

**ФЕДЕРАЛЬНОЕ ГОСУДАРСТВЕННОЕ БЮДЖЕТНОЕ УЧРЕЖДЕНИЕ
НАУКИ ЦЕНТРАЛЬНЫЙ ЭКОНОМИКО-МАТЕМАТИЧЕСКИЙ
ИНСТИТУТ РОССИЙСКОЙ АКАДЕМИИ НАУК**

На правах рукописи

ЖУКОВА ЛЮДМИЛА ВЯЧЕСЛАВОВНА

**ЭКСПРЕСС-АНАЛИЗ СООТВЕТСТВИЯ СОЦИАЛЬНО-ЭКОНОМИЧЕСКОГО
СОСТОЯНИЯ ПРЕДПРИЯТИЯ НОРМАТИВНЫМ ТРЕБОВАНИЯМ НА ОСНОВЕ
ОБЩЕДОСТУПНОЙ ИНФОРМАЦИИ**

Специальность:

**08.00.05 «Экономика и управление народным хозяйством», специализация:
«Экономика, организация и управление предприятиями, отраслями,
комплексами (промышленность)» (экономические науки)**

Диссертация

**на соискание ученой степени
кандидата экономических наук**

Научный руководитель:

**кандидат экономических наук,
доцент Богданова Т.К.**

Москва – 2022

Оглавление

Введение.....	5
Глава 1 . Модели оценки социально-экономического состояния объекта исследования в России и за рубежом.....	24
1.1 Проблемы управления финансово-хозяйственной деятельностью объекта исследования.....	24
1.2 Переход к цифровой экономике: использование общедоступных источников информации.....	27
1.3 Особенности современной системы контроля и надзора в РФ	31
1.3.1 Полномочия органов государственного контроля	31
1.3.2 Проблемы системы государственного и внутрифирменного контроля	33
1.4 Классификация подходов и моделей оценки социально – экономического состояния объекта исследования	37
1.5 Постановка проблемы исследования.....	43
Глава 2 . Концептуальная модель экспресс-анализа соответствия состояния объекта исследования требованиям контрольных и надзорных органов.....	48
2.1 Поэтапная схема построения интегрального показателя экспресс-анализа соответствия социально-экономического состояния объекта исследования заявленным требованиям	48
2.2 Типологизация источников общедоступной информации и методы их обработки.....	53
2.2.1 Типологизация источников общедоступной информации	53
2.2.2 Методы обработки общедоступной информации	59
2.3 Алгоритм построения интегрального показателя для проведения экспресс-анализа.....	61
2.3.1 Общая схема алгоритма построения интегрального показателя экспресс-анализа	61

2.3.2	Компонента 1. Вероятность финансового неблагополучия	65
2.3.3	Компоненты 2 и 3. Статус объекта исследования по масштабу и принадлежности к аномальной группе	68
2.3.4	Компоненты 4, 5 и 6. Медийная активность относительно объекта исследования, положительная и негативная тональность упоминаний об объекте исследования в интернет-источниках.	73
2.3.5	Компонента 7. Соответствие требованиям государственных органов.	75
2.4	Информационно-логическая модель экспресс-анализа соответствия состояния объекта исследования нормативам и регламентам со стороны контрольных и надзорных органов на основе общедоступной информации.....	78
Глава 3	Экспресс-анализ социально-экономического состояния предприятий из большой совокупности на соответствие нормативным требованиям.....	82
3.1	Экспресс-анализ соответствия социально-экономического состояния промышленного предприятия, являющегося элементом большой совокупности предприятий, нормативным требованиям департамента инвестиционной политики г. Москвы.	82
3.1.1	Общая поэтапная схема построения интегрального показателя экспресс-анализа соответствия социально-экономического состояния промышленного предприятия заявленным требованиям	83
3.1.2	Информационная база исследования.....	87
3.1.3	Оценка компонент интегрального показателя.....	90
3.1.4	Расчет интегрального показателя экспресс- анализа.....	99
3.1.5	Анализ полученных результатов.....	101
3.2	Экспресс-анализ соответствия социально-экономического состояния коммерческого банка нормативным требованиям ЦБ РФ.	104
3.2.1	Общая поэтапная схема построения интегрального показателя экспресс-анализа соответствия социально-экономического состояния коммерческого банка заявленным требованиям	104

3.2.2	Информационная база исследования.....	107
3.2.3	Оценка компонент интегрального показателя.....	108
3.2.4	Расчет интегрального показателя экспресс- анализа.....	111
3.2.5	Анализ полученных результатов.....	112
	Заключение	114
	Список использованной литературы.....	117
	Приложение 1. Типологизация источников общедоступной информации ..	127
	Приложение 2. Перечень используемых обозначений для базы данных промышленных предприятий	129
	Приложение 3. Значения среднего и вариации основных финансовых показателей для выборки промышленных предприятий за 2016-2018 гг.	131
	Приложение 4. Перечень предприятий группы 1. «Доминирующие предприятия».....	132
	Приложение 5. Сравнительная таблица средних значений финансовых показателей группы 2 и группы 3.	133
	Приложение 6. Результаты проверки непараметрического критерия Манна – Уитни о равенстве распределений между группами 2 и 3 для промышленных предприятий	135
	Приложение 7. Результаты проверки непараметрического критерия Манна – Уитни о равенстве распределений между подгруппами 1 и 2 группы 2 промышленных предприятий.....	136
	Приложение 8. Результаты проверки непараметрического критерия Манна – Уитни о равенстве распределений между подгруппами 1 и 2 группы 3 промышленных предприятий.....	137
	Приложение 9. Перечень используемых обозначений для базы данных коммерческого банка.....	138

Введение

Актуальность темы исследования

В современных экономических реалиях роль грамотного управления и контроля в деятельности предприятия любой отрасли экономики нельзя недооценивать. При эффективном стратегическом управлении уменьшаются риски сбоев в работе систем, экономятся человеческие и материальные ресурсы, повышается эффективность деятельности организации, ускоряется возврат средств и снижается уровень социальной напряженности.

Последние 10 лет глобальная и региональные экономики переживают множественные и разнообразные кризисы, возрастает нестабильность макроэкономической среды. Усиление политических разногласий между странами, регионами, государством и частным сектором является предсказуемой реакцией на финансовые кризисы на протяжении разных эпох и в разных странах. Это приводит к необходимости решения таких задач в управлении экономическими объектами как ускоренное реагирование на неожиданные события, мониторинг, оперативный прогноз с учетом информации из внешней среды. С учетом происходящих в последние годы изменений в экономике наблюдается пристальное внимание к способам и методам управления, а также к той информации, на основе которой принимаются решения. Чем подробнее и качественнее является используемая информация, тем точнее будут приняты на ее основе управленческие решения. О проблеме управления экономическими объектами в последние годы было написано достаточно много научных статей как российскими, так и зарубежными исследователями. Если рассматривать контрольную деятельность исполнительных органов власти Российской Федерации, то система контроля и надзора над экономическими объектами, для деятельности которых существуют нормативы и регламенты, также меняется: включаются

новые источники информации, внедряется мониторинг внешней среды, медийного пространства и другие источники открытых данных.

На протяжении долгих лет надзорный контроль осуществлялся на основании официальной статистической отчетности. В нормативных и регулятивных документах, в методиках оценки деятельности экономических объектов (промышленное предприятие, коммерческий банк, школа, регион) устанавливались только количественные показатели за определенный период времени, что могло привести к расхождению между фактическим и документальным состоянием некоторых объектов, к искажению действительности, а, в итоге, — к неверным управленческим решениям. С 2014 года органы государственной власти обращают большое внимание на социальный эффект от деятельности поднадзорных объектов. Число сайтов с отзывами, комментариями и текстовыми постами (официальными и неофициальными) растет. Возникла потребность давать оценку качественных характеристик деятельности объектов. Также возрастает скорость и динамика принятия решений, лаг в поступлении официальной отчетности иногда достигает полугода с момента окончания отчетного периода. Сегодня требуются уже новые источники информации для принятия оперативных решений в отношении поднадзорного экономического объекта. Проведение экспресс-анализа на основе общедоступных данных из сети интернет, позволяющего оценивать соответствие экономического объекта установленным документами регламентам и нормативам, является **актуальной задачей**. И, как следствие, возникает необходимость в разработке концептуального инструментария для проведения такого экспресс-анализа.

Степень разработанности проблемы исследования

За прошедшие 70 лет в экономических исследованиях, посвященных оценке состояния экономических объектов в разных сферах, было создано, апробировано и внедрено значительное число методов и моделей. Все модели

можно условно разделить на три группы: регрессионно-скоринговые, модели интегральных показателей и индикаторов состояния, а также машинного обучения. Для предприятий и организаций применяемые модели можно разделить на прогнозирующие вероятность банкротства и оценивающие общее состояние объекта исследования.

Еще с 1960-х годов в секторе промышленности для оценки вероятности банкротства предприятий с помощью интегрального показателя известна модель Э. Альтмана, построенная на основе множественного дискриминантного анализа (МДА). Эта модель затем была модифицирована различными учеными для разных особых случаев, в том числе для развивающихся стран, для предприятий, не торгующих на бирже, и для оценки состояния предприятий разной отраслевой принадлежности: Р. Таффлером и Г. Тишшоу, Д.А. Колесниковым, О. Зайцевой, А. Шереметом и др. Проблему банкротства предприятий с использованием регрессионных моделей (в том числе, моделей логистической регрессии) для прогнозирования вероятности банкротства решали Ч. Тингуан, О. Пирогова, Д. Олсон, М. Змиевский. Изначально в моделях вероятности банкротства использовались финансовые данные только из официальной внутренней отчетности предприятия. Современные авторы для оценки состояния экономического объекта предлагают методики на основе индикаторов и их аналогов на базе структурированной информации, публикуемой в открытых источниках

Многие отечественные и зарубежные ученые посвятили свои работы оценке состояния предприятий финансовой отрасли (банков). Это — модели оценки вероятности отзыва лицензии (дефолта банка); их описывали и исследовали: А.А. Пересецкий, А.М. Карминский, С.В. Головань, А.В. Костров, К.Л. Поляков, Я. Клаас, Р. Калембесе, В. Руссо, З. Ванг, Й. Динш, Дж. Уонсли, С. Годлевски, А. Карась, Б. Эндельман, А. Эстрелла, М. Макинен и др. Моделированием рейтингов банков занимались А.А. Пересецкий, А.М. Карминский, Е. Хромова, С.А. Айвазян, Э. Альтман, Х.

Риджкен, Е.Л. Головки, В.Г. Сидоров, М. Ливингстон, И. Дистингуин, Н. Ван Темп и др. Применение машинного обучения в моделировании состояния коммерческого банка рассматривалось в работах А.М. Карминского, Д.В. Исаева, А.Л. Богданова, Д. С. Биджояна, Л. Хутао, Й. Сан и др.

Упомянутые выше методы и подходы не дают возможности отслеживать и своевременно принимать меры для финансовой поддержки и защиты поднадзорного экономического объекта в случае наступления финансового или экономического кризиса. Как показали недавние кризисы в экономике РФ 2008 и 2015 гг., многие компании с устойчивым финансовым положением на начало кризиса все же не смогли его преодолеть, так как при классических методах оценки возникает риск не успеть принять соответствующие меры.

Рассмотренные выше работы опираются на данные официальной статистики или на структурированные данные из открытых источников. При этом используется одна общая модель для оценки интегрального показателя. Большинство моделей разработано для своей сферы деятельности (банковский сектор, промышленность, образование и т.п.), что не предполагает их применения для оценки других сфер. Это затрудняет создание общей автоматизированной системы мониторинга и контроля состояния экономического объекта. Также большая часть описанных моделей и подходов ориентирована на текущих или потенциальных владельцев, в то время как модели, направленные на требования регуляторов (отраслевые министерства, ЦБ РФ, руководство холдинга и т.п.), разработаны слабо.

Вопросы качества осуществления аудиторской деятельности в настоящее время приобретают высокую актуальность, так как стоимость аудита возрастает, количество внешних аудиторских организаций сокращается. Существующая в настоящее время схема организации контроля за предприятиями, в большинстве случаев, начинается с инициации внеплановой проверки вышестоящей организацией, или с наступления момента проведения периодической проверки. Информационная база совокупности предприятий,

подлежащих проверке, формируется из документов, предоставленных предприятием, форм, полученных в ответ на запрос аудиторов или статистической отчетности, предоставляемой в соответствии с существующим законодательством. Зачастую возникает существенный лаг между моментом запроса и непосредственным поступлением информации. Численность персонала контрольных и надзорных органов, осуществляющих проверку, часто в десятки раз меньше численности проверяемых предприятий

В отличие от рассмотренных выше публикаций, в работе предлагается концептуальная модель построения интегрального показателя для проведения экспресс-анализа соответствия социально-экономического состояния предприятия требованиям, сформулированным в нормативных документах и регламентах его деятельности, выпущенных соответствующими контрольными органами.

«Социально-экономическое состояние предприятия – комплексная характеристика экономической, финансовой, производственной, управленческой, социальной деятельности и их влияния друг на друга». **«Соответствие социально-экономического состояния объекта исследования** требованиям контрольных и надзорных органов — прямое или косвенное определение соблюдения требований, предъявляемых к объекту»¹. Термин «оценка соответствия» определен как «доказательство того, что заданные требования к продукции, процессу, системе, лицу или органу, выполнены». **«Экспресс-анализ** — быстровыполнимый процесс в масштабе реального времени, представляющий собой качественный и количественный анализ данных об объекте исследования на текущий момент»².

Экспресс-анализ проводится на основе общедоступных структурированных и неструктурированных данных, полученных из

¹ ГОСТ ISO/IEC 17000-2012 (ISO/IEC 17000:2004).

² Ефремова Т.Ф. Толковый словарь (Новый словарь русского языка. Толково-словообразовательный. В 2 т.). М.: Русский язык, 2000. 1084 с.

официальных и неофициальных интернет-ресурсов. Данные собираются как со статистических ресурсов, где они публикуются периодически, так и с открытых интернет-ресурсов, регулярно обновляемых с разной периодичностью. В отличие от используемых в течение многих лет структурированных данных, неструктурированные данные представляют собой объединение разрозненной и разнотипной информации. Это позволяет лицам, принимающим решение в контрольных и надзорных органах, или в контролирующей вышестоящей организации отслеживать на ранней стадии последствия принятых решений, что обеспечивает минимизацию финансовых, социальных и экономических рисков. Использование неструктурированных данных при оценке объектов исследования дает возможность вести непрерывный мониторинг текущего состояния этого объекта с учетом качественных характеристик (репутация, отзывы, тенденции) на основании данных, полученных бесконтактным методом из открытых источников.

Проведение экспресс-анализа в экономике и финансах стало возможным в связи с применением современных ЭВМ.

Таким образом, **актуальность исследования** обусловлена следующими факторами:

- необходимостью разработки гибкого инструментария для текущего мониторинга, анализа и прогноза социально-экономического состояния предприятия на основе общедоступных открытых данных;

- потребностью в текущей модели оценки, не зависящей от периодичности публикаций официальной статистической отчетности;

- необходимостью проведения экспресс-анализа соответствия социально-экономического состояния предприятия, удовлетворяющего нормативам и регламентам со стороны контрольных и надзорных органов с учетом количественных и качественных характеристик объекта исследования;

– необходимостью непрерывного мониторинга динамики изменения социально-экономического состояния подконтрольного объекта исследования на основе общедоступных данных.

Объект, предмет, цель и задачи исследования

Объект исследования соответствует следующим пунктам паспорта специальности 08.00.05 «Экономика и управление народным хозяйством», специализация: «Экономика, организация и управление предприятиями, отраслями, комплексами (промышленность)» ВАК Минобрнауки РФ:

- 1.1.1. Разработка новых и адаптация существующих методов, механизмов и инструментов функционирования экономики, организации и управления хозяйственными образованиями в промышленности
- 1.1.13. Инструменты и методы менеджмента промышленных предприятий, отраслей, комплексов.
- 1.1.17. Теоретические и методологические основы мониторинга развития экономических систем народного хозяйства.

Объект исследования — предприятие, являющееся элементом большой совокупности предприятий - объектов нормативного регулирования со стороны общего контрольного или надзорного органа.

Предмет исследования — социально-экономическое состояние предприятия, являющееся объектом нормативного регулирования со стороны контрольных и надзорных органов.

Цель диссертационного исследования — экспресс-анализ социально-экономического состояния предприятия, удовлетворяющего нормативам и регламентам со стороны контрольных и надзорных органов на основе общедоступной информации с использованием интегрального показателя.

Основными задачами исследования являются:

- 1) разработать концептуальную модель и общую схему ее реализации для экспресс-анализа соответствия социально-экономического состояния предприятия нормативам и регламентам со стороны контрольных и надзорных

органов на основе общедоступной информации с использованием интегрального показателя;

2) разработать алгоритм расчета интегрального показателя экспресс-анализа социально-экономического состояния предприятия с использованием расширенного набора аргументов функции на основе таблицы поиска;

3) разработать методику экспресс-анализа соответствия социально-экономического состояния предприятия, удовлетворяющего требованиям со стороны контрольных и надзорных органов на основе общедоступной информации, как главного инструмента предложенной концептуальной модели;

4) обосновать на реальных данных о промышленных предприятиях и финансовых организациях (банках) результативность предложенной концептуальной модели, разработанных алгоритмов и методики формирования набора компонент, для экспресс-анализа социально-экономического состояния объекта исследования;

Научная новизна

Предложена концептуальная модель и разработана на ее основе методика проведения экспресс-анализа соответствия социально-экономического состояния предприятия нормативным требованиям, предъявляемым к нему со стороны контрольных и надзорных органов, включающая типологизацию источников общедоступной информации, на основе которых формируются разнородные значения компонент, представляющие собой расширенный качественными характеристиками набор аргументов функции, преобразующей при помощи логических правил на основе таблицы поиска эти разнородные значения в значения интегрального показателя. Гибкость предложенного инструментария обеспечена оперативным изменением решающих правил, реализованных в таблице поиска, в соответствии с

изменением набора данных, полученных из общедоступных источников информации.

Научная новизна диссертационного исследования конкретизируется следующими положениями, выносимыми на защиту:

1. Предложена концептуальная модель как система взаимосвязанных понятий, характеризующих соответствие социально - экономического состояния предприятия нормативам и регламентам со стороны контрольных и надзорных органов, использующая, в отличие от известных моделей и методик, количественную и качественную внешнюю общедоступную информацию, дополняя разделы стратегического менеджмента, касающиеся оценки деятельности и корректировки в рамках стратегического контроля.

Предлагаемая концептуальная модель представляет собой гибкий инструментарий отслеживания динамики состояния подконтрольного предприятия, обеспечивающий своевременное принятие упреждающих мер, способствующих снижению рисков негативных социальных и экономических последствий.

Данная концептуальная модель предложена впервые.

2. Построена типологизация источников общедоступной информации, учитывающая разбиение источников информации на группы в зависимости от официальности публикуемой информации, характеристики достоверности предоставляемых данных, структурированности публикуемых данных, и периодичности обновления источников информации, отличительной особенностью которой является категория информационного ресурса: общедоступный источник информации или открытая база данных, с указанием достоверности данных.

3. Предложен расширенный набор аргументов функции для построения интегрального показателя экспресс-анализа соответствия социально-экономического состояния предприятия нормативам и регламентам на основе

внешних общедоступных данных, публикуемых регулярно и с разной периодичностью, позволяющих сформировать разные наборы компонент для расчета интегрального показателя с учетом специфических характеристик объекта исследования.

Отличительной особенностью предлагаемого набора является учет как количественной, так и качественной информации на основе структурированных и неструктурированных данных. Значение каждой компоненты, в зависимости от структурированности данных, рассчитывается с использованием соответствующего математического аппарата (логистическая регрессионная модель, кластерный анализ, семантический анализ, индексный метод).

4. Разработан новый алгоритм построения интегрального показателя экспресс-анализа соответствия социально-экономического состояния предприятия нормативным требованиям со стороны контрольных и надзорных органов на основе внешней общедоступной информации, публикуемой регулярно и с разной периодичностью, использующий таблицу поиска, пороговые значения в которой варьируются в зависимости от объекта и от информационной базы исследования.

Отличительной особенностью предложенного алгоритма является получение значений каждой компоненты из набора в значения интегрального показателя с помощью предложенной функции, заданной с помощью таблицы поиска и набора решающих правил. В отличие от применения весовых коэффициентов, использование таблицы поиска для свертки отдельных компонент в единый интегральный показатель позволяет учесть их нелинейное взаимодействие.

5. Разработано гибкое инструментальное решение предиктивной и описательной аналитики, обеспечивающее повышение результативности взаимодействия контрольных и надзорных органов с большой совокупностью предприятий, за счет сокращения сроков проверок, возможности

автоматизации проверки, и как следствие, уменьшение погрешности расчетов. Это инструментальное решение реализовано в виде информационно-логической модели, представляющей собой методику экспресс-анализа соответствия социально-экономического состояния каждого предприятия нормативам и регламентам.

Теоретическая и методологическая основа диссертационного исследования

Диссертационное исследование базируется на фундаментальных трудах зарубежных и российских ученых в области стратегического управления предприятиями.

В диссертационной работе использован следующий комплекс подходов и методов: достижения и методология теории вероятности, социально – экономической статистики, методы математического, эконометрического, интеллектуального анализа данных.

Информационная база диссертационного исследования

Результаты исследования апробированы на двух тематических примерах (табл. 1).

Таблица 1.

Математические и статистические методы, используемые для расчета компонент интегрального показателя

№ п/п	Экономический объект	Цель экспресс-анализа	Методы получения компонент
1.	Промышленные предприятия г. Москвы	Оценка потребности и целесообразности финансовой поддержки со стороны региональных органов власти	<ul style="list-style-type: none"> • логистическая регрессия, • кластерный анализ, • семантический анализ, • индексный метод
2.	Коммерческие банки	Оценка ЦБ надежности банка до момента нарушения банком нормативных значений регулятора	<ul style="list-style-type: none"> • логистическая регрессия, • кластерный анализ, • семантический анализ, • индексный метод

В качестве информационной базы исследования для апробации разработанной концептуальной модели использовались данные из открытых источников — с официальных и неофициальных сайтов ведомств, сайтов – агрегаторов отраслевой информации, сайтов рейтинговых агентств и электронных СМИ, сайтов экономических объектов.

Информационная база исследования для проведения экспресс-анализа соответствия социально-экономического состояния промышленного предприятия требованиям контрольного органа (Департамент инвестиционной и промышленной политики г. Москвы) состояла из данных 506 действующих промышленных предприятий, зарегистрированных в г. Москве. В качестве объясняющих переменных для оценки компонент интегрального показателя рассматривались различные финансовые и нефинансовые данные за 2016–2018 годы по формам № 1 и № 2 бухгалтерской отчетности (бухгалтерский баланс, отчет о финансовых результатах), данным о контрагентах с портала государственных закупок. Для каждого статичного показателя рассчитывалось его изменение по отношению к предыдущему периоду. Дополнительно использовались неструктурированные текстовые данные: отзывы о предприятиях как работодателях, отзывы о предприятии в электронных СМИ, т.е. качественные характеристики упоминаний названия предприятия в новостных лентах и сайтах выставок.

Для модели оценки степени близости коммерческого банка к неблагоприятному финансовому положению на основе нормативов и регламентов ЦБ РФ была использована обучающая выборка из 108 банков (действующих на февраль 2020 г.) и тестовая выборка, включающая три банка. В качестве объясняющих переменных для расчета компонент интегрального показателя использовались общедоступные данные о ежемесячных ставках по вкладам в банк для физических лиц за 01.2018–02.2022 гг. Дополнительно учитывались качественные характеристики деятельности: показатели рейтинговых агентств, общие характеристики банка по данным сайта Банка

России. Информация была собрана из восьми внешних источников: сайт ЦБ РФ; рейтинговые агентства; справочники и информационно-аналитические порталы. Требования к состоянию банка были взяты из нормативов регулятора – ЦБ РФ на основе отчетных форм 0409101, 0409102, 0409123 и 0409135 публичной финансовой отчетности.

Теоретическая значимость исследования состоит в разработке концептуальной модели построения гибкого инструментария на основе интегрального показателя экспресс-анализа соответствия социально-экономического состояния объекта исследования заявленным контрольными и надзорными органами регламентам и нормативам, опирающегося на общедоступные структурированные и неструктурированные данные, агрегирующего результаты оценки набора компонент. Раздел теории управления предприятием, касающийся оценки текущего социально-экономического состояния объекта исследования, дополняется учетом качественной характеристики в виде дополнительных компонент 4,5,6, входящих в интегральный показатель, оцениваемых методом семантического анализа текстовых данных, что способствует повышению качества оценки.

Практическая значимость диссертационного исследования заключается в создании целостного решения экспресс-диагностики социально-экономического состояния предприятия, включающего разработку алгоритма построения интегрального показателя с конечным числом категорий, агрегированного из набора разнородных компонент с помощью таблицы поиска. Результаты исследования могут быть интересны руководителям органов государственной власти, руководителям министерств и ведомств, занимающихся оценкой и управлением различными экономическими объектами (предприятиями, финансовыми организациями, образовательными учреждениями и т.п.), топ-менеджменту предприятий и организаций с целью улучшения качества текущего мониторинга,

стратегического и тактического планирования, а также своевременного выявления проблем в деятельности предприятия.

Работоспособность предложенной информационно-логической модели экспресс-анализа состояния объекта исследования подтверждается результатами ее апробации на выборке из 506 промышленных предприятий г. Москвы и 111 коммерческих банков России.

Апробация и внедрение результатов исследования

Достоверность полученных результатов, рекомендаций и выводов диссертационного исследования базируется на современной методологии научных исследований, подтверждается корректной постановкой задач и аргументированным обоснованием необходимости их решения, представительностью и достоверностью исходной информации. Обоснованность выводов подкреплена табличным материалом, графическими иллюстрациями, ссылками на соответствующие источники.

Результаты диссертационного исследования были представлены и получили положительную оценку на 3 международных научно-практических конференциях, 1 симпозиуме, 2 международных школах-семинарах и 1 научном семинаре:

- Доклад на тему «Концепция построения универсального комплексного индикатора для экспресс-оценки состояния экономического объекта» на научном семинаре «Проблемы моделирования и развития производственных систем». Москва, ЦЭМИ РАН, 3 декабря 2021г.

- Доклад на тему «Оценка банков в многокритериальной среде» на VII международной конференции «Актуальные проблемы системной и программной инженерии» (АПСПИ-2021)». Москва, МИЭМ НИУ ВШЭ, 15–19 ноября 2021 г.

- Доклад на тему «Методы расчета значений универсального комплексного индикатора для оценки социально-экономического состояния экономического объекта управления» на 44-е заседании Международной

научной школы-семинара «Системное моделирование социально-экономических процессов» имени академика С.С. Шаталина. Воронеж, ЦЭМИ РАН, 4–9 ноября 2021 г.

Доклад отмечен дипломом «За лучший доклад и высокий уровень результатов исследований в области управления экономическими системами».

- Доклад на тему «Экспресс-оценка состояния промышленных предприятий г. Москвы с использованием универсального комплексного индикатора» на XXII Всероссийском симпозиуме «Стратегическое планирование и развитие предприятий». Москва, ЦЭМИ РАН, 13–14 апреля 2021 г.

Доклад отмечен грамотой и признан лучшим в секции 5.

- Доклад на тему «Построение комплексного индикатора для оценки состояния российского коммерческого банка на основе структурированных и неструктурированных данных» на 43-ем заседании Международной научной школы-семинара «Системное моделирование социально-экономических процессов» имени академика С.С. Шаталина. Воронеж, ЦЭМИ РАН, 13–18 октября 2020 г.

- Доклад на тему «Использование методов Data Science в оценке степени загрязнения прибрежной зоны морского побережья курортных регионов» на Всероссийской конференции «Моря России: исследования береговой и шельфовой зон» (XXVIII Береговая конференция). Севастополь, ФГБУН ФИЦ МГИ, 21–25 сентября 2020 г.

- Доклад на тему: «Сравнительный анализ моделей прогностической аналитики в задачах классификации» на VI Международной конференции «Актуальные проблемы системной и программной инженерии (АПСПИ 2019)». Москва, МИЭМ НИУ ВШЭ, 12–14 ноября 2019 г.

Данная методология применяется в практической работе отдела больших данных «ЗАО ЕС-Лизинг» (**справка о внедрении**).

Публикации. Основные результаты и выводы диссертации опубликованы в 6 печатных работах общим объемом 5,17 п.л. в изданиях, входящих в перечень ведущих рецензируемых научных журналов и изданий, определяемый ВАК Минобрнауки РФ, а также имеющих индекс Scopus.

II. ОСНОВНОЕ СОДЕРЖАНИЕ ДИССЕРТАЦИОННОГО ИССЛЕДОВАНИЯ

Диссертация состоит из введения, 3 глав, 11 параграфов, заключения, списка использованной литературы, включающего 84 источника, 9 приложений, 18 таблиц и 11 рисунков. Общий объем работы составляет 142 страницы, 126 страниц без учета приложений.

Во введении обоснована актуальность темы диссертационного исследования, сформулированы цель и задачи исследования, которые необходимо решить для достижения поставленной цели, раскрыты научная новизна, теоретическая и практическая значимость результатов исследования.

В первой главе «Модели оценки социально-экономического состояния объекта исследования в России и за рубежом» дана характеристика моделей и подходов для оценки состояния различных объектов исследования. Сформулированы проблемы современной системы управления, контроля и надзора над экономическими объектами, рассмотрены источники общедоступной информации для проведения анализа. Проанализированы и классифицированы международные и российские модели и методы оценки социально-экономического состояния различных экономических объектов. Рассмотрены подходы к построению моделей интегральных показателей при оценке социально-экономического состояния экономического объекта. Сформулирована постановка проблемы исследования.

Во второй главе «Концептуальная модель экспресс-анализа соответствия социально-экономического состояния объекта исследования требованиям контрольных и надзорных органов» предложены концептуальная модель,

обобщенная схема реализации концептуальной модели, алгоритм расчета интегрального показателя на основе набора компонент и информационно-логическая модель экспресс-анализа соответствия состояния объекта исследования требованиям контрольных и надзорных органов на основе интегрального показателя.

Подробно описывается поэтапная схема построения интегрального показателя, состоящая из пяти этапов. Для определения значений интегрального показателя на основе различных эконометрических моделей и методов предлагаются методы расчета компонент, характеризующих требования контрольных органов к экономическому объекту, затем, с помощью таблицы поиска, все полученные разнородные значения компонент объединяются и определяется значение интегрального показателя, которое позволяет провести экспресс-анализ состояния объекта исследования. Рассматриваются и типологизируются источники информации для формирования базы данных исследования, на основе которой оцениваются компоненты интегрального показателя. Предлагается алгоритм формирования набора компонент, определяются методы оценки значений компонент. Дается характеристика информационно-логической модели построения интегрального показателя, объединяющей результаты различных моделей: логистической регрессионной модели, кластерного анализа, тематического моделирования.

В третьей главе «Экспресс-анализ социально-экономического состояния предприятий из большой совокупности на соответствие нормативным требованиям» проводится апробация разработанного нами инструментария на информационной базе исследования для двух эмпирических исследований на примере промышленных предприятий и финансовых организаций (коммерческих банков).

Для промышленных предприятий г. Москвы разработан инновационный инструментарий, позволяющий использовать современные технологии работы

с разнородными неструктурированными данными (сквозная технология цифровой экономики) для поддержки деятельности контрольных органов – аудиторов Контрольно-счетной палаты г. Москвы, руководства профильного Департамента инвестиционной и промышленной политики при проведении ими мероприятий, связанных с проверкой эффективности мер поддержки и стимулирования деятельности в сфере промышленности г. Москвы.

Предложена концептуальная модель и алгоритм построения интегрального показателя экспресс-анализа соответствия социально-экономического состояния промышленного предприятия нормативным требованиям Департамента инвестиционной политики г. Москвы на примере 506 промышленных предприятий, зарегистрированных в г. Москве, по открытым данным за 2016 - 2018 годы. На основе финансовых и нефинансовых показателей были оценены компоненты и рассчитан интегральный показатель с использованием методов эконометрики, кластерного, статистического и индексного анализа. На основе рассчитанного интегрального показателя были выработаны рекомендации о целесообразности субсидирования предприятий на следующий финансовый год в отсутствие официальной статистической отчетности.

Для коммерческого банка предложена концептуальная модель построения интегрального показателя экспресс-анализа соответствия социально-экономического состояния банка заявленным требованиям со стороны ЦБ РФ. Контрольным органом в данном случае является ЦБ РФ – надзорный орган в банковской сфере. В соответствии с требованиями ЦБ РФ о надежности банка, был получен набор компонент интегрального показателя и рассчитано его значение для каждого банка. Прогностическая способность построенной модели, апробированной на 108 банках обучающей и трех банках тестовой выборки по открытым данным за 2017–2019 год, была подтверждена согласованность результатов с их фактическим состоянием в марте 2020 года.

Были выработаны рекомендации для усиления надзора над некоторыми банками из выборки с целью предотвращения отзыва у них лицензии.

В заключении изложены основные научные результаты и выводы диссертационного исследования.

В 9 приложениях приведен графический и табличный материал, детализирующий результаты исследования.

Глава 1. Модели оценки социально-экономического состояния объекта исследования в России и за рубежом

1.1 Проблемы управления финансово-хозяйственной деятельностью объекта исследования

В современных экономических условиях роль грамотного управления в деятельности организации любого уровня нельзя недооценивать. При оптимальном управлении экономятся ресурсы, повышается эффективность деятельности, ускоряется возврат средств, и снижаются риски финансовых потерь. В последние годы наблюдается пристальное внимание к способам и методам управления, а также к информации, на основе которой принимаются решения. Чем подробнее и качественнее используемая информация, тем точнее принятые на ее основе управленческие решения. Управление не существует без мониторинга и контроля состояния объекта исследования [65]. Мониторинг и внутренний контроль – одна из управленческих функций, цель которой – предоставить своевременно информацию для лиц, принимающих решения, обеспечив полноту и достоверность этой информации. Как отмечают авторы недостаточно активное использование современных цифровых технологий приводит к непоступлению информации каждый день, качество мониторинга и контроля находится в прямой зависимости от методической поддержки, требующей постоянного совершенствования методов обработки информации.

О проблеме управления экономическими объектами в последние годы было написано достаточно много, как российскими, так и зарубежными исследователями [65,72]. Авторы выделяют значимые проблемы в контрольных функциях организации, таких как акцентирование проведения контроля только на аудите финансово-хозяйственной деятельности, превращение контрольной деятельности в перепроверку учетных операций, включая арифметический пересчет. В то время как оценка внешних не

количественных показателей, таких как уровень популярности, лояльность потребителей, качество объекта исследования остаётся вне внимания контрольных органов.

Проблемы совершенствования управления различными объектами были затронуты во многих научных публикациях последних 10 лет. Например, в статье [71] автором подняты вопросы совершенствования инструментария планирования социально-экономического развития и оценки экономических объектов в условиях активного формирования инфраструктуры информационного общества, проведена оценка влияния цифровых технологий на содержание управленческих задач - в процессе изменчивости объектов и субъектов управления в крупномасштабных организационных системах. Отмечается недостаточно развитый уровень вовлеченности больших данных в управление на тактическом уровне [58]. Автор предлагает стандартизацию технологий сбора, передачи, хранения и обработки больших объемов данных, построение компьютерных математических моделей обеспечения многоуровневой системы государственного стратегического планирования [38, 42]. В работе [79] автор, анализируя исследования в области управления и контроля, отмечает слабо развитую в настоящее время систему внутреннего контроля. Делается вывод, что руководство большинства фирм считает внедрение внутреннего контроля затратным и трудоемким, при этом они недооценивают всей важности этой системы, отмечается разница в стандартах контроля российских и зарубежных компаний.

Также проблемы управления предприятиями затронуты и в работе [78], где автор отмечает слабые межведомственные связи в работе контрольных органов. Анализируя проблемы современной системы управления в государственных органах власти, автор отмечает совпадение контроля и проверок установленных правил, показана разветвленная как по горизонтали, так и по вертикали система органов контроля, что затрудняет сопоставление результатов разных контрольных мероприятий. При этом возникает

дублирование отдельных полномочий, отмечается отсутствие единого подхода к понятию надзора и контроля, так как надзор представляет собой непрерывное наблюдение и проверку, но не является функцией управления.

Проводя исследование современных механизмов управления промышленными предприятиями, Клишкин Ю.Н. в своей работе [55] проводит исследование в исторической перспективе развития механизмов управления промышленными предприятиями и их результативности, где в качестве результатов деятельности рассматриваются патенты на изобретения. В частности, в связи с ростом количества нематериальных активов, учтенных в финансовой отчетности, для промышленных предприятий актуальной является задача совершенствования механизмов управления, учитывающих изменение структуры баланса, которые позволяли бы корректнее оценивать потребность промышленных предприятия в получении дополнительных средств. Автор отмечает проблему недооценки роли нематериальных активов промышленных предприятий в качестве инструментального решения в теории стратегического менеджмента.

Подводя итог, авторы едины в мнении, что требуется решать проблему несовершенства методологических аспектов системы мониторинга и контроля, например, с помощью системы показателей, обеспечивающих иерархичность целей по уровню государственного управления. Предлагается расширять систему социально – экономического мониторинга показателями качества и уровня жизни, и другими характеристиками социальной среды, особенно на уровне муниципалитетов. Таким образом, отмечается наличие проблем в управлении, требующих развития методологической базы исследований деятельности предприятий.

1.2 Переход к цифровой экономике: использование общедоступных источников информации

В развитых странах в последнее время ведется интенсивная работа над созданием инновационных систем управления и контроля за различными экономическими объектами: услугами и товарами, производством, предприятием, отраслью, государством, любыми другими предприятиями, как по отдельности, так и в различных сочетаниях. Такой интерес к контрольной деятельности вызван сильно изменившимися условиями функционирования различных объектов исследования. Динамика среды возрастает, стабильность внешней среды снижается, при этом требования к экономии ресурсов и качеству результатов той или иной деятельности непрерывно повышаются. Количество информации, которую необходимо обработать для принятия того или иного решения, стабильно увеличивается, одновременно ужесточаются требования к качеству, безопасности и актуальности этой информации.

Количество данных, хранящихся как во внутренней среде организации любого типа, так и во внешней среде, в сети интернет, возрастает быстрыми темпами. Единственным выходом из сложившейся ситуации является последовательная, но вместе с тем достаточно быстрая цифровизация всех отраслей и сфер экономики и общества. Процесс цифровизации не протекает одномоментно, он требует грамотного планирования и высокой степени управленческой гибкости, заложенной на стратегическом уровне.

В условиях стремительно изменяющихся отраслей экономики, включая ее цифровизацию, все сложнее оценить динамику среды, в которой ведется экономическая деятельность. Особенно важно отслеживать тенденции, связанные с социологическими и технологическими изменениями в этой среде, включая международные тенденции. Еще одной особенностью современной экономической деятельности является важнейшая роль личности как руководства, так и клиентов, инвесторов, поскольку само современное понятие «оказание услуг» основывается на предположении, что оба этих вида

участников оказания услуг: как поставщики, так и потребители – должны быть готовы к диалогу. Без этого качество услуг не может быть высоким.

Другими словами, резко возрастает сложность управления динамичной цифровой образовательной деятельностью. Соответственно, появляется необходимость в новых средствах управления на основе новых технологий сбора и обработки первичных данных обо всех аспектах образования, включая «человеческий фактор». На данный момент единственным технологическим средством, позволяющим построить инновационную систему управления, являются технологии обработки больших данных, использование которых предусматривается в процессе цифровой трансформации.

Согласно официальному определению из новейшего международного стандарта:

– ISO/IEC 20546:2019(en) Information technology — Big data — Overview and vocabulary – под большими данными понимаются большие наборы данных. Прежде всего, в характеристиках объема, разнообразия, скорости и/или изменчивости данных, которые требуют масштабируемой технологии для эффективного хранения, манипулирования, управления и анализа (Big Data – extensive datasets — primarily in the data characteristics of volume, variety, velocity, and/or variability — that require a scalable technology for efficient storage, manipulation, management, and analysis).

- Российской версии стандарта ISO/IEC 20546 под названием ГОСТ Р ИСО/МЭК 20546 Информационная технология. Большие данные.

Подобный подход обеспечивает возможность на основании данных, полученных бесконтактным методом из открытых источников:

- максимально быстро начать отслеживать нужные в данный момент показатели качества ведения экономической деятельности или любые другие стратегические показатели, включая их динамику;

- последовательно включать в систему прочие стратегические показатели, позволяющие выявлять и отслеживать различные тенденции, мотивации, поведение, стратегии участников рынков товаров и услуг;

- оперативно оценивать качественные изменения в деятельности предприятия и мер государственной образовательной политики для поддержки принятия управленческих решений;

У каждой технологии есть преимущества и вызовы. Это необходимо принимать во внимание при приведении проектов на стадии планирования и проектирования информационных услуг и сред, где используются эти технологии, чтобы минимизировать затраты и предотвратить инциденты.

В отличие от классических, много лет используемых, однородных структурированных данных, общедоступные данные можно намного более эффективно применять для достижения следующих целей:

- Сделать процесс получения информации более прозрачным, как и саму информацию.

- Получать по требованию и по мере необходимости дополнительную детальную информацию об эффективности того или иного конечного объекта исследования, например, промышленного предприятия, что стимулирует инновации и повышает качество оказания конечных и поддерживающих услуг.

- Получать по требованию и по мере необходимости дополнительную по отношению к официально проводимой, эффективную, актуальную и точную аналитику, чтобы минимизировать риски, отслеживать на ранней стадии последствия принятых решений и предотвращать проблемы за счет моделирования и прогнозирования ситуации.

- Обеспечивать соответствующий персонал стратегической, тактической и операционной информацией в режиме реального времени. Это не только позволяет более эффективно управлять рисками, но и может

привести к сокращению управленческого персонала при одновременном повышении качества оказываемых услуг.

Привлечение неструктурированных данных позволяет при оценке объектов исследования:

- эффективно собирать об объектах исследования актуальную открытую информацию из сети интернет;
- выявлять и анализировать тенденции, мотивации, поведение, стратегии участников рынков на основании данных, полученных бесконтактным методом из открытых источников;
- оперативно оценивать качественные изменения для выбранного рынка и мер государственной политики для ведения оперативной системы контроля.

Наиболее актуальными и значимыми в области разработки применения инновационных подходов и методов, в силу высокого влияния изменяющихся экономических и политических условий, являются исследования последних двадцати лет [19, 25].

Построение систем управления и поддержки принятия решений на основе использования объединенных в единое целое разрозненных источников данных позволяет конечным пользователям принимать наиболее эффективные решения. Именно такой подход является базой для роста бизнеса компаний и повышения уровня зрелости управленческой деятельности на всех уровнях, альтернативы ему на данный момент нет. Такой подход, в конечном итоге, создает условия для дальнейшего роста экономики в целом.

Неструктурированные данные могут быть представлены в разных видах – текст, новостные сообщения, комментарии, фотографии [37]. С развитием интернет – технологий их объемы растут, и информация на основе таких данных может помочь в анализе состояния объекта исследования.

Все больше данных появляется в общедоступном виде как открытые данные. **Открытые данные** — это информация о деятельности

государственных органов и органов местного самоуправления, размещенная в сети «Интернет» в виде массивов данных в формате, обеспечивающем их автоматическую обработку в целях повторного использования без предварительного изменения человеком (машиночитаемый формат), и на условиях ее свободного (бесплатного) использования [6].

Использование инновационных цифровых возможностей сбора и анализа общедоступной информации из сети интернет позволяет дополнительно анализировать характеристики качества работы различных предприятий и других объектов исследования. Подобный анализ на открытых данных можно проводить при помощи вспомогательного независимого инструментария оценки состояния объекта исследования, созданного на основе анализа больших объемов структурированных, слабоструктурированных и неструктурированных данных из открытых источников сети интернет, и сравнивать результаты с официальной методикой исследования по внутренним или официальным статистическим данным.

1.3 Особенности современной системы контроля и надзора в РФ

1.3.1 Полномочия органов государственного контроля

Контроль – неотъемлемая часть при реализации управленческие функции. Управленческие функции проявляются во взаимодействии объекта и субъекта управления, контрольные функции позволяют отслеживать качество управления.

Без грамотно организованной системы контроля, невозможно в полной мере реализовать управленческие функции на стратегическом и тактическом уровне. Контрольная и надзорная деятельность относительно состояния предприятия осуществляется, как правило, внешними по отношению к предприятию органами, такими как аудиторские компании и государственные органы контроля и надзора. Существенная часть контрольных функций в РФ реализуется в настоящий момент государственными органами.

Контрольными (надзорными) органами государственной власти являются наделенные полномочиями по осуществлению государственного контроля (надзора), муниципального контроля соответственно федеральные органы исполнительной власти, органы исполнительной власти субъектов Российской Федерации, органы местного самоуправления, а также в случаях, предусмотренных федеральными законами, государственные корпорации, публично-правовые компании [8]. Несмотря на значительное влияние этого вида государственного регулирования на предпринимательский климат, до сих пор ему уделялось мало внимания. Перечень таких полномочий закреплен в п. 2 ст. 26.3 Федерального закона от 6 октября 1999 г. № 184-ФЗ «Об общих принципах организации законодательных (представительных) и исполнительных органов государственной власти субъектов Российской Федерации» и содержит 95 полномочий, из которых 10 являются контрольными [8].

Все собственные контрольно-надзорные полномочия субъектов Российской Федерации (как указанные, так и не обозначенные в Законе № 184-ФЗ) могут быть разделены на три группы:

- I группа — контрольные полномочия, которые полностью подпадают под действие Федерального закона от 26 декабря 2008 г. № 294-ФЗ «О защите прав юридических лиц и индивидуальных предпринимателей при осуществлении государственного контроля (надзора) и муниципального контроля» [4];

- II группа — контрольные полномочия, которые подпадают под действие Закона № 294-ФЗ, но в отношении их федеральными законами могут быть установлены особенности.

- III группа — контрольные полномочия, которые хотя и осуществляются в отношении индивидуальных предпринимателей и юридических лиц, но исключены из сферы действия Закона № 294-ФЗ в силу специальной оговорки в законе или иных условий.

По характеру контроля можно выделить следующие направления контрольной деятельности [81]:

- проверка выполнения функций по аккумуляции, распределению и использованию финансовых ресурсов соответственно своей компетенции;
- проверка выполнения финансовых обязательств перед государством и коммерческими организациями;
- проверка правильности использования имущества, находящегося в ведении хозяйствующего субъекта;
- проверка соблюдения правил совершения финансовых операций, расчетов и хранения денежных средств предприятиями, организациями и учреждениями;
- выявление внутренних резервов производства;
- устранение и предупреждение нарушений финансовой дисциплины;

Государственный контроль, его организация, формы взаимодействия между участниками регулируются нормативными актами. Постоянно идет процесс совершенствования и повышения эффективности государственного контроля и надзора.

1.3.2 Проблемы системы государственного и внутрифирменного контроля

Государственные органы власти РФ в своей деятельности занимаются, в том числе, поддержкой и развитием промышленности в субъекте РФ, они осуществляют контроль состояния отрасли, мониторинг и предотвращение кризисных явлений или срывов в работе ключевых для региона организаций. Таким образом, предприятие как отдельный субъект выступает в качестве экономического объекта исследования со стороны государственных органов, на которое с помощью налоговых льгот и субсидий органы власти могут оказывать воздействие.

Для решения задачи, поставленной Правительством РФ по улучшению делового климата, одним из ключевых направлений поставлено

совершенствование контрольно-надзорной деятельности государства, предусматривающее ее упрощение, устранение избыточности процедур, а также ориентацию на минимизацию рисков причинения вреда общественно значимым предприятиям.

Вопросы качества осуществления аудиторской деятельности в настоящее время приобретают высокую актуальность, так как стоимость аудита возрастает, количество внешних аудиторских организаций сокращается. Существующая в настоящее время схема организации контроля за предприятиями, в большинстве случаев, начинается с инициации внеплановой проверки вышестоящей организацией, или с наступления момента проведения периодической проверки. Информационная база совокупности предприятий, подлежащих проверке, формируется из документов, предоставленных предприятием, форм, полученных в ответ на запрос аудиторов или статистической отчетности, предоставляемой в соответствии с существующим законодательством. Зачастую возникает существенный лаг между моментом запроса и непосредственным поступлением информации. Численность персонала контрольных и надзорных органов, осуществляющих проверку, часто в десятки раз меньше численности проверяемых предприятий. Существующую схему взаимодействия контрольных органов с предприятиями при реализации контрольных мероприятий можно представить в виде следующей последовательности (рис. 1).

Недостатками текущей схемы организации контроля за состоянием поднадзорных предприятий является отсутствие качественной оценки полученной информации, что не обеспечивает должного взаимодействия контрольных и надзорных органов с предприятиями. Также возможно возникновение в некоторых случаях запаздывания актуальной достоверной информации от предприятия, связанного с заполнением форм статистической отчетности, по сравнению со сроками для проведения проверок.

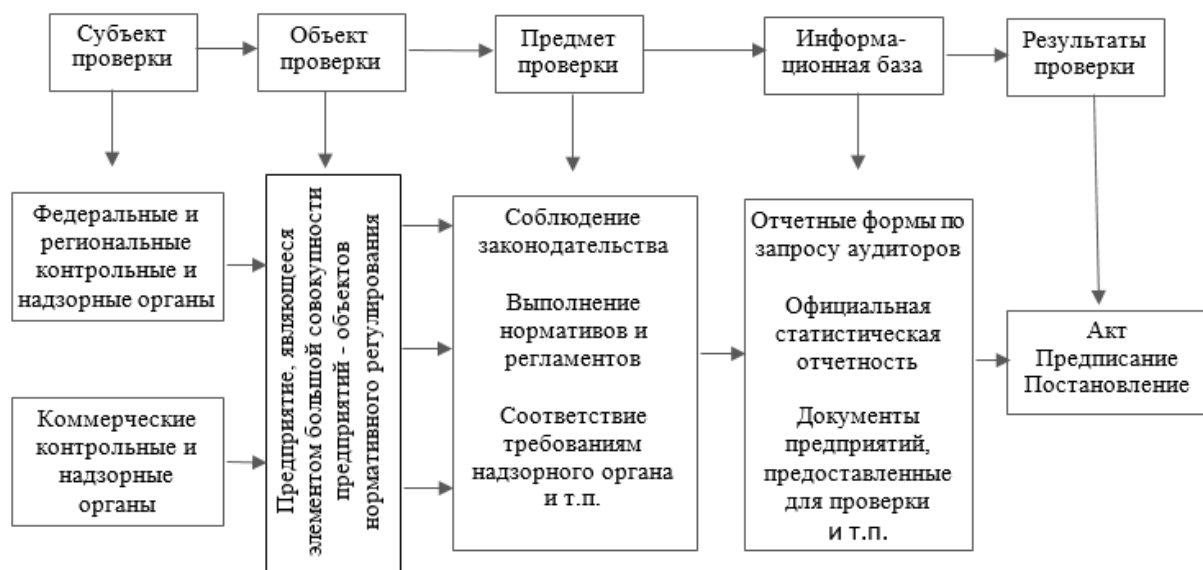


Рисунок 1. Текущая схема организации контрольных мероприятий

Для предприятий и организаций различных форм собственности, работающих в разных сферах экономики, государственные органы власти могут выдвигать обязательные требования к деятельности юридического лица в рамках законодательства РФ по следующим направлениям:

- осуществление государственного контроля за предпринимательской деятельностью, возникновением и ликвидацией фирм, выполнением ими разрешенных видов деятельности, соответствием фактической деятельности фирм видам, указанным при регистрации;
- получение от предпринимателей сведений бухгалтерского и финансового учета для ведения статистического учета и использования этих данных в целях регулирования предпринимательской деятельности;
- контроль налогообложения;
- лицензирование деятельности хозяйствующих субъектов; Лицензия является официальным документом, который разрешает указанный в нем вид деятельности в течение установленного срока;
- государственная регистрация деятельности;
- определение возможных организационно-правовых форм, которые могут принимать хозяйствующие субъекты;

- соблюдение норм экологического и санитарного надзора и контроля;
- соблюдение требований безопасности дорожного движения;
- соблюдение порядка установления цен на ряд товаров преимущественно имеющих жизненно важное для общества значение (топливо, энергия, потребительские товары первой необходимости, жилищно-коммунальные услуги и т. д.).

- своевременная передача форм государственного статистического наблюдения;

- соблюдение требований стандартизации и сертификации товаров и услуг;

- соблюдение государственных регламентов о рекламной деятельности;

- соблюдение норм, стандартов и требований к системе трудовых отношений в отношении работников;

Любые нарушения требований государственных органов могут приводить к штрафам, гражданскому и уголовному судопроизводству, прекращению деятельности, пеням, что влечет за собой и экономический ущерб всей деятельности организации или предприятия.

В процессе реализации контрольных и надзорных полномочий государственные органы власти сталкиваются с множеством проблем. На протяжении долгих лет надзорный контроль осуществлялся на основании официальной статистической отчетности. При возрастании скорости и динамики принятия решений лаг в поступлении официальной отчетности иногда достигает полугода с момента окончания отчетного периода.

Российскими авторами в научных исследованиях предлагаются различные модификации для совершенствования в сфере надзора и контроля. Так, в своей диссертационной работе [38] на соискание ученой степени Гайдук Е.А. предлагает в качестве нового подхода при принятии управленческих решений на муниципальном уровне – внедрение индикаторов оценки состояния системы управления, построенных на основе данных Федеральной

службы государственной статистики РФ, различных интернет-источников; периодических изданий, публикующих данные об экономических объектах, экспертные данные в виде текстовой информации. Гайдук Е.А. описывает применение балансового и оптимизационного моделирования для разработки и обоснования авторской методики комплексного балансово-оптимизационного моделирования и прогнозирования динамики показателей социально-экономического развития Московской области (МО). Предлагаемый подход обеспечивает решение проблемы трансформации существующей системы управления и контроля и развитие системы внутреннего контроля в условиях цифровизации экономики [43].

Таким образом, контрольные меры, принятые на основе официальной статистической информации, могут поступать с большим запаздыванием, ведь между окончанием отчетного периода и передачей официальных статистических данных в органы государственной власти субъекта может проходить от трех до восьми месяцев. Использование общедоступных данных позволяет получить сведения о подконтрольном объекте на текущий момент времени.

1.4 Классификация подходов и моделей оценки социально – экономического состояния объекта исследования

Рассмотрим подробнее понятие предприятия как объекта исследования. Термин «предприятие» является широко употребляемым термином для описания деятельности любых экономических систем [54]. При этом часто употребляются понятия «фирма» и «организация», имеющие свои отличительные особенности.

Согласно определению, приведенному в работе Клейнера Г.Б., «предприятие - это самостоятельный хозяйствующий субъект экономической деятельности. Фирма - это юридически оформленный субъект экономической деятельности, который может состоять из нескольких предприятий.

Следовательно, фирма и предприятие могут являться синонимами лишь в одном случае - при хозяйственном представлении фирмы в виде единственного предприятия.

Организация - это субъект социальной деятельности, который может и не выполнять хозяйственные функции, например, общественные объединения, добровольные тематические сообщества. Следовательно, синонимами понятия «организация» и «предприятие» могут выступать только в случае наличия явной хозяйственной деятельности организации» [40].

Таким образом, понятие «предприятия» является более обобщенным, именно этому понятию посвящено множество прикладных экономических теорий.

Целью настоящего исследования является оценка социально – экономического состояния предприятия, которая раскрывается в различных прикладных теориях предприятия: теория менеджмента, посвященная системам управления сотрудниками, современная развивающаяся теория стратегического менеджмента/маркетинга, посвященная классическим и современным концепциям управления для достижения оперативных, тактических и стратегических задач, а также теория финансового анализа, посвященная оценке экономического состояния объектов исследования на основе оценки и интерпретации ключевых показателей. Прикладной инструментарий для разработки подходов и методов анализа и управления предприятием раскрыт в эконометрических, математических моделях, теории игр, теории оптимального управления и др. [74].

Экономический инструментарий разработки моделей для оценки стратегий развития предприятия включает в себя финансовые модели, такие как модели Альтмана и его модификации, различного рода матричные подходы, такие как матрицы М. Портера, БКГ, И. Ансоффа, теорию системы сбалансированных показателей [66,45].

Инструментарий анализа предприятий развивается и современными авторами. Среди научных работ последних лет с описанием различных экономических и математических исследований наблюдается рост трудов, посвященных применению современных цифровых технологий обработки больших объемов структурированных, слабоструктурированных и неструктурированных данных открытых источников сети интернет, методов машинного обучения и искусственного интеллекта в моделях поддержки принятия решений.

В литературе описываются различные подходы и методы к решению задач поддержки принятия решений. К математическим методам относятся методы на основе нейросетей, машинного обучения, индексные методы. Их применение для задач управления рассмотрено во многих научных работах последних лет.

В литературе изложено множество подходов и инструментов подобного рода экономических и математических моделей [75]. Большинство из них базируются на эконометрическом исследовании или методах машинного обучения. Как правило, имеющиеся статистические данные разбиваются на группы (демографические, социальные, финансовые и т.д.), ранжируются, или каким-то образом объединяются в единый интегральный показатель, а факторам присваиваются веса. Зачастую результатом подобного исследования становится интегральный показатель (коэффициент), который удобен при сопоставлении объектов. Подобные инструменты в значительной степени опираются на внутренние данные или на имеющуюся статистическую базу.

Среди научных работ последних лет, посвященных различным экономическим и математическим исследованиям, все чаще делается акцент на применение современных цифровых технологий обработки больших объемов структурированных, слабоструктурированных и неструктурированных данных из открытых источников сети интернет,

методов машинного обучения и искусственного интеллекта в моделях поддержки принятия решений [43, 53].

В научных исследованиях многими авторами предлагаются различные экономико-математические модели, базирующиеся на официальной статистической информации [60]. Большинство из них представляют собой эконометрические модели или модели, использующие методы машинного обучения. Как правило, имеющиеся статистические данные разбиваются на группы (демографические, социальные, финансовые и т.д.), ранжируются, или каким-то образом объединяются в единый интегральный показатель, а факторам присваиваются веса. Зачастую результатом подобного исследования становится интегральный показатель (коэффициент), который удобен при сопоставлении объектов.

Подобные инструменты в значительной степени опираются на внутренние данные или на имеющуюся статистическую базу [41]. Оценивание состояния объекта исследования с помощью интегральных показателей, являющихся инструментами управления и контроля, активно используется во многих сферах экономики. Как правило, такие индикаторы строятся на внутренних данных, однако с ростом объемов доступной открытой информации появляются алгоритмы оценки состояния определенных объектов управления и на открытых структурированных данных. Недостатком этих моделей является их узкая специализация и привязка только к структурированным, а иногда и строго официальным данным, которые, как правило, имеют редкую периодичность издания. Это не позволяет отследить изменение состояния объекта в разные моменты времени.

Как правило, официальная статистическая отчетность публикуется в открытом доступе с определенной периодичностью и соответствующим лагом, что затрудняет оперативное реагирование в форс-мажорных ситуациях. Поэтому современные концепции управления предлагают к использованию и открытую информацию из сети интернет.

В [32] авторы предлагают определять вероятность отзыва лицензии (как одной из компонент оценки надежности банка) на основе публичной финансовой отчетности и учета волатильности внешней среды (изменение курсов валюты, стоимости одного барреля нефти и т.д. Отличие от этой работы в данной статье в том, что авторы в своей работе предлагают использовать структурированные и неструктурированные данные из открытых источников информации: оценки рейтинговых агентств, тональность новостей с упоминанием банка, рейтинги независимых российских рейтинговых агентств.

Подход к построению комплексного индикатора на основе использования открытых структурированных данных об объекте исследования в текущий момент времени для оценки степени износа береговой зоны предложен в работе [53]. В работе сформулирована модель расчета комплексного индикатора на основе структурированной информации об использовании населением территории в качестве необустроенных диких пляжей, мини отелей или мест общественного собрания, собираемой по публичным картам и сообществам. Одна из проблем, поднимаемая авторами данной статьи, - это проблема постоянного контроля над экономическими объектами, при использовании официальной статистической отчетности, выходящей с большой периодичностью и с большим лагом.

Например, в банковской сфере текущий контроль со стороны Центрального банка над другими коммерческими банками затруднен из-за отсрочки в поступлении сведений о прогнозе состояния финансов банка через некоторое количество месяцев с учетом предпосылок его политики ставок по вкладам.

Авторы предлагают привлекать открытые источники информации и на основе математической модели формировать комплексный индикатор с использованием таких данных [73]. Изложенный в статье подход позволяет получить на основе открытых данных оперативную информацию о состоянии

банка как объекта исследования, и ускорить принятие решений в его отношении. Результатом применения математической модели является вероятность отзыва лицензии (как одной из компонент модели оценки надежного банка) на основе публичной финансовой отчетности и учета волатильности внешней среды (изменение курсов валюты, стоимости одного барреля нефти и т.д.).

Для обработки методами машинного обучения, математической статистики и математическими моделями неструктурированная информация предварительно структурируется в виде набора некоторых показателей методами математической статистики, машинного обучения или эконометрики [71,25].

Например, в докладе на конференции «43-е заседание международной научной школы-семинара «Системное моделирование социально-экономических процессов»» 13-18 октября 2020 г. ЦЭМИ РАН, авторы для экспресс-анализа состояния банка, отличного от стабильного, предложили использовать в построении комплексного индикатора одну из компонент, рассчитанную на основе неструктурированных данных - текстовом содержимом новостей, выделяя из них негативные упоминания названия банка как отдельного признака, влияющего на значение оценки вероятности нестабильности его состояния [35].

В работе [53] авторами предложен подход к оценке экологического состояния побережья на основании исследования открытых данных о фактической застройке прибрежной территории, информации о функционировании гостиниц и других мест размещения без официальной регистрации. В работе обосновывается положительный эффект от учета таких данных на качество показателей загрязнения в дополнении к официальной статистике.

Обобщая результаты проведенного анализа научной литературы, можно выделить основные направления развития экономической теории предприятий:

1) обобщение разработанных в различных фундаментальных и прикладных теориях моделей и создание на этой основе общего подхода к анализу закономерностей с учетом отдельных специфических характеристик предприятий;

2) построение обобщенных подходов и моделей, объединяющих различные инструментальные решения, учитывающие широкий арсенал различных теоретических математических и статистических методов;

3) расширение набора исследуемых факторов при анализе социально - экономического состояния объекта исследования.

1.5 Постановка проблемы исследования

В данной работе предлагается дополнить существующие методологии управления использованием общедоступных данных, формированием разнородных критериев для решения проблемы отсутствия актуальной информации об объекте исследования.

Данная проблема может быть решена проведением экспресс-анализа текущего состояния объекта исследования, опирающегося на использование, как формализованной, так и неформализованной, структурированной и неструктурированной информации. Результатом экспресс-анализа является получение оценки состояния объекта исследования. Экспресс-анализ состояния объекта исследования – это предварительное исследование, во многом определяющее целесообразность проведения дальнейшего, более углубленного, анализа.

В отличие от рассмотренных выше работ автор диссертационного исследования предлагает концептуальную модель и методику построения интегрального показателя экспресс-анализа состояния объекта исследования

на основе открытых структурированных и неструктурированных данных, полученных с официальных и неофициальных сайтов. Данные собираются как со статистических ресурсов, где они публикуются периодически, так и с интернет - ресурсов, регулярно обновляемых с разной периодичностью.

Отличительной особенностью данной концептуальной модели является использование расширенного набора аргументов логической функции, которые преобразуются в значения интегрального показателя. Разработан алгоритм расчета интегрального показателя экспресс-анализа соответствия состояния объекта исследования заявленным требованиям со стороны государственных органов с целью выработки рекомендаций по совершенствованию деятельности контрольных органов.

Использование результатов анализа структурированных и неструктурированных данных из открытых источников сети интернет является наиболее полным и разносторонним способом для полноценного всестороннего анализа состояния экономического объекта исследования. Это позволяет получать объективную информацию о текущей ситуации, не подвергнутую промежуточной обработке и полученную на основании анализа большого количества разнообразных актуальных данных, которые хранятся в открытом доступе в любых источниках сети интернет. При необходимости результаты анализа внешних данных могут быть соотнесены с результатами аналогичной аналитической деятельности, проведенной на основе использования внутренних данных. Также результаты анализа по открытым общедоступным источникам могут дополнять официальные или внутренние данные в некоторых аспектах деятельности объекта исследования.

Предлагаемая концептуальная модель интегрального показателя основана на сбалансированной системе показателей (BSC), как концептуальном методе организации управления любой деятельностью. В основу этого метода положена система разработки и контроля ряда ключевых индикаторов, что позволяет принимать управленческие решения в

соответствии с принятой стратегией развития той или иной управляющей структуры, организации, предприятия. Этот подход позволяет уменьшить риски негативных последствий, своевременно приняв предиктивные меры, обеспечив гибкий режим отслеживания состояния объекта исследования.

Для построения компонент интегрального показателя предлагается использовать методы кластерного и регрессионного анализа, применяемые к количественным и качественным данным, характеризующих объект. Предлагаемая концептуальная модель учитывает комплексное нелинейное совместное влияние всех показателей на состояние объекта исследования.

Еще одно отличие этой концептуальной модели состоит в том, что в качестве отправной точки при формировании компонент интегрального показателя, предлагается учитывать требования контролирующих органов, в то время как в большинстве российских и зарубежных исследований индикаторы строятся для нужд инвесторов.

В работе предлагается информационно-логическая модель экспресс-анализа соответствия социально-экономического состояния предприятия нормативным требованиям со стороны контрольных и надзорных органов на основе общедоступной информации. Информационно - логическая модель построена на основе предложенной концептуальной модели, одной из важных особенностей которой является то, что она учитывает любые требования разных регуляторов, как количественных, так и качественных, предъявляемых к экономическим объектам разных типов (предприятиям, финансовым организациям, образовательным учреждениям и т. п.).

В рамках предлагаемого подхода формируются разные наборы компонент в зависимости от текущих требований контрольных органов. В результате агрегирования значений компонент с помощью разработанной таблицы ассоциаций рассчитывается значение интегрального показателя, являющегося основой экспресс-анализа. Одновременное использование структурированных и неструктурированных статистических данных

позволяет получить более точную качественную оценку объекта исследования с учетом изменений, которые еще не отражены в официальной статистической отчетности, предоставляемой с определенной периодичностью и неизбежным запаздыванием.

Результатом проведенного экспресс-анализа является оценка соответствия социально-экономического состояния предприятия нормативным требованиям со стороны регулятора. Выводы, сделанные на основе проведенного экспресс-анализа, служат обоснованием для принятия решения о необходимости более детального, углубленного анализа отдельных предприятий.

Важно оценить степень готовности организации к переменам и внедрению новых методологий. В мировой практике имеют место случаи, когда использование современных информационных технологий на основе общедоступных данных, включая принципиально новые средства аналитики, противоречат внутренней культуре организации и/или сложившемуся стилю руководства, которое может концентрироваться на соблюдении только внешних требований, зачастую не принимая во внимание необходимость использования передовых подходов и методов для организации некоторых аспектов внутреннего управления. В таких случаях использование новой концепции может стать бесполезным и привести к отказу от использования подобных технологий, чтобы не тратить лишние средства, либо к использованию технологии для создания имиджа ведения инновационной деятельности.

Предложенная концептуальная модель и схема формирования интегрального показателя экспресс-анализа соответствия социально-экономического состояния предприятия нормативным требованиям со стороны регулятора дает возможность оценить не только текущее, но и спрогнозировать будущее состояние такого объекта исследования.

Построенный интегральный показатель может быть использован:

- для оценки близости состояния объекта исследования к критическому уровню, установленному контролирующим органом;
- для анализа тенденции изменения состояния объекта исследования на основе структурированных и неструктурированных данных во времени;
- для оценки достоверности отчетных данных о состоянии объекта исследования;
- в качестве гибкого инструментария, позволяющего сравнивать состояние разномасштабных объектов исследования.

Глава 2. Концептуальная модель экспресс-анализа соответствия состояния объекта исследования требованиям контрольных и надзорных органов

2.1 Поэтапная схема построения интегрального показателя экспресс-анализа соответствия социально-экономического состояния объекта исследования заявленным требованиям

Методологический подход построения интегрального показателя для экспресс-анализа соответствия социально-экономического состояния объекта исследования нормативным требованиям со стороны контрольных и надзорных органов в качестве отправной точки предлагает учитывать требования контролирующих органов, в то время как в большинстве российских и зарубежных исследований оценка состояния объекта исследования осуществляется исходя из требований, предъявляемых к объекту его владельцами или инвесторами.

Концептуальная модель показывает принципиальную возможность поэтапного выполнения экспресс-анализа соответствия состояния объекта исследования заявленным требованиям (рис. 2).



Рисунок 2. Концептуальная модель построения экспресс-анализа соответствия социально-экономического состояния объекта исследования требованиям и нормативам

Взаимодействие контрольных и надзорных органов с предприятиями отражено в виде информации, на основе которой происходит экспресс-оценка.

Предлагаемый гибкий инструментарий экспресс-анализа для принятия управленческих решений, разработанный на основе концептуальной модели, представляет собой последовательность из пяти этапов, начиная от требований со стороны контрольных и надзорных органов, разработки и оценки набора компонент, характеризующих объект исследования, их агрегирование в единый интегральный показатель на основе таблицы истинности, и заканчивая мониторингом и ранжированием объектов исследования по результатам расчетов

Концептуальная поэтапная схема на основе концептуальной модели для построения экспресс-анализа соответствия состояния объекта исследования заявленным требованиям подробно представлена на рис. 3.



Рисунок 3. Концептуальная поэтапная схема проведения экспресс-анализа соответствия социально-экономического состояния объекта исследования требованиям и нормативам

Схема проведения экспресс-анализа соответствия социально-экономического состояния объекта исследования требованиям и нормативам контрольных и надзорных органов состоит из пяти этапов:

1 этап. На первом (начальном) этапе, на основе информации о требованиях от контролирующего органа к социально – экономическому состоянию объекта исследования, определяются компоненты.

Сами требования, регламенты, целевые значения показателей представлены в федеральных или региональных законодательных актах, других нормативных документах, прочих требованиях регуляторов, сформулированных и задокументированных. Для формализации и извлечения ценной информации из нормативных документов привлекаются формализованные алгоритмы аддитивной регуляризации тематических моделей, реализованные в виде набора программ на языке Python [62].

Контрольным органом может выступать как внешняя организация (ЦБ, отраслевое министерство), так и внутренний контрольный орган или отдел самой организации. Для контрольных органов внутри организации могут нормативы могут быть в виде внутренних распоряжений предприятия или организации. Исходя из совокупности информации о требованиях и нормативах к социально – экономической деятельности объекта исследования, формируются характеристики его деятельности, требующие оценки. Таким образом достигается оценка объекта под критерии контрольных и надзорных органов.

2 этап. На основе разработанных теоретических подходов и методов к анализу информации об объекте исследования с учетом целевых показателей, определяются компоненты индикатора. Каждая компонента оценивает различные аспекты деятельности объекта исследования – экономические, социальные, финансовые, технические и т.п.

В зависимости от вышеизложенного формируются источники данных: сайты объектов исследования, новостные источники, электронные площадки и агрегаторы информации, сайты государственных органов власти. Для построения составляющих интегрального показателя предлагается использовать методы кластерного и регрессионного анализа, применяемые к

количественным и качественным данным об объектах. С помощью математических методов оценивается значение каждой составляющей. В таблице 2 Таблица приведен перечень выделяемых возможных компонент, относящихся к четырем блокам (в соответствии с метриками, предложенными Робертом С. Капланом и Дэвидом П. Нортоном) типов входной информации для расчета компонент, типов переменных рассчитанного значения каждой компоненты, и методов оценки значений компонент [66].

Таблица 2.

Компоненты интегрального показателя и методы их оценки

№	Компоненты	Тип исходной информации	Тип переменной рассчитанного значения компоненты	Метод оценки
1 Блок компонент. Характеристика финансового состояния объекта исследования				
1	Вероятность финансового неблагополучия	структурированная	категориальная, порядковая	Логистическая регрессионная модель
2 Блок компонент. Статусная идентичность объекта исследования				
2	Статус объекта исследования по масштабу	структурированная	категориальная	Кластерный анализ
3	Статус объекта исследования по принадлежности к аномальной группе	структурированная	категориальная	Кластерный анализ
3 Блок компонент. Характеристика внешней информационной среды				
4	Медийная активность относительно объекта исследования	слабоструктурированные, квазиструктурированные и неструктурированные данные	количественная	Тематическое моделирование
5	Положительная тональность упоминаний об объекте исследования в интернет-источниках		количественная	Тематическое моделирование
6	Негативная тональность упоминаний об		количественная	Тематическое моделирование

№	Компоненты	Тип исходной информации	Тип переменной рассчитанного значения компоненты	Метод оценки
	объекте исследования в интернет-источниках			
4 Блок компонент. Нормативные требования к состоянию объекта исследования				
7	Соответствие требованиям государственных органов	структурированная	бинарная или категориальная	Статистический и индексный анализ

Под каждую компоненту формируются свои источники информации: сайты объектов исследования, новостные источники, электронные площадки и агрегаторы информации, сайты государственных органов власти и т.п.

3 этап. Построение интегрального показателя.

Для построения интегрального показателя на основе оцененных разнородных значений компонент интегрального показателя формируется подход к получению его выходного значения – относительного показателя (в процентах, в баллах, в относительной шкале сравнения). Для этого значения компонент в качестве аргументов подставляются в логическую функцию, агрегирующую различные значения компонент в единый интегральный показатель с помощью таблицы ассоциаций.

4 этап. Рейтингование / ранжирование объектов исследования на основе полученных значений комплексного индикатора.

Проводится ранжирование всех предприятий внутри полученных выше групп по степени перспективности их стратегического развития в течение следующего года. Одним из результатов выполнения этого этапа является список объектов исследования, отсортированный по убыванию значений интегрального показателя.

5 этап. Оценка соответствия объекта исследования заявленным требованиям на основе рейтинга / ранга. На этом этапе выполняется контроль

состояния объектов исследования, производится переоценка имеющейся информации, обновление данных и мониторинг за состоянием объекта исследования повторным выполнением этапов 2–4.

После этапа ранжирования производится оценка соответствия объекта исследования заявленным требованиям в соответствие с полученными рангами (рейтингами). Исходя из полученных значений на текущий момент, вырабатываются рекомендации для надзорных органов относительно объектов управления. По мере появления новой информации в течение периода мониторинга за группой объектов исследования, регулярно производится обновление данных из открытых источников, и на обновленных данных производится переоценка модели.

Также после поступления официальной статистической отчетности по результатам применения предлагаемого инструментального решения может проводиться процедура сверки результатов экспресс-анализа на основе общедоступных данных и результатов аудита на основании официальной статистической отчетности. Такое сопоставление позволит провести корректировку выбранных компонент для применения концептуальной модели на других данных, а также сделать вывод об отклонении фактических значений от статистических, и принять комплекс упреждающих мер в случае необходимости.

2.2 Типологизация источников общедоступной информации и методы их обработки

2.2.1 Типологизация источников общедоступной информации

Интернет – это виртуальное информационное пространство, уникальное и универсальное по своим свойствам и функциям. Это, прежде всего, средство открытого хранения и распространения информации: научной, деловой, познавательной и развлекательной. Общедоступная информация, размещенная в открытом доступе, может быть представлена как

официальными органами, ведущими статистический учет, так и сторонними ресурсами.

Вся общедоступная информация может быть представлена в виде данных различного типа. В настоящее время все существующие данные можно разделить на несколько типов по усилению их структуры [61]:

- 1) структурированные;
- 2) слабоструктурированные;
- 3) квазиструктурированные;
- 4) неструктурированные.

К структурированным данным относятся данные с самой высокой степенью упорядоченности, они сформированы определенным способом, имеют заданную структуру, и описывают конкретную предметную область. В совокупности это позволяет проводить автоматизированный, достоверный и глубокий анализ этих данных без долгой предварительной предобработки. Чаще всего такая информация представлена в виде таблиц.

Слабоструктурированными принято именовать данные, которые не соответствуют четкой структуре таблиц и отношений в БД, но при этом содержат в себе специальные разграничители (теги), которые позволяют семантически разделять весь объем данных. В качестве примера можно привести XML-документы.

Данные с изменчивым, не постоянным форматом относят к квазиструктурированным. Они требуют специальной предобработки различными математическими инструментами, больших временных затрат на их подготовку. Примером таких данных может служить страница сайта.

Данные без определенной формы и структуры относят к неструктурированным данным. На настоящий момент такой формат данных является преобладающим в связи с развитием информатизации населения, примерно 80% всей имеющейся на данный момент информации, являются

неструктурированной. Примером таких данных являются изображения, видео, аудио и текстовая информация из социальных сетей.

Данные в зависимости от их типа требуют своих методов предобработки и обработки. Большинство методов математической статистики и эконометрики основаны на анализе структурированной информации. Методы машинного обучения, нейронные сети позволяют анализировать слабоструктурированные, квазиструктурированные и неструктурированные данные, выделяя в них закономерности. Также при проведении различных процедур предобработки эти данные могут быть сведены к структурированным и включены в классические математические модели.

Если неструктурированные данные представлены текстом, то предварительная их обработка с помощью методов векторизации и классификации позволяет привести их к структурированному виду.

Информация для формирования базы исследования для проведения экспресс-анализа может быть получена из различных источников, отличающихся статусом, частотой обновления и степенью достоверности предоставляемой информации.

Автором предложена типологизация источников общедоступной информации с учетом степени надежности и достоверности источника, позволяющая реализовать концептуальную модель для построения интегрального показателя экспресс-анализа соответствия социально-экономического состояния предприятия нормативам и регламентам со стороны контрольных и надзорных органов.

Рассмотрим характеристики предлагаемой типологизации общедоступных данных.

Источник информации может быть представлен в нескольких видах. Официальные источники информации – это интернет-ресурсы, созданные государственными органами, деятельность и раскрытие информации опирается на соответствующее положение, существует регламент

относительно опубликованных данных, регламент их обновления, а также внутренняя система проверки достоверности и качества данных. Можно выделить агрегаторы и генераторы данных. Агрегаторы – это органы государственной статистики и другие государственные отделы и департаменты, которые на разных уровнях (федеральный, региональный) обобщают данные и поддерживают интернет-источники их публикации (gks.ru, data.mos.ru).

Официальный генератор данных – это государственная организация, которая, согласно постановлению о раскрытии информации, предоставляет имеющиеся у него сведения для пользования. Так, многие министерства публикуют в открытом доступе информацию о проводимых проверках, судебных процессах, задолженностях в отношении юридических лиц. Такие данные обладают высокой достоверностью и надежностью. Такие данные как правило, уже структурированы, они удобны для сбора и автоматизации и требуют минимальной предобработки.

Информация для проведения исследования также может быть взята с интернет-ресурса о каждом объекте исследования, в этом случае за достоверность отвечает сам владелец информации, что может приводить к снижению достоверности получаемых данных. В зависимости от вида экономической деятельности (обучение, производство, ресурсы и т.п.) существуют стандарты и требования к раскрытию информации. При наличии таких требований (например, ВУЗы), сбор упрощается и может быть автоматизирован, при отсутствии же таких требований к раскрытию данных они представляют собой разрозненные куски информации, не структурированные или слабоструктурированные, с разной степенью достоверности. Проблема достоверности открытых данных может быть решена на основе существующих методов и алгоритмов машинного обучения, использующих метрики качества данных [38].

Помимо официальных агрегаторов данных существуют и неофициальные, это ресурсы, на которых агрегирована информация по теме, и достоверность которой определяется самим сборщиком. К таким ресурсам относятся рейтинговые агентства, которые предоставляют обобщенную информацию по предприятиям, на основе которой делается анализ. Сертифицированные ЦБ РФ рейтинговые агентства предоставляют информацию согласно законодательству, поэтому информация обладает высокой достоверностью.

К неофициальным генераторам данных относятся площадки, публикующие информацию об объектах исследования, о различных аспектах – сайты объявлений, сайты торговых площадок, маркетплейсы и т.п. Владельцы таких сайтов проверяют поданную информацию, но за достоверность отвечают сами владельцы.

Информацию с качественной характеристикой об объекте исследования можно получить с различных тематических ресурсов, публикующих преимущественно неструктурированные текстовые данные. Такая информация требует особых способов сбора, агрегации и предобработки, но позволяет дополнить количественные характеристики объекта исследования его качественными динамическими показателями.

В приложении 1 представлена таблица типологизации источников общедоступной информации.

Предложенная типологизация источников информации позволяет сформировать базу данных исследования с разной достоверностью. Данные, полученные из официальных источников информации, а также от прямых поставщиков данных, обладают высокой степенью достоверности, хотя, как правило, низкую периодичность обновления (за счет процедуры верификации). Информация, получаемая от агрегаторов данных, может содержать ошибки, но обладает высокой вариабельностью и масштабом охвата качественных характеристик об объекте исследования.

В последние годы государственные органы власти обеспечивают доступ к данным ведомств согласно положению о раскрытии информации, что подтверждает текущую актуальность изменения существующих подходов и методик оценивания состояния предприятия к оценке на основе общедоступных данных. Достоверность этих данных подтверждается государственным органом власти, они соответствуют структуре, утвержденной в соответствующей методике, и, как правило, имеют несколько форматов, доступных для скачивания, что обеспечивает удобство и быстроту работы с ними.

Данные с интернет-источников рейтинговых агентств носят субъективный характер, поэтому рекомендуется выполнять проверку качества и согласованности таких данных. Так, существует перечень сертифицированных рейтинговых агентств, а также государственных органов, публикующих рейтинги, а также методику рейтингования по предприятиям:

1. ЦБ РФ, рейтинг коммерческих банков

2. «Эксперт РА» - российское кредитное рейтинговое агентство, включенное в реестр кредитных рейтинговых агентств Банка России. Занимается присвоением рейтингов, а также исследовательско-коммуникационной деятельностью с 1997 г. Публикует рейтинги коммерческих банков, предприятий, образовательных учреждений.

3. Международные рейтинговые агентства, аккредитованные ЦБ РФ: Standard & Poor's, Fitch Ratings, Moody's.

4. Российские рейтинговые агентства: НРА, «Рус-Рейтинг», АК&М и недавно созданная группой крупных акционеров при поддержке ЦБ АКРА).

Данные, поступающие из неофициальных источников позволяют дополнить структурированные количественные данные качественными характеристиками объектов исследования. Современный тренд в ИТ индустрии – рост объемов текстовой информации, так как текст остается одним из основных видов информации в большинстве электронных баз

данных и хранилищ. Развиваются и расширяются подходы и методы к обработке текстовой информации для формирования смысловых сущностей и показателей, характеризующих качественные характеристики объекта исследования. Поэтому привлечение неструктурированной информации – актуальный и современный тренд информационных технологий, обеспечивающий конкурентоспособность предлагаемой методики.

2.2.2 Методы обработки общедоступной информации

Вместе со стремительным накоплением информации быстрыми темпами развиваются и технологии анализа и обработки этих данных. Если еще десять лет назад было возможно, скажем, лишь сегментировать предприятия на группы со схожими предпочтениями, то теперь возможно строить модели для каждого предприятия в режиме реального времени, анализируя, например, его контракты на портале госзакупок. Современная модель должна учитывать как общие характеристики, так и индивидуальные для обособленной группы. В связи со стремительным развитием компьютерных технологий все чаще предлагается модифицировать и управлять моделью для тысяч объектов, имеющих общие черты, но и некоторые существенные отличия, анализируя множество их характеристик деятельности. Существует множество разных подходов к моделированию – эконометрический, при котором проверяются поставленные гипотезы, статистический, основанный на анализе данных, машинного обучения.

В последние годы активно развиваются особые методы анализа данных, так называемый Data mining (добыча данных, интеллектуальный анализ данных), использующий широкий потенциал современных методов классификации, моделирования и прогнозирования для выявления в изучаемом наборе данных новых сведений, которые не могут быть получены путем простого статистического анализа. Другими словами, методы Data mining являются средством извлечения из набора данных ранее неизвестных

скрытых взаимосвязей между переменными, характеризующих объект, предоставляя исследователю новые, в практическом смысле полезные и доступные для интерпретации знания, необходимые для принятия решений в различных сферах человеческой деятельности, в том числе и в менеджменте [14, 19].

Все методы принято разделять на два класса: дескриптивные методы и методы прогностической аналитики. Методы дескриптивного анализа, к которым относятся большинство моделей эконометрики, позволяют получить интерпретируемые (человеко-читаемые) модели, удобные для проверки гипотез относительно данных, зависимостей между факторами, их совокупного влияния и т.п. Для задач прогнозирования важнее точность модели, чем ее интерпретируемость. Поэтому задачи прогностической аналитики, реализуют как-правило с помощью методов машинного обучения (ММО), обладающими отличными прогностическими качествами.

Одним из перспективных направлений Data mining является предиктивная (или прогностическая) аналитика (Predictive analytics), которую используют для прогнозирования неизвестных событий в будущем. Ее основа - математическое моделирование, то есть создание модели, как набора правил, формул или уравнений, извлеченных из исходных данных и позволяющих генерировать предсказания на этой основе.

Распространенным методом прогностического моделирования является метод классификации, реализованный во многих компьютерных системах прогнозирования для оценки вероятности и риска возникновения события. Модели классификации используют значения одного или нескольких входных полей (независимых переменных или предикторов), чтобы предсказать значения одного или нескольких выходных целевых полей, или полей назначения (зависимых переменных, спрогнозированных моделью на основе анализа входных предикторов). Во многих областях экономики используют различные способы моделирования и алгоритмы машинного обучения:

деревья решений (алгоритмы дерева C&R, QUEST, CHAID, C5.0), регрессии (линейная, логистическая, обобщенная линейная и Кокса), нейронные сети, модели опорных векторов, Байесовские сети и др. [19-23].

Для текстовой информации в настоящий момент активно внедряется машинная обработка данных для применения семантического анализа [37,52,75,77].

2.3 Алгоритм построения интегрального показателя для проведения экспресс-анализа

2.3.1 Общая схема алгоритма построения интегрального показателя экспресс-анализа

Интегральный показатель представляет собой агрегирование значений компонент интегрального показателя, из которых и определяется его итоговое значение оценки соответствия социально-экономического состояния объекта исследования нормативам и регламентам со стороны контрольных и надзорных органов. Для этого значения компонент в качестве аргументов подставляются в функцию свертки, преобразующую различные значения компонент интегрального показателя в единое значение.

Рассмотрим такую функцию, заданную с помощью таблицы поиска. Пусть от контролирующего органа поступило n требований, сформулированных в n критериев, и для интегрального показателя сформировано k компонент I_j , где: $I_j \in X$, $j = [1, \dots, k]$, X – это множество всех значений, которые принимают компоненты интегрального показателя, обычно $X \in R$. При этом одно требование к состоянию объекта исследования от контролирующего органа может быть описано несколькими компонентами интегрального показателя, т.е. $k \geq n$. Значения интегрального показателя u_i принимает значения из некоторого заданного конечного дискретного множества Y , т.е. $u_i \in Y$, где: $i = [1, \dots, p]$. Как правило, для

удобства интерпретации, выбирают множество с небольшим количеством категорий – 2,3 или 5, но может быть использована и 100-бальная шкала.

Функция $\Psi(I_1, I_2, \dots, I_k)$ задается таблицей поиска (Табл.3), вводимой на основании важности критериев, полученных от контролирующего органа (Рис. 4).

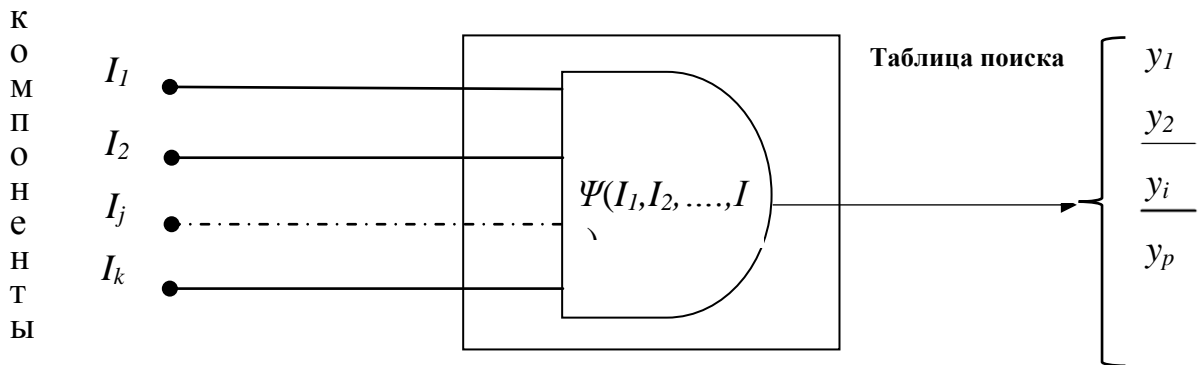


Рисунок 4. Схема формирования интегрального показателя с использованием функции, заданной с помощью таблицы поиска

Каждая компонента I_j может принимать q_j значений. Значения компоненты $I_j^{q_j} \in X_j$, где: $j = [1, \dots, k]$, $q_j = [1, \dots, Q_j]$, $X_j \in X$.

Общее количество комбинаций M всех возможных значений всех k компонент определяется следующим образом:

$$M = \prod_{j=1}^k Q_j \tag{1}$$

В общем случае таблица состоит из M строк и k столбцов. Первая строка (сочетание s_1) представляется первыми значениями каждой из k компонент I_j^1 . Последняя строка таблицы представляется максимальными значениями каждой из k компонент $I_j^{Q_j}$. Строки, находящиеся между 1-ой и последней строкой, могут формироваться различным образом, но так чтобы осуществлялся полный перебор всех возможных комбинаций значений, принимаемых каждой из компонент I_j .

Таблица 3.

Общий вид многомерной таблицы поиска значений интегрального показателя

s_m - номер сочетания, $s_m = [I_1, \dots, M]$	Компоненты интегрального показателя					Значение интегрального показателя
	I_1	...	I_j	...	I_k	
s_1	I_1^1	...	I_j^1	...	I_k^1	$y_{s_1} = \Psi(I_1^1, \dots, I_j^1, \dots, I_k^1)$ $y_{s_1} \in Y$
s_2	I_1^2	...	I_j^1	...	I_k^1	$y_{s_2} = \Psi(I_1^2, \dots, I_j^1, \dots, I_k^1)$ $y_{s_2} \in Y$
....
s_m	$I_1^{q_1}$...	$I_j^{q_j}$...	$I_k^{q_k}$	$y_{s_m} = \Psi(I_1^{q_1}, \dots, I_j^{q_j}, \dots, I_k^{q_k})$ $y_{s_m} \in Y$
....
s_M	$I_1^{Q_1}$...	$I_j^{Q_j}$...	$I_k^{Q_k}$	$y_{s_M} = \Psi(I_1^{Q_1}, \dots, I_j^{Q_j}, \dots, I_k^{Q_k})$ $y_{s_M} \in Y$

Использование таблицы поиска позволяет нивелировать возможные массовые специально организованные компании, порочащие репутацию объекта исследования, с помощью негативных высказываний в СМИ и интернет-ресурсах.

Значения переменных, измеренных в количественной шкале, переводятся в порядковые значения, на основе использования статистических характеристик разброса значений (межквартильный размах), а также логических правил разбиения, позволяющих определить точки разбиения значений компонент на категории (порог отсека). Для формирования порогов отсека значений компонент индикатора на категории применяется подход на основе математической статистики, оценивающий аномальность значений компонент относительно всей рассмотренной выборки с данными (табл. 4).

Таблица 4.

Пороги отсечения значений компонент интегрального показателя

№	Компоненты интегрального показателя	Тип значений компоненты	Порог отсечения значений компонент
1	Вероятность финансового неблагополучия объекта исследования	Количественная	Порог отсечения cutoff — число, при котором разница функций распределения $F_1(x)$ и $F_0(x)$ достигает максимума, такое что: $I_1 = 0$, если $P(Y = 1) < \text{cutoff}$; $I_1 = 1$, если $P(Y = 1) \geq \text{cutoff}$
2	Статус объекта исследования по масштабу	Категориальная	На основе функции принадлежности объекта x к кластеру m определяются значения компонент: $I_2 = m$, $m \in [1, \dots, M]$, $I_3 = c_l$, $c_l \in [1, \dots, M]$, m – номер кластера; c_l – кластер аномалий; M – общее количество полученных кластеров.
3	Статус объекта исследования по принадлежности к аномальной группе	Категориальная	
4	Медийная активность относительно объекта исследования	Количественная	Межквартильный размах IQR_n по выборке размера n : $IQR_n = Q_3 - Q_1$, где $F_n(Q_3) = 0.75$; $F_n(Q_1) = 0.25$; $F_n(x)$ — выборочная функция распределения. $I_k = 1$, если $X_k^j < 1.5 * IQR_n (X_k)$ $I_k = 2$, если $1.5 * IQR_n (X_k) \leq X_k^j \leq 3 * IQR_n (X_k)$ $I_k = 3$, если $X_k^j > 3 * IQR_n (X_k)$ $k = 4, 5, 6$ – номер компоненты; j – номер объекта исследования.
5	Положительная тональность упоминаний об объекте исследования в интернет-источниках	Количественная	
6	Негативная тональность упоминаний об объекте исследования в интернет-источниках	Количественная	
7	Соответствие требованиям государственных органов	Количественная или бинарная	Если требования директивны, то определяется бинарный показатель: $I_7 = 1$, если $X_7^j > 0$ (было хотя бы одно нарушение) $I_7 = 0$, если $X_7^j = 0$ (не было ни одного нарушения) Если требования не директивны, то есть, возможно продолжение деятельности при наличии нарушений, то: $I_7 = 1$, если $X_7^j \geq 1.5 * IQR_n (X_7)$ $I_7 = 0$, если $X_7^j < 1.5 * IQR_n (X_7)$

2.3.2 Компонента 1. Вероятность финансового неблагополучия

Эта компонента представляет собой вероятность наступления неблагоприятного финансового состояния объекта исследования (банкротство, отзыв лицензии по финансовым причинам, долговые обязательства). Для оценивания этой вероятности применяется регрессионная логистическая модель на основе данных финансовой отчетности, а также показателей их волатильности: стандартного отклонения и дисперсии, данных макроэкономических переменных, данных о государственных закупках в качестве поставщика или покупателя. В общем виде логистическая регрессионная модель принимает вид (2):

$$P(Y = 1|x, m, v) = \frac{1}{1 + e^{-z}} \quad (2)$$

$$z = \beta_0 + \sum \beta_i x_i + \sum \gamma_j m_j + \sum \varphi_k v_k$$

где $P(Y = 1|x, m, v)$ – условная вероятность неблагоприятного финансового состояния объекта исследования;

β_0 – константа;

x_i – переменные, характеризующие финансовое состояние объект исследования;

m_j – переменные, характеризующие внешнюю по отношению к объекту исследования среду (макроэкономические факторы);

v_k – неколичественные показатели деятельности объекта исследования;

$\beta_i, \gamma_j, \varphi_k$ – коэффициенты регрессии, которые должны быть оценены.

Так как значительная часть факторов, характеризующих объект исследования, измеряется по качественным шкалам — номинальным или порядковым, то при бинаризации этих факторов, т. е. превращении их в системы фиктивных переменных, необходимо добавить в спецификацию модели новые переменные, равные количеству категорий значений минус

один, выбранный за базовую. Для большинства реализаций логистической регрессии это единственно возможный вариант использовать как качественные, так и количественные переменные.

Если уровней достаточно много, и они хорошо представлены в тренировочном множестве, т.е. укрупнение уровней не имеет смысла или неадекватно решаемой задаче, это приводит к существенному росту ее размерности. В результате исследователь сталкивается с эффектом «проклятия размерности» (the curse of dimensionality) [67]. Таким образом, с ростом размерности задачи требуется адекватное увеличение числа измерений для сохранения точности оценки силы влияния каждого фактора. Поэтому предлагается ограничиться небольшим количеством категорий качественных переменных.

Метрики оценки качества классификации

Для оценки качества полученной модели чаще всего используются метрики – чувствительность, специфичность, ROC кривая и кросс-валидация. В результате обучения алгоритма классификации на обучающем множестве возникает классификация его объектов выбранным алгоритмом. Она может совпадать или не совпадать с истинным классом объектов в тренировочном множестве (табл. 5).

Таблица 5.

Таблица результатов классификации

Истинный класс\Прогноз	Positive (положительный ответ)	Negative (отрицательный ответ)
Positive (положительный ответ)	TP	FN
Negative (отрицательный ответ)	FP	TN

где P (positive) — количество положительных ответов (результатов наблюдений);

N (negative) — количество отрицательных ответов;

TP (true positive) — правильно распознанные положительные ответы;

FP (false positive) — ошибочно распознанные отрицательные ответы как положительные;

TN (true negative) — правильно распознанные отрицательные ответы;

FN (false negative) — ошибочно распознанные положительные ответы как отрицательные;

При этом выполняются следующие соотношения:

$$P = TP + FN;$$

$$N = TN + FP;$$

На основе этой таблицы определяются следующие характеристики качества классификатора:

Общая погрешность классификации (Error rate):

$$ER = \frac{FP+FN}{\Sigma}, \Sigma = TP + FN + FP + TN \quad (3)$$

Чувствительность (Sensitivity) $Sen = \frac{TP}{TP+FN}$ (4) — способность классификатора правильно классифицировать *Positive*.

Специфичность (Specificity) $Spe = \frac{TN}{TN+FP}$ (5) — способность классификатора правильно классифицировать *Negative*.

В связи со специфичностью иногда рассматривают $1-Spe$ — склонность классификатора принимать *Negative* за *Positive*.

График зависимости чувствительности Sen от $1-Spe$ носит название ROC-кривой, и является характеристикой прогностических способностей модели.

Техника кросс-валидации возникла как реакция на необходимость использования в исследовании одного множества объектов без возможности получения новых данных о них в новых условиях или данные о новых объектах. Это ситуация нередко складывается в тех отраслях знаний, где затруднено или невозможно проведение планируемых экспериментов, например, в эконометрике. В этом случае, эффективной стратегией проверки качества алгоритма классификации является разбиение случайным образом исходного

множества данных на несколько одинаковых по величине подмножеств (карманов) и выделение одного из них в качестве тестового множества. Остальные карманы включаются в тренировочное множество. Процедура повторяется перебором карманов в качестве тестовых множеств. Таким образом, например, при десяти карманах мы получим десять оценок качества классификации, что даст нам возможность подсчитать среднюю оценку качества и оценить его разброс. Эту процедуру можно повторять многократно, каждый раз по-новому выполняя разделение на карманы. Затем рассчитывается усредненное значение качества классификации, тем самым избегая переобучения.

2.3.3 Компоненты 2 и 3. Статус объекта исследования по масштабу и принадлежности к аномальной группе.

Эти компоненты представляет собой результаты кластеризации по определению принадлежности объекта исследования к одному из классов и позволяют учесть особенности всех объектов исследования данного типа по признакам местоположения, масштаба, вида деятельности и т.п. При этом учитывается специфика полученного кластера объектов, что позволяет более объективно оценить состояние предприятия относительно объектов из его класса.

Кластерный анализ является одним из способов систематизации и классификации объектов исследования. Он представляет собой набор методов, используемых для группировки объектов или событий в относительно однородные группы, которые называют *кластерами* (clusters). Объекты в каждом кластере должны быть похожи между собой и отличаться от объектов в других кластерах.

Кластерный анализ позволяет открыть в данных ранее неизвестные закономерности, которые практически невозможно исследовать другими способами и представить их в удобной для пользователя форме. Методы

кластерного анализа используются как самостоятельные инструменты исследований, так и в составе других средств Data mining (например, нейросетей). Кластерный анализ используют для решения нескольких задач: определить однородные (гомогенные) группы, рассортировать объекты по набору характеристик или выделить особенные специфические объекты. При кластеризации учитывают тип показателей группировки – количественные или качественные.

Сложный для решения вопрос кластерного анализа — это вопрос о количестве кластеров, которые следует выбрать. Здесь нет твердых правил, позволяющих быстро принять решение. Некоторые методы кластерного анализа требуют задания этого значения, тогда он выбирается на основе поставленной задачи, например, для удобства последующей интерпретации результата или исходя из экономических свойств данных, а некоторые методы позволяют определить наилучшее значение количества кластеров, при котором достигается высокая однородность, на основе оптимизации по какому-то критерию (например, критерий Акаике).

Таким образом, при задании количества кластеров до анализа есть возможность получить в результате группы с более высокой чем возможно неоднородностью, а при автоматической процедуре выбора числа кластеров – сложно интерпретируемый результат.

Результат кластеризации – получение групп с наличием общих черт. Описание и профилирование кластеров осуществляется на основе анализа кластерных центроидов. Центроиды представляют средние значения объектов, содержащиеся в кластере по каждой из переменных. Они позволяют выявлять индивидуальные характеристики кластера как группы объектов, присвоить ему номер или метку, сформировать описание и построить управленческую модель для полученных групп.

Подход к выбору метода кластеризации основан на анализе следующей информации:

- 1) поставленная управленческая задача;
- 2) тип анализируемых данных и показателей кластеризации;
- 3) объем исходных данных и имеющихся ресурсов для их обработки;
- 4) требования к интерпретируемости и визуализации результатов анализа.

Алгоритмы кластеризации подразделяются на два типа:

1. Иерархические методы;
2. Неиерархические методы.

Методы иерархической кластеризации бывают двух видов [30, 31]:

1. Агломеративные (объединяющие).

В этой категории методов происходит объединение исходных объектов и уменьшение количества кластеров [32]. Такой подход осуществляется «снизу-вверх»: создание небольших кластеров и объединение их в более крупные.

2. Дивизивные (разъединяющие).

Для алгоритмов дивизивного вида характерно начальное условие наличия одного кластера. Этот изначальный кластер делится на более мелкие кластеры [32]. Разъединяющие алгоритмы работают «сверху-вниз».

Недостатком этих методов является вычислительная сложность на данных большой размерности. Для иерархических методов кластеризации характерной особенностью является то, что наблюдения, попавшие однажды в кластер, при дальнейшем объединении (разъединении) объектов не могут переместиться в другой кластер, в отличие от неиерархических методов.

Неиерархические методы кластеризации часто называют методами *k*-средних. Основная отличительная идея неиерархических методов кластеризации – определение центра кластера и группировка всех объектов, находящихся на расстоянии от центра кластера в пределах заданного порогового значения [30, 31].

Суть этой группы методов - определить центр кластера, а в следующую очередь сгруппировать все объекты в пределах заданного от центра

порогового значения. Эти методы предполагают перебор всех значений для определения расстояния между объектами, они включают последовательный пороговый метод, параллельный пороговый метод и оптимизирующее распределение.

$$V = \sum_{i=1}^k \sum_{x_j \in S_i} (x_j - \mu_i)^2 \quad (6)$$

где k — число кластеров;

S_i — полученные кластеры, $i=1, 2, \dots, k$;

μ_i — центры масс векторов $x_j \in S_i$.

Метод оптимизирующего распределения будет иметь отличия от остальных пороговых методов в том, что при отнесении объекта к одному кластеру при нахождении объекта в пределах порогового значения впоследствии возможно включить его в другой кластер (перераспределить) с учетом порогового значения в целях оптимизации суммарного критерия, которым является среднее внутрикластерное расстояние, установленное для данного числа кластеров.

Для данных большой размерности с неизвестным числом кластеров часто применяют метод **BIRCH** (двухшаговая или двухступенчатая кластеризация), основанный на методе k -means (k -среднее). Он реализован в большинстве статистических пакетов, например, в IBM SPSS Statistics. Двухступенчатая кластеризация не требует задания количества кластеров, так как на первом шаге определяется оптимальное количество кластеров, а затем уже происходит разбиение на однородные группы. Данный метод позволяет анализировать большие объемы количественных данных и хорошо работает при небольших объемах памяти.

В этом алгоритме реализован двухэтапный процесс кластеризации. Первый этап заключается в формировании предварительного набора кластеров, т.е. определении количества кластеров. Следующий этап

заключается в применении к выявленным кластерам других алгоритмов кластеризации, которые были бы пригодны в работе с оперативной памятью.

Среди новых масштабируемых алгоритмов можно отметить также алгоритм **CURE** - алгоритм иерархической кластеризации, где понятие кластера формулируется с использованием концепции плотности, однако он реже встречается в сертифицированных статистических пакетах.

Качество кластеризации

Качество полученной кластеризации может быть оценено с помощью метрики, предлагаемой многими статистическими пакетами - силуэтной меры (2):

$$Sil = \frac{1}{N} \sum_{c_k \in C} \sum_{x_i \in c_k} \frac{b(x_i, c_k) - a(x_i, c_k)}{\max(a(x_i, c_k), b(x_i, c_k))} \quad (7)$$

где Sil – общее значение силуэтной меры кластеризации всех данных;

N – общее количество объектов в выборке;

C – множество всех кластеров;

c_k – k -ый кластер на множестве C ;

x_i – i -ый объект, $i \in [1, N]$;

$a(x_i, c_k) = \frac{1}{|c_k|} \sum_{x_j \in c_k} \|x_i - x_j\|$ - среднее расстояние от объекта $x_i \in c_k$ до других объектов x_j из этого кластера c_k (компактность);

$|c_k|$ - количество объектов в кластере c_k ;

$b(x_i, c_k) = \min_{c_l \in C \setminus c_k} \left\{ \frac{1}{|c_l|} \sum_{x_j \in c_l} \|x_i - x_j\| \right\}$ - среднее расстояние от объекта $x_i \in c_k$ до объектов x_j из другого кластера $c_l: k \neq l, k, l \in [1, C]$.

Силуэтная мера Sil принимает значения на отрезке от -1 до +1, где:

- 1 - все наблюдения расположены точно в центрах их кластеров;
- -1 - все наблюдения расположены в центрах некоторых других кластеров;
- 0 - наблюдения расположены в среднем на равных расстояниях от центра их кластера и центра ближайшего кластера.

2.3.4 Компоненты 4, 5 и 6. Медийная активность относительно объекта исследования, положительная и негативная тональность упоминаний об объекте исследования в интернет-источниках.

Оценка этих компонент представляет собой анализ неструктурированных или слабоструктурированных данных, преимущественно текстовых. Для проведения тематического моделирования по оценке значений компонент, характеризующих медийную активность и тональность упоминаний относительно объекта исследования, требуется предварительная предобработка этих данных – техническая и лингвистическая. На этапе предобработки из текста удаляются все предлоги и другие «лишние» слова, не несущие смысла, остальные слова приводятся именительный падеж, убирается окончание. Таким образом получается словарь слов.

Для использования текстовых данных в математических алгоритмах требуется предварительная предобработка этих данных, называемая векторизацией. Векторизация – это процесс конвертации текстовых документов в числовой вид. Существует множество разных алгоритмов векторизации, отличающихся размерностью обрабатываемых данных, требованиями к мощности инструментов. Реализованным во многих статистических пакетах и библиотеках (IBM SPSS Statistics, R, Gretel) является алгоритм Bag-of-Words (или «мешок слов»), широко используемый в работе с естественными языками. «Мешок слов» представляет собой векторное представление неупорядоченного набора слов в вектор размерности n [35,36,37,42]. Схематично алгоритм может быть представлен следующим образом.

Весь текст можно представить как набор слов и предлогов, предобработанных и преобразованных в виде отдельных термов (t_j), которые с помощью данного алгоритма переводятся в числовые данные из пространства R^n .

$$B: words \rightarrow R^n \quad (8)$$

$$B('some\ text\ in\ the\ Internet') = (w_{i,1}, w_{i,2}, \dots, w_{i,n}),$$

где t_j – терма j ;

$w_{i,j}$ – вес терма j в документе; вес документов нормируют так, чтобы $0 < w_{ij} < 1$, для $\forall i$;

n – количество термов в пространстве.

Документ в таком случае задаётся следующим образом:

$d = (w_1, w_2, \dots, w_{|V|})$, где d – вектор документа;

$|V|$ – количество уникальных термов в документе;

Вес терма можно задать несколькими вариантами:

1) Бинарным образом

$$w_i = \begin{cases} 1, & t_i \in d \\ 0, & t_i \notin d \end{cases} \quad (9)$$

2) По количеству вхождений терма

$w_i = n_i$, где n_i – количество вхождений терма в документ.

3) Частота терма (Term Frequency – TF).

$$w_i = tf(t_i, d) = \frac{n_i}{\sum_{k=1}^{|V|} n_k}, \quad (10)$$

где $tf(t_i, d)$ – частота терма;

n_i – количество вхождений терма в документ;

$\sum_{k=1}^{|V|} n_k$ – количество термов в документе.

4) Term Frequency – Inverse Document Frequency (TF-IDF).

Представление в виде двух параметров: $w_{ij} = tf_i * idf_i$, где tf_{ij} – это отношение числа термов t_i в документе d_j к общему числу термов в этом документе, а idf_i – число, обратное количеству документов, в котором встречается терм t_i . Таким образом, чем чаще слово встречается в этом документе, но реже встречается вообще во всех документах, тем больше вес этого терма в данном документе.

idf – обратная частота документа.

$$tf(t_i, d) = \frac{n_i}{\sum_{k=1}^{|V|} n_k} \quad (11)$$

$$idf(t, d) = \log \frac{|D|}{|d_i \supset t_i|} \quad (12)$$

$|d_i \supset t_i|$ – количество документов, в которых встречается t_i ;

$|D|$ – количество документов в корпусе.

Тогда вес считается следующим образом:

$$w_i = tf - idf(t_i, d, D) = tf(t_i, d) \times idf(t, d)$$

Для текстовых данных актуальной является задача определения тональности текста. Тональность – это эмоциональное отношение автора к некоторому объекту, выраженное в тексте [38,41].

Один из способов определения тональности предложения - поиск эмоциональной составляющей в тексте по ранее сформированным тональным словарям с применением анализа лингвистики. Использование готовых словарей к словарям, полученным на этапе предобработки, позволяет классифицировать текстовые единицы (предложения, слова) на 3 категории: амбивалентное, положительное и отрицательное.

Также оценку тональности методами машинного обучения и компьютерной лингвистики можно выполнять для векторизованных данных. Для них возможно применение семантического и математического анализа для выявления классов, кластеризации, корреляций и других видов зависимостей.

Для расчета значений компонент применяются методы статистики для обобщения информации об объекте исследования, например, с помощью прямого подсчета определяется общая тональность текста.

2.3.5 Компонента 7. Соответствие требованиям государственных органов.

Компонента представляет собой оценку количества нарушений в деятельности объекта исследования, в случае если заданы нормативные и пороговые значения со стороны контрольных или надзорных государственных органов власти.

Информация о имеющихся нарушениях может быть взята из общедоступных источников официальной информации:

1) Единый реестр проверок (ЕРП). Обеспечивает учёт проверок государственного и муниципального контроля (ч. 1 ст. 13.3 294-ФЗ). Контролю подвергаются индивидуальные предприниматели и юридические лица, Правила формирования и ведения единого реестра проверок заработали с 1 июля 2015 года путём введения в закон 294-ФЗ статьи 13.3. ЕРП содержит информацию о плановых и внеплановых проверках юридических лиц и индивидуальных предпринимателей. В информационной системе также есть сведения о результатах уже прошедших проверок, принятых мерах по пресечению или устранению последствий выявленных нарушений. Ведёт ЕРП Генеральная прокуратура. В ст. 13.3 294-ФЗ подробно описаны правила включения информации в ЕРП. В ней можно найти требования к порядку создания ЕРП, порядку присвоения учётного номера проверке, состав информации, которая включается в ЕРП и т.д. [4].

2) Сайт ЦБ РФ. Банк России устанавливает нормативы, которые обязана выполнять каждая кредитная организация в нашей стране. В случае несоблюдения нормативов регулятор может взыскать с кредитной организации штраф, ввести запрет на осуществление ею некоторых банковских операций (например, на прием вкладов от населения, назначить в банке временную администрацию), а в некоторых случаях даже отозвать у банка лицензию. Информация о значении нормативов по каждой кредитной организации публикуется на сайте ЦБ РФ.

3) На сайте Федеральной службы судебных приставов (ФССП России) опубликованы открытые данные, размещаемые в соответствии со статьей 6.1 Федерального закона от 02.10.2007 № 229-ФЗ «Об исполнительном производстве» Федеральная служба судебных приставов создает и ведет банк данных исполнительных производств в электронном виде. Согласно с «Порядком создания и ведения банка данных в исполнительном

производстве Федеральной службы судебных приставов в электронном виде» общедоступная часть Банка данных публикуется на официальном сайте ФССП России. Сведения о исполнительном производстве ФССП России, указанные в части 3 статьи 6.1 от 02.10.2007 № 229-ФЗ «Об исполнительном производстве», являются общедоступными до дня окончания или прекращения исполнительного производства, за исключением сведений о возвращении исполнительного документа взыскателю по основаниям, предусмотренным пунктами 3 и 4 части 1 статьи 46 Федерального закона, или об окончании исполнительного производства по основаниям, предусмотренным пунктами 6 и 7 части 1 статьи 47 Федерального закона, которые являются общедоступными в течение трех лет со дня окончания исполнительного производства [9].

4) Иные ведомственные и отраслевые источники информации, публикующие сведения о значении нормативов объекта исследования согласно распоряжениям и положениям, утвержденным соответствующими органами государственной власти.

Если требования директивны и предусмотрена блокировка (остановка) деятельности предприятий в случае присутствия нарушений, то значения компоненты 7 определяются как бинарный показатель, характеризующий наличие или отсутствие таких нарушений.

Если требования контрольных органов не приводят к немедленной остановке деятельности в случае нарушений, то значения компоненты 7 определяются как порядковый показатель числа таких нарушений.

2.4 Информационно-логическая модель экспресс-анализа соответствия состояния объекта исследования нормативам и регламентам со стороны контрольных и надзорных органов на основе общедоступной информации

Консолидированным представлением вышеизложенного является информационно-логическая модель экспресс-анализа соответствия социально-экономического состояния объекта исследования требованиям контрольных и надзорных органов, представленная на рисунке 5. На рис. 6 этап 3 представлен в общем виде. В рамках этого этапа производится оценивание компонент интегрального показателя и расчет значений самого интегрального показателя в зависимости от значений расчета каждой компоненты из набора. Этап 3 является ключевым в алгоритме расчета интегрального показателя, на рис. 6 представлена детальная информационно-логическая модель этого этапа.

Предлагаемая информационно - логическая модель представляет собой объединение алгоритма расчета отдельных компонент интегрального показателя с применением математических, эконометрических и статистических методов, характеристик входной и выходной информации на каждом этапе, и, собственно, алгоритма расчета значений интегрального показателя через свертку значений отдельных компонент, преобразуемых с помощью таблицы поиска.

Суть предложенной концептуальной модели заключается в решении следующих задач:

- Формирование компонент, формализующих характеристики, предъявляемые к объекту исследования, определение источника данных.
- Расчет компонент интегрального показателя согласно предлагаемым методам;

- Расчет с помощью таблицы поиска значения интегрального показателя, с учетом порогов отсечения значений компонент интегрального показателя;

Таким образом, предложенная информационно – логическая модель представляет собой гибкий инструментарий для решения проблемы оценки соответствия социально-экономического состояния предприятия требованиям контрольных и надзорных органов.

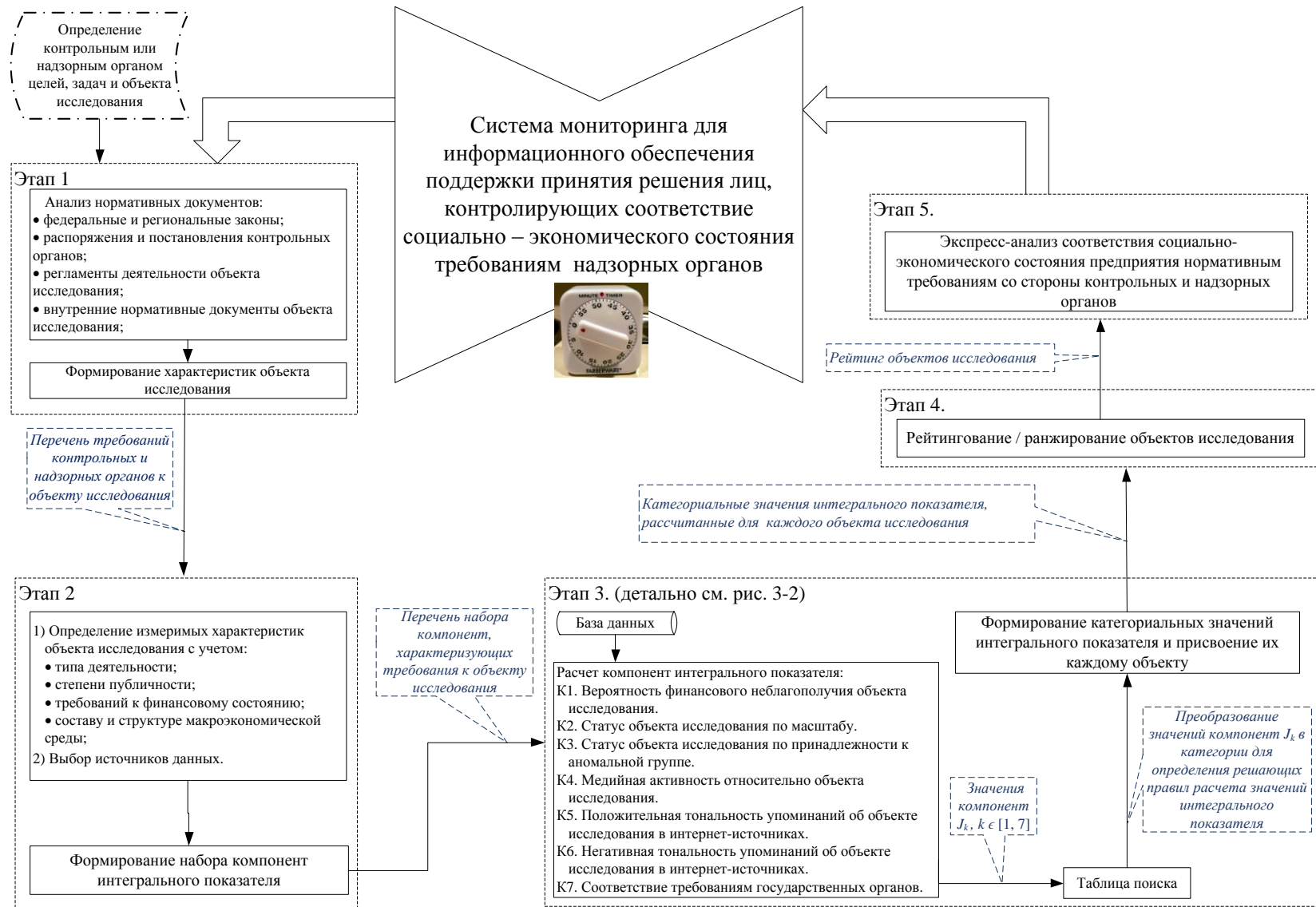


Рисунок 5. Информационно-логическая модель алгоритма расчета компонент интегрального показателя.

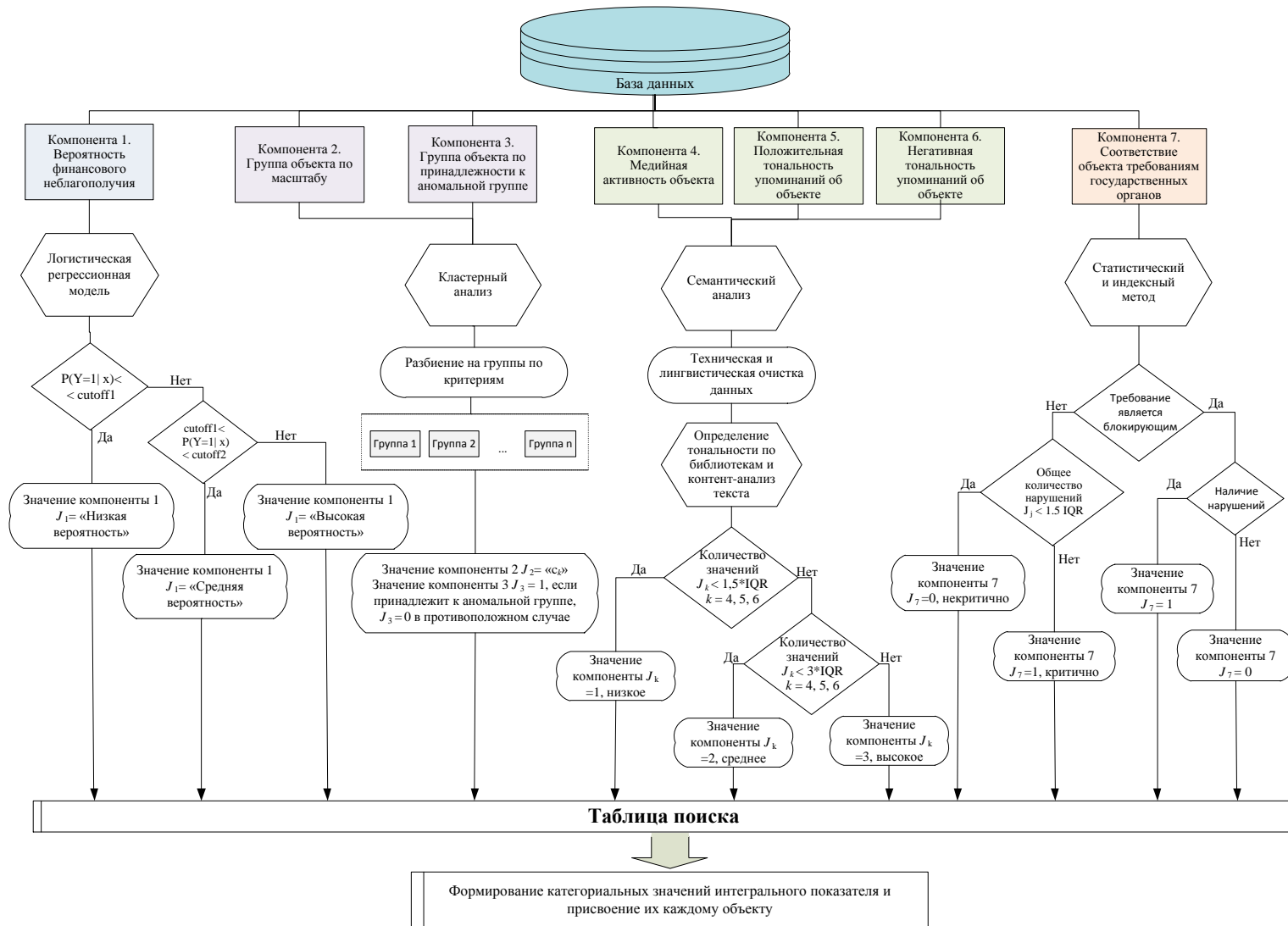


Рисунок 6. Информационно-логическая модель алгоритма расчета компонент интегрального показателя (этап 3).

Глава 3. Экспресс-анализ социально-экономического состояния предприятий из большой совокупности на соответствие нормативным требованиям

3.1 Экспресс-анализ соответствия социально-экономического состояния промышленного предприятия, являющегося элементом большой совокупности предприятий, нормативным требованиям департамента инвестиционной политики г. Москвы.

Одна из задач при принятии решения о стратегии развития предприятия, стоящая перед его топ-менеджментом, понять, насколько текущее социально-экономическое состояние предприятия соответствует требованиям, предъявляемым к нему надзорными и контрольными органами, с возможностью отслеживать это соответствие в текущий момент времени до выхода официальных статистических данных по общедоступным источникам информации. При этом также может быть решена задача дополнения внутренних данных внешними, описывающими качественные характеристики организации (ее медийный облик, репутацию как работодателя и т.п.).

С другой стороны, государственные органы власти субъектов РФ в своей деятельности занимаются, в том числе, поддержкой и развитием промышленности в субъекте РФ, они осуществляют контроль состояния отрасли, мониторинг и предотвращение кризисных явлений или срывов в работе ключевых для региона организаций. Таким образом, предприятие выступает в качестве объекта управления, надзора и контроля со стороны государственных органов власти, на которое органы власти субъекта могут оказывать воздействие с помощью налоговых льгот и субсидий.

Задача практического исследования – апробировать предлагаемую в диссертационной работе концептуальную модель экспресс-анализа соответствия социально-экономического состояния объекта исследования нормативам и регламентам со стороны контрольных и надзорных органов для

оценки соответствия социально – экономического состояния промышленного предприятия г. Москвы требованиям департамента инвестиционной политики г. Москвы о целесообразности финансовой поддержки со стороны государственных органов власти. Данное исследование направлено на создание и тестирование инструментального решения для лиц, принимающих решение в органах контроля и надзора. Основная задача такого инструмента – расширить методологию экспресс-анализа в отсутствие официальной статистической отчетности. Под объектами исследования понимаются предприятия из большой совокупности промышленных предприятий города Москвы и их деятельность, являющиеся как получателями, так и не получателями мер поддержки и стимулирования деятельности в сфере промышленности города Москвы.

3.1.1 Общая поэтапная схема построения интегрального показателя экспресс-анализа соответствия социально-экономического состояния промышленного предприятия заявленным требованиям

Адаптированный под поставленную задачу подход состоит из пяти этапов:

1 этап. Анализ требований законодательства, нормативных документов, прочих требований регуляторов.

В результате анализа постановления Правительства Москвы, федеральных, региональных и иных нормативных актов и распоряжений, формируются критерии для анализа состояния предприятия.

Нормативные требования к промышленному предприятию были составлены на основе следующих документов:

– Постановление правительства г. Москвы от 6 августа 2019 г. n 982-пп об утверждении порядков предоставления субсидий из бюджета города Москвы субъектам малого и среднего предпринимательства, осуществляющим предпринимательскую деятельность в рамках лицензионных договоров

(сублицензионных договоров), договоров коммерческой концессии (субконцессии) (в ред. от 30.12.2020 n 2405-пп) [1].

– Федеральный закон от 24.07.2007 N 209-ФЗ (ред. от 27.12.2018) «О развитии малого и среднего предпринимательства в Российской Федерации» [7].

– п.12. Плана работы Контрольно-счетной палаты Москвы на 2019 год;

– Постановление Правительства РФ от 04.04.2016 N 265 «О предельных значениях дохода, полученного от осуществления предпринимательской деятельности, для каждой категории субъектов малого и среднего предпринимательства» [2];

– Приказ Минпромторга России от 13.04.2015 N 798 (ред. от 30.07.2018) «Об утверждении Перечня организаций, оказывающих существенное влияние на отрасли промышленности и торговли» [3];

Выполнен содержательный анализ правовой, нормативной, отчетной и иной документации, доступной для автора из открытых источников. Это позволило оценить требования к промышленному предприятию и сформировать характеристики, формализующие эти требования со стороны контрольных и надзорных органов. По результатам проведения анализа был разработан перечень требований к предприятию, а также состав первичных показателей, необходимых для оценки эффективности деятельности потенциальных получателей субсидий и налоговых льгот.

Выбраны следующие характеристики деятельности промышленного предприятия – потенциального получателя финансовой поддержки:

- **финансово-хозяйственная деятельность объектов наблюдения.** Исследование основано на данных из бухгалтерской отчетности; о банкротстве; о деятельности предприятия в сфере государственных закупок;

- **надежность и репутация компании.** Исследование основано на данных о вакансиях и их характеристиках; отзывов о предприятиях как работодателях;

- **конкурентная среда и размер предприятия.** Исследование основано на данных о подобных предприятиях по финансовому положению;
- **инфопространство.** Социальный образ предприятия в электронных СМИ. Исследование основано на информационных сообщениях о предприятии в ленте новостей тематических мероприятий – выставок, объединений.

2 этап. На основе методов анализа информации о промышленных предприятиях с учетом целевых показателей определяется список характеристик и компонент (набора аргументов функции) интегрального показателя.

По результатам анализа данных из внутренних и внешних источников разработаны требования к сбору данных, сформирован перечень показателей, необходимый для решения задачи и определены источники информации (см табл. 6).

Таблица 6.

Источники информации для базы исследования

№	Данные	Источники
1.	Данные о финансово-экономических показателях бизнеса: виды деятельности, результаты деятельности; даты начала и окончания деятельности; текущий статус; данные из бухгалтерской (финансовой) отчетности; режим налогообложения; участие в торгах на порталах государственных закупок	Сайты объектов наблюдения; реестр сведений о фактах деятельности юридических лиц; федеральный реестр сведений о банкротстве; сервис проверки контрагентов, справочник юридических лиц и индивидуальных предпринимателей;
2.	Данные о конкурентной среде: характеристики технопарков; отраслевое направление деятельности объекта наблюдения; предприятия и их характеристики по месту нахождения объекта наблюдения	Сайты технопарков и других площадок – агрегаторов информации; справочники с наличием адресов, телефонов, часов работы предприятий и пр.
3.	Данные о рынке труда: вакансии, размещенные предприятием; отзывы об предприятии как работодателя; текучесть кадров, популярность работодателя, неблагонадежность работодателя согласно отзывам	Сервисы поиска работы и персонала. База резюме и вакансий

№	Данные	Источники
4.	Данные о публичной активности объектов наблюдения в г. Москве: упоминания наименований предприятий (в разрезе дат, источников, ключевых слов) и событий (форум, выставка, покупка оборудования, лизинг оборудования и т.п.) с упоминанием наименования предприятия в новостях; тематические тренды упоминаний;	Интернет-СМИ, социальные сети
5.	Статистические данные	Федеральная служба государственной статистики
6.	Данные из нормативно-правовых документов	Сервисы нормативно-правовой поддержки

3 этап. Построение интегрального показателя экспресс-анализа.

На основе выбранных характеристик объектов исследования выбраны 5 компонент интегрального показателя (табл. 7).

Таблица 7.

Компоненты интегрального показателя для промышленного предприятия

№	Компоненты	Промышленное предприятие
1	Вероятность финансового неблагополучия объекта исследования	Вероятность наступления неплатежеспособности предприятия ³
2	Статус объекта исследования по масштабу	Группа предприятий «крупное», «среднее», «малое»
3	Статус объекта исследования по принадлежности к аномальной группе	Подгруппа предприятий по уровню финансового состояния
4	Медийная активность относительно объекта исследования	Наличие упоминаний названия предприятия в новостной ленте промышленных выставок
5	Положительная тональность упоминаний об объекте исследования в интернет-источниках	-
6	Негативная тональность упоминаний об объекте исследования в интернет-источниках	-
7	Соответствие требованиям государственных органов	Совокупная характеристика участия на портале госзакупок

³ Богданова Т.К., Алексеева Ю.А. Прогнозирование вероятности банкротства предприятий с учетом изменения финансовых показателей в динамике. Бизнес-информатика, no. 1 (15), 2011, pp. 50-60.

Выбранные компоненты оцениваются с помощью математических, статистических и эконометрических методов на основе подхода и методов, изложенных в главе 2 диссертационного исследования. Затем рассчитывается интегральный показатель оценки соответствия социально-экономического состояния промышленного предприятия нормативам и регламентам контрольных органов г. Москвы с помощью составленной таблицы ассоциаций. Он принимает 3 значения:

- рекомендована финансовая поддержка на федеральном уровне (2),
- рекомендована финансовая поддержка на региональном уровне (1),
- не рекомендована финансовая поддержка (0).

Рассчитанные значения интегрального показателя носят рекомендательный характер.

4 этап. Рейтингование / ранжирование объектов исследования на основе полученных значений интегрального показателя.

На этом этапе происходит группировка промышленных предприятий на категории по целесообразности получения финансовой помощи.

5 этап. Оценка соответствия объекта исследования заявленным требованиям на основе рейтинга / ранга. Осуществляется переоценка имеющейся информации, обновление данных и мониторинг состояния предприятий повторным выполнением этапов 2–4.

3.1.2 Информационная база исследования

Предлагаемая в диссертационной работе концептуальная модель экспресс-анализа была апробирована для проведения экспресс-анализа соответствия социально-экономического состояния промышленного предприятия требованиям контрольного органа (Департамент инвестиционной и промышленной политики г. Москвы), база данных состояла из информации о 506 действующих промышленных предприятиях, зарегистрированных в г. Москве. В качестве объясняющих переменных для

оценки компонент интегрального показателя рассматривались различные финансовые и нефинансовые данные за 2016 - 2018 гг. по формам №1 и №2 бухгалтерской отчетности (бухгалтерский баланс, отчет о финансовых результатах), данным о контрагентах с портала госзакупок. Для каждого статичного показателя рассчитывалось его изменение по отношению к предыдущему периоду. Дополнительно использовались неструктурированные данные - текстовые: отзывы о предприятиях как работодателях, отзывы о предприятии в электронных СМИ, т.е. качественные характеристики упоминаний названия предприятия в новостных лентах и сайтах выставок.

Формирование перечня показателей, характеризующих эффективность деятельности потенциальных получателей субсидий и налоговых льгот основано на результатах анализа предметной области в части решения задачи.

Результаты формирования перечня показателей используются в построении интегрального показателя для оценки эффективности деятельности получателей субсидий и налоговых льгот по сравнению с остальными предприятиями промышленности по двум группам критериев:

- Экономические критерии;
- Критерии публичной активности.

Группа экономических критериев включает финансовые и экономические показатели бизнеса и показатели конкурентной среды промышленных организаций.

Группа критериев публичной активности включает показатели рынка труда, показатели публичной активности промышленных организаций в г. Москве и показатели, характеризующие общественное мнение по теме субсидий и теме налоговых льгот в г. Москве.

Перечень показателей информационной базы исследования промышленных предприятий, представлен в таблице 8.

Таблица 8.

Перечень показателей исследования

I. ЭКОНОМИЧЕСКИЕ КРИТЕРИИ			
1.	Финансовые и экономические показатели бизнеса		
	1	Наименование объекта наблюдения	
	2	Сайт объекта наблюдения	
	3	Основной вид деятельности объекта наблюдения	
	4	Текущий статус объекта наблюдения (значения: действующее, в процессе ликвидации, ликвидировано)	
	5	Год регистрации объекта наблюдения	
	6	Размер уставного капитала объекта наблюдения по данным ЕГРЮЛ	
	7	Среднесписочная численность работников у объекта наблюдения	
	8	Выручка объекта наблюдения за период 2016 - 2017 гг.	
	9	Прибыль (убыток) от продаж объекта наблюдения за период 2016 - 2017 гг.	
	10	Стоимость капитала объекта наблюдения за период 2016 - 2017 гг.	
	11	Валюта баланса объекта наблюдения за период 2016 - 2017 гг.	
	12	Величина результатов исследований и разработок, входящих в состав Внеоборотных активов объекта наблюдения за период 2016 - 2017 гг.	
	13	Величина основных средств объекта наблюдения за период 2016 - 2017 гг.	
	14	Величина запасов объекта наблюдения за период 2016 - 2017 гг.	
	15	Величина дебиторской задолженности объекта наблюдения за период 2016 - 2017 гг.	
	16	Величина нераспределенной прибыли (непокрытого убытка) объекта наблюдения за период 2016 - 2017 гг.	
	17	Величина краткосрочных обязательств по балансу у объекта наблюдения за период 2016 - 2017 гг.	
	18	Расчетный показатель: величина собственных оборотных средств (функционирующего капитал) в динамике	
	19	Расчетный показатель: фондоотдача в динамике	
	20	Расчетный показатель: ресурсоотдача в динамике	
	21	Расчетный показатель: рентабельность продаж в динамике	
	22	Расчетный показатель: валовая прибыль в динамике	
	23	Расчетный показатель: валовая рентабельность реализованной продукции в динамике	
	24	Год начала участия объекта наблюдения на госзакупках	
	25	Роль объекта наблюдения в госзакупках (значения: поставщик, заказчик)	
	26	Количество контрактов, заключенных объектом наблюдения за период 2016 - 2017 гг.	
	27	Сумма контрактов, заключенных объектом наблюдения за период 2016 - 2017 гг.	
2.	Конкурентная среда объектов наблюдения		
	28	Виды деятельности объекта наблюдения, указанный в ссылке на него	

		29	Количество организаций с аналогичным объекту наблюдения видом деятельности по отраслям и местоположению
		30	Количество организаций с аналогичным объекту наблюдения видом деятельности по отраслям и местоположению
II. КРИТЕРИИ ПУБЛИЧНОЙ АКТИВНОСТИ			
3	Рынок труда		
		31	Список мест размещений вакансий объектом наблюдения
		32	Общее количество вакансий, размещенных объектом наблюдения
		33	Количество отзывов по объекту наблюдения как о работодателе
		34	Популярность работодателя (значения: от 0 до 10)
		35	Неблагонадежность работодателя согласно отзывам
4	Публичная активность объектов наблюдения в г. Москве		
		36	Количество упоминаний наименований объектов наблюдения (в разрезе дат, источников, ключевых слов) в новостях
		37	Упоминание события (форум, выставка, покупка оборудования, лизинг оборудования и т.п.) в новостной ленте технопарка с упоминанием наименования резидента - объекта наблюдения
		38	Дата события из новостной ленты технопарка с упоминанием объекта наблюдения
		39	Количество упоминаний объектов наблюдений в новостных лентах с событиями (форумами, выставками) в г. Москве; (в разрезе дат, источников, ключевых слов)

На основе собранных данных рассчитаны производные показатели финансового анализа, полный перечень которых представлен в приложении 2. Рассчитаны абсолютные приросты, темпы роста, индексы финансового анализа за 2016, 2017 и 2018 гг.

3.1.3 Оценка компонент интегрального показателя

Компонента 1. Вероятность финансового неблагополучия объекта исследования - **Вероятность наступления неплатежеспособности предприятия.**

Для оценки вероятности наступления неплатежеспособности предприятия использована модель, предложенная в своей работе авторами Богдановой Т.К., Алексеевой Ю.А. [33].

Согласно предлагаемой модели, вероятность наступления неплатежеспособности предприятия может быть оценена с помощью уравнения логистической регрессии [33]:

$$P_j = \frac{1}{1 + e^{-(32,633 - 1,082 \cdot X_{1,j} - 6,932 \cdot X_{2,j} + 3,697 \cdot X_{3,j} - 5,712 \cdot X_{4,j} - 1,573 \cdot X_{5,j})}} \quad (13)$$

где P_j – вероятность банкротства предприятия j ;

$X_{1,j}$ – общая оборачиваемость активов предприятия j ;

$X_{2,j}$ – отношение чистой прибыли к активам предприятия j ;

$X_{3,j}$ – отношение краткосрочных и долгосрочных заемных средств к активам предприятия j ;

$X_{4,j}$ – отношение долгосрочных обязательств к активам предприятия j ;

$X_{5,j}$ – натуральный логарифм выручки предприятия j ;

Рассчитанные по данной модели значения вероятности банкротства, агрегированные по интервалам, усредненные за 2016 – 2017 гг, представлены на рис. 7.

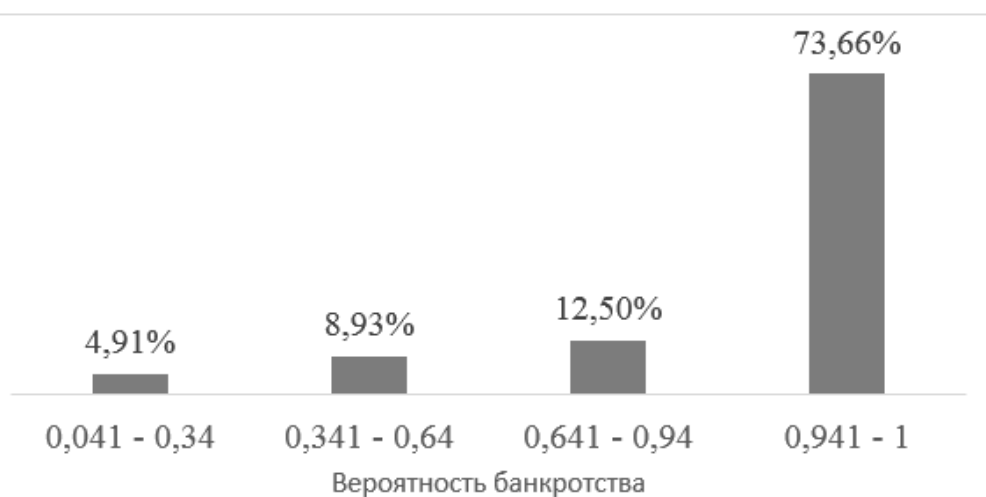


Рисунок 7. Распределение предприятий по вероятности банкротства

Так как среди всех рассматриваемых промышленных предприятий большая доля – это малые или микро предприятия, у которых долгосрочные обязательства или отсутствуют, или не превышают 10% от баланса.

Для расчета интегрального показателя вероятность банкротства была преобразована в категориальную бинарную переменную степени банкротства. Для характеристики состояния промышленных предприятий - устойчивости финансового положения - был выбран показатель изменения вероятности банкротства за год. В случае если вероятность банкротства снижается (прирост вероятности отрицательный), то имеет место улучшение общего финансового состояния предприятия.

Точкой разбиения (порогом отсечения) изменения вероятности банкротства за 2016-2017 гг. выбрано значение $-0,02$, отделяющее 12% верхних значений выборки от всех остальных (рис. 8). Это предприятия с сильно уменьшившейся за предыдущий год вероятностью банкротства относительно остальных предприятий.

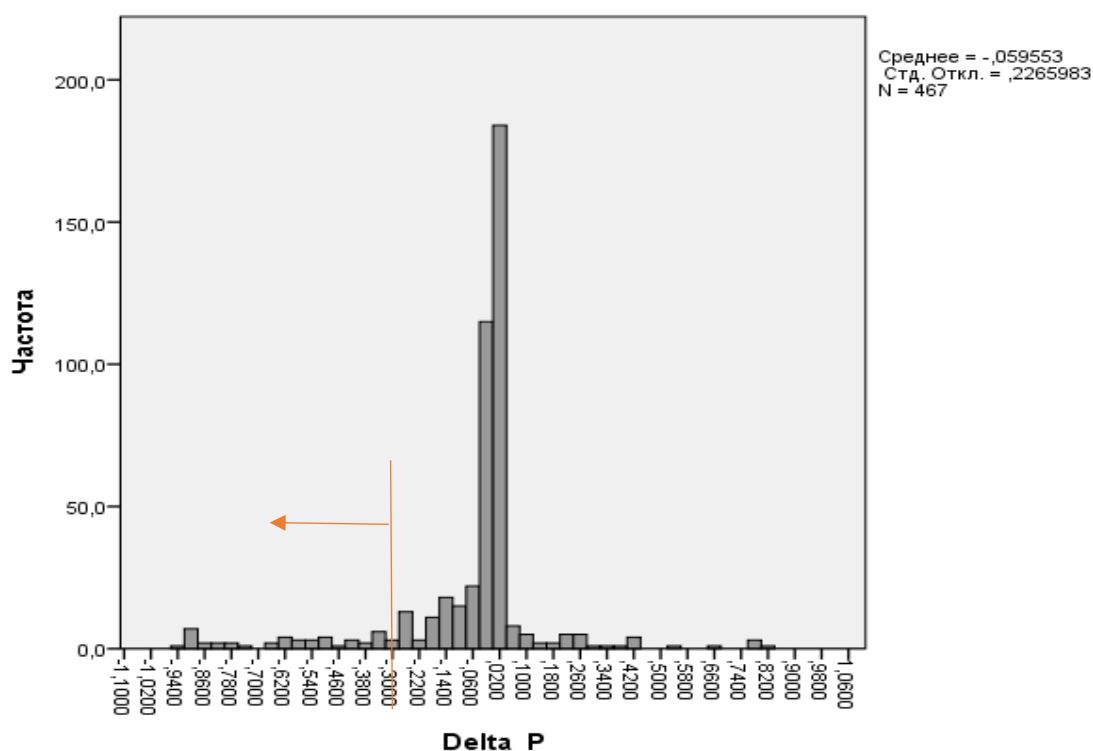


Рисунок 8. Распределение прироста вероятности банкротства промышленных предприятий 2017/2016г

Компонента 1 равна 1, то есть вероятность финансового неблагополучия низкая, по данным за 2016 – 2018 гг. для 27,6% промышленных предприятий.

Компонента 2. Статус объекта исследования по масштабу - **группа предприятий «крупное», «среднее», «малое».**

Промышленные предприятия г. Москвы, как видно из приложения 3, сильно различаются по масштабу деятельности. Для выявления однородных групп была проведена двухступенчатая кластеризация и группировка всех предприятий.

На первом этапе определялся статус предприятия методами статистических группировок выделены три группы предприятий по масштабу: 1-я группа **«Доминирующие компании», состоящая из 35 компаний** (приложение 4).

Группа доминирующих компаний получена путем отбора объектов наблюдения на основании прямых и косвенных критериев. Для этой группы «Доминирующие компании» невозможно с помощью экспресс – анализа оценить эффективность предоставления субсидий и налоговых льгот на уровне субъекта Федерации в связи с их масштабом в экономике региона.

Прямыми критериями отбора являются показатели выручки, прибыли (убытка) до налогообложения и уставного капитала по данным бухгалтерской (финансовой) отчетности за 2016 и за 2017 г.

Косвенными критериями отбора являются показатели влияния на отрасль согласно приказу Минпромторга России от 13.04.2015 № 798 (в ред. от 30.07.2018) «Об утверждении Перечня организаций, оказывающих существенное влияние на отрасли промышленности и торговли» и показатель максимальной суммы контакта согласно Федеральному закону от 05.04.2013 № 44-ФЗ (ред. от 01.05.2019) «О контрактной системе в сфере закупок товаров, работ, услуг для обеспечения государственных и муниципальных нужд» (с изм. и доп., вступ. в силу с 12.05.2019) [3].

2-я группа **«Предприятия среднего и малого бизнеса»**, включает следующие подгруппы. Группа формировалась по остаточному принципу из предприятий, не вошедших в первую и третью группу.

3-я группа «Микропредприятия».

Методика отбора в третью группу «Микропредприятия» заключалась в отборе объектов наблюдения:

- включенных в Единый реестр субъектов малого и среднего предпринимательства, согласно № 209-ФЗ от 24.07.2007 (ред. от 27.12.2018) «О развитии малого и среднего предпринимательства в Российской Федерации».

- у которых значение показателя выручки за период 2016–2017 гг. составило менее 10 млн руб. в год.

В приложении 5 и 6 приведены результаты сравнения распределения показателей между группами 2 и 3 по средним и медианным значениям, а также с помощью непараметрического критерия Манна – Уитни, подтверждающее значимость различий между группами, визуальные различия заметны при сравнении среднегодовых значений по подгруппам (рис. 9).



Рисунок 9. Сравнение средних значений по группе 2 – «среднее» и группе 3 – «малое» предприятие

Исследование показало обоснованность деления предприятий по масштабу, а компонента 2 может принимать одно из значений: «крупное», «среднее», «малое».

Компонента 3. Статус объекта исследования по принадлежности к аномальной группе - **подгруппа предприятий по уровню финансового состояния.**

Для второй и третьей группы предприятий выделены однородные по набору показателей, характеризующих финансовое состояние предприятия, подгруппы. Для выявления однородных подгрупп с неизвестным числом кластеров выбран метод BIRCH (двухшаговая кластеризация).

В качестве входных параметров кластеризации использовались следующие показатели по 471 предприятию:

- 1) Среднесписочная численность работников у объекта наблюдения
- 2) Выручка объекта наблюдения за период 2016 - 2017 гг.
- 3) Стоимость капитала объекта наблюдения за период 2016 - 2017 гг.
- 4) Валюта баланса объекта наблюдения за период 2016 - 2017 гг.
- 5) Величина дебиторской задолженности объекта наблюдения за период 2016 - 2017 гг.
- 6) Величина нераспределенной прибыли (непокрытого убытка) объекта наблюдения за период 2016 - 2017 гг.
- 7) Величина краткосрочных обязательств по балансу за период 2016 – 2017 гг.
- 8) Сумма контрактов, заключенных объектом наблюдения за период 2016 - 2017 гг.
- 9) Расчетный показатель: прирост валовой прибыли за 2016-2017 гг.
- 10) Расчетный показатель: валовая рентабельность реализованной продукции в динамике

Кластеризация выполнена с помощью программного продукта IBM SPSS Statistics. Результаты полученной кластеризации оценивались с помощью сравнения показателей по подгруппам:

- 1) показатели динамики (цепные и базисные) финансовых показателей внутри полученных групп,

- 2) показатели средних и медианных значений баланса,
- 3) показатели вариации основных показателей финансовой деятельности по годам.

В результате было получено, что наилучшее разделение - на 2 кластера по признаку финансового состояния – условные «лидеры», с хорошей динамикой финансовых показателей, и «аутсайдеры», имеющие некоторые негативные тенденции в ключевых показателях.

Промышленные предприятия, отнесенные к кластеру «Лидеры» отличаются большей долей основных средств и более быстрым оборотом дебиторской задолженности при взаимодействии с контрагентами, высокими значениями выручки (и прироста выручки за 2016-2017 гг.). Их финансовые показатели также превышают показатели «Аутсайдеров» (табл.9):

Таблица 9.

Сравнение финансовых показателей из группы «Лидеров» и «Аутсайдеров».

Наименование показателя	Группа 2. "Среднее"				Группа 3. "Малое"			
	"Лидеры"		"Аутсайдеры"		"Лидеры"		"Аутсайдеры"	
	Среднее значение	Медиана	Среднее значение	Медиана	Среднее значение	Медиана	Среднее значение	Медиана
Среднесписочная численность работников, чел.	116	59	80	29	9	7	8	3
Выручка, млн. р., 2016 г.	1 072	252	229	30	35	11	10	0
Выручка, млн. р., 2017 г.	1 134	252	249	38	48	15	17	1
Стоимость капитала, млн.р., 2016 г.	773	99	197	16	13	2	10	0
Стоимость капитала, млн.р., 2017 г.	934	145	324	18	18	4	12	0
Валюта баланса, тыс.р., 2016 г.	2 914 744	393 203	477 746	54 785	64 971	8 120	22 459	1 120
Валюта баланса, тыс.р., 2017 г.	3 044 491	476 129	377 366	71 922	101 818	17 805	37 288	4 673

Распределение количества предприятий по кластерам представлено на рисунке 10.

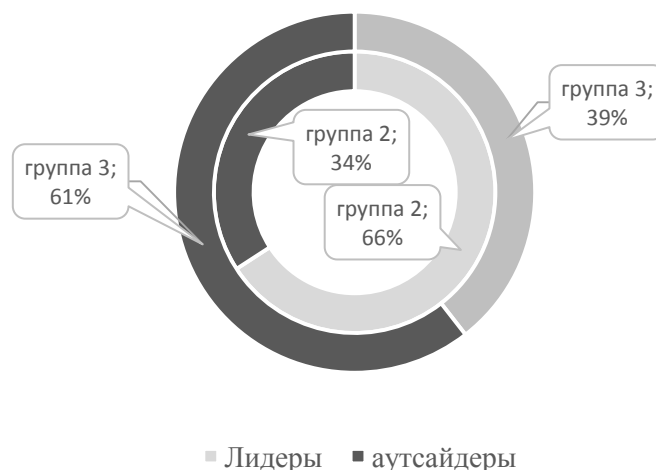


Рисунок 10. Распределение предприятий по группам внутри кластеров

Для подтверждения значимости результатов разбиения проверены статистические гипотезы о совпадении распределения показателей по группе получателей и не получателей финансовой помощи. Отвержение гипотезы о совпадении означает неравенство распределений и значимые различия по показателям (приложение 6 и 7).

Для группы 2 по признаку получения финансовой помощи зафиксирована статистически значимая разница по 33 показателям из 54, для группы 3 – 25 из 51 показателей.

Таким образом, **компонента 3** может принимать четыре значения:

Значения 2-1 и 2-2. Вторая группа предприятий – «Предприятия среднего и малого бизнеса» разделилась на два подкласса: «лидеры» и «аутсайдеры».

Значения 3-1 и 3-2. Третья группа – «Микропредприятия» также разделилась на два подкласса: «лидеры» и «аутсайдеры».

Компонента 4. Медийная активность объекта исследования - наличие упоминаний названия предприятия в новостной ленте промышленных выставок.

Эта компонента представляет собой оценку наличия упоминаний в новостных лентах:

- выставок (Экспо 2016, Экспо 2017, Экспо 2018, Технофорум 2016, Технофорум 2017, Текстильлегпром 2017);
- площадок промышленных предприятий г. Москвы (ОЭЗ Технополис «Москва», Технопарки г. Москвы);

Тематика новостных лент с упоминанием наименований предприятий – это успехи, достижения компании. В лентах упоминается о номинантах премий, о лучших достижениях, разработках, их наличие положительно может сказаться на финансовой стабильности предприятия в будущем. Поэтому наличие таких упоминаний положительно сказывается на имидже предприятия, и, как следствие, на его финансовом состоянии.

Примером такого упоминания являются:

«Компания-разработчик «Бутис» инновационных изделий для радиоаппаратуры предложила новый продукт – термокомпенсированные фильтры для навигационных моделей». Источник: сайт <https://technomoscow.ru/>, раздел: Новости.

Компонента 4 принимает два значения - 1 – есть положительные упоминания и 0 – нет положительных упоминаний.

Компонента 5. Соответствие требованиям государственных органов - Совокупная характеристика участия на портале госзакупок.

Для регистрации на портале государственных закупок компания должна подать документы, подтвердить свою деятельность и состоятельность в выполнении контрактов, что может служить критерием выполнения определенных требований.

Статистический анализ разницы среди подгрупп «лидеров» и «аутсайдеров» группы 2 и группы 3 промышленных показателей подтвердил значимость участия на портале государственных закупок (табл. 10).

Таблица 10.

Показатели по подгруппам промышленных предприятий

Наименование показателя	Группа предприятий			
	Группа 2		Группа 3	
	Лидеры	Аутсайдеры	Лидеры	Аутсайдеры
Количество предприятий	160	83	90	138
Вероятность банкротства, среднее	0,7	0,81	0,7	0,98
Зарегистрирован на портале, кол-во	153	25	82	2
Среднее количество контрактов	50	38	4	1
Средняя сумма контрактов на 1 предприятие за 2016-2017 гг, тыс.р.	24457	837	16710	652

Сравнительный анализ подтверждает значимые различия относительно участия на портале госзакупок. Компонента 5 принимает бинарные значения: **1** – есть статус участника на портале государственных закупок; **0** – нет статуса участника на портале государственных закупок;

3.1.4 Расчет интегрального показателя экспресс- анализа

Интегральный показатель оценки соответствия социально-экономического состояния промышленного предприятия нормативам и регламентам контрольных органов г. Москвы принимает 3 значения – рекомендована финансовая поддержка на федеральном уровне (2), рекомендована финансовая поддержка на региональном уровне (1) и не рекомендована финансовая поддержка (0).

Для получения этих значений интегрального показателя из разнородных пяти оцененных компонент с учетом выбранных порогов отсека была составлена таблица поиска (табл. 11).

Таблица 11.

Таблица поиска для расчета значений интегрального показателя оценки
состояния промышленного предприятия

№	Компонента 1. Вероятность ¹	Компонента 2. Статус группы ²	Компонента 3. Статус аномальности ³	Компонента 4. Медийная ⁴	Компонента 7. Регламенты ⁵	Значение интегрального показателя ⁶
1	1, 100% выборки	крупное	не оценивался	1	1, 100% выборки	2
2	1, 100% выборки	крупное	не оценивался	0	1, 100% выборки	2
3	1	средний	2-1	1	1	1
4	1	средний	2-2	1	1	1
5	1	средний	2-1	0	1	1
6	1	средний	2-1	1	0	1
7	1	средний	2-1	0	0	0
8	0	средний	2-1	1	1	1
9	0	средний	2-2	1	1	0
10	1	микро	2-1	1	1	1
11	1	микро	3-2	1	1	1
12	1	микро	3-1	0	1	1
13	1	микро	3-1	1	0	0
14	0	микро	3-1	1	1	0
15	0	микро	3-2	1	1	0
16	0	средний	2-1	1	0	0
17	0	средний	2-1	0	1	0
18	0	средний	2-1	0	0	0
19	1	средний	2-2	1	0	0
20	0	средний	2-2	1	0	0
21	1	средний	2-2	0	1	0
22	0	средний	2-2	0	1	0
23	1	средний	2-2	0	0	0
24	0	средний	2-2	0	0	0
25	1	микро	3-1	0	0	0
26	0	микро	3-1	0	1	0
27	0	микро	3-1	1	0	0
28	0	микро	3-1	0	0	0
29	0	микро	3-2	1	0	0
30	0	микро	3-2	0	0	0
31	1	микро	3-2	1	0	0
32	1	микро	3-2	0	0	0

1: 1, если изменение вероятности неплатежеспособности за 2017/2016 гг меньше (-0.02);

0, если изменение вероятности неплатежеспособности за 2017/2016 гг больше или равно (-0.02).

2: кластер 1 - «крупное», кластер 2 - «среднее», кластер 3 - «микро».

3: 2-1, 3-1 – подкластер «лидеры» в кластере «среднее» или «микро» предприятие соответственно;

2-2, 3-2 – подкластер «аутсайдеры» в кластере «среднее» или «микро» предприятие соответственно;

4: 1 – есть положительные упоминания;

0 – нет положительных упоминаний;

5: 1 – есть статус участника на портале государственных закупок;

0 – нет статуса участника на портале государственных закупок;

6: 0 - не рекомендована финансовая поддержка;

1 – рекомендована финансовая поддержка на региональном уровне;

2 – рекомендована финансовая поддержка на федеральном уровне;

Рассчитанные значения интегрального показателя носят рекомендательный характер.

Результаты классификации предприятий по типу финансовой помощи на основании расчета значений интегрального показателя для промышленных предприятий, представлен на рисунке 11.

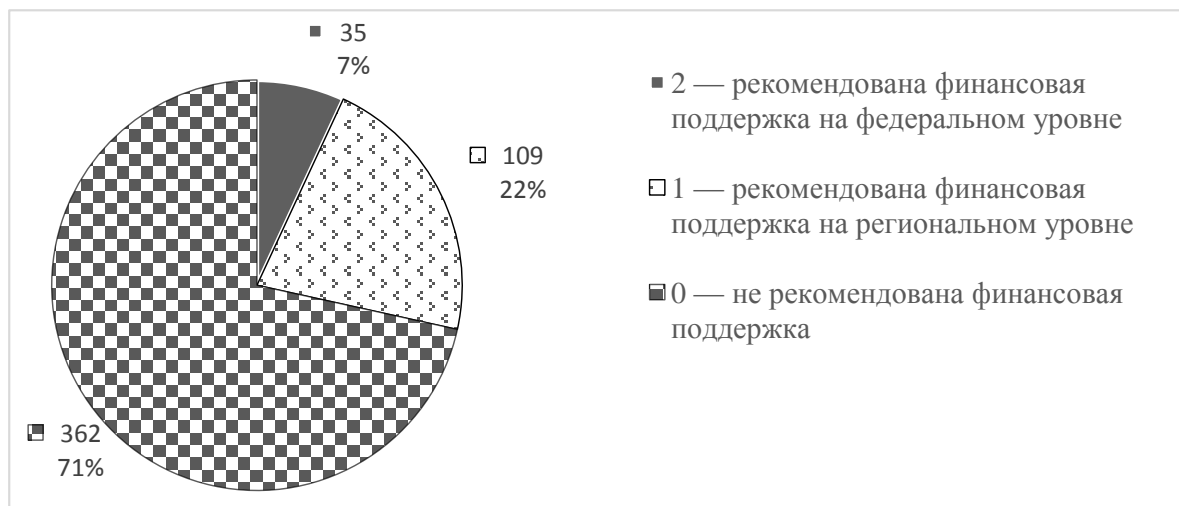


Рисунок 11. Результаты расчетов интегрального показателя для промышленных предприятий

3.1.5 Анализ полученных результатов

Сверка полученных значений для промышленных предприятий проводилась путем сопоставления текущего состояния объекта исследования

и динамики его основных финансовых показателей, построение тестовой выборки невозможно в силу невозвратности финансовой помощи.

По результатам экспресс-оценки соответствия социально-экономического состояния промышленного предприятия требованиям, сформулированным в нормативных документах и регламентах департамента инвестиционной политики г Москвы, были составлены профили предприятий, соответствующие значению 2 интегрального показателя - рекомендована финансовая поддержка на региональном уровне.

Профили предприятий на основе результатов экспресс-анализа получены значения порогов отсечения, соответствующих данным на 2017 – 2018 г., пример профиля получателей финансовой помощи из регионального бюджета для предприятий среднего и малого бизнеса представлен в табл. 12.

Таблица 12.

Профили промышленных предприятий среднего и малого бизнеса (кластер 2), которым рекомендована финансовая поддержка на региональном уровне

Наименование показателя	Значение показателя за 2016-2018 гг.
Доля кластера в общем объеме по численности предприятий	48% (243 из 506)
Доля компаний, которым рекомендована финансовая поддержка, в общем объеме численности предприятий кластера	28% (68 из 243)
Выручка от реализации, медиана млн. рублей	1056
Прирост валюты баланса, млн. рублей	105
Показатели участия в торгах на портале государственных закупок:	да, участвует
Средняя сумма контракта, млн. рублей	11
Среднее количество контрактов в год	47

Результаты качественного сопоставления значений интегрального показателя для трех промышленных предприятий, полученные на основе анализа значений интегрального показателя, представлены в таблице 13.

Таблица 13.

Результаты расчетов интегрального показателя для промышленных предприятий.

Значение интегрального показателя	Количество предприятий	Выборочное предприятие
2 – рекомендована финансовая поддержка на федеральном уровне	35 (6,9%)	АО ЦНИИ «ЦИКЛОН» - центральный научно-исследовательский институт «Циклон» - единственный сертифицированный в России разработчик и производитель неохлаждаемых тепловизоров и систем технического зрения на их основе. Успешное крупное предприятие, оказывающее влияние на бюджет города.
1 – рекомендована финансовая поддержка на региональном уровне	171 (33,7%)	ООО «СИМБИОТЕХ» - создание качественных и реалистичных симуляторов при РНИМУ им Н.И. Пирогова. Отмечается положительная тенденция в финансовых показателях, имеются признаки добросовестного партнера на госзакупках.
0 - не рекомендована финансовая поддержка	300 (59,4%)	Предприятия, имеющие не благоприятную динамику финансовых показателей относительно своего кластера, статус надежности которых не подтверждается внешними признаками (например, статусом на госзакупках).

Для промышленных предприятий проведен экспресс-анализ соответствия потребности и целесообразности для государственных органов власти в финансовой помощи промышленного предприятия г. Москвы на примере 506 промышленных предприятий, зарегистрированных в г. Москве, по открытым данным за 2016, 2017 и 2018 гг. Результаты экспресс-анализа соответствовали фактическим данным на следующий год о динамике предприятий.

Предложенная концептуальная модель и методика дает возможность получить основу для совершенствования стратегии развития предприятия.

3.2 Экспресс-анализ соответствия социально-экономического состояния коммерческого банка нормативным требованиям ЦБ РФ.

Предложенная концепция построения интегрального показателя для экспресс-анализа социально-экономического состояния предприятия, удовлетворяющего нормативам и регламентам со стороны контрольных и надзорных органов была апробирована для получения экспресс-анализа состояния коммерческого банка. Контрольным органом для решения данной задачи является ЦБ РФ - надзорный орган в банковской сфере.

Построение математической модели экспресс-анализа социально-экономического состояния коммерческого банка, удовлетворяющего нормативам и регламентам со стороны контрольных и надзорных органов, в соответствии с рассмотренной выше концепцией состоит из пяти этапов.

3.2.1 Общая поэтапная схема построения интегрального показателя экспресс-анализа соответствия социально-экономического состояния коммерческого банка заявленным требованиям

Этап 1. Формирование целевых показателей оценки состояния банка на основе требований к показателям ликвидности и надежности банка, полученных из нормативов и законодательных актов ЦБ РФ.

В соответствие с установленными ЦБ РФ нормативами и требованиями к коммерческому банку были сформированы следующие характеристики:

- вероятность финансового банкротства коммерческого банка;
- финансовая стабильность коммерческого банка;
- качественная оценка деятельности коммерческого банка экспертами - участниками рынка.

Этап 2. Определение компонент интегрального показателя экспресс-анализа состояния банка.

На основе формализации требований ЦБ были сформированы следующие характеристики банка:

- устойчивость финансового положения самого банка, его платежеспособность;
- устойчивость банка при изменении внешней среды;
- отсутствие предпосылок к банкротству банка по разным причинам;

Характеристика ЦБ РФ - вероятность финансового банкротства коммерческого банка.

Для определения этой характеристики рассчитывается компонента 1 и 4.

Компонента 1 - прогнозная вероятность наступления банкротства коммерческого банка.

Компонента 4 - соответствие требованиям государственных органов – оценивается выполнение обязательных нормативов ЦБ Н1 – Н7.

Характеристика ЦБ РФ - финансовая стабильность коммерческого банка.

Для определения этой характеристики рассчитывается компонента 2.

Компонента 2 – группы банков по финансовой стабильности, это косвенный признак нестабильности коммерческого банка, полученный на основе оценки принадлежности банка к группе «неблагополучных».

Характеристика ЦБ РФ – добросовестная деятельность банка. Для определения этой характеристики рассчитывается компонента 3.

Компонента 3 – Негативная тональность упоминаний об объекте исследования в интернет-источниках, представляющая собой наличие негативной оценки общего состояния банка со стороны участника рынка.

На основе выбранных характеристик объектов исследования выбраны компоненты интегрального показателя, представленные в табл. 14.

Таблица 14.

Компоненты интегрального показателя для коммерческого банка

№	Компоненты	Банк
1	Вероятность финансового неблагополучия объекта исследования	Вероятность банкротства банка (из-за отзыва лицензии по финансовым причинам)
2	Статус объекта исследования по масштабу	-
3	Статус объекта исследования по принадлежности к аномальной группе	Группы банков по финансовой стабильности
4	Медийная активность относительно объекта исследования	-
5	Положительная тональность упоминаний об объекте исследования в интернет-источниках	-
6	Негативная тональность упоминаний об объекте исследования в интернет-источниках	Наличие негативного упоминания на сайте экспертного агентства expert.ra
7	Соответствие требованиям государственных органов	Выполнение обязательных нормативов ЦБ Н1 – Н7

Этап 3. Построение интегрального показателя для экспресс-анализа социально-экономического состояния коммерческого банка, удовлетворяющего нормативам и регламентам со стороны контрольных и надзорных органов.

Информация была получена из восьми внешних источников: сайт ЦБ РФ; рейтинговые агентства; справочники и информационно-аналитические порталы. Требования к состоянию банка были взяты из нормативов регулятора – ЦБ РФ. Информация была взята из открытых источников (табл. 15).

Таблица 15.

Перечень источников общедоступной информации о коммерческом банке

№	Наименование внешнего источника	Тип источника
1	Сайт ЦБ РФ	официальный сайт ведомства (регулятора)
2	Яндекс	поисковая система
3	Руспрофайл - сервис проверки контрагентов	справочник/информационно-аналитический портал

№	Наименование внешнего источника	Тип источника
4	ООО ИА «Банки.ру»	справочник/информационно-аналитический портал
5	АО «Эксперт РА»	рейтинговое агентство
6	Финансовый онлайн-агрегатор Mainfin.ru	справочник/информационно-аналитический портал
7	Банкирос – финансовый сервис с информацией по банкам и их продуктам	справочник/информационно-аналитический портал
8	BankoDrom.ru	справочник/информационно-аналитический портал

Критерием проверки качества модели служил статус коммерческого банка на момент сбора информации, то есть являлся ли он действующим или ликвидированным.

Для получения значений интегрального показателя были рассчитаны компоненты I_1 , I_2 и I_3 , I_3 . Интегральный показатель принимает три значения:

- 1 – низкое, требуется пристальное внимание к банку,
- 2 – среднее, требуется внимание к банку,
- 3 – высокое, не требуется внимание к банку,

Этап 4 заключается в ранжирование коммерческих банков обучающей выборки на основе рассчитанных значений интегрального показателя. Результатом выполнения этого этапа является список коммерческих банков, отсортированных по убыванию значений интегрального показателя.

Этап 5. Апробация разработанной модели и мониторинг состояния коммерческих банков путем повторного выполнения этапов 3-4.

3.2.2 Информационная база исследования

Базой исследования для 111 коммерческих банков выступили основные показатели, характеризующие ставки банков по вкладам, взятые за период с 01.01.2018 по 01.02.2020 год, показатели рейтинговых агентств - с 2017 по 2020 год. Общие характеристики банка по данным сайта Банка России и других открытых источников - по состоянию на март 2020 года. Информация была собрана из восьми внешних источников: сайт ЦБ РФ; рейтинговые агентства; справочники и информационно-аналитические порталы.

Требования к состоянию банка были взяты из нормативов регулятора – ЦБ РФ. Первичные показатели: данные о различных ставках по депозитам на срок до 1 года; макроэкономические показатели – ключевая ставка, средние ставки по всем банкам РФ и по 30 крупнейшим банкам РФ; показатели, характеризующие состояние банка – его активы, значения нормативов (Н1 – Н7); показатели рейтинговых агентств за период с 2017 по 2020 годы; семантическая (текстовая) информация об упоминании банка – всего более 60 показателей. На их основе были построены расчетные показатели: разброс ставок относительно макроэкономических показателей – медиана и дисперсия разбросов, относительные показатели динамики ставок, их медианы и разброс, относительное и абсолютное место банка в рейтинге банков, негативное упоминание банка в новостной ленте.

На основе выбранных источников была сформирована база данных показателей, характеризующих деятельность коммерческого банка (приложение 8).

- На основе исходных данных были рассчитаны показатели центральной тенденции и разброса: медиана, дисперсия отклонений ставок по депозитным вкладам для каждого коммерческого банка относительно макроэкономических показателей: ключевой ставки, среднебанковской ставки по всем банкам, по 30 крупнейшим банкам России;

- относительные показатели динамики ставок, их медианы, дисперсии;
- относительное и абсолютное место банка в рейтинге банков;
- негативное упоминание банка в новостной ленте на сайте рейтингового агентства.

3.2.3 Оценка компонент интегрального показателя

Компонента 1 (I_1) – прогнозная вероятность наступления банкротства коммерческого банка.

Для получения оценки вероятности отзыва лицензии у коммерческого банка построена логистическая регрессионная модель на основе исходных данных и расчетных показателей. К числу таких показателей относятся: место банка в рейтинге, волатильность банковских ставок (медиана отклонений от ключевой и среднерыночной ставок), качественный расчетный показатель, характеризующий восприятие экспертным сообществом деятельности коммерческого банка, на основе обработки неструктурированных текстовых данных - упоминание о понижении рейтинга коммерческого банка в новостной ленте на сайте рейтингового агентства. Зависимая переменная Y - бинарная переменная, принимающая два значения - 0, если текущий статус банка - действующий, и 1 - если он ликвидирован.

Оцененная логистическая регрессионная модель имеет вид:

$$P(Y=1) = \frac{1}{1+e^{-z}} \quad (2)$$

где $z = -16.7 + 1.4955 * \text{Понижение} + 0,024 * \text{место в рейтинге} + 1,128 * \text{Медиана ставок} + 0,389 * \text{Волатильность ставки RI}$

Все коэффициенты модели значимы на 5% уровне, показатели чувствительности (т.е. доля ликвидированных банков) и специфичности (доля действующих банков) модели высокие, соответственно: 85,9% и 73,3%.

По результатам работы модели с помощью разбиения всех полученных значений на три интервала была сформирована **компонента 1** интегрального показателя, характеризующая вероятность отзыва лицензии в течение полугода после момента оценки модели:

- низкая - $P(Y=1) \leq 0,25$;
- средняя - $0,25 < P(Y=1) < 0,35$;
- высокая - $P(Y=1) \geq 0,35$.

Границы отсека выбирались на основе оптимизации чувствительности и специфичности следующим образом. Верхняя граница, равная 0,35, была получена при одновременном достижении максимальных

значений чувствительности и специфичности модели (85,9% и 73,3%). Нижняя граница, равная 0,25, соответствует значению специфичности, равному 80%.

Компонента 3 (I₂) - оценка принадлежности банка к кластеру «неблагополучных».

При кластеризации обучающей выборки методом BIRCH (двухшаговый метод) по основным финансовым показателям коммерческого банка, по данным о значении основных семи нормативов банков ($H_1 - H_7$), по рейтингу коммерческого банка в рейтинге банков рейтингового агентства и изменению этого рейтинга в ретроспективе, было получено 5 однородных кластеров.

Наиболее интересным оказался 1-ый кластер, составляющий 14% от обучающей выборки. В этом кластере присутствовало 30% ликвидированных банков. У всех банков этого кластера место в рейтинге банков находилось ниже 200. Также было отмечено снижение медианы нормативов каждого коммерческого банка за ретроспективный период, снижение медианы дисперсии ставок по депозитным вкладам за предыдущий период.

По результатам работы модели была сформирована ***компонента 2*** интегрального показателя - принадлежность к первому кластеру, как признак “неблагополучия” коммерческого банка.

Компонента 6 (I₃) – наличие негативной оценки деятельности банка со стороны участника рынка - экспертного агентства.

Эта компонента формируется как бинарный показатель принадлежности к одному из двух множеств. Компонента 3 принимает значение 1, если есть хотя бы одно негативное упоминание наименования коммерческого банка в новостной ленте на сайте рейтингового агентства за текущий год, и значение 0, если нет ни одного негативного упоминания.

Компонента 7 (I₄) – выполнение обязательных нормативов ЦБ Н1 – Н7.

Так как не выполнение обязательных нормативов ЦБ ведет к серьезным последствиям вплоть до приостановки лицензии банка, то эту компоненту можно считать бинарной:

0 – не было нарушений нормативов;

1 – были нарушения нормативов;

3.2.4 Расчет интегрального показателя экспресс- анализа

Из различных компонент интегрального показателя рассчитываются его итоговые значения с помощью таблицы поиска (табл. 16).

Таблица 16.

Таблица поиска для расчета значений интегрального показателя оценки состояния коммерческого банка

Компонента 1. Вероятность ^a	Компонента 3. Группа	Компонента 6. Медийная ^b	Компонента 7. Нормативы ^c	Значение интегрального показателя ^d
> 0,35	Кластер 1	0	0	1
> 0,35	Кластер 1	0	1	1
> 0,35	Кластер 1	1	0	1
> 0,35	Кластер 1	1	1	1
> 0,35	Кластер 2	0	0	2
> 0,35	Кластер 2	0	1	1
> 0,35	Кластер 2	1	0	1
> 0,35	Кластер 2	1	1	1
0,25 – 0,35	Кластер 1	0	0	2
0,25 – 0,35	Кластер 1	0	1	1
0,25 – 0,35	Кластер 1	1	0	1
0,25 – 0,35	Кластер 1	1	1	1
0,25 – 0,35	Кластер 2	0	0	3
0,25 – 0,35	Кластер 2	0	1	2
0,25 – 0,35	Кластер 2	1	0	2
0,25 – 0,35	Кластер 2	1	1	2
< 0.25	Кластер 1	0	0	3
< 0.25	Кластер 1	0	1	3
< 0.25	Кластер 1	1	0	3
< 0.25	Кластер 1	1	1	2
< 0.25	Кластер 2	0	0	3
< 0.25	Кластер 2	0	1	3
< 0.25	Кластер 2	1	0	3
< 0.25	Кластер 2	1	1	2

- a вероятность банкротства: $> 0,35$ – высокая, $0,25 - 0,35$ - средняя, $< 0,25$ – низкая;
 b **0** - не было негативных упоминаний; **1** – были негативные упоминания;
 c **0** – не было нарушений нормативов; **1** – были нарушения нормативов;
 d - **1** – низкое, требуется пристальное внимание к банку,
 2 – среднее, требуется внимание к банку,
 3 – высокое, не требуется внимание к банку,

3.2.5 Анализ полученных результатов

Апробация разработанной модели расчета интегрального показателя была проведена на тестовой выборке из трех коммерческих банков (табл. 17).

Таблица 17.

Значение интегрального показателя для коммерческих банков тестовой выборки

Наименование банка	Компонента 1. Вероятность	Компонента 3. Группа	Компонента 6. Медийная	Компонента 7. Нормативы	Значение (y_i) интегрального показателя
ПАО КБ ПФС-БАНК	Высокая (0,844)	Не принадлежит к кластеру 1	0	1	3 , требуется пристальное внимание к банку
АКБ ПЕРЕСВЕТ (ПАО)	Низкая (0,01)	Принадлежит к кластеру 1	1	0	2 , требуется внимание к банку
АО КБ СИТИБАНК	Низкая (0,007)	Не принадлежит к кластеру 1	0	0	1 , не требуется внимание к банку

Результаты апробации разработанной модели на тестовой выборке из трех коммерческих банков, как видно из таблицы 18, подтверждаются текущим статусом коммерческих банков на март 2020.

Таблица 18.

Сопоставление расчетного значения интегрального показателя и текущего статуса коммерческого банка

Наименование банка	Значение (y_i) интегрального показателя	Текущий статус коммерческого банка. Комментарий.
ПАО КБ ПФС-БАНК	3 , требуется пристальное внимание к банку	Лицензия отозвана в марте 2020г.
АКБ ПЕРЕСВЕТ (ПАО)	2 , требуется внимание к банку	Действующий.

		Наблюдается неоднократное нарушение обязательных нормативов регулятора в процессе санации банка
АО КБ «СИТИБАНК»	1, не требуется внимание к банку	Действующий

Таким образом, предложенная на основе анализа открытых структурированных и неструктурированных данных концепция экспресс-анализа состояния коммерческого банка, позволяет получить вероятность наступления неудовлетворительного состояния банка. Полученные результаты позволяют применить превентивные меры поддержки или защиты интересов вкладчиков со стороны надзорного органа (ЦБ РФ).

Таким образом, представленный подход на основе анализа открытых больших данных позволяет получить вероятность наступления состояния банка, отличного от стабильного, оценки на ближайшие полгода и применить предиктивные меры поддержки или защиты интересов вкладчиков со стороны надзорного органа (ЦБ РФ).

Заключение

Разработана концептуальная модель, схема ее реализации и алгоритм построения интегрального показателя, позволяющие получить экспресс-оценку состояния объекта исследования со стороны контрольных органов с целью предотвращения развития негативных событий.

Для решения задачи разработки инструментального решения проведения экспресс-анализа оценки соответствия социально – экономического состояния предприятия установленным документами регламентам и нормативам был разработан инновационный метод, позволяющий использовать современные информационные технологии работы с общедоступными данными.

В итоге проведения диссертационного исследования были получены следующие результаты:

1. Предложена концептуальная модель экспресс-анализа соответствия социально-экономического состояния предприятия нормативам и регламентам со стороны контрольных и надзорных органов, позволяющая своевременно принимать меры для предотвращения развития негативных событий, и оценена возможность ее информационного обеспечения.
2. Предложена типологизация источников общедоступной информации для формирования базы исследования с гарантированной степенью достоверности общедоступных данных, обеспечивающая возможность формирования компонент интегрального показателя для проведения экспресс-анализа социально-экономического состояния предприятия, являющегося элементом большой совокупности предприятий - объектов нормативного регулирования со стороны общего контрольного или надзорного органа.
3. Предложен набор компонент для расчета интегрального показателя, представляющий собой набор количественных и качественных характеристик объекта исследования, формируемых на основе требований, предъявляемых со стороны контрольных и надзорных органов. Данный подход позволяет

формализовать требования к предприятию со стороны контрольных или надзорных органов, тем самым обеспечивая создание механизма совершенствования их взаимодействия.

4. Разработан алгоритм построения интегрального показателя экспресс-анализа соответствия социально-экономического состояния объекта исследования нормативам и регламентам со стороны контрольных и надзорных органов на основе таблицы поиска, агрегирующей полученные значения набора компонент в конечный набор значений интегрального показателя.

5. Предложен алгоритм расчета пороговых значений для преобразования разнородных оцененных значений каждой компоненты в значения качественной дискретной шкалы, которые с использованием таблицы поиска агрегируются в значения интегрального показателя, измеренного в порядковой шкале с ограниченным количеством категорий. Полученные значения обеспечивают градацию предприятий по степени соответствия их социально-экономического состояния требованиям контрольных и надзорных органов, являясь инструментом мониторинга и раннего предупреждения негативных последствий развития событий.

6. Продемонстрировано действие предложенного гибкого инструментария на основе разработанной информационно-логической модели экспресс-анализа соответствия социально-экономического состояния промышленного предприятия нормативам и регламентам контрольных и надзорных органов г. Москвы для 506 промышленных предприятий. Даны рекомендации для предоставления финансовой поддержке со стороны федеральных и региональных органов власти. Получены профили промышленных предприятий по степени целесообразности предоставления финансовой поддержки. Определены пороговые значения для отклонения или поддержки выделения финансовой помощи каждому предприятию.

7. Обоснованы рекомендации для усиления надзора за группой банков на основе предложенной концептуальной модели и разработанной информационно-логической модели экспресс-анализа соответствия социально-экономического состояния коммерческого банка нормативам и регламентам ЦБ РФ для 111 коммерческих банков. Определены сочетания пороговых значений для выделенных кластеров банков, сегментирующие их на группы по текущему статусу (действующий, подлежащий дополнительной проверке из-за угрозы отзыва лицензии, или высока вероятность отзыва лицензии).

Список использованной литературы

Нормативно – правовые акты

1. Постановление правительства г. Москвы n 982-пп об утверждении порядков предоставления субсидий из бюджета города Москвы субъектам малого и среднего предпринимательства, осуществляющим предпринимательскую деятельность в рамках лицензионных договоров (сублицензионных договоров), договоров коммерческой концессии (субконцессии) от 06.08.2019 г. (в ред. от 30.12.2020 n 2405-пп).
2. Постановление Правительства РФ N 265 «О предельных значениях дохода, полученного от осуществления предпринимательской деятельности, для каждой категории субъектов малого и среднего предпринимательства» от 04.04.2016.
3. Приказ Минпромторга России N 798 «Об утверждении Перечня организаций, оказывающих существенное влияние на отрасли промышленности и торговли» от 13.04.2015 (ред. от 30.07.2018).
4. Федеральный закон N 294-ФЗ «О защите прав юридических лиц и индивидуальных предпринимателей при осуществлении государственного контроля (надзора) и муниципального контроля» от 26.12.2008.
5. Федеральный закон N 44-ФЗ «О контрактной системе в сфере закупок товаров, работ, услуг для обеспечения государственных и муниципальных нужд» от 05.04.2013 (ред. от 01.05.2019) (с изм. и доп., вступ. в силу с 12.05.2019).
6. Федеральный закон N 8-ФЗ «Об обеспечении доступа к информации о деятельности государственных органов и органов местного самоуправления» от 09.02.2009 (ред. от 30.04.2021).
7. Федеральный закон N 209-ФЗ «О развитии малого и среднего предпринимательства в Российской Федерации» от 24.07.2007 (ред. от 27.12.2018).

8. Федеральный закон № 184-ФЗ «Об общих принципах организации законодательных (представительных) и исполнительных органов государственной власти субъектов Российской Федерации» от 06.10.1999 г.
9. Федеральный закон № 229-ФЗ «Об исполнительном производстве» от 02.10.2007.

Научная литература:

10. Bang S., Son S., Roh H., Lee J.3 Bae S., Lee K., Hong C., Shin H. Quad-phased data mining modeling for dementia diagnosis // BMC Med Inform Decis Mak. – May, 2017. – 17 (Suppl.1). [Электронный ресурс]. URL: <https://bmcmedinformdecismak.biomedcentral.com/articles/10.1186/s12911-017-0451-3/> (Дата обращения: 24.02.2022). Doi: 10.1186/s12911-017-0451-3.
11. Berk R. Statistical Learning from a Regression Perspective. - New York, NY: Springer Verlag, 2008. – 358 p.
12. Bidzhoyan D., Bogdanova T. Russian banks credit risk stress-testing based on the publicly available data // Advances in Intelligent Systems and Computing. -2018. - Vol. 850. - P. 262–271.
13. Breiman L., Friedman J., Olshen R., C. Stone. Classification and Regression Trees. - Belmont, CA: Wadsworth International Group, 1984. – 368 p.
14. Dogan M.V., Beach S.R.H., Simons R.L., Lendasse A., Penaluna B., Philibert R.A. Blood-Based Biomarkers for Predicting the Risk for Five-Year Incident Coronary Heart Disease in the Framingham Heart Study via Machine Learning. Genes (Basel) // Genes, 2018. - 9(12). – P. 1–15. Doi: 10.3390/genes9120641.
15. Hastie T., Tibshirani R. and J. Friedman. The Elements of Statistical Learning. - New York: Springer. Verlag, 2009. – 200 p.
16. Hosmer D.W., Lemeshow S., Sturdivant R. Applied Logistic Regression. - Wiley. – 2013. – 558 p.
17. Liu, B. Sentiment Analysis and Subjectivity. - Handbook of Natural Language Processing, 2010. – 36 p. DOI: 10.1201/9781420085938-c26.

18. Manyika J. et al. Big data: The next frontier for innovation, competition, and productivity. - McKinsey Global Institute. [Электронный ресурс]/ URL: <https://www.mckinsey.com/business-functions/mckinsey-digital/our-insights/big-data-the-next-frontier-for-innovation> / (Дата обращения: 24.02.2022).
19. Naydenova E., Tsanas A., Howie S., Casals-Pascual C., De Vos M. The power of data mining in diagnosis of childhood pneumonia // *J R Soc Interface*. – 2016. - 13(120). – P 21-31. Doi: 10.1098/rsif.2016.0266.
20. Pang, B., Lee, L. Opinion Mining and Sentiment Analysis // *Foundations and Trends in Information Retrieval*. – 2008. - Vol. 2. - No 1-2. – P. 1-135.
21. Quinlan R. C4.5: Programs for machine learning. - Morgan Kaufmann Publishers, Inc. - 1993. – 312 p.
22. Shahmoradi L., Langarizadeh M., Pourmand G., Fard Z.A., Borhani A. Comparing Three Data Mining Methods to Predict Kidney Transplant Survival // *Acta Inform Med*. – 2016. - 24(5). – P. 322–327. Doi:10.5455/aim.2016.24.322–327.
23. Tavares M., Paredes S., Rocha T., Carvalho P., Ramos J., Mendes D., Henriques J., Morais J. Expert knowledge integration in the data mining process with application to cardiovascular risk assessment // *Conf Proc IEEE Eng Med Biol Soc*. – 2015.- P. 89-109. Doi: 10.1109/EMBC.2015.7318909.
24. Tian Zhang, Raghu Ramakrishnan, Miron Livny BIRCH: An Efficient Data Clustering Method for Very Large Databases // *ACM UKIGMOD International Conference on Management of Data*. – 1996. [Электронный ресурс]/ URL: <http://citeseer.ist.psu.edu/zhang96birch.html>/ (Дата обращения: 25.10.2021).
25. Tseng C.J., Lu C.J., Chang C.C., Chen G.D., Cheewakriangkrai C. Integration of data mining classification techniques and ensemble learning to identify risk factors and diagnose ovarian cancer recurrence // *Artif Intell Med*. – 2017. - № 78. P. 47–54. Doi: 10.1016/j.artmed.2017.06.003. Epub 2017 Jun 10.
26. Xu J., Yang P., Xue S., Sharma B., Sanchez-Martin M., Wang F., Beaty K.A., Dehan E. Parikh B. Translating cancer genomics into precision medicine with

artificial intelligence: applications, challenges and future perspectives // Hum Genet. – 2019. - 138(2). – P. 109–124. Doi: 10.1007/s00439-019-01970-5.

27. Zhukova L.V, Polyakov K.L. Comparative Analysis of Predictive Analytics Models in Classification Problems // Actual Problems of Systems and Software Engineering APSSE 2019 (Invited Papers). - Los Alamitos, Washington, Tokyo: IEEE Computer Society, 2019. - P. 162-169.

28. Аверчев И.В. Управленческий учет и отчетность. Постановка и внедрение / И.В. Аверчев. — М.: Вершина, 2006. — 512 с.

29. Айвазян С.А., Афанасьев М.Ю. Оценка мероприятий, направленных на управление факторами неэффективности производства // Прикладная эконометрика. 2007. - Т. 15. - № 3. - С. 43–55.

30. Айвазян С.А., Афанасьев М.Ю., Кудров А.В. Индикаторы основных направлений социально-экономического развития и их агрегаты в пространстве характеристик региональной дифференциации // Прикладная эконометрика. – 2019. - № 2 (54) – С. 51-69.

31. Андреева Н.А., Угримова С.Н. К вопросу применения статистических методов интегральной оценки эффективности системы управления промышленными предприятиями // Учет и статистика. – 2019. - № 1 (53). - С. 42–49.

32. Биджоян Д.С., Богданова Т.К. Концепция моделирования и прогнозирования вероятности отзыва лицензии российских банков // Экономическая наука современной России. – 2017. - Т. 79. - № 4. - С. 88–103.

33. Богданова Т.К., Алексеева Ю.А. Прогнозирование вероятности банкротства предприятий с учетом изменения финансовых показателей в динамике // Бизнес-информатика. – 2011.– Т1 (15).– С. 50-60.

34. Богданова Т.К., Жукова Л.В. Построение комплексного индикатора для оценки состояния российского коммерческого банка на основе структурированных и неструктурированных данных // Международная научная школа-семинар «Системное моделирование социально-

экономических процессов»», 43-е заседание 13-18 октября 2020 г. - ЦЭМИ РАН. – 2020. – С. 481–488.

35. Богданова Т.К., Жукова Л.В. Оценка состояния объекта управления на основе универсального комплексного индикатора с использованием структурированных и неструктурированных данных // Бизнес-информатика. – 2021ю - Т. 15. - № 2. - С. 21–33.

36. Большакова Е.И., Воронцов К.В., Ефремова Н.Э., Клышинский Э.С., Лукашевич Н.В., Сапин А.С. Автоматическая обработка текстов на естественном языке и анализ данных. - М: НИУ ВШЭ, 2017. - 269 с.

37. Бородин О.Н. Извлечение информации из полнотекстовых источников данных // Агроинженерия. - 2008. - №1. – С. 42-43.

38. Волков А.И., Рейнгольд Л.А. Открытые данные: проблемы и решения // Прикладная информатика – 2014. - №3 (51). - С. 5-12.

39. Гайдук Е.А. Диссертация на соискание ученой степени кандидата экономических наук «Применение балансового и оптимизационного моделирования при принятии управленческих решений на муниципальном уровне». - 2012. [Электронный ресурс] - / <https://www.dissercat.com/content/primenenie-balansovogo-i-optimizatsionnogo-modelirovaniya-pri-prinyatii-upravlencheskikh-res/> (Дата обращения: 11.12.2021).

40. Гальперин В.М., Игнатъев С.М., Моргунов В.И. Микроэкономика. - СПб.: Экономическая школа, 2004. - Т.1. – 348 с.

41. Головань С.В., Евдокимов М.А., Карминский А.М., Пересецкий А.А. Модели вероятности дефолта российских банков. II Влияние макроэкономических факторов на устойчивость банков // Российская экономическая школа. Препринт WP/2004/043. - 2004. - 25 с.

42. Горев А.И., Горева Е.Г. О применимости существующих алгоритмов обработки данных к Big data // МСМ. -2020. -№1 (53). – С. 139-143.

43. Громова С.В. Проблемы организации внутреннего контроля на предприятии // Молодой ученый. – 2016. - № 18 (122). - С. 240–242.
44. Демидова Л.А., Степанов М.А. Подход к решению задачи выявления структурных трансформаций в группах временных рядов // Cloud of science. - 2019. - №2. – С. 75-79.
45. Дихтль Е., Хершген Х. Практический маркетинг: учеб, пособие / пер. с нем. А.М. Макарова; под ред. И.С. Минко. - М.: Высшая школа, 1995. - 228 с.
46. Елисеева Е.Н. Финансовый инструментарий оценки несостоятельности промышленных предприятий // Регион: системы, экономика, управление. – 2019. - №3 (46).- С. 132-140.
47. Жукова Л.В. Экспресс-анализ состояния промышленных предприятий Москвы с использованием универсального комплексного индикатора. // Экономическая наука современной России. -2021. -Т.4 (95). -С. 89–96. DOI: 10.33293/1609-1442-2021-4(95)-89-97.
48. Жукова Л.В., Кирюшина А.А., Ковальчук И. М., Рузаева А. В. Повышение результативности системы дистанционного образования с помощью машинного обучения и технологии блокчейн // Прикаспийский журнал: управление и высокие технологии. -2018. -№ 2 (42). -С. 56-68.
49. Жукова Л.В., Кирюшина А.А., Ковальчук И.М., Чугунов В.Р. Применение методов Data Science и математической статистики для структурирования спроса и предложения на промышленные товары и услуги // Actual Problems of System and Software Engineering. Proceedings of the 6th International Conference Actual Problems of System and Software Engineering. Moscow, Russia, 12-14 November, - 2019. - Vol. 2514.- P. 333 -341.
50. Жукова Л.В., Поляков К.Л. Опыт моделирования вероятности кредитного дефолта клиентов микрофинансовых организаций (на примере одной МФО) // Экономический журнал Высшей школы экономики. - 2019. - Т. 4. - № 23. - С. 497-523, DOI: 10.17323/1813-8691-2019-23-4-497-523.

51. Карминский А.М., Костров А.В. Моделирование вероятности дефолта российских банков: расширенные возможности // Журнал новой экономической ассоциации. - 2014. - Т. 17. - № 1. - С. 64–86.
52. Касевич В.Б. Элементы общей лингвистики. - М.: Наука, 1977. - 183 с.
53. Кирюшина А.А, Жукова Л.В., Чикина Л.Г. Использование больших данных в оценке степени загрязнения прибрежной зоны морского побережья курортных регионов // Тезисы докладов Всероссийской научной конференции «Моря России: исследования береговой и шельфовой зон». - 2020. - С. 409-410.
54. Клейнер Г.Б. Эволюция и реформирование промышленных предприятий // Вопросы экономики. - 2000. - №5. - С. 62-74.
55. Климкин Ю.Н. О механизмах управления промышленными предприятиями // Региональные проблемы преобразования экономики. – 2018. - № 3 (89). - С. 19-25.
56. Краснянский М.Н. Обухов А. Д., Соломатина Е.М., Воякина А.А. Сравнительный анализ методов машинного обучения для решения задачи классификации документов научно-образовательного учреждения // Вестник Воронежского государственного университета. - 2018. – № 3. – С. 173-182.
57. Кричевский М.Л. Методы машинного обучения при выборе стратегии предприятия // Вопросы инновационной экономики. 2019. - Т.9. - № 1. С. 251-266.
58. Майер-Шенбергер В, Кукьер К. Большие данные: революция, которая изменит нашу жизнь, работу и мысли. - И: Манн, Иванов и Фербер, 2013. - 240 с.
59. Малых Н.И., Проданова Н.А. Современные подходы к оценке стоимости коммерческого банка // Статистика и экономика. - 2016. - №6. – С. 79-84.
60. Марков С.М., Маркова А.С. Целевые показатели (индикаторы) как критерии комплексной оценки эффективности реализации государственных программ поддержки малого и среднего предпринимательства // Вестник

Псковского государственного университета. Серия: Экономика. Право. Управление. - 2017. - №5. – С. 132 – 137.

61. Медведев Д.А. Большие данные: причины появления и как их можно использовать // Наука и образование сегодня. – 2019. - № 4 (39). - С. 14-16.

62. Милкова М.А., Неволин И.В., Пигорев Д.П. Современные методы извлечения ключевой информации из нормативных документов // Экономическая наука современной России. – 2021 – Т 2. – С. 101-114. [https://doi.org/10.33293/1609-1442-2021-2\(93\)-101-114](https://doi.org/10.33293/1609-1442-2021-2(93)-101-114).

63. Морозов А.Н. Альтернативные источники статистической информации как основа принятия политических решений // Вопросы государственного и муниципального управления. – 2018. - № 2. -С. 50-70.

64. Мрочковский Н.С. (2020). Комплексная оценка эффективности управления цифровизацией организаций // Инновации и инвестиции. – 2020. - №6. - С. 127–129.

65. Муллахметов Х.Ш. Проблемы организации управленческого контроля // Экономическая наука современной России. – 2007. - № 4. – С. 59-69.

66. Нортон Дэйвид П., Каплан Роберт С. Сбалансированная система показателей. - М: Олимп-Бизнес, 2008. – 304 с.

67. Олендерфер М. С., Блэшфилд Р. К. Кластерный анализ / Факторный, дискриминантный и кластерный анализ: пер. с англ.; Под. ред. И. С. Енюкова. — М.: Финансы и статистика, 1989. – 202 с.

68. 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. -С. 24-31.

69. Паклин Н. Алгоритмы кластеризации на службе Data Mining. BaseGroup Labs [Электронный ресурс] URL: <http://www.basegroup.ru/clusterization/datamining.htm/> (дата обращения: 17.07.2021).

70. Пересецкий А.А. Модели причин отзыва лицензии российских банков // Препринт WP/2010/085. - М.: РЭШ, 2010. – 26 с.
71. Пересецкий А.А. Эконометрические методы в дистанционном анализе деятельности российских банков. - М.: ВШЭ, 2012. – 191 с.
72. Писарева О.М. Анализ состояния и характеристика потенциала развития инструментария стратегического планирования в условиях цифровой трансформации экономики и управления // МИР (Модернизация. Инновации. Развитие). – 2018. - Т. 9. - № 4.– С. 502-529.
73. Полшков Ю.Н. Прикладные эконометрические методы анализа рисков при управлении хозяйственным комплексом региона и уровнем жизни его населения // Вестник Института экономических исследований. -2019. -№1 (13). –С. 12-18.
74. Попов Е.В., Симонова В.Л. Классификация теорий предприятия // Journal of new economy. – 2003. - №. 6.– С. 80-89.
75. Попова С.В., Ходырев И.А. Извлечение ключевых словосочетаний // Научно-технический вестник информационных технологий, механики и оптики. - 2012. - №1 (77). – С. 67-71.
76. Пузанов А.С., Трутнев Э.К., Маркварт Э., Попов Р.А., Сафарова М.Д. Стратегическое планирование и градорегулирование на муниципальном уровне. - М.: Издательский дом «Дело» РАНХиГС. 2017. - 354 с.
77. Савенков П.А. Использование методов и алгоритмов машинного обучения в системах поддержки принятия управленческих решений // Вестник науки и образования. - 2019. - №1-2 (55). – С. 23-25.
78. Семина Т.А. Анализ тональности текста: современные подходы и существующие проблемы // Социальные и гуманитарные науки. Отечественная и зарубежная литература. Сер. 6, Языкознание: Реферативный журнал. - 2020. - №4. – С. 198 – 210.

79. Соболевская Ю.В. Контрольные органы: уровни компетенции // Вопросы государственного и муниципального управления.- 2013.- № 4.- С. 113–130.
80. Турищева Т.Б. Формирование системы внутреннего контроля (СВК) в российских условиях. // Азимут научных исследований: экономика и управление. – 2017. - Т.6. - №.3(20). - С. 351–354.
81. Фролова Е.Е. Виды государственного контроля в рамках финансовой деятельности государства // Известия Байкальского государственного университета. – 2010.- №5 - С. 136-143.
82. Чугунов В.Р., Жукова Л.В., Ковальчук И.М., Ковалева А.С. Математические методы группирования данных для принятия управленческих решений в задачах планирования // Actual Problems of System and Software Engineering 2017. Proceedings of the 5th International Conference on Actual Problems of System and Software Engineering Supported by Russian Foundation for Basic Research. – 2017. – С. 333-341.
83. Шалымов Д.С. Алгоритмы устойчивой кластеризации на основе индексных функций и функций устойчивости // Санкт-Петербургский государственный университет. – 2008. – Т4. – С. 236–248.
84. Южаков В.Н., Добролюбова Е.И., Покида А.Н., Зыбуновская Н.В. Оценка результативности государственного контроля с позиции бизнеса: ключевые тенденции // Вопросы государственного и муниципального управления. – 2020. - № 2.- С. 32-53.

Приложение 1. Типологизация источников общедоступной информации

Наименование группы источника информации	Характеристика источника информации по достоверности	Структурированность информации	Обновление данных в источнике	Пример источника информации
Официальные генераторы и агрегаторы данных	Сайты федерального и региональных органов статистики; сайты министерств и ведомств, публикующие по положению о раскрытии информации тематические данные, достоверность которых подтверждается соответствующим органом государственной власти	Структурированные данные	Периодичность обновления — обычно 1 раз в квартал или реже	rosstat.gov.ru; zakupki.gov.ru; fssp.gov.ru; cbr.ru; wciom.ru
Сайты и страницы в социальных сетях объектов исследования	Сайты предприятий, организаций всех форм собственности, на которых они обязаны размещать информацию о своей деятельности. Достоверность информации, как правило, подтверждается только самим объектом исследования	Все типы данных	Постоянное обновление	technomoscow.ru; uniconf.ru; tinkoff.ru; 57.mskobr.ru
Неофициальные генераторы данных	Сайты организаций, занимающихся деятельностью, связанной с объектами исследования и публикующие данные о них в открытых	Преимущественно структурированные данные	Согласно утвержденной методике, обновление может проводиться как с заданной периодичностью, так и постоянно	cian.ru; hse.ru/rlms

Наименование группы источника информации	Характеристика источника информации по достоверности	Структурированность информации	Обновление данных в источнике	Пример источника информации
	источниках. Достоверность обеспечивается внутренним мониторингом и контролем информации			
Неофициальные агрегаторы данных.	Российские и международные агрегаторы данных, обычно предоставляющие их для научных и иных исследований. Достоверность обеспечивается внутренним мониторингом	Преимущественно структурированные данные	Обычно обновление проводится с периодичностью, соответствующей периодичности обновления официальных данных	bankodrom.ru; banki.ru; avtostat.ru; data.worldbank.org.
Неофициальные интернет-источники экспертных исследований	Российские и международные сайты экспертных организаций, рейтинговых агентств, персональные страницы признанных экспертов. Достоверность данных обеспечивается репутацией эксперта	Все типы данных	Обновление проводится согласно частным правилам источника информации	raexpert.ru; ra-national.ru; schooexpert.ra
Неофициальные общедоступные интернет-источники	Страницы в социальных сетях, блоги, комментарии под контентом, страницы неофициальных сообществ. Достоверность данных, как правило, не подлежит проверке	Преимущественно неструктурированные или слабоструктурированные данные	Постоянное обновление	moneyzz.ru; pedsovet.su

**Приложение 2. Перечень используемых обозначений для базы данных
промышленных предприятий**

Наименование	Наименование в БД	Тип данных
ИНН	INN	Строка
Адрес web-сайта	url	Строка
Направление деятельности	activity_area_official	Строка
Направление деятельности дополнительно	activity_area_official_2	Строка
Область применения продукции	products_scope	Строка
Адрес	address	Строка
ОГРН	OGRN	Строка
Статус юридического лица	artificial_body_status	Строка
Год регистрации	registration_year	Строка
Основная отрасль	main_industry	Строка
Уставный капитал (по данным ЕГРЮЛ)	registered_capital_EGRUL	Число
Уставный капитал, руб	registered_capital	Число
Режим налогообложение	taxation_scheme	Строка
Среднесписочная численность работников	employees	Число
Филиалы и представительства, количество	filiation	Число
Выручка, млн р, 2016	fin_earnings_2016	Число
Выручка, млн р, 2017	fin_earnings_2017	Число
Выручка, млн р, 2018	fin_earnings_2018	Число
Прибыль, млн р, 2016	fin_profit_2016	Число
Прибыль, млн р, 2017	fin_profit_2017	Число
Прибыль, млн р, 2018	fin_profit_2018	Число
Капитал, 2016	capital_2016	Число
Капитал, 2017	capital_2017	Число
Капитал, 2018	capital_2018	Число
Госзакупки, год начала участия	gov_year	Строка
Госзакупки, роль Поставщик	gov_supplier	Число
Госзакупки, роль Заказчик	gov_customer	Число
Госзакупки, количество контрактов 2016	gov_count_2016	Число
Госзакупки, количество контрактов 2017	gov_count_2017	Число
Госзакупки, количество контрактов 2018	gov_count_2018	Число
Госзакупки, сумма контрактов, р 2016	gov_value_2016	Число
Госзакупки, сумма контрактов, р 2017	gov_value_2017	Число
Госзакупки, сумма контрактов, р 2018	gov_value_2018	Число
БАЛАНС (актив), 2016	balance_2016	Число
БАЛАНС (актив), 2017	balance_2017	Число

Наименование	Наименование в БД	Тип данных
Внеоборотные активы Результаты исследований и разработок 2016	non_current_assets_results_2016	Число
Внеоборотные активы Результаты исследований и разработок 2017	non_current_assets_results_2017	Число
Внеоборотные активы Основные средства 2016	non_current_assets_fixed_2016	Число
Внеоборотные активы Основные средства 2017	non_current_assets_fixed_2017	Число
Оборотные активы Запасы 2016	current_assets_inv_2016	Число
Оборотные активы Запасы 2017	current_assets_inv_2017	Число
Оборотные активы Дебиторская задолженность 2016	current_assets_debt_2016	Число
Оборотные активы Дебиторская задолженность 2017	current_assets_debt_2017	Число
Капитал и резервы. Нераспределенная (непокрытый убыток) 2016	capital_undist_profit_2016	Число
Капитал и резервы. Нераспределенная (непокрытый убыток) 2017	capital_undist_profit_2017	Число
Краткосрочные обязательства 2016	short_term_obligation_2016	Число
Краткосрочные обязательства 2017	short_term_obligation_2017	Число
Общее количество вакансий	total_vacancies	Число
Количество вакансии рабочих	count_vac_workers	Число
Количество вакансии специалистов	count_vac_specialists	Число
Количество вакансии управленцев	count_vac_managers	Число
Количество вакансии без опыта работы или образования	count_vac_exp	Число
Раздел направления деятельности	section_activity_area	Строка
Направление деятельности по данным картографических служб	activity_area	Строка
Количество фирм в категории по г. Москве	count_activity_area	Число
Направление деятельности дополнительно	activity_area_2	Строка
Максимальная сумма контракта, руб	max_value	Число
Максимальная сумма контракта за 2016-2018	max_value_2016_2018	Число

Приложение 3. Значения среднего и вариации основных финансовых показателей для выборки промышленных предприятий за 2016-2018 гг.

Наименование показателя	Среднее значение	Диапазон	Стандартное отклонение
Выручка, млн.р., 2016	413,998	18000,000	1452,832
Выручка, млн.р., 2017	443,07	16006,10	1434,84
Чистая прибыль, млн.р., 2016	32,34	11747,18	477,44
Чистая прибыль, млн.р., 2016	-4,04	11492,76	505,28
Стоимость капитала, млн.р., 2016	302,49	42568,72	1871,31
Стоимость капитала, млн.р., 2016	378,84	54180,43	2467,62
Количество контрактов на портале госзакупок, 2016	40	999	118
Количество контрактов на портале госзакупок, 2017	36	943	100
Количество контрактов на портале госзакупок, 2018	19	427	48
Сумма баланса, тыс.р., 2016	976335	54973137	4541200
Сумма баланса, тыс.р., 2017	1070854	56155103	4556078
Основные средства, тыс.р., 2016	278740	36357243	1988450
Основные средства, тыс.р., 2017	341802	37618616	2135592
Текущие активы, запасы, тыс.р., 2016	250498	25197811	1571669
Текущие активы, запасы, тыс.р., 2017	211075	16270576	1071891
Дебиторская задолженность, тыс.р., 2016	77781	34702745	1484924
Дебиторская задолженность, тыс.р., 2017	18231	44683562	2005835
рентабельность_активов_2016	,00	,02	,00
рентабельность_активов_2017	,00	,02	,00
ресурсоотдача_2016	,00	,02	,00
ресурсоотдача_2017	,00	,04	,00
рентабельность_капитала_2016	,81	186,46	8,29
рентабельность_капитала_2017	1,57	391,50	19,58
дельта_выручка	30,23	10084,31	436,88
дельта_прибыль	-36,26	21281,01	946,87
дельта_капитал	75,00	26507,53	997,26

Приложение 4. Перечень предприятий группы 1. «Доминирующие предприятия»

№ п/п	Название резидента	ИНН
1	ОАО «Ударница»	7706042326
2	АО «Водный Стадион Спорт Инвест»	7736525151
3	ООО «Маппер»	7743859476
4	ЗАО "Партнер Ф"	7709239514
5	ПАО «НК «Роснефть» – МЗ «Нефтепродукт»,	7720065580
6	ОАО "Царицыно"	7724017435
7	ПАО "КМЗ"	7721024057
8	АО "ЛАБОРАТОРИЯ КАСПЕРСКОГО"	7713140469
9	АО "ОКБ "Кристалл"	7720015691
10	АО "Косметическое объединение "СВОБОДА"	7714078157
11	МОАО «Большевичка»	7708029923
12	ЗАО "Хлебозавод №22"	7731014604
13	АО "Биннофарм"	7735518627
14	ОАО "Кондитерский концерн "Бабаевский"	7708029391
15	АО Вимм-Билль-Дан	7713085659
16	ООО "ВКП ЛТ"	7736032029
17	ОАО «Черкизовский Мясоперерабатывающий Завод»	7718013714
18	ОАО Московская кондитерская фабрика "Красный Октябрь"	7706043263
19	ОАО "ФАБЕРЛИК"	5001026970
20	ПАО "МОСТОТРЕСТ"	7701045732
21	АО "АНГСТРЕМ"	7735010706
22	АО "МПО им. Румянцева"	7714081921
23	ОАО «Рот Фронт»»	7705033216
24	АО "Центральный научно-исследовательский институт "Циклон"	7718159209
25	ООО "Дымовское колбасное производство"	7731178578
26	ФГУП «Московский эндокринный завод»	7722059711
27	ФГУП "НАМИ"	7711000924
28	АО "Научно-Исследовательский Институт Систем Связи и Управления"	7728804257
29	АО "Федеральный центр науки и высоких технологий "Специальное научно- производственное объединение "Элерон"	7724313681
30	АО ГКБ «Вымпел» имени И.И. Торопова	7733546058
31	ЗАО "Рено Россия"	7709259743
32	ФГУП ВНИИА им. Духова	7707074137
33	АО "НПЦ газотурбостроения "Салют" (с 01.01.2019 Реорганизация путем присоединения к АО ОДК ИНН 7731644035)	7719409437
34	АО "ММП имени В.В. Чернышева"	7733018650
35	ОАО "ЦЛЗ"	5051000880

**Приложение 5. Сравнительная таблица средних значений финансовых показателей группы 2 и группы 3
результатов разбиения промышленных предприятий.**

	Группа предприятий					
	группа 2			группа 3		
	Среднее значение	Диапазон	Стандартное отклонение	Среднее значение	Диапазон	Стандартное отклонение
fin_earnings_2016	784,00	18.000,00	1.952,73	19,66	555,00	48,41
fin_earnings_2017	831,46	16.000,00	1.918,55	29,12	882,10	70,50
fin_profit_2016	61,15	11.747,18	664,03	1,64	87,33	7,66
fin_profit_2017	-10,28	11.492,76	706,11	2,52	137,28	14,01
capital_2016	576,11	42.568,72	2.577,60	10,88	519,07	42,68
capital_2017	725,33	54.180,43	3.413,00	14,12	513,41	45,78
balance_2016	2.034.047,09	54.973.137,00	6.468.965,05	39.240,08	2.936.897,00	204.528,13
balance_2017	2.181.218,64	56.155.016,00	6.422.177,60	62.760,63	4.902.674,00	336.148,83
non_current_assets_results_2016	16.595,28	610.365,00	59.242,64	1.237,44	55.238,00	6.061,50
non_current_assets_results_2017	20.404,97	409.725,00	58.796,94	1.523,63	84.103,00	8.377,87
non_current_assets_fixed_2016	547.829,02	36.357.243,00	2.782.142,39	6.960,29	427.427,00	40.366,58
non_current_assets_fixed_2017	663.323,67	37.618.616,00	2.970.260,33	12.662,33	605.038,00	64.299,51
current_assets_inv_2016	331.130,50	11.024.956,00	1.178.904,54	2.982,71	59.021,00	8.186,42
current_assets_inv_2017	359.855,67	12.923.364,00	1.347.542,95	5.288,90	133.541,00	15.386,39
current_assets_debt_2016	487.869,24	25.197.811,00	2.193.870,00	10.753,62	261.848,00	33.517,94
current_assets_debt_2017	402.347,13	16.270.576,00	1.483.462,44	15.270,92	305.364,00	38.966,09
capital_undist_profit_2016	152.033,07	34.702.745,00	2.094.671,81	2.785,97	219.303,00	15.677,32

capital_undist_profit_2017	33.843,83	44.683.562,00	2.823.306,71	2.248,15	275.720,00	18.679,08
short_term_obligation_2016	854.899,09	30.482.488,00	3.233.111,00	24.195,43	1.678.337,00	124.900,26
short_term_obligation_2017	833.286,68	28.024.583,00	3.051.810,78	28.927,94	981.651,00	99.527,51
gov_count_2016	50,88	999,00	134,82	8,44	110,00	18,29
gov_count_2017	46,39	943,00	114,08	7,89	84,00	16,00
gov_count_2018	24,45	427,00	55,37	5,78	90,00	13,34
gov_value_2016	6.103.875.941,73	827.000.000.000,00	68.187.849.819,93	10.515.143,15	121.500.000,00	20.310.273,38
gov_value_2017	531.427.802,77	15.869.671.000,00	1.704.980.538,47	17.021.456,39	327.900.000,00	46.501.383,74
gov_value_2018	18.998.144.216,80	2.775.000.000.000,00	227.310.727.740,01	14.072.146,87	168.200.000,00	30.964.853,14
коэффициент_оборота_ДЗ_2016	1.210,86	62.296,25	6.222,10	1.143,76	39.909,03	4.310,91
коэффициент_оборота_ДЗ_2017	684.227,61	133.782.470,73	9.530.911,96	1.548,62	109.315,41	8.818,06
средний_контракт_2016	22.268.448,38	2.552.469.135,80	169.359.515,75	729.818,35	60.750.000,00	4.411.234,29
средний_контракт_2017	10.510.023,18	662.517.178,00	48.942.255,93	816.246,63	65.580.000,00	4.859.005,29
прирост_выручка	49,72	10.084,31	607,28	9,46	423,00	34,47
прирост_прибыль	-71,10	21.281,01	1.318,53	,88	157,77	15,28
прирост_капитал	142,33	26.507,53	1.386,25	3,24	316,15	20,81

**Приложение 9. Перечень используемых обозначений для базы данных
коммерческого банка**

Показатели по ставкам на депозиты банков по данным из внешних источников взяты за период с 01.01.2018 по 01.02.2020 годы

№	Обозначение	Наименование	Содержание
1	R1	Maximum interest rate of the bank on-call rub deposit, %	Максимальная процентная ставка банка по вкладам в рублях согласно заключенным договорам с физическими лицами со сроком до востребования, % годовых
2	R2	Maximum interest rate of the bank for up to 90 days rub deposit , %	Максимальная процентная ставка банка по вкладам в рублях согласно заключенным договорам с физическими лицами на срок до 90 дней, % годовых
3	R3	Maximum interest rate of the bank for 91 to 180 days rub deposit , %	Максимальная процентная ставка банка по вкладам в рублях согласно заключенным договорам с физическими лицами на срок от 91 до 180 дней, % годовых
4	R4	Maximum interest rate of the bank for 181 days to 1 year rub deposit , %	Максимальная процентная ставка банка по вкладам в рублях согласно заключенным договорам с физическими лицами на срок от 181 дня до 1 года, % годовых
5	R5	Maximum interest rate of the bank extending beyond one year rub deposit, %	Максимальная процентная ставка банка по вкладам в рублях согласно заключенным договорам с физическими лицами на срок свыше 1 года, % годовых
		Средние ставки на депозиты банков по данным сайта Банка России за период с 01.01.2018 по 01.02.2020 годы	
6	D1	Weighted average interest rates on-call rub deposit for the whole Russia, %	Средневзвешенные процентные ставки по привлеченным кредитными организациями вкладам (депозитам) физических лиц в рублях (в целом по Российской Федерации) физических лиц, со сроком привлечения до востребования, % годовых
7	D2	Weighted average interest rates for 31 to 90 days rub deposit for the whole Russia, %	Средневзвешенные процентные ставки по привлеченным кредитными организациями вкладам (депозитам) физических лиц в рублях (в целом по Российской Федерации) физических лиц, со сроком привлечения от 31 до 90 дней, % годовых

№	Обозначение	Наименование	Содержание
8	D3	Weighted average interest rates for 91 to 180 days rub deposit for the whole Russia, %	Средневзвешенные процентные ставки по привлеченным кредитными организациями вкладам (депозитам) физических лиц в рублях (в целом по Российской Федерации) физических лиц, со сроком привлечения от 91 до 180 дней, % годовых
9	D4	Weighted average interest rates for 181 to 1 year rub deposit for the whole Russia, %	Средневзвешенные процентные ставки по привлеченным кредитными организациями вкладам (депозитам) физических лиц в рублях (в целом по Российской Федерации) физических лиц, со сроком привлечения от 181 дня до 1 года, % годовых
10	D5	Weighted average interest rates extending beyond one year rub deposit for the whole Russia, %	Средневзвешенные процентные ставки по привлеченным кредитными организациями вкладам (депозитам) физических лиц в рублях (в целом по Российской Федерации) физических лиц, со сроком привлечения свыше 1 года, % годовых
11	D6	Weighted average interest rates on-call rub deposit for 30 largest banks, %	Средневзвешенные процентные ставки по привлеченным кредитными организациями вкладам (депозитам) физических лиц в рублях (по 30 крупнейшим банкам) физических лиц, со сроком привлечения до востребования, % годовых
12	D7	Weighted average interest rates for 31 to 90 days rub deposit for 30 largest banks, %	Средневзвешенные процентные ставки по привлеченным кредитными организациями вкладам (депозитам) физических лиц в рублях (по 30 крупнейшим банкам) физических лиц, со сроком привлечения от 31 до 90 дней, % годовых
13	D8	Weighted average interest rates for 91 to 180 days rub deposit for 30 largest banks, %	Средневзвешенные процентные ставки по привлеченным кредитными организациями вкладам (депозитам) физических лиц в рублях (по 30 крупнейшим банкам) физических лиц, со сроком привлечения от 91 до 180 дней, % годовых
14	D9	Weighted average interest rates for 181 to 1 year rub deposit for 30 largest banks, %	Средневзвешенные процентные ставки по привлеченным кредитными организациями вкладам (депозитам) физических лиц в рублях (по 30 крупнейшим банкам) физических лиц, со сроком привлечения от 181 дня до 1 года, % годовых
15	D10	Weighted average interest rates	Средневзвешенные процентные ставки по привлеченным кредитными организациями

№	Обозначение	Наименование	Содержание
		extending beyond one year rub deposit for 30 largest banks, %	вкладам (депозитам) физических лиц в рублях (по 30 крупнейшим банкам) физических лиц, со сроком привлечения свыше 1 года, % годовых
		Макроэкономические показатели за период с 01.01.2018 по 01.02.2020 годы	
12	RF	Key rate of Central Bank, %	Ключевая ставка Банка России, % годовых
		Общие характеристики банка по данным сайта Банка России и других открытых источников по состоянию на март 2020 года	
13	N	Name of the bank	Краткое наименование банка в текстовом формате
14	F	Full name of the bank	Полное наименование банка в текстовом формате
15	Ro	Region of the bank	Наименование региона, в котором зарегистрирован банк в текстовом формате
16	Ren	Registry number of the bank	Регистрационный номер банка по Книге государственной регистрации кредитных организаций
17	Red	Registration date of the bank	Дата регистрации банка
18	L	License's status of the of bank	Отметка о наличии и состоянии лицензии банка в поле "Лицензия". Например в случае ликвидации запись: "Лицензия отозвана приказом Банка России ОД-6 от 10.01.2020", в случае действующей организации запись: "Генеральная лицензия на осуществление банковских операций (26.08.2016)".
19	No	Number of offices	Количество отделений и филиалов банка
20	Pi	Participation in DI	Отметка об участии банка в системе страхования вкладов. Значения: Да/нет
21	Cc	Charter capital, rub	Величина уставного капитала, рублей
22	Na	Net assets, million-rub	Активы-нетто - стоимость активов организации (с учетом нематериальных) за вычетом имеющихся у организации долговых обязательств. Показатель характеризует реальную стоимость имущества банка, млн. рублей.

№	Обозначение	Наименование	Содержание
23	NpD	Natural persons' deposits, million-rub	Вклады физических лиц – лицевые счета по учету денежных средств физических лиц, открытых на основании договора банковского счета, вклады (депозиты) и прочие привлеченные средства физических лиц, принятые кредитной организацией по договору банковского вклада (депозита), в том числе путем выдачи вкладчику именной сберегательной книжки, выпущенные кредитными организациями сберегательные сертификаты, а также неисполненные обязательства по депозитам и прочим привлеченным средствам физических лиц, резидентов и нерезидентов, в валюте Российской Федерации и иностранной валюте, код формы по ОКУД 0409101, млн. рублей
24	NpL	Natural persons' loans million-rub	Кредиты, предоставленные физическим лицам – задолженность (включая просроченную) физических лиц, резидентов и нерезидентов, перед кредитными организациями по кредитам и прочим средствам в валюте Российской Федерации и иностранной валюте, код формы по ОКУД 0409101, млн. рублей
		Значения обязательных нормативов банка по данным сайта Банка России, код формы 0409135, месячная за январь, апрель, июль, октябрь 2019 года; январь 2020 года	
25	H1.1	Basic Capital Adequacy Standard (H1.1) on months, %	Норматив достаточности базового капитала банка (H1.1) по месяцам, в процентах
26	H1.2	Capital adequacy standard (H1.2) on months, %	Норматив достаточности основного капитала банка (H1.2) по месяцам, в процентах
27	H1.0	Equity standard (H1.0) on months, %	Норматив достаточности собственных средств (капитала) банка (H1.0) по месяцам, в процентах
28	H1.4	Financial leverage standard (H1.4) on months, %	Норматив финансового рычага (H1.4) по месяцам, в процентах
29	H2	Instant liquidity standard (H2) on months, %	Норматив мгновенной ликвидности банка (H2) по месяцам, в процентах
30	H3	Current liquidity standard (H3) on months, %	Норматив текущей ликвидности банка (H3) по месяцам, в процентах

№	Обозначение	Наименование	Содержание
31	H4	Long-term liquidity standard (H4) on months, %	Норматив долгосрочной ликвидности банка (H4) по месяцам, в процентах
		Показатели рейтинговых агентств за период с 2017 по 2020 годы	
32	Rao	Ranked in overall	Место в общем рейтинге по ключевым показателям деятельности рассчитывается по методике Banki.ru с использованием отчетности кредитных организаций РФ, публикуемой на сайте Банка России
33	Pir	Place in the rating by year	Место в рейтинге по методологии Экперт РА присвоения рейтингов кредитоспособности банкам по годам
34	Rpr	Rating prediction in year	Прогнозное состояние банка по методологии Экперт РА присвоения рейтингов кредитоспособности банкам по годам
35	Dr	Date of rating assignment in year	Дата рейтингования по методологии Экперт РА присвоения рейтингов кредитоспособности банкам по годам
36	Nev	Number of events 'reduction'	Количество событий, связанных с понижением места в рейтинге по результатам контекстного поиска