

АЛГОРИТМ ИНТЕЛЛЕКТУАЛЬНОГО АНАЛИЗА РЕГИОНАЛЬНЫХ ДАННЫХ ОБ ИНВЕСТИЦИОННОМ РИСКЕ

Е. А. МАКАРОВА • Э. Р. ГАБДУЛЛИНА • М. М. ЮСУПОВ • Г. Р. ВАГАПОВА

Аннотация. В статье предложен алгоритм интеллектуального анализа региональных данных об инвестиционном риске, который предполагает проведение компонентного, кластерного и нейросетевого анализа для первичного формирования кластеров, а также построение трех вариантов деревьев решений для окончательного формирования кластеров регионов. Выявлены малочисленные кластеры регионов, характеризующиеся высоким уровнем экономического, социального и экологического риска. Анализ выявленных кластеров регионов позволяет сформулировать их характерные особенности, а также сформулировать правила оценки инвестиционного риска.

Ключевые слова: инвестиционный риск; интеллектуальный анализ данных; компонентный анализ; нейронные сети Кохонена; кластер; деревья решений.

ВВЕДЕНИЕ

Решением проблем анализа инвестиционной привлекательности регионов и предприятий занимаются многие исследователи. Одной из задач анализа инвестиционной привлекательности является задача формирования инвестиционного рейтинга регионов, которая играет важную роль в стимулировании развития не только регионов, но и страны в целом. Формирование адекватной оценки инвестиционного рейтинга является важным фактором, который позволяет провести сравнение регионов по многим показателям, что необходимо для выявления проблемных областей в управлении государством или регионами и определения направлений их развития. Знание инвестиционного рейтинга позволяет, с одной стороны, мотивировать регионы к конкуренции за улучшение своих показателей, а с другой стороны, обеспечить информационную поддержку органам власти при определении приоритетов для развития регионов. В целом оценка инвестиционного рейтинга регионов имеет важное значение как с позиций государственных органов власти при принятии решений по распределению бюджетных ресурсов, так и с позиций частных инвесторов при принятии решений о размещении капитала [1–4].

В настоящее время создано множество методик оценки инвестиционного рейтинга регионов, используемых различными рейтинговыми агентствами, которые занимаются формированием различных рейтингов [5–9]. Построение рейтингов ведется на основе статистических данных федеральных ведомств, таких как Госкомстат, ЕМИСС, Минфин России [1].

Согласно методике расчета рейтинга регионов, предложенной компанией «РАЭС-Аналитика», каждому региону присваивается рейтинг инвестиционной привлекательности, который представлен в виде индекса, определяющего соотношение между уровнем интегрального инвестиционного риска и величиной совокупного инвестиционного потенциала региона. Агрегирование данных производится с использованием экспертных оценок [5].

По методологии оценки инвестиционной привлекательности регионов национального рейтингового агентства (НРА) инвестиционная привлекательность регионов складывается из следующих семи факторов: географическое положение и природные ресурсы; трудовые ресурсы региона; региональная инфраструктура; внутренний рынок региона; производственный потенциал региональной экономики; институциональная среда и социально-политическая стабильность; финансовая устойчивость регионального бюджета и предприятий региона [6]. Для оценки перечисленных семи факторов инвестиционной привлекательности используется

набор из 57 показателей. Источники данных, используемые в методике НРА для оценки инвестиционной привлекательности региона, представлены статистическими данными, опросами предпринимательского сообщества, а также экспертными оценками. Методика предполагает проведение расчета интегрального индекса инвестиционной привлекательности для каждого региона, а также применение методов кластерного анализа для выделения пороговых значений индекса и распределения регионов по трем укрупненным категориям и девяти группам инвестиционной привлекательности. Однако агрегирование данных также производится с использованием экспертных оценок.

Возникает необходимость совершенствования процедуры формирования инвестиционного рейтинга регионов и разработки цепочки взаимосвязанного применения методов интеллектуального анализа данных [10–18].

На основе предложенной методологии интеллектуального управления макроэкономической системой разрабатывается система интеллектуального управления инвестиционными процессами, включающая в себя подсистему интеллектуального анализа данных и формирования инвестиционного рейтинга регионов [19, 20]. Предложенная процедура разработки инвестиционного рейтинга регионов включает в себя следующие этапы: формирование исходной выборки и ее структурирование путем выделения трех локальных выборок данных, отражающих инвестиционную привлекательность, инвестиционный потенциал и инвестиционный риск; проведение интеллектуального анализа данных (ИАД) по каждой выборке на основе применения методов компонентного, кластерного, нейросетевого и нейро-нечеткого анализа; интегрирование полученных данных с использованием компонентного анализа; формирование инвестиционного рейтинга регионов [20–26].

В статье рассматривается выполнение одного из этапов предложенной процедуры формирования инвестиционного рейтинга регионов, связанного с проведением интеллектуального анализа региональных данных об инвестиционном риске.

АЛГОРИТМ ИНТЕЛЛЕКТУАЛЬНОГО АНАЛИЗА РЕГИОНАЛЬНЫХ ДАННЫХ

Цель анализа заключается в выявлении кластеров регионов РФ, различающихся по показателям инвестиционного риска, который включает в себя экономические, социальные и экологические аспекты риска. Сформирована выборка, которая включает в себя признаки: преступления в экономической сфере на одну организацию (*pre_ud*); удельный вес убыточных организаций (*uo*), уровень безработицы (*ub*); уровень младенческой смертности (*ums*); сброс загрязненных сточных вод в водные объекты на душу населения (*sbd*).

Алгоритм интеллектуального анализа данных об инвестиционном риске предполагает выполнение следующих шагов: проведение компонентного анализа, построение диаграмм рассеивания и формирование первого варианта выделения кластеров регионов; проведение кластерного анализа, построение дендрограммы и формирование второго варианта выделения кластеров регионов; проведение нейросетевого анализа, построение самоорганизующихся карт Кохонена и формирование третьего варианта выделения кластеров регионов; построение трех вариантов деревьев решений на основе сформированных трех вариантов выделения кластеров, их анализ и выбор окончательного варианта формирования кластеров регионов.

КОМПОНЕНТНЫЙ И КЛАСТЕРНЫЙ АНАЛИЗ ДАННЫХ

На первом шаге алгоритма ИАД проводится компонентный анализ вначале полной выборки, а затем усеченной выборки. В полной выборке присутствуют все объекты (регионы), усеченная выборка формируется путем исключения ряда регионов, выделенных на основе результатов компонентного анализа полной выборки.

В результате проведения компонентного анализа полной выборки с использованием пакета StatGraphics построены три главные компоненты (ГК). В таблице 1 представлены весовые коэффициенты признаков для построенных ГК.

Таблица 1

Весовые коэффициенты главных компонент (полная выборка)

Название признака	Обозначение	Компонента F1	Компонента F2	Компонента F3
Преступления в экономической сфере на одну организацию	pre-ud	0,125142	0,65848	0,697219
Сброс загрязненных сточных вод в водные объекты на душу населения	sbd	-0,282313	-0,647969	0,530233
Уровень безработицы	ub	0,608948	-0,0526608	0,0445572
Уровень младенческой смертности	ums	0,532284	-0,0924827	-0,292185
Удельный вес убыточных организаций	uo	0,500495	-0,367714	0,381289

Показано, что главная компонента F1 в целом характеризует социально-экономические аспекты риска и определяется признаками: уровень удельного веса убыточных организаций (uo), младенческой смертности (ums), уровень безработицы (ub).

Построена диаграмма рассеивания, на которой выделены три кластера (рисунок 1). Кластер 1 объединяет в себе два региона, для которых характерен высокий инвестиционный риск (Республика Ингушетия, Чеченская Республика); регионы кластера 2 характеризуются довольно высоким инвестиционным риском (Республика Тыва, Республика Дагестан, Чукотский автономный округ, Еврейская автономная область). Кластер 3 включает все оставшиеся регионы, которые характеризуются средним и низким уровнем инвестиционного риска.

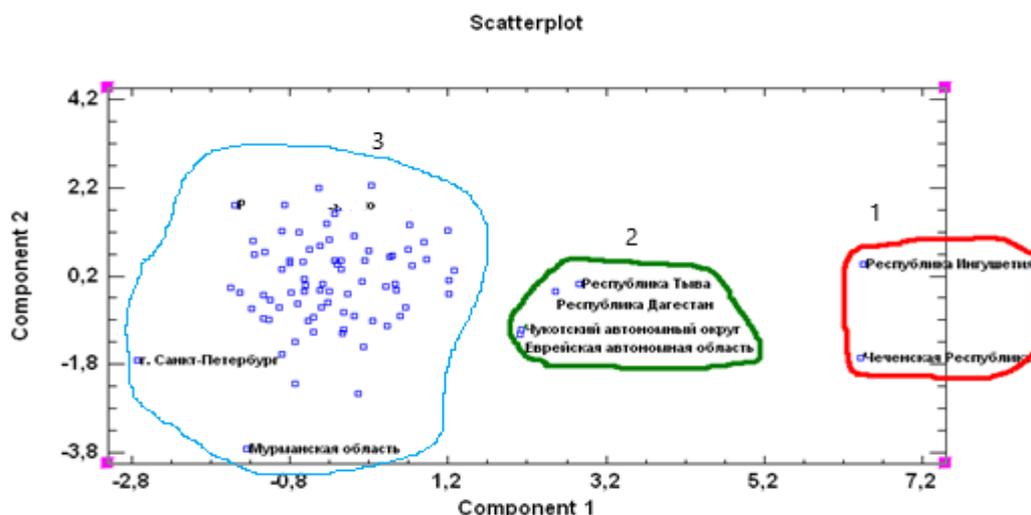


Рис. 1. 2D-диаграмма рассеивания в плоскости главных компонент F1 и F2 (полная выборка)

Компонента F2 характеризует экономический и экологический риск и определяется признаками: преступления в экономической сфере на одну организацию (pre_ud); сброс загрязненных сточных вод в водные объекты на душу населения (sbd). Наличие противоположных знаков у весовых коэффициентов признаков, определяющих ГК F2, трактуется как одновременное присутствие высокого (низкого) экономического риска и низкого (высокого) уровня экологического риска в характеристиках регионов, находящихся в крайней верхней (крайней нижней) зоне диаграммы. В области низких значений компоненты F2 находятся регионы, которые характеризуются низким экономическим риском и высоким экологическим риском, например, Мурманская область. Компонента F3 характеризует экономический и экологический риск (рисунок 2). Республика Башкортостан относится к кластеру 3, при этом занимает положение, близкое к низкому уровню социально-экономического риска, а также к среднему уровню экологического риска.

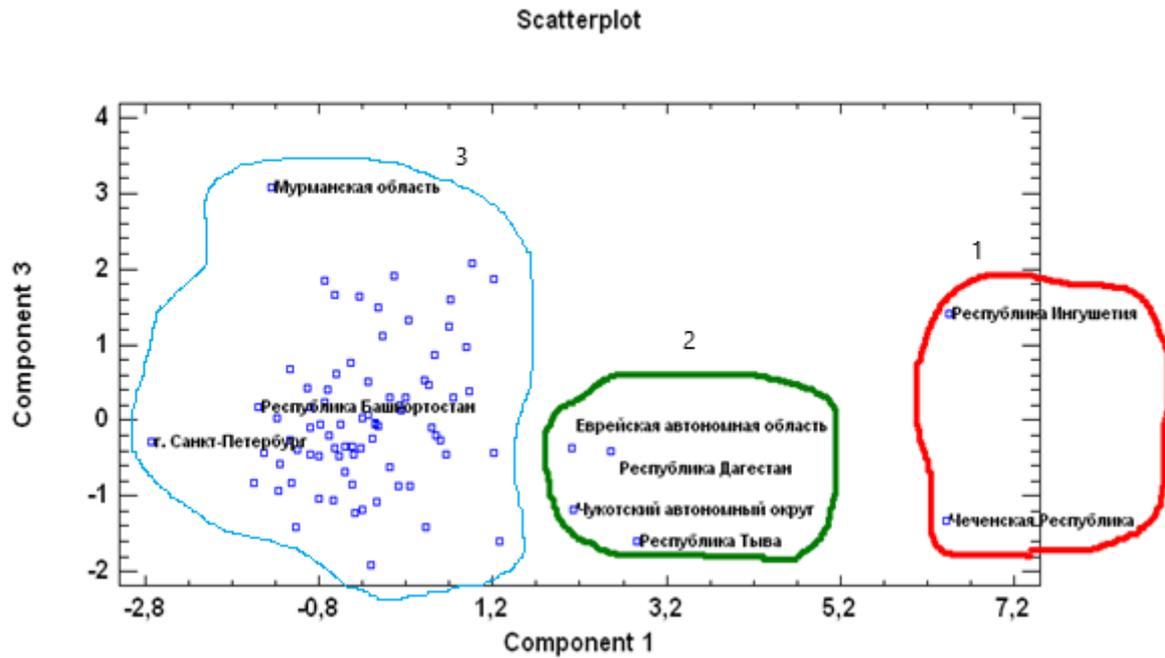


Рис. 2. 2D-диаграмма рассеивания в плоскости главных компонент F1 и F3

Поскольку наблюдается значительное скопление объектов в кластере 3, а объекты кластеров 1 и 2 занимают большую часть диаграммы рассеивания, то целесообразно удалить объекты первых двух кластеров из выборки.

Далее выполняется компонентный анализ данных для усеченной выборки. В таблице 2 представлены весовые коэффициенты признаков для трех построенных главных компонент.

Таблица 2

Весовые коэффициенты главных компонент (усеченная выборка)

Название признака	Обозначение	Компонента F1	Компонента F2	Компонента F3
Преступления в экономической сфере на одну организацию	pre-ud	0,414054	-0,278715	-0,729714
Сброс загрязненных сточных вод в водные объекты на душу населения	sbd	-0,209866	0,630398	0,113241
Уровень безработицы	ub	0,631138	0,218463	0,245296
Уровень младенческой смертности	ums	0,474432	-0,296395	0,57482
Удельный вес убыточных организаций	uo	0,401366	0,623971	-0,25319

Проведен анализ диаграммы рассеивания, на которой выделены кластеры 3, 4, 5, 6 и 7. (рисунок 3). Показано, что компонента F1 характеризует экономический и социальный риск. Довольно высоким уровнем экономического и социального риска обладают регионы кластера 7 (например, Карачаево-Черкесская Республика, Республика Северная Осетия-Алания). Средний уровень риска присущ регионам кластера 5 (например, Новгородская область, Курганская область). Низкий риск имеют регионы кластера 4 (например, Республика Башкортостан, Самарская область), г. Санкт-Петербург имеет очень низкий инвестиционный риск (кластер 3).

Компонента F2 характеризует сочетание влияния экономического и экологического рисков (одинаковые знаки коэффициентов). В области высоких значений компоненты F2 находятся регионы кластера 6, которые характеризуются высоким экономическим и экологическим риском одновременно (например, Мурманская область, Республика Калмыкия).

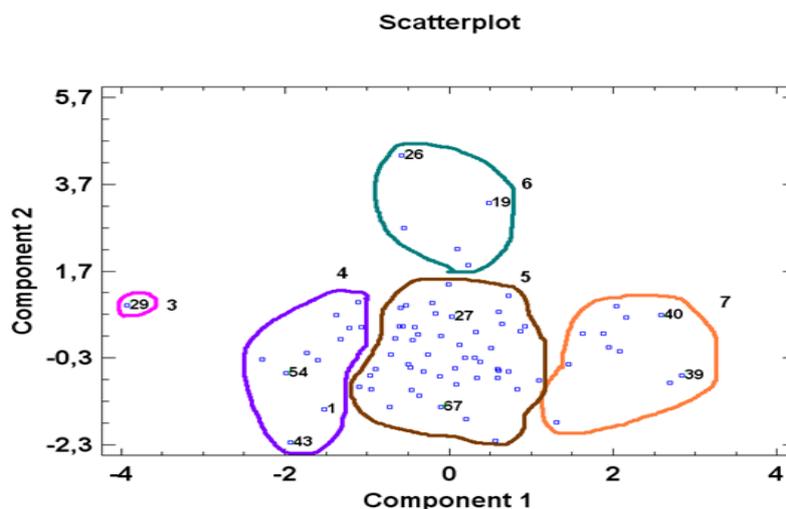


Рис. 3. 2D-диаграмма рассеивания в плоскости главных компонент F1 и F2 (усеченная выборка)

Компонента F3 характеризует экономический и социальный риск (рисунок 4). В области высоких значений компоненты находятся регионы, характеризующиеся высоким социальным риском, например, Ненецкий автономный округ, Карачаево-Черкесская Республика, при этом они имеют низкий экономический риск (весовой коэффициент имеет отрицательный знак). Регионы Амурская область, Республика Алтай имеют высокий экономический риск и низкий социальный риск.

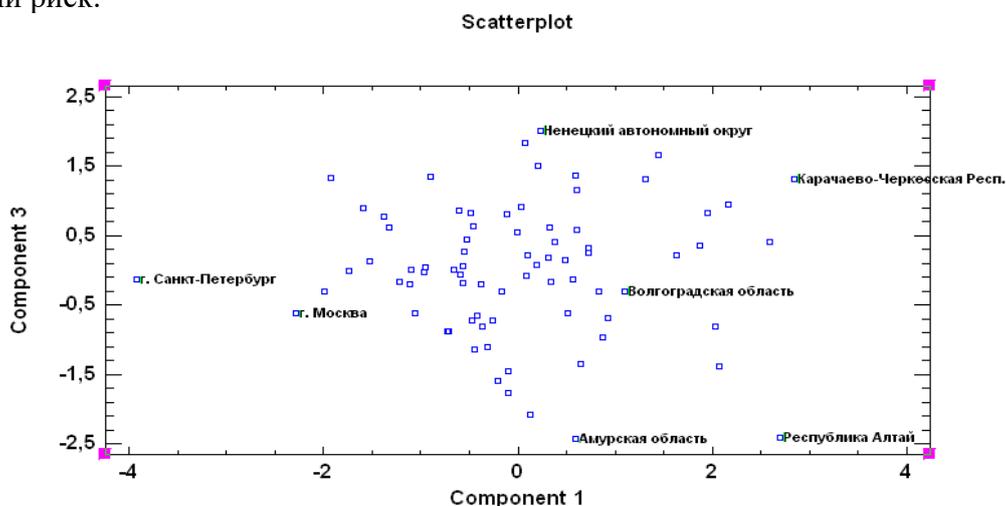


Рис. 4. 2D-диаграмма рассеивания в плоскости главных компонент F1 и F3 (усеченная выборка)

На втором шаге алгоритма ИАД выполняется кластерный анализ усеченной выборки. Все объекты разделились на 5 кластеров. Анализируя координаты центроидов построенных кластеров, можно сделать вывод о том, что значимыми признаками являются показатели преступления в экономической сфере на одну организацию, сброс загрязненных сточных вод в водные объекты на душу населения, уровень безработицы. На основе анализа координат центроидов сформулированы характеристики кластеров.

Сравнительный анализ состава кластеров, построенных по результатам компонентного и кластерного анализа, показал, что состав кластеров 1, 2, 4 и 6 одинаков за незначительными

исключениями. Однако два самых многочисленных кластера 5 и 7, выделенные в компонентном анализе и характеризующиеся средним и малым уровнями инвестиционной активности, в кластерном анализе разделены на четыре и два соответственно более мелких кластера. Кроме того, регионы Оренбургская и Амурская область, выделенные в отдельные кластеры в компонентном анализе, объединились с близлежащими кластерами.

ПОСТРОЕНИЕ КАРТ КОХОНЕНА

Третьим шагом алгоритма ИАД является проведение нейросетевого анализа с использованием нейронных сетей Кохонена для усеченной выборки.

По результатам обучения выделены пять кластеров. Самоорганизующиеся карты Кохонена (СОК), полученные в результате нейросетевого анализа, представлены на рисунке 5.

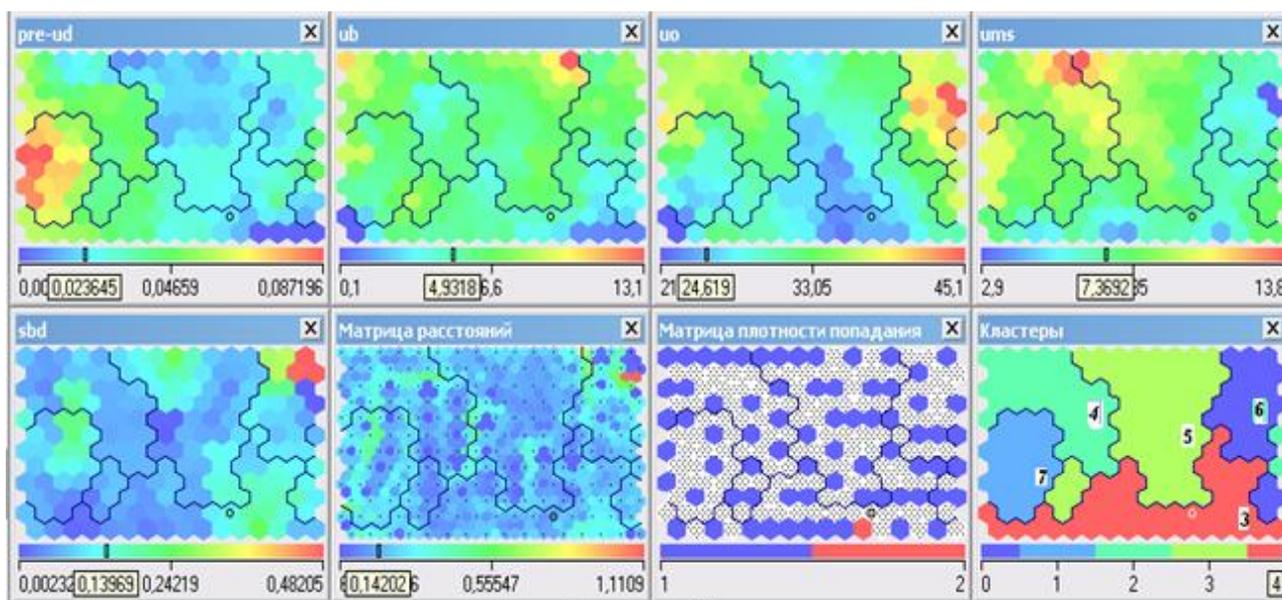


Рис. 5. Самоорганизующиеся карты Кохонена (усеченная выборка)

Анализ СОК позволяет сделать вывод о том, что признаки – преступления в экономической сфере на одну организацию (*pre_ud*); удельный вес убыточных организаций (*uo*); уровень безработицы (*ub*) – имеют достаточную дискриминантную силу, что подтверждается предыдущими видами анализа.

Для анализа характеристик кластеров произведен анализ соответствий кластеров, выделенных в компонентном и нейросетевом анализе, по их составу. На рисунке 5 введена двойная нумерация кластеров. Выявлены регионы с высоким уровнем экономического, экологического и социального риска. К ним относятся регионы кластеров 6 и 7. Кластер 3 также малочисленный и имеет низкий уровень риска, кластеры 4 и 5 характеризуются довольно низким и средним уровнем риска.

На основе данных о Ненецком АО проведен анализ с помощью диаграммы «Что-если». На рисунке 6 представлены данные о регионе и график зависимости принадлежности к кластеру от значений удельного веса убыточных организаций *uo*. Из рисунка видно, что Ненецкий АО находится в кластере 0 (кластер 6 по компонентному анализу), и удельный вес убыточных организаций $uo = 45,1$.

Построенная зависимость используется при формировании знаний о границах кластеров и условиях перехода из неблагоприятных кластеров в более благоприятные, характеризующиеся более низким уровнем риска. Для рассматриваемого региона требуется снизить удельный вес убыточных организаций до 34, чтобы перейти в кластер 4 (кластер 3 по компонентному анализу).

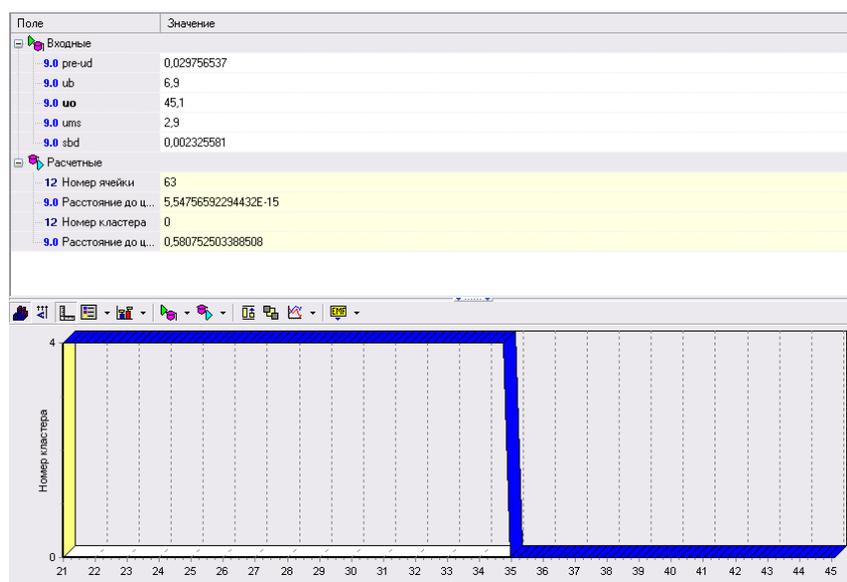


Рис. 6. Диаграмма «Что-если» (для Ненецкого АО)

ОБСУЖДЕНИЕ РЕЗУЛЬТАТОВ

Четвертым шагом алгоритма ИАД является построение трех деревьев решений по результатам проведения компонентного, кластерного анализа и построения карт Кохонена.

На рисунке 7 представлено дерево решений ДР1, построенное по результатам компонентного анализа.

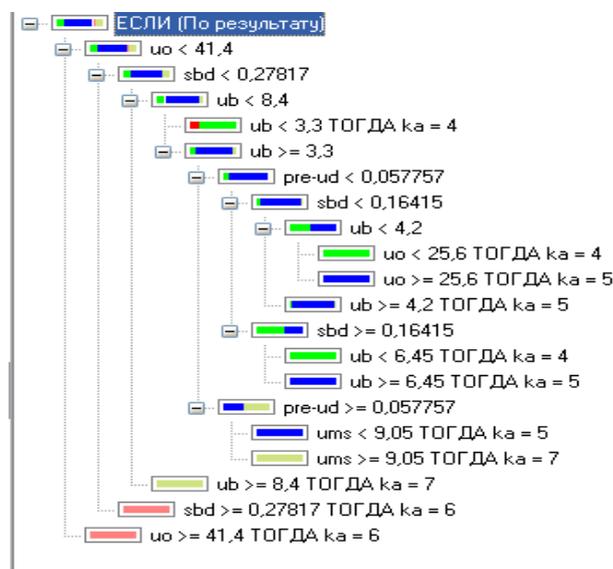


Рис. 7. Дерево решений ДР1, построенное по результатам компонентного анализа

По таблице сопряженности определены четыре ошибочно классифицированных региона: Санкт-Петербург, Тамбовская область, Республика Хакасия, Томская область.

Построены дерево решений ДР2 на основе результатов кластерного и дерево решений ДР3 на основе результатов нейросетевого анализа; количество ошибочно классифицированных объектов в них – 9 и 7 регионов соответственно. Выбор дерева решений и, следовательно, результатов формирования кластеров регионов сделан в пользу ДР1, построенного на основе результатов компонентного анализа. Такое решение обусловлено не только получением минимальной ошибки в ДР1, но и дополнительным анализом структуры построенных деревьев решений. При этом рассматривались вопросы соответствия переменных в первых трех

уровнях дерева переменным, наиболее значимым в компонентном анализе; сравнения количества извлеченных правил и глубины деревьев решений; интерпертируемости построенных правил.

Выводы

Таким образом, предложен алгоритм интеллектуального анализа региональных данных об инвестиционном риске, который предполагает проведение компонентного, кластерного и нейросетевого анализа для первичного формирования кластеров, а также построение трех вариантов деревьев решений для окончательного формирования кластеров регионов. Выявлены малочисленные кластеры регионов, характеризующиеся высоким уровнем риска по всем его составляющим, учитывающим экономические, социальные и экологические аспекты риска. Выявлены довольно малочисленные кластеры регионов с низким уровнем риска, которые имеют высокий уровень инвестиционного потенциала и инвестиционной активности. Оставшиеся регионы образовали несколько достаточно населенных кластеров, характеризующихся средним уровнем риска. Анализ выявленных кластеров регионов позволяет сформулировать их характерные особенности, а также сформировать правила оценки инвестиционного риска. Предложенный алгоритм интеллектуального анализа региональных данных об инвестиционном риске и результаты проведенного анализа целесообразно использовать в совокупности с алгоритмами и правилами оценки инвестиционной привлекательности и инвестиционного потенциала для формирования инвестиционного рейтинга регионов.

СПИСОК ЛИТЕРАТУРЫ / REFERENCES

1. Регионы России. Социально-экономические показатели: Статистический сборник. М.: Росстат, 2023. [[Regions of Russia. Socio-economic indicators. Statistical collection. Moscow: Rosstat, 2023. (In Russian).]]
2. Агеева Е. В., Баранов А. О. Прогноз развития инвестиционного комплекса России в 2022–2023 гг. // ЭКО. 2022. № 5. С. 111–130. [[Ageeva E. V., Baranov A. O. "Forecast for the development of the investment complex of Russia in 2022–2023" // IVF. 2022. No. 5. pp. 111–130. (In Russian).]]
3. Зубаревич Н. В. Регионы России в новых экономических условиях // Журнал Новой экономической ассоциации. 2022. № 3 (55). С. 226–234. [[Zubarevich N. V. "Regions of Russia in new economic conditions" // Journal of the New Economic Association. 2022. No. 3 (55), pp. 226–234. (In Russian).]]
4. Кузнецов В. И., Владимиров Н. А., Сычева М. А. О дифференциации регионов Российской Федерации по уровню инвестиционной привлекательности // Статистика и экономика. 2019. № 16. С. 25–33. DOI: <https://doi.org/10.21686/2500-3925-2019-2-25-33> [[Kuznetsov V. I., Vladimirov N. A., Sycheva M. A. "On the differentiation of regions of the Russian Federation according to the level of investment attractiveness" // Statistics and Economics. 2019. No. 16, pp. 25–33. (In Russian).]]
5. Методика составления рейтинга инвестиционной привлекательности регионов России компании «РАЭК-Аналитика». [Электронный ресурс] URL: <https://raex-rr.com/> [[Methodology for Compiling a Rating of Investment Attractiveness of Russian Regions by RAEX-Analytics. [Electronic resource] URL: <https://raex-rr.com/> (In Russian).]]
6. Национальное рейтинговое агентство. [Электронный ресурс]. URL: <https://www.ra-national.ru/> [[National Rating Agency. [Electronic resource]. URL: <https://www.ra-national.ru/> (In Russian).]]
7. Национальный инвестиционный рейтинг. [Электронный ресурс]. URL: https://asi.ru/government_officials/rating/ [[National Investment Rating. [Electronic resource]. URL: https://asi.ru/government_officials/rating/ (In Russian).]]
8. Кузнецова С. В., Смирнова О. А. Анализ подходов к оценке инвестиционной привлекательности региона // Современные наукоемкие технологии. Региональное приложение. 2016. № 1. С. 44–51. [[Kuznetsova S. V., Smirnova O. A. "Analysis of approaches to assessing the investment attractiveness of a region" // Modern Science-Intensive Technologies. Regional Application. 2016. No. 1, pp. 44–51. (In Russian).]]
9. Масленникова Е. В. Инвестиционная привлекательность региона: методики оценки и рейтинги инвестиционной привлекательности // Общество, экономика, управление. 2017. № 2. С. 37–40. [[Maslennikova E. V. "Investment attractiveness of the region: assessment methods and ratings of investment attractiveness" // Society, Economics, Management. 2017. No. 2, pp. 37–40. (In Russian).]]
10. Хомякова А. А., Мизгирев Л. С., Шергин В. В. Использование методов интеллектуального анализа данных в процессах управления инвестиционной привлекательностью региона // Известия высших учебных заведений. Серия: Экономика, финансы и управление производством. 2020. № 2. С. 14–23. [[Khomyakova A. A., Mizgirev L. S., Shergin V. V. "Using methods of intellectual data analysis in the processes of managing the investment attractiveness of the region" // News of Higher Educational Institutions. Series: Economics, Finance, and Production Management. 2020. No. 2, pp. 14–23. (In Russian).]]
11. Огородников П. И., Матвеева О. Б., Гусева Е. П. Компонентный анализ инвестиционной привлекательности сельскохозяйственных районов Центральной зоны Оренбургской области, их ранжирование и классификация // НАУКОВЕДЕНИЕ.

2017. Т. 9. № 6 [Электронный ресурс]. URL: <https://naukovedenie.ru/PDF/167EVN617.pdf> [[Ogorodnikov P. I., Matveeva O. B., Guseva E. P. "Component analysis of the investment attractiveness of agricultural areas of the Central zone of the Orenburg region, their ranking and classification" // SCIENCE. 2017. Vol. 9, No. 6 [Electronic resource]. URL: <https://naukovedenie.ru/PDF/167EVN617.pdf> (In Russian).]]
12. Казанская А. А., Мишура Л. Г. Использование машинного обучения в инвестиционной деятельности // Научный журнал НИУ ИТМО. Серия Экономика и экологический менеджмент. 2020. № 2. С. 23-34. [[Kazanskaya A. A., Mishura L. G. "Using machine learning in investment activities" // Scientific Journal of NRU ITMO. Series "Economics and Environmental Management". 2020. No. 2, pp. 23-34. (In Russian).]]
 13. Мусацков В. Ю., Афанасьева Г. А. Искусственные нейронные сети в экономике и в экологии // Проблемы региональной экологии. 2023. № 4. С. 51-55. [[Musatskov V. Yu., Afanasyeva G. A. "Artificial neural networks in economics and ecology" // Problems of Regional Ecology. 2023. No. 4, pp. 51-55. (In Russian).]]
 14. Бушуева Л. И., Кузнецов В. И., Попова Ю. Ф. Применение метода главных компонент для анализа ключевых показателей развития конкуренции на рынках товаров и услуг Северо-Западного федерального округа Российской Федерации // Корпоративное управление и инновационное развитие экономики Севера: Вестник Научно-исследовательского центра корпоративного права, управления и венчурного инвестирования Сыктывкарского государственного университета. 2023. Т. 3. № 2. С. 124-134. [[Bushueva L. I., Kuznetsov V. I., Popova Yu. F. "Application of the principal component's method for the analysis of key indicators of competition development in the markets of goods and services of the North-Western Federal District of the Russian Federation" // Corporate Governance and Innovative Development of the Economy of the North : Bulletin of the Research Center for Corporate Law, Management and Venture Investment of Syktvykar State University. 2023. Vol. 3, No. 2, pp. 124-134. (In Russian).]]
 15. Козак Е. Применение нейронных сетей в экономике // Экономика: вчера, сегодня, завтра. 2021. Т. 11. № 5А. С. 113-119. DOI: 10.34670/AR.2021.43.81.014 [[Kozak E. "Application of neural networks in economics" // Economics: Yesterday, Today, Tomorrow. 2021. Vol. 11, No. 5A, pp. 113-119. (In Russian).]]
 16. Трифонов Ю. В., Сочков А. Л., Соловьев А. Е. Оценка экономического потенциала регионов РФ на основе методологии нейросетевого кластерного анализа // Вестник Нижегородского университета им. Н. И. Лобачевского. Серия Социальные науки. 2021. № 3(63). С. 38-47. [[Trifonov Yu. V., Sochkov A. L., Solovyov A. E. "Assessment of the economic potential of regions of the Russian Federation based on the methodology of neural network cluster analysis" // Bulletin of the Nizhny Novgorod University. Series: Social sciences. 2021. No. 3(63), pp. 38-47. (In Russian).]]
 17. Серова Е. Г., Лукинский В. С. Методы и инструменты интеллектуального анализа данных в цифровой логистике и управлении цепями поставок // Логистика и управление цепями поставок. 2018. № 4(87). С. 73-80. [[Serova E. G., Lukinsky V. S. "Methods and tools for data mining in digital logistics and supply chain management" // Logistics and Supply Chain Management. 2018. No. 4(87), pp. 73-80. (In Russian).]]
 18. Василенко М. А., Кузина Е. Л., Тагильцева Ю. А., Прокопова А. М. Нейронные сети в решении экологических проблем развития транспортных услуг // Инновации и инвестиции. 2023. № 4. С. 281-284. [[Vasilenko M. A., Kuzina E. L., Tagiltseva Yu. A., Prokорова А. М. "Neural networks in solving environmental problems in the development of transport services" // Innovations and Investments. 2023. No. 4, pp. 281-284. (In Russian).]]
 19. Димов Э. М., Ильясов Б. Г., Макарова Е. А., Закиева Е. Ш., Ефтонова Т. А., Гиздатуллина Э. С. Методология системного динамического моделирования и управления функционированием многоотраслевого производственного комплекса в рамках воспроизводственного процесса макроуровня // Инфокоммуникационные технологии. 2018. Т. 16. № 1. С. 81-96. DOI 10.18469/ikt.2018.16.1.09. EDN ХТНВКН. [[Dimov E. M., Ilyasov B. G., Makarova E. A., Zakieva E. Sh., Eftonova T. A., Gizdatullina E. S. "Methodology of system dynamic modeling and management of the functioning of a diversified production complex within the framework of the macro-level reproduction process" // Infocommunication Technologies. 2018. Vol. 16, No. 1, pp. 81-96. (In Russian).]]
 20. Ильясов Б. Г., Димов Э. М., Макарова Е. А., Ефтонова Т. А. Нейросетевой анализ динамически неравновесных ситуаций взаимодействия секторов экономики с использованием имитационной модели // Инфокоммуникационные технологии 2016. Т. 14. № 3. С. 285-300. [[Ilyasov B. G., Dimov E. M., Makarova E. A., Eftonova T. A. "Neural network analysis of dynamically nonequilibrium situations of interaction between economic sectors using a simulation model" // Infocommunication Technologies Vol. 14, No. 3, pp. 285-300. (In Russian).]]
 21. Ильясов Б. Г., Макарова Е. А., Закиева Е. Ш., Бойцов А. Н. Алгоритм анализа данных об инвестиционных процессах в регионах РФ с использованием метода главных компонент // Современные наукоемкие технологии. 2023. № 6. С. 27-34. [[Ilyasov B. G., Makarova E. A., Zakieva E. Sh., Boytsov A. N. "Algorithm for analyzing data on investment processes in the regions of the Russian Federation using the principal component method" // Modern Science-Intensive Technologies. 2023. No. 6, pp. 27-34. (In Russian).]]
 22. Закиева Е. Ш. Методология поддержки принятия решений при управлении социальной системой на основе динамического моделирования и интеллектуальных технологий // Системная инженерия и информационные технологии. 2023. Т. 5. № 3 (12). С. 69-92. EDN: UWIPDO. [[Zakieva E. Sh. "Methodology for decision support in managing a societal system based on dynamic modeling and intelligent technologies" // System Engineering and Information Technologies. 2023. Vol. 5, No. 3 (12), pp. 69-92. EDN: UWIPDO. (In Russian).]]
 23. Ильясов Б. Г., Макарова Е. А., Закиева Е. Ш., Гиздатуллина Э. С. Оценка данных о доходах населения в региональном разрезе методом главных компонент // Экономика региона. 2019. Т. 15. Вып. 2. С. 601-617. [[Ilyasov B. G., Makarova E. A., Zakieva E. Sh., Gizdatullina E. S. "Assessment of data on population income in a regional context using the principal component method" // Economics of the Region. 2019. Vol. 15, issue 2, pp. 601-617. (In Russian).]]

24. Юсупов М. М., Макарова Е. А., Камаева Р. Р. Анализ дифференциации потребительских расходов домохозяйств на основе агент-ориентированного моделирования // Системная инженерия и информационные технологии. 2023. Т. 5. № 6(15). С. 57-66. EDN: PVCZYZ. [[Yusupov M. M., Makarova E. A., Kamaeva R. R. "Analysis of differentiation of household consumer expenditures based on agent-based modeling" // System Engineering and Information Technologies. 2023. Vol. 5, No. 6(15), pp. 57-66. EDN: PVCZYZ. (In Russian).]]
25. Шурьгин А. С., Макарова Е. А. Система агент-ориентированного моделирования функционирования кластеров предприятий с учётом налогообложения // Системная инженерия и информационные технологии. 2023. Т. 5. № 6(15). С. 24-31. EDN: TNDUPQ. [[Shurygin A. S., Makarova E. A. "System of agent-based modeling of the functioning of clusters of enterprises taking into account taxation" // System Engineering and Information Technologies. 2023. Vol. 5, No. 6(15), pp. 24-31. EDN: TNDUPQ. (In Russian).]]
26. Дегтярева И. В., Макарова Е. А., Павлова А. Н. Интеллектуальные алгоритмы анализа ситуаций и принятия решений при взаимодействии секторов экономики в неопределённых рыночных условиях на основе нейронных сетей Кохонена // Управление экономикой: методы, модели, технологии: Мат-лы XVI Межд. науч. конф. Уфа: РИК УГАТУ, 2016. С. 457-466. [[Degtyareva I. V., Makarova E. A., Pavlova A. N. "Intelligent algorithms for analyzing situations and making decisions in the interaction of economic sectors in uncertain market conditions based on Kohonen neural networks" // Economic Management: Methods, Models, Technologies: Proc. of the XVI Int. scientific conf. Ufa, 2016, pp. 457-466. (In Russian).]]

Поступила в редакцию 20 февраля 2024 г.

МЕТАДАННЫЕ / METADATA

Title: Algorithm for regional data mining about investment risk.

Abstract: The article proposes an algorithm for regional data mining about investment risk, which involves principal component, cluster and neural network analysis for the initial formation of clusters, as well as the construction of three decision tree variants for the final formation of clusters of regions. Numerous clusters of regions characterized by a high level of economic, social and environmental risk have been identified. The analysis of the identified clusters of regions allows us to formulate their characteristic features, as well as to form rules for assessing investment risk.

Key words: investment risk, data mining, principal component analysis, Kohonen neural networks, cluster, decision trees.

Язык статьи / Language: русский / Russian.

Об авторах / About the authors:

МАКАРОВА Елена Анатольевна

ФГБОУ ВО «Уфимский университет науки и технологий», Россия. Профессор каф. технической кибернетики. Дипл. инж.-системотехник (Уфимск. авиац. ин-т, 1982). Д-р техн. наук по управлению в соц. и экон. системах (Уфимск. гос. авиац. техн. ун-т, 2011). Иссл. в обл. интелл. поддержки принятия управл. решений в многосекторных макроэкономических системах. E-mail: ea-makarova@mail.ru ORCID: <https://orcid.org/0000-0003-0043-3693>

MAKAROVA Elena Anatolyevna

Ufa University of Science and Technologies, Russia. Professor of the Department of Technical Cybernetics. Dipl. Systems Engineer (Ufa Aviation Institute, MAKAROVA). Dr Tech Sci in management in social and economic systems (Ufa State Aviation Technical University, 2011). Research in the field of intellectual support for management decisions in macroeconomic systems.. E-mail: ea-makarova@mail.ru ORCID: <https://orcid.org/0000-0003-0043-3693>

ГАБДУЛЛИНА Эльвира Риятовна

ФГБОУ ВО «Уфимский университет науки и технологий», Россия. Доцент кафедры технической кибернетики. Иссл. в обл. интелл. поддержки принятия управл. решений в многосекторных макроэкономических системах E-mail: gabdullina.er@ugatu.su ORCID: <https://orcid.org/0000-0002-2421-1510>

GABDULLINA Elvira Riyatovna

Ufa University of Science and Technologies, Russia. Associate Professor of the Department of Technical Cybernetics. Research in the field of intellectual support for management decisions in macroeconomic systems. E-mail: gabdullina.er@ugatu.su ORCID: <https://orcid.org/0000-0002-2421-1510>

ЮСУПОВ Марат Маннурович

ФГБОУ ВО «Уфимский университет науки и технологий», Россия. Аспирант ин-та информатики, математики и робототехники. Иссл. в обл. динамического моделирования и принятия решений в социально-экономических системах E-mail: yusu-marat@yandex.com ail.com ORCID: <https://orcid.org/0009-0005-8308-7053>

YUSUPOV Marat Mannurovich

Ufa University of Science and Technologies, Russia. Postgraduate student, Institute of Informatics, Mathematics, and Robotics. Research in the field of dynamic modeling and decision-making in socio-economic systems. E-mail: yusu-marat@yandex.com ail.com ORCID: <https://orcid.org/0009-0005-8308-7053>

ВАГАПОВА Гульсум Рустамовна

ФГБОУ ВО «Уфимский университет науки и технологий», Россия. Магистрант ин-та информатики, математики и робототехники. Иссл. в обл. принятия решений при управлении инвестиционным развитием регионов E-mail: gulsumv@bk.ru ORCID: <https://orcid.org/0009-0005-2213-467X>

VAGAPOVA Gulsum Rustamovna

Ufa University of Science and Technologies, Russia. Master's student at the Institute of Informatics, Mathematics and Robotics. Research in the field of decision-making to manage investment development of regions. E-mail: gulsumv@bk.ru ORCID: <https://orcid.org/0009-0005-2213-467X>