистических исследований Кубанского государственного университета [Электронный ресурс]. Доступ по ссылке: http://statlab.kubsu.ru/sites/project_bank/project.php (дата обращения: 21.01.2015).

References

1. F.F. Abdullaev Maliy i sredniy biznes v sovremennoy Rossii// Diskussiya, 2013, №3(33). Dostupna po ssylke: <http://journal-discussion.ru/publication.php?id=144>.
2. Halafyan A.A. Statisticheskiy analiz dannyh. STATISTICA 6.0: Uchebnik. 3-e izd., ispr. i dop. – М.: Binom. 2007. 512 S.
3. Baza dannyh Kontragent. Dostupna po ssylke: <http://www.k-agent.ru/> (data obrashheniya: 26.12.2014).

4. Oficial'nyj sajt sistemy kompleksnogo raskrytiya informacii i novostej. Metodika rascheta finansovyh pokazatelej. http://www.skrin.ru/info_doc/docs/mehod.html
5. Konsul'tantPljus Buhgalterskij balans (forma 1) (OKUD 0710001), http://www.consultant.ru/law/ref/forms/buh_formy/buhgalterskii_balans
6. Konsul'tantPljus Otchet o pribyljah i ubytkah (forma 2) (OKUD 0710002), http://www.consultant.ru/law/ref/forms/buh_formy/otchet_o_pribylyah_i_ubytkah/
7. Halafyan A.A. Obzor metodov statisticheskogo analiza dannyh// Sajt laboratorii statisticheskikh issledovanij Kubanskogo gosudarstvennogo universiteta <http://www.statlab.kubsu.ru/node/4>
8. Shevchenko I.V. Sozdanie Virtual'noj klientskoj bazy dlja analiza kreditosposobnosti rossijskikh predpriyatij / I. V. Shevchenko, A.A. Halafyan, E.Ju. Pelipenko // Finansy i kredit. 2010. №1. S. 13 – 18.
9. StatSoft. Jelektronnyj uchebnik po statistike. Derev'ja klassifikacii <http://www.statsoft.ru/home/textbook/modules/stclatre.html>
10. Breiman L.J. Random Forests–Random Features // UC Berkeley, Statistics Department, Technical Report. 1999. N 567.]:
11. O vnesenii izmenenij v polozhenie Banka Rossii ot 26.03.2004 №254-P «O porjadke formirovaniya kreditnymi organizacijami rezervov na vozmozhnye poteri po ssudam, po ssudnoj i priravnennoj k nej zadolzhennosti»: ukazanie Banka Rossii ot 25.10.2013 №3098-U: prinjat soglasno protokolu zasedanija Soveta direktorov Banka Rossii ot 14.10.2013 № 22. Dostup iz sprav.-pravovoj sistemy «Konsul'tantPljus» (data obrashhenija: 13.09.2014).
12. StatSoft. Jelektronnyj uchebnik po statistike. Nelinejnoe ocenivanie <http://www.statsoft.ru/home/textbook/modules/stnonlin.html>
13. A.N.Gorban' Obobshhennaja approksimacionnaja teorema i vychislitel'nye vozmozhnosti nejronnyh setej //Sibirskij zhurnal vychislitel'noj matematiki. – 1998. – T.1, № 1. – S. 12-24.
14. Haykin, S., 1994, “Neural Networks: A Comprehensive Foundation”, Prentice-Hall, Inc., Upper Saddle River, N.J.
15. Fausett, L. V., 1994, “Fundamentals of Neural Networks: Architectures, Algorithms, and Applications”, Prentice-Hall, Inc., Englewood Cliffs, N.J.
16. Hertz, J., et.al., 1991, “Introduction to The Theory of Neural Computations”, Addison-Wesley, Redwood City, California.
17. Mirkes E.M. Nejroinformatika: Ucheb. posobie dlja studentov / E.M. Mirkes. Krasnojarsk: IPC KGTU, 2002, 347 s. Ris. 58, tabl. 59, bibliogr. 379 naimenovanij.
18. Sistema podderzhki prinjatija reshenija v sfere ocenki finansovogo sostojaniya predpriyatij MSB // Sajt laboratorii statisticheskikh issledovanij Kubanskogo gosudarstvennogo universiteta [Jelektronnyj resurs]. Dostup po ssylke: http://statlab.kubsu.ru/sites/project_bank/project.php (data obrashheniya: 21.01.2015).