

Технологии информационного общества

ПРОГНОЗИРОВАНИЕ БАНКРОТСТВА МЕТОДАМИ МАШИННОГО ОБУЧЕНИЯ

Статья рекомендована к публикации членом редакционного совета А.Н. Райковым 23.11.2020.

Доржиев Ардан Саянович

*Финансовый университет при Правительстве Российской Федерации, факультет информационных технологий и анализа больших данных, магистрант
Москва, Россия
a.s.dorzhev@mail.ru*

Аннотация

В своевременном прогнозировании банкротства заинтересованы многие стороны: акционеры, руководители и сотрудники самой компании, контрагенты, партнеры, кредиторы и, конечно, регуляторы и органы государственной власти. В статье представлены существующие модели прогнозирования банкротства: как традиционные статистические, так и современные, на основе машинного обучения, описаны ограничения в применении моделей, способы работы с несбалансированными данными, а также показано, в чем заключается превосходство современных методов прогнозирования над традиционными статистическими.

Ключевые слова

прогнозирование банкротства, Z-оценка, дискриминантный анализ, машинное обучение, нейронные сети, случайные леса, метод опорных векторов

Введение

Прогнозирование финансовой неустойчивости – важный компонент управления компанией. Из-за банкротства¹ акционеры теряют доходы, бизнес, нарушаются цепочки поставок, фискальные органы лишаются налоговых поступлений, госорганы фиксируют снижение экономического роста и повышение социальной напряженности, а работники вынуждены искать новую работу. Поэтому получение ясной картины финансового и имущественного состояния компаний уже много лет является целью специалистов самых разных областей знаний.

Компании активно внедряют современные технологии в основную деятельность, автоматизируя большое количество бизнес-процессов [1]. Это позволяет выстраивать, например, электронную экспертизу, с помощью которой возможно оперативно объединять множество экспертов из разных предметных областей для полного охвата рассматриваемой проблемы и принятия коллективного решения [41]. Электронная экспертиза также подразумевает взаимодействие людей с интеллектуальными системами, которые способны строить анализ и прогнозы на основе более широкого пространства переменных. Так, в частности, методы машинного обучения применяются уже в значительном количестве бизнес-задач [71], в том числе для интеллектуальной аналитики больших данных [14], которые компания аккумулирует, и для построения моделей прогнозирования.

Большинство таких моделей, которые будут рассмотрены далее, относятся к классу инерционных, то есть тех, которые предоставляют прогноз, основанный на продолжении некоторого тренда из предыдущих временных периодов. Однако риски банкротства могут также возникнуть из-за некорректного стратегического менеджмента. Такой менеджмент подразумевает разработку

¹ Для краткости в данной статье термин «банкротство» употребляется для обозначения состояния, при котором должник не способен рассчитаться со своими кредиторами

© Доржиев А.С., 2021. Производство и хостинг журнала «Информационное общество» осуществляется Институтом развития информационного общества.

Данная статья распространяется на условиях международной лицензии Creative Commons «Атрибуция — Некоммерческое использование — На тех же условиях» Всемирная 4.0 (Creative Commons Attribution – NonCommercial – ShareAlike 4.0 International; CC BY-NC-SA 4.0). См. <https://creativecommons.org/licenses/by-nc-sa/4.0/legalcode.ru>

долгосрочных целей и действий, которые позволят достичь более высоких результатов в будущем, например, стать лидирующей компанией в своей отрасли. Разрабатываемые при этом стратегии обычно носят амбициозный характер, поэтому цели компании в таком случае не направлены на пролонгацию сложившейся динамики. Правильный анализ стратегической ситуации также важен для прогнозирования банкротства. Существует множество методов для ее оценки, в том числе и на основе анализа больших данных [67]. Но это отдельная категория прогнозов, и в рамках данной работы рассмотрена не будет.

Данная работа посвящена различным методам машинного обучения для решения задачи прогнозирования банкротства и их сравнению с традиционными методами статистики.

Обзор методов прогнозирования

Один из первых подходов к анализу вероятности банкротства была Z-оценка Альтмана. В своем исследовании [29] он использовал мультивариативный дискриминантный анализ для изучения 66 компаний. Его Z-оценка представляет собой линейную комбинацию следующих показателей компании:

1. Текущие активы / общие активы.
2. Нераспределённая прибыль / общие активы.
3. Прибыль до уплаты налогов / общие активы.
4. Рыночная стоимость компании / общая сумма долга.
5. Продажи/ общие активы.

Альтман вывел формулу, которая учитывает показатели компании и выдает оценку, по которой определяется финансовое состояние компании, и то, насколько вероятно, что она станет банкротом.

В 1980 году американский специалист Ольсон стал одним из первых, кто применил метод логистической регрессии к оценке вероятности дефолта компании [61]. Его модель основана на анализе большего числа признаков, чем модель Альтмана:

1. Натуральный логарифм отношения совокупных активов к индексу-дефлятору валового внутреннего продукта (ВВП).
2. Совокупные обязательства.
3. Совокупные активы.
4. Рабочий капитал.
5. Краткосрочные обязательства.
6. Оборотные активы.
7. Чистая прибыль.
8. Чистая прибыль + Амортизация.

Модель Ольсона дает более интерпретируемый результат, который находится в промежутке между 0 и 1 и является вероятностью наступления дефолта. В дискриминантных моделях вероятность банкротства не имеет номинального значения. Поэтому, в логистических, в отличие от дискриминантных, отсутствуют «серые» зоны, то есть такие зоны, при попадании в которые невозможно однозначно сделать вывод о вероятности банкротства.

Дискриминантный анализ и логистическая регрессия использовались в дальнейшем широко [43] [73], но при этом прогнозные модели строились на относительно малых объемах данных [44]. Современные вычислительные мощности позволяют строить модели на анализе больших данных, используя машинное обучение [76].

Процесс создания прогнозных моделей на основе машинного обучения начинается с подготовки данных. Далее выбираются метрики для оценки результатов, подбираются алгоритмы и определяются их гиперпараметры, с помощью которых оптимизируется их работа. Этап работы с данными состоит не только из их очистки, при котором из выборки удаляются выбросы - значения, сильно выбивающиеся из общего тренда или распределения, но также здесь проводят нормализацию и стандартизацию для численных данных, кодирование для категориальных данных, например, перевод значений “М” и “Ж” признака “Пол” в числа 0 и 1 [31]. Это необходимо, так как большинство алгоритмов машинного обучения способно работать только с численными данными. Дополнительно из данных можно получать метаданные путем поиска комбинаций исходных в выборке. Процесс создания новых признаков из уже имеющихся называется feature engineering [39].

Когда данные обработаны, их разделяют на две группы: целевой признак, который содержит ответы, например, банкрот/не банкрот, и признаки, которые будут обрабатываться алгоритмом машинного обучения для прогнозирования целевого признака [56].

Для запуска, обучения и тестирования моделей всю выборку делят на несколько частей: тренировочные, проверочные и тестовые. С помощью первой обучают модель, вторая необходима для корректировки и поиска наилучших параметров алгоритма, а с помощью последней, которая представляет собой данные, которые модель ни разу не видела, оценивают работу построенной модели. Далее проводится поиск оптимального алгоритма и определение таких его параметров, которые позволяют получить наиболее точный прогноз.

Среди последних исследований по теме прогнозирования банкротства много работ посвящено сравнению стандартных моделей, основанных на методах статистики, и современных моделей на основе машинного обучения [6][32]. Так, например, в работе [32, с. 405-417] авторы исследовали показатели американских и канадских компаний с 1985 по 2013. Результаты исследования показали, что точность прогнозов, полученных с помощью машинного обучения в среднем выше на 10-20%, чем у моделей на основе дискриминантного анализа и логистической регрессии. При этом авторы отмечают, что для построения моделей они не проводили специальную работу над данными: стандартизацию или создание новых признаков, так как в рамках данного исследования им было важно изучить возможность применения алгоритмов обучения без обработки данных. Однако они подчеркивают, что тщательная работа с признаками – это самый перспективный способ улучшения прогнозных моделей на основе машинного обучения.

Один из вариантов работы с признаками был предложен в статье [53, с. 161-188], где австралийские специалисты Стюарт Джонс и Тим Ванг представили новый алгоритм машинного обучения «TreeNet». Имея в распоряжении данные по 4,5 миллионам компаний, они построили 3 модели: одну бинарную и два классификатора, которые определяют несколько состояний рассматриваемой компании: действующая, на стадии банкротства, обанкротившаяся, прекратившая свою деятельность не по причине банкротства. К стандартным показателям компании авторы дополнительно добавили внешние факторы риска: макроэкономические параметры (рост безработицы, инфляцию и т.д.), а также другие – нефинансовые показатели компаний (капитал компании на 1 сотрудника, доход на 1 сотрудника, период погашения дебиторской задолженности и т.д.). В результате, добавление новых признаков позволило существенно увеличить точность построенных прогнозных моделей.

Помимо расширения пространства признаков путем включения в модели дополнительных данных существует другой подход к работе с признаками, который был представлен в статье [37] и называется «label proportions». Он заключается в том, что данные приводятся в виде групп, для каждой из которых вычисляется доля каждого класса, например, банкрот или не банкрот. Таким образом, целью обучения является предсказать класс для каждой группы.

В рамках исследований по машинному обучению специалисты строят множество моделей, основанных на разных алгоритмах, чтобы понять, какой из них лучше подойдет для решения поставленной задачи. Наиболее популярными для задачи прогнозирования банкротства являются нейронные сети [28] [35] [45], случайный лес [6] [25] [32], метод опорных векторов [51] [64] [69], а также ансамбли алгоритмов [60] [72].

Приведенные алгоритмы могут быть полезны не только для создания прогнозных моделей, но и для обнаружения скрытых нелинейных связей между признаками, а также для построения более сложных алгоритмов. Так, в работе [28, с.110-122] приведено сравнение моделей на основе алгоритма AdaBoost, который представляет собой группу простых классификаторов, и нейронных сетей. В исследовании использовались данные европейских компаний, среди которых в качестве основных авторы выделили размер компании, ее финансовые показатели, структуру собственности и вид деятельности. В результате было показано, что с помощью алгоритма AdaBoost величина ошибки прогноза снижается почти на 30%.

Несмотря на широкое применение технологий анализа больших данных и машинного обучения существуют некоторые ограничения и сложности в создании интеллектуальных прогнозных моделей. Одной из проблем является несбалансированность данных в исследуемой выборке. Причина заключается в том, что компаний-банкротов намного меньше, чем финансово устойчивых [11]. Дисбаланс классов может привести к искажению метрик алгоритмов и неправильной интерпретации результатов исследования. В работе [11] описаны 2 способа сбалансировать классы в наборе данных. Первый заключается в сокращении большего по численности наблюдений класса и называется

“андерсемплингом”. Второй, соответственно “оверсемплинг”, предполагает копирование наблюдений миноритарного класса. Так, во многих исследованиях авторы обычно прибегают к одной из приведенных методик балансировки классов [6] [11] [30]. Однако, специалистами из Словакии был предложен еще один способ работы с диспропорцией данных. Они предложили модель с использованием метода одноклассовой классификации [77]. В этом случае обучение идет только на том классе, данных по которому больше, а аномалии, получаемые в результате работы модели, помечаются как другой класс.

Другим важным ограничением в применении и внедрении моделей машинного обучения является сложность алгоритмов, на которых они построены. Для не специалистов в области математики и науки о данных большинство алгоритмов машинного обучения – это черный ящик, которому на вход дают данные, а на выходе получают результат. Такие модели плохо интерпретируемы, и бизнес-заказчикам трудно полностью доверить им принятие решений, так как не до конца понятно, как они работают [62, с. 464–473].

Также, большинство прогнозных моделей, которые разработаны на основе данных бухгалтерского анализа, основаны на допущении, что публикуемая финансовая отчетность является полностью достоверной и справедливой. Но в действительности некоторые компании могут специально исказить истинную картину своего финансового состояния в целях сохранения инвестиционной привлекательности или изменения восприятия риска для партнеров. Так, в работе [50] предлагается способ анализа степени влияния таких манипуляций с данными на модели прогнозирования. Для решения этой задачи авторы дополнили модели новым признаком - мерой искажения финансовой отчетности. Эта мера основывается на показателе доходности активов компании. В результате, было показано, что классификаторы с этим новым признаком, который учитывает возможные искажения отчетности компании, дают более точный прогноз, чем те, которые основаны только на чистых данных.

Большинство работ по этой теме сосредоточены на оптимизации старых или создании новых методов достижения наибольшей точности прогнозных моделей. Однако, важно также учитывать и оценивать издержки, связанные с ошибками моделей прогнозирования. В исследовании [38, с. 612-630] представлена методология, основанная на гетерогенном выборе ансамбля алгоритмов и многоцелевой оптимизации затрат из-за ошибочной работы классификатора. Авторы проанализировали компании Франции, Бельгии, Италии и поделили их на группы по 7 отраслям для тестирования предложенного решения. Построив ансамбли классификаторов, отобранных с помощью методов многокритериальной оптимизации, специалисты добились уменьшения издержек от ошибок работы моделей.

Прогнозные модели применительно к российскому рынку

В России модели прогнозирования банкротства начинают развиваться только в 1990-е годы [6]. Первые исследования представляли собой попытки адаптировать к российскому рынку зарубежные модели, которые строились на анализе американских, канадских и европейских компаний [5]. Однако полученные таким образом модели обладали низкой прогнозной силой и требовали существенных дальнейших доработок [10], связанных с более правильным подбором переменных для анализа согласно российскому законодательству о банкротстве и системы бухгалтерского учета нашей страны [26].

При этом в большинстве российских исследований авторы стараются сосредоточить внимание на прогнозировании банкротства в рамках одной отрасли [21] или одного сегмента бизнеса [2], так как полагают, что для разных отраслей и типов компаний следует выбирать для анализа и разные наборы данных, методы и алгоритмы [22].

Также особую значимость при построении прогнозных моделей имеет и обработка внешних данных. Так, в работе [25] авторы проанализировали российские компании строительной, торговой отрасли и отрасли обрабатывающего производства для прогнозирования их финансового состояния. Им удалось выявить важные внешние факторы, а именно: прирост валового внутреннего продукта, индекс потребительских цен, прирост курса доллара США, ставка центрального банка, прирост индекса московской биржи, а также уровень инфляции и безработицы.

В исследовании [6], где представлен анализ компаний малого и среднего бизнеса, основной задачей было сравнение подходов к прогнозированию критического состояния, при котором компания не способна исполнить все свои обязательства и вынуждена стать банкротом. Исследователи проанализировали финансовые показатели и внешние нефинансовые факторы около 1 млн российских компаний из отрасли торговли, строительства, обрабатывающего производства и недвижимости. В результате, построив несколько прогнозных моделей с помощью методов дискриминантного анализа,

логистической регрессии, а также деревьев решений и алгоритма случайного леса, авторы обнаружили сильное влияние нефинансовых факторов таких, как отрасль, федеральный округ, где предприятие зарегистрировано, а также возраст самого предприятия.

Как отмечается в работе [10] российские исследования по данной теме сильно уступают зарубежным. Основным недостатком являются малые выборки исследуемых данных для моделей. В большинстве работ выборки имеют объем менее 1 тысячи наблюдений. Из-за этого становится неясно, как будет вести себя модель на больших данных, и какая у нее будет при этом точность. Также, изучив данные из базы СПАРК-Интерфакс за период с 2006 по 2010 гг., авторы обнаружили значительную долю на российском рынке компаний-однодневок. Добавление таких компаний в выборку для составления модели может негативно повлиять на ее точность. Другими отрицательными факторами, которые присутствуют на российском рынке, являются нелегальная деятельность компаний, случаи криминальных банкротств и плохое качество данных, которое заключается в том, что не все фирмы в полной мере предоставляют финансовые данные [6], а некоторые намеренно искажают их, чтобы извлечь выгоду [10] [50].

Таким образом, построение прогнозных моделей для российских компаний является актуальной задачей, так как не все зарубежные модели удается успешно адаптировать под специфику российской экономики и законодательства.

Заключение

В данной работе были представлены существующие подходы к задаче прогнозирования банкротства компаний, начиная от стандартных методов статистики и заканчивая современными алгоритмами машинного обучения. Как показали приведенные в работе исследования, применение машинного обучения позволило строить модели, предоставляющие более точный прогноз.

Задача прогнозирования является комплексной, поэтому пути решения заключаются не только в выборе подходящего алгоритма. Также важно дополнительно проводить работу с данными, включать в анализ внешние данные, генерировать метаданные путем извлечения ценных артефактов из существующих, также искать новые способы и приемы их обработки.

Компаниям важно правильно оценивать свои показатели, чтобы принимать эффективные управленческие решения. Также важно проверять тех, с кем предстоит сотрудничать, то есть иметь инструмент проверки финансового состояния партнёров, заемщиков, поставщиков и т. д. Поэтому, спрос на подобные модели классификации, прогнозирования, оценки будет только расти [69].

Литература

1. Ариничев И. В., Матвеева Л. Г., Ариничева И. В. Прогнозирование банкротства организации на основе метрических методов интеллектуального анализа данных // Journal of Economic Regulation (Вопросы регулирования экономики). – 2018. – Т. 9. – №. 1.
2. Ариничев И.В., Богдашев И.В. Оценка риска банкротства субъектов малого предпринимательства на основе методов машинного обучения // Вестник РУДН. Серия: Экономика. 2017. №2.
3. Водошнянова Д. В., Уродовских В. Н. Аналитический обзор методов прогнозирования вероятности банкротства предприятий // Концепт. – 2015. – №. 11.
4. Горбатков С. А., Белолищев И. И., Макеева Е. Ю. Выбор системы экономических показателей для диагностики и прогнозирования банкротств на основе нейросетевого байесовского подхода // Финансы: Теория и Практика. – 2013. – №. 4.
5. Демешев Б. Б., Тихонова А. С. Динамика прогнозной силы моделей банкротства для средних и малых российских компаний оптовой и розничной торговли // Корпоративные финансы. – 2014. – Т. 8. – №. 3.
6. Демешев Б. Б., Тихонова А. С. Прогнозирование банкротства российских компаний: межотраслевое сравнение // Экономический журнал Высшей школы экономики. – 2014. – Т. 18. – №. 3.
7. Евстропов М. В. Оценка эффективности моделей прогнозирования банкротства предприятий // Экономический анализ: теория и практика. – 2008. – №. 13.

8. Жеребилова А. Б. Методики предсказания банкротства организации в рамках построения прогнозно-функциональной системы // Государственное и муниципальное управление. Ученые записки СКАГС. – 2018. – №. 2.
9. Зеленков Ю. А., Федорова Е. А. Двух-шаговый метод построения неоднородного ансамбля классификаторов для прогнозирования банкротства.
10. Казаков А. В., Кольшкнн А. В. Разработка моделей прогнозирования банкротства в современных российских условиях // Вестник Санкт-Петербургского университета. Экономика. – 2018. – №. 2.
11. Карминский А. М., Бурехин Р. Н. Сравнительный анализ методов прогнозирования банкротств российских строительных компаний // Бизнес-информатика. – 2019. – Т. 13. – №. 3.
12. Макаров А. М., Гимазетдинов Д. Р., Кавченков Е. А. Выбор модели прогнозирования риска банкротства для российских предприятий, функционирующих в сфере внешнеэкономической деятельности // Вестник Удмуртского университета. Серия «Экономика и право». – 2017. – Т. 27. – №. 1.
13. Макеева Е. Ю., Аршавский И. В. Применение нейронных сетей и семантического анализа для прогнозирования банкротства // Корпоративные финансы. 2014. №4.
14. Матвеевский С. С. Опыт Японии по использованию аналитики больших данных для снижения кредитного риска при финансировании малых и средних предприятий // Вестник университета. – 2019. – №. 10.
15. Никифорова Н. А., Донцова Л. В., Донцов Е. В. Интеллектуальный анализ данных в моделировании финансового состояния предприятий // Финансовый журнал. – 2011. – №. 2.
16. Опекунов А. Н., Кузьмина М. Г. Принципы формирования моделей прогнозирования вероятности банкротства предприятий с использованием элементов машинного обучения // Модели, системы, сети в экономике, технике, природе и обществе. – 2019. – №. 4.
17. Опекунов А. Н., Никитина Н. В., Камардина Н. В. Реализация принципов машинного обучения при построении моделей прогнозирования банкротства предприятий // Проблемы развития предприятий: теория и практика. – 2019. – №. 1-1. – С. 152-156.
18. Синельникова-Мурылева Е. В., Горшкова Т. Г., Макеева Н. В. Прогнозирование дефолтов в российском банковском секторе // Экономическая политика. – 2018. – Т. 13. – №. 2.
19. Уродовских В. Н., Бахаева А. А. Об адекватности моделей оценки риска банкротства отечественных предприятий // Социально-экономические явления и процессы. – 2010. – №. 6.
20. Фёдорова Е. А., Гиленко Е. В., Довженко С. Е. Модели прогнозирования банкротства: особенности российских предприятий // Проблемы прогнозирования. – 2013. – №. 2.
21. Федорова Е. А., Довженко С. Е., Тимофеев Я. В. Какая модель лучше прогнозирует банкротство российских предприятий? // Экономический анализ: теория и практика. – 2014. – №. 41 (392).
22. Фёдорова Е. А., Довженко С. Е., Фёдоров Ф. Ю. Модели прогнозирования банкротства российских предприятий: отраслевые особенности // Проблемы прогнозирования. – 2016. – №. 3 (156).
23. Федорова Е. А., Лазарев М. П., Федин А. В. Прогнозирование банкротства предприятия с учетом факторов внешней среды // Финансовая аналитика: проблемы и решения. – 2016. – №. 42 (324).
24. Федорова Е. А., Мусиенко С. О., Федоров Ф. Ю. Анализ влияния внешних факторов на прогнозирование финансовой несостоятельности российских компаний // Вестник Санкт-Петербургского университета. Экономика. – 2020. – Т. 36. – №. 1.
25. Федорова Е. А., Мусиенко С. О., Федоров Ф. Ю. Прогнозирование банкротства субъектов малого и среднего предпринимательства в России // Финансы и кредит. – 2018. – Т. 24. – №. 11 (779).
26. Федорова Е. А., Чухланцева М. А., Чекизов Д. В. Оценка эффективности прогнозирования банкротства предприятий на основе российского законодательства // Финансы и кредит. – 2017. – Т. 23. – №. 13 (733).
27. Федотова О. В. Прогнозирование банкротства предприятия // Молодой ученый. – 2018. – №. 4. – С. 122-124.
28. Alfaro E., Garcia N., Gamez M., Elizondo D. Bankruptcy forecasting: An empirical comparison of AdaBoost and neural networks // Decision Support Systems 45 – 2008. – p. 110-122.
29. Altman E. Financial Ratios, Discriminant Analysis and the Prediction of Corporate Bankruptcy // The Journal of Finance 23 (4) – 1968. – p. 589-609.
30. Alrasheed D., Che D. Improving Bankruptcy Prediction Using Oversampling and Feature Selection Techniques // Proceedings on the International Conference on Artificial Intelligence (ICAI). – The

- Steering Committee of The World Congress in Computer Science, Computer Engineering and Applied Computing (WorldComp), 2018. – p. 440-446.
31. Antunes F., Ribeiro B., Pereira F. Probabilistic modeling and visualization for bankruptcy prediction // *Applied Soft Computing* 60 – 2017. – p. 831-843.
 32. Barboza F., Kimura H., Altman E. Machine learning models and bankruptcy prediction // *Expert Systems with Applications* 83 – 2017. – p. 405-417.
 33. Blanco-Oliver A., Irimia-Dieguez A., Oliver-Alfonso M., Wilson N. Improving Bankruptcy Prediction in Micro-Entities by Using Nonlinear Effects and Non-Financial Variables // *Czech Journal of Economics and Finance* 65 (2) – 2015. – p. 144-165.
 34. Bredart X. Bankruptcy Prediction Model Using Neural Networks // *Accounting and Finance Research* 3, No. 2 – 2014. – p. 124-128.
 35. Burnell P., Folarin B. The Impact of Neural Networks in Finance // *Neural Comput & Applic* 6 – 1997. – p. 193-200.
 36. Chen J. M. Models for Predicting Business Bankruptcies and Their Application to Banking and Financial Regulation // *Penn St. L. Rev.* – 2018. – T. 123. – p. 735.
 37. Chen Z., Chen W., Shi Y. Ensemble learning with label proportions for bankruptcy prediction // *Expert Systems With Applications* 146 – 2020.
 38. De Bock K., Coussement K., Lessmann S. Cost-sensitive business failure prediction when misclassification costs are uncertain: A heterogeneous ensemble selection approach // *European Journal of Operational Research* 285 – 2020. – p. 612-630.
 39. Gogas P., Papadimitriou T., Agrapetidou. Forecasting bank failures and stress testing: A machine learning approach // *International Journal of Forecasting* 34 – 2018. p. 440-455.
 40. Griffin J. M., Lemmon M. L. Book-to-market equity, distress risk, and stock returns // *The Journal of Finance.* – 2002. – T. 57. – №. 5. – p. 2317-2336.
 41. Gubanov D. et al. E-expertise: modern collective intelligence. – Springer International Publishing, 2014.
 42. Hong Hanh L. E., VIVIANI J. L. Predicting Bank Failure: Statistical Technique versus Intelligent Technique.
 43. Hardinata L., Warsito B. Bankruptcy prediction based on financial ratios using Jordan Recurrent Neural Networks: a case study in Polish companies // *Journal of Physics: Conf. Series* 41025 – 2018. – 012098
 44. Hillegeist S. A. et al. Assessing the probability of bankruptcy // *Review of accounting studies.* – 2004. – T. 9. – №. 1. – p. 5-34.
 45. Hosaka T. Bankruptcy prediction using imaged financial ratios and convolutional neural networks // *Expert Systems With Applications* 117 – 2019. – p. 287-299.
 46. Iturriaga F., Sanz. I. Bankruptcy visualization and prediction using neural networks: A study of U.S. commercial banks // *Expert Systems with Applications* 42 – 2015. - p. 2857-2869.
 47. Jardin P. Bankruptcy prediction and neural networks: The contribution of variable selection methods. // *Proceedings of the Second European Symposium on Time Series Prediction, Helsinki University of Technology, Porvoo, Finland* – 2008. - p. 271-284.
 48. Jardin P. Failure pattern-based ensembles applied to bankruptcy forecasting // *Decision Support Systems* 107 – 2018. – p. 64-77.
 49. Jardin P. Predicting bankruptcy using neural networks and other classification methods: The influence of variable selection techniques on model accuracy // *Neurocomputing* 73, issue 10-12 – 2010. – p. 2047-2060.
 50. Jardin P. Veganzones D., Severin E. Forecasting Corporate Bankruptcy Using Accrual-Based Models // *Comput Econ* 54 – 2019 – p. 7-43.
 51. Jayanthi. J, Suresh Joseph. K, Vaishnavi. J. Bankruptcy Prediction using SVM and Hybrid SVM Survey // *International Journal of Computer Applications* 33(7) – 2011. – p. 39-45.
 52. Jo N.O., Shin K.S. Bankruptcy Prediction Modeling Using Qualitative Information Based on Big Data Analytics // *J Intell Inform Syst.* 22(2) – 2016. – p. 33-56.
 53. Jones S., Wang T. Predicting private company failure: A multi-class analysis // *J. Int. Finance. Markets Inst. Money* 61 – 2019. – p.161-188.
 54. Kasgari A. A. et al. Prediction of bankruptcy Iranian corporations through artificial neural network and Probit-based analyses // *Neural Computing and Applications.* – 2013. – T. 23. – №. 3-4. – p. 927-936.
 55. Kasgari A. A. et al. The bankruptcy prediction by neural networks and logistic regression // *International Journal of Academic Research in Accounting, Finance and Management Sciences.* – 2013. – T. 3. – №. 4. – p. 146-152.

56. Kovacova M., Kliestikova J. Modelling bankruptcy prediction models in Slovak companies // SHS Web of Conferences 39 – 2017 – 01013
57. Mai F. et al. Deep learning models for bankruptcy prediction using textual disclosures // European journal of operational research. – 2019. – Т. 274. – №. 2. – С. 743-758.
58. Meese E. N., Viken T. Machine learning in bankruptcy prediction: utilizing machine learning for improved bankruptcy predictions in the Norwegian market with an emphasis on financial, management and sector statements: Master's thesis. – 2019.
59. Nagaraj K., Sridhar A. A predictive system for detection of bankruptcy using machine learning techniques // International Journal of Data Mining & Knowledge Management Process Vol. 5 No. 1. – 2015.
60. Nanni L., Lumini A. An experimental comparison of ensemble of classifiers for bankruptcy prediction and credit scoring // Expert Systems With Applications 36 – 2009. – p. 3028–3033.
61. Ohlson J. A. Financial ratios and the probabilistic prediction of bankruptcy // Journal of accounting research. – 1980. – p. 109-131.
62. Olson D., Delen D., Meng Y. Comparative analysis of data mining methods for bankruptcy prediction // Decision Support Systems 52 – 2012. – p. 464–473.
63. Pena T., Martínez S., Abudu B. Bankruptcy prediction: A comparison of some statistical and machine learning techniques // Computational Methods in Economic Dynamics. – Springer, Berlin, Heidelberg, 2011. – p. 109-131.
64. Petropoulos A., Siakoulis V., Stavroulakis V., Vlachogiannakis N. Predicting bank insolvencies using machine learning techniques // International Journal of Forecasting
65. Pompe P.M., Feelders A.J., Using Machine Learning, Neural Networks, and Statistics to Predict Corporate Bankruptcy // Microcomputers in Civil Engineering 12 – 1997. – p. 267–276.
66. Pozorska J., Scherer M. Company Bankruptcy Prediction with Neural Networks // Artificial Intelligence and Soft Computing – 2018. – p.183-189
67. Raikov A. Manufacturer's Strategic Risk Temperature Assessment with Convergent Approach, Cognitive Modelling and Blockchain Technology // IFAC-PapersOnLine. – 2019. – Т. 52. – №. 13. – С. 1289-1294.
68. Shi Y., Li X. An overview of bankruptcy prediction models for corporate firms: A systematic literature review // Intangible Capital 15(2) – 2019.
69. Shin K.S., Lee T.S., Kim H.J. An application of support vector machines in bankruptcy prediction model // Expert Systems with Applications 28 – 2005. – p. 127–135.
70. Soui M., Smiti S., Mkaouer M.W., Ejbali R. Bankruptcy Prediction Using Stacked Auto-Encoders // Applied Artificial Intelligence 2019. – p. 80-100.
71. Suss J., Treitel H. Predicting bank distress in the UK with machine learning // Bank of England Staff Working Paper No. 831 – 2019.
72. Tsai C.F., Wu J.W. Using neural network ensembles for bankruptcy prediction and credit scoring // Expert Systems with Applications 34 – 2008. – p. 2639–2649.
73. Tseng F. M., Lin L. A quadratic interval logit model for forecasting bankruptcy // Omega. – 2005. – Т. 33. – №. 1. – p. 85-91.
74. Wagenmans F. Machine learning in bankruptcy prediction: Master's thesis. – 2017
75. Wang G., Ma J., Yang S. An improved boosting based on feature selection for corporate bankruptcy prediction // Expert Systems with Applications 41 – 2014. – p. 2353–2361.
76. Wang N. Bankruptcy Prediction Using Machine Learning // Journal of Mathematical Finance 7 – 2017. – p. 908-918.
77. Zoricak M., Gnip P., Drotar P., Gazda V. Bankruptcy prediction for small- and medium-sized companies using severely imbalanced datasets // Economic Modelling 84 – 2020. – p. 165-176.

BANKRUPTCY PREDICTION USING MACHINE LEARNING METHODS

Dorzhiiev Ardan Sayanovich

*Financial University under the Government of the Russian Federation, Department of Information Technologies and Big Data Analysis, Master's degree student
Moscow, Russia
a.s.dorzhiiev@mail.ru*

Abstract

Bankruptcy prediction models are important to many stakeholders starting from associates up to shareholders, creditors and partners. The article contains an overview of both traditional statistic prediction methods and modern ones with the use of machine learning.

Keywords

bankruptcy prediction, Z-score, discriminant analysis, machine learning, neural networks, random forest, support vector machine

References

1. Arinichev I. V., Matveeva L. G., Arinicheva I. V. Prognozirovaniye bankrotstva organizatsii na osnove metricheskikh metodov intellektual'nogo analiza dannykh // Journal of Economic Regulation (Voprosy regulirovaniya ekonomiki). – 2018. – T. 9. – № 1.
2. Arinichev I.V., Bogdashev I.V. Ocenka riska bankrotstva sub'ektov malogo predprinimatel'stva na osnove metodov mashinnogo obucheniya // Vestnik RUDN. Seriya: Ekonomika. 2017. №2.
3. Vodop'yanova D. V., Urodovskih V. N. Analiticheskij obzor metodov prognozirovaniya veroyatnosti bankrotstva predpriyatij //Koncept. – 2015. – № 11.
4. Gorbatkov S. A., Belolipcev I. I., Makeeva E. YU. Vybor sistemy ekonomicheskikh pokazatelej dlya diagnostiki i prognozirovaniya bankrotstv na osnove nejrosetevogo bajesovskogo podhoda //Finansy: Teoriya i Praktika. – 2013. – № 4.
5. Demeshev B. B., Tihonova A. S. Dinamika prognoznoj sily modelej bankrotstva dlya srednih i malyh rossijskikh kompanij optovoj i roznicnoj torgovli //Korporativnye finansy. – 2014. – T. 8. – № 3.
6. Demeshev B. B., Tihonova A. S. Prognozirovaniye bankrotstva rossijskikh kompanij: mezhotraslevoe sravnenie //Ekonomicheskij zhurnal Vysshej shkoly ekonomiki. – 2014. – T. 18. – № 3.
7. Evstropov M. V. Ocenka effektivnosti modelej prognozirovaniya bankrotstva predpriyatij //Ekonomicheskij analiz: teoriya i praktika. – 2008. – № 13.
8. Zherebilova A. B. Metodiki predskazaniya bankrotstva organizatsii v ramkah postroeniya prognozno-funktional'noj sistemy //Gosudarstvennoe i municipal'noe upravlenie. Uchenye zapiski SKAGS. – 2018. – № 2.
9. Zelenkov YU. A., Fedorova E. A. Dvuh-shagovyj metod postroeniya neodnorodnogo ansamblya klassifikatorov dlya prognozirovaniya bankrotstva.
10. Kazakov A. V., Kolyshkin A. V. Razrabotka modelej prognozirovaniya bankrotstva v sovremennykh rossijskikh usloviyah //Vestnik Sankt-Peterburgskogo universiteta. Ekonomika. – 2018. – № 2.
11. Karminskij A. M., Burekhin R. N. Sravnitel'nyj analiz metodov prognozirovaniya bankrotstv rossijskikh stroitel'nykh kompanij // Biznes-informatika. – 2019. – T. 13. – № 3.
12. Makarov A. M., Gimazetdinov D. R., Kavchenkov E. A. Vybor modeli prognozirovaniya riska bankrotstva dlya rossijskikh predpriyatij, funkcioniruyushchih v sfere vneshneekonomicheskoy deyatel'nosti //Vestnik Udmurtskogo universiteta. Seriya «Ekonomika i pravo». – 2017. – T. 27. – № 1.
13. Makeeva E. YU., Arshavskij I. V. Primeneniye nejronnykh setej i semanticheskogo analiza dlya prognozirovaniya bankrotstva // Korporativnye finansy. 2014. №4.
14. Matveevskij S. S. Opyt YAponii po ispol'zovaniyu analitiki bol'shikh dannykh dlya snizheniya kreditnogo riska pri finansirovaniy malyh i srednih predpriyatij // Vestnik universiteta. – 2019. – № 10.
15. Nikiforova N. A., Doncova L. V., Doncov E. V. Intellektual'nyj analiz dannykh v modelirovaniy finansovogo sostoyaniya predpriyatij // Finansovyj zhurnal. – 2011. – № 2.

16. Opekunov A. N., Kuz'mina M. G. Principy formirovaniya modelej prognozirovaniya veroyatnosti bankrotstva predpriyatij s ispol'zovaniem elementov mashinnogo obucheniya // *Modeli, sistemy, seti v ekonomike, tekhnike, prirode i obshchestve*. 2019. №4.
17. Opekunov A. N., Nikitina N. V., Kamardina N. V. Realizaciya principov mashinnogo obucheniya pri postroenii modelej prognozirovaniya bankrotstva predpriyatij // *Problemy razvitiya predpriyatij: teoriya i praktika*. – 2019. – №. 1-1. – S. 152-156.
18. Sinel'nikova-Muryleva E. V., Gorshkova T. G., Makeeva N. V. Prognozirovaniye defoltov v rossijskom bankovskom sektore // *Ekonomicheskaya politika*. – 2018. – T. 13. – №. 2.
19. Urodovskih V. N., Bahaeva A. A. Ob adekvatnosti modelej ocenki riska bankrotstva otechestvennyh predpriyatij // *Social'no-ekonomicheskie yavleniya i processy*. – 2010. – №. 6.
20. Fyodorova E. A., Gilenko E. V., Dovzhenko S. E. Modeli prognozirovaniya bankrotstva: osobennosti rossijskih predpriyatij // *Problemy prognozirovaniya*. – 2013. – №. 2.
21. Fedorova E. A., Dovzhenko S. E., Timofeev YA. V. Kakaya model' luchshe prognoziruet bankrotstvo rossijskih predpriyatij? // *Ekonomicheskij analiz: teoriya i praktika*. – 2014. – №. 41 (392).
22. Fyodorova E. A., Dovzhenko S. E., Fyodorov F. YU. Modeli prognozirovaniya bankrotstva rossijskih predpriyatij: otraslevye osobennosti // *Problemy prognozirovaniya*. – 2016. – №. 3 (156).
23. Fedorova E. A., Lazarev M. P., Fedin A. V. Prognozirovaniye bankrotstva predpriyatija s uchedom faktorov vneshnej sredy // *Finansovaya analitika: problemy i resheniya*. – 2016. – №. 42 (324).
24. Fedorova E. A., Musienko S. O., Fedorov F. YU. Analiz vliyaniya vneshnih faktorov na prognozirovaniye finansovoj nesostoyatel'nosti rossijskih kompanij // *Vestnik Sankt-Peterburgskogo universiteta. Ekonomika*. – 2020. – T. 36. – №. 1.
25. Fedorova E. A., Musienko S. O., Fedorov F. YU. Prognozirovaniye bankrotstva sub"ektov malogo i srednego predprinimatel'stva v Rossii // *Finansy i kredit*. – 2018. – T. 24. – №. 11 (779).
26. Fedorova E. A., CHuhlanceva M. A., CHekrizov D. V. Ocenka effektivnosti prognozirovaniya bankrotstva predpriyatij na osnove rossijskogo zakonodatel'stva // *Finansy i kredit*. – 2017. – T. 23. – №. 13 (733).
27. Fedotova O. V. Prognozirovaniye bankrotstva predpriyatija // *Molodoj uchenyj*. – 2018. – №. 4. – S. 122-124.
28. Alfaro E., Garcia N., Gamez M., Elizondo D. Bankruptcy forecasting: An empirical comparison of AdaBoost and neural networks // *Decision Support Systems* 45 – 2008. – p. 110-122.
29. Altman E. Financial Ratios, Discriminant Analysis and the Prediction of Corporate Bankruptcy // *The Journal of Finance* 23 (4) – 1968. – p. 589-609.
30. Alrasheed D., Che D. Improving Bankruptcy Prediction Using Oversampling and Feature Selection Techniques // *Proceedings on the International Conference on Artificial Intelligence (ICAI)*. – The Steering Committee of The World Congress in Computer Science, Computer Engineering and Applied Computing (WorldComp), 2018. – p. 440-446.
31. Antunes F., Ribeiro B., Pereira F. Probabilistic modeling and visualization for bankruptcy prediction // *Applied Soft Computing* 60 – 2017. – p. 831-843.
32. Barboza F., Kimura H., Altman E. Machine learning models and bankruptcy prediction // *Expert Systems with Applications* 83 – 2017. – p. 405-417.
33. Blanco-Oliver A., Irimia-Dieguez A., Oliver-Alfonso M., Wilson N. Improving Bankruptcy Prediction in Micro-Entities by Using Nonlinear Effects and Non-Financial Variables // *Czech Journal of Economics and Finance* 65 (2) – 2015. – p. 144-165.
34. Bredart X. Bankruptcy Prediction Model Using Neural Networks // *Accounting and Finance Research* 3, No. 2 – 2014. – p. 124-128.
35. Burnell P., Folarin B. The Impact of Neural Networks in Finance // *Neural Comput & Applic* 6 – 1997. – p. 193-200.
36. Chen J. M. Models for Predicting Business Bankruptcies and Their Application to Banking and Financial Regulation // *Penn St. L. Rev.* – 2018. – T. 123. – p. 735.
37. Chen Z., Chen W., Shi Y. Ensemble learning with label proportions for bankruptcy prediction // *Expert Systems With Applications* 146 – 2020.
38. De Bock K., Coussement K., Lessmann S. Cost-sensitive business failure prediction when misclassification costs are uncertain: A heterogeneous ensemble selection approach // *European Journal of Operational Research* 285 – 2020. – p. 612-630.
39. Gogas P., Papadimitriou T., Agrapetidou. Forecasting bank failures and stress testing: A machine learning approach // *International Journal of Forecasting* 34 – 2018. p. 440-455.

40. Griffin J. M., Lemmon M. L. Book-to-market equity, distress risk, and stock returns // *The Journal of Finance*. – 2002. – Т. 57. – №. 5. – p. 2317-2336.
41. Gubanov D. et al. *E-expertise: modern collective intelligence*. – Springer International Publishing, 2014.
42. Hong Hanh L. E., VIVIANI J. L. Predicting Bank Failure: Statistical Technique versus Intelligent Technique.
43. Hardinata L., Warsito B. Bankruptcy prediction based on financial ratios using Jordan Recurrent Neural Networks: a case study in Polish companies // *Journal of Physics: Conf. Series* 41025 – 2018. – 012098
44. Hillegeist S. A. et al. Assessing the probability of bankruptcy // *Review of accounting studies*. – 2004. – Т. 9. – №. 1. – p. 5-34.
45. Hosaka T. Bankruptcy prediction using imaged financial ratios and convolutional neural networks // *Expert Systems With Applications* 117 – 2019. – p. 287-299.
46. Iturriaga F., Sanz. I. Bankruptcy visualization and prediction using neural networks: A study of U.S. commercial banks // *Expert Systems with Applications* 42 – 2015. - p. 2857-2869.
47. Jardin P. Bankruptcy prediction and neural networks: The contribution of variable selection methods. // *Proceedings of the Second European Symposium on Time Series Prediction*, Helsinki University of Technology, Porvoo, Finland – 2008. - p. 271-284.
48. Jardin P. Failure pattern-based ensembles applied to bankruptcy forecasting // *Decision Support Systems* 107 – 2018. – p. 64-77.
49. Jardin P. Predicting bankruptcy using neural networks and other classification methods: The influence of variable selection techniques on model accuracy // *Neurocomputing* 73, issue 10-12 – 2010. – p. 2047-2060.
50. Jardin P. Veganzones D., Severin E. Forecasting Corporate Bankruptcy Using Accrual-Based Models // *Comput Econ* 54 – 2019 – p. 7-43.
51. Jayanthi. J, Suresh Joseph. K, Vaishnavi. J. Bankruptcy Prediction using SVM and Hybrid SVM Survey // *International Journal of Computer Applications* 33(7) – 2011. – p. 39-45.
52. Jo N.O., Shin K.S. Bankruptcy Prediction Modeling Using Qualitative Information Based on Big Data Analytics // *J Intell Inform Syst.* 22(2) – 2016. – p. 33-56.
53. Jones S., Wang T. Predicting private company failure: A multi-class analysis // *J. Int. Finance. Markets Inst. Money* 61 – 2019. – p.161-188.
54. Kasgari A. A. et al. Prediction of bankruptcy Iranian corporations through artificial neural network and Probit-based analyses // *Neural Computing and Applications*. – 2013. – Т. 23. – №. 3-4. – p. 927-936.
55. Kasgari A. A. et al. The bankruptcy prediction by neural networks and logistic regression // *International Journal of Academic Research in Accounting, Finance and Management Sciences*. – 2013. – Т. 3. – №. 4. – p. 146-152.
56. Kovacova M., Kliestikova J. Modelling bankruptcy prediction models in Slovak companies // *SHS Web of Conferences* 39 – 2017 – 01013
57. Mai F. et al. Deep learning models for bankruptcy prediction using textual disclosures // *European journal of operational research*. – 2019. – Т. 274. – №. 2. – C. 743-758.
58. Meese E. N., Viken T. Machine learning in bankruptcy prediction: utilizing machine learning for improved bankruptcy predictions in the Norwegian market with an emphasis on financial, management and sector statements: Master's thesis. – 2019.
59. Nagaraj K., Sridhar A. A predictive system for detection of bankruptcy using machine learning techniques // *International Journal of Data Mining & Knowledge Management Process* Vol. 5 No. 1. – 2015.
60. Nanni L., Lumini A. An experimental comparison of ensemble of classifiers for bankruptcy prediction and credit scoring // *Expert Systems With Applications* 36 – 2009. – p. 3028-3033.
61. Ohlson J. A. Financial ratios and the probabilistic prediction of bankruptcy // *Journal of accounting research*. – 1980. – p. 109-131.
62. Olson D., Delen D., Meng Y. Comparative analysis of data mining methods for bankruptcy prediction // *Decision Support Systems* 52 – 2012. – p. 464-473.
63. Pena T., Martínez S., Abudu B. Bankruptcy prediction: A comparison of some statistical and machine learning techniques // *Computational Methods in Economic Dynamics*. – Springer, Berlin, Heidelberg, 2011. – p. 109-131.
64. Petropoulos A., Siakoulis V., Stavroulakis V., Vlachogiannakis N. Predicting bank insolvencies using machine learning techniques // *International Journal of Forecasting*

65. Pompe P.M., Feelders A.J., Using Machine Learning, Neural Networks, and Statistics to Predict Corporate Bankruptcy // *Microcomputers in Civil Engineering* 12 – 1997. – p. 267–276.
66. Pozorska J., Scherer M. Company Bankruptcy Prediction with Neural Networks // *Artificial Intelligence and Soft Computing* – 2018. – p.183-189
67. Raikov A. Manufacturer’s Strategic Risk Temperature Assessment with Convergent Approach, Cognitive Modelling and Blockchain Technology // *IFAC-PapersOnLine*. – 2019. – T. 52. – №. 13. – C. 1289-1294.
68. Shi Y., Li X. An overview of bankruptcy prediction models for corporate firms: A systematic literature review // *Intangible Capital* 15(2) – 2019.
69. Shin K.S., Lee T.S., Kim H.J. An application of support vector machines in bankruptcy prediction model // *Expert Systems with Applications* 28 – 2005. – p. 127–135.
70. Soui M., Smiti S., Mkaouer M.W., Ejbali R. Bankruptcy Prediction Using Stacked Auto-Encoders // *Applied Artificial Intelligence* 2019. – p. 80-100.
71. Suss J., Treitel H. Predicting bank distress in the UK with machine learning // *Bank of England Staff Working Paper No. 831* – 2019.
72. Tsai C.F., Wu J.W. Using neural network ensembles for bankruptcy prediction and credit scoring // *Expert Systems with Applications* 34 – 2008. – p. 2639–2649.
73. Tseng F. M., Lin L. A quadratic interval logit model for forecasting bankruptcy // *Omega*. – 2005. – T. 33. – №. 1. – p. 85-91.
74. Wagenmans F. Machine learning in bankruptcy prediction: Master’s thesis. - 2017
75. Wang G., Ma J., Yang S. An improved boosting based on feature selection for corporate bankruptcy prediction // *Expert Systems with Applications* 41 – 2014. – p. 2353–2361.
76. Wang N. Bankruptcy Prediction Using Machine Learning // *Journal of Mathematical Finance* 7 – 2017. – p. 908-918.
77. Zoricak M., Gnip P., Drotar P., Gazda V. Bankruptcy prediction for small- and medium-sized companies using severely imbalanced datasets // *Economic Modelling* 84 – 2020. – p. 165-176.