

DOI: 10.26794/2587-5671-2022-26-6-131-155

УДК 336.6(045)

JEL G33, G17

Факторы риска банкротства российских компаний

А.А. Жуков^а, Е.Д. Никулин^б, Д.А. Щучкин^с^{а,б} Санкт-Петербургский государственный университет, Санкт-Петербург, Россия;^с Санкт-Петербургский государственный университет телекоммуникаций им. профессора М.А. Бонч-Бруевича, Санкт-Петербург, Россия

АННОТАЦИЯ

Банкротство российских компаний в текущей ситуации стало частым явлением. Определение факторов риска банкротства позволяет прогнозировать перспективу развития бизнеса. Авторы поставили задачу определить относительную силу влияния отдельных финансовых и нефинансовых факторов на вероятность банкротства компании. Для изучения факторов риска проведен анализ 3184 крупных российских компаний (с выручкой более 2 млрд руб. в год и численностью персонала более 250 человек) различных отраслей, функционировавших в период 2009–2020 гг. Общее количество наблюдений составляет 38 208. Для анализа были выбраны 30 показателей, разделенных на пять групп: показатели рентабельности, ликвидности, оборачиваемости, финансовой устойчивости и общие (нефинансовые) показатели. Для исследования выбран один из методов машинного обучения — метод случайного леса. Выборка исследования включает компании семи отраслей, среди которых: обрабатывающие производства, торговля, строительство, электроэнергетика, добыча полезных ископаемых, производство сельскохозяйственной продукции, водоснабжение, а также прочих отраслей, в которые отнесены компании образования, здравоохранения, сельского хозяйства, гостиницы. Анализ проводился как в совокупности по всей выборке без распределения по отраслям, так и по выборкам компаний, распределенным отдельно по сферам производства, торговли и услуг. По выборке в целом тестируемая модель в 86% случаев корректно предсказала возможность банкротства компаний за рассматриваемый промежуток времени. Данный результат подтвердил высокую эффективность использования методов машинного обучения (и, в частности, алгоритма случайного леса) применительно к решению задачи прогнозирования банкротства компаний. По полученным данным сделан вывод, что для производственных и торговых компаний наибольшее влияние на вероятность банкротства оказывают показатели рентабельности, а для компаний сферы услуг — финансовой устойчивости. Решение задачи определения факторов риска банкротства российских компаний должно привести к сокращению числа предприятий-банкротов, что, в свою очередь, будет способствовать оздоровлению и развитию национальной экономики.

Ключевые слова: корпоративные финансы; крупные компании; бизнес; финансовый анализ; финансовая устойчивость; прогнозирование банкротства; факторы риска банкротства; методы машинного обучения; случайный лес

Для цитирования: Жуков А.А., Никулин Е.Д., Щучкин Д.А. Факторы риска банкротства российских компаний. *Финансы: теория и практика*. 2022;26(6):131-155. DOI: 10.26794/2587-5671-2022-26-6-131-155

ORIGINAL PAPER

Bankruptcy Risk Factors of Russian Companies

А.А. Zhukov^а, Е.Д. Nikulin^б, D.A. Shchuchkin^с^{а,б} St. Petersburg State University, St. Petersburg, Russia;^с Bonch-Bruevich St. Petersburg State University of Telecommunications, St. Petersburg, Russia

ABSTRACT

The bankruptcy of Russian companies in the existing environment has become rather common. Determination of bankruptcy risk factors allows predicting the prospects for business development. The authors set the task to determine the relative influence of individual financial and non-financial factors on the probability of a company's bankruptcy. To study risk factors, the authors analyzed 3184 large Russian companies (with revenues of more than 2 billion rubles per year and more than 250 employees) of various industries operating from 2009 to 2020. The total number of observations is 38,208. For analysis, 30 factors were selected and divided into five groups: profitability, liquidity, turnover, financial stability and general (non-financial) factors. For the study, one of the machine learning methods was used — the random forest method. The sample consists of companies from seven industries, including manufacturing, retail, construction, electric power, mining, agricultural production, and water supply, as well as other industries, which include companies in education, healthcare, agriculture, and hospitality. The analysis was carried out both in aggregate for the entire sample without being distributed by industry, and for samples distributed by

manufacturing, retail, and service industries. In the sample as a whole, the tested model in 86% of cases correctly predicted the possibility of a company going bankrupt for the period under review. This result confirmed that machine learning methods (in particular, the random forest algorithm) are highly effective in solving the problem of bankruptcy prediction for a company. Based on the data obtained, the paper concludes that profitability factors have the most significant impact on the probability of bankruptcy for manufacturing and retail companies. For service companies, it is financial stability factors. Solving the problem of determining the bankruptcy risk factors of Russian companies will ensure a reduction in the number of bankrupt enterprises, which, in turn, will contribute to the recovery and development of the national economy.

Keywords: corporate finance; large companies; business; financial analysis; financial stability; bankruptcy prediction; bankruptcy risk factors; machine learning methods; random forest

For citation: Zhukov A.A., Nikulin E.D., Shchuchkin D.A. Bankruptcy risk factors of Russian companies. *Finance: Theory and Practice*. 2022;26(6):131-155. DOI: 10.26794/2587-5671-2022-26-6-131-155

ВВЕДЕНИЕ

Проблематика оценки результативности деятельности компаний, а также прогнозирования перспектив развития бизнеса остро стоит перед различными экономическими агентами. Внутренним и внешним пользователям учетной информации важно понимать, на какие показатели необходимо в первую очередь ориентироваться при финансовом анализе деятельности бизнеса с целью прогнозирования риска его банкротства¹. Для рынка РФ такой анализ особенно актуален, так как, по данным Центра макроэкономического анализа и краткосрочного прогнозирования (ЦМАКП), с 2016 г. наметилась устойчивая тенденция по сокращению числа действующих организаций в экономике РФ². По данным отчета Центрального банка РФ, максимальное с 2015 г. количество сообщений о банкротстве достигло в 2017 г., превысив 11 000³.

Одной из первых математических моделей по оценке риска банкротства компаний, получивших широкое распространение на практике, является модель Э. Альтмана, предложенная им в 1968 г. [1]. Первая версия модели Альтмана была сформулирована с помощью метода *множественного дискриминантного анализа* (*multiple discriminant analysis*) данных по 66 американским компаниям, часть из которых за период наблюдения обанкротилась, а часть — продол-

жила свое существование. Основным достоинством модели, определившим ее практическую значимость, является разработанный на ее основе интегральный индекс финансового состояния компании (*Z-score*), позволяющий ранжировать организации по степени риска банкротства. В более поздних работах проводился межстрановой анализ модели Э. Альтмана и определялась применимость модели для различных рынков [2, 3]. Многие модели прогнозирования банкротства компании, появившиеся в 1970-х — 1980-х гг., были методологически основаны на модели Альтмана (см. например, модель Р. Таффлера [4]).

Важным этапом развития моделей прогнозирования банкротства компании стал выход в 1980-е гг. работ М. Змиевского [5] и Дж. Ольсона [6]. В работе М. Змиевского использовалась пробит-модель (*probit model*), в то время как подход Ольсона был основан на логит-модели (*logit model*). Метод логистической регрессии до сих пор используется во многих отечественных и зарубежных работах по данной проблематике (см., например, [7–10]).

Современные исследования, посвященные банкротству компаний, основаны на продвинутых методах статистического анализа данных, прежде всего, методах *машинного обучения* (*machine learning*) [11–15]. Основной причиной распространения данных методов в работах по анализу риска банкротства компаний является то, что они позволяют преодолеть недостатки регрессионных моделей, выражающиеся, в частности, в снижении прогностической силы этих моделей в случае наличия нелинейного характера связи между переменными [7].

Макроэкономические особенности страновых рынков, а также различия в действующем законодательстве и учетных стандартах неизбежно требуют адаптации зарубежных моделей прогнозирования банкротства под специфику определенной страны и/или разработку оригинальных моделей на основе данных по компаниям этой страны. Систематический анализ литературы, выполненный в 2018 г. А.В. Казаковым и А.В. Колышкиным на основе более чем 40 оте-

¹ В соответствии со ст. 2 «Закона о банкротстве» РФ, несостоятельность (банкротство) — «это признанная арбитражным судом неспособность должника в полном объеме удовлетворить требования кредиторов по денежным обязательствам, о выплате выходных пособий и (или) об оплате труда лиц, работающих или работавших по трудовому договору, и (или) исполнить обязанность по уплате обязательных платежей», ФНС РФ. URL: <https://www.nalog.gov.ru/rn77/taxation/bankruptcy/> (дата обращения: 17.02.2021).

² Программа фундаментальных исследований НИУ ВШЭ. Банкротства юридических лиц в России: основные тенденции. 19.01.2021. URL: <http://www.forecast.ru/default.aspx> (дата обращения: 25.05.2021).

³ Доклад о денежно-кредитной политике Центрального банка РФ, июль 2021. URL: https://cbr.ru/analytics/dkp/ddcp/longread_3_35/page/ (дата обращения: 05.04.2021).

чественных моделей прогнозирования банкротства [16], продемонстрировал в среднем «низкий» уровень работ по данному направлению. К числу недостатков существующих исследований авторы, в частности, относят небольшой размер выборки, по которой оценивались параметры моделей, и отсутствие тестовой выборки. Следует также отметить небольшое число публикаций по российским компаниям, в которых бы использовались современные методы машинного обучения [7]. Таким образом, актуальность проведения исследований по проблематике прогнозирования банкротства российских компаний сохраняется.

Настоящее исследование отличается от существующих работ по российскому рынку, в которых риск банкротства анализируется при помощи методов машинного обучения, по следующим параметрам. *Во-первых*, в центре внимания большинства исследований подобного рода (см., например, [7, 17–19]) находится задача сравнения *предсказательной силы* (*predictive power*) различных моделей — как правило, регрессионных моделей и моделей, основанных на машинном обучении. В то же самое время настоящая работа нацелена на решение другой задачи — в ней определяется *относительная сила влияния* отдельных финансовых и нефинансовых *факторов* на вероятность банкротства компании. Проводимый с помощью метода случайного леса анализ позволяет проранжировать рассматриваемые показатели с точки зрения приоритетности при оценке риска банкротства компании. *Во-вторых*, существующие исследования либо не предполагают сравнительного межотраслевого анализа (т.е. анализ проводится по выборке в целом) [18, 19], либо изучают только одну отрасль [7, 17]. В настоящей работе проводится анализ как для выборки в целом, так и для отраслевых подвыборок — компаний сферы производства, торговли и услуг. *В-третьих*, важным результатом данного исследования является определение промежутков значений соответствующих показателей, характеризующихся разной степенью риска банкротства компании — а именно «высокой», «средней» и «низкой». Данный результат позволяет в удобном для практического использования виде классифицировать компании в зависимости от степени вероятности их банкротства.

Таким образом, целью настоящей работы является определение основных факторов риска банкротства российских компаний.

Авторы анализируют данные по крупным российским компаниям (с выручкой более 2 млрд руб. в год и численностью персонала более 250 человек), принадлежащим к различным отраслям. Период наблюдения составляет 12 лет (с 2009 по 2020 г.). В выборку вошли данные по 3184 компаниям, т.е. общее количество

наблюдений составляет 38 208. Основным методом анализа является один из методов машинного обучения — *случайный лес* (*random forest*). В качестве потенциальных факторов банкротства проанализированы более 30 показателей, в том числе и нефинансовые (например, численность персонала или форма собственности). Финансовые показатели (факторы), в свою очередь, разделены на четыре типа: ликвидности, рентабельности, финансовой устойчивости, оборачиваемости.

ПРИМЕНЕНИЕ МЕТОДОВ МАШИННОГО ОБУЧЕНИЯ В ИССЛЕДОВАНИЯХ ПО БАНКОТСТВУ РОССИЙСКИХ КОМПАНИЙ

В последние годы в исследованиях, посвященных анализу риска банкротства компании, все чаще используются продвинутое методы статистического анализа данных — прежде всего, методы машинного обучения. Активное их применение в исследованиях по российским компаниям относится к 2010-м гг.

Основным результатом большинства проведенных исследований является вывод о том, что методы машинного обучения позволяют получать более точные прогнозы вероятности банкротства компании по сравнению с традиционными методами регрессионного анализа данных. Так, например, в рамках исследования Б.Б. Демешева и А.С. Тихоновой, проведенного в 2014 г., сравнивалась прогностическая сила различных статистических методов, включающих как традиционные, так и более современные подходы: логит- и пробит-моделей; моделей, основанных на дискриминантном анализе; деревьев классификаций, а также случайного леса. Выборка ограничивалась данными, собранным по компаниям среднего и малого бизнеса за 2011–2012 гг. По результатам исследования оказалось, что наиболее точным инструментом прогнозирования банкротства компании является такой метод машинного обучения, как случайный лес [20].

В исследовании [17] также сделан вывод, что методы машинного обучения характеризуются более высокой точностью прогноза вероятности банкротства. Его авторы, проанализировав данные по 5120 российским компаниям из отрасли производства и распределения электроэнергии, газа и воды за 2009 г., утверждают, что наиболее точным инструментом прогнозирования банкротства является нейронная сеть, которая показала более высокий процент корректных предсказаний банкротств по сравнению с дискриминантным анализом и логит-моделью.

Аналогичный вывод был сделан в исследовании А.М. Карминского и Р.Н. Бурехина [7]. Авторы сравнивали между собой прогностическую силу большого числа моделей, включая различные модификации

нейронной сети, деревьев классификации и случайного леса. На основе анализа данных по компаниям строительной отрасли России за 2011–2017 гг. было показано, что на рассматриваемой выборке наилучшими результатами обладает искусственная нейронная сеть с одним скрытым слоем и четырьмя нейронами.

В целом исследования показывают, что методы машинного обучения в целом и нейронные сети в частности могут значительно усилить анализ риска банкротства компании, однако эффективность работы этих методов зависит от качества и доступности исходных данных [21–23].

Ряд научных работ посвящен разработке новых алгоритмов/моделей для прогнозирования финансовой несостоятельности российских компаний на основе методов машинного обучения. Например, в исследовании [23] был проведен анализ финансовой устойчивости компаний обрабатывающих отраслей с помощью нейросетевой модели. И.В. Ариничев и И.В. Богдашев, используя деревья бинарной классификации, построили алгоритм для определения риска наступления банкротства компаний малого бизнеса [24]. В исследовании Е.Ю. Макеевой и И.В. Аршавского [22] сделан акцент на роли качественной информации при анализе риска банкротства компании. Данная информация напрямую не отражена в финансовых показателях компании, однако при этом присутствует в корпоративной годовой отчетности. На основе применения методов семантического анализа отчетности компании и ансамбля искусственных нейронных сетей сделан вывод о повышении предсказательной способности модели при включении в нее качественной информации.

Отдельно отметим исследование Е.А. Федоровой, С.О. Мусиенко и Ф.Ю. Федорова [19]. Это одна из немногих работ, в которой для российских компаний малого и среднего бизнеса с помощью методов машинного обучения уточняются нормативные значения показателей, фигурирующих в законодательстве по банкротству, а также в зарубежных исследованиях риска банкротства организаций. Главным образом, рассматривались показатели ликвидности и финансовой устойчивости. В работе использовались методы дерева решений (*decision tree*), случайного леса, бэггинга (*bagging*) и бустинга (*boosting*). Наилучшую прогностическую способность показал метод случайного леса.

Настоящее исследование также посвящено выявлению наиболее значимых факторов банкротства российских компаний и определению их критических значений. В отличие от [19], объектом исследования выступают российские компании крупного бизнеса различных отраслей. Рассматривается ши-

рокий набор как финансовых, так и нефинансовых факторов потенциального банкротства компании. В соответствии с целью статьи в качестве основного инструментария исследования выступает метод случайного леса.

МЕТОД СЛУЧАЙНОГО ЛЕСА

Случайный лес (*random forest*) — это алгоритм классификации, состоящий из множества деревьев решений. Данный алгоритм используется для классификации, кластеризации, селекции признаков и поиска аномалий. Также случайный лес определяет важность факторов, т.е. их влияние на процесс классификации. Таким образом, появляется возможность расположить факторы в порядке приоритетности [25].

По сравнению с регрессионными моделями, которые относительно чувствительны к выбросам, случайный лес является более устойчивым к этой проблеме. Еще одним преимуществом метода случайного леса является его меньшая, по сравнению с методом нейронных сетей, подверженность переобучению [7].

Базовой единицей алгоритма случайного леса является дерево решений. Оно представляет собой серию вопросов о входных данных, которые имеют ответы «да» или «нет». Вопросы задаются до тех пор, пока дерево не приходит к какому-либо решению. Случайный лес является алгоритмом машинного обучения, а его основное преимущество состоит в умении обрабатывать новые данные, которые ранее никогда не встречались. Поэтому для того, чтобы модель не переобучалась, используется не одно дерево решений, а случайный лес, состоящий из них. На *рис. 1* показан условный пример случайного леса. Желтым цветом отмечены решения, к которым приходят деревья.

Общий принцип алгоритма случайного леса заключается в том, что исследователь строит множество деревьев решений (деревьев классификации). Каждое дерево в случайном лесу возвращает прогноз результата, и результат с наибольшим количеством голосов становится прогнозом леса. Ключевой чертой является низкая корреляция между деревьями. Такой эффект обусловлен тем, что разные деревья отвечают на разные вопросы «да» или «нет», тем самым приходя к результату. Часть деревьев могут вести к ложному результату, а часть к корректному. Таким образом, деревья защищают друг друга от индивидуальных ошибок. Большое количество слабокоррелированных деревьев, анализирующих информацию совместно, будет превосходить любую из их отдельных составляющих [26].

ЭТАПЫ ЭМПИРИЧЕСКОГО ИССЛЕДОВАНИЯ

В соответствии с целью исследования были выполнены следующие этапы работы.

На *первом этапе* был определен предварительный набор количественных и качественных переменных для анализа. В список исследуемых факторов банкротства компании были включены переменные из различных отечественных и зарубежных исследований (см., например, [1, 6, 27–29]). Все переменные были разделены на пять категорий: *показатели рентабельности, ликвидности, оборачиваемости, финансовой устойчивости* и *общие (нефинансовые) показатели*. Более того, чтобы избежать «эффекта каннибализации», ряд переменных был удален, так как при анализе описательной статистики выяснилось, что они имели высокую корреляцию с другими переменными.

Итоговый состав переменных для алгоритма случайного леса с учетом исключения части переменных представлен в *табл. 1*.

На *втором этапе* с помощью метода случайного леса для компаний выборки были определены наиболее значимые (важные) факторы с точки зрения влияния на вероятность банкротства. Это было сделано с помощью библиотеки Scikit-learn, алгоритмы которой позволяют рассчитать относительную важность каждого фактора, т.е. вклад каждого показателя в прогноз. Для оценки значимости факторов случайный лес использует критерий Джини. Полученные значения оценок важности отдельных показателей были нормализованы таким образом, чтобы сумма всех оценок равнялась 1. Чем выше значение оценки значимости определенного фактора, тем больше его вклад в прогноз банкротства компании. Важность признаков была определена как для выборки в целом, так и для отраслевых подвыборок — для производственных компаний, компаний сферы услуг и торговых компаний.

На *третьем этапе* были определены пороговые значения показателей, оказавшихся по итогам анализа отдельных отраслей наиболее значимыми для оценки риска банкротства компании. Для этого использовался визуальный анализ модели классификации, построенной алгоритмом случайного леса, при помощи библиотеки PDPbox (Python). Данный результат позволяет по значению отдельных показателей разделить компании на три группы — с высоким, средним и низким риском банкротства.

Отметим, что в настоящей работе осуществлялось разделение выборки на *тестовую* и *тренировочную* подвыборки. С этой целью использовался модуль `train_test_split` библиотеки Scikit-learn. С помощью этого модуля были выделены из всех данных 20% тестовых значений и 80% тренировочных. Для того

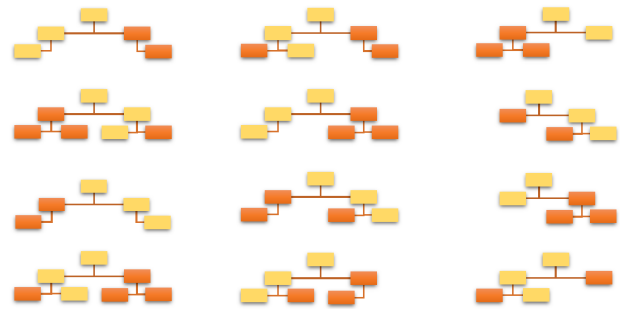


Рис. 1 / Fig. 1. Образец работы случайного леса / Random forest algorithm prototype

Источник / Source: составлено авторами / compiled by the authors.

чтобы тестовые значения затрагивали разные годы, функция `train_test_split` отбирает 20% данных для тестовой выборки *в случайном порядке*. Подобное разделение выборки приводит к тому, что модель обучается на одной части данных, а тестируется на другой. Это позволяет избежать эффекта переобучения модели и убедиться в том, что построенная модель не просто «запомнила ответы» в тренировочной выборке, но и с высокой точностью предсказывает результаты на тестовой части выборки.

ВЫБОРКА

В выборку были включены крупные российские компании. К среднему бизнесу согласно Единому реестру малого и среднего бизнеса относятся компании с численностью персонала до 250 человек и с годовым доходом до 2 млрд руб.⁴ Таким образом, в данном исследовании анализировались компании с выручкой свыше 2 млрд руб. в год и численностью персонала свыше 250 человек. Для сбора данных за 12 лет (с 2009 по 2020 г.) использовалась система СПАРК⁵. В выборку вошли данные по 3184 компаниям. Общее число наблюдений составило 38 208. Согласно собранным данным 68% компаний, вошедших в выборку, продолжали осуществлять операционную деятельность на протяжении всего периода наблюдения, а 32% — обанкротились.

Выборка исследования включает компании 7 отраслей: обрабатывающие производства, торговля, строительство, электроэнергетика, добыча полезных ископаемых, сельскохозяйственная продукция, водоснабжение, а также прочие отрасли, в которые попали компании образования, здравоохранения, сельского

⁴ Единый реестр малого и среднего бизнеса. URL: <https://ofd.nalog.ru/about.html?section=conditions> (дата обращения: 11.03.2021).

⁵ База данных СПАРК. URL: <http://www.spark-interfax.ru/> (дата обращения: 15.04.2021).

Таблица 1 / Table 1

Состав переменных исследования / Variables used in the study

Номер / Number	Тип показателя / Indicator type	Формула / Formula	Тип переменной / Variable type
1	Рентабельность	$\frac{EBITDA_{i,p}}{A_{i,p}},$ где $EBITDA_{i,p}$ – прибыль до налогов, процентов и амортизации компании i в год p; $A_{i,p}$ – общие активы компании i в год p	Количественная
2	Рентабельность	$\frac{Retained\ earnings_{i,p}}{A_{i,p}},$ где $Retained\ earnings_{i,p}$ – нераспределенная прибыль компании i в год p; $A_{i,p}$ – общие активы компании i в год p	Количественная
3	Рентабельность	$\frac{EBIT_{i,p}}{Rev_{i,p}},$ где $EBIT_{i,p}$ – прибыль до налогов и процентов компании i в год p; $Rev_{i,p}$ – общая выручка компании i в год p	Количественная
4	Рентабельность	$ROE_{i,p} = \frac{NI_{i,p}}{Equity_{i,p}},$ где $ROE_{i,p}$ – рентабельность капитала компании i в год p; $NI_{i,p}$ – чистая прибыль компании i в год p; $Equity_{i,p}$ – капитал компании i в год p	Количественная
5	Ликвидность	$\frac{CA_{i,p}}{STD_{i,p}},$ где $CA_{i,p}$ – текущие активы компании i в год p; $STD_{i,p}$ – краткосрочные обязательства компании i в год p	Количественная
6	Ликвидность	$\frac{Cash_{i,p}}{A_{i,p}},$ где $Cash_{i,p}$ – денежные средства компании i в год p; $A_{i,p}$ – общие активы компании i в год p	Количественная
7	Финансовая устойчивость	$\frac{FA_{i,p}}{CA_{i,p}},$ где $FA_{i,p}$ – внеоборотные активы компании i в год p; $CA_{i,p}$ – текущие активы компании i в год p	Количественная
8	Финансовая устойчивость	$\frac{EBITDA_{i,p}}{Int_{i,p}},$ где $EBITDA_{i,p}$ – прибыль до налогов, процентов и амортизации компании i в год p; $Int_{i,p}$ – проценты к уплате компании i в год p	Количественная

Продолжение таблицы 1 / Table 1 (continued)

Номер / Number	Тип показателя / Indicator type	Формула / Formula	Тип переменной / Variable type
9	Финансовая устойчивость	$\frac{WC_{i,p}}{LTD_{i,p}},$ где $WC_{i,p}$ — рабочий капитал компании i в год p ; $LTD_{i,p}$ — долгосрочные обязательства компании i в год p	Количественная
10	Финансовая устойчивость	$\frac{STD + LTD_{i,p}}{Equity_{i,p}},$ где $STD + LTD_{i,p}$ — общие обязательства компании i в год p ; $Equity_{i,p}$ — капитал компании i в год p	Количественная
11	Оборачиваемость	$\frac{Rev_{i,p}}{\overline{A_{i,p}}},$ где $Rev_{i,p}$ — выручка компании i в год p ; $\overline{A_{i,p}}$ — среднегодовые активы компании i в год p	Количественная
12	Оборачиваемость	$Inventories\ turnover_{i,p} = \frac{COGS_{i,p}}{Average\ inventories_{i,p}},$ где $COGS_{i,p}$ — себестоимость компании i в год p ; $Average\ inventories_{i,p}$ — среднегодовые запасы компании i в год p	Количественная
13	Оборачиваемость	$Fixed\ assets\ turnover_{i,p} = \frac{Revenue_{i,p}}{Average\ fixed\ assets_{i,p}},$ где $Revenue_{i,p}$ — выручка компании i в год p ; $Average\ fixed\ assets_{i,p}$ — среднегодовые внеоборотные активы компании i в год p	Количественная
14	Оборачиваемость	$Accounts\ receivable\ turnover\ period_{i,p} = \frac{Revenue_{i,p}}{Average\ receivables_{i,p}},$ где $Revenue_{i,p}$ — выручка компании i в год p ; $Average\ receivables_{i,p}$ — среднегодовая дебиторская задолженность компании i в год p	Количественная
15	Оборачиваемость	$\frac{Rev_{i,p}}{WC_{i,p}},$ где $Rev_{i,p}$ — общая выручка компании i в год p ; $WC_{i,p}$ — рабочий капитал компании i в год p	Количественная
16	Общий	$\ln\left(\frac{A_{i,p}}{GPN\ price\ index_p}\right),$ где $A_{i,p}$ — активы компании i в год p ; $GPN\ price\ index_p$ — индекс-дефлятор ВВП в год p	Количественная

Окончание таблицы 1 / Table 1 (continued)

Номер / Number	Тип показателя / Indicator type	Формула / Formula	Тип переменной / Variable type
17	Общий	$bin_{1,i,p}$ (принимает значение 1, если $NI_{i,p} + NI_{i-1,p} < 0$ и значение 0 в противном случае, где $NI_{i,p}$ – чистая прибыль компании i в год p)	Бинарная
18	Общий	$bin_{2,i,p}$ (принимает значение 1, если $Total liabilities_{i,p} > Total Assets_{i,p}$ и значение 0 в противном случае, где $Total liabilities_{i,p}$ – текущая задолженность компании i в год p , $Total Assets_{i,p}$ – общие активы компании i в год p)	Бинарная
19	Общий	$\frac{NI_{i,p} - NI_{i-1,p}}{ NI_{i,p} + NI_{i-1,p} },$ где $NI_{i,p}$ – чистая прибыль компании i в год p	Количественная
20–23	Общий	Численность персонала (4 переменные: 251–500 человек, 501–1000 человек, 1001–5000 человек, более 5000 человек)	Бинарная
24–27	Общий	Вид деятельности (4 переменные: производство, торговля, услуги, другие)	Бинарная
28–31	Общий	Организационно-правовая форма (4 переменные: ООО, ПАО, АО, иное)	Бинарная
32–36	Общий	Форма собственности (5 переменных: частная, иностранная, российская+иностранная, государственная/федеральная/муниципальная собственность, другие)	Бинарная

Источник / Source: составлено авторами по данным [1, 6, 27–29] / compiled by the authors according to [1, 6, 27–29].

хозяйства, гостиницы и др. Распределение компаний по отраслям представлено на диаграмме рис. 2.

Как видно из рис. 2, в выборке высока доля компаний из сектора торговли, а также обрабатывающей промышленности, которые вместе составляют 61% всей выборки. Также значительную долю составляют компании сферы строительства и электроэнергетики.

Для того чтобы проанализировать возможные различия в факторах риска банкротства компаний по отраслям, были выделены три крупные отраслевые группы компаний: производства, торговли и сферы услуг. Таким образом, выборка была разделена на три примерно равные части (рис. 3).

В выборке преимущественно присутствуют компании с численностью персонала в диапазоне от 1001 до 5000 человек — 41%. Больше 50% составляют компании с организационно-правовой формой ООО. Большинство компаний выборки являются частной собственностью граждан РФ — 69%. 13% компаний

принадлежат иностранным инвесторам (гражданам, государствам и юридическим лицам).

Описательная статистика переменных исследования представлена в табл. 2.

Как видно из табл. 2, в выборку вошли как прибыльные, так и неприбыльные компании, т.е. компании как с положительными, так и отрицательными значениями показателей рентабельности.

РЕЗУЛЬТАТЫ ИССЛЕДОВАНИЯ

Метод случайного леса был применен в работе для прогнозирования банкротства компаний на рассматриваемом промежутке времени, а также для определения факторов, в наибольшей степени влияющих на вероятность банкротства. Говоря о результатах, прежде всего необходимо отметить, что модель в 86% случаев верно предсказала, обанкротилась ли компания выборки за рассматриваемый период наблюдения или продолжила свое существование. Данный процент можно считать достаточно

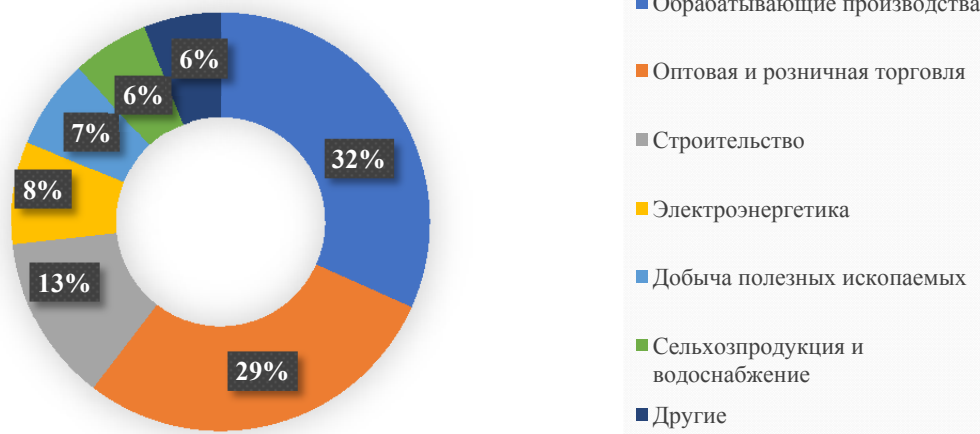


Рис. 2 / Fig. 2. Распределение компаний выборки по отраслям / Distribution of sample companies by industry

Источник / Source: составлено авторами / compiled by the authors.

высоким, так как в ряде российских работ точность прогнозирования методов машинного обучения варьируется от 73 до 90% [8, 31].

Далее с помощью критерия Джини была оценена важность каждой из рассматриваемых пяти групп показателей для прогнозирования банкротства по выборке в целом. На рис. 4 представлен график, который отражает среднюю значимость показателя из каждой категории. Поскольку на каждую группу приходилось разное число показателей, для оценки важности показателей каждой группы было рассчитано среднее влияние факторов в каждой группе на вероятность банкротства. Для этого общая оценка важности факторов, относящихся к отдельной группе, делилась на количество факторов этой группы в модели.

На рис. 4 видно, что наиболее значимые категории показателей с точки зрения прогнозирования вероятности банкротства российских компаний — это показатели рентабельности и финансовой устойчивости. Оценка их важности практически в два раза превышает оценку важности показателей ликвидности и оборачиваемости.

Относительная сила влияния отдельных показателей на вероятность банкротства приведена в табл. 3. В данной и последующих таблицах представлены наиболее значимые факторы, суммарная оценка важности которых по критерию Джини составляет 0,9.

Из табл. 3 видно, что суммарный вес показателей, входящих в группы рентабельности и финансовой устойчивости, составляет почти 53%, что подтверждает необходимость учета этих показателей при прогнозировании банкротства. Средняя сила влияния отдельного показателя в этих группах составляет 8 и 7% соответственно. Следует подчеркнуть, что и в классической модели Альтмана большинство показателей относились к рентабельности и финан-

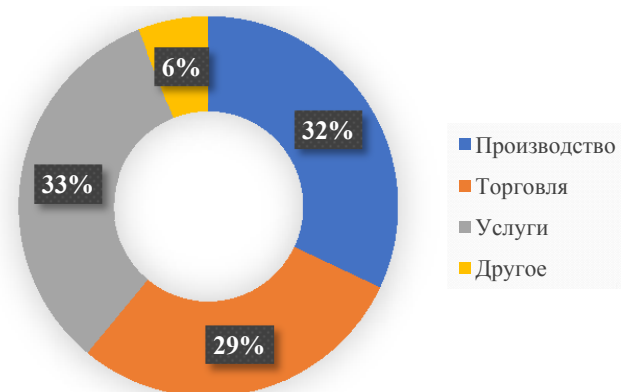


Рис. 3 / Fig. 3. Распределение компаний выборки по сфере деятельности / Distribution of sample companies by type of business operation

Источник / Source: составлено авторами / compiled by the authors.

совой устойчивости [1]. Полученный нами результат также согласуется с исследованием А.М. Карминского и Р.Н. Бурехина, которое показало, что показатели финансовой устойчивости и ликвидности являются наиболее значимыми для прогнозирования банкротства компании на российском рынке [7]. Б.Б. Демешев в 2014 г., анализируя строительную отрасль в России, также подчеркнул важность показателя рентабельности, который в его работе рассчитывался как отношение прибыли до процентов и налогов к общим активам компании [20].

Интересен также тот факт, что «общие» показатели (например, численность персонала, форма собственности и др.) оказались малозначимыми при прогнозировании вероятности банкротства. Следует отметить, что нефинансовые показатели также оказались малозначимыми и в некоторых других исследованиях по российским компаниям. Это, в частности, гово-

Таблица 2 / Table 2

Описательная статистика / Descriptive statistics

Показатель / Indicator	Тип показателя / Type of indicator	Выборочное среднее / Sample average	Среднеквадратическое отклонение / Standard deviation	Min	Max
$\frac{EBITDA_{i,p}}{A_{i,p}}$	Рентабельность	0,089	0,145	-1,029	1,213
$\frac{Retained\ earnings_{i,p}}{A_{i,p}}$	Рентабельность	0,214	0,361	-2,448	1,073
$\frac{EBIT_{i,p}}{Rev_{i,p}}$	Рентабельность	0,067	0,237	-3,936	2,635
$ROE_{i,p}$	Рентабельность	0,161	0,765	-10,211	9,333
$\frac{CA_{i,p}}{STD_{i,p}}$	Ликвидность	2,115	2,578	0,067	31,779
$\frac{Cash_{i,p}}{A_{i,p}}$	Ликвидность	0,049	0,084	0	0,749
$\frac{FA_{i,p}}{CA_{i,p}}$	Финансовая устойчивость	1,281	2,234	0	24,831
$\frac{EBITDA_{i,p}}{Int_{i,p}}$	Финансовая устойчивость	33,742	163,636	-336,7	2536
$\frac{WC_{i,p}}{LTD_{i,p}}$	Финансовая устойчивость	23,457	124,003	-967,833	1513
$\frac{STD + LTD_{i,p}}{Equity_{i,p}}$	Финансовая устойчивость	6,283	25,514	-273,306	364,871
$\frac{Rev_{i,p}}{A_{i,p}}$	Оборачиваемость	1,860	1,679	0	13,919

Окончание таблицы 2 / Table 2 (continued)

Показатель / Indicator	Тип показателя / Type of indicator	Выборочное среднее / Sample average	Среднеквадратическое отклонение / Standard deviation	Min	Max
$\text{Inventories turnover}_{i,p} = \frac{COGS_{i,p}}{\text{Average inventories}_{i,p}}$	Оборачиваемость	36,086	148,735	0	1954
$\text{Fixed assets turnover}_{i,p} = \frac{\text{Revenue}_{i,p}}{\text{Average fixed assets}_{i,p}}$	Оборачиваемость	40,239	174,057	0,001	2595
$\text{Accounts receivable turnover period}_{i,p} = \frac{\text{Revenue}_{i,p}}{\text{Average receivables}_{i,p}}$	Оборачиваемость	9,020	12,493	0	138,983
$\ln\left(\frac{A_{i,p}}{GPN \text{ price index}_p}\right)$	Общий	17,884	1,524	9,158	25,654
$\frac{NI_{i,p} - NI_{i-1,p}}{ NI_{i,p} + NI_{i-1,p} }$	Общий	0,003	0,605	-1	1

Источник / Source: составлено авторами / compiled by the authors.

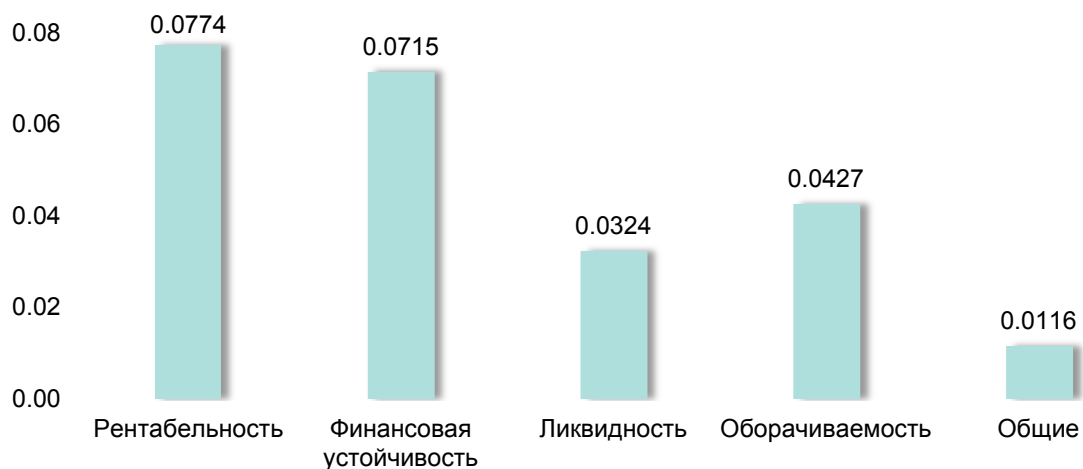


Рис. 4 / Fig. 4. Оценка важности групп факторов риска банкротства по алгоритму случайного леса / Evaluation of the importance of groups of bankruptcy risk factors according to the random forest algorithm

Источник / Source: составлено авторами по результатам исследования / compiled by the authors based on the research results.

Таблица 3 / Table 3

Оценка важности факторов риска банкротства по алгоритму случайного леса для всех компаний выборки / Evaluation of the importance of bankruptcy risk factors according to the random forest algorithm

Номер / Number	Тип показателя / Type of indicator	Название / Designation	Тип переменной / Type of variable	Важность / Importance
1	Рентабельность	$\frac{Retained\ earnings_{i,p}}{A_{i,p}}$	Количественная	0,235
2	Финансовая устойчивость	$\frac{STD + LTD_{i,p}}{Equity_{i,p}}$	Количественная	0,141
3	Общий	$\ln\left(\frac{A_{i,p}}{GPD_p}\right)$	Количественная	0,112
4	Оборачиваемость	$\frac{Rev_{i,p}}{WC_{i,p}}$	Количественная	0,073
5	Финансовая устойчивость	$\frac{FA_{i,p}}{CA_{i,p}}$	Количественная	0,052
6	Оборачиваемость	$\begin{aligned} Accounts\ receivable \\ turnover\ period_{i,p} &= \\ &= \frac{Revenue_{i,p}}{Average\ receivables_{i,p}} \end{aligned}$	Количественная	0,051
7	Оборачиваемость	$\begin{aligned} Inventories\ turnover_{i,p} &= \\ &= \frac{COGS_{i,p}}{Average\ inventories_{i,p}} \end{aligned}$	Количественная	0,046
8	Ликвидность	$\frac{Cash_{i,p}}{A_{i,p}}$	Количественная	0,039
9	Рентабельность	$\frac{EBITDA_{i,p}}{A_{i,p}}$	Количественная	0,037
10	Общий	ПАО	Бинарная	0,034
11	Оборачиваемость	$\frac{Rev_{i,p}}{A_{i,p}}$	Количественная	0,028
12	Рентабельность	$ROE_{i,p}$	Количественная	0,027
13	Ликвидность	$\frac{CA_{i,p}}{STD_{i,p}}$	Количественная	0,026

Окончание таблицы 3 / Table 3 (continued)

Номер / Number	Тип показателя / Type of indicator	Название / Designation	Тип переменной / Type of variable	Важность / Importance
14	Финансовая устойчивость	$\frac{WC_{i,p}}{LTD_{i,p}}$	Количественная	0,022
15	Оборачиваемость	$Fixed\ assets\ turnover_{i,p} = \frac{Revenue_{i,p}}{Average\ fixed\ assets_{i,p}}$	Количественная	0,015

Источник / Source: составлено авторами по результатам исследования / compiled by the authors based on the research results.

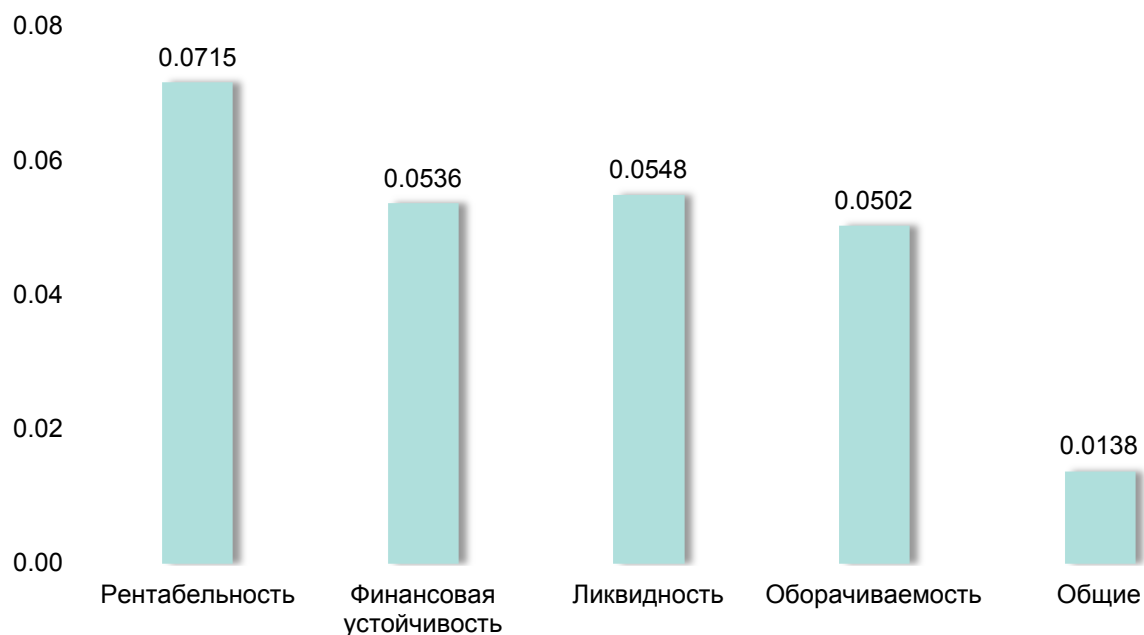


Рис. 5 / Fig. 5. Оценка важности групп факторов риска банкротства по алгоритму случайного леса для производственных компаний / Evaluation of the importance of groups of bankruptcy risk factors according to the random forest algorithm for manufacturing companies

Источник / Source: составлено авторами по результатам исследования / compiled by the authors based on the research results.

рит о том, что «большой размер и длительный срок работы компании на рынке не могут гарантировать устойчивость на российском рынке» [7].

Результаты для подвыборки производственных компаний в целом аналогичны результатам, полученным по выборке в целом (рис. 5). Вновь наиболее существенными группами показателей с точки зрения прогнозирования банкротства являются группы рентабельности и финансовой устойчивости. Однако также велика роль показателей ликвидности и оборачиваемости.

В табл. 4 показана значимость отдельных показателей при прогнозировании банкротства производственных компаний.

Как видно из табл. 4, по-прежнему среди наиболее значимых показателей присутствуют показатели отношения нераспределенной прибыли к активам и финансового рычага. Зарубежные исследователи также подчеркивают важность показателя отношения нераспределенной прибыли к активам, так как нераспределенная прибыль демонстрирует объем средств, оставшихся после расчетов со всеми поставщиками капитала [31].

Что касается компаний сферы торговли, то среди рассматриваемых групп показателей первое место по значимости при прогнозировании банкротства занимает рентабельность (рис. 6). Значимость показателей

Таблица 4 / Table 4

Оценка важности факторов риска банкротства по алгоритму случайного леса для производственных компаний / Evaluation of the importance of bankruptcy risk factors according to the random forest algorithm for manufacturing companies

Номер / Number	Тип показателя / Type of indicator	Название / Designation	Тип переменной / Type of variable	Важность / Importance
1	Рентабельность	$\frac{Retained\ earnings_{i,p}}{A_{i,p}}$	Количественная	0,182
2	Ликвидность	$\frac{Cash_{i,p}}{A_{i,p}}$	Количественная	0,089
3	Финансовая устойчивость	$\frac{STD + LTD_{i,p}}{Equity_{i,p}}$	Количественная	0,088
4	Оборачиваемость	$Inventories\ turnover_{i,p} = \frac{COGS_{i,p}}{Average\ inventories_{i,p}}$	Количественная	0,082
5	Оборачиваемость	$Accounts\ receivable\ turnover\ period_{i,p} = \frac{Revenue_{i,p}}{Average\ receivables_{i,p}}$	Количественная	0,067
6	Общий	$\ln\left(\frac{A_{i,p}}{GPD_p}\right)$	Количественная	0,061
7	Оборачиваемость	$\frac{Rev_{i,p}}{A_{i,p}}$	Количественная	0,047
8	Финансовая устойчивость	$\frac{FA_{i,p}}{CA_{i,p}}$	Количественная	0,045
9	Рентабельность	$\frac{EBIT_{i,p}}{Rev_{i,p}}$	Количественная	0,042
10	Рентабельность	$ROE_{i,p}$	Количественная	0,032
11	Рентабельность	$\frac{EBITDA_{i,p}}{A_{i,p}}$	Количественная	0,03
12	Оборачиваемость	$Fixed\ assets\ turnover_{i,p} = \frac{Revenue_{i,p}}{Average\ fixed\ assets_{i,p}}$	Количественная	0,029

Окончание таблицы 4 / Table 4 (continued)

Номер / Number	Тип показателя / Type of indicator	Название / Designation	Тип переменной / Type of variable	Важность / Importance
13	Финансовая устойчивость	$\frac{WC_{i,p}}{LTD_{i,p}}$	Количественная	0,028
14	Оборачиваемость	$\frac{Rev_{i,p}}{WC_{i,p}}$	Количественная	0,025
15	Общий	ПАО	Бинарная	0,023
16	Общий	Российская + иностранная собственность	Бинарная	0,02
17	Ликвидность	$\frac{CA_{i,p}}{STD_{i,p}}$	Количественная	0,019
18	Общий	Численность персонала 501–1000	Бинарная	0,019
19	Общий	Численность персонала 251–500	Бинарная	0,014
20	Общий	АО	Бинарная	0,01

Источник / Source: составлено авторами по результатам исследования / compiled by the authors based on the research results.

ликвидности усиливается по сравнению с производственными компаниями, что в целом соответствует специфике торговых предприятий.

В табл. 5 показана значимость отдельных показателей при прогнозировании банкротства торговых компаний.

Согласно представленным в табл. 5 результатам первые три места в списке наиболее важных с точки зрения прогнозирования банкротств показателей занимают соответственно показатели рентабельности (0,218), финансовой устойчивости (0,082) и ликвидности (0,079).

В сфере услуг среди рассматриваемых групп показателей коэффициенты финансовой устойчивости обладают наибольшей значимостью при прогнозировании банкротства с большим отрывом от других групп показателей (рис. 7).

В табл. 6 показана значимость отдельных показателей при прогнозировании банкротства торговых компаний.

Как видно из табл. 6, для компаний сферы услуг наибольшей значимостью при прогнозировании банкротства обладает показатель отношения внеоборотных активов к оборотным (0,165). Показатель соотношения заемного и собственного финансирования также обладает высокой значимостью (0,133). По-

казатель финансового левериджа имеет значительное влияние на вероятность банкротства, с повышением уровня долга компания рискует снизить свою финансовую устойчивость [32].

В дополнение к представленным результатам были получены пороговые значения финансовых показателей, разграничивающие интервалы, характеризующиеся разной степенью риска банкротства. С этой целью использовался визуальный анализ модели классификации, построенной алгоритмом случайного леса при помощи библиотеки PDPbox (Python). Библиотека позволяет построить графики частичной зависимости (partial dependence plots, или PDP-графики), которые отражают оценку влияния отдельных переменных на результат классификации. График PDP представляет собой ломаную линию. Хотя модель случайного леса не дает представление о направлении взаимосвязи признаков с результатом классификации, его можно увидеть на PDP-графиках при помощи знаков подобно знакам перед коэффициентами в регрессионных моделях.

При анализе финансовой несостоятельности PDPbox помогает проследить зависимость между отдельным фактором банкротства в модели классификации и возможным банкротством компании. Множество допустимых значений признака разбивается на три

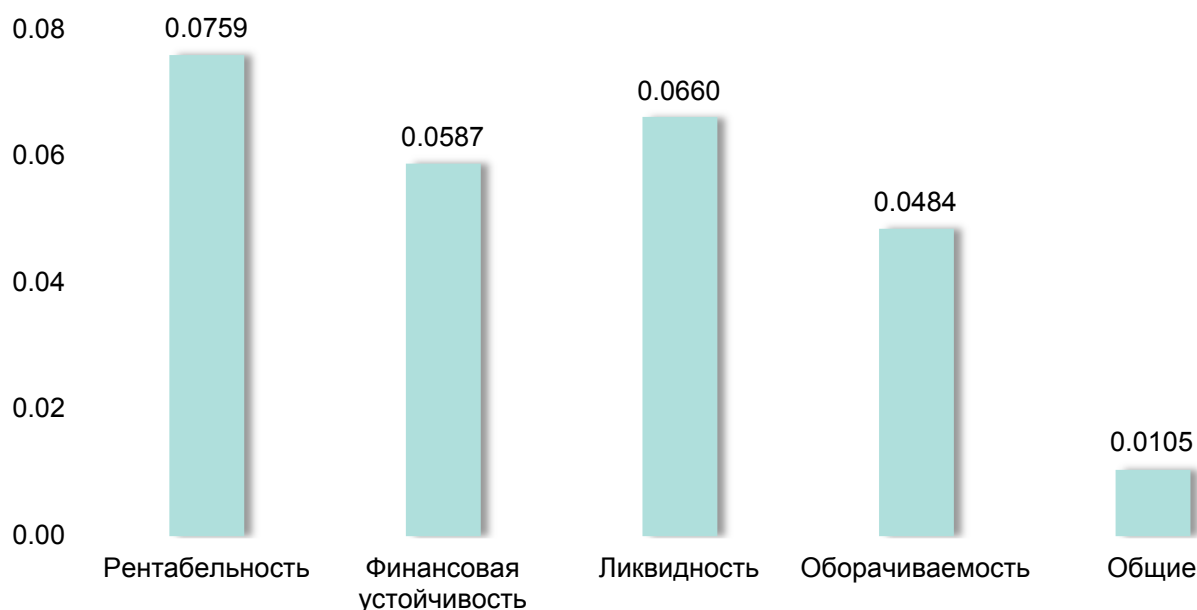


Рис. 6 / Fig. 6. Оценка важности групп факторов риска банкротства по алгоритму случайного леса для торговых компаний / Evaluation of the importance of groups of bankruptcy risk factors according to the random forest algorithm for retail companies

Источник / Source: составлено авторами по результатам исследования / compiled by the authors based on the research results.

Таблица 5 / Table 5

Оценка важности факторов риска банкротства по алгоритму случайного леса для торговых компаний / Evaluation of the importance of bankruptcy risk factors according to the random forest algorithm for retail companies

Номер / Number	Тип показателя / Type of indicator	Название / Designation	Тип переменной / Type of variable	Важность / Importance
1	Рентабельность	$\frac{Retained\ earnings_{i,p}}{A_{i,p}}$	Количественная	0,219
2	Финансовая устойчивость	$\frac{WC_{i,p}}{LTD_{i,p}}$	Количественная	0,082
3	Ликвидность	$\frac{Cash_{i,p}}{A_{i,p}}$	Количественная	0,079
4	Общий	$\ln\left(\frac{A_{i,p}}{GPD_p}\right)$	Количественная	0,075
5	Оборачиваемость	$\frac{Accounts\ receivable\ turnover\ period_{i,p}}{Revenue_{i,p}} = \frac{Average\ receivables_{i,p}}{Revenue_{i,p}}$	Количественная	0,064

Окончание таблицы 5 / Table 5 (continued)

Номер / Number	Тип показателя / Type of indicator	Название / Designation	Тип переменной / Type of variable	Важность / Importance
6	Финансовая устойчивость	$\frac{FA_{i,p}}{CA_{i,p}}$	Количественная	0,062
7	Оборачиваемость	$Inventories\ turnover_{i,p} = \frac{COGS_{i,p}}{Average\ inventories_{i,p}}$	Количественная	0,059
8	Ликвидность	$\frac{CA_{i,p}}{STD_{i,p}}$	Количественная	0,053
9	Оборачиваемость	$\frac{Rev_{i,p}}{A_{i,p}}$	Количественная	0,05
10	Оборачиваемость	$\frac{Rev_{i,p}}{WC_{i,p}}$	Количественная	0,048
11	Финансовая устойчивость	$\frac{STD + LTD_{i,p}}{Equity_{i,p}}$	Количественная	0,032
12	Рентабельность	$\frac{EBITDA_{i,p}}{A_{i,p}}$	Количественная	0,029
13	Рентабельность	$ROE_{i,p}$	Количественная	0,028
14	Рентабельность	$\frac{EBIT_{i,p}}{Rev_{i,p}}$	Количественная	0,027

Источник / Source: составлено авторами по результатам исследования / compiled by the authors based on the research results.

области — значения, характеризующиеся «высоким», «средним» и «низким» риском банкротства соответственно. Области соответствуют качественно разным группам финансового состояния. Для разбиения на области применяется эмпирический метод (elbow method), который позволяет найти «критические» точки на PDP-графиках. Эти точки (как и сами области) приведены в таблице для каждого признака по отдельности.

Визуально на PDP-графиках зависимости вероятности банкротства от значения фактора может наблюдаться заметное изменение «кривизны». Начиная с определенного значения фактора, вероятность банкротства существенно снижается или растет. Такое изменение «кривизны» анализируется при помощи значений косинусов углов между звеньями PDP-графика. Для определения пороговых значений, форми-

рующих границы интервалов, которые определяют качественно разные группы, выбираются два узла с наибольшим косинусом угла.

В табл. 7 представлены промежутки значений показателей каждой из групп, оказавшихся наиболее значимыми для прогнозирования банкротства, в зависимости от степени риска банкротства.

Представленные в табл. 7 результаты позволяют получить представление о том, какие промежутки значений финансовых показателей характеризуются различным риском банкротства для компании. Данный результат позволяет классифицировать компании в зависимости от степени вероятности их банкротства. Для получения более детальных выводов целесообразно провести подобный анализ для отдельных отраслей/подотраслей российской экономики.

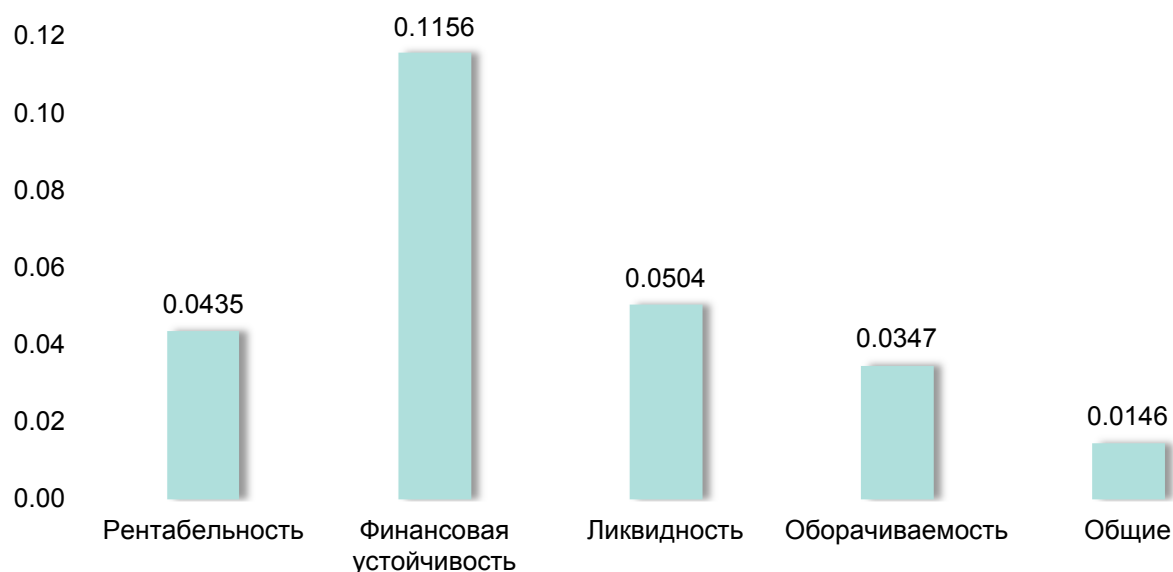


Рис. 7 / Fig. 7. Оценка важности групп факторов риска банкротства по алгоритму случайного леса для компаний сферы услуг / Evaluation of the importance of groups of bankruptcy risk factors according to the random forest algorithm for service companies

Источник / Source: составлено авторами по результатам исследования / compiled by the authors based on the research results.

Таблица 6 / Table 6

Оценка важности факторов риска банкротства по алгоритму случайного леса для компаний сферы услуг / Evaluation of the importance of bankruptcy risk factors according to the random forest algorithm for service companies

Номер / Number	Тип показателя / Type of indicator	Название / Designation	Тип переменной / Type of variable	Важность / Importance
1	Финансовая устойчивость	$\frac{FA_{i,p}}{CA_{i,p}}$	Количественная	0,165
2	Финансовая устойчивость	$\frac{STD + LTD_{i,p}}{Equity_{i,p}}$	Количественная	0,133
3	Общий	$\ln\left(\frac{A_{i,p}}{GPD_p}\right)$	Количественная	0,087
4	Рентабельность	$\frac{EBIT_{i,p}}{Rev_{i,p}}$	Количественная	0,077
5	Общий	ПАО	Бинарная	0,059
6	Оборачиваемость	$Inventories\ turnover_{i,p} = \frac{COGS_{i,p}}{Average\ inventories_{i,p}}$	Количественная	0,058
7	Ликвидность	$\frac{Cash_{i,p}}{A_{i,p}}$	Количественная	0,052

Окончание таблицы 6 / Table 6 (continued)

Номер / Number	Тип показателя / Type of indicator	Название / Designation	Тип переменной / Type of variable	Важность / Importance
8	Финансовая устойчивость	$\frac{WC_{i,p}}{LTD_{i,p}}$	Количественная	0,049
9	Ликвидность	$\frac{CA_{i,p}}{STD_{i,p}}$	Количественная	0,048
10	Рентабельность	$\frac{EBITDA_{i,p}}{A_{i,p}}$	Количественная	0,041
11	Оборачиваемость	$\begin{aligned} \text{Accounts receivable} \\ \text{turnover period} &= \\ &= \frac{Revenue_{i,p}}{Average\ receivables_{i,p}} \end{aligned}$	Количественная	0,039
12	Рентабельность	$\frac{Retained\ earnings_{i,p}}{A_{i,p}}$	Количественная	0,029
13	Рентабельность	$ROE_{i,p}$	Количественная	0,028
14	Оборачиваемость	$\frac{Rev_{i,p}}{A_{i,p}}$	Количественная	0,027

Источник / Source: составлено авторами по результатам исследования / compiled by the authors based on the research results.

ВЫВОДЫ

Оценка вероятности банкротства компании и определение факторов риска банкротства чрезвычайно важны для понимания перспектив развития бизнеса в любой отрасли. Прогнозирование вероятности банкротства компании традиционно осуществляется при помощи математических моделей, как правило, основанных на эконометрических методах или методах машинного обучения. Эти модели адаптируются под специфику рынков отдельных стран, что предполагает тестирование модели на данных, собранных по выборке компаний соответствующей страны.

Современные исследования показывают, что методы машинного обучения обеспечивают получение более высокой точности прогноза вероятности банкротства компании по сравнению с эконометрическими методами. При этом в большинстве работ по российским компаниям, основанных на методах машинного обучения, сравнивается прогностическая способность различных моделей, в то время как срав-

нительно мало внимания уделяется анализу отдельных факторов банкротства компании.

В рамках данного исследования решалась задача оценки относительной силы влияния отдельных факторов на вероятность банкротства крупных российских компаний различных отраслей при помощи одного из методов машинного обучения — алгоритма случайного леса. Рассматривались показатели по 3184 компаниям с 2009 по 2020 г. По выборке в целом тестируемая модель в 86% случаев корректно предсказала возможность банкротства компании за рассматриваемый промежуток времени. Данный результат подтвердил высокую эффективность использования методов машинного обучения (и, в частности, алгоритма случайного леса) применительно к решению задачи прогнозирования банкротства компании.

Проведенное исследование также показало, что факторы риска банкротства компаний существенным образом зависят от их отраслевой принадлежности:

1. Для производственных компаний следует в первую очередь обращать внимание на группу показате-

Таблица 7 / Table 7

Промежутки значений показателей, характеризующихся разной степенью риска банкротства компании / Intervals of indicator values characterized by different degrees of company bankruptcy risk

Показатель / Indicator	Промежуток значений показателя / Interval of indicator values		
	с высоким риском банкротства / with high bankruptcy risk	со средним риском банкротства / with medium bankruptcy risk	с низким риском банкротства / with low bankruptcy risk
Выборка в целом			
$\frac{Retained\ earnings_{i,p}}{A_{i,p}}$	Меньше 0,087	От 0,087 до 0,156	Выше 0,156
$\frac{STD + LTD_{i,p}}{Equity_{i,p}}$	Меньше 2,4 Выше 10,9	От 2,4 до 10,9	–
$\frac{Rev_{i,p}}{WC_{i,p}}$	Меньше 0,2	От 0,2 до 1,8	Выше 1,8
$\frac{Cash_{i,p}}{A_{i,p}}$	Меньше 0,003	От 0,003 до 0,02	Выше 0,02
Производственные компании			
$\frac{Retained\ earnings_{i,p}}{A_{i,p}}$	Меньше 0,04	От 0,04 до 0,28	Выше 0,28
$\frac{Cash_{i,p}}{A_{i,p}}$	Меньше 0,002	От 0,002 до 0,005	Выше 0,005
$\frac{STD + LTD_{i,p}}{Equity_{i,p}}$	Меньше 2,5 Выше 8	От 2,5 до 8	–
$\frac{Accounts\ receivable\ turnover\ period_{i,p}}{=}$ $= \frac{Revenue_{i,p}}{Average\ receivables_{i,p}}$	Меньше 1,89	От 1,89 до 4,73	Выше 4,73
Торговые компании			
$\frac{Retained\ earnings_{i,p}}{A_{i,p}}$	Меньше 0,12	От 0,12 до 0,27	Выше 0,27
$\frac{WC_{i,p}}{LTD_{i,p}}$	Меньше –0,5	От –0,5 до 1,6	Выше 1,6

Окончание таблицы 7 / Table 7 (continued)

Показатель / Indicator	Промежуток значений показателя / Interval of indicator values		
	с высоким риском банкротства / with high bankruptcy risk	со средним риском банкротства / with medium bankruptcy risk	с низким риском банкротства / with low bankruptcy risk
$\frac{Cash_{i,p}}{A_{i,p}}$	Меньше 0,0004	От 0,0004 до 0,004	Выше 0,004
$\frac{Accounts\ receivable\ turnover\ period_{i,p}}{=}$ $= \frac{Revenue_{i,p}}{Average\ receivables_{i,p}}$	Меньше 0,13	От 0,13 до 2,72	Выше 2,72
Сфера услуг			
$\frac{Fixed\ assets_{i,p}}{Current\ assets_{i,p}}$	Меньше 0,01	От 0,01 до 0,73	Выше 0,73
$\frac{EBIT_{i,p}}{Rev_{i,p}}$	Меньше 0,134	От 0,134 до 0,323	Выше 0,323
$\frac{Inventories\ turnover_{i,p}}{=}$ $= \frac{COGS_{i,p}}{Average\ inventories_{i,p}}$	Меньше 3,4	От 3,4 до 11,8	Выше 11,8
$\frac{Cash_{i,p}}{A_{i,p}}$	Меньше 0,001	От 0,001 до 0,01	Выше 0,01

Источник / Source: составлено авторами по результатам исследования / compiled by the authors based on the research results.

телей рентабельности. Например, согласно полученным результатам риск банкротства существенно возрастает при значениях отношения нераспределенной прибыли к активам меньше 4%.

2. Для торговых компаний необходимо, прежде всего, ориентироваться на показатели ликвидности и рентабельности. Отношение денежных средств к активам не должно опускаться ниже 0,04%, а отношение нераспределенной прибыли к активам — ниже 12%.

3. Для компаний сферы услуг приоритетными с точки зрения прогнозирования вероятности банкротства являются показатели финансовой устойчивости. В частности, низкий риск банкротства наблюдается в ситуации, когда отношение внеоборотных активов к оборотным превышает 73%.

Таким образом, в работе не только выявлены факторы, существенным образом влияющие на вероятность банкротства российских компаний в различных отраслях, но и определены «пороговые» значения этих показателей, при которых существенно увеличивается риск банкротства.

Результаты работы могут быть использованы как внутренними (менеджмент, совет директоров), так и внешними (аналитики, кредиторы и т.д.) стейкхолдерами для определения текущего финансового состояния компании, а также прогнозирования перспектив развития бизнеса.

К возможным направлениям будущих исследований можно отнести углубленный анализ факторов банкротства малых и средних российских предприятий. Также для получения более детальных выводов

целесообразно провести анализ факторов банкротства для отдельных отраслей/подотраслей российской экономики при помощи методов машинного обучения. Решение и определения факторов риска

банкротства российских компаний должно привести к сокращению числа предприятий-банкротов, что, в свою очередь, будет способствовать оздоровлению и развитию национальной экономики.

СПИСОК ЛИТЕРАТУРЫ

1. Altman E.I. Financial ratios, discriminant analysis and the prediction of corporate bankruptcy. *The Journal of Finance*. 1968;23(4):589–609. DOI: 10.1111/j.1540-6261.1968.tb 00843.x
2. Altman E.I., Fargher N., Kalotay E. A simple empirical model of equity-implied probabilities of default. *The Journal of Fixed Income*. 2011;20(3):71–85. DOI: 10.3905/jfi.2011.20.3.071
3. Altman E.I., Iwanicz-Drozdzowska M., Laitinen E., Suvas A. Distressed firm and bankruptcy prediction in an international context: A review and empirical analysis of Altman's Z-score model. *SSRN Electronic Journal*. 2014. DOI: 10.2139/ssrn.2536340
4. Agarwal V., Taffler R.J. Twenty-five years of the Taffler Z-score model: Does it really have predictive ability? *Accounting and Business Research*. 2007;37(4):285–300. DOI: 10.1080/00014788.2007.9663313
5. Zmijewski M. Methodological issues related to the estimation of financial distress prediction models. *Journal of Accounting Research*. 1984;22:59–82. DOI: 10.2307/2490859
6. Ohlson J. Financial ratios and the probabilistic prediction of bankruptcy. *Journal of Accounting Research*. 1980;18(1):109–131. DOI: 10.2307/2490395
7. Карминский А.М., Бурехин Р.Н. Сравнительный анализ методов прогнозирования банкротств российских строительных компаний. *Бизнес-информатика*. 2019;13(3):52–66. DOI: 10.17323/1998-0663.2019.3.52.66
8. Jones S., Hensher D.A. Predicting firm financial distress: A mixed logit model. *The Accounting Review*. 2004;79(4):1011–1038. DOI: 10.2308/accr.2004.79.4.1011
9. Хайдаршина Г.А. Совершенствование методов оценки риска банкротства российских предприятий в современных условиях. *Имущественные отношения в Российской Федерации*. 2009;(8):86–95.
10. Карминский А.М., Костров А.В., Мурзенков Т.Н. Моделирование вероятности дефолта российских банков с использованием эконометрических методов. Препринт WP7/2012/04. М.: Изд. дом ВШЭ; 2012. 64 с.
11. Behr A., Weinblat J. Default patterns in seven EU countries: A random forest approach. *International Journal of the Economics of Business*. 2017;24(2):181–222. DOI: 10.1080/13571516.2016.1252532
12. Li Y., Wang Y. Machine learning methods of bankruptcy prediction using accounting ratios. *Open Journal of Business and Management*. 2018;6(1):1–20. DOI: 10.4236/ojbm.2018.61001
13. Joshi S., Ramesh R., Tahsildar S. A bankruptcy prediction model using random forest. In: 2nd Int. conf. on intelligent computing and control systems (ICICCS). (Madurai, 14–15 June 2018). Piscataway, NJ: IEEE; 2018. DOI: 10.1109/ICCONS.2018.8663128
14. Денисов Д.В., Смирнова Д.К. Применение метода случайных лесов для оценки резерва произошедших, но еще не заявленных убытков страховой компании. *International Journal of Open Information Technologies*. 2016;4(7):45–50.
15. Груздев А.В. Метод случайного леса в скоринге. *Риск-менеджмент в кредитной организации*. 2014;(1):28–43.
16. Казаков А.В., Колышкин А.В. Разработка моделей прогнозирования банкротства в современных российских условиях. *Вестник Санкт-Петербургского университета. Экономика*. 2018;34(2):241–266. DOI: 10.21638/11701/spbu05.2018.203
17. Колышкин А.В., Гиленко Е.В., Довженко С.Е., Жилкин С.А., Чое С.Е. Прогнозирование финансовой несостоятельности предприятий. *Вестник Санкт-Петербургского университета. Экономика*. 2014;(2):122–142.
18. Федорова Е.А., Гиленко Е.В., Довженко С.Е. Модели прогнозирования банкротства: особенности российских предприятий. *Проблемы прогнозирования*. 2013;(2):85–92.
19. Федорова Е.А., Мусиенко С.О., Федоров Ф.Ю. Прогнозирование банкротства субъектов малого и среднего предпринимательства в России. *Финансы и кредит*. 2018;24(11):2537–2552. DOI: 10.24891/fc.24.11.2537
20. Демешев Б.Б., Тихонова А.С. Прогнозирование банкротства российских компаний: межотраслевое сравнение. *Экономический журнал Высшей школы экономики*. 2014;18(3):359–386.
21. Горбатков С.А., Белолипец И.И. Гибридный метод оценки риска банкротств на базе байесовского ансамбля нейросетей и logit-модели. *Науковедение*. 2013;(6). URL: http://naukovedenie.ru/PDF/25EVN_613.pdf

22. Макеева Е.Ю., Аршавский И.В. Применение нейронных сетей и семантического анализа для прогнозирования банкротства. *Корпоративные финансы*. 2014;8(4):130–141. DOI: 10.17323/j.jcfr.2073–0438.8.4.2014.130–141
23. Богданова Т.К., Шевгунов Т.Я., Уварова О.М. Применение нейронных сетей для прогнозирования платежеспособности российских предприятий обрабатывающих отраслей. *Бизнес-информатика*. 2013;(2):40–48.
24. 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;25(2):242–254. DOI: 10.22363/2313–2329–2017–25–2–242–254
25. Liaw A., Wiener M. Classification and regression by randomForest. *R News*. 2002;2(3):18–22. URL: <https://cogns.northwestern.edu/cbmj/LiawAndWiener2002.pdf>
26. Gepp A., Kumar K. Predicting financial distress: A comparison of survival analysis and decision tree techniques. *Procedia Computer Science*. 2015;54:396–404. DOI: 10.1016/j.procs.2015.06.046
27. Altman E.I., Sabato G. Modelling credit risk from SMEs: Evidence from the US market. *ABACUS: A Journal of Accounting, Finance and Business Studies*. 2007;43(3):332–357. DOI: 10.1111/j.1467–6281.2007.00234.x
28. Жданов В.Ю., Афанасьева О.А. Модель диагностики риска банкротства предприятий авиационно-промышленного комплекса. *Корпоративные финансы*. 2011;5(4):77–89. DOI: 10.17323/j.jcfr.2073–0438.5.4.2011.77–89
29. Drezner Z., Marcoulides G., Stohs M.H. Financial applications of a Tabu search variable selection model. *Journal of Applied Mathematics and Decision Sciences*. 2001;5(4):215–234. DOI: 10.1155/S 1173912601000165
30. Altman E.I., Sabato G., Wilson N. The value of non-financial information in SME risk management. *Journal of Credit Risk*. 2010;6(2):95–127. DOI: 10.21314/JCR.2010.110
31. Перерва О.Л., Степанов С.Е., Незимова С.С. Сравнение эконометрических моделей и методов бизнес-аналитики предсказания банкротства предприятий. *Науковедение*. 2017;9(6)1–9. URL: <https://naukovedenie.ru/PDF/82EVN 617.pdf>
32. Molina C.A. Are firms underleveraged? An examination of the effect of leverage on default probabilities. *The Journal of Finance*. 2005;60(3):1427–1459. DOI: 10.1111/j.1540–6261.2005.00766.x

REFERENCES

1. Altman E.I. Financial ratios, discriminant analysis and the prediction of corporate bankruptcy. *The Journal of Finance*. 1968;23(4):589–609. DOI: 10.1111/J.1540–6261.1968.TB 00843.X
2. Altman E.I., Fargher N., Kalotay E. A simple empirical model of equity-implied probabilities of default. *The Journal of Fixed Income*. 2011;20(3):71–85. DOI: 10.3905/jfi.2011.20.3.071
3. Altman E.I., Iwanicz-Drozdowska M., Laitinen E., Suvas A. Distressed firm and bankruptcy prediction in an international context: A review and empirical analysis of Altman's Z-score model. *SSRN Electronic Journal*. 2014. DOI: 10.2139/ssrn.2536340
4. Agarwal V., Taffler R.J. Twenty-five years of the Taffler Z-score model: Does it really have predictive ability? *Accounting and Business Research*. 2007;37(4):285–300. DOI: 10.1080/00014788.2007.9663313
5. Zmijewski M. Methodological issues related to the estimation of financial distress prediction models. *Journal of Accounting Research*. 1984;22:59–82. DOI: 10.2307/2490859
6. Ohlson J. Financial ratios and the probabilistic prediction of bankruptcy. *Journal of Accounting Research*. 1980;18(1):109–131. DOI: 10.2307/2490395
7. Karminsky A.M., Burekhin R.N. Comparative analysis of methods for forecasting bankruptcies of Russian construction companies. *Business Informatics*. 2019;13(3):52–66. DOI: 10.17323/1998–0663.2019.3.52.66 (In Russ.: *Biznes-informatika*. 2019;13(3):52–66. DOI: 10.17323/1998–0663.2019.3.52.66).
8. Jones S., Hensher D.A. Predicting firm financial distress: A mixed logit model. *The Accounting Review*. 2004;79(4):1011–1038. DOI: 10.2308/accr.2004.79.4.1011
9. Haydarshina G.A. Improvement of methods for assessing risk of bankruptcy of Russian enterprises in modern conditions. *Imushchestvennye otnosheniya v Rossiiskoi Federatsii = Property Relations in the Russian Federation*. 2009;(8):86–95. (In Russ.).
10. Karminsky A.M., Kostrov A.V., Murzenkov T.N. Modeling of the probability of default of Russian banks using econometric methods. Preprint WP7/2012/04. Moscow: NRU HSE; 2012. 64 p. (In Russ.).
11. Behr A., Weinblat J. Default patterns in seven EU countries: A random forest approach. *International Journal of the Economics of Business*. 2017;24(2):181–222. DOI: 10.1080/13571516.2016.1252532

12. Li Y., Wang Y. Machine learning methods of bankruptcy prediction using accounting ratios. *Open Journal of Business and Management*. 2018;6(1):1–20. DOI: 10.4236/ojbm.2018.61001
13. Joshi S., Ramesh R., Tahsildar S. A bankruptcy prediction model using random forest. In: 2nd Int. conf. on intelligent computing and control systems (ICICCS). (Madurai, 14–15 June 2018). Piscataway, NJ: IEEE; 2018. DOI: 10.1109/ICCONS.2018.8663128
14. Denisov D.V., Smirnova D.K. Application of random forest method to estimate the incurred but not reported claims reserve of an insurance company. *International Journal of Open Information Technologies*. 2016;4(7):45–50. (In Russ.).
15. Gruzdev A.V. Random forest method in scoring. *Risk-menedzhment v kreditnoi organizatsii*. 2014;(1):28–43. (In Russ.).
16. Kazakov A.V., Kolyshkin A.V. The development of bankruptcy prediction models in modern Russian economy. *Vestnik Sankt-Peterburgskogo universiteta. Ekonomika = St. Petersburg University Journal of Economic Studies (SUJES)*. 2018;34(2):241–266. (In Russ.). DOI: 10.21638/11701/spbu05.2018.203
17. Kolyshkin A.V., Gilenko E.V., Dovzhenko S.E., Zhilkin S.A., Choe S.E. Forecasting the financial insolvency of enterprises. *Vestnik Sankt-Peterburgskogo universiteta. Ekonomika = St. Petersburg University Journal of Economic Studies (SUJES)*. 2014;(2):122–142. (In Russ.).
18. Fedorova E.A., Gilenko E.V., Dovzhenko S.E. Models for bankruptcy forecasting: Case study of Russian enterprises. *Studies on Russian Economic Development*. 2013;24(2):159–164. (Russ. ed.: *Problemy prognozirovaniya*. 2013;(2):85–92.).
19. Fedorova E.A., Musienko S.O., Fedorov F. Yu. Prediction of bankruptcy of small and medium-sized business entities in Russia. *Finansy i kredit = Finance and Credit*. 2018;24(11):2537–2552. (In Russ.). DOI: 10.24891/fc.24.11.2537
20. Demeshev B., Tikhonova A. Default prediction for Russian companies: Intersectoral comparison. *Ekonomicheskii zhurnal Vysshei shkoly ekonomiki = The HSE Economic Journal*. 2014;18(3):359–386. (In Russ.).
21. Gorbakov S., Beloliptsev I. A hybrid method for estimating the risk of bankruptcies based on Bayesian neural network ensemble and the logit-model. *Naukovedenie*. 2013;(6). URL: <http://naukovedenie.ru/PDF/25EVN613.pdf> (In Russ.).
22. Makeeva E. Yu., Arshavsky I.V. Integration of neural networks and semantic interpretation for bankruptcy prediction. *Korporativnye finansy = Journal of Corporate Finance Research*. 2014;8(4):130–141. (In Russ.). DOI: 10.17323/j.jcfr.2073–0438.8.4.2014.130–141
23. Bogdanova T., Shevgunov T., Uvarova O. Using neural networks for solvency prediction for Russian companies of manufacturing industries. *Business Informatics*. 2013;(2):40–48. (In Russ.: *Biznes-informatika*. 2013;(2):40–48.).
24. Arinichev I.V., Bogdashev I.V. Estimation of bankruptcy risk of small business companies using methods of machine learning. *Vestnik Rossiiskogo universiteta druzhby narodov. Seriya: Ekonomika = RUDN Journal of Economics*. 2017;25(2):242–254. (In Russ.). DOI: 10.22363/2313–2329–2017–25–2–242–254
25. Liaw A., Wiener M. Classification and regression by randomForest. *R News*. 2002;2(3):18–22. URL: <https://cogns.northwestern.edu/cbm/LiawAndWiener2002.pdf>
26. Gepp A., Kumar K. Predicting financial distress: A comparison of survival analysis and decision tree techniques. *Procedia Computer Science*. 2015;54:396–404. DOI: 10.1016/j.procs.2015.06.046
27. Altman E.I., Sabato G. Modelling credit risk from SMEs: Evidence from the US market. *ABACUS: A Journal of Accounting, Finance and Business Studies*. 2007;43(3):332–357. DOI: 10.1111/j.1467–6281.2007.00234.x
28. Zhdanov V. Yu., Afanaseva O.A. Bankruptcy risk diagnostics model for aviation enterprises. *Korporativnye finansy = Journal of Corporate Finance Research*. 2011;5(4):77–89. (In Russ.). DOI: 10.17323/j.jcfr.2073–0438.5.4.2011.77–89
29. Drezner Z., Marcoulides G., Stohs M.H. Financial applications of a Tabu search variable selection model. *Journal of Applied Mathematics and Decision Sciences*. 2001;5(4):215–234. DOI: 10.1155/S1173912601000165
30. Altman E.I., Sabato G., Wilson N. The value of non-financial information in SME risk management. *Journal of Credit Risk*. 2010;6(2):95–127. DOI: 10.21314/JCR.2010.110
31. Pererva O.L., Stepanov S.E., Nezimova S.S. Comparison of econometric models and methods of business analytics for prediction of bankruptcy of enterprises. *Naukovedenie*. 2017;9(6):1–9. URL: <https://naukovedenie.ru/PDF/82EVN617.pdf> (In Russ.).
32. Molina C.A. Are firms underleveraged? An examination of the effect of leverage on default probabilities. *The Journal of Finance*. 2005;60(3):1427–1459. DOI: 10.1111/j.1540–6261.2005.00766.x

ИНФОРМАЦИЯ ОБ АВТОРАХ / ABOUT THE AUTHORS



Андрей Андреевич Жуков — студент программы магистратуры, Высшая школа менеджмента, Санкт-Петербургский государственный университет, Санкт-Петербург, Россия

Andrei A. Zhukov — Master program student, Graduate School of Management, St. Petersburg State University, St. Petersburg, Russia

<https://orcid.org/0000-0003-1932-0242>

Автор для корреспонденции / Corresponding author
andrey.zhukov399@gmail.com



Егор Дмитриевич Никулин — кандидат экономических наук, доцент, Высшая школа менеджмента, Санкт-Петербургский государственный университет, Санкт-Петербург, Россия

Egor D. Nikulin — Cand. Sci. (Econ.), Assoc. Prof., Graduate School of Management, St. Petersburg State University, St. Petersburg, Russia

<https://orcid.org/0000-0003-0475-3424>

nikulin@gsom.spbu.ru



Данил Андреевич Щучкин — студент программы магистратуры, Санкт-Петербургский государственный университет телекоммуникаций им. профессора М.А. Бонч-Бруевича, Санкт-Петербург, Россия

Danil A. Shchuchkin — Master program student, Bonch-Bruevich St. Petersburg State University of Telecommunications, St. Petersburg, Russia

<https://orcid.org/0000-0003-3346-1316>

da.shchuchkin@gmail.com

Заявленный вклад авторов:

А.А. Жуков — методологическая основа, сбор и анализ данных, обзор литературы, описание результатов.

Е.Д. Никулин — аннотация, введение, обзор литературы, выводы и рекомендации, методологическая основа.

Д.А. Щучкин — моделирование с помощью методов машинного обучения.

Authors' declared contribution:

A.A. Zhukov — developed methodological basis, data collection and analysis, literature review, description of the results.

E.D. Nikulin — wrote the abstract and introduction, literature review, general conclusions and recommendations, methodological basis.

D.A. Shchuchkin — modeling using machine learning methods.

Конфликт интересов: авторы заявляют об отсутствии конфликта интересов.

Conflicts of Interest Statement: The authors have no conflicts of interest to declare.

Статья поступила в редакцию 25.12.2021; после рецензирования 10.01.2022; принята к публикации 27.05.2022.

Авторы прочитали и одобрили окончательный вариант рукописи.

The article was submitted on 25.12.2021; revised on 10.01.2022 and accepted for publication on 27.05.2022.

The authors read and approved the final version of the manuscript.